Fourier analysis and cortical architectures: the exponential chirp transform *

Giorgio Bonmassar †
†Dept. Biomedical Engineering
Boston University
Boston, MA 02146
e-mail: giorgio@enga.bu.edu

Eric L. Schwartz ‡
‡Dept. Cognitive and Neural Systems
Boston University
Boston, MA 02146
e-mail: eric@thing4.bu.edu

Abstract

The use of visual representations in which pixel-size and local neighborhood topology are not constant is termed space-variant vision. This is the dominant visual architecture in all higher vertebrate visual systems, and is coming to play an important role in real-time active vision applications in the form of log-polar, foveating pyramid, and related approaches to machine vision.

The breaking of translation symmetry that is unavoidably associated with space-variant vision presents a major algorithmic complication for image processing. In this paper we use a Lie group approach to derive a kernel which provides a generalization of the Fourier Transform that provides a quasi-shift invariant [1] template matching capability in the distorted (range) coordinates of the space-variant mapping. We work out the special case of the log-polar mapping, which is the principle space-variant mapping in use; in this case, we call the associated integral transform the "exponential chirp

*Work supported by ARPA ANNT-ONR N00014-92-C-0119 and ONR MURI N60014-95-I-0409

[1] We use the term quasi-shift invariant to refer to the nonuniform sampling nature of space-variant maps like the log-polar mapping. Thus, for example, if an object is shifted across the aperture of a space-variant system, our methods will produce invariance up to, but not including, the application of a band-pass filter that is position dependent. This will be made clear in the text.

transform" (ECT). The method is, however, general for other forms of mapping, or warp, function.

Examples from the two-dimensional (image processing) log-polar transformation are presented along with the demonstration that the ECT preserves the foveating aspect of the space domain mapping, and therefore provides a quasi-shift-invariant realization for the applications of matched filter and phase-only filter. This work provides, for the first time, a conceptual basis for combining global spatial frequency methods with space-variant mappings in a way which is consistent with the anatomical fact that human vision, at the cortical level, takes place in log-polar coordinates.

Categories: Biological vision, real-time active vision, Low-Level Processing, Shape and Object Representation.

1 Introduction

This paper addresses a fundamental difficulty in performing frequency domain image processing on image architectures which are strongly space-variant, and, in particular, are described by the log-polar map, or one of its variants. 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 primate and human visual cortex (see (Schwartz, 1994) 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 continuum model for variable resolution, or foveating pyramid architectures in machine vision.

3. It provides a wide-angle yet high resolution (i.e. foveal) image format with attractive space-complexity. The compression relative to a comparable space-invariant format is up to four orders of magnitude for human vision (Rojer and Schwartz, 1990), and up to two or three orders of magnitude for currently realizable machine vision systems (Schwartz et al., 1995).

Real-time vision is an obvious area of application for this architecture, as the vertebrate visual system seems to suggest. Recently, a number of research groups have constructed machine vision systems based on this architecture (Weiman, 1989; Baloch and Waxman, 1991; Bederson et al., 1992; Engel et al., 1994; Sandini and Dario, 1989; Baron et al., 1995). However, application in machine vision has been impeded by the general lack of suitable image processing algorithms for the difficult application domain presented by the log-polar mapping. The present paper provides a possible solution to this quandary.

The specific problem addressed in this paper is to develop a generalized form of Fourier Transform which may be applied in the space-variant range coordinates of an image warp. The method used here is based on the use of Lie Group methods to formulate and then solve a specific partial differential equation (PDE), following the work of (Ferraro and Caelli, 1988) and (Rubinstein et al., 1991), who used these methods in the context of recognition of distorted patterns. In the present case, the solution of the PDE determined by the Lie Group approach produces a space-variant kernel which provides a shift-invariant property analogous to that of the conventional Fourier transform. Using the kernel that we have derived, which we call the exponential chirp kernel, it is possible to compute an integral transform with an image in, for example, log-polar coordinate form, and yet retain the advantageous shift (and also size and rotation) invariances associated with conventional Fourier analysis.

The log-polar (or more accurately, complex logarithmic) map is defined as follows:

$$w \;=\; K \log(z \;+\; a) \tag{1}$$

In this equation, K is an experimental constant which is only of relevance to the biological scale factor of a particular map, and will be dropped in the following discussion (see (Schwartz, 1994) for a review of estimates of this parameter), "a" a real constant (Rojer and Schwartz, 1990) that deals with the singularity on the origin, "z" represents visual pixel coordinates, and "w" represents log-polar coordinates. In the context of image processing, the log-polar coordinate transform must be adjoined to an algorithm for image warping, which expresses the space-variance of pixel size, as well

Figure 1: Space-invariant image.

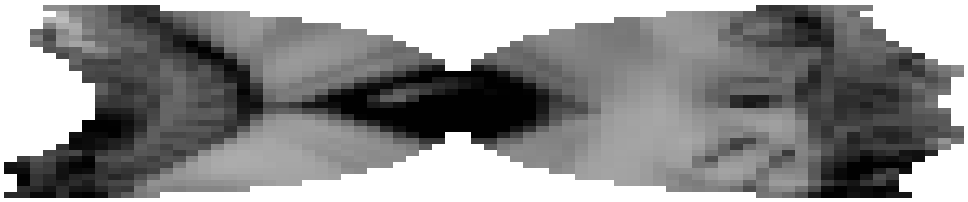

Figure 2: Log polar space-variant video frame, fixation point on left eye.

as pixel location. This can be stated as follows: We associate each pixel W from the range of the log-polar warp a set of domain pixels $f^{-1}(W)\epsilon Z$, where Z represents the set of domain pixels which would come, commonly, from a conventional t.v. image. Formally, using the notation χz to represent the location of a domain pixel,

$$f^{-1}(W) = Z \mid \log(\chi z + a)\epsilon W \tag{2}$$

Thus, each pixel W represents the support of a small group of pixels in the domain, whose size increases with increasing distance from the origin. The location of these pixels is expressed by the complex logarithmic coordinate change of eq.(1).

Figure 1 shows a conventional video frame.

Figure 2 shows a log polar mapping of Figure 1.

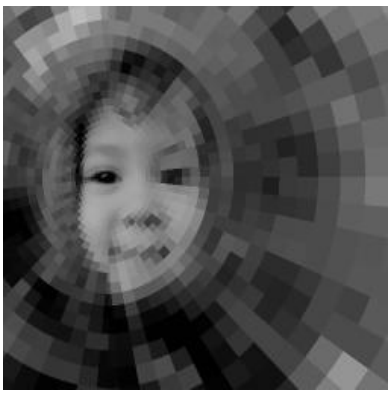

Figure 3: Inverse log polar space-variant video frame.

The log image has 2000 pixels, while the source image has roughly 256,000, yielding a compression of more than two orders of magnitude. Given the ability to point the sensor (i.e. active vision), and appropriate attentional algorithms to determine where to move the sensor, it is clear that this architecture could be extremely useful. Yet, the size and the "shape" of image features changes radically as the "fixation" point is moved, as is evident in the figure. In other words, the good news of SV vision is that enormous reductions in space-complexity can be achieved. The bad news is that even the simplest image processing tasks can become extremely difficult.

Historically, there have been three approaches to this problem. First, one might work in the inverse mapped log-polar image, as shown in Figure 3. This is the approach taken in machine vision in the context of "pyramid" algorithms. The foveating pyramid is a discrete approximation to the "retina" as in Figure 3. But, the number of pixels in this image is equal to that of Figure 1 (when represented in a device or memory with conventional constant pixel architecture) [2], therefore sacrificing one of the principle motivations of the method. Second, one might precede the log-polar mapping with a Fourier transform, i.e. use a Mellin-Fourier transform,

[2]In a full pyramid architecture, the number of pixels is 4/3 N, where N is the number of pixels in the original image (Burt and Adelson, 1981), while in a "foveating pyramid" (Burt, 1988) the number of pixels can approximate the small number used in the log-polar representation.

as has been pointed out during the past three decades, e.g. (Brousil and Smith, 1967; Casasent and Psaltis, 1976; Schwartz, 1977; Sheng and Arsenault, 1986). Third, one may use the connectivity graph of the transformed domain (Wallace et al., 1994). This approach cannot be generalized for frequency domain applications.

The Mellin-Fourier transform is equivalent to applying the operations of Fourier transform, log polar mapping, Fourier transform to the original image. The first Fourier transform is, up to a phase, translation invariant. The log-polar transform provides size and rotation invariance, up to a spatial shift (see (Schwartz, 1977)) and the final Fourier transform reduces this shift to a phase. The Mellin transform has generated a fair amount of attention as an example of a geometrically invariant integral transform method, and has been suggested as a basis for biological vision (Cavanagh, 1978), although this must be rejected (Schwartz, 1981) since it contradicts the fact that human vision is strongly space-variant.

At this point, it is important to point out a fact which has not always been emphasized. Although the Mellin transform technique does solve the problem of working in a space-variant domain such as the log-polar map, it does so, by, in a sense, "throwing the baby out with the bath water." The initial Fourier transform, which nullifies the unpleasant space-variance of the log-polar map, also removes one of the the main advantages for which the log-polar map was introduced (see (Schwartz, 1981)): the log-polar mapping used in the Mellin transform is in frequency space, and thus, rather than having a "foveating" or space-variant vision architecture, the Mellin transform provides a "fovea" in frequency space. The practical disadvantage of this is that it requires two-dimensional (FFT) transform of the full rank image. This is computationally unfavorable for many real-time vision applications, and obviates the space-complexity advantages outlined above for the log-polar mapping. Furthermore, the performance of the Mellin Transform in machine vision is compromised by its space-invariance: small targets have small signal to noise-ratio in the Mellin plane, and there is no possibility of "foveation" to improve this problem. As a result, the Mellin transform tends to perform very poorly in real imaging applications.

In the present paper, we derive, using Lie Group methods, the exponential chirp transform (ECT), which retains the favorable translation symmetry of the Mellin-Fourier approach, but does not require a transfor-

mation back into the original image plane, and, moreover, provides the possibility of "foveation", i.e. provides a space-variant generalization of the Mellin transform.

As outlined in eq.(2), we seek a form of space-variant kernel, which, when used in an integral transform with a space-variant image such as produced by the log-polar map, provides an analogue of the usual shift property that the Fourier transform provides in conventional space-invariant imaging. Specifically, we wish our kernel to have the property that the integral transform with the image (in the range coordinate system of the warped image) changes only by a phase when a shift is applied in the **domain** coordinates. Such a kernel would allow us to generalize Fourier Analysis to space-variant mappings, and would allow an efficient computation of the integral transform, since our range coordinate system has a very small number of pixels. We will now provide a brief review of Lie Group methods in partial differential equations, and we will then show a PDE system whose solution is the desired kernel. This method may be used in general to set up the PDE system which solves this problem for an arbitrary space-variant mapping. Then, we will demonstrate a solution to this PDE system for the particular case of the log-polar map. Finally, we show several examples in which we have applied template matching in a simple optical character recognition demonstration using this methodology, illustrating shift invariant template matching, performed with high efficiency in the log-polar image format.

2 Lie Group Theory

Consider the following invertible coordinate transformation in a two-dimensional space:

$$(\hat{x}, \hat{y}) = (\phi(x, y, \rho), \psi(x, y, \rho)) \tag{3}$$

If eq.(3) is a smooth one-dimensional manifold (e.g. curve parametrized by ρ) with smooth inverse map, such that:

$$(\phi(\hat{x}, \hat{y}, \rho_1), \psi(\hat{x}, \hat{y}, \rho_1)) = (\phi(x, y, \rho_2), \psi(x, y, \rho_2)) \tag{4}$$

where ρ_1 and ρ_2 are two parameters, then eq.(3) represents a one-parameter Lie group, parameterized by ρ; ρ_0 identifies the identity element (i.e. identity transformation):

$$(x, y) = (\phi(x, y, \rho_0), \psi(x, y, \rho_0)) \tag{5}$$

Lie Group theory provides a useful link between two very different areas of mathematics: group theory and differential-geometry.

We can express the coordinate transformation using the following Taylor's series, expanded at ρ_0:

$$(\hat{x}, \hat{y}) = (x, y) + (\alpha(x, y), \beta(x, y))\,[\rho - \rho_0] + \mathbf{o}([\rho - \rho_0]^2) \tag{6}$$

where

$$\begin{cases} \alpha(x, y) = \frac{\partial \phi(x,y,\rho)}{\partial \rho} \mid_{\rho=\rho_0} \\ \\ \beta(x, y) = \frac{\partial \psi(x,y,\rho)}{\partial \rho} \mid_{\rho=\rho_0} \end{cases} \tag{7}$$

Introducing the concept of *infinitesimal generator* (Ibragimov, 1994):

$$\mathbf{g} = \alpha(x, y)\frac{\partial}{\partial x} + \beta(x, y)\frac{\partial}{\partial y} \tag{8}$$

which can be used to generate a one-parameter group by direct solution of the differential equation expressed in eq.(7), taking eq.(5) as initial conditions. The elements of a Lie group may be also computed by the exponential map (see (Olver, 1993)): $\exp(\rho\mathbf{g})(x, y)$.

Particular attention has to be placed on the group's parameter, because it generates the generic element of the Lie group. In particular eq.(5) can be rewritten in this notation, where $\rho = 0$ determinates the identity element of the group:

$$\exp(0\,\mathbf{g})(x, y) = (x, y) \tag{9}$$

The infinitesimal generator $\mathbf{g}$ is:

$$\left.\frac{d}{d\rho}\right|_{\rho=0} [\exp(\rho \mathbf{g})(x,y)] = \mathbf{g}\,\Big|_{(x,y)} \tag{10}$$

Example 1 The generator $\mathbf{g_t} = \frac{\partial}{\partial x}$ generates the group of the x-axis translations:

$$(\hat{x}, \hat{y}) = (x + \rho, y)$$

Example 2 Considering the group of rotations in the plane:

$$(\hat{x}, \hat{y}) = (x\ \cos(\rho) - y\ \sin(\rho), x\ \sin(\rho) + y\ \cos(\rho))$$

according to eq.(7)

$$\begin{cases} \alpha(x,y) = \frac{\partial \phi(x,y,\rho)}{\partial \rho}\,|_{\rho=0} = -y \\ \\ \beta(x,y) = \frac{\partial \psi(x,y,\rho)}{\partial \rho}\,|_{\rho=0} = x \end{cases}$$

we obtain the infinitesimal generator $\mathbf{g_r}(x,y) = -y\,\frac{\partial}{\partial x} + x\,\frac{\partial}{\partial y}$

Example 3 The generator $\mathbf{g_d} = x\,\frac{\partial}{\partial x} + y\,\frac{\partial}{\partial y}$, using eq.(7) we obtain the following system of ordinary differential equations,

$$\begin{cases} \frac{\partial x}{\partial \rho} = x \\ \frac{\partial y}{\partial \rho} = y \end{cases}$$

which generates the group $(e^{\rho}\,x, e^{\rho}\,y)$ of exponential scalings.

3 The Exponential Chirp Transform (ECT)

We consider here the following two-dimensional log-polar (or complex logarithmic) transformation which is appropriate to model primate visual cortex (Rojer and Schwartz, 1990): $w = \log(z + a)$, where "a" is a definite positive parameter. [3] The transformation between spaces can therefore be written:

$$\begin{cases} \xi = \log \sqrt{(x+a)^2 + y^2} \\ \\ \eta = \arctan \dfrac{y}{x+a} \end{cases} \tag{11}$$

Our goal is to achieve "shift-invariance", like that of the usual Fourier transform, in the log-polar plane. Thus given an image in the domain of the polar-log mapping $s(\xi, \eta)$, we require a change of variables $(\xi, \eta) \rightarrow (f(\xi, \eta), g(\xi, \eta))$ so that translations in the original plane (x, y) will result only in phase factors for the following Fourier transform:

$$\iint_{\mathcal{D}} s(\xi, \eta)\, |J(\xi, \eta)| \exp[-2\,\pi\, j(f(\xi, \eta)\, k + g(\xi, \eta)\, h)]\, d\xi\, d\eta \tag{12}$$

where $|J(\xi, \eta)|$ is the determinant of the Jacobian of the transformation and $\mathcal{D} \equiv \{(\xi, \eta) : -\infty \leq \xi < +\infty$ and $\frac{-3\pi}{2} \leq \eta \leq \frac{\pi}{2}\}$. The following Lie group will enable us to find the two functions ($f(\xi, \eta)$ and $g(\xi, \eta)$) by solving a set of partial differential equations (PDE). These equations identify the group of symmetries with respect to simple translations in space. This analysis follows that of (Ferraro and Caelli, 1988; Rubinstein et al., 1991).

Theorem 1. (Olver, 1993) A given smooth function $f : M \rightarrow \Re$ is an *invariant* for a connected one-parameter group of transformations H acting on a manifold M (the condition invariance holds if $\forall(h, j)$ and $\forall(x, y) \in$

[3] The complex log transformation requires a branch cut, which is taken in this case to divide the plane into two parts ($Real(z) > 0$ and $Real(z) < 0$. Note that this is identical to (and in fact was motivated by) the anatomy of the brain: the two sides of this mapping are in direct correspondence with the two hemispheres of the brain. The visual cortex, which is of the form of a complex logarithmic mapping, is divided in this way for similar reasons. See (Schwartz, 1994) for discussion of the biological evidence for log-polar mapping in primate visual cortex.

$M : f((h,j)\cdot(x,y)) = f(x,y)$, where "$\cdot$" is the group operation) if and only if:

$$\mathbf{g}\left[f(x,y)\right] = 0 \tag{13}$$

where "$\mathbf{g}$" is the infinitesimal generator of H.
Proof. The necessity holds since, if f is an invariant then it is a constant for $\forall\rho$:

$$\frac{d}{d\rho}f(\exp(\rho\mathbf{g})(x,y)) = \left(\frac{\partial f}{\partial\hat{x}}\frac{\partial\hat{x}}{\partial\rho} + \frac{\partial f}{\partial\hat{y}}\frac{\partial\hat{y}}{\partial\rho}\right) = 0 \tag{14}$$

by using eq.(10) and eq.(7), with $\rho = 0$:

$$\frac{d}{d\rho}f(\exp(\rho\mathbf{g})(x,y)) = \mathbf{g}\left[f(x,y)\right] = 0 \tag{15}$$

which is clearly zero for $\forall(x,y) \in M$.

The condition is sufficient since eq.(13) holds for $\forall\rho$, then

$$\frac{d}{d\rho}f(\exp(\rho\mathbf{g})(x,y)) = 0 \tag{16}$$

which means that the function f is constant under the action of the one-parameter connected group $\exp(\rho\mathbf{g})$, where eq.(16) is defined. Then, by taking two points in M, (h,j) and (x,y): $f((h,j)\cdot(x,y)) = \text{const} = f(x,y)$, i.e. the invariance follows.□

A generic infinitesimal generator under a change of coordinates $(x = \phi(\xi,\eta), y = \psi(\xi,\eta))$ and inverse transformation $(\xi = \phi^{-1}(x,y), \eta = \psi^{-1}(x,y))$, has the form,

$$\begin{aligned}\mathbf{g} = &\left(\alpha(x,y)\tfrac{\partial\phi^{-1}(x,y)}{\partial x} + \beta(x,y)\tfrac{\partial\phi^{-1}(x,y)}{\partial y}\right)\Big|_{(x=\phi(\xi.\eta),y=\psi(\xi,\eta))}\tfrac{\partial}{\partial\xi} + \\ &\left(\alpha(x,y)\tfrac{\partial\psi^{-1}(x,y)}{\partial x} + \beta(x,y)\tfrac{\partial\psi^{-1}(x,y)}{\partial y}\right)\Big|_{(x=\phi(\xi.\eta),y=\psi(\xi,\eta))}\tfrac{\partial}{\partial\eta}\end{aligned} \tag{17}$$

where $\alpha(x,y)$ and $\beta(x,y)$ are the same as in eq.(8). The horizontal shift

infinitesimal generator (see example 1), under the logmap transformation is therefore,

$$\mathbf{g_t} = \exp(-\xi)\cos(\eta)\frac{\partial}{\partial \xi} - \exp(-\xi)\sin(\eta)\frac{\partial}{\partial \eta} \tag{18}$$

Using eq.(17) the infinitesimal generator for rotation (see example 2) is $\frac{\partial}{\partial \eta}$ and for exponential scaling is $\frac{\partial}{\partial \xi}$. The logmap transformation, therefore, changes the group exponential scalings and rotations into a group of simple horizontal and vertical translations. [4]
We are interested in "f" and "g" which satisfies the following system of partial differential equations, expressed by knowing that the generator of translations in $\Re^2$ is the vector $(1, 0)$ (see example 1 and eq.(13)[5]),

$$\begin{cases} \exp(-\xi)\cos(\eta)\frac{\partial f(\xi,\eta)}{\partial \xi} - \exp(-\xi)\sin(\eta)\frac{\partial f(\xi,\eta)}{\partial \eta} = 1 \\ \exp(-\xi)\cos(\eta)\frac{\partial g(\xi,\eta)}{\partial \xi} - \exp(-\xi)\sin(\eta)\frac{\partial g(\xi,\eta)}{\partial \eta} = 0 \end{cases} \tag{19}$$

It can be verified by direct substitution that the previous equation has a (particular) solution:

$$\begin{cases} f(\xi,\eta) = \exp(\xi)\cos(\eta) - a \\ g(\xi,\eta) = \exp(\xi)\sin(\eta) \end{cases} \tag{20}$$

The final result of this analysis is to write the explicit form of the exponential chirp transform (ECT), by combining eq.(20) and eq.(12),

$$\iint_{\mathcal{D}} s(\xi,\eta)\exp(2\xi)\exp[-2\pi j[k(\exp(\xi)\cos(\eta) - a) + l\ \exp(\xi)\sin(\eta)]\,d\xi\,d\eta \tag{21}$$

[4] It is important to note here that size and rotation symmetry is provided by the map $log(z + a)$ only in the limit that $a = 0$. Realistic models of primate anatomy require a small but finite a, so provide size and rotation scaling at regions larger than this value of a.

[5] It does not really matter if we use the vector $(1, 0)$ or $(0, 1)$, it will only differ by a transposition of the ECT matrix.

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

The ECT has the following properties:

1. Shift in the domain causes only a phase change in the complex amplitudes of the exponential chirp transform. This transform is done by using the exponential chirp kernel as in eq.(21), and the warped image, in the range coordinates.

2. This result may be extended to arbitrary mappings, by setting up the corresponding PDE, and solving it. In the present case, we have derived it for the class of log-polar mappings of the form $\log(z + a)$.

3. This mapping is correctly "foveating", i.e. space-variant, unlike the Mellin transform. It thus allows us to exploit the efficiency of computing in the typically small image sizes that result from space-variant imaging systems.

4 Results

We will now show some practical examples of application of the ECT, with the use of synthetic images of letters. Working in a distorted space, or in particular coping with a non-uniform sampling, requires special care in order to avoid aliasing. Basically both kernel and image have to be filtered before the actual integral transformation so that both would satisfy Shannon's sampling theorem. Regions near the fovea which are sampled with a higher sampling frequency require a less abrupt band-pass filtering than regions in the periphery which are sampled much more coarsely. Aliasing is avoided when the image is filtered with a position-dependent band-pass filter before sampling. The images of the letters in Figure 4 form an array of 256 by 256; the fovea is always centered in the center of the image with $R = 128$; the log-map images are 64 by 26 with parameter $a = 10.186$; therefore the polar log-map introduces a compression factor of approximately 40. The image of letters was initially filtered with an anti-aliasing non-uniform mean-filter. The image was divided into sectors with exponentially growing radius, and the polar-log representation was given by taking the average value in every sector.

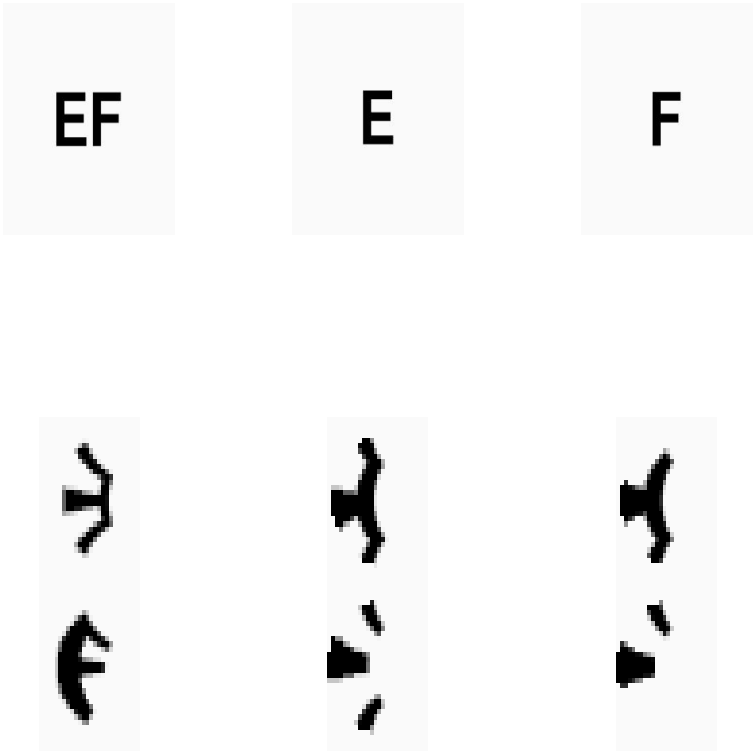

Figure 4: Image of two off-center letters (upper-left), and separately centered (upper-center and right) and their log-map representations (below).

Figure 5: Result of matching the letter E (left, the letter F in the right), using the exponential chirp transform together with a phase-only filter, in the presence of the "distractor" F (left, the letter E on the right).

The kernel of the ECT, in eq.(21), is clearly a complex function in which the frequency of the real and imaginary part grow exponentially with ξ. Therefore the kernel must be windowed to zero (on the (ξ, η) plane) when it reaches too high frequency according to the sampling used. This can be done using an ideal filter or by using one of the common windows (e.g. Bartlett, Hamming, Blackman, etc.), but in our examples we used ideal filters. The results in figure 5 show that the ECT is capable of locating a shifted copy of the letter 'F', even in the presence of the similar-looking letter 'E' and vice-versa. As we saw from the above examples, shift in the position correspond to changes in the phase of the ECT. A phase-only is used as a matching algorithm (see (**?**)). All computations are performed in log-polar coordinates. A second example shows the detection of a shifted letter F camouflaged by white noise, in which we show the robustness of the algorithm. The image plane is shown in Figure 6. A phase-only correlator, applied using the ECT for this problem, is shown in Figure 7. The frequency information present in the ECT, with uniform sampling in fre-

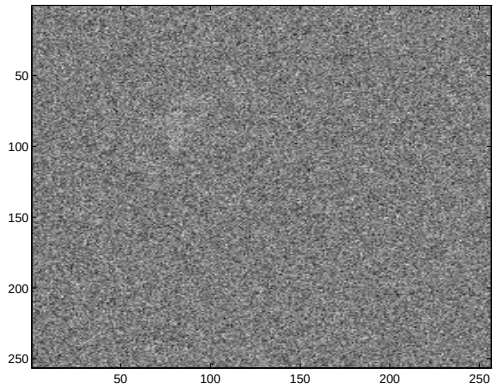

Figure 6: Letter F, shifted, in white noise.

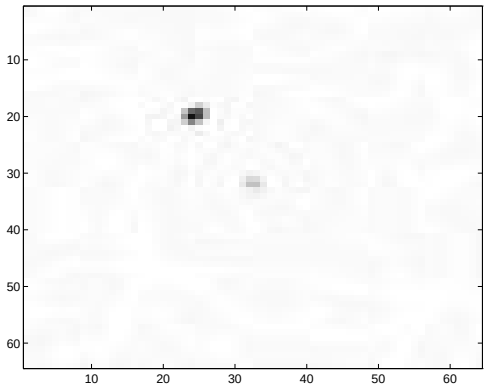

Figure 7: A phase-only applied to ECT of letter “F” in white noise.

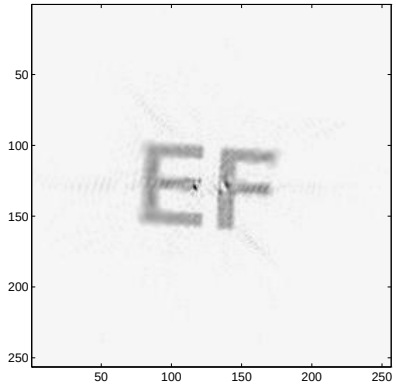

Figure 8: Example of inversion of the ECT by using the inverse Fourier transform.

quency, can be extracted reconstructing the image using the inverse FFT, as shown in figure 8.

5 Conclusions

In the present paper, we have shown the solution on how to combine the space-variant imaging properties of a mapping, such as the complex logarithm, which plays a major role both in biological vision and in real-time active vision applications, with the properties of the Fourier transform. We have used Lie Group methods to set up a partial differential equation whose solution is the kernel which gives us the desired invariance properties, in the range coordinates of the image warp.

In other work, (**?**), we have developed a highly optimized algorithm that allows computation of the ECT with complexity $O(N^2 \log(N))$ (i.e. the same as the FFT). From the point of view of real-time vision, this allows us to combine the space-efficiency of the log-polar mapping with the convenience of frequency domain template matching and image processing. We have estimated that the ECT can be performed at about 50-100 transforms/second, using a single DSP processor (e.g. TI320C40). This provides the possibility of performing shift, size and rotation[6] invariant

[6]Size and rotation invariance is achieved with the ECT by using a logmap in frequency

template matching in low-cost machine vision systems at rates which are comparable to optically based template matching! The practical importance of this work is threefold:

- A Space-Variant Generalization of the Mellin transform is described, which allows efficient image processing operations to be computed on log-polar image formats. This provides, for the first time, a possible reconciliation of the strongly space-variant nature of biological vision with the properties of global Fourier analysis.
- Lie Group methods are outlined which provide a method of generalizing this result to arbitrary space-variant mappings.
- The difficulty of performing efficient image processing operations on space-variant image architectures is solved.

6 Acknowledgement

The authors thank Prof. Alexander V. Levichev for reading, criticizing and suggesting many improvements on the manuscript.

References

Baloch, A. A. and Waxman, A. M. (1991). Visual learning: adaptive expections, and behavioural conditioning of the mobile robot mavin. *Neural Networks*, 4:271–302.

Baron, T., Levine, M. D., Hayward, V., and nand D. Grant, M. B. (1995). A biologically motivated robot eye system. *Proc. of the 8th Canadian Aeronautics and Space Insitute Conference on Astronautics, Ottawa, Ontario.*, 8:231–240.

(see (**?**)), so that changes in size or in rotation correspond to shifts in the ECT domain. Therefore we can infer the size and rotation of an object with respect to its template by performing template matching directly in the domain of the ECT.

Bederson, B., Wallace, R. S., and Schwartz, E. L. (1992). 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.

Brousil, J. K. and Smith, D. R. (1967). Threshold logic network for shape invariance. *IEEE transactions on electronic computers*, EC-16(6):818–828.

Burt, P. and Adelson, T. (1981). A laplacian pyramid for data compression. *IEEE Transactions on Communications*, 8:1230–1245.

Burt, P. J. (1988). Smart sensing within a pyramid machine. *IEEE Proceedings*, 76(8):1006–1015.

Casasent, D. and Psaltis, D. (1976). Position, rotation and scale-invariant optical correlation. *Applied Optics*, 15:1793–1799.

Cavanagh, P. (1978). Size and position invariance in the visual system. *Perception*, 7:167–177.

Engel, G., Greve, D., Lubin, J., and Schwartz, E. (1994). Space-variant active vision and visually guided robotics: Design and construction of a high-performance miniature vehicle. In *ICPR Proceedings*, ICPR-12, pages 487–490. International Conference on Pattern Recognition.

Ferraro, M. and Caelli, T. M. (1988). Relationship between integral transform invariances and lie group theory. *J. Opt. Soc. Am. A*, 5(5):738–742.

Ibragimov, N. H., editor (1994). *CRC handbook of Lie Group analysis of differential equations*, volume 1. CRC Press, Inc.

Olver, P. J. (1993). *Applications of Lie groups to differential equations*. Springer-Verlag.

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

Rubinstein, J., Segman, J., and Zeevi, Y. (1991). Recognition of distorted patterns by invariance kernels. *Pattern Recognition*, 24(10):959–967.

Sandini, G. and Dario, P. (1989). Active vision based on space-variant sensing. *Intl. Symp. on Robotics Research*.

Schwartz, E. L. (1977). Spatial mapping in primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetics*, 25:181–194.

Schwartz, E. L. (1981). Cortical anatomy, size invariance, and spatial frequency analysis. *Perception*, 10:455–468.

Schwartz, E. L. (1994). Computational studies of the spatial architecture of primate visual cortex: Columns, maps, and protomaps. In Peters, A. and Rocklund, K., editors, *Primary Visual Cortex in Primates*, volume 10 of *Cerebral Cortex*. Plenum Press.

Schwartz, E. L., Greve, D., and Bonmassar, G. (1995). Space-variant active vision: Definition, overview and examples. *Neural Networks*, 8:1297–1308.

Sheng and Arsenault (1986). Experiments on pattern recognition using invariant fourier-mellin descriptors. *J. Opt. Soc. Am. A*, 3(6):771–776.

Wallace, R., Ong, P.-W., Bederson, B., and Schwartz, E. (1994). Space variant image processing. *International Journal of Machine Vision*, 13(1):In press.

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