Geometric invariance in space-variant vision systems: the exponential chirp transform

Giorgio Bonmassar †

†Dept. Biomedical Engineering
Boston University
Boston, MA 02146

Eric L. Schwartz ‡

‡Dept. Cognitive and Neural Systems
Boston University
Boston, MA 02146

Abstract

In this paper, we outline a method to derive geometric invariance kernels which may be applied to a space-variant sensor architecture. The basic idea is to transform a kernel with desired symmetry properties (e.g. the Fourier Kernel) in the domain to the range of the transform. By combining this transformed kernel with the Jacobian of the transformation, we obtain a new integral transform, in the range , which has similar properties to the original transform. We illustrate this idea with a variant of the Mellin-Fourier transform, applied to an image which has been transformed by a log-polar mapping. The kernel obtained, which we call an "exponential chirp" has properties (unlike the Mellin-Fourier transform) which are both consistent with the spatial nature of human vision and can be applied directly in the space-variant image plane. We outline applications to visual template matching and auto-correlation, and show a one-dimensional example of a generalization of cepstral auto-correlation using this method.

1 Introduction

Space-variant vision refers to sensor architectures in which pixel size varies across the sensor plane [5].

Space-variant systems based on the log polar mapping have transformation properties that may be useful in simplifying the computation of optical flow and possibly size and rotation invariance [4, 6].

However, one of the major impediments to using a log-polar sensor architecture has been that the desired property of space-variance breaks translation symmetry in such a way that the simplest image processing operations become extremely complex.

Historically, there have been two different approaches to this problem. First, one might work in the inverse-mapped log-polar image. However, the number of pixels in this image is equal to that of the original uniformly sampled image, therefore sacrificing one of the principle motivations of the method. Second, one might use a Mellin-Fourier transform, as has been pointed out during the past three decades, e.g. [1]. The major computational drawback of the Mellin-Fourier approach is that it requires two-dimensional FFT of the full (i.e. the original uniformly sampled) image. The major biological drawback is that it requires a Fourier Transform of the image to occur prior to the complex logarithmic map of primary visual cortex, which is grossly inconsistent with the anatomy of the visual system. In the present paper, we describe a method, the exponential chirp transform (ECT), which retains the favorable translation symmetry of the Mellin-Fourier approach, but does not require a transformation back into the original image plane. [1]

1.1 One-dimensional case

Given a function $f(x)$ and an invertible transformation $w : x \rightarrow \xi$ the Fourier Transform of $f(x)$ is:

*Supported by ARPA ANNT Program, ONR Contract #N00014-92-C-0119. Giorgio Bonmassar supported by Consiglio Nazionale delle Ricerche (Italy), Bando b.s. n.203.07.23 del 30/06/92

[1] There are applications of Lie Group Theory to the generation of invariance kernels [3], which provide an invariance kernel for a stated symmetry. So, for example, if the desired symmetry is shift, these methods produce the trigonometric kernel as an invariance kernel. For more general symmetries, these methods are capable of producing a desired invariance kernel. However, in this context, this invariance kernel is applied in the image plane. Using the methods of the present paper, an invariance kernel could be transformed and applied under an arbitrary space-variant transformation.

$$\int_{-\infty}^{+\infty} f(x) \exp[-j\omega\, x]\, dx$$

It becomes, (using the Jacobian) in the ξ space:

$$\int_{-\infty}^{+\infty} f(x(\xi)) \frac{\partial x(\xi)}{\partial \xi} \exp[-j\, x(\xi)\, \omega]\, d\xi \qquad (1)$$

The kernel is:

$$K_{\mathrm{T}}(\xi, \omega) = \frac{\partial x(\xi)}{\partial \xi} \exp[-j\, x(\xi)\, \omega] \qquad (2)$$

The associated integral transform:

$$\int_{-\infty}^{+\infty} f(\xi) K_{\mathrm{T}}(\xi, \omega)\, d\xi \qquad (3)$$

is "invariant" up to a phase, under translation in the x domain. This invariance may be complicated by aliasing due to the detailed nature of the transformation w. Following [2], we will consider the one-dimensional transformation: [2]

$$\xi(x) = \begin{cases} log(x+a) & : \quad x \geq 0 \\ 2\, log(a) - log(-x+a) & : \quad x < 0 \end{cases} \qquad (4)$$

for which the kernel, as in eq.(2) is:

$$\begin{aligned} \exp[\xi - j(\exp(\xi) - a)\omega] \quad &: \quad \xi \geq log(a) \\ a^2 \exp[-\xi - j(a - a^2 \exp(-\xi))\omega] \quad &: \quad \xi < log(a) \end{aligned} \qquad (5)$$

The kernel found is reminiscent of a "chirp" with exponentially growing frequency and magnitude. Figure 1 illustrates the one-dimensional kernel, with anti-aliasing filtering: aliasing must be carefully handled, since the rapidly growing frequency of the kernel will eventually cause aliasing to occur.

1.2 Two-dimensional case

Given a function $f(x, y)$ and an invertible and differentiable transformation $w : (x, y) \rightarrow (\xi, \eta)$, the Fourier Transform of $f(x, y)$ in the (ξ, η) space is given by the following integral transform:

$$\int\int_{-\infty}^{+\infty} f(x(\xi, \eta), y(\xi, \eta)) K_{\mathrm{T}}(\xi, \eta, k, l)\, d\xi\, d\eta \qquad (6)$$

[2] This represents a logarithmic mapping 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 $||x|| >> a$.

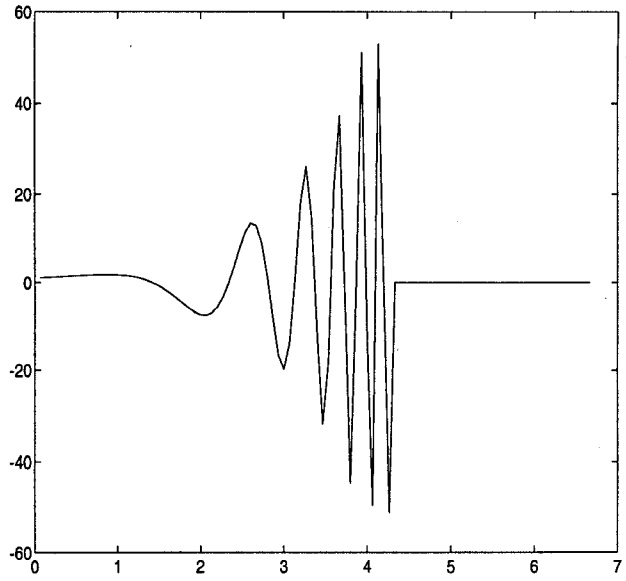

Figure 1: One-dimensional example of kernel with anti-aliasing filtering

where $K_{\mathrm{T}}(\xi, \eta, k, l)$ is the following kernel:

$$\begin{aligned} K_{\mathrm{T}}(\xi, \eta, k, l) &= J(\xi, \eta) \cdot \\ &\cdot \exp[-j(k \cdot x(\xi, \eta) + l \cdot y(\xi, \eta))] \end{aligned} \qquad (7)$$

where $J(\xi, \eta)$ is the Jacobian of the transformation. Note that we have achieved the objective of finding a kernel such that we can compare test data appearing at different positions in the image plane, using convolutions and other image processing operations performed only in the transformed coordinates. For the case of interest here, in which the transformation is a log polar mapping, this allows us to work with the log-polar image directly, despite the fact that an object is grossly deformed in size and shape in the log-polar plane, when it moves in the image plane. Moreover, since the log-polar plane is orders of magnitude smaller than the image plane, we can reap the benefits of the space-variant architecture and avoid the penalty of the lack of simple shift invariance. We consider here the following two-dimensional log-polar (or complex logarithmic) transformation [2]: $w = log(z + a)$, where "a" is a definite positive parameter. [3] The transformation

[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

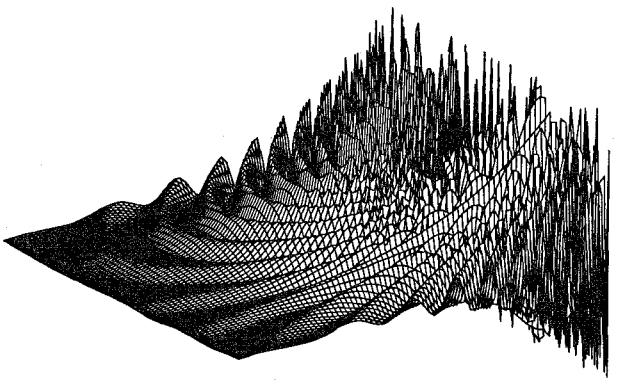

Figure 2: Example of a two-dimensional kernel for log-polar transformation

between spaces can therefore be written:

$$\begin{cases} \xi = log\sqrt{(x+a)^2+y^2} \\ \eta = \arctan\frac{y}{x+a} \end{cases} \quad (8)$$

so we can express x and y as:

$$\begin{cases} x = \exp(\xi)\cos(\eta) - a \\ y = \exp(\xi)\sin(\eta) \end{cases} \quad (9)$$

Eq.(7) becomes:

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

where $\mathcal{D} \equiv \{(\xi,\eta) : -\infty \leq \xi < +\infty \, and \, \frac{-3\pi}{2} \leq \eta \leq \frac{\pi}{2}\}$. Figure 2 illustrates the kernel in two-dimensions, its behavior is exponential both in frequency and amplitude along the ξ axis and is clearly oscillating along the η axis.

2 Results and Conclusions

We will now show how the use of these kernels can be used to perform shift invariance in the image plane of a map function which is strongly space-variant. These demonstrations are made using cepstral auto-correlation here, as an illustration of a template matching technique (although many others might have been chosen). The cepstrum is often used to detect a shifted copy of an object [4] and has been used for echo detection in one-dimensional signals and stereo estimation in two-dimensional signals (images). As an illustration, we have generated a Gaussian pulse, and have examined shifted copies of this pulse in both the original signal plane and after a logarithmic transformation provided by Equation(4).

In the first column of Figure 3, we show a pulse and its echo. The next column shows the pulse pair after logarithmic transformation. In the third column, we show the cepstrum of the pulse pair after logarithmic transformation (distorted pulse pair). The cepstrum of the original pulse pair (not shown) would be a delta function located at the position of the shift. The fourth column shows the cepstrum, obtained using the ECT (cepstral ECT), of the distorted pulse pair. Each row shows a progressively larger shift between the original pulse and its echo. As the shift becomes larger, the logarithmic transformation produces an increasing distortion of the second pulse. In the cepstrum of the log plane (column three of the figure), there is essentially no signal at the location of the shift, which is expected, since the shifted pulse is both narrower and of distorted shape: it should not match well with the original Gaussian template. In the last column, which shows the cepstral ECT as we have described in Equation (5), there is now a sharp peak at the expected position. (Note that the large peak at the origin is always present and would, in practice, be removed by thresholding). The cepstrum, applied to the log transformed space (third column), fails to detect a match essentially in every row. However, the transformed kernel, used in the cepstrum (fourth column) has a peak of signal to noise ratio that is perhaps 3:1. However, the peak is somewhat smaller than in the case of smaller shift (first row), and the peak is somewhat broadened. The behavior of the cepstral ECT (fourth column) can be explained using the concept of a linear position-varying system [7] due to the fact that the filter response depends on the position of the second pulse. The linear position-varying system is characterized by a position-varying transfer function $H(x,\omega)$.

In summary, we have shown a simple method to generate a set of kernels which can be applied to an image which has been re-mapped according to some invertible map function.

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.

[4] The cepstrum is the power spectrum of the log power spectrum

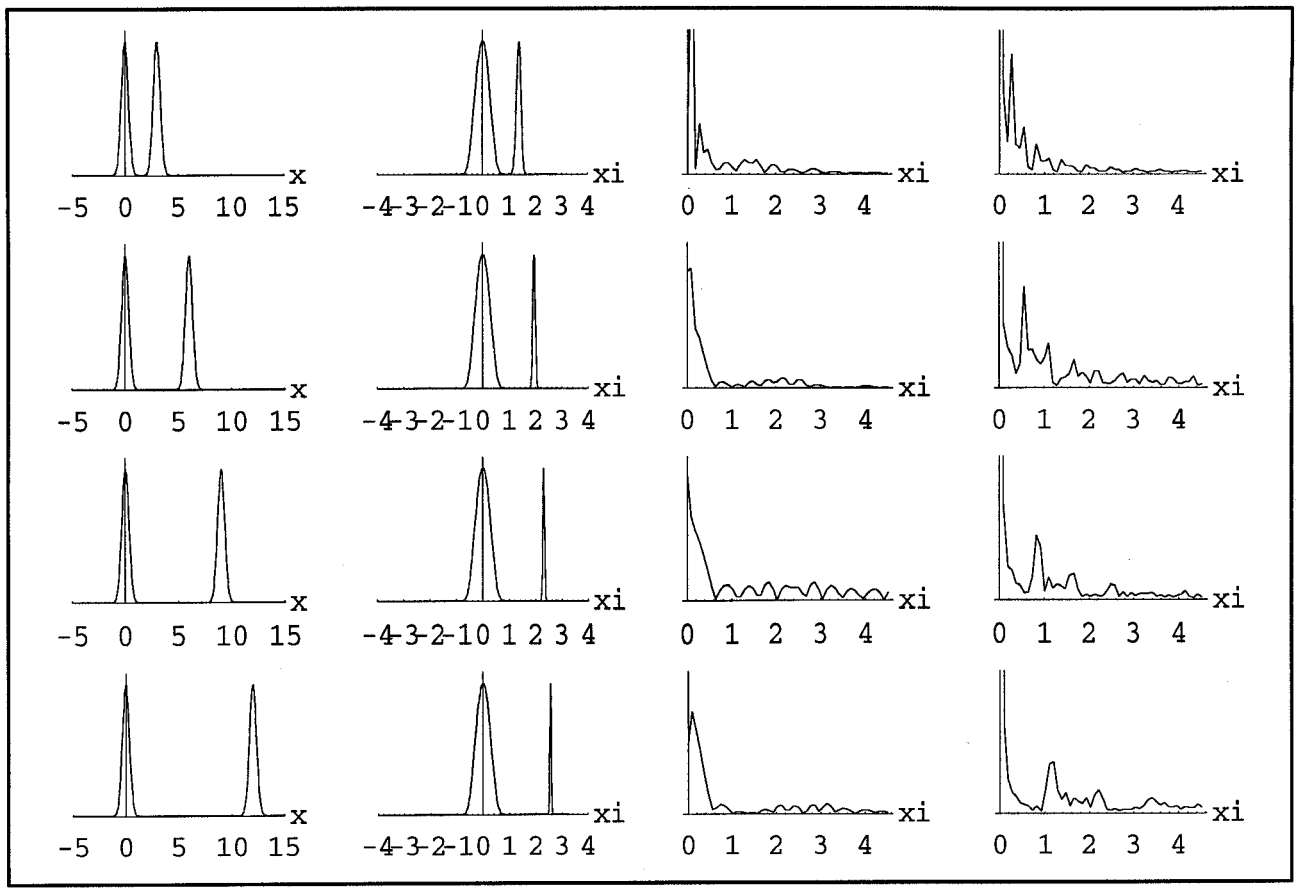

Figure 3: One dimensional example of the application of the method. In the first column are presented two Gaussians with logarithmic sampling, in the second column are the distorted Gaussians, in the third the cepstral autocorrelation and in the last column is the application of a generalization of cepstral auto-correlation using the method described in the text.

References

[1] J. K. Brousil and D. R. Smith. A threshold-logic network for shape invariance. *IEEE Transactions on Computers*, EC-16:818–828, 1967.

[2] 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.

[3] Jacob Rubinstein, Joseph Segman, and Yehoshua Zeevi. Recognition of distorted patterns by invariance kernels. *Pattern Recognition*, 24(10):959–967, 1991.

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

[5] E. L. Schwartz, B. Merker, E. Wolfson, and A. Shaw. Computational neuroscience: Applications of computer graphics and image processing to two and three dimensional modeling of the functional architecture of visual cortex. *IEEE Computer Graphics and Applications*, 8(4):13–28 (July), 1988.

[6] C. F. Weiman and G. Chaikin. Logarithmic spiral grids for image-processing and display. *Computer Graphics and Image Processing*, 11:197–226, 1979.

[7] L. A. Zadeh. A general theory of linear signal transmission systems. *J. Franklin Inst.*, 253:293–312, 1952.