Accurate Image Registration from Local Phase Information

Himanshu Arora, Anoop M. Namboodiri, and C.V. Jawahar
Center for Visual Information Technology, IIIT, Hyderabad, India
{ himanshu@research., anoop@, jawahar@}iiit.ac.in

Abstract

Accurate registration of images is essential for many computer vision algorithms for medical image analysis, super-resolution, and image mosaicing. Performance of traditional correspondence-based approaches is restricted by the reliability of the feature detector. Popular frequency domain approaches use the magnitude of global frequencies for registration, and are limited in the class of transformations that can be estimated. We propose the use of local phase information for accurate image registration as it is robust to noise and illumination conditions and the estimates are obtained at sub-pixel accuracy without any correspondence computation. We form an overdetermined system of equations from the phase differences to estimate the parameters of image registration. We demonstrate the effectiveness of the approach for affine transformation under Gaussian white noise and varying illumination conditions.

1. Introduction

Image Registration is a process of geometrically aligning two or more images obtained from different views [11]. In many computer vision and image processing applications, we need highly accurate image registration under differing noise and illumination conditions. Such applications include generation of super-resolution images from multiple images, where the output quality depends mostly on the registration accuracy. In large scale mosaicing, a small error in registration of two images can lead to large errors at later stage. In medical image analysis, registration accuracy is required to predict diseases based on the image comparison. Image understanding algorithms such as 3D reconstruction from videos also needs accurate registration.

Errors in image registration primarily result from the presence of noise and variations in illumination. Spatial domain techniques require robust estimation of certain primitives for correspondence. The registration accuracy is limited to the accuracy of the primitive detection algorithms and is also affected by noise and illumination changes. However, in the transform domain, there are methods that are robust to noise and illumination changes. Reddy and Chatterji [9] proposed an algorithm for image registration to recover the translation, rotation and scale parameters. Rotation and scale factors are estimated from the magnitude of the Fourier transform of the images in polar and logarithmic coordinates. Phase correlation is used for estimation of translation parameters. Fourier Mellin Transformation [5] could also be used for computing rotation, scale and translation parameters between image pairs. The FMT approach takes advantage of the fact that the changes in scale in spatial domain is equivalent to a phase change in the Mellin domain. Kumar et al. [8] has used the line patterns in texture to estimate up to affine transformation. Fourier based techniques for estimating projective and general image transformations are still missing.

In this paper, we make use of local phase computed from a finite impulse response (FIR) band pass filter to achieve registration with high accuracy. A phase-based approach has been chosen over the magnitude of the response of the filter because of its inherent stability [2]. Phase does not depend on the intensity levels of the image and hence the measurements are invariant to smooth shading and lighting variations. Explicit signal construction and feature detection are not required and yet the results are obtained at sub-pixel accuracy. Local phase information has been effectively used to solve similar problems such as stereo disparity computation [10] and optical flow [4] with high accuracy under noise conditions. The contributions of this work are summarized as:

- Accurate image registration algorithm with an order of magnitude better results compared to the conventional schemes.
- Correspondenceless approach with robustness to illumination variations and band-limited noise.
- Deviating from the popular use of magnitude, local phase is shown to be useful for accurate registration.
- A method for estimating local translation components and estimation of image registration parameters from these estimates. The results are shown for affine transformation, although it can be extended to any class of image transformations.

2. Computation of Local Phase Difference

Local phase information has been demonstrated to be useful in problems such as computation of stereo disparity [10] and optical flow field estimation [4]. To compute the depth of a scene point from an image pair, Sanger [10] uses the information that disparity is proportional to the difference in local phase. For optical flow field computation, Gautama and Hulle [4] tracked constant phase over time and showed it to be robust to Gaussian white noise. We extend the one-dimensional translation estimation in phase-based stereo [10] to two dimensions, and combine the translation components computed from phase to align the images iteratively.

Local Phase: Any bandpass filter, with a finite support, can be used for extracting the local phase in an image. Gabor filters [3] are commonly used as band pass filters as they achieve the theoretical minimum of product of spatial width and bandwidth for any complex valued linear filter. A smaller bandwidth allows accurate computation of local phase and smaller width is desirable for localization.

Mathematically, a Gabor filter is a multiplication of a complex harmonic function with a Gaussian envelope. In two dimensions, a normalized Gabor filter function [7] has an analytical form:

$$g(x, y) = \frac{f_1 f_2}{\pi\gamma\eta} e^{-((\frac{{f_1}^2}{\gamma^2})x'^2 + (\frac{{f_2}^2}{\eta^2})y'^2)} e^{j2\pi(f_1 x' + f_2 y')}, \tag{1}$$

where $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$, (f_1, f_2) is the central frequency of the filter, γ and η controls the spatial width and the bandwidth of the filter, θ is the orientation of the filter, and j is $\sqrt{-1}$.

We compute the local phase difference using the Gabor wavelet $g(x, y)$. Let $i_1(x, y)$ and $i_2(x, y)$ be two partially overlapping images. These images are convolved with $g(x, y)$ having central frequency (f_1, f_2), as:

$$s_m(x, y, f_1, f_2) = i_m(x, y) * g(x, y), \tag{2}$$

where $m \in \{1, 2\}$. The response has both real and imaginary components. The phase at location (x, y) corresponding to i_m is computed as:

$$\phi_m(x, y, f_1, f_2) = \text{arg}[s_m(x, y, f_1, f_2)], \tag{3}$$

where arg[ ] is the complex argument in $(-\pi, \pi]$. Therefore, the phase difference at the spatial location (x, y) as a function of (f_1, f_2) can be written as

$$\Delta\phi(x, y, f_1, f_2) = [\phi_2 - \phi_1]_{2\pi} \tag{4}$$

Confidence of the Phase Difference: Errors could be introduced in phase difference computation due to noise and the absence of the local frequencies with which the images are convolved. Sanger [10] has described the degree of match in the amplitude values as a confidence measure. The value of confidence is high if the amplitudes of the Gabor filter response at (x, y) in both the images are close. Let $|s_1|$ and $|s_2|$ be the amplitudes of the Gabor filter responses for the central frequency, (f_1, f_2). The confidence value is computed as:

$$r = min\left[\frac{|s_1|}{|s_2|}, \frac{|s_2|}{|s_1|}\right] \tag{5}$$

In addition, if the amplitude falls below a particular threshold, the confidence value is set to zero.

3. Registration from Local Phase

In our approach, the partially overlapping images are convolved with Gabor Filters to calculate the local phase differences. This is done at multiple frequencies so that if some of the local frequencies are absent they can be detected and removed from further consideration. We compute the local translational components at each spatial location from the phase difference estimated at least at two different frequencies. An overdetermined system of equations is formed from the local translation estimates, which is solved to estimate the image registration parameters, iteratively. For our algorithm, we define partially overlapping images as the image pair where in any small 2D window at location (x, y) the corresponding point lie within the cycle of the signal. This condition should hold true at most of the locations in the image for the algorithm to converge. In this Section we describe the computation of local translation parameters and suggest a mechanism for computing the registration parameters.

3.1. Local Translation Estimation

The translation between two 1D sinusoids can be accurately computed by measuring the phase difference at the same location of the sinusoid and then dividing it by the frequency of the signal (see Figure 1). The same concept

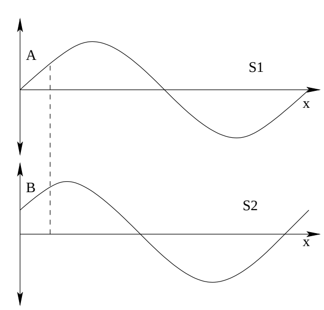


Figure 1. Computation of shift from two 1D signals as (*phase difference*/*frequency*) of the signal

is used to compute the translation components from two 2D sinusoidal signals. To estimate two parameters, we need the phase difference at least at two different frequencies. The computation of translation components can also be formulated on the basis of Fourier Shift theorem, according to which, a shift of $(\Delta x, \Delta y)$ in the spatial domain would produce a phase difference of $2\pi \times (\Delta x f_x + \Delta y f_y)$ at (f_1, f_2). i.e., if

$$i_2(x, y) = i_1(x + \Delta x, y + \Delta y), \tag{6}$$

then in Fourier domain at (f_1, f_2) the relationship is given by:

$$I_2(f_x, f_y) = I_1(f_x, f_y) e^{2\pi j (f_x \Delta x + f_y \Delta y)} \tag{7}$$

We compute the phase difference at multiple frequencies, (f_1, f_2), by using Gabor filters and form on over-determined system of equations in Δx and Δy. We choose only those estimates, where the confidence value is more than a given threshold. The usage of Gabor filters is based on the assumption that the phase output of the Gabor filter is linear as a function of spatial position [2]

3.2. Computing Registration Parameters

Given a set of point correspondences, the image transformation parameters can be estimated by solving an over-determined system of equations. Similarly, local translation components computed at various spatial locations can be thought as point correspondences with high accuracy. Given many accurate estimation of such pairs the image registration parameters can be computed accurately.

In this paper, we limit the class of registration algorithms to that of planar views related by affine transformation. We concentrate on affine transformation because most of the partial overlap can be approximated by affine transformations. Projective transformation is the most general form which relates the two views. It is represented by a 3 $\times$ 3 matrix M such that $x' = Mx$ where x and x' are the homogeneous coordinates of the two images. An affine transformation is a linear transformation in in-homogeneous coordinates followed by a translation and captures translation ,rotation, scaling and shearing in a plane. Similarity transformation is a subgroup of affine transformation, which captures only translation, rotation and scaling. Mathematically, under affine transformation two views of an object are related by

$$x = ax' + by' + c; \qquad y = dx' + ey' + f$$

At each location, we estimate the translation parameters, which is related to the correspondence of a point (x, y) in one image with (x', y') in the other. We form an overdetermined system of equations to solve for a, b, c, d, e and f.

3.3. Iterative Parameter Estimation

The local translation parameters, calculated at each spatial location, are approximately correct. This is because in a small window points need not be related by pure translation. Moreover, the two points need not lie within the cycle of the signal. However, over iterations, as the corresponding points will be closer, the effect due to these assumptions would be negligible. We iteratively update the transformation parameters till convergence.

Algorithm 1 Accurate Image Registration

Input: An image pair, i_1 and i_2.
Output: Parameter describing the geometric relationship between two images accurately.

1: Compute the approximate registration parameters.
2: **repeat**
3: Obtain the overlapping image pair using the current registration parameters.
4: Convolve both the images with a bank of Gabor filters and calculate the phase difference values.
5: Calculate the translation parameters at each location by solving for Δx and Δy from phase differences with sufficient confidence.
6: Form an over-determined system of equations using the translation estimates and solve it to update the registration parameters.
7: **until** convergence

We now explore the robustness of the phase information with respect to noise and illumination conditions.

3.4. Robustness of the Proposed Algorithm

Noise Tolerance: For band-limited noise, the error in the estimation is reduced by considering the phase output of those filters that do not allow the frequencies to pass through. This is done by assigning low scores to those phase difference estimates where the same signal is not present at the same location in the two images. A simple way of doing so is by comparing the amount of mismatch between the amplitude values.

Illumination Invariance: Illumination change in a small window at the same location in the two images can be modeled as the multiplication of the pixel values by a constant. The phase information computed at these two locations will remain unchanged as compared to the magnitude of the signal, which will be scaled by the illumination constant. This fact follows from equation 2 and 3.

4. Experimental Results

We perform experiments on synthetic and real images. On real images, we test and compare the performance of our algorithm under Gaussian white noise and varying illumination conditions with RANSAC [1] (RNSC), Fourier-Mellin Transform (FMT) [5] and an algorithm

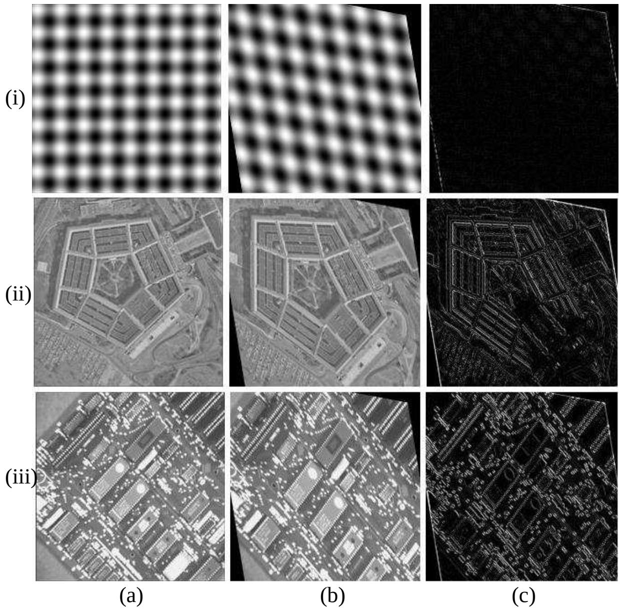


Figure 2. (a) and (b) are the images to be registered. (c) shows the absolute image difference after using our algorithm.

based on iterative minimization of sum of squared differences of the intensity values [6] using gradient descent (GD). RANSAC is robust in estimating the transformation parameters in the presence of outliers. Fourier-Mellin Transform is robust to noise and varying illumination conditions, though it can estimate only up to similarity transformation. In the absence of illumination variations in the images, an image registration algorithm based on minimization of the intensity value differences can register the images accurately.

To compare the performance of the algorithms, we generate the image pairs having known transformation parameters from synthetic and real images. We define the mean shift error, e, as the average distance between corresponding points after registration. Mathematically it can be expressed as:

$$e = \frac{1}{N} \sum_{(x,y)} d(M'[x\ y\ 1]^T, M[x\ y\ 1]^T), \qquad (8)$$

where $[x\ y\ 1]$ is the homogeneous representation of the image coordinate of i_1, M and M' are 3×3 transformation matrices representing the known and estimated image registration parameters, N is the total number of points and $d(,)$ is the Euclidean distance between two points. The mean shift error indicates how close the overlay is.

- **Sinusoidal Image:** First row of Figure 2 shows two synthetic 2D sinusoidal image pairs that are to be registered. We use four Gabor filters having central frequencies in the range 0.06 and 0.12. The algorithm converged in 4 iterations with a final mean shift error of 0.08 per pixel. Figure 2 (c)shows the absolute image difference after using our algorithm between the reference and the registered images.

- **Pentagon and Apple Chip Image:** The second and third rows of Figure 2 shows the pentagon and apple chip image pairs. Experimental steps are same as mentioned for sinusoidal image pairs. The algorithm converged in 7 steps and the mean of shift error after using our algorithm was 0.108.
  Gaussian White Noise: We add Gaussian white noise with zero mean and standard deviation varying from 0 to 6 in both the images. For image pairs related by an affine transformation, we compared our proposed approach with RANSAC and the gradient descent based image registration algorithm [6](GD). For images related by similarity transformation we compare the performance of our algorithm with Fourier Mellin Transformation (FMT). Table 1 summarizes the mean of shift error values computed.

	Affine			Similarity	
σ	Proposed	RANSAC	GD	Proposed	FMT
0	0.10	0.18	0.068	0.06	1.19
1	0.18	0.35	0.070	0.11	1.19
2	0.31	0.41	0.071	0.18	1.19
4	0.52	0.76	0.073	0.33	1.19
6	0.65	1.05	0.078	0.45	1.19

(a) Errors on Pentagon Image Pair

	Affine			Similarity	
σ	Proposed	RNSC	GD	Proposed	FMT
0	0.15	0.22	0.121	0.09	1.19
1	0.21	0.35	0.123	0.14	1.19
2	0.27	0.46	0.123	0.18	1.19
4	0.45	0.80	0.129	0.29	1.19
6	0.59	0.97	0.132	0.41	1.19

(b) Errors on Apple Chip Image Pair

Table 1. Comparison of the proposed scheme with other image registration algorithms under Gaussian white noise.

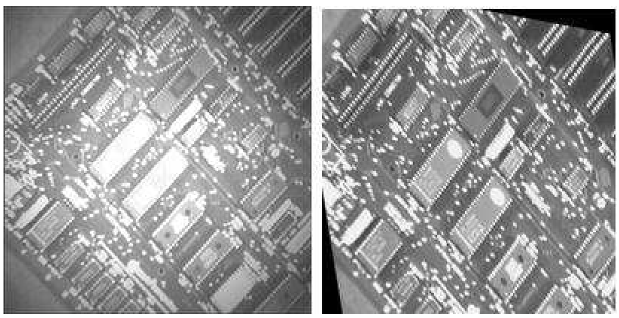
  Illumination Conditions: Varying illumination change in the whole image can be modeled by the multiplication of grey values by a constant, in a small window. Figure 3 shows the image pair with one of the images having smooth lighting in radially outward direction from the center. Table 2 summarizes the errors.

4.1. Discussions

For the images related by affine transformation, our algorithm performs better than RANSAC in presence of Gaussian white noise and varying illumination changes.

Image	Affine			Similarity	
	Prop.	RNSC	GD	Prop.	FMT
Pentagon	0.245	1.284	5.357	0.170	1.190
Apple	0.421	0.910	3.165	0.318	1.190

Table 2. Performance comparison under varying illumination conditions.



(a) (b)

Figure 3. Registration under smooth lighting conditions: (a) has radially decreasing illumination.

It performs better than GD in presence of illumination changes. However, in presence of Gaussian white noise, GD registered the images more accurately. Note that variations in pixel values at the center is more than any other location in the image under varying illumination conditions, while Gaussian white noise has zero mean and hence does not affect the minimum of mean squared error. For the images related by similarity transformation, our algorithm performs far better than FMT within the given noise limits and illumination conditions. FMT is very robust to noise and illumination conditions, but the accuracy is limited because of the detection of impulse response at non-integer locations and the factors involving coordinate transformations.

In short, we note that the proposed algorithm can handle differing noise and illumination changes simultaneously whereas existing approaches fails to perform well. Transform domain techniques are robust to these variations, but the class of image transformation that can be calculated are very limited. The proposed algorithm can estimate the local translation parameters as long as there are sufficient locations, where the corresponding points lie within a cycle of the signal. In real life scenarios this might mean that the corresponding pixel should be at the most 8-10 pixels apart. We can overcome this limitation by having a quick approximation of the registration parameters by using any existing image registration algorithm. The phase difference and translation parameters at different locations can be computed in parallel and hence the algorithm is well suited for parallel architectures.

5. Conclusions

We present a novel approach for accurate image registration by using the local phase information. Our approach does not use point correspondences and computes the registration parameters for partially overlapping images. It is robust to band-limited noise and illumination changes that are present in real world scenarios. Images can be registered accurately within a few iterations of the algorithm. Moreover, the proposed algorithm can run on parallel architectures. We have shown the results up to affine transformation, although the formulation can be extended to any class of image transformations. Experiments indicate that the registration can be achieved with sub-pixel accuracy, under noise and illumination changes.

References

[1] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981.

[2] D. J. Fleet and A. D. Jepson. Stability of phase information. *PAMI*, 15(12):1253–1268, December 1993.

[3] D. Gabor. Theory of communication. *J. IEE*, 93:429–457, 1946.

[4] T. Gautama and M. Van Hulle. A phase-based approach to the estimation of the optical flow field using spatial filtering. *IEEE Trans. Neural Networks*, 13(5):1127–1136, 2002.

[5] X. Guo, Z. Xu, Y. Lu, and Y. Pang. An application of fourier-mellin transform in image registration. In *CIT '05: Proceedings of the The Fifth International Conference on Computer and Information Technology*, pages 619–623, Washington, DC, USA, 2005. IEEE Computer Society.

[6] L. Ibanez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide*. Insight Software Consortium, 2005. http://www.itk.org.

[7] J.-K. Kämäräinen. *Feature Extraction Using Gabor Filters*. PhD thesis, Lappeenranta University of Technology, 2003.

[8] M. P. Kumar, S. Kuthirumunal, C. V. Jawahar, and P. J. Narayanan. Planar homography from fourier domain representation. In *Proceedings of SPCOM*, pages 560–564, Dec 2004.

[9] B. S. Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8):1266–1271, 1996.

[10] T. D. Sanger. Stereo disparity computation using Gabor filters. *Biological Cybernetics*, 59:405–418, 1988.

[11] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.