부드러운 카메라 움직임을 위한 EM 알고리듬을 이용한 삼차원 보정

(Structure and Motion Estimation with Expectation Maximization and Extended Kalman Smoother for Continuous Image Sequences)

서 용 덕 † 홍 기 상 ††
(Yongduek Seo) (Ki-Sang Hong)

요 약 이 논문은 카메라가 연속적으로 움직일 때 그 카메라로부터 얻은 동영상을 분석하여 카메라의 움직임에 대한 정보와 영상내의 구조물의 삼차원 정보를 계산하는 알고리듬에 대한 것이다. 일반적으로 불연속한 위치에서 얻은 영상의 집합으로부터 삼차원정보 및 카메라 정보를 얻는 경우에는 카메라의 움직임에 대한 제약조건이 필요 없지만, 비디오 카메라를 이용하여 동영상을 취득하는 경우에는 항상 카메라의 움직임이 부드러워야 한다는 조건이 따라 붙는다. 따라서, 이 논문에서는 '부드러운 움직임을 가지는 카메라'라는 제약조건을 포함하는 카메라 및 삼차원정보의 최적화 과정에 대하여 연구하였다. 목적하는 바를 얻기 위하여 Expectation-Maximization 방법을 사용하여 카메라의 움직임에 대한 모델 파라메터를 동시에 추정하였는데, 이를 위하여 Extended Kalman Filter 와 Extended Kalman Smoother를 적용하였다. 이 연구는 길이가 긴 비디오 영상열의 비젼 해석에 기본이 된다. 실제 영상을 이용하여 실험한 결과를 보였다.

키워드 : Expectation-Maximization, Kalman Filter, 삼차원 복원, 동영상 분석

Abstract This paper deals with the problem of estimating structure and motion from long continuous image sequences, applying the Expectation Maximization algorithm based on extended Kalman smoother to impose the time-continuity of the motion parameters. By repeatedly estimating the state transition matrix of the dynamic equation and the parameters of noise processes in the dynamic and measurement equations, this optimization gives the maximum likelihood estimates of the motion and structure parameters. Practically, this research is essential for dealing with a long video-rate image sequence with partially unknown system equation and noise. The algorithm is implemented and tested for a real image sequence.

Key words : Expectation-Maximization, Kalman Filter, Kalman smoother, structure and motion estimation, motion constrained optimization

1. Introduction

The problem of structure and motion computation has been a major subject of computer vision, and nowadays it becomes common to deal with a very long sequence of images captured in video rate in the applications like structure recovery from real video and video augmentation with graphic objects [1-7]. This tendency suggests that we need to pay more attention to recursive estimation algorithms, as mentioned in [8].

The research on the off-line structure and motion computation from image sequences has been devoted to the development of linear (initial) reconstruction algorithms with self-calibration methods followed by a bundle optimization. We give only a few recent papers among the abundance[9-16]. However, for the processing of (video-rate) sequ-

† 정 회 원 : 서강대학교 미디어공학과 교수
yndk@sogang.ac.kr
†† 종신회원 : 포항공과대학교 전자전기공학과 교수
honks@postech.ac.kr
논문접수 : 2003년 5월 9일
심사완료 : 2004년 10월 21일

ence of images, it is important to have a recursive estimation technique because of the size of the input data and parameters, and therefore various types of the Kalman filter (KF) have been adopted in the recursive computation studies[17-20].

An advantage of recursive estimation by KF over the batch bundle adjustment is that it can incorporate the smoothness (adjacency in time) of the motion parameters during the optimization, but on the contrary one needs to know the dynamic equation and the noise covariance of the state transition and measurement equations. In practice, the dynamics of the motion parameters are assumed to be constant velocity or constant acceleration, and the measurement function becomes nonlinear when the camera is assumed to be a pinhole projection one, which leads to the application of the Extended Kalman filter(EKF). For the application of EKF, the noise covariances and the initial states of the motion parameters with their covariances should be known {\em a priori}; usually, the values of noise covariances are set from the experience and the initial values are computed from a linear reconstruction of a few images sampled. In addition, a weak point of the EKF is that it proceeds only forward in time to estimate the parameters even in the off-line processing.

Hence, in order to overcome those problems and develop a systematic recursive optimization algorithm, we adopt the (fixed-interval) extended Kalman smoother (EKS) which entails forward and backward filtering over the whole image sequence. The forward filtering has the same procedure as EKF - estimation of the structure and motion parameters (SaM parameters) and their covariances, given initial states and covariances; The backward smoothing improves the output of the EKF by taking the {\em a posteriori} measurements additionally in the estimation of current states [21,22]. In particular, our EKS is called Rauch-Tung-Striebel (RTS) smoother[23].

Finally, we apply Expectation Maximization (EM) algorithm together with EKS to the update of the initial values of the SaM parameters, system noise covariance, measurement noise covariance, and the state transition matrix. The EM algorithm is a method for finding a mode of the proposed likelihood function through the iteration of expectation and maximization, and the most remarkable attribute of the EM algorithm is that it ensures an increase in the likelihood function at each iteration. Details can be found in [24,25]. A method of EM for learning (estimating) a linear dynamic system has been studied in [26]. Its extention to the system of linear dynamics and nonlinear measurement are given in [27] with experiments on robot arm mapping and classification with medical data. More general approaches can be found in [28].

Section 2 gives the problem formulation with the introduction to the EM algorithm with EKS. Section 3 shows how to compute the initial values for the optimization before we start EM iteration. Experimental results are given in Section 4 and the conclusion in Section 5. A version of this paper was reported in [29].

2. Notations and Problem Formulation

The coordinates of ith image point at time t in the image I_t is denoted by $m_{ti}=[x_{ti}, y_{ti}]^T$ and corresponding 3D coordinates $M_i=[X_i, Y_i, Z_i]^T$. The projection of M_i to m_{ti} is given by perspective projection through the 3×4 camera matrix C_t :

$$x_{ti}=\frac{C_t^{1,1}X_i+C_t^{1,2}Y_i+C_t^{1,3}Z_i+C_t^{1,4}}{C_t^{3,1}X_i+C_t^{3,2}Y_i+C_t^{3,3}Z_i+C_t^{3,4}},\quad y_{ti}=\frac{C_t^{2,1}X_i+C_t^{2,2}Y_i+C_t^{2,3}Z_i+C_t^{2,4}}{C_t^{3,1}X_i+C_t^{3,2}Y_i+C_t^{3,3}Z_i+C_t^{3,4}} \tag{1}$$

The camera matrix C_t is given by

$$C_t=K_t[R_t|T_t], \tag{2}$$

where K_t is the 3×3 calibration matrix, R_t the 3 ×3 rotation matrix, T_t the 3D translation vector of the camera at t. The calibration matrix is of the form

$$K_t=\begin{bmatrix} f_t & s & u_t \\ 0 & f_t & v_t \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

and, in this paper, we consider the calibration matrix of zero skew ($s=0$), without loss of generality.

2.2 Maximum Likelihood Estimation

Usual optimization algorithms like bundle adjustment assume that measured image coordinates z_{ti} are contaminated by zero-mean Gaussian noise η with 2×2 covariance matrix R:

$$z_{ti} = m_{ti} + \eta \tag{4}$$

The solution θ^* minimizes the probability density function of total projection error

$$p(z_{1:N}|R) = \prod_{t=1}^{N}\prod_{i=1}^{M} p(z_{ti}|R), \tag{5}$$

where $z_{1:N}$ is the vector of all the measurements.

Since we assume a smooth change of the camera parameters, the dynamic equation is modeled by a discrete linear system:

$$\theta_t = F\theta_{t-1} + q_t \tag{6}$$

where F is a time-invariant state transition matrix, q_t process noise of zero-mean Gaussian with covariance Q. Compared with the previous one, the likelihood of this system is denoted by

$$p(z_{1:N}|R, Q, F) \tag{7}$$

We call $\psi = \{R, Q, F, \theta_0, P_0\}$ the system parameters. When there is no assumption on the camera parameters, the measurement covariance R need not be known for the optimization as long as it is Gaussian because the log-likelihood for the optimization can be defined up to scale [30]. Even in the case of R, Q, and F being known, the optimization can be accomplished through the application of EKF. In practice, those parameters are set empirically from the experiments. In this paper, however, we do not rely on such empirical methods but apply EM algorithm to estimate the system parameters as well as the SaM parameters.

2.1 Extended Kalman Smoother

Eq. (6) together with initial SaM parameters θ_0 and the covariance P_0 defines the dynamic equation of our formulation. The measurement equation is given by

$$z_t = [x_{t1}, y_{t1}, \ldots, x_{tM}, y_{tM}]^T = h(\theta_t) + \eta_t \tag{8}$$

where $h(\cdot)$ is the $2M$ vector of corresponding non-linear equations of the form of Eq. (1).

Then, the forward filtering equations of the EKS are:

$$\begin{aligned}
\hat{\theta}_{t+1|t} &= F\hat{\theta}_t \\
P_{t+1|t} &= FP_tF^T + Q \\
K_{t+1} &= P_{t+1|t}H_{t+1}^T[R + H_{t+1}P_{t+1|t}H_{t+1}^T]^{-1} \\
\hat{\theta}_{t+1} &= \hat{\theta}_{t+1|t} + K_{t+1}[z_{t+1} - h(\hat{\theta}_{t+1|t})] \\
P_{t+1} &= P_{t+1|t} - K_{t+1}H_{t+1}P_{t+1|t}
\end{aligned}$$

where

$$H_{t+1} = \left.\frac{\partial h}{\partial \theta}\right|_{\hat{\theta}_{t+1|t}} \tag{9}$$

The backward smoothing is done by

$$\begin{aligned}
A_t &= P_tF^TP_{t+1|t}^{-1} \\
\hat{\theta}_{t|N} &= \hat{\theta}_t + A_t[\hat{\theta}_{t+1|N} - F\hat{\theta}_t] \\
P_{t|N} &= P_t + A_t[P_{t+1|N} - P_{t+1|t}]A_t^T \\
P_{t,t-1|N} &= P_tA_{t-1}^T + A_t[P_{t+1,t|N} - FP_t]A_{t-1}^T
\end{aligned}$$

with the initial values

$$\begin{aligned}
\hat{\theta}_{N|N} &= \hat{\theta}_N \\
P_{N|N} &= P_N \\
P_{N,N-1|N} &= [I - K_NH_N^T]FP_{N-1}
\end{aligned} \tag{10}$$

Therefore, the estimates (expected values) of the SaM parameters θ_t and their corresponding covariances P_t and cross-covariances $P_{t,t-1|N}$ can be computed when the system parameters $\psi = \{R, Q, F, \theta_0, P_0\}$ are given.

2.3 Expectation Maximization

Here, we briefly introduce the principle of EM method. The EM minimizes the likelihood function $p(z_{1:N}|\psi)$ by iterating the following two steps:

E-step: Estimate the optimal SaM states $\{\theta_t\}$ given the system parameters ψ.

M-step: Estimate new system parameters ψ that increases the probability $p(z_{1:N}|\psi)$ given the new SaM states.

The logarithm of the likelihood function can be written as

$$\ln p(z_{1:N}|\psi) = \ln p(\theta_{0:N}, z_{1:N}|\psi) - \ln p(\theta_{0:N}|z_{1:N}, \psi), \tag{11}$$

and expectation of the second term on the right hand side under the distribution $p(\theta_{0:N}|z_{1:N}, \psi_k)$ is maximized for ψ_k:

$$E[\ln p(\theta_{0:N}|z_{1:N}, \psi_k)] \geq E[\ln p(\theta_{0:N}|z_{1:N}, \psi)] \tag{12}$$

for any ψ, which is the key principle of the EM algorithm. Hence, the optimization requires the computation of the expectation of the first term of Eq. (11) in order to maximize the log-likelihood at each iteration. That is, the expectation of $\ln p(\theta_{0:N}, z_{1:N}|\psi_k)$ calculates the SaM states given ψ_k (E-step). The maximization finds new ψ_{k+1} that increases the log-likelihood (M-step), which can be done by differentiations of the expected log-likelihood with respect to ϕ. In brief, the maximization is done by the following equations:

$$\begin{aligned} F &= \Lambda\Delta^{-1} \\ Q &= \frac{1}{N}(\Gamma-\Lambda\Delta^{-1}\Gamma^T) \\ R &= \frac{1}{2MN}\sum_{t=1}^{N} Tr[\nu\nu^T + H_t P_{t|N} H_t^T] \end{aligned} \quad (13)$$

with initial value updates $\theta_0 = \hat{\theta}_{0|N}$ and $P_0 = P_{0|N}$, where Tr[] is the trace operator of a matrix, and

$$\begin{aligned} \nu &= z_t - h_t(\hat{\theta}_{t|N}) \\ \Gamma &= \sum_{t=1}^{N}(\hat{\theta}_{t|N}\hat{\theta}_{t|N^T} + P_{t|N}) \\ \Delta &= \sum_{t=1}^{N}(\hat{\theta}_{t-1|N}\hat{\theta}^T_{t-1|N} + P_{t-1|N}) \\ \Lambda &= \sum_{t=1}^{N}(\hat{\theta}_{t|N}\hat{\theta}^T_{t-1|N} + P_{t,t-1|N}) \end{aligned} \quad (14)$$

3. Initial Parameter Values

The focal length f is initially assumed to undergo the motion of constant acceleration with zero-mean Gaussian process noise. The dynamics of three rotation angles $\omega_X, \omega_Y, \omega_Z$, and three translation components t_X, t_Y, t_Z are all assumed to have the motion of constant acceleration with zero-mean Gaussian nose with there own variances. Thus, the camera motion components θ^C of the SaM state vector $\theta = [\theta^C, \theta^S]$ consist of

$$\theta^C = [f, \dot{f}, \ddot{f}, \omega_X, \dot{\omega}_X, \ddot{\omega}_X, \ \dots \ , t_Z, \dot{t}_Z, \ddot{t}_Z]^T \quad (15)$$

and the structure components

$$\theta^S = [X_1, Y_1, Z_1, \ \dots \ , X_M, Y_M, Z_M]^T \quad (16)$$

Hence, the discrete time dynamic equation of any motion component $p = [q, \dot{q}, \ddot{q}]^T$ is described by

$$p_t = F_p p_{t-1} + r^p_t, \quad (17)$$

where the 3×3 state transition matrix F_p and the error covariance $Q_p = E[r^p_t r^{p\,T}_t]$ are

$$F_p = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \quad Q_p = \sigma_p^2 \begin{bmatrix} T^5/20 & T^4/8 & T^3/6 \\ T^4/8 & T^3/3 & T^2/2 \\ T^3/6 & T^2/2 & T \end{bmatrix}, \quad (18)$$

where E is the expectation operator, T the sampling time interval and σ_p^2 the initial noise variance. Note that the maximization is done for each element of the matrix Q_p, not just for σ_p. For F_p, we update the three upper triangular terms in the EM iteration.

Note that the full transition matrix F becomes block diagonal

$$F = diag(F_f, F_{\omega_X}, \ \dots \ , F_{t_Z}, I_{3\times3}, \ \dots \ , I_{3\times3}) \quad (19)$$

since the transition of the structure components is identical. Likewise, the total covariance Q becomes a block-diagonal matrix; however, the elements corresponding to the structure components are zero since they are fixed in time.

Initially, the quantities σ_p are set empirically from the experience, which will be updated during the EM iteration.

The covariance of the structure components are obtained after the initial reconstruction, and the method is described in Section 3.1 with the computation method of initial reconstruction.

As described before, the Gaussian measurement noise is assumed to be the same for all the image coordinates and the variance is given by $R = \rho I$. We will update ρ during the maximization. Again, an initial estimate of ρ is computed in the next section.

3.1 Initial Reconstruction

We use point matches from image sequences for the reconstruction of SaM parameters. The initial reconstruction is done for a few images sampled from the whole sequence including the first three consecutive images. Projective reconstruction is obtained from the algorithm of [31] followed by projective bundle adjustment. Then, the parameters of Euclidean reconstruction are obtained linearly from the assumption of known principal points, skew and aspect ratio, refined by a linear iteration using dual-quadric projection equation as intermediate stage [32], and finally optimized by a Euclidean bundle adjustment minimizing the sum of re-projection errors:

$$J = \frac{1}{2N_I M}\sum_{t=1}^{N_I}\sum_{i=1}^{M}(\tilde{x}^2_{ti} + \tilde{y}^2_{ti}) \quad (20)$$

where N_I is the number of images used in the initial reconstruction and $(\tilde{x}_{ti}, \tilde{y}_{ti})$ is the re-projection error of the ith 3D point M_i in tth image. Then, using the reconstructed 3D coordinates of the image matches, we use the first three camera parameters to compute the initial estimates of the motion parameters:

$$\begin{aligned} \hat{\dot{p}}_t &= \hat{p}_{t+1} - \hat{p}_t \qquad \text{for } t = 0, 1 \\ \hat{\ddot{p}}_0 &= \hat{\dot{p}}_1 - \hat{\dot{p}}_0 \end{aligned} \quad (21)$$

The approximation of the covariance matrix of $M_i=[X_i, Y_i, Z_i]$ is given from the Jacobian and Hessian of J with the help of the implicit function theorem:

$$P_{M_i}=DRD^T, \quad D=-H^{-1}\frac{\partial g}{\partial z}, \quad H=\frac{\partial g}{\partial M_i}, \quad (22)$$

where $g=\partial J/\partial M_i$ [33]. Finally, the initial estimate of the measurement variance R is given approximately by $R=J(\hat{\theta})$ [30].

4. Experiments

We implemented our algorithm in MATLAB and tested on an image sequence of 399 frames taken from a real NTSC video. Because of interleaving in the original video frames, we divided each frame into two fields and made use of only even fields in the experiments. Thus, the sampling period is $T=1/30$ and the aspect ratio $\gamma=2$. Figure 1 shows the first and the last field images whose original size is 720×243. The lines of the black rectangles in the image were tracked and the corner points were extracted as the intersections of the corresponding lines. The geometric structure of the rectangles were utilized only in the corner point extraction and not used elsewhere.

Initial Euclidean reconstruction was computed as written in Section 3.1; the first three consecutive images and four other images sampled among the sequence ($N_I=7$, $M=32$) were given to projective reconstruction followed by flexible auto-calibration algorithms. Initial RMS re-projection error was approximately 0.23 pixel.

During the EM iteration, we checked the change of total RMS values, and the iteration stopped when the change after smoothing was small enough. Figure 2 shows the evolution of the two

Figure 1 The first and the last field images. The image scale has changed for display

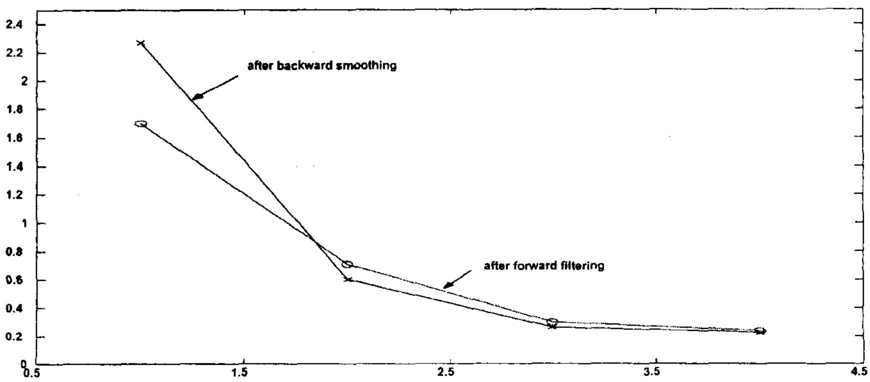

Figure 2 Total RMS's after forward filtering and backward smoothing. The graphs converged to 0.237 pixels for forward filtering and 0.224 pixels for backward smoothing, respectively

total RMS values (RMS after forward filtering and RMS after backward smoothing), which finished after four iterations.

Figure 3 shows two plots of RMS versus time resulted from the first and last EM iterations, respectively. Global RMS range has decreased apparently. On the left plot, the result of backward smoothing shows higher RMS than forward filtering, which was caused by the initial approximation of ψ and partly by the smoothablity of the Kalman smoother; During the forward filtering, the structure parameters were updated at each time step as well, which resulted in the reduction of the reprojection error in each time step. However, the state transition for the structure components are fixed as the identity matrix and corresponding process noise is assumed to be zero, because physically the structure is not moving. Hence through the backward smoothing, the structure parameters become not to be updated since the last updated structure parameters are the best estimates according to the *smoothability condition* [21]. Figure 4 shows this phenomenon. Notice the different behaviors of the changes of the X coordinate through the forward and backward recursion. Its variance shows the same pattern (right plot).

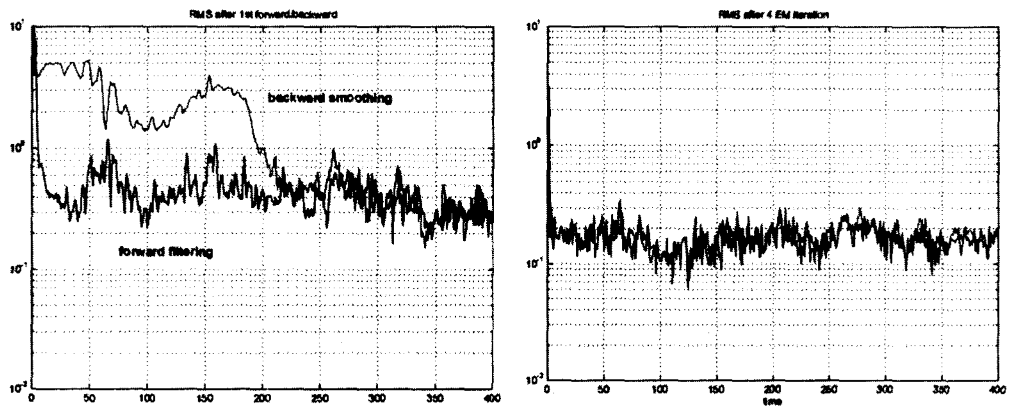

Figure 3 Plots of RMS versus time in log scale. Two plots are drawn in the same ordinate scale. Left: the result after the first forward and backward recursions. Right: after the final EM iteration. Notice that the global error range has decreased after EM iteration

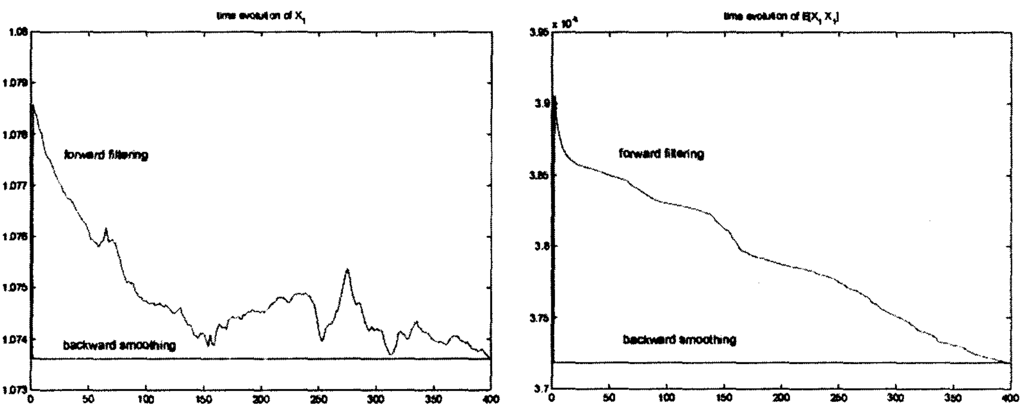

Figure 4 Left: Evolution of the coordinate X_1 through forward and backward recursion during the last EM iteration. Forward filtering updates the value through time, but Kalman smoother does not change it. Right: Evolution of the variance of X_1

Figure 5 shows the change of the focal length before and after the EM iteration. At the first forward/backward estimation, the estimates have rather large differences (up-left) with similar plot of variances (down-left). After the EM iteration, the estimates of focal length settled down (top-right) and their variances were also decreased.

Figures 6 and 7 show the estimates of rotation angles and translation components before and after EM iteration. The graphs of forward filtering in the left plots are not seen well in the right plot because of the overlapping with the graphs of backward smoothing. We can see that, after EM iteration, the estimates become more smooth and stable.

5. Conclusion and Future Works

We incorporated the smoothness constraint of the camera motion parameters into the computation of the structure and motion of a long image sequence captured in video rate. The application of extended Kalman smoother enabled us to impose the smoothness constraint in the optimization, and by the EM algorithm we could repeatedly update the noise covariances and state transition matrix of the extended Kalman smoother systematically. Our formulation resulted in accurate smooth estimation of motion and structure parameters by overcoming the weak-points of the bundle adjustment and the recursive estimation based on Kalman filtering, at the same time.

Our algorithm has considered the case where the input point matches are all given from the image sequence. Future extensions are 1) incorporating

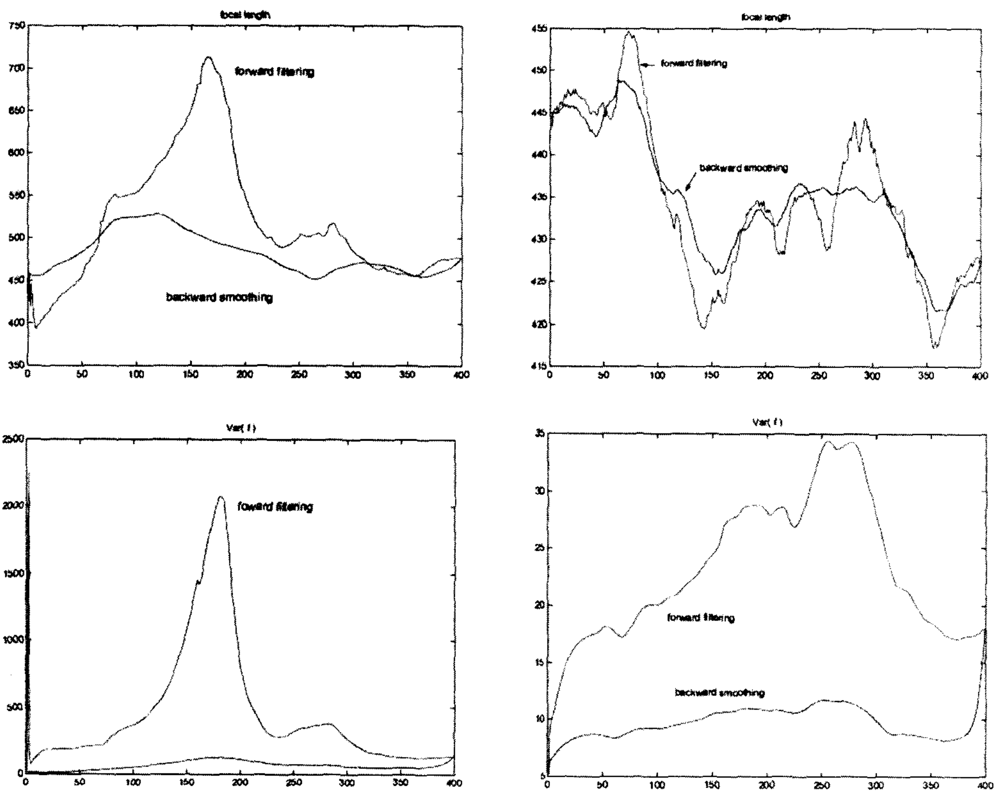

Figure 5 Change of the focal length through time. UP: Left results from the first EM iteration, and right from the last EM iteration. DOWN: Corresponding variances; Notice the different scales of ordinates in the plots. Initial forward filtering induced a large variance

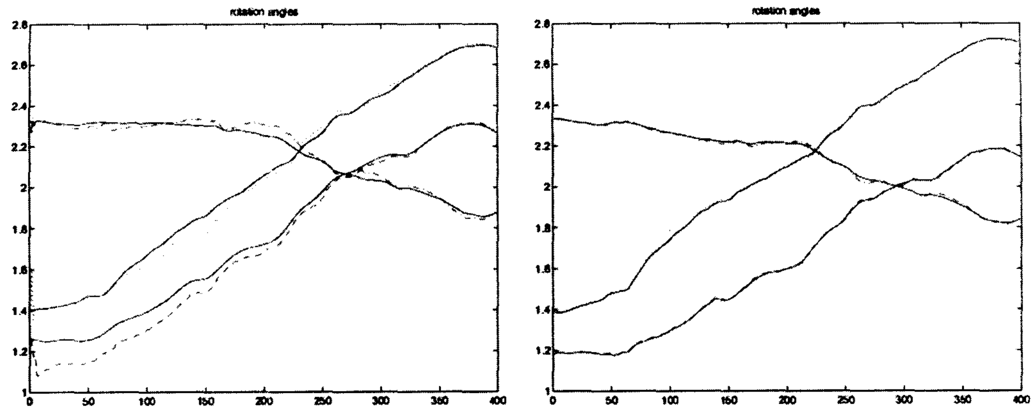

Figure 6 Estimates of rotation angles. Left: After first forward/backward recursion before the application of the EM iteration. Right: Results after the four EM iteration

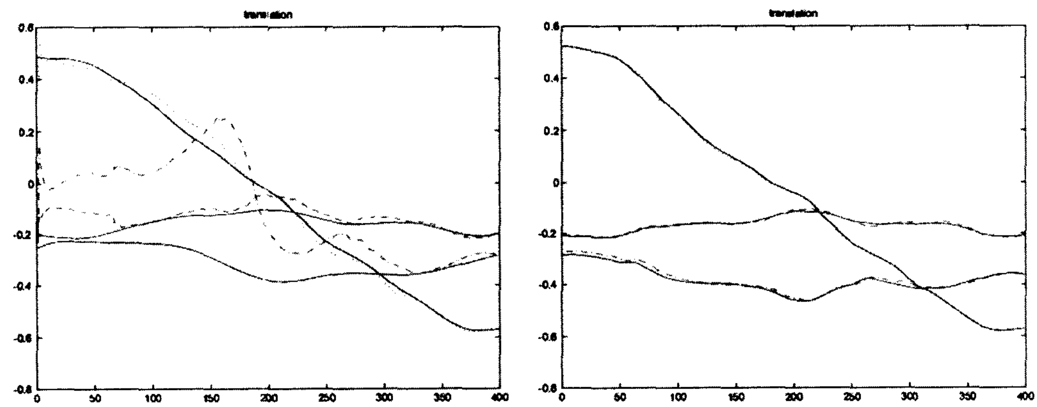

Figure 7 Estimates of three translation components. Left: After first forward/backward optimization before EM. Right: Results after the four EM iteration

line matches in the EM formulation, 2) dealing with the case that the input matches are partially known, and 3) establishing a formulation for a real-time implementation.

References

[1] D. Nister. Reconstruction from uncalibrated sequences with a hierachy of trifocal tensors. In Proc. European Conf. on Computer Vision, 2000.

[2] C.Matsunaga and K.Kanatani. Calibration of a moving camera using a planar patter: Optimal computation, reliability evaluation, and stabilization by model selection. In Proc. European Conf. on Computer Vision, 2000.

[3] G. Simon and M.-O. Berger. Registeration with a moving zoom lens camera for augmented reality applications. In Proc. European Conf. on Computer Vision, 2000.

[4] Kurt Cornelis, Marc Pollefeys, and Maarten Vergauwen adn Luc Van Gool. Augmented reality using uncalibrated video sequences. Second Workshop on Structure from Multiple Images of Large Scale Environments, In conjunction with ECCV 2000, Dublin, Ireland, 2000.

[5] G. Simon, A.W. Fitzgibbon, and A. Zisserman. Markerless tracking using planar structures in the scene. In IEEE and ACM International Symposium on Augmented Reality. IEEE Computer Society Press, 2000.

[6] V. Lepetit and M.-O. Berger. Handling occlusion in augmented reality systems: A semi-automatic method. IEEE and ACM International Symposium on Augmented Reality. IEEE Computer Society

Press, 2000.
[7] Y. Seo and K.-S. Hong. Weakly calibrated video-based augmented reality:embedding and rendering through virtual camera. In IEEE and ACM International Symposium on Augmented Reality. IEEE Computer Society Press, 2000.
[8] Philip F. McLauchlan. A batch/reccursive algorithm for 3D scene reconstruction. In Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2000.
[9] B. Triggs, P.F. Mc L auchlan, R.I. Hartley, and A.W. Fitzgibbon. Bundle adjustment - a modern synthesis. In Vision Algorithms: Theory and Practice, Lecture Notes in Computer Science 1883. Springer-Verlag, 1999.
[10] Ezio Malis and Roberto Cipolla. Multi-view constraints between collineations: applicatio to self-calibation from unknown planar structures. Proc. 6th European Conf. on Computer Vision, Dublin, Ireland. Springer-Verlag, 2000.
[11] Paulo Mendonca and Roberto Cipolla. A simple technique for self-calibration. In Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999.
[12] B. Triggs. Autocalibration from planar scenes. In Proc. of European Conf. Computer Vision, 1998.
[13] M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In Proc. 6th Int. Conf. on Computer Vision, Mumbai, India, 1998.
[14] B. Triggs. Autocalibration and the absolute quadric. In Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1997.
[15] A. Heyden and K. A strom. Euclidean reconstruction from constant intrinsic parameters. In ICPR'96, pages 339~343. IEEE Computer Society Press, 1996.
[16] M. Pollefeys, L. Van Gool, and M. Proesmans. Euclidean 3d reconstruction from image sequences with variable focal lengths. In Proc. 4th European Conf. on Computer Vision, Cambridge, UK, pages 31~42, 1996.
[17] P. F. McLauchlan and D. W. Murray. A unifying framework for structure and motion recovery from image sequences. In Proc. 5th Int. Conf. on Computer Vision, MIT, Boston, MA, pages 314~320. IEEE Computer Society Press, Los Alamitos, California, 1995.
[18] A. Azarbayejani and A.P. Pentland. Recursive estimation of motion, struture and focal length. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(6):562-575, 1995.
[19] Paul Beardsley, Andrew Zisserman, and David Murray. Sequential updating of projective and affine structure from motion, Technical Report OUEL 2070/95, Dept. of Engineering Science, University of Oxford, 1995.
[20] N. Cui, J. Weng, and P. Cohen. Recursive-batch estimation of motion and structure from monocular image sequences. Computer Vision, Graphics, and Image Processing, 59(2):154-170, 1993.
[21] A. Gelb, editor. Applied Optimal Estimation. MIT Press, 1974.
[22] P.S. Maybeck, editor. Stochastic models, estimation, and control VOLUME 2. Academic Press, 1979.
[23] H.E. Rauch, E. Tung, and C.T. Striebel. Maximum likelihood estimates of linear dynamic systems. AIAA Journal, 3(8):1445-1450, 1965.
[24] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society Series B, 39:1-38, 1977.
[25] Todd K. Moon. The expectation-maximization algorithm. IEEE Signal Processing Magazine, 13(6):47-60, 1996.
[26] Z.Ghahramani and G.E. Hinton. Parameter estimation for linear dynamical systems. Technical Report CRG-TR-96-2, University of Toronto, 1996.
[27] M.N. Nando de Freitas and Andrew Gee. Dynamic learning with the EM algorithm for neural networks. VLSI Signal Processing Systems, pages 119~131, 2000.
[28] Z. Ghahramani and M.J. Beal. Learning nonlinear dynamical systems using an em algorithm. In M.S. Kearns, S.A. Solla, and D.A. Cohn, editors, Advances in Neural Information Processing Systems, V.11. MIT Press, 1999.
[29] Y. Seo and K.S. Hong. Structure and motion estimation with expectation maximization and extended Kalman smoother for continuous image sequence, In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, Dec. 2001.
[30] K. Kanatani. Statistical Optimization for Geometric Computation: Theory and Practice. Elsevier Science, 1996.
[31] Anders Heyden, Rikard Berthilsson, and Gunnar Sparr. An iterative factorization method for projective structure and motion from image sequences. Image and Vision Computing, 17(5), 1999.
[32] Y. Seo, Anders Heyden, and R. Cipolla. A linear iterative method for auto-calibration using the DAC equation. In Proc. IEEE Conf. Computer Vision and Pattern Recognition, Dec. 2001.
[33] O. Faugeras. Three-Dimensional Computer Vision. MIT Press, Cambridge, Mass, 1993.
[34] A.J. Davison and N. Kita. Sequential localisation and map-building in computer vision and robotics. Second Workshop on Structure from Multiple

Images of Large Scale Environments, In conjunction with ECCV 2000, Dublin, Ireland, 2000.

서 용 덕

1992년 경북대학교 전자전기공학과 학사 1994년 포항공과대학교 대학원 전자전기공학과 석사. 2000년 포항공과대학교 전자전기공학과 박사. 1994년 4-5월 스웨덴 Lund University Mathematical Imaging Group 방문연구원, 2000년 4월-2001년 5월 영국 캠브리지대학 방문연구원. 현재 서강대학교 영상대학원 미디어공학과 교수. 관심분야는 Camera self-calibration, augmented reality, digital geometry processing

홍 기 상

1977년 서울대학교 전자공학사. 1979년 한국과학기술원 전기및전자공학과 석사 1984년 한국과학기술원 전기및전자공학과 박사. 1984년~1989년 한국 에너지연구소연구원. 1988년~1989년 Carnegie-Mellon 대학교 방문교수. 1986년~현재 포항공과대학교 교수. 관심분야는 합성개구레이더 영상처리, 영상처리 및 컴퓨터 시각, 가상현실, 패턴인식 등