SPECIFICS OF USING LIDAR AND IR-CAMERA FOR DETECTING OBSTACLES ON MARITIME VESSELS

Valentin Soloviev
Novia University of Applied Sciences
Abstract:
Development of autonomous vessels has raised the need to research and develop the sensor technology to be used in vessels. Now, mainly human eyes provide the information about objects close to vessels. If there is no crew onboard or their number is reduced, then there is a need to obtain that information by other means.

Sensors commonly used in vessels nowadays are limited in their use and there are a multiple of cases where they do not provide satisfactory accuracy. The purpose of our research was to determine usefulness of Lidar and Infrared-cameras in detecting objects and obstacles in a maritime environment. Most notably, both Lidar and IR-cameras are not dependent on light conditions, so they work during night as well as during day. This would allow increasing the accuracy of surveillance on ships.

Despite multiple upsides, there are cases where the data from these sensors can be unreliable. For example, Lidar can give false positives due to reflections from ice and IR-cameras can produce pictures where objects are too hard to distinguish from one another due to their similar temperatures. These difficulties can be compensated by using the data from multiple sensors together - ice can be distinguished from normal objects by its heat signature and Lidar can show when 2 objects are not the same by measuring how far they are. Overall, research shows that sensors provide a lot of important data and that this topic has a lot to promise.

This research work was done as a part of the ÄlyVESI - Smart City Ferries research, development and innovation project. The project aims to study the possibilities as well as test and develop new solutions and services for intelligent transport in urban waterways. The financing of the project is mainly from the ERDF. Additional financiers are the Finnish Transport Safety Agency and the cities of Helsinki and Espoo.
Table of Contents

1. Sensor Specifications 3
  1.1. Lidar Specification 3
  1.2. IR Camera Specification 4
2. Synchronizing Lidar Data with Video 5
  2.1. Calibration of Intrinsic Parameters 6
  2.2. Calibration of Extrinsic Parameters 7
  2.3. Reusability of Parameters 8
  2.4. Extendibility 9
3. Features of Lidar Data 10
  3.1. Reflections from water 10
  3.2. Reflections from ice 11
  3.3. Object Detection 12
  3.4. Losing objects between laser scans 14
  3.5. Reflective properties of surfaces 15
  3.6. Snow 15
  3.7. View angle variation 16
4. Features of data from Infrared Camera 17
  4.1. Contrast 18
  4.2. Reflections 20
  4.3. Living things are easy to detect 21
  4.4. Computer Vision algorithms are applicable 22
5. Further Research 22
1. Sensor Specifications

1.1. Lidar Specification

The basic working principle of Lidar-sensors is to illuminate the target with pulsed laser light and measure the reflected pulses. This way sensors are able to determine the distance to an object. By doing this hundreds of times, a 3D-map of the measured landscape is created.

Because the sensor is using laser to measure distances, Lidar sensors share many limitations in common with other optical measurement devices (such as video cameras). The two main features that make Lidar interesting are independency of light sources and accurate distance to measured objects.

Unlike normal video cameras, Lidar does not need an external light source to function properly, so there is no difference between using it during daytime or nighttime. The error of distance measurement is +/- 3cm (constant error that does not scale with distance), so the distances provided by Lidar are reliable enough for maritime purposes. Despite all the upsides, Lidar has four main factors that limit its usability:

- Lidar only sees distance and angle: Lidar is unable to discern the color of the point or whether it is connected to another point. Lidar is unable to determine whether there are two objects close together or if there is one object.
- False positives: Lidar does not see objects, it sees reflections. In case there is something tiny reflecting the laser beam, Lidar will see it as a point. Snow and rain will severely limit the visibility of Lidar.
- Low resolution: Lidar has much lower resolution than generic video cameras. When we get one or two points next to each other, it is quite hard to understand what type of an object Lidar is seeing.
- Accuracy varies due to surface: Since Lidar uses reflected laser to measure the distance to objects, it means that the objects’ reflective properties affect how well Lidar is able to see these objects.

For the purpose of this research, we have used SICK’s LD-MRS400001 Lidar. Unfortunately, the characteristics of this model do not align exactly with our needs, but it is good enough to allow us to see the basic behaviour of Lidar sensors. Two main features separate our Lidar from other models on the market with similar price:

- Higher range: SICK’s laser has the range of up to 300 meters (this depends on the reflective properties of the object). At the time of buying, there were no models on the market with higher reach for less than 10000 euros.
- Lower scanning frequency and resolution: We are only getting 5 measurements per second and each of those gives us at the most 800 points
(8 rows of 100 points, if there is something visible) at 110 degree FOV. Yet there are models for a similar price that have at least twice this resolution.

The price for a Lidar can vary from 200 euros for the cheapest models, to up to 80000 euros for the newest models. The range of these models also varies quite a lot, from 5 meters on some of the more specialized models to 300 meters or higher for models with the highest range. New Lidars are continuously being developed and prices on the older models are dropping. For example, recently Velodyne has dropped the price on their VLP-16 Lidar to 4000 euros and it produces up to 600000 points per second with a 360 degree FOV and 100 meter distance. Meanwhile, our model from SICK was priced at 10000 euros and had a resolution that is 15 times lower (although the distance was 3 times longer). Generally speaking, the price range of more affordable Lidars is between 2000 euros and 14000 euros, with the price depending mainly on the range (around 100 meters for cheaper models and 200-300 for more expensive ones), but the field of view and resolution also play a big role and those vary highly depending on the model.

1.2. IR Camera Specification

From a user’s perspective Infrared cameras work in a similar way as normal video cameras. The only difference is that instead of measuring the light from a reflected object, it detects the radiation in the long-infrared range of the electromagnetic spectrum. As long as there are no obstacles between the camera and the measured object, the camera can detect the temperature of that object.

The main advantage of thermal cameras for a maritime environment is that they do not need light to function. Because cameras detect heat of an object, they will function in nighttime as well as in daytime. In some cases IR cameras will even work better in nighttime, since lack of heat from the sun creates a better contrast between objects that are emitting heat and objects that are only reflecting heat.

Living organisms are quite noticeable in IR cameras, because they all emit heat, so detecting a lone person in water is a trivial task when using the camera. Generally, different vehicles are also easily visible in cameras as they tend to produce heat. However, it is important to note some limitations of IR cameras:

- Water and Glass reflect IR: There are surfaces that reflect infrared radiation. One such surface is water. In an image from an infrared camera, it can be easy to mistake reflections of a heat-emitting object for a real object.
- IR cameras require contrast: If the object has a similar temperature to it’s surroundings, then it becomes difficult to detect it on the picture from the IR camera.
- IR cameras do not show distance: Like in normal video cameras, IR images do not show distance to the object. For this reason it becomes hard to
distinguish one object from another if there are multiple objects next to each other (assuming they have similar temperatures).

For the purposes of this research, we used FLIR’s M232 camera. The camera was selected due to its availability, but works as a good representation for other models that are on the market. There are 2 characteristics that should be noted in the camera (these problems seem to be common for other similarly priced cameras as well):

- Narrow field of view: a 320x240 resolution and 24° x 18° field of view. This is quite small for practical purposes and would require to either constantly move the camera around or to set up multiple cameras. This is one of the main limitations of the cheaper thermal cameras, as thermal cameras with high resolution have a much higher price range.
- Small capture rate: Our camera captures the video at 9Hz. For maritime navigation, this is good enough and more expensive models have higher frame rates.

Flir is the main manufacturer of thermal cameras. Originally, these cameras were made for professional and military use, so the price range for them is above 200 000 euros and they are relatively difficult to obtain. Fortunately, due to advancements in the field, the variety has greatly increased and there are thermal cameras for maritime use under the 10 000 price range. Our FLIR M232 is one of such cameras with a price of 3500 euros.

The cheaper models differ mainly in framerate (9Hz, 25Hz and 30Hz) as well as the built-in AD-adapter or electric PAN and TILT. The models on the higher end of the price range have higher resolution, measurement distance, gyroscopic balancer, automatic tracking and integrability with radar-systems. Unfortunately, due to their military use, many of FLIR’s models are unavailable outside the USA or are otherwise more difficult to obtain.

2. Synchronizing Lidar Data with Video

To analyze the data that was captured by Lidar, we had to correlate this to the actual objects in the surroundings. For this purpose we put Lidar and a digital camera on top of each other. The data for both camera and Lidar was timestamped, so synchronization by time was trivial. The bigger problem was to overlay the Lidar data on top of the pictures to find out what it was actually seeing.

To overlay the data we used a process called “3D projection”. In our case this process consisted of three steps:

1. Determine Intrinsic Error Parameters of camera: First step is to determine how the camera distorts the image when it captures it.
2. Determine Extrinsic Error Parameters of camera: Second step is to find out the exact relation between camera and Lidar.
3. Project the points using a matrix with known errors: Use the function from OpenCV (open source library) to project 3d points from Lidar data onto a 2d pane of a camera image by applying error parameters from the 2 previous steps.

This synchronization is useful not only for visualization, but is also a central part for sensor fusion. By using the points from Lidar to determine distance to objects, we can significantly increase the accuracy of object detection algorithms from the video. Conversely, we can also use computer vision algorithms to detect objects in video and then use those objects to increase accuracy of clustering algorithms that are used for detecting objects from point-clouds in Lidar data. The exact approach to take is a topic for further research, but a first step of any fusion of Lidar and Video data is to project Lidar points onto the video.

2.1. Calibration of Intrinsic Parameters

To determine the distortion that the camera introduces to the image, we need to perform a calibration process. This is done by printing a calibration pattern and then taking 50 images of this pattern at different distances and positions from this pattern. Afterwards the program (all the source code and documentation for this are available in open source library: OpenCV) will calculate the distortions of the camera and give them as an output in a vector, which can then be used on the projection step.

This calibration step introduces the biggest error into the projection algorithm. If the calibration pattern is simply shown and held in front of the camera, then the distortion parameters will be inaccurate and the projection will not work on different distances and if the points are in different parts of the image. To calibrate properly, the image needs to be shown at multiple different distances, in all corners of the video and at different rotations.

Because we do not have the equipment to automate this, the end result will likely always have disproportionate amount of calibration pictures for some distance, which will introduce bias to the algorithm that tries to calculate distortion and that will introduce error into the end result.
An example of a simple checkered pattern for calibration of intrinsic error parameters.

2.2. Calibration of Extrinsic Parameters

Extrinsic parameters are essentially two 3d vectors. One vector represents the offset between the origin of Lidar points ([0,0,0] point for lidar) and origin of the camera. The second vector represents the rotation of Lidar relative to the camera. At first we measured the real rotation and offset between camera and Lidar, but that caused the errors to be too big. Due to the lack of proper equipment, (ideally this would require luminofor which is used to calibrate lasers) the calibration process was difficult. After we tried multiple different approaches with varying success, the following is the result we ended up with:

1. Take around 8 pictures with the camera and lidar with distinct landmarks that Lidar is able to see.
2. Write a program that is able to show all these pictures at the same time and overlay the Lidar points onto the pictures.
3. Add controls to the program so that we can change the 3d vectors for offset and rotation in realtime.
4. Change these vectors until all the 3d Lidar-points in the pictures are as close as possible to the real-life objects they were supposed to represent in the camera picture.
2.3. Reusability of Parameters

The parameters we get from the process above are applicable only to the same camera and only if it is not moved in relation to Lidar. If the camera is moved, then the second step needs to be repeated. If the camera is changed or the lens on the camera is changed, then the whole process must be repeated. It is notable that even if the new camera is set in exactly the same place as the old one, it is highly advisable not to skip the second step, as it is likely that there are differences in the second vector due to distortions. Another important point to remember is to have good pictures for calibration in the second step. The pictures we used were for distances above 10 meters. After we tested this setup indoors, we noticed that the system was not producing that good of a projection accuracy on close distances. Another instance was when we calibrated with pictures where all of the buoys and other points of interest were on the right. This led to biased parameters when we used these parameters on an actual video.
Most of the “landmarks” were on the right during calibration, so the points that are on the far left side of the picture got misplaced. If there is not enough variety of “landmarks” in the calibration picture, the error for cases not handled by them will increase.

The conclusion we came to was to calibrate the system with the pictures from the actually used cases rather than doing it locally. Additionally, it is important to have “landmarks” (the points of interest that are being used for calibration) to be at varying distances and preferably in all parts of the video frame.

2.4. Extendibility

The calibration process is applicable to any projection from 3d space to 2d space. If there are other sensors similar to Lidar, we can overlay their readings onto the video and if there are other cameras, we can overlay points onto them. Notably, it is possible to project Lidar points onto the pictures from IR cameras. The main challenge to doing that is calibrating the intrinsic parameters. Because normal calibration patterns cannot be used with IR (IR cameras cannot see pictures on paper), the process needs to be adapted.

We have not studied the topic sufficiently, but by adjusting the algorithms used for calibration and by creating a custom device that would have a visible pattern in the infrared spectrum, it should be possible to replicate the first step of calibration (the second step can be done in the same way as described in chapter 2.2). An alternative way to do this is to first find the extrinsic parameters and then to calibrate the intrinsic ones, then return to calibrating intrinsic and do that in a loop until the accuracy is sufficient.
Unfortunately, due to time constraints we have not been able to verify either of those approaches, so they are left to be topics of further research.

3. Features of Lidar Data

To examine the usefulness of Lidar, we built a sensor platform and put that platform on a ferry. By recording the data from Lidar for one and a half weeks, we were able to discern the main features that the measurements from Lidar possesses. In this chapter, we will discuss all of them and present our assumptions for these.

3.1. Reflections from water

Pillars and boats are properly visible on Lidar, but water is not.

In normal circumstances, water does not produce any reflections for Lidar to see (the reflection happens, but it is polarized and is not detected by Lidar). If there is a relatively calm water surface, Lidar will not see any reflections, so we do not get any false positives from directing Lidar at the water. Unfortunately, we were not able to perform measurements in weather with high waves or foam, so we were unable to determine whether that will introduce false positives.
3.2. Reflections from ice

Ice is a slightly difficult surface for Lidar. The reflection from ice depends on the angle that the Lidar’s laser hits the ice with as well as on the nature of the ice. From our measurements, it seemed that slush in water was not reflecting the signal, but chunks of ice were captured.

Ice and ice chunks on the right were visible to the lidar, but water and slush on the left were not reflected.

Our Lidar was set up in a way that it was looking mainly forward and only the lowest 2 lines of laser could hit the ice. Because of this, we had a line of false positives at distances of 40-50 meters where Lidar was seeing reflections from the ice. This line stands out from other measurements and has a constant distance (depends on the way the Lidar is set up), so filtering these points is a relatively simple task.
The lowest layer of scan hits the surface of broken ice and is reflected as an even line. Flat ice surface is not visible to Lidar.

3.3. Object Detection

From the objects that were detected during the measurement period, Lidar was able to find points in all of them. If the object was close (100 meters or closer), then it was easy for Lidar to find ships and buoys (unfortunately there was not much else to detect during winter). It was also easy to discern them from each other.

The island on the right has an irregular pattern of points.
A ship is visible as a big rectangular clump of points. Different types of ships might be hard to discern from one another, but they are clearly visible on the Lidar scanner.

Cardinal spar buoys are seen mainly as thin rectangular objects (as a line when they are farther and as a pyramid when closer). They are noticeably smaller than a ship, but a regular form makes them easy to recognize.

When the distance is longer (over 200 meters), the resolution of the Lidar starts being too low and it shows only one or two points for the objects it sees. If it was tracking older objects, then even 1 to 2 points is enough to tell us that “we still see that buoy that was near us a while ago”, but in case of discovering a new object, it becomes impossible to tell what we actually see.
At large distance we only get one point for objects and it is impossible to discern what the objects actually are.

3.4. Losing objects between laser scans

When the objects get far enough (or if they are small enough) they might slip between the net of laser scans. When we are on a moving vessel, this appears as one object sometimes disappearing from view for a bit and then reappearing again. This phenomenon depends on the resolution of the Lidar scanner, but it is not noticeable on small distances (below 50m) because the distance between adjacent laser scans is smaller than half a meter (average Lidar will have an even smaller error).
Two buoys on the left and in the middle (top picture) are visible in the first frame, but disappear in the second because they slip between the ‘net’ of lasers.

3.5. Reflective properties of surfaces

The quality and reliability of Lidar measurements is dependent on the material of the observed objects. In case of objects that reflect light well, we see them much farther away and more reliably. For example - buoys are visible at a more than 300-meter distance and it is easy to track them. On the other hand, an island covered in snow or even a harbor is becoming visible only at a 150-meter distance. There are also cases where a surface does not reflect the laser, so Lidar does not see it.

3.6. Snow

Like other optical devices, heavy rain or snow affects Lidar as well. We were not able to capture heavy rain or snow due to time constraints, but we were able to get light snow.

The snowflakes that Lidar captured were mainly at a 2-9 meter distance. We assume this to be caused by the size of snowflakes, but there does not appear to be any false positives of snowflakes at distances farther than 10 meters. These measurements add noise to the data, but depending on the strength of the snow, this type of noise can still be filtered.

The snow is constantly moving, so if there is an object behind the snowflakes, lasers will reach it and tell us that there is something behind all the snow. Of course this depends on how heavy the snow/rain is, so more research needs to be done in this direction.
Snow is visible as noisy points at a distance of 2-9 meters.

3.7. View angle variation

This last factor is not exactly the fault of Lidar, but is closely related. We noticed that the points located by Lidar vary depending on the movement direction of the ferry. When the ferry was going backwards (relative to view direction of Lidar) we saw the ice at a distance of around 50 meters, but when we were going forwards, this ice was not visible in our measurements.

Left picture: the ferry is going forward and is tilted up. Right picture: the ferry is returning and tilted down. Lidar sees ice as false positives in the latter case.
After looking through the data, we concluded that the reason for this peculiarity was our setup. When we were adjusting the angle for the Lidar, we tried to set it just slightly above the ice, so that we would not be getting continuous false positives from it. We did this calibration when the ship was moving forward. When the ship was reversing, the vertical angle of view had slightly shifted thus throwing our calibrations off. This was because ice under the ferry lifted it slightly at the “front” (the direction of movement).

4. Features of data from Infrared Camera

We did not have time to test the Infrared Camera as much as we wanted. We installed it on the aforementioned sensor platform and ran some tests to get a general picture, but there was no time to do a more in-depth study like with Lidar. The main issue was that we were not able to see how it behaves in rain or snow. That being said, we did manage to capture the key features and problems that are characteristic for IR cameras.
4.1. Contrast

Infrared cameras capture only the surface temperature of objects. If multiple objects have similar temperatures and are next to each other, it becomes hard to discern where one object ends and another starts. It can also be quite difficult to determine what is actually happening in the picture if the temperature of all objects is similar.

Picture from the IR camera. Seagulls and ducks sitting on the ice are clearly visible (bottom right), but the ferry on the right could be mistaken for something else. It is also not obvious where the ferry ends.
Most notable problem of bad contrast is when a big object in the background has a similar temperature to objects in the foreground. In that case, objects in the foreground will simply not be visible on the picture. In cases where these foreground objects are tracked, this would not be such a big problem, since they can still be somewhat discernible (even if temperatures are similar, they still have some variation). On the other hand, if we are only given a picture and need to detect objects on it, then it could be an impossible task in some cases.

*Three people in front of the building are barely visible on the camera, but due to heat they are completely invisible on the IR.*
4.2. Reflections

The water surface reflects IR so we will see reflections of what is above water when the camera looks at water. It is also notable that ice surfaces differ quite noticeably from water - it does not reflect IR as well, but is still showing up as a warm surface.

Heat from buildings is reflected on the water surface in the middle and the ice seems to be slightly warm.
4.3. Living things are easy to detect

Vehicles have some warm parts and they can be detected in thermal images, but that is not always the case. Living things on the other hand are always noticeable on thermal images. It is easy to detect people on the ship or even ducks on the ice.

Three people at the front of ferry are clearly visible. The fourth is harder to see, but is clearly visible as well. On the other hand, it is quite hard to find them in a normal video image.
4.4. **Computer Vision algorithms are applicable**

Because IR cameras produce a stream of images, we can apply computer vision algorithms to detect objects. Edge-detection algorithms can help us find proper objects even in cases where multiple objects have similar temperature next to each other (though it would still be difficult). More importantly, if the object is moving, we can track it, which would increase our accuracy noticeably. There is also an option to use machine-learning algorithms to train computers to discern different cases based on the temperature and situation.

5. **Further Research**

Due to constraints in budget and time, there were some cases that were not tested. The most important thing that was left untested was how much different weather conditions affect the data that IR cameras and Lidar receive. Getting measurements in heavy snow or heavy rain would be really useful for assessing the accuracy of the system. Also, the measurements were mainly done in a relatively calm environment, so doing additional measurements during summer, when there is a lot more traffic, would give a better understanding of how well the IR camera and Lidar can see smaller vessels and people in the water.

The research that was done is groundwork for the things that can be done with these sensors. One important topic for further research is fusion of different sensors. Currently we measured effectiveness of each sensor separately, but Lidar fits really well with IR cameras as it can compensate for situations where IR cameras do not discern objects. IR cameras on the other hand allows us to compensate for low visibility of cameras during night-time. Using the data from these sensors to compensate for each other would allow us to get accurate measurements as well as a lot of data on the observed objects.

In recent years, there has been noticeable advancements in radar-technology, so using short-range radars is another venue of research for increasing accuracy of autonomous systems on maritime vessels.