ULTRASONIC AND VIDEO DATA FUSION FOR MOBILE ROBOT NAVIGATION

A. Bonci, T. Leo, S. Longhi

Università di Ancona, Dipartimento di Elettronica e Automatica, via Brecce Bianche, 60131 Ancona.
fax: +390712204835
e-mail: bonci@ee.unian.it

Keywords: sensor fusion, mobile robots, ultrasonic transducers, uncertainty, robot vision.

Abstract

A multisensor fusion approach for improving the map-building capability of a mobile robot is presented in this paper. Ultrasonic and video data are used. A modelling technique for indoor environments based on line features extraction from video data and from range data is proposed. The Hough transform is considered for extracting lines from the occupancy grid and from video images.

1 Introduction

An autonomous vehicle requires perception of its local environment for both sensor based navigation and position estimation. Occupancy grids, based on ultrasonic range readings [6], [12], [1], provide a robust tool for building the local map of the environment for navigation. Map building has been addressed by many researchers, over the years two basic approaches to environment representation have been developed: grid-based modelling and feature-based modelling. In the first approach the workspace of the robot is decomposed into square areas denoted as cells. In each cell a value is stored that corresponds to the level of certainty that an obstacle exists within the cell area. Actually the main approaches for computing the certainty level of a cell are based on integration of ultrasonic readings using probabilistic [6], possibilistic [7], and evidential [14], techniques. An analysis of these methods with experimental results is reported in [5].

A characteristic of the structured environments is that objects tend to lie in straight lines. Indoors environments can be represented by a collection of line segments, representing the vertical surfaces of walls, doors, objects, etc. In the feature-based modelling approach line segments or surfaces are used for modelling indoor environment and for improving the estimated position of the mobile robot [15], [2]. Line segments can be founded in the occupancy grids as aligned cells of high probability of occupation. By considering a grid and its probabilities as an image with different levels of resolution it is possible to apply image processing algorithms to the detection of line features. Line segments can be also recovered from video images of the indoor scenes using well-known image processing tools. In this paper a method for improving accuracy and reliability on the extraction of line segments from occupancy grid based on a technique for fusing ultrasonic and video data is proposed. Low cost sonar sensors and CCD camera are used.

The paper is organised as follows: Sections 2 introduces the problem of modelling the environment using line features. Section 3 describes the proposed techniques for detecting line features. The integration of line features is described in Section 4. Experimental results are introduced and discussed in Section 5. Concluding remarks are explained in Section 6.
2 Problem statement

Two representations have been used for dynamic 2D environment modelling of a mobile robot using ultrasonic range readings: occupancy grids and parametric features. The occupancy grid method is well adapted to local path planning and reactive navigation using a variety of algorithms. The main drawbacks are the difficulty in using grids to improve the vehicle position estimation and the amount of computer memory needed for representing large environments. Parametric features describe the boundaries of free-space in terms of lines or surfaces defined by a list of parameters. Such a description is useful to the local path planning and to the position estimation. Unfortunately, noise and uncertainty in ultrasonic sensor signals make unreliable the process of grouping ultrasonic sensor readings in geometric primitives. The pre-filtering procedure mentioned in [1] reduces the wrong readings due to multiple reflections. The probabilistic mapping reduces uncertainty and noise on sonar measures by a recursive Bayesian updating. But also using these procedures small obstacles such as table legs are practically impossible to distinguish from noise. In general the use of complementary information given by different kind of sensors reduces the problems of grouping adjacent sensor measures for obtaining more reliable features. In this paper complementary information extracted by video measures will be integrated with sonar readings. In particular a method for modelling the robot environment by extracting parametric line features and the associated certainty level both from the occupancy grid and from video images. A matching between lines extracted from occupancy grid and from video images is performed for obstacles belonging to the part of the floor visible from the CCD camera (also called “visible space”). The computation of the visible space of the camera is performed by using the “viewing frustum” [4]. The pinhole model, for the video camera installed on the mobile robot, is considered. This model is well justified for modern solid-state cameras, especially in the context of mobile robots. The camera parameters are estimated by the algorithm proposed in [9]. In the pinhole model the viewing frustum is the projection of the image plane corners from the pinhole that is located one focal length behind the image plane. Pointing the camera down towards the floor, in the forward direction of the robot, the visible space on the ground floor is deduced by the projection of the frustum vertices on the floor plane.

3 Line features detection

The problem of line feature extraction is faced for both kinds of sensor data representations:

\[ w_i = [\Theta_i, P(\Theta_i)] \]  

where \( \Theta_i = (\theta_i, \rho_i) \) is a vector of the line parameters \( \theta_i \) and \( \rho_i \); \( \theta_i \) is the orientation angle of the \( i \)-th line, \( \rho_i \) is the distance between the origin and the \( i \)-th line and \( P(\Theta_i) \) is the evidence that exist a line feature having parameters \( \theta_i \) and \( \rho_i \).
occupancy grids and video images. The Hough transform (HT) is used for line features extraction.

3.1 The standard Hough transform

The standard HT [10], [8] is a well-known method for detecting line segments on grey level images. The HT is based on a function representing the polar equation of the line segment

\[ f(x, y, \rho, \theta) = x \cdot \cos \theta + y \cdot \sin \theta - \rho = 0. \]  

(2)

Where \( \rho \) is the distance between the origin of the considered reference frame and the line, \( \theta \) is the orientation of the normal to the line with the x-axis of the reference frame, and \( x, y \) are the co-ordinates, in the image space of the line feature points.

The HT is accomplished by first creating an accumulator array \( H(\theta, \rho) \), called Hough space, to represent each possible set of \( (\theta, \rho) \). Thus the Hough space is approximated by a discrete array. For each edge points \((x, y)\) in the image, detected by using some orthogonal differential operator such as Sobel operator [8], the parameters \((\theta, \rho)\) are estimated and quantized. Then, for each edge point, all corresponding \((\theta, \rho)\) elements in the accumulator are incremented accordingly. After the edge points processing, the accumulator array is searched for peaks. The peaks identify the parameters of the most likely lines. An extension of the HT is here proposed for calculating the probability of a \((\theta, \rho)\)-couple, which is the probability of a line, by using only the probability of edge points. In the occupancy grid an edge point is a grid cell with a high probability to be occupied, while in the video image an edge point is a pixel with a high probability to be an edge.

3.2 Finding line features in occupancy grids

During the robot exploration the cells of the occupancy grid are updated using the probabilistic signal level fusion of ultrasonic range data proposed in [1]. Line segments can be founded in the occupancy grid as aligned cells with high probability of occupation. By interpreting a grid and its probabilities as an image with different level of resolution (grey level) is possible to apply the HT for detecting straight lines. The probabilities of the lines can be computed as follows.

Given the probability of each grid-cell belonging to the line represented by the \((\theta, \rho)\)-couple, the proposed extension of the HT computes the probability of the line segment. This probability is computed by means of Bayesian and Soft Evidence theories [3]. In order to specify the computational algorithm, the following notations are introduced.

Given a line with parameters vector \( \Theta_G = (\theta, \rho) \), denote \( n \) the number of cells in the occupancy grid belonging to the line \( \Theta_G \) and \( c_i(\Theta_G) \), \( i=1,...,n \), the event “the i-th cell of the occupancy grid belonging to the line \( \Theta_G \) is occupied” and \( \tilde{c}_i(\Theta_G) \) the unsure event “the i-th cell of the occupancy grid belonging to the line with vector parameters \( \Theta_G \) is occupied with a proper uncertainty \( P(c_i(\Theta_G)|\tilde{c}_i(\Theta_G)) \)”. Therefore \( P(c_i(\Theta_G)|\tilde{c}_i(\Theta_G)) \) is the probability of event \( c_i(\Theta_G) \) given the evidence \( \tilde{c}_i(\Theta_G) \), which is the probability estimated by the sonar readings to the cell i-th of the occupancy grid and stored on the occupancy grid. In the following, the explicit dependence on \( \Theta_G \) in \( c_i(\Theta_G) \) and \( \tilde{c}_i(\Theta_G) \) has been dropped for simplicity of notation. Let \( P(c_i)=P(\neg c_i)=0.5 \) be the prior occupancy/non occupancy probability of the cell i-th and denote with \( P(\Theta_G) \) the prior probability of a line feature having parameters vector \( \Theta_G \). The Hough space \( H_G(\theta, \rho) \) is used for storing existence evidence of each detected line feature. The use of the HT is proposed for finding all the lines from the grid map and to store, for the i-th cell of the grid map belonging to the line, the map co-ordinates and the probability \( P(c_i|\tilde{c}_i) \). The existence evidence of a line feature with parameters vector \( \Theta_G \) depends from the evidence of all the cells \((x_i, y_i)\) satisfying equation (2), with \( x=x_i \) and \( y=y_i \). Therefore the probability of each line feature with parameter \( \Theta_G \) is the conditional probability of the line to the events \( \tilde{c}_1, \tilde{c}_2, ..., \tilde{c}_n \) and by Bayes theorem this probability has the form:
Each cell of the Hough space \( H_G \) (space with the probability \( \hat{\theta} \)) estimation of the line parameters \( \Theta \) are pixel and the available information is an \( \theta \) space. Assume \( \hat{\theta}, \hat{\rho}(\hat{\theta}) \) image. Denote the estimated line parameters from a pixel \( [11] \) and \( \hat{\rho} \) with the line probability computed using equation (3). Therefore a line segment on the grid map, with parameters \( \Theta \) is a local maximum in the Hough space \( H_V \) and \( \hat{\rho} \). Computing the contribution of each edge point to particular is proposed a Bayesian approach for propagation technique proposed in [11]. In

The proposed approach is based on the error uncertainty corrupting the image. A model for the uncertainty of line features extracted from digital images, based on a sensor data. A model for the uncertainty of line features extracted from digital images is different from the features probability in video images is different from features extracted from digital images, based on a Bayesian approach to the HT, is proposed.

### 3.2 Finding line features in video images

The extraction of line features from digital images using the HT has been extensively addressed in literature. The evaluation of the uncertainty joined to the line features extracted from a digital image is a key issue for accomplishing tasks as geometry extraction and combination of feature with other sensor data. A model for the uncertainty of line features extracted from digital images, based on a Bayesian approach to the HT, is proposed.

The considered approach for computing line features probability in video images is different from the approach proposed for the grid map. The difference is due to the available information on edge points used for computing the HT. In the occupancy grid, edge points are the grid cells and the available information is their occupancy probability \( P(\hat{\zeta} | \zeta) \). In the video image, edge points are pixel and the available information is an estimation of the line parameters \( \Theta_v \) based on gradient information and on a model of the uncertainty corrupting the image.

The proposed approach is based on the error propagation technique proposed in [11]. In particular is proposed a Bayesian approach for computing the contribution of each edge point to the Hough space \( H_V(\Theta, \rho) \) associated to the video image. Denote \( \Theta_v = (\hat{\rho}, \hat{\Theta}) \) a vector consisting of all possible quantized values of \( \Theta \) and \( \rho \) in the Hough space. Assume \( \hat{\Theta}_v = (\hat{\rho}, \hat{\Theta}) \) be a vector representing the estimated line parameters from a pixel [11] and

\[
P(\Theta_v | \hat{\zeta}_1, \hat{\zeta}_2, ..., \hat{\zeta}_n) = \frac{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | \Theta_v)}{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | -\Theta_v)} \times \frac{1}{1 + \frac{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | -\Theta_v)}{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | \Theta_v)}}
\]

(3)

Where the computation of the terms \( P(\hat{\zeta} | \Theta_v) \), \( P(\hat{\zeta} | -\Theta_v) \) are reported in Appendix A.

Each cell of the Hough space \( H_G(\Theta, \rho) \) is updated with the line probability computed using equation (3). Therefore a line segment on the grid map, with parameters \( \Theta_v \), is a local maximum in the Hough space with the probability \( P(\Theta_v | \hat{\zeta}_1, \hat{\zeta}_2, ..., \hat{\zeta}_n) \) greater than a probability threshold (often over 0.5).

The probabilistic update of \( H_V(\Theta, \rho) \) is accomplished only for lines \( \Theta_v \) with a non-zero probability. Denote \( e_i(\Theta_v) \) the evidence “the i-th edge pixel belong to the line feature with parameters \( \Theta_v \)”, where \( i=1, ..., m \) and \( m \) is the number of pixel in the video image belonging to the line \( \Theta_v \). Assume \( \hat{e}_i(\Theta_v) \) be the conditional independent pieces of evidence concerning \( \Theta_v \) and generated by a pixel whose probability \( P(\hat{e}_i | \Theta_v) \), is \( P(\hat{e}_i | \Theta_v) \).

Given the information \( \hat{e}_i(\Theta_v) \) about the pixel belonging to the line \( \Theta_v \), the posterior probability on \( \Theta_v \) has the following form, where the explicit dependence on \( \Theta_v \) of \( e_i(\Theta_v) \) and \( \hat{e}_i(\Theta_v) \) has been dropped for simplicity of notation:

\[
P(\Theta_v | \hat{\zeta}_1, \hat{\zeta}_2, ..., \hat{\zeta}_n) = \frac{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | \Theta_v)}{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | -\Theta_v)} \times \frac{1}{1 + \frac{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | -\Theta_v)}{P(\Theta_v) \prod_{i=1}^{n} P(\hat{\zeta}_i | \Theta_v)}}
\]

(5)

The term \( P(\hat{\zeta} | -\Theta_v) \) has the following form,

\[
P(\hat{\zeta} | -\Theta_v) = \frac{P(\hat{\zeta} | -\Theta_v) - P(\hat{\zeta} | \Theta_v)P(\Theta_v)}{1 - P(\Theta_v)}
\]

(6)

It is deduced by considering: Bayes theorem for \( P(\hat{\zeta} | -\Theta_v) \), the relation \( P(\hat{\zeta} | -\Theta_v) = 1 - P(\Theta_v | \hat{\zeta}) \) and the Bayes theorem for \( P(\Theta_v | \hat{\zeta}) \).

### 4 Fusion of line features

The proposed multisensor fusion process is based on an algorithm for fusing line features detected from sonar data and video data. The matching between the lines is performed for obstacles
belonging to the portion of the floor falling in the visible space of the camera. All the lines detected from the portion of the occupancy grid in the visible space can be matched with the lines detected from the video image by projecting the lines of the video image on the floor plane. The perspective projection of the lines by means of the pinhole camera model also allows updating the lines parameters of the Hough accumulator relative to the video image to the framework of the floor. The existence probability of the lines are stored in both Hough accumulators: $H_G(\theta, \rho)$ relative to lines detected from the occupancy grid and $H_V(\theta, \rho)$ relative to lines of the video image projected on the floor.

The proposed matching algorithm is based on the combination of the lines probability encoded in both Hough accumulators. In particular a Bayesian estimator is developed. For each $j$-th line feature, described by its vector of parameters $\Theta_j=(\theta_j, \rho_j)$ are provided two probabilistic estimates $P_G(\Theta_j | \theta_j)$ stored in $H_G$ and $P_V(\theta_j | \theta_j)$ stored in $H_V$, where $\theta_j$ is the estimation of the line parameters obtained by using the Occupancy Grid and $\theta_j$ is the estimation of the line parameters obtained by using the video image. Using the Bayes formula, the combined estimates $P(\theta_j | \theta_j \cup \theta_j)$ is given by

$$P(\theta_j | \theta_j \cup \theta_j) = \frac{P(\theta_j | \theta_j)P(\theta_j | \theta_j)}{\sum_{\theta_i} P(\theta_j | \theta_i)P(\theta_j | \theta_i)} \quad (7)$$

Where, using the Bayes formula, $P(\theta_j | \theta_j)$ is given in the following form:

$$P(\theta_j | \theta_j) = \frac{P(\theta_j | \theta_j)P(\theta_j)}{P(\theta_j)} \quad (8)$$

By substituting equation (8) in equation (7) the following combination formula for fusing the sonar and video data is obtained:

$$P(\theta_j | \theta_j \cup \theta_j) = \frac{P_V P_G}{P_G + \frac{P_V (1-P_G)}{1-P(\theta_j)}} \quad (9)$$

That is also known as Independent Opinion Pool [13].

5 Experimental results

The proposed approach has been tested in indoor environments by using a LabMate mobile base, that is a wheeled robot equipped with a proximity system composed of a half ring of 13 Polaroid ultrasonic sensors and with a low cost CCD webcam Philips PCVC 675K. Odometry sensors provide an estimation of the robot position. For the developed experiments the camera was installed in front of the vehicle and pointed down in the left side. Different experiments have been developed. In the following a sample of experimental test is reported. Figure 1 shows the occupancy grid map of an indoor environment built when the robot follows the trajectory from the white box to the circle. The map is represented in a grey scale of colours, which goes from black (the null occupancy probability) to white (the maximum occupancy probability). The video image, acquired when the circle in Figure 1 marks the robot position, is showed in Figure 2, where the line features extracted are also reported. The related lines probability stored in the Hough space $H_V$ is showed in Figure 3. The considered part of the occupancy grid when the robot is in the position marked by the circle in Figure 1 is showed in Figure 4. In the same figure the extracted line features are also reported.

The related lines probability stored in the Hough space $H_G$ is showed in Figure 5. As showed in Figures 2 and 4, and tested in a number of experiments, the HT results to an high number of overlapping lines that complicating the outcome. To extract line features separately from both sonar and video data is not enough for obtaining a reliable obstacle shape. In order to extract only the significant lines, a fusion procedure of the sonar data with the video data is used between the Hough accumulators. The result of the fusion procedure introduced in Section 4 is reported in Figure 6. The lines showed in Figure 6 specifies the shape of the obstacles (walls) having an high probability and the dashed line shows the camera visible space. In general the use of this procedure in the navigation module of a mobile vehicle increases the accuracy of environment knowledge with a consequent improvement of security and operating robustness of the system. The obstacles detected at the end of
the considered test are reported in Figure 7, where: real obstacles are reported by dashed lines and detected obstacles with continuous dark lines. This figure shows the satisfactory results obtained with the proposed fusion procedure. Similar results have been obtained with different trajectories and environments.

Fig. 1. Occupancy grid with: robot trajectory, robot starting position, robot current position (circle) and observed region (dashed line).

Fig. 2. Video image with extracted line features.

Fig. 3. Hough space $H_V$ storing the lines probability of the acquired image reported in Figure 2.

Fig. 4. Part of the occupancy grid falling in the camera visible space with the extracted lines.
6 Conclusions

The paper presents a fusion technique that deals with the map building issue. This technique is able to map large environment using line features; the resulting model is simple and accurate with minimum memory demand. Several experiments have showed that the proposed technique suffers if the sonar readings are completely unreliable but it work good in normal situations. The analysis of how the map building degrades if the environment is not highly structured will be investigated in further research activities. Other research efforts will be focused on the use of the low cost sensors and environmental modelling for solving localisation issues.

Appendix A

In equation (3) the unknown terms are within the product symbols, both terms can be computed using the following equations derived from the Bayes theory

\[
P(\hat{\theta} | \theta_i) = \frac{P(\theta_i | \hat{\theta}) \cdot P(\hat{\theta})}{P(\theta_i)} \tag{10}
\]

\[
P(\hat{\theta} | \neg \theta_i) = \frac{P(\neg \theta_i | \hat{\theta}) \cdot P(\hat{\theta})}{P(\neg \theta_i)} \tag{11}
\]

For the soft evidence [3], in equation (10) the term \( P(e_j | \hat{\theta}) \) has the form:

\[
P(\theta_j | \hat{\theta}) = P(e_j | \hat{\theta}) \cdot P(\theta_j | e_j) + P(\neg e_j | \hat{\theta}) \cdot P(\theta_j | \neg e_j) \tag{12}
\]

Where by Bayes theorem:
\[ P(\Theta_G | c_i) = \frac{P(c_i | \Theta_G) \cdot P(\Theta_G)}{P(c_i)} \]  
\[ P(\Theta_G | -c_i) = \frac{[1 - P(c_i | \Theta_G)] \cdot P(\Theta_G)}{P(c_i)} \]  
\[ P(-c_i | \hat{c}_i) = 1 - P(c_i | \hat{c}_i) \]  
Therefore it follows that
\[ P(\Theta_G | \hat{c}_i) = \frac{P(\Theta_G)}{P(c_i)} [1 + P(c_i | \Theta_G) \cdot [2P(c_i | \hat{c}_i) - 1] - P(c_i | \hat{c}_i)] \]  
(16)
Where \( P(c_i | \Theta_G) \) is experimentally estimated and obtained from the Bayesian equation
\[ P(c_i) = P(c_i | \Theta_G) \cdot P(\Theta_G) + P(c_i | -\Theta_G) \cdot P(-\Theta_G) \]  
(17)

In this way the term \( P(\hat{c}_i | \Theta_G) \) is computed making use of equations (12) and (16). In a similar way the term \( P(\hat{c}_i | -\Theta_G) \) is computed making use of dual equations to (12) and (16).

**Appendix B**

In a recent paper [11] a scheme for updating the Hough accumulator using the uncertainty contribution of each edge point have been introduced. In this scheme the uncertainty (the variance) of the estimated line parameters \( \theta \) and \( \rho \) are analytically computed for each edge point based on image noise, edge orientation estimation and parametric line representation. The line parameters uncertainty is used for evaluating the joint density distribution \( p(\hat{\theta}_v | \Theta_v) \): the likelihood of the all quantized values \( \Theta_v = (\theta, \rho) \) given the observed line parameters \( \hat{\theta}_v = (\hat{\theta}, \hat{\rho}) \). Authors have proposed to assume \( \hat{\theta}_v \) distributed as \( \hat{\theta}_v \sim N(\Theta_v, \Sigma_{\hat{\theta}_v}) \), where \( \Sigma_{\hat{\theta}_v} \) is the covariance matrix of \( \hat{\theta}_v \), and to increment the accumulator by the sum of the \( \log(p(\hat{\theta}_v | \Theta_v)) \) at each edge point. Here is proposed a new Bayesian approach for incrementing the Hough space and for computing the joint probability of \( \rho \) and \( \Theta \) for each line feature. In this scheme the uncertainty computation is based on the error propagation technique proposed by [11] but it changes for inferring the joint probability \( p(\Theta_v | \Theta_v) \) from the bivariate normal distribution

\[ p(\Theta_v | \Theta_v) = \frac{1}{2\pi\sigma_{\theta}} \exp \left\{ -\frac{1}{2} \frac{(\hat{\theta} - \theta)^2}{\sigma_{\theta}^2} \right\} \]  
(18)
The key issue of the problem is based on the property of \( \hat{\theta}_v \) to have a singular covariance matrix \( \Sigma_{\hat{\theta}_v} \). If the covariance matrix is singular, then \( \hat{\theta}_v \) is a singular or degenerate bivariate normal distribution. This mean that the probability density for \( \hat{\theta}_v \) is always concentrated in a subspace whose dimension is smaller than that of the space generated by \( \hat{\theta}_v \), therefore the probability density distribution \( p(\hat{\theta}_v | \Theta_v) \) cannot be directly computed. Fortunately it is possible to demonstrate, by using some properties of the bivariate joint distribution, that it is possible to transform equation (18) in the normal distribution of only one of its two random variables. In particular it is possible to reduce equation (18) as follows

\[ p(\hat{\theta}_v | \Theta_v) = \frac{1}{\sqrt{2\pi\sigma_{\theta}}} \exp \left\{ -\frac{1}{2} \frac{(\hat{\theta} - \theta)^2}{\sigma_{\theta}^2} \right\} \]  
(19)

By integrating equation (19) the probability that the edge pixel belong to the line \( \Theta \) is reduced to equation (4).

**References**


