Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

A practical iterative two-view metric reconstruction with uncalibrated cameras

ZHOU Yong-jun†, KOU Xin-jian

(*Schoo1 of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*)

†E-mail: yjzhou@sjtu.edu.cn

Received Jan. 10, 2007; revision accepted May 8, 2007

Abstract: This paper presents a practical iterative algorithm for two-view metric reconstruction without any prior knowledge about the scene and motion in a nonsingular geometry configuration. The principal point is assumed to locate at the image center with zero skew and the same aspect ratio, and the interior parameters are fixed, so the self-calibration becomes focal-length calibration. Existing focal length calibration methods are direct solutions of a quadric composed of fundamental matrix, which are sensitive to noise. A quaternion-based linear iterative Least-Square Method is proposed in this paper, and one-dimensional searching for optimal focal length in a constrained region instead of solving optimization problems with inequality constraints is applied to simplify the computation complexity, then unique rotational matrix and translate vector are recovered. Experiments with simulation data and real images are given to verify the algorithm.

Key words: Close-range photogrammetry, Computer vision, General relative orientation, Unit quaternion, Metric reconstruction
doi:10.1631/jzus.2007.A1614 **Document code:** A **CLC number:** P234.1

INTRODUCTION

This paper proposes an iterative metric reconstruction method from two uncalibrated images with constant intrinsic parameters without any prior knowledge about the object space. Metric reconstruction is the highest level reconstruction which can be obtained without any scene information (Fusiello, 2000). To upgrade metric to Euclidean reconstruction, it only needs 3D similarity transformations with a uniform scalar factor, which could be realized with a scalar bar or sufficient control points.

To achieve metric reconstruction, focal length and relative pose of two views should be recovered firstly, which was also called general relative orientation in (Pan, 1999). If focal length is known, then it becomes traditional relative orientation in photogrammetry community, so focal length calibration is essential. Focal length can be recovered from two uncalibrated views with closed-form approaches (Pan, 1999; Ueshiba and Tomita, 2003) or singular value decomposition (SVD) methods (Hartley, 1992; Newsam *et al*., 1996; Sturm, 2001). Hartley's algorithm is a direct algorithm based on the SVD of the fundamental matrix. Newsam *et al*.(1996) followed Hartley's SVD method, but it was considerably simple in that it did not require further solutions of any higher order systems. Pan *et al*.(1995) presented a closed-form expression for the focal lengths as roots of quadric polynomials composed of fundamental matrix elements, along with an iterative Least-Square Method (LSM) for refinement, but it proved easy to fail. Recently, more attention was paid to stability and singularity researches on focal length self-calibration (Ueshiba and Tomita, 2004; Lao *et al*., 2004; Sturm *et al*., 2005; Kanatani *et al*., 2006).

Among the existing approaches, most are direct solutions from fundamental matrix, which has a serious drawback: the result is very sensitive to the quality of the matches, in particular, the focal lengths are often computed to be imaginary due to matching inaccuracy (Sturm *et al*., 2005). Pan *et al*.(1995) presented an iterative LSM for refinement, whereas it is apt to fail when the initial values are not so close.

This paper proposes a new iterative algorithm which has two major advantages: firstly it uses quaternion representing rotations for iterations and has large convergence region and fast convergence rate; secondly it uses 1D searching for optimal focal length in a constrained region instead of inequality constrained LSM optimization. So this algorithm can improve the reconstruction accuracy compared with other approaches. To simplify our algorithm, it is under the premise that the two views have the same intrinsic parameters and that the skew equals 0, the aspect ratio equals 1 in this paper, which is applicable with the development of camera manufacturing technology (Huang *et al*., 2004).

PRELIMINARIES

Assuming $\boldsymbol{M}(X, Y, Z)$ is a 3D point in the object space, and $\boldsymbol{m}(u, v)$ is its 2D image point. If we use $\tilde{\boldsymbol{M}}(X,Y,Z,1)$ and $\tilde{\boldsymbol{m}}(u,v,1)$ to denote their homogeneous coordinates, the relationship between $\boldsymbol{M}$ and its image projection $\boldsymbol{m}$ can be given as the pinhole model:

$$\tilde{\boldsymbol{m}} = \lambda \boldsymbol{K}[\boldsymbol{R} \,|\, -\boldsymbol{R}\boldsymbol{t}]\tilde{\boldsymbol{M}} = \lambda \boldsymbol{P}\tilde{\boldsymbol{M}}, \tag{1}$$

where λ is an arbitrary scale factor, $\boldsymbol{P}$ is the projective matrix, the extrinsic parameters ($\boldsymbol{R}$, $\boldsymbol{t}$) are the rotation and translation vectors which relate the object coordinate system to the camera coordinate system, and $\boldsymbol{K}$ is the intrinsic matrix:

$$\boldsymbol{K} = \begin{pmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2}$$

where (u_0, v_0) are the image coordinates of the principal point in pixel units, f_x and f_y are the scale factors in u and v axes respectively (for convenience, we call f the focal length), and s is the skewness of the two image axes.

Rotation matrix $\boldsymbol{R}$ is a 3×3 matrix with 3 degrees of freedom (DOFs), which could be parameterized as Euler angles, axis-angle, quaternion, Rodrigues and skew-matrix (Förstner, 1999). Rotation matrix represented by unit quaternion is given as:

$$\boldsymbol{R}(\boldsymbol{q}) = \begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2+q_3q_0) & 2(q_1q_3-q_2q_0) \\ 2(q_1q_2-q_3q_0) & q_0^2-q_1^2+q_2^2-q_3^2 & 2(q_2q_3+q_1q_0) \\ 2(q_1q_3+q_2q_0) & 2(q_2q_3-q_1q_0) & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix}, \tag{3}$$

where $\boldsymbol{q}=(q_0\ q_1\ q_2\ q_3)$ is a unit quaternion satisfied with:

$$\|\boldsymbol{q}\| = \sqrt{q_0^2+q_1^2+q_2^2+q_3^2} = 1.$$

Let $\boldsymbol{m}(u, v)$ and $\boldsymbol{m}'(u', v')$ be 2D projections of $\boldsymbol{M}$ on left and right images respectively, the point correspondences should be satisfied with epipolar geometry constraints (Fig.1):

$$(\tilde{\boldsymbol{m}}')^{\mathrm{T}}\boldsymbol{F}\tilde{\boldsymbol{m}} = 0, \tag{4}$$

where $\boldsymbol{F}$ is a 3×3 matrix called fundamental matrix with rank($\boldsymbol{F}$)=2 and 7 DOF. If the two images captured by the same camera without zooming and focusing, i.e., the two images have the same intrinsic parameters, and the object coordinate system is aligned to the left camera coordinate system, with the notation ≅ implying equality up to scale, the projection relationship in Eq.(1) can be given as:

$$\begin{cases} \tilde{\boldsymbol{m}} \cong \boldsymbol{K}[\boldsymbol{I} \,|\, \boldsymbol{0}]\tilde{\boldsymbol{M}}, \\ \tilde{\boldsymbol{m}}' \cong \boldsymbol{K}[\boldsymbol{R} \,|\, -\boldsymbol{R}\boldsymbol{t}]\tilde{\boldsymbol{M}}. \end{cases} \tag{5}$$

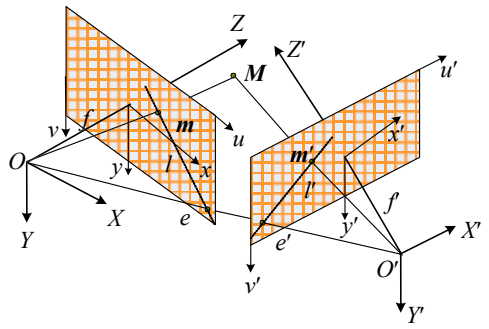

Fig.1 Two-view geometry with epipolar constraints

And $\boldsymbol{F}$ is the combination of extrinsic and intrinsic parameters, given by:

$$\boldsymbol{F} \cong \boldsymbol{K}^{-\mathrm{T}}\boldsymbol{B}\boldsymbol{R}\boldsymbol{K}^{-1}, \tag{6}$$

where $\boldsymbol{R}$ is the rotation matrix for the right ray and corresponding left ray, $\boldsymbol{B}$ is an anti-symmetric matrix defined by the translations of the two projective centers, also known as baseline (b_x, b_y, b_z):

$$\boldsymbol{B}=\begin{pmatrix}0 & -b_z & b_y\\ b_z & 0 & -b_x\\ -b_y & b_x & 0\end{pmatrix}. \tag{7}$$

TWO-VIEW FOCAL LENGTH CALIBRATION BY FUNDAMENTAL MATRIX FACTORIZATION

Form Eq.(4) and Eq.(6), the epipolar constraints may be expressed as:

$$\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{K}^{-\mathrm{T}}\boldsymbol{B}\boldsymbol{R}\boldsymbol{K}^{-1}\tilde{\boldsymbol{m}}'\cong 0. \tag{8}$$

Under the assumption that the principal point (u_0, v_0) is located at the image center, f_x=f_y=f and s=0, decompose $\boldsymbol{K}$ into the product of two matrices:

$$\begin{aligned}\boldsymbol{K}&=\begin{pmatrix}f & 0 & u_0\\ 0 & f & v_0\\ 0 & 0 & 1\end{pmatrix}\\ &=\begin{pmatrix}f_0 & 0 & u_0\\ 0 & f_0 & v_0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}c & 0 & 0\\ 0 & c & 0\\ 0 & 0 & 1\end{pmatrix}=c\boldsymbol{K}_0\bar{\boldsymbol{K}},\end{aligned} \tag{9}$$

where f_0 is a scalar parameter predetermined by the resolution of captured images, $\boldsymbol{K}_0$ is used to eliminate the matrix ill-conditions and instable computations, which is commonly called normalization, as the focal length is 10^3 or 10^4 times larger relative to other parameters, and c is a variable defined as c=f/f_0. If c is determined and f can be solved sequentially, letting

$$\bar{\boldsymbol{m}}=\boldsymbol{K}_0\tilde{\boldsymbol{m}},\quad \bar{\boldsymbol{m}}'=\boldsymbol{K}_0\tilde{\boldsymbol{m}}', \tag{10}$$

Eq.(8) can be given as:

$$\bar{\boldsymbol{m}}'^{\mathrm{T}}\bar{\boldsymbol{K}}^{-\mathrm{T}}\boldsymbol{B}\boldsymbol{R}\bar{\boldsymbol{K}}^{-1}\bar{\boldsymbol{m}}\cong 0, \tag{11}$$

where

$$\bar{\boldsymbol{K}}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1/c\end{pmatrix},\quad \bar{\boldsymbol{K}}^{-1}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & c\end{pmatrix}. \tag{12}$$

By normalizing $\boldsymbol{K}_0$ with image points, the fundamental matrix $\boldsymbol{F}$ converts to the semi-calibrated fundamental matrix $\bar{\boldsymbol{F}}$, Eq.(6) can be written as:

$$\bar{\boldsymbol{F}}\cong\bar{\boldsymbol{K}}^{-\mathrm{T}}\boldsymbol{B}\boldsymbol{R}\bar{\boldsymbol{K}}^{-1}. \tag{13}$$

Let the singular value decomposition of $\bar{\boldsymbol{F}}$ be given by:

$$\bar{\boldsymbol{F}}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}, \tag{14}$$

here $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthonormal matrices and $\boldsymbol{\Sigma}$=diag(σ_1 σ_2 σ_3) is a diagonal matrix (for rank deficiency of fundamental matrix, at least one of the three singular values are nearly zero). For convenience, assuming the last singular value σ_3=0, we can get focal length calibration equations as follows (Sturm *et al*., 2005):

$$\begin{aligned}&c^2[\sigma_1U_{31}U_{32}(1-V_{31}^2)+\sigma_2V_{31}V_{32}(1-U_{32}^2)]\\ &\quad+U_{32}V_{31}(\sigma_1U_{31}V_{31}+\sigma_1U_{32}V_{32})=0,\end{aligned} \tag{15}$$

$$\begin{aligned}&c^2[\sigma_1V_{31}V_{32}(1-U_{31}^2)+\sigma_2U_{31}U_{32}(1-V_{32}^2)]\\ &\quad+U_{31}V_{32}(\sigma_1U_{31}V_{31}+\sigma_2U_{32}V_{32})=0,\end{aligned} \tag{16}$$

$$\begin{aligned}&c^4[\sigma_1^2(1-U_{31}^2)(1-V_{31}^2)-\sigma_2^2(1-U_{32}^2)(1-V_{32}^2)]\\ &+c^2[\sigma_1^2(U_{31}^2+V_{31}^2-2U_{31}^2V_{31}^2)-\sigma_2^2(U_{32}^2+V_{32}^2-2U_{32}^2V_{32}^2)]\\ &+(\sigma_1^2U_{31}^2V_{31}^2-\sigma_2^2U_{32}^2V_{32}^2)=0.\end{aligned} \tag{17}$$

The above three equations are algebraically dependent. Two of these have trivial solutions. Without critical configuration, we can obtain the right solutions from quadratic Eq.(17).

QUATERNION-BASED ITERATIVE LSM FOR GENERAL RELATIVE ORIENTATION

Least-Square Method analysis

Sturm Algorithm first estimates the semi-calibrated fundamental matrix and then recovers focal length. It is a linear direct method. To recover the parameters explicitly at a high accuracy, LSM is often performed for refinement when more than six point correspondences are available. The cost function of least-square criteria is:

$$\sum_{i=1}^{m}\bar{\boldsymbol{m}}_i'^{\mathrm{T}}\bar{\boldsymbol{K}}^{-\mathrm{T}}\boldsymbol{B}\boldsymbol{R}\bar{\boldsymbol{K}}^{-1}\bar{\boldsymbol{m}}_i=\min. \tag{18}$$

Eq.(18) can be expressed in a standard optimization form:

$$\min_{\boldsymbol{\Theta}}\left\{f(\boldsymbol{\Theta})=\frac{1}{2}\sum_{j=1}^{m}\Phi_j^2(\boldsymbol{\Theta}):\boldsymbol{\Theta}\in\mathbb{R}^n\right\},\qquad(19)$$

where m is the total number of point correspondences, n is the number of unknown parameters, $\Phi(\boldsymbol{\Theta})$ is twice continuously differentiable function from $\mathbb{R}^n$. In this paper, $\boldsymbol{\Theta}$ are the unknowns including focal length, baseline and rotation angles. If we set b_x=1, then the independent unknowns are c, b_y, b_z, φ, ω, κ. When appropriate initial values of unknowns are given, the linear iterative method or nonlinear Gauss-Newton method and Levenberg-Marquardt method are often used for solving the least-square problem.

Denote $\nabla f(\boldsymbol{\Theta})$ and $\nabla f^2(\boldsymbol{\Theta})$ as the gradient and the Hessian of function $f(\boldsymbol{\Theta})$, respectively. From Eq.(19), it is easy to verify that:

$$\begin{cases}\nabla f(\boldsymbol{\Theta})=\sum_{j=1}^{m}\Phi_j(\boldsymbol{\Theta})\nabla\Phi_j(\boldsymbol{\Theta})=J^{\mathrm{T}}(\boldsymbol{\Theta})\Phi(\boldsymbol{\Theta}),\\ \nabla^2 f(\boldsymbol{\Theta})=\sum_{j=1}^{m}\nabla\Phi_j(\boldsymbol{\Theta})\nabla\Phi_j^{\mathrm{T}}(\boldsymbol{\Theta})+\sum_{j=1}^{m}\Phi_j(\boldsymbol{\Theta})\nabla^2\Phi_j^{\mathrm{T}}(\boldsymbol{\Theta})\\ \qquad\qquad=J^{\mathrm{T}}(\boldsymbol{\Theta})J(\boldsymbol{\Theta})+\sum_{j=1}^{m}\Phi_j(\boldsymbol{\Theta})\nabla^2\Phi_j^{\mathrm{T}}(\boldsymbol{\Theta}),\end{cases}\qquad(20)$$

where $\Phi(\boldsymbol{\Theta})=(\Phi_1(\boldsymbol{\Theta}),\ \ldots,\ \Phi_m(\boldsymbol{\Theta}))^{\mathrm{T}}$, and $J(\boldsymbol{\Theta})=(\nabla\Phi_1(\boldsymbol{\Theta}),\ \ldots,\ \nabla\Phi_m(\boldsymbol{\Theta}))^{\mathrm{T}}$ is the $m\times n$ Jacobi matrix of $\Phi(\boldsymbol{\Theta})$. The Hessian of such a least square error function is a combination of the first and second order terms:

$$\nabla^2 f(\boldsymbol{\Theta})=J^{\mathrm{T}}(\boldsymbol{\Theta})J(\boldsymbol{\Theta})+Q(\boldsymbol{\Theta}),$$

with $Q(\boldsymbol{\Theta})=\sum_{j=1}^{m}\Phi_j(\boldsymbol{\Theta})\nabla^2\Phi_j^{\mathrm{T}}(\boldsymbol{\Theta})$. In practice, the approximation of Hessian is used as: $H(\boldsymbol{\Theta})=J^{\mathrm{T}}(\boldsymbol{\Theta})J(\boldsymbol{\Theta})$, which is based on the premise that the first-order term eventually dominates the second-order term. Levenberg-Marquardt method uses the Gauss-Newton approximation of the Hessian with the kth search direction being defined as the solution of

$$[J^{\mathrm{T}}(\boldsymbol{\Theta}^k)J(\boldsymbol{\Theta}^k)+\mu_k\boldsymbol{I}]d^k=-J^{\mathrm{T}}(\boldsymbol{\Theta}^k)\Phi(\boldsymbol{\Theta}^k),$$

where, μ_k is a non-negative scalar, d^k is the solution of the constrained optimization sub-problem.

Whereas in our cases, the estimated focal length by Sturm Algorithm is about 20% offset from its true value, which belongs to the so-called large-residual problems, and the second-order term in Eq.(20) could not be neglected, so the commonly-used Levenberg-Marquardt method does not work well any more. Then an improved LSM is presented in this paper, which is improved in two aspects. Firstly we use quaternion to represent rotation matrix instead of Euler angles, and in iterative process, we do not estimate quaternion directly but estimate another three parameters. Experiment shows that it has a large convergence region and fast convergence rate. Secondly we use 1D searching for optimal focal length in a constrained region to avoid incorrect convergence and also simplify the optimization problem with inequality constraints.

Quaternion-based iterative Least-Square Method

Quaternions were used to represent vector rotation by Hamilton in 1843, and were introduced to photogrammetry and computer vision society by Horn (1987; 1991) and Zeng (1990). According to the definition of quaternion, Eq.(11) can be converted into quaternion form. Let $\boldsymbol{l}$=(0 x_l y_l c), $\boldsymbol{r}$=(0 x_r y_r c) represent the normalized coordinates for the left and right matches in quaternion form, and $\boldsymbol{b}$=(0 b_x b_y b_z), then Eq.(11) can be written as:

$$\boldsymbol{b}\cdot(\boldsymbol{l}\otimes\boldsymbol{r}')=0,\qquad(21)$$

where, $\cdot$ denotes quaternion dot product, $\otimes$ denotes quaternion cross product (refer to the Appendix), $\boldsymbol{r}'$=(0 x_r' y_r' z_r') is the right ray denoted by quaternion vector in left camera coordinate system. It can be rotated from right camera coordinate system by using quaternion as:

$$\boldsymbol{r}'=\boldsymbol{q}\otimes\boldsymbol{r}\otimes\boldsymbol{q}^*,\qquad(22)$$

For convenience, Eq.(19) can be expressed in function form:

$$\boldsymbol{F}(\boldsymbol{b},\boldsymbol{q},c)=\boldsymbol{b}\cdot(\boldsymbol{l}\otimes\boldsymbol{r}')=\boldsymbol{b}\cdot(\boldsymbol{l}\otimes\boldsymbol{q}\otimes\boldsymbol{r}\otimes\boldsymbol{q}^*).\qquad(23)$$

When more than 6 point correspondences are

available, the unknowns are the least-square solutions, and the least-square criteria can be expressed as:

$$\Phi(\boldsymbol{\Theta})=\sum_{i=1}^{m}F_i^2(c,\boldsymbol{b},\boldsymbol{q})=\min. \tag{24}$$

From Eq.(21), we can get the derivative of the focal length:

$$\frac{\partial F}{\partial c}=-x_r'z_lE_{31}-y_r'z_lE_{32}+(2cE_{33}-x_lE_{13}-y_lE_{23})z_r', \tag{25}$$

where, $\boldsymbol{E}=\boldsymbol{BR}$ is called Essential Matrix and E_{ij} denotes the ith row and jth column element of $\boldsymbol{E}$. Then we can obtain the derivative matrix for baseline neglecting the scalar part of quaternion:

$$\begin{aligned}\frac{\partial F}{\partial \boldsymbol{b}}&=(\boldsymbol{l}_v\times\boldsymbol{r}_v')=\boldsymbol{S}_{l_v}c\\&=\begin{pmatrix}0&z_l&-y_l\\-z_l&0&x_l\\y_l&-x_l&0\end{pmatrix}\begin{pmatrix}x_r'\\y_r'\\z_r'\end{pmatrix}=\begin{pmatrix}z_ly_r'-y_lz_r'\\x_lz_r'-z_lx_r'\\y_lx_r'-x_ly_r'\end{pmatrix},\end{aligned} \tag{26}$$

where $\boldsymbol{l}_v$, $\boldsymbol{r}_v'$ denote the vector parts of the corresponding quaternions, and $\boldsymbol{S}_{l_v}$ denotes the anti-symmetric matrix of $\boldsymbol{l}_v$.

And the derivative matrix for quaternion is:

$$\frac{\partial F(\boldsymbol{b},\boldsymbol{q})}{\partial \boldsymbol{q}}=\frac{\partial F(\boldsymbol{b},\boldsymbol{q})}{\partial \boldsymbol{r}'}\frac{\partial \boldsymbol{r}'}{\partial \boldsymbol{q}}, \tag{27}$$

where,

$$\frac{\partial \boldsymbol{r}'}{\partial \boldsymbol{q}}=\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{q}}\otimes\boldsymbol{r}'\otimes\boldsymbol{q}^*-\boldsymbol{q}\otimes\boldsymbol{r}'\otimes\frac{\partial \boldsymbol{q}^*}{\partial \boldsymbol{q}}. \tag{28}$$

Letting $\frac{\partial \boldsymbol{r}'}{\partial \boldsymbol{q}}=\dot{\boldsymbol{r}}'$, $\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{q}}=\dot{\boldsymbol{q}}$, $\frac{\partial \boldsymbol{q}^*}{\partial \boldsymbol{q}}=\dot{\boldsymbol{q}}^*$, Eq.(28) can be simplified as:

$$\dot{\boldsymbol{r}}'=\dot{\boldsymbol{q}}\otimes\boldsymbol{r}\otimes\boldsymbol{q}^*+\boldsymbol{q}\otimes\boldsymbol{r}\otimes\dot{\boldsymbol{q}}^*. \tag{29}$$

From Eq.(22), we can get:

$$\boldsymbol{r}=\boldsymbol{q}^*\otimes\boldsymbol{r}'\otimes\boldsymbol{q}. \tag{30}$$

Combining Eq.(29) and Eq.(30) yields:

$$\dot{\boldsymbol{r}}'=\dot{\boldsymbol{q}}\otimes\boldsymbol{q}^*\otimes\boldsymbol{r}'+\boldsymbol{r}'\otimes\boldsymbol{q}\otimes\dot{\boldsymbol{q}}^*. \tag{31}$$

According to the laws for quaternion calculus, $\dot{\boldsymbol{q}}\otimes\boldsymbol{q}^*$ and $\boldsymbol{q}\otimes\dot{\boldsymbol{q}}^*$ have the equal but negative sign scalar and vector parts, we can denote them as:

$$\dot{\boldsymbol{q}}\otimes\boldsymbol{q}^*=(s,\ \boldsymbol{v}),\quad \boldsymbol{q}\otimes\dot{\boldsymbol{q}}^*=(-s,\ -\boldsymbol{v}). \tag{32}$$

Combining Eq.(31) and Eq.(32), and letting $\boldsymbol{w}=2\boldsymbol{v}=(w_1\ w_2\ w_3)^{\mathrm{T}}$, we get:

$$\dot{\boldsymbol{r}}'=2\boldsymbol{v}\times\boldsymbol{r}'=\boldsymbol{w}\times\boldsymbol{r}'=-\begin{pmatrix}0&-z_r'&y_r'\\z_r'&0&-x_r'\\-y_r'&x_r'&0\end{pmatrix}\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}, \tag{33}$$

where,

$$\begin{aligned}\begin{pmatrix}s\\\boldsymbol{v}\end{pmatrix}&=\dot{\boldsymbol{q}}\otimes\boldsymbol{q}^*=\bar{\boldsymbol{Q}}\dot{\boldsymbol{q}}\\&=\begin{pmatrix}q_0&q_1&q_2&q_3\\-q_1&q_0&-q_3&q_2\\-q_2&q_3&q_0&-q_1\\-q_3&-q_2&q_1&q_0\end{pmatrix}\begin{pmatrix}\Delta q_0\\\Delta q_1\\\Delta q_2\\\Delta q_3\end{pmatrix}.\end{aligned} \tag{34}$$

Combine Eq.(32) and Eq.(33), and let

$$\boldsymbol{C}=2\begin{pmatrix}q_0&-q_3&q_2\\q_3&q_0&-q_1\\-q_2&q_1&q_0\end{pmatrix}. \tag{35}$$

As $\det(\boldsymbol{C})=1$, so $\boldsymbol{C}$ has inverse matrix as follows:

$$\begin{pmatrix}\Delta q_1\\\Delta q_2\\\Delta q_3\end{pmatrix}=\boldsymbol{C}^{-1}\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}, \tag{36}$$

then the final derivative matrix with quaternion vector part is:

$$\begin{aligned}\frac{\partial F(\boldsymbol{b},\boldsymbol{q})}{\partial \boldsymbol{q}}&=\frac{\partial F(\boldsymbol{b},\boldsymbol{q})}{\partial \boldsymbol{r}'}\frac{\partial \boldsymbol{r}'}{\partial \boldsymbol{q}}\\&=(\boldsymbol{b}_v\times\boldsymbol{l}_v)(\boldsymbol{w}\times\boldsymbol{r}')=-(\boldsymbol{b}_v\times\boldsymbol{l}_v\times\boldsymbol{w})\times\boldsymbol{r}'\\&=\begin{pmatrix}(b_xy_l-b_yx_l)y_r'+(b_xz_l-b_zx_l)z_r'\\(b_xy_l-b_yx_l)x_r'+(b_zy_l-b_yz_l)z_r'\\(b_yz_l-b_zy_l)y_r'+(b_xz_l-b_zx_l)x_r'\end{pmatrix}\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}.\end{aligned} \tag{37}$$

From Eqs.(25), (26) and (36), we can get the ith ($i\leq m$) row of Jacobi matrix for $\Phi(\boldsymbol{\Theta})$ [$\boldsymbol{\Theta}=(c, b_y, b_z, w_1, w_2, w_3)$]

$$\begin{cases} J(i,1)=-x_r'z_lE_{31}-y_r'z_lE_{32}+(2cE_{33}-x_lE_{13}-y_lE_{23})z_r', \\ J(i,2)=x_lz_r'-z_lx_r', \\ J(i,3)=y_lx_r'-x_ly_r', \\ J(i,4)=(b_xy_l-b_yx_l)y_r'+(b_xz_l-b_zx_l)z_r', \\ J(i,5)=(b_xy_l-b_yx_l)x_r'+(b_zy_l-b_yz_l)z_r', \\ J(i,6)=(b_yz_l-b_zy_l)y_r'+(b_xz_l-b_zx_l)x_r'. \end{cases} \tag{38}$$

If the focal length is known, then the quaternion-based linear LSM iterative algorithm can be formulated as follows:

(1) Find point correspondences and normalize it with given interior parameters;

(2) Let b_x=1, b_y=0, b_z=0, q_1=q_2=q_3=0, q_0=1 be the initial values. As the focal length is known, then the unknowns $\boldsymbol{\Theta}$=(c, b_y, b_z, w_1, w_2, w_3). Use Eq.(26) and Eq.(38) to get the Jacobi matrix;

(3) Use $\delta\boldsymbol{\Theta}=(J^{\mathrm{T}}J)J^{\mathrm{T}}\boldsymbol{L}$ to get the kth iteration corrections, where $\boldsymbol{L}$ is an $n\times1$ residual vector with $L_i=-\bar{\boldsymbol{m}}'\boldsymbol{B}\boldsymbol{R}\bar{\boldsymbol{m}}$;

(4) Use Eq.(36) to get the quaternion corrections;

(5) Update the baseline and quaternion, continue to the next iteration till convergence.

Constrained focal length search LSM

It seems that when the initial values are known, the general relative orientation problem can be easily solved like the traditional relative orientation. Pan *et al*.(1995) used traditional LSM for refinement and also gave two examples, but experiments indicate that they do not always work well, and are apt to converge to wrong values if the initial values are far from the true one. By further analyzing, the authors believe that there are two reasons inducing the wrong convergence, one is that f=0 is a global minimum (Fig.2, which is a simulation example with focal length of 1), the other is that the focal length ratio c is in the dominant direction and easily diverts to correct values in iterative linear LSM without any constraints. To avoid wrong convergence, two improvements are made on traditional LSM, one is using quaternion representing rotation matrix instead of Euler angles, the other is using 1D searching for optimal focal length in a constrained region.

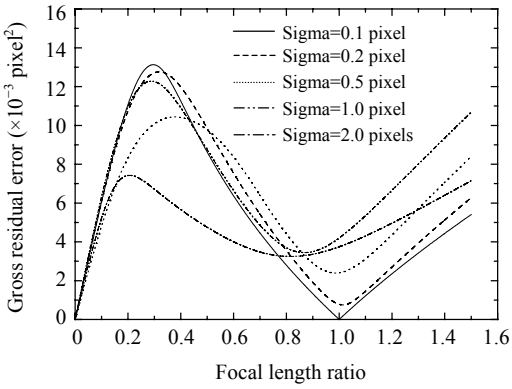

Fig.2 Cost function value with varied focal lengths

The focal length constraints should be given at an appropriate region. According to the results of numerous simulation experiments and theoretical analysis, it is practical to confine the focal length within the interval $0.6c_0$~$1.4c_0$, that is to say, the estimated focal length is within 40% of the initial value. Then the unconstrained LSM becomes constrained LSM as:

$$\begin{cases} \Phi(\boldsymbol{\Theta})=\sum_{i=1}^{n}F_i^2(c,\boldsymbol{b},\boldsymbol{q})=\min, \\ 0.6c_0\leq c\leq 1.4c_0, \end{cases} \tag{39}$$

where, c_0 is the initial value of the focal length. Eq.(39) is an inequality constrained optimization problem, and penalty functions are widely used to solve it, which is difficult for most photogrammetry engineers. So we use 1D searching for optimal focal length. This algorithm is based on the good convergence capability of quaternion-based LSM. With this method, the focal length is given in advance, and the above-mentioned quaternion-based LSM is performed to calculate the value of cost function until the minimum value is found in the given region. We use 1D golden section searching algorithm for finding focal length in the constrained region, which is an excellent method for finding a minimum along a line. Details of the algorithm are elaborated on by Yuan and Sun (2001).

Outline for two-view metric reconstruction

The procedure of the iterative two-view metric reconstruction can be given as follows:

(1) Use Sturm Algorithm to obtain the initial value of focal length;

(2) Use 1D golden section searching algorithm to find the optimal focal length and the rotational matrix $\boldsymbol{R}$, and translate vectors $\boldsymbol{t}$ simultaneously;

(3) From Eq.(5), we can get:

$$\boldsymbol{P} \cong \boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}], \quad \boldsymbol{P}' \cong \boldsymbol{K}[\boldsymbol{R} \mid -\boldsymbol{R}\boldsymbol{t}]. \tag{40}$$

Then the metric reconstruction $\boldsymbol{X}_m$ can be given as:

$$\begin{pmatrix} \tilde{\boldsymbol{m}} & \boldsymbol{0} & \boldsymbol{P} \\ \boldsymbol{0} & \tilde{\boldsymbol{m}}' & \boldsymbol{P}' \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \boldsymbol{X}_m \end{pmatrix} = 0, \tag{41}$$

where λ_1 and λ_2 are scale factors.

EXPERIMENTS

Simulation experiments

Experiments were used to simulate images captured by hand-held digital camera in general geometry configuration. Scenes with twenty 3D points are randomly generated within a 10 m×10 m×3 m region in front of the left image about 20 m ahead, projecting onto two images respectively and producing 20 point correspondences. The image size is 1280×1000 pixels, the ideal principal point locates at the image center, and the ideal foal length is 1500 pixels. The relative rotation matrix is based on φ, ω, κ Euler sequences. Experiment assumes that the baseline ratios are (1.0 0.5 0.2) and the Euler angles are (−10° 5° 2°). Experiments are used to verify: (1) the focal length accuracy between iterative algorithm and direct algorithm under different image noises, (2) the relationship between the principal point error and the focal length error, and (3) the relationship of the metric reconstruction error and the focal length error.

Fig.3 is the average focal length relative error in 500 times simulation experiments under different normal distribution noise levels. The horizontal axis is the standard deviation of the additive noise. Fig.3 indicates that the focal length is very sensitive to image noise, and that the quaternion-based iterative LSM (QILSM) can improve the estimated focal length accuracy over that of the direct methods.

Fig.4 indicates the focal length relative error when the principal point varies from ±50 pixels without image noises, and shows that the principal point offset has little influence on focal length estimation. Compared with image noise, it is reasonable to assume that the principal point locates at the image center in some application fields.

We used all the 20 points as control points for 3D similarity transformation (also called absolute orientation). The average error is defined as $\pm\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}/(\sqrt{3}n)$, where Δx, Δy, Δz are the residual errors after transformation. The results in Fig.5 indicate that the largest reconstruction error is

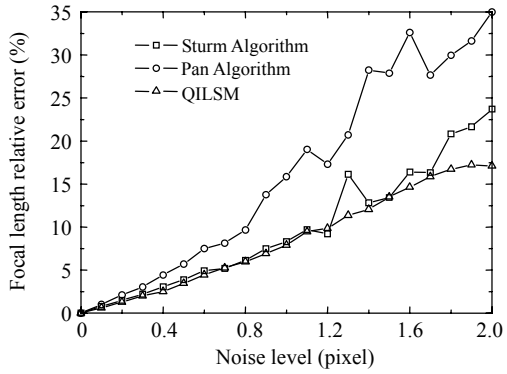

Fig.3 Focal length error under different noise levels

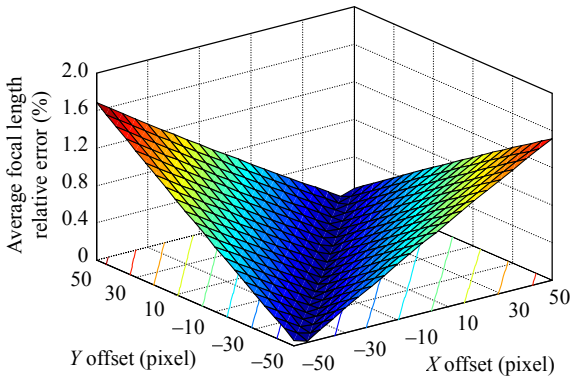

Fig.4 Focal length error with principal point offset

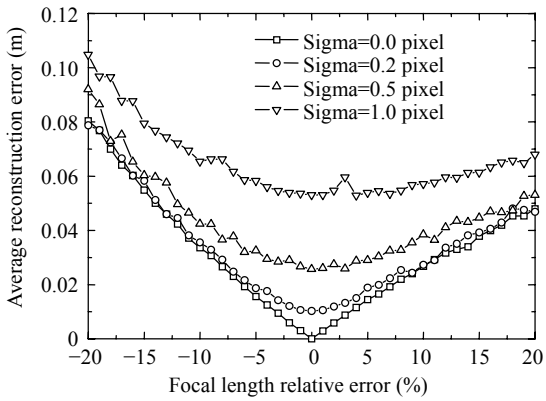

Fig.5 Reconstruction error with our practical algorithm

within ±0.1 m even if the focal length relative error is ±20%, and that the average error is ±3 cm when the focal length is within 10%.

Real images

The real data is the example downloaded from http://dprg.geomatics.ucalgary.ca/bundle.html, and was used for verifying the software BUNDLE (Version 1.5) developed by Digital Photogrammetry Research Group with University of Calgary, Canada. More than 20 images were captured by Kodak DCS260 camera with the same intrinsic parameters. The image resolution was 1536×1024 pixels. We select three of the overlapped images from the image sequences to test the algorithm. The points distribution could be seen as Fig.6. The camera intrinsic parameters were accurately calculated by bundle adjustment with consideration of the lens distortions, the principal point was (764.821, 509.368), and the focal length was 1703.489. In our algorithm we assume that the image center (768, 512) was the principal point and that f_0=2000. The results are shown in Table 1, from which we can see that the relative focal length error was about 10%, and that the reconstruction errors of the control points and check points were within 0.1 m. The image pairs of 2-3 and 1-3 have larger errors than pair 1-2. By further analysis, we found that it is the overlap and points distribution that induce the large errors. From Fig.6 we can see that the overlap area of 1-2 is much more than those of 1-3 and 2-3, so the overlap and points distribution can influence the reconstruction greatly.

Table 1 Reconstruction error of image pairs

Image pairs	Focal length (pixels)	Average error of control points (m)	Average error of check points (m)
1-2	1612.6	±0.0268	±0.0345
1-3	1526.3	±0.0638	±0.0726
2-3	1566.8	±0.0610	±0.0582

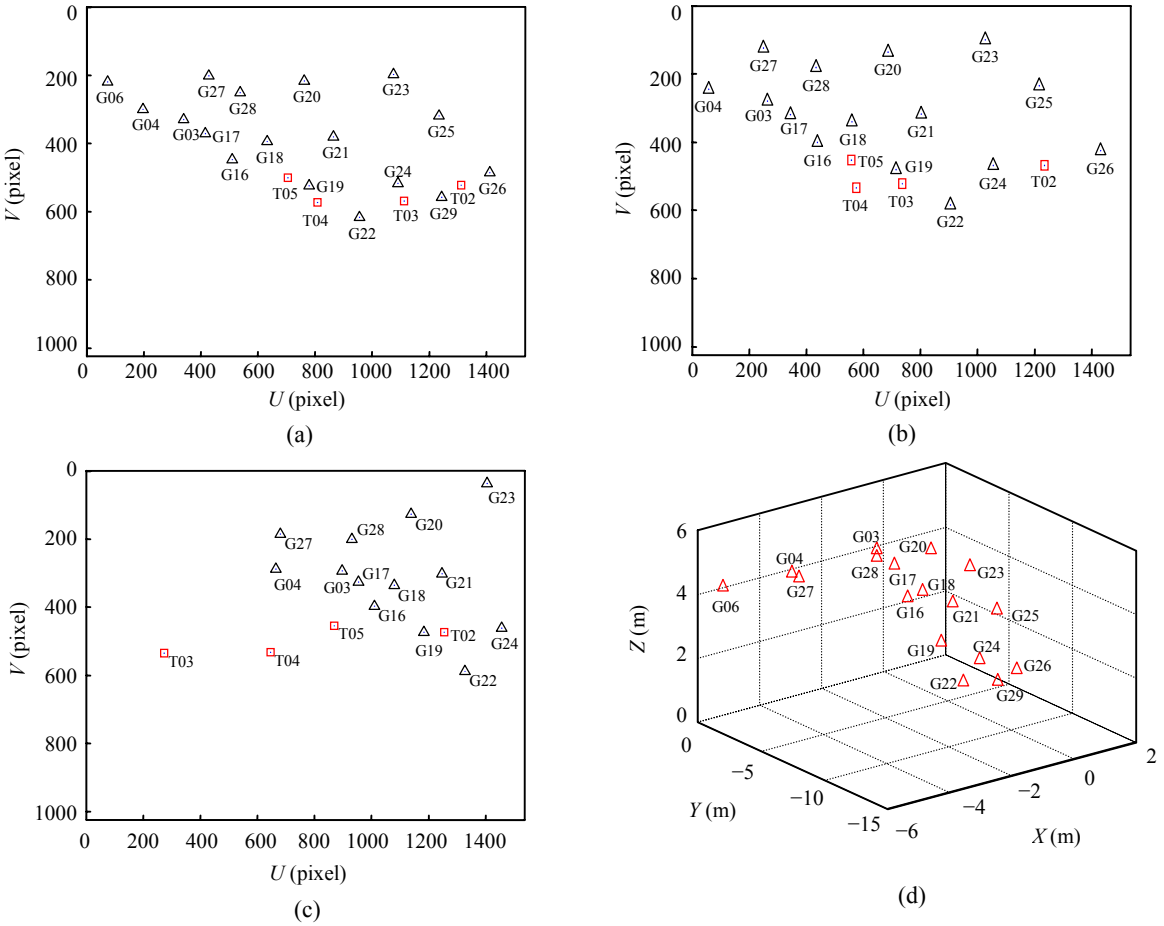

Fig.6 Distribution of the image points and check points. △: control points; □: check points
(a) Image 1; (b) Image 2; (c) Image 3; (d) Points distribution

CONCLUSION

This paper presents an iterative algorithm for metric reconstruction in the absence of pre-calibration or object-space control information with two views. It was assumed that the two images have equal focal length, and that the principal points were either known or determined from the fiducial marks. The direct solutions of focal length calibration in computer vision were used as the initial values, followed with 1D gold section searching for optimal focal length in a constrained region, and finally the relative poses of the two images were recovered for further metric reconstruction. For practical application, we also analyze the focal length accuracy influenced by the principal point error. Simulation and experiments indicated that the metric reconstruction results can be used as the initial values for bundle adjustment or other further refinements, and also can be used for low precision requirement situations. It also presents a prospect for close-range photogrammetry without professional metric cameras.

References

Förstner, W., 1999. On Estimating Rotations. Technical Report, T.U. München.

Fusiello, A., 2000. Uncalibrated Euclidean reconstruction: a review. *Image and Vision Computing*, **18**(6-7):555-563. [doi:10.1016/S0262-8856(99)00065-7]

Hartley, R., 1992. Estimation of Relative Camera Positions for Uncalibrated Cameras. Proc. 2nd European Conf. on Computer Vision, p.579-587.

Horn, B.K.P., 1987. Closed-form solution of absolute orientation using unit quaternion. *J. Opt. Soc. Am.*, **4**(4):629-642.

Horn, B.K.P., 1991. Relative orientation revisited. *J. Opt. Soc. Am.*, **8**(10):1630-1638.

Huang, F., Hu, Z.Y., Wu, Y.H., 2004. A new method on single view metrology. *Acta Automatica Sinica*, **30**(4):487-495.

Kanatani, K., Nakatsuji, A., Sugaya, Y., 2006. Stabilizing the focal length computation for 3D reconstruction from two uncalibrated views. *Int. J. Computer Vision*, **66**(2):109-122. [doi:10.1007/s11263-005-3952-y]

Lao, W., Cheng, Z., Kam, A.H., 2004. Focal Length Self-calibration Based on Degenerated Kruppa's Equations: Method and Evaluation. Int. Conf. on Image Processing. Singapore, p.3391-3394.

Newsam, G.N., Huynh, D.Q., Brooks, M.J., Pan, H.P., 1996. Recovering unknown focal lengths in self-calibration: an essentially linear algorithm and degenerate configurations. *Int. Arch. Photogram. Remote Sensing*, **XXXI**(B3):575-580.

Pan, H.P., Huynh, D.Q., Hamlyn, G., 1995. Two-image resituation: practical algorithm. *Proc. SPIE*, **2598**:174-190. [doi:10.1117/12.220898]

Pan, H.P., 1999. A direct closed-form solution to general relative orientation of two stereo views. *Digital Signal Processing*, **9**(3):195-221. [doi:10.1006/dspr.1999.0344]

Sturm, P., 2001. On Focal Length Calibration from Two Views. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, p.145-150.

Sturm, P., Cheng, Z.L., Chen, P.C.Y., Poo, A.N., 2005. Focal length calibration from two views: method and analysis of singular cases. *Computer Vision and Image Understanding*, **99**(1):58-95. [doi:10.1016/j.cviu.2004.11.002]

Ueshiba, T., Tomita, F., 2003. Self-calibration from Two Perspective Views under Various Conditions: Closed-form Solutions and Degenerate Configurations. Proc. Australia-Japan Advanced Workshop on Computer Vision. Adelaide, Australia, p.118-125.

Ueshiba, T., Tomita, F., 2004. A Closed-form Solution for a Two-view Self-calibration Problem under Fixation. Proc. 2nd Int. Symp. on 3D Data Processing, Visualization and Transmission, p.648-655. [doi:10.1109/TDPVT.2004.1335300]

Yuan, Y.X., Sun, W.Y., 2001. Optimization Theory and Methods. Science Press, Beijing, p.69-73 (in Chinese).

Zeng, Z.Q., 1990. A PC-based program of close range photogrammetry without approximate values. *J. Surveying and Mapping*, **19**(4):298-306 (in Chinese).

APPENDIX: QUATERNION ALGEBRA AND CALCULUS

A quaternion is a 4D extension to complex numbers. It proves that quaternion represents rotations and orientations in 3D. There are several methods we can use to represent quaternion. The two most popular notations are complex notation Eq.(A1) and 4D vector notation Eq.(A2).

$$\boldsymbol{q}=w+x\mathrm{i}+y\mathrm{j}+z\mathrm{k}, \tag{A1}$$

where w, x, y, z are real numbers, i, j, k are unit complexes, $\mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=-1$, $\mathrm{ij}=-\mathrm{ji}=\mathrm{k}$.

$$\boldsymbol{q}=[w\ \ \boldsymbol{v}], \tag{A2}$$

where $\boldsymbol{v}=(x, y, z)$ is called a vector part and w a scalar part.

Addition: $\boldsymbol{q}_1+\boldsymbol{q}_2=[w_1+w_2 \ \ \boldsymbol{v}_1+\boldsymbol{v}_2]$;

Multiplication:

$\boldsymbol{q}_1\otimes\boldsymbol{q}_2=(w_1w_2-\boldsymbol{v}_1\cdot\boldsymbol{v}_2, \ w_1\boldsymbol{v}_2+w_2\boldsymbol{v}_1+\boldsymbol{v}_1\times\boldsymbol{v}_2)$;

Multiplication represented by matrix:

$$\boldsymbol{q}_1\otimes\boldsymbol{q}_2=\boldsymbol{Q}_1\boldsymbol{q}_2=\begin{pmatrix} w_1 & -x_1 & -y_1 & -z_1 \\ x_1 & w_1 & -z_1 & y_1 \\ y_1 & z_1 & w_1 & -x_1 \\ z_1 & -y_1 & x_1 & w_1 \end{pmatrix}\begin{pmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

$$=\bar{\boldsymbol{Q}}_2\boldsymbol{q}_1=\begin{pmatrix} w_2 & -x_2 & -y_2 & -z_2 \\ x_2 & w_2 & z_2 & -y_2 \\ y_2 & -z_2 & w_2 & x_2 \\ z_2 & y_2 & -x_2 & w_2 \end{pmatrix}\begin{pmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{pmatrix};$$

Conjugate: $\boldsymbol{q}^*=[w \ -\boldsymbol{v}]$;

Norm: $N(\boldsymbol{q})=w^2+x^2+y^2+z^2$;

Inverse: $\boldsymbol{q}^{-1}=\boldsymbol{q}^*/N(\boldsymbol{q})$;

Unit quaternion: $\boldsymbol{q}$ is a unit quaternion with $N(\boldsymbol{q})=1$ and $\boldsymbol{q}^{-1}=\boldsymbol{q}^*$;

Rotation with quaternion: $\boldsymbol{r}'=\boldsymbol{q}\otimes\boldsymbol{r}\otimes\boldsymbol{q}^*$.