The Log-polar Image Representation in Pattern Recognition Tasks[*]

V. Javier Traver and Filiberto Pla

Dep. de Llenguatges i Sistemes Informàtics,
Universitat Jaume I, E12071-Castelló, Spain
{vtraver,pla}@uji.es

Abstract. This paper is a review of works about the use of the log-polar image model for pattern recognition purposes. Particular attention is paid to the rotation- and scale-invariant pattern recognition problem, which is simplified by the log-polar mapping. In spite of this advantage, ordinary translations become a complicated image transform in the log-polar domain. Two approaches addressing the estimation of translation, rotation and scaling are compared. One of them, developed by the authors, takes advantage of the principles of the active vision paradigm.

1 Introduction

Computer vision often looks at biology for inspiration. This is the case of log-polar images, which follows a foveal model found in some vertebrates, including humans [19]. After these neuro-physiological findings, researchers in artificial vision started to adopt it in their algorithms.

Two main areas where the log-polar model has been adopted are active vision [5] and pattern recognition [28], as it is summarized in the rest of this section. The novel contribution of this paper is as a survey of works on log-polar imagery for pattern recognition, with an emphasis paid in the comparison between a traditional well-known technique (Fourier-Mellin Transform) and an active-vision based approach (developed by the authors) for the problem of similarity motion estimation. After that, rotation and scaling invariances in the log-polar domain and some approaches using them are introduced in Section 2. The problem of dealing with conventional (cartesian) translations in log-polar images is also considered, and two algorithms that can cope with translations combined with scaling and rotation are discussed and compared. Then, in Section 3, applications using the log-polar transform and the benefits brought by it are briefly commented. Finally, at the light of the results surveyed throughout the paper, concluding remarks are drawn in Section 4.

---

[*] Research partly funded by *Conselleria d'Educació, Cultura i Cincia, Generalitat Valenciana*, under project CTIDIB/2002/333.

F.J. Perales et al. (Eds.): IbPRIA 2003, LNCS 2652, pp. 1032–1040, 2003.

Log-polar Images and Active Vision. On the one hand, active vision [1, 2] and foveal sensing are intimately linked to the extent that the term *space-variant active vision* has been coined in the past. One of the favorable properties that log-polar images bring to the field of active vision is a trade-off solution between large field of view, small image size and good resolution at the point of interest. Log-polar images are such that devote a high visual acuity in the center of the field of view, so that foveated objects can be perceived with very good quality. Because resolution decreases exponentially with eccentricity, the size of the log-polar image is small, so that active vision algorithms can exhibit real-time performance. The coarser resolution at the periphery can still be used to detect potentially interesting events that deserve further attention. The topology of log-polar images carry an implicit focus of attention that is particularly useful in vergence control [8] or tracking algorithms [25], because background information becomes less distracting in comparison with uniformly sampled images. Advantages of the log-polar geometry for time-to-impact computation [24], depth estimation [23], or motion stereo [3], among others, have also been studied.

Log-polar Mapping and Pattern Recognition. On the other hand, pattern recognition problems may also benefit from the log-polar representation. This paper focuses on the use of this image model to achieve rotation- and scale-invariant (RSI), or translation-, rotation- and scale-invariant (TRSI) object representations. It is worth noticing an important difference that usually arises between the log-polar images used in active vision applications and those used in recognition tasks. In robotics, there has been a trend to design and use true retina-like sensors (e.g. [29]), or at least, to simulate the log-polar images by software conversion (e.g. [4, 11, 25]) while using cartesian images for the only purpose of the transformation. In this case, we can speak of log-polar *images*. However, practitioners in pattern recognition usually approach the problem from a different perspective, in which it is more appropriate to speak of the log-polar *mapping*, because the transformation is used as a tool. This distinction is fundamental due to its practical consequences:

Image Size. Many recognition problems do not have hard-time constraints, so that devoting a lot of time to process large log-polar images is not really a concern. This is not the case, however, when fast computations are a must.

Biological Motivation. Because log-polar images are obtained as a result of a transformation, cartesian images are still available, and one can make use of both kinds of images, exploiting the best of both worlds. This advantage may be seen as a problem if one is interested in being biologically consistent, or when cartesian images are simply not available.

2 Achieving Invariances

Rotation and Scaling Invariances. It is well-known the fact that rotation and scaling become shifts with the log-polar transform. These properties derive from the topological nature of log-polar images. On the one hand, the *polar* geometry maps angles to shifts along the angular axis of the log-polar image. On the other hand, the *logarithmic* law —which governs the location of receptive fields away from the center of fixation— maps changes of scale to shifts along the radial axis of log-polar images.

Therefore, to estimate a rotation angle or a scale factor, one has only to estimate two shifts: an angular shift and a radial shift, from which the scaling and rotation angle can be derived straightforwardly using the underlying log-polar model. It is worth noticing that some, but not all, of the existing log-polar models actually possess these properties [25]. Therefore, the choice of a log-polar model should take this into account if RSI is a desirable feature.

Pattern Representation and Recognition. Edge invariance (a more suitable name to refer to the RSI of log-polar images) is the key feature exploited in [28] to represent pattern templates. To that end, it is central the *scale and rotation normalization*: rotation normalization is achieved by cyclically shifting the rows of the image by an amount corresponding to the angle of the major axis of the pattern. Scale normalization involves shifting the image down until there is an edge in the bottom row. This work is an enhancement of the normalization method first proposed in [16].

Fourier-Mellin Transform. The Fourier-Mellin Transform (FMT) is a method for making images rotation-, scale- and translation-invariant. The idea is to evaluate the Fourier Transform (FT), then the Log-polar Transform (LPT), and finally another Fourier Transform [20]. This approach relies on the shift theorem of the FT, on the edge-invariance property of the LPT, and on their combination on the appropriate data and in the right order. Through phase correlation, this system can not only determine whether a similar object is in both images, but also to quantify the scaling (α), rotation (ϕ) and translation (d_x, d_y). The overall steps needed in this process are illustrated in Fig. 1, while for specific details, the reader is addressed to Ref. [20].

Another interesting point in [20] is that complex numbers are used to represent color information. Phase correlation performed in this way can discriminate between the different colors of similarly shaped objects. Thus, the argument of the displacement peak —which is complex— is an angle whose value corresponds to the difference in color between the object in the reference image and that in the object image. The advantage of using complex color representation is that the color of the displaced object is calculated as part of the location procedure, with no extra processing.

A log-polar transform on the visibility image (the magnitude of the Fourier transform), is applied in [7], resulting in the so-called *log-polar visibility.* After that, scalings and rotations correspond to shifts. The approach is illustrated in a shift-, scale- and rotation-independent object recognition task. Even though the phase of the Fourier transform is interesting for image representation, an accurate description of the magnitude may also suffice to represent images. In

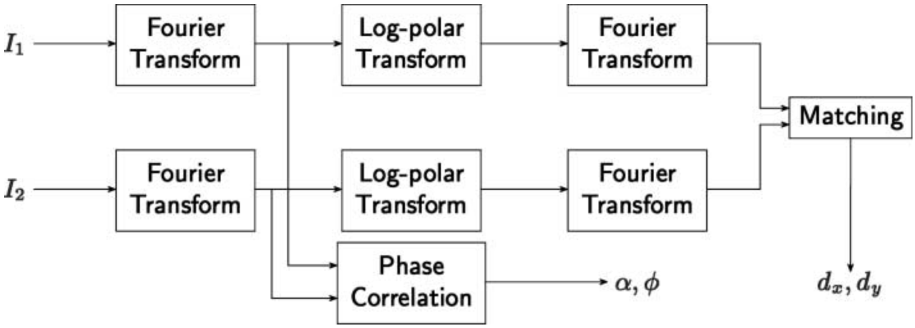

Fig. 1. Fourier-Mellin Transform. The input to the algorithm are two cartesian images I_1, I_2, and the output are the parameters α, ϕ, d_x, and d_y

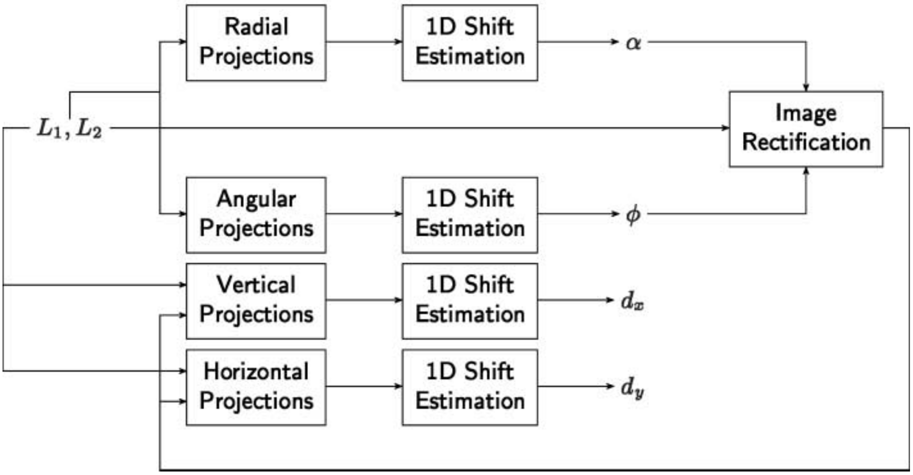

Fig. 2. Two-stage projection-based similarity estimation. The input to the algorithm are two log-polar images L_1, L_2, and the output are the motion estimates (α, ϕ, d_x, d_y)

addition, by using the magnitude of the Fourier transform, the information of the position of the object (which is associated with the phase) is lost, which is an advantage in this particular case.

Exponential Chirp Transform. As can be appreciated, the joint use of the Fourier and the log-polar transforms results in a powerful tool for general object recognition tasks. The main disadvantage of the FT is that the shift invariance does not hold when it is applied directly to a log-polar image. Thus, FMT resorts to cartesian images, which is not a feasible solution if only log-polar images are available, or biologically plausible solutions are sought.

Bonmassar and Schwartz introduced the Exponential Chirp Transform [6], which allows a Fourier-like transform be defined *directly* in the log-polar plane, thus having the benefits of the space-variant architecture and avoiding the drawback of the lack of simple shift invariance.

Active-Vision Approach. While estimating rotation and scaling is facilitated by the log-polar transform, the estimation of ordinary translation is more difficult in log-polar images [26]. The problem gets even more complicated when rotations, scaling and translations are combined.

As we have seen above, the FMT approaches the problem by using both cartesian and log-polar images. This, along with the fact that computationally expensive Fourier transforms are used, makes the FMT an inappropriate solution in many cases. By exploiting the advantages of active vision, another much simpler and efficient solution is proposed in [25]. The idea is that, if translation is small, rotation and scaling can be estimated quite easily. Then, after removing the effects of these transformations, translation can be estimated by some other simple procedure [26]. The assumption of small translations is valid under an active tracking scenario, where the motion of a target is compensated by dynamically moving the camera. Unlike the FMT, this algorithm, outlined in Fig. 2, makes direct and only use of log-polar images. One further interesting point in this proposal is that all four motion components (horizontal and vertical translation, as well as rotation and change of scale) are dealt with through a uniform approach: estimating the 1D shift between two one-dimensional signals. These 1D signals are image projections computed along different appropriate directions.

3 Log-polar Transform for Pattern Recognition

In this section, applications of the LPT are briefly commented. Besides the edge invariance, which the first part of this paper focuses on, other advantages that can be derived from the LPT are mentioned in the subsequent examples.

Straight lines and circles detection. Equations of straight lines and circles take a special form in the log-polar domain [27, 30]. In [30], a point-by-point complex product of the FFT (Fast FT) of the log-polar image with the FFT of a template line, after being mapped back to the spatial domain, yields strong convolution peaks for the most salient straight lines. A similar idea is used for circle detection. Weiman [27] shows that the parameter space of Hough transform has a coordinate system which is identical to the log-polar coordinate system, and stresses the advantages implied by this fact: it simplifies the computations for line recognition, eliminates slope quantization problems, and the so-called *log-Hough* transform is efficient in line and curves detection.

Features recognition by foveation. Lim *et al.*'s work [15] addresses the problem of foveation and feature (line, circle, and ellipse) recognition. An initial open-loop foveation generates a coarse movement (a saccade). Subsequently, a closed-loop foveation involves more accurate movements (micro saccades). What it is important in this work is the observation that when the foveation is accurate, the resultant pattern in log-polar space is always a horizontal line, regardless of the type of feature. This interesting result is used in the closed-loop foveation stage: if the foveation is not accurate, the log-polar mapping will be a slightly curved segment. Additionally, the deviation of this curve from the straight line indicates the amount of error.

Neural networks-based systems. The reduction of information achieved by the log-polar mapping is an attractive feature for researchers in neural networks (NNs), because it allows computationally feasible implementations of NNs for object recognition [18]. In addition, better recognition rates are reported when the input characters are represented by log-polar images than when they are represented by cartesian images because of the more favorable distribution of informative areas [17]. For the recognition of handwritten numerals, two NNs (one working with the input image in *cartesian* coordinates, and the other working in *log-polar* coordinates) are adopted in [14], the latter NN making rotational and scaling invariance possible. By using the interesting principle of *segmentation by recognition*, segmentation plays a key role in a successful recognition of connected cursive handwritten numerals or characters. In [12], a system for recognizing warning and caution traffic signs is described. Two NNs were used: one for color segmentation and the other for traffic sign invariant signature classification. The first 16 Fourier coefficients of the transformed log-polar images of the traffic signs were used as input for the second NN. The system is reported to achieve correct classifications in the presence of rather large noise levels.

Face detection/recognition. A system for face detection and recognition, which uses the log-polar mapping is described in [13]. The face detector encloses the face from the complex scene with a circular boundary and locates the position of the nose. The log-polar mapping is basically used for feature extraction in conjuction with PCA (Principal Component Analysis). The largest circle in the log-polar grid is adjusted to enclose the whole object, to obtain scale normalized feature vectors. Interestingly, the recognition rate is rather less sensitive to the log-polar image resolution than other methods compared in [10]. This robustness of log-polar images to variations in their resolution seems to be in accordance with the insensitiveness of the correlation measures to different image resolutions [4].

Contextualizing features. It is a point deserving some attention that the periphery in log-polar images has been used for two different —and seemingly contradictory— purposes. On the one hand, its coarse resolution allows that information in peripheral areas do not become too distracting, and the relevant information at the fovea becomes "dominating". On the other hand, even though the important information may reside in the high-resolved fovea, image data further away from the fovea may play an important "discriminative" role. In other words, it may help disambiguate the central information, by providing an economic (coarsely resolved) "context" (the surrounding information). To benefit from this more subtle, and less exploited capability, the LPT is found to be ideal in [9] in training NNs for facial features (eyes, nose and mouth) location, as well as in [21, 22], where an active face recognition system is described.

4 Final Remarks

The edge invariance property of log-polar images brings important advantages in rotation- and scale-invariant object recognition. Ordinary translations, which map to a complex transformation in the log-polar domain, can also be properly dealt with by proposed algorithms. The Fourier-Mellin transform, a standard tool to solve the more complex problem of rotation-, scale-, and translation-invariance, has important disadvantages (high computational requirements and the use of both cartesian and log-polar images) which may be overcome by approaches based on the principles of active vision. Other interesting properties, such as small processing time, appropriate coordinate system for feature detection, and inexpensive coarse-resolution context disambiguation, are also advantageous, as demonstrated in practical applications. Finally, it is evident that *active recognition* (pattern recognition under an active vision paradigm), and the use of log-polar imagery within it, is a promising framework to solve both old and new pattern recognition problems.

References

[1] J. Y. Aloimonos, I. Weiss, and A. Bandyopadhyay. Active vision. *Intl. Journal of Computer Vision*, pages 333–356, 1988. 1033

[2] D. H. Ballard. Animate vision. *Artificial Intelligence*, 48:57–86, 1991. 1033

[3] S. L. Bartlett, A. Hampapur, M. J. Huber, D. Kortenkamp, and S. Moezzi. Vision for mobile robots. In J. L. C. Sanz, editor, *Image Technology: Advances in Image Processing, Multimedia and Machine Vision*, pages 1–37. Springer, 1996. 1033

[4] A. Bernardino and J. Santos-Victor. Sensor geometry for dynamic vergence: Characterization and performance analysis. In *Workshop on Performance Characteristics of Vision Algorithms, ECCV*, 1996. (Also as TR 01/96 at VisLab, Lisbon, Portugal). 1033, 1037

[5] M. Bolduc and M. D. Levine. A review of biologically motivated space-variant data reduction models for robotic vision. *Computer Vision and Image Understanding (CVIU)*, 69(2):170–184, Feb. 1998. 1032

[6] G. Bonmassar and E. L. Schwartz. Space-variant Fourier analysis: The exponential chirp transform. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 19(10):1080–1089, Oct. 1997. 1036

[7] L. Capodiferro, R. Cusani, G. Jacovitti, and M. Vascotto. A correlation based technique for shift, scale and rotation independent object identification. In *Intl. Conf. on Acoustics, Speech, and Signal Processing*, volume 1, pages 221–224, 1987. 1034

[8] C. Capurro, F. Panerai, and G. Sandini. Vergence and tracking fusing log-polar images. In *Intl. Conf. on Pattern Recognition (ICPR)*, pages 740–744. IEEE, 1996. 1033

[9] P. Debevec. A neural network for facial feature location. C283 Course report, Berkeley University, `http://www.cs.berkeley.edu/~debevec/face_recognition_report.pdf`, Fall 1992. 1038

[10] K. Hotta, T. Kurita, and T. Mishima. Scale invariant face recogniton method using spectral features in log-polar image. In *Applications of Digital Image Processing XXII (SPIE Proc.)*, pages 33–43, 1999. 1037

[11] F. Jurie. A new log-polar mapping for space variant imaging. Application to face detection and tracking. *Pattern Recognition*, 32:865–875, 1999. 1033

[12] N. Kehtarnavaz and A. Ahmad. Traffic sign recognition in noisy outdoor scenes. In *IEEE Intl. Conf. on Intelligent Vehicles*, pages 460–565, 1995. 1037

[13] L. H. Koh, S. Ranganath, M. W. Lee, and Y. V. Venkatesth. An integrated face detection and recognition system. In *Intl. Conf. on Image Analysis and Processing (ICIAP)*, Venice, Italy, Sept. 1999. 1037

[14] S. Lee and T. Horprasert. Recognition of handwritten connected numerals based on dual cooperative neural network. In *Intl. Conf. on Neural Networks*, Perth, Australia, Dec. 1995. 1037

[15] F. L. Lim, G. A. W. West, and S. Venkatesh. Use of log polar space for foveation and feature recognition. *IEE Proc. Vis. Image Signal Process.*, 144(6):323–331, Dec. 1997. 1037

[16] L. Massone, G. Sandini, and V. Tagliasco. 'Form-Invariant' topological mapping strategy for 2D shape recognition. *Computer Vision, Graphics, and Image Processing*, 30:169–188, 1985. 1034

[17] Z. Mikrut and G. Augustyn. Influence of the object representation on the results of characters recognition in the car's licence plates. In *5th Conf. on Neural Networks and Soft Computing*, Zakopane, Poland, June 2000. 1037

[18] Z. Mikrut and B. Czwartkowski. Log-Hough space as input for a neural network. In *4th Conf. on Neural Networks and their Applications*, Zakopane, Poland, May 1999. 1037

[19] E. L. Schwartz. Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception. *Biological Cybernetics*, 25:181–194, 1977. 1032

[20] A. L. Thornton. *Color Object Recognition Using a Complex Colour Representation and the Frequency Domain.* PhD thesis, Department of Engineering, The University of Reading, May 1998. 1034

[21] M. Tistarelli. Active/space-variant object recognition. *Image and Vision Computing (IVC)*, 13(3):215–226, Apr. 1995. 1038

[22] M. Tistarelli and E. Grosso. Active vision-based face authentication. *Image and Vision Computing (IVC)*, 18:299–314, 2000. 1038

[23] M. Tistarelli and G. Sandini. Estimation of depth from motion using an anthropomorphic visual sensor. *Image and Vision Computing (IVC)*, 8(4):271–278, Nov. 1990. 1033

[24] M. Tistarelli and G. Sandini. On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 15:401–410, 1993. 1033

[25] V. J. Traver. *Motion Estimation Algorithms in Log-polar Images and Application to Monocular Active Tracking.* PhD thesis, Dep. Llenguatges i Sistemes

Informàtics, Universitat Jaume I, Castellón (Spain), Sept. 2002. 1033, 1034, 1036

[26] V. J. Traver and F. Pla. Dealing with 2D translation estimation in log-polar imagery. *Image and Vision Computing (IVC)*, 21(3):145–160, Feb. 2003. 1036

[27] C. F. R. Weiman. Polar exponential sensor arrays unify iconic and Hough space representation. In *SPIE Conf. on Intelligent Robots and Computer Vision VIII: Algorithms and Techniques*, volume 1192, Philadelphia, Nov. 1989. 1036

[28] J. C. Wilson and R. M. Hodgson. Log-polar mapping applied to pattern representation and recognition. *Computer Vision and Image Processing*, pages 245–277, 1992. 1032, 1034

[29] R. Wodnicki, G. W. Roberts, and M. D. Levine. A foveated image sensor in standard CMOS technology. In *Custom Integrated Circuits Conf.*, Santa Clara, May 1995. 1033

[30] D. Young. Straight lines and circles in the log-polar image. In *British Machine Vision Conference*, Bristol, UK, Sept. 2000. 1036