# A COMMON FRAMEWORK FOR STEERABILITY, MOTION ESTIMATION, AND INVARIANT FEATURE DETECTION

*Yacov Hel-Or*

Hewlett-Packard Labs Israel
Technion City, Haifa, 32000, Israel

*Patrick C. Teo*

Department of Computer Science
Stanford University, Stanford, CA 94305

## 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.

## 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 origi-

nal image. Note that this represetation is, of course, not complete. The original image cannot be reconstruted from its first derivatives.

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 [9, 12]. 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]. 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 [6]. Recently, Michaelis and Sommer [7] also suggested a method for deriving steerable filters using a Lie group theoretic approach.

| Group | Operator | Generator | Equivariant Function Space |
|-------|----------|-----------|----------------------------|
| $x$-translation | $g_{t_x}(\tau)\, f = f(x+\tau, y)$ | $L_{t_x} = \frac{\partial}{\partial x}$ | $\{\psi_p(y)x^p e^{\alpha x}\}$ for $0 \leq p \leq m$. |
| $x$-scaling | $g_{s_x}(\tau)\, f = f(e^\tau x, y)$ | $L_{s_x} = x\frac{\partial}{\partial x}$ | $\{\psi_p(y)x^\alpha (\ln x)^p\}$ for $0 \leq p \leq m$. |
| Rotation | $g_r(\tau)\, f = f(x\cos\tau + y\sin\tau,\ -x\sin\tau + y\cos\tau)$ | $L_r = -x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x} = \frac{\partial}{\partial \theta}$ | $\{\psi_p(r)e^{ik\theta}\}$ |
| Uniform scaling | $g_s(\tau)\, f = f(e^\tau x, e^\tau y)$ | $L_s = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} = r\frac{\partial}{\partial r}$ | $\{\psi_p(\theta)r^\alpha (\ln r)^p\}$ for $0 \leq p \leq m$. |

Table 1: Several examples of one parameter groups, their generators, and associated equivariant function spaces. $\psi_p$ are arbitrary functions, $\alpha \in \mathbf{C}$ and $k \in \mathbf{Z}$. In the rotation and uniform scaling examples, $(r, \theta)$ are the polar coordinates of the image.

## 2. BACKGROUND ON LIE GROUPS

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

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

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

A family of transformations $\{g(\tau)\}$ parameterized by $\tau_1, \ldots, \tau_k$ 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.

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. The infinitesimal actions of a $k$-parameter group are a set of differential operators $\{L_i \mid i = 1 \ldots k\}$, called the *generators* of the group, corresponding to derivatives of the transformation at the identity with respect to each parameter $\tau_i$ in turn; i.e.,

$$\left.\frac{d\hat{f}}{d\tau_i}\right|_{\tau=0} = L_i\, \hat{f} \quad \text{where} \quad L_i = \left.\left(\frac{\partial x}{\partial \tau_i}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \tau_i}\frac{\partial}{\partial y}\right)\right|_{\tau=0}$$

The $k$ generators provide a basis for the $k$-dimensional tangent space $\mathcal{G} = \{\tau_1 L_1 + \cdots + \tau_k L_k | \tau \in \mathbf{R}^k\}$. There is a correspondence between a $k$-parameter Lie group and its $k$-dimensional tangent space in the form of the exponential map:[1]

$$g(\tau)\, f(x,y) = e^{\tau_1 L_1} \cdots e^{\tau_k L_k}\, f(x,y). \tag{1}$$

The notation $e^{\tau_i L_i}$ represents the series expansion $e^{\tau L_i} = I + \tau_i L_i + \frac{1}{2!}\tau_i^2 L_i^2 + \cdots$, which is an infinite sum of differential operators [2]. The exponential map generates a group

similar to the original group up to a change of parameterization. Examples of common one-parameter groups and their generators are given in Table 1.

## 3. EQUIVARIANT BASIS FUNCTIONS

In this section, we identify the functions that are steerable under different transformation groups. Before describing these functions, we formalize the notion of steerability with a definition.

**Definition 1 (Steerability)** : *A function $f(x,y) : \mathbf{R}^2 \mapsto \mathbf{R}$ is* steerable *under a $k$-parameter Lie transformation group $G$ if any transformation $g(\tau) \in G$ of $f$ can be written as a linear combination of a fixed, finite set of* basis *functions $\{\phi_i(x,y)\}$:*

$$g(\tau)\, f(x,y) = \sum_{i=1}^n \alpha_i(\tau)\, \phi_i(x,y) = \alpha^T(\tau)\Phi(x,y)$$

The functions $\alpha_i$ are known as the *steering functions* of $f$ associated with the basis $\{\phi_i\}$ and depend solely on the transform parameters. Without loss of generality, we assume that $n$ is the minimum number of basis functions required and these basis functions are linearly independent. Clearly, the set of basis functions required to steer a given function is not unique; any (non-singular) linear transformation of the set of basis functions could also be used.

If a function $f$ is steerable with a set of basis functions $\Phi$, then each of the basis functions $\phi_i$ are themselves steerable with the same basis functions. This is true since each basis function can be rewritten as a linear combination of transformed replicas of $f$ (chosen to be linearly independent). Thus, transforming a basis function is equivalent to linearly combining the set of transformed replicas of $f$, which are themselves steerable.

Since steerability of the given function $f$ implies steerability of its basis functions $\phi_i$ as well, it is more natural to express steerability in terms of a function space, i.e. in terms of the space spanned by the basis functions $\{\phi_i\}$.

**Definition 2 (Equivariant Function Space)** :
*An $n$-dim. function space $\mathcal{F} = span\{\phi_1, \ldots, \phi_n\}$ is* equivariant *under a $k$-parameter Lie transformation group $G$ if every $\phi_i$ is steerable with respect to the basis $\{\phi_1, \ldots, \phi_n\}$, i.e., there is a matrix function $A(\tau)$, called the* **interpolation matrix** , *such that:*

$$g(\tau)\Phi(x,y) = A(\tau)\Phi(x,y) \quad \text{for all} \quad g(\tau) \in G$$

*This equation is called the* interpolation equation.

---

[1] This is only true for group elements whose Taylor expansions converge, and for elements within the connected component containing the identity. In this paper, we consider only such transformation groups for which convergence holds.

From the definition it follows that an equivariant function space is a function space that is closed under the associated transformation group. More generally, any function $f \in \mathcal{F}$, such that $f = \sum c_i \phi_i = \mathbf{c}^T \Phi$ is steerable by steering the basis of $\mathcal{F}$:

$$g(\tau)f = \mathbf{c}^T A(\tau)\Phi.$$

As a result, any function $f$ is steerable under a $k$-parameter transformation group if and only if it belongs to some function space that is also equivariant under the same transformation group.

For example, consider the function space $\mathcal{F}_\theta = \mathrm{span}\{\cos\theta, \sin\theta\}$ under the one-parameter group of rotations: $g_r(\tau)f(\theta) = f(\theta - \tau)$. It is easy to verify the following two identities:

$$\begin{aligned}
\cos(\theta - \tau) &= \cos\tau\cos\theta + \sin\tau\sin\theta, \\
\sin(\theta - \tau) &= -\sin\tau\cos\theta + \cos\tau\sin\theta.
\end{aligned}$$

Thus, rotated versions of any basis function in $\mathcal{F}_\theta$ can always be expressed as linear combinations of the basis functions. Hence, any $f \in \mathcal{F}_\theta$ is steerable under the rotation group.

## 4. CONSTRUCTION OF EQUIVARIANT SPACES

In the previous section, equivariant function spaces were defined to be closed under the associated transformation group. Because we are dealing with Lie transformation groups, the closure of a function space under $g(\tau)$ can be reformulated, more simply, in terms of the group generators $\{L_1, \ldots, L_k\}$. This approach is based on the seminal work of S. Amari [1] who originally proposed it in the context of invariant feature detection in pattern recognition.

**Theorem 1 (Interpolation Equation)** :
*The function space $\mathcal{F} = \mathrm{span}\{\phi_1, \cdots, \phi_n\}$ is equivariant under the transformation group $G$ if and only if $\mathcal{F}$ is closed under the action of each generator $L_i$ of $G$. That is, $g(\tau)\Phi = A(\tau)\Phi$ if and only if there is a set of $n \times n$ matrices $\{B_1, \ldots, B_k\}$ such that:*

$$L_i\,\Phi = B_i\,\Phi \quad for\ all \quad i = 1, \ldots, k$$

*In particular, the interpolation matrix can be written as follows:*

$$A(\tau) = e^{\tau_k B_k} \cdots e^{\tau_1 B_1} \quad such\ that \quad g(\tau)\Phi = A(\tau)\Phi$$

**Proof 1** : The proof of this theorem can be found in [5]. □

Theorem 1 provides a recipe for verifying whether a space spanned by a set of functions $\{\phi_i\}$ is equivariant, and if it is, derives the interpolation matrix $A(\tau)$. Several examples of the equivariant function spaces for one parameter transformation groups are listed in Table 1. Examples of equivariant function spaces for multi-parameter transformations can be found in [5].

## 5. COMMON FRAMEWORK

The importance of steerable functions stems from the property of superposition of linear systems. Hence, any linear operation $\mathcal{L}\{*\}$ applied to a transformed version of an equivariant basis set can be expressed as a linear combination of the operation applied separately to the original basis functions. Thus:

$$\mathcal{L}\{g(\tau)\Phi\} = \mathcal{L}\{A(\tau)\Phi\} = A(\tau)\mathcal{L}\{\Phi\}. \quad (2)$$

This property can be exploited in different problems in computer vision, namely: steerable filter design, motion estimation, and invariant feature detection.

### 5.1. Steerable Filters

Steerability of filter kernels 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 $\langle g(\tau)f, s \rangle = \sum_i a_i(\tau) \langle b_i, s \rangle$ where $f$ is the kernel of the filter and $\{b_i\}$ are some basis filters.

In Freeman and Adelson [3], 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 [4, 3, 13, 11].

Because convolution is a linear operation, it is easy to see that each kernel $f$ belonging to an equivariant function space s.t. $f = \mathbf{c}^T \Phi$ is steerable via the interpolation equation: $\langle g(\tau)f, s \rangle = \mathbf{c}^T A(\tau) \langle \Phi, s \rangle$. Thus, steerability can be viewed as a forward problem within the framework as defined in Equation 2. From a set of measured outputs $\mathcal{L}\{\Phi\}$ (the outputs of the basis filters), we compute $\mathcal{L}\{g(\tau)\Phi\}$ for each transformation $g(\tau)$ using $\mathcal{L}\{g(\tau)\Phi\} = A(\tau)\mathcal{L}\{\Phi\}$.

### 5.2. Motion Estimation

Motion estimation, on the other hand, is an inverse problem within the framework. Since a transformation of the image can be carried out by inversely transforming the motion estimation filters, this problem can be considered in terms of the filters. Typically, the same set of motion estimation filters are applied to the original and the transformed images. From the outputs of these filter, we would like to determine the parameters $\tau$ of the transformation $g(\tau)$. Within the framework of Equation 2, from two sets of filter outputs corresponding to $\mathcal{L}\{\Phi\}$ and $\mathcal{L}\{g(\tau)\Phi\}$, our goal is to compute $g(\tau)$. Thus, it is clear that equivariant basis functions are ideal motion estimation filters.

Assume that an equivariant basis set $\phi_i$, $i = 1, \cdots, n$ is chosen to find the $k$-parameter transformation between two images. The first step in the motion estimation process is to apply these functions to the two images obtaining two sets of measured features:

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

Since the functions are equivariant and an inner-product is a linear operation, the measured outputs are related by the interpolation equation:

$$\hat{\mathbf{o}} = e^{B_1 \tau_1} e^{B_2 \tau_2} \cdots e^{B_k \tau_k} \mathbf{o} = A(\tau)\mathbf{o}. \quad (3)$$

The motion parameters of the associated images is obtained by solving for $\tau$ in this system. In general, this system of equations is non-linear; thus, gradient descent minimization is required. Assume that we have an initial guess $\tau^0$ of the motion parameters. Expanding $A(\tau)$ about the initial guess and substituting the linear terms of this expansion into Equation 3 yields a linear system:

$$\hat{\mathbf{o}} = A^0 \mathbf{o} + [B_1 A^0 \mathbf{o}, B_2 A^0 \mathbf{o}, \cdots, B_k A^0 \mathbf{o}](\tau - \tau^0)^T$$

where $A^0 = A(\tau^0)$. Defining:

$$\mathbf{a} \doteq (\hat{\mathbf{o}} - A^0 \mathbf{o}) \quad ; \quad \mathbf{b} \doteq [B_1 A^0 \mathbf{o}, B_2 A^0 \mathbf{o}, \cdots, B_k A^0 \mathbf{o}] \, ,$$

the solution for $\tau$ in the least-squares sense gives:

$$\tau = \tau^0 + (\mathbf{b}^T \mathbf{b})^{-1} \mathbf{b}^T \mathbf{a} \ . \tag{4}$$

This process is repeated with the current solution serving as the new guess until convergence. The convergence of this process to the correct solution, however, is not guaranteed as it depends on the accuracy of the initial guess.

### 5.3. 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 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 function space $\mathcal{F}$ to best characterize the pattern, and (2) determine invariants within this finite (and possibly small) dimensional space.

Since the function space is equivariant, the outputs of applying these functions as filters to transformed images satisfy the relation in Equation 3. This equation is a parametric description of a $k$-dimensional manifold in $R^n$ ($n$ is the number of basis functions and $k$ is the number of transformation parameters). Two sets of outputs, computed from a pair of images related by a transformation in the group, lie on the same $k$-dimensional manifold. Hence, functions that are constant over each manifold are invariant under the transformation; i.e., $h(\mathbf{o}) = h(\hat{\mathbf{o}}) = h(A(\tau)\mathbf{o})$ for $\hat{\mathbf{o}}, \mathbf{o}$ 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 $A(\tau)$ 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 a 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 [10]. In our case, this implies that a function $h(\hat{\mathbf{o}}) = h(e^{\tau_k B_k} \cdots e^{\tau_1 B_1} \mathbf{o})$ is invariant under $g(\tau)$ if and only if

$$L_i h(\hat{\mathbf{o}}) = B_i \hat{\mathbf{o}} \cdot \nabla h = 0 \tag{5}$$

where $\nabla h = (\frac{\partial h}{\partial \hat{o}_1}, \ldots, \frac{\partial h}{\partial \hat{o}_n})^T$ and $1 \leq i \leq k$. A good review on how to solve such a system of PDE's can be found in [8].

### 6. 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 and motion estimation was discussed as the dual to the steerability problem. The common framework facilitates the transfer of results between the different problems more readily. The treatment of motion estimation and invariant feature detection within a common framework may facilitate a novel integrations of the two.

### 7. REFERENCES

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

[2] 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.

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

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

[5] Y. Hel-Or and P. Teo. A common framework for steerability, motion estimation and invariant feature detection. Technical Report STAN-CS-TN-96-28, Stanford University, 1996.

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

[7] M. Michaelis and G. Sommer. A Lie group-approach to steerable filters. *Pattern Recognition Letters*, 16(11):1165–1174, November 1995.

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

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

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

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

[12] 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.

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