Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT
Signal Processing 86 (2006) 2304–2317
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Optimal estimation of line segments in noisy lidar data

Andreas Kapp[a,*], Lutz Gröll[b]

[a] *University of Karlsruhe (TH), Institute of Measurement and Control, 76128 Karlsruhe, Germany*
[b] *Forschungszentrum Karlsruhe, Institute for Applied Computer Science, P.O. Box 3640, 76021 Karlsruhe, Germany*

Received 28 February 2005; accepted 17 October 2005
Available online 6 December 2005

Abstract

Lidar data usually is obtained by independently measuring distance r and angle φ. Therefore, measurements of r and φ are statistically independent. However, in most approaches measurements in x and y are assumed to be uncorrelated thus not taking properly into account the noise characteristic.

This article investigates the application of least squares (LS), total least squares (TLS), mixed-LS–TLS (MTLS), structured total least norm (STLN) and maximum-likelihood (ML) estimators to the problem of estimating line segments in noisy lidar data and compares their performance from a theoretical point of view. This analysis is supported by simulation results. A new approach of estimating an arbitrary line segment without the need of parametric constraints is proposed.

Keywords: Least squares; Total least squares; Maximum likelihood; Optimal estimation; Line fitting

1. Introduction

Sensors such as lidar, radar and sonar can be used to build maps of the environment required for navigation, cartography and tracking.

Fig. 1 depicts a typical constellation: within the sensor's field of view there is a car and a wall. All objects are assumed to be unmoved. The contours of both objects can be approximated by straight line segments.

In many cases it is sufficient to model the (two-dimensional (2D)) environment piece-wise by straight lines. The parameters of these lines have to be extracted from the sensor data in a somehow optimal way.

Most often TLS or orthogonal regression and their variants are used to estimate lines in noisy lidar data [1–3]. Wijesoma [4] proposes a robust Eigenvector technique which is very similar to TLS. Refs. [5,6] apply the Hough transform in a slightly modified form which can cope with outliers. Crowley [7] uses a kind of weighted orthogonal regression with the weights derived from the covariance matrix of the linearized error model.

However, apart from [7] these techniques do not take into account the special noise characteristic of the data given by lidar, radar or sonar. These sensors measure distance and angle independently. This leads to the polar coordinates of the data being statistically independent while the Cartesian coordinates are statistically dependent (see Section 3).

*Corresponding author. Tel.: +49 721 608 3794; fax: +49 721 661 874.
E-mail addresses: kapp@mrt.uka.de (A. Kapp), Lutz.Groell@iai.fzk.de (L. Gröll).

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2005.10.016

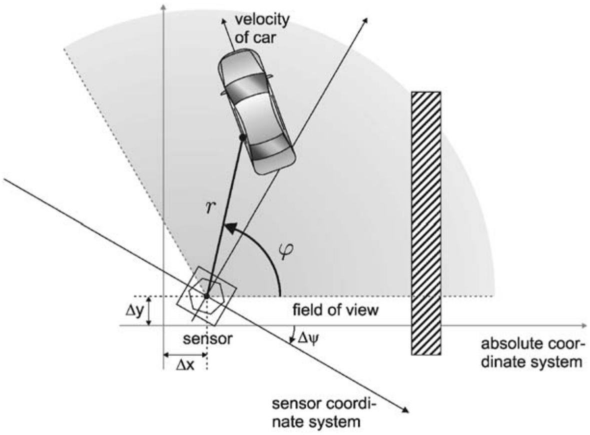

Fig. 1. Measurement configuration.

Therefore, estimators specially designed to take into account the statistical dependence outperform the approaches listed above.

Our approach follows the idea of first modeling the sensor's error and then deriving an estimator perfectly suited to this model. It is very similar to the one given by Pfister et al. [8]. They take into account both errors in distance and angle. Then they linearize with respect to (w.r.t.) deviations of the angle since these deviations are assumed to be very small. They compute each point's uncertainty based on this linearized error model and use these uncertainties as weights for an ML formulation.

In contrast to Pfister's approach we start our consideration by modeling the radial errors in an ML formulation. We neglect the errors in angle completely since they are very small. Thus, the problem becomes more pleasant since this simplification erases many local minima. An advantage of Pfister's work lies in its choice of parametrization of the line since all lines are possible. Our approach cannot cope directly with all lines due to the parametrization chosen. But we can overcome this limitation by rotating the coordinate system before applying the ML estimator which does not affect the optimality of the estimate.

The article is organized as follows: in Section 2, we give a short survey of possible parametrizations for lines highlighting their advantages and disadvantages. Section 3 introduces the noise model. Section 4 describes the different approaches. In Section 5, we give statistical properties of the LS estimator and the ML estimator. Simulation results in Section 6 and the conclusions in Section 7 finish this article. All proofs are given in Appendix A.

In the following all coordinates are given in the sensor's coordinate system since there is no need for global coordinates in this context. However, when both translation (Δx, Δy) and rotation $\Delta\psi$ between sensor and global coordinates are known it is easy to perform this transformation.

2. Parametrizations of line segments

The task of estimating line segments naturally falls into two parts:

(1) Estimation of the line's parameters.
(2) Estimation or determination of the endpoints.

Following the noise characteristic given in Section 3 it is obvious to get the endpoints by cutting the estimated line with the outermost (radial) rays since in many cases noise on the angle can be neglected compared to noise on the radius, respectively (resp.) distance. Therefore, we concentrate in the following on the estimation of the line's parameters and consider the task of determining the endpoints afterwards as solved. Since a line in 2D has a total of two degrees of freedom it can be described by a two-parametric model or a model with more than two parameters and additional constraints. The difficulty is to parametrize every possible line within the model chosen. We now give common parametrizations in Cartesian coordinate systems:

(I) *Explicit form*: $y = \theta_1 x + \theta_2$.
(II) *Hessian normal form*: $x\cos\theta_1 + y\sin\theta_1 - \theta_2 = 0$.
(III) *Constrained Hessian normal form*: $\theta_1 x + \theta_2 y - \theta_3 = 0$, s.t. $\theta_1^2 + \theta_2^2 = 1$.
(IV) *General form with normalization*: $\theta_1 x + \theta_2 y + \theta_3 = 0$, s.t. $\theta_1^2 + \theta_2^2 + \theta_3^2 = 1$.
(V) For the sake of completeness we list other common parametrizations of straight lines: Two-point form, point-slope form, intercept form and point-vector form, see [9] for details.

Model I is the most common one (Fig. 2). It cannot cope with lines parallel to the y-axis since θ_1 would be infinite. Model II describes all possible lines using only two parameters but is nonlinear in nature. Model III avoids nonlinearity by extension to three parameters and a quadratical constraint. In the case of Model II this constraint is fulfilled implicitly since $\cos^2\theta_1 + \sin^2\theta_1 = 1$ holds. Model IV is also a linear