AUTONOMOUS GUIDED VEHICLE

Sensor analysis

Magnus Nørgaard[⋆◇]
Niels Kjølstad Poulsen[⋆]
Ole Ravn[◇]

[⋆]Department of Matematical Modellering, building 321
[◇]Department of Automation, building 326

The Technical University of Denmark
DK-2800 Lyngby

1998-02-03 16.02

1 Introduction

This report documents the sensor analysis phase of the STVF funded project: *Event based Kalman Filtering*. The characteristics of different sensors relevant to the Autonomous Guided Vehicle (AGV) project at the Department of Automation, DTU, are investigated. Currently the AGV is equipped with three types of sensors: a wheel encoder on each of the two driving wheels, two cameras mounted on pan-tilt units, and an ultrasound tranciever/receiver. Models of the sensors are derived and an attempt is made to establish useful stochastic descriptions of the uncertainty of the sensors. Such stochastic models of the sensors and their uncertainty are necessary for data fusion purposes.

2 The Wheel Encoder

Two identical DC-motors are used for driving the two front wheels of the AGV. Each of these motors are supplied with a digital encoder. The encoders work as reset integrators with quantization: the encoder generates a pulse each time the motor shaft has rotated a certain angle and these pulses are summed up by a counter. When the number of counted pulses has been read the counter is reset to zero. The encoder thus provides a quantisized measure of the rotation angle of the motor shaft between sampling instants. If slipping is disregarded and the circumference of the wheel is known this angle relates directly to the driven distance of the corresponding wheel. By dividing the rotated angle in one sampling period with the time of the period a measure of the average velocity of the motor shaft is obtained.

2.1 Encoder Model

Usually the manufacturer of the encoder specifies the number of pulses that the encoder will generate per rotation. If the *encoder gain* is introduced as

$$k_{enc} = \frac{\text{pulses per rotation}}{2\pi} \; [pulses/rad] \tag{1}$$

the output of the encoder is thus given by

$$z(t_{k+1}) = int\left(\int_{t_k}^{t_{k+1}} k_{enc}\theta(t)dt\right) \tag{2}$$

where $\theta(t)$ is the rotation of the motor shaft measured in radians and int is the integer value of the integral.

2.2 Quantification of the uncertainty

When the encoders are used for estimating the driven distance of each wheel the numbers read from the counter at each sampling instant are simply added. For sensor fusion purposes it is practical to convert the resolution of the encoders to a resolution in terms of the distance driven by the wheel:

$$h_{enc} = \frac{r}{k_{enc}N} \; [meters/pulse] \tag{3}$$

r is the radius of the wheel and N is the gear between motor shaft and wheel axle.

According to [Lev82] the measurement error due to the quantization can be assumed to be white and Gaussian distributed with mean value $-\frac{h_{enc}}{2}$ and variance $\frac{h_{enc}^2}{12}$.

3 The Ultrasound Tranceiver/Receiver

This is not yet fully integrated on the AGV and thus it is not documented in this report.

4 The Camera as a Sensor

The camera is by far the most versatile of the sensors used on IAU's autonomous guided vehicle. From the same digital image it is possible extract completely different features and furthermore the same features can often be attained in quite different ways. Currently the vehicle applies digital image processing for three different tasks: to track a tape course laid out on the floor, to detect the opening angle of a door and to detect artificial guide marks on the walls. One of the near future challenges is to be able to select and recognize natural guide marks in the environment, such as lamps, door knobs, switches, etc.

Since it is possible to perform several types of analyses on the same image it is quite common to consider the outcome of a particular image processing algorithm as a "virtual" sensor. In this case the

sensor model is of course highly dependent on the selected image processing algorithm.

Some general characteristics of the camera sensor that are relevant in a multi sensor data fusion context have been listed below:

- There will often be significant time delays due to the scanning time required to acquire the image and the time required for the processing of the image.
- Sometimes the response time of an image analysis algorithm is unknown. One example is when the algorithm requires an unknown number of iterations to complete. Another is when the algorithm fails to find the features it is looking for. When the response time is unknown it is difficult to comply with real-time demands.
- The accuracy of the result depends on the knowledge of various *extrinsic* and *intrinsic* parameters. The extrinsic parameters are the camera's height above the floor, its pan and tilt angles and its displacement relative to a well-defined point on the vehicle or in a world fixed frame. What the set of intrinsic parameters consists of depends on how the actual camera hardware is modeled. Some examples are focal length, the position of the center point and distortion coefficients. It is possible to measure the extrinsic parameters by hand, but usually the accuracy is enhanced if they are determined through an appropriate calibration procedure. The intrinsic parameters are always determined by calibration. See for example [Sør94] and [Lar97].

The rest of this section will summarize how a 2D camera image relates to the physical 3D world. Most or all of the image processing is usually carried out in a 2D coordinate system, but when the important features have been extracted it is necessary to transform this information to an appropriate 3D coordinate system. The following sections will treat the tracking of the tape course (section 4.2) and the guide mark detection (section 4.3). The actual image processing details will not be covered. The attention is restricted to the issue of how the detected features are related to the physical world.

4.1 Relating the Camera Image to the Physical World

Depending on the task to be solved different coordinate systems can be used for describing the physical world in its three dimensions. Since the camera image is two-dimensional an a'priori knowledge about the physical world must be available if it should be possible to carry out the transformation from 2D to 3D. Sometimes this knowledge is most easily expressed in a coordinate system fixed to the body of the AGV while at other times it is more natural to use a coordinate system fixed to the camera.

In the low-level control of the AGV such as the mode where it follows a tape course on the floor, the absolute position and orientation (*pose*) of the vehicle is of little concern. What is needed in this case is a measure of what the current pose is relative to the tape course. In order to relate the images to the physical world it is utilized that the tape is laying on the floor. On the other hand, when the guide marks are used for determining position and orientation of the vehicle it is assumed that the optical center (i.e., the focal point) of the camera has the same height over the floor as the center of the guide mark. For this task it is quite convenient to first use the camera coordinate system to determine the pose relative to the guide mark and subsequently utilize the knowledge about the position of the guide mark in a world-fixed coordinate system to estimate the absolute pose of the vehicle.

To describe the physical world consider first a coordinate system fixed to the vehicle as shown in Fig. 1. This coordinate system will in the following be referred to as the AGV-coordinate system. A similar coordinate system was used for deriving a dynamic model of

the vehicle [NPR97], except that a z-coordinate perpendicular to the floor has been introduced in this case. Before it is discussed how to relate the two-dimensional screen-coordinate system in which the camera operates to a point in this coordinate system it is useful to consider the opposite transformation. I.e., studying the principle of how an image is acquired.

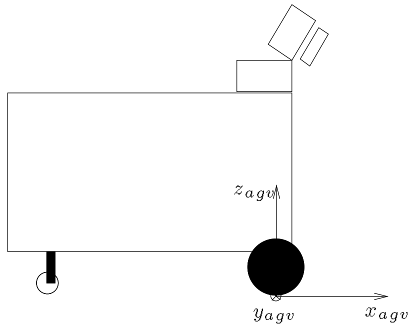

Figure 1. *The AGV-coordinate system. The origin is situated in the vehicle's point of rotation.*

It is convenient to break the transformation from AGV-coordinates to screen coordinates up into a number of intermediate transformations. These transformations depend on the applied hardware and on the type and geometry of the camera system. [Sør94] provides an excellent treatment of the subject and his approach will be used here.

As one of the intermediate stages it is common to introduce a so-called *camera coordinate system* which has its origin in optical center, whose x and y-axis point horizontal and vertical in the image plane and with a z-axis that is perpendicular to this plane. In this coordinate system the *image plane* has its origin in $(0, 0, f)$, where f is the focal length. The transformation from camera coordinates to screen coordinates depends on how the camera is modeled and on the *intrinsic parameters* related to the selected type of camera

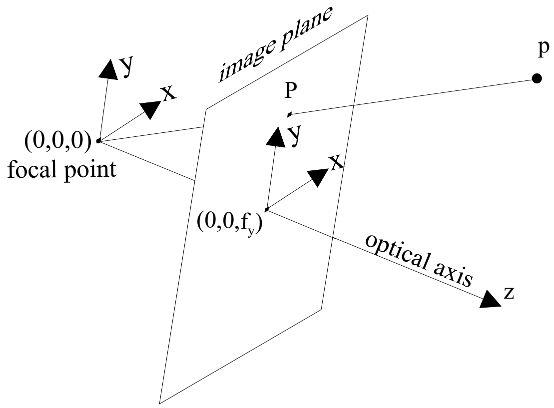

Figure 2. *The camera and the image coordinate systems. The camera coordinate system has its origin in the focal point. In camera coordinates the image coordinate system has its origin in* $(0, 0, f)$.

model. [Sør94] assumes the so-called *pin-hole model* and he furthermore assumes that the lens causes a radial disturbance, only. Before showing the transformations required to get from camera coordinates to screen coordinates the transformation from AGV coordinate system to camera coordinate system will be treated.

4.1.1 From AGV-coordinates to screen coordinates

AGV coordinates to camera coordinates

The camera is mounted on a pan-tilt head which currently can be located in three different positions on the roof of the vehicle. Since the camera may frequently be changed and the mounting also is somewhat ad-hoc, it is convenient to introduce a coordinate system with the same base vectors as the camera coordinate system but with origin in the tilt base instead of in the optical center of the camera. To carry out the transformation from AGV coordinates to these tilt base coordinates a translation of the origin and two rotations are required.

The first rotation is due to the panning of the pan-tilt head and it is thus around the z_{agv}-axis. Figure 3 shows the rotated coordinate system, here specified by the axes x_{pan}, y_{pan} and z_{pan}. The tilting of

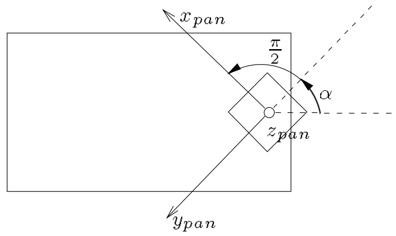

Figure 3. *Coordinate system with origin in the pan base. α is the pan angle.*

the camera corresponds to a rotation around the x_{pan}-axis. This is shown in Figure 4. Notice that a rotation is defined to be positive if

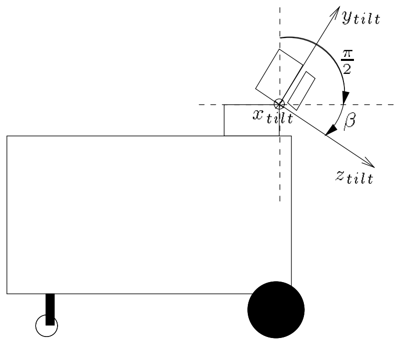

Figure 4. *Coordinate system with origin in the tilt base. β is the tilt angle.*

it is anti-clockwise when standing at infinity on the axis of rotation and looking towards the origin of the coordinate system.

Let the following geometric quantities be given:

α : The pan angle (in radians).

β : The tilt angle (in radians).

Δx_t : The location of the tilt (=pan) axle on the x_{agv}-axis. $\Delta x_t > 0$ if the pan-tilt head is mounted in front of the y_{agv}-axis.

Δy_t : The location of the tilt (=pan) axle on the y_{agv}-axis. $\Delta y_t > 0$ if the pan-tilt head is mounted in the left side of the vehicle.

Δz_t : The location of the tilt axle on the z_{agv}-axis. Since $z_{agv} = 0$ on the floor this distance will always be positive.

O_t : The tilt base.

O_{agv} : The origin of the AGV coordinate system.

Let i_{agv}, j_{agv}, and k_{agv} be the base vectors of the AGV coordinate system, and let p be a point which in this system has the coordinates $(x_{agv}, y_{agv}, z_{agv})$. That is,

$$\overrightarrow{O_{agv}p} = x_{agv} i_{agv} + y_{agv} j_{agv} + z_{agv} k_{agv} \tag{4}$$

The vector from the origin of the tilt base coordinate system to the point p can be written as the sum

$$\overrightarrow{O_t p} = \overrightarrow{O_t O_{agv}} + \overrightarrow{O_{agv} p} \tag{5}$$

where

$$\overrightarrow{O_t O_{agv}} = \Delta x_t i_{agv} + \Delta y_t j_{agv} + \Delta z_t k_{agv} \tag{6}$$

Determination of the coordinates for p in the tilt base coordinate system is thus a matter of describing the vector $\overrightarrow{O_t p}$ in composants of the base vectors of the camera coordinate system:

$$\overrightarrow{O_t p} = x_t i + y_t j + z_t k \tag{7}$$

Consider first the coordinate system with origin in the tilt base obtained by rotating the (x_{agv}, y_{agv}) plane around the z_{agv}-axis as

shown in Figure 3. Described in this coordinate system the base vectors $(i_{agv}, j_{agv}, k_{agv})$ are given by

$$\begin{bmatrix} i_{agv} & j_{agv} & k_{agv} \end{bmatrix} = R_{z,\tilde{\alpha}} = \begin{bmatrix} c_{\tilde{\alpha}} & s_{\tilde{\alpha}} & 0 \\ -s_{\tilde{\alpha}} & c_{\tilde{\alpha}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

The shorthand notation $\tilde{\alpha} = \alpha + \frac{\pi}{2}$, $c_{\tilde{\alpha}} = \cos\tilde{\alpha}$, and $s_{\tilde{\alpha}} = \sin\tilde{\alpha}$ has been introduced for simplification. The tilt base coordinate system is now obtained by rotating this coordinate system the angle $\tilde{\beta} = \beta + \frac{\pi}{2}$ around the x_{pan}-axis. In this coordinate system the base vectors $(i_{agv}, j_{agv}, k_{agv})$ are given by

$$\begin{bmatrix} i_{agv} & j_{agv} & k_{agv} \end{bmatrix} = R_{x,\tilde{\beta}} R_{z,\tilde{\alpha}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{\tilde{\beta}} & s_{\tilde{\beta}} \\ 0 & -s_{\tilde{\beta}} & c_{\tilde{\beta}} \end{bmatrix} R_{z,\tilde{\alpha}} \tag{9}$$

$$= \begin{bmatrix} c_{\tilde{\alpha}} & s_{\tilde{\alpha}} & 0 \\ -s_{\tilde{\alpha}} c_{\tilde{\beta}} & c_{\tilde{\alpha}} c_{\tilde{\beta}} & s_{\tilde{\beta}} \\ s_{\tilde{\alpha}} s_{\tilde{\beta}} & -c_{\tilde{\alpha}} s_{\tilde{\beta}} & c_{\tilde{\beta}} \end{bmatrix} \tag{10}$$

The tilt base coordinates (x_t, y_t, z_t) for the point p are thus determined by

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = R_{x,\tilde{\beta}} R_{z,\tilde{\alpha}} \left(\begin{bmatrix} x_{agv} \\ y_{agv} \\ z_{agv} \end{bmatrix} - \begin{bmatrix} \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{bmatrix} \right) \tag{11}$$

Finally, a translation of the origin from the tilt base to the optical center is made. This translation depends on where the camera has its optical center and on how the camera is mounted to the pan-tilt unit. The x-coordinate of the origin is unaffected if the camera is mounted symmetrically over the tilt base. Define the following distances, which must measured with a ruler unless the shape of the camera is well-known:

Δx : The x-coordinate of the tilt base. 0 if the camera is mounted symmetrically over the tilt base, which is usually the case.

Δy : The y-coordinate of the tilt base.

Δz : The z-coordinate of the tilt base.

The complete coordinate transform can now be written

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta y \\ \Delta z \end{bmatrix} + R \left(\begin{bmatrix} x_{agv} \\ y_{agv} \\ z_{agv} \end{bmatrix} - \begin{bmatrix} \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{bmatrix} \right) \tag{12}$$

where

$$R = R_{x,\tilde{\beta}} R_{z,\tilde{\alpha}} = \begin{bmatrix} c_{\tilde{\alpha}} & s_{\tilde{\alpha}} & 0 \\ -s_{\tilde{\alpha}} c_{\tilde{\beta}} & c_{\tilde{\alpha}} c_{\tilde{\beta}} & s_{\tilde{\beta}} \\ s_{\tilde{\alpha}} s_{\tilde{\beta}} & -c_{\tilde{\alpha}} s_{\tilde{\beta}} & c_{\tilde{\beta}} \end{bmatrix} \tag{13}$$

Camera coordinates to undistorted image coordinates

The projection of p onto the 2D image plane results in the (undistorted) image coordinates:

$$\frac{X_u}{f} = \frac{x}{z} \quad \Leftrightarrow \quad X_u = f\frac{x}{z} \tag{14}$$

$$\frac{Y_u}{f} = \frac{y}{z} \quad \Leftrightarrow \quad Y_u = f\frac{y}{z} \tag{15}$$

It is often found that it is convenient to use the pixels height (distance between pixels in the Y-direction) as the unit on the axes in the image plane. In the vidicon tube the height of a pixel is the distance between two scan lines while on the CCD-chip it is the distance between two neighboring CCD elements. Let f_y denote the focal length measured in pixels

$$f_y = \frac{f}{d_y} \tag{16}$$

and introduce instead

$$\tilde{X}_u = f_y \frac{x}{z} \tag{17}$$

$$\tilde{Y}_u = f_y \frac{y}{z} \tag{18}$$

Undistorted image coordinates to distorted image coordinates

The deviation from the pin-hole model is referred to as the *distortion*. The distortion can be modeled with varying degrees of sophistication. [Sør94] recommends using a relatively simple distortion model and shows that this model yields good results in practice. The first approximation he makes is to consider so-called radial distortion only. The principle is shown in Figure 5. Let r_d denote the distance from origin to distorted point and r the distance from origin to the ideal position according to the pin-hole model. D_r denotes the distortion.

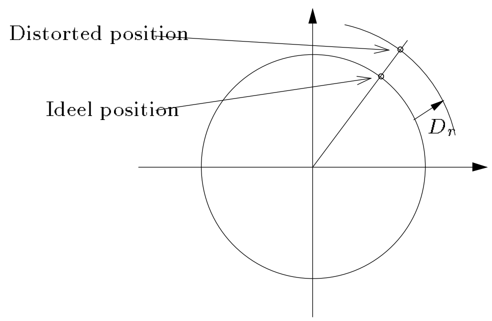

Figure 5. *The radial distortion represents the camera's deviation from the pin-hole model.*

$$r = r_d + D_r \tag{19}$$

The distortion is approximated by the following expression ([Sør94])

$$D_r = \kappa_1 r_d^3 \tag{20}$$

which results in the following *undistorted* image coordinates

$$X_u = X_d\left(1+\frac{D_r}{r_d}\right) = X_d\left(1+\kappa_1 r_d^2\right) \tag{21}$$

$$Y_u = Y_d\left(1+\frac{D_r}{r_d}\right) = Y_d\left(1+\kappa_1 r_d^2\right) \tag{22}$$

where

$$r = \sqrt{X_u^2+Y_u^2} \tag{23}$$

$$r_d = \sqrt{X_d^2+Y_d^2} \tag{24}$$

Or, when using pixel units

$$\tilde{X}_u = \tilde{X}_d\left(1+\frac{\tilde{D}_r}{\tilde{r_d}}\right) = \tilde{X}_d\left(1+\tilde{\kappa}_1 \tilde{r_d}^2\right) \tag{25}$$

$$\tilde{Y}_u = \tilde{Y}_d\left(1+\frac{\tilde{D}_r}{\tilde{r_d}}\right) = \tilde{Y}_d\left(1+\tilde{\kappa}_1 \tilde{r_d}^2\right) \tag{26}$$

where

$$\tilde{r} = \sqrt{\tilde{X}_u^2+\tilde{Y}_u^2} \tag{27}$$

$$\tilde{r_d} = \sqrt{\tilde{X}_d^2+\tilde{Y}_d^2} \tag{28}$$

Unfortunately it is difficult to determine the distorted coordinates from the shown relations since r_d depends on the distorted coordinates and thus is unavailable. To determine the distorted coordinates it would be necessary to solve a third order equation and pick the right solution, which can be quite complicated. [LL95] proposes two approaches to the problem. Either one can use a few Newton-Raphson iterations to find the coordinates numerically or one can use a simple approximation. The approximation is based on the observation that r and r_d usually will deviate by 1% or less. Thus it is reasonable to assume that

$$1+\tilde{\kappa_1} r^2 \approx 1+\tilde{\kappa_1} r_d^2 \tag{29}$$

since $\tilde{\kappa}$ is in the range $10^{-8}-10^{-6}$ for the cameras normally used with IAU's autonomous guided vehicle. Using this approximation the following relations can be used to determine the distorted coordinates

$$\tilde{X}_d \approx \frac{\tilde{X}_d}{(1+\tilde{\kappa}_1\tilde{r}^2)} \tag{30}$$

$$\tilde{Y}_u \approx \frac{\tilde{Y}_d}{(1+\tilde{\kappa}_1\tilde{r}^2)} \tag{31}$$

As shown in [LL95] the error that one makes with this approximation will be less than 0.2 pixels for a 512×512 camera with $\tilde{\kappa}_1 \approx 10^{-7}$. The interested reader is referred to [LL95] for details on how to use the Newton-Raphson method to obtain a better accuracy.

It should be noted that the radial distortion sometimes is modeled differently. The model discussed in this section is used in order to facilitate the transformation from screen to physical world. Other approaches are occasionally used to alleviate the opposite transformation. See [Sør94].

Distorted image coordinates to screen coordinates

By convention the screen coordinate system differs from the image coordinate system in that the axes are reversed and the origin is in the upper left corner of the screen. The relation between the distorted image coordinates and the screen coordinates is given by

$$X_f = -\frac{X_d}{s_x d_x} + C_x \tag{32}$$

$$Y_f = -\frac{Y_d}{d_y} + C_y \tag{33}$$

The aspect ratio is usually introduced as the ratio between the magnification in the Y and X directions

$$a_s = \frac{s_x d_x}{d_y} \tag{34}$$

In which case one can write

$$X_f = -\frac{\tilde{X}_d}{a_s} + C_x \tag{35}$$

$$Y_f = -\tilde{Y}_d + C_y \tag{36}$$

4.1.2 From screen coordinates to AGV-coordinates

The opposite transformation is obtained by inverting each of the steps outlined above. As explained previously the step from 2D to 3D is particularly difficult and requires a certain knowledge about the physical world.

Screen coordinates to distorted image coordinates

The transformation from screen coordinates to distorted image coordinates is straightforward

$$\tilde{X}_d = a_s(C_x - X_f) \tag{37}$$

$$\tilde{Y}_d = C_y - Y_f \tag{38}$$

Distorted image coordinates to undistorted image coordinates

The transformation from distorted to undistorted image coordinates was shown previously (25), (26)

$$\tilde{X}_u = \tilde{X}_d\left(1 + \tilde{\kappa}_1 \tilde{r_d}^2\right) \tag{39}$$

$$\tilde{Y}_u = \tilde{Y}_d\left(1 + \tilde{\kappa}_1 \tilde{r_d}^2\right) \tag{40}$$

where

$$\tilde{r_d} = \sqrt{\tilde{X}_d^2 + \tilde{Y}_d^2} \tag{41}$$

Undistorted image coordinates to AGV coordinates

If it is assumed that the z_{agv}-coordinate is known the most convenient way of carrying out the transformation is to first determine the two ratios x/z and y/z and from these determine x_{agv} and y_{agv}.

x/z and y/z are easily determined from the undistorted image coordinates

$$\frac{x}{z} = \frac{\tilde{X}_u}{f_y} \tag{42}$$

$$\frac{y}{z} = \frac{\tilde{Y}_u}{f_y} \tag{43}$$

Determination of x_{agv} and y_{agv} is slightly more involved. Recall that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta y \\ \Delta z \end{bmatrix} + R\left(\begin{bmatrix} x_{agv} \\ y_{agv} \\ z_{agv} \end{bmatrix} - \begin{bmatrix} \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{bmatrix}\right) \tag{44}$$

Let

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{45}$$

and let $z_{agv} = z'_{agv}$ be the known coordinate in the AGV-coordinate system. Then (44) can be written

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r_{11}x_{agv} + r_{12}y_{agv} \\ r_{21}x_{agv} + r_{22}y_{agv} \\ r_{31}x_{agv} + r_{32}y_{agv} \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} \tag{46}$$

where

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta y \\ \Delta z \end{bmatrix} + \begin{bmatrix} r_{13} \\ r_{23} \\ r_{33} \end{bmatrix} z'_{agv} - R\begin{bmatrix} \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{bmatrix} \tag{47}$$

This system of equations can now be solved for x_{agv} and y_{agv} given that the two ratios x/z and y/z are known

$$\left\{\begin{array}{rcl} \dfrac{x}{z} & = & \dfrac{r_{11}x_{agv} + r_{12}y_{agv} + \gamma_1}{r_{31}x_{agv} + r_{32}y_{agv} + \gamma_3} \\[2ex] \dfrac{y}{z} & = & \dfrac{r_{21}x_{agv} + r_{22}y_{agv} + \gamma_2}{r_{31}x_{agv} + r_{32}y_{agv} + \gamma_3} \end{array}\right. \tag{48}$$

$\Updownarrow$

$$\left\{\begin{array}{rcl} \left(r_{11} - r_{31}\frac{x}{z}\right) x_{agv} + \left(r_{12} - r_{32}\frac{x}{z}\right) y_{agv} & = & \frac{x}{z}\gamma_3 - \gamma_1 \\[2ex] \left(r_{21} - r_{31}\frac{y}{z}\right) x_{agv} + \left(r_{22} - r_{32}\frac{y}{z}\right) y_{agv} & = & \frac{y}{z}\gamma_3 - \gamma_2 \end{array}\right. \tag{49}$$

$\Updownarrow$

$$\boxed{\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = \begin{bmatrix} r_{11} - r_{31}\frac{x}{z} & r_{12} - r_{32}\frac{x}{z} \\ r_{21} - r_{31}\frac{y}{z} & r_{22} - r_{32}\frac{y}{z} \end{bmatrix}^{-1} \begin{bmatrix} \frac{x}{z}\gamma_3 - \gamma_1 \\ \frac{y}{z}\gamma_3 - \gamma_2 \end{bmatrix}} \tag{50}$$

By carrying out the inversion and the multiplication and inserting the undistorted image coordinates, the following is obtained

$$\begin{bmatrix} x_{agv} \\ \\ y_{agv} \end{bmatrix} =$$
$$\frac{1}{n}\begin{bmatrix} f_y(\gamma_1 r_{22} - \gamma_2 r_{12}) + (\gamma_2 r_{32} - \gamma_3 r_{22})\tilde{X}_u + (\gamma_3 r_{12} - \gamma_1 r_{32})\tilde{Y}_u \\ \\ f_y(\gamma_2 r_{11} - \gamma_1 r_{21}) + (\gamma_3 r_{21} - \gamma_2 r_{31})\tilde{X}_u + (\gamma_1 r_{31} - \gamma_3 r_{11})\tilde{Y}_u \end{bmatrix}$$

where

$$n = f_y(r_{21}r_{12} - r_{11}r_{22}) + (r_{22}r_{31} - r_{21}r_{32})\tilde{X}_u + (r_{11}r_{32} - r_{12}r_{31})\tilde{Y}_u \tag{51}$$

If the matrix elements $\{r_{ij}\}$ are substituted for their actual trigonometric functions, n can be simplified to

$$n = -f_y \cos\tilde{\beta} - \tilde{Y}_u \sin\tilde{\beta} \tag{52}$$

4.1.3 From screen coordinates to AGV coordinates - a simplified procedure

If the distortion is neglected it is straightforward to rewrite the complete transformation from screen coordinates to AGV coordinates as a single expression

$$\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = \frac{1}{-f_y \cos\tilde{\beta} - C_y \sin\tilde{\beta} + Y_f \sin\tilde{\beta}} \begin{bmatrix} a_1 + b_1 X_f + c_1 Y_f \\ a_2 + b_2 X_f + c_2 Y_f \end{bmatrix} \tag{53}$$

$$\begin{array}{rcl} a_1 & = & f_y(\gamma_1 r_{22} - \gamma_2 r_{12}) + a_s C_x(\gamma_2 r_{32} - \gamma_3 r_{22}) + C_y(\gamma_3 r_{12} - \gamma_1 r_{32}) \\ & = & f_y(\gamma_1 r_{22} - \gamma_2 r_{12}) - C_x b_1 - C_y c_1 \\ b_1 & = & -a_s(\gamma_2 r_{32} - \gamma_3 r_{22}) \\ c_1 & = & -(\gamma_3 r_{12} - \gamma_1 r_{32}) \\ a_2 & = & f_y(\gamma_2 r_{11} - \gamma_1 r_{21}) + a_s C_x(\gamma_3 r_{21} - \gamma_2 r_{31}) + C_y(\gamma_1 r_{31} - \gamma_3 r_{11}) \\ & = & f_y(\gamma_2 r_{11} - \gamma_1 r_{21} - C_x b_2 - C_y c_2 \\ b_2 & = & -a_s(\gamma_3 r_{21} - \gamma_2 r_{31}) \\ c_2 & = & -(\gamma_1 r_{31} - \gamma_3 r_{11}) \end{array}$$

γ_1, γ_1 and γ_1 are found from (47).

4.2 Line following

When the distortion is neglected a straight line on the floor will obviously correspond to a straight line on the screen. In order to determine how the parameter form describing a straight line on the screen transforms to the AGV-coordinate system introduce the following expression for a straight line

$$\begin{bmatrix} X_f \\ Y_f \end{bmatrix} = \begin{bmatrix} X_f^0 \\ f_y \cot\tilde{\beta} + C_y \end{bmatrix} + t \begin{bmatrix} v_1 \\ \frac{1}{\sin\tilde{\beta}} \end{bmatrix} \tag{54}$$

This somewhat peculiar parameter form has been selected to obtain a simpler expression for the line in AGV-coordinates

$$\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = \frac{1}{t}\begin{bmatrix} a_1 + b_1 X_f^0 + c_1(f_y \cot\tilde{\beta} + C_y) + t\left(b_1 v_1 + \frac{c_1}{sin\beta}\right) \\ a_2 + b_2 X_f^0 + c_2(f_y \cot\tilde{\beta} + C_y) + t\left(b_2 v_1 + \frac{c_2}{sin\beta}\right) \end{bmatrix}$$
$$= \begin{bmatrix} b_1 v_1 + \frac{c_1}{sin\beta} \\ b_2 v_1 + \frac{c_2}{sin\beta} \end{bmatrix} + \frac{1}{t}\begin{bmatrix} a_1 + b_1 X_f^0 + c_1(f_y \cot\tilde{\beta} + C_y) \\ a_2 + b_2 X_f^0 + c_2(f_y \cot\tilde{\beta} + C_y) \end{bmatrix}$$

When $\tilde{\beta} = p\pi$, $p = 0, \pm 1, \pm 2, ...$, corresponding to the situation where the optical axis lies in the (x_{agv}, y_{agv}) plane, a different approach must be taken. Let

$$\begin{bmatrix} X_f \\ Y_f \end{bmatrix} = \begin{bmatrix} X_f^0 \\ Y_f^0 \end{bmatrix} + t\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{55}$$

then

$$\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = \pm\frac{1}{f_y}\begin{bmatrix} a_1 + b_1 X_f^0 + c_1 Y_f^0 \\ a_2 + b_2 X_f^0 + c_2 Y_f^0 \end{bmatrix} \pm \frac{t}{f_y}\begin{bmatrix} b_1 v_1 + c_1 v_2 \\ b_2 v_1 + c_2 v_2 \end{bmatrix} \tag{56}$$

where $-$ is used when $p = 0, \pm 2, \pm 4, ...$ and $+$ is used when $p = 1, \pm 3, \pm 5, ...$

The expression in (55) is not applicable when horizontal lines on the screen are considered. In this case it is more useful to apply:

$$\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = \frac{1}{n^0}\begin{bmatrix} a_1 + c_1 Y_f^0 \\ a_2 + c_2 Y_f^0 \end{bmatrix} + \frac{1}{n^0}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} X_f^0 \tag{57}$$

where

$$n^0 = -f_y \cos\tilde{\beta} - C_y \sin\tilde{\beta} + Y_f \sin\tilde{\beta} \tag{58}$$

It is in this case necessary to impose the constraint

$$n^0 \neq 0 \qquad \Leftrightarrow \qquad Y_f^0 \neq f_y \cot\tilde{\beta} + C_y \tag{59}$$

4.2.1 Determine the image of the nearest line segment

In order to build a simulator for the line tracking it is necessary to determine the intersection between a chosen line in the digitized image and the course on the floor.

Assume that the course is constituted strictly by segments of straight lines and that they are represented in a database by a starting-point $(x^i_{agv,0}, y^i_{agv,0})$ and and an end-point $(x^i_{agv,1}, y^i_{agv,1})$, such that the equation for line segment i can be written in parameter form as

$$\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = \begin{bmatrix} x^i_{agv,0} \\ y^i_{agv,0} \end{bmatrix} + t\begin{bmatrix} x^i_{agv,1} - x^i_{agv,0} \\ y^i_{agv,1} - y^i_{agv,0} \end{bmatrix}, \qquad t \in [0,1] \tag{60}$$

the intersection (t, X_f) between line segment i and the line corresponding to the horizontal line, $Y_f = Y_f^0$ on the screen. This is found as

$$\begin{bmatrix} x^i_{agv,0} \\ y^i_{agv,0} \end{bmatrix} + t\begin{bmatrix} x^i_{agv,1} - x^i_{agv,0} \\ y^i_{agv,1} - y^i_{agv,0} \end{bmatrix} = \frac{1}{n^0}\begin{bmatrix} a_1 + c_1 Y_f^0 \\ a_2 + c_2 Y_f^0 \end{bmatrix} + \frac{1}{n^0}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} X_f \tag{61}$$

If the two lines are not parallel, in which case they do not intersect, this has the solution

$$\begin{bmatrix} t \\ X_f \end{bmatrix} = \begin{bmatrix} x^i_{agv,1} - x^i_{agv,0} & -\frac{b_1}{n^0} \\ y^i_{agv,1} - y^i_{agv,0} & -\frac{b_2}{n^0} \end{bmatrix}^{-1} \left(\frac{1}{n^0}\begin{bmatrix} a_1 + c_1 Y_f^0 \\ a_2 + c_2 Y_f^0 \end{bmatrix} - \begin{bmatrix} x^i_{agv,0} \\ y^i_{agv,0} \end{bmatrix}\right)$$

4.3 Pose Estimation Using Guide Marks

When used with autonomous vehicles a *guide mark* should be a pattern that is relatively easy to detect, which can easily be distinguished from all other objects in the environment and which can provide an accurate estimate of the position and orientation of the vehicle (the so-called *pose*). In the project of [oH93] a guide mark with these properties has been selected and image processing algorithms for detecting it have been developed. The guide mark is shown in Fig. 6a.

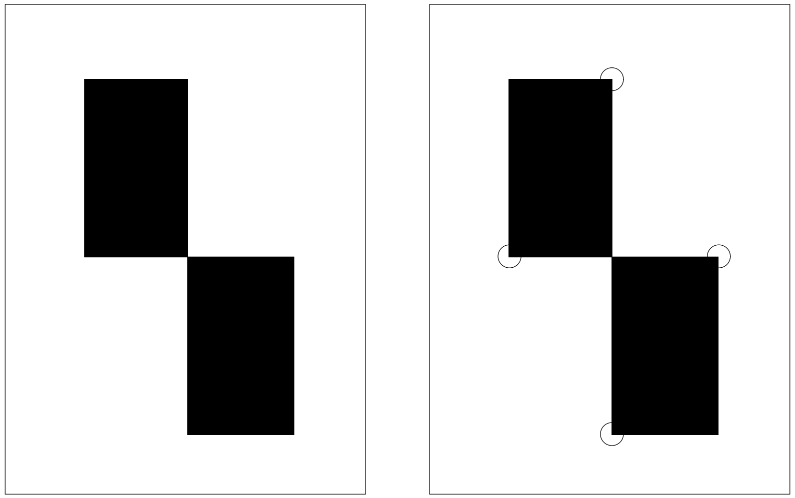

Figure 6. *a: The guide mark. b: The four detected points*

The objective of the developed image processing is to detect the 4 points marked with circles in Fig. 6b. When these have been found the pose can be determined through a series of geometrical considerations. While the image processing procedure developed in [oH93] appears to be working quite efficiently and robust, the expressions they propose for determining the pose from the four detected points are not sufficiently accurate unless the vehicle is far away from the guide mark. Also they have not taken into account that the camera can be panned and tilted arbitrarily. For this reason more generally applicable estimates of the pose will be derived below. All expressions are exact and only based on the assumption that the guide mark center has the same height over the floor as the focal point of the camera. Naturally, this will not be satisfied in practice when it is possible to tilt the camera arbitrarily. However, the small discrepancy in height that might occur will typically lead to minor errors only and thus it is reasonable to neglect it.

The approach taken in order to determine the pose is similar in spirit

to the one described in [oH93]. This is not based on the conventional coordinate transform technique most often used for relating camera images to the physical world, but instead on "ray optical" geometry.

As an intermediate state in determining the pose it is convenient to determine the 3 quantities shown in Fig. 7.

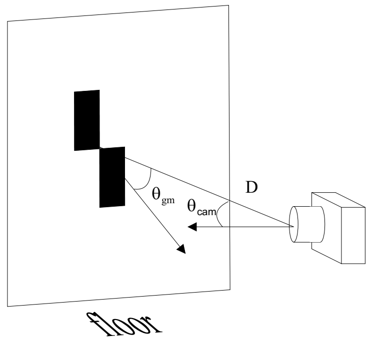

Figure 7. *D is the distance between focal point and guide mark center, θ_{gm} is the angle between the normal to the guide mark and the line between the focal point and the guide mark center (the line-of sight) and θ_{agv} is the angle between the optical axis and line-of-sight.*

Unless otherwise noted, all operations carried out in the following will take place in the camera coordinate system.

4.3.1 Distance from focal point to guide mark center (D)

Since the focal point is assumed to have the same height over the floor as the guide mark center, the distance D can obviously be determined by considering the right triangle shown in Fig. 8

$$D = \frac{H_{mrk}}{2\tan(\frac{\psi}{2})} \tag{62}$$

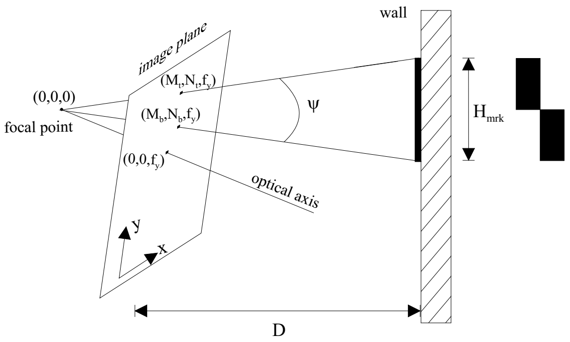

Figure 8. *Guide mark and vehicle viewed from the side.* (M_t, N_t, f_y) *is the image of the upper most point of four detected points in the guide mark and* (M_b, N_b, f_y) *is the image of the bottom point. The points showed in the drawing are specified in pixel units in the camera coordinate system*

As it appears from the drawing, the angle ψ can also be found as the angle between the two vectors

$$\mathbf{v_t} = \begin{bmatrix} M_t \\ N_t \\ f_y \end{bmatrix} \qquad \mathbf{v_b} = \begin{bmatrix} M_b \\ N_b \\ f_y \end{bmatrix} \tag{63}$$

and it is thus it is given by

$$\cos(\psi) = \frac{\mathbf{v_t} \cdot \mathbf{v_b}}{|\mathbf{v_t}||\mathbf{v_b}|} \tag{64}$$

4.3.2 Angle between normal to guide mark and line-of-sight (θ_{gm})

The magnitude of the angle θ_{gm} is determined by

$$\cos(\theta_{gm}) = \frac{D'}{D} \tag{65}$$

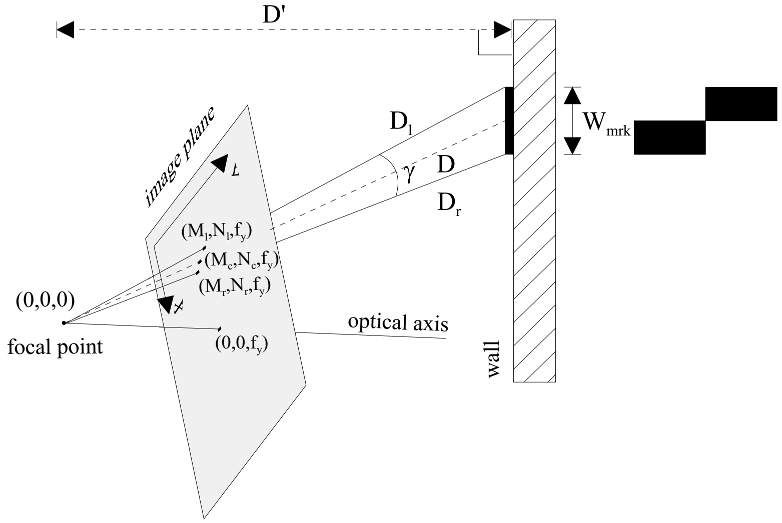

Figure 9. Guide mark and vehicle viewed from above. (M_l, N_l, f_y) is the left point in the guide mark, (M_r, N_r, f_y) is the right point and (M_c, N_c, f_y) is the guide mark center. D' is altitude of the triangle connecting the focal point and left and right point.

where D is known from (62). Unfortunately this expression cannot provide the sign of the angle since it cannot distinguish between whether the camera is to the left or to the right of the guide mark. This issue will be dealt with in a moment. First the distance D' will be determined. Consider first the triangle obtained by connecting left and right side of the guide mark and the point f. Combining the following two formulas for the area of the triangle

$$A = \frac{1}{2}W_{mrk}D' \tag{66}$$

$$A = \frac{1}{2}D_r D_l \sin(\gamma) \tag{67}$$

leads to

$$D_r D_l = \frac{D'W_{mrk}}{\sin(\gamma)} \tag{68}$$

This is result is now inserted in the following expression, which has been obtained by applying the law of cosines

$$W_{mrk}^2 = D_r^2 + D_l^2 - 2D_r D_l \cos(\gamma) \tag{69}$$

giving

$$D_r^2 + D_l^2 - W_{mrk}^2 = 2\frac{D'W_{mrk}}{\tan(\gamma)} \tag{70}$$

To remove the two unknown distances D_r and D_l from the expression it is necessary to utilize the information that the line from the point f to the center of the guide mark has the length D. The following is a general formula used for determining the median of a triangle

$$D = \frac{1}{2}\sqrt{2D_r^2 + 2D_l^2 - W_{mrk}^2} \tag{71}$$

Squaring left and right side of this formula and reordering the terms gives

$$D_r^2+D_l^2-W_{mrk}^2 = 2D^2-\frac{1}{2}W_{mrk}^2 \quad \Leftrightarrow \quad D' = \left(\frac{D^2}{W_{mrk}} - \frac{1}{4}W_{mrk}\right)\tan(\gamma) \tag{72}$$

The angle γ is determined in the same way as ψ was determined in (64)

$$\cos(\gamma) = \frac{\mathbf{v_r} \cdot \mathbf{v_l}}{|\mathbf{v_r}||\mathbf{v_l}|} \tag{73}$$

where

$$\mathbf{v_r} = \begin{bmatrix} M_r \\ N_r \\ f_y \end{bmatrix} \qquad \mathbf{v_l} = \begin{bmatrix} M_l \\ N_l \\ f_y \end{bmatrix} \tag{74}$$

As was mentioned above the sign of θ_{gm} has to be determined with some care. Let θ_{gm} be defined positive when measuring clockwise from the normal of the guide mark when viewed from above. That is, $\theta_{gm} > 0$ when $D_l < D_r$ and vice versa. Applying the law of sines three times result in the following relations

$$\left\{ \begin{array}{rcl} \frac{\sin\alpha_l}{D_r} & = & \frac{\sin\alpha_r}{D_l} \\ \sin\alpha_l & = & D\frac{\sin\gamma_l}{W_{mrk}/2} \\ \sin\alpha_r & = & D\frac{\sin\gamma_r}{W_{mrk}/2} \end{array} \right. \tag{75}$$

$$\Rightarrow \qquad D_l\frac{2D}{W_{mrk}}\sin\gamma_l = D_r\frac{2D}{W_{mrk}}\sin\gamma_r \qquad \Leftrightarrow \qquad D_l\sin\gamma_l = D_r\sin\gamma_r \tag{76}$$

In order to determine the two angles γ_r and γ_l the law of sines is now applied four times to the triangle shown in Fig. 10a+b.

$$\left\{ \begin{array}{rcl} \frac{\sin\beta_r}{d_{cf}} & = & \frac{\sin\gamma_r}{d_{rc}} \\ \frac{\sin\beta_r}{d_{lf}} & = & \frac{\sin\gamma}{d_{rl}} \end{array} \right. \qquad \Rightarrow \qquad d_{lf}\frac{\sin\gamma}{d_{rl}} = d_{cf}\frac{\sin\gamma_r}{d_{rc}} \tag{77}$$

$$\left\{ \begin{array}{rcl} \frac{\sin\beta_l}{d_{cf}} & = & \frac{\sin\gamma_l}{d_{lc}} \\ \frac{\sin\beta_l}{d_{rf}} & = & \frac{\sin\gamma}{d_{rl}} \end{array} \right. \qquad \Rightarrow \qquad d_{rf}\frac{\sin\gamma}{d_{rl}} = d_{cf}\frac{\sin\gamma_l}{d_{lc}} \tag{78}$$

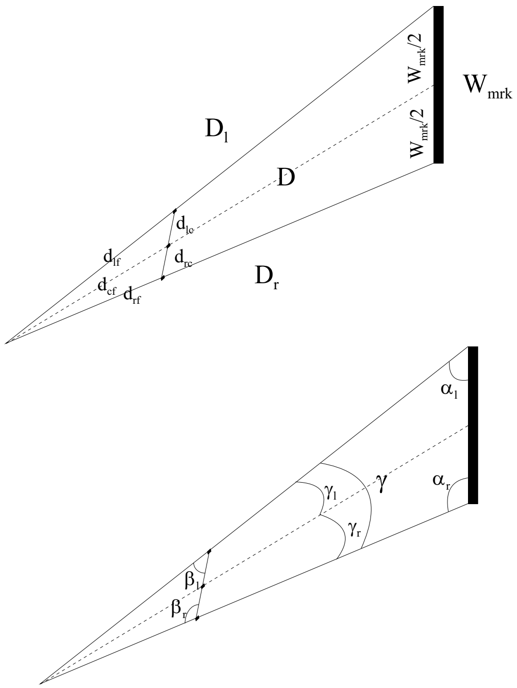

Figure 10. *Determining the sign of the angle θ_{gm}. a: Distances. b: Angles*

The two results are then combined into

$$\frac{\sin\gamma_r}{d_{rc}d_{lf}} = \frac{\sin\gamma_l}{d_{lc}d_{rf}} \tag{79}$$

which, when compared to (76), gives the following relation

$$D_l d_{lc} d_{rf} = D_r d_{rc} d_{lf} \tag{80}$$

in which case it is obvious that

$$\theta_{gm} > 0 \quad\Leftrightarrow\quad D_l < D_r \quad\Leftrightarrow\quad d_{lc}d_{rf} > d_{rc}d_{lf} \tag{81}$$

$$d_{rf} = |v_r| \tag{82}$$

$$d_{lf} = |v_l| \tag{83}$$

$$d_{lc}^2 = (M_c - M_l)^2 + (N_c - N_l)^2 \tag{84}$$

$$d_{rc}^2 = (M_c - M_r)^2 + (N_c - N_r)^2 \tag{85}$$

The image of the center point (M_c, N_c) is not available directly, but is easily determined as the intersection of the line going through the images of the top and bottom points and the line going through the images of the left and right points.

$$\begin{bmatrix} M_t \\ N_t \end{bmatrix} + t\begin{bmatrix} M_b - M_t \\ N_b - N_t \end{bmatrix} = \begin{bmatrix} M_l \\ N_l \end{bmatrix} + s\begin{bmatrix} M_r - M_l \\ N_r - N_l \end{bmatrix} \tag{86}$$

$\Updownarrow$

$$\begin{bmatrix} M_b - M_t & M_r - M_l \\ N_b - N_t & N_r - N_l \end{bmatrix}\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} M_l - M_t \\ N_b - N_l \end{bmatrix} \tag{87}$$

If this is solved for t, the image of the center can be determined by

$$(M_c, N_c) = (M_l + t(M_b - M_t), N_l + t(N_b - N_t)) \tag{88}$$

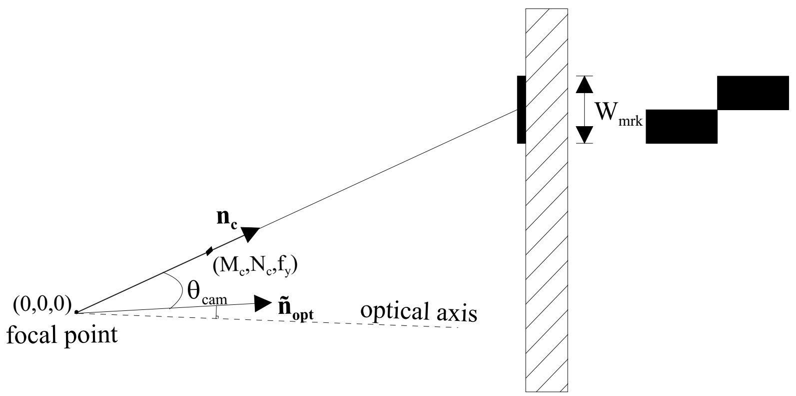

Figure 11. *θ_{cam} is the angle between line-of-sight to the guide mark center and the horizontal projection of optical axis.*

4.3.3 Angle between line-of-sight to guide mark center and horizontal projection of the optical axis (θ_{cam})

Let $\mathbf{n_c}$ denote the direction vector for line-of-sight to the guide mark center and let $\tilde{\mathbf{n}}_{\mathbf{opt}}$ denote the direction vector for the horizontal projection of the optical axis. With these definitions the magnitude of the angle θ_{cam} can be determined from the relation

$$\cos\theta_{cam} = \tilde{\mathbf{n}}_{\mathbf{opt}} \cdot \mathbf{n_c} \tag{89}$$

The direction vector for line-of-sight to the guide mark center is given by

$$\mathbf{n_c} = \frac{\begin{bmatrix} M_c \\ N_c \\ f_y \end{bmatrix}}{\sqrt{M_c^2 + N_c^2 + f_y^2}} \tag{90}$$

The normal to the image plane is

$$\mathbf{n_{opt}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{91}$$

When the camera is not tilted this vector is obviously horizontal. In general the direction vector for the horizontal projection of the optical axis ($\tilde{\mathbf{n}}_{\mathbf{opt}}$) is obtained by rotating the normal to the image plane with the opposite tilt angle around the x-axis. That is,

$$\tilde{\mathbf{n}}_{\mathbf{opt}} = \begin{bmatrix} 0 \\ \sin\beta \\ \cos\beta \end{bmatrix} \tag{92}$$

If θ_{cam} is defined such that it is positive when the optical axis is to the right of the guide mark center when viewed from above, θ_{cam} is determined as follows:

$$\theta_{cam} = \begin{cases} \arccos(\tilde{\mathbf{n}}_{\mathbf{opt}} \cdot \mathbf{n_c}) & , \quad M_c > 0 \\ -\arccos(\tilde{\mathbf{n}}_{\mathbf{opt}} \cdot \mathbf{n_c}) & , \quad M_c < 0 \end{cases} \tag{93}$$

4.3.4 Estimating the pose

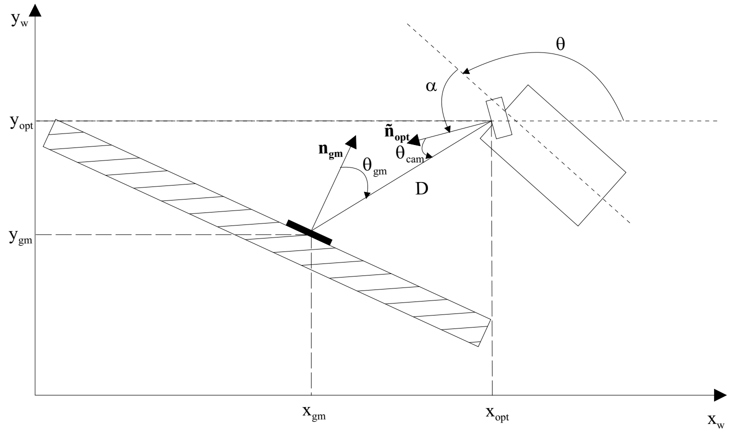

Figure 12. *Determination of the vehicle's world coordinates when D, θ_{gm} and θ_{agv} are known.*

Let $\mathbf{n_{gm}}$ denote the normal to the guide mark. The (x,y)-coordinates for the optical center in the world coordinate system (i.e., the projection of focal point on the floor) is given by (see Fig. 12)

$$\begin{bmatrix} x_{opt} \\ y_{opt} \end{bmatrix} = D \begin{bmatrix} \cos(\theta_{gm}) & \sin(\theta_{gm}) \\ -\sin(\theta_{gm}) & \cos(\theta_{gm}) \end{bmatrix} \mathbf{n_{gm}} + \begin{bmatrix} x_{gm} \\ y_{gm} \end{bmatrix} \tag{94}$$

where (x_{gm}, y_{gm}) the guide mark center specified in world coordinates. When this is available the orientation of the vehicle can be determined by

$$\tan(\hat{\theta} + \theta_{cam} + \alpha) = \frac{y_{gm} - y_{opt}}{x_{gm} - x_{opt}} \tag{95}$$

The result of the function $arctan(x/y)$ is limited to the interval $[-\pi/2; \pi/2]$, which means that the function will produce valid results only when $x \geq 0$. To overcome this problem the so-called "four-quadrant inverse tangent function", $arctan2$, will be used used. $arctan2$ is defined by

$$\arctan 2(x, y) = \begin{cases} \arctan(y, x) & , x \geq 0 \\ \arctan(y, x) + \pi & , x < 0, y \geq 0 \\ \arctan(y, x) - \pi & , x < 0, y < 0 \end{cases} \tag{96}$$

Using $arctan2$ to determine $\hat{\theta}$ gives

$$\hat{\theta} = \arctan 2 \left(\frac{y_{gm} - y_{opt}}{x_{gm} - x_{opt}} \right) - \theta_{cam} - \alpha \tag{97}$$

The origin of the AGV-coordinate system has the world-coordinates $(x_{w,0}, y_{w,0}, 0)$ which is the point on the floor immediately below the point of rotation. The z-axis is the same in two coordinate systems, but the (x, y) plane is rotated with the angle θ. The transformation from world coordinates to AGV-coordinates thus takes the form

$$\begin{bmatrix} x_{agv} \\ y_{agv} \end{bmatrix} = R_\theta \left(\begin{bmatrix} x_w \\ y_w \end{bmatrix} - \begin{bmatrix} x_{w,0} \\ y_{w,0} \end{bmatrix} \right) \tag{98}$$

where

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{99}$$

In this case estimates of the angle θ and the world coordinates for the point on the floor immediately below the focal point are available (i.e., an estimate of the (x_w, y_w)-coordinates for the point are available). The AGV-coordinates for the focal point is of course easy to determine using (12)

$$\begin{bmatrix} x_{agv,opt} \\ y_{agv,opt} \\ z_{agv,opt} \end{bmatrix} = \begin{bmatrix} \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{bmatrix} - R^T \begin{bmatrix} 0 \\ \Delta y \\ \Delta z \end{bmatrix} \tag{100}$$

An estimate of the point of rotation can thus be found using

$$\begin{bmatrix} \hat{x}_{w,0} \\ \hat{y}_{w,0} \end{bmatrix} = \begin{bmatrix} x_{opt} \\ y_{opt} \end{bmatrix} - \hat{R_\theta}^T \begin{bmatrix} x_{agv,opt} \\ y_{agv,opt} \end{bmatrix} \tag{101}$$

where

$$\hat{R_\theta} = \begin{bmatrix} \cos\hat{\theta} & \sin\hat{\theta} \\ -\sin\hat{\theta} & \cos\hat{\theta} \end{bmatrix} \tag{102}$$

4.3.5 Quantification of the Uncertainty on the Pose Estimate

The accuracy of the pose estimate $(\hat{x}, \hat{y}, \hat{\theta})$ depends on how well several other parameters are known

H_{mrk}, W_{mrk}	Height and width of guide mark.
α, β	Pan and tilt angle.
f_y, a_s, κ_1	Focal point, aspect ratio, distortion
C_x, C_y	Screen coordinate of the optical axis' intersection with the image plane.
$P_{gm}, \mathbf{n_c}$	Guide mark center and normal to guide mark in world coordinates
P_{opt}	Focal point in AGV-coordinates

It is assumed that the uncertainty of these parameters amounts to a measurement error that can be referred to the identified guide mark points and that this error is white and Gaussian distributed

$$P \in N(\hat{P}, Q_{pix}), \qquad P = M_t, M_b, M_l, M_r, N_t, \dots \quad (103)$$

$$Q_{pix} = \begin{bmatrix} a_s\sigma^2_{pix} & 0 \\ 0 & \sigma^2_{pix} \end{bmatrix} \quad (104)$$

It is further assumed that this in turn implies that the pose estimate is white and Gaussian distributed.

In the Master's thesis [Bak97] an attempt was made to derive analytical expressions for the variance of the pose estimate that depends on the pixel variance, σ^2_{pix}, the variance of the pan angle, σ^2_{α}, and the variance of the tilt angle, σ^2_{β}. The present subsection is more or less identical to Appendix B of this thesis.

Variance of D

Recall that D is found from the following expression

$$D = \frac{H_{mrk}}{2\tan(\frac{\psi}{2})} \quad (105)$$

The angle ψ will in general be fairly small ($\approx 0.4rad$ for $D = 1m$) and a reasonable simplification is thus

$$\begin{aligned} D &\approx \frac{H_{mrk}}{\psi} && (106) \\ &= \frac{H_{mrk}}{\arccos\left(\frac{M_tM_b+N_tN_b+f_y^2}{\sqrt{M_t^2+N_t^2+f_y^2}\sqrt{M_b^2+N_b^2+f_y^2}}\right)} && (107) \\ &= f_D(P_t, P_b) && (108) \end{aligned}$$

where

$$P_t = (M_t, N_t) \qquad \text{and} \qquad P_b = (M_b, N_b) \quad (109)$$

The variance of D is now determined by linerizing D around P_t og P_b using a first order Taylor approximation.

$$V\{D\} = [\nabla_{P_t,P_b} f_D] \begin{bmatrix} Q_{pix} & 0 \\ 0 & Q_{pix} \end{bmatrix} [\nabla_{P_t,P_b} f_D]^T \quad (110)$$

D is not differentiable for $M_t = M_b \wedge N_t = N_b$ since $\arccos(x)$ is differentiable only for $x \in]-1, 1[$. This case is, however, an unlikely outcome of the image processing. Under the assumption that the derivative of D exists, the following is obtained

$$\begin{aligned} V\{D\} = {} & \frac{D^4}{H^2_{mrk}((M_t^2+N_t^2+f_y^2)(M_b^2+N_b^2+f_y^2)-(M_tM_b+N_tN_b+f_y^2)^2)} \cdot \\ & \left(\left(M_b - \frac{(M_tM_b+N_tN_b+f_y^2)}{M_t^2+N_t^2+f_y^2}M_t\right)^2 a_s^2 + \left(M_t - \frac{(M_tM_b+N_tN_b+f_y^2)}{M_b^2+N_b^2+f_y^2}M_b\right)^2 a_s^2 \right. \\ & \left. + \left(N_b - \frac{(M_tM_b+N_tN_b+f_y^2)}{M_t^2+N_t^2+f_y^2}N_t\right)^2 + \left(N_t - \frac{(M_tM_b+N_tN_b+f_y^2)}{M_b^2+N_b^2+f_y^2}N_b\right)^2\right)\sigma^2_{pix} \end{aligned} \quad (111)$$

Variance of $|\theta_{gm}|$

Recall that $|\theta_{gm}|$ is given by

$$|\theta_{gm}| = \arccos\left(\left(\frac{D^2}{W_{mrk}} - \frac{W_{mrk}}{4}\right)\frac{\tan(\gamma)}{D}\right) \quad (112)$$

Since γ is small ($\approx 0.2rad$, for $d = 1m$) and $\frac{D^2}{W_{mrk}} \gg \frac{W_{mrk}}{4}$ ($0.9m \gg 0.025m$, for $D = 1m$), the following simplification is reasonable

$$\begin{aligned} |\theta_{gm}| &\approx \arccos\left(\frac{D}{W_{mrk}}\gamma\right) \\ &= \arccos\left(\frac{H_{mrk}}{W_{mrk}}\frac{\gamma}{\psi}\right) \\ &= f_{\theta_{gm}}(D, P_r, P_l) && (113) \end{aligned}$$

where

$$P_l = (M_l, N_l) \qquad \text{and} \qquad P_t = (M_t, N_t) \tag{114}$$

A linearization is now made around P_r, P_l og D, where by

$$\mathrm{V}\{\theta_{gm}\} \;=\; [\nabla_{D,P_r,P_l} f_{\theta_{gm}}] \begin{bmatrix} \mathrm{V}\{D\} & 0 & 0 \\ 0 & Q_{pix} & 0 \\ 0 & 0 & Q_{pix} \end{bmatrix} [\nabla_{P,P_r,P_l} f_{\theta_{gm}}]^T$$

When $M_l = M_r$, $M_t = M_b \wedge N_t = N_b$ (unlikely in practice) or $\psi = \frac{H_{mrk}}{W_{mrk}}\gamma$ then the derivative of $|\theta_{gm}|$ does not exist. Under the assumption that the derivative does exists, the following expression can be obtained

$$\begin{aligned} \mathrm{V}\{\theta_{gm}\} \;=\; & \\ & \frac{\gamma^2}{W_{mrk}^2 - D^2\gamma^2}\mathrm{V}\{D\} + \\ & \frac{D^2}{((M_r^2+N_r^2+f_y^2)(M_l^2+N_l^2+f_y^2)-(M_rM_l+N_rN_l+f_y^2))(W_{mrk}^2-D^2\gamma^2)} \cdot \\ & \left(\left(M_r - \frac{M_rM_l+N_rN_l+f_y^2}{M_l^2+N_l^2+f_y^2}M_l\right)^2 a_s^2 + \left(M_l - \frac{M_rM_l+N_rN_l+f_y^2}{M_r^2+N_r^2+f_y^2}M_r\right)^2 a_s^2 + \right. \\ & \left. \left(N_r - \frac{M_rM_l+N_rN_l+f_y^2}{M_l^2+N_l^2+f_y^2}N_l\right)^2 + \left(N_l - \frac{M_rM_l+N_rN_l+f_y^2}{M_r^2+N_r^2+f_y^2}N_r\right)^2 + \right) \sigma_{pix}^2 \end{aligned} \tag{115}$$

Variance of $|\theta_{cam}|$

Recall that $|\theta_{cam}|$ is given by:

$$\begin{aligned} |\theta_{cam}| &= \arccos\left(\frac{N_c\sin(\beta) + f_y\cos(\beta)}{\sqrt{M_c^2 + N_c^2 + f_y^2}}\right) \\ &= f_{\theta_{cam}}(P_c) \end{aligned} \tag{116}$$

The guide mark center, $P_c = (M_c, N_c)^T$, is determined by interpolation of the image coordinates of the four guide mark points and are assumed to obey the same distribution. Let the variance of the tilt

angle, β, be denoted, σ_β^2. An expression for the variance of θ_{cam} can then be obtained after a linearization of (116) around P_c and β

$$\mathrm{V}\{\theta_{cam}\} \;=\; [\nabla_{P_c,\beta} f_{\theta_{cam}}] \begin{bmatrix} Q_{pix} & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix} [\nabla_{P_c,\beta} f_{\theta_{cam}}]^T$$

In this case there are problems with differentiability when $M_c = 0 \wedge N_c = 0 \wedge \beta = 0$, that is, when the image if the guide mark center is located in the center of the image plane and the camera tilt angle, β, is 0. Otherwise

$$\begin{aligned} \mathrm{V}\{\theta_{cam}\} \;=\; & \frac{(N_c\sin\beta + f_y\cos\beta)^2 M_c^2}{(M_c^2+N_c^2+f_y^2)^3\left(1 - \frac{(N_c\sin\beta+f_y\cos\beta)^2}{(M_c^2+N_c^2+f_y^2)}\right)} a_s^2\sigma_{pix}^2 + \\ & \frac{\left(\frac{\sin\beta}{\sqrt{M_c^2+N_c^2+f_y^2}} - \frac{(N_c\sin\beta+f_y\cos\beta)N_c}{(M_c^2+N_c^2+f_y^2)^{3/2}}\right)^2}{1 - \frac{(N_c\sin\beta+f_y\cos\beta)^2}{(M_c^2+N_c^2+f_y^2)}} \sigma_{pix}^2 + \\ & \frac{(N_c\cos\beta f_y\sin\beta)^2 M_c^2}{(M_c^2+N_c^2+f_y^2)^3\left(1 - \frac{(N_c\sin\beta+f_y\cos\beta)^2}{(M_c^2+N_c^2+f_y^2)}\right)} \sigma_\beta^2 \end{aligned} \tag{117}$$

Variance of $\hat{\theta}$

Recall that $\hat{\theta}$ is given by the following expression

$$\begin{aligned} \hat{\theta} &= \arctan 2\left(\frac{y_{ngm}\cos\theta_{ngm} - x_{ngm}\sin\theta_{gm}}{x_{ngm}\cos\theta_{ngm} + y_{ngm}\sin\theta_{gm}}\right) - \theta_{agv} - \alpha \\ &= f_\theta(\theta_{gm}, \theta_{cam}, \alpha) \end{aligned} \tag{118}$$

where $\mathbf{n_{gm}} = [x_{ngm}, y_{ngm}]^T$ is the normal to the guide mark and α is the pan angle.

Obviously, the derivative of $f_\theta(\cdot,\cdot,\cdot)$ does not exist for

$$x_{ngm}\cos\theta_{gm} + y_{ngm}\sin\theta_{gm} = 0. \tag{119}$$

which can occur in two different cases

1. $x_{ngm} = 0 \wedge \theta_{gm} = 0$
2. $y_{ngm} = 0 \wedge |\theta_{gm}| = \frac{\pi}{2}$

Assuming that the derivative exists, a linearization around θ_{gm}, θ_{agv} and α leads to

$$\sigma_{\hat{\theta}}^2 = [\nabla_{\theta_{gm},\theta_{cam},\alpha} f_\theta] \begin{bmatrix} V\{\theta_{gm}\} & 0 & 0 \\ 0 & V\{\theta_{cam}\} & 0 \\ 0 & 0 & \sigma_\alpha^2 \end{bmatrix} [\nabla_{\theta_{gm},\theta_{cam},\alpha} f_\theta]^T$$

This can be simplified to

$$\sigma_{\hat{\theta}}^2 = V\{\theta_{gm}\} + V\{\theta_{cam}\} + \sigma_\alpha^2 \tag{120}$$

Variance of $\hat{x}_{w,0}$ and $\hat{y}_{w,0}$

Recall from previously that $\hat{x}_{w,0}$ and $\hat{y}_{w,0}$ are given by

$$\begin{aligned} \hat{x}_{w,0} &= D(x_{ngm}\cos\theta_{gm} + y_{ngm}\sin\theta_{gm}) + x_{gm} \\ &\quad -x_{agv,opt}\cos\hat{\theta} + y_{agv,opt}\sin\hat{\theta} \qquad (121) \\ \hat{y}_{w,0} &= D(y_{ngm}\cos\theta_{gm} - x_{ngm}\sin\theta_{gm}) + y_{gm} \\ &\quad -x_{agv,opt}\sin\hat{\theta} - y_{agv,opt}\cos\hat{\theta} \qquad (122) \end{aligned}$$

where $(x_{gm}, y_{gm})^T$ is the guide mark center specified in world coordinates, $(x_{ngm}, y_{ngm})^T$ is the normal to the guide mark and $(x_{agv,opt}, y_{agv,opt})^T$ is the (x, y) coordinate of the focal point in the AGV-coordinate system.

By linerization, the following expression is obtained

$$\begin{aligned} V\left\{\begin{bmatrix} x_{w,0} \\ y_{w,0} \end{bmatrix}\right\} &= [\nabla_{D,\theta_{gm},\theta} f_{x_{w,0},y_{w,0}}] \begin{bmatrix} V\{D\} & 0 & 0 \\ 0 & V\{\theta_{gm}\} & 0 \\ 0 & 0 & \sigma_{\hat{\theta}}^2 \end{bmatrix} \\ &\quad \times [\nabla_{D,\theta_{gm},\theta} f_{x_{w,0},y_{w,0}}]^T. \qquad (123) \end{aligned}$$

Although it is not a particularly well-motivated assumption (the motivation is that the expressions are *extremely* complicated) it is for the sake of simplicity assumed that the covariance between $\hat{x}_{w,0}$ and $\hat{y}_{w,0}$ can be neglected. The expressions for the variance of $\hat{x}_{w,0}$ and $\hat{y}_{w,0}$ takes the form

$$\begin{aligned} \sigma_x^2 &= (\sin\hat{\theta} x_{agv,opt} + \cos\hat{\theta} y_{agv,opt})^2 \sigma_{\hat{\theta}}^2 + \\ &\quad D^2(-\sin\theta_{gm} x_{ngm} + \cos\theta_{gm} y_{ngm})^2 V\{\theta_{gm}\} \\ &\quad (\cos\theta_{gm} x_{ngm} + \sin\theta_{gm} y_{ngm})^2 V\{D\} \qquad (124) \\ \sigma_x^2 &= (-\cos\theta x_{agv,opt} + \sin\theta y_{agv,opt})^2 \sigma_{\hat{\theta}}^2 + \\ &\quad D^2(\cos\theta_{gm} x_{ngm} + \sin\theta_{gm} y_{ngm})^2 V\{\theta_{gm}\} \\ &\quad (-\sin\theta_{gm} x_{ngm} + \cos\theta_{gm} y_{ngm})^2 V\{D\} \qquad (125) \end{aligned}$$

4.3.6 Implementation of the Covariance Calculations

[Bak97] contains some considerations on the practical implementation of the variance calculations. Since the expressions for determination of the variances are based on a long sequence of approximations already, one might as well take the full step and substitute some of the more complex calculations by look-up tables.

In [Bak97] an experiment was made with the MATLAB simulator which revealed the following properties:

1. $V\{D\}$ depends mainly on D.
2. $V\{\theta_{gm}\}$ depends mainly on D and θ_{gm}.
3. $V\{\theta_{cam}\}$ is nearly constant.

It was furthermore shown that the following formula could be used with good approximation for the Phillips camera in its present mount

$$V\{D\} = (0.0518D - 0.0003)^4 \tag{126}$$

Table 4.3.6 was created to simplify the determination of the variance $V\{\theta_{gm}\}$. The average of the spread is specified in intervals of D and θ_{gm}.

D [m]	θ_{gm} [°]					
	[3.6-18[	[18-32.4[	[32.4-46.8[	[46.8-61.2[	[61.2-75.6[	[75.6-90]
[0.250-0.925[	1.2413	0.5663	0.3661	0.2862	0.2287	0.2030
[0.925-1.600[	2.6342	1.1853	0.7622	0.5703	0.4815	0.4647
[1.600-2.275[	4.2521	1.7865	1.1373	0.8710	0.7388	0.6831
[2.275-2.950[	5.5758	2.3924	1.5434	1.1704	0.9717	0.9194
[2.950-3.625[	7.0179	3.0955	1.8964	1.4636	1.2375	1.1541
[3.625-4.300[	7.5092	3.5112	2.3251	1.7414	1.5013	1.3914
[4.300-4.975[	9.3853	4.3272	2.6543	2.0852	1.7728	1.6315
[4.975-5.650[	11.1270	4.5287	3.1251	2.3361	1.9784	1.8364
[5.650-6.325[	12.9823	5.3413	3.3831	2.6859	2.2953	2.0887
[6.325-7.000]	16.5801	6.6609	3.9816	2.9932	2.4952	2.3291

The spread of θ_{cam} is in order of magnitude of $0.02°$ and is thus neglected since it in general will vanish when compared to the spread of θ_{gm}.

The covariance matrix $R_k = \begin{bmatrix} \sigma^2_{\hat{x}_{w,0}} & 0 & 0 \\ 0 & \sigma^2_{\hat{y}_{w,0}} & 0 \\ 0 & 0 & \sigma^2_{\hat{\theta}_{w,0}} \end{bmatrix}$ is determined according to the following procedure

1. Determine $V\{D\}$ by evaluating (126)
2. Determine $V\{\theta_{gm}\}$ from look-up table.
3. Determine R_k by evaluating (120), (124), (125).

References

[Bak97] M. Bak. Sensor fusion using kalman filering. Master's thesis, IMM/IAU, DTU, 1997. (In Danish).

[Lar97] David H. Larson. Image based calibration toolbox. Technical report, IAU, 1997.

[Lev82] M. Levin. *Computer Based Control Systems*. IAU, DTU, 1982. (In Danish).

[LL95] Thomas Dall Larsen and Jacob Lildballe. Vision based robot control. Master's thesis, Department of Automation, DTU, 1995. (in Danish).

[NPR97] Magnus Nørgaard, Niels K. Poulsen, and Ole Ravn. Autonomous guided vehicle - modelling. Technical report, Dep. of Mathematical Modelling, DTU, July 1997.

[oH93] Mortensen og Henriksen. Natural and artificial guide marks in AGV navigation. Master's thesis, IAU, DTU, 1993. (In Danish).

[Sør94] Allan Theill Sørensen. *Analysis of Real-Time Vision in Control and Automation Systems*. PhD thesis, IAU, DTU, 1994. (In Danish).

A Sensor Constants

Encoder constants:

Pulses/rotation, 800

Phillips camera constants:

Focal length, f_y	756 pixels
Radial distortion coefficient, κ_1	5.85×10^{-7}
Aspect ratio, a_s	1.015
Optical axis' intersection	
with the image plane, (C_x, C_y)	$(240.34, 231.52)$
Resolution	512×512 pixels

Phillips camera mounted to PTU unit:

Tilt base coordinates in camera coordinate system

Δ_y -0.093m;

Δ_z -0.050m;

Center mount of PTU unit:

Tilt base coordinates in AGV-coordinate system

Δ_{xt} -0.010m

Δ_{yt} 0.0m;

Δ_{zt} 0.635m;