A Dual Log-polar Map Rotation and Scale-Invariant Image Transform

Gang-Hwa Lee[1] and Suk-Gyu Lee[2,#]

1 School of EECS, Yeungnam University, 214-1, Dae-dong, Gyongsan, Gyongbuk, Korea, 712-749
2 Dept. of Electrical Engineering, Yeungnam University, 214-1, Dae-dong, Gyongsan, Gyongbuk, Korea, 712-749
Corresponding Author / E-mail: sglee@ynu.ac.kr, TEL: +82-53-810-2487, FAX: +82-53-810-4767

KEYWORDS : Dual log-polar map, Fourier-Mellin transform, Invariance

The Fourier-Mellin transform is the theoretical basis for the translation, rotation, and scale invariance of an image. However, its implementation requires a log-polar map of the original image, which requires logarithmic sampling of a radial variable in that image. This means that the mapping process is accompanied by considerable loss of data. To solve this problem, we propose a dual log-polar map that uses both a forward image map and a reverse image map simultaneously. Data loss due to the forward map sub-sampling can be offset by the reverse map. This is the first step in creating an invertible log-polar map. Experimental results have demonstrated the effectiveness of the proposed scheme.

Manuscript received: October 18, 2007 / Accepted: July 13, 2008

1. Introduction

The recognition of a target object can be an important consideration in machine vision and image processing. Methods that will still distinguish the target object despite any translational, rotational, or scale changes it may experience are required. One such method is the invariant image transform. The Fourier-Mellin transform,[1-6] which is a combination of the Fourier and Mellin transforms, provides the necessary invariant properties and requires the use of a log-polar map.[7] Use of the log-polar map for image registration in the recovery of large-scale similarities has been described previously.[8]

There have been many approaches to Fourier-Mellin transforms including the optical implementation[1-4] and the digital implementation.[5,6] However, most research has focused on the confirmation and/or theoretical enhancement of the well-known Fourier and Mellin transformations.

The Fourier transform of a translated image gives the same Fourier magnitude as the untranslated image. Only their Fourier phases are different. We have translational invariance if we consider only Fourier magnitudes. If an image rotates, the Fourier transform of the rotated image also rotates by the same amount. This gives the rotational invariance. The final and most problematic invariance is scale invariance, which requires the Mellin transform. Unfortunately, the Mellin transform is equivalent to the Fourier transform of the log-scaled radial (or modulus) variable. This log-125 polar map causes considerable data loss at the edges of the image due to the nonlinear sub-sampling.

The implementation aspects of the invariant image transform using a log-polar map have been described.[7] This approach succeeded in implementing the Fourier-Mellin transform by a log-polar map. Unfortunately, this watermarking application did not include direct application of frequency domain watermarking due to the non-invertible sub-sampling of the log-polar map. In a log-polar map, we first convert the rectangular coordinates of an image to polar coordinates and then logarithmically sample the radial variable of the image. This results in dense sampling in the vicinity of the image center. At the edges of the image, however, the sampling becomes very coarse and therefore data are lost. Once data are lost, they can never be recovered. Thus, a log-polar map is not invertible.

In this paper, we propose a dual log-polar map to help make a log-polar map invertible. The dual log-polar map has two log-polar maps in parallel: a forward-image and a reverse-image log-polar map. The forward-image map is the same log-polar map as that implemented by O'Ruanaidh and Pun.[7] However, the reverse-image map uses a reverse image prepared from a copy of the original image by rotating each quadrant of the original image through 180°. Consequently, the loss of information at the edges of the image in the forward map is compensated by the reverse map that loses data in the middle of the image. We believe that this approach can solve the inherent sub-sampling problem of a log-polar map. We describe the advantages and limitations of this dual log-polar map algorithm in the following sections.

In this paper, we propose a dual log-polar map designed to create an invertible log-polar map, and discuss the results of analysis of the reverse-image log-polar map. Simulation results confirmed the feasibility of the proposed dual log-polar map.

Section 2 describes the translation-, rotation-, and scale-invariance. Section 3 develops the proposed algorithm of the dual log-polar map, and explains the experimental results. Section 4 contains the conclusion and discusses areas for further study.

2. Invariant Image Transform (Fourier-Mellin Transform)

The invariant transform can be accomplished by moment

invariants[9] and/or Fourier-Mellin transformation. In this paper, we focus on the transformational method. Figure 1 shows the original forward and reverse images. The reverse image is constructed from the forward (original) image by rotating each quadrant by 180°.

Fig. 1(a) Original "Lena" forward image

Fig. 1(b) "Lena" reverse image

2.1 Translation, Rotation, and Scale Invariance

2.1.1 Translational Invariance

When an image $f(x,y)$ is translated to $f(x-x_0, y-y_0)$, the Fourier transform is

$$F_T(u,v) = e^{-j(ux_0+vy_0)}F(u,v), \tag{1}$$

where $F(u,v)$ and $F_T(u,v)$ are the Fourier transforms of $f(x,y)$ and $f(x-x_0,y-y_0)$, respectively. By taking the magnitudes of $F(u,v)$ and $F_T(u,v)$,

$$|F_T(u,v)| = |e^{-j(ux_0+vy_0)}F(u,v)| = |F(u,v)|, \tag{2}$$

the translational invariance is obtained theoretically. In the discrete case, translation becomes a circular shift of the image.

While Eq. (2) does represent translational invariance, there is one practical problem. As the center of the image does not usually coincide with the image centroid, finding the center of the image without human intervention is not a simple matter. We must assume that the image center is known *a priori*; otherwise, we must have some means of finding the image center. There is one method for this that uses the central moment:[9]

$$(\bar{x},\bar{y}) = \left(\frac{\iint xf(x,y)dxdy}{\iint f(x,y)dxdy}, \frac{\iint yf(x,y)dxdy}{\iint f(x,y)dxdy} \right), \tag{3}$$

where $\iint f(x,y)dxdy$ is the zero-order moment (total image power), and the numerator entries are the first-order moments of the x- and y-coordinates, respectively. The only problem with this approach is that an image that is circularly shifted by the image centroid is not pleasing to the human eye.

2.1.2 Rotational Invariance

When an image $f(x,y)$ is rotated by θ_0, the rotated image is expressed by

$$f_R(x,y) = f(x\cos\theta_0 - y\sin\theta_0,\ x\sin\theta_0 + y\cos\theta_0), \tag{4}$$

and its Fourier transform is

$$F_R(u,v) = F(u\cos\theta_0 - v\sin\theta_0,\ u\sin\theta_0 + v\cos\theta_0). \tag{5}$$

The rotation angle θ_0 is preserved exactly in both the spatial and Fourier domains.

2.1.3 Scale Invariance

If we assume that the x- and y-coordinates of the image have the same scale factor β, then

$$f(\beta x,\beta y) \quad \Leftrightarrow \quad F_S(u,v) = \frac{1}{\beta^2}F(\frac{u}{\beta},\frac{v}{\beta}). \tag{6}$$

The real difficulty of this form for $\beta < 1$ is that the sub-sampling of $f(x,y)$ causes a change of information in the discrete implementation. However, as bilinear or other interpolation methods can be used, we may assume that the data loss is not severe. The scaling of the Fourier coefficient, $1/\beta^2$, is not a problem on its own, and in fact may be used to determine the scale factor by comparing the magnitudes of both transforms, $F(u,v)$ and $F_S(u,v)$.

When all three variances are combined, the corresponding Fourier transform becomes

$$|F_{T,R,S}(u,v)| = \frac{1}{\beta^2}|F(\frac{u}{\beta}\cos\theta_0 - \frac{v}{\beta}\sin\theta_0,\ \frac{u}{\beta}\sin\theta_0 + \frac{v}{\beta}\cos\theta_0)|. \tag{7}$$

Note that if the image center is known precisely and the translational invariance is implicit, then the modulus operation $|\cdot|$ is not necessary.

2.2 Polar Converted Image

If the polar converted image is denoted by $P(r,\theta) = f(x,y)$, then it is easy to show that Eqs. (5) and (6) can be combined into

$$P_{R,S}(r,\theta) = P(\beta r,\theta+\theta_0). \tag{8}$$

If the translational invariance is not implicit, $P(r,\theta) = |F(u,v)|$, and Eq. (8) must be interpreted as

$$P_{R,S}(r,\theta) = \frac{1}{\beta^2}P(\frac{r}{\beta},\theta+\theta_0). \tag{9}$$

2.3 Mellin Transform

The Mellin transform can be derived from the two-sided Laplace

transform

$$M_{P(\rho,\theta)}(s) = \int_{-\infty}^{\infty} e^{-s\rho} P(\rho,\theta)\, d\rho . \quad (10)$$

Let $e^{\rho} = r$, then $P(\rho) = P(\ln r)$ and $dr = rd\rho$. Then, Eq. (10) becomes

$$M_{P(\ln r,\theta)}(s) = \int_{0}^{\infty} r^{-s-1} P(\ln r,\theta)\, dr , \quad (11)$$

where $s = \sigma + j\omega$ and $r > 0$.
When $P(r,\theta) = f(x,y)$, the scaled function $P_S(\rho,\theta) = P(\ln \beta r,\theta)$ has the Mellin transform

$$M_{P(\ln \beta r,\theta)}(s) = \int_{0}^{\infty} r^{-s-1} P(\ln \beta r,\theta)\, dr = \beta^{s} M_{P(\ln r,\theta)}(s). \quad (12)$$

As $|\beta^{s}|_{s=j\omega} = |e^{s\ln\beta}|_{s=j\omega} = 1$, we obtain the required scale invariance of

$$|M_{P(\ln \beta r,\theta)}(s)| = |M_{P(\ln r,\theta)}(s)|. \quad (13)$$

Equations (10) and (13) indicate that we can implement the Mellin transform of $P(\ln r,\theta)$ by the Fourier transform of $P(\rho,\theta)$.

If the translational invariance is not implicit, the polar form of $P(r,\theta) = f(x,y)$ becomes $P(r,\theta) = \frac{1}{\beta^2} |F(\frac{u}{\beta},\frac{v}{\beta})|$, and the scaled function becomes $P_S(\rho,\theta) = \frac{1}{\beta^2} P(\ln\frac{r}{\beta},\theta)$. To achieve scale invariance, we require that $|M_{P(\ln \beta r,\theta)}(s)| = |\beta^{-s-2} M_{P(\ln r,\theta)}(s)|$. The conditions for this are either $\sigma = -2$ or $\beta = 1$ and $\omega = \frac{n\pi}{\ln\beta}$, $n = 0,1,2,\ldots$. Note that $\beta = 1$ means no scale change. In this case, it is very difficult to implement the scale invariance due to the discrete nature of the frequency ω. Without the scale term $1/\beta^2$, we achieve scale invariance and it can be resolved by normalization later.

2.4 Fourier-Mellin Transform

When we perform the Fourier transform of the rotated and scaled polar image of Eq. (8) in both $\ln r$ and θ coordinates, we obtain the two-dimensional Fourier transform

$$FM_{P(\rho,\theta)}(\omega_\rho,\omega_\theta) = \int_{0}^{2\pi}\int_{0}^{\infty} e^{-j(\omega_\rho \rho + \omega_\theta \theta)} P(\rho,\theta)\, d\rho d\theta . \quad (14)$$

As the rotation- and scale-changed image has the transform

$$FM_{R,S}(\omega_\rho,\omega_\theta) = e^{-j\omega_\rho \ln\beta} e^{-j\omega_\theta\theta_0} FM(\omega_\rho,\omega_\theta), \quad (15)$$

the final invariance is obtained by taking its Fourier magnitude as

$$|FM_{R,S}(\omega_\rho,\omega_\theta)| = |e^{-j\omega_\rho \ln\beta} e^{-j\omega_\theta\theta_0} FM(\omega_\rho,\omega_\theta)| = |FM(\omega_\rho,\omega_\theta)|. \quad (16)$$

If the image is not translationally implicit, Eq. (16) must be read as

$$|FM_{T,R,S}(\omega_\rho,\omega_\theta)| = |e^{j\omega_\rho \ln\beta} e^{-j\omega_\theta\theta_0} \frac{1}{\beta^2} FM(\omega_\rho,\omega_\theta)| = \beta^{-2} |FM(\omega_\rho,\omega_\theta)|, \quad (17)$$

and the following normalization step is required

$$\frac{|FM_{T,R,S}(\omega_\rho,\omega_\theta)|}{|FM_{T,R,S}(0,0)|} = \frac{|FM(\omega_\rho,\omega_\theta)|}{|FM(0,0)|}. \quad (18)$$

3. Dual Log-polar Map

A logarithmic r-coordinate conversion of the polar converted image is required to implement Eq. (16). This process is accompanied by the serious problem of data loss at the edges of the image due to the logarithmic sub-sampling. To avoid this, we use another log-polar map of the reverse image in parallel. The subsequent switching of the recovered image will compensate in large part for the lost data.

3.1 Discrete Fourier Transform of Reverse Image

The reverse image of Figure 1(b) can be interpreted as the two consecutive operations of $f[((-x,-y))]_N$ followed by $f[((x-\frac{N}{2},y-\frac{N}{2}))]_N$. The first operation $f[((-x,-y))]_N$ is the 180° rotation of $f(x,y)$ as the discrete Fourier transform (DFT) is circularly periodic. The second operation is the cross-diagonal swapping of the image quadrants; the first quadrant swaps with the third, and the second with the fourth.

When the DFT of $f(x,y)$ is $F(u,v)$, the DFT properties[10] of the two operations $f[((-x,-y))]_N$ and $f[((x-\frac{N}{2},y-\frac{N}{2}))]_N$ are, respectively,

$$f[((-x,-y))]_N \Leftrightarrow F^{*}(u,v), \quad (19)$$

and

$$f[((x-\frac{N}{2},y-\frac{N}{2}))]_N \Leftrightarrow (-1)^{(u+v)} F(u,v). \quad (20)$$

As indicated by Eqs. (19) and (20), the recovery of the forward image from the Fourier transform of the reverse image can be obtained by the inverse Fourier transform of $(-1)^{(u+v)} F^{*}(u,v)$.

3.2 Implementation Results

We produced log-polar maps of the forward and reverse images using the algorithm of O'Ruanaidh and Pun.[7] Figures 2 and 3 show the log-polar map images of Figs 1 (a) and (b), respectively.

Fig. 2 Log-polar map of the forward image

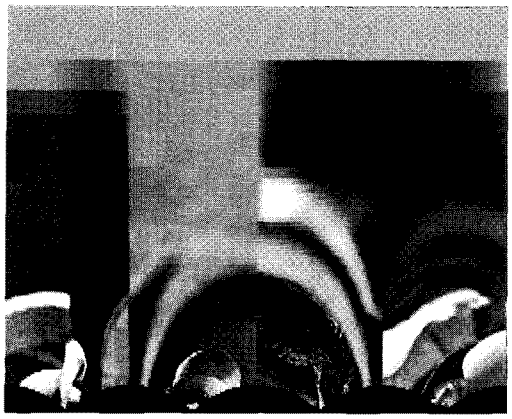

Fig. 3 Log-polar map of the reverse image

The resolution of both figures in the log-polar maps is 256×1.414 in the r-coordinate and 256 in the θ-coordinate. In both figures, the horizontal axis is the θ-coordinate, and the vertical axis is the r-coordinate. The upper horizontal line is the log-polar map of the single point at the image center. The four cross-points on the bottom line correspond to the four corners of the image.

Figures 4 and 5 show the recovered images (inverse log-polar mapped images) of the forward and the reverse images, respectively. Figure 4 clearly shows the sub-sampling effect at the edges of the forward image.

Fig. 4 Forward recovery from the inverse log-polar map of the forward image. The sub-sampling effect at the edge of the image is quite obvious

Fig. 5 Reverse recovery from the inverse log-polar map of the reverse image. The sub-sampling effect in the middle of the image is obvious

However, in Fig. 5, the sub-sampling effect occurs in the middle of the image. This is why we prepared the reverse image for the reverse log-polar map. With this arrangement, we have two log-polar recoveries: the forward map recovery that is most accurate inside, and the reverse recovery that is most accurate outside.

Figure 6 shows the switched selection of the recovered images. The inner image is selected from the forward recovery and the outer image from the reverse recovery. The switching occurs at the perimeter of the diamond-shaped switch box expressed as the absolute value of rectangular coordinates by

$$|x|+|y| \le d_r, \tag{21}$$

where d_r is half of the image length.

If Eq. (21) is satisfied, we choose forward recovery; otherwise, we choose reverse recovery. In Fig. 6, an artifact of the diamond-shaped switch box can be clearly seen; it is especially evident in the background white rod on the left-hand side of the image. The reason for this is that the circular polar orientations are in opposite directions in each map.

Fig. 6 Switched selection of recovered images. The inner image is selected from the forward recovery and the outer image from the reverse recovery using a diamond-shaped switch box

Fig. 7 Switched selection of recovered images. The inner image is selected from the forward recovery and the outer image from the reverse recovery using a circular switch box

We used a circular switch box to remedy the artifacts created by the diamond-shaped switch box. This is a very natural choice considering the nature of polar coordinates. In this particular image, a switch circle with a diameter of two thirds of the image length gives good results. Equation (22) expresses the circular switching formula

$$x^2 + y^2 \le \left(\frac{2}{3} d_r\right)^2, \tag{22}$$

where the switching radius $\frac{2}{3} d_r$ is close to $\sqrt{2}/2\, d_r$.

Figure 7 shows such a switching strategy. The analytical method of minimizing the distance measure of the switched and original images, using the switch shape as a parameter, could also be used.

Continuing work on this topic involves refinement of a polar conversion grid for both forward and reverse images, and the study of the feasibility of multiple (*e.g.*, 4) log-polar maps using 2 forward and 2 reverse images each offset by half a quadrant width from the others. This should reduce the discontinuity in the middle of the image fairly well.

3.3 Comparison of Computational Complexity

The cost of the proposed dual method is the increased computational complexity. It is well known that a 1-dimensional fast Fourier transform requires only half of $N \log_2 N$ multiplications.[10] As a 2-dimensional FFT can be computed by a series of 1-dimensional FFT pairs, the total number of multiplications required is $N \log_2 N$, where N is the number of pixels in the image.
The log-polar conversion was performed for 4 quadrants separately. Each quadrant conversion requires $(N/2)^2$ multiplications as the operation requires the calculation of log-polar sub-sampling grids. As a result, for forward recovery, log-polar and inverse log-polar conversion require $2 \cdot 4 \cdot (N/2)^2$, and FFT and inverse FFT require $2 \cdot N \log_2(N)$ multiplications.

Preparation of the reverse image requires 180° rotation of each quadrant for the reverse recovery. This requires 4 half-image-sized FFTs and results in $4 \cdot (N/2) \log_2 (N/2)$ additional multiplications, ignoring the multiplication by –1. The dual log-polar method requires roughly twice the computational complexity of the single forward map alone.

The running time for the MATLAB program to produce Fig. 2 is 3.04 s, and that for Fig. 3 is 2.9 s on a Pentium PC, including the display of the image. The inverse operation takes approximately the same time. The interpolation complexity is larger in the forward image map.

3.4 Application Areas

Application of the proposed scheme is possible on any system that experiences translational, rotational, and scale changes. In particular, the scheme is well-suited to a system that requires the inverse transform to recover the original images. A watermarking scheme that authenticates the copyright of the original image sources is one good example.[7]

Another possible application in machine vision is the identification of rotation angles. Figure 8 shows the recovery of the

Fig. 8 Forward recovery of rotated (15°) and scaled (0.9) image

Fig. 9 Reverse recovery of rotated (15°) and scaled (0.9) image

"Lena" image that has been rotated 15° and scaled by 0.9. The image suffers the problem of polar conversion, which causes a sawtooth image boundary. This phenomenon is an open problem in any procedure involving rotation.

In this case, we can use the reverse recovery as shown in Fig. 9; the rotated image boundary is fairly straight. A stereoscopic vision system to obtain depth information[11] may be used with the proposed scheme.

4. Conclusions

In this paper, we proposed a dual log-polar map as a prototype for a true translational-, rotational-, and scale-invariant transform. The compensation for the lost data in the forward log-polar map is achieved using the reverse log-polar map. The reverse map does not mean a simple reversal of image direction. It satisfies the Fourier transform properties that are necessary for the invariant transform.
The invariant image transform is purely theoretical in the continuous spatial domain, but has several fundamental problems: it loses data due to sub-sampling, it is not reversible due to the Fourier magnitude, and it always requires interpolation. The experimental results showed that the proposed algorithm works quite well in providing a solution to the first of these problems.

Further studies to continue this research are refinement of the polar grid for both log-polar maps and consideration of the proposed algorithm for machine vision applications.

REFERENCES

1. Casasent, D. and Psaltis, D., "Position, rotation, and scale invariant optical correlation," Applied Optics, Vol. 15, No. 7, pp. 1795-1799, 1976.
2. Casasent, D. and Psaltis, D., "New Optical Transforms for Pattern Recognition," Proceedings of the IEEE, Vol. 65, No. 1, pp. 77-84, 1977.
3. O'Toole, R. K. and Stark, H., "Comparative study of optical-digital vs all-digital techniques in textural pattern recognition," Applied Optics, Vol. 19, No. 15, pp. 2496-2506, 1980.
4. Sheng, Y. and Duvenoy, J., "Circular-Fourier-radial-Mellin transform descriptors for pattern recognition," Journal of Optical Society of America, Vol. 3, No. 6, pp. 885-888, 1986.
5. Feng, L. and Brandt, R. D., "Towards absolute invariants of images under translation, rotation, and dilation," Pattern Recognition Letter, Vol. 14, Issue 5, pp. 369-379, 1993.
6. Brandt, R. D. and Feng, L., "Representations that uniquely characterize images modulo translation, rotation, and scaling,"

Pattern Recognition Letters, Vol. 17, Issue 9, pp. 1001-1015, 1996.

7. O' Ruanaidh, J. J. K. and Thierry, P., "Rotation, scale and translation invariant spread spectrum digital image watermarking," Signal processing, Vol. 66, Issue 3, pp. 303-317, 1998.
8. Zokai, S. and Wolberg, G., "Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations," IEEE Transactions on Image Processing, Vol. 14, Issue 10, pp. 1422-1434, 2005.
9. Hu, M. K., "Visual pattern recognition by moment invariants," IRE Transactions on Information Theory, Vol. 8, Issue 2, pp. 179-187, 1962.
10. Oppenheim, A. V. and Schafer, R. W., "Discrete-time Signal Processing," Prentice Hall, pp. 530-547, 1989.
11. Fu, K. S., Gonzalez, R. C. and Lee, C. S. G., "Robotics; Control, Sensing, Vision, and Intelligence," McGraw-Hill, pp. 325-328, 1987.