Robust solution to three-dimensional pose estimation using composite extended Kalman observer and Kalman filter

H.D. Taghirad   S.F. Atashzar   M. Shahbazi

Advanced Robotics and Automated Systems (ARAS), Faculty of Electrical and Computer Engineering,
K. N. Toosi University of Technology, Iran
E-mail: Taghirad@kntu.ac.ir

Abstract: Three-dimensional (3D) pose estimation of a rigid object by only one camera has a vital role in visual servoing systems, and extended Kalman filter (EKF) is vastly used for this task in an unstructured environment. In this study, the stability of the EKF-based 3D pose estimators is analysed in detail. The most challenging issue of the state-of-the-art EKF-based 3D pose estimators is the possibility of its divergence because of the measurement and model noises. By analysing the stability of conventional EKF-based pose estimators a composite technique is proposed to guarantee the stability of the procedure. In the proposed technique, the non-linear-uncertain estimation problem is decomposed into a non-linear-certain observation in addition to a linear-uncertain estimation problem. The first part is handled using the extended Kalman observer and the second part is accomplished by a simple Kalman filter. Finally, some experimental and simulation results are given in order to verify the robustness of the method and compare the performance of the proposed method in noisy and uncertain environment to the conventional techniques.

1   Introduction

Estimation of the position and orientation of a rigid object from vision has a crucial role in robotics and computer vision [1]. In robotics, it plays a vital function in visual servoing systems especially position based-eye in hand visual servoing [2, 3]. It is known that at least six non-linear independent equations are required to determine the six DOF relative pose between the camera and the object [4]. In order to solve this set of equations, at least three non-aligned feature points on the object and the geometric equations between the camera frame and the object frame in addition to the two-dimensional (2D) projection of the feature points in the image plane are required [4]. In order to solve 3D pose estimation problem, extended Kalman filter (EKF) was first proposed by Wu [5], and subsequently, enhanced for pose-based visual servoing by Wilson [6]. The EKF has the ability of noise, disturbance and uncertainty rejection. Additionally, it reinforces the accuracy and speed of the estimation process, which is a stringent requirement in online applications [7], and furthermore, by using the prediction ability of the EKF, a dynamic pre-selection technique for projected features in the image plane can be established. This ability can sensibly reduce the elapsed time and failures in image processing and feature extraction procedures in the 3D pose estimators [8, 9]. In order to enhance the accuracy of the EKF-based pose estimators, adaptive extended Kalman filter (AEKF) is proposed in [10] as the base of a predictive 3D

pose estimator. In this technique, an adaptation law is utilised in order to tune the covariance matrices of the filter measurement and the model noises. Another constructive alternative is the iterative extended Kalman filter (IEKF) approach, which is utilised in [11, 12] for 3D pose estimation. Recently, a composite EKF-based algorithm is proposed in [3], which is called iterative adaptive extended Kalman filter (IAEKF), in order to fuse the advantages of AEKF and IEKF for enhancing the stability and performance. Although, using all these approaches (IEKF, AEKF and IAEKF) the stability of the estimation procedure is considerably enhanced; however the convergence of the usual EKF-based techniques is not guaranteed. It should be noted that, despite the radiant effectiveness of the EKF, it is not an optimal filter, opposed to regular Kalman filter. Consequently, linearisation may cause divergence of the estimation error when the assumption of local linearity is not met [12, 13]. This issue is the most challenging problem of the EKF-based pose estimators. Reif *et al*, have shown that, EKF remains stable if the system satisfies the non-linear observability rank conditions as well if the norm of measurement noise and initial guess error are small [14]. However, a thorough stability analysis of the EKF-based pose estimation has not fully discussed in the literature. In most previous studies it is just recommended to select a sufficient high sampling rate for enhancing the accuracy of linearisation and decreasing the diverging effect of the uncertainties [3, 6]. However, in practical applications the sampling time is restricted by the camera frame rate and

image processing required time. In [15] it is shown that the EKF-based pose estimators have high potential of divergence under high speed and rough motions.

In this paper first, the stability of the EKF-based pose estimators is thoroughly studied, and it is shown that, the estimation error are bounded, provided that certain conditions and assumptions are satisfied. These suppositions are related to the level of measurement noise, intensity of disturbance, size of model uncertainties and properness of initial guess. As a consequence it is approved that, if the target has a rough motion, or the vision system failed to find the correct projected features at some instants, or the initial guess is not accurate enough estimation error may become unbounded. To remedy this problem, a new technique is introduced by using a composition of Kalman filter and extended Kalman observer (EKO), by which the estimation error remains bounded in present of measurement noise, faults and process uncertainties. In fact in this technique, the non-linear-uncertain estimation problem is decomposed into a non-linear-certain observation in addition to a linear-uncertain estimation problem. The first part is handled by an EKO and the second part is accomplished by a simple Kalman filter. Furthermore, it is proved that, the proposed method can robustly estimate the positions of rigid bodies with almost no restriction on the process and observation noises. In addition to that, the vision faults can be detected and marked as incorrect measurement data. This feature is considerably functional in practical applications. Finally, the performance of the proposed method is verified by some simulations and experiments.

This paper is organised as follows. In Section 2 the pose estimation problem has been reviewed. In Section 3, the conventional EKF-based pose estimators are discussed. In Section 4 the convergence conditions of conventional EKF-based pose estimators is introduced. Subsequently, Section 5 illustrates details of the proposed method for pose estimation and its stability in an uncertain situation, and Section 6 contains some experimental results.

2 Pose estimation modelling

Camera is a media, which maps 3D reference space into 2D image coordinate (image plane). In field of pose estimation, projected 2D coordinates of some feature points, attached on the object, is used to calculate the 3D coordinates of the object represented in the reference frame. The case where the camera is mounted on the robot's end-effectors, which is called an eye in hand configuration, is considered in this paper. In this case, reference frame is attached on the camera. Each 3D pre-known feature point $p_i^o = (x_i^o, y_i^o, z_i^o); i = 1, \ldots, n, n \geq 3$ expressed in object frame has a coordinate in the camera frame, which is represented by: $p_i^C = (x_i^C, y_i^C, z_i^C)$. The relationship between p_i^C and p_i^o can be defined by a rotation matrix $\boldsymbol{R}$ and a translation matrix $\boldsymbol{t}$ as follows

Q2

$$p_i^C = \boldsymbol{R}(\alpha, \beta, \gamma) \cdot p_i^o + \boldsymbol{t} \qquad (1)$$

In which, $p_i^C = (X_i^C, Y_i^C, Z_i^C)$ and $p_i^o = (X_i^o, Y_i^o, Z_i^o)$. Fig. 1 represents the corresponding frames of the pose estimation procedure. Consider α, β, γ as the roll, pitch and yaw orientation angles, respectively, and the translation is represented by
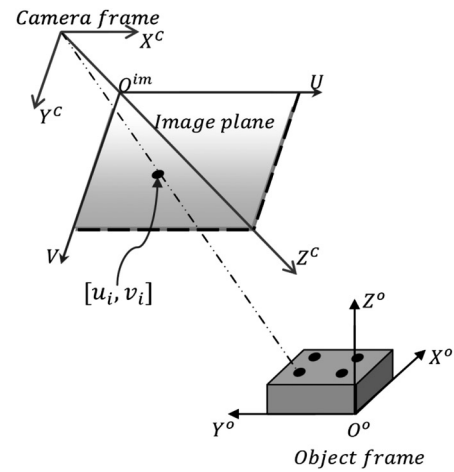
$$t = [X \quad Y \quad Z]^T \qquad (2)$$

Fig. 1 *Camera, object and image frames*

In fact $\boldsymbol{R}(\alpha, \beta, \gamma)$ and t introduce the relative position and orientation of the object frame with respect to the camera frame. The final goal of the pose estimation procedure is to find the corresponding relative parameters using the projections of the feature points in the image plane, which are denoted by $p_i^{im} = (u_i, v_i); i = 1, \ldots, n$. Using a pin-hole camera model, the relationship between p_i^{im} and p_i^C can be formulated as follows

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = K \cdot \begin{bmatrix} \dfrac{X_i^C}{z_i^C} & \dfrac{Y_i^C}{z_i^C} \end{bmatrix}^T \qquad (3)$$

$\boldsymbol{K}$ is the calibration coefficients vector, where

$$K = [\boldsymbol{K}_x \quad \boldsymbol{K}_y]^T = \begin{bmatrix} \dfrac{F}{P_x} & \dfrac{F}{P_x} \end{bmatrix}^T \qquad (4)$$

In this notation F is the focal length and P_x and P_y are the inter-pixel spacing parameters along u_i and v_i. By folding the $3n$ (1) into the pin-hole camera model (3), a set of $2n$ non-linear equations is achieved ('n' is the number of the feature points). In these non-linear equations, six components of the relative pose are unknown parameters, which should be estimated or calculated. In order to solve this set of non-linear equations and achieve a unique solution, at least four non-aligned points on the object are required [16, 17]. There are some numerical methods designed to solve this set of equations. However, most numerical methods have huge computational cost and need considerable elapsed time, which is not bearable for real-time applications. Additionally, numerical techniques are extremely sensitive to feature extraction error, noise and disturbance, which may cause insufficient results. In contrast, EKF provides a fast and flexible solution, which can easily deal with redundant measurement information for enhancing the estimation results accuracy. Also EKF has the ability of noise, disturbance and uncertainty rejection and it can enhance the precision of estimation process [18, 19].

3 Conventional EKF-based pose estimators

The most famous real-time method which can deal with the solution of the non-linear, uncertain and noisy pose estimation problem is the EKF. In conventional techniques

first, a discrete model of the relative motion between the camera and the target is introduced to incorporate into the discrete Kalman formulation [7]. In the literature, it is shown that an acceptable assumption is to use a constant velocity motion model, which assumes invariable relative velocity in each period of time 'T' [6, 10]. This assumption is drivable if the period of time is small enough, and moreover, the relative motion is roughly smooth. This model can be developed as follows [6]

$$S_k = A \cdot S_{k-1} + v_k \tag{5}$$

where

$$A = \text{diagonal} \left\{ \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \cdots, \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \right\}_{12 \times 12} \tag{6}$$

and $S = [X\ \dot{X}\ Y\ \dot{Y}\ Z\ \dot{Z}\ \alpha\ \dot{\alpha}\ \beta\ \dot{\beta}\ \gamma\ \dot{\gamma}]^T$ is the state vector, which contains relative pose parameters as well as their first-order derivatives. Additionally, v denotes the model uncertainty which is assumed to be described by a zero mean Gaussian noise. The output of this system is

$$Z_k = G(S_k) + W_k \tag{7}$$

$$G(S_k) = [u_1 v_1 u_2 v_2 u_3 v_3 u_4 v_4]_k^T \tag{8}$$

Subsequently, it can be concluded that

$$G(S_k) = K \cdot \left[\frac{X_1^C}{Z_1^C}, \frac{Y_1^C}{Z_1^C}, \frac{X_2^C}{Z_2^C}, \frac{Y_2^C}{Z_2^C}, \frac{X_3^C}{Z_3^C}, \frac{Y_3^C}{Z_3^C}, \frac{X_4^C}{Z_4^C}, \frac{Y_4^C}{Z_4^C} \right]_k^T \tag{9}$$

In this formulation X_i^C/Z_i^C and Y_i^C/Z_i^C are calculated from the state vector S_k via (1). In the above equations, W_k is the measurement noise, which is related to the camera characteristics and the vision accuracy. W_k is assumed to be expressed by a zero-mean Gaussian noise. Z_k is a vector of the normalised coordinates of feature points in the image plane. In order to apply the EKF for pose estimation, output equations should be linearised around the current state at each sample time, considering the first-order approximation of the Taylor series. This approximation is the most challenging issue in convergence analysis of the EKF. The linearisation of the system equations, results in a non-optimal estimation procedure [12]. There are some vital conditions, which should be met to guarantee the

Table 1 Qualitative comparison of the proposed method
Q3 characteristics to IAEKF method [3]

	IAEKF (proposed in [3])	Proposed method (EKO plus KF)
stability	lower (not garanteed)	higher (proved)
precision	Higher, using adaptation law	lower
mathematical proof of stability	×	✓
acceleration estimation ability	×	✓
vision fault detection	×	✓

convergence of the estimation error, which is explained in detail in the next section. In the next step, the EKF formulation is implemented as follows:

First step: Prediction

$$\begin{cases} \hat{S}_{k,k-1} = A \cdot \hat{S}_{k-1,k-1} & (10) \\ P_{k,k-1} = A \cdot P_{k-1,k-1} A^T + Q_{k-1} & (11) \end{cases}$$

Second step: Linearisation

$$H_k = \frac{\partial G(S)}{\partial S} \bigg|_{S = \hat{S}_{k,k-1}} \tag{12}$$

Third step: Calculating Kalman gain

$$K = P_{k,k-1} H_k^T (R_k + H_k P_{k,k-1} H_k^T)^{-1} \tag{13}$$

Fourth stage: Estimation and correction

$$\begin{cases} \hat{x}_{k,k} = \hat{x}_{k,k-1} + K(Z_k - G(\hat{x}_{k,k-1})) & (14) \\ P_{k,k-1} = A \cdot P_{k-1,k-1} A^T + Q_{k-1} & (15) \end{cases}$$

In this formulation, Q and R represent the process and measurement noise covariance matrices. Selection of Q and R are done by designer insight based on the accuracy of the extracted features and the profile of the object movement. Some adaptation laws have been designed to deal with this issue [7, 20].

4 Stability analysis of conventional EKF-based pose estimation

The most serious problem of EKF-based pose estimators is its divergence possibility in an uncertain structure [12], which may extremely endanger the performance of the estimation procedure. In general, extended Kalman procedure is a linearised structured observer, which has been used either to 'observe' the states of a non-linear system in a certain structure (EKO), or to 'estimate' the states in an uncertain structure (EKF). In this section, first the stability conditions of the extended Kalman in both certain and uncertain situations are given. Furthermore, the introduced stability conditions are developed to extract the stability conditions of conventional EKF-based pose estimators. In the literature, it has been shown that, EKF is stable provided that the stability conditions of EKO are met and the process and measurement noise are small enough [14]. These conditions are introduced as follows:

4.1 Stability of EKO and EKF

Consider a non-linear stochastic system

$$z_{n+1} = f(z_n, x_n) + G_n \omega_n \tag{16}$$

$$y_n = h(z_n) + D_n v_n \tag{17}$$

where $n \in N_0$ is the discrete time, $z_n \in R^q$ is the state, $x_n \in R^p$ is the input and $y_n \in R^m$ is the output. Also, v_n, ω_n, are R^k and R^l valued uncorrelated zero-mean white noise processes, with identity co-variance and D_n, G_n time-varying matrices of size $m \times k$ and $q \times l$, respectively. Consider a discrete time EKF, that is given by the following coupled difference equations:

- Difference equation for state estimate

$$\hat{z}_{n+1} = f(\hat{z}_n, x_n) + K_n(y_n - h(\hat{z}_n)) \qquad (18)$$

where $\hat{z}_n$ denotes the estimated states and K_n is the Kalman observer matrix gain.

- Riccati difference equation

$$P_{n+1} = A_n P_n A_n^T + Q_n - K_n(C_n P_n C_n^T + R_n)K_n^T \qquad (19)$$

where A_n and C_n are achieved from linearisation as follows

$$A_n = \frac{\partial f}{\partial z}(\hat{z}_n, x) \qquad (20)$$

$$C_n = \frac{\partial h}{\partial z}(\hat{z}_n) \qquad (21)$$

the Kalman gain is obtained from

$$K_n = A_n P_n C_n^T (C_n P_n C_n^T + R_n)^{-1} \qquad (22)$$

In this expression, Q_n is a time-varying symmetric positive definite $q \times q$ matrix and R_n is a time-varying positive definite $m \times m$ matrix. Q and R matrices are usually considered as approximations of the covariance matrices for the process and measurement noises. In [14] it has been shown that, estimation error remains bounded if the following assumptions are met:

- Assume there are positive numbers $\bar{a}, \bar{c}, \alpha, \beta, \vartheta$ and γ such that the following bounds on various matrices are satisfied for every $n \geq 0$

$$\begin{cases} \|A_n\| \leq \bar{a} & (23) \\ \|C_n\| \leq \bar{c} & (24) \\ \alpha I \leq P_n \leq \beta I & (25) \\ \gamma I \leq Q_n & (26) \\ \vartheta I \leq R_n & (27) \end{cases}$$

A_n is non-singular for every $n \geq 0$.

- Assume there are positive real numbers $\epsilon_\varphi, \epsilon_\vartheta, K_\varphi, K_\vartheta > 0$ such that the non-linear functions φ, ϑ are bounded by

$$\|\varphi(z, \hat{z}, x)\| \leq K_\varphi \|z, \hat{z}\|^2 \qquad (28)$$

$$\|\vartheta(z - \hat{z})\| \leq K_\vartheta \|z - \hat{z}\|^2 \qquad (29)$$

in which

$$\varphi(z_n, \hat{z}_n, x_n) = f(z_n, x_n) - f(\hat{z}_n, x_n) - A_n(z_n - \hat{z}_n) \qquad (30)$$

and

$$\vartheta(z_n, \hat{z}_n) = h(z_n) - h(\hat{z}_n) - C_n(z_n - \hat{z}_n) \qquad (31)$$

Furthermore, $z, \hat{z} \in R^p$ with $\|z - \hat{z}\| \leq \epsilon_\varphi$ and $\|z - \hat{z}\|^2 \leq \epsilon_\vartheta$. Then, the estimation error is exponentially bounded in sense of mean square error and bounded by unit probability. In a certain structure in which the states are observed using EKO, the above equations are the 'only' conditions which ensure the stability of estimation process. However, in an uncertain structure, some additional assumptions should be considered to guarantee the stability of EKF. The EKF is stable provided that

-
$$\|\xi_0\| \leq \epsilon, \text{ for some } \epsilon > 0,$$
which ξ_0 is the initial guess error $\qquad (32)$

- Covariance matrices of the noise terms are bounded via

For some $\lambda > 0$: $\begin{cases} G_n G_n^T \leq \lambda I & (33) \\ D_n D_n^T \leq \lambda I & (34) \end{cases}$

Remark 1: In the literature there are some results given which can reduce the difficulties of the EKF stability analysis [14, 21]. These results can be used to evaluate the behaviour of EKF in pose estimation problem.

- Inequality (25) is closely related to the observability and detectability of linearised system. Moreover, it has been shown that this inequality hold, if the non-linear observability conditions are met.
- Inequality (28) and (29) hold if f and h are twice differentiable.

4.2 Stability for EKF-based pose estimators

According to the previous section there are some conditions that should be met in order to ensure the convergence of the EKF. In this section, we focus on the stability of the EKF-based pose estimators. These conditions have to be crucially met in order to use EKF in practical visual servoing systems.

Consider (6), according to this fact that, T is the sample time, which is bounded, it is clear that:

- $\|A_n\|$ is bounded, which means there is a positive number $\bar{a}$ by which $\|A_n\| \leq \bar{a}$.
- A_n is non-singular.

Furthermore, considering (9) and (12): h_{ij} is the ijth element of H and g_{ij} is the ijth element of G. If $Z \neq 0$ then g_{ij} and h_{ij} posses some harmonics, which are smooth and bounded, and

- $\|H_n\|$ is bounded.
- G is twice differentiable.

In the previous section, it has been mentioned that, inequality (25) is satisfied if the linearised system is observable at each epoch. Note that, If v_i is the ith eigenvector of A_n, it can be shown that $C_n \cdot v_i \neq 0$ provided that, $Z \neq 0$. Consequently, if the camera and object have a non-zero distance through Z axis in the camera frame, which is a logical assumption in pose estimation and visual servoing systems, the observability conditions are easily met. Consequently, if the initial guess is accurate enough, and if the measurement noise is relatively small and furthermore, if the process uncertainty is adequately small according to the target movement, pose estimation procedure becomes 'stable'.

Unfortunately in real applications, these assumptions are not easily met, and the process may lead to a divergent pose estimator. This fact is the challenging concern in employing EKF in pose estimation in practice. As it is mentioned before, some improvements have been used in order to deal with this problem in pose estimation issue. IEKF is one of the well-known improved formats, which has been employed in order to estimate the position of

rigid objects [12]. In this approach, measurement function is iteratively linearised considering the novel estimated pose at each sample time. This procedure enhances the accuracy of linearisation and decreases the uncertainty level (as an unwanted result of linearisation error). Consequently, the stability margin of the filter is enlarged to some extent; however, no guarantee for the convergence of the estimation process of this technique is given in the literature. In the next section, a new scheme is proposed to robustly estimate the position and orientation of an object in an uncertain structure and in the presence of measurement noise and model uncertainty.

5 Proposed Kalman-based approach

In this section, the structure of the proposed method will be illustrated. In this approach, the pose estimation algorithm has been divided into a two-step process, which includes a 'static pose estimation procedure' in addition to a 'pose correction and derivatives estimation'. First, at each iteration time the static position is estimated with bounded error using an EKO. In fact, at this stage a virtually noiseless condition is provided for the estimation process, to guarantee the stability of the static pose estimator. Next, the estimated pose will be additionally corrected, and its derivatives will be estimated using a Kalman filter. Since the estimated static position has bounded error, the stability conditions for the second step are met, and hence, the overall estimation error remains bounded. Fig. 2 illustrates one epoch of the proposed method. The formulation of the 'static pose estimator' is the same as the conventional KF-based pose estimation algorithms; however, their applications are different. The static estimator iterates at each sample time, consequently requires inner-sampling calculations. The function of the static estimator is to calculate an approximation of the current target's location at each overall period and feed this approximation to the KF. The function of KF is to estimate the pose, velocity and acceleration via noisy position using a constant acceleration model approximation.

5.1 First stage: static pose estimation

The input set of the static pose estimator includes the projection of the feature points on the image plane. The EKO's loop reiterates for several epochs in each 'overall **Q4** sample time' (The overall sample time is the sample time of whole estimation procedure, which is determined by the camera rate, image-processing rate and computations elapsed time.) using the same input for epochs. Therefore at each sample time the determined locations of the features,

is fed to EKO estimator iteratively, in which a constant input is set for the static estimator. Consequently after some internal iteration EKO estimates the static position of the object with the assumption of zero velocity and is no input from the target acceleration. Hence, EKO suffers from no uncertainties in modelling, and the variation of its input is zero because of the constant velocity model. However, this model suits to our assumption of the target motion. Moreover, the input data are constantly fed to EKO during each overall sample time, and therefore, EKO also suffers from no measurement noise. This calculated pose is not accurate then, since the inputs of the static estimator are not accurate considering the overall measurement noise of the estimation. In order to deal with this issue and more importantly estimate the velocity and acceleration of the object, the EKO output is fed into a KF which operates in real-time.

5.2 Second stage: pose correction and pose derivatives estimation (KF)

The second stage of the estimation procedure includes correction of the given static pose, considering the estimation of the target velocity and acceleration. The core of this process is based on KF structure, which operates in an uncertain environment considering the model uncertainty and measurement noises. A constant acceleration model has been used as an approximation of the target movement. This model is more general and more accurate compared to constant velocity models, and is able to reduce the process uncertainties and can further enhance the precision of the estimated results. Using a constant acceleration model leads to an estimate of the acceleration of the target, which can be further used in the control strategy or in the trajectory generation of the robot. A discrete dynamic model of the relative motion between camera and target can be developed as

$$S_k = A \cdot S_{k-1} + v_k \tag{35}$$

where, the system matrix A is augmented from six triple integrator model as

$$A = \text{diagonal} \left\{ \begin{bmatrix} 1 & T & 0.5T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}, \ldots, \begin{bmatrix} 1 & T & 0.5T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \right\}_{18 \times 18} \tag{36}$$

and, $S = [x \ \dot{x} \ \ddot{x} \ y \ \dot{y} \ \ddot{y} \ z \ \dot{z} \ \ddot{z} \ \alpha \ \dot{\alpha} \ \ddot{\alpha} \ \beta \ \dot{\beta} \ \ddot{\beta}]^T$, is the state vector which contains relative motion
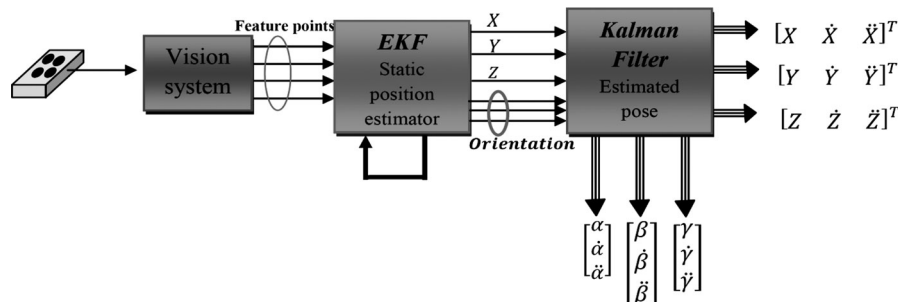

Fig. 2 *Flow chart of the proposed method, only on epoch is demonstrated*

(linear and angular) parameters of the camera and the target. Additionally, v is the model uncertainty vector, which is assumed to be described by zero-mean Gaussian noise. It has been mentioned that, the noisy location of the target is measured at this stage. Consequently, the output formulation of this system is

$$Z_k = CS_k + W_k \qquad (37)$$

in which (see (38))

And W_k is the measurement noise, which is related to the static pose estimator precision. W_k is assumed to be expressed by a zero-mean Gaussian noise. Z_k is the relative location between the target and the camera. Finally, incorporating this model in the simple linear KF formulation, the desired states (position and orientation, velocities and accelerations) are estimated during the object relative motion. Note that Q and R, which may represent the covariance matrices of the process and measurement noise, respectively, should be tuned appropriately to achieve a desired result. These matrices may be tuned adaptively via an adaptation law [7, 22].

5.3 Vision fault detection

As it is mentioned before, the main role of the vision system is to extract the coordinates of the feature points in the image plane. Therefore one of the most crucial sources of inaccuracy in pose estimation algorithms and visual servoing systems is the measurement errors [23], which may destabilise conventional estimation techniques. A large number of attempts have been proposed in the literature to cope with this problem and try to recover the performance in presence of inaccuracy of some parameters like camera calibration factors [24–26]. It has been illustrated that, the convergence of the proposed method is not related to the accuracy of the measurement. However, in some situations, when the vision algorithm fails to find the correct locations or an acceptable approximation of the feature points, a considerable disturbance is imposed to the estimation process, which may degrade the transient performance of pose estimation. In this paper, this category of measurement error has been denoted as vision faults, and one of the important capabilities of the proposed method is the ability of vision fault detection. In our proposed method, because of the two-step structure of the procedure, the vision faults cannot destabilise the estimation procedure; however, they can impose significant transient errors on the static estimator. Note that, the proposed method is able to detect the faults occurred in the vision system using some threshold on the Euclidean norm of the 'innovation' term of the static estimator. Considering (14), the difference between the approximated and observed coordinates of the feature points in the image-plane: $(Z_k - G(\hat{x}_{k,k-1}))$ can be named
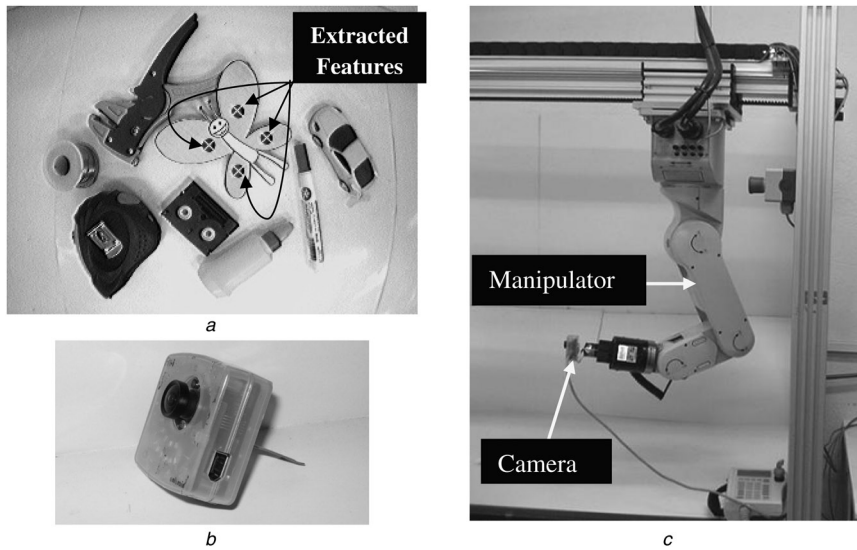
Fig. 3 *Experimental setup*
a Object's environment
b Used camera
*c* Manipulator for generating the cyclic motions

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \qquad (38)$$

as the innovation term of the estimation process. The approximation is obtained using the estimated pose, geometric equations between the camera and the object frame and the camera pin-hole model. This norm can provide a suitable performance measurement of the estimation. In no fault case, the estimation is stable and the Euclidean norm of the innovation is near zero. If the extracted coordinates of the features in the image plane is contaminated with considerable disturbance, this norm is certainly bounded with a non-zero finite limit. Therefore using a suitable threshold for the innovation norm enables the estimator to detect any possible vision fault. In such cases it is advised to use the predicted position from the previous sample time, in order to not distribute the errors caused by the vision fault into the whole estimation.

6 Simulation and experimental results

The efficiency and the performance of the proposed pose estimation algorithm are verified through some simulations and then two experiments in this section. Note that, in these simulations and experiments, the covariance matrices of the process and noise are assumed to be fixed and no adaptation law is employed.
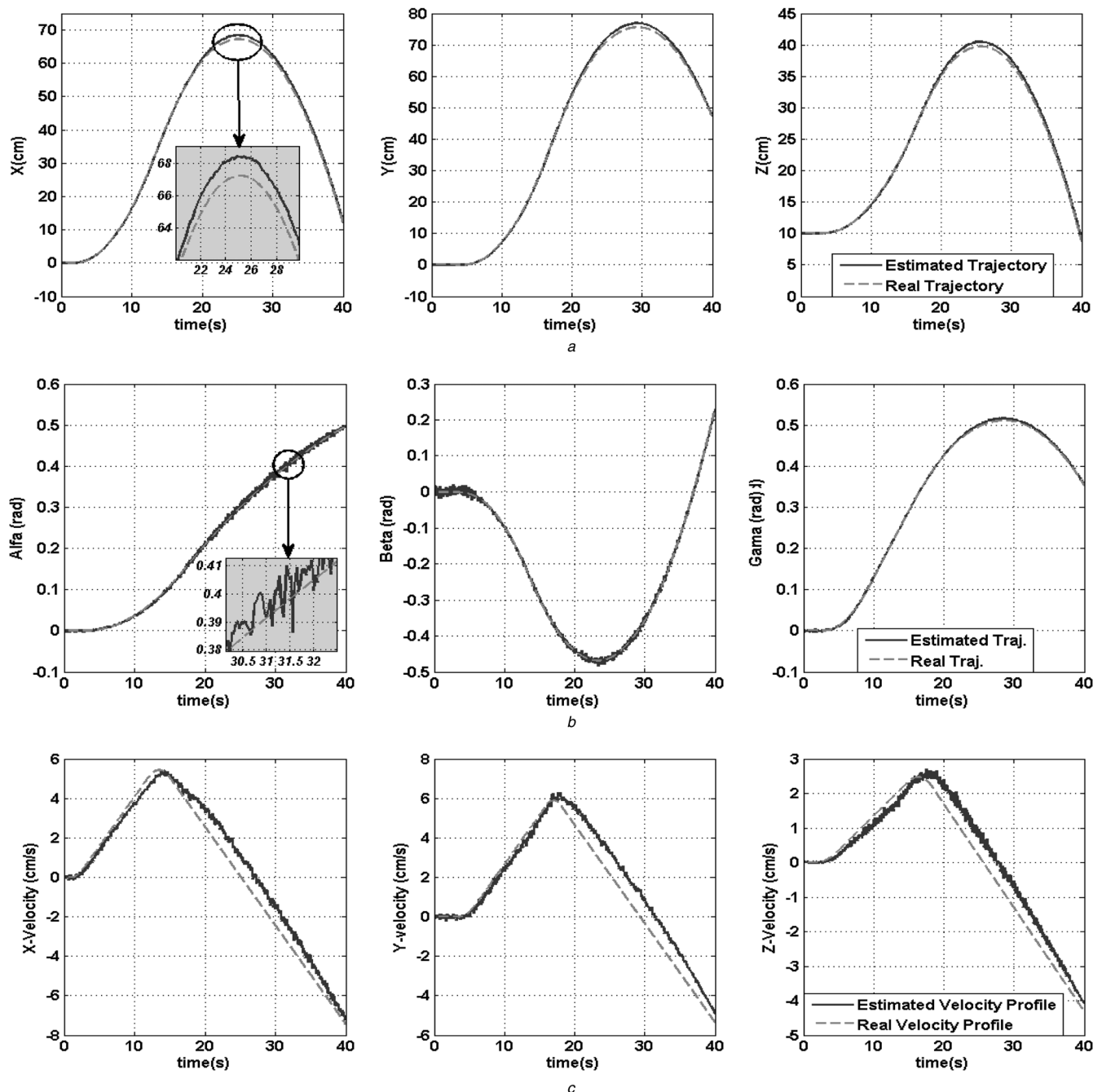
Fig. 4 *Simulation results for conventional EKF in a low-level noise condition*

a Position estimation
b Orientation estimation (Euler angles)
c Velocity estimation in *X*, *Y*, *Z*-axes
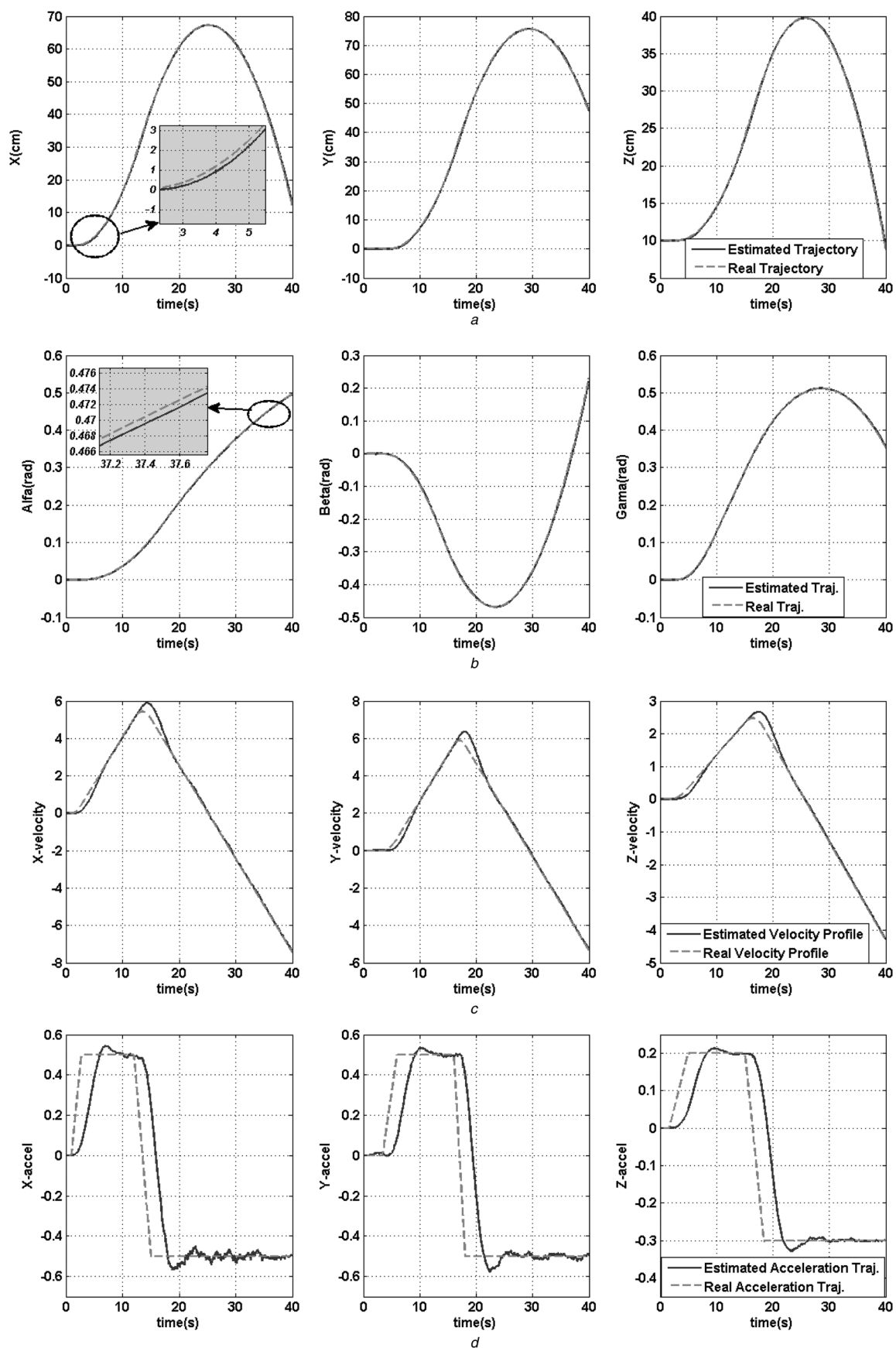
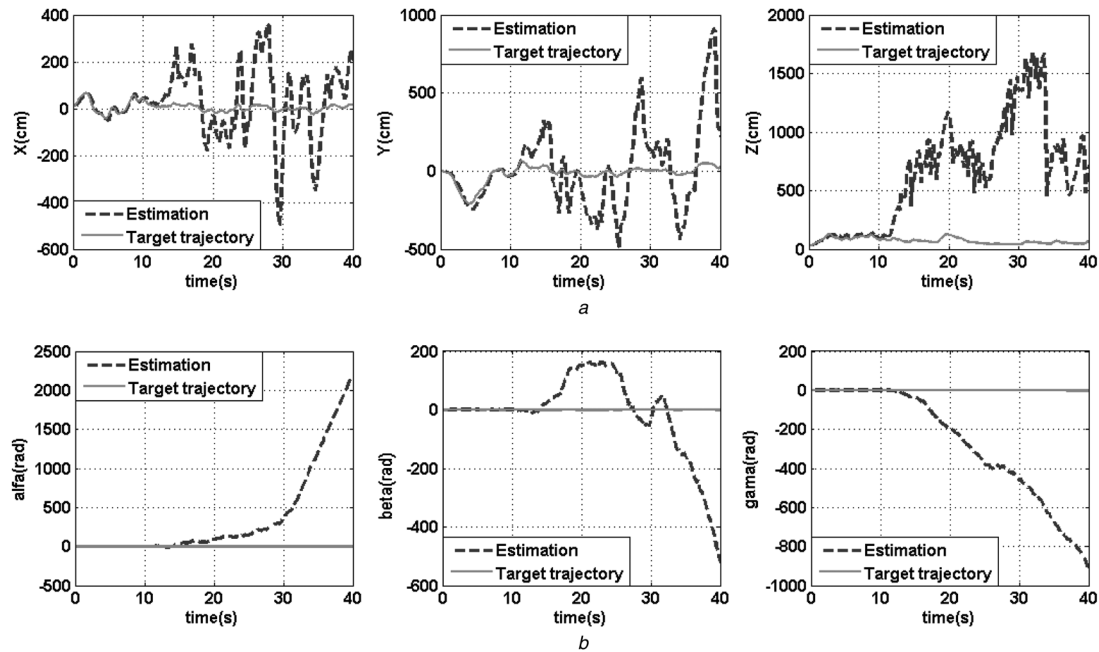**Fig. 5** *Simulation results for our proposed method in a low-level noise condition*

a Position estimation
b Orientation estimation (Euler angles)
c Velocity estimation in *X*, *Y*, *Z*-axes
d Acceleration estimation in *X*, *Y*, *Z*-axes

6.1 System set-up

The hardware set-up is composed by a PC equipped with a Pentium IV (1.84 GHz) processor and a 512 MB of Ram in addition the camera from Unibrain Company with 30 fps frame rate and a wide lens with 2.1 mm focal length. A 5DOF robot is utilised to generate the relative cyclic trajectories in the experiments.
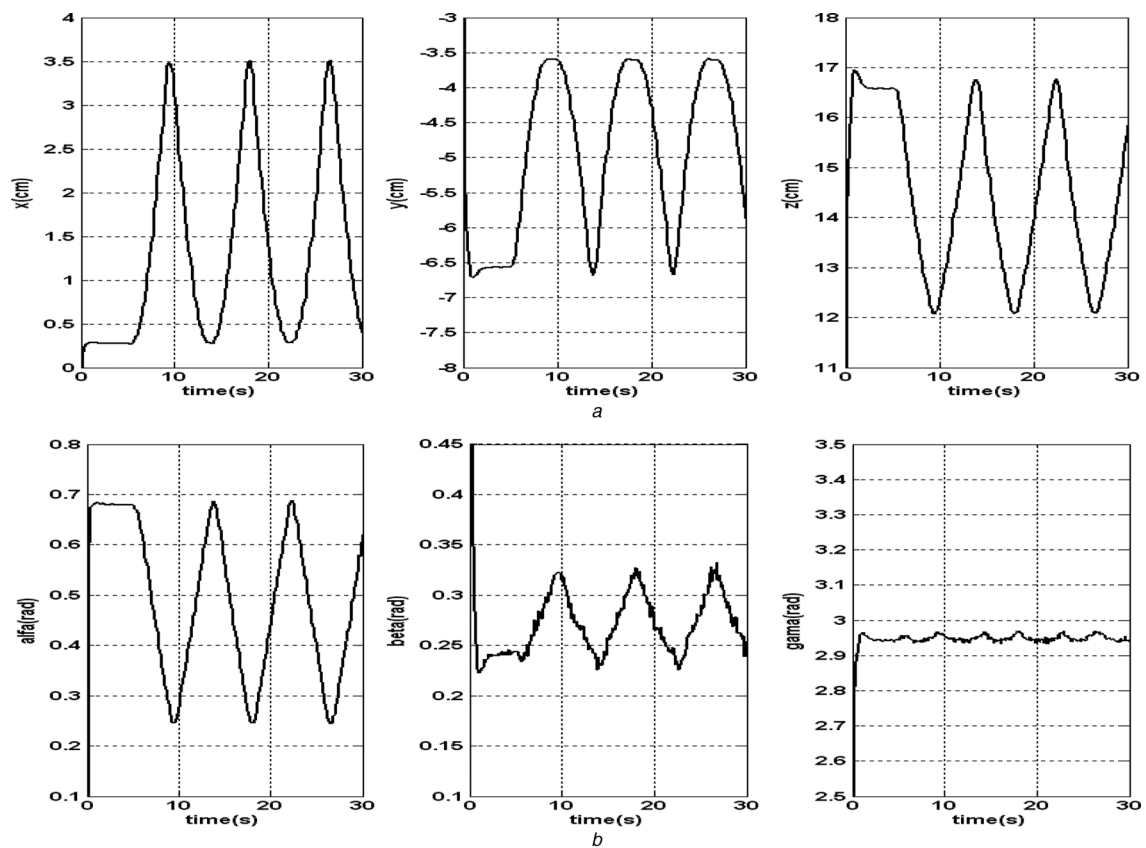

Fig. 6 *Simulation results for controversial EKF in a low-level noise condition*

a Position estimation
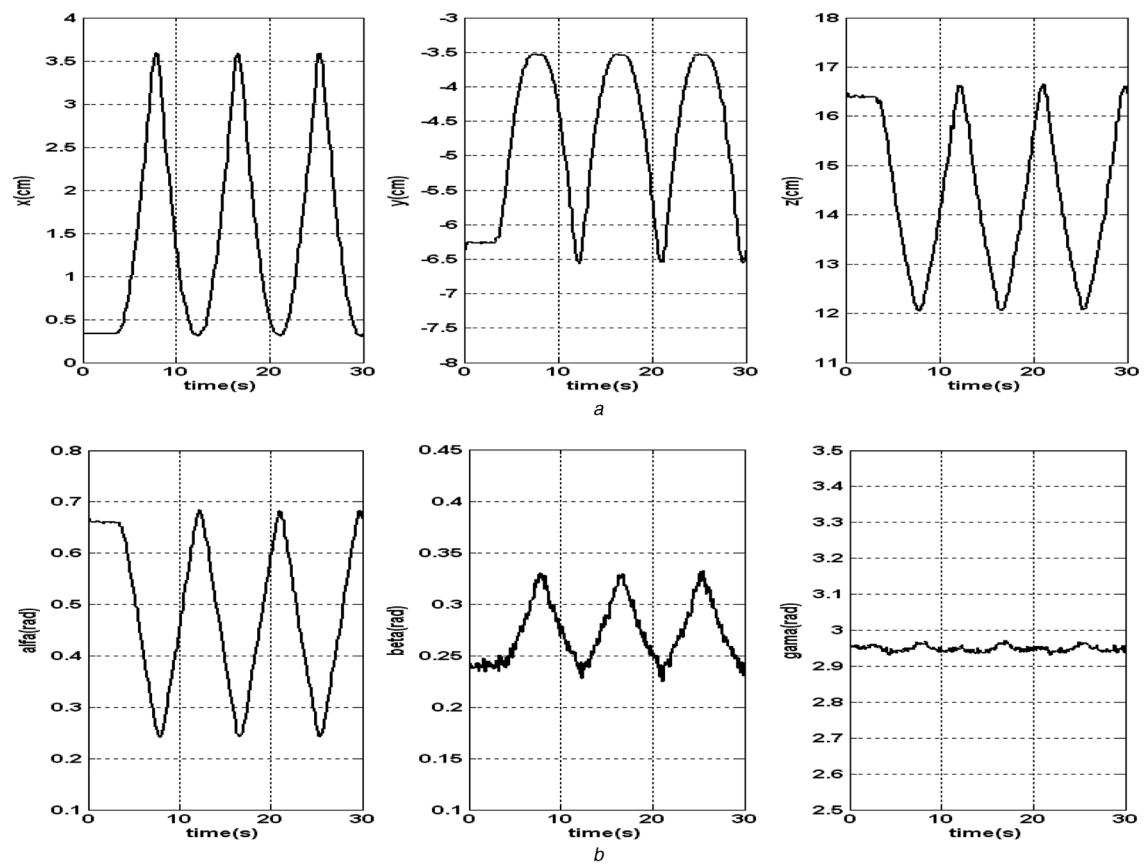*b* Orientation estimation
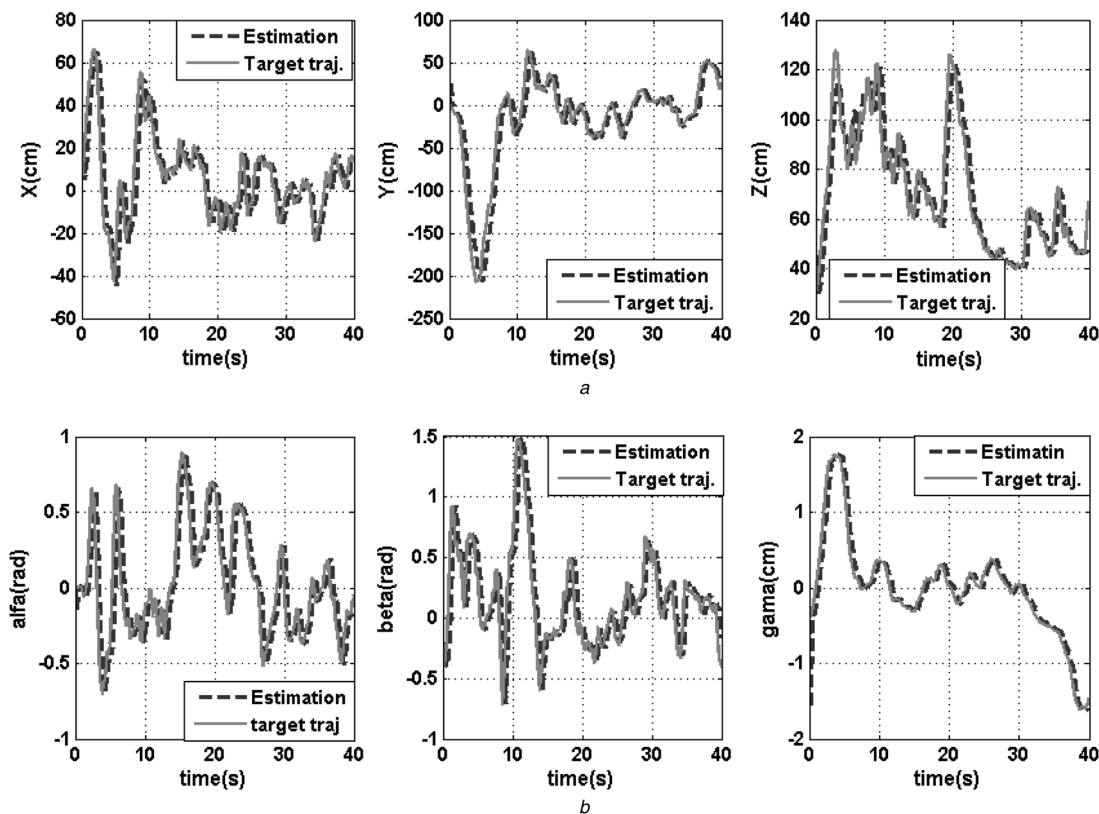
Fig. 7 *Simulation results for our proposed method in a high-level noise condition*

a Position estimation
b Orientation estimation

Fig. 8 *Experiment no. 1, results for controversial method in a low-level noise condition*

a Position estimation
b Orientation estimation

Fig. 9 *Experiment no. 1, results for proposed method in a low-level noise condition*

a Position estimation
b Orientation estimation

This robot is a Mitsubishi manipulator model RV-2AJ, which is shown in the experimental setup Fig. 3. Software is implemented in MATLAB, which includes image processing, feature extraction and pose estimation algorithms. Since the maximum frame rate of the camera is 30 fps and the software is implemented by the MATLAB, the sample time is limited to about 0.1 s.

6.2 Low noise simulation

It can be inferred from literature that EKF has an acceptable performance at low noise conditions [14]. Accordingly, as the first evaluation simulation, the behaviour of conventional EKF-based pose estimator is compared to our proposed method in the present of low level measurement and process noise. In this simulation, the variance of the measurement noise is assumed to be '0.5', the target motion is relatively smooth and the sample time is considered relatively small '0.04 s', in order to ensure a low level of process noise considering the 'constant velocity' model.

Figs. 4 and 5 represent two trajectory estimation results, respectively, using conventional EKF against our proposed method. In these figures the actual motion of the target is plotted against the estimated values, for six pose variables. Moreover, the ability of the velocity and acceleration estimation has been evaluated. It should

be noted that in the conventional method the acceleration profile cannot be estimated. In these simulations, both methods estimate the trajectory of the target with bounded error. Although it is clear that, the proposed method exhibits a superior behaviour with more accurate estimation and the ability of acceleration estimation. This simulation clearly illustrates the effectiveness of the proposed method in order to estimate the position, velocity and acceleration of the target via one camera.

6.3 High noise simulation

The main goal of the following simulation is to evaluate the capability of the proposed method at a high noise level. In this simulation the variance of the measurement noise is assumed to be '5', the target has a coarse motion in 40 s, and the sample time is set more realistically to 0.2 s, in order to provide a relatively uncertain situation for the estimation process. Figs. 6 and 7 illustrate the pose estimation results. In this simulation it is observed that the proposed method can effectively estimate the target motion with an acceptable error. However, the conventional EKF-based pose estimator is unable to estimate the motion of the target with bounded error. This simulation clearly illustrates the capability of the proposed method in stable estimation of the target motion profile at high noise level.
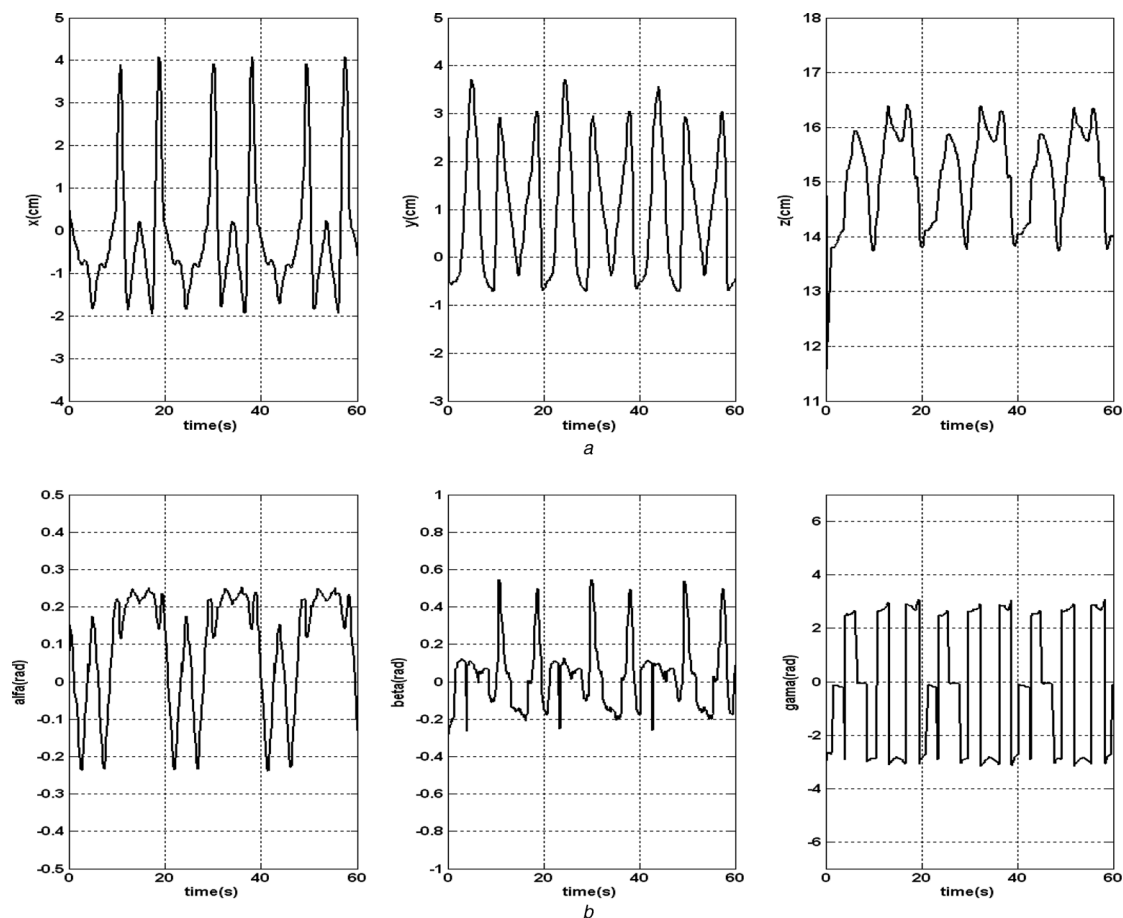
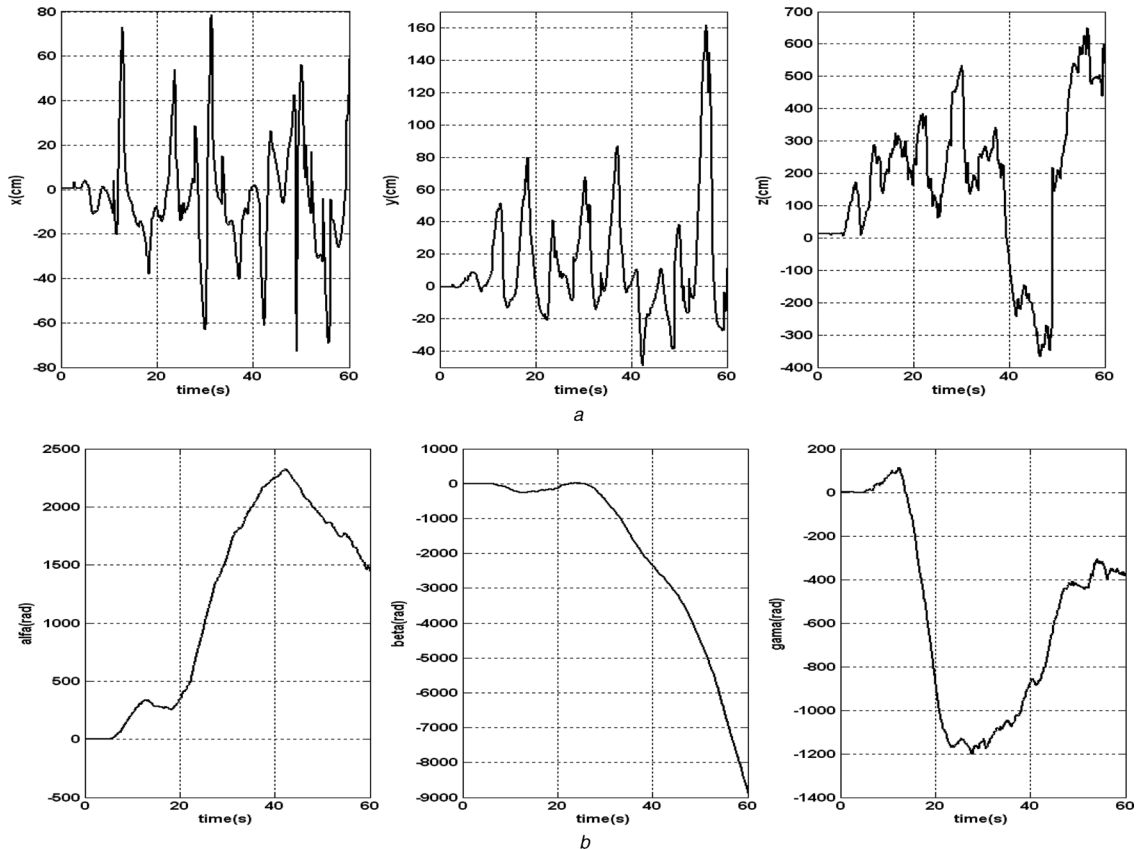**Fig. 10** *Experiment no. 2, results for proposed method in a high-level noise condition*
a Position estimation
b Orientation estimation

6.4 Smooth trajectory experiment

In the experiments given here and in the next subsection the real behaviour of the proposed method is evaluated and compared to that of the conventional EKF in the present of actual uncertainties and noises. The level of measurement noises and model uncertainties can be changed through the object background and the motion profile. The required trajectories are generated by a robot which moves the camera relative to the fix object. The camera was installed on the end-effector of the robot. In this experiment, both methods were used to estimate a special smooth trajectory. Obviously smoothness leads to a low level of the model uncertainty since the assumption of constant acceleration and constant velocity at each sample time is better satisfied. It should be noted that, in this experiment the object's environment is very clean with a fixed bright colour in order to ensure a low level of measurement noise, concerning the high accuracy of the feature extraction method in this situation. Figs. 8 and 9 illustrate the results of this experiment. As it is expected, both methods can estimate the smooth path with a bounded error. In fact, in this experiment, the upper bound of the model uncertainties and measurement noises are within an acceptable value and consequently the convergence conditions of EKF have been met and the conventional EKF method can estimate the relative trajectory in a stable convergence. Additionally, it is observed that the convergence of the proposed method is not linked to the upper bounds of the measurement and process noise.

6.5 Rough trajectory experiment

In the final experiment, a coarse and cyclic path has been generated using the robot manipulator. Both methods are implemented for estimation of this relative coarse motion. In this experiment, the object background is untidy and contains a wide variety of colourful objects as shown in Fig. 3. Consequently, the accuracy of the feature extraction method in finding the coordinates of the feature points in the image plane is degraded. Moreover, the possible fault occurrence in vision process is considerably increased, and therefore in the above scenario a relatively high measurement noise is generated. Furthermore, this type of rough trajectory motion drastically degrades the assumption of constant velocity model, which causes a significant process uncertainty. As a conclusion, this experiment is designed to evaluate the behaviour of both methods in high level noise situation. Figs. 10 and 11 illustrate the results of this experiment. As it is expected, in this experiment the proposed method was able to estimate the relative trajectory in a stable manner. However, in this noisy situation the conventional method is unable to converge to a proper target position estimate with a bounded estimation error. This experiment has been designed to evaluate the real behaviour of the estimation methods in a more realistic industrial environment with fewer constraints, and it has been shown that the proposed method exhibit a promising behaviour compared to that of conventional technique.

Fig. 11 *Experiment no. 2, results for conventional EKF method in a high-level noise condition*

a Position estimation
b Orientation estimation

7 Conclusions

In this paper, a new robust Kalman-based technique has been proposed in order to deal with the non-linear problem of pose estimation. The most challenging issues for using the EKF in pose estimation is its divergence possibility on the estimation error, which can extremely degrade the estimation process performance. The most appealing property of the proposed technique is the robust estimation ability, despite of the measurement data and process model precision in an industrial environment. In this method non-linear estimation problem in uncertain situation is decomposed into two parts. At each sample time, initially an inaccurate static position is calculated using an EKO and then the corrected position, velocity and acceleration terms are estimated by a Kalman filter. It has been shown that, the stability conditions of the conventional EKF-based pose estimation, are directly related to the boundaries of the measurement noise and the model uncertainties, which are directly related to the vision accuracy and the motion profile of the target. Consequently there is no guarantee to ensure the convergence of conventional methods in general. In contrast it is illustrated that the proposed method is stable in uncertain situations by using a two-stage algorithm, which isolates the non-linearity problem from the uncertainty issue. In order to have a quantitative comparison of the proposed method to the state-of-art EKF-based pose estimation methods, the characteristics of the proposed method is compared to that of IAEKF [3]. As it can be seen in this table our proposed method will never suffer from instability, since its stability conditions can be easily satisfied through the two-stage algorithm. Although IAEKF may possess better estimation performance because of its incorporated adaptation law, its general convergence is not guaranteed since its stability is not proven. Furthermore, the proposed method has the ability of velocity and acceleration estimation, and fault detection possibility, owing to its inherent structure, whereas in the other methods these abilities are not directly accessible.

8 References

1 Wilson, W.: 'Visual servoing control of robots using Kalman filter estimates of robot pose relative to work-pieces', in Hashimoto, K. (Ed.): 'Visual servoing' (World Scientific, 1994), pp. 71–104

2 Nickels, K., DiCicco, M., Bajracharya, M., Backes, P.: 'Vision guided manipulation for planetary robotics – position control', *Robot. Auton. Syst.*, Accepted in 2009, in the press

Q5

3 Janabi-Sharifi, F., Marey, M.: 'A Kalman-filter-based method for pose estimation invisual servoing', *IEEE Trans. Robot.*, 2010, **26**, (5)

Q6

4 Marchand, É., Chaumette, F.: 'Feature tracking for visual servoing purposes', *Robot. Auton. Syst.*, 2005, **52**, (1), pp. 53–70

5 Wu, J.J., Rink, R.E., Caelli, T.M., Gourishankar, V.G.: 'Recovery of the 3D location and motion of a rigid object through camera image (An extended Kalman filter approach)', *Int. J. Comput. Vis.*, 1988, **3**, pp. 373–394

6 Wilson, W.J.: 'Relative end effector control using Cartesian position based visual servoing', *IEEE Trans. Robot. Autom.*, 1996, **12**, pp. 684–689

7 Lippiello, V., Siciliano, B., Villani, L.: 'Adaptive extended Kalman filter for visual motion estimation of 3D objects', *Control Eng. Pract.*, 2007, **15**, pp. 123–134

8 Janabi-Sharifi, F., Wilson, W.J.: 'Automatic selection of image features for visual servoing', *IEEE Trans. Robot. Autom.*, 1997, **13**, pp. 890–903

9 Lippiello, V., Siciliano, B., Villani, L.: 'A new method of image features pre-selection for real-time pose estimation based on Kalman filter'. Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, 2002, pp. 372–377

10 Ficoceli, M., Janabi-Sharifi, F.: 'Adaptive filtering for pose estimation in visual servoing'. Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2001, pp. 19–24

11 Lefebvre, T., Bruyninckx, H., De Schutter, J.: 'Kalman filters for nonlinear systems: A comparison of performance', *Int. J. Control*, 2004, **77**, (7), pp. 639–653

12 Shadman, A., Janabi-sharifi, F.: 'Sensitivity analysis of EKF and iterated EKF pose estimation for position-based visual servoing'. Proc. IEEE Conf. on Control Applications, 2005, pp. 755–760

13 Julier, S.J., Uhlmann, J.K.: 'Unscented filtering and non-linear estimation', *Proc. IEEE*, 2004, **92**, (3), pp. 401–422

14 Reif, K., Gunther, S., Yaz, E.: 'Stochastic stability of the discrete-time extende Kalman filter', *IEEE Trans. Autom. Control*, 1999, **44**, (4), pp. 714–728

15 Ficocelli, M., Janabi-Sharifi, F.: 'Adaptive filtering for pose estimation in visual servoing'. Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Maui, HI, 2001, pp. 19–24

16 Sim, T.P., Hong, G.S., Lim, K.B.: 'Modified smith predictor with dementhon-horaud pose estimation algorithm for 3D dynamic visual servoing', *Robotica*, 2002, **20**, pp. 615–624

17 Yuan, J.S.-C.: 'A general photogrammetric method for determining object position and orientation', *IEEE Trans. Robot. Autom.*, 1989, **5**, pp. 129–142

18 Wang, J., Wilson, W.: '3D relative position and orientation estimation using Kalman filter for robot control'. Proc. IEEE Int. Conf. on Robotics and Automation, 1996, pp. 2638–2645

19 Hutchinson, S., Hager, D., Corke, P.I.: 'A tutorial on visual servo control', *IEEE Trans. Robot. Autom.*, 1996, **12**, (5), pp. 651–670

20 Lippiello, V., Villani, L.: 'Managing redundant visual measurements for accurate pose tracking', *Robotica*, 2003, **21**, pp. 511–519

21 Bougerol, P.: 'Kalman filtering with random coefficients and contractions', *SIAM J. Control*, 1993, **31**, pp. 942–959

22 Myers, K.A., Tapley, B.D.: 'Adaptive sequential estimation with unknown noise statistics', *IEEE Trans. Autom. Control*, 1976, **21**, pp. 520–523

23 Kyrki, V., Kragic, D., Christensen, H.I.: 'Measurement errors in visual servoing', *Robot. Auton. Syst.*, 2006, **54**, (10), pp. 815–827

24 Sebastián, J.M., Pari, L., Angel, L., Traslosheros, A.: 'Uncalibratedvisual servoing using the fundamental matrix', *Robot. Auton. Syst.*, 2009, **57**, (1), pp. 1–10

25 Tsai, C.-Y., Song, K.-T., Dutoit, X., VanBrussel, H., Nuttin, M.: 'Robust visual tracking control system of a mobile robot based on a dual-Jacobian visual interaction model', *Robot. Auton. Syst.*, 2009, **57**, (6–7), pp. 652–664

26 Fioravanti, D., Colombo, C., Allotta, B.: 'Self-calibrated visual servoing with respect to axial-symmetric 3D objects', *Robot. Auton. Syst.*, 2009, **57**, (4), pp. 451–459