

Pipeline Landmark Detection for Autonomous Robot Navigation using Time-of-Flight Imagery

Jens T. Thielemann, Gøril M. Breivik and Asbjørn Berge
SINTEF ICT
P.O.Box 124 Blindern, N-0314 Oslo, Norway
{jtt,gbre,asbe}@sintef.no

Abstract

3D imaging systems provide valuable information for autonomous robot navigation based on landmark detection in pipelines. This paper presents a method for using a time-of-flight (TOF) camera for detection and tracking of pipeline features such as junctions, bends and obstacles. Feature extraction is done by fitting a cylinder to images of the pipeline. Data in captured images appear to take a conic rather than cylindrical shape, and we adjust the geometric primitive accordingly. Pixels deviating from the estimated cylinder/cone fit are grouped into blobs. Blobs fulfilling constraints on shape and stability over time are then tracked. The usefulness of TOF imagery as a source for landmark detection and tracking in pipelines is evaluated by comparison to auxiliary measurements. Experiments using a model pipeline and a prototype robot show encouraging results.

1. Introduction

Sewer systems are run in most municipalities in the world. Power plants and factories use pipelines for carrying oil and gas. Both are based on pipeline systems having the risk of being troubled by cracks, corrosion, erosion and external damage. Inspection of pipeline systems is crucial for failure detection.

With the exception of research robots, current sewer inspection systems are based on remotely operated cable-tethered robots with on board video camera systems [1]. Pipeline inspection in general can be greatly simplified using autonomous robots navigating based on information from the environment. This encourage development of navigation systems for robots in pipelines.

We are currently developing an in-pipe robot consisting of a two meter train of ten independent modules that mimics snakelike maneuverability. The final robot will be able to move vertically and horizontally through pipelines of differ-

ent diameters. Navigation of the robot will be done based on landmarks and a coarse pipeline map, thus a vision system for landmark and obstacle detection and tracking is needed.

Previous work on navigation in pipelines has applied a variety of sensors. Ahrary et al [1] describes a method based on stereo camera and laser scanner data, doing distance measurement by stereo matching. Stereo cameras are also used in the MAKRO project [10]. The autonomous robot KURT [6] does landmark detection based on ultrasound data in order to recognize different types of pipe crossings.

Stereo cameras introduce physical restrictions on the robot due to the need for camera separation. Further, stereo cameras depend on feature matching from both cameras for range estimation. This produces a rather sparse and unevenly distributed data set. Laser range sensors measure distances at a coarse grid across the range sensor field-of-view, also providing sparse data sets. A drawback of ultrasound is the limitation in resolution.

Most autonomous robot navigation systems attempt to solve two simultaneous tasks; exploring an unknown environment by building a map, and determining the robot position in this map. This process is known as simultaneous localization and mapping (SLAM). SLAM has been applied to both indoor [5] and outdoor [8] navigation. Despite significant progress in the area, challenges still remain according to recent surveys, see [3] and [11]. However, pipelines are very constrained environments, and a coarse predefined map of pipeline landmarks coupled with a landmark detection method is deemed sufficient for robot localization and navigation.

Time-of-flight (TOF) cameras are a promising source of 3D imagery for landmark and obstacle detection in pipelines. The recently developed TOF cameras have a great potential for mobile robot applications. They have been used for mobile robotics mapping and navigation in real world environments [7], and obstacle detection and segmentation for navigation in factory-like environments [2]. Applications of TOF cameras in pipeline environments have

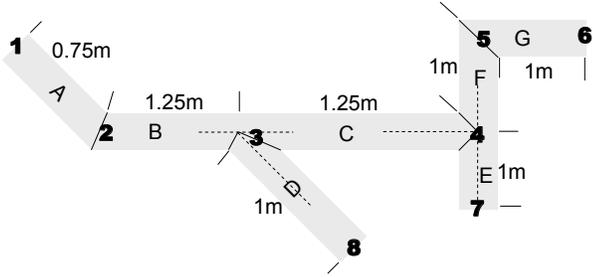


Figure 1: Sketch of the pipeline system used for experiments. The robot starts at waypoint 1 and moves to waypoint 4. Waypoint 2 is a 45° bend, waypoint 3 is a Y-junction, waypoint 4 is a T-junction and waypoint 5 is a 90° bend.

to our knowledge not been published in the literature.

In this paper, we study the application of TOF cameras in pipelines and evaluate its usefulness as a sensor for landmark and obstacle detection. We hypothesize that significant deviations from the idealized, cylindrical shape of a pipeline either are landmarks like branches and junctions, or obstacles like internal deposits. We also evaluate using a cone as an alternative geometric primitive for deviation measurements. Ranges to detected landmarks are validated by comparison to auxiliary measurements.

To evaluate the method experimentally, a robot prototype was built using a Lego Mindstorms NXT set. The TOF camera, a Mesa SwissRanger SR-3000, was mounted on the front of the robot, tilted slightly upward. A pipeline model consisting of 40 cm diameter sewer pipelines was used for experiments. Figure 1 shows a sketch of the pipeline system, including 45° and 90° bends plus Y- and T-junctions. During experiments reported in this paper the robot moved from waypoint 1 to 4.

This paper is structured as follows. Our proposed method for landmark detection and tracking is presented in section 2. Results are evaluated and related to auxiliary measurements in section 3. Section 4 concludes the paper and gives a brief outlook on the obtained results.

2. Methodology

The proposed method for feature extraction is based on first fitting a cylinder or cone to the measured 3D data points, and subsequently detecting deviations between the fitted geometric primitive and the data at hand. The cylinder/cone fitting is done partly to ease later feature detection, partly to create a reference coordinate system in which feature location is easily interpretable.

2.1. Cylinder fitting

We consider it reasonable to believe that the view-direction of the camera in most cases will be rather parallel to the cylinder axis. Therefore, we have chosen to parameterize the cylinder to be fitted by the point where the cylinder axis intersects the camera’s focal plane, and the point where it intersects a second plane parallel to the focal plane one meter ahead of the camera. The two intersection points are specified in planar coordinates, and with radius this becomes five parameters. This parameterization has the benefit having few continuous parameters of the same scale, which is beneficial for later optimization.

The cost function used for model fitting is inspired by the MLESAC [12] criterion. This criterion is designed to reward good coherence between model and data, whilst simultaneously ignoring gross outliers coming from e.g. measurements of a wall in the end of a pipeline. The criterion is given by

$$R = \sum_i \rho(d_i) \quad \rho(e) = \begin{cases} e^2 & e^2 < t^2 \\ t^2 & e^2 > t^2 \end{cases} \quad (1)$$

where R is our cost function, d_i is the squared distance between data point i and the closest point on the cylinder candidate, and t is an acceptable deviation threshold for gross outliers.

Initial experiments indicate that measurement points that are either far away from the camera or have low signal amplitude are probably incorrect. When evaluating the cost function, such points are ignored by using fixed amplitude and distance thresholds to suppress them.

The optimization is done using the Nelder-Mead method [9]. This is initialized the first time using a cylinder axis coinciding with the camera axis, and subsequently by reusing the result from the previous cylinder fitting. The parameter reuse provides additional stability and faster convergence during fitting.

2.2. Cone fitting

Initial experiments using cylinder fitting encourage a model adjustment, as the acquired data within the range of interest appear to take a more conic than cylindrical shape. The cone fitting model is based on the same principles as the cylinder fitting, with the exception that an additional parameter specifying cone steepness also is estimated.

Experiments show that the cone model can lead to unlikely cone configurations in cases where the view of the pipe is very limited. A sanity check on cone steepness is therefore done on the cone parameters. In case of rejection, the cylinder parameters are used for further analysis.

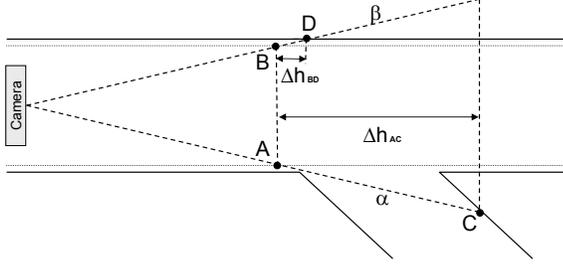


Figure 2: Illustration of feature detection process. Thick lines: Actual data. Thin lines: Fitted cylinder/cone. α, β : Viewlines/pixels. A, B : Expected intersection points between viewlines α, β and fitted cylinder/cone. C, D : Actual data points along viewlines. Viewline α is detected as a possible feature candidate due to the large distance Δh_{AC} between A and C . Viewline β is classified as not being a feature candidate due to the short distance Δh_{BD} between B and D .

2.3. Comparison with synthetic image

The estimated cone parameters are used for synthesizing an image representing the camera view expected if the current view of the pipe is featureless. The synthesization is done by first estimating unit vectors specifying the view direction of each camera pixel, and subsequently calculating the intersection point between these vectors and the fitted cone.

For ease of reporting later feature position, we have chosen to use cylindrical coordinates (θ, r, h) , describing angle, radius and range, for further processing. Origo is defined as the intersection of cylinder axis and the camera's focal plane.

The synthesised image is compared with the actual data, using the h -component of the cylindrical coordinates as basis for comparison, see Figure 2. Using the h -coordinate was in initial experiments found to give better results than e.g using the r -coordinate. Pixel regions with sufficiently large deviation between model and actual data are assumed to be possible landmarks or obstacles and are processed further.

2.4. Feature extraction

Pixels in the camera data that deviate more than a given threshold from the synthetic cone image are segmented out and grouped into blobs. These blobs are filtered for size and shape using blob area and compactness, defined by $\text{Perimeter}^2 / 4\pi \text{Area}$, to remove unlikely candidates, since most of the artifacts are assumed to be smaller and have a lower compactness than the landmarks.

In order for the detected features to become useful for a higher-level navigation system, the algorithm needs to re-

port the distance and position of the features. This is done by picking the blob pixel from the synthetic cone image that has the lowest h -coordinate, and using the pixel to determine the feature range and position. Using this pixel to specify the feature detection point, is considered by us less ambiguous than possible alternatives. For instance, instead using the mean distance to the detected object will shift the estimated feature position as the feature gradually comes more into view, making measurements less precise.

2.5. Feature tracking

The range to each detected feature is tracked over time, partly in order to remove spurious features. Features are grouped over time by selecting the closest neighbor in (h, θ) -space for each detection point found in subsequent images. The r -component is not taken into account, as this parameter does not provide any additional information on changes over time in a pipeline of constant diameter. Features appearing in less than five consecutive images are considered outliers.

2.6. Comparison with speed measurements

In order to determine feature detection and tracking ability, we compare our results with manual speed measurements. We estimate the speed of the robot during its traversal by measuring the distance between junctions and the traversal time between these junctions. As the robot has constant motor power we further assume constant speed between junctions.

3. Experimental results and discussion

Range images obtained in the illustrated pipe model using the prototype robot was segmented by the described method. A simple landmark detector in the segmented images was used to track distance to landmark over time. Tracking results were compared to auxiliary timing of the robot path through the pipeline.

3.1. Cylinder model versus cone model

Experiments were performed to decide the best model for description of the range data. The initial assumption was that recorded data would correspond to a cylinder. By visual inspection of recorded range data we observed that the images had some defects giving a field taking a slightly conic shape.

Range images were recorded in a pipeline without any junctions or obstacles. Thus, one would expect that all observed defects in the recorded data are due to secondary reflections from the pipeline, lens scattering, poor camera calibration and further camera imperfections. Figure 3 shows a plot of the captured range image, seen from above. Here,

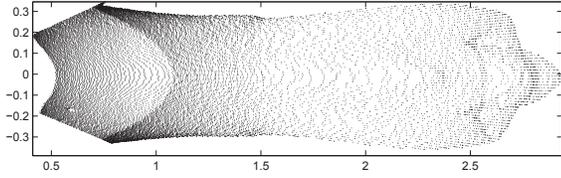


Figure 3: TOF camera image of a pipeline without junctions or obstacles, seen from above. All units in meters. The density of range measurements decreases markedly at a distance of 1.5 meters. Note the conic shape of the range data, discernible from the camera and up to 1.25 meters. Note also that the measurements indicate a pipeline diameter of 60 cm, while the pipeline diameter used in the experiment was 40 cm. The defects in the recorded data are assumed to be due to poor camera calibration, secondary reflections, lens scattering and further camera imperfections.

the recording defects are evident, as the pipeline diameter was 40 cm, whilst measurements indicate a pipeline diameter up to 60 cm.

Inspection of the range data reveals that the data initially narrows slightly with increasing range. Note also that the density of range measurements decreases with increasing distance from the camera, limiting the useful distance range. From the camera and up to a range of 1.25 meters a conic shape can be clearly discerned.

Both the cylinder model and the cone model were fit to the data captured. During data fitting, data points with an intensity below 3000 and a range beyond 1.5 meters were ignored. The hard threshold t^2 for outlier data in the optimization criterion in Equation 1 was set to one cm in both cases.

Illustrated in Figure 4a is the best cylinder fit for the measured data points. The cylinder fit points slightly off axis. Visual inspection of Figure 4b indicates that the cone model is a better fit for the region of interest up to one meter.

In particular, the cone model is much better able to estimate the correct pipeline axis when landmarks are present. Landmarks cause a lack of points on at least one side of the pipeline, resulting in a cylinder aligned only along the opposite pipeline wall. Due to the conic shape of the measurements, the alignment will be off-axis. The cone model has an additional degree of freedom that better represents the captured data, and thus allows for more accurate axis alignment even in the presence of landmarks.

3.2. Extraction of Y- and T-junction landmarks

Deviations from the cone model of the range image were used for segmentation and tracking of landmarks. Reported from the feature tracker is distance to all plausible landmarks in the frame.

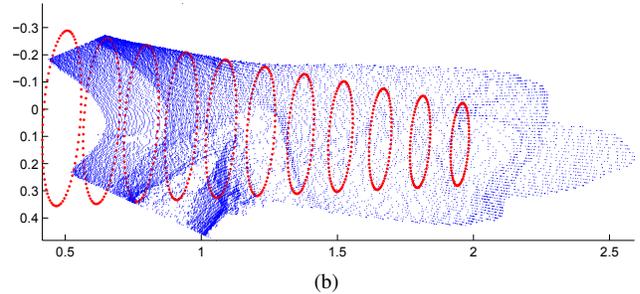
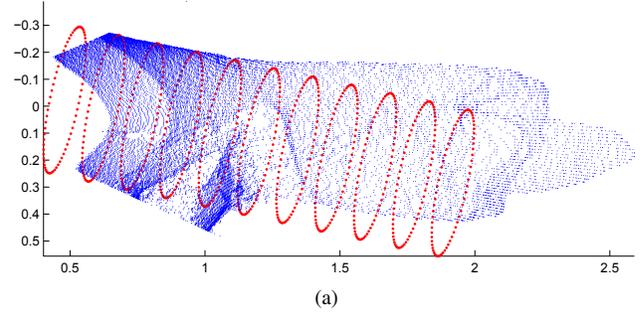
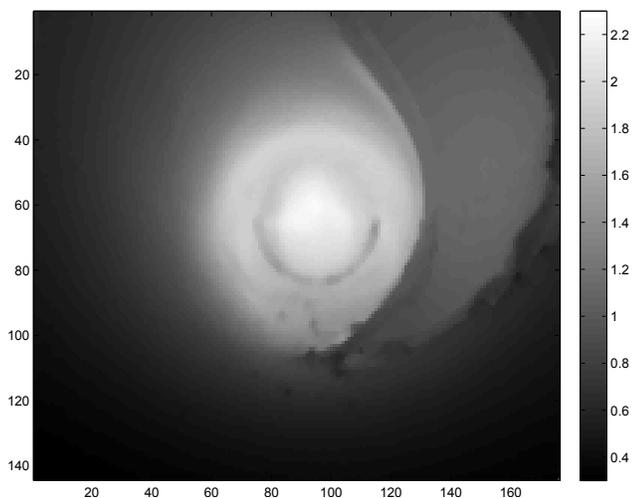


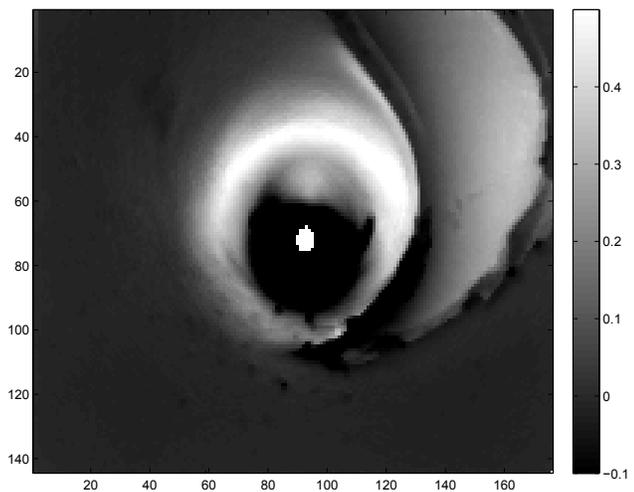
Figure 4: Cylinder (a) and cone (b) fit to the TOF range image of pipeline with opening, top-view. Camera positioned at B in Figure 1, looking towards C. Opening towards pipe D visible at approximately 0.75-1.25 m distance. All units in meters. Fitted model indicated with solid points. Only data up to one meter is used for further deviation measurements. Unevenly distributed data measurement points result in a best fitting cylinder slightly off-axis. The cone model handles this better and estimates the pipeline axis more accurately.

In the experiment, TOF images were captured at a rate of 13 frames per second while the robot was following a pre-programmed path. Speed varied during the experiment, but was assumed constant in the part where the robot was moving toward the Y- and T-junctions. Hence, the measured distance to detected landmarks should decrease roughly linearly with time.

The cone model was fit to the range in each frame in the video sequence. If the estimated cone converged to a point closer than two meters ahead of the camera, a fitted cylinder model was used instead. The data set was subsampled by a factor ten to speed up processing. As illustrated in Figure 2, deviations were measured as the projected distance of the observed range onto the fitted cone axis. Deviations above 20 cm from the cone fit were classified as landmark candidates, and segmented out. The feature was classified as a landmark if the pixel area of the deviating segment totaled more than 300 pixels and its compactness was less than four. Landmarks more than one meter from the camera were ignored. These parameters were derived experimentally.



(a) Raw TOF range image of Y-junction, indicating measured distance, i.e. the h -coordinate, in each pixel in meters. Observe that the landmark is clearly discernible in the right of the image.

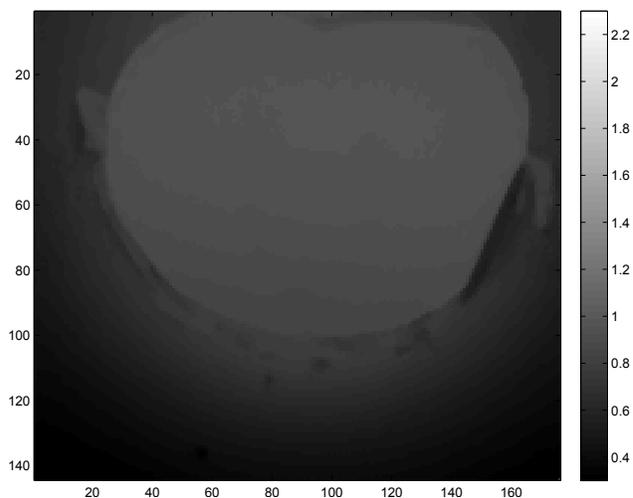


(b) Deviation measure image of Y-junction obtained by subtracting expected range measurements of a synthesized cone from the raw TOF range image measurements. The Y-junction in the right of the image clearly deviates from the fitted cone.

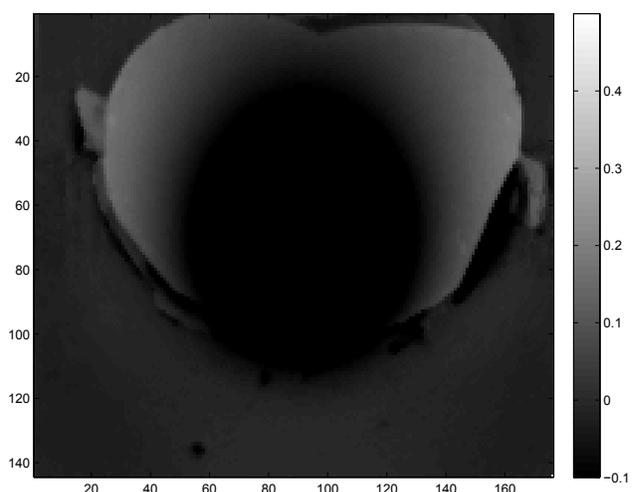
Figure 5: Distance and deviation measure images of the Y-junction.

A range image showing the Y-junction can be seen in Figure 5a. In each pixel measured distance, i.e. the h -coordinate, is indicated in meters. The junction is clearly distinguishable in the right part of the image. By subtracting the cone fit from the observed data, a deviation measure for each pixel can be obtained. This is shown in Figure 5b. For pixels where we observe the pipeline walls, deviations would be expected to be close to zero.

Figure 6a illustrates the range image of the T-junction as seen from a distance of approximately 0.7 meters. The



(a) Raw TOF range image of T-junction, from a distance of roughly 0.7 meters, indicating measured distance, i.e. the h -coordinate, in each pixel in meters. Note that the crossing pipe wall can be observed as a large area with a distance of roughly one meter in the center of the image.



(b) Deviation measure image of the T-junction, obtained by subtracting expected range measurements of a synthesized cone from the raw TOF range image measurements. The left and right turns of the pipe are easily discernible. Note the occurrence of negative deviation in the center of the image, this is due to the cone model penetrating the crossing pipe wall.

Figure 6: Distance and deviation measure images of the T-junction.

crossing pipe wall can be seen as a large area of having roughly a distance of one meter to the camera. Comparison of the cone model to observed data gives the deviation measure image illustrated in Figure 6b. The left and right turns of the pipe are easily discernible, and can be detected and tracked in the same way as the Y-junction. Note the occurrence of negative deviation in the center of the image since the cone model penetrates the crossing pipe wall.

3.3. Tracking of Y- and T-junction landmarks

The proposed landmark tracking procedure was applied to data recorded when the robot traversed the path from waypoints 1 to 4 in Figure 1. The initial pipe segment between 1 and 2 is too short to obtain a reasonable amount of data to model the cone fit. The landmark at waypoint 2, a 45 degree bend cannot be reliably detected in the available experiments and is excluded from the reported results. Detection and modeling of bends is a topic for future work.

Around three seconds into the path, the camera is aligned along the pipeline from waypoint 2 to 4, and detection and tracking of landmarks are performed while the robot is moving. Illustrated in Figure 7a is the measured distance to the Y-junction landmark as a function of time, marked with crosses. The solid line in Figure 7a is the estimated distance to landmark, based on auxiliary speed measurement of the traversal between waypoints in the pipeline model. Positioning of the reference line with regard to feature points has been done using least squares, whilst ensuring preservall of auxiliary speed measurement.

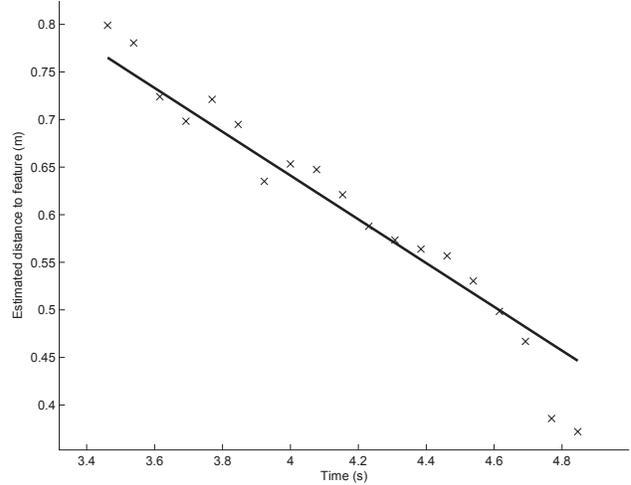
Similar results for the approach to the T-junction, along the pipeline from waypoint 3 to 4, can be seen in Figure 7b. The image contains two candidate landmarks to be tracked, the left and right turn, marked with different symbols, circles and crosses. The solid line indicates estimated distance to landmark, based on timing of traversal between waypoints. One can observe that, as the robot approaches the wall of the crossing pipeline, relatively sparse amounts of range data along the pipeline direction is available. In presence of sparse range data, the cone model becomes slightly unstable and the cylindrical model is used. This is seen as a slight jump in distance measurements observed around 14.5 seconds into the robot path.

4. Conclusions and outlook

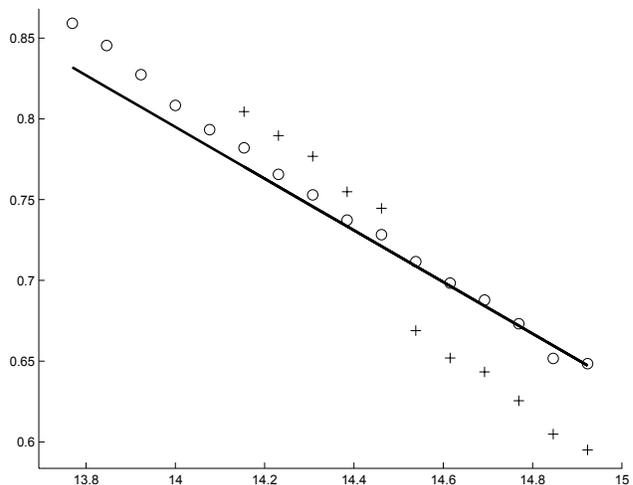
The time-of-flight camera has been shown to be a promising sensor for our autonomous robot navigation needs. Landmarks necessary for navigation are demonstrated easily detectable and trackable using a straight-forward model.

Our current method for feature detection is sensitive to situations where very little of the current pipe is visible, i.e. when the camera has almost entered a T-junction or has just started its traversal. In such cases, there are few data points to provide a reliable estimate of cylinder parameters. Future work will focus on improving the analysis of the deviations, reducing the requirement for an accurate model fit, possibly also improving the results by measuring deviation in cartesian as opposed to cylindrical coordinates.

An improved camera calibration would probably give better experimental results. It is possible to improve the model fitting and image synthetization by using more



(a) Illustrated with crosses are tracked distances to the Y-junction in each frame. The junction is detected around 3.5 seconds into the robot path, and stays in the camera field of view for approximately 1.5 seconds until distance to landmark is closer than 0.35 meters. Indicated with a solid line is the distance to target calculated by auxiliary timing of the traversal of the path.



(b) The T-junction has two simultaneous targets at approximately the same distance from the robot, first detected around 13.5 seconds into the path. Illustrated with solid lines are distances to targets based on auxiliary timing of the traversal of the path. The slight jump in distance measurements around 14.5 seconds can be attributed to the feature tracker switching from cone model to cylinder model, and a lack of reliable range data when the robot moves increasingly closer to the pipeline wall in the T-junction.

Figure 7: Measured distances to tracked landmarks

knowledge of the imaging process. For instance, some of the effects described in [4] could be incorporated, using a forward model instead of the described reverse-modeling of the effects.

The estimated range to detected landmarks or obstacles is yet to be verified with better precision. While the measured speed of approaching features seem to be correct, fur-

ther work is necessary to verify the absolute positioning ability of the TOF camera. Secondary sensors, such as accelerometer, gyroscope, contact sensors and odometry data, can provide additional information for tracking of complicated robot movements horizontally and vertically. Fusion of such sensor data for model verification and improvement would be an interesting topic for further studies. A more sophisticated algorithm for differentiation and classification of obstacles from landmarks should also be developed.

We found that in a pipeline environment, TOF cameras can provide continuous 3D data of sufficiently high quality. Where stereo cameras and laser range sensors provide rather sparse data sets, the large data sets and data quality from TOF cameras simplifies further analysis work. While future work on robustness is necessary, basic landmark detection based on TOF imagery has been demonstrated.

References

- [1] A. Ahrary, L. Tian, S.-I. Kamata, and M. Ishikawa. Navigation of an autonomous sewer inspection robot based on stereo camera images and laser scanner data. *International Journal on Artificial Intelligence Tools*, 16(4):611–625, 2007.
- [2] R. Bostelman, T. Hong, and R. Madhavan. Obstacle detection using a time-of-flight range camera for automated guided vehicle safety and navigation. *Integrated Computer-Aided Engineering*, 12(3):237–249, 2005.
- [3] Z. Jia, A. Balasuriya, and S. Challa. Recent developments in vision based target tracking for autonomous vehicles navigation. *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, pages 765–770, 2006.
- [4] T. Kavli, T. Kirkhus, J. Thielemann, and B. Jagielski. Modeling and compensating measurement errors caused by scattering in time-of-flight cameras. To appear in *SPIE Optics + Photonics*, San Diego, August 10 - 14, 2008., 2008.
- [5] S. Kim and S.-Y. Oh. SLAM in indoor environments using omni-directional vertical and horizontal line features. *Journal of Intelligent & Robotic Systems*, 51(1):31–43, 2008.
- [6] F. Kirchner and J. Hertzberg. A prototype study of an autonomous robot platform for sewerage system maintenance. *Autonomous Robots*, 4(4):319–331, 1997.
- [7] S. May, B. Werner, H. Surmann, and K. Pervolz. 3D time-of-flight cameras for mobile robotics. *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 790–795, Oct. 2006.
- [8] M. Miettinen, M. Ohman, A. Visala, and P. Forsman. Simultaneous localization and mapping for forest harvesters. *Robotics and Automation, 2007 IEEE International Conference on*, pages 517–522, 10-14 April 2007.
- [9] J. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.
- [10] E. Rome, J. Hertzberg, F. Krichner, U. Licht, and T. Christaller. Towards autonomous sewer robots: the MAKRO project. *Urban Water*, 1(1):57–70, 1999.
- [11] S. Thrun. *Robotics and Cognitive Approaches to Spatial Mapping*, chapter Simultaneous Localization and Mapping, pages 13–41. Springer Verlag, 2008.
- [12] P. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Journal of Computer Vision and Image Understanding*, 78(1):138–156, 2000.