January 1996 Report No. STAN-CS-TN-96-28

A Common Framework for Steerability, Motion Estimation and Invariant Feature Detection

by

Yacov Hel-Or and Patrick C. Teo

Department of Computer Science

Stanford University
Stanford, California 94305

A Common Framework for Steerability, Motion Estimation and Invariant Feature Detection

Yacov Hel-Or[†] and Patrick C. Teo[‡]

[†]NASA Ames Research Center
Moffett Field, CA 94035-1000
toky@white.Stanford.EDU

[‡]Department of Computer Science
Stanford University, Stanford, CA 94305
teo@white.Stanford.EDU

Abstract

Many problems in computer vision and pattern recognition involve groups of transformations. In particular, motion estimation, steerable filter design and invariant feature detection are often formulated with respect to a particular transformation group. Traditionally, these problems have been investigated independently. From a theoretical point of view, however, the issues they address are related. In this paper, we examine the relationships between these problems and propose a theoretical framework within which they can be discussed in concert. This framework is based on constructing a natural representation of the image for a given transformation group. Within this framework, many existing techniques of motion estimation, steerable filter design and invariant feature detection appear as special cases. Furthermore, several new results are direct consequences of this framework. First, a canonical decomposition of all filters that can be steered with respect to any one-parameter group and any multi-parameter Abelian group is proposed. Filters steerable under various subgroups of the affine group are also tabulated. Second, two approximation techniques are suggested for filters that cannot be steered exactly. Approximated steerable filters can also be used for motion estimation. Third, within this framework, invariant features can easily be constructed using traditional techniques for computing point invariance.

Categories: Low-Level Processing, Pattern Analysis, Motion Analysis.

1 Introduction

In computer vision, the problems of steerability, motion estimation and invariant feature detection have usually been investigated independently. One reason for this could be that the intended practical applications of each are vastly different. From a theoretical point of view, however, these three problems address similar core issues. In this paper, we examine these common issues and propose a theoretical framework within which they can be discussed in concert.

In this framework, we are concern with images undergoing some transformation (translation, rotation, affine etc.). The main idea underlying this framework is to find an efficient representation of an image with respect to a given group of transformations. The representation is efficient in the sense that it is simple (linear and finite), and that transformations in the group can both be identified and applied directly to the representation. The representation need not be complete, i.e. the image need not be reconstructible from the set of features. For example, consider the group of all rotations about a given point. A possible representation of an image which is efficient with respect to this group is the horizontal and vertical directional derivatives at that point. This representation is finite (two dimensional) and linear (since directional derivative is a linear operator). It is efficient with respect to rotations since any directional derivative can be calculated from the horizontal and vertical derivatives, i.e. the representation of a rotated image can be reconstructed from the representation of the original image.

Such efficient representations are useful in motion estimation. Because the transformation is detectable from the representation, one can estimate the motion of the image from this lower-dimensional representation instead of from the image directly. Practically, one first computes the outputs of a set of filters from the original and the transformed images, and then estimates the motion from these two sets of measurements. Likewise, functions which are invariant under image transformations can be defined directly over the representation. Since the representation is finite-dimensional, methods for computing invariants with features like points can be employed. Furthermore, because the dimension of the new representation is finite, it is possible to generate all independent image invariants with respect to the given transformation.

Lie group theory has been used extensively in constructing geometric invariants [20, 23, 13]. It is a useful theory because it relates the possibly nonlinear transformation group to a linear vector space called the tangent space, which is the infinitesimal action of the group about its identity. Using this connection, many theorems about the group itself can be proven via simpler proofs in terms of its tangent space. The application of Lie theory to the design of steerable filters or to motion estimation has not, however, been as widespread. The framework proposed in this paper is based on several results from Lie theory, so the families of transformations treated are those that form Lie groups. This, however, is not too restrictive since many transformations of interest to computer vision are Lie groups. Examples include: image translation, rotation, scaling and affine transformation. These transformations may either be global, i.e. acting over the entire image, or local, as in the computation of optical flow.

Several others have also used Lie group theory in a similar context. Amari originally proposed the construction of such efficient representations for invariant feature detection via feature normalization [1, 2]. This work applies and extends his idea to the problems of steerability and motion estimation, and suggests a framework encompassing all three problems. Furthermore, the treatment of invariance within this framework is more general than Amari's feature normalization technique. Lenz also recognized the usefulness of finite-dimensional function spaces that are closed under some transformation and applied the idea to several computer vision applications including pattern detection [15].

The rest of the paper is organized as follows. Section 2 provides a brief introduction to Lie group theory. Following that, Sections 3 and 4 outline the framework and detail several important theorems. In Section 5, examples concerning one-parameter groups are given along with a canonical decomposition of efficient representation spaces for one-parameter groups. Section 6 describes the framework in the context of multi-parameter groups. Next, Section 7 suggests two approximation techniques that are useful when no efficient representation exists. This is followed, in Section 9, by a description of how invariants may be constructed within the framework.

2 Background on Lie Groups

Lie groups are often encountered as families of transformations acting on a signal. In this paper, we consider, primarily, the families of transformation groups acting on real-valued, two-dimensional images. We assume that these images are non-zero only within a bounded region and denote them by $s(x, y) : \mathbf{R}^2 \mapsto \mathbf{R}$. We describe each family of transformations by *operators* $\{g(\boldsymbol{\tau})\}$, where $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_k) \in \mathbf{R}^k$ are parameters of the transformation. For example, consider the family of one-dimensional translations of an image in the x-direction:

$$\hat{s}(\hat{x}, \hat{y}) = g_{t_x}(\tau)\ s(x, y) = s(x - \tau, y)$$

where τ denotes the amount of translation. In words, the operator $g_{t_x}(\tau)$ acts on the original image $s(x, y)$ to yield a new translated image $\hat{s}(\hat{x}, \hat{y}) = s(x - \tau, y)$.

A family of transformations $\{g(\boldsymbol{\tau})\}$ parameterized by $\boldsymbol{\tau}$ over some predefined range is a Lie group if: (1) it satisfies the group conditions of closure under composition, associativity, inverse and the existence of an identity, and (2) the maps for inverse and composition are smooth. Thus, the family of translations forms a Lie group: First, every translation operator $g_{t_x}(\tau)$ has an inverse, namely, $g_{t_x}(\iota(\tau))$ where $\iota(\tau) = -\tau$. Since $\iota(0) = 0$, $g_{t_x}(0)$ is also the identity operator. Second, composition of two operators can be described by a third operator which also belongs to the same family, i.e. $g_{t_x}(\tau_a) g_{t_x}(\tau_b) = g_{t_x}(\rho(\tau_a, \tau_b))$ where $\rho(\tau_a, \tau_b) = \tau_a + \tau_b$. In addition, composition is associative, that is to say, $g_{t_x}(\tau_a)(g_{t_x}(\tau_b) g_{t_x}(\tau_c)) = (g_{t_x}(\tau_a) g_{t_x}(\tau_b))\ g_{t_x}(\tau_c)$; as such, $\rho(\tau_a, \rho(\tau_b, \tau_c)) = \rho(\rho(\tau_a, \tau_b), \tau_c)$. Finally, both the inverse map, $\iota(\tau)$, and the composition map, $\rho(\tau_a, \tau_b)$ are smooth. The dimension of the parameter space of a Lie transformation group may be different from the dimension of the image space upon which it acts. Here, the family of translations in the x-direction forms a one-parameter Lie group ($\tau \in \mathbf{R}$) while the space upon which it acts is two-dimensional ($(x, y) \in \mathbf{R}^2$).

Another familiar family of transformations that is also a Lie group is the group of rotations in the plane $g_r(\tau)$ such that $\hat{s}(\hat{x}, \hat{y}) = g_r(\tau)\ s(x, y) = s(x \cos\tau - y \sin\tau, x \sin\tau + y \cos\tau)$. It is straightforward to check that the necessary conditions, verified in the previous example, are also satisfied here.

Lie groups are rich in structure and many properties of the group can be discerned by studying the properties of infinitesimal actions of the group. In the following, infinitesimal actions of a group are defined and elaborated. We consider first one-parameter groups and then extend our explanation to multi-parameter groups.

One-parameter Groups Given a one-parameter transformation group parameterized by τ, the infinitesimal transformation of an image $s(x, y)$ about the identity ($\tau = 0$) is defined using Leibnitz's chain rule:

$$\frac{d}{d\tau}\left(g(\tau)\ s\right)\Big|_{\tau=0} = \frac{d\hat{s}}{d\tau}\Big|_{\tau=0} = \left(\frac{\partial x}{\partial \tau}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \tau}\frac{\partial}{\partial y} + \frac{\partial}{\partial \tau}\right)\Big|_{\tau=0} \hat{s}.$$

The differential operator on the right hand side of the equation is called the (infinitesimal) *generator* of the transformation and is denoted by L, i.e.

$$L = \left(\frac{\partial x}{\partial \tau}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \tau}\frac{\partial}{\partial y} + \frac{\partial}{\partial \tau}\right)\Big|_{\tau=0} \tag{1}$$

The set of elements $\mathcal{G} = \{\tau L \mid \tau \in \mathbf{R}\}$ forms a one-dimensional vector space called the *tangent space* of the group where L can be thought of as a one-dimensional basis vector.

There is a strong connection between the tangent space and the Lie group from which it was derived. Namely, each element $g(\tau)$ of the group can be generated by an element in the tangent space, $\tau L \in \mathcal{G}$, via the exponential map:[1]

$$g(\tau)\ s(x,y) = e^{\tau L}\ s(x,y)\ . \tag{2}$$

The notation $e^{\tau L}$ represents the series expansion $e^{\tau L} = I + \tau L + \frac{1}{2!}\tau^2 L^2 + \cdots$, which is an infinite sum of differential operators [3]. This is a rather surprising result since the operator $g(\tau)$ can transform the image in highly nonlinear ways while $\mathcal{G}$ is simply a linear vector space.

Recall the group of translations in the x-direction presented earlier. The derivative of the transformation about the identity is

$$\frac{d}{d\tau}(g_{t_x}(\tau)s)\bigg|_{\tau=0} = \frac{\partial(x-\tau)}{\partial\tau}\bigg|_{\tau=0}\frac{\partial}{\partial x}\ s = -\frac{\partial}{\partial x}\ s$$

and hence its generator is $L_{t_x} = -\frac{\partial}{\partial x}$. Using the exponential map suggested in Equation 2, we find that

$$\begin{aligned} g_{t_x}(\tau)\ s &= e^{\tau L_{t_x}}\ s \\ &= (1 - \tau\tfrac{\partial}{\partial x} + \tfrac{1}{2!}\tau^2\tfrac{\partial^2}{\partial x^2} + \cdots)\ s \\ &= s - \tau\tfrac{\partial s}{\partial x} + \tfrac{1}{2!}\tau^2\tfrac{\partial^2 s}{\partial x^2} + \cdots \end{aligned}$$

which is exactly the Taylor expansion of $s(x-\tau,y)$ about $\tau = 0$. Further examples of one-parameter groups and their generators are given in Table 1 of Section 5.

Multi-parameter Groups The situation with multiple-parameter Lie groups is analogous. The generators of a multi-parameter group are the set of differential operators $\{L_i \mid i = 1\ldots k\}$ corresponding to derivatives of the transformation at the identity with respect to each parameter τ_i in turn, i.e.

$$\frac{d\hat{s}}{d\tau_i}\bigg|_{\boldsymbol{\tau}=\mathbf{0}} = L_i\ \hat{s}$$

where

$$L_i = \left(\frac{\partial x}{\partial\tau_i}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\tau_i}\frac{\partial}{\partial y} + \frac{\partial}{\partial\tau_i}\right)\bigg|_{\boldsymbol{\tau}=\mathbf{0}}.$$

[1] To be precise, this is only true for group elements sufficiently close to the identity element so that their Taylor expansions converge, and for elements within the connected component containing the identity. In this paper, we consider only transformation groups with one connected component and for which convergence also holds.

The k generators provide a basis for the k-dimensional tangent space $\mathcal{G} = \{\tau_1 L_1 + \cdots + \tau_k L_k | \boldsymbol{\tau} \in \mathbf{R}^k\}$.[2] As before, there is a correspondence between a k-parameter Lie group and its k-dimensional tangent space in the form of the exponential map:

$$g(\boldsymbol{\tau})\ s(x,y) = (\prod_{i=1}^{k} e^{\tau_i L_i})\ s(x,y) = e^{\tau_1 L_1} \cdots e^{\tau_k L_k}\ s(x,y). \tag{3}$$

Although the exponential map provides a correspondence between every operator in the Lie group and every element in its tangent space, the parameterization of the group generated by the exponential map may be different from that of the original group. Furthermore, the choice of ordering the individual exponential maps in Equation 3 is arbitrary while different ordering give rise to a different group parameterization. Hence, the exponential map generates a group similar to the original group up to a change of parameterization. For example, consider the two parameterizations of the two-parameter affine group acting solely on the x coordinate:

$$\begin{aligned} g_1(\tau_1, \tau_2)\ s(x,y) &= s(e^{\tau_1} x - \tau_2, y), \\ g_2(\tau_1, \tau_2)\ s(x,y) &= s(e^{\tau_1}(x - \tau_2), y). \end{aligned}$$

Both yield the same generators:

$$L_{\tau_1} = x \frac{\partial}{\partial x}, \qquad L_{\tau_2} = -\frac{\partial}{\partial x}$$

thus, having the same exponential map. However, this is not a problem as we are often interested in the group of transformations and not the particular parameterization of it. Furthermore, we can easily reparameterize the generated group using the original parameterization.

With multi-parameter groups, if we vary a single parameter τ_i and keep the others fixed, we get a one-parameter group of transformations $\{g_i(\tau_i)\}$ that is a *subgroup* of the original k-parameter group. Hence, by varying each of the k different parameters separately, we can construct k different one-parameter subgroups. When every element from one subgroup commutes with every element from a second subgroup, i.e. $g_i(\tau_i)\ g_j(\tau_j) = g_j(\tau_j)\ g_i(\tau_i)$ for all τ_i, τ_j, the two subgroups are said to commute with each other. Two subgroups commute if and only if their Lie bracket vanishes, i.e. $[L_i, L_j] \doteq L_i L_j - L_j L_i = 0$ [3]. When two subgroups commute, exponentiating their respective generators can be done in either order, i.e. $e^{\tau_i L_i} e^{\tau_j L_j} = e^{\tau_j L_j} e^{\tau_i L_i} = e^{\tau_i L_i + \tau_j L_j}$. This is not true for non-commuting subgroups. A multi-parameter group for which all its one-parameter subgroups commute is called an

[2] Loosely speaking, the linear independence of the k generators is assured if the k-parameter group from which it was derived cannot be replaced by another with fewer parameters [3].

Abelian group. The two-parameter group of the previous example is not Abelian since

$$[L_{\tau_1}, L_{\tau_2}] = -x\frac{\partial}{\partial x}(\frac{\partial}{\partial x}) + \frac{\partial}{\partial x}(x\frac{\partial}{\partial x}) = \frac{\partial}{\partial x} \neq 0.$$

Hence, as demonstrated earlier, $e^{\tau_1 L_1} e^{\tau_2 L_2} \neq e^{\tau_2 L_2} e^{\tau_1 L_1} \neq e^{\tau_1 L_1 + \tau_2 L_2}$.

On the other hand, for the one-parameter transformation groups listed in Table 1, the pairs, $\{g_{t_x}, g_{t_y}\}, \{g_{t_x}, g_{s_y}\}, \{g_{t_y}, g_{s_x}\}, \{g_{s_x}, g_{s_y}\}, \{g_r, g_s\}$, are commutative.

3 Equivariant Feature Spaces

In computer vision, a common method of extracting features from an image is via the inner-product of the image with some function $\phi(x,y)$:

$$f = < \phi, s > \doteq \int\int \phi(x,y)\overline{s(x,y)}\, dxdy$$

where $\overline{s(x,y)}$ denotes the complex conjugate of $s(x,y)$. We assume that our images are real-valued so the bar may be omitted. We also assume that our images are non-zero only over a bounded region and denote the vector space of these images by S. We refer to $\phi(x,y)$ as the *measuring function* and f as the corresponding *measured feature*. In order to consider several measuring functions in tandem, we introduce the vectorial function $\Phi(x,y) \doteq (\phi_1(x,y), \ldots \phi_n(x,y))^T$ and define the inner-product between a vectorial measuring function $\Phi(x,y)$ and an image $s(x,y)$ by

$$\mathbf{f} = < \Phi, s > \doteq (< \phi_1, s >, \ldots, < \phi_n, s >)^T$$

where $\mathbf{f}$ is an n-dimensional feature vector. Assuming that the functions $\{\phi_i\}$ are linearly independent, they form a basis for an n-dimensional function space called the *measuring space* that is denoted by span(Φ). The choice of Φ as basis functions for the measuring space is not unique as any other $\Phi' = C^T\Phi$, where C is a non-singular matrix, will also span the same measuring space. Likewise, the *feature space* F is the n-dimensional vector space of measured features $\mathbf{f}$. The inner-product of the signal with the vectorial measuring function Φ is a linear mapping from the space of images S to the space of features F. If the dimension of the feature space is lower than that of the image space, then the mapping is many-to-one, i.e. many images could yield the same feature vector.

For a given transformation group $g(\boldsymbol{\tau})$, the set of images obtained by transforming $s(x,y)$ with every member of the group is known as the image orbit $O_g(s)$, i.e. $O_g(s) = \{\hat{s}(\hat{x},\hat{y}) \mid \exists \boldsymbol{\tau} \text{ s.t } \hat{s}(\hat{x},\hat{y}) = g(\boldsymbol{\tau})\ s(x,y)\}$. When an image is transformed, the measured features of the new image will, in general, be different from those of the original. However, if the measuring functions are chosen appropriately (with respect to the given transformation group), then although the new measured features will still be different, they can be interpolated *exactly* from the original features. In this case, the transformation can be applied directly in the feature space F without the necessity of transforming the image itself.

Denoting $\hat{\mathbf{f}} = \langle \Phi, g(\boldsymbol{\tau})\, s \rangle$ as the n-dimensional feature vector of the transformed image, we restate this property more formally as:

Definition 1 (Equivariant Feature Space) *A feature space F is* **equivariant** *under a k-parameter transformation group $g(\boldsymbol{\tau})$ if there exists $A(\boldsymbol{\tau})$, a matrix of functions in $\boldsymbol{\tau}$, known as the* **interpolation matrix**, *such that*

$$\hat{\mathbf{f}} = A(\boldsymbol{\tau})\, \mathbf{f}$$

for each $\boldsymbol{\tau} \in \mathbf{R}^k$ and for every image s. This equation is called the **interpolation equation**.

Note, that if the original feature $\mathbf{f}$ is a complete representation of the image s, i.e. it is possible to reconstruct the image s from $\mathbf{f}$, then F is equivariant regardless of the choice of measuring functions. However, in this paper we are interested in partial representations where $\mathbf{f}$ is of small dimensionality and cannot reconstruct $s(x, y)$.

The concept of equivariant feature spaces is particularly relevant to computer vision. Given an equivariant feature space F, several problems, notably, steerability, motion estimation and invariant feature detection, can be expressed in a common framework:

Steerability. Steerability is a property associated with a filter when the outputs of transformed replicas of its kernel, can be interpolated exactly from a fixed set of basis filter outputs. Formally, a filter is steerable if $\langle g(\boldsymbol{\tau})\psi, s \rangle = \sum_i a_i(\boldsymbol{\tau})\, \langle b_i, s \rangle$ where ψ is the kernel of the filter and $\{b_i\}$ are some basis filters. In Freeman and Adelson [9], for example, the authors described a method of computing the output of a rotated filter from a linear combination of the outputs of specially chosen basis filters. Other authors have also put forward techniques for designing steerable filters [10, 9, 25, 22]. Since a transformation of the kernel can be carried out by inversely transforming the image (see below), it is easy to see that each measuring function ϕ_i in Φ associated with an equivariant feature space F is steerable via the interpolation equation: $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\, \mathbf{f}$. Thus, steerability can be viewed as a forward problem within the framework. From a set of measured features $\mathbf{f}$ (the outputs of the basis filters), we compute $\hat{\mathbf{f}}$ for each transformation $g(\boldsymbol{\tau})$ and for any image s.

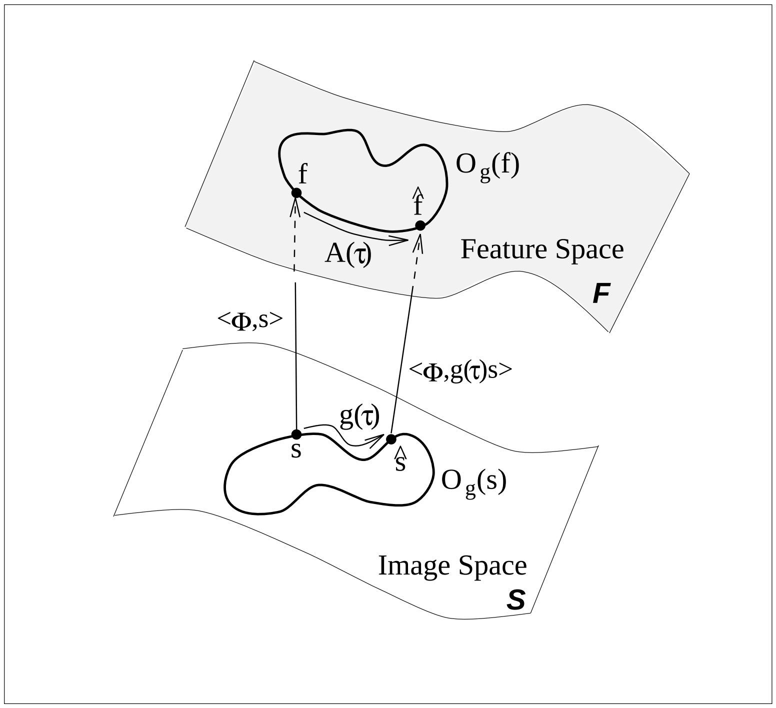

Figure 1: Correspondence between orbits in image space and orbits in feature space. This is possible only if the feature space is equivariant under the transformation group.

Motion Estimation. Motion estimation, on the other hand, is an inverse problem within the framework.[3] Given $\mathbf{f}$ and $\hat{\mathbf{f}}$, measurements made from the original and transformed images respectively, we would like to determine the parameters $\boldsymbol{\tau}$ of the transformation $g(\boldsymbol{\tau})$. Again, if the family of transformations is a Lie group, an equivariant space F is useful to this extent as it relates any two sets of measurements in the space, e.g. $\mathbf{f}$ and $\hat{\mathbf{f}}$, by the interpolation equation $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\,\mathbf{f}$. With this relation, one can then solve for the transform parameters from the two sets of measurements. Besides being forward and inverse problems, there is a small technical difference between the steerability property and motion estimation. In the case of steerability, the transformation occurs on the filter while motion estimation computes the transformation that takes place on the image.

Invariant Feature Detection. An invariant feature or pattern detector indicates the presence (or absence) of a particular pattern in an image regardless of how the image has been transformed. For example, an edge detector should be able to detect the presence of an edge independent of the orientation of the edge in the image. The straightforward

[3]By motion estimation, we do not restrict ourselves only to the detection of infinitesimal changes as is usually implied by the term motion; instead, we consider finite amounts of transformation as well.

approach to this problem is to directly determine filter kernels that are invariant to the given transformation and then use their outputs to identify the pattern. Alternatively, the problem can be approached in two stages: (1) construct a large enough equivariant space F to best characterize the pattern, and (2) determine invariants within this finite (and possibly small) dimensional space. Since the feature space is equivariant, measured features of transformed versions of the pattern can be interpolated, i.e. $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\,\mathbf{f}$. This equation is a parametric description of a k-dimensional manifold in the feature space F. An implicit representation of this manifold that is independent of $\boldsymbol{\tau}$ is clearly invariant under the transformation. Thus, constructing invariants over the feature space amounts to implicitizing the interpolation equation.

4 Equivariant Measuring Spaces

In the previous section, the importance and relevance of equivariant feature spaces were presented. In this section, we develop a framework for constructing measuring spaces whose corresponding feature spaces are equivariant under a given transformation group. This framework is based on the seminal work of S. Amari [1, 2] who originally proposed it in the context of invariant feature detection in pattern recognition.

Throughout the rest of the paper, we will assume that the group of transformations acts on the image. This is true in the case of motion estimation and invariant feature detection. Before proceeding, we make the following observation: the inner-product between a transformed image $g(\boldsymbol{\tau})s(x,y)$ and some measuring function $\phi(x,y)$ can be rewritten as the inner-product between the original image $s(x,y)$ and an appropriately transformed measuring function $\bar{g}(\boldsymbol{\tau})\ \phi(x,y)$:[4]

$$\langle \phi, g(\boldsymbol{\tau})\ s\rangle \;=\; \langle \bar{g}(\boldsymbol{\tau})\ \phi, s\rangle\,. \tag{4}$$

The transformation $\bar{g}$ is known as the *conjugate* of g. It is easy to verify that if g is a Lie group, then $\bar{g}$ is also a Lie group. As such, we denote its generator by $\bar{L}$ and refer to it as the conjugate generator of g. Table 1 lists several common one-parameter groups and their conjugate generators. With steerability, it is not necessary to introduce the conjugate of a group as the transformation is applied directly onto the measuring function itself and not onto the image. The derivation of conjugate generators is explained in Appendix A.

In the previous section, a feature space is defined to be equivariant under a given transfor-

[4]Again, we restrict ourselves to groups (up to a re-parameterization) that can be generated by the exponential map.

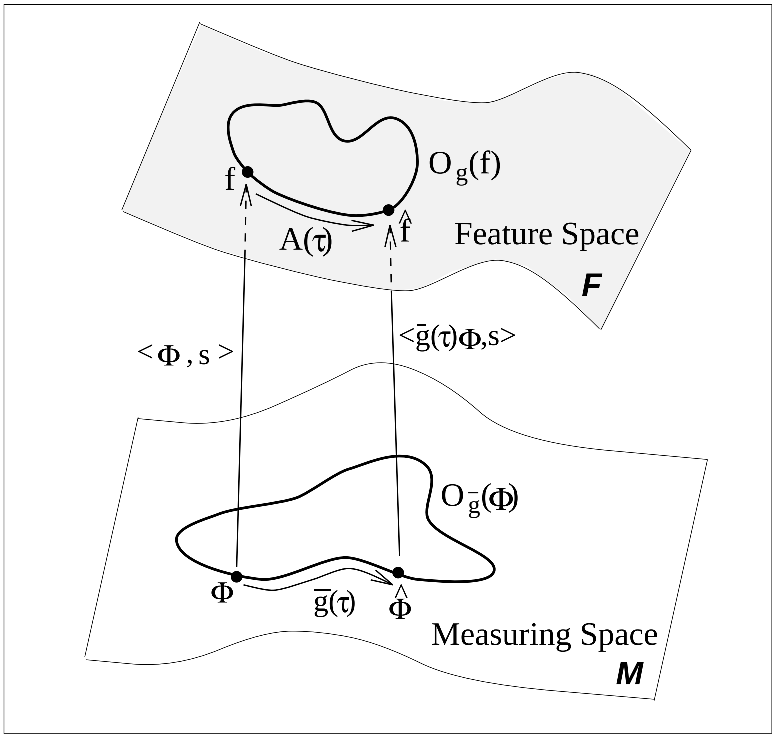

Figure 2: Correspondence between orbits in image space and orbits in feature space. This is possible only if the feature space is equivariant under the transformation group.

mation if any two feature vectors are related by the interpolation equation. In the following theorem, the conditions required for a feature space to be equivariant are given in terms of its measuring space.

Theorem 1 (Equivariant Measuring Space, Amari '78 [2]) *: A feature space F is equivariant with respect to a transformation group $g(\boldsymbol{\tau})$ if and only if there exists a square matrix $A(\boldsymbol{\tau})$ such that*

$$\bar{g}(\boldsymbol{\tau})\;\Phi = A(\boldsymbol{\tau})\;\Phi \qquad \forall \boldsymbol{\tau} \in \mathbf{R}^k.$$

In this case, span(Φ), i.e. the function space spanned by the elements of Φ, forms an **equivariant measuring** *space.*

Proof 1 : If $\bar{g}(\boldsymbol{\tau})\;\Phi = A(\boldsymbol{\tau})\;\Phi$, then

$$\begin{aligned}
\hat{\mathbf{f}} &\doteq \langle \Phi, g(\boldsymbol{\tau})\; s\rangle \\
&\doteq \langle \bar{g}(\boldsymbol{\tau})\;\Phi, s\rangle \\
&= \langle A(\boldsymbol{\tau})\;\Phi, s\rangle \\
&= A(\boldsymbol{\tau})\,\langle \Phi, s\rangle \\
&= A(\boldsymbol{\tau})\,\mathbf{f}.
\end{aligned}$$

The necessity direction of the theorem can be verified in a similar manner. However, this direction follows from the definition of measured features, i.e. from the linearity of their construction. If non-linear operations are permitted, it may be possible to find equivariant feature space (i.e. features satisfying $\hat{\mathbf{f}} = A\mathbf{f}$) despite the unsuitability of the necessity condition. We further explore this point later on in this section. □

The last theorem states a very useful characteristic of an equivariant feature space: that its associated measuring space is closed under the conjugate transformation. It makes sense to say that the measuring space $span(\Phi)$ is equivariant because any element in the space can be written as linear combinations of ϕ_i and since each ϕ_i is closed under the conjugate transformation, any linear combination of ϕ_i is also closed under the conjugate transformation. Because we are dealing with Lie transformation groups, the closure of the measuring space under $\bar{g}(\boldsymbol{\tau})$ can be reformulated, more simply, in terms of its set of generators $\{\bar{L}_1, \ldots, \bar{L}_k\}$.

Theorem 2 (Interpolation Equation) : *The measuring space* $span(\Phi)$ *is equivariant under the group* $g(\boldsymbol{\tau})$ *if and only if* Φ *is closed under the action of each conjugate generator* $\bar{L}_i$ *of* g*, i.e.*

$$\bar{g}(\boldsymbol{\tau})\ \Phi = A(\boldsymbol{\tau})\ \Phi \qquad \textit{if and only if} \qquad \bar{L}_i\ \Phi = B_i\ \Phi \quad \forall\, i = 1, \ldots, k$$

for some set of $n \times n$ *matrices* $\{B_1, \ldots, B_k\}$*. In particular, the interpolation matrix can be written as follows:*

$$A(\boldsymbol{\tau}) = e^{\tau_k B_k} \cdots e^{\tau_1 B_1}.$$

Proof 2 : Let $\hat{\Phi}(\boldsymbol{\tau}) = \bar{g}(\boldsymbol{\tau})\ \Phi$, the (conjugate) transformed measuring functions. Since $\bar{g}(\boldsymbol{\tau})$ is a Lie group, it follows from the exponential map in Equation 3 that

$$\begin{aligned}
\hat{\Phi}(\tau_1, \ldots, \tau_k) &= e^{\tau_1 \bar{L}_1} \cdots e^{\tau_k \bar{L}_k}\ \Phi \\
&= (1 + \tau_1 \bar{L}_1 + \cdots) \cdots (1 + \tau_k \bar{L}_k + \cdots)\ \Phi \\
&= (1 + \tau_1 \bar{L}_1 + \cdots) \cdots (1 + \tau_k B_k + \cdots)\ \Phi \\
&= (1 + \tau_1 \bar{L}_1 + \cdots) \cdots (1 + \tau_{k-1} \bar{L}_{k-1} + \cdots)\ e^{\tau_k B_k}\ \Phi \\
&= (1 + \tau_1 \bar{L}_1 + \cdots) \cdots\ e^{\tau_k B_k}\ (1 + \tau_{k-1} \bar{L}_{k-1} + \cdots)\ \Phi \\
&= (1 + \tau_1 \bar{L}_1 + \cdots) \cdots\ e^{\tau_k B_k}\ e^{\tau_{k-1} B_{k-1}}\ \Phi \\
&\ \vdots \\
&= e^{\tau_k B_k} \cdots e^{\tau_1 B_1}\ \Phi,
\end{aligned}$$

in which the substitution $(\bar{L}_i)^m \Phi = (B_i)^m \Phi$ is used repeatedly. It can easily be verified that $\bar{L}_i\ \Phi = B_i\ \Phi$ implies $(\bar{L}_i)^m \Phi = (B_i)^m \Phi$ via the linearity of the differential operator

1. Derive the conjugate generators $\bar{L}_1, \ldots, \bar{L}_k$ of the given transformation group $g(\boldsymbol{\tau})$.
2. Verify for each generator $\bar{L}_i$ that

$$\bar{L}_i\, \Phi = B_i\, \Phi$$

where B_i is some $n \times n$ matrix.
3. If so, the measuring space span(Φ) is equivariant and the interpolation equation is simply

$$\hat{\mathbf{f}} = A(\boldsymbol{\tau})\, \mathbf{f}$$

where the interpolation matrix

$$A(\boldsymbol{\tau}) = e^{\tau_k B_k} \cdots e^{\tau_1 B_1}.$$

Figure 3: Recipe for verifying that the function space of span(Φ) is equivariant. If so, the interpolation matrix $A(\boldsymbol{\tau})$ is also derived.

$\bar{L}_i$. The order in which the generators $\bar{L}_i$ are applied is arbitrary. However, as pointed out in Section 2, the order will determine the parameterization of the generated group. Another way of proving this direction of the theorem is by solving the differential equation, $\bar{L}_i \Phi \doteq \frac{d\hat{\Phi}}{d\tau_i}\Big|_{\boldsymbol{\tau}=0} = B_i \Phi$, for $\hat{\Phi}$. Conversely, if $\hat{\Phi} = e^{\tau_k B_k} \cdots e^{\tau_1 B_1}\ \Phi$, taking derivatives with respect to τ_i (about $\boldsymbol{\tau} = 0$) on both sides of the equation yields the system of equations $\bar{L}_i\ \Phi = B_i\ \Phi$.

Since each ϕ_i in Φ is closed under the action of every generator $\bar{L}_i$, any linear combination of ϕ_i is also closed. Hence, the function space span(Φ) is also closed under the action of every generator $\bar{L}_i$. $\square$

Theorem 2 provides a recipe for verifying that the space spanned by a set of functions $\{\phi_i\}$ is an equivariant measuring space, and if it is, derives the interpolation matrix $A(\boldsymbol{\tau})$. Figure 3 summarizes the procedure. Unfortunately, the construction of all possible n-dimensional equivariant measuring spaces is not as methodical in general. For one-parameter groups, however, the construction is straightforward and will be treated extensively in the next section.

The following are corollaries that can be used to construct more complicated equivariant spaces from existing ones. Their validity can easily be verified.

Corollary 1 : *If Φ is a vector of n equivariant measuring functions, then $P\Phi$, where P is a non-singular $n \times n$ matrix, is also a vector of equivariant measuring functions. Hence, two vectors of measuring functions Φ_1, Φ_2 share the same equivariant space if and only if they can be related by a non-singular $n \times n$ matrix P such that $\Phi_1 = P\Phi_2$.*

Corollary 2 : *If Φ_1 and Φ_2 are vectors of equivariant measuring functions with respect to the same transformation group, then the space spanned by their direct sum $\Phi_1 \oplus \Phi_2$ (i.e. the concatenation of the two vectors) is also equivariant.*

Corollary 3 : *If Φ_1 and Φ_2 are vectors of equivariant measuring functions with respect to the same transformation group, then the space spanned by the Kronecker product of the two vectors of functions $\Phi_1 \otimes \Phi_2$ (i.e. the pairwise products of measuring functions from Φ_1 and Φ_2) is also equivariant.*

The last corollary presents a powerful way to enrich an equivariant measuring space. For example, assume $\Phi_1 = (1, x)^T$ spans an equivariant measuring space with respect to some transformation group. Using Corollary 3 (and assuming $\Phi_1 = \Phi_2$), it is immediate to conclude that $(1, x, x^2, \cdots)^T$ also spans an equivariant space. This corollary stems from the linearity of the equivariant equation; Since $\bar{g}\Phi_i = A_i\Phi_i, \quad i = 1, 2$, it follows that

$$\bar{g}(\Phi_1\Phi_2^T) = A_1\Phi_1\Phi_2^T A_2^T \quad . \tag{5}$$

The entries of matrix $\Phi_1\Phi_2^T$ are the elements of the Kronecker product $\Phi_1 \otimes \Phi_2$. Since Equation 5 is linear in $\Phi_1\Phi_2^T$, it is possible to rewrite it as $\bar{g}\Phi_{1,2} = A_{1,2}\Phi_{1,2}$, where $A_{1,2}$ is some interpolation matrix and $\Phi_{1,2}$ is a vector function composed of entries from $\Phi_1\Phi_2^T$. Thus, $\Phi_{1,2}$ spans an equivariant measuring space as well. $\square$

The linearity of the equivariant equation leads to a similar result with respect to the feature space:

Corollary 4 : *If $\mathbf{f}_1$ and $\mathbf{f}_2$ are features of two equivariant feature spaces with respect to the same transformation group, then the feature space formed by $\mathbf{f}_1 \otimes \mathbf{f}_2$ is also equivariant.*

Proof : Similar to corollary 3, the proof follows from the linearity of the interpolation equation: $\hat{\mathbf{f}}_i = A_i\mathbf{f}_i, \quad i = 1, 2.$ $\square$

This corollary expands the variety of the equivariant feature spaces even further. It enables (in some cases) to perform non-linear operations on the image and yet to remain in an equivariant feature space.

5 Equivariant Spaces for One-Parameter Groups

In the previous section, the conditions that are required for a measuring space to be equivariant under a transformation group were stated. In this section, we attend to the construction of all possible equivariant spaces with respect to any one-parameter transformation group. First, we provide examples of several equivariant measuring spaces. After that, we show that any one-parameter group can be reparameterized to appear as a group of translations in the new parameterization. Finally, we propose a canonical decomposition of all the measuring spaces equivariant under the translation group (and correspondingly under any one-parameter group that has been appropriately reparameterized).

5.1 The Translation Group

Consider the group of one-dimensional image translations in the x-direction: $\hat{s}(\hat{x}, \hat{y}) = g_{t_x}(\tau)\ s(x, y) = s(x - \tau, y)$ whose conjugate generator $\bar{L}_{t_x} = \frac{\partial}{\partial x}$. An n-dimensional measuring space Φ is equivariant with respect to $g_{tx}(\tau)$ if $\bar{L}_{tx}\ \Phi = \frac{\partial}{\partial x}\ \Phi = B\ \Phi$ for a given $n \times n$ matrix B. The general solution to this differential equation is

$$\Phi(x, y) = e^{Bx}\ \Phi(0) \tag{6}$$

where $\Phi(0)$ is the value of Φ at $x = 0$. Actually, the product of $\Phi(x, y)$ with any function solely in y leaves it equivariant; however, without loss of generality, we refer to $\Phi(x, y)$ only as $\Phi(x)$. Since $\Phi(0)$ can be arbitrary chosen, any element in the column space of e^{Bx} is a possible solution. We will denote this by $\Phi(x) \in \mathcal{R}(e^{Bx})$ where $\mathcal{R}$ refers to the column space of the matrix e^{Bx}. Regardless of the choice of $\Phi(0)$, the interpolation equation is the same, i.e. $\hat{\mathbf{f}} = e^{B\tau}\mathbf{f}$; only the measuring functions associated with the features will be different.

In the following examples, we present different choices for the matrix B and derive the corresponding measuring functions.

Example 1 : Consider the simplest case where B is a 1×1 matrix, i.e. $B = [\lambda]$ where λ is a scalar value (which may be complex). From Equation 6, the space of measuring functions is: $\phi(x) = ae^{\lambda x}$, where a is some scalar value (the value at $\phi(0)$). The interpolation equation in this feature space is $\hat{f} = e^{\tau\lambda} f$. An alternative way to show the above result is by expanding the definitions of $\hat{f}$ and f. Recall that $f = \langle \phi, s \rangle = \int \int ae^{\lambda x}\ s(x, y)\ dxdy$. Therefore,

$$
\begin{aligned}
\hat{f} &= \langle \phi, \hat{s} \rangle \\
&= \int \int ae^{\lambda x}\ s(x - \tau, y)\ dxdy \\
&= \int \int ae^{\lambda(u+\tau)}\ s(u, y)\ dudy \\
&= e^{\lambda\tau} \int \int ae^{\lambda u}\ s(u, y)\ dudy \\
&= e^{\lambda\tau} f.
\end{aligned}
$$

When λ is purely imaginary, the measuring functions are complex exponentials. In phase-based motion estimation, the parameter τ is regarded as the difference in phase. Fleet and Jepson [8] proposed an accurate method of measuring disparity by estimating the difference in phase between two (windowed) complex exponentials.

Example 2 : Now, let

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

In this case, the solution to Equation 6 implies that

$$\Phi(x) \in \mathcal{R}(e^{Bx}) = \mathcal{R}\left[\begin{pmatrix} e^{\lambda_1 x} & 0 \\ 0 & e^{\lambda_2 x} \end{pmatrix}\right] \qquad \text{and} \qquad \hat{\mathbf{f}} = e^{\tau B}\ \mathbf{f} = \begin{pmatrix} e^{\lambda_1 \tau} & 0 \\ 0 & e^{\lambda_2 \tau} \end{pmatrix} \mathbf{f}.$$

Simoncelli et. al. [25] proposed a criterion for shiftability in position that decomposes the filter into a set of complex exponentials (via the Fourier decomposition). In this example, it would correspond to B being a diagonal matrix with unique and purely imaginary λ's.

Example 3 : Let

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

In this case, the measuring function and interpolation equation are

$$\Phi(x) \in \mathcal{R}(e^{Bx}) = \mathcal{R}\left[\begin{pmatrix} 1 & x & \frac{1}{2!}x^2 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}\right] \qquad \text{and} \qquad \hat{\mathbf{f}} = e^{\tau B}\ \mathbf{f} = \begin{pmatrix} 1 & \tau & \frac{1}{2!}\tau^2 \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{pmatrix} \mathbf{f}$$

respectively. This example produces the moment filters which are used in many applications involving invariant feature detection [12] and motion estimation [30].

5.2 The Rotation Group

Another commonly encountered one-parameter transformation group is the group of rotations in the plane:

$$g_r(\tau)\ s(x,y) = s(x\cos\tau - y\sin\tau, x\sin\tau + y\cos\tau)$$

where τ represents the angle of rotation. The conjugate generator of the rotation group is: $\bar{L}_r = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$. It is easy to see that if we represent the image $s(x,y)$ in polar coordinates (r,θ), then rotation becomes similar to translation: $g_r(\tau)s(r,\theta) = s(r,\theta-\tau)$. In these coordinates, the conjugate generator is $\bar{L}_r = \frac{\partial}{\partial\theta}$. Therefore, as before, an n-vector of measuring functions $\Phi(r,\theta)$ is equivariant with respect to $g_r(\tau)$ if it satisfies the equation

$$\bar{L}_r\ \Phi \doteq \left.\frac{\partial\Phi}{\partial\theta}\right|_{\theta=0} = B\ \Phi$$

where B is an $n\times n$ matrix. The general solution to the above equation is simply

$$\Phi(\theta) = e^{B\theta}\Phi(0)$$

where $\Phi(0)$ is the value of $\Phi(\theta)$ at $\theta=0$. Since $\Phi(0)$ is arbitrarily chosen, $\Phi(\theta)\in\mathcal{R}(e^{B\theta})$.

Example 4 : In this example, we show that a vector of measuring functions is equivariant with respect to rotation and derive its interpolation matrix. Let $\Phi(x,y)$ be a vectorial function containing the spatial derivatives of a Gaussian in the x- and y- directions:

$$\Phi(x,y) = \begin{pmatrix} \frac{\partial}{\partial x}e^{-(x^2+y^2)/2} \\ \frac{\partial}{\partial y}e^{-(x^2+y^2)/2} \end{pmatrix} = \begin{pmatrix} -xe^{-(x^2+y^2)/2} \\ -ye^{-(x^2+y^2)/2} \end{pmatrix} = \begin{pmatrix} -r\cos(\theta)e^{-r^2/2} \\ -r\sin(\theta)e^{-r^2/2} \end{pmatrix}.$$

Applying the generator $\bar{L}_r = \frac{\partial}{\partial\theta}$ to Φ, we get

$$\bar{L}_r\ \Phi = \begin{pmatrix} r\sin(\theta)e^{-r^2/2} \\ -r\cos(\theta)e^{-r^2/2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\ \Phi = B\ \Phi.$$

Thus, the elements of $\Phi(x,y)$ span a measuring space whose interpolation function is

$$\hat{\mathbf{f}} = e^{\tau B}\ \mathbf{f} = \begin{pmatrix} \cos(\tau) & -\sin(\tau) \\ \sin(\tau) & \cos(\tau) \end{pmatrix}\mathbf{f}.$$

This is an example of the steerable filters suggested by Freeman and Adelson [9].

5.3 Canonical Coordinates of One-Parameter Groups

The construction of a set of equivariant measuring functions depends on the existence of a solution to the system of partial differential equations $\bar{L}\ \Phi = B\ \Phi$. It was shown that for translations and planar rotations, solutions exist for any given matrix B. In this section, we show that solutions exist for *any* one-parameter transformation group. The simplest way to show this is via a reparameterization of the current coordinates into some canonical coordinates where solutions are known to exist. For any one-parameter transformation group $g(\tau)$, there exists a change of coordinates such that the group resembles a translation in the new parameterization [3]. Hence, given an image $s(x, y)$, one can determine a change of coordinates $s(\eta(x, y), \xi(x, y))$ such that

$$g(\tau)\ s(\eta, \xi) = s(\eta - \tau, \xi).$$

Segman et. al. [23] used this reparameterization to construct invariant kernels for pattern recognition. Ferraro and Caelli [5] also used this method in a similar context and suggested its relevance to biological vision.

Since the group operation is the same as one-dimensional translation, the equivariant condition with respect to the canonical coordinates is also the same, i.e.

$$\bar{L}_{\eta,\xi}\ \Phi(\eta, \xi) \doteq \frac{\partial}{\partial \tau}\Phi(\eta, \xi)\bigg|_{\tau=0} = B\ \Phi(\eta, \xi).$$

Therefore, its equivariant spaces also resemble the equivariant spaces for translation (up to a change of coordinates).

Example 5 : In Section 5.2, polar coordinates were used for the group of rotations in the plane. It is easy to show that polar coordinates are the canonical coordinates for this group. Recall the change of coordinates from Cartesian to polar:

$$\eta = \arctan(y/x) = \theta \quad ; \quad \xi = \sqrt{x^2 + y^2} = r.$$

Rotating an image $s(x, y)$ in Cartesian coordinates is the same as translating the image in polar coordinates: $g_r(\tau)\ s(\eta, \xi) = s(\eta - \tau, \xi)$ where $\tau \in [0, 2\pi)$.

Example 6 : Consider next the one-parameter group of scaling in the x direction, i.e. $g_{s_x}(\tau)\ s(x, y) = s(e^{-\tau}x, y)$ where $e^{-\tau}$ ensures that the scaling constant is always positive. The canonical coordinates of this transformation group are obtained by the following coordinate changes:

$$\eta = \ln(x)\ ;\ \xi = y.$$

In this case,

$$g_{s_x}(\tau)\ s(\eta,\xi) = s(\ln(e^{-\tau}x),\xi) = s(\ln(x)+\ln(e^{-\tau}),\xi) = s(\eta-\tau,\xi)$$

which is a translation in the new coordinate system. Suppose now that

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \Phi(\eta) \in \mathcal{R}(e^{B\eta}) = \mathcal{R}\left[\begin{pmatrix} 1 & \eta & \frac{1}{2!}\eta^2 \\ 0 & 1 & \eta \\ 0 & 0 & 1 \end{pmatrix}\right]$$

as in Example 3 of Section 5.1 with measuring space in η coordinates. After a change of coordinates, the measuring space in x coordinates is

$$\Phi(x) \in \mathcal{R}\left[\begin{pmatrix} 1 & \ln x & \frac{1}{2!}(\ln x)^2 \\ 0 & 1 & \ln x \\ 0 & 0 & 1 \end{pmatrix}\right].$$

5.4 Canonical Decomposition of One-Parameter Equivariant Spaces

For any one-parameter group, the n-vector of equivariant measuring functions Φ depends on the apriori choice of the $n \times n$ matrix B. However, the same function space, span(Φ), may be generated by different B matrices. The following theorem provides an equivalence condition among the various B matrices that generate the same equivariant measuring space.

Theorem 3 : *Let Φ_1, Φ_2 be two n-vectors of equivariant measuring functions (with respect to the same one-parameter group) and B_1, B_2 are such that: $\bar{L}\ \Phi_1 = B_1\Phi_1$ and $\bar{L}\ \Phi_2 = B_2\Phi_2$, then*

$$\Phi_1 = P\ \Phi_2 \qquad \textit{iff} \qquad B_1 = PB_2P^{-1}.$$

for any non-singular $n \times n$ matrix P.

In words, two vectors of equivariant measuring functions Φ_1, Φ_2, with respect to the same group, span the same function space if and only if their corresponding matrices B_1, B_2 are *similar*. Hence, it suffices to examine all matrices B that are unique up to a similarity transformation. The proof of the above theorem and the decomposition of all the equivariant measuring spaces into canonical classes can be found in [11]. It is shown there that any measuring function $\Phi(x)$ that is equivariant under x-translation can be represented by a direct sum of equivariant measuring functions Φ_{J_i} where

$$\Phi_{J_i} = (e^{\lambda_i x}, xe^{\lambda_i x}, x^2 e^{\lambda_i x}, \cdots, x^{n-1}e^{\lambda_i x})^T$$

Group	Operator	Generator	Conjugate Generator	Equivariant Measuring Space
Brightness scaling	$g_b(\tau)\, s(x,y) = e^{\tau} s(x,y)$	$L_b = I$	$\bar{L}_b = I$	all the functions
x-translation	$g_{t_x}(\tau)\, s(x,y) = s(x-\tau, y)$	$L_{t_x} = -\frac{\partial}{\partial x}$	$\bar{L}_{t_x} = \frac{\partial}{\partial x}$	$\{\psi_p(y) x^p e^{\alpha x}\}$ for $0 \le p \le m$.
x-scaling	$g_{s_x}(\tau)\, s(x,y) = s(e^{-\tau}x, y)$	$L_{s_x} = -x\frac{\partial}{\partial x}$	$\bar{L}_{s_x} = I + x\frac{\partial}{\partial x}$	$\{\psi_p(y) x^{\alpha} (\ln x)^p\}$ for $0 \le p \le m$.
Rotation	$g_r(\tau)\, s(x,y) = s(x\cos\tau - -y\sin\tau, x\sin\tau + y\cos\tau)$	$L_r = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} = -\frac{\partial}{\partial \theta}$	$\bar{L}_r = -x\frac{\partial}{\partial y} +y\frac{\partial}{\partial x} = \frac{\partial}{\partial \theta}$	$\{\psi_p(r) \theta^p e^{\alpha\theta}\}$ for $0 \le p \le m$.
Uniform scaling	$g_s(\tau)\, s(x,y) = s(e^{-\tau}x, e^{-\tau}y)$	$L_s = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y} = -r\frac{\partial}{\partial r}$	$\bar{L}_s = 2I + x\frac{\partial}{\partial x} +y\frac{\partial}{\partial y} = I + r\frac{\partial}{\partial r}$	$\{\psi_p(\theta) r^{\alpha} (\ln r)^p\}$ for $0 \le p \le m$.

Table 1: Several examples of one parameter groups, their generators, conjugate generators, and associated equivariant measuring spaces. In the rotation and uniform scaling examples, (r, θ) are the polar coordinates of the image.

for some λ_i. When λ_i is zero, the equivariant measuring space is spanned by the first n moments. Alternatively, when n is one and λ_i is purely imaginary, the space is spanned by the complex exponentials, which are also the Fourier basis functions. Since any one-parameter transformation group can be put into its canonical coordinates (where the group operation becomes a translation in these new coordinates), the decomposition of equivariant measuring spaces for translation applies directly to all other one-parameter transformation groups (after reparameterization) as well. Table 1 is a summary of several common one-parameter groups and their equivariant measuring spaces.

6 Equivariant Spaces for Multi-Parameter Groups

Unfortunately, there is no systematic way to construct general n-dimensional equivariant spaces for multi-parameter groups. With one-parameter groups (in their canonical coordinates), solutions to the system of partial differential equations exist for arbitrary choices of B. Unlike one-parameter groups, arbitrary choices of B_i for multi-parameter groups will often not yield solvable systems of differential equations. However, for Abelian multi-parameter groups a categorization of the equivariant spaces similar to that for one-parameter groups can be carried out. In the following, the categorization of equivariant spaces for Abelian multi-parameter groups is presented. After that, two techniques for handling non-Abelian multi-parameter groups are suggested.

Abelian Multi-Parameter Groups When the multi-parameter group is Abelian, there exists a reparameterization of the group so that the group action is equivalent to independent translations in the new parameterization [3, 23, 5]. Formally, for any two-parameter Abelian group, there exists a reparameterization of the image $s(\eta(x,y),\xi(x,y))$ so that

$$g(\tau_1,\tau_2)\ s(\eta,\xi) = s(\eta-\tau_1,\xi-\tau_2).$$

Segman and Zeevi in [23] describes a constructive way of determining this canonical reparameterization. In the new parameterization, the equivariant space for the two-parameter group is simply the product of the equivariant spaces for each one-parameter translation group:

$$\mathrm{span}(\Phi(\eta,\xi)) = \mathrm{span}(\ \eta^p e^{\alpha\eta}\ ) \otimes \mathrm{span}(\ \eta^q e^{\beta\xi}\ ) = \mathrm{span}(\ \eta^p \xi^q e^{\alpha\eta+\beta\xi}\ )$$

for $0 \leq p \leq m$ and $0 \leq q \leq l$. Note that multi-parameter groups acting on a two-dimensional image with more than two parameters are necessarily not Abelian as there are only two independent translations in an image.

<u>**Example 7**</u> **:** Consider the group of rotation and uniform scaling made up of the two one-parameter subgroups $g_r(\tau_1)$ and $g_s(\tau_2)$ from Table 1 in Section 5. The conjugate generators for these groups are $\bar{L}_r = -x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x}$ and $\bar{L}_s = 2I + x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$ respectively. Recall that two one-parameter groups are Abelian if their generators commute, i.e.

$$\begin{aligned}
[\bar{L}_r,\bar{L}_s] &= [-x\tfrac{\partial}{\partial y} + y\tfrac{\partial}{\partial x}, 2I + x\tfrac{\partial}{\partial x} + y\tfrac{\partial}{\partial y}] \\
&= [-x\tfrac{\partial}{\partial y}, +x\tfrac{\partial}{\partial x}] + [-x\tfrac{\partial}{\partial y}, +y\tfrac{\partial}{\partial y}] + [y\tfrac{\partial}{\partial x}, +x\tfrac{\partial}{\partial x}] + [y\tfrac{\partial}{\partial x}, +y\tfrac{\partial}{\partial y}] \\
&= (-x^2\tfrac{\partial^2}{\partial y\partial x} + x\tfrac{\partial}{\partial y} + x^2\tfrac{\partial^2}{\partial x\partial y}) + (-xy\tfrac{\partial^2}{\partial y^2} - x\tfrac{\partial}{\partial y} + xy\tfrac{\partial^2}{\partial y^2}) + \\
&\quad (xy\tfrac{\partial^2}{\partial x^2} + y\tfrac{\partial}{\partial x} - xy\tfrac{\partial^2}{\partial x^2}) + (y^2\tfrac{\partial^2}{\partial x\partial y} - y\tfrac{\partial}{\partial x} - y^2\tfrac{\partial^2}{\partial y\partial x}) \\
&= 0.
\end{aligned}$$

The reparameterization that makes $g_r(\tau_1)$ and $g_s(\tau_2)$ act like translations on the image is:

$$\begin{aligned}
\eta(x,y) &= \arctan(y/x) = \theta \\
\xi(x,y) &= \ln(\sqrt{x^2+y^2}) = \ln(r)
\end{aligned}$$

Hence, the equivariant spaces for rotation and scaling are:

$$\mathrm{span}(\ \ln(r)^p e^{\beta\ln(r)}\ ) \otimes span(\ \theta^q e^{\alpha\theta}\ ) \quad \text{for}\ \ 0 \leq p \leq m\ \ \text{and}\ \ 0 \leq q \leq l\ \ .$$

Non-Abelian Multi-Parameter Groups For multi-parameter groups that are not Abelian, there are no reparameterizations such that the group behaves like the group of independent translations in the new parameterization. One way to approach the problem is to start with the largest Abelian subgroup of the multi-parameter group for which the equivariant spaces can be constructed. The rest of the subgroups impose constraints on the equivariant space by way of the differential equations: $\bar{L}_i\Phi = B_i\Phi$. Thus, the equivariant measuring space for the multi-parameter group can be constructed by successively constraining the equivariant space of the largest Abelian subgroup.

Example 8 : Consider the multi-parameter group made up of translations in the x and y directions together with the group of rotations, i.e. g_{t_x}, g_{t_y} and g_r respectively. The largest Abelian subgroup is the two-parameter group of translations in the x and y directions. The equivariant space for this group is: $\text{span}(\Phi) = \text{span}(x^p y^q e^{\alpha x+\beta y})$ for $0 \leq p \leq m$ and $0 \leq q \leq l$. The group of rotations yields the additional constraint that $\bar{L}_r\Phi = B_r\Phi$ where $\bar{L}_r = -x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x}$. By observation, we can rule out the exponentials $e^{\alpha x+\beta y}$ (i.e. $\alpha = \beta = 0$) since applying $\bar{L}_r$ to each term raises the power of the monomial factor by one each time; repeated application of the conjugate generator will raise the power without bound. Applying $\bar{L}_r$ to the monomial $x^p y^q$, however, raises the power in one variable and decreases the power in the other. Successive applications will result in one of the variables being reduced to zero. Hence, $\{x^p y^q\}$ is an equivariant space under this group where $0 \leq p+q \leq m$ for some m.

Throughout the paper, equivariant measuring spaces are constructed in two steps: (1) B_i matrices are selected, (2) equivariant spaces are derived using the exponential map. Alternatively, one could begin by selecting an interpolation matrix and deriving from it the systems of differential equations. The interpolation matrix, parameterized by the group parameters $\boldsymbol{\tau}$, describes a family of matrices that together with matrix multiplication forms a matrix group [15]. This matrix group is known as the (linear) representation of the transformation group. In many cases, such representations have been derived [26]. Given the interpolation matrix $A(\boldsymbol{\tau})$, the interpolation equation (in terms of the measuring function) is:

$$\bar{g}(\boldsymbol{\tau})\,\Phi = A(\boldsymbol{\tau})\,\Phi.$$

Taking derivatives on both sides with respect to the parameters $\boldsymbol{\tau}$ at $\boldsymbol{\tau} = 0$ results in the following k systems of differential equations:

$$\bar{L}_i\,\Phi = B_i\,\Phi \qquad \text{where} \qquad B_i = \frac{\partial}{\partial \tau_i} A(\boldsymbol{\tau})\bigg|_{\boldsymbol{\tau}=0}.$$

Thus, the matrices B_i are determined from the interpolation matrix A. Solving these systems of differential equations for Φ yields a function space that is equivariant under the group.

Subsequently, the corollaries at the end of Section 4 can be used to create larger equivariant spaces.

Example 9 : Consider the two-parameter group of translations and scalings in the x-direction:

$$g(\tau_1, \tau_2)\ s(x, y) = e^{-\tau_1}\ s(e^{-\tau_1} x - \tau_2, y).$$

An interpolation matrix A for this group is the following:

$$A(\tau_1, \tau_2) = \begin{pmatrix} e^{-\tau_1} & \tau_2 \\ 0 & 1 \end{pmatrix}.$$

Notice that composition of two transformations $g(\tau_1^2, \tau_2^2)\ g(\tau_1^1, \tau_2^1)$ resembles the multiplication of the two corresponding interpolation matrices $A(\tau_1^2, \tau_2^2)\ A(\tau_1^1, \tau_2^1)$:

$$g(\tau_1^2, \tau_2^2)\ g(\tau_1^1, \tau_2^1) = g(\tau_1^1 + \tau_1^2, e^{-\tau_1^2}\tau_2^1 + \tau_2^2)$$

and

$$A(\tau_1^2, \tau_2^2)\ A(\tau_1^1, \tau_2^1) = A(\tau_1^1 + \tau_1^2, e^{-\tau_1^2}\tau_2^1 + \tau_2^2).$$

This is not a mere coincidence; in fact, all interpolation matrices are related to their groups in a similar way.[5] From the interpolation matrix $A(\tau_1, \tau_2)$ the conjugate generators can be derived:

$$\begin{aligned} \bar{L}_1 &\doteq \frac{\partial}{\partial \tau_1} A(\tau_1, \tau_2)\bigg|_{\tau_1=\tau_2=0} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = B_1 \\ \bar{L}_2 &\doteq \frac{\partial}{\partial \tau_2} A(\tau_1, \tau_2)\bigg|_{\tau_1=\tau_2=0} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = B_2 \end{aligned}$$

Subsequently, equivariant measuring functions are the solutions of the following systems of differential equations:

$$\begin{aligned} \bar{L}_1\ \Phi &= x\tfrac{\partial}{\partial x}\ \Phi &= B_1\ \Phi, \\ \bar{L}_2\ \Phi &= \tfrac{\partial}{\partial x}\ \Phi &= B_2\ \Phi. \end{aligned}$$

One solution to these equations is $\Phi(x) = (x, 1)^T$ where again each entry of Φ can be multiplied by a function solely in y. Using the corollaries at the end of Section 4, larger

[5] Mathematically speaking, interpolation matrices are matrix representations of their groups, i.e. there is a mapping π such that $\pi(\bar{g}(\boldsymbol{\tau})) = A(\boldsymbol{\tau})$ between the transformation group and the matrix group such that

$$\pi(g(\tau_1^2, \ldots, \tau_k^2)\ g(\tau_1^1, \ldots, \tau_k^1)) = \pi(g(\tau_1^2, \ldots, \tau_k^2))\ \pi(g(\tau_1^1, \ldots, \tau_k^1)).$$

This mapping is called a homomorphism. If, in addition, the mapping is one-to-one, then it is called an isomorphism.

Groups (# parameters)	Equivariant Measuring Space
x, y-translation (2)	$\{x^p y^q e^{\alpha x+\beta y}\}$ for $0 \le p \le m$ and $0 \le q \le l$.
x, y-scaling (2)	$\{x^\alpha y^\beta \ln(x)^p \ln(y)^q\}$ for $0 \le p \le m$ and $0 \le q \le l$.
Rotation Uniform-scaling (2)	$\{(\ln r)^p \theta^q e^{\alpha\theta+\beta ln(r)}\}$ for $0 \le p \le m$ and $0 \le q \le l$.
x, y-translation Rotation (3)	$\{x^p y^q\}$ for $0 \le p+q \le m$.
x, y-translation x, y-scaling (4)	$\{x^p y^q\}$ for $0 \le p \le m$ and $0 \le q \le l$.
x, y-translation x, y-scaling Rotation (5)	$\{x^p y^q\}$ for $0 \le p+q \le m$.

Table 2: Several examples of multi-parameter groups and their equivariant measuring spaces.

equivariant spaces can be constructed. In particular, repeated application of the Kronecker product of $\Phi(x)$ with itself shows that $\{x^p\}$ for $0 < p < m$ is a measuring space, where m is some integer.

Table 2 is a summary of several common multi-parameter groups and their equivariant measuring spaces.

7 Approximating Equivariant Spaces

Often, one may be interested in designing a filter kernel ψ with some desired properties. In order to be able to predict the outputs of ψ under some given transformation group, ψ must be included in some equivariant measuring space. If there is no such equivariant space, the filter kernel can first be approximate by a set of equivariant measuring functions:

$$\psi \approx \sum c_i \phi_i = \mathbf{c}^T \Phi \ .$$

Then, a transformed version of ψ is approximated via the interpolation equation:

$$\bar{g}\psi \approx \mathbf{c}^T A \Phi \ .$$

A compactly-supported function is a function that is non-zero only over some compact region of its domain, and zero everywhere else. A non-compact transformation group refers to a group whose parameter space is non-compact. For example, the group of translations is non-compact since its parameter space is $\mathbf{R}$ while the group of rotations whose parameter space is S^1 is compact. For compactly-supported functions, there are no finite-dimensional function spaces that can be used to steer these functions under a non-compact transformation group. The simple example of steering a raised cosine under translation is illustrative of this point: in order to steer a raised cosine under translation, an infinite number of raised cosines are needed.

Fortunately, if only local steerability is desired, i.e., only a limited range of transformation is necessary, then a finite number of functions might be sufficient to steer a compactly-supported function. We suggest two possible approximation methods:

Function Approximation. As before, the function to be steered is first approximated using an appropriate equivariant function space. This approximation is then steered by steering the basis functions spanning the space. However, since only local steerability is desired, the domain over which the function is approximated need only be a subset of its actual domain:

$$\left|\psi(x) - \mathbf{c}^T\ \Phi(x)\right|^2 < \varepsilon \quad \text{for} \quad x \in R_{\text{approx}}$$

where ε is some small value. the size of R_{approx} depends on the range of parameter space over which local steerability is expected.

Intuitively, we need to approximate the function over a large enough subset of its domain such that all transformed replicas of it will also be adequately approximated. For example, consider the problem of steering a one-dimensional raised cosine under translation. The raised cosine is compactly-supported over the interval $[-1, 1]$. The range of translations over which it is to be steered is $[-1, 1]$. Thus, the union of the support of all possible translated raised cosines is $[-2, 2]$. We refer to this interval as the *integration region* as this is the (fixed) interval of integration for a corresponding steerable filter. Clearly, then, the original raised cosine needs to be well approximated over this interval $[-2, 2]$. Unfortunately, approximating it over this interval is not enough. When the raised cosine is translated to the left by -1, for example, the interval $[2, 3]$ (the right tail) of the original raised cosine's domain enters the integration interval. If the original raised cosine is poorly approximated in this region, then the interval $[1, 2]$ of this translated raised cosine will be poorly approximated as well. The same holds when the raised cosine is translated to the right by 1. Hence, the original raised cosine needs to be well approximated over the interval $[-3, 3]$. We refer to this interval as the *approximation region*. The integration region is a subset of the approximation region; the compact support of the original function is, in turn, a subset of the integration region. Figure ?? illustrates the approximation and integration regions for a one-dimensional function steered under translation.

The integration and approximation can also be defined mathematically. We assume that the transformations are smooth and locality of steerability implies steerability within a compact region of parameter space $G \subset \mathbf{R}^k$. Let R_f be the compact support of the original function outside of which it is zero. The integration region is therefore:

$$R_{\text{int}} = \bigcup_{\boldsymbol{\tau} \in G} \bar{g}(\boldsymbol{\tau})\ R_f$$

where the union is taken over the compact region of parameter space. The application of the group operator to the region R_f produces the corresponding region of the transformed function. The approximation region is defined in terms of the inverse of the conjugate group:

$$R_{\text{approx}} = \bigcup_{\boldsymbol{\tau} \in G} g(\boldsymbol{\tau})\ R_{\text{int}}$$

For example, assume the permissible transformations consist of all the translations $g_{t_x}(\tau)$ where $\tau \in [-d, d]$ (see Figure 4). In this case, the integration interval will be $R_{\text{int}} = [a-d, b+d]$ and the approximation interval is $R_{\text{approx}} = [a-2d, b+2d]$. In the case where the translation parameter is $\tau \in [0, d]$, the integration interval will be $R_{\text{int}} = [a, b+d]$ and the approximation interval $R_{\text{approx}} = [a-d, b+d]$. Note that if the equivariant measuring functions are cyclic over R_{approx} it is possible to take advantage of the repetitive behavior of the approximated function $\mathbf{c}^T\Phi$. Replicas of the approximated function are tiled in the spatial domain and problem occurs when a neighbor replica is transformed into the integration interval. To overcome this problem, it is enough to approximate the filter kernel only over the integration interval.

Using the approximation $\psi(x) \approx \mathbf{c}^T\Phi$, approximated over R_{approx}, the transformed function $\bar{g}\ \psi(x)$ is now well approximated over the integration interval as long as the transformation parameters are in G. Therefore it is enough to calculate the inner-product $\langle \bar{g}\psi, s\rangle$ within this range:

$$\begin{aligned} \textstyle\int_{-\infty}^{\infty}(\bar{g}(\boldsymbol{\tau})\ \psi(x))s(x)\ dx &= \textstyle\int_{R_{\text{int}}}(\bar{g}(\boldsymbol{\tau})\ \psi(x))s(x)\ dx \\ &\approx \mathbf{c}^T A(\boldsymbol{\tau})\ \textstyle\int_{R_{\text{int}}}\Phi(x)s(x)\ dx \\ &= \mathbf{c}^T A(\boldsymbol{\tau})\mathbf{f} \end{aligned}$$

where the first equality holds since $\bar{g}(\boldsymbol{\tau})\ \psi(x)$ is zero outside R_{int} for all $\boldsymbol{\tau} \in G$. This implies that the measuring functions $\{\phi_i\}$ can be treated as compactly supported functions whose values are zero outside the integration interval. Note, that although the function is approximated over R_{approx} it is enough to store the approximation only over R_{int}.

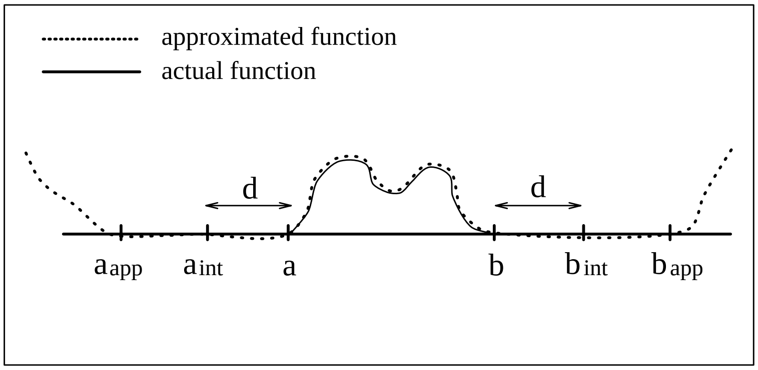

Figure 4: A function (solid line) with a compact support $R_f = [a, b]$. The function is approximated (dotted line) over a wider interval $R_{\text{approx}} = [a_{app}, b_{app}]$ outside of which the approximation might be poor. The integration of this approximation is performed over the integration interval $R_{\text{int}} = [a_{int}, b_{int}]$. Note that if the approximated function is x-translated by $-d$ to d units, the approximation is always precise within the integration interval.

Figure 5 shows several translations of a Gaussian having s.t.d.=3 (upper row) and their corresponding approximations using 10 sets of measuring functions $\{\sin(kx), \cos(kx)\}$ (lower row). The measuring functions are equivariant with respect to x-translation and the translations of the approximated Gaussian were calculated using the interpolation equation. The Gaussian has a standard-deviation (s.t.d.) of 3 units and the translations were performed over a bounded range of parameters (from 0 to 15 units, i.e. up to 5 s.t.d's). The Gaussian is approximated over an interval such that all translations of the Gaussian are effectively zero outside this interval. In this case, the approximation interval was $[0, 60]$ since we used the cyclic behavior of the measuring functions. As demonstrated by the figure the real translations and the approximated translations are quite similar to each other. However, the quality of the approximations depends on several factors:

- The number of equivariant measuring functions used to approximate the filter kernel. It is obvious that the larger the number of measuring functions used, the better the approximation. Figure 6 shows the approximation of the same translated Gaussian, now using only 5 sets of measuring functions. The deficiency in the approximation quality is visible.
- The support of the filter kernel in the canonical coordinates (where the transformation resembles a translation). If the support is small, a large number of measuring filters is required in order to obtain a reasonable approximation. This stems from the fact that functions that are equivariant with respect to a non-compact transformation group have infinite support (even if we use them as compact support functions). Therefore, a

large number of such measuring functions are necessary in order to compose a narrow support filter kernel. This is well knows in the Fourier domain where the support of the transform is inversely proportional to the function's support. Figure 7 shows the approximations of a translated Gaussian using the same number of measuring functions as in Figure 5 (i.e. 10 sets of $\{\sin(kx), \cos(kx)\}$). However, in this case the s.t.d. of the Gaussian is 1 unit. The degradation in the approximations is apparent.

- The amount of transformation. The measuring functions approximate the filter kernel in its original position: $\psi \approx \mathbf{c}^T\Phi$. Since the measuring functions are equivariant, $\bar{g}(\mathbf{c}^T\Phi) = \mathbf{c}^T A\Phi$, i.e. the approximation $\mathbf{c}^T\Phi$ is precisely transformed by the interpolation equation. Therefore, the function $d = \psi - \mathbf{c}^T\Phi$ which is the difference between the real kernel and its approximation, is transformed as well:

  $$\bar{g}d = \bar{g}\psi - \mathbf{c}^T A\Phi \quad .$$

  As a result, the behavior of d depends solely on the characteristics of the transformation. If the conjugate transformation ensures that the energy of a transformed function remains unchanged, the quality of the approximation (i.e. $\int(\bar{g}d)^2 dxdy$) will be constant. This is the case for the the transformations listed in Table 1.

  The constant quality of the approximation is demonstrated in Figure 8. The two graphs show the sum of squared-differences (s.s.d.) between the original Gaussian and its approximation versus various translations. The graphs show the s.s.d. for several cases. In the right graph the s.s.d. is plotted for several Gaussians having different s.t.d. values. All of them were approximated with 4 harmonics. The left graph plots the s.s.d. for different numbers of harmonics. In all cases, the calculated s.s.d. is constant for all permissible transformations. This behavior is demonstrated as well in Figure 10 which shows the s.s.d. of a Gaussian under scaling (see Figure 9 for details). Recall that the s.s.d. values under scaling transformation $e^\tau s(e^\tau x, y)$ are constant due to the the coefficient multiplication which compensates for the changes in the image energy. Note, however, the differences between the graphs of Figure 10 and those of Figure 8. First, the s.s.d. plots starts to increase at high scale factor values. This result hints that the approximation interval that has been used is too small with respect to high scale factor values. Expanding it will flatten the s.s.d. plots. Second, changing the s.t.d. does not influence the approximation quality as it does in the translation case. The reason for this phenomena is that changing the s.t.d. does not influence the support of the function in the canonical coordinates $(\ln x)$. Hence, the differences in the approximation are not significant.

The necessity to bound the range of transformations does not arise if the transformation group is compact. Rotation is an example of such a group since the rotation parameter is bounded in $[0, 2\pi)$. In a similar manner, a rotation of any polar separable function can be interpolated by approximating its angular component by an equivariant space of sinusoids in angular coordinates. This was recently proposed by Simoncelli and Farid [24] as a way of constructing steerable wedge filters.

Figure 11 reports the error in numerically approximating translates of a Gaussian by a set of singular vectors using the method suggested by Perona [22]. The total error associated with this scheme is less than the function approximation technique; singular vectors are computed so as to minimize the total error $\int(\bar{g}\psi - \mathbf{c}^T\Phi)^2 dx\ dy\ d\tau$. However, the maximum error introduced for this technique can be higher than the above, especially for cases where a large range of transformations is permissible. This can be seen by comparing Figure 11 with Figure 8. Furthermore, in the function approximation method, the interpolation function can be derived analytically. In Perona's scheme, however, the interpolation function are computed numerically. This is a critical issue in motion estimation problems. Using an analytical interpolation equation, motion parameters can be solved analytically. In the numerical approach, however, the motion parameters should be found using a search scheme.

The constant quality of the approximated filter kernel is a big advantage of the function approximation method. Using it, one can estimate the deviation of the approximation from the original filter kernel under any (permissible) transformation. If a higher accuracy is required, additional measuring functions can be added to the approximation. The disadvantage of this technique is twofold: First, it cannot be applied if an equivariant measuring space does not exist for the given transformation. Second, it may be necessary to use a large number of measuring functions in order to get a reasonable approximation. The approximation technique described in the next subsection tries to overcome these disadvantages.

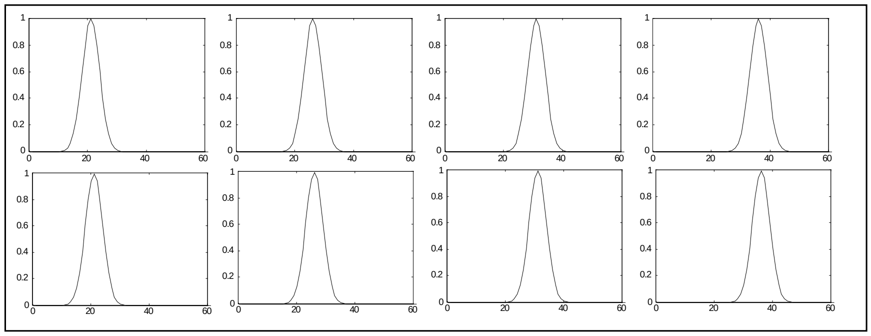

Figure 5: Upper row: Several translates of a Gaussian having s.t.d.=3. Lower row: The translated Gaussians as approximated using 10 sets of $\{\sin(kx), \cos(kx)\}$ with compact support.

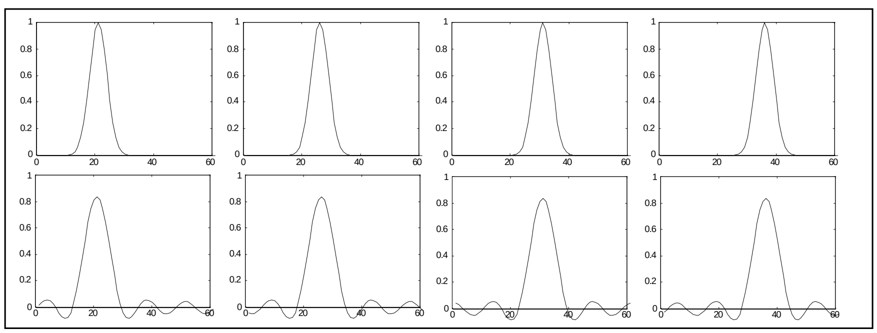

Figure 6: Upper row: Several translates of a Gaussian having s.t.d.=3. Lower row: The translated Gaussians as approximated using 5 sets of $\{\sin(kx), \cos(kx)\}$ with compact support.

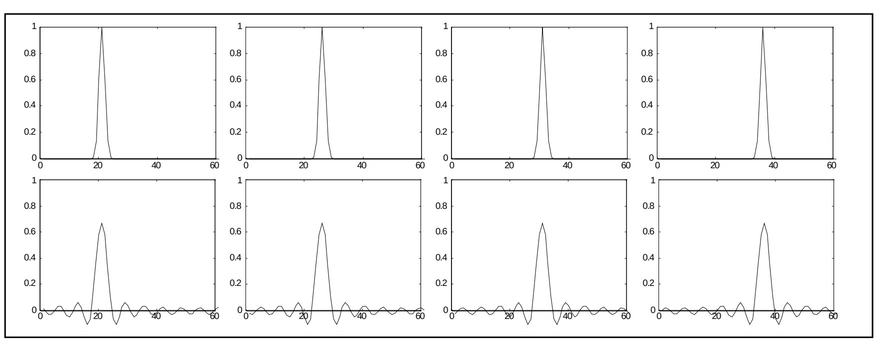

Figure 7: Upper row: Several translates of a Gaussian having s.t.d.=1. Lower row: The translated Gaussians as approximated using 10 sets of $\{\sin(kx), \cos(kx)\}$ with compact support.

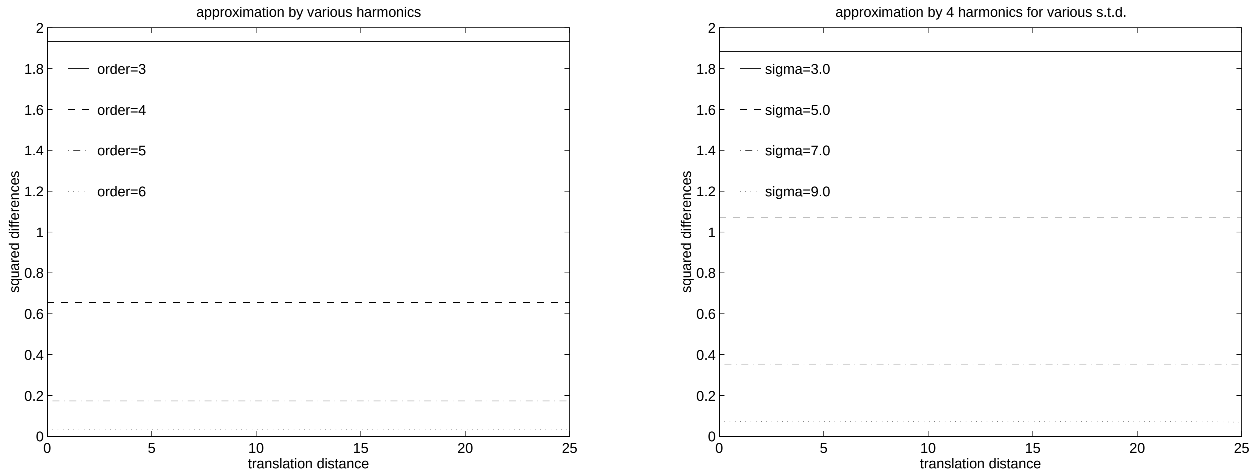

Figure 8: Error in approximating the Gaussian with $\{\sin(kx), \cos(kx)\}$. In the left graph "order" indicates the number of measuring functions used in the approximation. In the right graph, "sigma" refers to the standard deviation of the Gaussian.

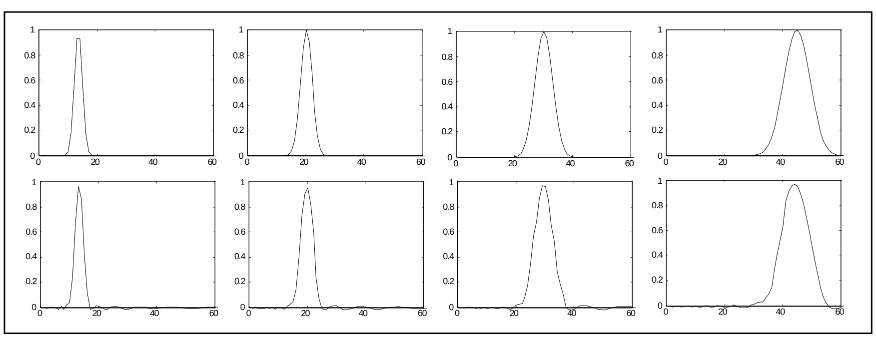

Figure 9: Upper row: A Gaussian with $s.t.d = 3$ and its scalings relative to the origin. The scale factors are (from left to right) $e^{-0.8}, e^{-0.4}, 1, e^{0.4}$. Lower row: The Gaussian and its scalings as approximated with 15 sets of $\{\sin(k \ln x), \cos(k \ln x)\}$ with compact support.

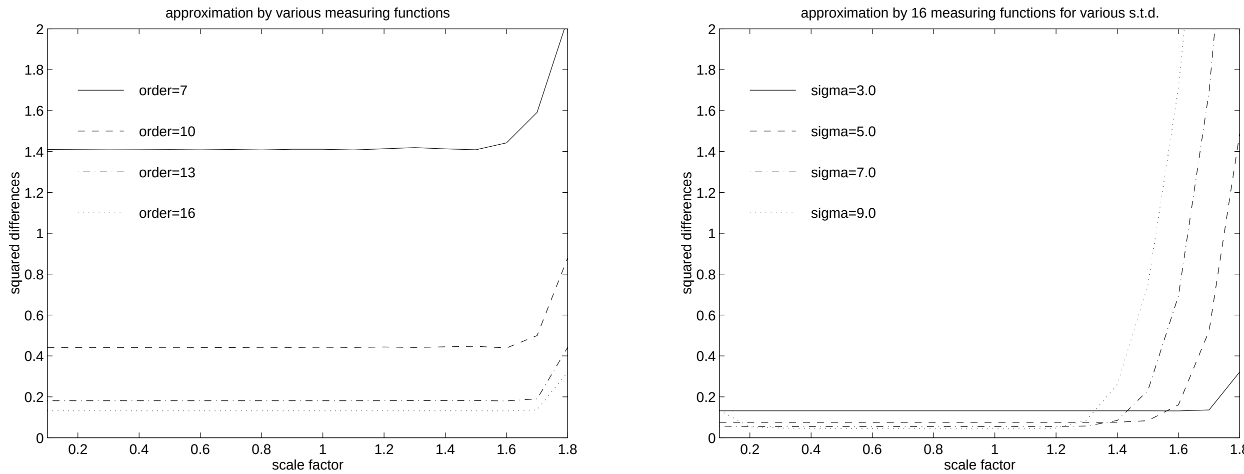

Figure 10: Error in approximating the Gaussian with a set of $\{\sin(k \ln x), \cos(k \ln x)\}$. In the left graph "order" indicates the number of measuring functions used in the approximation. In the right graph, "sigma" refers to the standard deviation of the Gaussian.

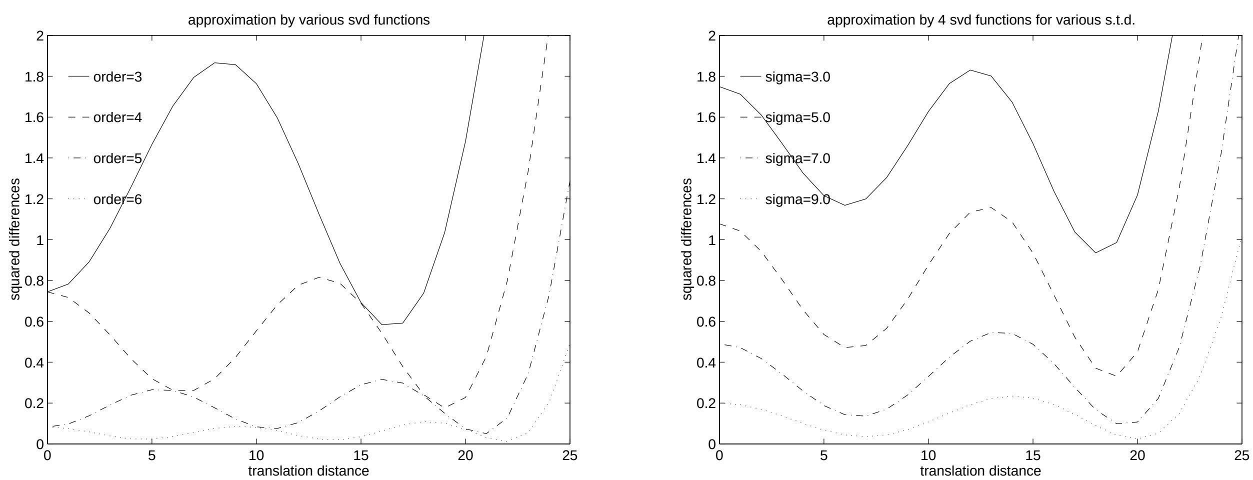

Figure 11: Error in numerically approximating translates of a Gaussian by a set of singular vectors that correspond to the n^{th} largest singular values of the singular value decomposition. In the left graph, "order" refers to the number of singular vectors used. In the right graph, "sigma" represents the standard deviation of the Gaussian.

Interpolation Approximation. In contrast with the previous method which approximates the filter kernel with an equivariant measuring space, the following technique approximates the equivariant space itself. If Φ is a vector of measuring functions, it is equivariant if $\bar{L}_i \Phi = B_i \Phi$. Approximating the equivariant measuring space means that the above equality is not precise i.e. $\bar{L}_i \Phi \approx B_i \Phi$. On one hand we aim to minimize the equivariant approximation error $E_{space} = |\bar{L}_i \Phi - B_i \Phi|^2$ but on the other hand we would like to minimize the kernel function approximation $E_{func} = |\mathbf{c}^T \Phi - \psi|^2$. In the previous technique, E_{space} was zero since we approximated ψ with a fully equivariant measuring space. In the following technique E_{func} is zero, however, the equivariance criterion is relaxed. What is the consequence of approximating equivariant spaces? Since $\bar{L}_i \Phi \approx B_i \Phi$ it follows that $\bar{g}_i \Phi \approx e^{B_i \tau_i} \Phi$, meaning that the interpolation equation is now approximated and not precise.

Assume at this point that we are dealing with a one parameter transformation group $g(\tau)$ with the conjugate generator $\bar{L}$. We construct a vector of measuring functions as follows:

$$\Phi = (\psi, \bar{L}\psi, \bar{L}^2\psi, \ldots, \bar{L}^{n-1}\psi)^T$$

Using it, the equivariant equation can be approximated:

$$\bar{L}\Phi \approx B\Phi \quad \text{where} \quad B = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

The approximation error stems from the n^{th} equation where the term $\bar{L}^{n-1}\psi$ is equated to zero which is not necessarily true. The matrix B is in Jordan form having one $n \times n$ block with a corresponding eigenvalue $\lambda = 0$. The interpolation equation for Φ is approximated using B:

$$\bar{g}\Phi \approx e^{B\tau}\Phi \quad \text{where} \quad e^{B\tau} = \begin{pmatrix} 1 & \tau & \frac{1}{2!}\tau^2 & . \\ & 1 & \tau & . \\ & & . & . \\ & & & 1 \end{pmatrix},$$

The first equation in this system is the interpolation equation for ψ: $\bar{g}(\tau)\psi \approx (e^{B\tau}\Phi)_1$ where subscript 1 refers to the first entry in the vector. It is easy to see that this equation gives the n^{th} order Taylor expansion of $\bar{g}(\tau)\psi$ about $\tau = 0$. Therefore, similar to the expansion, the higher the order (dimension of Φ) the better the approximation.

The situation with multi-parameter groups is analogous. The vectorial function Φ is composed of the following sets of measuring functions:

$$S_0 \ = \ \psi$$

$$
\begin{aligned}
S_1 &= \{\bar{L}_i S_0\} \\
&\vdots \\
S_{n-1} &= \{\bar{L}_i S_{n-2}\}
\end{aligned}
$$

and the associated B_i's are generated accordingly. Care should be taken in the multi-parameter as well as in the one-parameter groups that the set of generated measuring functions will not include linearly dependent functions.

Figure 12 shows the s.s.d. of a translated Gaussian using several of its derivatives. The error is extremely small for translations of several units, but increases rapidly for farther translations. Therefore, this technique is appropriate only for small transformations. In fact, translations of a kernel function with finite support cannot be approximated beyond the range of support of the original function; all the derivatives outside the support vanish, and any linear combination of these derivatives is zero. Therefore the radius of convergence of the Taylor expansion for such a function is at most the range of the support. The Gaussians presented in our examples are subject to the same problem due to their similarity to compact support functions. As a result, the approximation error increases for smaller s.t.d. as shown in Figure 12.

The Gaussian and its derivatives have been widely used in computer vision. In motion estimation, Manmatha and Oliensis [18] suggested a method for extracting the local affine deformations of an image using Gaussians and its derivatives. Recently, Liu et al. [16] proposed a method of estimating optical flow using Hermite polynomials (which are derivatives of a Gaussian) in three dimensions. Earlier on, the usefulness of Gaussians and its derivatives in representing local geometry have also been recognized [14, 27, 28].

In the following examples, interpolation equations for the filter kernel $\psi = \cos(y)G(\sigma)$ were approximated, where G is a two-dimensional Gaussian. In these examples $\sigma = 2.5$. The interpolation equation was calculated for a 5-parameter transformation including: x-scaling, y-scaling, x-translation, y-translation and rotation. All the generators $\bar{L}_i$ applied to ψ yield a linear sum of terms of the form $\{x^k y^l \cos(y)G(\sigma), x^k y^l sin(y)G(\sigma)\}$. In these examples Φ was composed of such terms up to the 3^{rd} order. All together, 18 terms were used in the measuring vector. Figures 13-17 show several one parameter transformation sub-groups. Figure 18 demonstrates interpolations for a two-parameter sub-group.

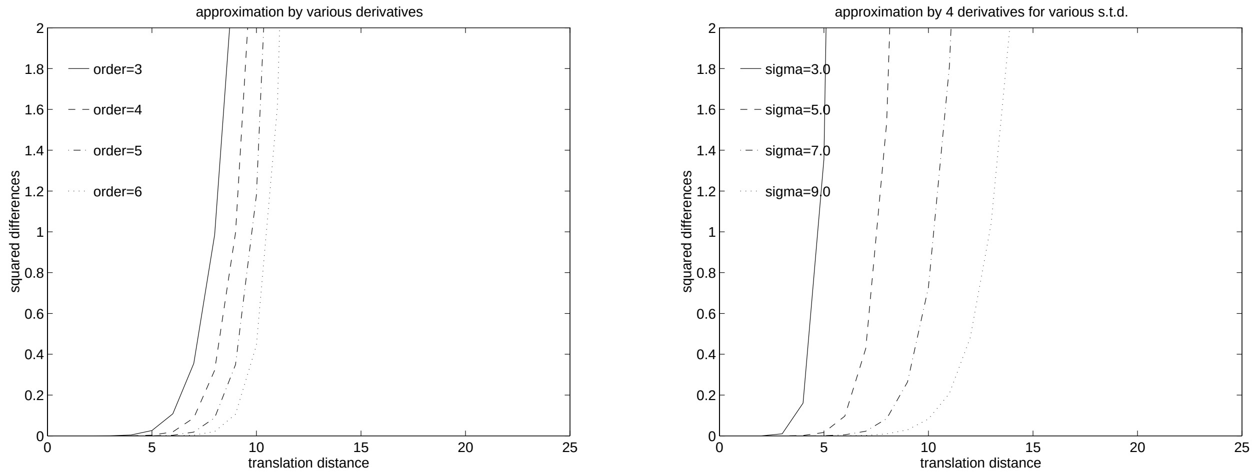

Figure 12: Error in approximating the Gaussian with several of its derivatives. In the left graph "order" indicates the highest derivative used in the approximation. In the right graph "sigma" refers to the standard deviation of the Gaussian.

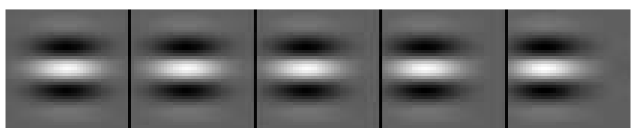

Figure 13: The function $\psi = \cos(x)G(2.5)$ and several of its translations as approximated by the interpolation approximation. The translations are (left to right) 0, 0.75, 1.5, 2.25, and 3 units in the x direction.

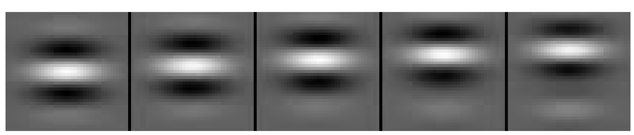

Figure 14: The function $\psi = \cos(x)G(2.5)$ and several of its translations as approximated by the interpolation approximation. The translations are (left to right) 0, 0.75, 1.5, 2.25, and 3 units in the y directions.

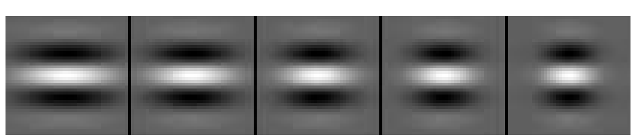

Figure 15: The function $\psi = \cos(x)G(2.5)$ and several of its scalings as approximated by the interpolation approximation. The scale factors are (left to right) $e^{-0.3}$, $e^{-0.15}$, 1, $e^{0.15}$, and $e^{0.3}$ in the x coordinate.

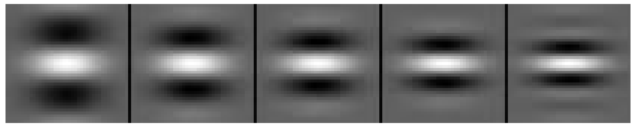

Figure 16: The function $\psi = \cos(x)G(2.5)$ and several of its scalings as approximated by the interpolation approximation. The scale factors are (left to right) $e^{-0.3}$, $e^{-0.15}$, 1, $e^{0.15}$, and $e^{0.3}$ in the y coordinate.

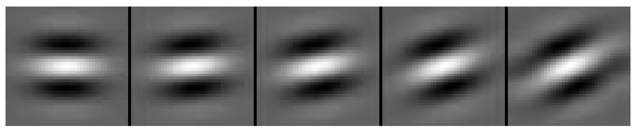

Figure 17: The function $\psi = \cos(x)G(2.5)$ and several of its rotations as approximated by the interpolation approximation. The rotations are (left to right) 0.1, 0.2, 0.3 and 0.4 radians.

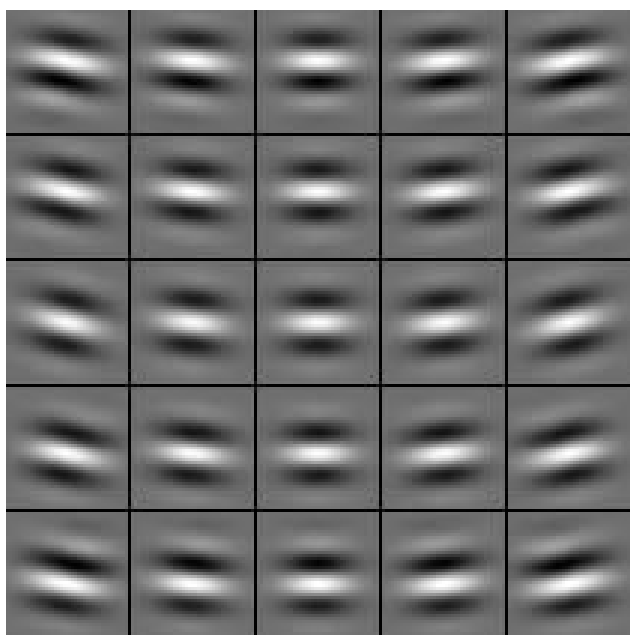

Figure 18: The function $\psi = \cos(x)G(2.5)$ and several of its two-parameter transformations $\bar{g}_r(\tau_1)\bar{g}_{ty}(\tau_2)\psi$. The transformations are rotations of (left to right) -0.2, -0.1, 0, 0.1, 0.2 radians and y-translations of (bottom up) -1.5, -0.75, 0, 0.75, 1.5 units.

8 Motion Estimation using Equivariant Spaces

In our framework, the steerability problem is expressed as a forward (or interpolation) problem where measurable features of a transformed image are interpolated from a set of features measured in the original image. This interpolation is performed using the interpolation equation $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\mathbf{f}$ where $\mathbf{f}$ and the motion parameters $\boldsymbol{\tau}$ are known. This forward problem is easy to solve once the interpolation equation is constructed; it only requires the substitution of the known parameters into the interpolation equation. Motion estimation, on the other hand, is regarded as an inverse (or estimation) problem where the motion parameters are estimated given measured features $\mathbf{f}$ and $\hat{\mathbf{f}}$ from the original and the transformed images. The complexity and robustness of the estimation depends on the nature of the interpolation matrix $A(\boldsymbol{\tau})$ and in most cases requires a non-linear minimization process.

Motion estimation, in this framework, is not restricted to infinitesimal changes between images; we consider finite transformations as well. However, as stated, we are dealing only with transformations that are Lie groups. In the cases where the entire motion of the image cannot be modeled by a transformation group, motion estimation is applied to local neighborhoods of the image. The local neighborhood is commonly called the estimation window within which the motion is assumed to be characterized by a particular transformation group. This transformation group is the motion model of the window. For every estimation window in the image, we try to fit motion parameters with respect to the motion model.

8.1 The Estimation Window

The size of the estimation window plays an important role in the estimation process. In order to construct a filter kernel whose support is appropriate for the estimation window, several equivariant measuring functions are linearly composed as explained in the previous section. If the real motion of the image is included in the motion model then the size of the window can be expanded and a small number of equivariant measuring functions are required for constructing the filter kernel. However, in most cases, the actual motion of the image is more complex than any transformation group, and the estimation window must be small. Recall, that the smaller the window size, the larger the number of measuring functions required.

Another decision which one should make is what transformation group to choose as the motion model. If the group has a large number of parameters then complex motions can be approximated by it, and accordingly, the estimation window can be expanded. However, the more flexible the transformation group, the more restricted its equivariant measuring functions. For example, the equivariant measuring functions for the translation+rotation group are the monomials $\{x^p y^q\}$ (see Table 2). Composing a filter kernel with these functions may required many measuring functions. However, this group has three degrees of freedom and is more flexible in approximating complex motions than the translation group alone.

In Phase-based motion estimation techniques [6, 8, 7] the translations between two images are estimated by the phase change of two quadrature pairs applied to the two images. In practice, two equivariant measuring functions having $B = \mathrm{diag}(\lambda, \bar{\lambda})$ are used. These measuring functions have infinite support. Since the real transformation is more complex than simple translation, the phase-based approaches bound the support of the measuring function by a Gaussian envelope (estimation window) with the hope that this will not influence the estimation results. However, it is impossible to contract the width of the Gaussian envelope too much due to what is called the "window problem". The "interpolation equation", in this case, translates the modulation of the measuring functions but leaves its envelope unchanged. This, of course, adds error to the true interpolation equation which increases as we contract the envelope. In our scheme, however, the envelope as well as the modulation are transformed by the interpolation equation. Therefore, the window problem is not introduced and the size of the approximation window can be adjusted as required. The price we pay for contracting the approximation window is that many equivariant measuring functions must be used to compose the narrow support filter kernel.

8.2 Solving for the Motion Parameters

Assume that a set of filter kernels $\psi_i, \quad i = 1, \cdots, k$ are chosen to find a one-parameter transformation between two images. These kernels are approximated by a linear set of equivariant measuring functions Φ (possibly having limited supports) as explained in the previous section:

$$\psi_i = \mathbf{c}_i^T \Phi \quad .$$

The first step in the motion estimation process is to apply the measuring functions to the two images obtaining two sets of measured features:

$$\mathbf{f} = \langle \Phi, s(x,y) \rangle \quad ; \quad \hat{\mathbf{f}} = \langle \Phi, g(\tau) s(x,y) \rangle \quad .$$

Since the measuring functions are equivariant the measured features are related by the interpolation equation: $\mathbf{c}_i^T \hat{\mathbf{f}} = \mathbf{c}_i^T e^{B\tau} \mathbf{f} \quad i = 1, \cdots, k$. Composing these equations for all the kernels gives a system of k equations:

$$C\hat{\mathbf{f}} = Ce^{B\tau}\mathbf{f} \quad \text{where} \quad C = [\mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_k}]^T \quad . \tag{7}$$

The motion parameter of the associated estimation window is obtained by solving for τ in this system, . In general, this is a non-linear system and some gradient minimization process must be used. Assume that we have an initial guess τ^0 of the motion parameter. Expanding $e^{B\tau}$ about the initial guess gives:

$$e^{B\tau} = e^{B\tau^0} + (\tau - \tau^0)Be^{B\tau^0} + O(\tau - \tau^0)^2$$

Substituting the linear terms of this expansion into Equation 7 yields a linear system:

$$C\hat{\mathbf{f}} \approx C(e^{B\tau^0} + (\tau - \tau^0)Be^{B\tau^0})\mathbf{f}$$

Defining:

$$\mathbf{a} \doteq C(\hat{\mathbf{f}} - e^{B\tau^0}\mathbf{f}) \quad ; \quad \mathbf{b} \doteq CBe^{B\tau^0}\mathbf{f} \quad \text{such that} \quad \mathbf{a} \approx \mathbf{b}(\tau - \tau^0)$$

the solution for τ using standard least-squares minimization gives:

$$\tau = \tau^0 + \frac{\mathbf{b}^T\mathbf{a}}{\mathbf{b}^T\mathbf{b}} \quad . \tag{8}$$

This process is repeated with the current solution serving as the new guess until convergence. However, the convergence of this process to the correct solution is not guaranteed and it depends on the quality of the initial guess. The same solution can be applied to multi-parameter groups as well. The multi-parameter version of Equation 7 is:

$$C\hat{\mathbf{f}} = C\Pi\mathbf{f} \tag{9}$$

where $\Pi \doteq \Pi_{i=1}^{k} e^{B_i\tau_i}$. If for any $X(\boldsymbol{\tau})$ the notation X^0 refers to $X(\boldsymbol{\tau} = \boldsymbol{\tau}^0)$, the least-squares minimization of Equation 9 gives:

$$\boldsymbol{\tau} = \boldsymbol{\tau}^0 + (\mathbf{b}^T\mathbf{b})^{-1}\mathbf{b}^T\mathbf{a} \quad .$$

where in this case:

$$\mathbf{a} = C(\hat{\mathbf{f}} - \Pi^0\mathbf{f}) \quad \text{and} \quad \mathbf{b} = C[B_1\Pi^0\mathbf{f}, \cdots, B_k\Pi^0\mathbf{f}] \quad .$$

8.3 Uniqueness of the Solution

In many cases, the motion parameters $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_k)$ can not be estimated uniquely, even theoretically. This occurs when the interpolation equation $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\mathbf{f}$ is not monotonic, i.e. different transformation parameters might give the same value for $\hat{\mathbf{f}}$. As mentioned above the interpolation matrix A is a matrix representation of the transformation group. This means that there is a mapping π between the transformation group and the matrix group such that

$$\pi(\bar{g}(\boldsymbol{\tau})) = A(\boldsymbol{\tau}) \quad .$$

If this mapping is one-to-one (isomorphism), the estimated transformation is unique. If the mapping is many-to-one (homomorphism), the estimation is not unique and the interpretation of the multiple solutions is ambiguous. In the case of homomorphism, we must apply heuristic considerations and choose one solution among many possible. However, the quality of the heuristic decision will improve if we reduce the ambiguity of the solution.

For example, assume we use two equivariant measuring functions $\Phi_0 = (\sin(k_0 x), \cos(k_0 x))^T$ to find the x-translation between two images. Since the B matrix associated with Φ_0 is

$$B = \begin{pmatrix} 0 & k_0 \\ -k_0 & 0 \end{pmatrix}$$

the interpolation equation relating the measured features is:

$$\hat{\mathbf{f}} = A(\tau)\mathbf{f} = e^{B\tau}\mathbf{f} \quad \text{where} \quad e^{B\tau} = \begin{pmatrix} \cos(k_0\tau) & \sin(k_0\tau) \\ -\sin(k_0\tau) & \cos(k_0\tau) \end{pmatrix} \quad .$$

The mapping from $g_{t_x}(\tau)$ to $A(\tau)$ is a homomorphism; The motion parameter τ is defined on the real line and mapped into a compact range: $\tau \rightarrow \tau \bmod \frac{2\pi}{k_0}$. Therefore, in this case, $A(\tau) = A(\tau + \frac{2\pi n}{k_0})$ and the solution for the motion parameter can be found up to modulus $\frac{2\pi}{k_0}$. However, the ambiguity of the solution can be reduced if we choose measuring functions with lower frequencies $\Phi_1 = (\sin(k_1 x), \cos(k_1 x))^T$ where $k_1 < k_0$. This indeed will broaden the distance between two possible solutions, however, the robustness of the solution will decrease. To see that, recall that around the correct solution: $\frac{d\hat{\mathbf{f}}}{d\tau} = B\hat{\mathbf{f}}$. Using this equation and assuming a small change in $\hat{\mathbf{f}}$ is denoted by $\delta\hat{\mathbf{f}}$, we obtain: $\delta\hat{\mathbf{f}} = B\hat{\mathbf{f}}\delta\tau$. This relation shows that the error in the approximated τ for a fixed error in $\hat{\mathbf{f}}$ is inversely proportional to the norm[6] of B. Therefore, if Φ_1 is chosen instead of Φ_0, the ambiguity of the solution will be reduced, but its sensitivity to noise will be more severe.

[6]The norm of a matrix B is $\lambda_{max}(B^T B)$ where λ_{max} refers to the highest eigen-value.

To overcome this trade off, it is possible to include several frequencies in the measuring functions and solve the entire system simultaneously. For example, in our case we compose a larger set of measuring functions: $\Phi = (\Phi_0^T, \Phi_1^T)^T$. It is easy to see that the interpolation matrix for Φ is a result of a mapping $\tau \rightarrow \frac{2\pi}{gcd(k_0,k_1)}$ where $gcd(k_0, k_1)$ stands for the *greatest common divisor* of k_0 and k_1. Of course, $\mathrm{gcd}(k_0, k_1) \leq k_0, k_1$ so the ambiguity of the solution is reduced without deteriorating the robustness of the solution.

9 Invariants in Equivariant Feature Spaces

Consider an n-dimensional equivariant feature space F and a k-parameter transformation group such that $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\,\mathbf{f}$ where $\boldsymbol{\tau} \in \mathbf{R}^k$. Recall that an orbit of an image $O_g(s)$ is defined by $O_g(s) = \{g(\boldsymbol{\tau})\ s \mid \boldsymbol{\tau} \in \mathbf{R}^k\}$ which is the set of images obtained by transforming the original image in all possible ways. Likewise, the orbit of a feature is defined by $O_g(\mathbf{f}) = \{A(\boldsymbol{\tau})\,\mathbf{f} \mid \boldsymbol{\tau} \in \mathbf{R}^k\}$. For a k-dimensional group and n measuring functions, this orbit forms a k-dimensional manifold (or surface) in the n-dimensional feature space. Indeed, there is a whole family of k-dimensional orbits filling the n-dimensional feature space, one for each class of images whose members are transformed versions of each other.

Invariant Functions. Two features in the feature space, computed from a pair of images related by a transformation in the group, lie on the same k-dimensional manifold.[7] Hence, functions of features that are constant over each manifold are invariant under the transformation, i.e. $h(\hat{\mathbf{f}}) = h(\mathbf{f}) = c$ for $\hat{\mathbf{f}}, \mathbf{f}$ on the same manifold. In general, determining functions which are invariant over arbitrary families of manifolds is difficult. However, the manifolds in equivariant feature spaces are far from arbitrary. This is because the matrix $A(\boldsymbol{\tau})$ is actually a k-dimensional matrix group, i.e. a group whose elements are matrices and whose composition and inverse operators are matrix multiplication and inverse respectively. As a result, we can employ yet another theorem from Lie theory which states that a function is invariant under a transformation group if and only if applying any infinitesimal generator of the group to it results in zero identically [21]. In our case, this implies that a function $h(\hat{\mathbf{f}}) = h(e^{\tau_k B_k} \cdots e^{\tau_1 B_1}\ \mathbf{f})$ is invariant under $g(\boldsymbol{\tau})$ if and only if

$$\bar{L}_i\ h(\hat{\mathbf{f}}) = B_i\ \hat{\mathbf{f}} \cdot \nabla h = 0 \tag{10}$$

[7]Two completely different images could have features on the same manifold or even the same features. The likelihood of this depends on how well the measuring space approximates both images. This is not a problem if we can be certain that the two images will be transformed versions of each other, which is true in several applications.

where $\nabla h = (\frac{\partial h}{\partial \hat{f}_1}, \ldots, \frac{\partial h}{\partial \hat{f}_n})^T$ and $1 \leq i \leq k$. A very good review about how to solve such a set of PDEs using differential forms can be found in [19].

Another way to approach the problem of constructing invariants in the feature space is through implicit representations of the feature orbits. The interpolation function $\hat{\mathbf{f}} = A(\boldsymbol{\tau})\,\mathbf{f}$ can be seen as a parametric description of the feature manifold $O_g(\mathbf{f})$ where $\boldsymbol{\tau}$ are the parameters. An implicit representation of this manifold gives a description which is independent of the parameters and thus invariant with respect to the transformation. The manifold $O_g(\mathbf{f})$, in this case, is represented by a set of functions $\{p_k(\hat{\mathbf{f}}) = 0\}$ whose variety coincides with the manifold itself. In particular, $p_k(\hat{\mathbf{f}}) = p_k(\mathbf{f})$ and thus these functions are invariant under $g(\boldsymbol{\tau})$. Actually, any function $h(p_1(\hat{\mathbf{f}}), p_2(\hat{\mathbf{f}}), \cdots)$ is constant over the orbit $O_g(\mathbf{f})$ and therefore invariant under the transformation group as well.

There are several techniques for implicitizing parametric descriptions. For example, a set of polynomials can be implicitized by constructing the Groebner bases of their ideal with a particular ordering [4]. The use of Groebner bases to generate invariants in computer vision have also been recently suggested by Werman and Shashua [29]. Looking at invariants as implicit representations of manifolds also allows one to determine the total number of independent invariants that can be generated. Since $O_g(\mathbf{f})$ is a k-dimensional manifold in an n-dimensional space, $n - k$ implicit equations are required to describe the manifold. This corresponds to the maximum number of independent invariants that can be generated. Note, that the space dimension, n, must be bigger or equal than k, the dimension of the manifold embedded in it. However, if $n = k$, the invariants $h(\hat{\mathbf{f}})$ that are constant over the feature manifold $O_g(\mathbf{f})$ are constant all over the n-dimensional space, and thus uninteresting. Therefore, generating interesting invariant for a k-dimensional group requires as least $k + 1$ measured features.

Example 10 : Consider the measuring space $\Phi(x) = (\frac{x^2}{2!}, x, 1)^T$ that is equivariant under translation such that $\bar{L}_{t_x}\,\Phi = B\Phi$ where

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \hat{\mathbf{f}} = \begin{pmatrix} 1 & \tau & \frac{\tau^2}{2!} \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{pmatrix} \mathbf{f} = A(\tau)\,\mathbf{f}.$$

By Equation 10, a function $h(\hat{f}_1, \hat{f}_2, \hat{f}_3)$ is invariant under translation if and only if

$$B\,\hat{\mathbf{f}} \cdot \nabla h = \hat{f}_2 \frac{\partial h}{\partial \hat{f}_1} + \hat{f}_3 \frac{\partial h}{\partial \hat{f}_2} = 0.$$

Since we are dealing in a one dimensional manifold in a three dimensional space, two independent solutions exist. These are the two functions: $h_1(\hat{\mathbf{f}}) = \hat{f}_3$ and $h_2(\hat{\mathbf{f}}) = \hat{f}_1\hat{f}_3 - \frac{1}{2}\hat{f}_2^2$.

Actually, any function $h^\star(h_1, h_2)$ is invariant with respect to translation. It is straightforward to verify that h_2 is an invariant:

$$\begin{aligned}
\hat{f}_1 \hat{f}_3 - \tfrac{1}{2}\hat{f}_2^2 &= f_3(f_1 + \tau f_2 + \tfrac{\tau^2}{2!} f_3) - \tfrac{1}{2}(f_2 + \tau f_3)^2 \\
&= f_3(f_1 + \tau f_2 + \tfrac{\tau^2}{2!} f_3) - \tfrac{1}{2}(f_2^2 + \tau^2 f_3^2 + 2\tau f_2 f_3) \\
&= f_3 f_1 - \tfrac{1}{2} f_2^2 \; .
\end{aligned}$$

Example 11 : Let $\Phi(x) = (\cos kx, \sin kx)^T$ be a measuring space that is equivariant under translation such that $\bar{L}_{t_x}\, \Phi = B\Phi$ where

$$B = \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix}.$$

By Equation 10, a function $h(f_1, f_2)$ is invariant under translation if and only if

$$B\, \hat{\mathbf{f}} \cdot \nabla h = -k\hat{f}_2 \frac{\partial h}{\partial \hat{f}_1} + k\hat{f}_1 \frac{\partial h}{\partial \hat{f}_2} = 0.$$

Solving, we get $h(\hat{f}_1, \hat{f}_2) = h^\star(\hat{f}_1^2 + \hat{f}_2^2)$ for any function $h^\star$, in particular for the identity function. Hence, we have verified that the sum of squares of the inner-product of any pair of the Fourier basis functions with a signal is invariant under translation. Furthermore, all invariants can be written in terms of $\hat{f}_1^2 + \hat{f}_2^2$.

Example 12 : One of the simplest ways to construct an invariant is to construct two closely related measuring spaces Φ_1, Φ_2 using matrices B_i and $-B_i^T$ such that $\bar{L}_i\, \Phi_1 = B_i\, \Phi_1$ and $\bar{L}_i\, \Phi_2 = -B_i^T\, \Phi_2$. Recall that the interpolation matrix for Φ_1, Φ_2 are

$$\begin{aligned}
A_1(\boldsymbol{\tau}) &= e^{\tau_k B_k} \cdots e^{\tau_1 B_1}, \\
A_2(\boldsymbol{\tau}) &= e^{-\tau_k B_k^T} \cdots e^{-\tau_1 B_1^T}
\end{aligned}$$

respectively. As a result, the inner-product of the feature vectors $\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2$ corresponding to Φ_1, Φ_2 is an invariant:

$$\begin{aligned}
\hat{\mathbf{f}}_1^T\, \hat{\mathbf{f}}_2 &= (A_1(\boldsymbol{\tau})\, \mathbf{f}_1)^T\, A_2(\boldsymbol{\tau})\, \mathbf{f}_2 \\
&= (e^{\tau_k B_k} \cdots e^{\tau_1 B_1}\, \mathbf{f}_1)^T\, e^{-\tau_k B_k^T} \cdots e^{-\tau_1 B_1^T}\, \mathbf{f}_2 \\
&= \mathbf{f}_1^T e^{\tau_1 B_1^T} \cdots e^{\tau_k B_k^T}\, e^{-\tau_k B_k^T} \cdots e^{-\tau_1 B_1^T}\, \mathbf{f}_2 \\
&= \mathbf{f}_1^T\, \mathbf{f}_2
\end{aligned}$$

where the third equality follows from the identity $(e^A)^T = e^{A^T}$. This method is similar to the technique of generating invariant kernels suggested by Segman et. al. [23].

Numerous techniques for computing invariants on features like points and derivatives have been proposed by researchers in the past (e.g. [20]). In general, many of these techniques can be readily applied to construct invariants on equivariant feature spaces since these spaces are finite-dimensional. The method suggested above is very similar to the one used by Moons et al. [19] to construct invariants on points and derivatives. Another simple method of constructing polynomial invariants was recently proposed by Keren [13] in which instead of seeking to determine all the possible invariants, the author describes a procedure for symbolically deriving polynomial invariants of a given order. The method can also be employed in this context to derive polynomial invariants over feature vectors $\mathbf{f}$ (or even over prolongations, i.e. multiple feature vectors).

Invariant Feature Detection. Invariant feature or pattern detectors are used to identify specific patterns like edges and corners in an image independent of some family of image transformations. Within the framework, image invariants are computed in two-stages:

1. A set of equivariant measuring functions is chosen so that their inner-products with the given pattern will yield a characteristic signature that can be used to identify or discriminate it from other patterns.
2. A sufficient number of independent invariant functions over the feature space are computed so that the characteristic signature can be identified regardless of the pattern transformation.

Generating invariants in this manner has the advantage that one can construct equivariant measuring spaces that are rich enough to fully characterize a given pattern. This is done independent of the invariant functions which are only derived later. Since the dimension of the feature space is finite and relatively small, we can easily compute all the invariants associated with the given equivariant measuring space. Furthermore, traditional point-based techniques for computing invariance can also be used.

10 Equivariance in Point Coordinates

Throughout this paper measured features are calculated from grey scales values of an image. In many cases however, the available information in an image is a set of point coordinates rather than grey-scale values. The presented framework for steerability, motion estimation, and invariants can easily be generalized to include features measured from point coordinates.

In order to see the connection between a signal $s(x,y)$ and point coordinates, assume that each point $\mathbf{p}_i = (x_i, y_i)$ is represented by a delta function located at (x_i, y_i):

$$\delta_i(x,y) = \delta(x - x_i, y - y_i) \ \ .$$

Following this representation, a "signal" is a sum of delta functions representing the entire set of points:

$$s(x,y) = \sum_i \delta_i(x,y)$$

The rest of the treatment is similar to the signal case in all aspects. For example, a measured feature is calculated as follows:

$$\langle \Phi(x,y), s(x,y) \rangle = \int\int \Phi(x,y) s(x,y) dx dy = \int\int \Phi(x,y) \sum_i \delta_i(x,y) dx dy = \sum_i \Phi(x_i, y_i) \ .$$

Using the point based "signal", it is possible to find invariants and steerable functions for a set of points undergoing some transformation. With this approach, the motion parameters between two sets of points can be estimated without the necessity of finding correspondence between the sets (i.e. mutual matching between points in the first image to points in the second image). Clearly, the correspondence problem is difficult to solve (having exponential complexity), and it is an important advantage if it is possible to avoid it. The idea of treating point and line coordinates similar to grey scale values has already been suggested by Manmatha [17].

11 Conclusions

We have presented a common theoretical framework for steerable filter design, motion estimation and invariant feature detection based on the theory of Lie groups. Within the framework, the notion of steerability is extended to arbitrary transformation groups. Furthermore, a canonical decomposition of all finite-dimensional steerable bases for any one-parameter and any multi-parameter Abelian transformation group was proposed. The completeness of the canonical decomposition implies that the steerability of any filter with respect to such groups depends on whether it can be described in terms of the canonical bases. Filters steerable under various subgroups (not necessarily Abelian) of the affine group were also provided. Two methods for approximating the steerability of functions over a restricted range of transformations were suggested to deal with filters that cannot be steered exactly. In all cases, the interpolation functions are analytic and can easily be derived.

Motion estimation was discussed as dual to the steerability problem. In the presented framework the so called "window problem" does not arise, therefore, the neighborhood within which the motion parameters are calculated can be arbitrarily small. Guidelines for ensuring the robustness and the uniqueness of the estimated solution were detailed.

Using the framework, image invariants are computed in two stages: (1) a finite-dimensional equivariant measuring space is constructed, (2) invariants over the corresponding equivariant feature space are derived. Since invariants are computed over the finite-dimensional feature space, point-based techniques for computing invariants can be employed.

Finally, a common framework for steerable filter design, motion estimation and invariant feature detection facilitates the transfer of results between the different problems more readily. Indeed, the framework presented in this paper draws from the results of several different areas. The treatment of motion estimation and invariant feature detection within a common framework may facilitate a novel integrations of the two.

12 Acknowledgements

The authors would like to acknowledge David Heeger, Guillermo Sapiro, Carlo Tomasi, Brent Beutter, Misha Pavel, Shimon Edelman, Al Ahumada, and Peter Olver for their helpful comments, suggestions and discussions during the course of this work.

References

[1] S. Amari. Invariant structures of signal and feature spaces in pattern recognition problems. *RAAG Memoirs*, 4:19–32, 1968.

[2] S. Amari. Feature spaces which admit and detect invariant signal transformations. In *Proc. 4th Int. Joint Conf. Pattern Recognition*, pages 452–456, 1978.

[3] A. Cohen. *An introduction to the Lie theory of one-parameter groups; with applications to the solution of differential equations*. D. C. Heath & Co., Boston; New York, 1911.

[4] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer-Verlag, Berlin; New York, 1992.

[5] M. Ferraro and T. Caelli. Lie transformation groups, integral transforms, and invariant pattern recognition. *Spatial Vision*, 8(1):33–44, 1994.

[6] D. Fleet. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, 1990.

[7] D. Fleet. Disparity from local weighted phase-correlation. In *Proc. IEEE Int'l Conf. on Systems, Man, and Cybernetics: Humans, Information and Technology*, pages 48–54, New York, NY, 1994.

[8] D. Fleet and A. Jepson. Phase-based disparity measurement. *CVGIP: Image Understanding*, 53(2):198–210, 1991.

[9] W. Freeman and E. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.

[10] G. Granlund and H. Knutsson. *Signal processing for computer vision.* Kluwer Academic Publishers, Boston, 1995.

[11] Y. Hel-Or and P. Teo. Canonical decomposition of steerable functions. *submitted to the Journal of Mathematical Imaging and Vision, also in the CVPR 96, pp. 809-816.*

[12] M. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory*, February:179–187, 1962.

[13] D. Keren. Using symbolic computation to find algebraic invariants. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16(11):1143–1149, 1994.

[14] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

[15] R. Lenz. *Group theoretical methods in image processing.* Springer-Verlag, Berlin; New York, 1990.

[16] H. Liu, T. Hong, M. Herman, and R. Chellappa. A reliable optical flow algorithm using 3-d hermite polynomials. Technical Report CS-TR-3291, University of Maryland, 1994.

[17] R. Manmatha. A framework for recovering affine transforms using points, lines or image brightnesses. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 141–146, Seattle, WA, 1994.

[18] R. Manmatha and J. Oliensis. Extracting affine deformations from image patches i: finding scale and rotation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 754–755, New York, NY, 1993.

[19] T. Moons, E. Pauwels, L. Van Gool, and A. Oosterlinck. Foundations of semi-differential invariants. *International Journal of Computer Vision*, 14:25–47, 1995.

[20] J. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision.* The MIT Press, Cambridge, MA, 1992.

[21] P. Olver. *Equivalence, Invariants, and Symmetry.* Cambridge University Press, Cambridge, 1995.

[22] P. Perona. Deformable kernels for early vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(5):488–499, 1995.

[23] J. Segman, J. Rubinstein, and Y. Zeevi. The canonical coordinate method for pattern deformation: theoretical and computational considerations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(12):1171–1183, 1992.

[24] E. Simoncelli and H. Farid. Steerable wedge filters. In *Proc. Int. Conf. on Computer Vision*, pages 189–194, Boston, MA, 1995.

[25] E. Simoncelli, W. Freeman, E. Adelson, and D. Heeger. Shiftable multiscale transforms. *IEEE Trans. Information Theory*, 38(2):587–607, 1992.

[26] J. Talman. *Special functions; a group theoretic approach.* W. A. Benjamin, Inc., New York, NY, 1968.

[27] P. Werkhoven and J. J. Koenderink. Extraction of motion parallax structure in the visual system i. *Biological Cybernetics*, 63:185–191, 1990.

[28] P. Werkhoven and J. J. Koenderink. Extraction of motion parallax structure in the visual system ii. *Biological Cybernetics*, 63:193–199, 1990.

[29] M. Werman and A. Shashua. The study of 3d-from-2d using elimination. In *Proc. Int. Conf. Computer Vision*, pages 473–479, Boston, MA, 1995.

[30] Y. Xiong and S. Shafer. Moment filters for high precision computation of focus and stereo. In *Proc. Int. Conf. on Intelligent Robots and Systems*, pages 108–113, 1995.

A Conjugate Generators

The conjugate generators of a group G is the set of generators $\{\bar{L}_i\}$ of the conjugate group $\bar{G}$ such that the following identity holds:

$$\langle \phi, g(\boldsymbol{\tau})\ s \rangle = \langle \bar{g}(\boldsymbol{\tau})\ \phi, s \rangle$$

The operator g is a member of the group G while $\bar{g}$ is a member of the conjugate group $\bar{G}$. The right hand side of the above equation can be derived from the left by a change of variables which involves inverting the operator $g(\boldsymbol{\tau})$. However, since G is a Lie group, the following theorem shows that the conjugate generators can be obtained directly from the generators of G.

Theorem 4 (Conjugate Generators) *: Let $\{L_i\}$ be the generators of the transformation group G. The differential operators $\{\bar{L}_i\}$ satisfying*

$$\langle \phi, L_i\ s\rangle = \left\langle \bar{L}_i\ \phi, s\right\rangle$$

are the conjugate generators of the group, i.e. are the generators of the conjugate group $\bar{G}$.

Proof 4: Since G is a group, we can rewrite the action of $g(\boldsymbol{\tau})$ on $s(x,y)$ via the exponential map:

$$\begin{aligned}
\langle \phi, g(\tau_1,\ldots,\tau_k)\ s\rangle &= \left\langle \phi, e^{\tau_1 L_1}\cdots e^{\tau_k L_k}\ s\right\rangle \\
&= \left\langle \phi, (1+\tau_1 L_1+\cdots)\ e^{\tau_2 L_2}\cdots e^{\tau_k L_k}\ s\right\rangle \\
&= \left\langle (1+\tau_1 \bar{L}_1+\cdots)\ \phi, e^{\tau_2 L_2}\cdots e^{\tau_k L_k}\ s\right\rangle \\
&= \left\langle e^{\tau_1 \bar{L}_1}\ \phi, e^{\tau_2 L_2}\cdots e^{\tau_k L_k}\ s\right\rangle \\
&\ \ \vdots \\
&= \left\langle e^{\tau_k \bar{L}_k}\cdots e^{\tau_1 \bar{L}_1}\ \phi, s\right\rangle
\end{aligned}$$

where the substitution $\langle \phi, (L_i)^m\ s\rangle = \left\langle (\bar{L}_i)^m\ \phi, s\right\rangle$ is used repeatedly. The differential operators $\{\bar{L}_i\}$ are the generators of the group $e^{\tau_k \bar{L}_k}\cdots e^{\tau_1 \bar{L}_1}$ which is, by definition, the conjugate group $\bar{G}$. Hence, $\{\bar{L}_i\}$ are also the conjugate generators of G. □

The conjugate generators $\bar{L}_i$ can be derived from their corresponding generators L_i using the two identities:

1. $\langle \phi, c\ s\rangle = \langle c\ \phi, s\rangle$ for any function c in the integration variables.
2. $\left\langle \phi, \frac{\partial}{\partial x}\ s\right\rangle = -\left\langle \frac{\partial}{\partial x}\ \phi, s\right\rangle$ and, similarly, $\left\langle \phi, \frac{\partial}{\partial y}\ s\right\rangle = -\left\langle \frac{\partial}{\partial y}\ \phi, s\right\rangle$.

The first identity is obvious from the definition of the inner-product. The second identity can be proven using integration by parts:

$$\begin{aligned}
\left\langle \phi, \frac{\partial}{\partial x}\ s\right\rangle &= -\int\int \phi(x,y)(\frac{\partial}{\partial x}\ s(x,y))\ dxdy \\
&= \int \phi(x,y)s(x,y)|_{-\infty}^{\infty}\ dy - \int\int(\frac{\partial}{\partial x}\ \phi(x,y))s(x,y)\ dxdy \\
&= -\int\int(\frac{\partial}{\partial x}\ \phi(x,y))s(x,y)\ dxdy \\
&= -\left\langle \frac{\partial}{\partial x}\ \phi, s\right\rangle
\end{aligned}$$

where the third equality holds because the signal $s(x,y)$ is bounded.

The conjugate generators for the transformations listed in Table 1 follow immediately:

1. $\langle \phi, L_{t_x} s\rangle = \left\langle \phi, -\frac{\partial}{\partial x} s\right\rangle = \left\langle \frac{\partial}{\partial x}\phi, s\right\rangle$ and hence $\bar{L}_{t_x} = \frac{\partial}{\partial x}$.
   By the same derivation it follows that $\bar{L}_{t_y} = \frac{\partial}{\partial y}$.

2. $\langle \phi, L_r s \rangle = -\left\langle \phi, y\frac{\partial}{\partial x}s \right\rangle + \left\langle \phi, x\frac{\partial}{\partial y}s \right\rangle = -\left\langle y\phi, \frac{\partial}{\partial x}s \right\rangle + \left\langle x\phi, \frac{\partial}{\partial y}s \right\rangle = \left\langle \frac{\partial}{\partial x}(y\phi), s \right\rangle - \left\langle \frac{\partial}{\partial y}(x\phi), s \right\rangle$ and hence $\bar{L}_r = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$

3. $\langle \phi, L_{s_x} s \rangle = -\langle \phi, s \rangle - \left\langle \phi, x\frac{\partial}{\partial x}s \right\rangle = -\langle \phi, s \rangle - \left\langle x\phi, \frac{\partial}{\partial x}s \right\rangle = -\langle \phi, s \rangle + \left\langle \frac{\partial}{\partial x}(x\phi), s \right\rangle = -\langle \phi, s \rangle + \left\langle (\phi + x\frac{\partial}{\partial x}\phi), s \right\rangle = \left\langle x\frac{\partial}{\partial x}\phi, s \right\rangle$ and hence $\bar{L}_{s_x} = x\frac{\partial}{\partial x}$.
   By the same derivation it follows that $\bar{L}_{s_y} = y\partial_y$.