

Image Registration in Hough Space Using Gradient of Images

Ramtin Shams

Research School of Information Sciences and Engineering (RSISE)
The Australian National University (ANU)
Canberra, ACT 0200
ramtin.shams@anu.edu.au

Nick Barnes and Richard Hartley

The Australian National University (ANU) and NICTA*
Canberra, ACT 0200
nick.barnes@nicta.com.au

Abstract

We present an accurate and fast method for rigid registration of images with large non-overlapping areas using a Hough transformation of image gradients. The Hough space representation of gradients can be used to separate estimation of the rotation parameter from the translation. It also allows us to estimate transformation parameters for 2D images over a 1D space, hence reducing the computational complexity. The cost functions in the Hough domain have larger capture ranges compared to the cost functions in the intensity domain. This allows the optimization to converge better in the presence of large misalignments. We show that the combination of estimating registration parameters in the Hough domain and fine tuning the results in the intensity domain significantly improves performance of the application compared to the conventional intensity-based multi-resolution methods.

1. Introduction

The aim of registration is to find the optimal transformation, which best aligns two or more images by minimizing some cost function. Registration is a fundamental task frequently encountered in various domains, including surveillance, remote sensing, environmental monitoring, mapping, and medical imaging [2].

Given a suitable cost function, registration is an N -dimensional optimization problem, where N is the num-

ber of parameters that comprise the transformation space. Typically, at each iteration of the optimization algorithm, the moving (or floating) image I^M is transformed and the cost function is calculated against the fixed (or reference) image I^F . The optimal transformation T_{opt} is determined by finding the transformation parameters that minimize the cost function

$$T_{opt} = \underset{T}{\operatorname{argmin}} -S(I^F; I^{M(T)}), \quad (1)$$

where, $-S$ is the cost function (S is the similarity measure), T is the transformation, and $I^{M(T)}$ is the transformed moving image. Image transformation and cost function computation are time consuming. It is hence desirable to reduce the number of iterations required by the optimization algorithm to converge.

The cost function should ideally be smooth, have a large capture range and a single minimum that coincides with the transformation parameters that align the images. However, in practice, cost functions may contain several local minima and have a limited capture range that can prevent the optimization algorithm from converging to the global minimum, resulting in mis-registration.

The performance of the registration task can be enhanced by improving the shape of the cost function or by proper initialization of the optimization algorithm to avoid local minima and start within the attraction range of the cost function. The latter usually entails obtaining estimates of the registration parameters. We achieve this by transforming image gradients to Hough space and using a mutual information (MI) cost function.

MI-based cost functions that operate on pixel intensities have been very successful in registration of multi-modal images in the medical imaging domain [9]. This is due to the

*NICTA is a research centre funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council, through Backing Australia's Ability and the ICT Research Centre of Excellence programs.

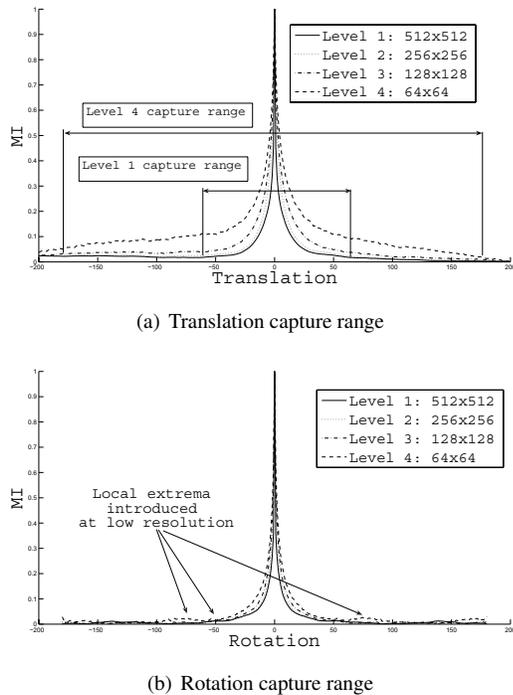


Figure 1. Sample MI functions computed for the image shown in Fig.5(a). (a) Translation capture range improves at lower resolutions. (b) Rotation capture range is almost unchanged at lower resolutions and local extrema are introduced.

fact that the cost function does not require a linear relationship between the corresponding intensities, is very accurate, can be used automatically, does not rely on features, and can be virtually applied to any type of image. However, one often overlooked aspect is the limited capture range of MI-based methods [13]. This may not be a major issue if the misalignment between the images is small, but limits the applicability of the method as a general purpose registration technique. Multi-resolution methods provide some relief and extend the capture range of translation parameters (Fig.1(a)), but are less effective for the rotation or scale (Fig.1(b)). They can also introduce additional local extrema at lower resolutions (Fig.1(b)). The lower resolutions have also been reported to be more sensitive to the initial order of the parameters that are being optimized [8].

In [10] and [7], authors show that the shape of the pixel intensity-based MI function can be improved by introducing a correction factor based on the co-occurrence of image gradients and mutual information (MI) of image gradients, respectively. The improved robustness of the registration task, using these methods, comes at the cost of in-

creasing the computational complexity of the cost function by introducing the image gradient calculation at each step. Our method, however, improves performance and computational efficiency at the same time.

2. Approach and Contribution

We consider the class of proper rigid transformations (proper rotation and translation). We improve accuracy, robustness and efficiency of the registration task by acquiring a close initial estimate of the registration parameters.

We calculate image gradients first and assume that each gradient vector represents a directional line that can be parametrized by its angle ϕ and its orthogonal distance from the origin ρ as shown in Fig. 2. The Hough transformation of the gradient vectors, obtained in this manner, provides a spatially relaxed function of the image shape that we use for registration of the images (more on this in Section 3). We show that this transformation separates estimation of the rotation parameter from translation and scale. Rotation, in the image domain, transforms to a circular shift in the Hough domain, which is invariant to the translation and scale between the images. This allows us to obtain a robust estimate of the rotation parameter quickly by performing an exhaustive search on a 1-parameter cost function that operates on a 1D data-set.

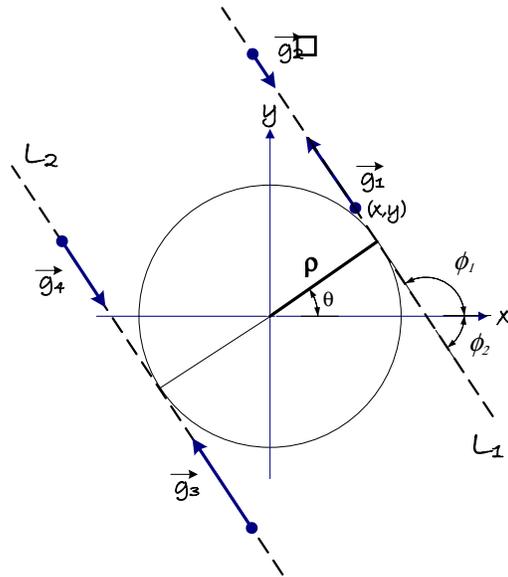


Figure 2. Gradient vectors in Cartesian coordinates and corresponding Hough space parameters. Each gradient vector is parametrized with its orthogonal distance ρ from the origin and its angle ϕ .

Given the rotation parameter, we obtain estimates of the translation parameters by optimizing a similarity measure along a number of 1D lines parallel to the ρ axis in Hough space. We back-project the cost functions into a 2D translation space to obtain a robust estimate of the x and y translations.

Finally, we use estimated parameters to initialize an optimization algorithm that operates on a conventional pixel intensity-based cost function in a multi-resolution set-up to achieve sub-pixel accuracy.

Our method has a number of benefits such as reduction in dimensionality of the problem, increased efficiency, applicability to multi-modal images, applicability to images with large intensity variances, and tolerance to large non-overlapping areas. The combination of these properties and use of the Hough transform of gradient images rather than the intensity images differentiates our work from existing approaches such as the gradient field method in [6], intensity-based Hough registration in [3], and the gradient-intensity method in [13]. The rotation estimation in [13] is robust w.r.t to non-overlapping areas but the centroid-based translation estimations are sensitive to non-overlapping areas of the image.

In Section 4, we show that our method significantly improves performance while maintaining a high level of accuracy and computational efficiency. The parameter estimations are robust w.r.t non-overlapping areas. We achieve 99% and 95% success rate (for two sets of experiments) on a wide range of images with a large dynamic range, whereas under the same conditions the conventional multi-resolution method achieves 24% and 15% success rate.

3. Method

3.1 Image Gradients

Let \mathbf{I} be an image and \mathbf{K}_x and \mathbf{K}_y be differentiating kernels in x and y directions, the gradient image is then calculated using (2) and (4).

$$\mathbf{G}_x = \mathbf{K}_x \otimes \mathbf{I}, \quad \mathbf{G}_y = \mathbf{K}_y \otimes \mathbf{I}, \quad (2)$$

$$r_g(x, y) = \sqrt{g_x(x, y)^2 + g_y(x, y)^2}, \quad (3)$$

$$\phi_g(x, y) = \arctan \frac{g_y(x, y)}{g_x(x, y)}, \quad (4)$$

where \mathbf{G}_x and \mathbf{G}_y are image gradient matrices in x and y directions, r_g and ϕ_g are the gradient magnitude and phase at position (x, y) , respectively, and \otimes denotes convolution.

Note that the gradient calculations only need to be done once for each image and as such, we can use a more accurate filter with wider support without affecting the overall efficiency of the algorithm. We use a 7×7 derivative

of a Gaussian filter (which has the desired effect of differentiating the smoothed image [5]) with standard deviation $\sigma = 1.5$. The separability property of differentiating Gaussian kernels was used to speed up the convolutions.

3.2 Gradient Image Transformation in Hough Space

Image gradients can be fully described by their spatial location (x, y) , magnitude and phase. We assume that each gradient represents a line L parallel to the direction of the gradient vector which also passes through the gradient's spatial location as shown in Fig. 2. Each gradient vector is parametrized with its phase (angle of L) and the distance of L from the origin. This parametrization of the gradient field, drops the gradient magnitude and relaxes the spatial dependency of the gradient vector (all parallel gradient vectors on L are represented with the same parameters). We will show how this parametrization allows us to break down rigid registration, which is normally a 3-parameter optimization problem, to three simpler 1-parameter optimizations.

A gradient vector at spatial coordinates (x, y) with angle ϕ (as shown in Fig. 2) can be represented in Hough space by (ρ, ϕ) , where

$$\rho = y \cos \phi - x \sin \phi. \quad (5)$$

Gradient parametrization is different from the standard Hough parametrization for lines and takes direction into account. In gradient parametrization, ρ is a signed real number and the gradient angle ϕ is used instead of the perpendicular line angle θ . Using a signed ρ prevents formation of unwanted local minima in the cost function, that would otherwise be created due to ambiguity of the standard Hough parametrization of the lines that ignores direction.

The similarity (rigid + scale) transformation of an image can be written as

$$\mathbf{x}' = s\mathbf{R}\mathbf{x} + \mathbf{t}, \quad \text{or} \quad (6)$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = s \begin{bmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (7)$$

where t_x and t_y are translation parameters, γ is the rotation angle, and s is scale. The transformation will result in the following change in Hough space

$$\begin{aligned} \rho' &= y' \cos \phi' - x' \sin \phi', & (8) \\ \rho' &= s[y \cos(\phi' - \gamma) - x \sin(\phi' - \gamma)] + \underbrace{t_y \cos \phi' - t_x \sin \phi'}_{\rho_t(\phi')}. & (9) \end{aligned}$$

We now re-organize (9) so that it is comparable with the line equation in (5)

$$\frac{\rho' - \rho_t(\phi')}{s} = y \cos(\phi' - \gamma) - x \sin(\phi' - \gamma), \quad (10)$$

by comparing (10) with (5) we have

$$\begin{bmatrix} \rho' \\ \phi' \end{bmatrix} = \begin{bmatrix} s\rho + \rho_t(\phi + \gamma) \\ \phi + \gamma \end{bmatrix} \quad (11)$$

Equation (11) specifies a similarity transformation in Hough space. One immediate benefit is the separation of the rotation parameter from the translation and scale. Additionally, transformation in Hough space is computationally efficient as it consists of a shift along the ϕ axis and a multiplication and a constant shift for each value of ϕ along the ρ axis.

We now define the Hough transform function $\Gamma(\rho, \phi)$, as the number of gradient vectors at a particular direction ϕ and at a distance ρ from the origin. In practice, we need a discrete version of $\Gamma(\cdot)$. We use the *uniform volume histogram* [12], which has been shown to be robust w.r.t. gradient calculation errors and noise and results in smoother cost function for image registration purposes [12]. We will discuss how $\Gamma(\cdot)$ can be used for image registration in the next section.

3.3. Estimation of Transformation Parameters

Using (11) we have

$$\Gamma'(\rho, \phi) = \Gamma(s\rho + \rho_t, \phi + \gamma), \quad (12)$$

$$\sum_{\rho} \Gamma'(\rho, \phi) = \sum_{\rho} \Gamma(s\rho + \rho_t, \phi + \gamma), \quad (13)$$

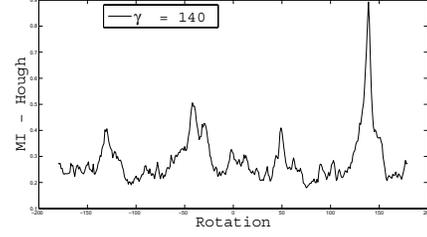
$$\text{or } \overline{\Gamma'}(\phi) = \overline{\Gamma}(\phi + \gamma), \quad (14)$$

where $\Gamma'(\cdot)$ is the gradient function of the transformed image and $\overline{\Gamma}(\cdot)$ and $\overline{\Gamma'}(\cdot)$ are the average of gradient functions along the ρ axis.

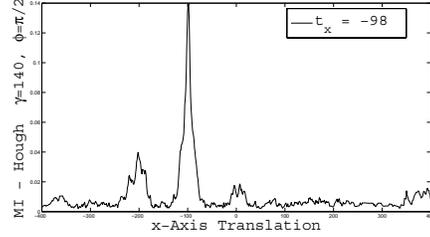
Equations (12) and (13) allow us to decouple estimation of the transformation parameters. It can be readily seen that rotation reduces to a shift along the ϕ axis in the Hough domain and is independent of the scale and translation. As such, finding the rotation parameter between the images reduces to a 1-parameter optimization problem over the rotation dynamic range of $(-\pi, \pi]$

$$\hat{\gamma} = \underset{\gamma}{\operatorname{argmin}} -MI(\overline{\Gamma'}(\phi); \overline{\Gamma}(\phi + \gamma)), \quad (15)$$

where MI is the mutual information similarity measure [9]. In practice, due to the small overhead in calculating the 1D cost function, we use an exhaustive optimization strategy by calculating the cost function at 1° intervals and choosing



(a) 1D MI function (rotation)



(b) 1D MI function (x-axis translation)

Figure 3. 1D MI functions are shown for images of Fig.5(a). The misalignment between the images is $[-98.03 \ -97.00 \ 140.52^\circ]$. Our method estimates the transformation parameters as $[-98 \ -97 \ 140^\circ]$. Note that the shape of MI functions are unimportant at this stage, since we are performing an exhaustive search to estimate the parameters.

the rotation parameter which results in the minimum cost. Resolving γ is an important step, since multi-resolution methods perform much better if the images are rotationally aligned.

We can further exploit Hough space transformation to quickly estimate translation parameters. For rigid registration ($s = 1$), we pick a number of lines parallel to the ρ axis and calculate MI for all values of ρ_t in the dynamic range at 1 pixel intervals. MI calculation is very efficient since we are computing the similarity for two 1D data-sets (two lines in Hough space). Let us pick $\phi = \pi/2$ line for which $\rho_t = t_x$. As shown in Fig. 3(b), the MI function attains its maximum at t_x . Similarly, we could estimate t_y by looking at $\phi = 0$ line. However, this may not always be robust as demonstrated in Fig. 4. We can improve the robustness of the estimation by looking at the similarity functions of other lines, which provide combined estimates of t_x and t_y ($\rho_t(\phi) = t_y \cos \phi - t_x \sin \phi$).

Let us treat the MI function as the probability distribution of the random variable ρ_t . For each line $\phi = \alpha$ we have

$$\Pr(\rho_t(\alpha)) \propto MI(\Gamma'(\rho, \alpha); \Gamma(\rho + \rho_t(\alpha), \alpha + \hat{\gamma})). \quad (16)$$

Equation (16) may be thought of as a projection of the probability distribution $\text{Pr}(\rho_t)$ in the direction determined by $\phi = \alpha$. Our aim is to find the best (most probable) estimate of the translation parameters t_x and t_y given a number of projections of $\text{Pr}(\rho_t)$ in different directions. This can be achieved by a process of back-projection analogous to that used in computing the inverse Radon transform. We back-project $\text{Pr}(\rho_t)$ for a number of lines, and find the estimate of translation parameters where the back-projected probability density function achieves its maximum, as shown in Fig. 4. In practice no more than 10 lines were required for achieving robust results.

3.4. Final Optimization

As the final step, we switch back to the intensity domain to obtain more accurate results. We use a multi-resolution Gaussian pyramid which is optimized using the Simplex [11] method. The multi-resolution optimization is initialized with the estimated parameters. The λ parameter which determines the volume of the initial simplex is set to a small value (10 pixels along the translation axis and 10° along the rotation axis) to allow the optimization to converge quickly.

4. Experiments

We tested our method on aerial images from Google Earth [1]. We used 20 images in our experiments selected from four categories: landmark, urban, rural and natural. *Landmark* images include a prominent man-made structure, *urban* images are taken from city centers and mostly contain rectilinear structures, *rural* images contain few buildings and are otherwise occupied by natural scenery, and *natural* images contain no artificial structures or roads. One sample registration per category is shown in Fig. 5.

In total, we performed 8000 registrations in two sets of experiments by generating 200 random transformations for each image within the angular dynamic range of $(-\pi, \pi]$. The translational dynamic range was set to $[-100, 100]$ pixels for the first experiment and increased to $[-150, 150]$ for the second experiment. The original images were 1024×1024 , we cropped them to 512×512 around the center to ensure that non-overlapping areas are occupied by real data from the original images. The minimum overlap between the images was 50% and the average overlap was 75% for the entire set.

We compared our method with the conventional multi-resolution method in the pixel intensity domain. We used MI as the similarity measure and optimized the cost function using the Simplex method.

The registration errors were calculated by selecting 10,000 control points uniformly distributed across the image and computing the average distance between the trans-

formed control points under the calculated transformation and the actual transformation.

5. Discussion

5.1. Accuracy and Performance of the Registrations

The success rate of our method was 99% for the first experiment and 95% for the second experiment. Whereas the performance of the conventional multi-resolution method was 24% and 15%, respectively. The performance of our method decreases as we increase the translation dynamic range, this is due to the reduction in overlap between the images and deterioration of the rotation estimate. The poor performance of the conventional method is due to its inability to cope with the large capture range of the experiments. Our method performed very well for overlaps as low as 50%, for all groups of images. The landmark group continued to perform well below 50% overlap. Apparently, the combination of natural scenery and a few distinct artificial structures provides good conditions for registration with our method. Other groups showed only a slightly lower performance. We also tested our method with images taken with a handheld camera and without a tripod. The images taken in this way contain non-rigid deformations, however our method is able to find a close rigid approximation as shown in Fig. 5(e) and Fig. 5(f).

Both Hough-based and intensity-based methods registered images very accurately, whenever they were initialized within the capture range of their respective cost functions. The average registration error was 0.42 pixels for our method. This was expected, since MI-based cost functions are known to be accurate [4].

5.2. Capture Range of the Hough Method

Cost functions based on Hough gradient functions exhibit larger translational capture range compared to intensity-based cost functions. The capture range is typically more than twice, as shown in Fig. 7.

The larger capture range can be attributed to the spatially relaxed formulation of the gradient function Γ . The gradients can be freely displaced along the line that passes through them without affecting the gradient function. For example, in Fig. 2, placing g_1 anywhere on L_1 does not change Γ .

5.3. Similarity Registration

It is possible to use our Hough registration method for similarity registration. The translation and scale can be estimated using a method similar to that described in Section

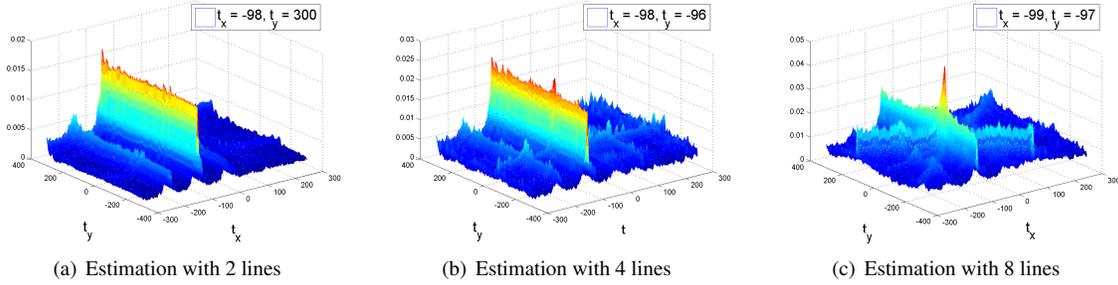


Figure 4. Robustness of estimating translation parameters improves by adding more lines. (a) With 2 lines, t_y is estimated incorrectly. (b) With 4 lines both parameters are correctly estimated but there are other high peaks. (c) With 8 lines, a single dominant peak is formed at the correct alignment.

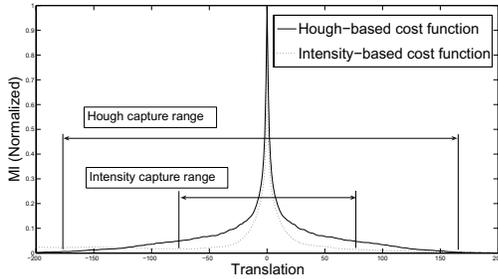


Figure 7. Hough method has a larger translational capture range, due to the spatially relaxed formulation of the cost function.

3.3 for estimation of the translation parameters, however this time an exhaustive 2D search needs to be performed. Alternatively, one can optimize for two translations and the scale parameter, once the rotation is resolved, using a local optimization algorithm such as Simplex:

$$\hat{t}_x, \hat{t}_y, \hat{s} = \underset{t_x, t_y, s}{\operatorname{argmin}} -MI(\Gamma'(\rho, \phi); \Gamma(\rho', \phi')), \quad (17)$$

$$\rho' = s\rho + t_y \cos(\phi + \hat{\gamma}) - t_x \sin(\phi + \hat{\gamma}), \quad (18)$$

$$\phi' = \phi + \hat{\gamma}. \quad (19)$$

We performed some preliminary experiments with both methods. In our experience, the second method which uses local optimization is less computationally expensive and performs better.

References

- [1] *Google Earth*. Google Inc., <http://earth.google.com/>, 2006.
- [2] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, Dec. 1992.
- [3] S. Chunhavitayatera, O. Chitsobhuk, and K. Tongprasert. Image registration using Hough transform and phase correlation. In *Int. Conf. on Advanced Communication Tech. (ICACT)*, pages 973–977, Feb. 2006.
- [4] A. Cole-Rhodes, K. Johnson, J. LeMoigne, and I. Zavorin. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Trans. on Image Processing*, 12(12):1495–1511, Dec. 2003.
- [5] D. A. Forsyth and J. Ponce. *Computer Vision, A Modern Approach*. Prentice Hall, Pearson Education, Inc., Upper Saddle River, New Jersey, 2003.
- [6] J. Gluckman. Gradient field distributions for the registration of images. In *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, volume 2, pages 691–694, Sept. 2003.
- [7] J. Liu, J. Tian, and Y. Dai. Multi-modal medical image registration based on adaptive combination of intensity and gradient field mutual information. In *IEEE Int. Conf. of Engineering in Medicine and Biology Society (EMBS)*, pages 1429–1432, Aug. 2006.
- [8] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 16(2):187–198, Apr. 1997.
- [9] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Med. Image Anal.*, 2(1):1–36, 1998.
- [10] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Trans. on Med. Imaging*, 19(8):809–814, Aug. 2000.
- [11] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, Cambridge, second edition, 1992.
- [12] R. Shams, R. A. Kennedy, P. Sadeghi, and R. Hartley. Gradient intensity-based registration of multi-modal images of the brain. In *Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, Rio de Janeiro, Brazil, Oct. 2007.
- [13] R. Shams, P. Sadeghi, and R. A. Kennedy. Gradient intensity: A new mutual information based registration method. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR) Workshop on Image Registration and Fusion*, Minneapolis, MN, June 2007.

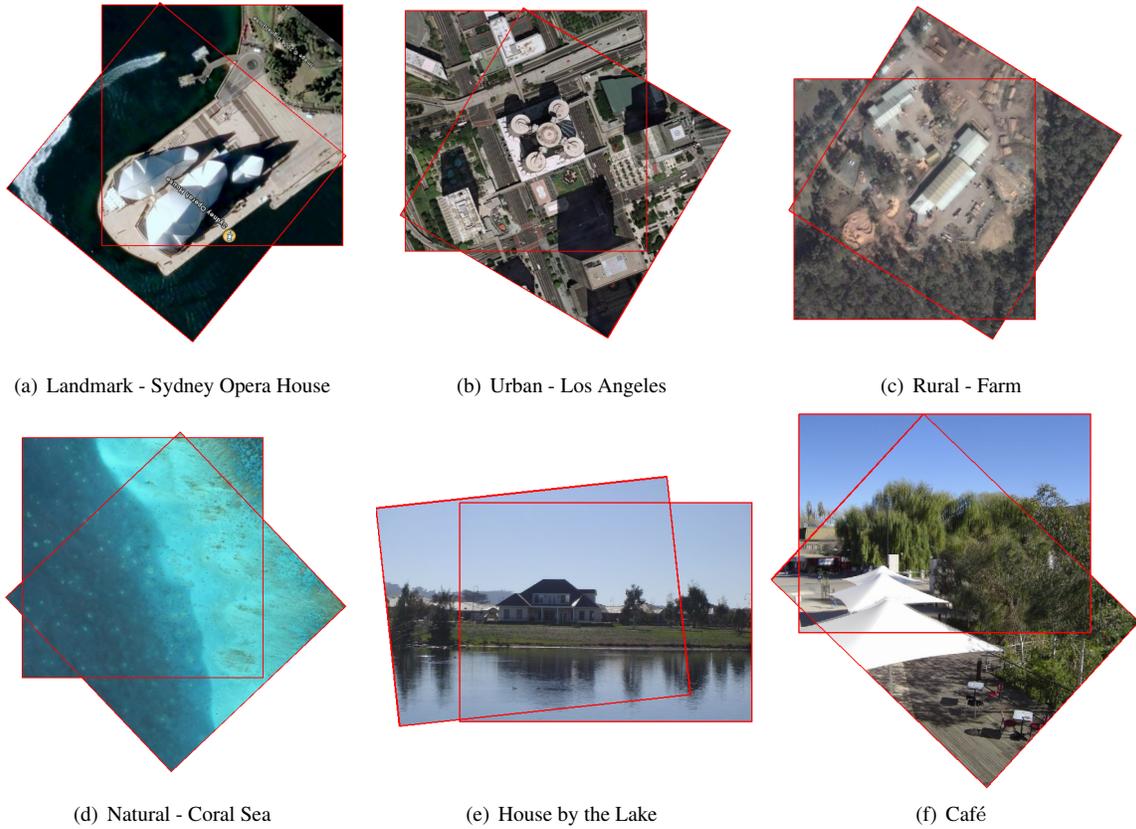


Figure 5. (a-d) Examples of registered images in each category. (e-f) Pictures taken manually with a handheld camera. There are some affine and perspective distortions between the images, however our method can find a close rigid approximation.

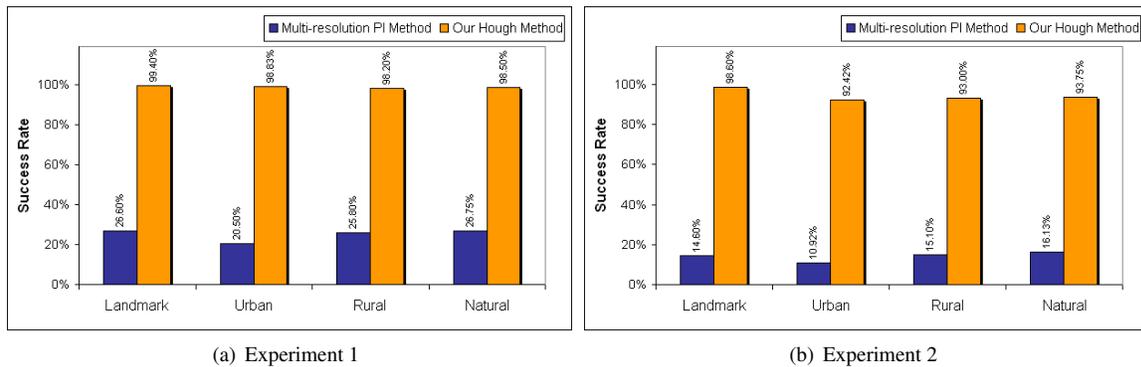


Figure 6. Superior performance of our registration method compared to the conventional multi-resolution approach.