Recursive estimation of tri-focal tensors

B. Tech. Project Report

Submitted in partial fulfillment of the requirements
for the degree of

Bachelor of Technology
in
Electrical Engineering

by

S. Dheeraj Prasad
Roll No: 00007026

under the guidance of

Prof. S. Chaudhuri

Department of Electrical Engineering
Indian Institute of Technology, Bombay
Mumbai

Acknowledgments

I would like to thank my guide, Prof. S. Chaudhuri for his constant guidance and support without which this B. Tech project would not have been possible. A constant motivating factor, he helped me at every faltering step and provided me with fresh and innovative solutions for the same.

S. Dheeraj Prasad

Contents

Chapter 1

Introduction

Most models employed for recursive scene computation (such as those in [7], [8], [9], [10] and [11]) use correspondences over two consecutive frames and model them in terms of the motion parameters between the two frames. If one considers correspondences over three consecutive frames, one can come up with a much more robust model for scene computation. The better quality for estimation can be attributed to the fact that we have a larger data set to estimate a particular parameter. One method of doing so is by using tri focal tensors for estimation since the tri focal tensors help establish a relationship between correspondences over three scenes. The object of this project is to develop a model for recursive estimation of tri-focal tensors over frames that would give rise to a very robust model for scene computation. A linear predictive model is aimed at in order to employ the Kalman filter.

The report is organised as follows: Chapter 2 deals with a brief introduction to tri focal tensors and the mathematics that would be required in developing the model for recursive estimation of tri focal tensors. Chapter 3 deals with the formulation of the model. Chapter 4 deals with the results obtained on simulating the model. Chapter 5 deals with concluding remarks about the project and chapter 6 deals with improvements that can be explored in the future.

Chapter 2

Tri focal tensors: An introduction

As discussed in [1] the tri-focal tensor plays an analogous role in three views to the fundamental matrix in two views. It helps establish incidence relations between three different views of a particular scene. In our project we are dealing with point correspondences in each frame as depicted in the figure below:

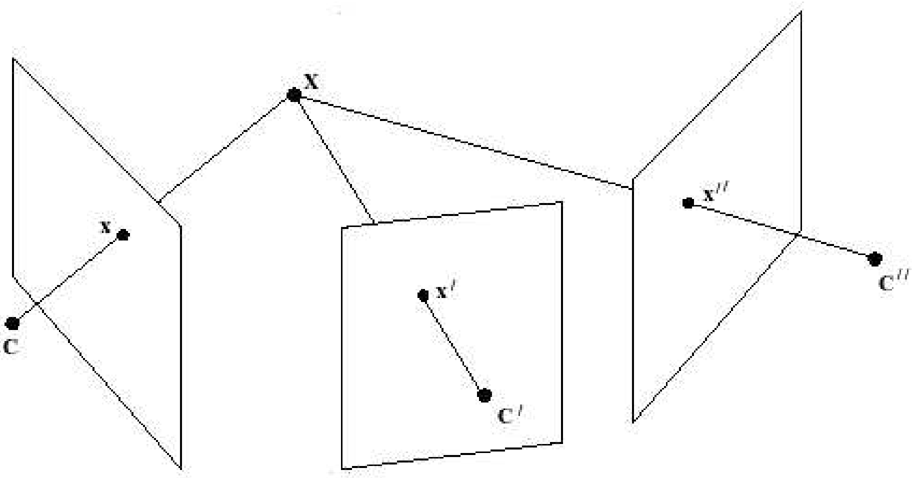

Figure 2.1: point-point-point correspondence

If a 3D point $\mathbf{X}$ projects to points $\mathbf{x}$ = (x,y,1), $\mathbf{x}'$ = (x′,y′,1) and $\mathbf{x}''$ = (x″,y″,1) respectively in the three different frames as shown in figure 2.1 then the tri focal tensor $\mathbf{T}$ can be used to express the incidence relationship for these points as

$$x^i(x'^j\epsilon_{jpr})(x''^k\epsilon_{kqs})\mathbf{T}_i^{pq} = \mathbf{0}_{rs} \qquad i,j,k,p,q,r,s = 1,2,3 \tag{2.1}$$

where

$$\epsilon_{ijk} = \begin{cases} 0 & \text{unless i,j and k are distinct} \\ +1 & \text{if ijk is an even permutation of 123} \\ -1 & \text{if ijk is an odd permutation of 123} \end{cases} \tag{2.2}$$

From equation (2.1) one notices that we can get 9 different equations or *trilinearities*, however only 4 out of these trilinearities are linearly independent. Hence for every given point correspondence over 3 frames it is possible to get a set of 4 independent trilinearites. These trilinearities can be written as follows:

$$\text{x[x'x"}\mathbf{T}_1^{33} - x''\mathbf{T}_1^{13} - x'\mathbf{T}_1^{31} + \mathbf{T}_1^{11} + y[x'x''\mathbf{T}_2^{33} - x''\mathbf{T}_2^{13} - x'\mathbf{T}_2^{31} + \mathbf{T}_2^{11}$$
$$+\text{[x'x"}\mathbf{T}_3^{33} - x''\mathbf{T}_3^{13} - x'\mathbf{T}_3^{31} + \mathbf{T}_3^{11} = 0 \quad (2.3)$$

$$\text{x[x'y"}\mathbf{T}_1^{33} - y''\mathbf{T}_1^{13} - x'\mathbf{T}_1^{32} + \mathbf{T}_1^{12} + y[x'y''\mathbf{T}_2^{33} - y''\mathbf{T}_2^{13} - x'\mathbf{T}_2^{32} + \mathbf{T}_2^{12}$$
$$+\text{[x'y"}\mathbf{T}_3^{33} - y''\mathbf{T}_3^{13} - x'\mathbf{T}_3^{32} + \mathbf{T}_3^{12} = 0 \quad (2.4)$$

$$\text{x[y'x"}\mathbf{T}_1^{33} - x''\mathbf{T}_1^{23} - y'\mathbf{T}_1^{31} + \mathbf{T}_1^{21} + y[y'x''\mathbf{T}_2^{33} - x''\mathbf{T}_2^{23} - y'\mathbf{T}_2^{31} + \mathbf{T}_2^{21}$$
$$+\text{[y'x"}\mathbf{T}_3^{33} - x''\mathbf{T}_3^{23} - y'\mathbf{T}_3^{31} + \mathbf{T}_3^{21} = 0 \quad (2.5)$$

$$\text{x[y'y"}\mathbf{T}_1^{33} - y''\mathbf{T}_1^{23} - y'\mathbf{T}_1^{32} + \mathbf{T}_1^{22} + y[y'y''\mathbf{T}_2^{33} - y''\mathbf{T}_2^{23} - y'\mathbf{T}_2^{32} + \mathbf{T}_2^{22}$$
$$+\text{[y'x"}\mathbf{T}_3^{33} - y''\mathbf{T}_3^{23} - y'\mathbf{T}_3^{32} + \mathbf{T}_3^{22} = 0 \quad (2.6)$$

Now it has been shown in [1],[3] and[4] that the tri focal tensor can be calculated upto a scale factor, hence we assume the term $\mathbf{T}_3^{33}$ to be equal to 1. Since the tri focal has 27 elements this scaling reduces the number of unknown elements in the tri focal tensor to 26. The fact that every point correspondence over 3 scenes gives rise to 4 independent equations implies that we would require at least 7 point correspondences to calculate the tri focal tensor to the chose scale. Thus if we are given 7 point correspondences with no four of them being coplanar we can evaluate the tri focal tensor to the required scale factor.
We thus use the following method for evaluating the tri focal tensor from 7 given point correspondences over 3 scenes. In the equations (2.3), (2.4), (2.5) and (2.6) we substitute the value for $\mathbf{T}_3^{33}$ as 1 and can form 28 equations for the given 7 points as

$$C_1 T = C_2 \tag{2.7}$$

where the above is formed by rearranging equations (2.3), (2.4), (2.5) and (2.6). C_2 is a 28×1 vector that contains the negative of the coefficients of $\mathbf{T}_3^{33}$. T is a 28×1 vector that contains all the elements of the tri focal tensor $\mathbf{T}$ except for $\mathbf{T}_3^{33}$. C_1 is the coefficient matrix obtained from equations (2.3),(2.4),(2.5) and (2.6).

Chapter 3

Model for recursive estimation of tri focal tensors

As stated earlier we intend to use the Kalman filter for recursive estimation of tri focal tensors. We shall first look at the mathematical model of the Kalman filter discussed in [2] that is to be used for the estimation process. Later we shall discuss the method for evaluating the parameters of this mathematical model for estimating tri focal tensors.

3.1 Kalman filtering

The operation of the Kalman filter can be considered as a predictor. Let us consider that the current time of measurement is t. The noisy measurement being made at time t is denoted by y(t) and the corresponding actual state is denoted by x(t). At time t we receive the previous filtered estimate $\hat{x}(t-1|t-1)$ and the covariance $\tilde{P}$. We now need to get the best possible estimate of the state x(t) using the previous t-1 data samples.

We can refer to the first phase of the algorithm as the *prediction phase* of the algorithm. Once the prediction is made we obtain the estimate $\tilde{x}(t|t-1)$ and the associated error covariance $\tilde{P}(t|t-1)$. After we have made these predictions we calculate the error covariance $R_e(t)$ and the Kalman gain K(t). When we obtain the measurement y(t) at time t we determine the innovation e(t). After this we enter the phase that can be termed as the *correction phase* of the algorithm. In the correction stage, we correct the state based on the new measurement made and this is called the innovation. The old predicted state $\tilde{x}(t|t-1)$ is used to form the filtered or corrected state estimate $\tilde{x}(t|t)$ and $\tilde{P}(t|t)$. After this has been done the innovation is weighed by the Kalman gain K(t) to correct the old state estimate predicted by $\tilde{x}(t|t-1)$. The idea behind this is that we cannot entirely depend on the measurement or the prediction and hence we need to choose a weighted mean between the two. The whole process involves finding the correct weight in order to optimize the estimation process. We also correct the associated error covariance. This process involves taking into account the process noise covariance R_w and measurement noise covariance R_v Once this is done the algorithm repeats for time $t+1$. The algorithm can be summarized as follows:

Prediction:

$$\tilde{x}(t|t-1) = A(t-1)\tilde{x}(t-1|t-1) + B(t-1)u(t-1) = W(t-1)w(t-1) \qquad (3.1)$$

$$\tilde{P}(t|t-1) = A(t-1)\tilde{P}(t-1|t-1)A'(t-1) + W(t-1)R_w(t-1)W'(t-1) \tag{3.2}$$

Measurement Model

$$y(t) = C(t)x(t) + v(t) \tag{3.3}$$

Innovation

$$e(t) = y(t) - \tilde{y}(t|t-1) = y(t) - C(t)\tilde{x}(t|t-1) \tag{3.4}$$

$$R_e(t) = C(t)\tilde{P}(t|t-1)C'(t) + R_v(t) \tag{3.5}$$

Gain

$$K(t) = \tilde{P}(t|t-1)C'(t)R_e^{-1}(t) \tag{3.6}$$

Correction

$$\tilde{x}(t|t) = \tilde{x}(t|t-1) + K(t)e(t) \tag{3.7}$$

$$\tilde{P}(t|t) = [I - K(t)C(t)]\tilde{P}(t|t-1) \tag{3.8}$$

3.2 Calculation of parameters for the Kalman filter model

Let us consider 4 consecutive frames corresponding to the motion of a particular scene, say F_1,F_2,F_3 and F_4. Let the tri focal tensor $\mathbf{T}_1$ and $\mathbf{T}_2$ correspond to the frames (F_1, F_2, F_3) and (F_2, F_3, F_4) respectively. It can be seen that both the tri focal tensors $\mathbf{T}_1$ and $\mathbf{T}_2$ have 2 common frames i.e F_2 and F_3. Hence we would like to exploit this common information and use it to predict $\mathbf{T}_2$ from $\mathbf{T}_1$. Now we should note that the process of using the Kalman filter is equivalent to a minimum square error method[2]. This is most effective for linear predictive systems and hence we aim at deriving a linear relationship between $\mathbf{T}_2$ from $\mathbf{T}_1$.

Let us first define certain parameters of the Kalman filter model in terms of the different image frames. At a given time t we have x(t-1) and x(t) are 27 element columns vector representing the tri focal tensor $\mathbf{T}_1$ and $\mathbf{T}_2$ and y(t) is a 14 element column vector representing the points in the frame F_4 such that

$$x(t-1)(l,1) = \mathbf{T}_i^{jk} \qquad l = 9 \times (i-1) + 3 \times (j-1) + k \tag{3.9}$$

$$x(t)(m,1) = \mathbf{T}_i^{jk} \qquad m = 9 \times (i-1) + 3 \times (j-1) + k \tag{3.10}$$

$$y(t) = [\ x_1''' \ y_1''' \ x_2''' \ y_2''' \ x_3''' \ y_3''' \ x_4''' \ y_4''' \ x_5''' \ y_5''' \ x_6''' \ y_6''' \ x_7''' \ y_7'''\ ]^T \tag{3.11}$$

We have chosen $R_v = I_{27\times27}$, $R_w = I_{14\times14}$ and $B(t) = \mathbf{0}_{27\times27}$. We shall now evaluate the value of C(t) and A(t). We shall first do so for C(t) and then for A(t), the reason being A(t) can be calculated under two particular cases, i.e. uniform motion between frames and non uniform motion between frames.

To evaluate the value of C(t) we need to derive equations for expressing the points x_i''' and y_i''' (i=1,2, … 7) in terms of the tri focal tensor $\mathbf{T}_2$. We can rewrite equations (2.3), (2.4), (2.5) and (2.6) as

$$[x''x''T_1^{33} - x'T_1^{13} + y'x''T_2^{33} - y'T_2^{13} + x''T_3^{33} - T_3^{13}]x''' = x'x''T_1^{31} - x'T_1^{11} + x''y'T_2^{31} - y'T_2^{11} + x''T_3^{31} - T_3^{11} \tag{3.12}$$

$$[x''x''T_1^{33} - x'T_1^{13} + y'x''T_2^{33} - y'T_2^{13} + x''T_3^{33} - T_3^{13}]y''' = x'x''T_1^{32} - x'T_1^{12} + x''y'T_2^{32} - y'T_2^{12} + x''T_3^{32} - T_3^{12} \tag{3.13}$$

$$[x''y''T_1^{33} - x'T_1^{23} + y'y''T_2^{33} - y'T_2^{23} + y''T_3^{33} - T_3^{23}]x''' = x'y''T_1^{31} - x'T_1^{21} + y''y'T_2^{31} - y'T_2^{21} + y''T_3^{31} - T_3^{21} \tag{3.14}$$

$$[x''y''T_1^{33} - x'T_1^{23} + y'y''T_2^{33} - y'T_2^{23} + y''T_3^{33} - T_3^{23}]y''' = x'y''T_1^{32} - x'T_1^{22} + y''y'T_2^{32} - y'T_2^{22} + y''T_3^{32} - T_3^{22} \tag{3.15}$$

Now let the co-efficients of x_i''' in equation (3.12) and (3.14) be a_i^1 and a_i^2 respectively. Similarly the co-efficients of y_i''' in equation (3.13) and (3.15) are a_i^1 and a_i^2 respectively. We use the equation pairs (3.12) and (3.13) or (3.14) and (3.15) to write the relation between y(t) and x(t) in the form

$$C_1 \times y(t) = C_2 \times x(t) \tag{3.16}$$

where C_1 is a 14×14 matrix and C_2 is a 14×27 matrix. Now we need C_1 to be invertible in order to reduce equation (3.16) to the form of (3.3). But from simulation it was noticed that if C_1 is formed using only the equation pair (3.12) and (3.13) or the equation pair (3.14) and 3.15), C_1 is not full rank(when noise was not added to the points). In order to eliminate this problem a value λ is chosen between 0 and 1 and $\lambda \neq 0.5$ and we form equations E_1 and E_2 such that

$$E_1 : \lambda \times (3.12) + (1 - \lambda) \times (3.14) \tag{3.17}$$

$$E_2 : (1 - \lambda) \times (3.13) + \lambda \times (3.15) \tag{3.18}$$

We get 2 such equations for every point (x_i''', y_i'''), (i=1,2,. . .,7) and we can use these equations to evaluate C_1. C_1 is no more rank deficient. The rank deficiency is removed in this case because for a given pt $(x_i''', y_i'''$ the coefficients in E_1 and E_2 are now $\lambda \times a_i^1 + (1 - \lambda) \times a_i^2$ and $\lambda \times a_i^2 + (1 - \lambda) \times a_i^1$ respectively. In the initial case the coefficients were equal and this resulted in the rank deficiency. In order to ensure full rank we make sure that $\lambda \neq 0.5$. Now that C_1 is full rank we can calculate C(t) by the following relation

$$C(t) = C_1^{-1} C_2 \tag{3.19}$$

We now come to the evaluation of A(t). Since the tri focal tensor is dependent on the relative motion between frames we can consider 2 separate cases. The first and trivial case is when the motion between all the frames is the same. In such a case the tri focal tensor for any given set of 3 consecutive frames will be the same and in this case A(t) works out to be $I_{27 \times 27}$. We shall now consider the non trivial case where the motion is non uniform between the frames.

In order to predict $\mathbf{T}_2$ from $\mathbf{T}_1$ we need to identify what are the common parameters in calculating these 2 tensors. The answer to this is the motion parameters between the frames

F_2 and F_3. We thus try to express these 2 tensors in terms of the motion parameters between F_2 and F_3 and consequentially relate them. Before we do so let us define certain parameters or variables that we shall use in the deduction of the above mentioned expressions.

Let the rotation matrices (of order 3×3) between the frames (F_1, F_2), (F_2, F_3) and (F_3, F_4) be α, β and γ respectively. Similarly the translation matrices between the above frame pairs are $t1$, $t2$ and $t3$ respectively. We can thus express $\mathbf{T}_1$ and $\mathbf{T}_1$ in terms of β and $t2$ as follows. (let $\mathbf{T}_1$ and $\mathbf{T}_1$ be represented by L and M respectively)

$$L_i^{11} = (\alpha_{1i}t1_2 - \alpha_{2i}t1_1)\beta_{12} + (\alpha_{1i}t1_3 - \alpha_{3i}t1_1)\beta 13 + \alpha_{1i}t2_1 \tag{3.20}$$
$$L_i^{12} = (\alpha_{1i}t1_2 - \alpha_{2i}t1_1)\beta_{22} + (\alpha_{1i}t1_3 - \alpha_{3i}t1_1)\beta 23 + \alpha_{1i}t2_2 \tag{3.21}$$
$$L_i^{13} = (\alpha_{1i}t1_2 - \alpha_{2i}t1_1)\beta_{32} + (\alpha_{1i}t1_3 - \alpha_{3i}t1_1)\beta 33 + \alpha_{1i}t2_3 \tag{3.22}$$
$$L_i^{21} = (\alpha_{2i}t1_1 - \alpha_{1i}t1_2)\beta_{11} + (\alpha_{2i}t1_3 - \alpha_{3i}t1_2)\beta 13 + \alpha_{2i}t2_1 \tag{3.23}$$
$$L_i^{22} = (\alpha_{2i}t1_1 - \alpha_{1i}t1_2)\beta_{21} + (\alpha_{2i}t1_3 - \alpha_{3i}t1_2)\beta 23 + \alpha_{2i}t2_2 \tag{3.24}$$
$$L_i^{23} = (\alpha_{2i}t1_1 - \alpha_{1i}t1_2)\beta_{31} + (\alpha_{2i}t1_3 - \alpha_{3i}t1_2)\beta 33 + \alpha_{2i}t2_3 \tag{3.25}$$
$$L_i^{31} = (\alpha_{3i}t1_1 - \alpha_{1i}t1_3)\beta_{11} + (\alpha_{3i}t1_2 - \alpha_{2i}t1_3)\beta 12 + \alpha_{3i}t2_1 \tag{3.26}$$
$$L_i^{32} = (\alpha_{3i}t1_1 - \alpha_{1i}t1_3)\beta_{21} + (\alpha_{3i}t1_2 - \alpha_{2i}t1_3)\beta 22 + \alpha_{3i}t2_2 \tag{3.27}$$
$$L_i^{3i} = (\alpha_{3i}t1_1 - \alpha_{1i}t1_3)\beta_{31} + (\alpha_{3i}t1_2 - \alpha_{2i}t1_3)\beta 32 + \alpha_{3i}t2_3 \tag{3.28}$$

$$i = 1, 2, 3$$

$$M_i^{11} = +(\gamma_{12}t2_2 + \gamma_{13}t2_3 + t3_1)\beta_{1i} - \gamma_{21}t2_1\beta_{2i} - \gamma_{13}t2_1\beta_{3i} \tag{3.29}$$
$$M_i^{12} = +(\gamma_{22}t2_2 + \gamma_{23}t2_3 + t3_2)\beta_{1i} - \gamma_{22}t2_2\beta_{2i} - \gamma_{23}t2_1\beta_{3i} \tag{3.30}$$
$$M_i^{13} = +(\gamma_{32}t2_2 + \gamma_{33}t2_3 + t3_3)\beta_{1i} - \gamma_{31}t2_3\beta_{2i} - \gamma_{33}t2_1\beta_{3i} \tag{3.31}$$
$$M_i^{21} = -\gamma_{11}t2_2\beta_{1i} + (\gamma_{11}t2_1 + \gamma_{13}t2_3 + t3_1)\beta_{2i} - \gamma_{13}t2_2\beta_{3i} \tag{3.32}$$
$$M_i^{22} = -\gamma_{21}t2_2\beta_{1i} + (\gamma_{21}t2_1 + \gamma_{23}t2_3 + t3_2)\beta_{2i} - \gamma_{23}t2_2\beta_{3i} \tag{3.33}$$
$$M_i^{23} = -\gamma_{31}t2_2\beta_{1i} + (\gamma_{31}t2_1 + \gamma_{33}t2_3 + t3_3)\beta_{2i} - \gamma_{33}t2_2\beta_{3i} \tag{3.34}$$
$$M_i^{31} = -\gamma_{1i}t2_3\beta_{1i} - \gamma_{12}t2_3\beta_{2i} + (\gamma_{11}t2_1 + \gamma_{12}t2_2 + t3_1)\beta_{3i} \tag{3.35}$$
$$M_i^{32} = -\gamma_{2i}t2_3\beta_{1i} - \gamma_{22}t2_3\beta_{2i} + (\gamma_{21}t2_1 + \gamma_{22}t2_2 + t3_2)\beta_{3i} \tag{3.36}$$
$$M_i^{33} = -\gamma_{3i}t2_3\beta_{1i} - \gamma_{32}t2_3\beta_{2i} + (\gamma_{31}t2_1 + \gamma_{32}t2_2 + t3_3)\beta_{3i} \tag{3.37}$$

$$i = 1, 2, 3$$

Now using the above equations we can formulate an equation of the type of equation (3.38) by eliminating β and $t2$. Now A_1 is invertible and hence we can get $\mathbf{T}_2$ in terms of $\mathbf{T}_1$ as shown in the equation. Now it can be seen that there are many ways of expressing $\mathbf{T}_2$ in terms of the parameters of β and $t2$. We choose A_2 in such a way that the coefficients of $t2_1$, $t2_2$ and $t2_3$ are 0. The justification for this step has been given later.

$$\mathbf{T}_1 = A_1 \times X_{motion} \quad \mathbf{T}_2 = A_2 \times X_{motion} \quad \Longrightarrow A(t) = A_2 \times A_1^{-1}$$

$$where \qquad X_{motion} = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \beta_{31} \\ \beta_{32} \\ \beta_{33} \\ t2_1 \\ t2_2 \\ t2_3 \end{bmatrix} \tag{3.38}$$

With this we have laid down methods for evaluation the parameters used in the mathematical model defined for the Kalman filter. W(T) is taken as $I_{27\times27}$. It should be noted that the motion parameters used in estimating the tensor have been assumed to have been given. In practice it might be possible to retrieve the motion parameters but the translation can only be calculated to a scale[6]. This is a reason for setting the co-efficients of $t2_1$, $t2_2$ and $t2_3$ to 0 because they can be calculated only to a particular scale. The rotation between frames however can be calculated accurately. Thus the current model for estimating the tri focal tensors assumes that the motion parameters are known to the user and methods for accurately calculating these parameters from the image frames may be explored later.

Chapter 4

Results

4.1 Details about the program code

A code was written in MATLAB implementing the model developed in the previous chapter, for the recursive estimation of tri focal tensors. The code assumes that the motion can be non uniform between certain frames, however the degree of non uniform motion is restricted to a certain level, i.e linearly varying rotational and translational acceleration . The code asks the user to enter

- The axis of rotation around which the scene rotates
- The initial rotational velocity of the scene: ϕ_0
- The change in ϕ_0 per frame: ϕ_v
- The change in ϕ_v per frame: ϕ_a
- The initial translation velocity of the scene : T_0
- The change in T_0 per frame: T_v
- The change in T_v per frame: T_a

Once the above parameters are generated the program generates 7 random points in 3D space and generates further frames(the number is also inputted from the user) from this set of points by using the motion parameters defined by the user. The tri focal tensors for consecutive frames are calculated and then the mathematical model developed in the previous chapter is used for estimating tri focal tensors. An initial error covariance is assumed for the tri focal tensor corresponding to the first 3 image frames. The error covariance is then expressed in terms of the Frobenius norm of the error covariance matrix. It is expected that the error covariance should reduce with time (or consecutive frames in this case). Hence we expect our graphs of "Frobenius norm of error covariance matrix" *vs* "frame" to be a graph that falls down, depicting reduction in error.

4.2 Details about the results of the program

The program was first tested for the trivial case of uniform (i.e. same) motion between all frames. In this case the tri focal tensor remained constant over the frames. This was tested in the absence of noise in the scene points. The reason for doing so is because in the presence of noise one would not expect the tensor to remain constant over the consecutive frames. Here A(t) (refer: eqn(3.1)) was taken to be $I_{27\times 27}$. The fact that the tri focal tensor was correctly evaluated was checked by calculating the motion parameters between 2 consecutive frames, assuming that the motion between the other 2 consecutive frames corresponding to the tensor, are known. The code was then tested for non uniform motion between frames. In this case also the motion parameters were correctly evaluated from the tensors thus indicating the correct evaluation of the tensors. The results of the simulations are at the end of the chapter.

In the uniform case the error covariance falls and that is as expected. However it is noticed that for the non uniform case in which T_0 parameters are less than or equal to 1 the solution appears to diverge, while the error reduces in the other cases. This is not expected as the error is supposed to reduce with the *passage* of frames. One probable reason for this error is the calculation of A(t). It is noticed that even when the translation and rotation is constant over the frames A(t) does not come out to be identity matrix, though the tensors $\mathbf{T}_1$ and $\mathbf{T}_2$ are equal with a difference in the 15^{th} decimal only. This problem has not been resolved so far.

4.3 Sample outputs

4.3.1 $A(t) = I_{27\times 27}$

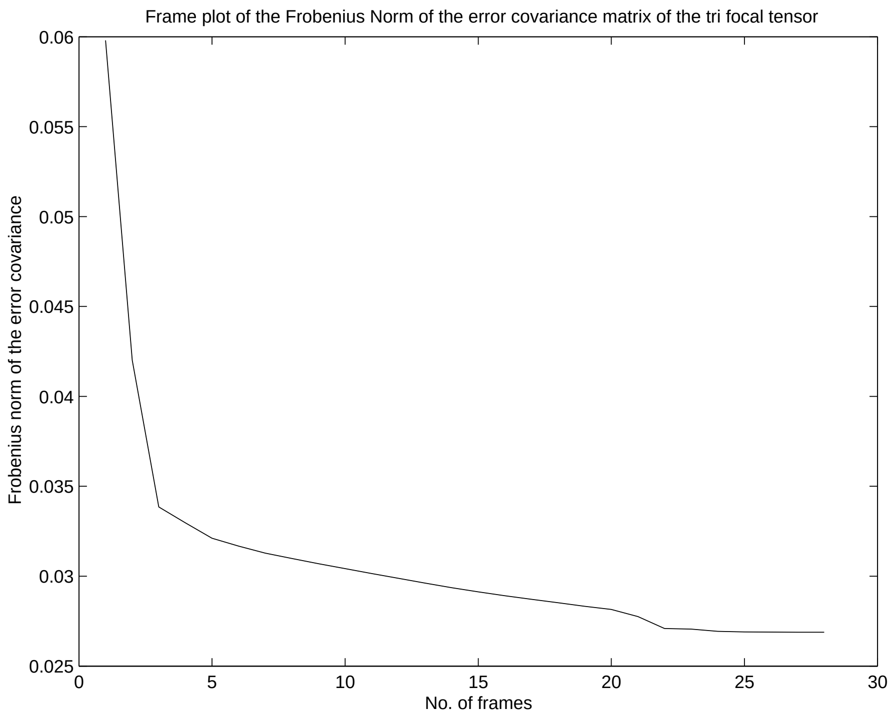

Figure 4.1:
$(n1, n2, n3) = (0, 1, 0) \quad (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0), \quad (T_0, T_v, T_a) = ([2, 4, 5]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

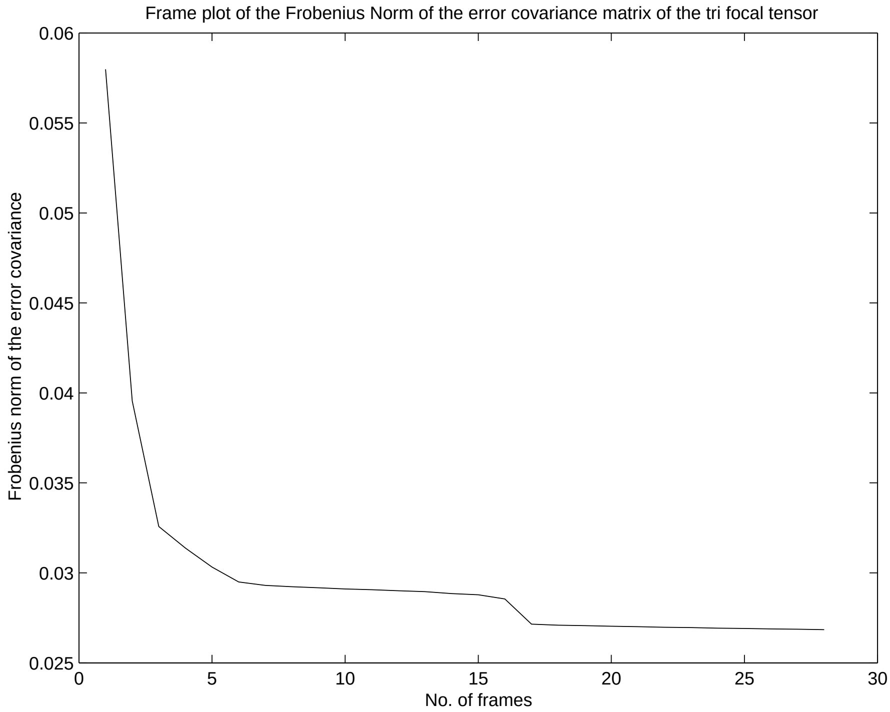

Figure 4.2:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0),\ (T_0, T_v, T_a) = ([1, 1, 1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

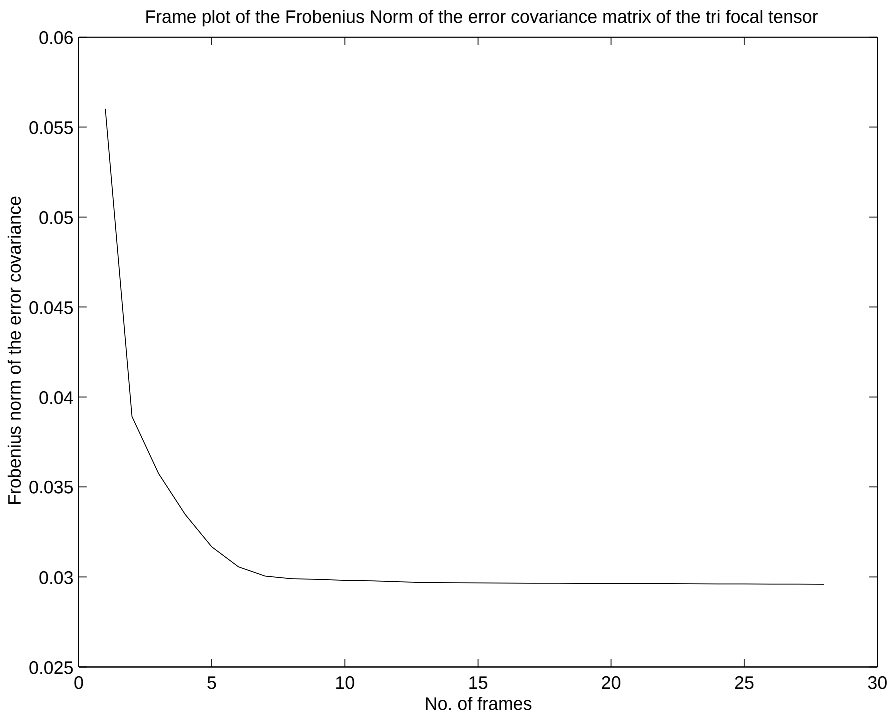

Figure 4.3:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0),\ (T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

4.3.2 $A(t)$ is calculated from the trilinearities

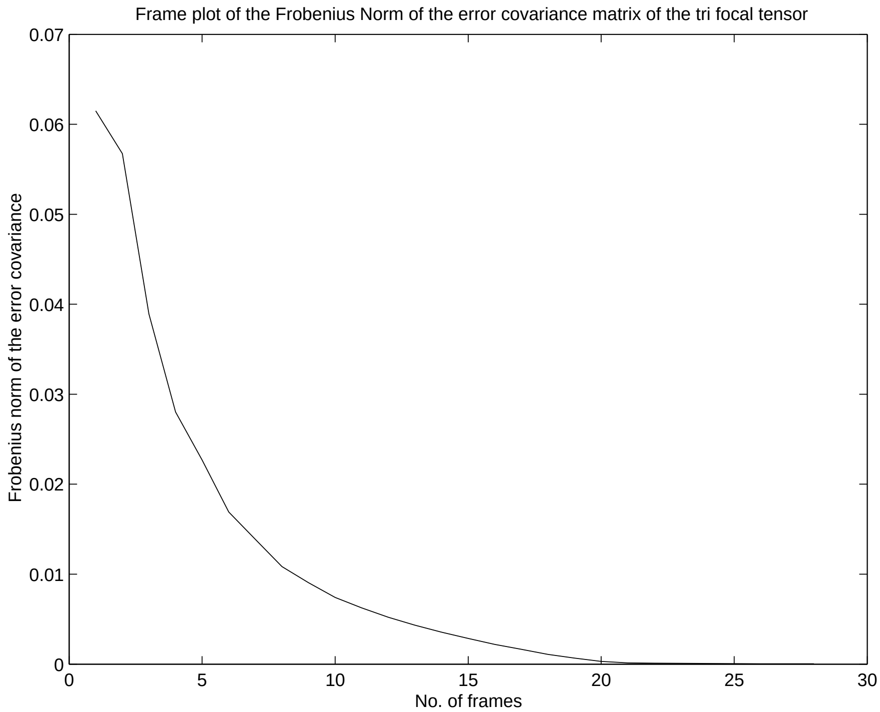

Figure 4.4:
$(n1, n2, n3) = (0, 1, 0) \quad (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0), \quad (T_0, T_v, T_a) = ([2, 4, 5]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

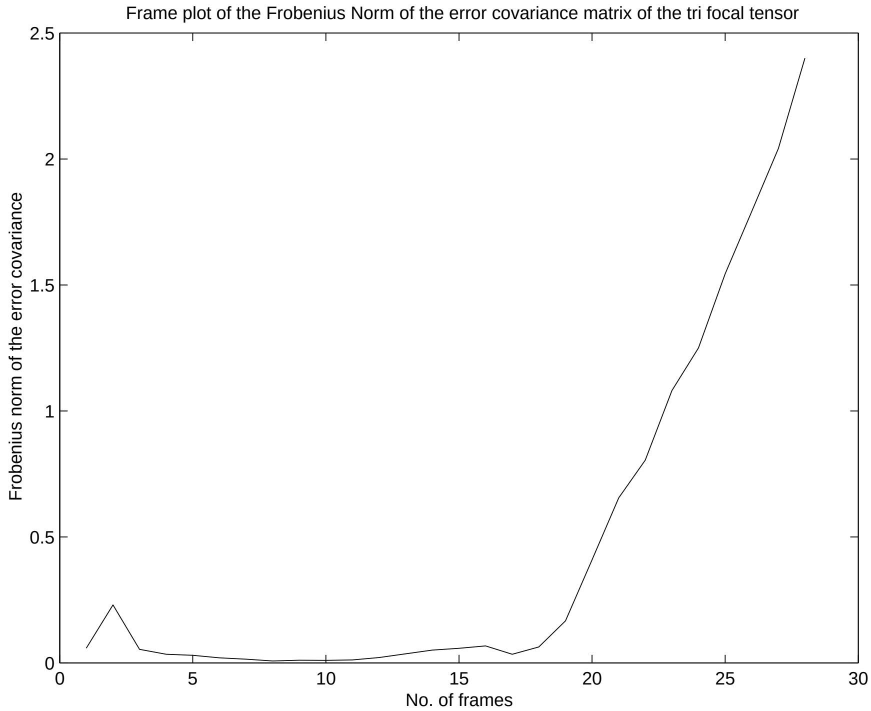

Figure 4.5:
$(n1, n2, n3) = (0, 1, 0) \quad (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0) \quad (T_0, T_v, T_a) = ([1, 1, 1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

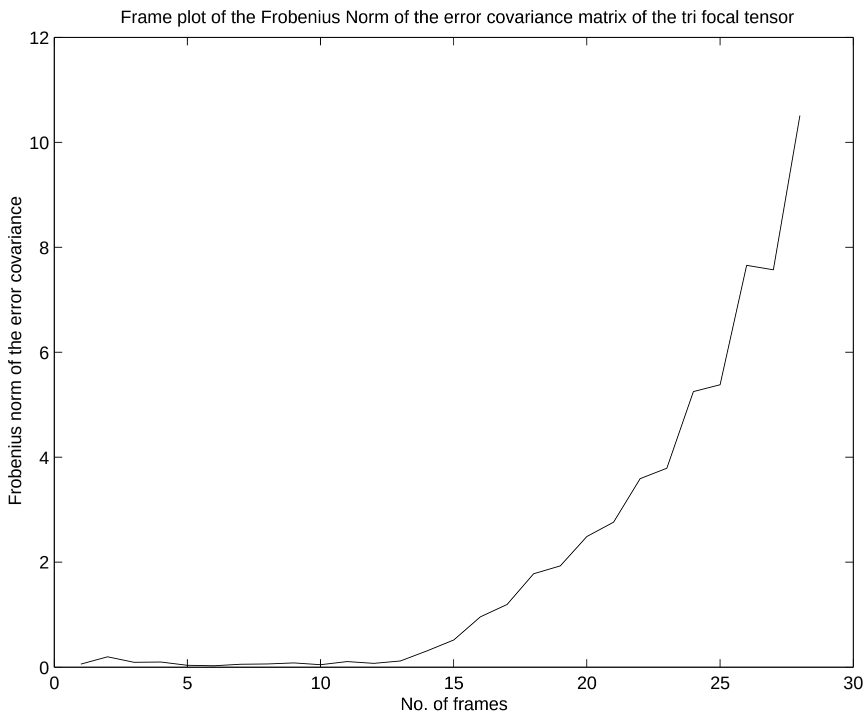

Figure 4.6:
$(n1, n2, n3) = (0, 1, 0)$, $(\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$, $(T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

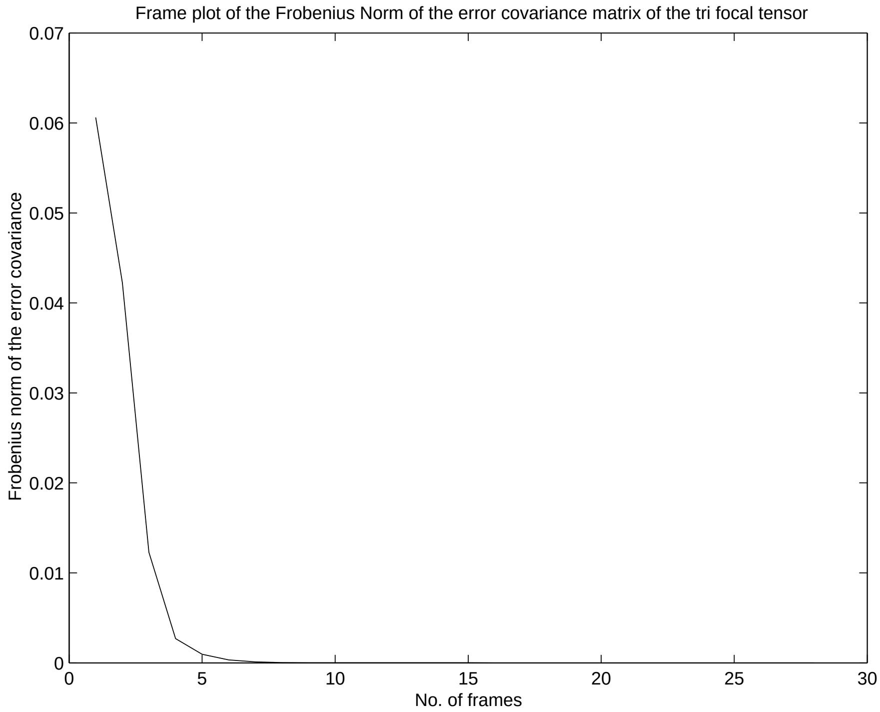

Figure 4.7:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0),\ (T_0, T_v, T_a) = ([2, 4, 5]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

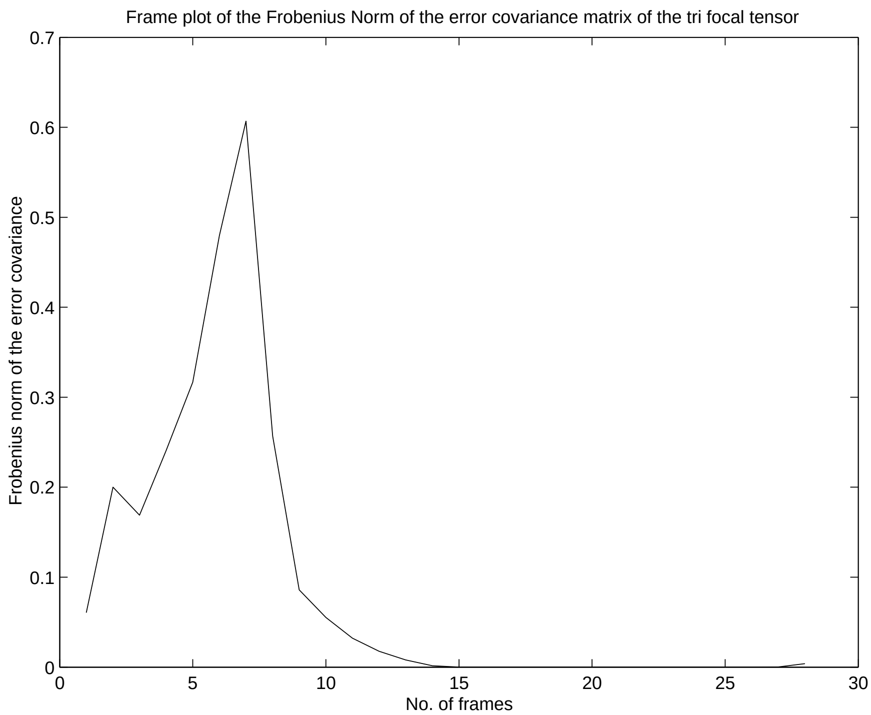

Figure 4.8:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0),\ (T_0, T_v, T_a) = ([1, 1, 1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

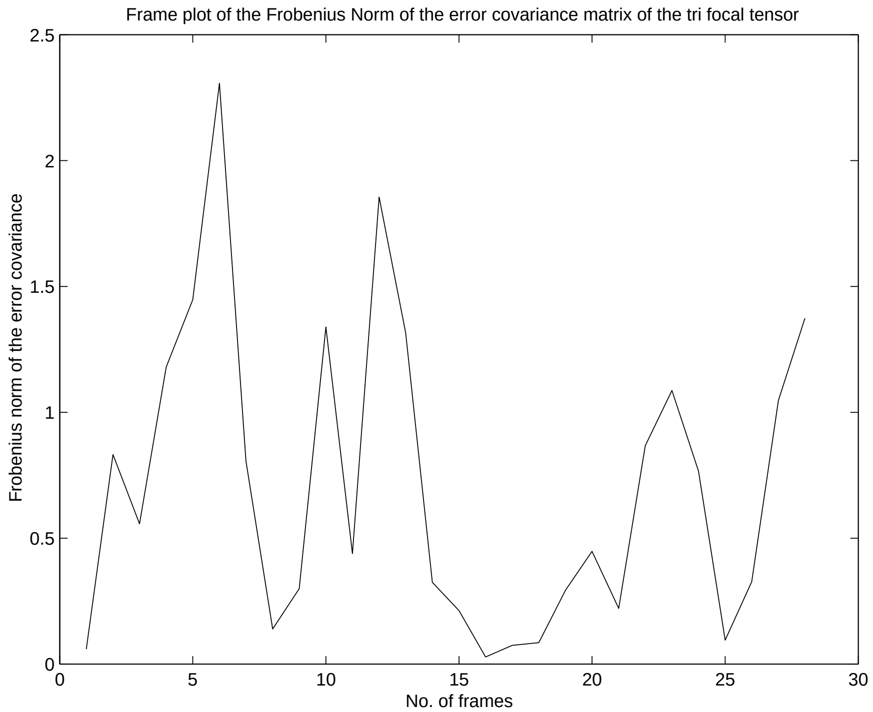

Figure 4.9:
$(n1, n2, n3) = (0, 1, 0)$,$(\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0)$
$(T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

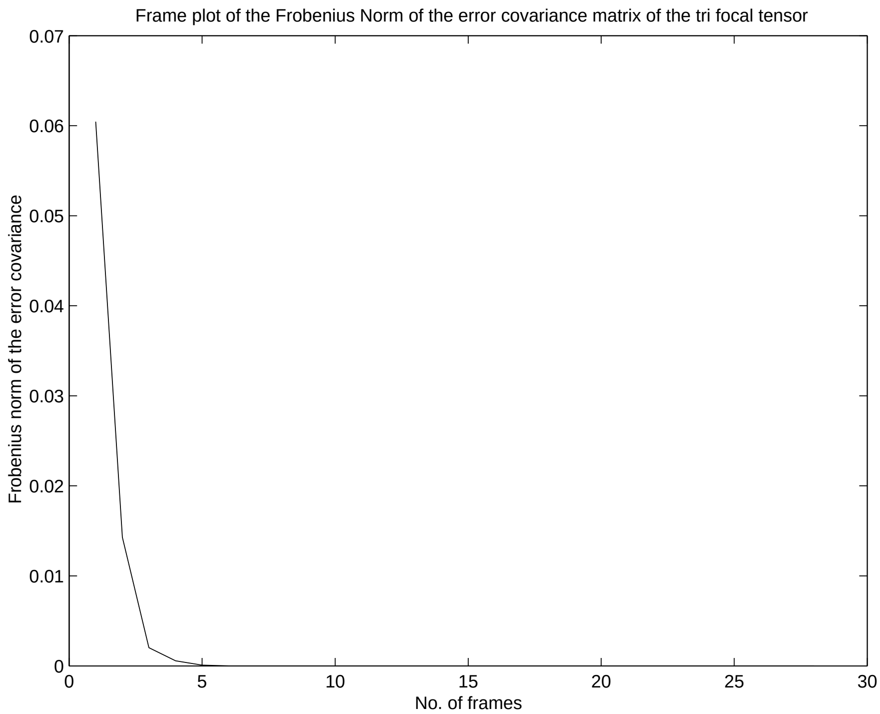

Figure 4.10:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0.05),\ (T_0, T_v, T_a) = ([2, 4, 5]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

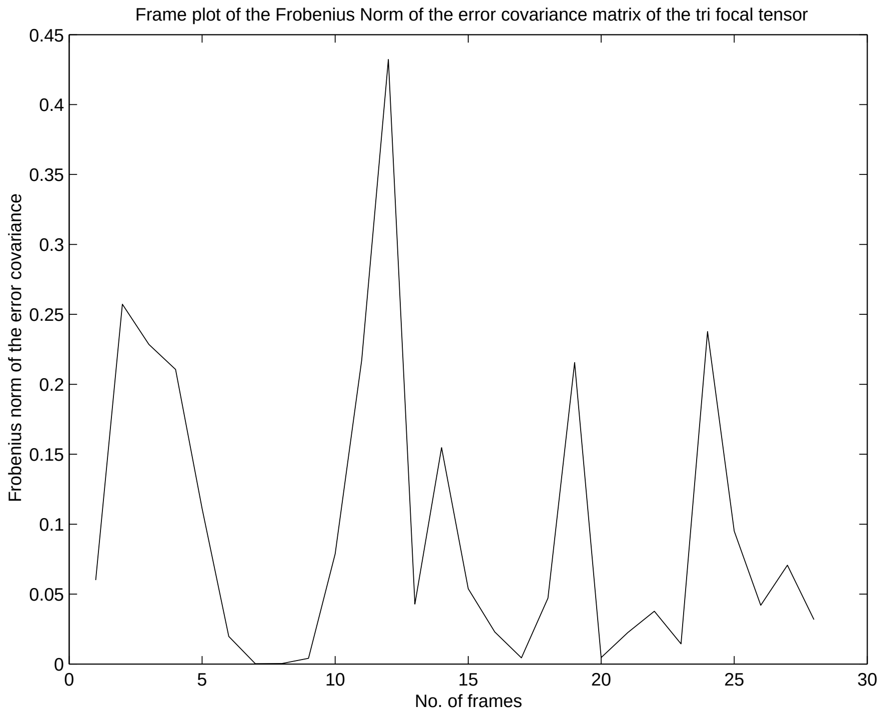

Figure 4.11:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0.05),\ (T_0, T_v, T_a) = ([1, 1, 1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

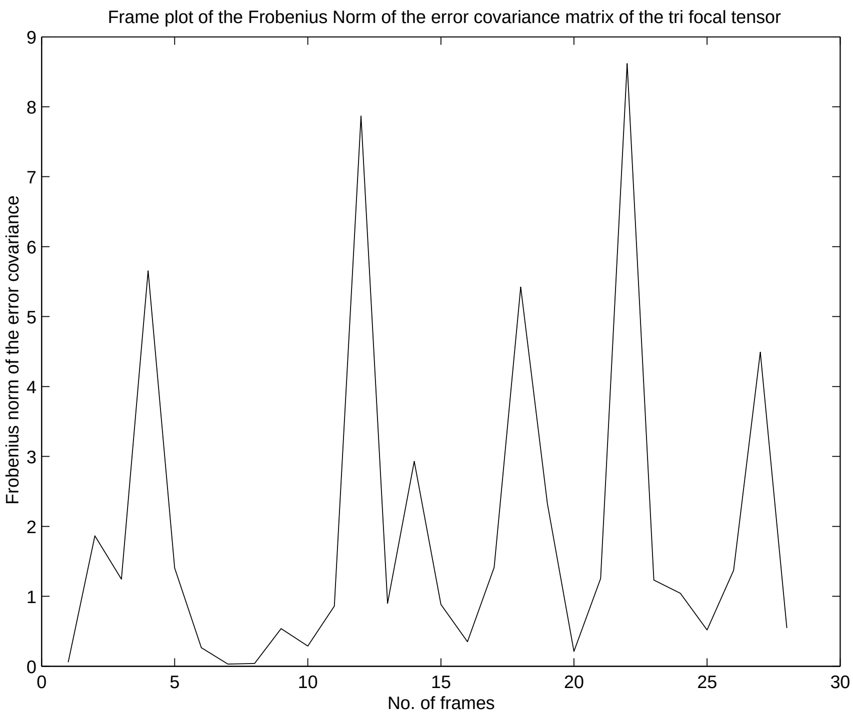

Figure 4.12:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0.05)$
$(T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0, 0, 0]^T, [0, 0, 0]^T)$

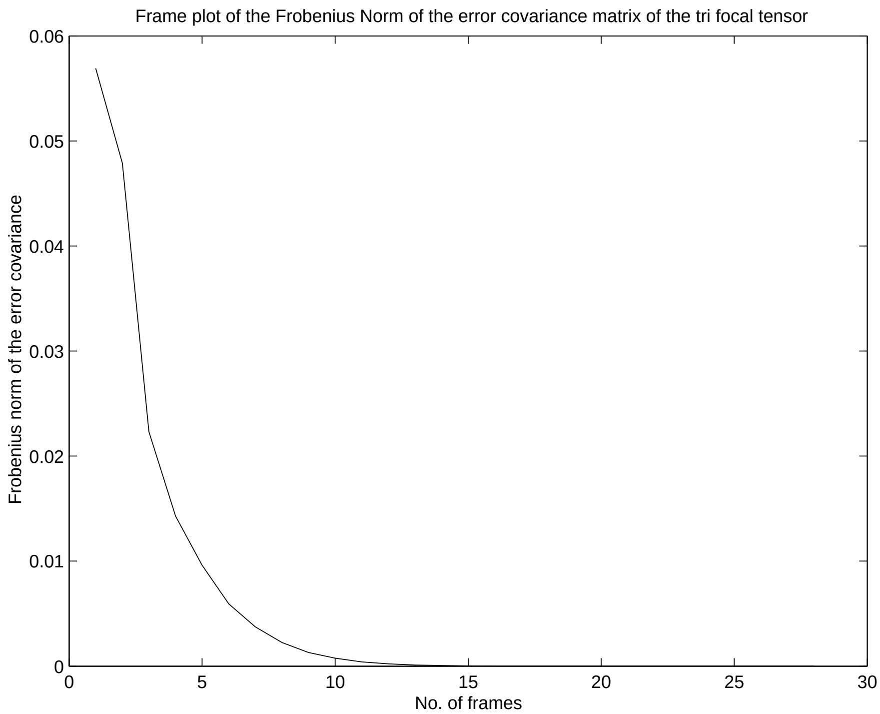

Figure 4.13:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$
$(T_0, T_v, T_a) = ([2, 4, 5]^T, [0.1, 0.1, 0.1]^T, [0, 0, 0]^T)$

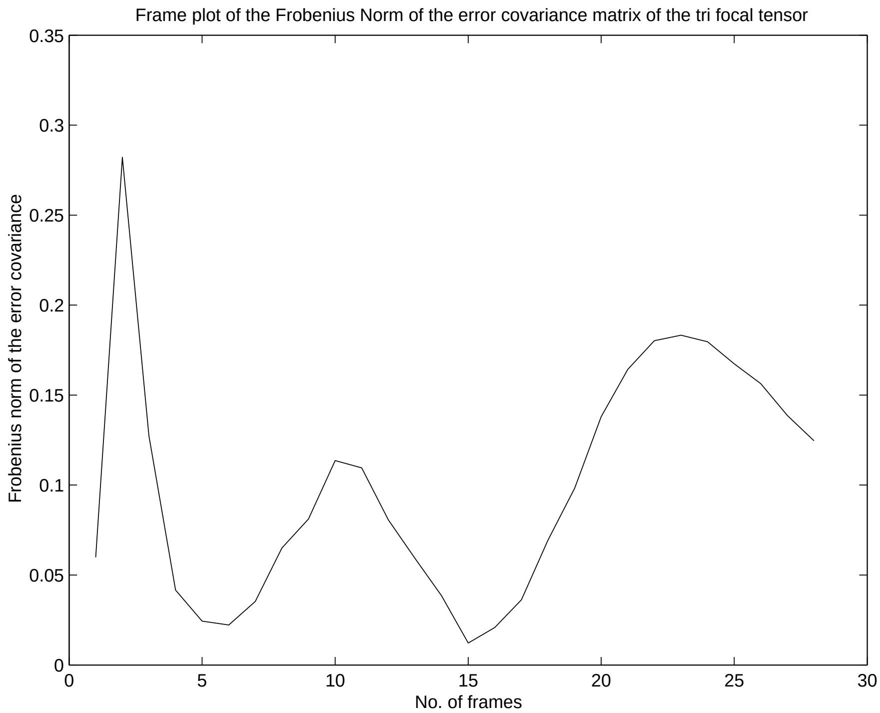

Figure 4.14:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$
$(T_0, T_v, T_a) = ([1, 1, 1]^T, [0.1, 0.1, 0.1]^T, [0, 0, 0]^T)$

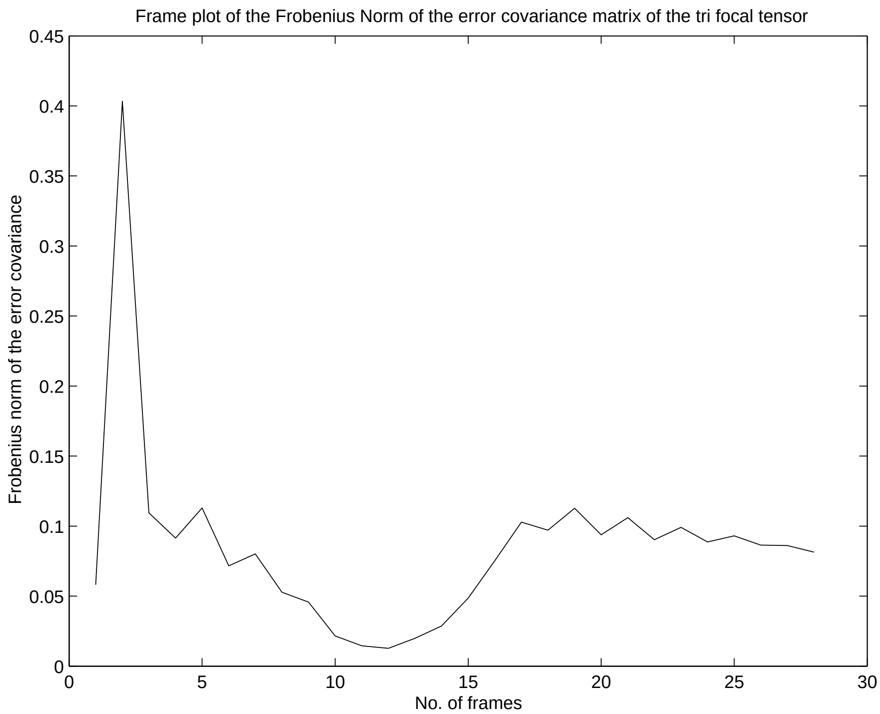

Figure 4.15:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$
$(T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0.1, 0.1, 0.1]^T, [0, 0, 0]^T)$

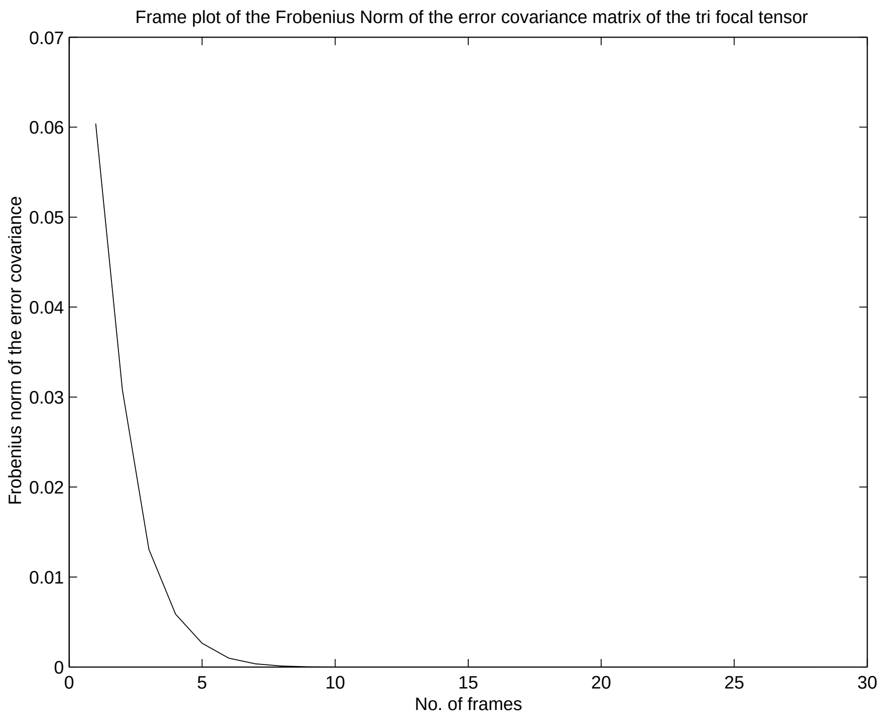

Figure 4.16:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$
$(T_0, T_v, T_a) = ([2, 4, 5]^T, [0.1, 0.3, 0.4]^T, [0.01, 0, 01, 0.01]^T)$

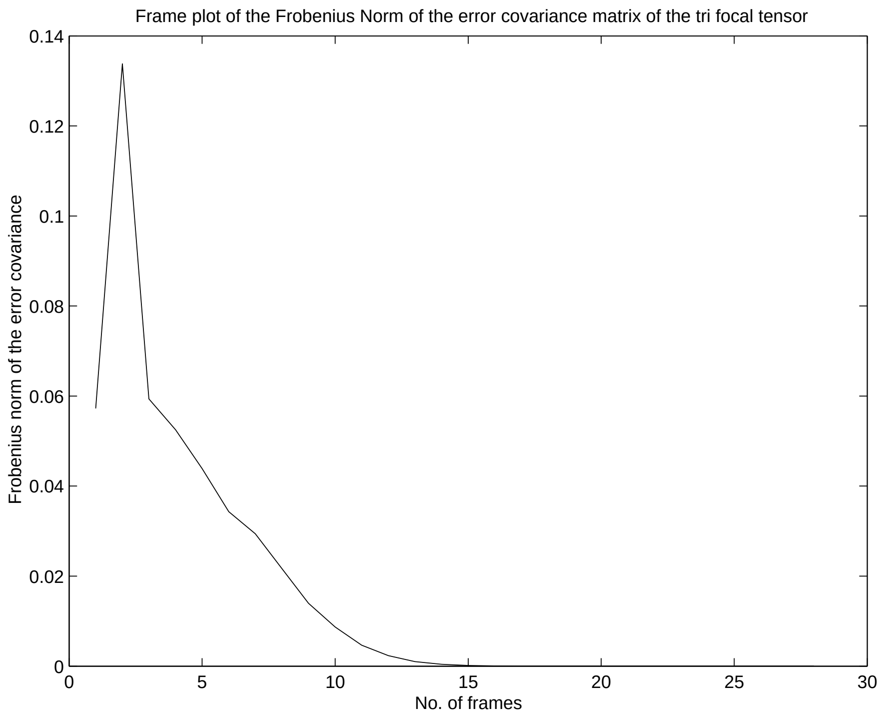

Figure 4.17:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$
$(T_0, T_v, T_a) = ([1, 1, 1]^T, [0.1, 0.3, 0.4]^T, [0.01, 0.01, 0.01]^T)$

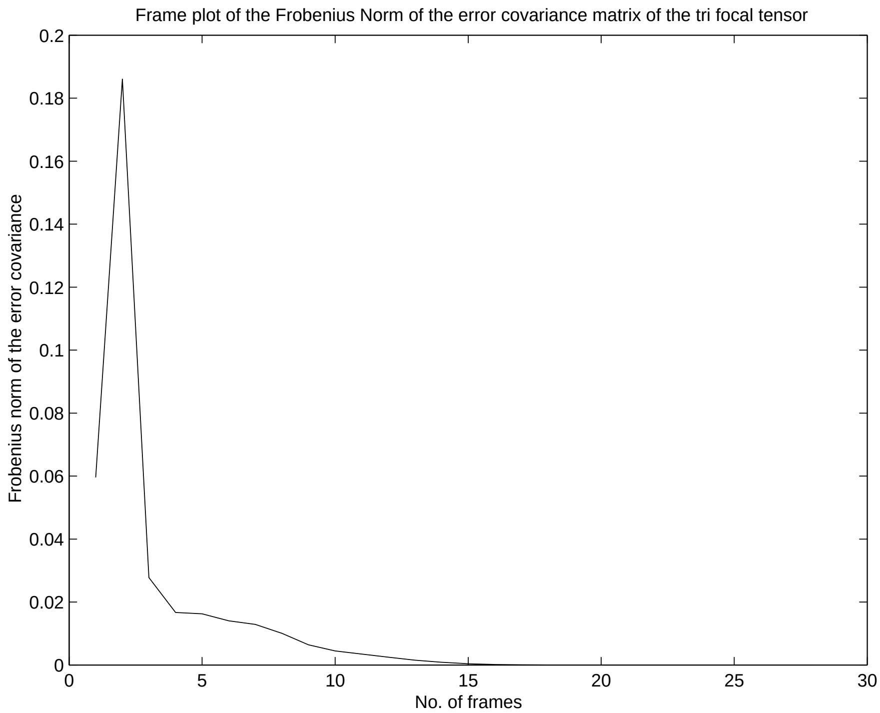

Figure 4.18:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0, 0)$
$(T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0.1, 0.3, 0.4]^T, [0.01, 0.01, 0.01]^T)$

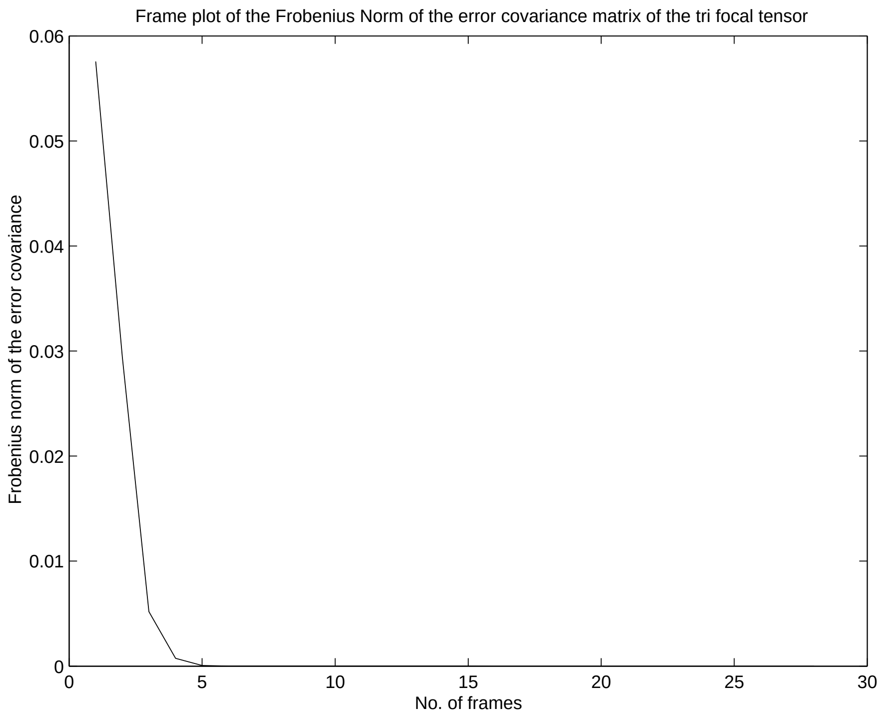

Figure 4.19:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0.05)$
$(T_0, T_v, T_a) = ([2, 4, 5]^T, [0.1, 0.3, 0.4]^T, [0.01, 0, 01, 0.01]^T)$

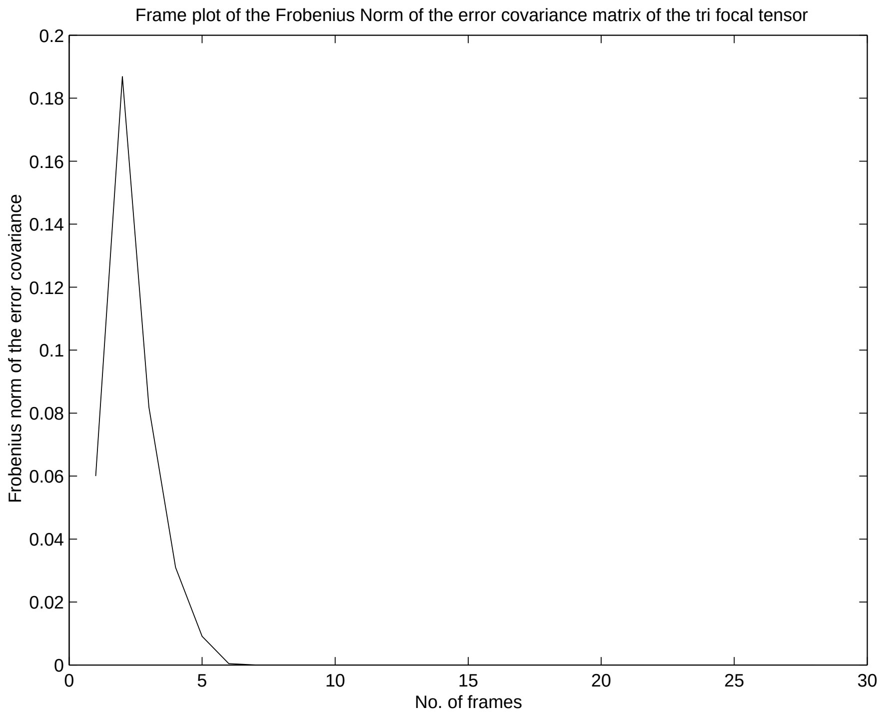

Figure 4.20:
$(n1, n2, n3) = (0, 1, 0),\ (\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0.05)$
$(T_0, T_v, T_a) = ([1, 1, 1]^T, [0.1, 0.3, 0.4]^T, [0.01, 0.01, 0.01]^T)$

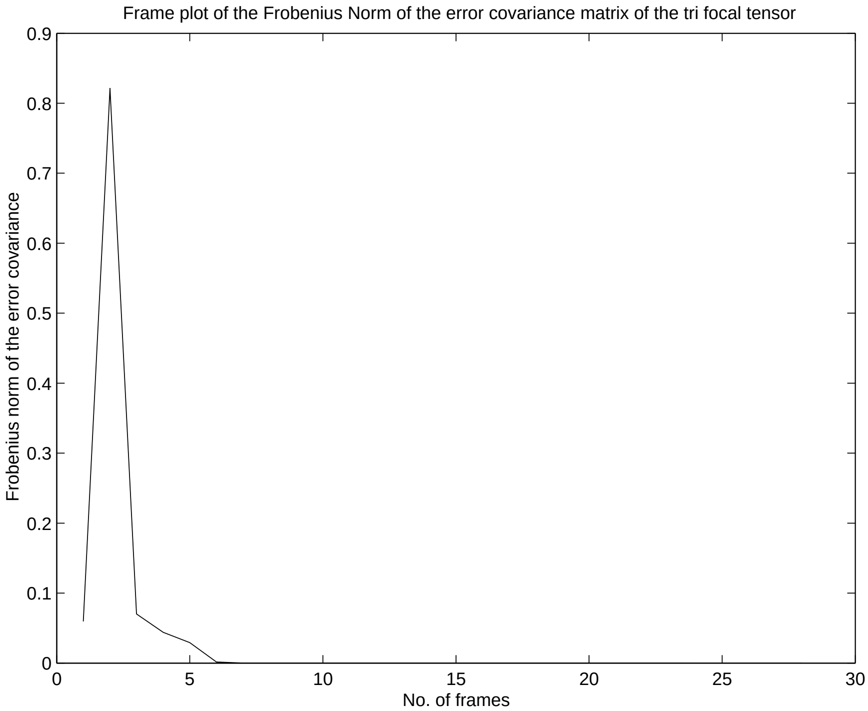

Figure 4.21:
$(n1, n2, n3) = (0, 1, 0)$, $(\phi_0, \phi_v, \phi_a) = (\frac{\pi}{34}, 0.1, 0.05)$
$(T_0, T_v, T_a) = ([0.1, 0.1, 0.1]^T, [0.1, 0.3, 0.4]^T, [0.01, 0.01, 0.01]^T)$

Chapter 5

Conclusion

The results show that model for recursively estimating tri focal tensors is yet not fully functional for all the cases and that it is erroneous for a certain set of inputs. However at the same time the error seems to fall *very fast* in the other cases. This is a positive result as the very idea of estimating tri focal tensors was to exploit correspondences over 3 frames in order to get a more robust model for scene computation. As of now the reconstruction of scene has not been done because the model is still incomplete. Once the model has been tested and verified for synthetic images the results can be tried on actual images to see if the estimation of tri focal tensors *actually* does give better results than using correspondences over 2 image frames.

Chapter 6

Future work

The following are a few suggestions that one could keep in mind while developing the model in future, in order to make it more effective.

- The current model uses the linear model for the Kalman filter. One can probably come up with better results by using the iterated extended Kalman filter.
- As of now while calculating C(t) for the measurement equation of the Kalman filter C(t) is meant to be the co-efficient of a tri focal tensor but it itself has parameters consisting of terms of the tri focal tensor which is being estimated. If possible C(t) should be evaluated in such a way that it does not contain any parameters of the tensor that is being evaluated. If it is not possible, then it becomes even more imperative to use the iterated extended Kalman filter to estimate the tensor.
- A method must be developed for calculating all the motion parameters from the image frames. It should be noted that if the motion parameters are given for the first 3 frames of the process, the motion parameters can be accurately calculated for the remaining frames.

Bibliography

[1] Hartley, R. and Zisserman, A. "Multiple View Geometry in Computer Vision" *Cambridge University Press*, 2000

[2] Candy, J.V "Signal Processing: The Model Based Approach" *McGraw Hill International Editions*, 1986.

[3] Hartley, R. LInes and points in three views and the tri-focal tensor *International Journal of Computer Vision, 22(2)* 1997, pp125-140

[4] Shashua, A. "Algebraic Functions for Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 17, No. 8*, August 1995, pp779-789.

[5] Rousso, B.; Avidan, S.; Shashua, A.; Peleg, S.; "Robust recovery of camera rotation from three frames" *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on , 18-20 June 1996* pp796-802

[6] Tsai, R.; Huang, T.; "Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces" *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No1. January 1984*

[7] Broida, T.J. and Chellappa, R. "Estimation of Object Motion Paramaters from Noisy Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAMI-8, No. 1*, 1986, pp90-99.

[8] Weng, J., Huang, T.S. and Ahuja, N. "3-D Motion Estimation, Understanding and Prediction from Noisy Image Sequences", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-9, No. 3*, 1987, pp370-389

[9] Young, G.S. and Chellappa, R. "3-D Motion Estimation using a Sequence of Noisy Stero Images", 1988

[10] Broida, T.J., Chandrashekhar, S. and Chellappa, R. "Recursive 3-D Motion Estimation from a Monocular Image Sequence" *IEEE Transactions on Aerospace and Electronic Systems Vol, 26, No. 4*, 1990, pp639-656

[11] Chaudhuri, S. and Karandikar, S.S. "Recursive Methods for the Estimation of Rotation Quaternions", *IEEE Transactions on Aerospace and Electronic Systems, Vol 32, No. 2*, 1996, pp845-854