Visualization of pallets

Article · October 2006

DOI: 10.1117/12.684677

Roger Bostelman
National Institute of Standards and Technology

Tsai Hong
National Institute of Standards and Technology

Presented at the SPIE Optics East 2006 Conference, Boston, MA, USA, October 1-4, 2006

Visualization of Pallets

Roger Bostelman, Tsai Hong, Tommy Chang

National Institute of Standards and Technology,
100 Bureau Drive, MS 8230
Gaithersburg, MD 20899 U.S.A.
Email: roger.bostelman, tsai.hong, or tommy.chang@nist.gov,
Tel: 301-975-3426, Fax: 301-921-6165

ABSTRACT

The National Institute of Standards and Technology (NIST) has been studying pallet visualization for the automated guided vehicle (AGV) industry. Through a cooperative research and development agreement with Transbotics, an AGV manufacturer, NIST has developed advanced sensor processing and world modeling algorithms to verify pallet location and orientation with respect to the AGV. Sensor processing utilizes two onboard AGV, single scan-line, laser-range units. The “Safety” sensor is a safety unit located at the base of a forktruck AGV and the “Panner” sensor is a panning laser-ranger rotated 90 degrees, mounted on a rotating motor, and mounted at the top, front of the AGV. The Safety sensor, typically used to detect obstacles such as humans, was also used to detect pallets and their surrounding area such as the walls of a truck to be loaded with pallets. The Panner, was used to acquire many scan-lines of range data which was processed into a 3D point cloud and segment out the pallet by a priori, approximate pallet load or remaining truck volumes. A world model was then constructed and output to the vehicle for pallet/truck volume verification. This paper will explain this joint government/industry project and results of using LADAR imaging methods.

Keywords: 4D/RCS, LADAR, pallet visualization, sensor processing, world modeling

1. INTRODUCTION

The National Institute of Standards and Technology (NIST) has been working with Transbotics, an AGV company, on a joint project toward automated truck loading. One of the first published papers on a similar project was ROBOLIFT by Garibotto et el.[1]. Garibotto et el used a traditional vision based approach. Their primary sensor was a black and white camera. One drawback of using a camera is the reliance on well-controlled lighting condition. With recent advances in LADAR (laser detection and ranging) technology, it is possible to operate in complete darkness and invariant lighting.

Sensor processing utilizes two onboard AGV, single scan-line LADAR units. One is a Safety unit located at the base, front of the forktruck AGV and typically used to detect obstacles such as humans, and slow or stop the vehicle dependent upon obstacles distance from the vehicle. The second, panning LADAR, or “Panner,” was rotated 90 degrees, mounted on a rotating motor for panning, and mounted at the top, front of the AGV. The rotating motor allows the LADAR to quickly pan 60° to acquire many scan-lines of range data.

In order to achieve automated truck loading, two major tasks were identified.

1. Location and orientation of palletized goods needs to be determined and/or verified.
2. Truck container needs to be checked to ensure proper loading, as well as safety.

The Panner sensor processing algorithm combines the many single scan-lines into a 3D point cloud and segments out the palletized good by a priori, approximate pallet load or remaining truck volumes. The Safety unit monitors and tracks the interior of the truck during the loading process.

This paper describes LADAR sensor processing developed during the joint effort with Transbotics. The paper first outlines the overall sensor processing objective, specification and requirement. Details of the LADAR sensors and their data are presented next followed by the sensory processing algorithms that make up the majority of the paper. Finally the results and future development are discussed.

2. OBJECTIVES

In order to achieve the goal of automated truck loading, there are two main objectives. First, the forktruck AGV must be able to verify and locate the palletized good. For example, it must verify that the palletized good is at the correct location and orientation to be picked up. Second, the forktruck AGV must be able to verify the truck container, e.g., making sure the container has enough room for loading.

2.1 Palletized goods

In general, various palletized goods may be located throughout a warehouse. It would be hopeless if the warehouse were to totally lack any structure whatsoever – it would be chaotic. Some structure and prior knowledge, no matter how vague or generic, must exist. They are listed below:

- Palletized goods are box-shaped (plastic-wraped stack of cartons).
- Palletized goods has different, but fixed depth and width. (i.e. 16 cm x 19 cm (40 in x 48 in))
- All palletized goods have the same height but they may be stacked.
- Palletized goods are either placed flatly on the floor or on some supporting surface parallel to the floor (e.g., conveyer).
- Palletized goods may be missing from an expected location
- There is no prior knowledge about the background scene.

With the forktruck AGV stopped in front of a pallet (see Figure 1), the following attributes are measured:

- Skew angle (relative to AGV heading)
- left corner
- width
- height
- depth (when possible)

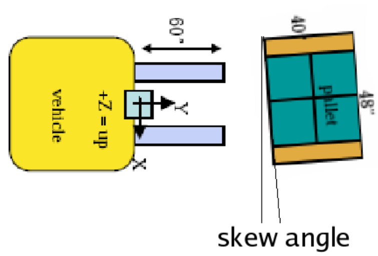

Figure 1: Forktruck AGV and palletized good.

Since there is no prior knowledge about the background in the scene, it could be any background. In general, the background is often cluttered with junk and other man-made objects. This kind of background is typical in a warehouse. The sensor processing algorithm must be robust enough not to confuse palletized goods with other objects in the background. Also, the algorithm should avoid false-positives where it is better to not detect a pallet rather than to incorrectly detect something else.

2.2 Truck container measurement

Palletized goods are carried and loaded into the truck container. The truck container must be verified first before the first load.

- container door may be partially open
- container may not be empty
- container may be missing

Various volumetric and orientation parameters are measured with the AGV stopped just outside and facing the interior of the truck container (see Figure 2). They are:

- skew angle (relative to AGV heading)
- left corner
- width (distance between the two side walls)
- height (defined as lowest overhanging object, if any)
- depth (to the truck front wall or closest object)

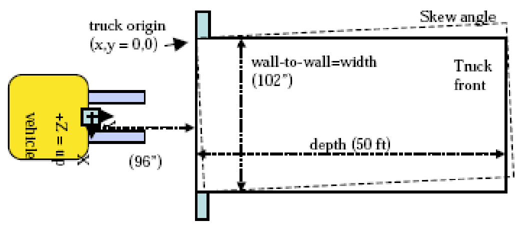

Figure 2: Forktruck AGV outside the truck container.

Once the AGV is inside the truck container (see Figure 3), the following measurements are made:

- distance to let wall
- distance to right wall
- distance to left pallet row
- distance to right pallet row

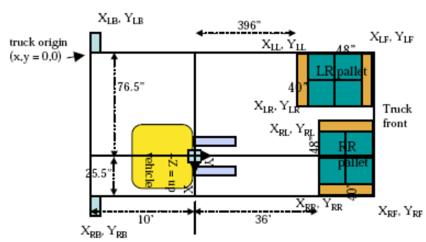

Figure 3: Forktruck AGV inside the truck container.

3. LADAR SENSORS AND CONFIGURATIONS

A Sick laser scanner, model LMS291[1], is rotated 90° mounted (scanner is vertical) on a panning device. Another Sick, model S3000, is mounted on the base of the forktruck AGV. The S3000 LADAR is also used as a safety sensor. Both the LMS291 and the S3000 return range measurements. However, the S3000 has extra hardware functionalities suitable for acting as a safety device.

Subsections below refer to the Sick LMS291 simply as the "panning LADAR" and Sick S3000 as the "base LADAR".

3.1 Panning LADAR

Various field-of-view (FOV) and resolution settings were evaluated. A final setting was chosen to balance the trade-off between the data acquisition time and density of the data. The panning device has its field-of-view set to 60°, while the Sick LMS291 was set to 100°. Both devices were set at 1° resolution. This setting produced a range image of 61 x 101 pixels, covering a 60^{o} x 100^{o} FOV.

---

[1] *NIST does not endorse products discussed within this paper nor manufacturers of these products. Products mentioned are for information purposes only and are not expressed as an endorsement for them or their manufacturer.*

The range information allows reconstructing the scene in various visual forms such as pseudo-colored range image and a 3D point cloud. The panning LADAR was mounted on the main mast of the forktruck AGV. It moves up and down with the fork mast. Figure 4 shows an example of a range image and figure 5 shows a 3D point cloud of the same scene.

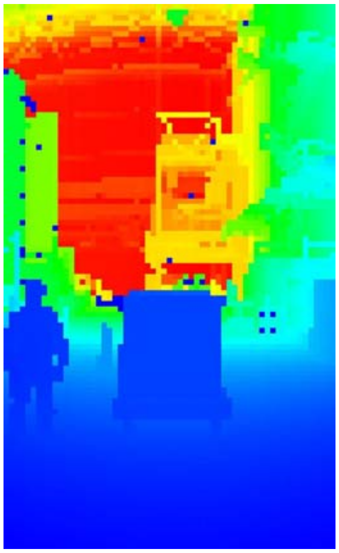

Figure 4: Example of range image, pseudo-color coded. blue=close red=far away

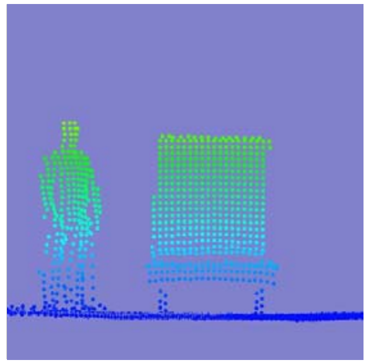

Figure 5: The corresponding 3D point cloud in orthographic view. Points shown are less than 3 meters away and lower than 10 meters in height from the sensor.

3.2 Base LADAR

The Sick S3000 scans the area just above the floor with a 190° FOV and at 0.25° resolution. This produced high density 761 readings. The S3000 is mainly used as a safety unit and typically used to detect obstacles such as humans, and slow or stop the vehicle dependent upon obstacles distances to the vehicle. However, when the vehicle stops, we can use the S3000 to measure the truck container such as distance to the: left wall, right wall, left pallet row, and right pallet row in the truck. Figure 6 shows a sample scan of the inside of a truck with a pallet in the back left corner.

In order to robustly measure those properties, the algorithm uses the prediction of the state of those properties to verify and measure the state of those properties. The algorithm is outlined as followings:

1. Convert range data into corresponding Cartesian X, Y in sensor frame.
2. The system predicts the locations of four corners of the truck in the sensor coordinate e.g. (x1, y1), (x2, y2), (x3, y3), (x4, y4).
3. Apply Line Hough Transform to the data points.
4. Find the closest lines in the Hough Space which are matched to the line for the left wall, right wall and the back wall of the truck.
5. If any wall of the truck is matched to the prediction, the distance from the center of the sensor (0, 0) to the wall is computed.
6. If the back wall of the truck is matched to the prediction, the distance to the back wall of the truck is computed.
7. For points inside of the box which are bounded by the three walls, compute the distance to the three walls.
8. Group all the points which have the distance to the left wall that are smaller than the distance to the right wall as the left pallet and group all the points which have the distance to right wall that are smaller than the distance to the left wall as the right pallet.

9. Find the minimum y distance from the sensor for the left pallet group and the right pallet group. The minimum y distance will be used to determine the distance to the left and right pallets.

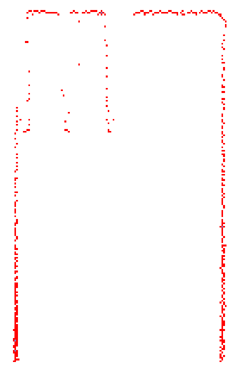

Figure 6: Typical scan of base LADAR (Sick S3000) inside the truck container. Note a pallet at the left corner.

4. DETECTING AND MEASURING PALLETIZED GOODS

Since all palletized goods (subsequently referred to as boxes) are placed on some supporting surface that are parallel to the floor (e.g., a conveyer), the four side facets of a box are perpendicular to the floor. It makes sense to try to first detect the side facets and then identify and predict the location of the box.

A combination of bottom-up (segmentation) and top-down (model based) approach is taken. The algorithm is outlined below:

Bottom-up (segmentation):

1. Convert raw range data into corresponding Cartesian X,Y,Z in vehicle frame.
2. Extract and retain all points that lie on any vertical plane.
3. Group vertical points into corresponding vertical plane.

Top-down (model-based and prediction):

1. Perform bottom-up process only on specific Region of Interest, as specified by prior knowledge (prediction).
2. After bottom-up processing, retain only vertical planes that satisfy target box's dimension constrain.

Bottom-up Step 2 and 3 involve image segmentation that applies cascaded the Hough transformation. The combination of these 2 steps can also be considered a typical divide-and-conquer approach.

Following subsections describe each step in more detail.

4.1 Bottom-up step1: convert raw range data into corresponding Cartesian X,Y,Z coordinate

There are three coordinate frames:

1. LADAR (Xa, Ya)
2. Sensor (Xs, Ys, Zs)
3. Vehicle (Xv, Yv, Zv)

The first step in the bottom-up processing is to take range data in the LADAR frame and transform them into the corresponding Vehicle frame. Figures 7 and 8 graphically depict the LADAR and Vehicle frames.

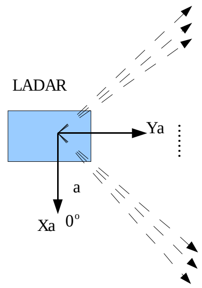

Figure 7:LADAR frame (Xa,Ya). Angle 'a' goes from [45° to 145°] ==>100° fov.

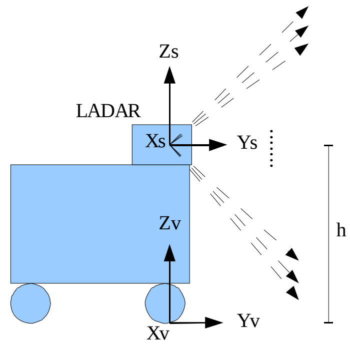

Figure 8: Sensor frame (Xs,Ys,Zs) and Vehicle frame (Vx,Vy,Vz). X-axes point to the right of the vehicle.

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = T \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix} r \begin{bmatrix} \cos(a) \\ \sin(a) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix}$$

$r = range\ value,$

$a = LADAR\ beam\ angle$

$h = sensor\ mounting\ height$

$T = sensor\ orientation,\ transformation = [RotXs][RotYs][RotZs]$

Here RotXs is the tilt effect, RotZs is the yaw effect and RotYs is the roll effect. It is assumed that there is no roll effect. i.e., RotYs = identity matrix. The sensor has a varying tilt mount (i.e., AGV's fork mast tilts, depending on operation mode), and the yaw effect, RotZs, is a function of panning angle. Specifically:

$$RotXs = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & -\sin(t) \\ 0 & \sin(t) & \cos(t) \end{bmatrix}, RotYs = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, RotZs = \begin{bmatrix} \cos(p) & -\sin(p) & 0 \\ \sin(p) & \cos(p) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$t = LADAR\ tilt\ respect\ to\ \ frame\ Xs, Ys, Zs.\ (t > 0\ \ tilting\ upward)$

$p = 90^o\ -\ reading\ \ from\ the\ panner\ device\ (p > 0\ \ panning\ toward\ left)$

4.2 Bottom-up step 2: extract and retain all points that lie on any vertical plane

Since boxes are placed flatly on some supporting surface parallel to the floor, the four side facets of the box are assumed perpendicular to the floor. In addition, the sensor has no roll component, a column in the LADAR image projects to a vertical stripe in the 3D scene (see Figure 9).

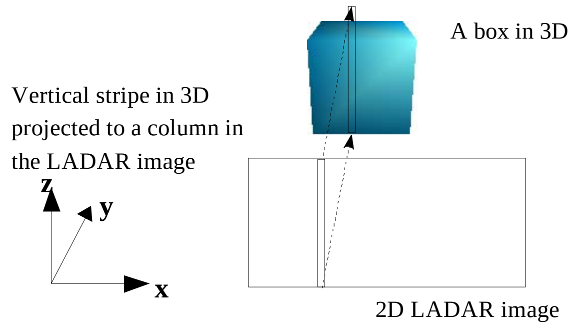

Figure 9: Vertical stripe in 3D space projects to a column in 2D LADAR image.

By using this fact, this step partitions the point cloud according to columns in the image. Each column in the image has only a small number of points. Some of these 3D points are collinear, as in the case when they lay on the side facet of a box. However, walls and other man-made objects also have linear surfaces – a case to be handled in step 3 (below). The Hough Transform is used to efficiently find lines in a set of points. But first, the 3D points are converted to 2D by throwing out the X component.

4.3 Bottom-up step 3: group vertical points into corresponding vertical plane

By the end of Step 2, some columns in the LADAR image have sets of points that form vertical lines in 3D. These lines come from walls, side facets of the box and other man-made objects in the background. It is also possible to have a 3D vertical line that does not lie on the surface of a particular object. This degenerated case happens when there are overhanging objects, as in vertical line 1 shown in Figure 10.

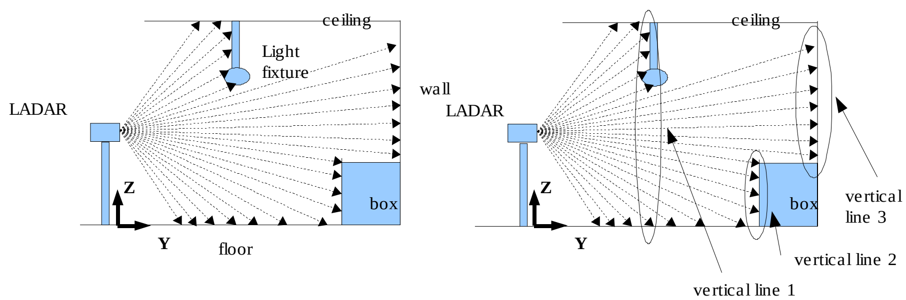

Figure 10: Points forming vertical lines in a column of pixels in the LADAR image. Vertical line 1 is undesired because it groups 2 distinct objects together (lighting fixture and floor).

Degenerated lines are further broken into segments. This is done by measuring 'holes' in the line and breaking the line at each hole location. Figure 11 below illustrates this process in detail.

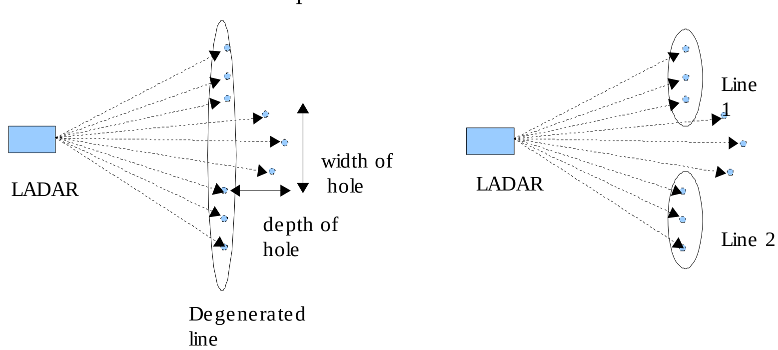

Figure 11: Breaking a degenerated line into segments.

In step 3, each grouped vertical line from step 2 is projected on to the X-Y plane, forming a top-down view of the scene. A vertical line in Y-Z plane becomes a single point in the X-Y plane, as illustrated in Figure 12.

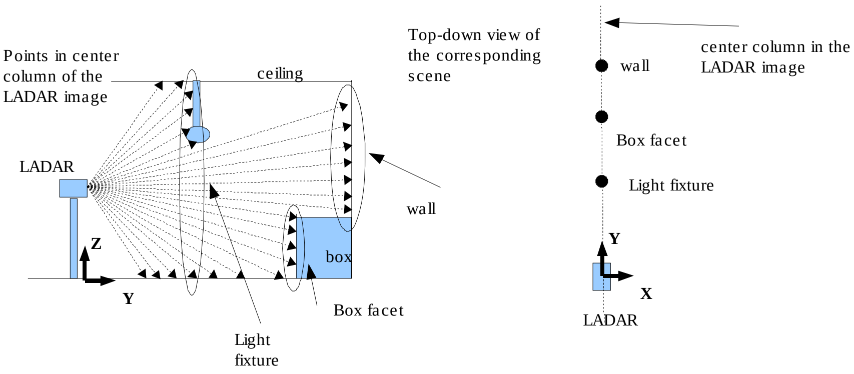

Figure 12: A vertical line in the Y-Z plane projects to a single point on the X-Y plane.

Grouping of vertical points into a corresponding surface is done by projecting all vertical lines onto the X-Y plane and then performing a Hough Transform again – thus cascading two stages of Hough Transforms (first stage was performed in step2).

Since each vertical line is a collection of vertical points, and by grouping vertical lines, points belonging to the same surface/plane are grouped together. Figure 13 shows the grouping of points into a surface/plane. After the 2nd Hough Transform, we have a group of vertical planes. After the first Hough Transform in step 2, a similarly degenerated case can occur. However, it can be handled in a similar way (see Figure 11).

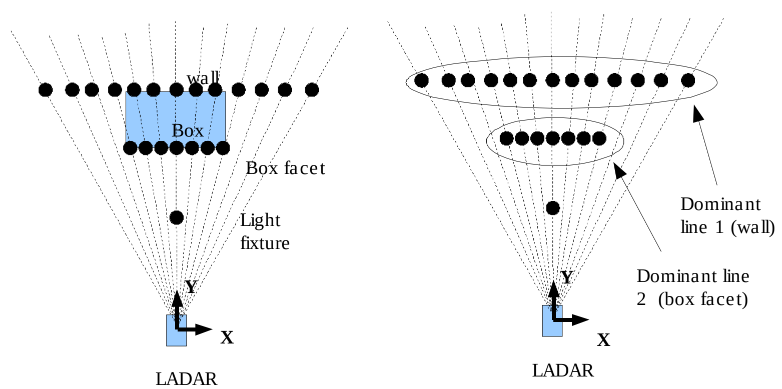

Figure 13: Hough Transform detecting dominant lines. Each line now represents a grouping of vertical points on the same surface or facet. Note that length of Line 2 fits the Box's width constrain.

4.4 Top-down step 1: perform bottom-up processing only on Regions of Interest

A world model is kept current about the location of every box. At any given time, the forktruck AGV (or the centralized controller) has some knowledge about the not-yet-verified boxes. This knowledge is used as a prediction, refining the bottom-up processing to focus only on specific regions in the LADAR image. The world model is updated with any correction from the verification/measurement result. The world model prediction includes: corner, height, width, skew angle and depth of the box.

4.5 Top-down step 2: retain only vertical planes that satisfy constraints

After the bottom-up processing, only a few vertical planes truly belong to target boxes. In order for a vertical plane to be a valid box facet, it must satisfy the following constraints:

1. The width of the plane is either the width or the depth of the box (within pre-specified tolerance).
2. The height of the plane must satisfy (within pre-specified tolerance) the expected height of the box.

Once a valid box facet is determined, the left corner location is easily deduced.

5. MEASURING TRUCK CONTAINER

As described in previous sections, there are two scenarios involving truck volume (subsequently referred to as truck) – inside and outside.

Before entering the truck, the forktruck AGV needs to verify the state of the truck, such as its orientation (skew angle) and door position (open or closed). Once inside the truck, the AGV needs to constantly make sure its pre-programmed path isn't obstructed. For inside the truck, only the base LADAR is used.

For outside the truck, the same bottom-up processing (described in the previous section) involving only the panning LADAR is performed.

5.1 Outside: truck width and skew measurements

The algorithm is outlined below:

1. Group the point cloud into 3D vertical columns. Each column represents a group of points forming a 3D vertical column. The top-down view of this grouping is shown in Figure 14.
2. The algorithm attempts to find the two most dominant lines (defined as lines having the most points) in this top-down view. This search is guided by the predicted truck skew angle (i.e., +/- 15.0° around predicted truck skew).
3. Once the two dominant lines are extracted, they are assumed to be the side walls of the truck. The shortest distance between them is determined. This is the truck width measurement.
4. The truck skew angle is determined as the angle of the better-fitted (less fitting error) side wall. This angle is measured relative to the vehicle frame.

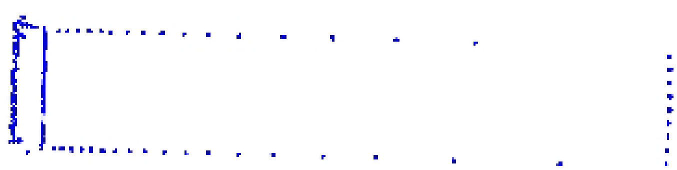

Figure 14: Top-down view of a typical panning LADAR truck scan. The truck door is to the left and the front wall of the truck is to the right. Raw scan data provided by Transbotics.

5.2 Outside: truck depth measurement

The truck depth is measured by the closest flat vertical plane facing the forktruck AGV. The algorithm takes the top-down view as in Figure 14 and finds the closest "unbreakable line." A line is "unbreakable" when there are no other points that lie behind it. Figures 15 and 16 below show the detected unbreakable line in cyan color. The search for this line is guided by the measured skew angle, i.e., +/- 2° search window around measured truck skew angle.

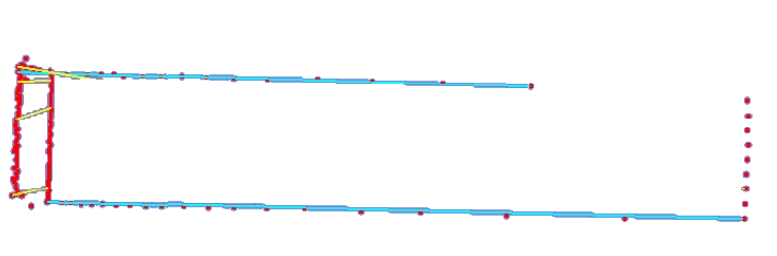

Figure 15: Detected truck side walls are marked in cyan color, they corresponds to two most dominant lines. Measured skew angle =-1.5°(with respect to the forktruck AGV not shown). Measured width = 2.5m.

Figure 16: The detected front wall of the truck is shown in cyan color on the right. Depth of the truck measured 16.45 m from the forktruck AGV. Note the truck door was partially detected as an overhanging object.

5.3 Outside: truck height measurement

The truck height measurement is achieved by removing all points on or close to the truck's side and front walls. The height is defined as the lowest point inside the truck. This can be easily done as the side walls and the front wall have already been determined from the previous step. Figure 17 shows the point cloud with truck's sides and front walls removed. Green points are points on the ceiling. Blue points are on the floor. Red points are the highest points. The point cloud is viewed at an angle.

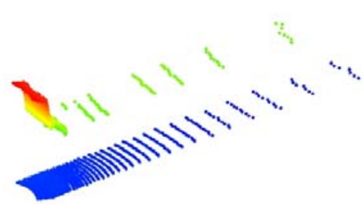

Figure 17: Truck height. Measured as the lowest point = 2.10m.

In some cases, the truck floor may have a tilt angle with respect to the building floor. The sensor may have tilt mount that is not calibrated. In both cases, the floor appears to be tilted and the height measurement will not be correct. However, surface fitting on the floor data can compensate for this error.

5.4 Inside truck measurement

The general inside truck measurement algorithm is outlined below and is similar to the algorithms already described in detail in previous sections:

1. Segment linear features using Hough Transform and Least Squares line fit.
2. Find truck walls and boxes (palletized goods) using the rule-based constraints method with prediction and focus of attention. Specifically, for each linear feature, determine and verify:
    - the angle of the line is within the predicted angle
    - the end points of the line is within the predicted depth
    - the length of the line is within the predicted length

6. RESULTS AND CONCLUSION

Initial testing results have been promising. However, various pieces are not yet fully integrated into the overall automated truck loading system. Figures 18, 19, and 20 show raw LADAR images and the intermediate processing results. After completing the process, only pixels belonging to the target (in this case, a double stacked pallet and load) are marked.

We have described a sensor processing algorithm using cascaded Hough Transforms to effectively and robustly group 3D point cloud into vertical planes. The sensory processing algorithm described in this paper works well for the domain of the problem i.e. box-like object and vertical surface. Although the algorithm requires a zero-roll sensor mount, this is not unreasonable. In the future, the authors would like to investigate an approach to perhaps generalize the algorithm to handle generic geometric objects such as spheres, cones or trihedral solids by considering the Generalized Hough Transform [4].

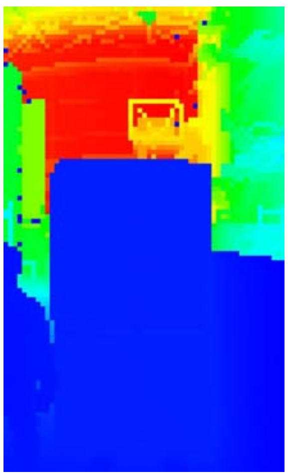

Figure 18: Panning LADAR range image.

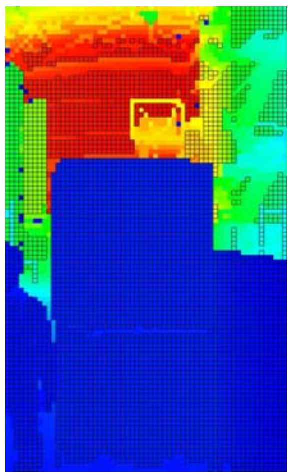

Figure 19: The marked pixels belonging to any vertical plane.

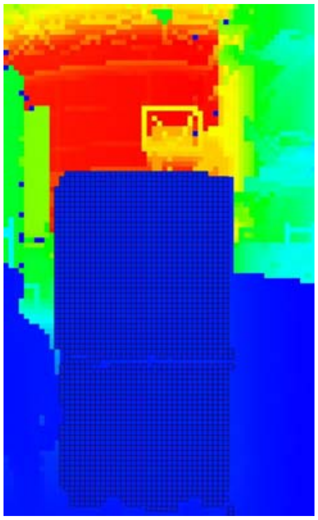

Figure 20: Marked pixels belong to the detected pallet stack.

ACKNOWLEDGEMENTS

The authors would like to thank Transbotics for jointly participating in this cooperative research and development project.

REFERENCE

1. Garibotto, G., Masciangelo, S., Bassino, P., Coelho, C., Pavan, A., Marson, M., “Computer Vision Control of an Intelligent Forklift Truck”, IEEE Intelligent Transportation System, 1997, pg 589-594.
2. Chang, T., Hong, T., Legowik, S., Abrams, M., "Concealment and Obstacle Detection for Autonomous Driving," Proceedings of the Robotics & Applications 1999 Conference, Santa Barbara, CA, October 1999.
3. Bock, Rudolf., “Hough Transform.” [online] available <http://rkb.home.cern.ch/rkb/AN16pp/node122.html> , April 7, 1998.
4. Fisher, S., Perkins, S., Walker, A., and Wolfart, E., “Hough Transform.” [online] available <http://homepages.inf.ed.ac.uk/rbf/HIPR2/hough.htm>, 2003.