Provably stable nonlinear and adaptive observers for dynamic structure and motion estimation

Ola Dahl and Anders Heyden

Applied Mathematics, Malmö University, Sweden

Abstract. Structure and motion estimation from long image sequences is a hard problem, especially when it comes to proving stability and convergence of different methods. We propose a novel approach based on nonlinear and adaptive observers based on a dynamic model of the motion. The estimation of the three-dimensional position and velocity of the camera as well as the three-dimensional structure of the scene is done by observing states and parameters of a nonlinear dynamic system, containing a perspective transformation in the output equation, often referred to as a perspective dynamic system. The paper presents a stability analysis for a class of observers for the estimation of position. The analysis provides conditions for convergence, and insight into feasible motions. An advantage of the proposed method is that it is filter-based, i.e. it provides an estimate of structure and motion at each time instance, which is then updated based on a novel image in the sequence. Finally, the performance of the proposed method is shown in simulated experiments.

Key words: Dynamic vision, perspective dynamic systems, nonlinear observers, adaptive observers, stability analysis

1 Introduction

One of the central problems in computer vision is the recovery of structure and motion from image sequences. Most approaches to this problem are batch-methods, where first all images are gathered and then the calculations are performed on all the data. These methods are usually based on multi-view tensors and nonlinear least squares optimization, see [1] for an overview of these approaches. However, some attempts have been made to develop recursive methods (in the sense of processing images as they becomes available and always having an estimate of motion and structure at hand), e.g. [2, 3] and also some related work in the area of automatic control, e.g. [4]. The main motivation for developing recursive methods is to being able to use them in real-time applications, where on-line structure and motion estimation is essential.

Another aspect of structure and motion estimation methods is the possibility to prove convergence and stability of the proposed algorithms. Very few attempts in this direction has been made, especially in the case of long image sequences. Recent work on applying convex optimization techniques to different sub-problems has being able to show that global optima are obtained, e.g. [5, 6]. However, this has not been shown in the case of both structure and motion estimation for longer image sequences. Other results concerning convergence of batch-methods have been reported in [7] for a variant of the factorization method and in [8] for still another batch method. Within the area of automatic control, observability and controllability of dynamic perspective systems have been studied in e.g. [9].

In order to apply nonlinear adaptive observers, a dynamic model of the motion of the camera is introduced. This formulation turns the structure and motion problem into a problem of observing states and parameters in the resulting dynamic perspective system. Structure and motion estimation without knowledge of motion parameters can be considered the most challenging case, and is described e.g. in [10], where structure-independent motion estimation is performed using a dynamic system, and in [11], where structure estimation is treated. References [10] and [11] present algorithms for estimating structure as well as motion using e.g. implicit extended Kalman filters. The algorithms are verified experimentally but it is difficult to establish analytical results regarding convergence and stability. A specific class of algorithms for structure estimation, where available values for angular and linear velocities are used and where position is estimated, can be formulated as *nonlinear observers*. This kind of observers are described e.g. in [12–22], which present estimators for *structure only* using different kinds of nonlinear observers, providing various analytical results regarding stability, and simulation examples for illustration of observer performance.

This paper describes how a parametrization of the underlying dynamic system can be used to formulate estimation problems for *structure as well as motion*, and how the so obtained problem formulations can be used for the derivation of estimators, using available methods from nonlinear and adaptive control. Problem formulations for different estimation tasks are presented, and observers are derived and illustrated using simulation examples. We also derive analytical results regarding stability and convergence for the case of structure estimation.

The possibility of using nonlinear and adaptive observers to handle simultaneous structure and motion estimation has been previously treated in [23], however using a different approach involving stereo vision.

Sections 2 and 3 review the derivation of the perspective dynamic system under consideration and describe the dynamic vision parametrization. In Section 4 it is shown how the parametrization can be used to obtain a common problem formulation for several different structure and/or motion estimation tasks. In Section 5 a nonlinear observer structure is selected, and nonlinear and adaptive observers for the parameterized perspective dynamic system are presented. The stability analysis is presented in Section 6. The performance of the derived observers is then illustrated in Section 7 using simulation examples.

2 Dynamic perspective system

The dynamic system parametrization is derived from a dynamic system which is obtained, as is commonly done, from a description involving coordinate systems

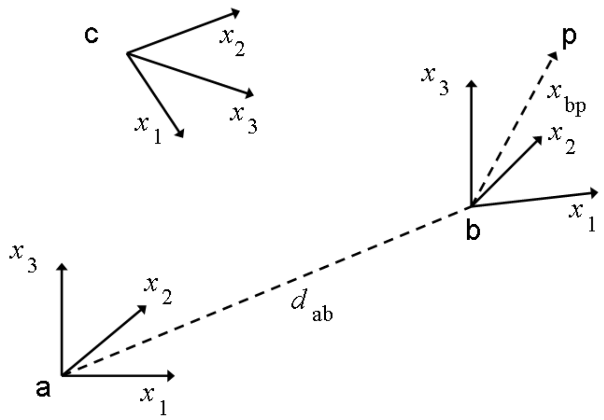

Fig. 1. Coordinate systems used for describing the position of a 3D point p, belonging to an observed object.

for the observed object and for the camera. For the purpose of clarity we also employ an inertial coordinate system when defining a specific motion.

The inertial coordinate system is denoted the a-system. The object coordinate system is denoted the b-system, and is attached to the observed object which is assumed to be a rigid body. The object may be stationary or moving. The camera coordinate system is referred to as the c-system, and is considered attached to a possibly moving camera. The coordinate systems are illustrated in Fig. 1, where a selected point p on the observed object is also shown.

A system of differential equations for the motion of the point p can be derived. Introducing the notation d_{aba} for the coordinates of the vector d_{ab} in Fig. 1 when expressed using the orientation of the a-system, and the notation x_{bpb} for the coordinates of the vector x_{bp} in Fig. 1 when expressed using the orientation of the b-system, we get, using a rotation matrix R_{ab} which expresses the relative orientation between the two coordinate systems, that $x_{apa} = d_{aba} + R_{ab}x_{bpb}$. Similarly, the coordinates of the vector x_{bp} can be expressed using the orientation of the c-system using a rotation matrix R_{cb}, as

$$x_{cpc} = d_{cbc} + R_{cb}x_{bpb} \ . \tag{1}$$

Define the skew-symmetric matrix $S(v)$ associated with a vector $v \in \mathbb{R}^3$, using a cross-product with an arbitrary vector u as

$$S(v)u = v \times u \ . \tag{2}$$

Introduce also the angular velocity vector ω_{cbc} and the matrix $S(\omega_{cbc}) = \dot{R}_{cb}R_{cb}^T$. Differentiating (1) with respect to time and using the orthogonality property $R_{cb}^T R_{cb} = \mathrm{I}$, together with the assumption of rigid body motion which implies $\dot{x}_{bpb} = 0$, then results in

$$\dot{x}_{cpc} = S(\omega_{cbc})x_{cpc} + \dot{d}_{cbc} - S(\omega_{cbc})d_{cbc} \ . \tag{3}$$

The relation (3) holds for an arbitrary point p, and since the position of p as seen from inside the body, i.e. x_{bpb}, does not appear in (3), we obtain differential equations for the motion of N points x^i_{cpc} with $i \in \{1, 2 \dots N\}$, as

$$\dot{x}^i_{cpc} = S(\omega_{cbc})x^i_{cpc} + \dot{d}_{cbc} - S(\omega_{cbc})d_{cbc} \ . \tag{4}$$

The camera model used here is a frontal pinhole imaging model [24] with an image plane parallel to the x_1-x_2-plane of the c-system, a focal length f, an optical center which coincides with the origin of the c-system, a camera transformation matrix $C \in \mathbb{R}^{2\times 2}$ and an offset vector $\delta \in \mathbb{R}^{2\times 1}$. Introducing the vectors $y = \begin{pmatrix} y_1 \ y_2 \end{pmatrix}^T$ and $\xi = \begin{pmatrix} \xi_1 \ \xi_2 \end{pmatrix}^T = \begin{pmatrix} \frac{x_{cpc,1}}{x_{cpc,3}} \ \frac{x_{cpc,2}}{x_{cpc,3}} \end{pmatrix}^T$ and defining $C_f = f \cdot C$, this results in 2D image coordinates expressed in vector form as

$$y = C_f \xi + \delta \ . \tag{5}$$

For the purpose of deriving a dynamic system, for which observers can be constructed, introduce the simplified notation

$$x = x_{cpc}, \ d = d_{cbc}, \ \omega = \omega_{cbc}, \ \xi = \begin{pmatrix} \frac{x_1}{x_3} \ \frac{x_2}{x_3} \end{pmatrix}^T \ . \tag{6}$$

Further, define the matrix A and the vector b as

$$A = S(\omega), \quad b = \dot{d} - Ad \tag{7}$$

with the mapping S as defined in (2). Combining equation (3) with the output vector y given by the camera model (5) then results in the system

$$\begin{aligned} \dot{x} &= Ax + b \\ y &= C_f \xi + \delta \ . \end{aligned} \tag{8}$$

Extending the model (8) to describe the motion and observation of multiple points x^i on the same rigid object results in

$$\begin{aligned} \dot{x}^i &= Ax^i + b \\ y^i &= C_f \xi^i + \delta \end{aligned} \quad , \quad i \in \{1, 2 \dots N\} \ . \tag{9}$$

Note that the model parameters A, b, C_f and δ, as a result of the rigid body assumption and the use of a single camera, are common to all the points x^i.

3 Dynamic vision parametrization

Given x from (8), introduce the scalar parameter γ and the vector z by

$$\gamma = \frac{1}{\sqrt{x^T x}}, \quad z = \gamma x \ . \tag{10}$$

It can be seen from (10) that z is the unit vector in the direction of the 3D position x, and also that γ is the inverted distance to the feature point under consideration, i.e. the 3D point with coordinates given by x.

Differentiating γ in (10) with respect to time using (8) and the fact that $x^T Ax = 0$ since A is skew-symmetric, gives

$$\dot{\gamma} = -\gamma^2 z^T b\ . \tag{11}$$

Combining (8) with the definition of z in (10) and using (11), we further have that $\dot{z} = Az + b\gamma - z(z^T b)\gamma$. Observing that ξ, according to (6) and by the definition of z in (10), also can be expressed as $\xi = \left(\frac{z_1}{z_3}\ \frac{z_2}{z_3}\right)^T$, the dynamic system (8) can be formulated as

$$\begin{aligned} \dot{z} &= Az + (I - zz^T)b\gamma \\ y &= C_f \xi + \delta\ . \end{aligned} \tag{12}$$

Now assume that the camera is *calibrated*, i.e. that C_f and δ in (5) are known. Also assume that C_f is invertible. Since y is measured, these assumptions imply that ξ can be assumed known. By (10) and the definition of ξ in (6) the vector z can also be expressed as

$$z = \frac{1}{\sqrt{\xi_1^2 + \xi_2^2 + 1}} \left(\xi_1\ \xi_2\ 1\right)^T\ . \tag{13}$$

Thus, since ξ is assumed to be a known measurement signal also z can be assumed known. Combining (11) with the first equation in (12), and introducing

$$g_0(z) = I - zz^T \tag{14}$$

a dynamic system can be formulated as

$$\begin{aligned} \dot{z} &= Az + g_0(z)b\gamma \\ \dot{\gamma} &= -\gamma^2 z^T b\ . \end{aligned} \tag{15}$$

Similarly, for the motion of more than one point a dynamic system corresponding to (9) is obtained as

$$\begin{aligned} \dot{z}^i &= Az^i + g_0(z^i)b\gamma^i \\ \dot{\gamma}^i &= -(\gamma^i)^2 (z^i)^T b \end{aligned}\ , \quad i \in \{1, 2 \ldots N\}\ . \tag{16}$$

Equation (10) together with (15) and its multipoint version (16), constitute the desired dynamic vision parametrization. It is referred to as a parametrization rather than e.g. a coordinate transformation, since the transformation from x to z is not invertible. Instead the vector z, which, assuming a calibrated camera, is measurable according to (13), can be regarded as an alternative form of image coordinates. More specifically, the parametrization (10) can be interpreted as a projection onto a spherical image surface. This type of projection is used also in [19], where a different type of structure estimator is investigated, and where motion estimation is not considered.

4 Estimation Problem Formulations

Introduce a measurable vector $\eta \in \mathbb{R}^N$ together with a vector of unknown parameters $\theta \in \mathbb{R}^M$. A dynamic system, where the vectors η and θ together constitute the state vector, will be used as the basis for different estimation problems. The dynamic system is written as

$$\begin{aligned} \dot{\eta} &= \psi(\eta) + \phi^T(\eta)\theta \\ \dot{\theta} &= \mu(\eta, \theta) \end{aligned} \tag{17}$$

where ψ, ϕ and μ are vector-valued functions, determined from the particular estimation problem under consideration. The matrix ϕ will be denoted the *regressor* matrix.

Structure estimation - Given the motion parameters A and b, a problem of structure estimation can be formulated as the task of estimating the quantities γ^i, $i \in \{1, \dots, N\}$ in (16), which then give estimates for the three-dimensional position x for each observed feature point, using (10).

In order to relate this problem to the framework provided by (17), let the measurable vector η and the unknown parameter vector θ be given by

$$\eta = \left((z^1)^T \; (z^2)^T \dots (z^N)^T\right)^T, \quad \theta = \left(\gamma^1 \; \gamma^2 \dots \gamma^N\right)^T \tag{18}$$

Introducing a block diagonal $3N \times 3N$-matrix $\bar{A}$ with block diagonal elements A and comparing (16) with (17) it can be seen that the function $\psi(\eta)$ in (17) in this case is described by $\psi(\eta) = \bar{A}\eta$. It can further be seen that the regressor matrix $\phi(\eta)$ is given by a block diagonal matrix $\phi(\eta) = \text{diag}\left(g_0(z^i)b\right)$, $1 \le i \le N$ and that the function $\mu(\eta, \theta)$ defining the parameter dynamics is

$$\mu(\eta, \theta) = \left(-(\gamma^1)^2(z^1)^T b \dots -(\gamma^N)^2(z^N)^T b\right)^T \,. \tag{19}$$

Estimation of structure and angular velocity - Assuming a known linear velocity vector b, the problem of estimating the angular velocity ω as well as the structure parameters γ^i, $i \in \{1, \dots, N\}$, is considered. The angular velocity ω is related to the skew-symmetric matrix A in (16) according to (7). Using $\omega \times z = -z \times \omega$ which implies $S(\omega)z = -S(z)\omega$, the system (16) can be written as

$$\begin{aligned} \dot{z}^i &= -S(z^i)\omega + g_0(z^i)b\gamma^i \\ \dot{\gamma}^i &= -(\gamma^i)^2(z^i)^T b \end{aligned} \quad , \quad i \in \{1, 2 \dots N\} \,. \tag{20}$$

The parameter vector θ in (17) is here given by $\theta = \left(\omega^T \, \gamma^1 \, \gamma^2 \dots \gamma^N\right)^T$. Again using the measurable vector η as defined in (18), and comparing (17) and (20), we get $\psi(\eta) = 0$ and the regressor matrix

$$\phi(\eta) = \begin{pmatrix} -S(z^1) & g_0(z^1)b & \dots & \dots & 0 \\ -S(z^2) & 0 & g_0(z^2)b & \dots & 0 \\ \vdots & \vdots & \dots & \ddots & \vdots \\ -S(z^N) & 0 & \dots & \dots & g_0(z^N)b \end{pmatrix}^T$$

Assuming that the angular velocity evolves according to a dynamic system of the form $\dot{\omega} = \mu_\omega(\omega)$, the parameter dynamics $\mu(\eta, \theta)$ in (17) can be expressed as

$$\mu(\eta,\theta) = \left(\mu_\omega(\omega)^T \; -(\gamma^1)^2(z^1)^T b \; \ldots \; -(\gamma^N)^2(z^N)^T b\right)^T$$

Estimation of structure and linear velocity - Assuming a known angular velocity ω, the problem of estimating the linear velocity b as well as the structure parameters γ^i, $i \in \{1,\ldots,N\}$, can be formulated using (17), however augmented with an additional *scaling condition.* The need for the scaling condition can be seen from (16), where for each i, the linear velocity b only appears in the product $\gamma^i b$. Consequently, γ^i and b cannot be distinguished without the use of additional constraints. Introducing $\beta^i = \gamma^i b$, and assuming that the vector b evolves according to a dynamic system $\gamma^i \dot{b} = \mu_\beta(\beta^i)$, the dynamic system (16) is reformulated, using the equation for $\dot{\gamma}^i$ in (16), as

$$\begin{aligned} \dot{z}^i &= Az^i + g_0(z^i)\beta^i \\ \dot{\beta}^i &= -((z^i)^T\beta^i)\beta^i + \mu_\beta(\beta^i) \end{aligned} \quad , \quad i \in \{1, 2 \ldots N\} \; . \tag{21}$$

Define normalized values of the parameters γ^i as

$$\alpha^i = \frac{\gamma^i}{\gamma^1}, \quad i \in \{2, \ldots, N\} \; . \tag{22}$$

The problem of estimating the linear velocity b as well as the structure parameters γ^i, $i \in \{1,\ldots,N\}$ can be formulated using two dynamic systems having the form (17). The first dynamic system is formulated for the purpose of estimating β^1, and the second dynamic system is formulated for the purpose of estimating α^i, $i \in \{2,\ldots,N\}$. For this estimation task, the relation $\beta^i = \gamma^i b = \frac{\gamma^i}{\gamma^1}\gamma^1 b = \alpha^i\beta^1$ is used to reformulate (21) for the purpose of estimating α^i, $i \in \{2,\ldots,N\}$, given an estimate of β^1.

The parameters to be estimated, i.e. the quantities β^1 and α^i, $i \in \{2,\ldots,N\}$, need to be combined with a scaling condition, for the purpose of computing the structure parameters γ^i. A scaling condition can be derived e.g. from assumed knowledge of the distance between two object points. Assuming a distance d between two points x^1 and x^2, i.e. $(x^1 - x^2)^T(x^1 - x^2) = d^2$, we get, using $z^i = \gamma^i x^i$ according to (10) together with (22), that

$$(z^1 - \frac{1}{\alpha^2}z^2)^T(z^1 - \frac{1}{\alpha^2}z^2) = (\gamma^1)^2 d^2 \tag{23}$$

Since z^1 and z^2 are measurable, equation (23) shows that γ^1 can be computed, given an estimated value of α^2 and an assumed value of the distance d. The remaining γ^i, $i \in \{2,\ldots,N\}$ can then be computed using (22).

Structure and motion estimation Following the strategy described above, where two dynamic systems on the form (17) are used, together with a scaling condition for the purpose of estimating structure and linear velocity assuming knowledge of the angular velocity, a formulation for estimation of angular and linear velocity as well as structure can be derived. Here we choose to extend the first dynamic system for the purpose of estimating also angular velocity. The parameter vector becomes $\theta_1 = \left(\omega^T \ (\beta^1)^T\right)^T$ and the regressor matrix is given by $\phi_1(\eta_1) = \left(-S(z^1) \ g_0(z^1)\right)^T$. The second dynamic system uses the estimated β^1 as well as the estimated ω which then replaces a known angular velocity with an estimated angular velocity. Finally, a scaling condition e.g. (23) is used for the computation of γ^1, and hence the remaining γ^i, $i \in \{2, \dots, N\}$ can be computed using (22).

5 Nonlinear and adaptive observers

Introduce a matrix F which is Hurwitz, and a symmetric positive definite matrix Q. A symmetric positive definite matrix P can then be computed as the unique solution to the Lyapunov equation [25],

$$F^T P + PF = -Q \ . \tag{24}$$

Introducing the estimated quantities $\hat{\eta}$ and $\hat{\theta}$, an estimator for (17) can then be formulated as

$$\begin{aligned} \dot{\hat{\eta}} &= \psi(z) + F(\hat{\eta} - \eta) + \phi^T(\eta)\hat{\theta} \\ \dot{\hat{\theta}} &= -\phi(\eta)P(\hat{\eta} - \eta) + \mu(\eta, \hat{\theta}) \ . \end{aligned} \tag{25}$$

The estimator (25) constitutes an extension of the estimator presented in [26], which in turn is based on [27]. The extension is here due to θ not being a constant parameter, as is assumed in [26]. Therefore, the second equation in (25) contains a correction term $\mu(\eta, \hat{\theta})$ which is not present in the estimator in [26].

The estimator (25) is formulated with reference to the dynamic system (17). Section 4 describes how the dynamic system (17) can be used as the basis for formulating different estimation problems in dynamic vision. The estimator (25) can therefore be used as a common estimator structure for the problems considered, by selecting the vectors η and θ and the functions ψ, ϕ and μ as described in Section 4.

The matrices F and Q can be regarded as *tuning parameters* for the estimator (25).

6 Stability analysis

A stability analysis for the case of structure estimation is presented. The analysis shows asymptotic stability of the estimator, hence giving a proof of convergence in the sense that if the initial structure estimate is close enough to its true value, the estimate will converge as the time $t \to \infty$.

For the case of structure estimation the dynamic system (17) is given by (15). The observer (25) is rewritten, using the notation $\hat{z}$ and $\hat{\gamma}$ for $\hat{\eta}$ and $\hat{\theta}$

respectively, as

$$\begin{aligned}\dot{\hat{z}} &= F\tilde{z} + Az + g_0(z)b\hat{\gamma} \\ \dot{\hat{\gamma}} &= -b^T g_0(z)^T P\tilde{z} - \hat{\gamma}^2 z^T b\end{aligned} \tag{26}$$

Introducing the estimation errors $\tilde{z} = \hat{z} - z$ and $\tilde{\gamma} = \hat{\gamma} - \gamma$, and using (15) and (26) and the relation $\gamma^2 - \hat{\gamma}^2 = -\tilde{\gamma}(\tilde{\gamma} + 2\gamma)$, gives the error equations

$$\begin{aligned}\dot{\tilde{z}} &= F\tilde{z} + g_0(z)b\tilde{\gamma} \\ \dot{\tilde{\gamma}} &= -b^T g_0(z)^T P\tilde{z} - \tilde{\gamma}(\tilde{\gamma} + 2\gamma)z^T b\end{aligned} \tag{27}$$

The error equations are considered asymptotically stable if $\tilde{z} \to 0$ and $\tilde{\gamma} \to 0$ as $t \to \infty$. We will show stability of the error equations by Lyapunov's indirect method, where stability of a nonlinear system is deduced from the stability of a linear system, obtained from a *linearization* of the nonlinear system [25]. Linearization of (27) gives

$$\begin{aligned}\dot{\tilde{z}} &= F\tilde{z} + g_0(z)b\tilde{\gamma} \\ \dot{\tilde{\gamma}} &= -b^T g_0(z)^T P\tilde{z} - 2\gamma\tilde{\gamma}z^T b\end{aligned} \tag{28}$$

As a first step, we show that $\tilde{z} \to 0$ and that $\tilde{\gamma}$ is bounded. Introduce the scalar function

$$V(\tilde{z}, \tilde{\gamma}) = \frac{1}{2\gamma^4}\left(\tilde{z}^T P\tilde{z} + \tilde{\gamma}^2\right) \tag{29}$$

Differentiating (29) along the trajectories of (28), gives

$$\dot{V}(\tilde{z}, \tilde{\gamma}) = -\frac{2}{\gamma^5}\dot{\gamma}\left(\tilde{z}^T P\tilde{z} + \tilde{\gamma}^2\right) + \frac{1}{2\gamma^4}\left(2\tilde{z}^T P\dot{\tilde{z}} + 2\tilde{\gamma}\dot{\tilde{\gamma}}\right) \tag{30}$$

which using the second equation in (15) together with (28) and (24), becomes

$$\dot{V}(\tilde{z}, \tilde{\gamma}) = \tilde{z}^T\left(\frac{2}{\gamma^3}z^T bP - \frac{1}{2\gamma^4}Q\right)\tilde{z} \tag{31}$$

Assuming that the matrices Q and P are such that the matrix

$$\frac{2}{\gamma^3}z^T bP - \frac{1}{2\gamma^4}Q \tag{32}$$

is negative definite for the object motions considered, i.e. for the range of values for z, b and γ, we get $\dot{V}(\tilde{z}, \tilde{\gamma}) \leq 0$ from which we can conclude that $\tilde{z} \to 0$ and that $\tilde{\gamma}$ is bounded, e.g. [25, 28]. In order to show that also $\tilde{\gamma} \to 0$, a reasoning inspired by a stability proof in [28] will be used. First, a function $\varphi(t)$ is defined as

$$\varphi(t) = \frac{1}{2}(\tilde{\gamma}(t+T)^2 - \tilde{\gamma}(t)^2) \tag{33}$$

The function $\varphi(t)$ is bounded since $\tilde{\gamma}$ is bounded. The time derivative $\dot{\varphi}(t)$ becomes

$$\dot{\varphi}(t) = \tilde{\gamma}(t+T)\dot{\tilde{\gamma}}(t+T) - \tilde{\gamma}(t)\dot{\tilde{\gamma}}(t) = \int_t^{t+T} \frac{d}{d\tau}(\tilde{\gamma}(\tau)\dot{\tilde{\gamma}}(\tau))d\tau \tag{34}$$

The integral in (34) can be rewritten, using (28), resulting in

$$\begin{aligned}\dot{\varphi}(t) = &-\int_t^{t+T} \tilde{\gamma}(\tau)b(\tau)^T g_0(z(\tau))^T P g_0(z(\tau))b(\tau)\tilde{\gamma}(\tau)d\tau \\ &-\int_t^{t+T} \tilde{\gamma}(\tau)b(\tau)^T g_0(z(\tau))^T PF\tilde{z}(\tau)d\tau - \int_t^{t+T} \frac{d}{d\tau}(\tilde{\gamma}(\tau)b(\tau)^T g_0(z(\tau))^T P)\tilde{z}(\tau)d\tau \\ &-\int_t^{t+T} \frac{d}{d\tau}(2\gamma(\tau)\tilde{\gamma}(\tau)z(\tau)^T b(\tau))d\tau\end{aligned} \tag{35}$$

If we now assume a *persistency of excitation condition*, also denoted PE condition, i.e. that for all t and T, there is a positive number k such that

$$\int_t^{t+T} b(\tau)^T g_0(z(\tau))^T g_0(z(\tau))b(\tau)d\tau \geq kI \tag{36}$$

we can show that $\tilde{\gamma} \to 0$ by the following reasoning. The PE condition (36) implies that the first integral in (35) fulfils

$$\int_t^{t+T} \tilde{\gamma}(\tau)b(\tau)^T g_0(z(\tau))^T P g_0(z(\tau))b(\tau)\tilde{\gamma}(\tau)d\tau > 0 \tag{37}$$

Assume now, contradictory to what we would like to show, that it is not the case that $\tilde{\gamma} \to 0$. We then see that, for t large enough, the first and last integrals in (35) will dominate over the two remaining integrals, which both tend to zero since $\tilde{z} \to 0$. Hence it is possible to achieve, possibly by selecting Q such that the eigenvalues of P are sufficiently large (which can be done, according to e.g. [29]), that for t large enough, we get $\dot{\varphi}(t) < 0$ which contradicts the fact that $\varphi(t)$ is bounded. Hence, $\tilde{\gamma} \to 0$.

To summarize, we have shown asymptotical stability of the error equations (27) by using the corresponding linearized equations (28), provided the matrix (32) is negative definite for the motions considered and also that the PE condition (36) holds.

The PE condition (36) can be interpreted in terms of unfavourable motions, using the following reasoning. From (14) and the observation that z by its definition (10) is a vector of unit length, it can be seen that if $b \mid\mid z$ over some time interval, the integrand of (36) will be identically zero over that interval, and the PE condition will not be fulfilled. Except for the case $b = 0$, this means that the translational velocity of the observed point relative to the camera should not be directed along a straight line connecting the camera center with the current measurement z during any time interval. That the case $b \mid\mid z$ results in an unfavorable motion from an observability point of view can also be seen directly from (15), since such a b would disrupt the influence of γ on the z-dynamics, and thus render the parameter identification process infeasible, since then changes in the parameter γ cannot be observed through the dynamics of the available signal z. It can also be seen that $b \mid\mid z$ implies that provided $b_3 \neq 0$, it holds that $\begin{pmatrix} y_1 & y_2 \end{pmatrix}^T = \begin{pmatrix} b_1/b_3 & b_2/b_3 \end{pmatrix}^T$, which is the *focus of expansion* mentioned as an unobservable case also for the observers presented in e.g. [13, 15, 16, 20].

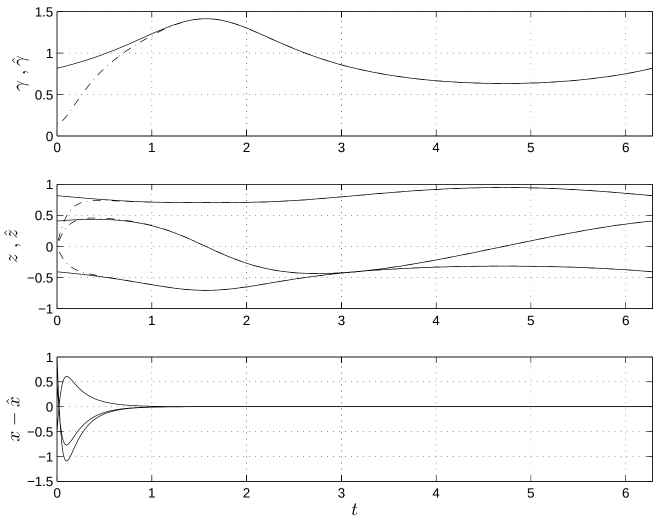

Fig. 2. Structure estimation results, using a circular camera motion. True (solid) and estimated (dash-dotted) values of γ (top) and z (middle). Estimation error for x (bottom).

7 Simulation examples

The performance of the estimator (25) is demonstrated in simulation examples. In the simulations, we use $F = -10 \cdot I$ and $Q = 750 \cdot I$ as the matrices employed to determine the matrix P from (24). A *structure estimation* example is first presented, using the estimator (25) applied to the dynamic system (17). The camera is moving in a circle of radius 1 in the x_1-x_2-plane of the inertial coordinate system, around an observed object located at the center of the circle. The optical axis of the camera, i.e. the x_3-axis, is kept oriented towards the origin of the inertial system. The object is assumed fixed, and we can therefore for simplicity define the inertial system to coincide with the object coordinate system. In the dynamic systems formulation, this motion can be described using the parameter vectors $\omega = \begin{pmatrix}-1 \; 0 \; 0\end{pmatrix}^T$ and $b = \begin{pmatrix}0 \; -1 \; 0\end{pmatrix}^T$. A single feature point is observed, with an initial position selected as $x_0 = \begin{pmatrix}-0.5 \; 0.5 \; 1\end{pmatrix}^T$. The chosen initial position represents a point which is not located at the center of the circular motion. The estimation results are shown in Fig. 2. As can be seen in Fig. 2, the estimates converge nicely towards the true values. The simulation illustrates how structure is estimated for a motion giving rise to time-varying γ and z. A similar simulation example using the same observer, where also noise is added, can be found in [30].

A *structure and motion estimation* example is presented, using the estimator (25) applied to the dynamic system (17), for estimation of angular and linear velocity as well as three-dimensional position. The observed object contains two feature points, executing a periodic motion used also in [23], governed by the

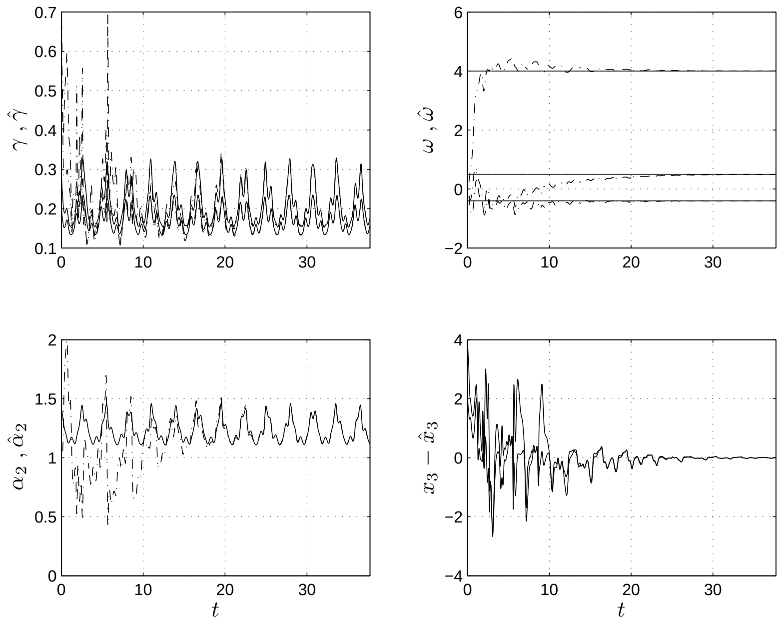

Fig. 3. Structure and motion estimation results, for a periodic object motion as in [23]. True (solid) and estimated (dash-dotted) values of γ (upper left), α^2 (lower left) and ω (upper right). The depth errors are shown in the lower right plot.

parameter vectors $\omega = \begin{pmatrix}-0.4 \; 0.5 \; 4\end{pmatrix}^T$ and $b = \begin{pmatrix}0 \; 2\pi\sin(2\pi t) \; 2\pi\cos(2\pi t)\end{pmatrix}^T$. The initial values for the object points are given by $x_0 = \begin{pmatrix}2 \; 2 \; 4 \; 2 \; 2 \; 2\end{pmatrix}^T$, with a distance $d = 2$ between the two points. The estimation results are shown in Fig. 3, where it can be seen in the lower right plot that the three-dimensional position is recovered. It can also be seen, in the upper right plot, that the estimated angular velocity converges to its correct value. The upper left plot shows that the estimated parameters $\hat{\gamma}$ converge to the actual parameters γ for the two points, as is the case for α^2 as seen in the lower left plot.

8 Conclusions

Estimation of 3D structure and motion from 2D images can be performed using a dynamic systems formulation, where nonlinear and adaptive observers can be utilized for estimation of states and parameters. In this paper we have demonstrated how a single parametrization of the underlying perspective dynamic system can be used for formulation of estimation problems for structure as well as motion, and also how a common observer structure can be used for the estimation problems considered.

Observers are derived for estimation of structure for known motion, for estimation of angular or linear velocity and structure, and for estimation of angular and linear velocity as well as structure. The observers estimating linear velocity use a scaling condition, which resolves the well-known scale ambiguity of the 3D reconstruction formulation. A stability analysis is presented, showing convergence of the structure estimator. Simulations are presented, illustrating for the case of estimating linear and angular velocity as well as structure, how a nonlinear observer can be used for motion estimation as well as recovery of three-dimensional position in a monocular vision system, using measurements from two-dimensional images.

The proposed nonlinear observer is able to estimate structure and motion, using as few as two feature points. This, however, requires that certain dynamic properties of the angular and linear velocities are known. The observer thus demonstrates a trade-off compared to a more computer vision oriented approach, where no specific assumptions regarding the motion dynamics are required, but instead additional feature points are needed.

References

1. Hartley, R., Zisserman, A.: Multiple View Geometry. Cambridge University Press (2003)
2. Soatto, S., Perona, P.: Reducing structure from motion: A general framework for dynamic vision, part 1: Modeling. IEEE Transactions on Pattern Analysis and Machine Intelligence **20**(9) (September 1998)
3. Azarbayejani, A., Pentland, A.P.: Recursive estimation of motion, structure, and focal length. IEEE Transactions on Pattern Analysis and Machine Intelligence **17**(6) (1995) 562–575
4. Ghosh, B.K., Inaba, H., Takahashi, S.: Identification of riccati dynamics under perspective and orthographic observations. IEEE Transactions on Automatic Control **45**(7) (July 2000) 1267–1278
5. Kahl, F.: Multiple view geometry and the L_∞-norm. In: International Conference on Computer Vision, IEEE Computer Society Press (2005) 1002–1009
6. Sim, K., Hartley, R.: Recovering camera motion using linfty minimization. In: Computer Vision and Pattern Recognition. Volume 1., IEEE Computer Society Press (2006) 1230–1237
7. Mahamud, S., Hebert, M., Omori, Y., Ponce, J.: Provably-convergent iterative methods for projective structure from motion. In: Computer Vision and Pattern Recognition. Volume 1., IEEE Computer Society Press (2001) 1018–1025
8. Oliensis, J.: Multiframe structure from motion in perspective. In: International Workshop on Representation of Visual Scenes, IEEE Computer Society Press (1995) 77–84 In conjunction with ICCV 2005.
9. Ghosh, B.K., Loucks, E.P., Jankovic, M.: An introduction to perspective observability and recursive identification problems in machine vision. In: Proceedings of the 33rd Conference on Decision and Control. (December 1994)
10. Soatto, S., Frezza, R., Perona, P.: Motion estimation via dynamic vision. IEEE Transactions on Automatic Control **41**(3) (March 1996)
11. Soatto, S.: 3-d structure from visual motion: Modeling, representation and observability. Automatica **33**(7) (1997) 1287–1312
12. Matthies, L., Kanade, T., Szeliski, R.: Kalman filter-based algorithms for estimating depth from image sequences. International Journal of Computer Vision **3** (1989) 209–236

13. Jankovic, M., Ghosh, B.K.: Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear observer. Systems & Control Letters **25** (1995) 63–73
14. Matveev, A., Hu, X., Frezza, R., Rehbinder, H.: Observers for systems with implicit output. IEEE Transactions on Automatic Control **45**(1) (January 2000) 168–173
15. Chen, X., Kano, H.: A new state observer for perspective systems. IEEE Transactions on Automatic Control **47**(4) (April 2002) 658–663
16. Dixon, W.E., Fang, Y., Dawson, D.M., Flynn, T.J.: Range identification for perspective vision systems. IEEE Transactions on Automatic Control **48**(12) (December 2003) 2232–2238
17. Dahl, O., Nyberg, F., Holst, J., Heyden, A.: Linear design of a nonlinear observer for perspective systems. In: Proc. of ICRA'05 - 2005 IEEE Conference on Robotics and Automation. (April 2005)
18. Abdursul, R., Inaba, H., Ghosh, B.K.: Nonlinear observers for perspective time-invariant linear systems. Automatica **40** (2004) 481–490
19. Ma, L., Chen, Y., Moore, K.L.: Range identification for perspective dynamic systems with 3d imaging surfaces. In: American Control Conference. (June 2005)
20. Karagiannis, D., Astolfi, A.: A new solution to the problem of range identification in perspective vision systems. IEEE Transactions on Automatic Control **50**(12) (December 2005) 2074–2077
21. Gupta, S., Aiken, D., Hu, G., Dixon, W.E.: Lyapunov-based range and motion identification for a nonaffine perspective dynamic system. In: American Control Conference. (June 2006)
22. Martino, D.D., Germani, A., Manes, C., Palumbo, P.: Design of observers for systems with rational output function. In: Proceedings of the 45th Conference on Decision and Control. (December 2006)
23. Chen, X., Kano, H.: State observer for a class of nonlinear systems and its application to machine vision. IEEE Transactions on Automatic Control **49**(11) (November 2004) 2085–2091
24. Ma, Y., Soatto, S., Košecká, J., Sastry, S.S.: An Invitation to 3-D Vision. Springer-Verlag (2004)
25. Khalil, H.K.: Nonlinear Systems. Pearson Education, Inc. (2000)
26. Teel, A., Kadiyala, R., Kokotovic, P., Sastry, S.: Indirect techniques for adaptive input-output linearization of non-linear systems. Int. Journal of Control **53**(1) (1991) 192–222
27. Kudva, P., Narendra, K.S.: Synthesis of an adaptive observer using Lyapunov's direct method. Int. Journal of Control **18**(6) (1973) 1201–1210
28. Marino, R., Tomei, P.: Nonlinear Control Design - Geometric, Adaptive and Robust. Prentice Hall (1995)
29. Komaroff, N.: Simultaneous eigenvalue lower bounds for the lyapunov matrix equation. IEEE Transactions on Automatic Control **33**(1) (January 1988)
30. Dahl, O., Nyberg, F., Heyden, A.: Nonlinear and adaptive observers for perspective dynamic systems. In: American Control Conference. (July 2007)