VISUAL MOTION TRACKING WITH FULL ADAPTIVE EXTENDED KALMAN FILTER: AN EXPERIMENTAL STUDY ⋆

Vincenzo Lippiello, Bruno Siciliano, and Luigi Villani

PRISMA Lab
Dipartimento di Informatica e Sistemistica
Università degli Studi di Napoli Federico II
Via Claudio 21, 80125 Napoli, Italy
{*vincenzo.lippiello,siciliano,lvillani*}*@unina.it*

Abstract:
An algorithm for real-time estimation of the pose of a moving object of known geometry is considered. The algorithm is based on a discrete-time Extended Kalman Filter which computes the object pose on the basis of visual measurements of the object features. The robustness of the algorithm with respect to measurement noise and modelling errors is improved by considering a full adaptive version of the Extended Kalman Filter. A complete experimental study is presented to test the performance and feasibility of the approach.

Keywords: Pose Estimation, Vision, Motion Tracking, Visual Servoing, Adaptive Extended Kalman Filter.

1. INTRODUCTION

The autonomy of a robotic system operating in unstructured environments can be significantly enhanced if a visual system is adopted to achieve direct measurements of the task in progress.

Since visual measurements are affected by significant noise and disturbances, the accurate estimation of the position and orientation of a target object in motion may be a difficult task. For this reason, the Extended Kalman Filter (EKF) is usually adopted to achieve noise rejection and enhance estimation accuracy (Wang and Wilson, 1992).

If the quality of the camera sensors is good, the illumination of the scene is stable, and the model of the object is accurate, then the use of a classic formulation of the EKF may guarantee satisfactory results. In fact, in the above hypothesis, it is reasonable to assume that the statistics of the state noise and of the observation noise are known a priori and remain constant, as required by the EKF. On the other hand, if one ore more of the above conditions are not verified, it may be convenient to adopt an Adaptive Extended Kalman Filter (AEKF) (Myers and Tapley, 1976). Various formulations of AEKF are available, which address the problem of the real-time adaptation of the statistical parameters of the covariance matrices of the state and observation noise (Jetto *et al.*, 1999; Ficocelli and Janabi-Sharifi, 2001).

In the work (Lippiello *et al.*, 2004), an AEKF has been presented, based on the visual motion estimation algorithm proposed in (Lippiello and Villani, 2003). The adaptive law is inspired to the heuristic approach of (Myers and Tapley, 1976). The novelty of the algorithm mainly concerns

⋆ This work was supported by CNR and MIUR.

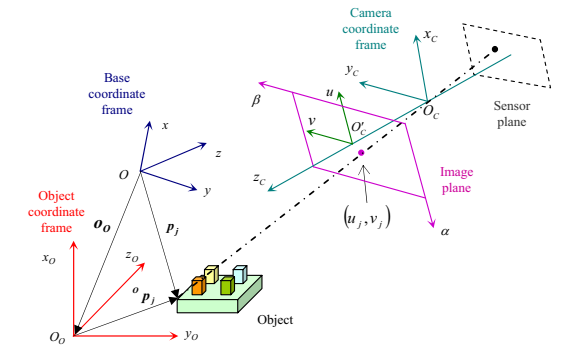

Fig. 1. Pin-hole model of the camera and reference frames

with the adaptive law for the observation noise statistics, which is suitably designed for a visual motion estimation problem based on the use of a variable set of image features.

This paper is aimed at presenting a complete experimental study to demonstrate the effectiveness of the adaptive approach proposed in (Lippiello *et al.*, 2004). A number of case studies are presented, to compare the EKF versus the AEKF with different object trajectories as well as to appreciate the effects of the update laws for the state and observation noise covariance matrices separately.

2. MODELING

Consider the pin-hole model of a video camera fixed with respect to a base coordinate frame O–xyz represented in Fig. 1. Let O_c–$x_c y_c z_c$ be a frame attached to the camera (camera frame), with the z_c-axis aligned to the optical axis and the origin in the optical center. A superscript will be used to denote the reference frame of a variable, whenever different from the base frame.

The sensor plane is parallel to the $x_c y_c$-plane at a distance $-f_e$ along the z_c-axis, where f_e is the effective focal length of the camera lens. The image plane is parallel to the $x_c y_c$-plane at a distance f_e along the z_c-axis. The intersection of the optical axis with the image plane defines the principal optic point O'_c, which is the origin of the image frame O'_c–uv whose axes u and v are taken parallel to the axes x_c and y_c, respectively.

A point P with coordinates ${}^c\boldsymbol{p} = \begin{bmatrix} {}^cx & {}^cy & {}^cz \end{bmatrix}^{\mathrm{T}}$ in the camera frame is projected onto the point of the image plane of coordinates

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{f_e}{{}^cz} \begin{bmatrix} {}^cx \\ {}^cy \end{bmatrix}, \tag{1}$$

that can be expressed in terms of number of pixels as

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} s_u & 0 \\ 0 & s_v \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \tag{2}$$

being $\begin{bmatrix} \alpha_0 & \beta_0 \end{bmatrix}^{\mathrm{T}}$ the coordinates of the point O'_c, whereas s_u and s_v are the row and column scaling factors, respectively.

The position and orientation of a frame O_o–$x_o y_o z_o$ attached to the target object with respect to the base frame can be expressed in terms of the vector of the origin $\boldsymbol{o}_o = \begin{bmatrix} x_o & y_o & z_o \end{bmatrix}^{\mathrm{T}}$ and of the rotation matrix $\boldsymbol{R}_o(\boldsymbol{\phi}_o)$, where $\boldsymbol{\phi}_o = \begin{bmatrix} \varphi_o & \vartheta_o & \psi_o \end{bmatrix}^{\mathrm{T}}$ is the vector of the Roll, Pitch and Yaw angles.

Consider m feature points of the object. The coordinate vector ${}^c\boldsymbol{p}_j$ of the feature point P_j $(j = 1, \ldots, m)$ can be expressed as

$${}^c\boldsymbol{p}_j = \boldsymbol{R}_c^{\mathrm{T}} \left(\boldsymbol{o}_o - \boldsymbol{o}_c + \boldsymbol{R}_o(\boldsymbol{\phi}_o) {}^o\boldsymbol{p}_j \right) \tag{3}$$

where $\boldsymbol{o}_c$ and $\boldsymbol{R}_c$ are, respectively, the position vector and the rotation matrix of the camera frame with respect to the base frame, whereas ${}^o\boldsymbol{p}_j$ is the coordinate vector of P_j expressed in the object frame. Notice that ${}^o\boldsymbol{p}_j$ is a constant vector that is assumed to be known, since it can be computed from a CAD model of the object. Moreover, the quantities $\boldsymbol{o}_c$ and $\boldsymbol{R}_c$ are constant, because the camera is assumed to be fixed to the workspace, and can be computed through a suitable calibration procedure.

By folding the $3m$ equations (3) into (1) and (2), a system of $2m$ nonlinear equations is achieved. The equations depend on the measurements of the m feature points in the image plane of the camera, while the six components of the vectors $\boldsymbol{o}_o$ and $\boldsymbol{\phi}_o$ are the unknown quantities to be estimated. To achieve a solution at least three non-aligned points (six independent equations) are required.

The computation of the solution is nontrivial. The Kalman filter provides a computationally tractable solution, which can also incorporate and exploit redundant measurement information.

3. EXTENDED KALMAN FILTER

In order to estimate the pose of the object, a discrete-time state space model of the object motion has to be considered. The state vector of the model is chosen as the (12×1) vector

$$\boldsymbol{w} = \begin{bmatrix} x_o & \dot{x}_o & y_o & \dot{y}_o & z_o & \dot{z}_o & \varphi_o & \dot{\varphi}_o & \vartheta_o & \dot{\vartheta}_o & \psi_o & \dot{\psi}_o \end{bmatrix}^{\mathrm{T}}. \tag{4}$$

For simplicity, the object velocity is assumed to be constant over one sample period T. The corresponding dynamic modelling error can be considered as an input disturbance $\boldsymbol{\gamma}_k$. The discrete-time dynamic model can be written as

$$\boldsymbol{w}_k = \boldsymbol{A}\boldsymbol{w}_{k-1} + \boldsymbol{\gamma}_k \tag{5}$$

where the state transition matrix $\boldsymbol{A}$ is a constant (12×12) block diagonal matrix of the form

$$\boldsymbol{A} = \mathrm{diag} \left\{ \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \ldots, \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \right\}.$$

The outputs of the Kalman filter are chosen as the vectors of the normalized coordinates of the m feature points in the image plane of the camera

$$\boldsymbol{\zeta}_k^u = \begin{bmatrix} \frac{u_1}{f_e} & \cdots & \frac{u_m}{f_e} \end{bmatrix}_k^{\mathrm{T}} \tag{6a}$$

$$\boldsymbol{\zeta}_k^v = \left[\frac{v_1}{f_e} \;\cdots\; \frac{v_m}{f_e}\right]_k^{\mathrm{T}}. \tag{6b}$$

In view of (1), the corresponding output model can be written in the form

$$\boldsymbol{\zeta}_k^u = \boldsymbol{g}^u(\boldsymbol{w}_k) + \boldsymbol{\nu}_k^u \tag{7a}$$
$$\boldsymbol{\zeta}_k^v = \boldsymbol{g}^v(\boldsymbol{w}_k) + \boldsymbol{\nu}_k^v \tag{7b}$$

where $\boldsymbol{\nu}_k^u$ and $\boldsymbol{\nu}_k^v$ are the observation noise vectors for the u and v components of the normalized image plane, and $\boldsymbol{g}^u(\boldsymbol{w}_k)$, $\boldsymbol{g}^v(\boldsymbol{w}_k)$ are defined as

$$\boldsymbol{g}^u(\boldsymbol{w}_k) = \left[\frac{{}^c x_1}{{}^c z_1} \;\cdots\; \frac{{}^c x_m}{{}^c z_m}\right]_k^{\mathrm{T}} \tag{8a}$$

$$\boldsymbol{g}^v(\boldsymbol{w}_k) = \left[\frac{{}^c y_1}{{}^c z_1} \;\cdots\; \frac{{}^c y_m}{{}^c z_m}\right]_k^{\mathrm{T}}. \tag{8b}$$

The coordinates of the feature points ${}^c\boldsymbol{p}_j$ in (8) are computed from the state vector $\boldsymbol{w}_k$ via (3).

The components of $\boldsymbol{\gamma}_k$, $\boldsymbol{\nu}_k^u$ and $\boldsymbol{\nu}_k^v$ are considered as independent, non-stationary, Gaussian, white noise sequences with the statistical properties

$$\mathrm{E}[\boldsymbol{\gamma}_k] = \boldsymbol{q}_k, \mathrm{E}[\boldsymbol{\nu}_k^u] = \boldsymbol{r}_k^u, \mathrm{E}[\boldsymbol{\nu}_k^v] = \boldsymbol{r}_k^v \tag{9a}$$
$$\mathrm{E}[(\boldsymbol{\gamma}_k - \boldsymbol{q}_k)(\boldsymbol{\gamma}_l - \boldsymbol{q}_l)^{\mathrm{T}}] = \boldsymbol{Q}_k\delta_{kl} \tag{9b}$$
$$\mathrm{E}[(\boldsymbol{\nu}_k^u - \boldsymbol{r}_k^u)(\boldsymbol{\nu}_l^u - \boldsymbol{r}_l^u)] = \boldsymbol{R}_k^u\delta_{kl} \tag{9c}$$
$$\mathrm{E}[(\boldsymbol{\nu}_k^v - \boldsymbol{r}_k^v)(\boldsymbol{\nu}_l^v - \boldsymbol{r}_l^v)] = \boldsymbol{R}_k^v\delta_{kl} \tag{9d}$$

where $\mathrm{E}[\cdot]$ indicates the statistical mean operator and δ is the Kroneker symbol.

Since the output model is nonlinear in the system state, the so-called Extended Kalman Filter (EKF) must be adopted. An iterative formulation of the EKF for the state vector of the system defined by (5) and (7) may be found in (Lippiello and Villani, 2003).

4. ADAPTIVE EXTENDED KALMAN FILTER

If a high-quality camera sensor is used, the illumination of the scene is stable, and the velocity of the tracked object is nearly constant, then it is possible to use constant statistical parameters with optimal results. On the other hand, if these conditions are not satisfied, it may be convenient to update in real time the statistical parameters $\{\boldsymbol{q}_k, \boldsymbol{Q}_k, \boldsymbol{r}_k^u, \boldsymbol{r}_k^v, \boldsymbol{R}_k^u, \boldsymbol{R}_k^v\}$. This leads to the Adaptive Extended Kalman Filter (AEKF) presented in (Lippiello *et al.*, 2004).

The basic hypothesis for this approach is the constant value of the statistical parameters over N sample times (Myers and Tapley, 1976).

Since not all the visual features are always available during the motion and their location into the scene is strongly variable, it may be reasonable to assume the statistics of the observation noise to be equal for all the measurements of the feature points in the scene at time k. Hence the quantities $\{\boldsymbol{r}_k^u, \boldsymbol{r}_k^v, \boldsymbol{R}_k^u, \boldsymbol{R}_k^v\}$ are replaced by the quantities $\{r_k^u\boldsymbol{\imath}_m, r_k^v\boldsymbol{\imath}_m, \sigma_k^{u\,2}\boldsymbol{I}_m, \sigma_k^{v\,2}\boldsymbol{I}_m\}$, where $\boldsymbol{\imath}_m$ indicates an $(m\times 1)$ vector of components equal to 1 and $\boldsymbol{I}_m$ indicates the $(m \times m)$ identity matrix. Moreover, the samples of the observation noise sequences $\boldsymbol{\nu}_i^u$ ($\boldsymbol{\nu}_i^v$) are independent for $i = 1, \ldots, N$ and have a gaussian distribution with mean $r^u\boldsymbol{\imath}_m$ ($r^v\boldsymbol{\imath}_m$) and variance $\sigma^{u\,2}\boldsymbol{I}_m$ ($\sigma^{v\,2}\boldsymbol{I}_m$), where r^u, r^v, σ^u and σ^v are constant over N sample times.

In view of the nonlinear relation (7), an intuitive approximation of the observation noise sample vectors at time k is given by the quantities

$$\boldsymbol{\rho}_k^u = \boldsymbol{\zeta}_k^u - \boldsymbol{g}^u(\boldsymbol{w}_{k,k-1}) \tag{10a}$$
$$\boldsymbol{\rho}_k^v = \boldsymbol{\zeta}_k^v - \boldsymbol{g}^v(\boldsymbol{w}_{k,k-1}) \tag{10b}$$

which can be considered as independent and identically distributed over N samples. It can be shown (Myers and Tapley, 1976) that an unbiased estimator for r^u and r^v can be taken as

$$\hat{r}^u = \frac{1}{N}\sum_{i=1}^{N}\overline{\rho}_i^u, \quad \hat{r}^v = \frac{1}{N}\sum_{i=1}^{N}\overline{\rho}_i^v, \tag{11}$$

where $\overline{\rho}_i^u$ and $\overline{\rho}_i^v$ are scalar quantities equal to the mean values of the components of the vectors $\boldsymbol{\rho}_i^u$ and $\boldsymbol{\rho}_i^v$ respectively. Moreover, an unbiased estimator for $\sigma^{u\,2}$ and $\sigma^{v\,2}$ may be obtained as

$$\hat{\sigma}^{u\,2} = \frac{1}{m(N-1)}\sum_{i=1}^{N}\Big\{\|\boldsymbol{\rho}_i^u - \hat{r}^u\boldsymbol{\imath}_m\|^2 + - \frac{N-1}{N}\mathrm{tr}(\boldsymbol{\Gamma}_i^u)\Big\} \tag{12a}$$

$$\hat{\sigma}^{v\,2} = \frac{1}{m(N-1)}\sum_{i=1}^{N}\Big\{\|\boldsymbol{\rho}_i^v - \hat{r}^v\boldsymbol{\imath}_m\|^2 + - \frac{N-1}{N}\mathrm{tr}(\boldsymbol{\Gamma}_i^v)\Big\} \tag{12b}$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

In view of the linear dynamic state relation at time k given by (5), an intuitive approximation for the state noise vector at time k is

$$\boldsymbol{\varrho}_k = \boldsymbol{w}_k - \boldsymbol{A}\boldsymbol{w}_{k,k-1}, \tag{13}$$

which may be considered independent and identically distributed over N samples. It can be shown that an unbiased estimator for the mean value $\boldsymbol{q}$ of the state noise may be obtained as

$$\hat{\boldsymbol{q}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\varrho}_i, \tag{14}$$

while an unbiased estimator for the covariance matrix $\boldsymbol{Q}$ is given by

$$\hat{\boldsymbol{Q}} = \frac{1}{N-1}\sum_{i=1}^{N}\left\{(\boldsymbol{\varrho}_i - \hat{\boldsymbol{q}})(\boldsymbol{\varrho}_i - \hat{\boldsymbol{q}})^{\mathrm{T}} - \frac{N-1}{N}\boldsymbol{\Delta}_i\right\}, \tag{15}$$

where $\boldsymbol{\Delta}_i = \boldsymbol{A}\boldsymbol{P}_{i,i-1}\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{P}_{i,i}$.

In sum, the equations (10)–(15) provide a heuristic unbiased estimator for the statistical parameters of an EKF used for visual motion estimation, on the assumption that the last N samples are statistically independent and identically distributed. The above algorithm can be formulated in an iterative limited memory format, as reported in (Lippiello *et al.*, 2004).

5. VISUAL ESTIMATION ALGORITHM

The accuracy of the estimate provided by the Kalman filter depends on the number of the available feature points. Inclusion of extra points may improve the estimation accuracy but increases the computational cost. It has been shown that a number of feature points between four and six, if properly chosen, may represent a good trade-off (Wang and Wilson, 1992). Automatic selection algorithms have been developed to find the optimal feature points, see e.g., (Janabi-Sharifi and Wilson, 1997). An efficient selection technique presented in (Lippiello and Villani, 2003) is adopted here, which exploits the prediction of the object pose provided by the Kalman filter to perform a pre-selection of the points that are visible at the next sample time.

6. EXPERIMENTS

The experimental set-up (see Fig. 2) is composed by a PC equipped with a MATROX Genesis board, a SONY 8500CE B/W camera, and a COMAU Smart 3S robot. The MATROX board is used as frame grabber and for a partial image processing (e.g., windows extraction from the image).

The robot is used to move an object in the visual space of the camera; thus the object position and orientation with respect to the base frame of the robot can be computed from joint position measurements via the direct kinematic equation. To test the accuracy of the estimation provided by the Kalman filter, the camera was calibrated with respect to the base frame of the robot.

The camera resolution is 576×763 pixels and the nominal focal length of the lenses is 16 mm. The camera is disposed at a distance of about 130 cm from the object. The sampling time used for estimation is limited by the camera frame rate, which is about 26 fps. A simple neon illumination has been used; hence, during the object motion, the local illumination conditions of the windows of the image plane selected for feature extraction are quite variable due to reflections or shadows.

The image features are the corners of the object. The object used in the experiment has 40 corners, which are all candidate for feature extraction.

Two different case studies are considered. The first one is aimed at comparing the performance of the EKF to that of the AEKF considering

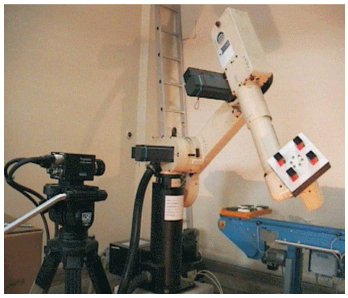

Fig. 2. Experimental setup

different object trajectories. The second one is aimed at evaluating the effects of the update laws for the matrix $\boldsymbol{Q}$ and $\boldsymbol{R}$ separately. Notice that the adoption of the AEKF in lieu of the EKF causes only a modest increase of the computational cost that, in terms of overall processor time, is about 16 percent.

For both EKF and AEKF, the initial value of the matrix $\boldsymbol{P}_{1,0}$ has been chosen as the null matrix; moreover, the initial value of the state vector $\boldsymbol{w}_{1,0}$ have been set null for the velocity components, while the pose components has been roughly estimated through direct measurements. A diagonal covariance matrix $\boldsymbol{Q}$ has been chosen both in the non-adaptive and in the adaptive case.

The values of the statistical parameters used for the EKF are set as initial values for the AEKF; they are: $\hat{r}_0^u = \hat{r}_0^v = 0$, $\hat{\sigma}_0^{u2} = \hat{\sigma}_0^{v2} = 9.0$, $\hat{\boldsymbol{q}}_0 = \boldsymbol{0}$, and $\hat{\boldsymbol{Q}}_0 = \text{diag}\{0,5,0,5,0,5,0,20,0,20,0,20\} \cdot 10^{-6}$. The physical dimensions are: pixel and pixel^2, respectively, for the mean and variances of the observation noise; mm, mm/s, rad and rad/s for the components of the mean of the state noise; mm^2, (mm/s)^2, rad^2 and (rad/s)^2 for the corresponding covariances.

The initial values of the observation noise covariances have been evaluated during the camera calibration procedure while the initial values of the state noise covariances have been set on the basis of the velocity range of the object trajectories. These values have been further tuned on the basis of a set of experiments carried out using the EKF, to achieve satisfactory tracking performance.

Notice that all the elements of the covariance matrix $\hat{\boldsymbol{Q}}_0$ corresponding to the position components of the state have been considered initially zero for the AEKF and constantly zero for the EKF. Fynally, $N_q = N_r = 30$.

6.1 First case study: EKF vs. AEKF

Three different object trajectories are considered.

- **TrajP:** the object position varies according to the time history reported on the left of Fig. 3 while the object orientation is left constant. The norm of the maximum linear velocity is about 10 cm/s.

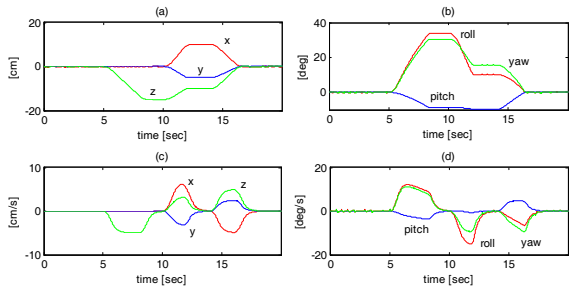

Fig. 3. Trajectory **TrajPO**. Position vector (a), RPY angles (b), linear velocity (c), time derivative of RPY angles (d).

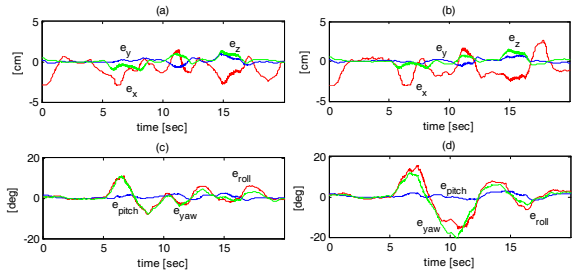

Fig. 4. Estimation errors for **TrajPO**. AEKF: position (a) and orientation (c) errors; EKF: position (b) and orientation (d) errors

- **TrajO:** the object position is left constant while the object orientation varies according to the time history reported on the right of Fig. 3. The norm of the maximum velocity for the RPY angles is about 20 deg/s.
- **TrajPO:** the object position and orientation vary according to the time history of Fig. 3.

Notice the trajectory **TrajPO** is the composition of **TrajP** and **TrajO**, but the resulting trajectories of the feature points on the image plane are different in the three cases. Hence they represent a significant test base to make a comparison.

The results of the experiments are summarized in Table I, where the mean value and the standard deviation of the norm of the position error $\boldsymbol{e}_P$ and orientation error $\boldsymbol{e}_O$ are reported, both in the case of EKF and AEKF. It can be seen that the use of the AEKF allows a general improvement of the tracking performance, especially for the orientation error components.

Table 1. Comparison of the estimation errors in the first case study.

		EKF		AEKF	
		Mean	**StD**	**Mean**	**StD**
TrajP	$\boldsymbol{e}_P$ [mm]	13.71	11.16	12.59	10.89
	$\boldsymbol{e}_O$ [deg]	3.50	3.20	2.29	1.58
TrajO	$\boldsymbol{e}_P$ [mm]	11.23	8.66	9.15	6.13
	$\boldsymbol{e}_O$ [deg]	5.95	5.26	4.78	3.33
TrajPO	$\boldsymbol{e}_P$ [mm]	13.58	8.70	11.92	8.06
	$\boldsymbol{e}_O$ [deg]	7.19	6.86	3.65	3.43

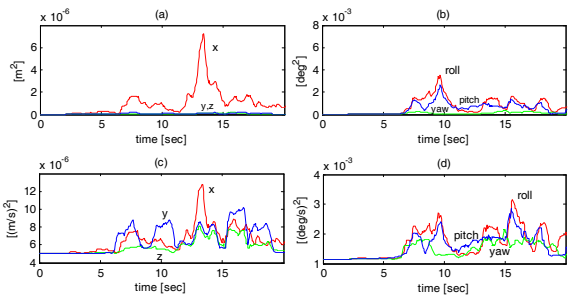

Fig. 5. Elements of the state noise covariance matrix for **TrajPO** in the first case study. Position (a), orientation (b), linear velocity (c), rotational velocity (d).

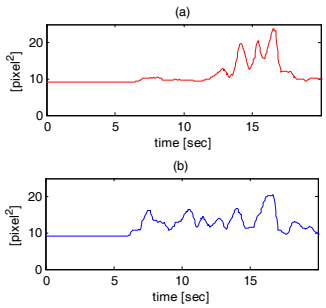

Fig. 6. Observation noise variance for **TrajPO** in the first case study. $\hat{\sigma}^{u2}$ (a), $\hat{\sigma}^{v2}$ (b)

For the trajectory **TrajPO**, the time history of the estimation errors are shown in Fig. 4. It can be seen that the initial values of the errors are the same in the case of the EKF and the AEKF. In particular, there is an initial offset for the x-component of the position error, due to the initial misalignment between the real initial position and the initial value of the Kalman filter. This error is initially recovered by the Kalman filter during the first 5 s, in the absence of motion. During the motion, the position errors keep limited for all the components, but is higher for the x-component. In fact, the x-axis is aligned to the optical axis of the camera, and thus the evaluation of the x-component of the object position, for a single-camera system, is more sensitive to measurements and modelling errors with respect to the position components lying on the image plane (Wang and Wilson, 1992). In general, the peaks on the errors happen when the velocity is higher, due to the modelling error for the EKF. These errors are partially recovered by the AEKF, in view of the adaptive law of the covariance matrix $\boldsymbol{Q}$.

The time histories of some of the statistical parameters which are updated on-line in the AEKF are also reported. In particular, the time histories of the elements of the (diagonal) covariance matrix of the state noise are shown in Fig. 5, while the time history of the observation noise for the u and v components are shown in Fig. 6. It can

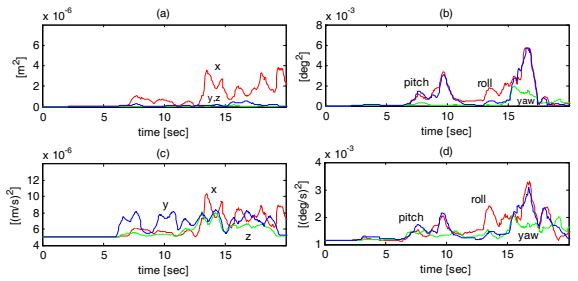

Fig. 7. Elements of the state noise covariance matrix for the AEKF-Q in the second case study. Position (a), orientation (b), linear velocity (c), rotational velocity (d).

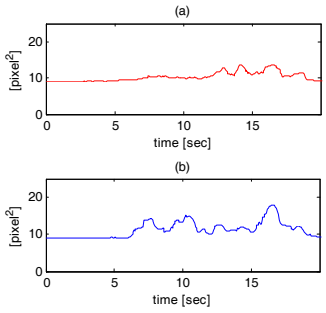

Fig. 8. Observation noise variance for the AEKF-R in the second case study. $\hat{\sigma}^{u2}$ (a), $\hat{\sigma}^{v2}$ (b).

be observed that all the updated parameters keep limited values; moreover, it can be recognized that there exists a correlation between the values of the elements of the $\boldsymbol{Q}$ matrix and the object trajectory. In particular, the pick values of the elements of $\boldsymbol{Q}$ corresponding to the position and orientation components can be related to the pick values of the linear and angular velocity of the object. Analogously, the pick values of the elements of $\boldsymbol{Q}$ corresponding to the linear and angular velocity can be related to the object accelerations.

6.2 Second case study: Updating Q and/or R

In this case study, only the trajectory **TrajPO** has been used. The EKF has been compared to two partial adaptive versions of the EKF: the AEKF-Q, where only matrix $\boldsymbol{Q}$ is updated, and AEKF-R, where only matrix $\boldsymbol{R}$ is updated.

The results are summarized in Table II in terms of the mean value and the standard deviation of the norm of the position error $\boldsymbol{e}_P$ and orientation error $\boldsymbol{e}_O$ (the values for the EKF and AEKF are the same reported in Table I for the trajectory **TrajPO**). It can be seen that, using the AEKF-Q, good results can be achieved with respect to the EKF, but worse than the complete AEKF, especially for the orientation error. On the other hand, the AEKF-R allows improving all the errors except the standard deviation of the mean value of the position error. This means that the main role to guarantee good tracking performance is played by the matrix $\boldsymbol{Q}$; however, a further improvement can be achieved if both $\boldsymbol{Q}$ and $\boldsymbol{R}$ are updated.

Table 2. Comparison of the estimation errors in the second case study.

		Mean	**StD**
EKF	$\boldsymbol{e}_P$ [mm]	13.58	8.70
	$\boldsymbol{e}_O$ [deg]	7.19	6.86
AEKF-Q	$\boldsymbol{e}_P$ [mm]	11.94	8.42
	$\boldsymbol{e}_O$ [deg]	6.00	4.29
AEKF-R	$\boldsymbol{e}_P$ [mm]	12.01	9.64
	$\boldsymbol{e}_O$ [deg]	5.82	5.56
AEKF	$\boldsymbol{e}_P$ [mm]	11.92	8.06
	$\boldsymbol{e}_O$ [deg]	3.65	3.43

The time histories of the elements of the (diagonal) covariance matrix of the state noise for AEKF-Q are shown in Fig. 7 while the time history of the observation noise for AEKF-R are shown in Fig. 8.

7. CONCLUSION

In this paper an experimental study on a full adaptive EKF for visual motion estimation is presented. The experimental results confirm the effectiveness of the full AEKF. In fact, the effects on the estimation error of the modelling error and of the measurement noise are reduced with respect to the non-adaptive formulation, at the expense of a small increase of computational load.

REFERENCES

Ficocelli, M., and F. Janabi-Sharifi (2001). Adaptive filtering for pose estimation in visual servoin. *2001 IEEE/RSJ Int. Conf. on Int. Robots and Systems*, Maui, HI, 19–24.

Janabi-Sharifi, F., and W.J. Wilson (1997). Automatic selection of image features for visual servoing. *IEEE Trans. on Robot. and Automat.*, **13**, 890–903.

Jetto, L., S. Longhi and G. Venturini (1999). Development and experimental validation of an adaptive extended Kalman filter for the localization of a mobile robot. *IEEE Trans. on Robot. and Automat.*, **15**, 219–229.

Lippiello, V., and L. Villani (2003). Managing redundant visual measurements for accurate pose tracking. *Robotica*, **21**, 511-519.

Lippiello, V., B. Siciliano and L. Villani (2004). Visual motion estimation of 3D objects: An adaptive extended Kalman filter approach. *2004 IEEE/RSJ Int. Conf. on Int. Robots and Systems*, Sendai, Japan, 957–962.

Myers, K.A., and B.D. Tapley (1976). Adaptive sequential estimation with unknown noise statistics. *IEEE Trans. on Automat. Contr.*, **21**, 520–523.

Wang, J., and W.J. Wilson (1992). 3D relative position and orientation estimation using Kalman filter for robot control. *1992 IEEE Int. Conf. on Robot. and Automat.*, Nice, France, 2638–2645.