IMAGE “GROUP-REGISTRATION” BASED ON REPRESENTATION THEORY

Lamia Ben Youssef
GRIFT, Lab. Cristal, ENSI
Campus Universitaire de la Manouba, 2010 Manouba, Tn

Faouzi Ghorbel
GRIFT, Lab. Cristal, ENSI
Campus Universitaire de la Manouba, 2010 Manouba, Tn

Keywords: Representation theory, homogeneous space, generalized correlation, group-registration, similarities, projective transformations.

Abstract: The general principle of a matching algorithm is to optimize a criterion that furnishes a measure of the similarity between two images for a given space of geometrical transformations. In this work, we propose a framework based on a similarity measure – the generalized correlation – built in a systematic way from the links between a features space and a group of transformations modeled by an action group.
Using results from representation theory, we can extend the correlation function to any homogeneous space with a transitively acting group. When the generalized Fourier transform exists, the group-based correlation can be expressed in a spectral space and it becomes possible to implement fast algorithms for its computation. Two important examples in the field of image processing are then detailed: the similarity group (rotation and scaling) on gray-level shapes from 2D images and the 3D rigid motion group (rotation and translation) followed by a plan projection.

1 INTRODUCTION

Cross correlation techniques, under different names and guises, constitute one of the most popular means for handling with problem of motion analysis and detection. Cross correlation is the basis for such methods as matched filtering, modulation and phase transfer functions, maximum likelihood type methods in statistical setting. The popularity of cross correlation methods can be explained by their simplicity, amenability to digital implementation, and robustness in presence of noise. Such techniques are well known to be efficient for registration process in the case of motions described by translation of the plane (Brown, 1992). However, the inability of conventional cross correlation to deal with rotation and dilation has brought about a number of other approaches from invariant pattern recognition.

Image registration is one of the oldest and fundamental processing to check image similarity based on a quantitative measure and to locate a template at the place of its optimum value. There have been several approaches for image registration, such as feature-, brightness- and aggregated measure-based ones among them color or gray-level histograms (Zitovà and Flusser, 2003).

Many non-brightness feature-based approaches have been proposed to robust image matching, among them the generalized Hough transform proposed by (Ballard, 1981) is one of the famous methods. Another exemple is given by the indexing approaches (Y. Lamdan, 1988) that adopt combinations of geometric features, such as edges or line segments, as key for indexing objects and searching them to pick up candidates. But we cannot expect a reliable matching result without some preprocessing technics –which are subject to instability– to detect features of interest. Moment-based approaches (S.X. Liao, 1996) and color or histogram indexing methods (M.J. Swain, 1991) can be used for rotated patterns, but they are computationally expensive and not really robust in case of brightness change and occlusion.

The correlation technics can be used to answer these two difficulties once geometrical distortions are integrated in the registration process. Pinstov (Pintsov, 1989) has suggested the use of a cross correlation function operating with three parameters: a picture, a template and the planar Euclidian group. This correlation, which involves a parametric tem-

Ben Youssef L. and Ghorbel F. (2006).
IMAGE “GROUP-REGISTRATION” BASED ON REPRESENTATION THEORY.
In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 317-322
DOI: 10.5220/0001373903170322

plate (like line, circle ...), appears to be a particular case of the well-known Radon transform. Segman (Segman, 1992; Segman and Zeevi, 1993; Rubinstein et al., 1991) has shown that the Fourier transform written with an appropriate kernel could be seen as a cross correlation function adapted to the general affine transformations.

In the present work, we extend Segman's model to more general group of transformations such as projective, non-uniform dilatation, and also to any templates, e.g. non-parametric shapes (planar or 3D objects). Based on representation theory, we propose a framework (Ben Youssef, 2004) for studying image deformations applicable not only in the plane but also in other domains like the sphere. This framework involves three steps:

- Identify the domain of definition of the signal as a homogeneous space and the group acting on it.
- Check whether an invariant measure exists for the acting group.
- Compute the transformation (group action) from the adapted correlation.

The remaining of the paper is organized as follows. Sec. 2 presents a short introduction to representation theory, in particular the generalized correlation and the convolution theorem is detailed. Sec. 3 deals with four interesting cases for image processing, i.e. the planar and vectorial similarity groups, the motion and the projective groups. Several experiments on real images illustrate the different cases and confirm the numerical feasibility of the methodology.

2 GROUP BASED CORRELATION

This section is not meant as a formal enumeration of the assumptions we make. It is a rather intuitive description of what is required to apply the recipe of a generalized Fourier transform. We assume that the reader is familiar with the concept of group. Let us live with an intuitive definition of a Lie group: its elements are on smooth manifold and that the group operation and the inversion are smooth maps (Miller and Younes, 2001). The real line and the circle are Lie groups with respect to addition and well known matrix Lie groups are the general linear group of square invertible matrices, the rotation groups $SO(n)$, and the Lorentz group.

A group $\mathcal{G}$ is acting on a space X when there is a map $\mathcal{G} \times \mathrm{X} \longrightarrow \mathrm{X}$ such that the identity element of the group leaves X as is and a composition of two actions has the same effect as the action of the composition of two group operations (associativity). For example, the isometry group $SE(2)$ acts on the plane $\mathcal{R}^2$. The rotation group $SO(3)$ can act on the sphere S^2. The set of all $gx \in \mathrm{X}$ for any $g \in \mathcal{G}$ is called the orbit of x. If the group possesses an orbit, that means for any $a, b \in \mathrm{X}, ga = b$ for a $g \in \mathcal{G}$, then the group action is called transitive. For example, there is always a rotation mapping one point on the sphere to another. If a subgroup H of $\mathcal{G}$ fixes a point $x \in \mathrm{X}$ then H is called the isotropy group. A typical example of an isotropy group is the subgroup $SO(2)$ of $SO(3)$ acting on the north-pole of a sphere.

A space X with a transitive Lie group action $\mathcal{G}$ is called homogeneous space. If the isotropy group is H, it is denoted with $\mathcal{G} = H$. The plane $\mathcal{R}^2$ is the homogeneous space $SE(2) = SO(2)$. The sphere S^2 is the homogeneous space $SO(3) = SO(2)$. Images are usually defined on homogeneous spaces and their deformations are the group actions. The question is now, for which groups we can explicit a correlation function and does a Fourier transform exist? The answer requires, first, to be able to integrate on the group and on the homogeneous space and, second, to find the Fourier basis analogous to $e^{ix\omega}$ on the real line (Miller and Younes, 2001).

We consider a simple two dimensional translation registration problem. The classical correlation :

$$\mathcal{C}(u,v) = \iint_{\mathcal{R}^2} f(x,y)\, h(x-u, y-v)\, dxdy, \quad (1)$$

is extended in a natural way to include rotations and dilations of the template object. Essentially, we rotate, and dilate the template object, overlap it with the image and compute an overlap area (weighted by the intensity value at each pixel) with the proper normalization.

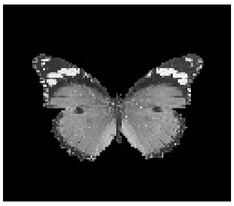
$$\mathcal{C}_{ex}(\alpha,\beta) = \iint_{\mathcal{R}^2} f\begin{pmatrix} x \\ y \end{pmatrix} h\begin{pmatrix} \alpha x \cos\beta - \alpha y \sin\beta \\ \alpha x \sin\beta + \alpha y \cos\beta \end{pmatrix}\, dx\, dy.$$

The function $\mathcal{C}_{ex}$ has a maximum that will determine the scale and the orientation of a target. One limitation of this method is its incapacity to detect a reference target at different scale (Fig 1). Therefore, we need to determine a correlation type depending on transformation parameters.

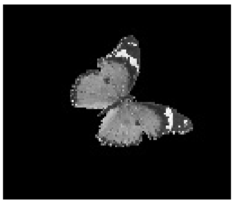
2.1 Generalized Correlation Function

Let $L(\mathrm{X}, d\mu)$ denotes the set of functions which have a mean defined as $\int_{\mathrm{X}} |f(x)|^2\, d\mu(x)$ where μ is an invariant measure. The correlation of functions f_1 and f_2 for all $g \in \mathcal{G}$ with respect to group $\mathcal{G}$ is given by

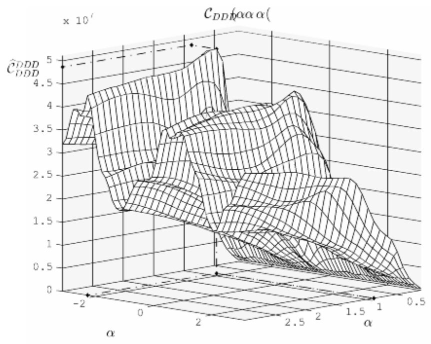
$$(f_1 \oplus f_2)(g) = \int_{\mathrm{X}} f_1(g^{-1}s) f_2(s) \chi(g) d\mu(s) \quad (2)$$



(a) Butterfly im.



(b) Transformed im.



(c) Correlation surface

Figure 1: Classical correlation surface.

If μ is also invariant under $\mathcal{G}$, then (2) becomes

$$(f_1 \oplus f_2)(g) = \int_{\mathrm{X}} f_1(g^{-1}s)\, f_2(s) d\mu(s). \tag{3}$$

In this case, the correlation function has the properties:

1. The invariance of the Haar measure on X under the transformation group $\mathcal{G}$ allow this following operation:
$$\int_{\mathrm{X}} f_1(g^{-1}s)\, f_2(s)\, d\mu(s) = \int_{\mathrm{X}} f_1(s)\, f_2(gs)\, d\mu(s) \tag{4}$$
2. If f_1 is the translated version of f_2 by g_0 with respect to $\mathcal{G}$, then the correlation in Eq. (3) becomes:
$$(f_1 \oplus f_2)(g) = \Delta\left(g_0^{-1}\right)(f_2 \oplus f_2)(g)(g_0 g), \tag{5}$$
where $\Delta(g_0^{-1})$ is the modulus function of group $\mathcal{G}$. For Abelian group $\Delta(g_0^{-1}) = 1$, then $C_{f_1\, f_2}(g) = C_{f_2\, f_2}(g_0 g)$ and $C_{f_1\, f_1}(g) = C_{f_2\, f_2}(g)$.

2.2 Correlation Theorem

The correlation theorem is useful, in part because it gives a way to simplify many calculations. Correlation can be very difficult to calculate directly, but is often much easier to calculate using Fourier transforms and multiplication. The correlation theorem can be stated in words as follows: the Fourier transform of a correlation integral is equal to the product of the complex conjugate of the Fourier transform of the first function and the Fourier transform of the second function.

Let $\mathcal{G}$ be a locally compact Abelian group with an invariant measure $d\mu$. The Fourier transform maps correlation into multiplication:

$$\mathcal{F}\,(f_1 \oplus f_2) = \mathcal{F}(f_1).\, \mathcal{F}^*(f_2). \tag{6}$$

Then generalized correlation in the Fourier domain is given by

$$(f_1 \oplus f_2) = \mathcal{F}^{-1}\Big(\mathcal{F}(f_1).\, \mathcal{F}^*(f_2)\Big). \tag{7}$$

When generalized Fourier transform is

$$\mathcal{F}_f(\lambda) = \hat{f}(\lambda) = \int_{\mathcal{G}} f(x)\,[T_\lambda(x)]\; d\mu(x), \tag{8}$$

its inverse when exists is explicit:

$$f(x) = \int_{\hat{\mathcal{G}}} \hat{f}(\lambda)\,[T_\lambda(x)]^{-1}\; d\mu(\lambda). \tag{9}$$

3 CORRELATION ON PARTICULAR GROUPS

The image plane X is identified with the Euclidean space $\mathbb{R}^2$.

3.1 Planar Radial Scaling and Angular Rotation

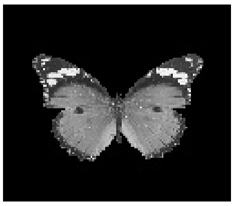
We first consider the group of radial dilation and angular rotation $\mathcal{N}$, group of elements (r, θ) where $r \in \mathbb{R}_+^*$ and $\theta \in \mathbb{S}^1$. It is well known that the invariance measure under the action of $\mathcal{N}$ is $\frac{dr}{r}\, d\theta$. X in polar co-ordinate system is identified to $\mathcal{G}$. According to Eq. (3), correlation of f and $h \in L^2\left(\mathbb{R}_+^* \times \mathbb{S}_1\right)$ for all $(\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{S}_1$ is:

$$C_{fh}(\alpha, \beta) = \int_0^{+\infty}\int_0^{2\pi} h\left(\frac{r}{\alpha}, \theta - \beta\right)\, f(r, \theta)\, \frac{dr}{r} d\theta. \tag{10}$$

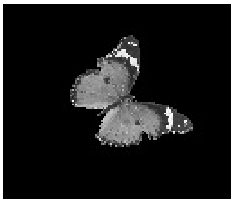
Fig. 2 illustrates the correlation surface between two images of the same object (a butterfly), but with different size and orientation ($\rho = 1.33$ and $\theta = 60^o$). The similarity between the two images is measured by detecting the maximum of the output correlation given by Eq. (10). The position of this maximum gives the parameters of the image transformation.

To calculate this correlation function, we can use the correlation theorem for the Fourier-Mellin transform (Ghorbel, 1994; Derrode and Ghorbel, 2001) which is given by

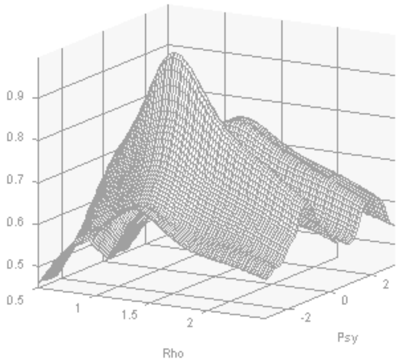
$$\mathcal{M}_{f_\sigma}(k, v) = \frac{1}{2\pi}\int_0^{\infty}\int_0^{2\pi} f(r, \theta)\, r^{iv}\, e^{-ik\theta}\, d\theta \frac{dr}{r}. \tag{11}$$



(a) Butterfly im.



(b) Transformed im.



(c) Correlation surface

Figure 2: Example of a scale and rotation correlation surface.

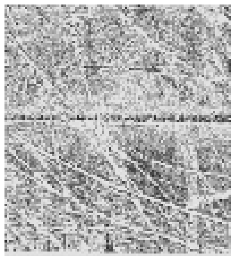
3.2 Motion Group

Let $\mathrm{M}(2)$ be the group of rigid motions of the plane over $\mathbb{R}^2$, i.e. the group of elements $g = (\mathbf{A}, \mathbf{b})$ where $\mathbf{A} \in \mathrm{SO}(2)$ and $\mathbf{b} \in \mathbb{R}^2$. Let $\mathrm{H} = \mathrm{SO}(2)$ be the transitive subgroup of $\mathrm{M}(2)$ over $\mathbb{S}^1$. By Mackey decomposition theorem, every element of $\mathrm{M}(2)$ may be represented in the form $g = (\mathbf{A}, \mathbf{b}) = (\mathbf{I}, \mathbf{b}) \cdot (\mathbf{A}, 0)$, with $\mathbf{I}$ the unitary matrix. Hence, each element of the left coset $g\mathrm{H}$ may be uniquely represented by a point $\mathbf{x} = \mathbf{b}$ of the plane $\mathbb{R}^2$. We obtain the action of $\mathrm{M}(2)$ on $\mathbb{R}^2$ by

$$\mathbf{A}\mathbf{x}+\mathbf{b} = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

The invariant measure on $\mathbb{R}^2$ (Lebesgue measure) is invariant under the action of rigid motions. With respect to the lift previously defined, $\mathrm{M}(2)$ is a product of $\mathbb{R}^2$ by the compact space $\mathbb{S}^1$ so that square summable functions with respect to the Lebesgue measure on $\mathbb{R}^2$ lift $\mathrm{M}(2)$ into square summable functions with respect to the Haar measure. These two facts allow to define the correlation function on the plane (Gauthier et al., 1991). Therefore, for two functions $f, h \in L^2 \left(\mathbb{R}^2, d\mathbf{x}\right)$ and for $d\mu(\mathbb{R}, \mathbf{x}) = d\theta d\mathbf{x}$, we have

$$C_{fh}(\mathbf{A}, \mathbf{b}) = \int_{\mathbb{R}^2} h(\mathbf{x}) f\left(\mathbf{A}\mathbf{x} + \mathbf{b}\right) d\mathbf{x}. \tag{12}$$



(a) Sea floor image.

(b) Trawl traces.

Figure 3: Detection of trawl traces on sonar seafloor image.

3.3 Similarity Group

We now consider the group of similarity GS, group of elements $(a, \mathbf{A}, \mathbf{b})$ where, $a \in \mathbb{R}_+^*$, $\mathbf{A} \in \mathrm{SO}(2)$ and $\mathbf{b} \in \mathbb{R}^2$. Let $\mathcal{N} = \mathbb{R}_+^* \times \mathbb{S}^1$ be the transitive subgroup of GS (see section 3.1). By Mackey decomposition theorem, each element of GS may be represented in the form $g = (a, \mathbf{A}, \mathbf{b}) = (1, I, \mathbf{b}).(a, \mathbf{A}, 0)$. Hence each element of the left coset $g\mathrm{H}$ may be uniquely represented by a point $\mathbf{x} = \mathbf{b}$ of the plane $\mathbb{R}^2$. The multiplication group is given by

$$g_1 \cdot g_2 = \left(a_1 \cdot a_2, (\mathbf{A}_1 \mathbf{A}_2)^T, a_1 \cdot \mathbf{A}_1^T b_2 + b_1\right).$$

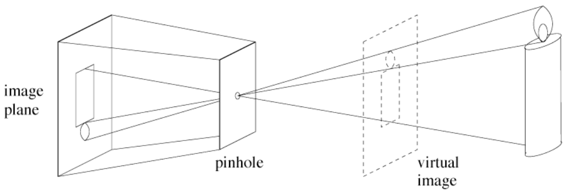
The invariant measure on $\mathbf{A}$ is invariant relatively to translations and, therefore, should be proportional to the Lebesgue measure. Such a measure cannot, however, be invariant to homothetic transformations $x \mapsto ax$ and, consequently, there is no measure on $\mathbf{A}$ invariant under GS. However there exists a relatively invariant measure on the homogeneous space given by the differential of Eq. $d(g\mathbf{x}) = \mathbf{J}_g.d\mathbf{x}$ ($\mathbf{J}$ is the Jacobean of g). According to Eq. (2) correlation of two functions f and h defined on $L^2 \left(\mathbb{R}^2, \mathbf{J}d\mathbf{x}\right)$ is

$$C_{fh}(a, \mathbf{A}, \mathbf{b}) = \int_{\mathbb{R}^2} f\left(\mathbf{x}\right) h\left(a\mathbf{A}\mathbf{x} + \mathbf{b}\right) \; \mathbf{J}d\mathbf{x} \tag{13}$$

Fig. 3 shows a typical application of the correlation given by Eq. (13). The goal was to identify circular or rectilinear sand structures of variable sizes on the sea ground. These structures come from various human activities, such as the use of explosive devices, trawling or boat anchoring within the seagrass bed. This correlation measure algorithm allowed to recognize and detect traces of trawls independently of their position, orientation and distance to the sonar head.

3.4 Projective Group

We begin by considering the process of image formation when a scene is viewed through a camera. In particular we will consider image formation through a **pinhole camera**. A pinhole camera is a box in which



(a) Sea floor image.

Figure 4: The pinhole imaging model.

one of the walls has been pierced to make a small hole through it.

Assuming that the hole is indeed just a point, exactly one ray from each point in the scene passes through the pinhole and hits the wall opposite to it. This results in an inverted image of the scene, as can be seen in figure 4.

Let us begin by considering a mathematical description of the imaging process through this idealized camera; we will consider issues like lens distortion subsequently. The pinhole camera or the projective camera images the scene by applying a perspective projection that maps the 3D space to a 2D plane. The camera has the pinhole located at the origin and the image plane consists of points $(x_l, 1, x_3)$ identified with complex numbers $x_3 + i\,x_1$. Thus, the image of an object is obtained by projecting the object points into the image plane according to $\frac{x_3+i\,x_1}{x_2}$. The camera is embedded into the complex plane:

$$\mathbb{C}^2 = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} | z_1 = x_2 + i\,y, z_2 = x_3 + i\,x_1 \right\} \tag{14}$$

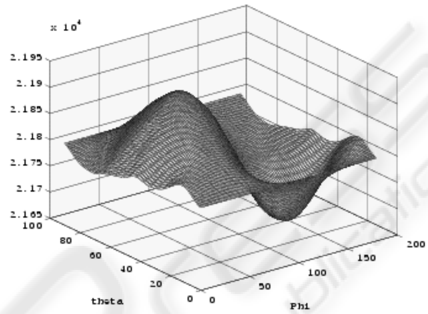
Geometry of the image plane $\mathbb{C}$, homogeneous under the action of $SL(2,\mathbb{C})/\{\pm I\}$ by linear-fractional transformations, can be dually described as follows:

1. $\mathbb{C}$ is the complex projective line with the group of projective transformations $PSL(2,\mathbb{C}) = SL(2,\mathbb{C})/\{\pm I\}$. Thus, the image projective transformations acting on the points of the image plane of the conformal camera can be identified, with projective geometry, to the one-dimensional complex line (Cox et al., 1996).
2. $\mathbb{C}$ is the Riemann sphere since, under stereographic projection, we have the isomorphism $\mathbb{C} \cong S_{(0,1,0)}$. The group $PSL(2,\mathbb{C})$ acting on $\mathbb{C}$ consists of the bijective meromorphic mappings of $\mathbb{C}$ (Turski, 2004).

The following subgroup of $SL(2,\mathbb{C})$ acting on the image plane gives all basic classes of the image projective transformations of planar objects. l'eq. suivante):

$$\mathbf{SU}(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} |\ |\alpha|^2 + |\beta|^2 = 1 \right\} \tag{15}$$

(a) Image. (b) Perspective im.



(c) Correlation surface

Figure 5: Perspective projection estimation with correlation.

generates image projective transformations (with conformal distortions) by first rotating the projection in the camera of an image on $S_{(0,1,0)}$ and then projecting it back to the image plane. This follows from the fact that there is one-to-two correspondence between the group of rotations $S0(3)$ and $\mathbf{SU}(2)$-the universal double covering group of $S0(3)$.

Or, on the sphere, we can explicit correlation for functions defined on $L^2(S^2)$

$$(f^1 \otimes f^2)(g) = \int_{S^2} f^1(\omega) f^2(g^{-1}\omega)\, d\mu(\omega), \tag{16}$$

where $d\mu(\omega) = \sin(\theta) d\theta d\phi$ is the rotation-invariant volume measure on the sphere. The Fourier transform also exists and is given by the spherical Fourier transform:

$$\hat{f}(l,m) = \int_{S^2} f(\theta,\phi)\, \bar{U_l^m}(\theta,\phi)\, \sin(\theta)\, d\theta d\phi. \tag{17}$$

Fig. 5 shows an example of image matching based on correlation measure (16).

4 CONCLUSION

In this paper, the problem of registering two images or recovering a template in a complex scene is formulated according to the generalized correlation approach. From the representation theory, we have proposed a correlation function adapted to the group of transformations and the data space when it is identified to a homogeneous space. Interesting cases for image processing, i.e. the planar and vectorial similarity groups, the motion and an interesting subgroup projective transformations, are detailed and illustrated with real images.

In future work, we plan to provide an in-depth study and comparison of the algorithms behavior with respect to robustness, computation time and achievability in real contexts. An other objective will be to propose a correlation function on the whole group $SL(2, \mathbb{C})$, providing a way to register images for all kinds of projective transformations.

REFERENCES

Ballard, D. (1981). Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111122.

Ben Youssef, L. (2004). *Corrélation sur les Groupes pour l'Analyse des Formes et l'Estimation du Mouvement, Application aux Images Sonar*. Phd thesis, Univ. de Bordeaux 3. in french.

Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376.

Cox, D., Little, J., and O'Shea, D. (1996). *Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer- Verlag.

Derrode, S. and Ghorbel, F. (2001). Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description. *Computer Vision and Image understanding*, 33(1):57–78.

Gauthier, J., Bornard, G., and Silbermann, M. (1991). Motions and pattern analysis : Harmonic analysis en motion groups and their homogeneous spaces. *IEEE trans. on Systems, Man, and Cybernetics*, 21(1):159–172.

Ghorbel, F. (1994). A complete invariant description for gray-level images by the harmonic analysis approach. *Pattern Recognition Letters*, 15:1043–1051.

Miller, M. and Younes, L. (2001). Group actions, homeomorphisms, and matching: A general framework. *J. of Computer Vision*, 41(1):6184.

M.J. Swain, D. B. (1991). Color indexing. *Internat. J. Comput. Vision*, 7(1):1132.

Pintsov, D. (1989). Invariant pattern recognition, symmetry, and Radon transforms. *J. of the Optical Society of America A*, 6(10):1544–1554.

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

Segman, J. (1992). Fourier cross correlation and invariance transformation for an optimal recognition of functions deformed by affine groups. *J. of the Optical Society of America A*, 9:895–902.

Segman, J. and Zeevi, Y. (1993). Image analysis by wavelet-type transforms: Group theoretic approach. *Mathematical Imaging and Vision*, 3(1):51–77.

S.X. Liao, M. P. (1996). On image analysis by moments. *IEEE Trans. Pattern Anal. Machine Intell*, 18(3):254266.

Turski, J. (2004). Geometric Fourier analysis on the conformal camera for active vision. *Society for industrial and applied mlathematics*, 46(2):230–255.

Y. Lamdan, H. W. (1988). Geometric hashing: a general and efficient model-based recognition scheme. *Proc. ICCV*, page 238249.

Zitovà, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000.