STEERABLE FILTERS AND INVARIANT RECOGNITION IN SPACETIME

Reiner Lenz

Dept. EE., Linköping University,S-58183 Linköping, Sweden

ABSTRACT

The groups which have received most attention in signal processing research are the affine groups and the Heisenberg-Weyl group related to wavelets and time-frequency methods. In low-level image processing the rotation-groups $\mathrm{SO}(2)$ and $\mathrm{SO}(3)$ were studied in detail. In this paper we argue that the Lorentz group $\mathrm{SO}(1,2)$ provides a natural framework in the study of dynamic processes like the analysis of image sequences. We summarize the connection between the group $\mathrm{SO}(1,2)$ and the groups $\mathrm{SU}(1,1)$ and $\mathrm{SL}(2,\mathbb{R})$ and give an overview over their representations. We show that their representation theory is in parts similar to the corresponding theory for the three-dimensional rotation group. The main differences between the compact groups (like $\mathrm{SO}(2)$ and $\mathrm{SO}(3)$) is however that the Fourier transforms for these groups involves infinite-dimensional representations and that the finite-dimensional represenstations are no longer unitary. In the signal processing context this means that the filter vectors computed by finite-dimensional steerable filter systems no longer transform as unitary vector transformations under the symmetry operations in $\mathrm{SO}(1,2)$.

1. INTRODUCTION

Group theoretical methods are receiving increasing interest in signal processing and computer vision in recent years. The most commonly used groups in signal processing are the affine and the Heisenberg-Weyl groups due to their links to wavelet[3] and time-frequency analysis [1]. In image processing and computer vision the rotation groups in two- and three dimensions and projective groups are studied. Some applications of finite groups were also considered [4, 5].

In this paper we want to point out the usefulness of $\mathrm{SL}(2,\mathbb{R})$ the group of real 2×2 matrices with determinant one and briefly describe some Fourier-type integral transforms related to this group. We will describe them in a way which is typical for a whole class of such results. This will be briefly mentioned at the end of the paper. For a detailed discussion the reader is referred to the literature [2, 8].

The connection between $\mathrm{SL}(2,\mathbb{R})$ and a set of integral transforms (which include the fractional Fourier transform) has been known for a long time [9] and its applications in optics are also well-established [6, 7]. Here we will establish its relation to the processing of image sequences. We will investigate the construction of steerable filter systems and briefly describe a Fourier transform related to this group. These results are the simplest examples of a general methodology [2] which is well-established in harmonic analysis but up to now they seem to have found no applications in signal processing.

2. INVARIANCE AND STEERABLE FILTERS

We assume that all functions are elements of the Hilbert space H with the scalar product $\langle .,. \rangle$. For the filter f and the signal s the result of filtering s with f is $\langle f,s\rangle$. For a finite set of filter functions $f_1,\ldots,f_K$ we collect the filter results $\langle f_1,s\rangle,\ldots,\langle f_K,s\rangle$ in a vector and we denote also this vector by $\langle f,s\rangle$. For a transformation g operating on the elements of the Hilbert space we write s^g for the transformed signal and likewise for the filters. Typical examples are the time-shifted signals $s(x-g)$ and the rotated signal $s(R^{-1}x)$ where R is a rotation matrix and x the position vector. The goal of invariance methods in image processing and pattern recognition is to find non-trivial filter functions f such that $\|\langle f,s^g\rangle\| = \|\langle f,s\rangle\|$ for all g under investigation. Usually these transformations form a group G. A variant is to find filter functions and matrices $T(g)$ such that $\langle f,s^g\rangle = T(g)\langle f,s\rangle$.

Steerable filter systems are in a sense the dual objects: here the problem is to find filter systems $f_1,\ldots,f_K$ such that $\langle f^g,s\rangle = T(g)\langle f,s\rangle$ for a group G : an infinite number of filter functions can thus be computed as linear combinations of a finite number of filter results computed from the raw data. For unitary matrices $T(g)$ we find that the length of the vectors $\langle f^g,s\rangle$ is invariant under the operations g. The length of the vector is thus a G-invariant measurement and the relation between f^g and f is encoded in the transform matrix $T(g)$.

3. THE GEOMETRY OF SPACE-TIME

Consider the space (t,x,y) where t denotes time and (x,y) are coordinates in 2-D space. A sequence of 2-D images is thus a function of (t,x,y). In the analysis of image sequences it is often assumed that the (t, x, y) space carries the usual geometry of 3-D space. In many applications (like tracking) it is however known that the quantity of interest has a maximal propagation speed. Normalizing the time axis accordingly we find that an event at the origin can only influence events in the time-cone

$$\mathfrak{C} = \left\{(t, x, y) : t^2 - x^2 - y^2 \geq 0, t \geq 0\right\}$$

The natural metric in (t, x, y) space is thus $\|(t, x, y)\| = t^2 - x^2 - y^2$ and the natural group connected to it is the Lorentz group $\mathrm{SO}(1, 2)$ of all 3×3 matrices with determinant one which preserve the scalar product:

$$\langle (t, x, y), (t', x', y') \rangle = tt' - xx' - yy'$$

We may also require that the future cone $t \geq 0$ is mapped into itself.

The time-cone which is the union of the level sets

$$\mathfrak{C}_\rho = \left\{(t, x, y) : t^2 - x^2 - y^2 = \rho, t \geq 0\right\}$$

and each $\mathfrak{C}_\rho$ is invariant under $\mathrm{SO}(1, 2)$. We can therefore split the investigation of functions on $\mathfrak{C}$ into a "radial" and an "angular" part. We write thus f as $f(\rho, \alpha, \beta)$ where ρ describes which hyperboloid $\mathfrak{C}_\rho$ is used and (α, β) are coordinates on this $\mathfrak{C}_\rho$. In the following $\rho = 1$ is assumed.

4. DIFFERENT REALIZATIONS

Problems involving $\mathrm{SO}(1, 2)$ are often easier to analyze in terms of other, related groups. We mention here two of them:

$\mathrm{SL}(2, \mathbb{R})$ The real 2×2 matrices with determinant one

$\mathrm{SU}(1, 1)$ The group of complex 2×2 matrices with determinant one which preserve the scalar product: $\langle (x, y), (x', y') \rangle = x\overline{x}' - y\overline{y}'$

The mappings between $\mathrm{SO}(1, 2)$ and $\mathrm{SL}(2, \mathbb{R})$ and $\mathrm{SU}(1, 1)$ are constructed as follows:

For a point (t, x, y) define the hermitian matrix h as

$$h = \begin{pmatrix} t & x + iy \\ x - iy & t \end{pmatrix}$$

and for $g \in \mathrm{SU}(1, 1)$ define:

$$\tilde{h} = \begin{pmatrix} \tilde{t} & \tilde{x} + i\tilde{y} \\ \tilde{x} - i\tilde{y} & \tilde{t} \end{pmatrix} = ghg^*$$

where g^* is the conjugate complex of the transposed matrix g. The resulting matrix $\tilde{h}$ is hermitian and the determinants of h and $\tilde{h}$ are equal. For $(\tilde{t}, \tilde{x}, \tilde{y})$ we get thus another point in (t, x, y) space and the map $(t, x, y) \mapsto (\tilde{t}, \tilde{x}, \tilde{y}) = \Phi((t, x, y))$ is an element in $\mathrm{SO}(1, 2)$. The map $g \mapsto \Phi(g)$ is a map from $\mathrm{SU}(1, 1)$ to $\mathrm{SO}(1, 2)$ which preserves the group operation. The elements g and $-g$ define the same Lorentz transformation and $\mathrm{SU}(1, 1)$ covers $\mathrm{SO}(1, 2)$ twice.

Next define the matrices

$$\tau = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \text{ and } \tau^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$$

Elementary matrix multiplication shows that $\tau g \tau^{-1}$ is an element in $\mathrm{SL}(2, \mathbb{R})$ for all $g \in \mathrm{SU}(1, 1)$.

The group $\mathrm{SO}(1, 2)$ acts on the hyperboloids $\mathfrak{C}_\rho$. For a 2×2 matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and a complex variable z we define Mz as

$$Mz = \frac{Az + B}{Cz + D}$$

If $M \in \mathrm{SU}(1, 1)$ then it can be shown that $Mz : \mathfrak{E} \to \mathfrak{E}$ and $Mz : \mathfrak{L} \to \mathfrak{L}$ where

$$\mathfrak{E} = \left\{z = x + iy : x^2 + y^2 < 1\right\}$$

is the interior of the unit disk and

$$\mathfrak{L} = \left\{z = x + iy : x^2 + y^2 = 1\right\}$$

is the unit circle.

For the matrices in $\mathrm{SL}(2, \mathbb{R})$ it can be shown that they act on the upper half-plane: $\mathfrak{H} = \left\{x + iy : y > 0\right\}$ and on the real line $\mathbb{R}$.

The groups $\mathrm{SO}(1, 2), \mathrm{SU}(1, 1)$ and $\mathrm{SL}(2, \mathbb{R})$ are thus essentially the same object and they operate on the hyperboloid, the unit disc, the unit circle, the upper half-plane and the real line.

On the hyperboloid, unit disc and the upper half plane the coordinates can be introduced as follows:

Take the point $(1, 0, 0)$ on the hyperboloid. It is a fixpoint for the subgroup $\mathrm{SO}(2)$ of 2-D spatial rotations. All points can be reached from $(1, 0, 0)$ by a transformation of the form $ka \in \mathrm{SO}(1, 2)$ where k is another 2-D rotation and a is a hyperbolic rotation:

$$k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \text{ and } \begin{pmatrix} \cosh\tau & 0 & \sinh\tau \\ 0 & 1 & 0 \\ \sinh\tau & 0 & \cosh\tau \end{pmatrix}$$

In group theoretical notation this becomes:

$$\mathfrak{C}_1 = \mathrm{SO}(1, 2)/\,\mathrm{SO}(2) \text{ or } \mathrm{SO}(1, 2) = KAK$$

where $K = \mathrm{SO}(2)$ is the 2-D rotation group and A is the group of hyperbolic rotations. This is known as the "Cartan decomposition" of $\mathrm{SO}(1, 2)$. The "polar coordinates" of the point $ka(1, 0, 0)$ are given by the parameters τ and φ of a and k. Written in usual functional form they are: $(\cosh\tau, \sinh\tau\sin\varphi, \sinh\tau\cos\varphi)$.

For the group $\mathrm{SU}(1,1)$ the origin is $(0,0)$ the origin of the unit disc and K and A consist of the matrices:

$$\begin{pmatrix} \mathrm{e}^{i\frac{\varphi}{2}} & 0 \\ 0 & \mathrm{e}^{-i\frac{\varphi}{2}} \end{pmatrix} \in K \text{ and } \begin{pmatrix} \cosh\frac{t}{2} & i\sinh\frac{t}{2} \\ -i\sinh\frac{t}{2} & \cosh\frac{t}{2} \end{pmatrix} \in A$$

The coordinates are $\tanh(\tau/2), \mathrm{e}^{i\varphi}$.

For $\mathrm{SL}(2,\mathbb{R})$ the origin is the point i on the imaginary axis and the matrices are

$$\begin{pmatrix} \cos\frac{\varphi}{2} & \sin\frac{\varphi}{2} \\ -\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix} \text{ and } \begin{pmatrix} \mathrm{e}^{\frac{t}{2}} & 0 \\ 0 & \mathrm{e}^{-\frac{t}{2}} \end{pmatrix}$$

Another useful decomposition is the "Iwasawa decomposition" $\mathrm{SO}(1,2) = ANK$ with A and K as defined in the Cartan decomposition and N given by

$$\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R}),$$

$$\begin{pmatrix} 1+i\frac{t}{2} & \frac{t}{2} \\ -\frac{t}{2} & 1-i\frac{t}{2} \end{pmatrix} \in \mathrm{SU}(1,1)$$

and

$$\begin{pmatrix} 1+\frac{t^2}{2} & t & -\frac{t^2}{2} \\ t & 1 & -t \\ \frac{t^2}{2} & t & 1-\frac{t^2}{2} \end{pmatrix} \in \mathrm{SO}(1,2).$$

Using the natural parameterizations of the groups N and A we get a second type of coordinate system on the hyperboloid, the unit disc and the upper half-plane: identify a point with $a(t)k(\varphi)0$ (with 0 the origin) and use (t,φ) as coordinate vector. Sometimes it is easier to use $\mathrm{SO}(1,2) = KAN$ instead, also this is known as Iwasawa decomposition.

5. REPRESENTATIONS OF $\mathrm{SL}(2,\mathbb{R})$

We now consider only $\mathrm{SL}(2,\mathbb{R})$ since the corresponding results for $\mathrm{SU}(1,1)$ and $\mathrm{SO}(1,2)$ are easily established via the mappings described above.

A (finite dimensional) representation of a group is a mapping from the group to a set of matrices which preserve multiplication, i.e.:

$$\rho : G \to GL(V);\ \text{such that: } \rho(g_1 g_2) = \rho(g_1)\rho(g_2) \quad (1)$$

where $\rho(g_1)\rho(g_2)$ is the ordinary matrix multiplication and $GL(V)$ denotes the space of all invertible mappings from the vector space V into itself. For infinite-dimensional spaces V the elements in $GL(V)$ are no longer matrices but general linear operators.

A representation is unitary if all the operators $\rho(g)$ are unitary and it is irreducible if the space V contains no non-trivial invariant subspaces.

We describe now the representations of $\mathrm{SL}(2,\mathbb{R})$. They can be constructed by using representations of a subgroup of the original group and lifting them to the representations of the full group. The construction starts with the Iwasawa decomposition KAN. Since A is commutative the representations of A are one-dimensional and are given by

$$\begin{pmatrix} \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{pmatrix} = \mathrm{sgn}(\alpha)^{\varepsilon}|\alpha|^{l}$$

where $\varepsilon = \{0,1\}$ and l is a complex number. It can be shown that these are the only representations of the larger group of triangular matrices and thus:

$$\chi_\sigma\left(\begin{pmatrix} \alpha & \beta \\ 0 & \frac{1}{\alpha} \end{pmatrix}\right) = \mathrm{sgn}(\alpha)^{\varepsilon}|\alpha|^{l} \text{ with } \sigma = (\varepsilon, l)$$

Next define for the triangular matrix $p = \begin{pmatrix} \alpha & \beta \\ 0 & \frac{1}{\alpha} \end{pmatrix}$ the modular function $\Delta(p) = |\alpha|^{-2}$ and the space $\mathcal{E}^{\sigma}$ as the space of functions on $\mathrm{SL}(2,\mathbb{R})$ which satisfy the condition:

$$f(gp) = \Delta^{1/2}(p)\chi_\sigma^{-1}(p)f(g)$$

for all $g \in \mathrm{SL}(2,\mathbb{R})$ and all triangular p.

For a general element $g \in \mathrm{SL}(2,\mathbb{R})$ we have the Iwasawa decomposition $kan = kp$ and therefore

$$f(g) = f(kp) = \Delta^{1/2}(p)\chi_\sigma^{-1}(p)f(k)$$

This shows that for given σ the functions in $\mathcal{E}^{\sigma}$ are defined by their values on K. The representation $\mathcal{U}^{\sigma}$ is now defined by:

$$\mathcal{U}^{\sigma}(g)f(h) = \mathcal{U}^{\sigma}_{g}\, f(h) = f(g^{-1}h)$$

These are the induced representations. For $\varepsilon = 0$ the functions in $\mathcal{E}^{\sigma}$ are even functions on the rotation subgroup K and for $\varepsilon = 1$ they are odd functions.

For the rotation $k(\varphi) \in K$ and a given value of ε define the functions

$$e_n^{\varepsilon}(k(\varphi)) = \mathrm{e}^{i(2n+\varepsilon)\varphi/2}$$

These functions belong to $\mathcal{E}^{\sigma}$ and they define this space completely. For a matrix $a = \begin{pmatrix} \mathrm{e}^{\tau/2} & 0 \\ 0 & \mathrm{e}^{-\tau/2} \end{pmatrix}$ we compute now the effect of the representation on such a basis element:

$$\langle e_m, \mathcal{U}^{\sigma}_{a}\, e_n^{\varepsilon}(k(\theta))\rangle = \frac{1}{4\pi}\int_0^{4\pi} e_m(\theta)\,\mathcal{U}^{\sigma}_{a}\, e_n^{\varepsilon}(k(\theta))\, d\theta =$$

$$(-1)^{m-n}\,\mathfrak{P}^{-\frac{1}{2}(l+1)}_{m+\frac{1}{2}\varepsilon, n+\frac{1}{2}\varepsilon}(\cosh\tau)$$

where $\mathfrak{P}$ is the "Jacobi function". A special case are the "Legendre functions" defined as

$$\mathfrak{P}_{\zeta}(\cosh\tau) = \mathfrak{P}^{\zeta}_{0,0}(\cosh\tau) =$$

$$\frac{1}{2\pi}\int_0^{2\pi}(\cosh\tau+\sinh\tau\cos\theta)^{\zeta}\,d\theta$$

and the "associated Legendre functions":

$$\mathfrak{P}_{\zeta}^{m}(\cosh\tau)=\frac{\Gamma(\zeta+m+1)}{\Gamma(\zeta+1)}\,\mathfrak{P}_{m,0}^{\zeta}(\cosh\tau)$$

In the case where $l+\varepsilon=2k-1$ is an odd integer define the spaces:

$$\mathcal{F}_{+}^{\sigma}=\{f=\sum_{n=-\infty}^{-k}a_n e_n^{\varepsilon}\},\mathcal{F}_{-}^{\sigma}=\{f=\sum_{n=k-\varepsilon}^{\infty}a_n e_n^{\varepsilon}\}$$

The properties of these representations are summarized in the next theorem:

Theorem. *1. $\mathcal{U}^{\sigma}$ is unitary if and only if σ is a pure imaginary number*

2. If $l+\varepsilon$ is not an odd integer then $\mathcal{U}^{\sigma}$ is irreducible.

3. Assume $l+\varepsilon$ is an odd integer: then $\mathcal{F}_{+},\mathcal{F}_{-}$ are invariant and $\mathcal{F}_{+}\cap\mathcal{F}_{-}=\{0\}$ if and only if $l+1>0$. In this case $\mathcal{F}_{+},\mathcal{F}_{-}$ are both irreducible.

4. If $\mathcal{F}_{+}\cap\mathcal{F}_{-}\neq\{0\}$ then $\mathcal{F}_{+}\cap\mathcal{F}_{-}$ is irreducible and $\mathcal{U}^{\sigma}$ induces irreducible representations on $\mathcal{F}_{+}/(\mathcal{F}_{+}\cap\mathcal{F}_{-})$ and $\mathcal{F}_{-}/(\mathcal{F}_{+}\cap\mathcal{F}_{-})$

The last theorem shows especially that there are no finite dimensional unitary representations. For steerable filter systems this has the following consequence: There are no non-trivial finite-dimensional steerable filter systems for which the length of the filter vector is an invariant under transformations from $\mathrm{SO}(1,2)$.

6. FOURIER TRANSFORM

For the three-dimensional rotation group it is well-known that the spherical harmonics form a complete orthonormal system for functions on the sphere. The spherical functions for the Lorentz group are the associated Legendre functions and there is an abstract Fourier transform for the groups $\mathrm{SO}(1,2)$, $\mathrm{SU}(1,1)$ and $\mathrm{SL}(2,\mathbb{R})$.

For the group $\mathrm{SL}(2,\mathbb{R})$ the corresponding integral transforms are given as follows: denote by $k(\theta)$ the rotation matrix $\begin{pmatrix}\cos\theta/2 & \sin\theta/2\\ -\sin\theta/2 & \cos\theta/2\end{pmatrix}$ and by $h(\alpha)$ the hyperbolic rotation $\begin{pmatrix}\cosh\alpha/2 & \sinh\alpha/2\\ \sinh\alpha/2 & \cosh\alpha/2\end{pmatrix}$. A point $z\in\mathfrak{H}$ is then given by $k(\theta)h(\alpha)i$ and (θ,α) are its coordinates. On $\mathfrak{H}$ there is a line element $\frac{(dx^2+dy^2)^{1/2}}{y}$. If f is a function on $\mathfrak{H}$ which only depends on α then f can be analyzed via the "Mehler transform:"

$$\hat{f}(\kappa)=2\pi\int_0^{\infty}f(\cosh\alpha)\,\mathfrak{P}_{-1/2+i\kappa}(\cosh\alpha)\sinh(\alpha)\,d\alpha$$

which can be inverted as:

$$f(\cosh\alpha)=\frac{1}{2\pi}\int_0^{\infty}\hat{f}(\kappa)\,\mathfrak{P}_{-1/2+i\kappa}(\cosh\alpha)\kappa\tanh\pi\kappa\,d\kappa$$

For arbitrary functions on the half-plane $\mathfrak{H}$ the corresponding transformation is as follows:

$$\tilde{f}_{\kappa}(x+iy)=\int_{\mathfrak{H}}\mathfrak{P}_{-1/2+i\kappa}(\cosh\alpha)f(\xi+i\eta)\frac{d\xi d\eta}{\eta^2}$$

where α is the hyperbolic distance between $x+iy$ and $\xi+i\eta$. This can be inverted as:

$$f(x+iy)=\frac{1}{2\pi}\int_0^{\infty}\tilde{f}_{\kappa}(x+iy)\kappa\tanh\pi\kappa\,d\kappa$$

This shows that in a complete description of a function on the upper half-plane all representations with index $\kappa\in\mathbb{R}^+$ are involved.

7. CONCLUSIONS

We reviewed some of the facts from the representation theory of the group $\mathrm{SO}(1,2)$ which is the natural symmetry group for systems in which propagation speed has an upper limit. We described these results in a framework which can be generalized to the general theory of harmonic analysis on symmetric spaces. Other problems which can be studied along the same lines are problems involving the sphere (with symmetry group $\mathrm{SO}(3)$) and three-dimensional space-time, i.e. time varying volumes with symmetry group $\mathrm{SL}(2,\mathbb{C})$.

8. REFERENCES

[1] Leon Cohen. *Time Frequency Analysis.* Signal Processing Series, Prentice Hall, 1995.

[2] H. Dym and H. P. McKean. *Fourier Series and Integrals.* Academic Press, 1972.

[3] M. Holschneider. *Wavelets, an analysis tool.* Clarendon Press, Oxford, 1995.

[4] R. Lenz. Investigation of receptive fields using representations of dihedral groups. *J. Vis. Comm. Im. Rep.*, 6(3), 209–227, September 1995.

[5] P. Meer, R. Lenz, and S. Ramakrishna. Efficient invariant representations. Int. J. Comp. Vis. (accepted).

[6] J. Sanchez Mondragon and K. B. Wolf. *Lie Methods in Optics, Proc. CIFMO-CIO Workshop* Springer, 1986.

[7] Olle Seger and R. Lenz. Modeling of the point spread function of laser scanning microscopes using canonical transforms. *J. Vis. Comm. Im. Rep.*, 3(4), 1992.

[8] N.Ja. Vilenkin and A.U. Klimyk. *Representation of Lie groups and special functions.* Kluwer, 1991.

[9] K. B. Wolf. *Integral transforms in Science and Engineering.* Plenum, 1979.