ELSEVIER Linear Algebra and its Applications 351–352 (2002) 701–717

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Canonical forms and parameter identification problems in perspective systems ☆

Satoru Takahashi *, Bijoy K. Ghosh

Department of Systems Science and Mathematics, Washington University, St Louis, MO 63130, USA

Received 20 February 2001; accepted 21 October 2001

Submitted by J. Rosenthal

Abstract

We introduce canonical forms for perspective dynamical systems under the action of a perspective group and illustrate their application to parameter identification with the aid of a single *charged coupled device* camera. We show that the parameters in the canonical form can be identified uniquely using an Extended Kalman Filter. © 2002 Elsevier Science Inc. All rights reserved.

Keywords: Perspective dynamical systems; Canonical forms; Kronecker index; Parameter identification; Extended Kalman Filter

1. Introduction

An important and somewhat difficult problem in machine vision is to identify parameters of motion dynamics from observing projection of feature points on the image plane observed over time. The difficulty arises from the fact that the observation function is rational (perspective) and the underlying sensor (a *charged coupled device* camera) is noisy. In fact, not all the parameters are identifiable and in [2–4], the identifiable parameters have been characterized via orbits of a *perspective group*. In this paper, we introduce a suitable canonical form (see [7,8,10–13,15,16]), and illustrate that these canonical forms can be used to identify parameters up to orbits of the underlying perspective group or a suitable subgroup of this group as the case may

☆ This research is supported in part by NSF under grant no. 9720357.

* Corresponding author.

E-mail addressess: saru@zach.wustl.edu (S. Takahashi), ghosh@netra.wustl.edu (B.K. Ghosh).

PII: S0024-3795(01)00557-2

be. To illustrate the main concept, we consider a homogeneous finite dimensional system of the form

$$\begin{aligned}\dot{\mathscr{X}} &= \mathscr{A}^{\mathrm{T}}\mathscr{X}, \quad \mathscr{X}(0) = \mathscr{X}_0,\\ \mathscr{Y} &= \mathscr{C}\mathscr{X},\end{aligned} \tag{1.1}$$

where $\mathscr{X} \in \mathbb{RP}^{n-1}$ and $\mathscr{Y} \in \mathbb{RP}^{m-1}$, the real projective spaces of dimensions $n-1$ and $m-1$, respectively. The underlying parameter identification problem of interest is described as follows.

Problem 1. Assume that we observe $\mathscr{Y}(t)$ in an unspecified interval $[0, T)$, the problem is to determine the extent to which the parameters $\mathscr{A}^{\mathrm{T}}$, $\mathscr{C}$ and the initial condition $\mathscr{X}_0$ are identifiable from the observed data.

In order to show connection between Problem 1 and a specific parameter identification problem in machine vision, we assume that a rigid object is moving in space and it is observed with the aid of a charged coupled device camera. Let us assume that (X, Y, Z) are the three coordinates of a feature point on the rigid object. Since the object is moving, we assume that the feature point on the object satisfies a differential equation of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}X\\Y\\Z\end{pmatrix} = \begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}X\\Y\\Z\end{pmatrix} + \begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}. \tag{1.2}$$

The motion dynamics (1.2) has already been introduced earlier in the literature [6] and is described as *affine dynamics*. In general, if the matrix $A = \{a_{ij}\}$ is assumed to be skewsymmetric, i.e. if $a_{ii} = 0$, $i = 1, 2, 3$, and if $a_{ij} = -a_{ji}$, the affine equation (1.2) reduces to what is called a *rigid motion dynamics*. In order to see the connection between the affine/rigid dynamics (1.2) and the homogeneous system (1.1), we homogenize the state variable (X, Y, Z) as

$$X = \frac{X_1}{W_1}, \quad Y = \frac{Y_1}{W_1}, \quad Z = \frac{Z_1}{W_1}. \tag{1.3}$$

$[X_1, Y_1, Z_1, W_1]^{\mathrm{T}}$ defines a point in $\mathbb{RP}^3$ in homogeneous coordinates. The affine dynamics (1.2) can be rewritten as the following homogeneous system [6] in $\mathbb{RP}^3$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}X_1\\Y_1\\Z_1\\W_1\end{bmatrix} = \begin{pmatrix}a_{11} & a_{12} & a_{13} & b_1\\ a_{21} & a_{22} & a_{23} & b_2\\ a_{31} & a_{32} & a_{33} & b_3\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{bmatrix}X_1\\Y_1\\Z_1\\W_1\end{bmatrix}. \tag{1.4}$$

We shall now describe the observation equation associated with (1.4). Recall that the feature point (X, Y, Z) is observed with the aid of a *charged coupled device* camera. There are two different projection models well known, see [3,14] for details. In the *perspective projection model*, it is assumed that the point (X, Y, Z) is projected as

$$(X, Y, Z) \longmapsto \left(\frac{X}{Z}, \frac{Y}{Z}\right), \quad Z \neq 0, \tag{1.5}$$

whereas in the *orthographic projection model* the point (X, Y, Z) is projected as

$$(X, Y, Z) \longmapsto (X, Y). \tag{1.6}$$

If we assume that $[y_1 \;\; y_2 \;\; y_3]^{\mathrm{T}}$ is the homogeneous coordinate vector of the observed output, it would follow that

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{pmatrix} \theta & 0 & 0 & 0 \\ 0 & \theta & 0 & 0 \\ 0 & 0 & \theta & 0 \end{pmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix} \tag{1.7}$$

under the *perspective projection model* and

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{pmatrix} \theta & 0 & 0 & 0 \\ 0 & \theta & 0 & 0 \\ 0 & 0 & 0 & \theta \end{pmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix} \tag{1.8}$$

under the *orthographic projection model* where θ is an arbitrary nonzero parameter. The homogeneous equation (1.4) together with the homogeneous observation function (1.7) or (1.8) gives rise to an example of a homogeneous dynamical system of the form (1.1). Other examples of homogeneous dynamical systems have been described in [3].

Perspective problems have already been considered in the literature by Kanatani [14] where the goal is to identify parameters of motion using a charged coupled device camera. Generically it is known [4] that parameters can be identified up to orbits of a perspective group. To illustrate the action of the perspective group on the dynamical system (1.1), note that one can change basis in the state space and consider a new state variable $\mathscr{Z} = P\mathscr{X}$. This would scale the pair $(\mathscr{A}^{\mathrm{T}}, \mathscr{C})$ to a new pair $(P\mathscr{A}^{\mathrm{T}}P^{-1}, \mathscr{C}P^{-1})$ where P is an $n \times n$ nonsingular matrix. In addition to the above described $\mathrm{GL}(n)$ action, there are two additional actions described as follows. For $\lambda, \mu \in \mathscr{R}$, $\mu \neq 0$, we scale the pair $(\mathscr{A}^{\mathrm{T}}, \mathscr{C})$ to a new pair $(\lambda I + \mathscr{A}^{\mathrm{T}}, \mu\mathscr{C})$. The new resulting output of (1.1) is given by

$$\mathscr{Y} = \mu\mathscr{C}\mathrm{e}^{A^{\mathrm{T}}t}\mathscr{X}_0 = (\mu\mathrm{e}^{\lambda t})\mathscr{C}\mathrm{e}^{A^{\mathrm{T}}t}\mathscr{X}_0 = \mu\mathrm{e}^{\lambda t}\mathscr{C}\mathscr{X}. \tag{1.9}$$

In homogeneous coordinates, (1.9) is same as the output of (1.1).

The perspective group is defined as the direct product of $\mathrm{GL}(n)$, $\mathbb{R}$ and $\mathbb{R}^+$ (the set of nonzero real numbers), where $\mathbb{R}$ is a group under addition of real numbers and $\mathbb{R}^+$ is a group under multiplication of real numbers. The scalings on the parameters $(\mathscr{A}^{\mathrm{T}}, \mathscr{C})$ as a result of the action of the perspective group is described as follows:

$$\begin{array}{rccc} \mathscr{X}_1: & \mathrm{GL}(n) \times \mathscr{P} & \longrightarrow & \mathscr{P} \\ & (P, \mathscr{A}^{\mathrm{T}}, \mathscr{C}) & \longmapsto & (P\mathscr{A}^{\mathrm{T}}P^{-1}, \mathscr{C}P^{-1}) \end{array} \tag{1.10}$$

$$\begin{array}{rccc} \mathscr{X}_2: & \mathbb{R} \times \mathscr{P} & \longrightarrow & \mathscr{P} \\ & (\lambda, \mathscr{A}^{\mathrm{T}}, \mathscr{C}) & \longmapsto & (\lambda I + \mathscr{A}^{\mathrm{T}}, \mathscr{C}) \end{array} \tag{1.11}$$

$$\begin{array}{llll} \mathscr{X}_3: & \mathbb{R}^+ \times \mathscr{P} & \longrightarrow & \mathscr{P} \\ & (\mu, \mathscr{A}^{\mathrm{T}}, \mathscr{C}) & \longmapsto & (\mathscr{A}^{\mathrm{T}}, \mu\mathscr{C}) \end{array} \tag{1.12}$$

where $\mathscr{P}$ denotes the pairs of matrices $(\mathscr{A}^{\mathrm{T}}, \mathscr{C})$ on which the perspective group acts. As a result of the action of the perspective group on the parameter space, the space is split up into orbits.

It may be trivially verified that parameters in the same orbit of the perspective group cannot be observed using (1.7) or (1.8). Furthermore it has been shown by Ghosh et al. [2–4] that generically the orbits of the perspective group are indeed observable, i.e. produces distinct outputs for at least some interval of time. The generic assumption is hard to check and tied to verifying the minimality of the homogeneous dynamical system. When $m \geqslant n$, one can obtain an explicit description of the generic set that guarantees minimality of (1.1) as a homogeneous system (see [5]). When $m < n$, which is perhaps the case most often, one would rely on a suitable *realization theory* for homogeneous systems and this problem is connected to the *rational exponential interpolation problem* described in [6]. In order to actually carry out the problem of observing the orbits of the perspective group, it is essential to parameterize the orbits by describing a chart with an appropriate set of coordinates and construct an observer that computes an estimate of the coordinates. In this paper we consider primarily the parameterization problem but also illustrate our results with observers that are derived from Extended Kalman Filters.

2. Background and motivation

The class of motion dynamics that we consider throughout this paper is a quadratic extension of the affine flow described by (1.2) and is given by the following:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} f_1 & f_2 & f_3 & 0 & 0 & 0 \\ 0 & f_1 & 0 & f_2 & f_3 & 0 \\ 0 & 0 & f_1 & 0 & f_2 & f_3 \end{pmatrix}\begin{pmatrix} X^2 \\ XY \\ XZ \\ Y^2 \\ YZ \\ Z^2 \end{pmatrix}, \tag{2.1}$$

where let $A = (a_{ij}),\ b = (b_1\ \ b_2\ \ b_3)^{\mathrm{T}}$ and $f^{\mathrm{T}} = (-f_1\ \ -f_2\ \ -f_3)$, $i, j = 1, 2, 3$. The dynamical system (2.1) is a Riccati dynamics in $\mathbb{R}^3$, which has been illustrated in [2]. It is a class of *quadratic motion models* more general than a *rigid and affine flow* which preserves *shape*, i.e. the shape of a planar surface remains planar, although the distance between two points on the plane may not remain constant. An important question, that is of interest in machine vision is to ask the following.

Problem 2. Consider the dynamical system (2.1) and assume that the state vector (X, Y, Z) is observed by the perspective projection (1.5) or the orthographic projection (1.6), to what extent are the motion parameters and the initial conditions $X(0), Y(0), Z(0)$ identifiable from the perspective and orthographic observations described in (1.5), (1.6) over a given time interval $[0, T), T > 0$?

The Riccati dynamics (2.1) can be easily homogenized in the notation described in (1.3) as the following:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix} = \begin{pmatrix} A & b \\ f^{\mathrm{T}} & 0 \end{pmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix}. \tag{2.2}$$

The observation functions (1.5), (1.6) have already been defined in homogeneous coordinates by (1.7), (1.8). If only the first component of the observation function (1.5) is available, we can write

$$\begin{bmatrix} y_1 \\ y_3 \end{bmatrix} = \begin{pmatrix} \theta & 0 & 0 & 0 \\ 0 & 0 & \theta & 0 \end{pmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix}. \tag{2.3}$$

Note that the pairs (2.2), (1.7) and (2.2), (2.3) are examples of homogeneous systems of the form (1.1). If we consider orthographic projection (1.6) instead of the perspective projection (1.5) then the homogeneous observation function (1.7) has to be changed (1.8).

The contributions of this paper are now discussed. We consider homogeneous dynamical system (2.2) under various observation functions of the form (1.7), (2.3) or (1.8). In Section 3, we assume that the coordinate system with respect to which the dynamics (2.2) has been defined is unknown. This in turn implies that the matrix $\mathscr{C}$ in (1.1) using which the observation function has been defined is unknown. This would be referred to as the case of unknown camera calibration. In Section 4, we assume that the position of the camera in the above coordinate system is known and that the optical axis of the camera is precisely along the Z coordinate. In Sections 3 and 4 we obtain a suitable canonical form that would describe the orbit under action of the perspective group. Using the obtained canonical form, we solve the orbit identification problem using Extended Kalman Filters.

3. The case of unknown camera calibration

For the case when the calibration parameter is not known we consider the dynamical system (1.1) together with observation function of the form (1.7) or (2.3). We assume that $\mathscr{A}^{\mathrm{T}}$ is an arbitrary 4×4 matrix and $\mathscr{C}$ is an arbitrary 3×4 or 2×4 matrix depending upon the choice of the observation functions (1.7) or (2.3).

3.1. Parameterization of the orbit under the perspective group

We begin this section by considering the case $n = 4$, $m = 2$ and assume that the associated Kronecker indices for the pair $\mathscr{A}^{\mathrm{T}}, \mathscr{C}$ are given by $\kappa_1 = \kappa_2 = 2$ (see [10,11] for a definition of Kronecker indices). First of all we define the matrix $\mathscr{B} = \mathscr{C}^{\mathrm{T}}$. Assuming the Kronecker pair (2,2), it follows that via the group action described in (1.10), the pair $\mathscr{A}, \mathscr{B}$ would take the form

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 0 & 0 & 1 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & \gamma \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.1}$$

where the parameters $\alpha_i, \beta_i, i = 1, 2, 3, 4$, and γ are arbitrary. Writing $\mathscr{B}$ as $(b_1 \;\; b_2)$ we assume that

$$\begin{aligned} \mathscr{A}^2 b_1 &= \gamma_1 b_1 + \gamma_2 b_2 + \gamma_3 \mathscr{A} b_1 + \gamma_4 \mathscr{A} b_2, \\ \mathscr{A}^2 b_2 &= \delta_1 b_1 + \delta_2 b_2 + \delta_3 \mathscr{A} b_1 + \delta_4 \mathscr{A} b_2. \end{aligned}$$

The pair (3.1) can be reduced to

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ \gamma_1 & \gamma_3 & \delta_1 & \delta_3 \\ 0 & 0 & 0 & 1 \\ \gamma_2 & \gamma_4 & \delta_2 & \delta_4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.2}$$

where we assume that the pair in (3.2) is the representation of $\mathscr{A}, \mathscr{B}$ with respect to the basis given by

$$\left\{ \mathscr{A} b_1 - \gamma_3 b_1 - \gamma_4 b_2,\; b_1,\; \mathscr{A} b_2 - \delta_3 b_1 - \delta_4 b_2,\; b_2 \right\}.$$

If we now consider the group action (1.11) on the pair (3.2), we can assume that the trace of the matrix $\mathscr{A}$ is 0. Thus we can replace δ_4 by $-\gamma_3$. Therefore a canonical form for the perspective system (1.1) is given by

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} 0 & \gamma_1 & 0 & \gamma_2 \\ 1 & \gamma_3 & 0 & \gamma_4 \\ 0 & \delta_1 & 0 & \delta_2 \\ 0 & \delta_3 & 1 & -\gamma_3 \end{pmatrix}, \quad \mathscr{C} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.3}$$

i.e.,

$$\begin{aligned} &\dot{X}_1 = \gamma_1 Y_1 + \gamma_2 W_1, \qquad \dot{Y}_1 = X_1 + \gamma_3 Y_1 + \gamma_4 W_1, \\ &\dot{Z}_1 = \delta_1 Y_1 + \delta_2 W_1, \qquad \dot{W}_1 = \delta_3 Y_1 + Z_1 - \gamma_3 W_1, \\ &z = Y_1 / W_1. \end{aligned}$$

Therefore, for the Kronecker index pair (2, 2) we have a total of seven parameters. If we choose $n = 4$ and $m = 2$ and assume that the Kronecker indices are $\kappa_1 = 3$ and $\kappa_2 = 1$, a suitable canonical form is given by

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & \alpha_1 & \beta_1 \\ 1 & 0 & \alpha_2 & \beta_2 \\ 0 & 1 & \alpha_3 & \beta_3 \\ 0 & 0 & \alpha_4 & \beta_4 \end{pmatrix}, \quad \mathscr{C} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & \gamma & 1 \end{pmatrix},$$

where the parameters $\alpha_i, \beta_i, i = 1, 2, 3, 4$, and γ are arbitrary. Denoting, as before, the columns of $\mathscr{B} := \mathscr{C}^{\mathrm{T}}$ by $(b_1 \;\; b_2)$, the Kronecker structure implies that the vectors b_1, b_2, $\mathscr{A}b_1$ and $\mathscr{A}^2 b_1$ are independent and

$$\begin{aligned} \mathscr{A}^3 b_1 &= \gamma_1 b_1 + \gamma_2 b_2 + \gamma_3 \mathscr{A} b_1 + \gamma_4 \, \mathscr{A}^2 b_1, \\ \mathscr{A} b_2 &= \delta_1 b_1 + \delta_2 b_2 \end{aligned}$$

for some scalar constants $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \delta_1, \delta_2$ that can be determined from the pair $\mathscr{A}, \mathscr{B}$. If we consider a basis given by

$$\left\{\mathscr{A}^2 b_1 - \gamma_4 \mathscr{A} b_1 - \gamma_3 b_1,\ \mathscr{A} b_1 - \gamma_4 b_1,\ b_1,\ b_2\right\}$$

it is easy to see that the pair $\mathscr{A}^{\mathrm{T}}, \mathscr{C}$ is given by

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & \gamma_1 & \gamma_2 \\ 1 & 0 & \gamma_3 & 0 \\ 0 & 1 & \gamma_4 & 0 \\ 0 & 0 & \delta_1 & -\gamma_4 \end{pmatrix}, \quad \mathscr{C} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.4}$$

and the perspective system is given by

$$\begin{aligned} &\dot{X}_1 = \gamma_1 Z_1 + \gamma_2 W_1, \quad && \dot{Y}_1 = X_1 + \gamma_3 Z_1, \\ &\dot{Z}_1 = Y_1 + \gamma_4 Z_1, && \dot{W}_1 = \delta_1 Z_1 - \gamma_4 W_1, \\ &z = Z_1/W_1. \end{aligned}$$

There are a total of five parameters for the Kronecker index pair $(3, 1)$. Finally if we now choose $n = 4$ and $m = 3$ and assume that the Kronecker indices are $\kappa_1 = 2$, $\kappa_2 = 1$, $\kappa_3 = 1$, it is possible to show that a suitable canonical form is given by

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} 0 & \alpha_1 & \beta_1 & \gamma_1 \\ 1 & \alpha_2 & \beta_2 & \gamma_2 \\ 0 & \alpha_3 & \beta_3 & \gamma_3 \\ 0 & \alpha_4 & \beta_4 & \gamma_4 \end{pmatrix}, \quad \mathscr{C} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \eta_1 & 1 & 0 \\ 0 & \eta_2 & \eta_3 & 1 \end{pmatrix},$$

wherein, like before, the parameters $\alpha_i, \beta_i, \gamma_i, i = 1, 2, 3, 4$, and $\eta_j, j = 1, 2, 3$, are arbitrary. Denoting the columns of $\mathscr{B} := \mathscr{C}^{\mathrm{T}}$ by $(b_1 \;\; b_2 \;\; b_3)$, the Kronecker structure implies that the vectors b_1, b_2, b_3 and $\mathscr{A}b_1$ are independent and

$$\begin{aligned} \mathscr{A}^2 b_1 &= \delta_1 b_1 + \delta_2 b_2 + \delta_3 b_3 + \delta_4 \mathscr{A} b_1, \\ \mathscr{A} b_2 &= \xi_1 b_1 + \xi_2 b_2 + \xi_3 b_3, \\ \mathscr{A} b_3 &= \zeta_1 b_1 + \zeta_2 b_2 + \zeta_3 b_3 \end{aligned}$$

for some scalar constants $\delta_1, \delta_2, \delta_3, \delta_4, \xi_1, \xi_2, \xi_3, \zeta_1, \zeta_2, \zeta_3$ that can be determined from the pair $\mathscr{A}, \mathscr{B}$. If we consider a basis given by

$$\{\mathscr{A}b_1 - \delta_4 b_1,\ b_1,\ b_2,\ b_3\},$$

it follows quite easily that the pair $\mathscr{A}^{\mathrm{T}}, \mathscr{C}$ is given by

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} 0 & \delta_1 & \delta_2 & \delta_3 \\ 1 & \delta_4 & 0 & 0 \\ 0 & \xi_1 & \xi_2 & \xi_3 \\ 0 & \zeta_1 & \zeta_2 & -\delta_4 - \xi_2 \end{pmatrix}, \quad \mathscr{C} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.5}$$

and the perspective system is given by

$$\begin{aligned} \dot{X}_1 &= \delta_1 Y_1 + \delta_2 Z_1 + \delta_3 W_1, \quad \dot{Y}_1 = X_1 + \delta_4 Y_1, \\ \dot{Z}_1 &= \xi_1 Y_1 + \xi_2 Z_1 + \xi_3 W_1, \\ \dot{W}_1 &= \zeta_1 Y_1 + \zeta_2 Z_1 - (\delta_4 + \xi_2) W_1, \\ z_1 &= Y_1 / W_1, \quad z_2 = Z_1 / W_1. \end{aligned}$$

Note that for the Kronecker index (2, 1, 1) there are a total of 9 parameters. From the above remark one obtains the following result.

Theorem 3 (Canonical form under unknown camera calibration). *Let us consider the perspective dynamical system* (1.1) *for the Kronecker indices* $(2, 2)$, $(3, 1)$ *and* $(2, 1, 1)$, *respectively. A canonical form for each of the three cases is given by* (3.3), (3.4) *and* (3.5), *respectively.*

Remark 4. Note that for each of the above three cases, the group action (1.12) can be used to scale the matrix $\mathscr{C}$ to a suitable canonical structure, viz. a matrix with unit norm.

3.2. Parameter Identification using an Extended Kalman Filter

Parameter estimation problems have already been studied using an *Extended Kalman Filter* by many researchers in the past, for motion and shape estimation under perspective projection. We would like to refer to [1,9] for a general introduction about these filters. In this section we mainly discuss the case of estimating the seven parameters in (3.3). Discretizing the dynamical system (1.1) described by (3.3) we obtain the following discretized dynamical system:

$$\begin{gathered} \mathscr{X}_{k+1} = \mathscr{A}_d \mathscr{X}_k, \quad z_k = Y_{1k} / W_{1k}, \\ \mathscr{A}_d = \begin{pmatrix} 1 & \gamma_1 T & 0 & \gamma_2 T \\ T & 1 + \gamma_3 T & 0 & \gamma_4 T \\ 0 & \delta_1 T & 1 & \delta_2 T \\ 0 & \delta_3 T & T & 1 - \gamma_3 T \end{pmatrix}, \end{gathered} \tag{3.6}$$

where $\mathscr{X}_k = (X_{1k} \;\; Y_{1k} \;\; Z_{1k} \;\; W_{1k})^{\mathrm{T}}$ and where T is the sampling time. Note that the parameters to be estimated are $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \delta_1, \delta_2, \delta_3$. The three coordinates of a point observed by the camera are given by

$$X_k = X_{1k}/W_{1k}, \quad Y_k = Y_{1k}/W_{1k}, \quad Z_k = Z_{1k}/W_{1k}. \tag{3.7}$$

We now assume that, at each time instant k, a set of n points are observed with coordinates $(X_k^i \;\; Y_k^i \;\; Z_k^i)$ for $i = 1, \ldots, n$. The coordinates are defined as in (3.7) for each of the n points. We consider the state vector θ_k to be given by

$$\theta_k = (\gamma_1 \;\; \gamma_2 \;\; \gamma_3 \;\; \gamma_4 \;\; \delta_1 \;\; \delta_2 \;\; \delta_3 \;\; X_k^1 \;\; Y_k^1 \;\; Z_k^1 \;\; \cdots \;\; X_k^n \;\; Y_k^n \;\; Z_k^n)^{\mathrm{T}} \in \mathbb{R}^{3n+7},$$

and the associated state equation is given by

$$\theta_{k+1} = F(\theta_k). \tag{3.8}$$

In fact, it follows from (3.2) that we need only the observation values Y_k^i. Let Y_k^i be observed with an additive noise, and let $\mathscr{Z}_k \in \mathbb{R}^n$ be the observation vector. Then

$$\mathscr{Z}_k = H\theta_k + w_k, \tag{3.9}$$

where $\{w_k\} \in \mathbb{R}^n$ is assumed to be a Gaussian white noise with mean zero and covariance R, i.e. $\mathrm{E}\{w_k w_l^{\mathrm{T}}\} = R\delta_{kl}$, δ_{kl} indicates Kronecker's delta and $H = [O_{n,7} \; \Theta_{n,3n}] \in \mathbb{R}^{n\times(3n+7)}$. Here $\Theta_{n,3n}$ is a suitable matrix and its detailed structure is omitted. The EKF is described as follows:

$$\begin{aligned}
\hat{\theta}_{k+1|k+1} &= F(\hat{\theta}_{k|k}) + \mathscr{K}_{k+1}\big[\mathscr{Z}_{k+1} - HF(\hat{\theta}_{k|k})\big], \\
\mathscr{K}_{k+1} &= Q_{k+1|k}H^{\mathrm{T}}\big[HQ_{k+1|k}H^{\mathrm{T}} + R\big]^{-1}, \\
Q_{k+1|k} &= F_k(Q_{k|k-1} - K_k H Q_{k|k-1})F_k^{\mathrm{T}}, \\
F_k &= \left.\frac{\partial F(\theta)}{\partial \theta^{\mathrm{T}}}\right|_{\theta=\hat{\theta}_{k|k}}.
\end{aligned} \tag{3.10}$$

Likewise an EKF for the discretization of the dynamical system (1.1) from the system matrices given by (3.4) and (3.5) may be obtained analogously. Simulation studies are performed for the following parameters. For the Kronecker indices (2, 2), (3, 1) and (2, 1, 1), respectively, parameters in (3.3), (3.4) and (3.5) are taken as

$$\begin{aligned}
&\text{Case } (2,2): \quad (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \delta_1, \delta_2, \delta_3) \\
&\qquad\qquad = (-0.2, 0.1, 1, 0.5, 1, 0.3, 1.5), \\
&\text{Case } (3,1): \quad (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \delta_1) \\
&\qquad\qquad = (-0.2, 0.1, 1, 0.5, 1), \\
&\text{Case } (2,1,1): \quad (\delta_1, \delta_2, \delta_3, \delta_4, \xi_1, \xi_2, \xi_3, \zeta_1, \zeta_2) \\
&\qquad\qquad = (-0.2, 0.1, 0.3, 1, 0.4, -0.3, 0.5, -1, 0.8).
\end{aligned} \tag{3.11}$$

Moreover we set the sampling period $T = 0.01$, the number of points is chosen as $n = 3$ and

$$\begin{aligned}
&\big(X_0^1 \;\; Y_0^1 \;\; Z_0^1\big) = (0.05 \;\; 0.3 \;\; -0.1), \quad \big(X_0^2 \;\; Y_0^2 \;\; Z_0^2\big) = (0.3 \;\; 0.1 \;\; -0.06), \\
&\big(X_0^3 \;\; Y_0^3 \;\; Z_0^3\big) = (-0.2 \;\; 0.1 \;\; -0.1).
\end{aligned}$$

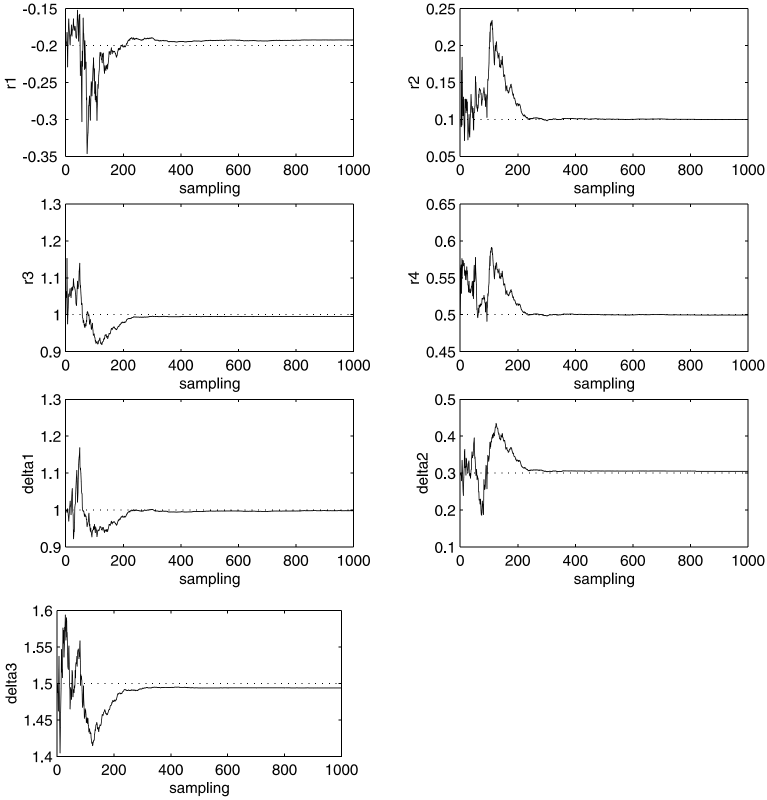

Fig. 1. Kronecker index (2, 2) case.

For each tuples of Kronecker indices (2, 2), (3, 1) and (2, 1, 1), the noise variances are chosen to be $R = \sigma^2 I_3$ and $R = \sigma^2 I_6$ with $\sigma = 0.0001$, respectively. Now using the EKF described by (3.10), parameters in (3.3), (3.4) and (3.5) have been estimated and plotted in Figs. 1, 2 and 3. The dotted lines denote the true values in (3.11) and the solid lines represent the estimated values using the EKF. It may be noted from the figures that the initial conditions for the parameters chosen during the simulation of the EKF are close to the true values. Otherwise for various other choices of the initial conditions, the parameters do not converge to the true value but either maintains a fixed bias or diverges.

4. The case of known camera calibration

In the case when the calibration parameter is assumed to be known one could suppose that the perspective dynamical system is given by

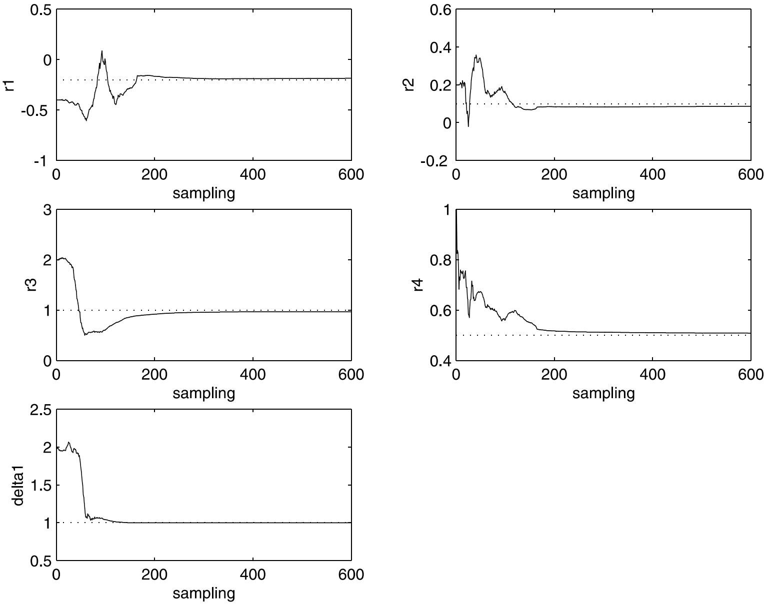

Fig. 2. Kronecker index (3, 1) case.

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix} = \mathscr{A}^{\mathrm{T}} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ W_1 \end{bmatrix} \tag{4.1}$$

with the homogeneous observation function being (1.7) or (1.8) where we define

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} A & b \\ f^{\mathrm{T}} & d \end{pmatrix},$$

and it is assumed that Trace $\mathscr{A} = 0$.

4.1. Parameterization of the orbit under perspective projection

For the perspective dynamical system (4.1), (1.7), it follows from Ghosh et al. [2] that one can consider a subgroup of the perspective group described in (1.10), (1.11) and (1.12) where P is given by

$$P^{-1} = \begin{pmatrix} p_{11} & 0 & 0 & 0 \\ 0 & p_{11} & 0 & 0 \\ 0 & 0 & p_{11} & 0 \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix}. \tag{4.2}$$

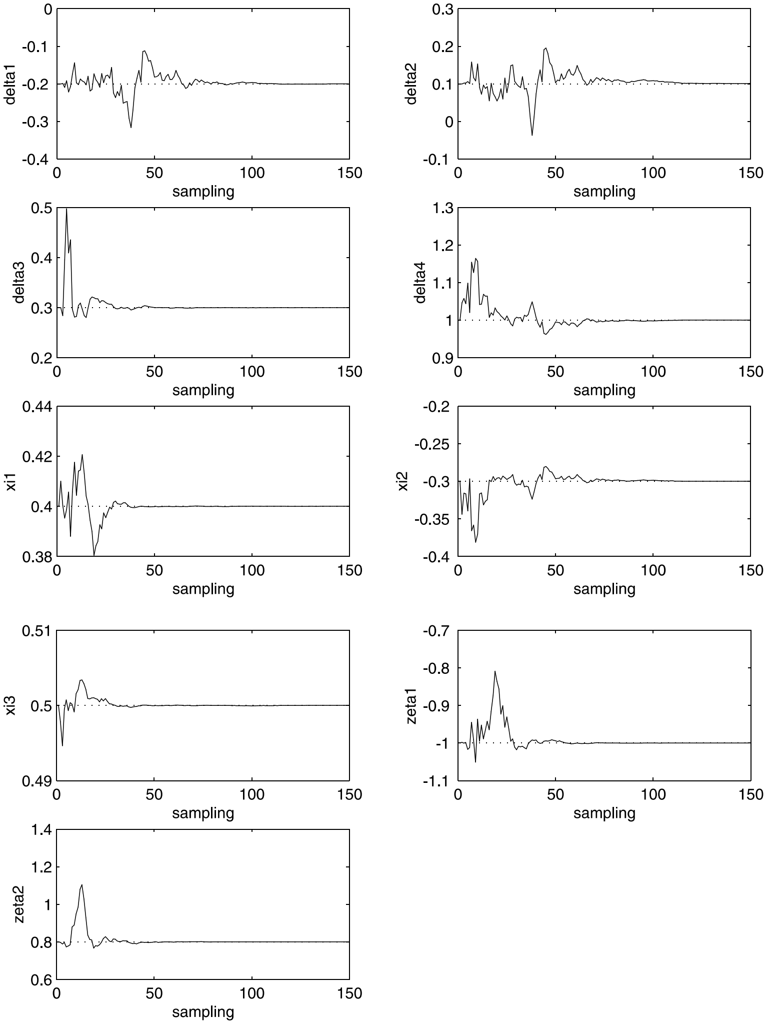

Fig. 3. Kronecker index (2, 1, 1) case.

In general the action of P changes the structure of the associated matrix $\mathscr{C}$, but P of (4.2) preserves the structure. We now have the following result.

Theorem 5 (Canonical form under known camera calibration, perspective projection case). *Consider the perspective dynamical system* (4.1) *with the observation function* (1.7) *and assume that* $b \neq 0$. *Under the perspective group action* (1.10), (1.11) *and* (1.12) *using P of* (4.2), *a canonical form in the case of known camera calibration is given by the subset* $\mathscr{S}$ *defined as follows*:

$$\mathscr{S} = \left\{ \begin{pmatrix} A & b \\ f^{\mathrm{T}} & d \end{pmatrix}; \textit{trace } A + d = 0, \right.$$
$$\left. b^{\mathrm{T}} A = 0 \textit{ and } b_i = 1 \ (i = 1 \textit{ or } 2 \textit{ or } 3) \right\}. \tag{4.3}$$

Proof. From the assumption $b \neq 0$ we assume that $b_3 \neq 0$ without any loss of generality. Then we may choose $p_{11} = b_3$ and $p_{44} = 1$. The scaling on the matrix $\mathscr{A}^{\mathrm{T}}$ described by the group action (1.10) is given by

$$\begin{aligned} A &\longmapsto A^* \triangleq A + b^*\rho, & b &\longmapsto b^* \triangleq (b_1/b_3 \quad b_2/b_3 \quad 1)^{\mathrm{T}}, \\ f^{\mathrm{T}} &\longmapsto b_3 f^{\mathrm{T}} - \rho A - (\rho b^* - d)\rho, & d &\longmapsto -(\rho b^* - d), \end{aligned}$$

where ρ denotes the vector $(p_{41} \ \ p_{42} \ \ p_{43})$. Here we can choose $p_{4i} = -b^{*\mathrm{T}} A_i / \|b^*\|^2$, $i = 1, 2, 3$, where A_i is the ith column vector of A. It follows that the vector ρ is unique and $b^{*\mathrm{T}} A^* = 0$. Furthermore in order to get trace $A^* - (\rho b^* - d)$ to be 0 we apply the group in (1.11) with $\lambda = (-a^*_{11} - a^*_{22} - a^*_{33} + (\rho b^* - d))/4$, where a^*_{ii}, $i = 1, 2, 3$, denotes a diagonal element of A^*. The cases $b_1 \neq 0$ or $b_2 \neq 0$ can be described likewise. $\square$

The canonical form (4.3) when $b_3 = 1$ is given by

$$\mathscr{A}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ -b_1 a_{11} - b_2 a_{21} & -b_1 a_{12} - b_2 a_{22} & -b_1 a_{13} - b_2 a_{23} & 1 \\ -f_1 & -f_2 & -f_3 & -a_{11} - a_{22} + b_1 a_{13} + b_2 a_{23} \end{pmatrix}, \tag{4.4}$$

where $\mathscr{C}$ is given in (1.7).

4.2. *Parameterization of the orbit under orthographic projection*

Finally we consider the perspective dynamical system (4.1), (1.8). The associated subgroup of the perspective group (described in (1.10), (1.11) and (1.12)) is given by the following special structure of P:

$$P^{-1} = \begin{pmatrix} p_{11} & 0 & 0 & 0 \\ 0 & p_{11} & 0 & 0 \\ p_{31} & p_{32} & p_{33} & p_{34} \\ 0 & 0 & 0 & p_{11} \end{pmatrix}. \tag{4.5}$$

We now describe the following.

Theorem 6 (Canonical form under known camera calibration, orthographic projection case). *Consider the perspective dynamical system* (4.1) *with the observation function* (1.8) *and assume that the vector c defined as* $c = (a_{13}\ a_{23}\ \ -f_3)^{\mathrm{T}} \neq 0$. *Under the perspective group action* (1.10), (1.11) *and* (1.12) *using P of* (4.5), *a canonical form is described as follows*:

$$\mathscr{S} = \left\{ \begin{pmatrix} A & b \\ f^{\mathrm{T}} & d \end{pmatrix}; \textit{trace } A + d = 0, \right.$$
$$\left. c^{\mathrm{T}} \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ -f_1 & -f_2 & d \end{pmatrix} = 0 \textit{ and } c_i = 1\ (i = 1 \textit{ or } 2\ \textit{ or } 3) \right\},$$

where c_i is the ith *element of c.*

Proof. The proof is analogous to the proof of Theorem 5 and is omitted. □

4.3. Parameter identification using an EKF

In this section, we consider parameter identification under perspective projection for the dynamical system (4.1), (1.7). We assume (4.4) as the structure of the matrix $\mathscr{A}^{\mathrm{T}}$ and (1.7) for the matrix $\mathscr{C}$. The dynamics is discretized with the sampling time T as

$$\mathscr{X}_{k+1} = \mathscr{A}_d \mathscr{X}_k, \quad z_{1k} = X_{1k}/Z_{1k}, \quad z_{2k} = Y_{1k}/Z_{1k}, \tag{4.6}$$

where $\mathscr{X}_k = (X_{1k}\ \ Y_{1k}\ \ Z_{1k}\ \ W_{1k})^{\mathrm{T}}$. $\mathscr{A}_d$ is a suitable discretized matrix of (4.4) and its detailed structure is omitted. We define the three observation variables given by

$$\bar{x}_k = X_{1k}/Z_{1k}, \quad \bar{y}_k = Y_{1k}/Z_{1k}, \quad \bar{z}_k = W_{1k}/Z_{1k}.$$

The parameters to be estimated are $a_{11}, \ldots, a_{23}, b_1, b_2, -f_1, -f_2, -f_3$. We consider the state vector θ_k to be given by

$$\theta_k = (a_{11}\ \cdots\ a_{23}\ b_1\ b_2\ -f_1\ -f_2\ \ -f_3\ \bar{x}_k^1\ \bar{y}_k^1\ \bar{z}_k^1\ \cdots\ \bar{x}_k^n\ \bar{y}_k^n\ \bar{z}_k^n)^{\mathrm{T}} \in \mathbb{R}^{3n+11},$$

where a set of n points are observed with coordinates $(\bar{x}_k^i\ \ \bar{y}_k^i\ \ \bar{z}_k^i)$ for $i = 1, \ldots, n$. It follows from (4.6) that the observation values are $\bar{x}_{ik}$ and $\bar{y}_{ik}$. Let $\bar{x}_{ik}$ and $\bar{y}_{ik}$ be observed with an additive noise, and let $\mathscr{Z}_k \in \mathbb{R}^{2n}$ be the observation vector. Then we define the dynamical system as (3.8), (3.9) where $\{w_k\} \in \mathbb{R}^{2n}$ and $H = [O_{2n,11}\ \Theta_{2n,3n}] \in \mathbb{R}^{2n\times(3n+11)}$. $\Theta_{2n,3n}$ is described by 0, 1 and its structure is also omitted. An EKF for the dynamical system (4.1) described by the parameter matrix (4.4) is obtained analogous to that obtained in (3.10).

Simulation is performed for the following parameters. Let us assume that the parameters in (4.4) be given by

$$\begin{aligned}&(a_{11},\ a_{12},\ a_{13},\ a_{21},\ a_{22},\ a_{23},\ b_1,\ b_2,\ -f_1,\ -f_2,\ -f_3)\\ &\quad = (-0.2,\ 0.1,\ 0.3,\ 0.5,\ 0.4,\ -0.3,\ -0.1,\ -0.5,\ -0.2,\ -0.1,\ -0.4).\end{aligned}$$

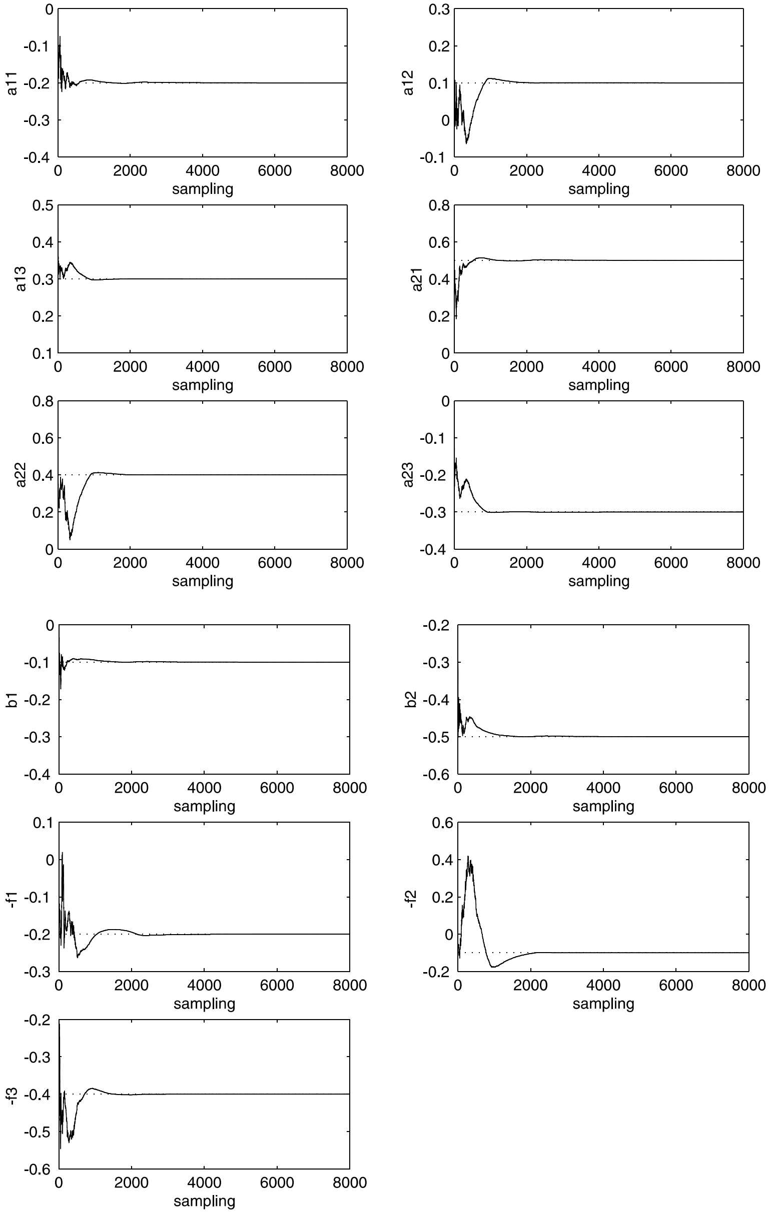

Fig. 4. Estimation of parameters in (4.4).

We set the sampling period $T = 0.001$, and the number of points is taken as $n = 3$. The initial positions of the three points are chosen as

$$\begin{pmatrix}\bar{x}_0^1 & \bar{y}_0^1 & \bar{z}_0^1\end{pmatrix} = (0.2 \;\; 0.1 \;\; -0.2), \quad \begin{pmatrix}\bar{x}_0^2 & \bar{y}_0^2 & \bar{z}_0^2\end{pmatrix} = (0.3 \;\; 0.2 \;\; 0.1),$$
$$\begin{pmatrix}\bar{x}_0^3 & \bar{y}_0^3 & \bar{z}_0^3\end{pmatrix} = (-0.1 \;\; 0.3 \;\; -0.1).$$

Let the noise variance be chosen to be $R = \sigma^2 I_6$ with $\sigma = 0.0001$, respectively. Now using the EKF denoted by (3.10), parameters in (4.4) have been estimated and plotted in Fig. 4. Let the dotted lines denote the true values and let the solid lines indicate the estimated values. The simulation of the EKF is shown to be close to the true value.

5. Conclusion

In this paper, we have introduced canonical forms for perspective dynamical systems obtained under the action of a perspective group. Two cases are analyzed, one in which the observation matrix is unknown and the other case in which this matrix is assumed to be fixed and known. In the former case we show that the problem is related to the classical Kronecker indices. We also observe in this case that because the calibration parameters are not known a priori the associated perspective group describes orbits of “large” dimension. Hence the number of identifiable parameters are quite small (7, 5 and 9 in our examples) compared to the total number of parameters. In the latter case, when the calibration parameters are known, 11 of the possible 16 parameters under perspective projection are identifiable and has been identified in our example. We parameterize orbits of the perspective dynamical systems, and show via simulation that the parameters can be identified using an Extended Kalman Filter. In this way, we provide a geometric framework to study the problem of parameter identification of a linear dynamical system with perspective and orthographic observations.

References

[1] C.K. Chui, G. Chen, Kalman Filtering with Real-Time Applications, Springer, Berlin, 1998.
[2] B.K. Ghosh, H. Inaba, S. Takahashi, Identification of Riccati dynamics under perspective and orthographic observations, IEEE Trans. Automat. Control 45 (7) (2000) 1267–1278.
[3] B.K. Ghosh, M. Jankovic, Y.T. Wu, Perspective problems in system theory and its application to machine vision, J. Math. Systems, Estimation and Control 4 (1) (1994) 3–38.
[4] B.K. Ghosh, E.P. Loucks, A perspective theory for motion and shape estimation in machine vision, SIAM J. Control Optim. 33 (5) (1995) 1530–1559.
[5] B.K. Ghosh, E.P. Loucks, A realization theory for perspective systems with applications to parameter estimation problems in machine vision, IEEE Trans. Automat. Control 41 (12) (1996) 1–17.
[6] B.K. Ghosh, C.F. Martin, Homogeneous dynamical systems theory, IEEE Trans. Automat. Control (in press).
[7] M. Hazewinkel, Moduli and canonical forms for linear dynamical systems III: the algebraic–geometric case, in: C. Martin, R. Hermann (Eds.), Geometric Control Theory, Math. Sci. Press, Brookline, MA, 1977.

[8] M. Hazewinkel, R.E. Kalman, On invariants canonical forms and moduli for linear constant finite-dimensional dynamical systems, Lecture Notes in Econom. – Math. System Theory, 131 (1976) 48–60.
[9] T. Kailath, Lecture on Wiener and Kalman Filtering, Springer, New York, 1981.
[10] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980.
[11] R.E. Kalman, Kronecker invariants and feedback, in: L. Weiss (Ed.), Ordinary Differential Equations, Academic Press, New York, 1972, pp. 459–471.
[12] R.E. Kalman, On partial realizations, transfer functions and canonical forms, Acta Polytech. Scand. 31 (1979) 9–32.
[13] R.E. Kalman, M. Arbib, P. Falb, Topics in Mathematical System Theory, McGraw-Hill, New York, 1965.
[14] K. Kanatani, Group-Theoretical Methods in Image Understanding, Springer, Berlin, 1990.
[15] C.F. Martin, R. Hermann, Application of algebraic geometry to system theory: the McMillan degree and Kronecker indices of transfer functions as topological and holomorphic invariants, SIAM J. Control 16 (1978) 743–755.
[16] V.M. Popov, Invariant description of linear time-invariant controllable systems, SIAM J. Control 10 (1972) 252–264.