

# Requirements for Motion Estimation in Image Sequences for Traffic Applications

M.B. van Leeuwen, F.C.A. Groen, A. Dev  
University of Amsterdam  
Department of Computer Science  
1098 SJ Amsterdam, The Netherlands

## Abstract

*In traffic applications research in devices that can increase the safety and comfort of vehicles is an important topic. In this paper a digital rear-view mirror is presented that helps the driver to analyse situations on the road behind him. From the observations from a camera the optical motion in the image plane can be estimated. Based on this motion estimation the real-world motion of the vehicles behind our car is interpreted. The motion interpretation problem is very sensitive for errors introduced in the motion estimation. This paper describes a method that shows how accurate the motion estimation must be to enable the required accuracy of the motion interpretation. This will be done by means of simulation experiments for characteristic situations (vehicle approaching, retreating or shifting lane).*

## 1. Introduction

### 1.1. Moving vehicles around you

Far over a 90% of road traffic accidents are attributable to human error. To increase the safety on the road, research is done in a diversity of assistance systems to increase the safety and comfort of vehicles. Important topics are for example automatic cruise control and obstacle avoidance systems [5]. On highways, many accidents occur when a car shifts lane. The work described in this paper is part of a project to develop a digital rear-view mirror. This device should inform the driver whether it is safe to shift lane or not.

To retrieve information about the environment of a vehicle a diversity of sensing devices (based on radar or vision) can be used. A small overview of the methods investigated so far can be found in [8]. Like in [8], we use a camera-based sensor to collect the information. One of the main advantages of vision compared to radar is that it doesn't affect its environment in any way; it is a passive sensor.

The sensory part of our device observes the behaviour of other vehicles behind the car that limit our motion possibilities. From the image sequences provided by the camera the behaviour of other vehicles must be derived. The images are the projection of the 3D-scene behind the car on the 2D-image plane of the camera. The optical motion in this plane contains information about the real motion of the cars. In literature this problem is generally addressed in two steps:

**motion estimation** In this part of the motion estimation algorithm the image displacements between successive image frames are computed.

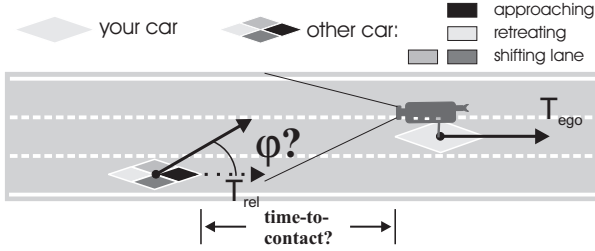
**motion interpretation** In the second part the 2D-image velocity provided by the first part of the motion algorithm is used to estimate parameters of the (3D-) real-world motion.

To solve the motion estimation problem, discrete features are matched in successive frames or the complete optical flow field is computed. A geometric model of image formation is used for the motion interpretation. In the application we are looking at, the motion of the objects is in or parallel to the plane in which the camera is moving (vehicles cannot fly and we assume a flat road). This simplifies our motion estimation problem and restricts the theory explained in this paper to this class of motion problems. However, the theory will still be useful for motion problems encountered when dealing with mobile robots.

In the past, research on motion estimation has concentrated mainly on issues of existence and uniqueness [1] [2] [4] [9] [13]. Since most uniqueness aspects of the problem are now well understood motion estimation research has shifted its focus on the robustness issue. The motion interpretation part of the motion algorithms appears to be very sensitive for errors introduced in the motion estimation part [7]. This robustness problem can be seen as the sensitivity of the geometric projection model for errors introduced in the motion estimation part.

## 1.2. Problem definition

In our paper we will investigate the sensitivity of our application to errors introduced in the motion estimation. We will try to find an answer to the question how accurate the motion estimation should be to enable a satisfying motion interpretation. The typical motion problem in traffic applications is illustrated in figure 1. A camera mounted on the back of a vehicle is used to collect images of the scene behind the car. From these images we want to estimate the number of vehicles present in the scene and their position and motion relative to our vehicle.



**Figure 1. How many vehicles are driving behind you, what is their position and what is their motion relative to your own vehicle?**

From the angle  $\varphi$  between the ego-motion of the camera and the relative motion of another vehicle we can classify its behaviour (vehicle stays, is approaching or retreating, or shifts lane). The time it takes for a vehicle to reach the camera is used to identify possible dangerous situations. This motion parameter will be referred to as the time-to-contact (TTC).

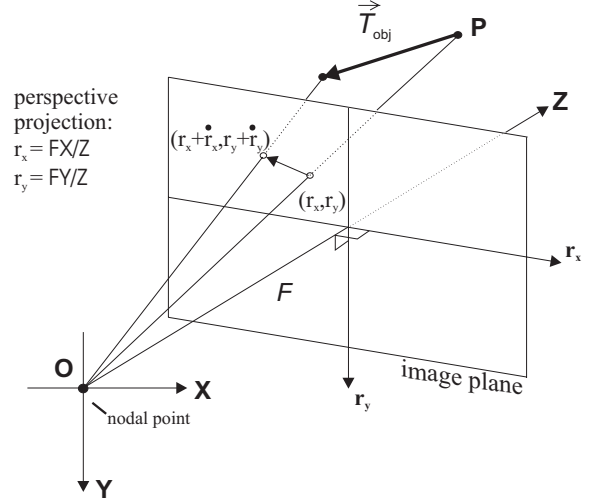
We assume the vehicles in the 3D-scene to be rigid objects and to be moving in or parallel to the plane in which the camera is moving (a flat road). We only consider translational motions. The images obtained with the camera will be affected by the ego-motion of our vehicle. We will assume the translational ego-motion to be known with a specific accuracy ([12]). The results of the motion estimation problem are provided by an optical flow estimation method described in [6].

The geometric model to solve the motion analysis problem is introduced in section 2. The sensitivity of this model will be investigated for two simplified cases in section 3. Once we have formulated the sensitivity of the projection model mathematically we will look what these relations mean. In section 4 we will show some simulation results. We will discuss the meaning of this work for our application and conclude this paper in section 5.

## 2. The projection model

In this section we will derive a projection model for several situations. This projection model approximates the relation between the motion as observed by the camera to the real-world motion. Different geometric projection models have been employed. The most general model is based on perspective projection ([7]). The camera model is illustrated in figure 2. The perspective projection  $(r_x, r_y)$  of a point  $(X, Y, Z)$  in the environment on the image plane is given by

$$r_x = FX/Z \quad \text{and} \quad r_y = FY/Z \quad (1)$$



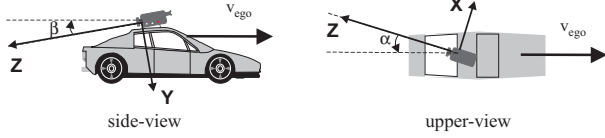
**Figure 2. Perspective projection in case of a static camera.**

Here, the constant  $F$  denotes the focal length of the camera. Previous work [3] [10] [14] [15] has shown how geometric distortions of the camera can affect the observation of the motion and explains how to correct for it. When an object point moves with a translational motion  $\vec{T}_{rel} = [t_{rx}, t_{ry}, t_{rz}]^T$  relative to the nodal point of the camera, its motion on the image plane is given by

$$\dot{r}_x = \frac{F t_{rx} - r_x t_{rz}}{Z} \quad \text{and} \quad \dot{r}_y = \frac{F t_{ry} - r_y t_{rz}}{Z} \quad (2)$$

The parameters  $\dot{r}_x$  and  $\dot{r}_y$  are the projection of the motion on the image plane. Consider the camera moving with constant translational motion given by  $\vec{T}_{ego} = [t_{ex}, t_{ey}, t_{ez}]^T$ , relative to the static part (background) of the scene. The coordinate frame is attached to the nodal point of the camera and the z-axis matches the viewing direction.

We introduce 2 angles ( $\alpha$  and  $\beta$ ) to define the viewing direction of the camera in relation to the driving direction of the vehicle. The choice of these angles is illustrated in figure 3. These angles will be chosen in such a way that the camera observes the appropriate part of the scene behind the vehicle. These parameters might vary in time due to several reasons, like for example irregularities in the surface of the road. As mentioned earlier we will assume the translational ego-motion (and thus the angles  $\alpha$  and  $\beta$ ) to be known with a specific accuracy.



**Figure 3. The angles  $\alpha$  and  $\beta$  define the relation between the viewing direction of the camera and the direction of driving of the vehicle.**

To investigate the sensitivity of the motion parameters (angle  $\varphi$  and the TTC) for variations in the estimated observed motion ( $\hat{r}_x, \hat{r}_y$ ) and variations in the viewing direction of the camera ( $\alpha$  and  $\beta$ ) we have to derive the following relations:

$$\varphi = \angle(\vec{T}_{ego}, \vec{T}_{rel}) = f(F, r_x, r_y, \dot{r}_x, \dot{r}_y, \alpha, \beta) \quad (3)$$

$$TTC = g(\vec{T}_{ego}, \vec{T}_{rel}) = g(F, r_x, r_y, \dot{r}_x, \dot{r}_y, \alpha, \beta) \quad (4)$$

The appendix shows how the relations can be derived for the motion parameter  $\varphi$  for the case  $\beta = 0$ . The relations for  $\varphi$  for the situation  $\beta \neq 0$  and the relations for the TTC can be found in [11]. It is important to notice how these relations depend on the action in the scene. They depend on the relative speed of the objects, their position, etc. To gain insight in the robustness problem, we have to analyse characteristic situations. These situations are 'car shifting lane', 'car approaching' and 'car retreating'. In case of approaching vehicles (with a possible lane shifting steering action), we find the following expression for the motion parameter  $\varphi$ :

$$\varphi|_{\beta=0} = \arctan\left(\frac{r_y \dot{r}_x - r_x \dot{r}_y}{F \dot{r}_y}\right) - \alpha \quad (5)$$

$$\varphi|_{\alpha=0} = \arctan\left(\frac{[r_y - F \tan(\beta)] \dot{r}_x - r_x \dot{r}_y}{F \dot{r}_y \sqrt{1 + \tan^2(\beta)}}\right) \quad (6)$$

### 3. Sensitivity of the projection model

From equations 5 and 6 we can derive the sensitivity for errors in the parameters  $\dot{r}_x, \dot{r}_y, \alpha$  and  $\beta$ . This is shown in the appendix for  $\beta = 0$ . If  $\beta = 0$  and we observe the motion of an approaching vehicle, we get the following expressions for the sensitivity of  $\varphi$  (expressed in coordinates  $(X, Y, Z)$ ):

$$\left. \frac{\partial \varphi}{\partial \dot{r}_x} \right|_{\beta=0} = \frac{Z t_{rz}}{F (t_{rx}^2 + t_{rz}^2)} \quad (7)$$

$$\left. \frac{\partial \varphi}{\partial \dot{r}_y} \right|_{\beta=0} = \frac{Z (Z t_{rx} - X t_{rz})}{Y F (t_{rx}^2 + t_{rz}^2)} \quad (8)$$

$$\left. \frac{\partial \varphi}{\partial \alpha} \right|_{\beta=0} = -1 \quad (9)$$

Some conclusions concerning the sensitivity of the motion parameter  $\varphi$  for errors in the estimation of the parameters  $\dot{r}_x, \dot{r}_y, \alpha$  and  $\beta$  (ref. appendix and [11]) are:

- The sensitivity of  $\varphi$  for errors in  $\hat{r}_x$  or  $\hat{r}_y$  is inversely proportional to the focal length of the camera.
- The sensitivity of  $\varphi$  for errors in  $\hat{\alpha}$  or  $\hat{\beta}$  does not depend on the focal length.
- The sensitivity of  $\varphi$  for errors in  $\hat{r}_x$  is not a function of the coordinates  $X$  and  $Y$ . In other words: the sensitivity of  $\varphi$  for errors in the x-component of the optical flow does not depend on the position in the image plane where the flow was estimated.
- The sensitivity of  $\varphi$  for errors in  $\hat{r}_y$  or  $\hat{\beta}$  does depend on the position in the image plane where the flow was estimated. The sensitivity for errors in these estimations will be large when the plane in which the estimated motion occurs is close to the plane in which the nodal point of the camera is moving (infinity in case the motion occurs in the same plane).
- The sensitivity of  $\varphi$  for errors in  $\hat{r}_x$  or  $\hat{r}_y$  increases when the approaching vehicle is further away.
- The sensitivity of  $\varphi$  for errors in  $\hat{r}_x$  or  $\hat{r}_y$  decreases when the relative motion of the vehicle in the negative z-direction increases.
- In case of approaching or retreating vehicles, the sensitivity of  $\varphi$  for errors in  $\hat{\beta}$  does not depend on the distance of the vehicle. However for the situation of a vehicle that shifts lane, this sensitivity increases when the lane shift occurs farther away from the camera.

Equations 7, 8 and 9 should be used to calculate worst case scenarios for  $\varphi$ , using (combinations of)

$$\Delta\varphi_{\dot{r}_x} = \left| \frac{\partial\varphi}{\partial\dot{r}_x} \right| \cdot \Delta\dot{r}_x \quad , \quad \Delta\varphi_{\dot{r}_y} = \left| \frac{\partial\varphi}{\partial\dot{r}_y} \right| \cdot \Delta\dot{r}_y \quad (10)$$

$$\Delta\varphi_\alpha = \left| \frac{\partial\varphi}{\partial\alpha} \right| \cdot \Delta\alpha \quad , \quad \Delta\varphi_\beta = \left| \frac{\partial\varphi}{\partial\beta} \right| \cdot \Delta\beta \quad (11)$$

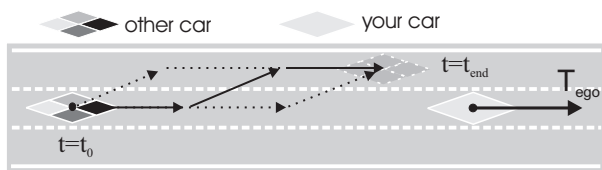
The same goes for the sensitivity of the motion parameter TTC.

#### 4. Simulation results

The previous approach enables one to investigate the robustness of the motion estimation algorithm for traffic applications. Simulations are required to describe characteristic situation of the application. The typical traffic situations (like illustrated in figure 4) are analysed by simulation experiments. The conclusions in the previous section showed that the distance between the plane in which the camera is moving and the plane in which the observed motion occurs should be as large as possible. In our case the camera is placed on top of the vehicle (1.5m above the surface of the road). Therefore it is wise to consider the motion of points of the observed vehicle close to the surface of the road. We have chosen to look at points around the licence plate of the observed vehicle. The vehicles behind us are represented as a moving plane perpendicular to the motion direction of the camera. This plane of size (h<sub>xw</sub>=0.15x1.7m) moves at an average height of 0.35m above the surface of the road.

The experiments give an answer to the following problems:

1. Given the accuracy of a solution for the correspondence problem, what error in the estimation of the motion parameters is to be expected for specific situations.
2. Given a specific situation, how accurate should the estimations of the correspondence problem be in order to enable estimations of the motion parameters with satisfying accuracy.



**Figure 4. Example of the simulation for the characteristic situation 'vehicle in same lane shifts to the left'.**

The results of simulations for the first class of problems depend heavily on the situation one is looking at. For examples of these results we refer to [11].

The results of the second class are of a more general character. The following figures show the simulation results for three different situations. In each simulated situation the camera is moving with 30m/s (108km/h) in the negative z-direction ( $\alpha = \beta = 0$ ). Furthermore  $F=5\text{mm}$  and the pixel size equals  $33\mu\text{m}$ .

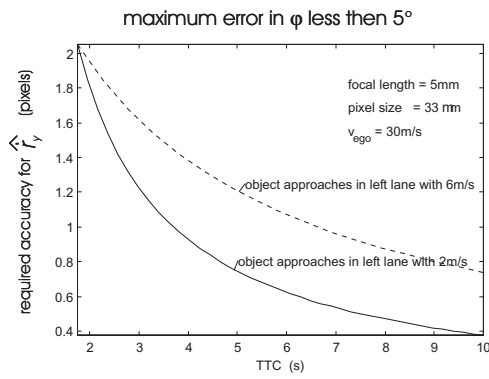
Figure 5: This figure is based on the simulation of a vehicle that approaches our vehicle in the left lane with constant speed. The results are shown for vehicles moving with 32m/s (115km/h) and 36m/s (130km/h). The figure shows the required accuracy for the estimation of the parameter  $\dot{r}_y$  that is necessary to estimate the motion parameter  $\varphi$  with an error less than  $5^\circ$ . This required accuracy is plotted as a function of the time-to-contact. For example, one might be interested in a solution to estimate the motion parameter  $\varphi$  for objects approaching with 2 m/s (or more) from the moment they are within 4 seconds from contact. The figure shows that for the simulated situation one needs to estimate the optical flow in the y-direction with approximately 1 pixel accuracy in order to estimate the angle  $\varphi$  with the required accuracy.

Figure 6: This figure is based on the simulation of a vehicle that first approaches in the right lane, than shifts left to reach the left lane and finally overtakes in the left lane. The results for lane shifts that are shown are initiated at four different distances (: 15m, 20m, 25m and 30m). The approaching speed of the vehicles is 32m/s (115km/h). During the lane shift an extra component is added of 0.9m/s (3km/h) in the x-direction. Again the relation between the required accuracy in  $\hat{r}_y$  and the motion parameter  $\varphi$  is shown.

Figure 7: This figure is based on the simulation of a vehicle that approaches in the same lane with 32m/s (115km/h) and 36m/s (130km/h). In this figure the relation between the required accuracy in  $\hat{r}_y$  and the TTC is given under the condition that the maximum allowed error in the estimation of the motion parameter TTC does not exceed 10% of its true value.

#### 5. Conclusions

The figures give an indication of the required accuracy in the estimations of the correspondence problem. However, the results heavily depend on the situation one is looking at. For example the focal length and pixel size of the camera, the height at which the camera is placed, the position of the points under consideration at the observed vehicle, etc. One can use the theory explained in this paper to gain insight in the influence of different parameters on the sensitivity problem. This insight will, together with restrictions following from the specific application, provide enough information to derive the relations like illustrated in the figures 5, 6 and 7.

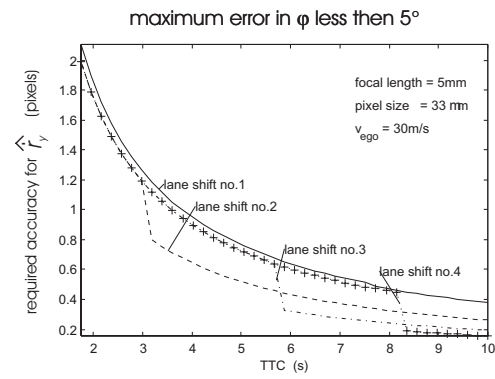


**Figure 5. Simulated situation: 'vehicle approaches in left lane'.** The figure illustrates which accuracy is required for  $\hat{r}_y$  in order to estimate the angle  $\varphi$  with a satisfying accuracy.

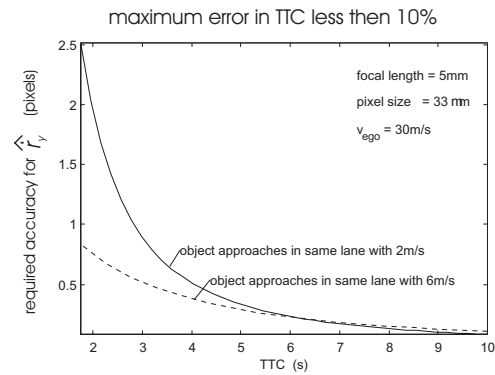
We used simulations to gain insight in the sensitivity problem. Next we will test this theory on real-world image sequences. The theory presented in this paper only considers vehicles driving on a straight road. We will extend this work for the case of curved parts of the road. We will also look at the influence of rotations of the camera around the y-axis (e.g. due to irregularities in the surface of the road).

## References

- [1] G. Adiv. Inherent ambiguities in recovering 3-d motion and structure from a noisy flow field. In *Proc. IEEE Conf. on Comp. Vision and Pattern Rec.*, pages 70–77, 1985.
- [2] F. Bergholm. Motion from flow along contours: A note on robustness and ambiguous cases. *Int. J. of Computer Vision*, 3:395–415, 1988.
- [3] H. A. Beyer. An introduction to photogrammetric camera calibration. Invited paper, seminaire Orasis, St. Malo, sept. 24-27, 1991.
- [4] T. Brodsky, C. Fermüller, and Y. Aloimonos. Directions of motion fields are hardly ever ambiguous. *Int. J. of Computer Vision*, 26(1):5–24, 1998.
- [5] Demo98. Automated vehicle guidance. Demo'98, Rijnwoude, The Netherlands, June 15-19, 1998.
- [6] A. Dev. *Visual Navigation on Optical Flow*. PhD thesis, University of Amsterdam, 1998.
- [7] C. Fermüller and Y. Aloimonos. Qualitative egomotion. *Int. J. of Computer Vision*, 15:7–29, 1995.
- [8] A. Giachetti, M. Campani, and V. Torre. The use of optical flow for road navigation. *IEEE Trans. on Robotics and Automation*, 14(1):34–48, Feb. 1998.
- [9] B. K. P. Horn. Motion fields are hardly ever ambiguous. *Int. J. of Computer Vision*, 1:259–274, 1987.
- [10] S. Kuijpers. Robot navigatie op basis van optic flow. MSc thesis, University of Amsterdam, 1998.



**Figure 6. Simulated situation: 'vehicle shifts lane(s)'.** The figure illustrates which accuracy is required in  $\hat{r}_y$  in order to the angle  $\varphi$  with a satisfying accuracy.



**Figure 7. Simulated situation: 'vehicle approaches in same lane'.** The figure illustrates which accuracy is required in  $\hat{r}_y$  in order to estimate the TTC with a satisfying accuracy.

- [11] M. B. van Leeuwen and F. C. A. Groen. Requirements for motion estimation in image sequences for traffic applications. Technical Report CS-99-01, University of Amsterdam, 1999.
- [12] T. Y. Tian, C. Tomasi, and D. J. Heeger. Comparison of approaches to egomotion computation. In *Proc. of Computer Vision and Pattern Rec.*, 1996.
- [13] R. Y. Tsai and T. S. Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 6(1):13–27, 1984.
- [14] P. R. Wolf. *Elements of Photogrammetry*. McGraw-Hill book company, 1986.
- [15] H. J. Woltring. Planar control in multi-camera calibration for 3d-gait studies. *J. Biomechanics*, 13:39–48, 1980.

In this appendix the relation  $\varphi = \angle(\vec{T}_{ego}, \vec{T}_{rel}) = f(F, r_x, r_y, \dot{r}_x, \dot{r}_y, \alpha, \beta)$  is derived. This relation is given by the geometric model formulised by equations 1 and 2. From this relation, the sensitivity of  $\varphi$  for errors in the estimations of  $\dot{r}_x$ ,  $\dot{r}_y$ ,  $\alpha$  and  $\beta$  can be derived.

Assuming a flat straight road and a constant translational motion of the camera, the motion of an object behind you will be in a plane parallel to the surface of the road. For the case  $\beta = 0$  (see figure 3) the relative motion of an object point in the y-direction will be equal to zero. Using the equations from 2 we can find the following expression for the ego motion and relative motion:

$$\vec{T}_{rel} = Z \cdot \left[ \frac{r_y \dot{r}_x - r_x \dot{r}_y}{F r_y}, 0, -\frac{\dot{r}_y}{r_y} \right]^T \quad \text{and} \quad \vec{T}_{ego} = t_{ez} \cdot [\sin(\alpha), 0, \cos(\alpha)]^T \quad (12)$$

From these vectors the direction of the relative motion of the object points can be derived. The angle  $\varphi$  equals the angle between the vectors

$$[-\sin(\alpha), 0, -\cos(\alpha)]^T \quad \text{and} \quad \text{sign}(r_y \dot{r}_y) \cdot \left[ \cos(\alpha) \frac{r_y \dot{r}_x - r_x \dot{r}_y}{F \dot{r}_y}, 0, -\cos(\alpha) \right]^T \quad (13)$$

Notice that  $\text{sign}(Z) = +1$  and  $\text{sign}(t_{ez}) = -1$  for our application. Previous relations can be simplified depending on the situation we are looking at. From these results, the angle  $\varphi$  can be derived. Equation 13 shows that the estimation of the motion parameter will heavily depend on the situation we are looking at. In this appendix we will use equation 13 to derive the expression for the angle  $\varphi$  in case of approaching objects.

We will consider  $0 \leq \alpha < +\pi/2$ . It is easy to show that  $(\beta = 0) \text{sign}(\dot{r}_y r_y) > 0$  in case of approaching vehicles and  $\text{sign}(\dot{r}_y r_y) < 0$  in case of retreating vehicles. Using relation 13 we find the following expression for  $\varphi$ :

$$\varphi|_{\beta=0} = \arctan\left(\frac{r_y \dot{r}_x - r_x \dot{r}_y}{F \dot{r}_y}\right) - \alpha \quad (14)$$

From 14 we can derive the sensitivity for errors in the parameters  $\dot{r}_x$ ,  $\dot{r}_y$  and  $\alpha$ :

$$\left. \frac{\partial \varphi}{\partial \dot{r}_x} \right|_{\beta=0} = \frac{r_y F \dot{r}_y}{(F \dot{r}_y)^2 + (r_y \dot{r}_x - r_x \dot{r}_y)^2}, \quad \left. \frac{\partial \varphi}{\partial \dot{r}_y} \right|_{\beta=0} = \frac{-r_y F \dot{r}_x}{(F \dot{r}_y)^2 + (r_y \dot{r}_x - r_x \dot{r}_y)^2}, \quad \left. \frac{\partial \varphi}{\partial \alpha} \right|_{\beta=0} = -1 \quad (15)$$

Inserting the relations for perspective projection (: equations 1 and 2) in previous equations results in

$$\left. \frac{\partial \varphi}{\partial \dot{r}_x} \right|_{\beta=0} = \frac{-Z t_{rz}}{F (t_{rx}^2 + t_{rz}^2)} \quad \text{and} \quad \left. \frac{\partial \varphi}{\partial \dot{r}_y} \right|_{\beta=0} = \frac{Z (X t_{rz} - Z t_{rx})}{F Y (t_{rx}^2 + t_{rz}^2)} \quad (16)$$

Analogous the following relations can be derived for the case  $\alpha = 0$ . For the derivation of these relations see [11].

$$\left. \frac{\partial \varphi}{\partial \dot{r}_x} \right|_{\alpha=0} = \frac{F \dot{r}_y [r_y + F \tan(\beta)] \sqrt{1 + \tan^2(\beta)}}{\left( F \dot{r}_y \sqrt{1 + \tan^2(\beta)} \right)^2 + ([r_y + F \tan(\beta)] \dot{r}_x - r_x \dot{r}_y)^2} \quad (17)$$

$$\left. \frac{\partial \varphi}{\partial \dot{r}_y} \right|_{\alpha=0} = -\frac{F \dot{r}_x [r_y + F \tan(\beta)] \sqrt{1 + \tan^2(\beta)}}{\left( F \dot{r}_y \sqrt{1 + \tan^2(\beta)} \right)^2 + ([r_y + F \tan(\beta)] \dot{r}_x - r_x \dot{r}_y)^2} \quad (18)$$

$$\left. \frac{\partial \varphi}{\partial \beta} \right|_{\alpha=0} = -\frac{F \dot{r}_y [(r_y \dot{r}_x - r_x \dot{r}_y) \tan(\beta) - F \dot{r}_x] \sqrt{1 + \tan^2(\beta)}}{\left( F \dot{r}_y \sqrt{1 + \tan^2(\beta)} \right)^2 + ([r_y + F \tan(\beta)] \dot{r}_x - r_x \dot{r}_y)^2} \quad (19)$$