Fourier Based Recovery of Anisotropic Scaling Parameters in Images

R. Gonzalez
NICTA and Griffith University,
Gold Coast Campus, QLD, Australia,

Abstract— **This paper presents a novel Fourier based approach for recovering the transformation parameters of images that have been translated, globally scaled, rotated and stretched along an arbitrary vector. The proposed method first recovers the global rotation and scaling parameters. Then Hough based techniques are used for recovering the magnitude and orientation of the stretch vector. Theory, methodology and preliminary experimental results are presented.**

Keywords—image registration, affine, linear transform, fourier

I. INTRODUCTION

Fourier-based techniques often provide an efficient, non-iterative and robust solution to the problem of estimating the transformation parameters of images that have been subject to similarity transformations. In combination with a log-polar transform, Phase Correlation can recover a wide range of global rotation, scale and translation parameters [1]. While the global nature of the Fourier transform makes these methods relatively insensitive to localised noise and aberrations it renders them unsuitable for recovery of elastic deformation parameters. While similarity transforms preserve the lengths and angles within an image, affine transformations only preserve the parallelism of lines, making recovery of their transformation parameters problematic for Fourier based methods. This paper extends to application of Fourier based methods to the recovery of linear transformation parameters. Linear transforms are a superset of similarity transforms that include differential (or anisotropic) scaling and shearing. There is little published work regarding the application of Fourier based techniques to the estimation linear or affine transform parameters.

Previous approaches for recovering of linear or affine transformation parameters avoid the use of frequency domain techniques all together as being unworkable in practice [2]. Although used for a different application, Fourier methods were historically used for estimating the spatial orientation of uniformly textured surface planes. The respective affine parameters were by obtained comparing the frequency components, peaks or the second order moments of Fourier transforms of two patches on the textured surface [3, 4]. The only other Fourier based method for estimating affine parameters between images in the literature approaches it as a nonlinear minimization problem using radial projections of the squared Fourier magnitudes of the images [5]. In contrast this paper presents a simple closed-form extension of Phase Correlation for the estimation of linear transform parameters. Section two presents the theoretical concepts; section three discusses methodology, with experimental results presented in section four.

II. THEORY

Let $f_1(x,y)$ and $f_2(x,y)$ be two images where $[x,y]\in\Re^2$ and let $f_2(x,y)$ be a version of $f_1(x,y)$ that is stretched by a factor of z along an arbitrary orientation denoted by φ then uniformly scaled by a factor s, rotated by angle θ and translated by vector T(x_0, y_0), such that $f_1(x,y)$ is related to $f_2(x,y)$ by the following:

$$f_2(x, y) = R(\varphi)\, S(\mathrm{z},1)\, R(-\varphi)\, S(\mathrm{s},\mathrm{s})\, R(\theta)\, f_1(x, y) + T \tag{1}$$

Where: $$S(\mu,\eta)=\begin{bmatrix} \mu & 0 \\ 0 & \eta \end{bmatrix} \qquad R(\alpha)=\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \tag{2}$$

Expanding (1) gives $f_2(x, y) = f_1(j, k)$ where:

$$\begin{aligned} j &= s(z+1)\cdot\left(x\cos(\theta)-y\sin(\theta)\right)+s(z-1)\cdot\left(x\cos(\theta-2\varphi)-y\sin(\theta-2\varphi)\right)+x_0 \\ k &= s(z+1)\cdot\left(x\sin(\theta)+y\cos(\theta)\right)-s(z-1)\cdot\left(x\sin(\theta-2\varphi)+y\cos(\theta-2\varphi)\right)+y_0 \end{aligned} \tag{3}$$

Taking the magnitude of the Fourier transform of both sides of (1) such that $G_i = | F_i |$ and using the linearity, shifting, scaling and rotation theorems gives:

$$G_2(u, v) = \frac{1}{|s|}F\begin{bmatrix} \frac{(z+1)}{s}\left(u\cos\theta-v\sin\theta\right)+\frac{(z-1)}{s}\left(u\cos\left(\theta-2\varphi\right)-v\sin\left(\theta-2\varphi\right)\right), \\ \frac{(z+1)}{s}\left(u\sin\theta+v\cos\theta\right)-\frac{(z-1)}{s}\left(u\sin\left(\theta-2\varphi\right)+v\cos\left(\theta-2\varphi\right)\right) \end{bmatrix} \tag{4}$$

Denoting the Cartesian coordinates (u, v) on the RHS of (4) in terms of the polar coordinates (r, ϕ) by substituting for $u = r\cos\phi$ and $v = r\sin\phi$, and ignoring the $1/|s|$ scaling factor gives:

$$G_2(u,v) = F\begin{bmatrix} \frac{(z+1)}{s}\left(r\cos\phi\cos\theta - r\sin\phi\sin\theta\right) \\ +\frac{(z-1)}{s}\left(r\cos\phi\cos(\theta-2\varphi) - r\sin\phi\sin(\theta-2\varphi)\right), \\ \frac{(z+1)}{s}\left(r\cos\phi\sin\theta + r\sin\phi\cos\theta\right) \\ -\frac{(z-1)}{s}\left(r\cos\phi\sin(\theta-2\varphi) + r\sin\phi\cos(\theta-2\varphi)\right) \end{bmatrix} \tag{5}$$

Applying the product to sum trigonometric identities to (5) results in:

$$G_2(u,v) = F\begin{bmatrix} \frac{(z+1)}{s}\left(r\cos(\phi+\theta)\right) + \frac{(z-1)}{s}\left(r\cos(\phi+\theta-2\varphi)\right), \\ \frac{(z+1)}{s}\left(r\sin(\phi+\theta)\right) - \frac{(z-1)}{s}\left(r\sin(\phi+\theta-2\varphi)\right) \end{bmatrix} \tag{6}$$

Mapping expression (6) into the polar plane (r, ϕ) where $\boldsymbol{r} = \sqrt{u^2 + v^2}$, and $\phi = \arctan(v/u)$ results in:

$$G_2(r,\phi) = G_1\left(r \cdot \frac{\sqrt{2\beta}}{s},\ \phi + \tau\right) \tag{7}$$

where $$\beta = \left(z^2+1\right) + \left(z^2-1\right)\cos\left(2\left[\theta+\phi-\varphi\right]\right) \tag{8}$$

and $$\tau = \arctan\left(\frac{\left(\frac{z+1}{z-1}\right)\sin(\theta+\phi) - \sin(\theta+\phi-2\varphi)}{\left(\frac{z+1}{z-1}\right)\cos(\theta+\phi) + \cos(\theta+\phi-2\varphi)}\right) - \phi \tag{9}$$

Taking the logarithm of the first variable in expression (7) converts the multiplications into additions to give:

$$G_2\left(\ln[r],\phi\right) = G_1\left(\ln[r] - \ln[s] + \tfrac{1}{2}\ln[2\beta],\ \phi + \tau\right) \tag{10}$$

Now let $\beta = H * \psi$ so that:

$$H = \left(z^2-1\right) \quad \& \quad \psi = \cos\left(2\left[\theta+\phi-\varphi\right]\right) + \frac{z^2+1}{z^2-1} \tag{11}$$

Now (10) can be rewritten as:

$$G_1\left(\ln[r] - \ln[s] - \ln[4] + \tfrac{1}{2}\ln[2H] + \tfrac{1}{2}\ln[2\psi],\ \phi+\tau\right) \tag{12}$$

Now let $\rho = \ln[r]$, $k = \ln[s] + ½ \ln[2H]$ $b = ½ \ln[2\psi]$ and relabeling these relationships as $h_2(\omega, \phi) = G_2(\rho, \phi)$ and $h_1(\omega, \phi) = G_1(\rho+k+b, \phi+\tau)$ gives the Fourier transform of h_2 as

$$\begin{aligned} \mathcal{F}\left\{h_2(\omega,\phi)\right\} &= H_2(\varepsilon,\alpha) \\ &= \frac{1}{\sqrt{2\pi}} \iint G_1(\omega,\phi) \cdot e^{-i2\pi\left(\varepsilon(\omega+k)+\alpha(\theta+\tau)\right)} \\ &\quad \times e^{-i2\pi\cdot\varepsilon\cdot b}\, d\omega d\phi \end{aligned} \tag{13}$$

Which can be simplified to:

$$H_2(\varepsilon,\alpha) = H_1(\varepsilon,\alpha) \cdot e^{-i(\varepsilon k+\alpha\tau)} \int e^{-i2\pi\cdot\varepsilon\cdot b}\, d\phi \tag{14}$$

Substituting for b into the remaining integral in (14) results in it having the approximate form of an n^{th} order Bessel function of the first kind denoted as $J_n(m)$. Accordingly (14) can be rewritten as:

$$H_2(\varepsilon,\alpha) = H_1(\varepsilon,\alpha) \cdot e^{-i(\varepsilon k+\alpha\tau)} \cdot J_n(m) \tag{15}$$

Taking the normalized cross power spectrum of (14) factors out the phase difference since the magnitude of a complex exponential is simply the radius of a unit circle:

$$\frac{H_1(u,v) \cdot H_2^*(u,v)}{\left|H_1(u,v) \cdot H_2^*(u,v)\right|} = e^{-i(\varepsilon k+\alpha\tau)} \cdot J_n(m) \tag{16}$$

Finally, the Fourier transform of a Bessel function of the first kind is a ring delta function. Convolving it with a Dirac delta at (k, τ) shifts the ring's origin, so taking the inverse Fourier transform of (16) gives:

$$\mathcal{F}^{-1}\left\{e^{-i(\varepsilon k+\alpha\tau)} \cdot J_n(m)\right\} = \delta(\varepsilon+k, \alpha+\tau) * Ring(m) \tag{17}$$

Now since ε and α are respectively the transformed logarithm of the global scale factor s and the rotation angle θ we can directly estimate the global scale and rotation from the ring delta's location on the correlation surface.

The magnitude of the stretch z can also be calculated from the radius of ring delta, as it is a function of z. Unfortunately, the orientation of the stretch φ, given by the phase of the ring cannot be isolated from this formulation. Instead, note that if β is a constant and letting $k = \ln[\beta/s]$, the integral in (13) can be eliminated leaving only the Dirac delta in (17). This is the case if either $z = 1$ or if the Fourier integral in (13) is only evaluated for a fixed value of ϕ, in which case it reduces to a 1D transform only in terms of ω. As the offset of the Dirac delta will be dependent on $\cos(\phi+\theta-\varphi)$, stepping through different values of ϕ from 0 to 2π the offset of the Dirac Delta will trace out a sinusoid. Since θ can be obtained from (17) and ϕ is predetermined, only φ remains to be calculated from the phase of the sinusoid.

III. METHOD DESCRIPTION

The recovery of the global scale and rotation parameters between a reference image and its transformed pair begins with obtaining a translation invariant representation of the images by making use of the Fourier shift theorem. The result is then remapped into a log-polar plane that converts rotation and global scaling into shifts. The conversion of a image $f(x,y)$ from the Cartesian space into the log-polar domain $g(\rho, \phi)$ 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 relationship:

$$g(\rho,\phi) = f\left(N/2 + \beta^{y}\cdot\cos(\alpha), N/2 + \beta^{y}\cdot\sin(\alpha)\right) \quad (18)$$

Where the variables $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are given by:

$$\alpha = \frac{\pi x}{N} \quad \& \quad \beta = e^{\log(N/2)/N/2} \quad (19)$$

If the stretch magnitude z is not unity, calculating the phase correlation of the remapped images as described in [6] will result in a ring delta located somewhere on the correlation surface. Locating the ring delta on the correlation surface can be performed via a modified Hough Circle transform (HCT).

The first stage in the ring detection is to reduce the noise in correlation surface image via local noise averaging. Each point in the correlation surface is renormalized according to the mean local noise as per equation (20).

$$m(x,y) = \frac{1}{4r^2}\sum_{i=-r}^{+r}\sum_{j=-r}^{+r} f(x+i, y+j) \qquad f'(x,y) = \left|\frac{f(x,y)}{m(x,y)}\right| \quad (20)$$

Due to the overall low signal to noise ratio the performance of the standard HCT [7] is sub optimal. Using the HCT to obtain a shortlist of the most likely ring locations and analyzing their statistics can achieve better detection accuracy. In this case the highest product of the ring's size and squared mean of the perimeter is then taken as the ring location [8]. Using polar coordinates (r, θ) the circumference of a circle centered at (ρ, ϕ) with radius r is given by:

$$K(\theta) = \rho\cos(\theta-\phi) + \sqrt{r^2 - \rho^2\sin^2(\theta-\phi)} \quad (21)$$

The ring's mean, where η is the number of points of support in the ring can be calculated as:

$$\mu = \frac{1}{\eta}\oint_K f(\theta)\cdot d\theta \quad (22)$$

The centroid of the ring (k, τ) gives the rotation and average scaling between the two images. The average scale factor S can be recovered as:

$$S = N^{\frac{1}{N}k} \quad (23)$$

The magnitude of any stretch can be recovered (where N is half the number of points in the final Fourier transform): from the ring's radius R, in a similar manner:

$$Z = N^{\frac{1}{N}(R+1)} \quad (24)$$

Given the angular component τ of the coordinates of the centre of the ring delta and assuming the stretch ratio from (21) is within the range 1<= z <=2, the rotation angle θ can be recovered as the first order approximation of (9):

$$\tau = \arctan\left(\tan(\theta+\phi)\right) - \phi \quad (25)$$

and

$$\Theta = \frac{\pi\cdot\tau}{N} \quad (26)$$

To recover the orientation of the stretching one can either first inverse transform the target image to eliminate it's rotation and have $\theta = 0$, otherwise the θ will need to be subtracted from the phase estimate when it is obtained.

The process of recovering the phase begins with the log-polar transformed images. Rather than calculate the normalized cross power spectrum for each ϕ which implies each image column, we segment the image into bands sufficiently narrow as to safely assume that the variation in ϕ is negligible within each band. While wider bands improve the resulting signal to noise ratio, they also reduce angular resolution. The normalized cross power spectrum is calculated for each pair of corresponding bands in the two images. This results in a set of correlation peaks, one for each band. The correlation peaks will form a sampled sinusoid of the form given by $\sqrt{\beta}$ from (8), the parameters of which could be easily recovered using of a least squares or similar curve fitting approach. In practice, these methods are highly unreliable for estimating the phase of the sinusoid due to the low signal to noise ratio in the correlation surface. A general Hough transform based approach is better suited for this task [9, 10]. A slightly different formulation is presented here for detecting the resulting sinusoid from that in the literature. Given a fixed-period sinusoid described in Cartesian space in terms of (x, y) as:

$$y = A\sin(\omega x + \phi) + c \quad (27)$$

It can be parameterized in terms of (A, ϕ, c) as:

$$A = \frac{Y - c}{\sin(\omega x + \phi)} \quad (28)$$

Each point in (x,y) now results in a family of inverse sine functions, each one defined over (A, ϕ) for each value of c. The corresponding set of curves for each sample of the sinusoid in (x,y) is accumulated in (A, ϕ, c) space and the location of the resulting maxima gives the parameters of the original sinusoid. Noise reduction by pre-filtering can be used to improve the reliability of correctly detecting the sinusoid.

The various steps in this process are depicted in Fig.1

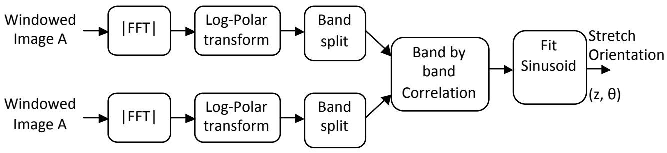

Fig. 1. Calculation of transformation parameters

IV. Experimental Results

The performance of the proposed scheme was evaluated for a wide range of transformation parameters. A total of 432 image pairs were generated by stretching a 1024x1204 source image by factors from 1.1 to 1.4 in steps of 0.1 along different orientations in 5 degree steps from 0 to 180 degrees, after first performing global scaling by factors of 1.0, 1.1 and 1.2.

The accurate recovery of the global scale across this range of parameters decreases as the amount of stretching increases, as shown in in Figure 2. This plots the average scale factors reliably recovered for different amounts of stretching. A given scale factor was deemed to be reliably detected if the average error at a given amount of stretching for all rotation angles was below 15%. This error threshold was chosen rather arbitrarily since error distribution tends to be a step function: where the rings are correctly identified the error is typically below 15%, otherwise the average error is above 40%.

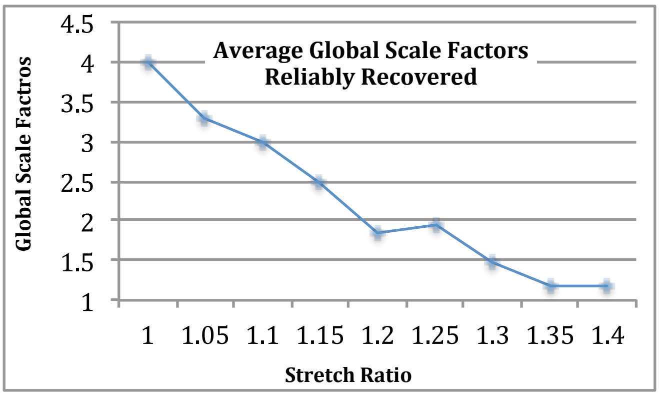

Fig. 2. Recovery of scale factors for different amounts of stretching

In all of these cases the orientation was recovered from the phase of the sinusoid with an average absolute error of 5.69% and a median absolute error of 0.93%. This is depicted in figures 3 and 4 showing the absolute orientation error in degrees for different stretch factors at different orientations and the average absolute error for all orientations respectively. The fluctuations in the curved in fig. 3 are due resolution limiations and noise. The magnitude of the stretching was recovered from the radius of the ring delta with an average absolute error of 6.42% and a median absolute error of 3.02%.

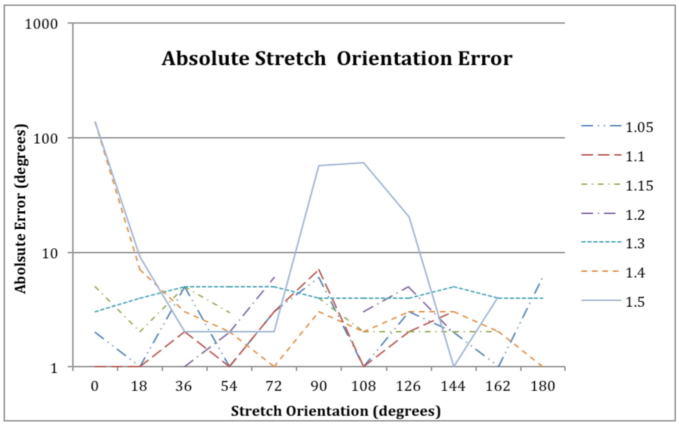

Fig. 3. Absolute orientation recovery errors for different stretch factors

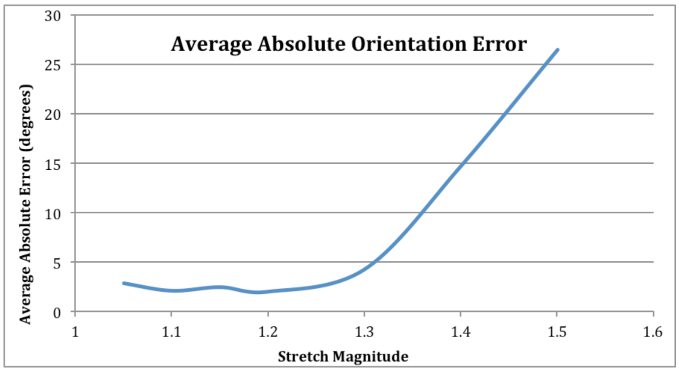

Fig. 4. Average absolute orientation recovery errors for all orientations

Figure 5 shows an image pair where the source image on the left has been globally scaling by a factor of 1.1 and stretched by a factor of 1.3 at angle of -40 degrees from the vertical as shown on its right. The orientation and magnitude of the stretch were correctly recovered respectively as -40 degrees 1.384 times. The deviation in the magitude is due to the quantisation introduced by the Hough space. The panel on the right is a plot the sinusoid formed by the aggregation of the correlation surfaces generated from the image bands.

Fig. 5. Global scaling at 1.1x and stretching by 1.3 at -40°

Figure 6 shows a resulting image after rotating the original by -50 degrees, then stretching it by a factor of 1.30 at an orientation of 25 degrees from the vertical. The proposed method correctly recovered the rotation at -50 degrees and the magnitude and orientation of the stretch at 1.27x and 33 degrees respectively. The centre panel is the ring delta resulting from the correlation with its corresonding sinusoid on its right.

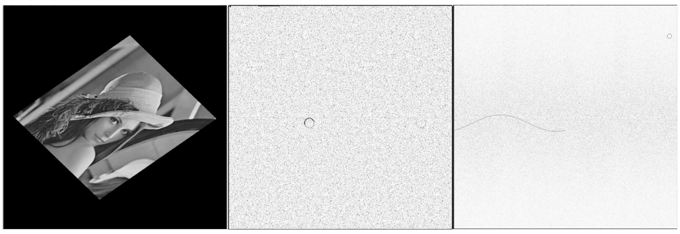

Fig. 6. Rotation of -50° followed by 1.3x stretching at 25°

Figure 7 shows a resulting image after stretching it by a factor of 1.40 at an orientation of -45 degrees from the vertical. The orientation and magnitue of the stretch were recovered as -51 degrees and 1.384 respetively.

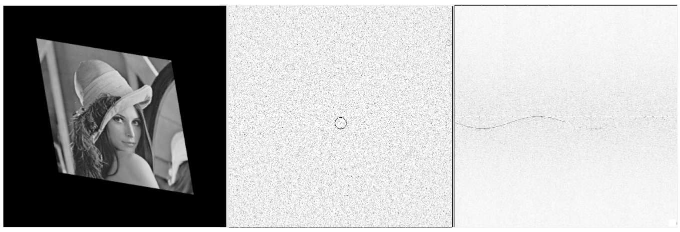

Fig. 7. Stretching at -45° by a factor of 1.40x

Figure 8 shows a resulting image after stretching it by a factor of 1.40 at an orientation of 75 degrees from the vertical. The orientation and magnitue of the stretch were recovered as 78 degrees and 1.325 respetively.

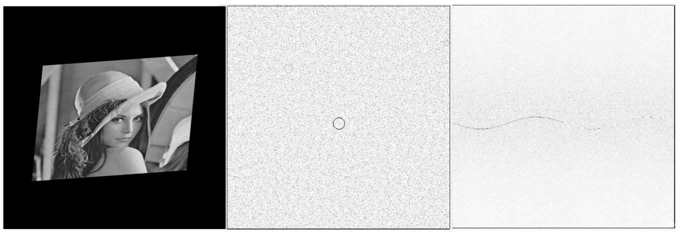

Fig. 8. Stretching at 75° by a factor of 1.40x

V. Conclusions

This paper has presented a novel, straightforward method for estimating the transformation parameters of images that have undergone global scaling, rotation and arbitrary stretching for image registration. Based on combining Phase Correlation with the Hough Transform, it provides noise resilience and constant processing time. While the distortion introduced by stretching limits the upper range of global scaling that could otherwise be recovered, it is able to recover stretch parameters over all possible orientations with reasonable accuracy.

References

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

[2] S. Zokai, G. Wolberg, “Image Registration Using Log-Polar Mappings for Recovery of Large-Scale Similarity and Projective Transformations.” IEEE Trans. Image Processing, vol.14, no.10, pp.1422-1434, 2005

[3] B.J. Super, A.C. Bovik, “Shape from texture using local spectral moments”, *IEEE Trans. PAMI* Apr 1995, Vol.17, No. 4, pp 333–343.

[4] J. Krumm, “Shape from periodic texture using the spectrogram”, *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92*, 15-18 Jun 1992, pp 284 - 289

[5] L. Lucchese, S. Leorin, GM. Cortelazzo, “Estimation of two-dimensional affine transformations through polar curve matching and its application to image mosaicking and remote-sensing data registration.” *IEEE Trans. Image Process*. 2006 Oct; Vo.15, No. 10, pp.3008-19.

[6] R. Gonzalez, “Improving Phase Correlation for Image Registration”, Proceedings of Image and Vision Computing New Zealand 2011, p.488-493, Nov. 29 Dec 1., http://www.ivs.auckland.ac.nz/ivcnz2011/proceedings.pdf

[7] R.Duda, and P.Hart, “Use of the Hough Transform to Detect Lines and Curves in Pictures”, *Communications of the ACM* 15, pp: 11–15, 1975.

[8] R.Gonzalez, “Robust Ring Detection In Phase Correlation Surfaces” to appear in the Proceedings of the International Conference on Digital Image Computing: Techniques and Applications (DICTA2013), Hobart Australia, November 2013.

[9] Zou Changchun; Shi Ge; , "A Hough transform-based method for fast detection of fixed period sinusoidal curves in images," Signal Processing, 2002 6th International Conference on , vol.1, no., pp. 909- 912 vol.1, 26-30 Aug. 2002

[10] K. Glossop, P. J. G. Lisboa, P. C. Russell, A. Siddans, and G. R. Jones, “An implementation of the hough transformation for the identification and labelling of fixed period sinusoidal curves,” *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 96–100, Apr. 1999