INNOVATIVE MECHANICAL ENGINEERING **ISSN 2812-9229 (Online)**
University of Niš, Faculty of Mechanical Engineering
VOL.2, NO 1, 2023, PP. 95 - 103

Original scientific paper *

DISTANCE ESTIMATION MODEL FOR THERMAL VISION SYSTEMS USING GAUSSIAN PROCESS REGRESSION

Emina Petrović[1], Ivan Ćirić[1], Milan Pavlović[2], Vlastimir Nikolić[1]

[1] Faculty of Mechanical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia

[2] Academy of Technical and Educational Vocational Studies - Section Niš, Aleksandra Medvedeva 20, 18000 Niš, Serbia,

Abstract. *An obstacle detection system (ODS) that can operate in a challenging environment and with limited visibility is the crucial element of autonomous systems. Distance estimation of the objects (obstacles, humans) plays a very important role in each ODS. In this paper, to estimate the distance between the camera and objects, a Gaussian processes regression (GPR) model was proposed. GPR is a machine learning method that is based on Bayesian and statistical learning theory. GPR is an effective method for processing data and predicting/estimating information that is adaptable to complex regression problems such as high dimensions, small samples, and non-linearity. The proposed model's data were collected for training and testing using the Smart Automation of Railway Transport (SMART) onboard obstacle detection system developed to achieve autonomous train operation (ATO). The bounding box features of the recognized object are used as input data, while output data was obtained by measuring distances form the camera and humans involved in the experiment. The GPR results are compared with measured distances and distance estimations obtained by image-plane homography. The proposed method was tested with the thermal camera and in impaired visibility scenario, but the presented methodology could be applied for various vision sensors.*

Key words: *Gaussian process regression, Thermal camera, Distance estimation, Obstacle detection system (ODS), Impaired visibility condition*

1. INTRODUCTION

Distance estimation of objects (obstacles, humans) is essential in many applications, including mobile robotics, autonomous vehicles, and Intelligent Transportation Systems (ITS). The main component of autonomous systems is a machine vision system with

***Received: January 21, 2023 / Accepted May 18, 2023.**
Corresponding author: Emina Petrović
Faculty of Mechanical Engineering, University of Niš, Serbia
emina.petrovic @masfak.ni.ac.rs

hardware and software solutions that can detect obstacles. This Obstacle Detection System (ODS) should be able to work in a challenging environment with low visibility [1].

One thing that all ODSs have in common is that they typically work best during daylight hours. Different types of illumination, dim lighting, or poor visibility brought on by bad weather significantly reduce or, in some cases, fail to produce satisfactory detection results. Due to its operating range being in the invisible infrared region of the spectrum, a thermal imaging system can be used to deal with these extremely specific conditions. Thermal imaging systems have not been the subject of many studies [2]. The techniques currently in use for distinguishing between living and non-living objects with thermal imaging systems are still under development. When a thermal imaging system is mounted on a moving platform or vehicle, the situation is even worse.

In [3], a sophisticated control of a mobile robot for human recognition in an indoor setting where infrared (IR) cameras are used to take pictures, and each frame includes intelligent segmentation and classification of any regions of interest that are found.

A method for real-time human detection using an autonomous mobile platform with an IR camera mounted on, is presented in [4], while in [5], a stereo system for detecting pedestrians that is mounted on a test vehicle with two far-infrared cameras is described.

Broggi et. al [6], presented a system for the detection of pedestrians that is equipped with a monocular IR camera mounted on a test vehicle, a car. To find pedestrians, they used multi-resolution localization of warm, symmetrical objects of defined size and aspect ratio. The proposed system was successful in detecting one or more pedestrians in experiments, but only when they were within a range of 7 to 43.5 meters.

Forth et. al [7] presented a thermal camera that is mounted on the train's roof with the intention of analyzing a potentially dangerous scenario that could arise while trains are operating at night or in adverse weather, particularly when some objects are placed on rail tracks.

In [8], an artificial intelligence approach for long range object (obstacles) distance estimation that improve estimation of a distance between the camera and an imaged object using image-plane homography is presented. The technique makes use of an artificial neural network that lowers the estimation error based on experimental data collected, as well as the homography between two planes - the image plane and the rail tracks plane.

In this study, the results for distance estimation obtained using the thermal camera's image-plane homography and those obtained using Gaussian processes regression are compared (GPR). The GPR training data correspond to those from [1, 2, 8]. The bounding box characteristics of the identified object are used as input data, and distance measurements between the camera and the test objects are used to obtain the output data.

2. Gaussian Process Regression

Vision sensors can generate massive observational data. The problems that may occur include the missing data, outliers excluding, perceived sensing failure, etc. Gaussian Process Regression (GPR) method is an effective method to process data and predict information even in cases when the observational data are in the form of nonstationary time series [9].

Gaussian regression process (GRP) is a new machine learning method based on Bayesian and statistical learning theory [10]. Gaussian processes became popular in the fields of machine learning and data analysis due to their flexibility and robustness [11]. They are already successfully applied to a wide range of applications, such as modelling dynamical systems [12], latent models for dimensionality reduction [13], prediction [14], or visual tracking [15].

The GRP is often called a nonparametric model, because it does not rely on a finite set of parameters meaning that it is not bounded by any single function, instead it can calculate a probability distribution over every possible function that may fit the input data [16].

In the function-space view, the Gaussian process is defined as any set of finite random variables which have a joint (multivariate) Gaussian distribution, whose properties are determined by the mean function $\mu(\boldsymbol{x})$ and covariance function $k(\boldsymbol{x}, \boldsymbol{x}')$.

$$f(x) \sim GP(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \tag{1}$$

The covariance function k (**x**, **x'**) is also known as "kernel function" and can be expressed in several differential forms. The kernel is parameterized by a set of kernel parameters commonly known as hyperparameters [10].

One of the main uses of the previous model is to predict $f(\bar{x}) = \bar{f}$, with some estimated confidence, for some new input data point $\bar{x}$. This is equivalent to calculating the conditional distribution $\bar{f}|\boldsymbol{x}, \bar{x}, \boldsymbol{y}$.

For given data set $D = \{(x_i, y_i) \mid i = 1, \dots, n\}$, where $\boldsymbol{x} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \in \mathbb{R}^d$ is matrix of input data, $\boldsymbol{y} = \{\boldsymbol{y}_1, \dots, \boldsymbol{y}_n\} \in \mathbb{R}$ is output data vector, and d is dimension of the input data, it is easy to show that the conditional distribution (likelihood) of $\boldsymbol{y}$ is given by

$$\boldsymbol{y} \sim N(\boldsymbol{m}(\boldsymbol{x}), \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}) + \sigma_n^2 \boldsymbol{I}_n) \tag{2}$$

Where $\boldsymbol{m}(x)$ is the mean vector, $\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x})$ is the corresponding covariance matrix, σ_n^2 is the positive noise covariance and $\boldsymbol{I}_n$ is an identity matrix.

The conditional distribution of predicted function value conditioned on observed training outputs $\boldsymbol{y}$ is expressed as:

$$\bar{f}|\boldsymbol{x}, \bar{x}, \boldsymbol{y} \sim N\left(\bar{f}^*, \ \bar{\sigma}^2\right) \tag{3}$$

where the predicted mean $\bar{f}^*$ and covariance $\bar{\sigma}^2$ of the predicted value $\bar{f}$ at the test point $\bar{x}$ are given by:

$$\bar{f}^* = \boldsymbol{K}(\boldsymbol{x}, \bar{x})[\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}) + \sigma_n^2 \boldsymbol{I}_n]^{-1} \boldsymbol{y} \tag{4}$$

$$\bar{\sigma}^2 = \boldsymbol{K}(\bar{x}, \bar{x}) - \boldsymbol{K}(\boldsymbol{x}, \bar{x})[\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}) \sigma_n^2 \boldsymbol{I}_n]^{-1} \boldsymbol{K}(\boldsymbol{x}, \bar{x})^T \tag{5}$$

$\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x})$ denotes the $n \times n$ symmetric positive definite covariance matrix evaluated at all pairs of training points $\boldsymbol{x}$. $\boldsymbol{K}(\boldsymbol{x}, \bar{x})$ denotes the covariance matrix of the new input test point $\boldsymbol{x}^*$ and all the training points $\boldsymbol{x}$, and $\boldsymbol{K}(\bar{x}, \bar{x})$ denotes the covariance matrix evaluated at the test point $\bar{x}$.

One can notice that the Gaussian process regression is defined by two functions: the mean function and the covariance function.

The mean function generates the expected value, and the covariance function defines how much the data are smoothed in estimating the unknown function given that the variables included in the model come from a joint, multivariate Gaussian distribution [14].

As already mentioned, the reliability of Gaussian process regression is dependent on how well we select the covariance function - kernel, so it is necessary to optimize the values of the hyperparameters in kernel, $\boldsymbol{\theta}$.

Based on Bayes' theorem, the hyperparameters are usually computed by maximizing the marginal likelihood $\log p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})$.

$$\log p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta}) = -\frac{1}{2}\boldsymbol{y}^T[\boldsymbol{K}(\boldsymbol{x},\boldsymbol{x}) + \sigma_n^2]^{-\mathbf{1}}\boldsymbol{y} - \frac{1}{2}\log|\boldsymbol{K}(\boldsymbol{x},\boldsymbol{x}) + \sigma_n^2| - \frac{n}{2}\log 2\pi \quad (6)$$

3. DATA ACQUISITION

As mentioned earlier, in this paper we used GPR to improve distance estimation results obtained by image-plane homography with the thermal camera presented in [1, 2]. The distance estimation results presented in [2] were used as a starting point for the research presented in this paper.

In [1, 2] the image-plane homography is used to detect obstacles in railways. The homography was used for mapping the image plane and the corresponding world plane. As the result of this mapping, the world 3D coordinates of each point in the imaged world plane can be calculated. The data acquisition for training and testing was done using the SMART onboard obstacle detection system for autonomous train operation [17].

The problems that may occur, for example, the image distortion, or the fact that rail tracks can have up to 2% inclination and declination, cause that the accuracy of distance estimation varies. Also, imprecise coordinates of the object bounding box recognized on or near the rail tracks further affect the accuracy of object's distance estimation.

Therefore, Gaussian processes regression (GPR) model was used because can make predictions/estimations incorporating prior knowledge based on data set with varying, inhomogeneous noise variance, and it can provide measurement uncertainty over predictions.

The experiment was done on a static test-stand on a level crossing to view the straight rail tracks in the length of about 1200 m. The thermal camera was used to record the rail tracks scene with the objects in low-light (night) conditions. For the purposes of the experiment, humans acted as moving obstacles near the rail tracks. They were positioned on some distances from the camera stand that were pre-measured and marked on the rail track using both a laser distance meter and GPS. During the experiment, humans imitated static obstacles on the rail tracks located at the distances of 950m, 940m, 930m etc., up to 50m from the camera test-stand. The field tests were performed on a Serbian railway test-site (at the location of Babin potok village, near the city of Prokuplje) approved for the experimental use by the Serbian Railway authorities.

To prepare distance estimation training and testing data, the raw experimental data was processed using image processing algorithms developed and tested in [1, 2] The bounding box of the detected object of the recognized object are used as an input data, while output data were obtained by measuring distances form camera and objects involved in the experiment.

4. RESULTS

The data set prepared using the experimental data explained in the previous chapter was used for training, validation, and testing of the GPR model. The data set contained a set of 273 samples.

The estimation results achieved by the homography based method are obtained using X_L and Y_L coordinates in pixels of the centre of the bounding box edge that is in the rail track plane in the real world.

“Obstacles” was detected every 10 metres on a range from 50m to 950m. The measured distances form camera and objects present the target values y.

In this research, along with the X_L and Y_L coordinates, the bounding box width w and height h are also used as inputs.

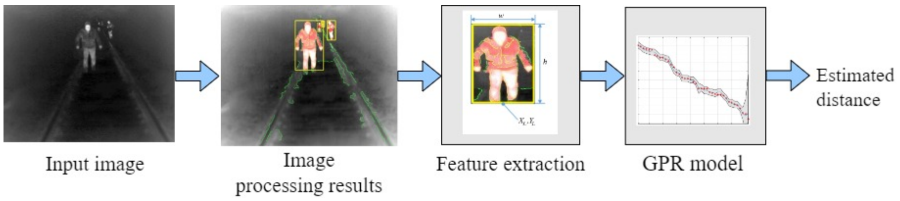

Fig. 1 Workflow of proposed system

The 80% of data (232 samples) were randomly taken for training and model validation, while the 20% of data (41 samples) were used for testing.

The training set was used to fit and estimate prediction error for model selection (validation), while test set was used to assess the generalization error i.e. how accurately the model can predict outcome values for previously unseen data.

As mentioned earlier, Gaussian process regression is defined by two functions: the mean function and the covariance (kernel) function.

In this research, the mean function in this model was set to be a constant value.

The covariance function used is so called “squared exponential” kernel function. The squared exponential kernel is defined as:

$$k(x_i, x_j|\theta) = \sigma_f^2 exp\left[\frac{-(x_i - x_j)^T(x_i - x_j)}{2\sigma_l^2}\right] \quad (7)$$

Where, x_i and x_j are values in the input space, σ_f and σ_l denotes the parameters of the kernel function - the signal variance and the length scale.

Cross validation was done by splitting the training data into 5 randomly chosen subsets (or folds) of roughly equal size. One of these subsets is used for validation of the model trained using the rest of the subsets. This process is repeated 5 times so that each subset is used exactly once for validation.

Optimization of the parameters that control the GPR algorithm’s behaviour (hyperparameters) was done using Bayesian optimization. The optimized hyperparameter for this GP was found to be $\boldsymbol{\theta} = [\sigma_l \quad \sigma_f] = [26.9509 \quad 226.8127]$.

To compare the performance of the obtained model and homography based estimated distance [1], Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2) were used.

These criteria can be calculated by the following equations:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (9)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (10)$$

Where y_i is actual output, $\hat{y}_i$ is predicted output, and $\bar{y}$ represents the mean of y_i. Number of samples is defined as n.

R^2 presents a measure of how close the data is to the fitted regression line. The closer values of R^2 are to 1, the better the proposed regression model is.

The performance metrics are calculated for both, homography and GPR based estimation, and the results are presented in Table 1.

Table 1 Performance comparation for Homography [1,2] and GPR based distance estimation methods

	MSE	MAE	RMSE	R^2 score
Homografy	271500	450	521.06	0.7795
GPR	384.31	14.342	19.604	0.9954

The homography based distance estimation method has been applied on the non-calibrated thermal image.

The camera lens and sensors produce distortion. For industrial use, these distorted images need to be calibrated. The calibration can also be performed on the thermal images, but for the long-range thermal camera used, it's very expensive and complicated [1]. Therefore, thermal image distortion introduces nonlinearity that the homography method cannot deal with, but GPR provides a mechanism by which these uncertainties can be quantified thoroughly, hence this method became an attractive alternative to the more traditional approach.

Fig. 2 illustrates the predicted response of our proposed model plotted against the measured distances, while Fig. 3 shows the regression results within the 95% confidence interval.

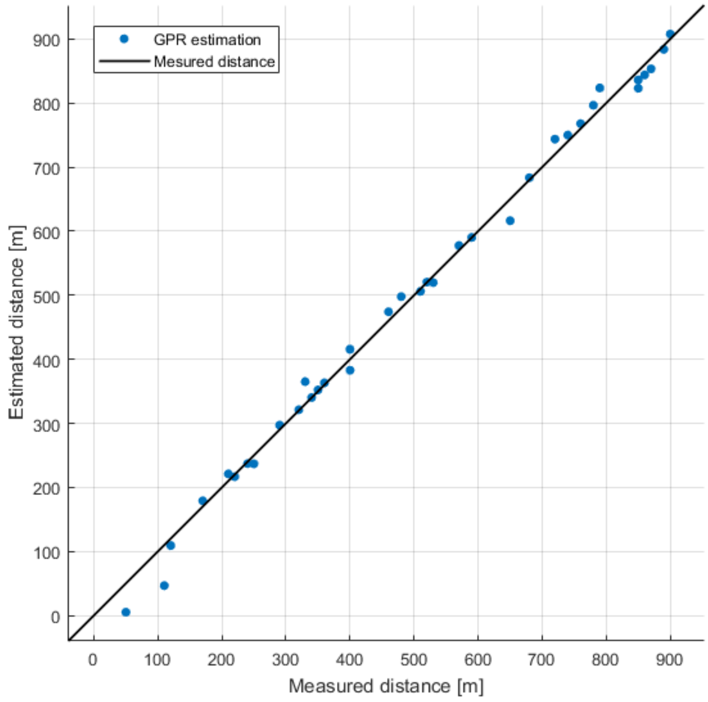

Fig. 2 Distance estimation results for the test dataset

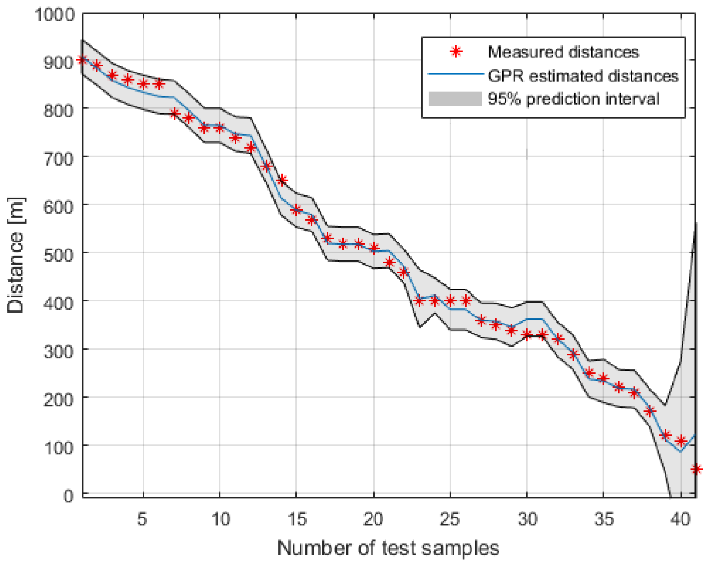

Fig. 3 The regression results with a 95% confidence interval

5. CONCLUSION

In this study, the results of object distance estimation using the Gaussian process regression are compared with those obtained using image-plane homography and measured distances. The proposed method was tested with a thermal camera and in conditions of poor visibility. The evaluation results show that GRP can significantly improve the distance estimation of thermal imaging system even when a vision system is mounted on a moving platform or vehicle. Although the presented distance estimation algorithm was tested with a thermal camera, it can be used with all vision sensors in a variety of fields where accurate perception is necessary, including robotics applications, autonomous vehicles, and Intelligent Transportation Systems.

Acknowledgement: *This research was financially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Contract No. 451-03-9/2021-14/200109). Data used in this paper were acquired during the realization of the HORIZON 2020 S2R-OC-IP5-01-2015 project "Smart Automation of Railway Transport – SMART".*

REFERENCES

1. M. G. Pavlović et al., "Advanced thermal camera-based system for object detection on rail tracks," Thermal Science, vol. 22, pp. S1551–S1561, 2018, doi: 10.2298/TSCI18S5551P.
2. M. G. Pavlović, "Application of intelligent machine vision systems for autonomous train operation," PhD thesis, 2020.
3. I. T. Ćirić et al., "Thermal vision based intelligent system for human detection and tracking in mobile robot control system," Thermal Science, vol. 20, no. suppl. 5, pp. 1553–1559, 2016, doi: 10.2298/TSCI16S5553C.
4. A. Fernández-Caballero, J. C. Castillo, J. Martínez-Cantos, and R. Martínez-Toms, "Optical flow or image subtraction in human detection from infrared camera on mobile robot," Rob Auton Syst, vol. 58, no. 12, pp. 1273–1281, Dec. 2010, doi: 10.1016/j.robot.2010.06.002.
5. M. Bertozzi, A. Broggi, C. Caraffi, M. del Rose, M. Felisa, and G. Vezzoni, "Pedestrian detection by means of far-infrared stereo vision," Computer Vision and Image Understanding, vol. 106, no. 2–3, pp. 194–204, May 2007, doi: 10.1016/J.CVIU.2006.07.016.
6. A. Broggi, A. Fascioli, M. Carletti, T. Graf, and M. Meinecke, "A multi-resolution approach for infrared vision-based pedestrian detection," IEEE Intelligent Vehicles Symposium, Proceedings, pp. 7–12, 2004, doi: 10.1109/IVS.2004.1336347.
7. F. Z. A. Forth, "Night-vision device for railway vehicles for improving safety | Eurailpress Archive." https://eurailpress-archiv.de/SingleView.aspx?show=22079&lng=en (accessed Jan. 20, 2023).
8. I. Ćirić, M. Pavlović, M. Banić, M. Simonović, and V. Nikolić, "AI Powered Obstacle Distance Estimation for Onboard Autonomous Train Operation", doi: 10.17559/TV-20210223081612.
9. L. Kang, R. S. Chen, N. Xiong, Y. C. Chen, Y. X. Hu, and C. M. Chen, "Selecting Hyper-Parameters of Gaussian Process Regression Based on Non-Inertial Particle Swarm Optimization in Internet of Things," IEEE Access, vol. 7, pp. 59504–59513, 2019, doi: 10.1109/ACCESS.2019.2913757.
10. C. E. Rasmussen and C. K. I. Williams, "Gaussian Processes for Machine Learning", Accessed: Jan. 20, 2023. [Online]. Available: www.GaussianProcess.org/gpml
11. Andreas. Svensson, T. B. Schön, Fredrik. Lindsten, C. Edward. Rasmussen, and Uppsala universitet Teknisk-naturvetenskapliga vetenskapsområdet., "Machine learning with state-space models, Gaussian processes and Monte Carlo methods".
12. M. P. Deisenroth, R. Turner, M. F. Huber, U. D. Hanebeck, and C. E. Rasmussen, "Robust Filtering and Smoothing with Gaussian Processes," IEEE Trans Automat Contr, vol. 57, no. 7, pp. 1865–1871, Mar. 2012, doi: 10.1109/TAC.2011.2179426.
13. X. Jiang, J. Gao, T. Wang, and L. Zheng, "Supervised latent linear Gaussian process latent variable model for dimensionality reduction," IEEE Trans Syst Man Cybern B Cybern, vol. 42, no. 6, pp. 1620–1632, 2012, doi: 10.1109/TSMCB.2012.2196995.

14. H. S. Grewal, M. S. Chacko, S. Ahmad, and H. Jin, "Prediction of the output factor using machine and deep learning approach in uniform scanning proton therapy," J Appl Clin Med Phys, vol. 21, no. 7, pp. 128–134, Jul. 2020, doi: 10.1002/ACM2.12899.
15. L. Zheng, M. Tang, and J. Wang, "Learning Robust Gaussian Process Regression for Visual Tracking," 2018.
16. S. Ambikasaran, D. Foreman-Mackey, L. Greengard, D. W. Hogg, and M. O'Neil, "Fast Direct Methods for Gaussian Processes," IEEE Trans Pattern Anal Mach Intell, vol. 38, no. 2, pp. 252–265, Mar. 2014, doi: 10.48550/arxiv.1403.6015.
17. http://smart.masfak.ni.ac.rs/ (accessed Jan. 20, 2023).