Improving Phase Correlation for Image Registration

Ruben Gonzalez

Institute for Integrated and Intelligent Systems
Griffith University, Gold Coast Campus, QLD, Australia, and
Queensland Research Laboratory, NICTA, Australia
R.Gonzalez@gu.edu.au

Abstract

Phase correlation is a well-known technique for image registration that is robust to noise and operates in constant time. Scale and rotation invariance can be achieved through means of a log-polar transformation. Unfortunately this method has been historically shown to be unable to handle large rotation or scaling factors, making it unsuitable for many image registration tasks. This paper presents a novel phase correlation based technique that is shown to outperform the current state of the art image registration methods in terms of being able to recover larger rotation and scaling factors and with reduced computational requirements.

Keywords: Image Registration; phase correlation, log-polar transform, Global Motion Estimation

1 Introduction

The task of aligning two or more images of the same scene taken from different viewpoints is known as image registration. This may involve correcting for different scaling factors or rotation angles, or the images may be shifted relative to each other or have some other projective distortion. Image registration is a basic step for a wide range of image processing and computer vision applications. Techniques proposed to address this problem fall into two broad categories; those that characterise the problem as one of feature matching or that of correlation.

Feature matching methods essentially consist of detecting features in images that can be matched with corresponding features in the other images from which a transformation model can be estimated. Feature matching methods typically resort to iterative algorithms to establish feature correspondences across images and reject outliers based on some measure of consensus. The alternatives for performing each of the steps involved in feature matching based registration have been noted by Brown [1] and Zitova et al. [2].

Instead of attempting to match a set of specific features across images, correlation methods approach the problem by attempting to match image textures (areas). This matching is typically performed using either cross-correlation in the spatial domain, or phase-correlation (PC) in the frequency domain. The main advantages of the correlation methods over feature matching approaches are constant processing time and insensitivity to noise. The main disadvantage of correlation-based image registration methods has been their inability to handle images exhibiting large rotation or scale factors.

Correlation based registration is often performed in the Fourier Domain using the convolution theorem as it is more efficient to perform multiplication in the Fourier Domain rather than cross-correlation in the spatial domain. While phase-correlation by itself only permits translation to be estimated via the Fourier Shift property, remapping images from Cartesian to log-polar coordinates as in the Fourier-Mellin Transform [3] permits rotation and scale parameters to be estimated from the resulting images.

Unfortunately, phase correlation has been shown to be unable to recover larger scale factors. It has even been suggested that "FFT based schemes are unable to handle real-world registration problems" [4]. The cause of this problem, border effects and rotation-induced aliasing can be reduced via windowing and filtering, but these have the side effect of removing the frequencies that phase correlation requires to perform well [5]. As such, the state of the art image registration approaches do not rely on phase correlation. One method using hierarchical, nonlinear, least squares optimization (HNLSO) [4] avoids the use of the FFT altogether. Yet while it is robust, it is also fairly slow. A more recent solution making use of normalized gradient correlation (NGC) [6] while considerably faster than HNLSO and providing comparable performance is slower than PC.

In this paper we present a novel phase correlation based approach to image registration that outperforms the state of the art methods and requires only a total of seven FFTs to recover the rotation, scale factors and translation parameters. This compares favourably with NGC that requires fourteen real and complex FFTs in total, eight for recovering the rotation and scale and another six FFTs for the translation.

2 Theory

The required theory for registering images that are translated, rotated and scaled with respect to each other via phase correlation is presented in this section.

2.1 Translation Estimation

Let $f_1(x,y)$ and $f_2(x,y)$ be two images where $[x, y] \in \Re^2$, that are displaced from each other by some arbitrary translation and attenuated by some given amount, that is f_1 and f_2 are related by the following transformation:

$$f_2(x,y) = a \cdot f_1(x + t_x, y + t_y) \quad (1)$$

According to the shift theorem, taking the Fourier transform of both results in the following relationship:

$$\mathcal{F}_2(u,v) = a\mathcal{F}_1(u,v) \cdot e^{-j(ut_x + vt_y)} \quad (2)$$

Since the magnitude of a complex exponential is simply the radius of a unit circle, the exponential can be factored out by calculating the normalised cross-power spectrum and substituting for $\mathcal{F}_2$.

$$\frac{a \cdot \mathcal{F}_1(u,v) \cdot \mathcal{F}_2^*(u,v)}{\left|a \cdot \mathcal{F}_1(u,v) \cdot \mathcal{F}_2^*(u,v)\right|} = e^{-j(ut_x + vt_y)} \quad (3)$$

The Fourier transform pair of the resulting complex exponential is a Dirac delta located at the offset of the first image relative to the second.

$$\mathcal{F}^{-1}\left\{ e^{-j(ut_x + vt_y)} \right\} = \delta(x + t_x, y + t_y) \quad (4)$$

2.2 Rotation and Scale Estimation

Let $f_1(x,y)$ and $f_2(x,y)$ be two images as before where f_2 is a translated, scaled and rotated replica of f_1 by some scale factor s, and rotation angle θ such that f_1 and f_2 are related by the following transformation:

$$f_2(x, y) = f_1(j, k) \quad (5)$$

Where:
$$\begin{aligned} j &= s \cdot x\cos\theta - s \cdot y\sin\theta - x_0 \\ k &= s \cdot x\sin\theta - s \cdot y\cos\theta - y_0 \end{aligned} \quad (6)$$

Taking the Fourier transform of both sides of (5) and using the shifting, scaling and rotation theorems:

$$\begin{aligned} \mathcal{F}_2(u, v) = \frac{1}{|s|} e^{-j(ux_o + vy_o)} \\ \times\ \mathcal{F}_1(u/s \cos\theta + v/s \sin\theta, \\ u/s \cos\theta - v/s \sin\theta) \end{aligned} \quad (7)$$

Taking the magnitude of both sides of (7) such that $G_i = |\mathcal{F}_i|$ and ignoring the weighting factor $1/|s|$ gives:

$$\begin{aligned} G_2(u, v) = G_1(u/s \cos\theta + v/s \sin\theta, \\ u/s \cos\theta - v/s \sin\theta) \end{aligned} \quad (8)$$

Denoting the Cartesian coordinates (u, v) on the RHS of (8) in terms of the polar coordinates (r, φ) by substituting for $u = r \cos\varphi$ and $v = r \sin\varphi$, gives:

$$\begin{aligned} G_2(u, v) = G_1(r/s \cos\theta \cos\varphi + r/s \sin\theta \sin\varphi, \\ r/s \sin\theta \cos\varphi + r/s \cos\theta \sin\varphi) \end{aligned} \quad (9)$$

Applying the product to sum trigonometric identities results in the following relationship:

$$G_2(u, v) = G_1(r/s \cos(\varphi + \theta), r/s \sin(\varphi + \theta)) \quad (10)$$

Mapping expression (10) into polar coordinates (r, φ) where $r = \sqrt{u^2 + v^2}$ and $\varphi = \arctan(v/u)$ results in:

$$G_2(r, \varphi) = G_1(r/s, \varphi + \theta) \quad (11)$$

Taking the logarithm of the first variable converts the division into a subtraction:

$$G_2(\log r, \varphi) = G_1(\log r - \log s, \varphi + \theta) \quad (12)$$

Substituting $\rho = \log r$ & $z = \log s$ into this equation:

$$G_2(\rho, \varphi) = G_1(\rho - z, \varphi + \theta) \quad (13)$$

The rotation and scaling difference between G_1 and G_2 can now be recovered in the polar coordinate system using phase correlation as if it were translation in Cartesian space using the standard phase correlation method outlined in section 2.1.

3 Improving Phase Correlation

Estimating the rotation and scaling parameters from two images requires the images to have the same centre of rotation and scaling, as both are sensitive to the local origin around which they take place. This is achieved by taking the magnitude of the Fourier transform of both images to obtain a translation invariant representation as the first step. To reduce edge effects, a Hamming window is typically applied to the source images prior to the Fourier Transform.

The resulting images are then remapped into the log-polar domain, such that rotation becomes horizontal shifts and scaling becomes vertical shifts. Following the work reported in [6] the conversion from Cartesian to Log-Polar form is performed using bilinear interpolation by resampling the magnitude of the N x N FFTs on an N/2 x N/2 log-polar grid using the following mapping of a source image $f(x,y)$ in the Cartesian space into the log-polar domain $g(\rho, \varphi)$.

$$g(\rho,\varphi) = f\left(N/2 + \beta^y \cdot \cos(\alpha), N/2 + \beta^y \cdot \sin(\alpha)\right) \quad (14)$$

Where the variables α and β are given by:

$$\alpha = \frac{\pi x}{N} \quad \& \quad \beta = e^{\log(N/2)/N/2} \quad (15)$$

Finally the inverse Fourier transform of the normalised cross-power spectrum is calculated to obtain the location of the Dirac delta:

$$PC(x,y) = \mathcal{F}^{-1}\left\{ \frac{G_1(\rho,\varphi) \cdot G_2^*(\rho,\varphi)}{\left|G_1(\rho,\varphi) \cdot G_2^*(\rho,\varphi)\right|} \right\} \quad (16)$$

The location of the peak indicates the rotation and scaling relation between the source images. The rotation angle θ and the scale S are recovered as:

$$\theta = \frac{\pi \varphi}{N} \quad \& \quad S = N^{\log_N \rho} \tag{17}$$

The rotation angle obtained is subject to a 180° ambiguity. This can be resolved by taking the magnitude of the FFT of the first source image after it has been scaled and rotated by (S, θ) and then also by (S, 180+θ). The resulting image that best matches the magnitude of the FFT transform of the second source image has the correct rotation angle. After obtaining the rotation and scale factors, standard phase correlation is applied to the transformed replica of the first image and the second image to resolve the translation component.

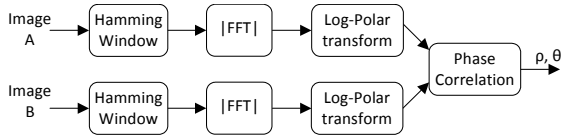

Fig. 1 Standard PC for Scale and Rotation Recovery

While this standard formulation (Fig. 1) works well for low scaling factors of up to about 1.8 [3] it fails at higher scaling factors as has been thoroughly documented [4, 5, 6]. In contrast, non-PC state of the art registration methods such as NGC [6] can readily recover scale factors of up to 6.0.

One reason for the shortcoming is because unlike the more uniformly random energy distribution of the original images, the energy distribution in log-polar transformed images tends to approximate a Dirichlet kernel centred about zero. The resulting DC term and neighbouring low frequency components create a secondary peak in the phase correlation result that competes with the scale and rotation-induced peak. This is not a problem with high-resolution images but becomes more serious at lower resolutions. This competing peak can be eliminated by using a narrow stop band high pass filter as defined by (18) to set the relevant low frequency components to zero.

$$HPF(x,y) = 1.0 - rect\left(x\,{}^{N}\!/_{64}, y\,{}^{N}\!/_{64}\right) \tag{18}$$

Even after removing the DC term, most of the energy remains in the low to mid frequency components, resulting in a Gaussian like distribution that is largely unaffected by rotation and scaling. This gives rise to common phase components that compete with the rotation and scale induced phase components to reduce the effective signal to noise ratio. This problem cannot be resolved by widening the stop band of the high pass filter since that would attenuate not only the unwanted common components but also the discriminating phase components. Instead, this problem is best mitigated as shown in Fig. 2 through the means of spectral whitening prior to the phase correlation such as via the application of a logarithmic function or by taking the square root of the spectral image. In the case of a logarithmic function such as $H_i(\rho, \varphi) = \log [G_i(\rho, \varphi)]$, then:

$$PC(x,y) = \mathcal{F}^{-1}\left\{\frac{H_1(\rho,\varphi)\cdot H_2^*(\rho,\varphi)}{\left|H_1(\rho,\varphi)\cdot H_2^*(\rho,\varphi)\right|}\right\} \tag{19}$$

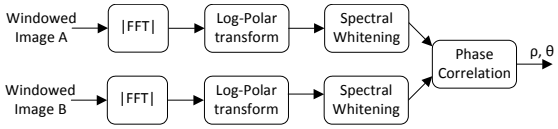

Fig. 2 Improved PC for Scale and Rotation Recovery

4 Experimental Evaluation

The performance of the PC based image registration with the proposed improvements was extensively evaluated using three different datasets. Two were photographically captured and the third dataset consisted of geometrically transformed replicas of three different 512x512 images (Lena, goldhill and washsat). Each of the three test images in the third dataset was replicated a total of 9000 times using a different combination of translation, scaling and rotation factors. This resulted in 27,000 different image pairs where the scale ranged from 0.0625 to 7.25 in 35 logarithmically uniform steps. The rotation ranged from ± 180 degrees in 15-degree steps and ten different translation vectors were used based on the scale factors.

The experiments were carried out to determine three characteristics of each method being evaluated. These were the (i) the average percent scale factor error, the (ii) average angular error in degrees and the (iii) maximum range of scale factors that could be reliably recovered across the entire dataset for each method. For the purpose of these experiments a given scale factor was deemed to be reliably detected if the method being assessed gave an average error of 25% or less across all rotation angles and translation vectors. The value of 25% was chosen rather arbitrarily since error distribution tends to be a step function: where the scale factors can be recovered the error is typically below 2% and where they cannot the average error is around 70%.

4.1 Generated Dataset Results

The first experiment using the generated dataset evaluated the performance of standard phase correlation as shown in fig. 1 using a preliminary Hamming window prior to the initial Fourier transform, without suppressing the DC term and related frequencies. Of the 27,000 image pairs in this dataset it was found that only those with scale factors from 0.69 to 1.31 were reliably resolved. The average scaling error, across the entire dataset was 49% and the average rotation error was 40.5 degrees.

The effectiveness of zeroing the DC term using a high pass filter (18) was then evaluated. This almost doubled the performance since scaling factors from 0.3125 to 2.56 were reliably recovered. The average scaling and rotation errors for the entire dataset were also reduced to 32% and 26.6 degrees respectively.

The effect of spectral whitening was then investigated in place of the high pass filter through the use of a log function (19). This further improved the recovery of scale factors to those within the range of 0.25 to 3.18. It also reduced the average scaling error over all possible image pairs down to 29% and the average rotation error down to 19.0 degrees.

Inspection of the resulting phase correlation image at this point revealed that at higher scale factors the amplitude of the correlation peak is reduced relative to that at lower scale factors causing it to be swamped by mid to low frequency noise. To combat this a modified form of unsharp masking was used to emphasize the high frequency components. This entailed subtracting a smoothed version of the phase correlation image from itself by multiplying it with the following filter:

$$HFE(x,y) = 1 - 0.85 * e^{-4(2y-N)^2/N^2} \quad (20)$$

The application of this filter (fig. 3) significantly increased the scale factors reliably detected to those within the range of 0.187 to 5.06. The average scale error over all test cases was further reduced to 13% and the average rotation error down to 9.3 degrees. This is an encouraging result given that it is across a range of scaling factors from 16:1 to 1:7.25.

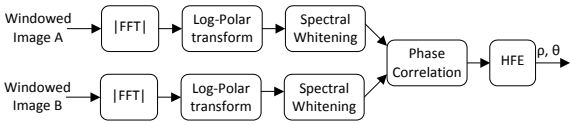

Fig. 3 PC with High Frequency Emphasis Filter

Table 1 outlines the average scaling and angular errors for each of the four phase correlation variants evaluated. It shows the 400% improvement in performance that was achieved over the base case using spectral whitening in conjunction with the modified unsharp masking.

Table 1 Average Error For All Methods

	Ave. Scaling Error (%)	Ave. Angular Error (degrees)
Standard PC	48.8%	40.5°
PC with HPF	32.2%	26.6°
Log-PC	29.0%	19.0°
Log-PC + HEF	13.3%	9.3°

The sensitivity of the proposed method to the image resolution was next evaluated. The same 27,000 pairs of images were subsampled at reduced resolutions of 256x256 and 128x128. The results in table 2 and fig. 3 confirm that in common with other Fourier based schemes, the performance of the method reduces with decreasing resolution. It was also found that at resolutions of 128x128 and below, recovery of the rotation and scale decreased as the amount of translation exhibited between each image pair increased. This did not occur at higher resolutions.

Table 2 Resolution Dependence of Improved Method

	128x128	256x256	512x512
Scale range	0.3-2.25	0.24 – 4.1	0.18 – 5.1
Scale error	55.3%	20.9%	13.3%
Angular error	53.4°	21.3°	9.3°

Fig. 4 The average scale error (bars) and angular error (line) obtained at different resolutions.

4.2 Comparative Assessment

The start of the art in image registration is arguably the NGC method that boasts recovery of scale factors up to 6.0. The NGC method has been extensively compared to the well known SIFT method as well as other recent methods such as those described in [4] and [7]. Accordingly the NGC results can be used as a benchmark to assess the performance of the improved phase correlation relative to not just NGC but also to the techniques that it has been assessed against.

To perform an accurate comparative assessment against NGC and the proposed method used the same software routines to perform functions that were common to both. These included FFT related routines, the log-polar transform and other utility functions. While the implementation of the NGC used here followed the description given in [6], it is likely that the two implementations differ in some small details. It is likely that any differing numerical results for the NGC method reported from other published ones are due to these differences.

The INRIA dataset [8] and [9] was used to assess the implementation of NGC and to benchmark the performance of the improved phase correlation method. The results of this assessment shown in Table 3 displays the maximum scale factors recovered by each method. The RHS column contains the experimentally obtained results for NGC, with previously reported results in parentheses. As the rotation angles recovered were essentially the same

for both methods they have been omitted. Scale is the limiting factor in the performance of both methods.

Allowing for rounding errors the experimental results obtained for NGC coincide fairly well with the published ones. The maximum reported scale value reported for NGC could not be replicated in five of the fifteen cases (Belledone, Crolle, Resid, Inria model). There appears to be some error in the reported values for the INRIA series as the ground truth was given as 4.03 and the reported maximum scale recovered was 3.91 whereas the experimentally recovered maximum scale using NGC was 5.88. Additionally, while NGC was able to recover the scale factor of 4.35 for the Enismag series of images it was not able to recover lower scale factors of 3.57 or 2.94. Similarly for the Boat series NGC was able to recover a scale factor of 4.35 but not the scale of 4.01.

Table 3 Scale Factors Recovered for INRIA Dataset

Name	**Proposed**	**NGC**
Asterix	5.78	5.78 (5.78)
Belledone	3.22	3.23 (5.57)
BIP	3.73	3.73 (3.73)
Boat	4.35	4.35 (4.26)
Crolle	4.00	2.70 (3.97)
East Park	5.88	5.88 (5.78)
East South	5.00	5.26 (5.18)
Enismag	4.55	4.35 (4.76)
INRIA	5.88	5.88 *(4.03)*
INRIA model	4.00	4.00 (4.82)
Laptop1	6.25	6.25 (6.22)
Laptop2	1.48	1.47 (1.51)
Resid	5.88	4.76 (5.85)
Van Gogh	**4.97**	3.36 (3.38)
UBC	2.89	2.89 (2.89)

In comparison the improved phase correlation was able to match the experimental results of NGC in all but one case (East South), and outperformed it in four cases (Crolle, Ensimag, Resid and Van Gogh). In the case of the Van Gogh series of images it outperformed the NGC results by 50%.

The limiting factor in the recovery of scale factors with the majority of the images in the INRIA dataset was the limited range of scale factors available within the dataset itself. Additionally, the scaling factor interval was quite large in many of the series of images, permitting only a course grain evaluation. Accordingly, a second dataset was used containing a much wider set of scale factors. This also ensured that the particular sensitivities of each registration method to the specific characteristics of any one dataset did not skew the performance results.

The second dataset consisting of 240 photographs [10] was captured with a 6.0 megapixel, 12x optical zoom Minolta DiMAGE Z6 digital camera. Each series of images in this dataset contained a range of rotation angles and scale factors above 10:1. Each series consisted on average of 25 individual images, giving an average scaling increment of about 10% for each image in the series. The images were subsampled to 1024 x 768. The registration results with this dataset using NGC and improved phase correlation are given in **Table 4**. The maximum ground truth scale factor for each image series is given in parenthesis following the image name.

Some examples of the performance of the improved phase correlation method using this dataset are shown in figs. 5-7. The enlarged image at the bottom of each figure shows the result of registering the pair of images at the top. This registered image shows the reference image (top left), rotated, scaled, translated and superimposed over its pair, according to the recovered transformation parameters.

Using this dataset, the improved phase correlation method outperformed NGC in 80% of tests by a good margin. The performance of NGC was equalled in the remaining 20% of tests. While scaling is not the limiting factor in this dataset neither method was able to recover the full range of available scale factors.

Table 4 Scale Factors Recovered for Second Dataset

Name (max scale)	**Proposed**	**NGC**
Painting (11.1)	**7.14**	3.57
Tractor (11.1)	**7.14**	6.25
Crane (11.4)	6.22	6.22
HiRise (11.2)	**5.18**	3.47
TwoLights (11.1)	**3.03**	2.49
Plant (10.8)	**7.20**	4.21
Steps (10.9)	**5.92**	4.59
Tower (16.0)	**4.16**	3.18
Uni (10.6)	**6.22**	5.31
Vending (10.8)	6.07	6.07

Fig. 5 Registered tractor image at $S = 7.14$ & $\theta = 12°$

Fig. 6 Registered 'Plant' Image at $S = 7.2$ & $\theta = 51°$.

Fig. 7 Registered 'Crane' Image at $S = 6.2$ & $\theta = 70°$.

5 Conclusions

While the theoretical advantages of phase correlation are well recognised, it has historically been unable to handle large rotation and scaling factors. This has led researchers in the past to look for better performing alternatives such as HNLSO and NGC. This paper has presented simple improvements to phase correlation that enable it to outperform the state of the art image registration methods by a significant margin. The proposed improvements permit it to recover a wide range of rotation angles and scaling factors up to 7.0 without a significant computational increase. All the traditional benefits of standard phase correlation are also retained including constant processing time and robustness to noise.

6 References

[1] L.G. Brown. "A survey of image registration techniques." *ACM Computing Surveys*, vol.24, no.4 pp.325-376, Dec. 1992.

[2] B. Zitová, J. Flusser: "Image registration methods: a survey." *Image Vision Computing*, vol. 21, no.11, pp.977-1000, 2003.

[3] B.S. Reddy, B.N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration", *IEEE Transactions on Image Processing*, vol.5, no. 8, pp.1266–1271, 1996.

[4] S. Zokai, G. Wolberg, "Image Registration Using Log-Polar Mappings for Recovery of Large-Scale Similarity and Projective Transformations." *IEEE Transactions on Image Processing,* vol.14, no.10, pp.1422-1434, 2005.

[5] H.S. Stone, B. Tao, M. Mcguire, "Analysis Of Image Registration Noise Due To Rotationally Dependent Aliasing" *Journal of Visual Communication and Image Representation* - JVCIR , vol. 14, no. 2, pp. 114-135, 2003

[6] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, T. Stathaki, "Robust FFT-Based Scale-Invariant Image Registration with Image Gradients," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.32, no.10, pp. 1899-1906, Oct. 2010

[7] W. Pan, K. Qin, and Y. Chen, "An Adaptable-Multilayer Fractional Fourier Transform Approach for Image Registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 3, pp. 400-413, Mar. 2009.

[8] Image Database, 2010, http://www.robots.ox.ac.uk/vgg/research/affine

[9] Image Database, 2010, http://lear.inrialpes.fr/people/mikolajczyk/

[10] Image Database, 2011 http://www.griffith.edu.au/professional-page/ruben-gonzalez