ICPR 2000 in Brcelona, Spain, Vol.1, pp.406-410 (2000)

Recursive factorization method for the paraperspective model based on the perspective projection

J. Fujiki

Electrotechnical Laboratory
Tsukuba, Ibaraki, Japan

T. Kurata

Electrotechnical Laboratory
Tsukuba, Ibaraki, Japan

Abstract

The factorization method, which allows us to reconstruct the motion of the camera and shape of the object simultaneously from multiple images, provides high stability in numerical computations and satisfactory results. To apply this method to real-time processing, the recursive factorization method has been proposed. However, factorization method based on the affine projection has a limitation in reconstruction accuracy, and to achieve accurate reconstruction, the motion should be restricted. To overcome this problem, we present a recursive factorization method for the paraperspective model based on the perspective projection. The present method is far superior to other ones, in that it not only achieves accurate Euclidean reconstruction in a short time but also provides high stability in numerical computations. Moreover, the method produces stable reconstruction in almost all cases even if some images contain errors because all images are treated as uniformly as possible.

1 Introduction

Recovering both motion and shape simultaneously from multiple images is an important and essential task in the field of computer vision. The factorization method based on the affine projection [7, 5] is an excellent method which provides high stability in numerical computations and relatively high quality of reconstruction. The method has been expanded recursively [4, 2] for application to real-time processing. However, factorization method on the affine projection has a limitation in reconstruction accuracy because the camera model is approximated linearly, and to achieve accurate reconstruction, the motion should be restricted.

To attain this, the factorization method based on the perspective projection has been presented[1, 6, 8], which requires estimation of the parameter called *projective depth* because the measurement matrix contains the projective depth. Christy and Horaud[1] computed the projective depth iteratively via iterative estimation of Euclidean shape. This method (referred to as the CH method throughout this paper) is equivalent to iterative estimation of the image coordinates by the paraperspective projection from those by the perspective projection. This method achieves accurate reconstruction of motion and shape, but is useful only calibrated camera because Euclidean shapes are required in iterative computations. Strum and Triggs[6] computed the projective depth in advance via epipolar geometry without performing iterative computation. However, the projective depth estimated via epipolar geometry is sensitive to measurement errors for feature points. Ueshiba and Tomita[8] estimated the measurement matrix containing projective depths as its elements using an evaluation function that treats all images as uniformly as possible. In this method, however, convergence of the iterative computation to estimate the measurement matrix containing projective depths takes a long time.

On the other hand, McLauchlan et al.[3] presented a recursive method for estimating motion and shape other than by factorization, through both affine projection and perspective projection using a variable state-dimension filter. However, reconstruction is very sensitive to the latest images.

In this paper, we present a recursive factorization method for the paraperspective model based on the perspective projection for the calibrated camera. The presented method allows fast and accurate Euclidean reconstruction. To realize accurate reconstruction, the CH method, which estimates the image coordinates of the paraperspective projection from the image coordinates of the perspective projection, is used. To reduce the processing time, recursive factorization based on the affine projection[2], which extracts motion information by using principal component analysis (PCA), is applied. The extraction of motion information by PCA treats all images as uniformly as possible and so the method gives stable reconstruction even if some images contain errors. Although reconstruction may be less than stable than by the batch process factorization method[7, 5], the reduction of processing time more than compensates for this stability.

Note that we use a calibrated camera throughout this paper. In this camera, an orthogonal relationship is established between the optical axis and the image plane, the intersection of them is the origin of the image coordinates and the aspect ratio is 1.

2 Factorization method for the paraperspective projection

This section outlines the factorization method for the paraperspective projection by Poelman and Kanade[5].

Let l be the focal length, $C_f = (\boldsymbol{i}_f, \boldsymbol{j}_f)^{\mathrm{T}}$ be an orthonormal basis of the f-th image plane, $\boldsymbol{k}_f$ be a unit vector along the camera's optical axis for the f-th image plane, $\boldsymbol{t}_f$ be a position vector of the camera center for the f-th image plane, $\boldsymbol{s}_p$ be a position vector of the p-th feature point in the world coordinates, and $C_f = (\boldsymbol{i}_f, \boldsymbol{j}_f, \boldsymbol{k}_f)^{\mathrm{T}}$ be the *camera coordinate system* of the f-th image plane ($f = 1, \ldots, F; p = 1, \ldots, P$).

Generally, the direction of the paraperspective projection is parallel to the direction passing through the center-of-mass of the object and the center of the camera. In this paper, we set the direction of the paraperspective projection as parallel to the direction passing through $\boldsymbol{s}_*$, which is a specific feature point, and the center of the camera to apply the CH method.

Let $\boldsymbol{x}_{fp}^{\mathrm{para}}$ be the image coordinates of the p-th feature point on the f-th image by the paraperspective projection. We use the relative coordinates from the specific feature point to simplify the relationship between the world coordinates of the feature points and the image coordinates of them. To treat the relative cooridnate simply, we define the quantities as $\boldsymbol{s}_p^* = \boldsymbol{s}_p - \boldsymbol{s}_*, \boldsymbol{t}_f^* = \boldsymbol{t}_f - \boldsymbol{s}_*, \boldsymbol{x}_{fp}^{*\mathrm{para}} = \boldsymbol{x}_{fp}^{\mathrm{para}} - \boldsymbol{x}_{f*}^{\mathrm{para}}$ where $\boldsymbol{s}_p^*$ and $\boldsymbol{t}_f^*$ are the relative world coordinates of $\boldsymbol{s}_p$ and $\boldsymbol{t}_f$ from the specific feature point $\boldsymbol{s}_*$, and $\boldsymbol{x}_{fp}^{*\mathrm{para}}$ are the relative image coordinates of $\boldsymbol{x}_{fp}^{\mathrm{para}}$ from the image coordinates of the specific feature point $\boldsymbol{x}_{f*}^{\mathrm{para}}$.

By using these quantities, the relationship between the world coordinates of the feature points and the image coordinates of them are given by

$$\boldsymbol{x}_{fp}^{*\mathrm{para}} = \lambda_{f*}^{-1}(l\,\overline{C}_f - \boldsymbol{x}_{f*}^{\mathrm{para}}\boldsymbol{k}_f^{\mathrm{T}})\boldsymbol{s}_p^*, \quad \lambda_{f*} = -\boldsymbol{k}_f^{\mathrm{T}}\boldsymbol{t}_f^* \quad (1)$$

where λ_{f*} is the depth of the specific feature point $\boldsymbol{s}_*$ of the f-th image.

By rearranging the equation (1) through f and p, we obtain the relationship

$$W^* = MS^* \quad (2)$$

where,
$W_f^* = (\boldsymbol{x}_{f1}^{*\mathrm{para}}, \ldots, \boldsymbol{x}_{fP}^{*\mathrm{para}})$, $W^* = (W_1^{*\mathrm{T}}, \ldots, W_F^{*\mathrm{T}})^{\mathrm{T}}$, $M_f = \lambda_{f*}^{-1}(l\,\overline{C}_f - \boldsymbol{x}_{f*}^{\mathrm{para}}\boldsymbol{k}_f^{\mathrm{T}})$, $M = (M_1^{\mathrm{T}}, \ldots, M_F^{\mathrm{T}})^{\mathrm{T}}$, $S^* = (\boldsymbol{s}_1^*, \ldots, \boldsymbol{s}_P^*)$. W^*, M and S^* are called *measurement matrix*, *motion matrix* and *shape matrix*.

By decomposing W^* as MS^*, we can reconstruct both the motion and shape. The factorization method performs the decomposition as described below.

First, W^* is temporarily decomposed as

$$\underset{(2F\times P)}{W^*} = \underset{(2F\times 3)}{\widehat{M}}\ \underset{(3\times P)}{\widehat{S}^*} = (\widehat{M}_1^{\mathrm{T}}, \ldots, \widehat{M}_F^{\mathrm{T}})\widehat{S}^*. \quad (3)$$

Generally, the singular value decomposition (SVD) is used for this temporal decomposition because SVD can estimate the least-square estimation of the rank 3 measurement matrix from the rank 4 or higher measurement matrix due to measurement errors and linear approximate errors of the model. At this point, the motion and shape have been affine reconstructed.

Second, computing a non-singular matrix A which satisfies $M = \widehat{M}A$, $S^* = A^{-1}\widehat{S}^*$ to achieve Euclidean reconstruction.

The constraints to compute A are represented as

$$\widehat{M}_f Q \widehat{M}_f^{\mathrm{T}} = M_f M_f^{\mathrm{T}} = \lambda_{f*}^{-2}(l^2\,\mathrm{I}_2 + \boldsymbol{x}_{f*}^{\mathrm{para}}\boldsymbol{x}_{f*}^{\mathrm{para\,T}}) \quad (4)$$

where $Q = AA^{\mathrm{T}}$. The constraints are called *metric constraints*. The metric constraints (4) are the linear homogeneous equation for six independent components out of 3×3 positive symmetric matrix Q, and Q is determined uniquely by removing a given multiple of freedom from three or more different images[5].

After computing Q, the general solution of A is derived as $A = L^{\mathrm{T}}U(^{\forall}U \in \mathrm{O}(3))$ where LL^{T} is the Cholesky decomposition of Q. Since U corresponds to the freedom for selecting the world coordinate system, there are substantially two sets of different solutions, which are mirror-symmetric to each other, depending on the positive/negative characteristic of det U.

After these steps, the extrinsic parameters $C_f, \boldsymbol{t}_f^*$ and λ_{f*} are reconstructed from the motion matrix $M_f = (\boldsymbol{m}_f, \boldsymbol{n}_f)^{\mathrm{T}}$ as shown below:

$$\lambda_{f*} = \left(\frac{\det\left(l^2\,\mathrm{I}_2 + \boldsymbol{x}_{f*}^{\mathrm{para}}\boldsymbol{x}_{f*}^{\mathrm{para\,T}}\right)}{\det(M_f M_f^{\mathrm{T}})}\right)^{1/4}, \quad (5)$$

$$C_f = \widetilde{A}_f^{-1}\widetilde{M}_f,\ \boldsymbol{t}_f^* = -(\lambda_{f*}/l)C_f^{\mathrm{T}}(\boldsymbol{x}_{f*}^{\mathrm{para\,T}}, l)^{\mathrm{T}} \quad (6)$$

where $A_f = \lambda_{f*}^{-1}(l\,\mathrm{I}_2, -\boldsymbol{x}_{f*}^{\mathrm{para}}) = (\boldsymbol{\alpha}_f, \boldsymbol{\beta}_f)^{\mathrm{T}}$,
$\widetilde{M}_f = (M_f^{\mathrm{T}}, \boldsymbol{m}_f \times \boldsymbol{n}_f)^{\mathrm{T}}$ and $\widetilde{A}_f = (A_f^{\mathrm{T}}, \boldsymbol{\alpha}_f \times \boldsymbol{\beta}_f)^{\mathrm{T}}$.

3 Christy-Horaud method

This section outlines the CH method [1] which can estimate the image coordinates of the paraperspective projection from those of the perspective projection in the framework of the factorization method.

Let $\boldsymbol{x}_{fp}^{\mathrm{per}}$ be the image coordinate of the p-th feature point of the f-th image by the perspective projection. We use the relative coordinates from the specific feature point as same reason as paraperspective projection. Therefore, we define the quantities as $\boldsymbol{x}_{fp}^{*\mathrm{per}} = \boldsymbol{x}_{fp}^{\mathrm{per}} - \boldsymbol{x}_{f*}^{\mathrm{per}}$ where $\boldsymbol{x}_{fp}^{*\mathrm{per}}$ are the relative image coordinates of $\boldsymbol{x}_{fp}^{\mathrm{per}}$ from the image coordinates of the specific feature point $\boldsymbol{x}_{f*}^{\mathrm{per}}$. By using these quantities, the relationship between the world coordinates of the feature points and the image coordinates of them are give by

$$\boldsymbol{x}_{fp}^{*\mathrm{per}} = \lambda_{fp}^{-1}(l\,\overline{C}_f - \boldsymbol{x}_{f*}^{\mathrm{per}}\boldsymbol{k}_f^{\mathrm{T}})\boldsymbol{s}_p^*,\ \lambda_{fp} = \lambda_{f*} + \boldsymbol{k}_f^{\mathrm{T}}\boldsymbol{s}_p^* \quad (7)$$

where λ_{fp} is called *projective depth* [8]. Because there holds $\boldsymbol{x}_{f*}^{\mathrm{para}} = \boldsymbol{x}_{f*}^{\mathrm{per}}$, the relationship between the image coordinates of the paraperspective projection and those of the perspective projection are given by

$$\boldsymbol{x}_{fp}^{*\mathrm{para}} = \mu_{fp}^*\boldsymbol{x}_{fp}^{*\mathrm{per}}, \quad \mu_{fp}^* = 1 + \lambda_{f*}^{-1}\boldsymbol{k}_f^{\mathrm{T}}\boldsymbol{s}_p^* \quad (8)$$

where μ_{fp}^* is a relative projective depth of the p-th feature point of the f-th image with respect to the specific feature point $*$.

Now, we define the measurement matrix containing projective depths as

$$W^* = \begin{pmatrix} \mu^*_{11}\boldsymbol{x}^*_{11} & \dots & \mu^*_{1P}\boldsymbol{x}^*_{1P} \\ \vdots & \ddots & \vdots \\ \mu^*_{F1}\boldsymbol{x}^*_{F1} & \dots & \mu^*_{FP}\boldsymbol{x}^*_{FP} \end{pmatrix} \quad (9)$$

where $\boldsymbol{x}^*_{fp} = \boldsymbol{x}_{fp} - \boldsymbol{x}_{f*}$ are observed quantities.

The algorithm for the CH method is as shown below.

(1) Let $\mu^*_{fp} = 1$ and perform Euclidean reconstruction with the factorization method as shown in section 2. Name one reconstruction of the motion and shape as $M^{(+)}, S^{*(+)}$ and its mirrored reconstruction of the motion and the shape as $M^{(-)}, S^{*(-)}$ respectively. Update $\mu^{*(\pm)}_{fp}$ by equation (8) for each solution (hereafter, the same symbols are used in the same manner).

(2) Update the measurement matrix W^*.

(3) Perform Euclidean reconstruction with the factorization method as shown in section 2.

(4) Update $M^{(\pm)}, S^{*(\pm)}$ by choosing the consistent solution from a pair of Euclidean reconstructions.

(5) Update $\mu^{*(\pm)}_{fp}$ by equation (8) for each solution.

(6) If $\mu^{*(\pm)}_{fp}$ does not converge, return to **(2)**.

(7) Select the solution which most closely matches the observed image coordinates out of $M^{(+)}, S^{*(+)}$ and $M^{(-)}, S^{*(-)}$.

4 Recursive factorization method based on the perspective projection

This section first describes the main processes in the presented method and then the algorithm for the method.

4.1 Fixing the world coordinates

Since reconstruction of the motion and shape involves the freedom of selecting the world coordinate system, the same motion and shape may be differently represented when different world coordinate systems are selected. Therefore, when considering the recursive factorization method, the world coordinate system must be fixed. For the present method, the world coordinate system is fixed by estimating the orthogonal matrix that connects two shape matrices.

Let $S^{(G)}_{\mathrm{ref}}$, $S^{(G)}_{\mathrm{ob}}$ be the shape matrix on the reference world coordinate system and that on an observed world coordinate system respectively (symbol $^{(G)}$ denotes that the center of mass of the feature points which comprise the shape matrix has been translated to the origin). The least-square estimation of the orthogonal matrix $\mathcal{E}$ which transforms the shape matrix on the observed world coordinate system into the shape matrix on the reference world coordinate system, that is, the orthogonal matrix $\mathcal{E}$ which satisfies $S^{(G)}_{\mathrm{ref}} = \mathcal{E}S^{(G)}_{\mathrm{ob}}$, is represented as $\mathcal{E} = UV^{\mathrm{T}}$ where UDV^{T} is the SVD of $S^{(G)}_{\mathrm{ref}}S^{(G)\mathrm{T}}_{\mathrm{ob}}$. This result can be given straightforwardly by using Lagrange multiplier method.

4.2 Occlusion process

If any feature point is occluded, only the world coordinates consisting of feature points actually observed are updated using the measurement matrix containing only feature points actually observed. The occluded feature point, if any, is not updated for the world coordinates.

4.3 Addition of the new feature point

When a new feature point is observed in two or more images with four or more points of which the world coordinates are known, then its world coordinates are determined.

We assume that the new feature point is observed through the image set $\mathcal{F} = \{f_1, \dots, f_J\}$. Let $\boldsymbol{X}^*_{\langle\mathcal{F}\rangle,\mathrm{new}} = (\mu^*_{f_1,\mathrm{new}}\boldsymbol{x}^{*\mathrm{T}}_{f_1,\mathrm{new}}, \dots, \mu^*_{f_J,\mathrm{new}}\boldsymbol{x}^{*\mathrm{T}}_{f_J,\mathrm{new}})^{\mathrm{T}}$ be the arranged image coordinates of the new feature point containing the relative projective depth and let $W^*_{\langle\mathcal{F}\rangle}$, $M_{\langle\mathcal{F}\rangle}$ and $S^*_{\langle\mathcal{F}\rangle}$ be the measurement matrix, the motion matrix and the shape matrix of the observed feature points through $\mathcal{F}$ respectively, the transformation from the image coordinates into the world coordinates is represented as $\mathcal{T}_{\langle\mathcal{F}\rangle} = [M^{\mathrm{T}}_{\langle\mathcal{F}\rangle}M_{\langle\mathcal{F}\rangle}]^{-1}M^{\mathrm{T}}_{\langle\mathcal{F}\rangle}$. Therefore, we can estimate the world coordinates of the new feature point by the following process:

(1) Let $\mu^*_{f_j,\mathrm{new}} = 1$.

(2) Update $\boldsymbol{s}^*_{\mathrm{new}}$ by $\boldsymbol{s}^*_{\mathrm{new}} = \mathcal{T}_{\langle\mathcal{F}\rangle}\boldsymbol{X}^*_{\langle\mathcal{F}\rangle,\mathrm{new}}$.

(3) Update $\mu^*_{f_j,\mathrm{new}}$ by $\mu^*_{f_j,\mathrm{new}} = 1 + \lambda^{-1}_{f_j*}\boldsymbol{k}^{\mathrm{T}}_{f_j}\boldsymbol{s}^*_{\mathrm{new}}$.

(4) If $\mu^*_{f_j,\mathrm{new}}$ does not converge, return to **(2)**.

4.4 Extraction of motion information

Each row of the measurement matrix represents the projection of the shape matrix along each row of the motion matrix. Therefore, the reliability of the reconstruction of the shape observed in some direction may be given as the covariance matrix of each row of the motion matrix. Thus, by compressing motion information while maintaining the covariance matrix by applying PCA to the motion matrix, we can compress the measurement matrix with no loss of reliability of reconstruction of the shape.

Let $F\Lambda E$ be the SVD of M, Λ represents the principal component (PC) of M and each row of E represents the PC vector of M according to the PCA. Therefore, each row of $\mathcal{M} = \Lambda E$ and each row of M have the same covariance matrix. Thus, we can consider that the motion information contained in 3×3 matrix $\mathcal{M}$ is identical to that in $2F \times 3$ matrix M, that is, $\mathcal{M}$ is the extraction of the motion information contained in M. We call M a *PC motion matrix*. Here, in relation to the PC motion matrix, the measurement matrix (called a *PC measurement matrix*) and metric constraints are represented as $\mathcal{W}^* = \mathcal{M}S^*$ and $\mathcal{M}\mathcal{M}^{\mathrm{T}} = \Lambda^2$.

4.5 Algorithm

(1) Perform Euclidean reconstruction using the CH method with $k(\geq 3)$ images and compute the PC motion matrix as shown in section **4.4**. The world coordinate system describing the reconstruction is defined as the reference world coordinate system.

(2) Perform the $f(> k+1)$-th *recursive measurement matrix* $W^*_{[f]}$, which consists of the $f-1$-st PC measurement matrix $\mathcal{W}^*_{[f-1]}$ and the f-th measurement matrix containing the projective depths, from the feature points observed in the f-th image (each image coordinate of the recursive measurement matrix is the relative coordinate from the specific feature point).

(3) Perform Euclidean reconstruction using the CH method and compute the motion matrix $M_{[f]}$ and the shape matrix $S^*_{[f]}$ in the reference world coordinate system as shown in section **4.1**. In this case, the f-th motion matrix M_f is the last two rows of $M_{[f]}$. Note that only W^*_f is updated in iterative computation on CH method and $\mathcal{W}^*_{[f-1]}$ is not updated.

(4) Any occlusion is processed and new feature points are added as shown in section **4.2** and **4.3**. At this point, reconstruction of the motion and shape using the f-th image is completed.

(5) Motion information is extracted as shown in section **4.4**.

5 Experiment

We use synthetic data to evaluate the presented method with the recursive factorization method based on the affine projection[2].

5.1 Data generation

Ninety-two feature points are generated on the surface of an opaque sphere of 150mm radius (the left side of Fig. 1). The focal length of the camera is 84mm converted to its 35mm camera equivalent. Each pixel is a square of side 8μm. We generate 121 images of which the camera coordinate matrix is set to $C_f^{\mathrm{T}} = R_y(\pi f/60)$. The depth of the center of the sphere changes monotonously from 1200mm to 800mm. The size of each image is 640×480 pixels and the image coordinates of the center of the sphere change from $(80, 80)$ to $(10, 10)$ monotonously (the center and the right side of Fig. 1). The image coordinates are sampled by sub-pixel order. We take 10 images for the initial estimation. We add new feature points when the new feature points are observed through 10 successive images. In the CH method, μ^*_{fp} has regarded to converged when its the change is less than 10^{-4}. In this case, the number of iterations is four or five times.

5.2 Error measurement

We evaluate the reconstruction error of the motion and shape of the presented method (PERRFM) with the recursive factorization method based on the affine projection[2] (PARARFM). Figure 2 compares the shapes reconstructed by the presented method and those by the recursive factorization method based on the affine projection[2].

Figure 3 shows the shape reconstruction error. The upper side shows when the sphere is transarent, that is, all feature points are observed through all images and the lower side shows when the sphere is opaque, that is, the occlusion process and addition of new feature points are considered. We also show the shape reconstruction error of the batch factorization method based on the affine projection [5] (PARA) and the batch CH method (PER) for all feature points observed through all images. The shape reconstruction error is given by

$$||S^{(\mathrm{G})^{\mathrm{true}}} - S^{(\mathrm{G})^{\mathrm{estimated}}}|| \cdot ||S^{(\mathrm{G})^{\mathrm{true}}}||^{-1} \times 100(\%).$$

Figure 4 shows the motion reconstruction error of $\boldsymbol{i}_f$. The upper side shows transarent sphere and the lower side shows opaque sphere. We also show the shape reconstruction error of the batch factorization method based on the affine projection [5] and the batch CH method for all feature points observed through all images. The reconstruction errors of $\boldsymbol{j}_f$ and $\boldsymbol{k}_f$ are not shown because they behave in the same way as $\boldsymbol{i}_f$. The motion reconstruction error is measured as the angle between the true direction and estimated direction.

As Fig. 3 and Fig. 4 show, the reliability of the motion and shape reconstruction by this method is far superior to that by the batch factorization method based on the affine projection [5] and to that by the recursive factorization method based on the affine projection [2]. Accurate reconstruction can be achieved even if occlusions exist and new feature points are added.

When we add the Gaussian noise of which covariance is two pixels, the reconstruction error for opaque shere increases only 0.3%.

5.3 Computational time

The presented method takes about one second to process the data given in section **5.1**. However, the computation time required for update depended largely on the number of observed feature points at each moment instead of the total number of feature points. Therefore, in comparing the computation time of the presented method with that of other methods, we assume that all feature points are observed in each image. We used Pentium II 450MHz PC for computation. The number of images used was fixed at 121, while the number of feature points was increased from 20 to 100. Under such conditions, computation times were measured. The processing time for selecting and tracking features points was not included in the computation times. The computation time for the batch factorization method is the sum of computation times required for individual batch processing of each image. Figure 5 shows that the presented method has a reasonable computational time. Although the computational time of the presented method is longer than that of the recursive factorization method based on the affine projection [2] due to the use of the CH method, the high accuracy of reconstruction of the presented method more than compensated for the longer computational time.

6 Conclusion

In this paper, We present the factorization method for paraperspective based on the perspective projection, which estimates the motion of the camera and the shape of the object every time an image is taken. The presented method achieves high-accuracy Euclidean reconstruction compare with the factorization methods based on the affine projection and achieve high-speed processing compare with the batch factorization methods. The presented method also provides higher stability in numerical computation and might give a stable reconstruction in almost all cases.

References

[1] S. Christy and R. Horaud, "Euclidean reconstruction: from paraperspective to perspective," Proc. 4th ECCV, vol.2, pp.129–140, 1996.

[2] J. Fujiki, T.Kurata and M.Tanaka, "Iterative factorization method for object recognition," Proc. SPIE98, San Diego, Vision Geometry VI, vol.3454-18, pp.192–201, July 1998.

[3] P.F. McLauchlan and D.W. Murray, "A unifying framework for structure and motion recovery from image sequences," Proc. 5th ICCV, pp.314–320, 1995.

[4] T. Morita and T. Kanade, "A sequential factorization method for recovering shape and motion from image streams," IEEE Trans. PAMI, vol.19, no.8, pp.858–867, 1997.

[5] C.J. Poelman and T. Kanade, "A paraperspective factorization method for shape and motion recovery," IEEE Trans. PAMI, vol.19, no.3, pp.206–218, 1997.

[6] P. Sturm and B. Triggs, "A factorization based algorithm for multi-image projective structure and motion," Proc. 4th ECCV, vol.2, pp.709–720, 1996.

[7] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method," IJCV, vol.9, no.2, pp.137–154, 1992.

[8] T. Ueshiba and F. Tomita, "A factorization method for projective and Euclidean reconstruction from multiple perspective views via iterative depth estimation," Proc. 5th ECCV, vol.1, pp.296–310, 1998.

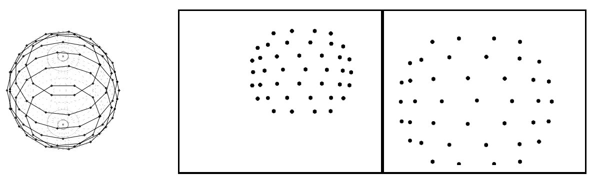

Figure 1: Feature points (left), 1st frame (center), 121st frame (right)

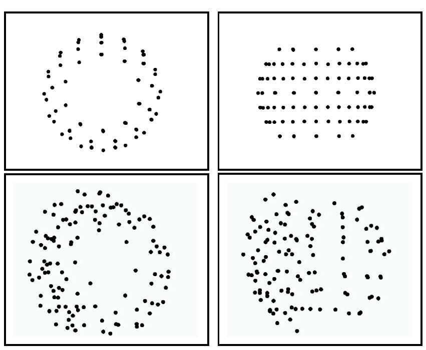

Figure 2: Reconstruction shape by the presented method (upper) and the recursive factorization method based on the affine projection (lower). The left images are viewed from the pole and the right images are viewed from the equator respectively.

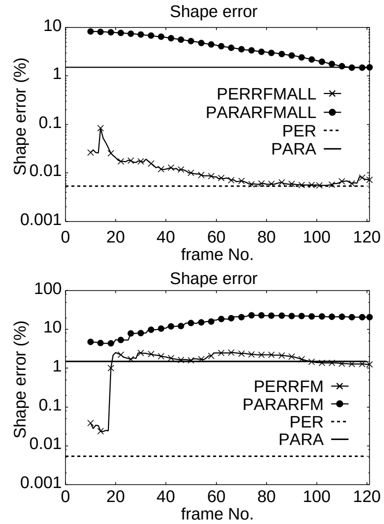

Figure 3: Shape reconstruction error of transparent sphere(upper) and opaque sphere(lower).

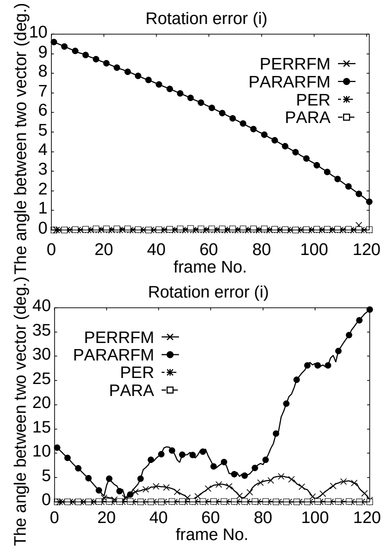

Figure 4: Motion reconstruction error of transparent sphere(upper) and opaque sphere(lower).

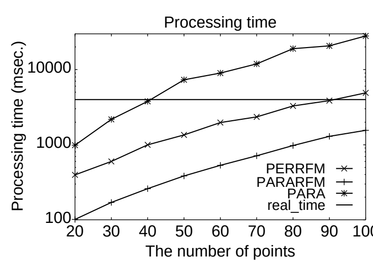

Figure 5: Computational time