Navigation of an autonomous sewer inspection robot based on stereo camera images and laser scanner data

Ahrary Alireza, Tian Li, Kamata Sei-Ichiro, Ishikawa Masumi

International Journal on Artificial Intelligence Tools

Volume 16, Number 4, Page Range 611-625, Year 2007-08-01

URL: http://hdl.handle.net/10228/2412
doi: 10.1142/S0218213007003461
NAVIGATION OF AN AUTONOMOUS SEWER INSPECTION ROBOT BASED ON STEREO CAMERA IMAGES AND LASER SCANNER DATA

ALIREZA AHRARY1,3, LI TIAN2, SEI-ICHIRO KAMATA2 and MASUMI ISHIKAWA3

1FAIS-Robotics Research Institute
2-1 Hibikino, Wakamatsu-ku, Kitakyushu 808-0196, Japan
2Graduate School of Information, Production and System, Waseda University
2-7 Hibikino, Wakamatsu-ku, Kitakyushu 808-0135, Japan
3Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu 808-0196, Japan
ali@ksrp.or.jp, tianli@ruri.waseda.jp, kam@waseda.jp, ishikawa@brain.kyutech.ac.jp

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)

Sewer environment is composed of cylindrical pipes, in which only a few landmarks such as manholes, inlets and pipe joints are available for localization. This paper presents a method for navigation of an autonomous sewer inspection robot in a sewer pipe system based on detection of landmarks. In this method, location of an autonomous sewer inspection robot in the sewer pipe system is estimated from stereo camera images. The laser scanner data are also used to ensure accurate localization of the landmarks and reduce the error in distance estimation by image processing. The method is implemented and evaluated in a sewer pipe test field using a prototype robot, demonstrating its effectiveness.

Keywords: Autonomous; sewer inspection; robot navigation; stereo matching; laser scanner; sewer pipe system.

1. Introduction

Pipe walls in sewer systems are prone to be damaged due to aging, traffic and chemical reactions, causing rainwater and groundwater seep into pipe systems. Regional city government reports1 state that this inflow amounts to approximately 30 percent of the total water flow. In addition to the inflow, outflow from damaged systems also occurs, contaminating the surrounding environment2–4.

Currently, inspection is carried out using a cable-tethered robot with an onboard video camera system. An operator remotely controls the movement of the robot and the video system. By this video-supported visual inspection, any noticeable damages or abnormalities are recorded. The reliability of this system depends on the experience of an operator. Therefore, the current manual inspection system is tedious and erroneous.
To ameliorate this difficulty, an autonomous and real-time sewer inspection robot is desired for identifying and extracting faulty areas. One of the central issues in developing an autonomous sewer robot is its navigation. Detecting landmarks such as manholes, inlets and joints in a sewer pipe system is an important task in navigation. In this paper, we propose a method for detecting landmarks and estimating the location of a mobile robot based on stereo camera images and laser scanner data.

Various methods have been proposed for the detection of landmarks. Hertzberg and Kirchner conducted experiments on navigation of a sewer robot in a dry sewer test field at GMD, Sankt Augustin. They succeeded in detecting local features by ultrasonic sensors. The detected landmarks were then scanned by a different, pivoted ultrasonic transducer. Finally, local features obtained by the scanner are classified using a specially trained artificial neural network. Although this method has a high rate of classification (75 out of 81 samples), it lacks the capability of locating an individual manhole.

Schonherr et al. proposed a pivoted ultrasonic transducer that permanently scans the walls of the upper half section of pipes for detecting inlet. This procedure is time consuming. Paletta et al. proposed a method for detecting inlets from grayscale images taken by an onboard CCD camera. Although it involves a time consuming training phase, the detection of inlets by trained neural networks is fast.

In a real sewer pipe system, a robot is required to explore unknown sewer area during the training phase. This property of time consuming training is undesirable from a practical point of view. To automatically detect damages in sewer pipes, Campbell et al. and Clarke proposed pipe profiling methods for 3D reconstruction. Using sensor devices mounted on a teleoperated sewer robot, data evaluation is performed in a stationary surveillance vehicle.

Our method uses stereo camera images and laser scanner data for detecting landmarks. A sewer inspection robot is capable of self localization, which is impossible by the above methods. In our method, the stereo images are captured by a stereo camera at the front of a robot. A fast and accurate stereo matching measure named a linear computation (hereafter referred to as LC) computes the distance between a sewer inspection robot and a landmark based on the stereo images. We refine the distance by a threshold elimination function. If the distance is less than 10cm, we check laser scanner data to detect the exact location of the landmark. Finally, information on the locations of a robot and landmarks are used for robot navigation.

The paper is organized as follows. Section 2 gives overview of our sewer pipe system. Then we present our method in detail in Section 3. Section 4 shows experimental results by our method. Section 5 provides conclusions.
2. Sewer Pipe System

Sewers mainly consist of smooth, uniform cylindrical or elliptical pipes. Fig. 1 illustrates the sewer pipe construction with the manhole over each pipe-bend, the inlets from houses, and the joints connecting two pipes.

For navigation in a sewer pipe system, a robot needs distinctive landmarks. Elements like manholes, inlets and pipe joints are well suited for this purpose. As shown in Fig. 2, manholes are always located in the upper part of the sewer image, while joints and inlets are always in the central and upper half of the sewer image. This facilitates the extraction of their positions from the image. The fact that the maximum length of a plastic or ceramic pipe is less than 2m makes capturing landmarks still easier. The resulting map has descriptions on manholes, inlets, pipe joints, the distance between them and so forth.

Fig. 1. Landmarks in the sewer pipe system.

Fig. 2. Locations of the landmarks in the sewer images.
2.1. The autonomous sewer inspection robot, KANTARO

In the conventional sewer inspection, commercially available robots are capable of
moving inside of straight pipes, but not capable of passing through pipe-bends such
as “curves”, “joints” and “T-junctions.” Therefore, the design of a robot capable of
moving inside of straight pipes and passing through pipe-bends is expected to bring
great progress in sewer inspection industry.

We have designed a compact and novel moving mechanism for sewer inspection
robots. The design is based on passive adaptation of robot wheels to the bends
in the pipe. This is accomplished by proper adjustment of wheel orientation and
passive damping of springs. A single robot with this new moving mechanism can
load a controller, electronic devices, motors and sensors, and is capable of passing
through any pipe bends, moving inside of different sized pipes and avoiding obstacles
without intelligent control using sensors. Fig. 3 shows the overview of the prototype
autonomous sewer inspection robot, KANTARO.

![KANTARO prototype](image)

Fig. 3. The prototype autonomous sewer robot, KANTARO.

2.2. Vision system

Our sewer inspection robot is designed for checking a sewer pipe system based on
sensory data such as stereo camera images and laser scanner data. Obtained data
on the sewer system can be used for inspection and navigation. The environment in
the pipe is less affected by changes of external light. Lighting condition is controlled
by an array of white LEDs around the camera to acquire the best quality image.

2.3. Laser scanner

It is difficult to use the conventional laser scanner on mobile robots such as KAN-
TARO because of its size and weight. Therefore, we design the new laser scanner in
Fig. 4 for Kantaro. The sensor is mounted on the rear-center of KANTARO and IR beam is emitted from the top of a rotary segment. The microcontroller calculates the distance, the laser scanner angle and so on. Data are transmitted to a computer at the sampling speed of 10KHz via the network with the maximum transmission rate of 1Mbps. Fig. 5 illustrates the errors in distance measurement. Table 1 describes the specification of the 2D laser scanner.

![The newly developed laser scanner.](image1)

![Errors in distance measurement.](image2)

The diameter of the typical manhole is larger than that of the connected pipe by 50 cm, and inlet size is between 10 and 15 cm. Sewer landmarks are not uniformly distributed along the pipe perimeter. Rather, they appear in the upper part, left side and right side of pipes. This allows to restrict the search to the angle width of 15 degrees for each candidate as shown in Fig. 6, and to design the size and position of each scanning window. By full rotation of the laser scanner, the mirror rotates 360 degrees and the distance is measured at the right, top and left scanning windows.
Table 1. Specification of the Laser Scanner.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning directions</td>
<td>360 degree</td>
</tr>
<tr>
<td>Scanning speed</td>
<td>0-1800 (rpm)</td>
</tr>
<tr>
<td>Distance range</td>
<td>70-190 (mm)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>±1 (mm)</td>
</tr>
<tr>
<td>Beam radius</td>
<td>0.5 (mm)</td>
</tr>
<tr>
<td>Transmission method</td>
<td>Full duplex serial transmission (2Mbps)</td>
</tr>
<tr>
<td>Sampling speed</td>
<td>More than 10 KHZ</td>
</tr>
<tr>
<td>Measured signals output</td>
<td>Distance, Scanning angle</td>
</tr>
<tr>
<td>weight</td>
<td>200 g</td>
</tr>
<tr>
<td>Size</td>
<td>37 × 48 × 166 (mm)</td>
</tr>
<tr>
<td>Power</td>
<td>±12V (0.5A) + 5V (1A)</td>
</tr>
</tbody>
</table>

Fig. 6. Scanning directions.

3. Methodology

The proposed method in Fig. 7 is composed of two stages: estimation of the location of a mobile robot and detection of landmarks. Firstly, a stereo sewer image is captured by the stereo camera on KANTARO. Secondly, we extract two rectangular Region of Interest (ROI) images for the manholes, inlets and joints, and detect edges in each ROI image (3.1). Thirdly, a fast and accurate stereo matching measure (LC) is applied to these ROI images to estimate the distances between the robot and landmarks (3.2). Fourthly, we refine the estimated distances using a threshold elimination function (3.3). If the refined distance is less than 10 cm, we check the laser scanner data to detect the landmark. Finally, the refined estimates of location of the robot and the landmarks are used for navigation (3.4).

3.1. Edge detection in extracted Rectangular ROI Image

In sewer pipe systems, manholes, inlets or joints are always located in particular parts of the sewer image; the manholes and inlets are always in the upper part of the sewer image, and the joints are always in the central part of the image. Fig.
Fig. 7. Overview of the proposed method.

8 illustrates the original image and the extracted rectangular ROIs of manholes, inlets and joints.

Fig. 8. Original image and the extracted rectangular ROI images. (a) The rectangular ROI image for the manhole and inlet. (b) The rectangular ROI image for the joint.

After the extraction of ROI, we use the canny edge detector to get the edges in ROI. The Canny edge detection algorithm is known to be the optimal edge detector in terms of error rate. It introduced two kinds of threshold for the gradient with hysteresis, i.e., the lower threshold \( T_1 \) and the higher threshold \( T_2 \). The values of the threshold are determined based on the following requirements. The first requirement is that edges in images should not be overlooked and that non-edges should not be detected as edges. The former corresponds to the error of the second kind, and the latter corresponds to the error of the first kind in mathematical statistics. The second requirement is that the edges be well localized. In other words, the distance between the detected edge and the actual edge be minimum. The third requirement is to produce only one response to each edge. This is included because...
the first two are not enough to eliminate multiple responses to an edge.

Based on these requirements, the canny edge detector firstly filters the image to eliminate noise. It then finds the image area with large spatial derivatives. The algorithm then tracks along these regions and any pixel in an edge list with the gradient larger than the higher threshold is classified as a valid edge. Pixels connected to a valid edge and with the gradient larger than the lower threshold are also classified as an edge. The small value of the higher threshold tends to increase the number of spurious and undesirable edge fragments, hence should be avoided.

3.2. Distance measurement by stereo matching

An application of the conventional gradient measure for stereo matching to our sewer image generates disparities, which provide the distance map based on the geometry of stereopsis. The resulting distance map is represented by the brightness of pixels; the larger the distance is, the darker the pixel is. Due to the shadow, however, the conventional gradient measure tends to produce fringes of the feature pixels in the distance map as in Fig. 9.

To overcome this difficulty, we propose a new measure which is partly similar to the conventional gradient measure. To compare locations in two images, most existing methods depend on similarity reflecting the resemblance of the corresponding locations of two images, and sometimes on distinctiveness reflecting the likelihood of the correctness of the match. Our measure combines the similarity and the distinctiveness into a single measure of matching. Since only the horizontal shifts need to be considered in the stereo matching, the differences of the brightness between horizontally neighboring pixels are calculated. The differences between two neighboring pixels in left-eye and right-eye images are

\[ D_L(x, y) = I_L(x, y) - I_L(x-1, y) \]

and

\[ D_R(x, y) = I_R(x, y) - I_R(x-1, y), \]

respectively.

We calculate the sum of their absolute values,

\[ C_d(x, y) = |D_L(x, y)| + |D_R(x + d, y)|, \]

to represent the distinctiveness, and the minus of the absolute value of their difference,

\[ G_d(x, y) = -|D_L(x, y) - D_R(x + d, y)|, \]

to represent the similarity at the displacement, \( d \). We define the matching measure as the sum of these two terms:

\[ E_d(x, y) = C_d(x, y) - G_d(x, y). \]

In summary, for a given displacement, \( d \), the matching measure, \( E_d \), is:

\[ E_d(x, y) = |D_L(x, y)| + |D_R(x + d, y)| - |D_L(x, y) - D_R(x + d, y)| \tag{1} \]

To find the best match for an isolated pixel, we maximize \( E_d \) with respect to \( d \) under consideration. It is not hard to understand why our method generates strong responses at the feature pixels. If the matching pixel is a non-feature one, the difference of brightness between horizontally neighboring pixels is small. As a result, both \( D_L(x, y) \) and \( D_R(x + d, y) \) have small values. Hence, \( E_d \) also has a small value. On the contrary, if the matching pixel is a feature one, it has a large difference of brightness horizontally. This will result in a large value of \( E_d \). Therefore, this method is expected to work well in the matching of feature pixels.
Fig. 9 demonstrates that the resulting distance map by the proposed measure provides better view of the feature pixels than that by the conventional gradient refined measure. The distance errors for ten pairs of stereo images using the gradient and LC in the feature pixels is shown in Fig. 10. The distance errors using our LC measure are significantly smaller than those by the conventional gradient measure in the feature pixels.

3.3. Refinement of the measured distance

Fig. 11 illustrates two linear fits of the number of edge points by our method. For computing the linear fits, we use a plenty of training images. Firstly, the edges of training images are extracted by the Canny edge detector using two sets of thresholds for the gradient, \((T_{11}, T_{21})\) and \((T_{12}, T_{22})\). True distances between the camera and landmarks are measured manually, and are used for ground truth. Then, we make two linear fits between the number of edge points extracted by two sets of thresholds for the gradient and the corresponding true distance. Finally, a shorter
distance $\tau_1$ is estimated by the linear fit in Fig. 11(a), given the number of edge points. Similarly, a longer distance $\tau_2$ is estimated by the linear fit in Fig. 11(b), given the number of edge points. These shorter and longer distances play a major role in the subsequent computation.

We can estimate the distance between the robot and each landmark by the stereo matching. We use the following threshold elimination function to obtain refined distance by averaging estimated distances.

\[
\rho(x) = \begin{cases} 
  x & \text{if } \tau_1 \leq x \leq \tau_2 \\
  \tau_2 & \text{if } x > \tau_2 \\
  \tau_1 & \text{if } x < \tau_1
\end{cases}
\]  

(2)

Suppose the estimated distances between the robot and features of a landmark are $D = \{d_1, \ldots, d_{N_f}\}$, where $N_f$ is the number of feature pixels in ROI image. The refined distance, $d_r$, is given by:

\[
d_r = \frac{1}{N_f} \sum_{i=1}^{N_f} \rho(d_i)
\]  

(3)

3.4. Detection of landmarks by a laser scanner

At the last stage, we compute the distance between the robot and each landmark. In other words, when the robot reaches one of the landmarks, the distance becomes zero and the system can detect it. However, the computed distance based on stereo image still has some error, hence the laser scanner data are applied to improve its estimation. If the refined distance is less than 10cm, we regard that the robot is close enough to the corresponding landmark. Then, the system checks the laser scanner data for still better estimation.

As shown in Fig. 6, the laser scanner has three scanning windows at the right, top and left. During one rotation of the laser scanner, the IR beam passes three windows and provides three distance values. Table 2 gives estimation of landmarks based on the status of the distance measurements. Our robot is designed for the pipes with the diameter ranging from 250 to 350mm. When the robot is moving
inside the sewer pipe, the distance measurement is between 125mm and 175mm. The distance value larger than 175mm is regarded as a distance measure at the landmarks.

### Table 2. Estimation of the type of landmarks based on measurements.

<table>
<thead>
<tr>
<th>Status No.</th>
<th>Distance value larger than 175mm</th>
<th>Estimated landmark type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>none</td>
<td>no landmark</td>
</tr>
<tr>
<td>1</td>
<td>at the top window</td>
<td>manhole</td>
</tr>
<tr>
<td>2</td>
<td>at the left window</td>
<td>inlet at the left</td>
</tr>
<tr>
<td>3</td>
<td>at the right window</td>
<td>inlet at the right</td>
</tr>
<tr>
<td>4</td>
<td>at the top and at the right or left window</td>
<td>manhole with junction</td>
</tr>
<tr>
<td>5</td>
<td>at the top, at the right window, and at the left window</td>
<td>robot is outside of the pipe</td>
</tr>
</tbody>
</table>

Fig. 12 illustrates typical places where the KANTARO detects the manhole and inlet in moving through the pipe with diameter 300mm. The rotation starts from the horizontal line (counterclockwise).

![Fig. 12. Typical landmarks detected. (a) The manhole. (b) Inlet at the right side range of scanning.](image)

### 4. Experimental results

To evaluate the effectiveness of the proposed method, we did a series of experiments by running the KANTARO in the sewer test field (Fig. 13) at FAIS-Robotics Research Institute (RRI). The sewer test field is made by plastic pipes with the diameter ranging from 250 to 300mm.

#### 4.1. Self localization

This test demonstrates the accuracy of self localization by capturing stereo images with size of 640×480 pixels such as shown in Fig. 14. We used the lower threshold
$T_1 = 50$ and the higher threshold $T_2 = 200$ for edge detection in each ROI images (3.1). Then the distance of feature pixels computed by LC measure (3.2). Next, we used two different sets of thresholds value for the gradient, $(T_{1_1} = 50, T_{2_1} = 150)$ and $(T_{1_2} = 50, T_{2_2} = 250)$ for computing the linear fit, as shown in Fig. 11 (3.3). Finally, the distance error between refind distance, $d_r$, and the true distance between the camera and a point under consideration in the sewer defined as the following:

$$\text{error} = |d_r - d'| \quad (4)$$

<table>
<thead>
<tr>
<th>Images type</th>
<th>The number of images</th>
<th>Rate(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manhole</td>
<td>32</td>
<td>96.87</td>
</tr>
<tr>
<td>Joint</td>
<td>26</td>
<td>88.46</td>
</tr>
<tr>
<td>Inlet</td>
<td>12</td>
<td>91.66</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>92.33</td>
</tr>
</tbody>
</table>

Table 3 presents the percentage of estimated distances with error less than 5cm. On the average, about 92% of input sewer images provide the distances with small errors. Although the unstable illumination condition in the real sewer pipe system may effect a large errors in the estimation of the distance. We still have to perform further experiments to fine tune the values of parameters used for edge detection to reduce the errors.

4.2. Navigation through sewer test field

This test demonstrates KANTARO’s ability in moving through a sewer pipe to the goal by using the proposed method. This experiment requires two basic perceptual abilities: detecting the landmarks and self localization. Both must be executed with high precision, because misperception such as turning action in straight pipe causes damage to KANTARO.
Fig. 14. Samples of stereo images captured by KANTARO during its movement in the sewer test field.

Fig. 15. Examples of test navigation routes.

The mission of reaching the goal starting from the initial position is given by a human operator. Fig. 15 provides two different routes in navigation.

When the robot moves along the Route 1, it makes a turning action at the first detected manhole with junction. Due to the position of the laser scanner on KANTARO, the manhole is detected when the robot fully entered into the manhole. At this moment, the robot can’t turn, because the front wheel passed the junction entrance. To solve this problem, we modified the action as follows; if the robot enters into the manhole and is to turn it moves backward by 40cm and turns.(Fig. 16).

This mission is successfully achieved in our sewer test field with the maximum speed of 15cm/s. Fig. 17 shows the GUI for navigation during the movement of KANTARO in the sewer test field.

5. Conclusions

We have proposed a method for navigation in sewer pipe systems using the robot platform, KANTARO. Our method uses stereo camera images and laser scanner data for detecting landmarks. It is capable of self localization of the sewer inspec-
Fig. 16. KANTARO turning in the manhole.

Fig. 17. GUI for navigation.

tion robot, which cannot be done by the conventional methods. The results of self localization experiment shows high performance in providing the appropriate distance. We also design the new mobile laser scanner for KANTARO. The locations of landmarks in sewer pipe system are estimated successfully based on measurements. The laser scanner is fast enough to continuously scan relevant pipe sections in the presence of landmarks, while the KANTARO moves at ordinary inspection speed of up to 15cm/s. Also moving the KANTARO in our sewer test field by using the proposed method was done successfully.

Acknowledgment
This research was supported by the Organization for Small and Medium Enterprises and Regional Innovation, JAPAN (SMRJ).

References
1. Material provided by Sewer System Administration Dept., City of Kitakyushu, Japan.