1998-02

Real-Time Restoration of Images Degraded by Uniform Motion Blur in Foveal Active Vision Systems

Real-Time Restoration of Images Degraded by Uniform Motion Blur in Foveal Active Vision Systems

Giorgio Bonmassar and Eric L. Schwartz

February 1998

Technical Report CAS/CNS-98-008

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
677 Beacon Street
Boston, MA 02215

Real-Time Restoration of Images Degraded by Uniform Motion Blur in Foveal Active Vision Systems.

Giorgio Bonmassar, Eric L. Schwartz

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. Recently we have 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 uniform camera motion.

I. Introduction

Active vision systems in practical contexts must contend with image blur caused by camera and object motion. This is especially true of systems used on mobile robotic or vehicle based platforms. Image restoration is a computationally intensive process which is problematic in real-time systems. Partly because of these observations, several VLSI chips have recently been introduced for motion compensation [1], [2]. 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.

Recent algorithms to compensate for motion blur have been based on fast transforms such as the DCT (Discrete Cosine Transform) [3]. These suffer from the computational complexity of transforming and inverse transforming the original full image. Such operations cannot be performed in real-time on a typical workstation. The same limitations affect methods based on the pseudo-inverse, SVD [4], Wiener filtering methods [5] and Wavelet based methods [6].

The present paper is based on the use of a different image format, the log-polar format, shown in Fig. 1. The advantage of this type of non-uniform sampling [7] is that it allows a *local* high-resolution and a large field of view but with a small number of sample points compared to the original image. Such a format is favorable for active-vision tasks such as object tracking [8], [9], [10]. Image formats based on the log-polar transformation are a continuous version (in different coordinates) of the more familiar image pyramid format. Until recently, the pyramidal format has been the multi-resolution architecture of choice, due in part to the general lack of image processing tools to be applied to the log-polar sampling.

This situation has changed recently with the introduction of a new form of Fourier Transform, which allows a fast algorithm comparable in complexity to the 2D FFT to be applied to foveal image formats such as that provided by the log-polar transform. This algorithm, the fast Exponential Chirp Transform (FECT) [11], allows the use of traditional spatial frequency based methods to be applied to foveal log-polar formats. 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 perform real-time, full-frame motion de-blur of space-variant images using minimal storage and conventional (e.g. current generation Pentium) hardware.

II. Previous Work

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

1. It provides a wide-angle yet high resolution (i.e. foveal) image format with attractive space-complexity. The compression (or more accurately decimation) relative to a comparable space-invariant format is up to two or three orders of magnitude for currently realizable machine vision systems [12].
2. It provides a continuum model for variable resolution, or foveating pyramid architectures in machine vision, which is an alternative to the dyadically sampled (e.g. 5 level) pyramid architecture familiar in machine vision.

We now briefly summarize the Exponential Chirp Transform. Full details are provided in [11]. Real-time vision is an obvious area of application for this architecture, which has been studied by several different laboratories in recent years [13], [14], [15], [16], [17]. The log-polar mapping ($w(\xi,\eta) = \log(z(x,y)+a), z(x,y) =$

G. Bonmassar is with the Dept. of Cognitive and Neural Systems, Boston University, Boston, MA 02215, USA. E-mail: giorgio@engc.bu.edu

E. L. Schwartz is with the Dept. of Cognitive and Neural Systems, Boston University, Boston, MA 02146, USA. E-mail: eric@thing4.bu.edu

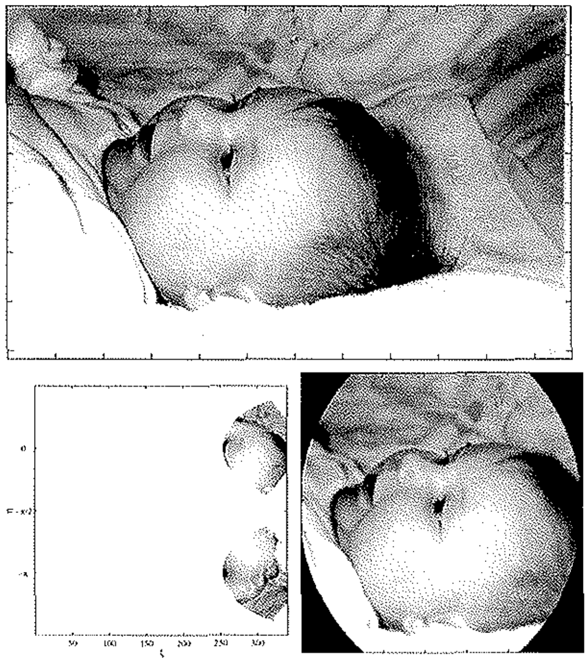

Fig. 1. Images of Caterina. At the top is a traditional uniformly sampled frame (588 × 358), in the bottom left is its log-polar representation (358 × 340, $a = 56.98$ and $R = 170$) and in the bottom right is the space-variant image in uniformly sampled coordinate space (340 × 340). This last image 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.

$x + iy$) can be defined as [1]:

$$\begin{aligned} \xi &= \log \sqrt{(x+a)^2 + y^2} \\ \eta &= \arctan \frac{y}{x+a} \end{aligned} \tag{1}$$

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

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

where $|J(\xi,\eta)|$ represents the determinant of 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:

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

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

$$\begin{aligned} k(\gamma,\theta) &= e^{\gamma} \cos(\theta) \\ h(\gamma,\theta) &= -e^{\gamma} \sin(\theta) \end{aligned} \tag{4}$$

[1]This represents a mapping of the image plane (x, y) to a log-polar domain (ξ, η) in which the singularity at the origin is removed by defining two separate branches, using some finite and positive "a" to provide a linear map for $\|x\| << a$ and becomes smoothly logarithmic for $\|x\| >> a$. Furthermore we are presenting the ECT for just one hemifield ($x \ge 0$), the computation of the other hemifield is straightforward.

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

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

$$e^{-2\pi j e^{(\gamma+\xi)}\cos(\theta+\eta)}\, d\xi\, d\eta \tag{5}$$

We refer to the numerical implementation of eq.(5) as the FECT. The details of the FECT can be found in [11]. A very important property of eq.(5), is the simplification of rotation and size transforms into shifts. This occurs because, by introducing polar coordinates:

$$\begin{array}{ll} x = \rho\cos(\theta) & y = \rho\sin(\theta) \\ k = \Omega\cos(\Theta) & h = \Omega\cos(\Theta) \end{array} \tag{6}$$

and by direct substitution of eq.(6) in the expression of the Fourier transform, we can write:

$$f(\rho,\theta+\theta_0) \Leftrightarrow F(\Omega,\Theta+\theta_0) \tag{7}$$

where $F(\Omega,\Theta)$ is the Fourier transform of $f(\rho,\theta)$ in polar coordinates. Therefore rotation by an angle θ_0 in spatial coordinates corresponds to rotation in frequency coordinates of the same angle θ_0. Furthermore, for the scaling property:

$$f(c_x\cdot x, c_y\cdot y) \quad \Leftrightarrow \quad \frac{1}{|c_x\cdot c_y|}F\left(\frac{k}{c_x},\frac{h}{c_y}\right) \tag{8}$$

scaling in frequency maps into inverse of scaling in space. Given that the images considered are real (i.e. the Fourier Transform is Hermitian), both rotation and size scale in frequency are centered with respect to the origin. Applying eq.(1) to eq.(8) and eq.(6), it is easy to show that rotations and scalings are mapped into simple shifts.

III. 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 and more robust approach is given by Wiener filtering. Following [5], we can express such filtering with the following relation:

$$F(h,k) = \frac{H^*(h,k)S_{uu}(h,k)}{|H(h,k)|^2\, S_{uu}(h,k) + S_{\eta\eta}(h,k)} \tag{9}$$

$F(h,k)$ is the transfer function of the desired de-blurring kernel, $H(h,k)$ is the motion vector frequency representation (see Fig. 2); $S_{uu}(h,k)$ 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.(9) can be well approximated by[2]:

$$F(h,k) \approx \frac{H^*(h,k)}{|H(h,k)|^2 + \epsilon} \tag{10}$$

The application of eq.(10) results in a net restoration of the original motion blurred (space-variant) images as shown in Fig. 3. This figure shows the effect of motion blur for five different motion vectors (see Fig. 2). 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. 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, 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.

IV. Motion vector estimation.

The primary difficulty with blind image restoration, or joint blur identification, is the estimation of the source of blur. Although many recent contributions [18], [19] are focused on space-variant motion[3] identification, we are here mainly interested in estimating the motion vector of a uniform image motion. An active camera system assumes constant velocity, in first approximation, for high shutter speed. Furthermore, the

[2] Considering ϵ as a small number

[3] A space-variant motion vector describes independent motion of multiple objects in a scene.

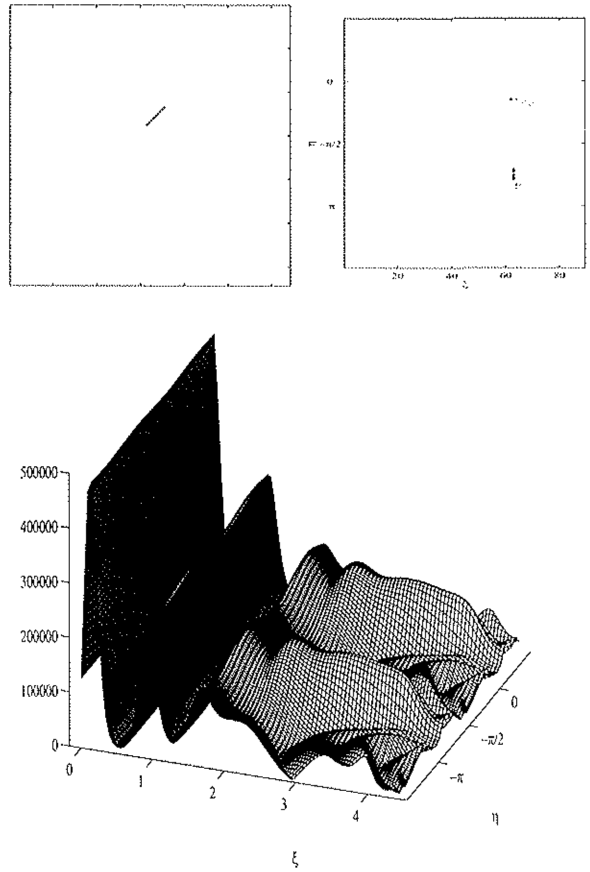

Fig. 2. An example of a motion vector. In the top left image, a motion vector is shown in normal Cartesian image coordinates. In the top right is the vector's log-polar representation. The bottom shows the magnitude of the FECT of the same vector.

scene is not expected to change its luminosity from frame to frame. Mathematically, the problem can be expressed in the following way[4]:

$$f_1(x, y) = s(x - d_x, y - d_y) + w_1(x, y) \tag{11}$$

$$f_2(x, y) = s(x, y) + w_2(x, y) \tag{12}$$

where $f_1(x, y)$ and $f_2(x, y)$ are the noisy representations of the actual frame and the previous one, d_x and d_y are the components of the displacement vector, and $w_1(x, y)$ and $w_2(x, y)$ is the Gaussian zero-mean additive white noise. In reality, each frame is affected by motion blur:

$$f(x, y) = s(x, y) * d(x, y) + w(x, y) \tag{13}$$

where "$*$" stands for convolution, and $d(x, y)$ is the Point Spread Function (PSF) of the blur operator. In practice $d(x, y)$ is related to the components of the displacement vector (d_x, d_y): during the camera motion the sensor's image points are integrating the light intensity, over the shutter opening time. The length and orientation of $d(x, y)$ is estimated by taking advantage of the properties of the log-polar sampling in frequency of the FECT. According to such properties, scalings or rotations are mapped into shifts.

A reference blur operator $d_0(x, y)$ is considered, and the correlation algorithm is used to estimate the desired point spread function $d(x, y)$:

$$c(x, y) = \log\left(\left|\tilde{F}(h, k)\right|\right) \log\left(\left|\tilde{D}_0(h, k)\right|\right) \tag{14}$$

where $\tilde{F}(h, k)$ is the FECT of the frame considered and $\tilde{D}_0(h, k)$ is the FECT of the reference blur operator. In frequency, the convolution operator in eq.(13), is a simple multiplication operation, and logarithm transforms this multiplication into a sum. Correlation known, to be very robust even in the presence of additive

[4]For an overview of the subject of standard motion estimation techniques, see [20].

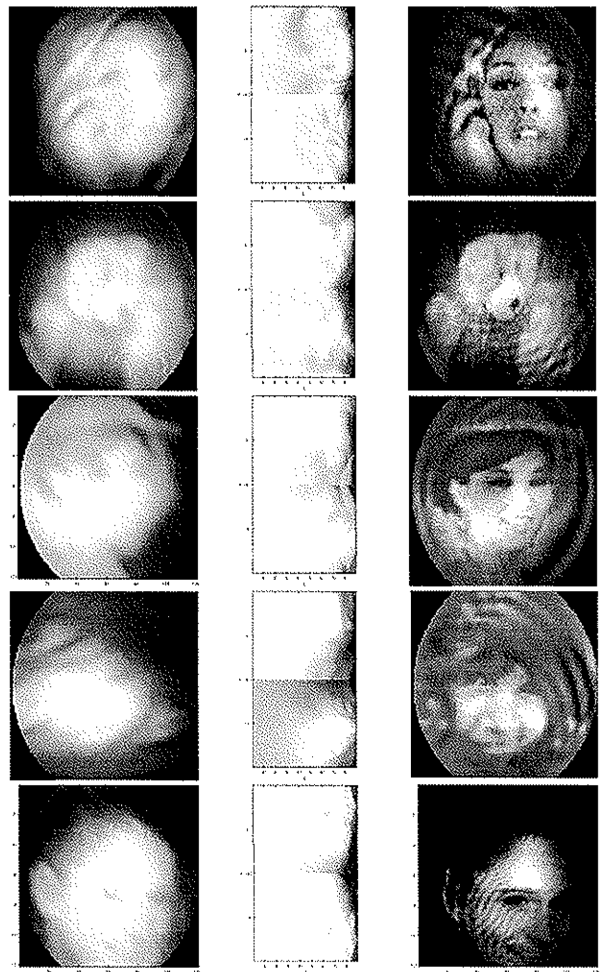

Fig. 3. Five examples of the application of Wiener filtering. In the first column are the motion blurred space-variant images uniformly sampled, i.e. by taking the logmap of the image and applying an inverse-logmap algorithm (cubic spline interpolation). In the second column is shown their log-polar map representation. Finally, in the third column is the result of the application of Wiener filtering: the images are 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 in the very near future.

noise [21], is used to find the position of $\log\left(\left|\tilde{D}_0(h,k)\right|\right)$ in $\log\left(\left|\tilde{F}(h,k)\right|\right)$. Figure 4 illustrates the method discussed so far. From left to right: the logarithm of the absolute value of the FECT of the original image (a mandrill), the logarithm of the absolute value of the FECT of a blurred image, the logarithm of the absolute value of the FECT of the reference blur operator, and finally the cross-correlation in eq.(14). This last plot has angles in the horizontal axis and logarithm of scales in the vertical, returning the orientation and scaling which $d(x, y)$ has with respect to $d_0(x, y)$.

V. 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 [5], since we are using a space-variant image representation. 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 [11]. In the present paper, the FECT and inverse FECT (IFECT) have been used for Wiener filtering. The main advantage over traditional methods is the very low complexity of the procedure. Yet another important aspect is the low storage of motion vectors

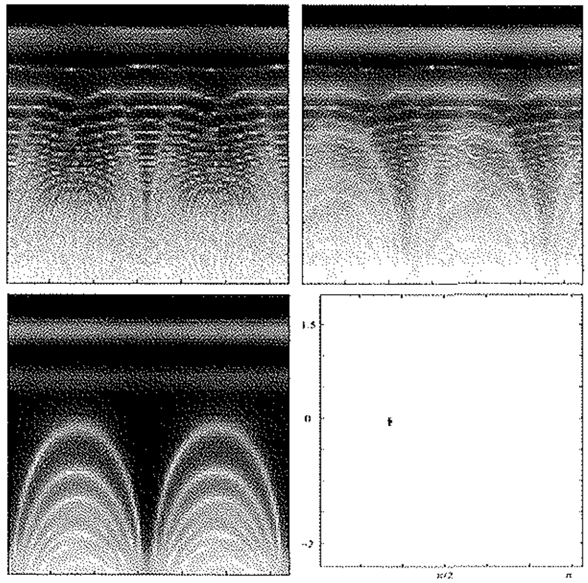

Fig. 4. Example which shows how the estimation of a motion vector is estimated: the logarithm of the absolute value of the FECT of the image of a mandrill (top left), the image of a blurred version of the mandrill (top right), the motion vector taken as reference (bottom left), and the cross-correlation plane (bottom right) where the motion plane is estimated. In this plane it is easy to notice that the peak value is located at 45° angle and -0.7 log-scale factor (a scaling of $\frac{1}{2}$).

needed[5]. The method presented is therefore of interest for practical application in active vision systems.

Acknowledgments

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

References

[1] 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*, vol. 31, no. 11, pp. 1733–1739, Nov. 1996.

[2] 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*, vol. 36, no. 10, pp. 1317–1325, Oct. 1989.

[3] 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*, vol. 7, no. 3, pp. 468–476, June 1997.

[4] 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*, vol. 13, no. 3, pp. 464–469, Mar. 1996.

[5] Anil K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1989.

[6] Nadia Baaziz and Claude Labit, "Multiconstraint wiener-based motion compensation using wavelet pyramids.," *IEEE Transactions on image processing.*, vol. 3, no. 5, pp. 688–692, May 1994.

[7] A.J. Jerry, "The shannon sampling theorem-its various extensions and applications: a tutorial review," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.

[8] Eric L. Schwartz, "Space-variant active vision and visually guided robot vehicles," in *Proceedings of Government Microelectronics Conference*. GOMAC, 1995.

[9] H. P. Rotstein and E. Rivlin, "Optimal servoing for active foveated vision," in *Computer Vision and Pattern Recognition*, San Francisco, California, June 18-20 1996, pp. 177–182.

[10] 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*, vol. 16, pp. 205–228, 1995.

[5] The motion vectors can be computed directly, but this is not convenient if a system which operates in *real-time* is needed.

[11] Giorgio Bonmassar and Eric Schwartz, "Space-variant fourier analysis: the exponential chirp transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 10, pp. 1080–1090, Oct. 1997.

[12] Eric Schwartz, Douglas Greve, and Giorgio Bonmassar, "Space-variant active vision: Definition, overview and examples," *Neural Networks*, vol. 8, no. 7/8, pp. 1297–1308, 1995.

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

[14] 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*, vol. 4, pp. 271–302, 1991.

[15] Benjamin Bederson, Richard S. Wallace, and Eric L. Schwartz, "A miniaturized active vision system," in *11th IAPR International Conference on Pattern Recognition*, The Hague, Netherlands, August 1992, vol. B of *Specialty Conference on Pattern Recognition Hardware Architecture*, pp. 58–62.

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

[17] 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.*, vol. 8, pp. 231–240, 1995.

[18] H. J. Trussell and S. Fogel, "Identification and restoration of spatially variant motion blurs in sequential images.," *IEEE transactions on image processing*, vol. 1, no. 1, pp. 123–126, Jan. 1992.

[19] Nader M. Namazi and David W. Foxall, "On the convergence of the generalized maximum likelihood algorithm for nonuniform image motion estimation.," *IEEE Transactions on image processing.*, vol. 1, no. 1, pp. 116–119, Jan. 1992.

[20] Hans Georg Musmann, Peter Pirsch, and Hans-Joachim Grallert, "Advances in picture coding," *Proceeding of the IEEE*, vol. 73, no. 4, pp. 523–548, Apr. 1985.

[21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.