PERIODICA POLYTECHNICA SER. MECH. ENG VOL. 37, NO. 4, PP. 237–250 (1993)

IMAGE POSITIONING AND CAMERA CALIBRATION WITH A GROUP THEORY BASED TECHNIQUE

Igor Lvovich EROSH* and Irine Gennadievna GLADKOVA**

Aviational Instrument Making Institute, S.-Petersburg[1]
* USSR, Saint-Petersburg
Smolenskaya Nabereznaya 1–537.
Phone: 352–32–09.
** USSR, Saint-Petersburg
Leninskij Prospect 92/1–593.
Phone: 143–17–45.

Received: September 8, 1993

Abstract

Group transformation theory and group representation theory are very effective tools for motion estimation and deriving invariants of an image. We study the properties of the projection system which enables the description of 2D images, their motions and geometrical distortions in terms of group transformations. We consider the decomposing image function in the special orthonormal basis containing Krestenson functions. Two types of transformation are investigated: rotation and scaling of an image.

Keywords: invariant, group transformations, image distortion, image motion.

Introduction

The subject of this paper is referred to as motion analysis and camera geometrical calibration. Motion analysis capabilities are required for robotic visual guidance systems, visual inspection systems, autonomous aircraft landing and navigation systems and for many other image processing and computer vision applications. We measure 'motion' by the estimation of the parameters of a transformation that relates scene object or (perhaps as an intermediate step image) point locations. We use the geometrical model of the perspective projection, (p-p) transform described by SCHALKOFF (1989). Our technique is in accord with the approaches based on the application of the algebraic forms for pattern recognition (HU, 1962; SADJADI and HALL, 1980). Unlike the approaches using the statistical properties of an image (or — the properties of an ensemble of images) and performing the spatiotemporal interpolation of image perturbations, in our technique the motion is estimated through the determination of precise geometric

[1]The paper was written when Gladkova I. had the scholarship at the Technical University of Budapest (at the Institute of Precision Mechanics and Optics).

parameter values. The mechanism of our approach contains 4 main steps: (1) The description of image transformations and geometrical distortions in terms of group theory and group representation theory. (2) The translation of an image function is defined in an Euclidean space — 'f' to that is defined on the correspondent group tranformation elements set — 'F'. (We discuss the choice of the special projection system in the image plane which enables the performance of the translation: $f \to F$) (3) The decomposing of the original and perturbed image functions f and F in the orthonormal basis proposed by VILENKIN (1965). To be concise in our paper, we do not explore this basis extensively. We must only mention that good results are available when Krestenson transformations are considered. (4) The matching of f and F in the spectral space with analysis of the correspondent phases and modules.

Key Concepts of Group Transformation and Group Representation Theories Applied to Image Perturbations

Recalling the p-p based geometric model we consider the motion of an object point in physical coordinates from x_0 to x_0'. This yields a 3-D motion vector $(x_0' - x_0)$. This 3-D motion vector has a 2-D image plane projection: the form $b(x_i) = x_i - x_i'$ identifies the image plane motion. The 2-D affine transform is used to model small image perturbations (SCHALKOFF, 1989; EROSH and MOSCALEV, 1985). For an image function of the form $f(x)$ where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1}$$

the general affine transformed version of this function is denoted by

$$f'(x_i) = f(x_i'), \tag{2}$$

where

$$x_i' = Ax_i + b. \tag{3}$$

Thus, the affine transform represents a linear transformation of the image plane onto itself. Expanding *Eq.* 3:

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{4}$$

In *Eq.* 3, A is the homogeneous affine transform matrix, and b is the translating vector. Two well-known versions of the homogeneous affine transform $(b = 0)$ are discussed in our paper.

Magnification or dilation:

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \tag{5}$$

and rotation about the origo with an angle θ:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{6}$$

Note that $b = 0$ if global coordinate system (e. g. the coordinate system of the scene) is coincident with the image plane – centric systems. Basically, the origin of the image – plane coordinate system is placed in the 'center of mass' and moment invariance properties are used to obtain this coincidence (HU, 1962; EROSH and MOSCALEV, 1985).

The case of *Eq.* 5 is available when optics is not tuned and raster geometric distortions take place. Linear operator A describes the element of the group of the hyperbolic rotation in the case of *Eq.* 5 and the element of the group of the trigonometric rotation in the case of *Eq.* 6.

Let us denote a transformation element by g, the set of transformation elements by G, $g \in G$, group operation by $*$. G is considered to be a group if conditions $(7')$-$(7''')$ are fulfilled.

1. Association of the group operation:

$$(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3). \tag{$7'$}$$

2. Existence of the unique element $e \in G$, called the unitary element so that

$$\forall \quad g \in G \quad g * e = e * g = g \text{ is true.} \tag{$7''$}$$

3. Existence of the unique inverse element $g \in G$ so that

$$\forall \quad g \in G \quad \exists \quad g^{-1} \in G|, \quad g * g^{-1} = e \text{ is true.} \tag{$7'''$}$$

If the condition $(7'''')$ is fulfilled, a group of transformations is called commutative:

4. Commutative law:

$$\forall \quad g_1, g_2 \in G \quad g_1 * g_2 = g_2 * g_1. \tag{$7''''$}$$

Both groups are commutative.

Linear operator establishes the correspondence between old $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and new $\begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix}$ coordinates of the perturbed point location.

Discrete trigonometric rotation is defined by discrete parameter θ_k:

$$\theta_k = (2\pi/n) \times k \tag{8}$$

with $k \leq n$.

If we regard the group element g_α as the image rotation with angle $\alpha = 2\pi/n$, we can denote:

$$g_2 = g * g = g^2$$
$$\dots$$
$$g_k = g * g * \dots * g = g^k.$$

Hence,

$$g^n = e. \tag{9}$$

According to the group representation theory (VILENKIN, 1965), the commutative group has a one-dimensional representation Γ. Let us consider:

$$\begin{cases} \Gamma(e) = 1; \\ \Gamma(g) = \gamma; \\ \Gamma(g^k) = (\Gamma(g))^k = \gamma^k. \end{cases} \tag{10}$$

From (10) we have:

$$\gamma = \exp(j \cdot 2\pi k/n), \qquad k = 0, 1, \dots, n-1. \tag{11}$$

If we set a number of frequencies $\omega = 0, 1, \dots, n-1$, we can get the set of the linear representations. Thus, for $n = 5$ we have:

$\Gamma\omega\backslash$	e	g	g^2	g^3	g^4
Γ_0	1	1	1	1	1
Γ_1	1	ψ_1	ψ_2	ψ_3	ψ_4
Γ_2	1	ψ_2	ψ_4	ψ_1	ψ_3
Γ_3	1	ψ_3	ψ_1	ψ_4	ψ_2
Γ_4	1	ψ_4	ψ_3	ψ_2	ψ_1

The element ψ_α of the representation $\Gamma\omega$ can be written in the form:

$$\psi_\alpha = \exp\left(j\frac{2\pi\alpha}{n}\right), \tag{12}$$

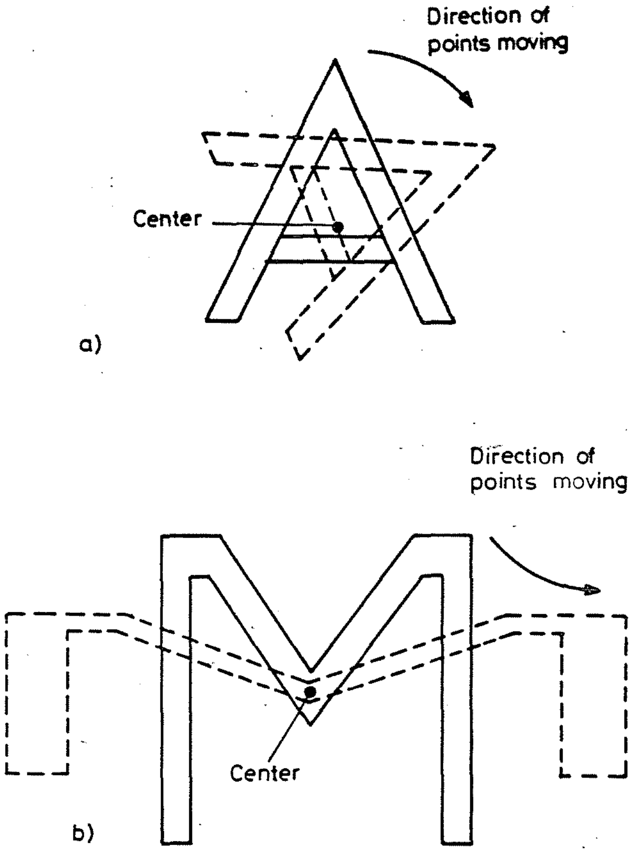


Fig. 1. Image transformations examples
a) case of trigonometric rotation
b) case of hyperbolic rotation

with $\alpha = \omega k \mathrm{mod}(n)$.

The representation space is an orthogonal direct sum of the non-equivalent in pairs, finite-dimensional unitary irreducible representations:

$$\Gamma(k) = \sum_{0}^{n} \Gamma\omega(k). \tag{13}$$

The result of a trigonometric rotation is shown in *Fig. 1a*.

The result of a hyperbolic rotation is plotted in *Fig. 1b*. The linear operator directing the group transformations depends on one parameter —

α (EROSH and MOSCALEV, 1985). The representation of the hyperbolic rotation has some singularities. Since the group is non-compact, the representation space is the continuous direct sum of the nonequivalent in pairs, irreducible representations:

$$\Gamma(t) = \int_{\infty}^{\infty} \Gamma\omega(t)dt \tag{14}$$

with ω — frequency.

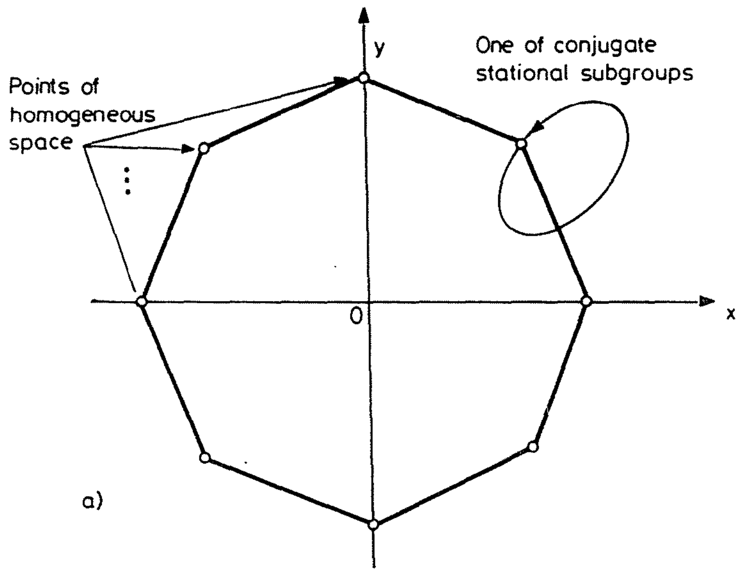
Design of Special Projection System Allowing Simple Description of Image Transformations

To perform the translation of the function defined in the Euclidean space into the function defined on the group elements, set of the correspondence between classes of the conjugate stational subgroups and the points of the homogeneous space is established. By *Fig. 2a,b* related to the case of the trigonometric rotation one can understand the basic principles of the translation. By choosing the homogeneous space (the circle from the set of concentric circles) we obtain the class of conjugate stational subgroups related to the space points. Having projected the image to that defined on the homogeneous space points, we get the image function defined on the group elements set.

The projection lines belong to the set of the homogeneous spaces of the group Q such that $\forall \quad g \in Q$, $g \in G$ (G — the current group) it is true: $q * g = g * q$ and any two points of Euclidean space can be transformed into each other by an element belonging to the group $\{q * g\}_{q,g}$.

Similarly, the homogeneous space for the group of hyperbolic rotation consists of the equilateral hyperbolas (*Fig. 2b*).

VILENKIN (1965) proved that the complete set of the non-equivalent in pairs finite dimensional irredicable representations of the group elements constitutes the orthogonal basis for decomposing the function defined on the group elements set. He also showed that the group parameters can be estimated using the spectral representation of the function. We use his conclusions for our purposes of the motion estimation and raster distortion correction.



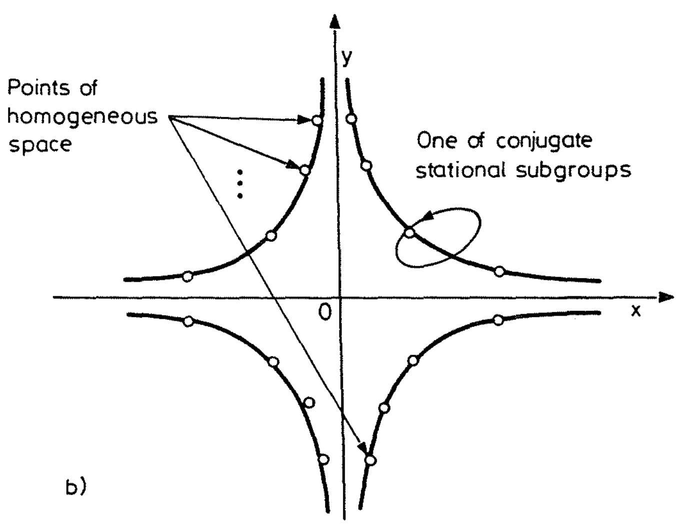


Fig. 2. Projection space
a) case of trigonometric rotation
b) case of hyperbolic rotation

Practical Procedures for Deriving Invariants and Parameters Estimation of Image Positions

Let us consider the industrial situation when objects are transported by the assembling table or by conveyer. The objects can rotate and can be overtuned??. This situation often takes place in robotic assembling production or when the visual inspection is performed.

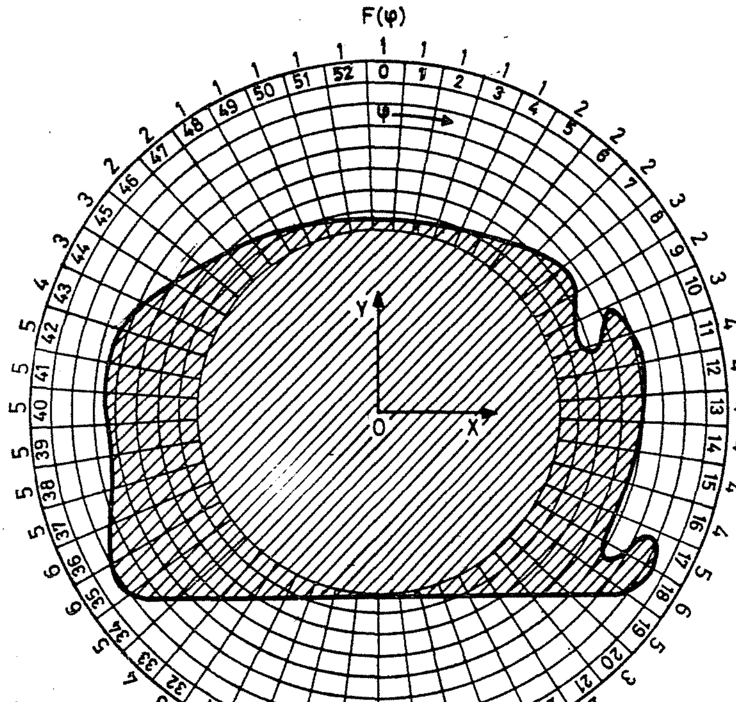


Fig. 3. Projection of image function onto 53-sided regular polygon

In *Fig. 3* one can see the projection of the object onto the n-sided regular polygon, $n = 53$, O is the 'center of mass' and the center of the circles constituting the homogeneous space (*Fig. 2a*). The shape of the object along the line of projection $\mathcal{L}k$ is characterized by the specially defined function $\Phi(\mathcal{L}k)$. This function may be defined as the integral sum of the non-zero values of image function along the line of projection:

$$\Phi k = \Phi(\mathcal{L}k) = \int_{\mathcal{L}k} f(x_i) \tag{15}$$

or for instance, as the maximum length of the non-zero continuous region of image function along the line of projection, etc. Let us assign the value of function $\Phi(\mathcal{L}k)$ to the n-sided regular polygon. In *Fig. 3* binary image is used and function Φ is defined according to (14). Thus, we perform the translation and obtain the image function defined on group elements set: $\Phi(k)$. Object rotations yield the discrete rotations of the n-sided regular polygon and the automorphism is available. The above expression (11) describes the system of the group elements, being itself an orthonormal basis for a spectral representation of the image function $\Phi(k)$, k — group parameter. We use a prime number to choose n — thus, we obtain the singular value of the angle Θk. The system (11) (if n is a prime number) coincides with the system of the Krestenson functions (EROSH, 1981).

Spectral representation for $\Phi(k)$ is:

$$\begin{gathered} S(\omega) = \sum_{k=0}^{n-1} \Phi(k) \exp\left(j\frac{2\pi}{n}\omega k\right) = \\ = \sum_{k=0}^{n-1} \Phi(k) \cos\left(\frac{2\pi}{n}\omega k\right) + j\sum_{k=0}^{n-1} \Phi(k) \sin\left(\frac{2\pi}{n}\omega k\right) = \\ = A(\omega) + jB(\omega), \end{gathered} \tag{16}$$

with ω, $k = 0, 1, 2, \ldots, n-1$, ω — frequency, k — discrete angle of rotation. Spectrum (16) has a complex value and is characterized by modules:

$$|S(\omega)| = \sqrt{S(\omega)\overline{S(\omega)}} = \sqrt{A^2(\omega) + B^2(\omega)} \tag{17}$$

and phase:

$$\varphi = \operatorname{arctg}\left[B(\omega)/A(\omega)\right] \tag{18}$$

Modulus is invariant under image rotation and is used for the identification of the objects in the camera field of view. Phase contains the information about object rotation related to some normal position. Let us have the spectrum (16) for the base position: $S_0(\omega) = A_0(\omega) + jB_0(\omega)$. Then the phase of the normal position is defined as follows:

$$\varphi_0 = \operatorname{arctg}[B_0(\omega)/A_0(\omega)]. \tag{19}$$

Note that the frequencies in (19) and (18) are equal. Single-value calculation of φ, φ_0 is possible using the signs of the real and imaginable parts in (18), (19).

The rotation of the object for discrete angle k_1 yields the circle offset of the function $\Phi(k)$ for k_1 takes: $\Phi(k+k_1)$, $k_1 = 0, 1, \ldots, n-1$. The relationship between the spectral representations of $\Phi(k)$ and $\Phi(k+k_1)$ is given by

$$S_{\Phi(k+k_1)}(\omega) = S_{\Phi(k)}(\omega) \exp\left(-j\frac{2\pi}{n}\omega k_1\right) = \\ = |S(\omega)| \exp\left(j\frac{2\pi}{n}(\varphi_0 - \omega k_1)\right). \tag{20}$$

From (20) we have:

$$\frac{2\pi}{n}(\varphi_0 - \omega k_1) = \varphi\frac{2\pi}{n}. \tag{21}$$

(21) yields:

$$k_1 = \frac{\varphi - \varphi_0}{\omega} \mod (n). \tag{22}$$

In (22) n is the modulus of the congruence.

When robot teaching is performed, we define the frequency ω_m corresponding to the maximum value of the spectral representation of the function $\Phi(k)$, this frequency should be used at the working stage. It is also necessary to calculate the phase of the spectral representation of function $\Phi(k)$ being acquired in the normal position of a scene and chosen by a teacher.

The Practical Analysis of Image Distortion Parameters

If optics of a sensor is not tuned, different geometric distortions arise. Examples of geometric distortions are shown in *Fig. 4*. To achieve the geometric distortion correction, we require to model the appropriate transformation of an image. There are lots of techniques, described in literature (SCHALKOFF, 1989; KARARA, 1980) that use the set of correspondent points to approximate the transformation by different polynomial models. We extend our group representation-based approach to the problem of the geometric distortion correction. In the paper we consider the case of the scale transformations (*Fig. 4b*), which are described by the group of the hyperbolic rotation. As with the case of an image rotation, 4 entities are required. We already pointed out the properties of the group of hyperbolic

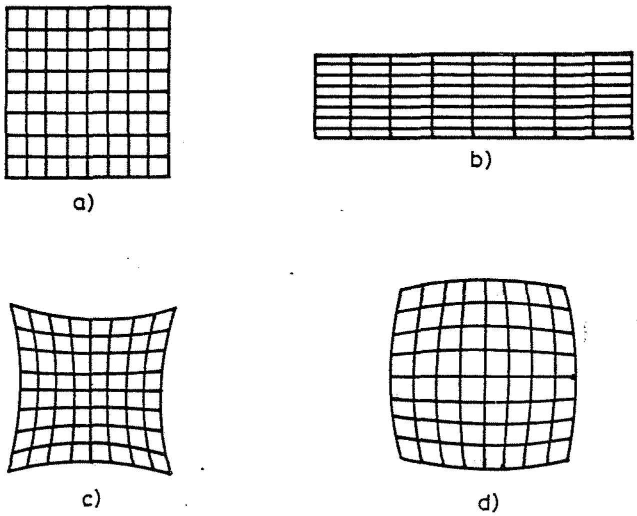


Fig. 4. Examples of raster geometric distortions
a) original image
b) scale distortions
c) pincushion distortion
d) barell distortion

rotation and considered the homogeneous space and the projection system for it (*Fig. 1b, 2b*).

The representation space of the hyperbolic rotation is the continuous direct sum of the non-equivalent subspaces yielding the spectral representations of the image function defined on group elements set in the form of Fourier integral:

$$\Phi(\alpha) = \int_{-\infty}^{\infty} F(\lambda) e^{j\lambda\alpha} d\lambda, \tag{23}$$

$$F(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\alpha) e^{-j\lambda\alpha} d\alpha. \tag{24}$$

In industrial systems we use the test objects and consider finite-dimensional, discrete image function $\Phi(k)$ of them. Given the step $\Delta\alpha$ for the group parameter α we obtain the discrete parameter α_k:

$$\alpha_k = k\Delta\alpha \tag{25}$$

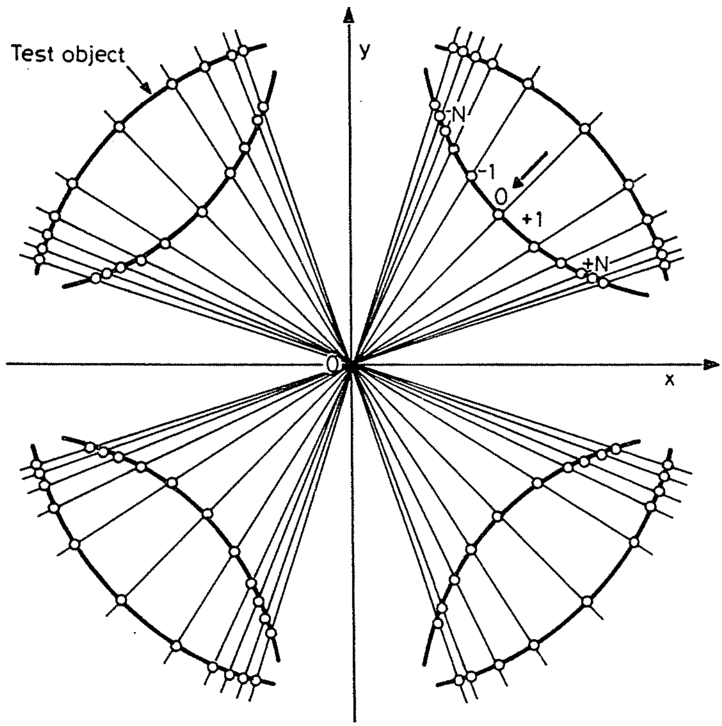


Fig. 5. Projection of image function onto hyperbolas

and define the bounds for k : $k = \overline{-N, N}$. Thus, the field of view (*Fig. 5*) is restricted. Practically, the values of N, $\Delta\alpha$ should be chosen so that in any case of a distortion parameter k is an integer. Given λ_n equal with

$$\lambda_n = \lambda_0 + n\Delta\lambda, \qquad n = \overline{-N, N} \tag{26}$$

we establish the relationship:

$$\hat{F}(\lambda_n) = \Delta\alpha \sum_{k=-N}^{N} \Phi(\alpha_k) e^{-j2\pi\lambda_n\alpha_k}, \tag{27}$$

$$\Phi(\alpha_k) = \Delta\lambda \sum_{k=-N}^{N} F(\lambda_n) e^{j2\pi\lambda_n\alpha_k}. \tag{28}$$

(27) and (28) are the discrete versions of (23) and (24). Since function $\hat{F}(\lambda_n)$ has a period $1/\Delta m$, we may demand: $|\lambda_n| < 1/2\Delta m \ \forall \ n \in$

$[-N, N]$ and use for λ_n the interval of change $[0, 1/\Delta m]$ instead of interval $[-1/2\Delta m,\ 1/2\Delta m]$. Given $\lambda_n = 1$, we obtain $0 < \Delta\alpha \leq 1$. From (27) and (25) we have:

$$\hat{F}(1) = \Delta\alpha \sum_{k=-N}^{N} \Phi(\alpha_k) e^{-j2\pi\Delta\alpha_k}. \tag{29}$$

There is the relationship between spectral characteristics of original (prime 0) and distorted (prime k_0) images, with $\lambda_n = 1$:

$$\hat{F}^{k_0}(1) = \hat{F}^0 e^{j2\pi\Delta\alpha_{k_0}}. \tag{30}$$

The difference of phases φ and φ^0 is used to obtain group parameter:

$$k_0 = \varphi(1) - \varphi^0(1). \tag{31}$$

The number of projection lines in a sector $[-N, N]$ should be enough to acquire explicit image functions of the test object used for calibration. Let us have $(2P + 1)$ lines to obtain the origin image function with required explicity. So we can calculate the sector β_0 (*Fig. 5*):

$$\beta_0 = \frac{\pi}{2} - 2\,\mathrm{arctg}\left(e^{-2\Delta\alpha_p}\right). \tag{32}$$

When the test object is distorted, the corresponding points of hyperbola shift up or down for T positions and sector β_0 is expanded to that defined as follows:

$$\beta = \frac{\pi}{2} - 2\,\mathrm{arctg}\left(e^{-2\Delta\alpha_{(p+T)}}\right). \tag{33}$$

The overall number of points is equal to :

$$2N + 1 = 2(p + T) + 1. \tag{34}$$

The values of p and T are chosen according to the explicit of calibration we want to obtain.

Conclusions

The computational procedure used to obtain both motion and geometric distortion parameters with the above mentioned techniques is often referred to as camera calibration. The general calibration scheme representing group theory-based approach is shown in *Fig. 6*. We use the test

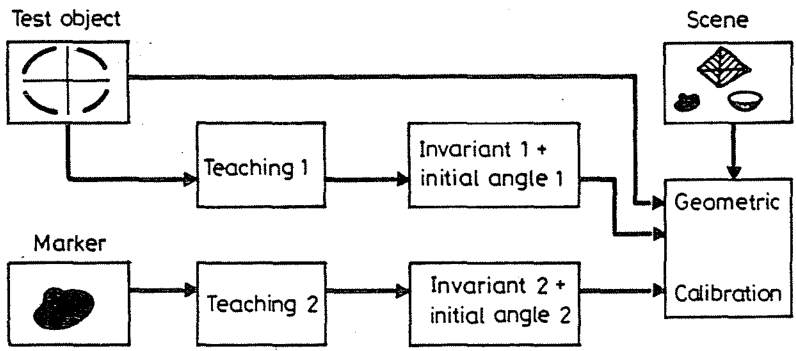


Fig. 6. General scheme of calibration

image consisting of noncontinuous concentric circles for geometric distortion calibration.

There are different applications where we have achieved good result in the explicity using the described technique: — In the land remote sensing one can estimate the distances between important points matching the acquired image and the map in the spectral space by (20), (27). — In the robotic assembling manufacturing the position of objects can be defined by calculation the phase according to (22).

References

1. SCHALKOFF, R. L. (1989): Image Processing and Computer Vision. USA, Wiley.
2. HU, M. K. (1962): Visual Pattern Recognition by Moment Invariants, *IEEE Trans. Inform. Theory*, Vol. IT–8, pp. 1979–1987.
3. SADJADI, F. A. — HALL, E. L. (1980): Three Dimensional Moment Invariants, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI–2, pp. 127–136.
4. VILENKIN, N. (1965): Specialnye funkcii i teoriya predstavlenij grup (Special Functions and Group Representation Theory). Moskva, Nauka (in Russian).
5. EROSH, I. L. — MOSCALEV, E. S. (1985): Adaptivnye robototehnicheskiye sistemy (Adaptive Robotic Systems). S.-Petersburg (Leningrad). (in Russian).
6. EROSH, I. L. (1981): Primeneniye preobrasovanij Krestensona dlya vychisleniya parametrov polozeniya po ploskim proekciyam (Application of Krestenson Transformations for Estimation of Position Parameters Using Flat Projections). *Technicheskaya cybernetica*, N 3, pp. 46–52. (in Russian).
7. KARARA, H. M. (1980): Non-Metric Cameras. In Atkinson, K. B. (Ed.). Developments in Close-Range Photogrammetry. London, Applied Science Publishers.