A NEW RELIABILITY MEASURE FOR ESSENTIAL MATRICES SUITABLE IN MULTIPLE VIEW CALIBRATION

Jaume Vergés-Llahí
ATR Intelligent Robotics and Communication Laboratories, Kyoto 619-0288, Japan
jverges@atr.jp

Daniel Moldovan
NICT Universal Media Center and ATR CIS Laboratories, Kyoto 619-0288, Japan
danielm@atr.jp

Toshikazu Wada
Dept. of Computer & Communication Science, Wakayama University, Wakayama 640-8510, Japan
twada@ieee.org

Keywords: Epipolar geometry, reliability measure, essential matrix, camera-dependency graph.

Abstract: This paper presents a new technique to recover structure and motion from a large number of images acquired by an intrinsically calibrated perspective camera. We describe a method for computing reliable camera motion parameters that combines a camera–dependency graph, which describes the set of camera locations and the feasibility of pairwise motion calculations, and an algorithm for computing the weights on the edges of this graph. A new criterion for evaluating the reliability of the essential matrices thus produced with respect to the epipolar constraint is here introduced. It is composed of two main elements, namely, the uncertainty of the renormalization process by which the essential matrix is derived and the error between the estimated matrix and its decomposition into the motion parameters of translation and rotation. Experimental results show that there exists a clear correlation between the proposed reliability measure and the error in the estimation of such motion parameters. The performance of the proposed method is demonstrated on a sequence of short base-line images where it is made clear that the strategy based on the shortest paths in terms of unreliability provides remarkably superior results to those obtained from the paths of consecutive camera locations.

1 INTRODUCTION

The purpose of calibration from multiple views consists in recovering the spacial location of a certain set of points along with the determination of the position of the camera from where these points were viewed. Despite the problem of obtaining such a 3D structure from the motion of a camera has been extensively studied for the last two decades (Hartley and Zisserman, 2003) it is remarkable that any previous approaches based their strategies on choosing *consecutive* views rather than on a more advantageous combination, that is, the one that would provide the least recovery error. It seems therefore quite advisable to try to establish a feasible mechanism for indirectly estimating this error as a mean of selecting the best among all such combinations.

This paper is focused on the proposal and justification of a new measure of reliability that captures the error of the recovered 3D structure and the camera movement between a pair of views which is suitable for carrying out a multiple view calibration. The main idea is to employ this measure for the purpose of evaluating beforehand an estimation of the validity of a certain pair of views and collecting such information in the process of selection of the best combination of views that provides the least recovery error between two given camera locations.

In general, it is known that the accuracy of the estimation of both the motion parameters and the recovered structure may greatly degrade when an increasing number of closely consecutive views are added into the computations. This is mainly due to inaccuracies accumulated throughout as well as to short baselines. Nevertheless, the larger the baseline is, the better the accuracy should become. Hence, the pairs of views that must be taken into account in the computations should be as apart each other as possible to improve the recovery results. In the purpose of efficiently considering all possible combinations of cam-

Vergés-Llahí J., Moldovan D. and Wada T. (2008).
A NEW RELIABILITYMEASURE FOR ESSENTIAL MATRICES SUITABLE IN MULTIPLE VIEWCALIBRATION.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 114-121
DOI: 10.5220/0001078301140121
Copyright © SciTePress

era locations, we suggest the usage of the graph structure we call *Camera–Dependency Graph* (CDG).

The CDG is composed of a set of nodes representing each view of the scene taken from a different location of the camera, whereas the weight on the edges corresponds to the degree of reliability of the pair of views being connected. Since the measure proposed in this paper will be shown to closely correlate with the recovery error, it is suggested that the most trustworthy sequence of views, in other words, the one with the smallest error, would be obtained by selecting the path in CDG that minimizes the total amount of *unreliability*, since it is that measure which is a indirect estimator of the recovery error.

The characterization of the unreliability of a camera pair is carried out by estimating the uncertainty of its epipolar constraint, i.e., the relative position and orientation of the camera. This is accomplished by way of two partial error estimations. The first one encompass the error produced in the iterative *correction* by which the essential matrix is obtained. The second comes from the *decomposition* of this matrix into a translation vector and a rotation matrix. In this paper we will show how the combination of these two values correlates with the recovery error in most of cases. As a consequence, the proper selection of views based on such a measure will improve the accuracy of the recovered structure and motion.

This paper is organized as follows. First, a review of some previous works in a similar problem is carried out, followed by the description and justification of the measure of unreliability proposed here. Afterwards, the experimental section will be described as well as the results obtained for the purpose of confirming our claims. This section will focus in two aspects, namely, the proof of the correlation between our criterion and the recovery error, and the usage of CDG as a route through substantially better multiple view calibrations. Finally, the conclusions drawn from the obtained results will be discussed along with the future work necessary to fulfil this research.

2 PREVIOUS WORK

The only attempt to our knowledge of evaluating the epipolar constraint quality to estimate a multiple view reconstruction is that of Martinec & Pajdla (Martinec and Pajdla, 2006). They introduced a so called *reliability–importance* matrix in which the *reliability* is based on the number of supporting inliers and the *importance*, on finding the shortest paths in a graph induced by a known epipolar geometry. Comparatively, our work employs a different approach for the *unreliability* which estimates how close the epipolar constraint is fulfilled by the resulting motion parameters as a combination of the uncertainty of the essential matrix and its decomposition error.

Chronologically, multiple view reconstruction was approached for the first time by Tomasi and Kanade (Tomasi and Kanade, 1992) that used factorization on affine cameras. An extension for perspective cameras was given later in (Sturm and Triggs, 1996). Perspective effect was handled using both epipolar geometry (Sturm and Triggs, 1996; Schaffalitzky and Zisserman, 2002; Martinec and Pajdla, 2005) and trifocal tensor (Fitzgibbon and Zisserman, 1998). In all these methods, points need to be visible in at least three views so as to glue partial reconstructions. Otherwise, a sequence of independently computed fundamental matrices or trilinear tensors might be optimally consistent with the image data, but not necessarily consistent with a unique camera trajectory. This is an important constraint on views.

The study of the essential matrix as a method of determining the epipolar geometry was initially performed in (Longuet-Higgins, 1981) and later generalized in (Luong and Faugeras, 1996) by the introduction of fundamental matrix when internal camera parameters were unknown. Two different methods for estimating the stability of fundamental matrix were introduced in (Csurka et al., 1997), namely, a statistical one and an analytical one. The first procedure yielded better results in case the noise level of data was known, despite this is not the usual case besides being computational expensive, while the second method performed better if the noise was moderate.

A different approach was introduced by Kanatani in (Kanatani, 2000). Starting from the same linear hypothesis describing the epipolar constraint, he derived a nonlinear optimization method whose optimal unbiased estimate was computed based on an iterative process of *renormalization* without enforcing the rank constraint. The obtained solution was afterwards *corrected* in order to fulfil that constraint. Experiments indicated that the obtained estimates were in the vicinity of the theoretical accuracy bound. This work is the origin of our work, which has been extended to encompass more complex calibrations described by the paths in CDGs.

3 CAMERA-DEPENDENCY GRAPH (CDG)

In this section we introduce the new concept of *Camera–Dependency Graph* (CDG) suitable to compute the external camera parameters[1] between any two camera locations as a path of intermediate positions. Specifically, a CDG is a graph $G = (\mathcal{V}, \mathcal{E})$ where the set of nodes $\mathcal{V}$ represents camera locations and the set of edges $\mathcal{E}$ relates two positions whenever the calculation of their relative movement is feasible, i.e., when enough common points can be seen from the two positions. Consequently, the complete movement between two camera positions is a concatenation of the intermediate displacements expressed as a path in CDG, as shown in Figure1.

The accuracy of the results greatly depends on how paths are selected from CDG. As mentioned above, the recovery error is greatly dependent on the *baseline* distance between successive camera locations. Besides, the amount of error also accumulates and the total accuracy decreases as the number of intermediate positions increases. Therefore, in order to improve the total performance of the multiple view calibration two main strategies can be attempted, i.e., using locations with larger baselines and reducing the number of intermediate positions, especially those with worse estimates.

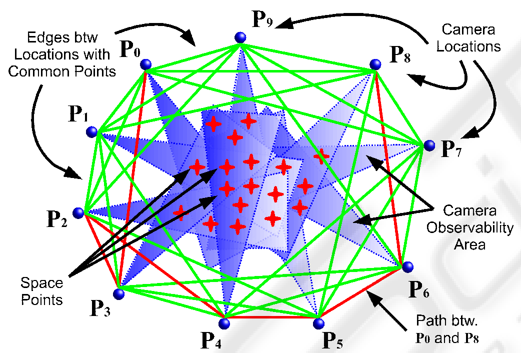


Figure 1: Camera-Dependency Graph (CDG). Graph of dependencies where nodes are camera locations and edges connect cameras sharing enough common points inside their observable areas.

Our approach selects a combination of intermediate views that connects two camera locations with the purpose of reducing the recovery error as much as possible. The *shortest* path in a CDG where edge weights are unreliabilities of pairwise camera motion is chosen. Since the shortest path corresponds to the combination having the smallest summation of unreliabilities and these values correlate with the error, the resulting camera movement and 3D data will consequently present a much lower amount of error compared to any other feasible path.

Some additional issues must be coped with in order to perform in practice such calculations. First, an algorithm to find the shortest paths that also fulfil a number of conditions is necessary. A path is only feasible if it always has enough common points visible from any three successive positions in it. This is equivalent to the existence of positions forming *triangles* and the whole path being *triangle-connected*, a property any path in a CDG must fulfil. Besides, the task of combining the pairwise displacements along a path in order to attain the complete movement must also be carefully addressed. Nonetheless, the description of these algorithms are out of the scope of this paper and will not be addressed here.

[1]Rotation matrix and translation vector.

4 UNRELIABILITY MEASURE OF ESSENTIAL MATRICES

In this section we define our measure of *unreliability* based on the epipolar constraint encompassed by means of the essential matrices $\mathbf{G}$ that will be used to form the weights in CDGs. Our starting point is the approach by Kanatani (Kanatani, 1996; Kanatani, 2000), where a theoretical accuracy bound on fundamental matrices is described. Despite that in our approach the fundamental matrix has been turned into an essential matrix, yet the same theory holds here. We basically quantify the error made in the two processes employed to compute an estimate of the movement parameters, namely, the *renormalization* error, coping with the error during the least–squares fitting of $\mathbf{G}$, and the *decomposition* error, accounting for the error carried out in the decomposition of $\mathbf{G}$ into its translation vector and rotation matrix.

4.1 Renormalization Error

The uncertainty of an estimate $\mathbf{G}$ is measured from the actual $\bar{\mathbf{G}}$ by the *covariance tensor* $\mathcal{V}[\mathbf{G}] = E[\mathcal{P}((\mathbf{G} - \bar{\mathbf{G}}) \otimes (\mathbf{G} - \bar{\mathbf{G}}))\mathcal{P}^{\top}]$, where $E[\circ]$ denotes expectation. The operator $\otimes$ stand for the tensor product among matrices, that is, if $\mathbf{A} = (A_{ij})$ and $\mathbf{B} = (B_{ij})$, the $(ijkl)$ element of their tensor product is $A_{ij}B_{kl}$. For tensors $\mathcal{P} = (P_{ijkl})$ and $\mathcal{T} = (T_{ijkl})$, the product $\mathcal{P}\mathcal{T}\mathcal{P}^{\top}$ is a tensor whose $(ijkl)$ element are $\sum_{m,n,p,q=1}^{3} P_{ijmn}P_{klpq}T_{mnpq}$, whereas the $(ijkl)$ elements of tensor $\mathcal{P}$ is given by $P_{ijkl} = \delta_{ij}\delta_{kl} - \bar{G}_{ij}\bar{G}_{kl}$, being δ_{ij} the Kronecker's delta.

There exists (Kanatani, 2000) a *theoretical lower bound* (TLB) on the covariance tensor $\mathcal{V}[\mathbf{G}]$ which represents an accuracy bound in the form

$$\mathcal{V}[\mathbf{G}] \succ \frac{\varepsilon^2}{N} \left(\mathcal{P}^S \, \mathcal{G} \, \mathcal{P}^S\right)_7^{-} \tag{1}$$

where $\mathcal{T} \succ \mathcal{S}$ for tensors $\mathcal{T}$ and $\mathcal{S}$ means that $(\mathcal{T} - \mathcal{S})$ is a positive semi–definite tensor, and the operation $(\circ)_r^-$ denotes the Moore–Penrose's inverse of rank r. The $(ijkl)$ element of the tensor $\mathcal{P}^S = (P^S_{ijkl})$ in Eq. (1) is given by $P_{ijkl} = \delta_{ij}\delta_{kl} - (\bar{G}^\dagger_{ij}\bar{G}^\dagger_{kl})/\|\bar{\mathbf{G}}^\dagger\|^2$, where $\bar{\mathbf{G}}^\dagger$ is the cofactor matrix of $\bar{\mathbf{G}}$.

On the one hand, from the renormalization step, which employs the unbiased least–squares eigenvalue fitting algorithm to approximate $\mathbf{G}$ (Kanatani, 2000), the *minimum* residual $J = (\mathbf{G}_9; \mathcal{G}\mathbf{G}_9)$ is extracted, providing an estimate of the squared noise level $\varepsilon^2 = J/(1-8/N)$ after renormalization. $\mathbf{G}_9$ is the eigenmatrix with the smallest eigenvalue of tensor $\mathcal{G}$. The covariance tensor $\mathcal{V}[\mathbf{G}]$ of the estimate $\mathbf{G}$ is then

$$\mathcal{V}[\mathbf{G}] = \frac{\varepsilon^2}{N}(\mathcal{G})_8^- \qquad (2)$$

where the estimate $\mathcal{G}$ is computed from *eigenvalues* λ_i and *eigenmatrices* $\mathbf{G}_i$ obtained in the renormalization algorithm as $\mathcal{G} = \sum_{i=1}^{8} \lambda_i \mathbf{G}_i \otimes \mathbf{G}_i$.

On the other hand, the *Root Mean Square* error (RMS) of $\mathbf{G}$ is defined as $rms[\mathbf{G}] = (E[\|\mathcal{P}(\mathbf{G} - \bar{\mathbf{G}})\|^2])^{1/2}$ and there exists a relation between this measure of accuracy and the covariance tensor $\mathcal{V}[\mathbf{G}]$ given by the trace of a tensor $\mathcal{T}$ as

$$rms[\mathbf{G}] \geq \sqrt{tr(\mathcal{V}[\mathbf{G}])} \qquad (3)$$

where $tr(\mathcal{T}) = \sum_{i,j=1}^{3} T_{ijij}$. Therefore, putting Eq. (1) and Eq. (3) together and writing them in terms of their eigenvalues, we obtain that

$$rms[\mathbf{G}] \geq \sqrt{\frac{\varepsilon^2}{N}\sum_{i=1}^{8}\frac{1}{\lambda_i}} \geq \sqrt{\frac{\varepsilon^2}{N}\sum_{i=1}^{7}\frac{1}{\lambda'_i}} = \varepsilon_r \qquad (4)$$

where λ_i are the eigenvalues of tensor $\mathcal{G}$ while λ'_i are these of tensor $\mathcal{P}^S\mathcal{G}\mathcal{P}^S$. The *renormalization* error ε_r is then defined as the lower bound in Eq.(4).

The relation between $rms[\mathbf{G}]$ and the TLB shows that renormalization attains this bound when higher order terms of noise are omitted (Kanatani, 2000). Hence, in practice this bound is a good approximation for the estimation of the error of the essential matrix. Nevertheless, if any further step is involved in the obtaining of the movement parameters, as it is in our case, a complementary measure is needed.

4.2 Decomposition Error

Two further steps are required to obtain the estimate of the translation $\mathbf{t}$ and the rotation $\mathbf{R}$ from an essential matrix $\mathbf{G}$. First, a geometric *correction* of $\mathbf{G}$ previously computed by renormalization to make it decomposable into the form $\mathbf{G} = \mathbf{t} \times \mathbf{R}$. Second, the *decomposition* into its movement parameters.

The first process is a Newton iteration based on a linear approximation of the decomposability constraint that can be carried out up to the same level of error attained in the renormalization. The decomposition itself is a robust method that provides a translation $\mathbf{t}$ being the unit eigenvector of matrix $\mathbf{G}\mathbf{G}^\top$ and a rotation $\mathbf{R} = \mathbf{V}\,diag(1,1,det(\mathbf{V}\mathbf{U}^\top))\mathbf{U}^\top$, where $\mathbf{V}$ and $\mathbf{U}$ come from the SVD of matrix $-\mathbf{t} \times \mathbf{G}$.

Remarkably, this method always provides a decomposable solution, since $\mathbf{R}$ is an *exact* rotation matrix. Furthermore, the vector $\mathbf{t}$ is always very close to the valid solution. Both facts are true even if $\mathbf{G}$ is not decomposable (Kanatani, 1996). Consequently, the decomposition can be seen as an ultimate stage in the optimal correction of the essential matrix $\mathbf{G}$ obtained by renormalization, producing an improved result.

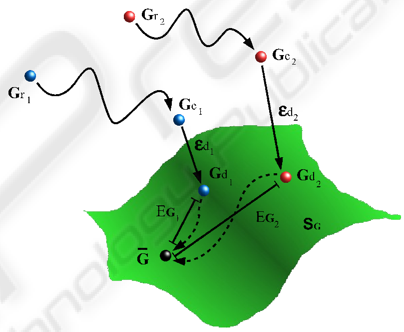


Figure 2: Decomposition Error. Symbols $\mathbf{G}_r$, $\mathbf{G}_c$, and $\mathbf{G}_d$ correspond to essential matrices after renormalization, correction, and decomposition, respectively. $E_\mathbf{G}$ and ε_d stand for recovery and decomposition errors. $S_\mathbf{G}$ is the manifold of decomposable matrices and $\bar{\mathbf{G}}$ represents the true one.

As a way to estimate the error in the calculation of the movement parameters $E_\mathbf{G}$, we suggest to measure how far the matrix $\mathbf{G}$ is from being truly decomposable. Therefore, the *decomposition* error is defined as

$$\varepsilon_d = \|\mathbf{G}_c - \mathbf{G}_d\| \qquad (5)$$

where $\mathbf{G}_c$ comes from the *correction* and $\mathbf{G}_d = \mathbf{t} \times \mathbf{R}$, being $\mathbf{t}$ and $\mathbf{R}$ obtained from $\mathbf{G}_c$ by *decomposition*. Our claim is that the farther a matrix $\mathbf{G}$ is from the decomposability constraint, the greater its decomposition differs from the real one and, therefore, the less reliable the matrix becomes, that is, if $\varepsilon_{d_1} \leq \varepsilon_{d_2}$ then $E_{\mathbf{G}_1} \leq E_{\mathbf{G}_2}$, as depicted in Fig.2.

4.3 Unreliability Measure

In Sec. 5.1 we show that in practice there are some cases where there exists a clear correlation between ε_d and the *recovery* error defined as $E_\mathbf{G} = \|\bar{\mathbf{G}} - \mathbf{G}_d\|$, where $\bar{\mathbf{G}}$ is the true essential matrix and $\mathbf{G}_d$ is the one obtained after decomposition. In other cases, the correlation is clearer with the renormalization error ε_r.

In order to improve the correlation with respect to $E_{\mathbf{G}}$ in any situation, the two previous measures are combined in one single value called *unreliability* $\nu_{\mathbf{G}}$ defined as follows

$$\nu_{\mathbf{G}} = \varepsilon_d \cdot \varepsilon_r = \|\mathbf{G}_c - \mathbf{G}_d\| \cdot \sqrt{\frac{\varepsilon^2}{N} \sum_{i=1}^{7} \frac{1}{\lambda'_i}} \qquad (6)$$

The value $\nu_{\mathbf{G}}$ corresponds to the weights on the edges of CDG and the paths, representing sequences of camera locations, will be selected to be the shortest ones in terms of this measure of unreliability. Posteriorly, these paths are used to recover the complete camera movement between to given positions. As said, since $\nu_{\mathbf{G}}$ correlates with the recovery error $E_{\mathbf{G}}$ and the paths thus obtained have the least possible unreliability, it follows that the movement recovered from these paths will have less error than other kind of feasible paths as shown in Sec. 5.2.

5 EXPERIMENTS AND RESULTS

This section describes the data employed and the experiments carried out, as well as the results obtained, in order to show the feasibility of the CDG framework based on the unreliability measure defined before as a way to perform multiple view calibrations.

The goals of the experiments are, first, to establish the correlation between the unreliability $\nu_{\mathbf{G}}$ of the estimated $\mathbf{G}$ and the recovery error $E_{\mathbf{G}}$ so as their use can be considered equivalent. Second, $\nu_{\mathbf{G}}$ and the CDG derived are applied to recover the camera movements and the 3D structure employing the sets of image points from a (generated) sequence of locations along a circular trajectory of the camera as data. Our aim is to display the recovery results attained by using the shortest path in terms of the unreliability $\nu_{\mathbf{G}}$ are better than those of the usual path of consecutive camera locations.

5.1 Description of the Data

Both the ease of obtaining a sufficient number of data to perform a generous and varied number of experiments in order to prove our claims, the ability of controlling the all the setting factors and the noise levels, as well as the necessity of having a precise ground truth to compare our results with, have compelled us to generate the spacial data that fit our requirements at the first stages of our research.

The data consist of a set *space* points $\{\mathbf{r}_\alpha\}_{\alpha=1,\ldots,N}$ randomly generated inside the region determined by

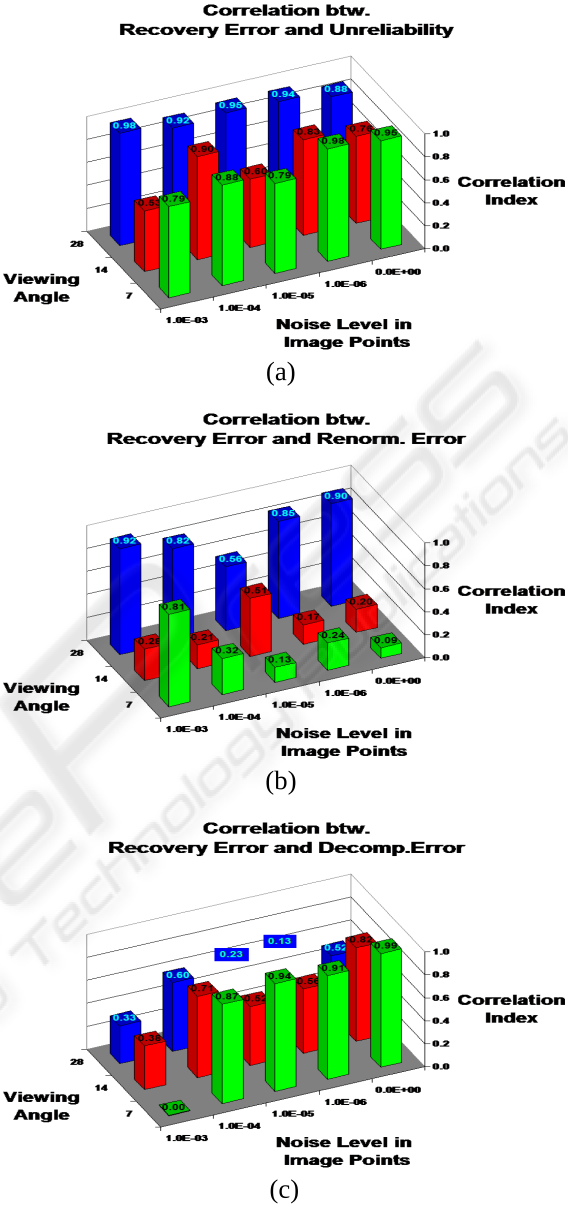


Figure 3: Experiment results (I). Correlation indices I_ρ between recovery error $E_{\mathbf{G}}$ and (a) unreliability $\nu_{\mathbf{G}}$, (b) renormalization error ε_{ren}, and (c) decomposition error ε_{dec}.

two concentric spheres of radius R_{max} and R_{min}, respectively, being $R_{min} = k_1 \cdot R_{max}$. M camera localizations were computed in a circular trajectory around the spheres at a distance $D_{cam} = k_2 \cdot R_{max}$, separated by intervals of γ_t degrees. The orientation of the camera plane is orthogonal to the radial direction.

The set of *image* points for each camera position was generated projecting space points by means of the perspective camera model and adding two kinds of perturbations afterwards. First, the camera viewing angle $\gamma_v \in \{7^o, 14^o, 28^o\}$ permits to limit the positions observing the same common points. Second, an amount of noise $\varepsilon_i \in \{0.0, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}\}$ was also added to image points to simulate the error appearing in the process of point extraction.

Constants used in the settings of each experiment are $N = 500$, $k_1 = 90\%$, $k_2 = 3$, $\gamma = 5^o$, and $M = 360^o/5^o = 72$, respectively. In total, there were $\#\{\gamma_v\} \cdot \#\{\varepsilon_i\} = 3 \cdot 5 = 15$ sets of experiments, where image points were affected by different noise levels and viewing angles while both 3D points and camera localization remained constant (ground truth).

5.2 Unreliability vs. Recovery Error

In order to evaluate the suitableness of the unreliability measure $\nu_{\mathbf{G}}$ defined in Sec. 4, for each of the previous data sets, the correlation between $\nu_{\mathbf{G}}$ and the recovery error $E_{\mathbf{G}}$ was computed, along with the renormalization error ε_r and the decomposition error ε_d, defined in Eq. (4) and Eq. (5), respectively.

Each *feasible*[2] pair of camera locations form an edge e_{ij} in CDG. Once the corresponding essential matrix $\mathbf{G}_{ij}$ for this edge and its decomposition into corresponding $\mathbf{t}_{ij}$ and $\mathbf{R}_{ij}$ were obtained using the algorithm in Sect.4, the values $\nu_{\mathbf{G}_{ij}}$, $\varepsilon_{r_{ij}}$, and $\varepsilon_{d_{ij}}$ were computed, as well as the recovery error $E_{ij} = \|\bar{\mathbf{G}}_{ij} - \mathbf{G}_{ij}\|$. Notice that $\mathbf{G}_{ij} = \mathbf{t}_{ij} \times \mathbf{R}_{ij}$.

For any viewing angle γ_v and image noise ε_i an *index of correlation* I_ρ was computed between error $E_{\mathbf{G}}$ and each one of the previous accuracy measures – $\nu_{\mathbf{G}}$, ε_r, and ε_d – as the *mean* value of all the partial correlations $\{\rho_i, i = 1, \ldots, N\}$ obtained as follows. A correlation ρ_i is calculated by taking the node n_i as the *origin* and employing the corresponding recovery error and the accuracy measures to the rest of nodes n_j, $i \neq j$, to compute a correlation coefficient. That is, if $X_i \in \{E_{\mathbf{G}_i}\}$ and $Y_i \in \{\nu_{\mathbf{G}_i}, \varepsilon_{r_i}, \varepsilon_{d_i}\}$, where $E_{\mathbf{G}_i} = \{E_{\mathbf{G}_{ij}}\}$, $\nu_{\mathbf{G}_i} = \{\nu_{\mathbf{G}_{ij}}\}$, $\varepsilon_{r_i} = \{\varepsilon_{r_{ij}}\}$, and $\varepsilon_{d_i} = \{\varepsilon_{d_{ij}}\}$ with $j = 1, \ldots, N$, then $\rho_i = \rho(X_i, Y_i) = cov(X_i, Y_i)/(\sigma_{X_i} \cdot \sigma_{Y_i})$. Indices I_ρ as a function of γ_v and ε_i are plotted in Fig. 3.

The results show that the correlation index between the unreliability $\nu_{\mathbf{G}}$ and the recovery error $E_{\mathbf{G}}$ is higher than either ε_r or ε_d alone. Besides, the values of this index is pretty high and stable against noise in the image plane and variations in the viewing angle. Moreover, a reciprocal behaviour of ε_r and ε_d is exhibited, that is, ε_r presents a higher correlation when that of ε_d is lower, and vice versa. As a consequence of such results, we state that the unreliability measure $\nu_{\mathbf{G}}$ defined in Eq. (6) is a useful and robust indirect estimate of the recovery error $E_{\mathbf{G}}$ in general.

[2]Sharing enough observable common points.

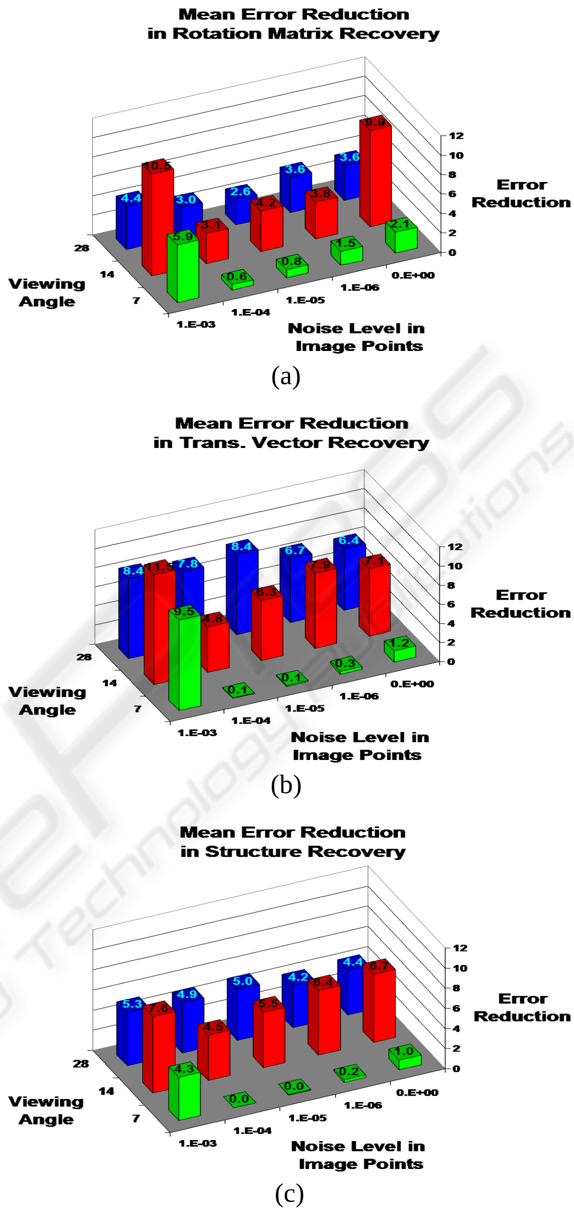
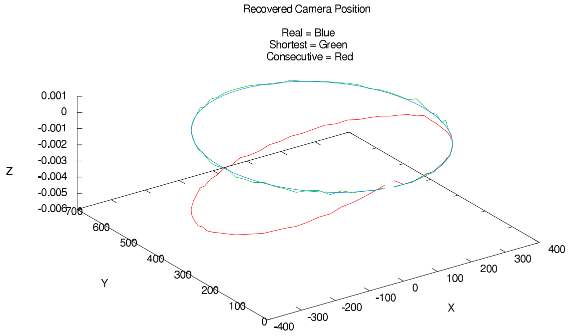


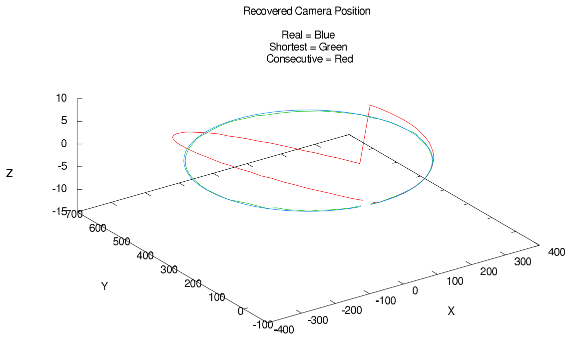
Figure 4: Experiment results (II). Mean error reduction between consecutive and the shortest paths of (a) rotation, (b) translation, and (c) 3D reconstruction, respectively.

5.3 Shortest vs. Consecutive Paths

The objective of this section is to demonstrate the suitability of employing the shortest path in a CDG based on the unreliability $\nu_{\mathbf{G}}$ for recovering the movement parameters corresponding to the camera locations along the aforementioned circular trajectory as well as the 3D space positions of the sets of image points. Apart from the shortest path of unreliabilities $\nu_{\mathbf{G}}$ between essential matrices, the more usual path of *consecutive* camera locations was also taken

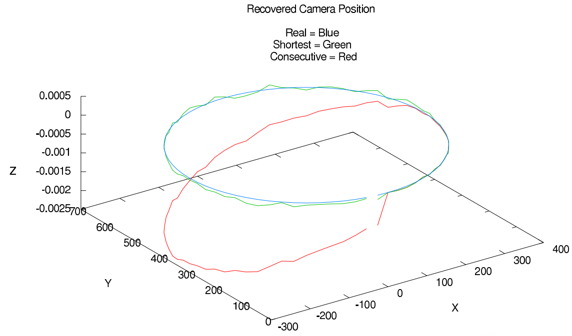


(a)

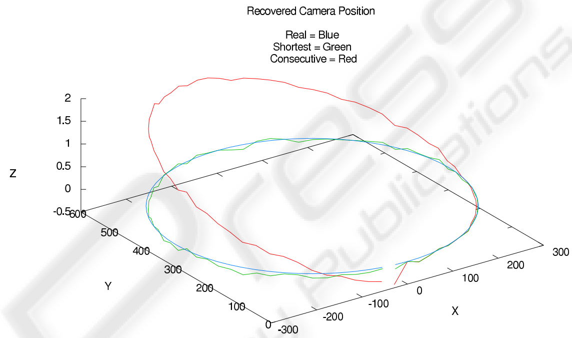


(b)

Figure 5: Some recovery results (I). Camera location in case $\gamma_v = 14^o$ and (a) $\varepsilon_i = 10^{-6}$, and (b) $\varepsilon_i = 10^{-3}$, respectively. Blue lines are the true camera trajectories, while green lines are the camera trajectory recovered using shortest paths, and the red ones corresponds to the camera trajectory recovered using consecutive paths. Units are pixels.



(a)



(b)

Figure 6: Some recovery Results (II). Camera location in case $\gamma_v = 28^o$ and (a) $\varepsilon_i = 10^{-6}$, and (b) $\varepsilon_i = 10^{-3}$, respectively. Blue lines are the true camera trajectories, while green lines are the camera trajectory recovered using shortest paths, and the red ones corresponds to the camera trajectory recovered using consecutive paths. Units are pixels.

into account in such a task. The error between the obtained results and the actual ground truth was calculated afterwards in order to make comparisons. In other words, the translation error $E_{\mathbf{t}}$, rotation error $E_{\mathbf{R}}$, and space points error $E_{\mathbf{r}}$ are calculated.

This process took at every step a location as the origin and computed the set of paths to the rest of them. The procedure was repeated then varying the origin to cover all possible camera locations and the mean values for all the previous error magnitudes were computed. Fig. 4 pictures the reduction in the amount of error when the shortest path to recover the motion parameters was used instead of the consecutive path. The error reduction was obtained dividing the error of a consecutive path and of a shortest path connecting the same origin and final locations.

The use of the shortest path definitely reduced the total amount error in the calculations of both camera movements ($\mathbf{t}$ and $\mathbf{R}$) and 3D structure, especially when γ_v was wider. This is because it is possible in that case to find paths which jump to more separate locations, providing as a consequence larger baselines that increases the accuracy. In case of narrower γ_v, this advantage may not exist and a worse result may appear in few cases due to some outlier locations, as it happens when $\gamma_v = 7^o$ and $\varepsilon_i \in \{0.0, 10^{-6}, 10^{-5}, 10^{-4}\}$. On the other hand, if $\gamma_v = 7^o$ and $\varepsilon_i = 10^{-3}$, the reduction is very big because the consecutive path provided a very poor result.

In Fig. 5 and Fig. 6 we plot some results depicting the shape of the actual camera trajectory along with the two kinds of trajectories recovered using the shortest paths and the consecutive paths. In both groups of plots, we selected two instances corresponding to two levels of image point noise, i.e., $\varepsilon_i = 10^{-6}$ and 10^{-3}. Due to the obvious space limitations it is in fact impossible to show all the results obtained for all the possible combinations of viewing angle and amount of error in image points. Hence, only these two examples of reconstructed trajectories have been selected to illustrate the performance of our approach.

Therefore, our aim is to display that, first, the trajectories recovered using the shortest–path strategy were substantially closer to the real one, and, second, how the error accumulated by the consecutive path growing as successive camera locations went farther from the origin[3]. Moreover, it can be observed that such an error grew accordingly to the amount of noise added to the image points. So in Fig. 5 (a) the scale of the error is smaller than that in Fig. 5 (b). The same can be stated from Fig. 6 (a) and Fig. 6 (b).

It is evident that the recovered trajectory deviates from the true one as the location goes farther from the origin. The scale in coordinates Z is not the same as in coordinates X and Y in order to show such deviation and can be seen as the error in this direction since the true value is $Z = 0.0$. Errors in directions X and Y are more difficult to plot here because their sizes is smaller compared to the range of these coordinates.

Finally, while the order of the errors produced by consecutive paths in coordinates Z were around 10^{-3} in Fig. 5 (a) and Fig. 6 (a), it was considerably larger in Fig. 5 (b) and Fig. 6 (b), i.e., between 1 and 10, which is around one thousand times bigger. This is similar to the order differences in the noise level present in the image points existing in these figures. Consequently, whereas the error in the consecutive–path trajectories had the same order as the image points noise, the error of the shorted–path trajectories was far smaller as can be seen in the depicted examples, being these trajectories really close to the ground truth. Moreover, the viewing angle γ_v also reduces the error of the recovered trajectories nearly to one half when $\gamma_v = 28^o$ with respect to the case of $\gamma_v = 14^o$.

6 CONCLUSIONS

We presented in this paper a new method for multiple view reconstruction based on the definition of an *unreliability* measure that is shown to indirectly estimate the recovery error. Experiments exhibited a clear correlation between our criterion and the error in the estimation of the motion parameters provided by the essential matrix computation and decomposition into translation and rotation. In addition to this, the concept of *Camera–Dependency Graph* (CDG) was introduced consisting of a graph where nodes represents camera positions and edges the feasibility of computing an essential matrix between such locations.

By employing a CDG whose weights are composed of the unreliability measures we could obtain a better result for the motion parameters estimation whenever the shortest paths in the CDG were employed rather than the usual paths of consecutive camera locations. It was proven that the reduction in the recovery error was larger in the case of using shortest–path trajectories than using consecutive paths. Besides, it was also shown by some examples how the better performance of our approach can be appreciated in the precision of the recovered trajectories.

This method can be used in applications that involve dense sequences of images, like those from autonomous robot navigation, estimation of camera trajectories or relative position, as well as for 3D point recovery. The future work will consist in applying this approach to problems such as simultaneous localization and mapping, or robot navigation, as an alternative way to increase the precision of these tasks.

[3] Position $(0.0, 0.0, 0.0)$ in both groups of images.

ACKNOWLEDGEMENTS

The research described in this paper has been funded by the Kankenhi No.19700188.

REFERENCES

Csurka, G., Zeller, C., Zhang, Z., and Faugeras, O. (1997). Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*.

Fitzgibbon, A. W. and Zisserman, A. (1998). Automatic camera recovery for closed or open image sequences. In *ECCV, vol. I, pp. 311-326*.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision.* Cambridge University, 2nd edition.

Kanatani, K. (1996). *Statistical optimization for geometric computation: Theory and Practice.* North-Holland.

Kanatani, K. (2000). Optimal fundamental matrix computation: Algorithm and reliability analysis. In *Proc. 6th Symp. Sensing via Image Inf.*

Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature, 293*, pages 133–135.

Luong, Q.-T. and Faugeras, O. (1996). The fundamental matrix: theory, algorithms, and stability analysis. *Int. J. Comput. Vision, 17-3*.

Martinec, D. and Pajdla, T. (2006). 3d reconstruction by gluing pair-wise euclidean reconstructions. In *3DPVT*.

Martinec, D. and Pajdla, T. (San Diego, USA, 2005). 3d reconstruction by fitting low-rank matrices with missing data. In *CVPR, vol. I, pp. 198-205*.

Schaffalitzky, F. and Zisserman, A. (2002). Multiview matching for unordered image sets. In *ECCV*.

Sturm, P. and Triggs, B. (1996). A factorization based algorithm for multi-image projective structure and motion. In *ECCV96(II), pp. 709-720*.

Tomasi, C. and Kanade, T. (Nov. 1992). Shape and motion from image streams under orthography: a factorization method. In *IJCV, 9(2): 134-154*.