International Journal of Computer Vision, [Volumn Number], 1–9 (1998)

Real-Time Motion Blur Estimation and Restoration in Foveal Active Vision Systems.

BONMASSAR GIORGIO

Department of Biomedical Engineering, Boston University, Boston MA 02215

ERIC L. SCHWARTZ

Department of Cognitive and Neural Systems, Boston University, Boston MA 02215

Received January 29, 1998;

Abstract. Foveated, log-polar, or space-variant image architectures provide a high resolution and wide field workspace, while providing a small pixel computation load. These characteristics are ideal for mobile robotic and active vision applications. A common problem in these application areas is image blur and motion artifact. Recently, there has been described a generalization of the Fourier Transform (the fast exponential chirp transform) which allows frame-rate computation of full-field 2D frequency transforms on a log-polar image format. In the present work, we use Wiener filtering, performed using the Exponential Chirp Transform, on log-polar (foveated) image formats to de-blur images which have been degraded by camera motion or other sources of blur. Since the exponential chirp transform provides size and rotation invariant properties, we only need to pre-compute and store a single motion vector template, exploiting the size and rotation properties of the chirp transform. This provides real-time performance without requiring a large storage of pre-computed templates. Examples of this processing on different

samples from a data base of images shows effective image restoration, with frame-rate performance on modest hardware (e.g. a Pentium or SHARC DSP processor). This work shows that the fast exponential chirp transform provides a simple algorithm with low spatial complexity for full frame, real-time image restoration of foveal imagery.

Keywords: Logpolar Mapping, Motion Blur, Space-Variant Image Processing, Image Restoration, Real-Time Imaging

1. Introduction

The popolarity of Active Vision System has increased over the past decade[5, 10], finding large application in many fields of practical importance, nevertheless it must often contend with image blur generated by camera and object motion. All systems that require both active vision and real-time imaging, such as mobile robotics or vehicle navigation based platforms, have implicitly to adopt a motion blur restoration strategy. Algorithms for such restoration are computationally very intensive and problematic to implement in real-time systems. Recently several VLSI and LSI chips have been introduced [20, 12], for motion estimation and compensation. In the present paper we present an algorithmic solution which does not require any special external hardware, for use in applications in which space-variant, or foveating image architecture is the basis of the active vision system. Recently, the authors have devellopped a *new* transform based on the Fourier transform: the *Exponential Chirp Transform* [6]. In this paper this transform is applied to perform real-time, full-frame motion de-blur of space-variant images using minimal storage and conventional (e.g. current generation Pentium) hardware.

Fast transforms, such as the DCT (Discrete Cosine Transform), have been recently proposed to devellop algorithms for motion blur compensation [11]. Such transforms are applied on the entire full-image, therefore are computationally very intense and cannot be performed in real-time using conventional workstations. The same limitations affect methods based on the pseudo-inverse, SVD

[8], and Wiener filtering methods [9] which are the traditional algorithms for motion compensation.

Motion de-blur of images with pyramidal representation, often proposed as ideal format for image transmission, is done using a compensation operator Wavelet based [1]. All methods based on image transformation incur a complexity which is at least linear with the image size, but most commonly the complexity is order $O(\frac{N_1 N_2}{2} \log(N_1 N_2))$, where N_1 and N_2 are the image sizes. A high complexity in general precludes real-time implementation without special purpose hardware.

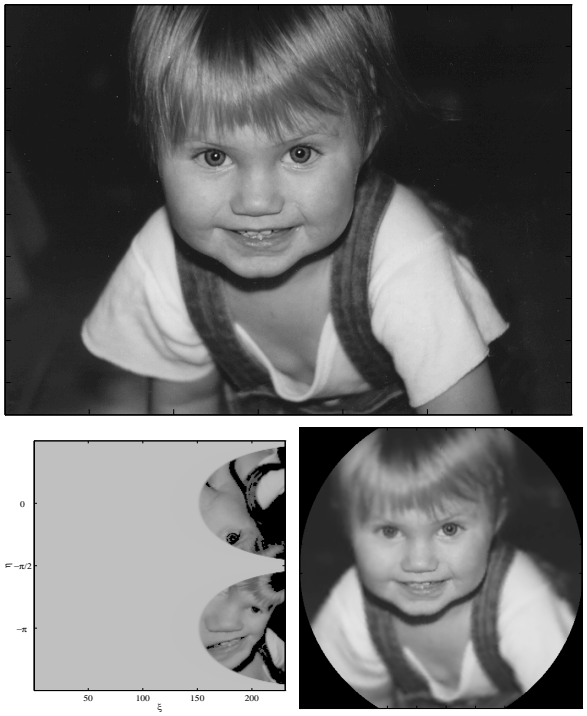

Fig. 1. Images of Caterina. At the top is a traditional uniformly sampled frame (588 × 358), in the bottom left is its log-polar representation (258×, $a = 40$ and $R =$) and in the bottom right is the space-variant image in uniformly sampled coordinate space (340 × 340). This last image is obtained by interpolation of the log-polar image, using a cubic spline algorithm. The notable aspect of this transformation is the very small amount of pixels needed to represent a log-polar image.

The images which are used principally in the present paper, are space-variant with high (*local*) resolution in the center (or *fovea*) and a linearly decreasing resolution with the distance from the center, such non-uniform sampling is called *log-polar*. The advantage of this type of image format is that it combines high-resolution and large field of view with a small number of overall pixels compared with the original image. We firmly believe that this format is ideal for active-vision tasks such as object tracking [19]. Furthermore this type of foveated vision, inspired by the human visual system, is often used in active vision systems for optimal servoing [15, 13]. The log-polar image format, apart from its axis representation[1], is known to be a continuos version image pyramid format, which is widely more adopted. The main disadvantage of the log-polar mapping is tied to

the high difficulty in its use: virtually no image processing tools are available. This lack of image processing tools has led to the somewhat paradoxical fact that the apparent architecture used in all higher vertebrate image processing has been largely unrepresented in machine vision applications.

Recently, with the authors introduction of a new transform, *the Exponential Chirp Transform*, it is possible to apply, in a relatively simply manner, virtually all frequency domain techniques to log-polar sampled images. The algorithm used to compute such tranform, the Fast Exponential Chirp Transform (FECT), has the same complexity of the ordinary two-dimensional FFT: $O\left(N_1 N_2 \log(N_1 N_2)\right)^2$. One of the principle applications motivating the development of the FECT was to allow template matching techniques such as cross-correlation (i.e. space-variant cross-correlation) directly in frequency, in log-polar image formats. In the present paper, we show that the FECT with a Wiener filtering based restoration algorithm provides the ability to estimate and correct motion blur in real time.

1.1. Previous Work

The log-polar map is of interest in computer vision for three major reasons:

1. It has been shown to be a good approximation to the image format used in the primate and human visual cortex (see [18] for a review of the mathematical characterization of the anatomical structure of V-1), and would seem to provide advantages to machine vision which are similar to those that are already understood to apply to human vision.
2. It provides a wide-angle yet high resolution (i.e. foveal) image format with attractive space-complexity. The compression (or decimation) relative to a comparable space-invariant format is up to four orders of magnitude for human vision [14], and up to two or three orders of magnitude for currently realizable machine vision systems [17].
3. It provides a continuum model for variable resolution, or foveating pyramid architectures in machine vision, which is an alternative to the coarsely sampled (e.g. 5 level) pyramid architecture familiar in machine vision.

Real-time vision is an obvious area of application for this architecture, which has been studied by several different laboratories in recent years [21, 2, 4, 7, 16, 3]. The log-polar mapping can be defined as:

$$\xi(x) = \begin{cases} \log(x+a), x \geq 0 \\ 2\,\log(a) - \log(-x+a), \\ x < 0 \end{cases} \quad (1)$$

The Fourier Transform under an arbitrary smooth change of coordinates $(x, y) \rightarrow (\xi, \eta)$ can be expressed as follows:

$$\iint_{-\infty}^{+\infty} f(x(\xi,\eta), y(\xi,\eta))|J(\xi,\eta)| \cdot e^{-2\pi j(k\cdot x(\xi,\eta)+h\cdot y(\xi,\eta))} \, d\xi\, d\eta \quad (2)$$

where $J(\xi, \eta)$ represents the Jacobian of the transformation.

By using the log-polar mapping expressed by eq.(1) in the previous equation, we obtain an expression for the Fourier Transform applied in log-polar coordinates:

$$\iint_{\mathcal{D}} f(\xi,\eta) e^{2\xi} e^{-2\pi j(k(e^{\xi}\cos(\eta)-a)+h\,e^{\xi}\sin(\eta))} \, d\xi\, d\eta \quad (3)$$

where $\mathcal{D} \equiv \{(\xi, \eta) : -\infty \leq \xi < +\infty \text{ and } \frac{-3\pi}{2} \leq \eta < \frac{\pi}{2}\}$.

Eq.(3) constitutes the frequency representation of a log-polar mapped image. Unfortunately its numerical complexity, $O(N_1^2 N_2^2)$, is too high for a real-time implementation. By introducing a log-polar sampling in frequency, eq.(3) can be written as a two-dimensional complex cross-correlation, which can be computed very efficiently using FFTs. The frequency log-polar re-mapping is given by:

$$k(\gamma,\theta) = e^{\gamma}\,\cos(\theta) - b$$

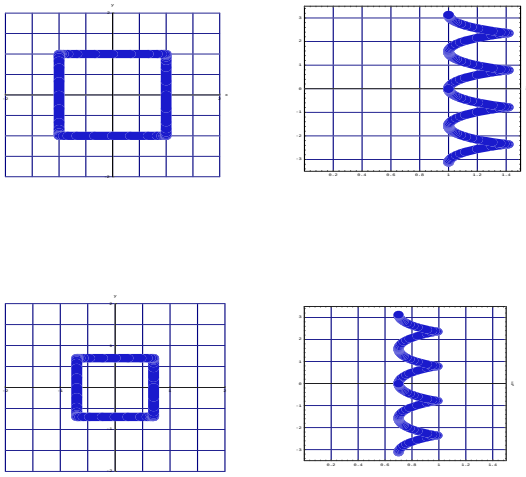

Fig. 2. The *scaling* property of the log-polar function $w = \log(z)$. Resizing a square in the Cartesian domain (left, top and bottom plots) correspond to a shift along the horizontal axis (right, top and bottom plots).

$$h(\gamma,\theta) = -e^{\gamma}\,\sin(\theta) \qquad (4)$$

Thus, the Exponential Chirp Transform can be written as the following cross-correlation:

$$F(\gamma,\theta) = e^{2\pi jak(\gamma,\theta)} \iint_{\mathcal{D}} (f^*(\xi,\eta)$$

$$e^{2\xi - 2\pi bje^{\xi}\cos\eta})^* e^{-2\pi je^{(\gamma+\xi)}\cos(\theta+\eta)}\,d\xi\,d\eta \qquad (5)$$

We refer to the numerical implementation of eq.(5) as the FECT. The details of the FECT can be found in [6].

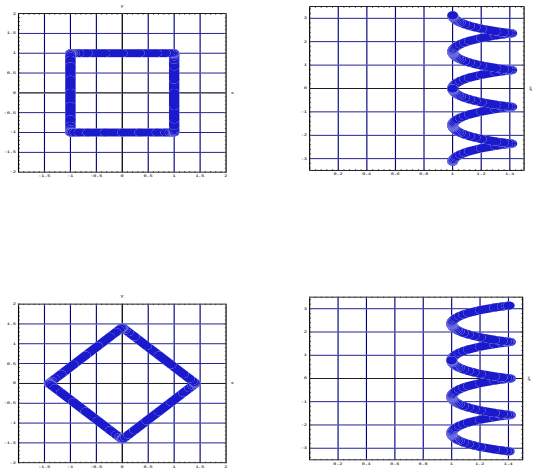

Fig. 3. The *rotation* property of the log-polar function $w = \log(z)$. Rotating a square in the Cartesian domain (left, top and bottom plots) correspond to a circular-shift along the vertical angular axis (right, top and bottom plots).

2. Wiener filtering.

De-blurring an image can be performed by using the inverse filter, which is the division of the image frequency transform by the blur transfer function. However, this method is extremely noisy due to the possible zeros or small values in the blur-vector frequency representation. An alternative

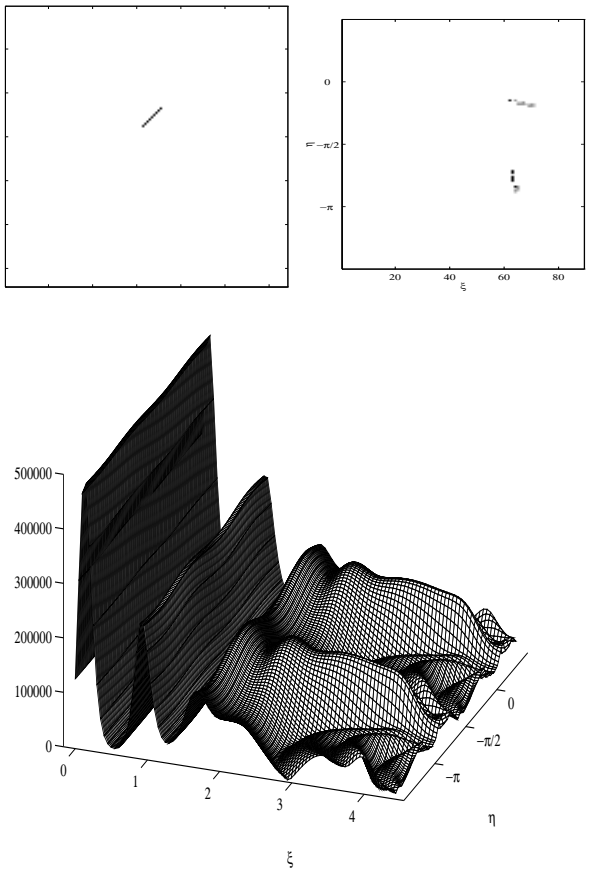

Fig. 4. An example of a motion vector. In the top left image, a motion vector is shown in normal Cartesian image coordinates. In the top right is the vector's log-polar representation. In the bottom, is the magnitude of the FECT of the same vector.

and more robust approach is given by Wiener filtering. Following [9], we can express such filtering with the following relation:

$$F(h,k) = \frac{H^*(h,k)S_{uu}(h,k)}{|H(h,k)|^2 S_{uu}(h,k) + S_{\eta\eta}(h,k)} \quad (6)$$

Where $F(h,k)$ is the transfer function of the desired de-blurring kernel, $H(h,k)$ is the motion vector frequency representation (see fig.(4)), $S_{uu}(h,k)$ is the image power spectral density (PSD), and $S_{\eta\eta}(h,k)$ is the stationary uncorrelated noise PSD. The algorithm presented in this paper uses the frequency components estimated with the FECT. When the noise spectrum becomes small, eq.(6) can be well approximated by[3]:

$$F(h,k) \approx \frac{H^*(h,k)}{|H(h,k)|^2 + \epsilon} \quad (7)$$

The application of eq.(7) results in a net restoration of the original motion blurred (space-variant) images as shown in fig. 5. This figure shows the effect of motion blur for five different motion vectors (see fig. 4). In an actual active vision system, the FECT offers the advantages of speed but also of lower storage space, by reducing the number of motion vectors needed for de-blurring. This is because rotations in the original Cartesian domain correspond to circular shifts in the log-polar domain, see fig. 3. Circular shifts of complex images are not computationally demanding, especially due to the fact that the size of the images are reduced given the log-polar representation. Similarly, size scaling in the original Cartesian domain corresponds to linear shift in the log-polar domain, see fig. 3, which is also not computationally expensive. Because of these symmetry properties of the FECT, only one (transformed) motion vector image needs to be stored in memory.

3. Summary and Conclusions

This paper describes a Wiener filter approach for restoration of images degraded by motion blur. This approach significantly differs from the traditional algorithm [9], since we are using a space-variant image-format. This representation, i.e. the log-polar mapping, allows a great reduction (i.e. decimation) of the number of pixels used, maintaining the local high-resolution of the original image and its original field of view.

The FECT transforms log-polar images into frequency space[6]. In the present paper, the FECT

and inverse FECT (IFECT) have been used in order to apply Wiener filtering. The very low-complexity of the procedure as well as the simplicity of the algorithm described are the main advantages over the traditional methods. Yet another important aspect is the low storage of motion vectors needed[4]. The method presented is therefore potentially of great interest for practical application in active vision systems.

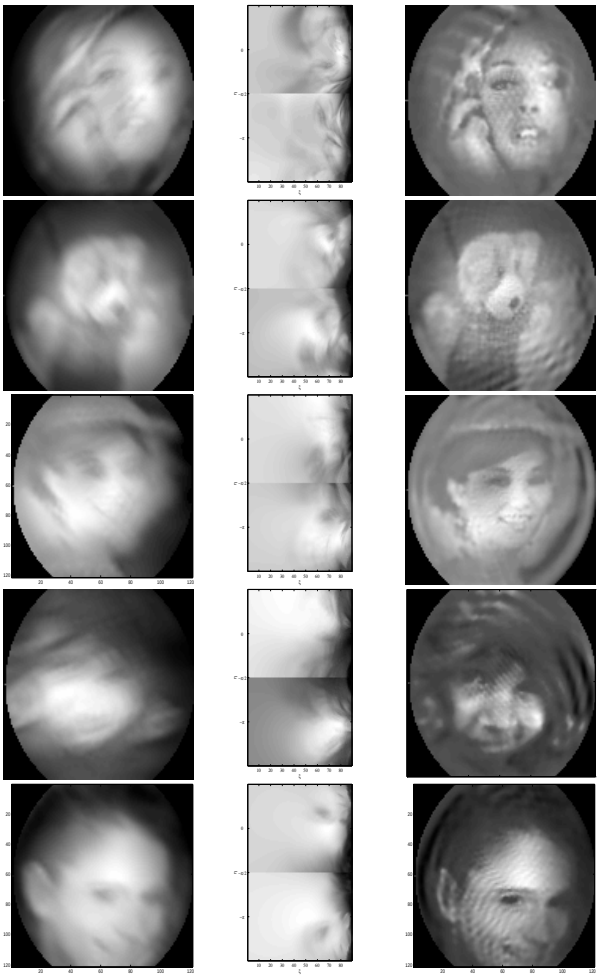

Fig. 5. Five examples of the application of Wiener filtering. In the first column are the motion blurred space-variant images uniformly sampled, i.e. by taking the logmap of the image and applying an inverse-logmap algorithm (cubic spline interpolation). In the second column is shown their log-polar map representation. Finally in the third column, is the result of the application of Wiener filtering, the images are then sharper and without motion blur. The five rows represent the effect blurring of five different motion vectors: woman $(45°, 10)$, teddy $(-45°, 10)$, woman $(135°, 14)$, toddler $(200°, 14)$, and woman $(-44°, 10)$. The ringing artifact in the deblurred images on the right is due to the overshoot of the motion vector in the FECT space. We will remove this artifact by in the very near future.

Acknowledgments

The present study was supported by grants (DARPA-ONR) N00014-96-C-0178 and (ONR-MURI) N00014-95-1-0409.

Notes

1. The log-polar images are here represented with a log-radius on the abscissa and a polar representation on the ordinate.
2. It is importante to notice that this time N_1 and N_2 are the number of vertical and horizontal points of a log-polar image representation, these number are usually much smaller than the dimensions of the original image.
3. Considering ϵ as a small number
4. The motion vectors can be computed directly, but this is not convenient if a system which operates in *real-time* is needed.

References

1. Nadia Baaziz and Claude Labit. Multiconstraint wiener-based motion compensation using wavelet pyramids. *IEEE Transactions on image processing.*, 3(5):688–692, May 1994.
2. Aijza A. Baloch, Allen M. Waxman, Aijza A. Baloch, and Allen M. Waxman. Visual learning: adaptive expections, and behavioural conditioning of the mobile robot MAVIN. *Neural Networks*, 4:271–302, 1991.
3. T. Baron, M. D¿ Levine, V. Hayward, and M. Bolduc nand D. Grant. A biologically motivated robot eye system. *Proc. of the 8th Canadian Aeronautics and Space Insitute Conference on Astronautics, Ottawa, Ontario.*, 8:231–240, 1995.
4. Benjamin Bederson, Richard S. Wallace, and Eric L. Schwartz. A miniaturized active vision system. In *11th IAPR International Conference on Pattern Recognition*, volume B of *Specialty Conference on Pattern Recognition Hardware Architecture*, pages 58–62, The Hague, Netherlands, August 1992.
5. Andrew Blake and Alan Yuille. *Active Vision*. Artificial Intelligence. MIT Press, Cambridge, Mass., 1992.
6. Giorgio Bonmassar and Eric Schwartz. Space-variant fourier analysis: the exponential chirp transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(10):1080–1090, October 1997.
7. Gil Engel, Douglas Greve, Joe Lubin, and Eric Schwartz. Space-variant active vision and visually guided robotics: Design and construction of a high-performance miniature vehicle. In *ICPR Proceedings*, ICPR-12. International Conference on Pattern Recognition, 1994.
8. D. A. Fish, J. Grochmalicki, and E. R. Pike. Scanning singular-value-decomposition method for restoring of images with space-variant blur. *J. Opt. Soc. Am. A*, 13(3):464–469, March 1996.
9. Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, 1989.
10. E. P. Krotkov. *Active Computer Vision by Cooperative Focussing and Stereo*. Springer-Verlag, NY, 1989.
11. Neri Merhav and Vasudev Bhaskaran. Fast algorithms for dct-domain image down-sampling and for inverse motion compensation. *IEEE transactions on circuits and systems for video technology*, 7(3):468–476, June 1997.

12. Kun min Yang, Ming ting Sun, and Lancelot Wu. A family of vlsi designs for the motion compensation block-matching algorithm. *IEEE transactions on circuits and systems*, 36(10):1317–1325, October 1989.

13. D. W. Murray, K. J. Bradshaw, P. F. McLauchlan, I. D. Reid, and P. M. Sharkey. Driving saccade to pursuit using image motion. *International Journal of Computer Vision*, 16:205–228, 1995.

14. A. S. Rojer and E. L. Schwartz. Design considerations for a space-variant visual sensor with complex-logarithmic geometry. *10th International Conference on Pattern Recognition, Vol. 2*, pages 278–285, 1990.

15. H. P. Rotstein and E. Rivlin. Optimal servoing for active foveated vision. In *Computer Vision and Pattern Recognition*, pages 177–182, San Francisco, California, June 18-20 1996.

16. Giuolio Sandini and Paolo Dario. Active vision based on space-variant sensing. *Intl. Symp. on Robotics Research*, 1989.

17. Eric Schwartz, Douglas Greve, and Giorgio Bonmassar. Space-variant active vision: Definition, overview and examples. *Neural Networks*, 8(7/8):1297–1308, 1995.

18. Eric L. Schwartz. Computational studies of the spatial architecture of primate visual cortex:columns, maps, and protomaps. In Alan Peters and Kathy Rocklund, editors, *Primary Visual Cortex in Primates*, volume 10 of *Cerebral Cortex*. Plenum Press, 1994.

19. Eric L. Schwartz. Space-variant active vision and visually guided robot vehicles. In *Proceedings of Government Microelectronics Conference*. GOMAC, 1995.

20. Kazuhito Suguri, Toshiro Minami, Hiroaki Matsuda, Ritsu Kusaba, Toshio Kondo, Ryota Kasai, Takumi Watanabe, Hidenori Sato, Nobutarou Shibata, Yakata Tashiro, Takaaki Izouka, Atsushi Shimizu, and Hiroshi Kotera. A real-time motion estimation and compensation lsi with wide search range from mpeg2 video encoding. *IEEE journal of solid-state circuits*, 31(11):1733–1739, November 1996.

21. Carl F. R. Weiman. Tracking algorithms using log-polar mapped image coordinates. *SPIE Proceedings on Intelligent Robots and Computer Vision VIII*, 1192, 1989.