

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300

Open access books available

130,000

International authors and editors

155M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.

For more information visit [www.intechopen.com](http://www.intechopen.com)



# Projective Registration with Manifold Optimization

Guangwei Li<sup>1) 2)</sup>, Yunpeng Liu<sup>1)</sup>, Yin Jian<sup>3)</sup> and Zelin Shi<sup>1)</sup>

1) *Shenyang Institute of Automation, CAS  
P.R.China*

2) *Qingdao University  
P.R.China*

3) *The Research Institute on General Development and Argumentation of Air Force  
Beijing, P.R.China*

## 1. Introduction

Image registration is of interest to scientists and engineers of various fields: computer vision, pattern recognition, and robotics ( Trucco & Plakas, 2006 ). One of the remarkable problems both in theory and in technique is how to cope with the dynamic geometric-warps. For small objects and large camera-to-scene, i.e., the background appearance of the target is enough far, the projective transformation effect of the target is negligible. A satisfactory tracking result can be achieved by means of approximating the geometric warps with the affine transformation. However, in many special and practical applications, such as image mosaics in computer graphics and vision guidance in the military field, the projective transformation should be considered. It is well known that the projective transformation exactly models the motion relationship between images of the identical planar object scene, and can describe the pan and the tilt of the camera, which the affine transformation cannot do (Mann & Picard, 1997). The projective map model has eight independent parameters which have widely varying sensitivities and the transformation is highly nonlinear (Michael Gleicher, 1997), all of which affect to design the registration algorithms with efficiency, accuracy and robustness.

Lucas-Kanade image registration method was first proposed in 1981. Within the classical space transformation-based tracking framework proposed by Hager et al (Hager & Belhumeur, 1998), a projective image registration approach based on the matrix parameterization was presented (Buenaposada & Baumela, 2002) in terms of the forward-addition algorithm on the vector space, which is called the VECTOR-GN algorithm in this paper . The inverse-composition algorithm was proposed (Baker & Matthews, 2004) not only to compute the Hessian matrix and the gradient matrix offline but also to improve the efficiency by improving the iterative structure. However, these strategies can not utilize the projective Lie group structure sufficiently and leave room to improve the performance of the image registration algorithms.

The geometric optimization algorithm based on the manifolds, a novel approach to solve the constrained problem, was proposed in the 1970s-1980s ( Gabay, 1982). The fundamental idea of this approach is to regard the constrained sets as one underlying manifold and to exploit the geometry of the underlying parameter space. That is to develop a strategy which views constrained problems as to be equivalent to unconstrained problems posed on the constraint sets. Philosophically, this approach is geometrically more intuitive, conceptually simpler and mathematically more elegant. Various problem have been solved by applying the optimization algorithms on various manifolds. Especially, the optimization algorithms based on Lie Groups (Owren & Welfert, 1996) and Riemannian manifolds (Yaguang, 1999) have already been applied to robot control and machine learning. Recent years have also witnessed the rich achievements in the fields of signal processing, computer vision and pattern recognition. Smith generalized some optimization algorithms on the vector space to Riemannian manifolds and studied the adaptive filter problem of nonlinear signal (Smith, 1993). Yean considered the optimization algorithm on  $SE(n, R)$  about the 2D-3D pose estimation in computer vision (Yean, 2005). Grenander proposed his famous General Pattern theory, of which the deformable template idea (Grenander, et al., 1998) is that the object is represented by the template and the infinite varieties of the pose and location associated with its occurrences are represented via transformations which act on the template. These transformations form one transitive group acting on the space of all possible transformed templates, which becomes a Lie group orbit. Hence, the problem on the automated target recognition and tracking switches into the parameter optimization problem on the Lie Group manifolds.

Exploiting the deep connection between the Lie group and its associated Lie algebra which is called the Lie group exponential map, the geometric optimization approach based on Lie Groups theory switches the constrained nonlinear problems into equivalently unconstrained problems, thereby significantly reducing the computational complexity. A novel homography-based target image registration and tracking approach based on the Lie algebra parameterization which is called the LEXP-GN algorithm in this paper was proposed (Eduardo & Jaime, 2007). The performance, such as the tracking precision and rate, is better than that of the tracking method based on the matrix parameterization.

Noticeably, there exists a bi-invariant Riemannian metric on a compact Lie group (such as  $SO(n,R)$ ) and the geodesic through the identity element of group is one-parameter group. Hence, the Lie group exponential map agrees with the Riemannian exponential map. However, a noncompact Lie group (such as  $SE(n,R)$ ,  $SL(n,R)$  and  $GA(n,R)$ ) has not a bi-invariant Riemannian metric and the Riemannian exponential map based on the geodesic is usually different from the exponential map based on the Lie group structure. Therefore, the geometric optimization algorithms on the noncompact Lie groups based on the Lie group exponential map have its limitations. To our knowledge, it seems that there is not very much research on the noncompact Lie group optimization. Mahony and Manton provided an instructive interpretation of the Newton optimization method on the noncompact Lie groups from the Cartan-Schouten connection views of Riemannian geometry (Mahony & Manton, 2002). However, the Newton method needs to compute the complicated Hessian matrix and is usually not feasible to be applied to the real time application.

The core of our registration algorithm is the optimization problem on the special linear group  $SL(3,R)$ . Recently, Seok, et al studied the optimization algorithm on  $SL(3,R)$  about the medical images registration problems (Seok, et al, 2007). Based on the Riemannian

exponential map obtained from the geodesic equation, we propose a second-order efficient target tracking algorithm within the intrinsic geometric optimization framework. The comparative experiments with VECTOR-GN and LEXP-GN indicate the improvement on the tracking rate and precision.

The rest of the chapter is organized as follows. After a brief introduction to the Lie group exponential map and the Riemannian exponential map, the connection between them is studied on section 2, with the geometric optimization framework. Section 3 investigates the second-order projective image registration algorithm based on the Riemannian exponential map within the intrinsic optimization framework. Some comparative results are shown for illustration and verification in section 4. Finally, section 5 concludes the investigation and proposes some further work. Some necessary supplementary material will be given in section 6.

## 2. Mathematical background

The exponential map and the intrinsic geometric optimization algorithm build the basis for our efficient projective registration method. The tools used here come primarily from Lie group and Riemannian geometry. To enable further discussion, we need to take a small detour into geometry on them. Further information can be found in the famous textbooks ( Helgason, 1978; Berger, 2003).

### 2.1 Lie group exponential map

A Lie group is a group endowed with the smooth manifold structure, and its group multiplicative operation is denoted by  $\circ$ . The tangent space at the identity element  $e$  of Lie group  $M$  is denoted by  $T_e M$ . Let assume  $m \in M, X_e \in T_e M$ . The left-invariant vector field  $X$ , which determines a left-invariant flow  $\varphi_x(t, m) = X_t(m)$ , can be obtained by the left-translation  $X_m = (L_m)_* X_e$ . The one-parameter group, an integral curve at  $e$ , is denoted by  $\gamma_x(t)$ . The vector space  $(T_e M, [\cdot, \cdot])$  equipped with a bilinear bracket operation is a Lie algebra denoted by  $\Lambda(M)$ . It is known that the left-invariant vector field, the left-invariant flow, the tangent space at the identity element, one-parameter group and the Lie algebra are equivalent in essence.

**Definition 1.** Lie group exponential map  $\text{Lexp} : \Lambda(M) \times \mathbb{R} \rightarrow M, (X, t) \rightarrow \text{Lexp}(tX) = \gamma_x(t)$

For convenience, we usually define the Lie group exponential map,  $\text{Lexp} : \Lambda(M) \rightarrow M$ , as follows

$$\text{Lexp}(X) = \gamma_x(1) \quad (1)$$

**Lemma 1.** There exists an open neighborhood  $W$  of 0 in Lie algebra  $\Lambda(M)$  and an open neighborhood  $U$  of  $e$  in  $M$  such that  $\text{Lexp}$  is an analytic diffeomorphism of  $W$  onto  $U$ .

From Lemma 1, we can define its inverse function known as the logarithm map which returns  $X = \log y$  such that  $\text{Lexp}(X) = y$ . (See Fig. 1).

The space of all  $n \times n$  nonsingular real matrices forms a Lie group, called the general linear group denoted by  $GL(n, \mathbb{R})$ . Its algebra is usually denoted by  $\mathfrak{gl}(n, \mathbb{R})$ , the set of all real

square matrices. Being a sub-group of  $GL(n, R)$ , the special linear group  $SL(n, R)$  is the space of all real  $n \times n$  matrices  $H$  satisfying  $\det H = 1$ . Its Lie algebra denoted by  $\mathfrak{sl}(n, R)$  consists of the real matrices of trace zero. What we concern in this paper is  $SL(3, R)$  whose Lie algebra is  $\mathfrak{sl}(3, R)$  with the following basis vectors.

$$\begin{aligned}
 e_1 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & e_2 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} & e_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & e_4 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 e_5 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & e_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & e_7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} & e_8 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}
 \end{aligned}$$

For matrix Lie groups, the group operation is matrix multiplication. The Lie bracket operation is  $[A, B] = AB - BA$  and the Lie exponential map of a matrix  $A \in \mathfrak{gl}(n, R)$  is computed by the formula

$$L \exp A = \exp A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \quad (2)$$

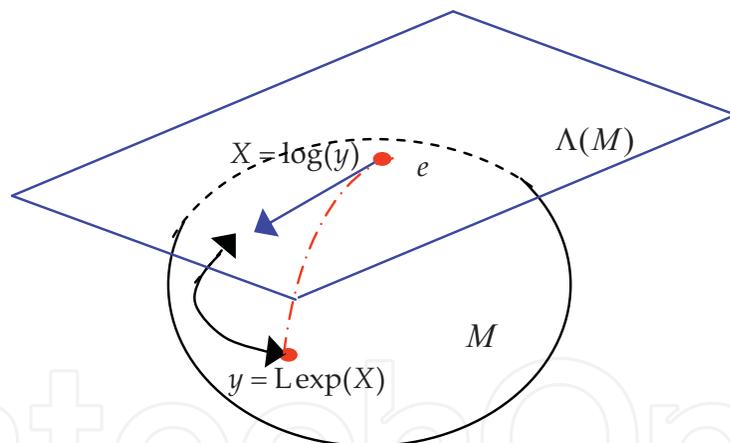


Fig. 1. Lie group exponential map and its inverse map

## 2.2 Riemannian exponential map

Let  $M$  be a smooth manifold of  $m$  dimensions. For every point  $p \in M$ , if an Euclidean inner product  $g(p) = \langle, \rangle: T_p M \times T_p M \rightarrow R$  is assigned on its tangent space,  $(M, g)$  is called a Riemannian manifold of  $m$  dimensions and  $g$  is called its Riemannian metric

Let  $M$  be a smooth manifold of  $m$  dimensions. For every point  $p \in M$ , if an Euclidean inner product  $g(p) = \langle, \rangle: T_p M \times T_p M \rightarrow R$  is assigned on its tangent space,  $(M, g)$  is called a Riemannian manifold of  $m$  dimensions and  $g$  is called its Riemannian metric.

Let  $\gamma : [a, b] \rightarrow M$  be a smooth curve in  $M$ . Its length is defined as

$$L(\gamma) = \int_a^b |\dot{\gamma}(t)| dt = \int_a^b \sqrt{\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle} dt \quad (3)$$

A smooth curve  $\gamma(t)$  is called a geodesic on Riemannian manifold if the family of tangent vector field  $\dot{\gamma}(t)$  is parallel with respect to  $\gamma(t)$ . The geodesic has an important property that it is the minimal length curve of the following energy function

$$E(\gamma(t)) = \int_a^b |\dot{\gamma}(t)|^2 dt \quad (4)$$

Let  $\gamma(t) = \gamma(t; p, v)$ ,  $t \in [0, 1]$  is a geodesic on Riemannian manifold satisfying  $\gamma(0) = p$ ,  $\dot{\gamma}(0) = v$ , and  $B_p(\varepsilon)$  be an open ball on  $T_p(M)$  whose center is origin and radius is  $\varepsilon$ . The Riemannian exponential map at  $p$  is defined as follows.

**Definition 2.** Riemannian exponential map:  $\text{Rexp}_p : B_p(\varepsilon) \rightarrow M$ ,  $\text{Rexp}_p(tv) = \gamma(t; p, v)$ .

In what follows, we denote the map  $\text{Rexp}_e$ , the Riemannian exponential map from the tangent space  $T_e M$  at the identity element  $e$  of the Lie group  $M$ , by  $\text{Rexp}$ ,

$$\text{Rexp}(v) = \text{Rexp}_e(v) = \gamma(1) \quad (5)$$

**Lemma 2.** Let  $(M, g)$  be a Riemannian manifold of dimensions. For any point  $p \in M$ , there exists an open neighborhood  $V$  of origin such that  $\text{Rexp}$  is an analytic diffeomorphism of  $V$  onto  $U = \text{Rexp}(V)$ .

From Lemma 2, if we define  $\text{Rexp}_p(v) = q$ , its inverse function known as the logarithm, can be defined as  $v = \text{Rlog}_p(q)$  such that,  $\gamma(0) = p, \gamma(1) = q$  and  $\dot{\gamma}(0) = v$ . (See Fig. 2)

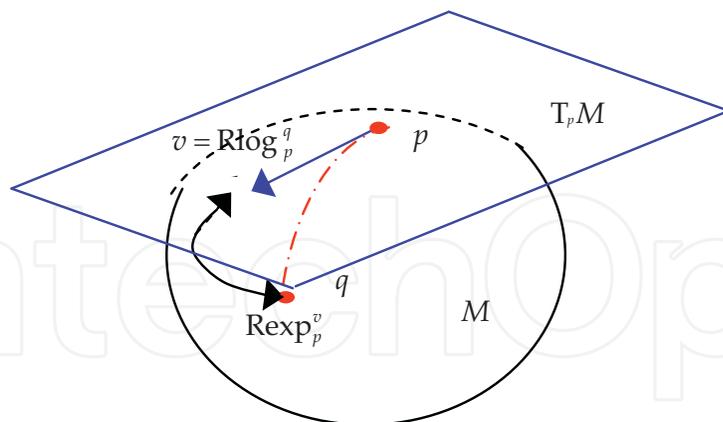


Fig. 2. Riemannian exponential map and its inverse map

For  $\text{SL}(n, \mathbb{R})$ ,  $\text{Rexp}(v)$  is one-to-one on  $\|v\| < \frac{1}{2}$  (Begelfor & Werman, 2005).

### 2.3 Relationship between Lie group exponential and Riemannian map

**Definition 3:** Let  $G$  be a Lie group,  $g$  be a Riemannian metric on  $G$ ,  $g$  is a left-invariant (right-invariant) Riemannian metric if its left-translation  $L_g$  (right-translation  $R_g$ ) is an isometric transformation on Riemannian manifolds. A metric is bi-invariant if and only if the metric is both left-invariant and right-invariant.

**Lemma 3.** There exists a bi-invariant Riemannian metric on a compact Lie group and the geodesics at the identity element  $e$  are one-parameter subgroups.

Lemma 3 shows that the Riemannian exponential map at the identity defined by the bi-invariant Riemannian metric agrees with the Lie group exponential map, that is, for any tangent  $v \in \Lambda(M)$

$$\text{Lexp}(v) = \text{Rexp}(v) \quad (6)$$

Let us consider the general linear group  $GL(n, \mathbb{R})$ . For every point  $p$  on the tangent space  $T_p M$ , inner product  $\langle \cdot, \cdot \rangle$  is defined as follows

$$\langle A, B \rangle = \text{Tr}(AB^T) \quad (7)$$

where  $A, B \in T_p M$ , and  $\text{Tr}(\cdot)$  is the trace of a matrix. Hence, the length of a tangent vector is defined

$$\|v\|^2 = \text{Tr}(vv^T) \quad (8)$$

Now, we begin to consider the Riemannian exponential map on a noncompact Lie group. First, we propose the following theorem on the minimal geodesics with respect to the right-invariant Riemannian metric on general linear group manifolds. Please pay attention to the point  $g$  on the Riemannian manifold, which is not the Riemannian metric.

**Theorem 1.** Let  $g(t)$ ,  $t \in [0, 1]$  be a minimal geodesic connecting  $g, h \in GL(n, \mathbb{R})$ . The tangent vectors minimizing the energy function  $\int_0^1 \|v(t)\|^2 dt$  and satisfying  $dg(t)/dt = v(t)g(t)$ ,  $g(0) = g$ ,  $g(1) = h$  are the solutions of the following matrix differential equation

$$\begin{aligned} dv(t)/dt &= v(t)v^T(t) - v^T(t)v(t) \\ &= [v(t), v^T(t)] \end{aligned} \quad (9)$$

A proof of Theorem 1 is given in 6.1.

From Theorem 1, it follows that the equation of the minimal geodesic on  $GL(n, \mathbb{R})$  is

$$g(t) = \exp((v(0) + v(0)^T)t) \exp(-v(0)^T t) g \quad (10)$$

Therefore, the Riemannian exponential map, the equation of the minimal geodesic equation on  $GL(n, R)$  which emanates from the identity with a velocity  $v$  is expressed

$$\text{Rexp}(v) = \exp((v^T + v))\exp(-v^T) \quad (11)$$

Now, we consider the exponential map on the subgroups of  $GL(n, R)$ . The special orthogonal matrix  $SO(n, R)$  is a compact group, and its algebra  $so(n, R)$  are skew symmetric matrices with zero trace. It follows that, for every  $v \in so(n, R)$ , we have  $v = -v^T$ , yielding

$$\text{Rexp}(v) = \exp((v^T + v))\exp(-v^T) = \exp(v) \quad (12)$$

We can see that the Riemannian exponential map is the same as the Lie group exponential map for the compact groups, which completely consists with the Lemma 3. For the noncompact group  $SE(n, R)$ , its geodesics can be obtained by lifting the geodesics from  $SO(n, R)$  and  $R^n$  (Zefran & Kumar, 1998). However, for the noncompact group  $SL(n, R)$ , every  $v \in sl(n, R)$ ,  $\text{Rexp}(v) \neq \text{Lexp}(v)$ . We construct the optimization algorithm with the Riemannian exponential map  $\text{Rexp}(v) = \exp(v^T + v)\exp(-v^T)$ ,  $v \in sl(3, R)$  in this paper

#### 2.4 Framework for geometric optimization

If a Lie group is embedded in Euclidean space to be a sub-manifold, the optimization problem on it can often become a classical constrained optimization. The conventional approach of dealing with the structure of the group is to use Lagrange multipliers. Based on the geometric optimization theory, we use local canonical coordinates to represent parameters and intrinsically take care of the geometric structure of Lie Groups to allow the use of unconstrained optimization routines (Vercauteren & Malis, 2007).

Let  $x$  be a point in the neighborhood of  $t \in M$ . From Theorem 1, there exists  $\omega = \sum_i^n v_i e_i \in \Lambda(M)$  such that

$$x = t \circ \text{Lexp}(\omega) = t \circ \text{Lexp}\left(\sum_i^n v_i e_i\right) \quad (13)$$

where  $v = (v_1, v_2, \dots, v_n)^T$ ;  $e_i (i = 1, \dots, n)$  is the basis of Lie algebra  $\Lambda(M)$ . Then, the Taylor series of a smooth function  $\varphi(\cdot)$  on Lie group  $M$  is obtained

$$\varphi(t \circ \text{Lexp}(\omega)) = \varphi(t) + J_t^\varphi v + \frac{1}{2} v^T H_t^\varphi v + O(\|v\|^3) \quad (14)$$

where  $[J_t^\varphi]_i = \frac{\partial}{\partial v_i} \varphi(t \circ \text{Lexp}(\omega))|_{v=0}$  and  $[H_t^\varphi]_{ij} = \frac{\partial^2}{\partial v_i \partial v_j} \varphi(t \circ \text{Lexp}(\omega))|_{v=0}$

The Taylor series (14) allows us to construct various optimization algorithms on Lie groups by generalizing algorithms on vector space. For example, the classical Newton-Raphson method adopts the following intrinsic update step

$$t \leftarrow t \circ \text{Lexp}(\omega) \quad (15)$$

where  $v$  solves  $H_t^\varphi v = -[J_t^\varphi]^T \varphi(t)$ .

Unfortunately, in many cases, the Hessian matrix  $H_t^\varphi$  is often difficult or impossible to compute. Even worse is that the convergence problem may arise when it is not definite positive. Hence, Benhimane et al constructed the intrinsic Gauss-Newton algorithm by preserving the linear part and discarding the quadric item of Taylor series ( Benhimane & Malis 2007). Motivated by the second-order minimization method based on the Lie algebra parameterization, we take place of the Lie group exponential map with the Riemannian exponential map to construct an efficient second-order minimization algorithm based on the geodesics on manifolds, which is called REXPP-ESM algorithm in this paper. This algorithm can also find back the Hessian matrix information discarded by LEXP-GN algorithm within the intrinsic optimization framework to further improve the registration performance. The performance of the REXPP-ESM will be explained in 4.2.

Lemma 4. Any manifold of dimension  $d$  can be embedded in  $E^{2d+1}$  (Berger, 2003).

From Lemma 4, we know that for a Lie group, there naturally exists an embedding map  $\pi: G \rightarrow R^n$ ,  $t \rightarrow \pi(t)$  such that it is a sub-manifold in Euclidean space. Especially, the spatial transform groups (e.g. rigid body, affine, projective.) used in the target recognition and tracking are often represented and computed by common matrixes in homogeneous coordinates.

### 3. Projective registration with manifold optimization

#### 3.1 Problem Statement

Suppose the camera is not be calibrated and the tracked object has a flat appearance. When the target is moving in the space, the relation between images can be described by projective transformation. The projective transform group is the group of the matrices of the form

$T = \begin{bmatrix} R & t \\ v & 1 \end{bmatrix}$ , where  $R$  is a  $2 \times 2$  nonsingular matrix,  $t$  is a column vector for the translation

and  $(v, 1)^T$  is the projection of the line at infinity. We choose the scale factor to normalize the projective group matrices  $T$  such that the determinants of  $T$  are equal to 1. Then the matrices  $T$  belong to the special linear group  $SL(3, R)$ . This normalization cannot change the degree of parametric freedom and is reasonable in real applications [21].

From Lemma 4, we can suppose the homogeneous coordinate of point  $p$  be  $(x, y, 1)^T$  and the embedding map in Euclidean space of  $SL(3, R)$  be  $\pi: t \rightarrow \pi(t)$ . Define a group action from  $SL(3, R)$  on  $p: w: SL(3, R) \times p \rightarrow p$ . The projective transformation is represented as follows

$$w(\pi(t))(p) = \frac{1}{a_{31}x + a_{32}y + a_{33}} \begin{bmatrix} a_{11}x + a_{21}y + a_{13} \\ a_{21}x + a_{22}y + a_{23} \\ a_{31}x + a_{32}y + a_{33} \end{bmatrix} \quad (16)$$

Let  $I(p)$  be the brightness value of the template and  $I(w(t)(p))$  be the intensity of projective-transformed target in the input image. The algorithm assumes the gray value is invariable at the same target position in two consecutive frames and calculates the projective transformation parameters to know the current position where the target is in the current image by solving the following function

$$\arg \min \|I(w(\pi(t))(p)) - I(p)\|^2 \quad (17)$$

### 3.2 Optimum parameters for projective registration

To solve the projective parameters is in fact to perform optimization on  $SL(3, R)$ . Based on the Lie algebra parameterization,  $\omega = \sum_{i=1}^8 v_i e_i$ , where  $v = (v_1, \dots, v_8)^T$  is the incremental motion parameter vector, we take place of the Lie group exponential map with the Riemannian exponential map and the nonlinear least-squared optimization problem can switch into

$$\arg \min \|I \circ t \circ \text{Rexpp}(\sum_{i=1}^n v_i e_i)(p) - I(p)\|^2 \quad (18)$$

Let  $f_p(t \circ \text{expp}(\omega)) = I \circ t \circ \text{expp}(\omega)(p) - I(p)$ . we get

$$\begin{aligned} \|f_p(t \circ \text{Rexpp}(\omega))\|^2 &= \|I \circ t \circ \text{Rexpp}(\omega)(p) - I(p)\|^2 \\ &\approx \left\| I \circ t(p) - I(p) + J_t^{f_p} \Big|_{v=0} v + \frac{1}{2} v^T H_t^{f_p} v + O(\|v\|^3) \right\|^2 \end{aligned} \quad (19)$$

We pay attention to fact that when the images are aligned with the optimal spatial transformation in target tracking, the template and the warped image as well as their gradient should be very close to each other, i.e.  $\nabla_p I \circ t^* \approx \nabla_p I$ . An efficient tracking algorithm will be constructed by utilizing this information to recover the information discarded with Gauss-Newton method by means of expanding the Jacobian matrix at the optimal transformation  $t^*$ , hence avoiding computing the Hessian matrix at the same time. A first-order Taylor series around 0 of  $J_t^{f_p}$  in (20) can lead to

$$v^T H_t^{f_p} = J_t^{f_p}(v) - J_t^{f_p}(0) + O(\|v\|^2) \quad (20)$$

Incorporating this expression into (19), we can get a true second-order approximation

$$\|f_p(t \circ \text{Rexpp}(\omega))\|^2 \approx \left\| I \circ t(p) - I(p) + \frac{1}{2} (J_t^{f_p}(v) + J_t^{f_p}(0))v + O(\|v\|^3) \right\|^2 \quad (21)$$

The following is to compute the Jacobian matrix  $J_t^{f_p}(0)$  and  $J_t^{f_p}(v)$  corresponding to the derivative of at 0 and  $v$ .

$$\begin{aligned}
 J_t^{f_p}(0) &= \frac{\partial I \circ t \circ \text{Rexpp}(\sum_{i=1}^n v_i e_i)}{\partial v^T} \Big|_{v=0} \\
 &= \frac{\partial I \circ t(q)}{\partial q^T} \Big|_{q=p} \cdot \frac{\partial w(t,p)}{\partial t^T} \Big|_{t=\pi(\text{Id})} \cdot \left[ \frac{\partial \pi(\text{Rexpp}(\sum_{i=1}^8 v_i e_i))}{\partial v_1} \Big|_{v_1=0}, \dots, \frac{\partial \pi(\text{Rexpp}(\sum_{i=1}^8 v_i e_i))}{\partial v_8} \Big|_{v_8=0} \right] \quad (22) \\
 &= \nabla_p^T(I \circ t) \frac{\partial w(t,p)}{\partial t^T} \Big|_{t=\pi(\text{Id})} \cdot [\pi(e_1) \ \dots \ \pi(e_8)] \\
 &= \nabla_p^T(I \circ t) J^{w_p} e_\pi
 \end{aligned}$$

For every line of the matrix  $J_t^{f_p}(0)$ ,  $[\nabla_p^T(I \circ t)]_{1 \times 3}$  is corresponding to the spatial derivative of the current warped image using the projective transformation  $t$ ;  $[J^{w_p}]_{3 \times 9}$  is the Jacobian matrix for projective transformation (13), and  $[e_\pi]_{9 \times 8}$  is the Jacobian matrix where  $\pi(e_i)$  is the matrix  $e_i$  reshaped as a vector ( the entries are picked line per line ). Id is identical transformation. The two Jacobians  $J^{w_p}$  and  $e_\pi$  are constants to be computed once and for all while the Jacobian  $J_t^{f_p}(0)$  has to be computed at each iteration since it depends on the updated value of projective parameters.

However, the Jacobian matrix  $J_t^{f_p}(v)$  is complicated and usually depends on  $t$ . Hence, we do not directly compute  $J_t^{f_p}(v)$ . If replacing the gradient of the optimally warped image  $I \circ t^* = I \circ t \circ \text{expp}(v_i^*)$  by its equivalent gradient of the template image, we can get a simple linear approximation of  $J_t^{f_p}(v_i^*) \cdot v_i^*$  as follows

$$J_t^{f_p}(v_i^*) \cdot v_i^* \approx \nabla_p^T I \cdot J^{w_p} \cdot e_\pi \cdot v_i^* \quad (23)$$

A proof of ( 23 ) is give in 6.2.

Let  $J_t$  be the following matrix

$$J_t = -\frac{1}{2} (\nabla_p^T I + \nabla_p^T (I \circ t)) J^{w_p} e_\pi \quad (24)$$

By incorporating (24) into (21), we have

$$\|f_p(t \circ \text{Rexpp}(\omega_v^*))\|^2 \approx \|I \circ t(p) - I(p) + J_t v_i^* + O\| v_i^* \|^3 \|^2 \quad (25)$$

This cost function has a local or global minimum at  $v$

$$v_i^* = J_t^+(I \circ t(p) - I(p)) \quad (26)$$

where  $J_i^+$  is the pseudo-inverse of  $J_i$ . Hence, the intrinsic iterative update is

$$t = t \circ \text{Rexp}(\sum_{i=1}^8 v_i e_i) \quad (27)$$

## 4. Experimental Results and Analysis

### 4.1 Experimental Results

To validate the feasibility and efficiency of our algorithm, we compare our REXPP-ESM algorithm with VECTOR-GN algorithm and LEXP-GN algorithm. All the algorithms are implemented in matlab and tested in the computer with Intel PIV 2.4GHZ and 512 Memory. Since the 8 parameters in the projective warp have different units, we compute the RMS (root-mean-square) error of the corresponding points between the template and target image rather than the RMS of parameters. In addition, it should be emphasized that neither preliminary image filtering nor multi-scale pyramid implementations nor other robust techniques has been used for this evaluation.



Fig. 3. Input image and template. (a) Input image. (b) Template.

Experiment 1: We utilize the experiment data provided by Baker etc. in CMU and the same experiment setting to compare the three algorithms ([http://www.ri.cmu.edu/people/bakers\\_simon.html](http://www.ri.cmu.edu/people/bakers_simon.html)). We experimented with the image in Fig. 3(a) and manually selected a  $100 \times 100$  pixel template (see Fig. 3(b)) in the center of the image. We randomly perturbed the four corner points of the template 1000 times with additive white Gaussian noise of a certain standard variance  $\sigma$  from one pixel to ten pixels and fitted for the projective warp parameters that these perturbed points define (for each standard variance, we generated 100 randomly inputs). We say that an algorithm converged if the RMS error in the canonical point locations is less than 3.0 pixels after 15 iterations. We computed the percentage of times that each algorithm converged for each standard variance. The results are shown in Fig. 4(a) that shows when the perturbation to the canonical point locations is less than about 3.0 pixels, all the three algorithms converge almost always. With the increase of the  $\sigma$ , the frequency of convergence for LEXP-GN algorithm rapidly decreases. While  $\sigma = 10$ , the frequency of convergence for VECTOR-GN algorithm, LEXP-GN algorithm and our REXPP-ESM algorithm is 30%, 49% and 60% respectively. For 100 times experiments of  $\sigma = 6$ , all experiment test data are shown in Fig. 4(b). Our REXPP-ESM algorithm requires 8 iterations to coverage while LEXP-GN requires 9 iterations and VECTOR-GN requires 14 iterations.

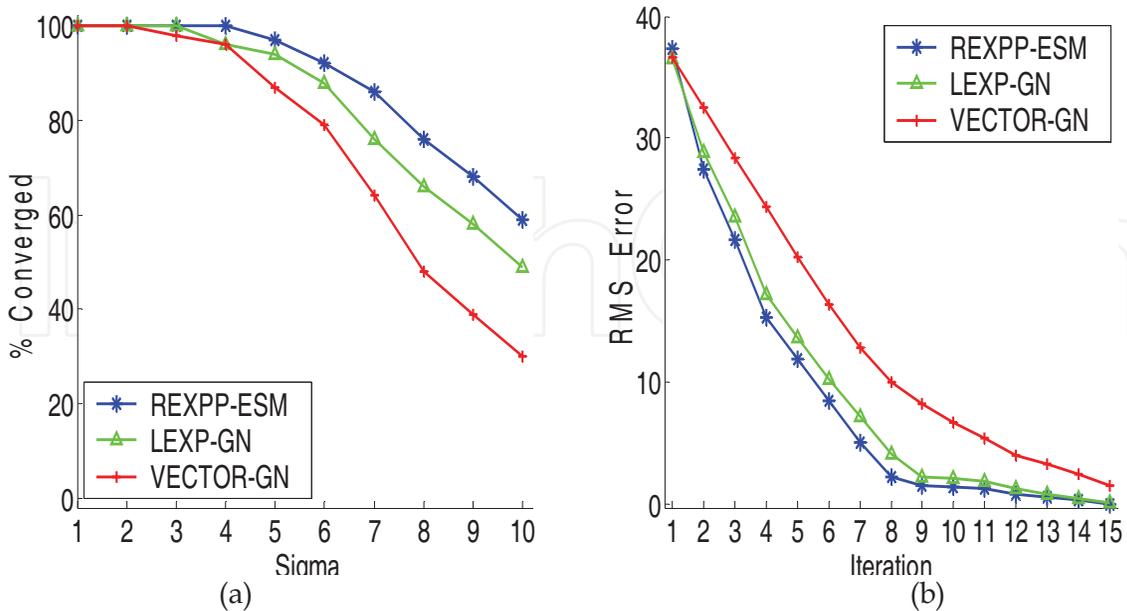


Fig. 4. Comparison between REXPP-ESM and other two methods. (a) Frequency of convergence. (b) Average convergence rates

Experiment 2: We show three experiment results for three typical image sequences using the three algorithms. The size of each frame in the first sequence is  $512 \times 512$  (<http://esm.gforge.inria.fr/ESMdownloads.html>) and the size of the tracked region is  $150 \times 150$ . The tracked region shows projective transformation. The VECTOR-GN algorithm, LEXP-GN algorithm and REXPP-ESM algorithm converge after 11, 7 and 6 iterations respectively. Fig. 5 shows some tracking results of them. The 161st and 187th frames have larger deformation. The VECTOR-GN algorithm cannot converge and the tracker slides off the tracking region. LEXP-GN algorithms cannot lock the tracking region at 161st frame. However, our REXPP-ESM algorithm can be implemented very well on all the frames. The second virtual house sequence includes one hundred still frames (<http://vasc.ri.cmu.edu/idb/html/motion/index.htm>). The size of each frame is  $512 \times 480$ . The tracked target is the window of the house and the size of template is  $52 \times 40$ . The tracked region shows larger projective transformation. The VECTOR-GN algorithm, LEXP-GN algorithm and REXPP-ESM algorithm converge after 7, 5 and 4 iterations respectively. Fig. 6 shows some tracking results of them. The sequences after 80th frame have larger deformation. The VECTOR-GN algorithm cannot converge and LEXP-GN algorithms cannot lock the tracking region at 91st frame either. However, our EXPP-ESM algorithm can be implemented very well on all the frames. The third car sequence contains 150 frames. The size of each frame is  $768 \times 576$ . The tracked target is the back of the running car and the tracked region shows larger enlargement warp. The VECTOR-GN algorithm, LEXP-GN algorithm and REXPP-ESM algorithm converge after 6, 4 and 3 iterations respectively. Fig. 7 show some tracking results of them. The sequences after 130th frame have larger deformation. The VECTOR-GN algorithm cannot converge and the tracker suffers from the drifts from the tracking region. LEXP-GN algorithm cannot lock the tracking region at 150th

frame. However, our REXPP-ESM algorithm can be implemented very well on all the frames. The table 1 summarizes the comparative performance of the three algorithms.

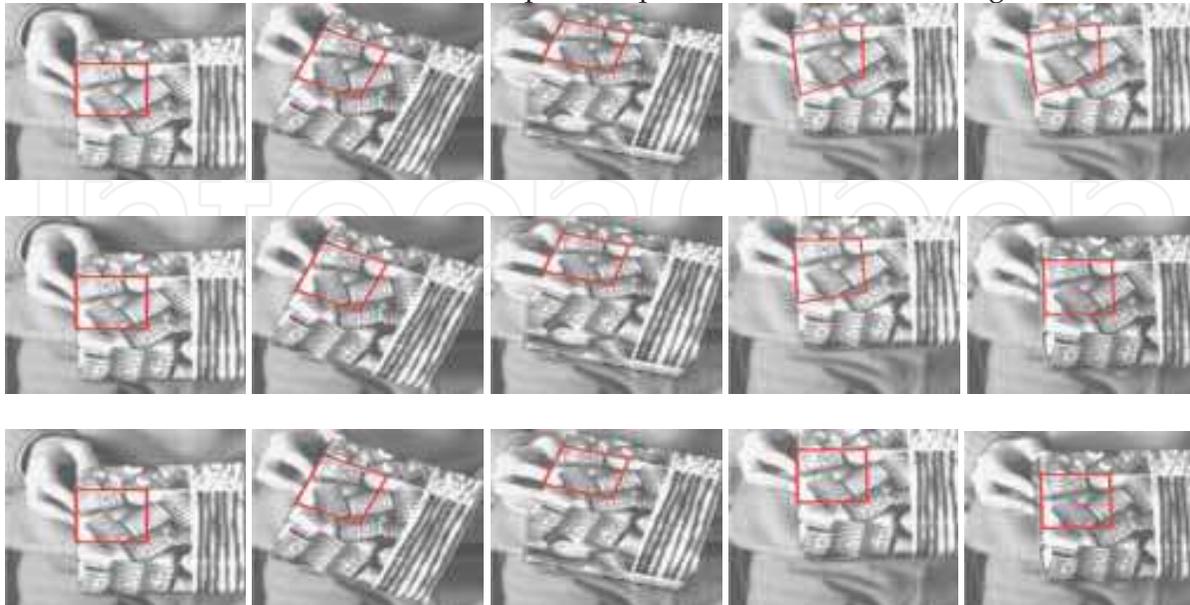


Fig. 5. Comparison of the VECTOR-GN (*first row*), LEXP-GN (*second row*) and REXPP-ESM (*third row*). The sequences contain 200 frames. From left to right in each column is No. 1, 50, 100, 161, 187 frame in wooden box sequences. See text for details.

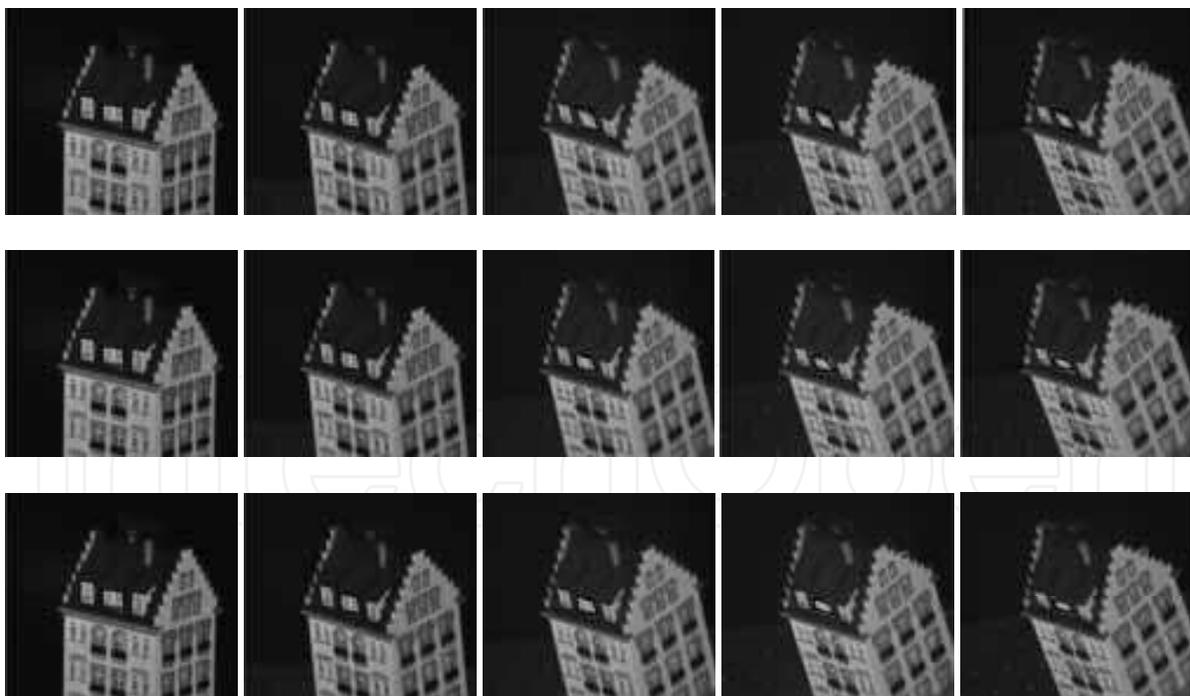


Fig. 6. Comparison of the VECTOR-GN (*first row*), LEXP-GN (*second row*) and REXPP-ESM (*third row*). The sequences contain 100 frames. From left to right in each column is No. 31, 60, 80, 90, 100 frame in virtual houses box sequences. See text for details.



Fig. 7. Comparison of the VECTOR-GN (*first row*), LEXP-GN (*second row*) and REXPP-ESM (*third row*). The sequences contain 150 frames. From left to right in each column is No. 1, 60, 130, 150 frame in virtual houses box sequences. See text for details.

#### 4.2 Analysis

From the results of the two experiments, we conclude that our algorithm utilizes the minimum geodesics and avoids computing Hessian matrix can make our REXPP-ESM algorithm is much superior to the VECTOR-GN algorithm and is slightly better than the LEXP-GN algorithm in the convergent frequency and convergence rate. Firstly, it is evident that the VECTOR-GN algorithm performs not well because it can not exploit the projective parameters intrinsic manifold structure. Secondly, it should be noted that when the distance between the two points are much close to the identity element, the REXPP-ESM performance is almost identical with the LEXP-GN because now the geodesic on  $SL(3,R)$  can be replaced by Lie exponential map. This can be obtained from the Fig. 4(a) of the first experiment where its data are synthesized when the perturbation is very small, namely, the projective warp is not remarkable. Although we adopt the second-order optimization, our REXPP-ESM doesn't perform much better than LEXP-GN. In real video sequences, although the deformation of the two continuous frames is usually not big, the drawbacks of the VECTOR-GN algorithm make it very easy to get local minimum while LEXP-GN and our REXPP-ESM perform well similarly. When the warps are bigger on some frames, our REXPP-ESM performs better than LEXP-GN, especially on the coverage rate. The reason for this is that our REXPP-ESM algorithm marches along the shortest distance during the optimization process than that of LEXP-GN. We confirm that some deep theory on Riemannian geometry should be introduced to explain it and leave it to work in the future.

sequence	algorithm	number of average iteration	Average coverage rate
wooden box	VECTOR-GN	11	89%
	LEXP-GN	7	96%
	REXPP-ESM	6	99.5%
virtual house	VECTOR-GN	7	90%
	LEXP-GN	5	95%
	REXPP-ESM	4	99.5%
car	VECTOR-GN	6	86%
	LEXP-GN	4	98%
	REXPP-ESM	3	100%

Table 1. Comparative performance of the three algorithms

## 5. Conclusion

We have presented the Riemannian exponential map on the noncompact Lie groups based on the minimum geodesics and constructed the registration and tracking algorithm without computing the Hessian matrix. The experimental results compared with the classical vector space algorithm and the Gauss-Newton optimization algorithm based on the Lie group exponential map show that the accuracy and the convergent rate demonstrate some evident improvements. It is emphasized that both Lie group exponential map and the Riemannian exponential map are based on the local linearization and are easy to diverge if the initial value and the iterative step size are not chosen improperly. Besides, we also investigate that the representative methods of the projective parameters have important effect on the experiment results and should be considered seriously.

## 6. Appendix

### 6.1 Proof of theorem 1

Assume  $\eta(t) \in gl(n, R)$ . Perturbing the shortest length curve  $g(t)$  with  $\exp(\varepsilon\eta(t))$  along  $\eta(t)$  leads to

$$g(t, \varepsilon) = \exp(\varepsilon\eta(t))g(t) = (\text{id} + \varepsilon\eta(t))g(t) + o(\varepsilon) \quad (28)$$

where the perturbation has the property that  $\eta(0) = 0$ ,  $\eta(1) = 0$  and that the boundary conditions remain satisfying at  $g(0, \varepsilon) = g$ ,  $g(1, \varepsilon) = h$ .

From (9), we can draw (29) and (30). On one side

$$d(g(t) + \varepsilon\eta(t)g(t))/dt = v(t)g(t) + \varepsilon\dot{\eta}(t)g(t) + \varepsilon\eta(t)v(t)g(t) \quad (29)$$

on the other side

$$d(g(t) + \varepsilon\eta(t)g(t))/dt = v(t)g(t) + \varepsilon v(t)\eta(t)g(t) + \varepsilon dv(t, \varepsilon)/d\varepsilon g(t) + o(\varepsilon) \quad (30)$$

Comparing ( 29 ) with ( 30 ) gives the equation

$$dv(t, \varepsilon)/d\varepsilon = \dot{\eta}(t) - [v(t), \eta(t)] \quad (31)$$

The minimal geodesic  $g(t)$  should satisfy to minimize the energy function  $\int_0^1 \|v(t)\|^2 dt$

$$\begin{aligned} & \frac{d}{d\varepsilon} \int_0^1 \|v(t) + \varepsilon \cdot dv(t, \varepsilon)/dt\|^2 dt \Big|_{\varepsilon=0} \\ &= 2 \int_0^1 \langle v(t), dv(t, \varepsilon)/dt \rangle dt \Big|_{\varepsilon=0} \\ &= 2 \int_0^1 \langle v(t), \dot{\eta}(t) \rangle dt + 2 \int_0^1 \langle [v(t), v^T(t)], \eta(t) \rangle dt \\ &= 2 \int_0^1 \langle -\dot{v}(t) + [v(t), v^T(t)], \eta(t) \rangle dt = 0 \end{aligned} \quad (32)$$

with using integration by parts and the boundary conditions on the perturbation  $\eta(0) = 0$  and  $\eta(1) = 0$ . Since this is zero over all perturbations  $\eta(t)$ , then we have

$$\begin{aligned} \dot{v}(t) &= v(t)v^T(t) - v^T(t)v(t) \\ &= [v(t), v^T(t)] \end{aligned} \quad (33)$$

End

## 6.2 Proof of (23)

Paying attention to the definition of the directional derivative and using compound derivative chain rule, we have

$$\begin{aligned} & J_t^{f_p}(v_t^*) \cdot v_t^* \\ &= \partial(I \circ t \circ \text{Rexpp}(v_t^*)(q))/\partial q^T \Big|_{q=\text{Rexpp}(-v_t^*) \circ \text{Rexpp}(v_t^*)(p)} \\ & \quad \cdot \partial w(s, p)/\partial s^T \Big|_{s=\pi(\text{Rexpp}(-v_t^*) \circ \text{Rexpp}(v_t^*))} \cdot \partial \pi(\text{Rexpp}(-v_t^*) \cdot \text{Rexpp}(v)) / \partial v^T \Big|_{v=v_t^*} \cdot v_t^* \\ &= \nabla_p^T(I \circ t^*) \cdot \partial w(s, p)/\partial s^T \Big|_{s=\pi(\text{Rexpp}(-v_t^*) \circ \text{Rexpp}(v_t^*))} \cdot \partial \pi(\text{Rexpp}(-v_t^*) \cdot \text{Rexpp}(v_t^* + t \cdot v_t^*)) / \partial v^T \Big|_{t=0} \\ &= (\nabla_p^T I + \varepsilon) \cdot J^{w_p} \cdot \partial \pi(\text{Rexpp}(tv_t^*)) / \partial t \Big|_{t=0} \\ &\approx \nabla_p^T I \cdot J^{w_p} \cdot e_\pi \cdot v_t^* \end{aligned} \quad (34)$$

where  $\varepsilon$  is a noise term of image to be discarded.

## 7. References

- Buenaposada, J. M. & Baumela, L. (2002). Real-time tracking and estimation of plane pose. In: *Proceedings of 16th International Conference on Pattern Recognition*. Canada: IEEE Computer Society Press, 2: 697–700.

- Baker, S & Matthews, M. (2004). Lucas-Kanade 20 years on: a unifying framework. *International Journal of Computer Vision*, 56(3):221–255.
- Berger, M. (2003). *A panoramic view of Riemannian geometry*. Berlin: Springer.
- Begelfor, E. & Werman, M. (2005). M. How to put probabilities on homographies. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 27(10):1666–1670.
- Benhimane S, & Malis E. (2007). Homography-based 2D visual tracking and servoing. *The International Journal of Robotics Research*, 26(7): 661–676.
- Eduardo, Jaime O A. (2007). Lie algebra approach for tracking and 3D motion estimation using monocular vision. *Image and Vision Computing*, 25(6): 907–921
- Gabay, D. (1982). Minimizing a differentiable function over a differential manifold. *Journal of optimization theory and applications*. 37(2): 117–217.
- Grenander, Ulf.; Miller, M. I. & Srivastava A. (1998). Hilbert–Schmidt lower bounds for estimators on matrix Lie groups for ATR. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 20(8): 790–802.
- Hager, G. D. & Belhumeur, P. N. (1998). Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 20(10):1025–1039
- Helgason, S. (1978). *Differential geometry, Lie Groups, and symmetric spaces*. Academic Press, 98–104.
- Mann, S & Picard, W. (1997). Video orbits of the projective group: a simple approach to featureless estimation of parameters. *IEEE Trans on Image Processing*, 6(9):1281–1295
- Michael Gleicher. (1997). Projective registration with difference decomposition. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and pattern Recognition*.
- Mahony, R. & Manton, J. H. (2002). The geometry of the Newton method on noncompact Lie Group. *Journal of Global Optimization*, 23(3-4): 309–327.
- Owren, B. & Welfert, B. (1996). The Newton iterations on Lie Groups. Technical Report Numerics.
- Smith, S. T. (1993). Geometric Optimization methods for adaptive filtering, *PhD Thesis*, Cambridge: University of Harvard.
- Seok, L. Minseok, C & Kim Hyungmin et al. Geometric direct search algorithms for image registration. *IEEE Trans. On Image Processing*, 2007, 16(9): 2215–2227.
- Trucco, E. & Plakas, K. (2006). Video tracking: a concise survey. *IEEE Trans on Oceanic Engineering*, 31(2): 520–529.
- Vercauteren, T.; Pennec, X., & Malis, E., et al. (2007). Insight into efficient image registration techniques and the demons algorithm. Proc. IPMI'07. Volume 4584 of LNCS. (July 2007) 495–506.
- Wang, H.C. (1969). Discrete nilpotent subgroups of Lie groups. *Journal of Differential Geometry* 3, pp. 481-492.
- Yaguang, Y. (1999). Optimization on Riemannian manifold. *Proceedings of the 38th Conference on Decision & Control*. USA: 1: 888–893.
- Yean, L. P. (2005). Geometric optimization for computer vision, *PhD Thesis*, Canberra: Australian National University.

Zefran, M & Kumar, V. & Croke, C. B. (1998). On the generation of smooth three-dimensional rigid body motions. *IEEE Transactions on Robotics and Automation*, 1998, 14(4):576 – 589

IntechOpen

IntechOpen



## **Pattern Recognition**

Edited by Peng-Yeng Yin

ISBN 978-953-307-014-8

Hard cover, 568 pages

**Publisher** InTech

**Published online** 01, October, 2009

**Published in print edition** October, 2009

For more than 40 years, pattern recognition approaches are continually improving and have been used in an increasing number of areas with great success. This book discloses recent advances and new ideas in approaches and applications for pattern recognition. The 30 chapters selected in this book cover the major topics in pattern recognition. These chapters propose state-of-the-art approaches and cutting-edge research results. I could not thank enough to the contributions of the authors. This book would not have been possible without their support.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Guangwei Li, Yunpeng Liu, Yin Jian and Zelin Shi (2009). Projective Registration with Manifold Optimization, Pattern Recognition, Peng-Yeng Yin (Ed.), ISBN: 978-953-307-014-8, InTech, Available from: <http://www.intechopen.com/books/pattern-recognition/projective-registration-with-manifold-optimization>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen