POSITION CALIBRATION OF A MOBILE ROBOT BASED ON 3D VISION

Niramon Ruangpayoongsak[1]
Hubert Roth[1]
Rudolf Schwarte[2]

[1]University of Siegen, Institute of Automatic Control Engineering,
[2]Institut für Nachrichtenverarbeitung (INV),
University of Siegen, 57068 Siegen, Germany
Phone: 0049-271-740-4439
niramon.r@uni-siegen.de, hubert.roth@uni-siegen.de
rudolf.schwarte@uni-siegen.de

Abstract: This paper proposes a new approach to mobile robot navigation using 3D vision and artificial landmarks. The Photo Mixer Device (PMD) technology, working on the principle of time of flight (TOF), provides the capability for the position calibration of a mobile robot in an indoor environment. In this work, the PMD camera, which provides a distance value for each pixel, results in an innovation in the design of landmarks and it enables 3D object recognition without the requirement of gathering an image database using upper and lower pixel processing. Parallel to global localization, the robot and its heading position obtained from landmark recognition is calibrated and hence accumulated errors from dead reckoning and gyro drift are eliminated.

Keywords: Mobile Robots, Navigation, Time Of Flight (TOF), Object Recognition, Localization

1. INTRODUCTION

Up to the present, the localization of Autonomous Guided Vehicles (AGV) has relied on various techniques using on-board sensors, such as infrared sensors, sonar sensors, laser range finders, encoders, compass sensors, motion sensors, and vision sensors such as charge coupled devices (CCD) cameras. These sensors vary in their strengths and weaknesses (Nehmzow, 2003). A popular localization technique called simultaneous localization and mapping (SLAM) combines sensor data such as a range finder and dead reckoning to provide good results. There are also many other ways of positioning a robot, for example by using vision based techniques.

One option is to use set position landmarks. Two types of landmarks are widely used, natural and artificial landmarks. Natural landmarks are objects that already exist in the environment and are not intentionally created only for the purpose of robot sensation. Artificial landmarks are designed by the user for object recognition and are located in the environment at a specified position. Artificial landmarks can be 2D or 3D objects and are sometimes differentiated by color. The landmark properties must be adapted for perception by the equipped sensors. The most popular vision sensor for object recognition is a CCD camera that produces gray-scale images. Even though there are many algorithms available for image processing and object recognition using statistic, probability and artificial intelligence approaches, the image processing methodology suited to 3D vision cameras has to provide distance value as an output.

This paper proposes a new method of mobile robot navigation using a Photo Mixer Device (PMD) camera and a 3D artificial landmark. This innovative artificial landmark and object recognition approach uses the ability of the PMD camera to measure the depth or distance to an object directly without having to process the gray-scale value of 16x16 pixels. The outputs from landmark recognition are then further exploited as inputs for the position calibration process.

This paper is organized as follows: first, section 2 describes a mobile experimental robot for locomotion and intelligent navigation (MERLIN) to be used as a platform; section 3 discusses the PMD camera principle; section 4 explains position calibration using the PMD camera and 3D artificial landmarks; and section 5 presents experimental results, followed by a conclusion, section 6.

2. THE MOBILE ROBOT

The mobile experimental robot for locomotion and intelligent navigation (MERLIN) was designed and has been implemented for a broad spectrum of indoor and outdoor tasks. It uses standardized functional modules like sensors, and actuators and a radio link or wireless LAN for communication. It is equipped with a gyroscope, magnetic encoders, ultrasonic sensors, bumpers, and a 3D compass for navigation. Figure 1 shows MERLIN with a PMD camera on board. The robot is controlled by an 80C167 CR 16 bit-processor for sensor and actuator interfaces and a PC104 that runs on Linux OS and 933 MHz CPU with a wireless LAN.

As MERLIN is an experimental mobile robot platform, the joystick and trajectory teleoperations are controlled by internet. The java client using TCP/IP communication is run on a client PC and java robot server is run on a PC104. The wireless LAN is used for data communication between the client running GUI and the server running a data transfer interface from robot to client PC via socket. MERLIN is also tele-operated via joystick control with automatic obstacle avoidance and haptic interface (as described in Kuhle *et al.*, 2004). Because of the innovation of the PMD camera, obstacle avoidance is developed using Fuzzy logic (Roth *et al.*, 2003a, b). Current localization uses dead reckoning, using magnetic encoders on the front steering wheels and kalman filtering. Error reduction in localization has been investigated.

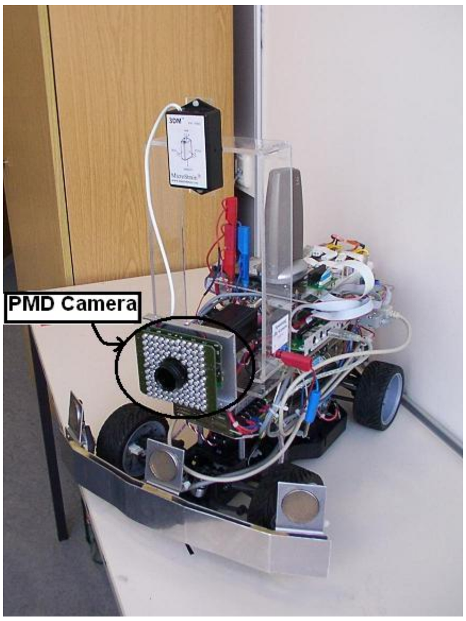

Fig. 1: MERLIN with PMD Camera

3. PHOTO MIXER DEVICE (PMD) CAMERA

The PMD camera has emerged from the new PMD technology. It provides ready-to-use depth information or three-dimensional vision without having to do gray-scale image processing. Each of the 16x16 pixels on the PMD image is the measured distance value and its detection area covers not only altitude like a laser scanner does, but it also covers the vertical and horizontal detection area with the cone volume of vision, like a CCD camera. A description of the PMD camera system and hardware components are presented below.

3.1. PMD Camera System

The principle behind a 3D TOF imaging system is shown in Fig. 2 schematically. An object is illuminated by a high frequency modulated light source. The reflected light signal is compared with an electric reference signal. All presented systems are equal in the fundamental function of photo-detection and signal processing. Imaging starts with the detection of light, wideband amplifying, signal conversion and quantization. Each step has its own error source (for example, noise) that is transferred through the whole system (Xu, *et al.*, 1998). The light that is backscattered from the object, however, is directly sensed and demodulated in the same area by the Photonic Mixer Device (PMD) array. The depth information of the scene is therefore acquired in pixels using the correlation results of the received optical signal and the demodulation signal in parallel for the complete matrix.

3.2 PMD Camera Hardware Components

Fig. 3 shows the elements of the 16x16 pixel camera. The IR-LEDs are arranged around an opening where the receiving optic is placed. The LED-array and the optics are fixed on a plate that also carries the PMD-array. This plate is screwed onto the camera case which contains the system boards with power supply, CPU and CPLD. Considering the known optic every pixel has its own lens coverage. When a target enters the field of view, the camera locates the object, by measuring its relative position according to the position of the pixel in the matrix and the measured distance value. The camera operates at the frame rate of 3 Hz and the half angle (α) of the detection area is 11 degrees. The maximum detection distance is 15 meters and gray-scale output can also be obtained. It is necessary to designate the 3D landmarks as to the reflection of light on different materials and the colours of an object, and the area of detection cone

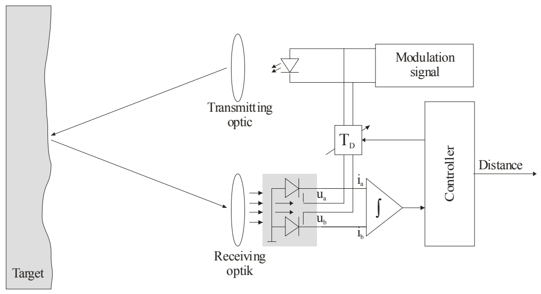

Fig. 2: Schematic PMD TOF setup

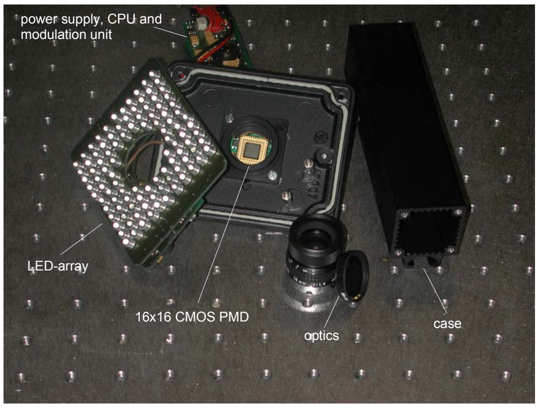

Fig. 3: Photography of 3D camera components

covered by the camera. This is described below in section 4.

4. CALIBRATION OF POSITION BASED PMD CAMERA AND 3D ARTIFICIAL LANDMARK

The position calibration technique used can be broken down into 3 sequential procedures: design of the artificial landmark, landmark recognition and position prediction and update. During navigation, the last two procedures run without the first procedure. However, the landmark design procedure is detailed below for further use in related fields.

4.1 Design of artificial landmark

The design of artificial landmarks material, colour, shape and size into consideration. The material used must reflect enough light and thus it should not be black in colour. The landmark should have a geometrical shape such that the edges and corners can easily be determined from the 16x16 pixel image. The landmark should not be too big or too small in order that the objects lie within the detection area of the camera at a distance of one to two meters. A sample landmark is shown in Fig. 4. On the RHS, the image is plotted using the distance value from the PMD camera. The black pixels represent a large distance value (far away). White pixels represent the surface area of the landmark, which is closer than the background (black). The blue pixels represent a distance value between white and black. There are some pixels on the edges and corners which are a mixture of white and blue. This shows light reflected at the edges; some of it scatters away and some returns to the camera.

The landmark is divided into two parts, the upper and lower part. The upper part is always symmetrical in shape, such as the cylinder shown in Fig. 4. This part must be mounted on the top of the lower part by referring to the centre of gravity in order that the image taken of this part is always the same for all directions. This is designed to simplify the object recognition strategy. The lower part always has a rectangular shape used to determine the distance and rotational angle position. Hence, each combination of upper and lower part produces a landmark that can be distinguished by the camera.

Here, instead of database images, model images are generated for a matching process in object recognition. Thus, an image database is not necessary. This design can also be applied to similar applications. The great advantage of this design is that the large database for various rotational angles and distances is reduced by keeping individual databases for the lower and upper parts and doing the model matching separately.

4.2 Landmark recognition

Most image processing literature refers to gray-scale images obtained from a CCD camera and implemented for higher resolution than 16x16 pixels. These techniques operate either directly with the pixel values and are called spatial domain techniques or with frequency and are called frequency domain techniques. Unfortunately, these techniques are not directly applicable to the existing PMD camera used here due to low resolution and distance value output. At the early stage, the time-based filtering technique is still necessary for reducing noises, although the camera provides ready-to-use output.

Image filtering Powerful kalman filtering is implemented here because of its well- known high performance. kalman filtering techniques use a system model, process noises, measurement noises,

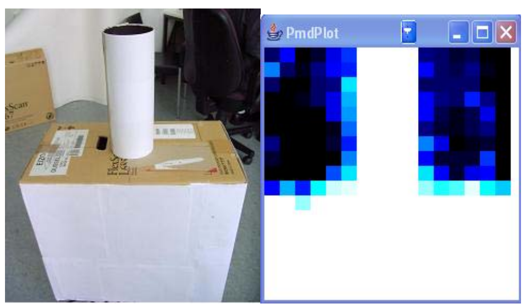

Fig. 4: An example of a 3D landmark (left) and image from PMD distance value (right)

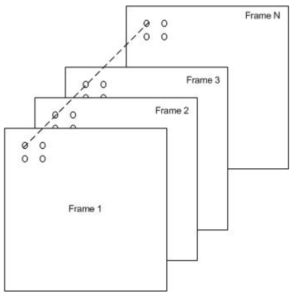

Fig. 5: Image frames consisting of state variables

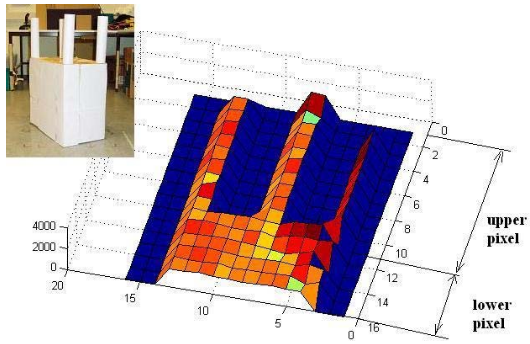

Fig. 6: Upper and lower pixel division

and the result from recursive computation yields the convergence of the self-adaptive kalman gain and estimated state variable (Welch and Bishop, 2002). Two steps of the discrete kalman filter algorithm are implemented here for estimation and update by using equations (1-2) and (3-5), respectively.

$$xe_k^- = Axe_{k-1} + Bu_{k-1} \tag{1}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{2}$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{3}$$

$$xe_k = xe_k^- + K_k (z_k - H\, xe_k^-) \tag{4}$$

$$P_k = (I - K_k H) P_k^- \tag{5}$$

At the first iteration $k = 1$, the value of xe_{k-1} and P_{k-1} are previously initialized. At state k, the a priori estimate state variable xe_k^- is calculated from the transition matrices A and B, the previous estimate state variable xe_{k-1} and previous control variable u_{k-1}. Then, the a priori estimate error P_k^- is predicted from the previous estimate error P_{k-1} and process noise covariance matrix Q. After that the update process is continually performed by calculating kalman gain K_k from P_k^-, the output matrix H and measurement noise covariance matrix R. The kalman gain is immediately used again for the calculation of the estimate state variable at the current iteration, xe_k, which is based on the measurement value z_k, and the predicted state variable xe_k^-. Last, the estimate error P_k is updated and the process is repeated again until the last specified number of iteration is reached.

In order that the PMD image can be filtered, the the picture is taken when the camera position is static and the filter is done at each individual pixel over the sampling period (frames) as shown in Fig 5. The sample data for each pixel and each frame are gathered as state variables, which is a scalar equal to the distance value of each pixel. Thus, the transition matrix A is a scalar equal to 1 and matrix B is zero. Equation (1) is thus simplified to $xe_k- = xe_{k-1}$.

3D Object Recognition In order to support the proposed landmark design, image recognition is also divided into two parts: upper and lower pixel processing. The first step is to recognize the distance and angle position, and the second is to classify landmarks. In Fig. 6, the upper pixel shows the shape of the individual landmark; whereas, the lower pixel shows the rectangular shape. The recognition process shown in Fig. 7 starts from lower pixel processing of the two images, the filtered image and the generated model image. Two outputs, rotational angle and distance positions, are generated. The distance measurement is used further for upper pixel processing and the landmark class is obtained last.

Model Image Generation As shown in Fig. 8, the coordinate of camera and landmark are defined as $P(x_c, y_c, z_c)$ and $P(x', y', z')$, respectively. To generate the model image, the 3D space is simplified into 2D space by showing the top view of these coordinates. When considering the landmark that consists of 4 edges, 4 linear equations representing a rectangular box are constructed from the relationship of four corner coordinates as in equations (6-7).

$$a_i = (y_i - y_{i+1})/(x_i - x_{i+1}) \tag{6}$$

$$b_i = y_i - ax_i \tag{7}$$

The camera equations are generated by

$$y_i = a_c x_i + b_c \tag{8}$$

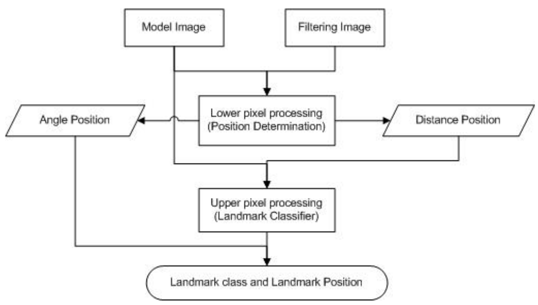

Fig. 7: PMD image recognition processes

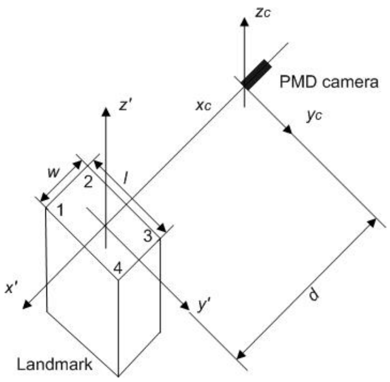

Fig. 8: Camera and landmark coordinates

Where a_i, a_c, b_i and b_c are linear equation coefficients of the corner i and camera vision shown in Fig. 9, respectively. To convert a landmark coordinate to a camera coordinate, $P(x', y', z')$ to $P(x_c, y_c, z_c)$, the rotation with angle δ and translation with distance d are included using equation (9). Note that z and y axes of both coordinates are equivalent.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} \cos\delta & \sin\delta & 0 \\ -\sin\delta & \cos\delta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + \begin{bmatrix} d \\ 0 \\ 0 \end{bmatrix} \quad (9)$$

The generated model image is the image of distance position obtained from the intersection of an object and camera equations as shown in Fig. 9. After the object is projected onto the camera coordinate, the intersection can be found by replacing equation (7) in (8). Note that for upper part model image generation, the same method is implemented but equations are different, depending on the shape of landmarks.

Model Matching In this process, there are two sub-processes, the upper and lower pixel matching processes. To match the lower pixel model, distance position D and rotational angle position δ are varied from minimum to maximum value in iterative loop. At each position, the model image is generated and the filtered image is matched to this model using a minimum distance classifier (Gonzalez and Woods, 1992). Similarly, when processing the upper pixels, the landmark classifier generates the model of the upper part and searches for the matched model.

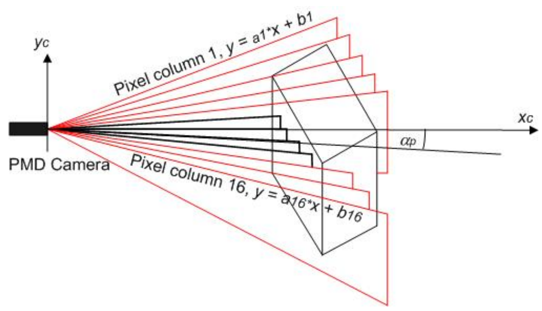

Fig. 9: Camera equation generation

4.3 Position prediction and update

The idea behind doing position calibration is to reduce the localization errors that occur, such as the heading and distance errors regarding slipped encoders, accumulation heading error from the drifting of the gyroscope. Global localization can be improved by correcting the position when the artificial landmark is observed and recognized. Based on the preliminary known landmark position, D and δ referred to by the camera coordinate are previously obtained from the image recognition process. As shown in Fig. 10, the vectors of robot position V_r is calculated from the landmark position V_o and the output of landmark recognition V_d by using

$$V_r = V_o + V_d. \quad (10)$$

In the figure, the global, camera, and robot coordinates are denoted by (X_g, Y_g), (X_c, Y_c), and (X_r, Y_r), respectively. δ is the rotational angle between the camera and object coordinates assumed to be the global coordinate. The camera heading angle β is the angle between the robot and the camera headings and α is the angle of the robot heading referred to by the global y axis Y_g. β is previously known by orientation and positioning of the camera motor, which is mounted on the robot. If β equals zero, the camera heading always points to the center of gravity of the object. With the a priori known object position $P(X_o, Y_o)$ on the global coordinate and the rotational angle

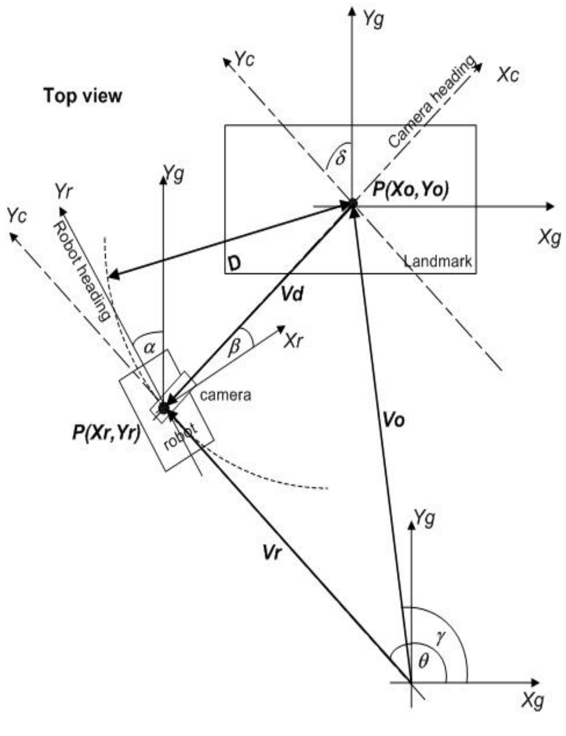

Fig. 10: Defined coordinates and related angles

δ and D obtained from the previous section, the robot position $P(X_r, Y_r)$ using the relationship in equation (10), is calculated by

$$\begin{bmatrix} X_r \\ Y_r \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \end{bmatrix} + \begin{bmatrix} D\cos(180+\delta) \\ D\sin(180+\delta) \end{bmatrix} \qquad (11)$$

and the robot heading α, the robot angle position, and the landmark angle position, referred to by the global coordinate are obtained from equations (12-14), respectively.

$$\alpha = \delta - \beta. \qquad (12)$$

$$\gamma = \tan^{-1}(Y_o/X_o) \qquad (13)$$

$$\theta = \tan^{-1}(Y_r/X_r) \qquad (14)$$

where X_r, Y_r, X_o and Y_o are position coordinates of the robot and landmark on the global coordinates, respectively.

5. EXPERIMENTAL RESULTS

The methodology of landmark recognition has been tested off-line with four landmark classes. The generated model image and filtered image are well matched according to the restricted low resolution and noises criteria. The design strategy can distinguish the rotational angle positions at 0, 45 and 90 degrees at distance positions of 80, 100, and 120 cm successfully. The matching process selects one of the above positions that best fit to the filtered image. Fig. 11 shows the matching results of the tests. The time duration for position calibration depends mainly on the image filtering process that requires at least 50 frames.

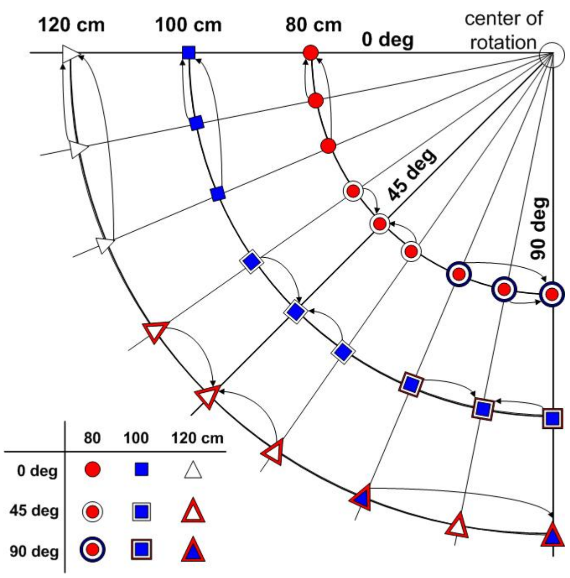

Fig. 11: Matching results

6. CONCLUSION

This paper proposes a new approach to mobile robot navigation using 3D vision and artificial landmarks. This requires the design and implementation of an artificial landmark, landmark recognition and position calibration. The results show that the design algorithms can be applied under 16x16 resolutions and noises constraints and the time consumption is mainly that of the filtering process. Further developments should occur in the implementation of a higher resolution camera.

7. REFERENCES

Gonzalez, R. C. and Woods, R. E. (1992). *Digital Image Processing*, p. 580, Addision-Wesley, USA.

Kuhle, J.; Roth, H.; Ruangpayoongsak, N. (2004). *MOBILE ROBOTS and airships in a multi-robot team*, The 1st IFAC Symposium on Telematics Applications in Automation and Robotics, Helsinki University of Technology, Finland, pp. 67-72.

Nehmzow, U. (2003). *Mobile Robotics: A Practical Introduction*, 2nd Edition, Springer-Verlag, London.

Roth, H.; Schwarte, R.; Ruangpayoongsak, N.; Kuhle, J.; Albrecht, M.; Grothof, M.; Heß, H. (2003a). *3D Vision Based on PMD-Technology for Mobile Robots,* Aerosense - Technologies and Systems for Defense & Security 2003, SPIE Conference, Orlando, Florida, paper no. 5803-66.

Roth, H.; Schwarte, R.; Ruangpayoongsak, N.; Kuhle, J.; Albrecht, M.; Grothof, M.; Heß, H. (2003b). *3D Vision Based on PMD-Technology and Fuzzy logic control for Mobile Robots,* Second International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, Antalya, Turkey, pp. 145 – 153.

Welch, G. and Bishop G. (2002). *An Introduction to the Kalman Filter*, UNC-Chapel Hill, TR 95-041, USA.

Xu, Z.; Schwarte, R.; Heinol, H.; Buxbaum, B.; Ringbeck, T.(1998). *Smart pixel - photonic mixer device (PMD), New system concept of a 3D-imaging camera-on-a-chip*, 5th International Conference on Mechatronics and Machine Vision in Practice, Nanjing, pp. 259-264.