Fuzzy Scene Matching in LADAR Imagery

J. Marjamaa, O. Sjahputera, J.M. Keller, P. Matsakis
Department of Computer Engineering and Computer Science
University of Missouri-Columbia, Columbia, MO 65211
{KellerJ, MatsakisP}@missouri.edu

***Abstract*-In computer vision, object recognition and spatial relationship evaluation are useful in creating a scene identification system. This paper examines object pair matching within LADAR (Laser Radar) imagery. By manipulating the three-dimensional data contained within a LADAR range image, it is possible to create a version of the scene as seen from above. In the transformed view, a fuzzy region represents each object, and the spatial relationship between two objects is represented by a force histogram. Matching two scenes then comes down to comparing force histograms. Each comparison gives a degree of similarity between the two object pairs considered, as well as an assessment of the pose parameters. These values can finally be combined to find a correct scene matching. Experiments on synthetic data as well as on real data demonstrate the applicability of our approach.**

I. INTRODUCTION

In order to identify a scene, it is useful to examine the components that make up the scene, as well as the relationships between these components. Earlier work in scene matching, using gray scale correlation, or Fourier space correlation [1] lack the power to use high-level concepts such as objects' spatial relationships. These systems, while very fast [2], do not expect the dramatic difference that can occur in most practical 3D applications, and require similar scenes to start with. Existing high-level systems such as those using descriptions [3] or relation graphs [4] do not use three-dimensional information and can not identify the same scene as viewed from two dramatically different angles. In this paper, we tackle this problem, and we examine scene matching in LADAR (Laser Radar) imagery. By manipulating the three-dimensional data contained within a LADAR range image, we create a version of the scene as seen from above. Rather than using model-based reconstruction such as [5], which would not be able to handle all of the possible variations of objects that might occur in real world scenes, systems that require multiple view angle sources [6], or systems that make assumptions based on interpolations about occluded regions [7], the areas of uncertainty in the transformed view are filled in with fuzzy regions based on what is known about the scene.

The force histogram introduced by Matsakis and Wendling in [8] is a valuable tool in representing the relative position between two 2D objects. In a preliminary study [9], we showed that this tool could be used to determine the similarities of object spatial relationships, and to identify objects transformed by rotation and dilation (i.e., scaling). However, when comparing different bird' s-eye views of some scene, the approach presented in [9] basically ignores distortions caused by declination and three-dimensional objects. An in-depth study was conducted in [10]. To identify relationships, a search is done through rotation space, dilation space, and also stretching (i.e., declination) space. This constitutes a vast improvement, but dramatic distortions caused by 3D objects still cannot be handled.

In this paper, segmented LADAR images are used in conjunction with the range data. The goal is to reconstruct each scene so that it is independent of the third dimension and the tilt (or declination) of the camera. In other words, the three-dimensional information is used to determine the declination angle, and to transform the segmented scene to a position as viewed from above. Matching is then performed by comparing the relationships between the objects in the transformed scenes. For rotations and scaling, we use the same search methods described in [10].

II. TRANSFORMING A SCENE TO A DECLINATION INDEPENDENT VIEWPOINT

To achieve the goals of this study, it was necessary to modify the images so that the comparisons between two scenes would be simplified. The most obvious way to achieve this was to eliminate the angle by which the scenes are declined so that two images only differ in their rotation and scaling. We used a set of range data from the power plant at China Lake, CA. This data set was provided by the Naval Air Warfare Center (NAWC), and generated by a laser radar system mounted on a surveillance plane flying above the power plant complex. Since the range data, from which the segmented images were created, was available, the images were modified using the three-dimensional information contained within to create a version of the image as seen from above.

The range data had many artifacts and noise points that would potentially throw off the results of any transformation. A median filter with a three by three window cleans the range data sufficiently. This corrects inaccuracies in the range data, but there are still problems with the segmented data.

The scenes used in this experiment were hand-segmented. The LADAR range data was processed to be easier to segment, by "lighting" the scenes. It makes the scenes look more natural to the human eye and easier to segment. A normal for each pixel is calculated from the three-dimensional positions in the range data of the pixel, the pixel above, and the pixel to the right of it. The dot product of this normal, and a light vector are then calculated to find an intensity value for the pixel. The result of this filter is seen in Fig. 1. The median and "lighting" filters are followed by the hand segmentation resulting in a rough approximation of the objects' edges.

To transform the segmented scenes, it is necessary for the boundaries of the segmented image, and the range data to be aligned. If this were not done correctly, some pixels would be separated from their objects. Fig. 2 demonstrates this problem. Fig. 2a shows what might be a single column of one of the scene images. Mapped onto the column are the labels resulting from hand segmentation. The X'ed blocks represent background pixels, and the O'ed blocks represent the object. Empty blocks are not visible from the original point of view, and are therefore not labeled. If no correction were performed on the boundaries, Fig. 2c results. From the original view, everything seems to be labeled correctly, but when the image is rotated to an overhead view, the inaccuracies in the segmentation become apparent. Notice the one O'ed block that becomes separated from the object. This block will cause inaccuracies in later procedures.

To correct this problem, range information is used. First, a search is performed on the segmented image to find the back edges of all objects. Next, these edges are adjusted to correlate with the best possible "real" edge on the range data by shifting the edge in one of the four cardinal directions. This does not eliminate all of the error caused by inaccurate segmenting, but does clean up the resulting image nicely.

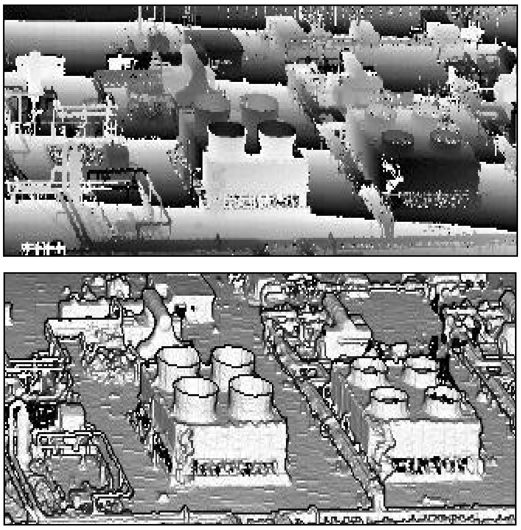

Fig. 1. The first image shows a range image from LADAR data. The second has had "lighting" applied to it.

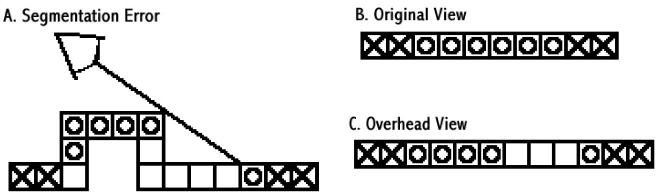

Fig. 2. A demonstration of segmentation problems.

A Hough-like transform is used to determine the angle of declination from the range data. An accumulator representing the different degrees of declination is created. We assume that the most commonly found angle between two data points in all of the columns of the data is the declination angle. Thus, the method only works on data that represents a relatively flat landscape with objects. The arctangent of the difference in range over the distance between a pair of data points in a given column provides the angle of the two points in space, which is used to increment the accumulator. After each pair of points in each column has been processed, the maximum value of the accumulator is used to indicate the most common declination angle in the entire scene. Since the images were mostly level, it was not necessary to find an angle to describe the "roll" of the scene, although a similar method could be used.

After the angle of declination has been determined, the labels of the segmented images are mapped to the three-dimensional positions of the range data. These positions are then rotated by the angle of declination. The intensity values of the three-dimensional points are projected down, so that the resulting image is the seen as viewed from above. The image resulting from this process has many gaps in it, due to a general spreading out of the points and three-dimensional objects that obscured points in the original image.

III. Reconstructing the Image

Once an accurate overhead image was created, it was possible to begin filling the "holes" caused by the rotation. There are two types of holes created. The first is a hole caused by the general spreading of the pixels caused by the rotation. The second is a hole caused by an object obscuring either another object, or the background. These two types of holes are handled differently to create an accurate description of the overhead view of the scene. If a single column of the image is examined, the type of hole can be determined by the two bounding pixels. If the two pixels have the same label, the hole type is caused by general spreading, or an object obscuring itself. Thus it is safe to assume that all of the pixels between the two have the same values as the bounding pixels. In the second case, the two bounding pixels belong to two different objects (or an object and the background). Then, there is really no way to determine what would lie between the two objects. For this reason, we have decided to "fill" these gaps by constructing fuzzy regions. To best represent the uncertainty, two types of boundaries were considered. If a hole occurs on the front side (the side closest to the camera) of an object, and there is significant evidence that that side is obscured by another object closer to the camera, then the fuzzy region associated with the object is constructed with the front-end membership function, scaled to fit the length of the uncertainty area (Fig. 3b). Hence, the uncertainty area is filled in smoothly, and the obscured object is "reconstructed" in a coherent manner. If a hole occurs on the backside of the object, and if enough evidence exists to support the conclusion that the object stretches out into the uncertainty area, then the fuzzy region is constructed using the back-end membership function (Fig. 3a). Fig. 4 shows an example of object reconstruction.

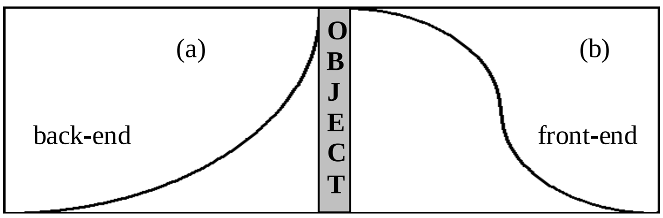

Fig. 3. Two membership functions used to construct the fuzzy regions. a) The back-end membership function. b) The front-end function.

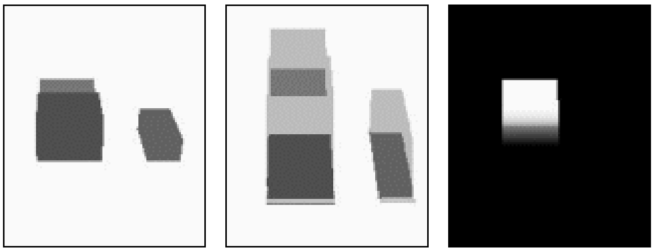

Fig. 4. The first image (left) shows the original scene. The second image (middle) shows the scene after being transformed to an overhead view (the light gray regions are the areas of uncertainty). The final image (right) shows the fuzzy region used to represent the furthest most object (the lighter the pixel, the more it belongs to the object).

IV. Object Pair Relation Analysis and Scene Matching

Once the two scenes to match have been transformed to a declination independent angle, the histograms of constant and gravitational forces [8] are calculated for each individual fuzzy region pair. At this point, when comparing two histograms not coming from the same scene, the rotational difference and scaling ratio of the images can be assessed and varied to maximize a similarity measure (for our experiments, we used the classic sigma-count of the intersection over the union). Rotational difference is calculated as the distance between the centroids of the two histograms, and can be varied by shifting one of the histograms along the horizontal axis. The scaling ratio is calculated by comparing the histogram averages, and can be varied by stretching one of the histograms vertically. Details can be found in [10].

For the experiments, it was assumed that the object contents of each scene were the same, thus for a scene with four objects, 720 possible scene matches must be considered. For each possible matching, an overall “matching degree” is derived from the computed maximum similarity measures, scale ratios, and rotational differences.

V. Results

For the first set of experiments, several artificial three-dimensional scenes of buildings on a flat plane were created to demonstrate the functionality of the system, as well as some of its strengths and weaknesses. The first pair of scenes to be compared demonstrates the merits of reconstructing the front edge of an object using a fuzzy region. The two scenes are displayed in Fig. 5. Note that the objects have been labeled consistently. For each scene, the declination angle is recovered and the overhead view is computed (see Sections 2 and 3). When the objects' edges are *not* reconstructed (i.e., no fuzzy regions), the similarity measures using the histograms of gravitational forces are found to be as in Table 1. When the objects are reconstructed using fuzzy regions, there is a very noticeable improvement, as shown by Table 2.

Table 1 has one very notable mistake. As seen in row 0_2, the pair was identified as 0_1. This mistake is due to the fact that the occlusion of object 2 makes it look like object 1. In the results based on the fuzzy reconstruction of the occluded region, the correct pairs are matched. In all experiments, the results of the histograms of constant and gravitational forces were quite comparable.

The images used in the second experiment are those displayed in Fig. 6. Object 0 has a structure that is obscured in the first image (left), but visible in the second (right). The pairs’ similarities in the crisp case were found to be as in Table 3. Table 4 shows that with the fuzzy regions there is again a notable improvement.

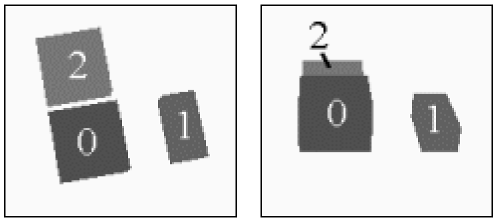

Fig. 5. These two images are of the same scene.
The first has a rotation of 10 degrees and a declination of 10 degrees.
The second has a rotation of 0 degrees and a declination of 60 degrees.

TABLE I
SIMILARITY MEASURES OF OBJECT PAIRS IN FIG. 5 USING CRISP REGIONS.

	0_1	0_2	1_2
0_1	0.956	0.73	0.663
0_2	0.974	0.717	0.677
1_2	0.495	0.349	0.705

TABLE II
SIMILARITY MEASURES OF OBJECT PAIRS IN FIG. 5 USING FUZZY REGIONS.

	0_1	0_2	1_2
0_1	0.966	0.715	0.679
0_2	0.618	0.875	0.425
1_2	0.645	0.455	0.911

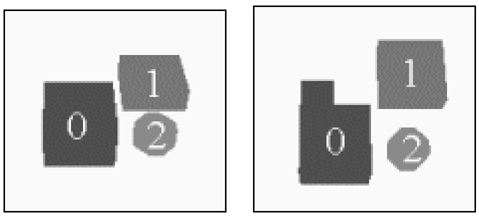

Fig. 6. These two images are of the same scene.
The first has a rotation of 0 degrees and a declination of 50 degrees.
The second has a rotation of 0 degrees and a declination of 20 degrees.

TABLE III
SIMILARITY MEASURES OF OBJECT PAIRS IN FIG. 6 USING CRISP REGIONS.

	0_1	0_2	1_2
0_1	0.755	0.853	0.862
0_2	0.681	0.969	0.788
1_2	0.832	0.828	0.963

TABLE IV
SIMILARITY MEASURES OF OBJECT PAIRS IN FIG. 6 USING FUZZY REGIONS.

	0_1	0_2	1_2
0_1	0.921	0.814	0.903
0_2	0.806	0.943	0.808
1_2	0.847	0.773	0.97

There is an incorrectly matched object pair in row 0_1 of Table 3. This is corrected when the occluded back-ends of the objects are reconstructed. Hence, based on this experiment, there are cases where there is merit to using a fuzzy region to fill in the uncertainty area on the back-end of an object. However, we have found that in cases in which there is no evidence that there are structures being obscured behind the object, this assumption would decrease the performance of the system as a whole.

In a final experiment, we examine how well the system performs on a real world image. The images are those displayed in Fig. 7. Only the four labeled objects were considered. The shapes of the objects and the distance between a given pair in these scenes vary greatly. By "eliminating" the declination, these factors are ameliorated. Note that even though the results with fuzzy regions are calculated, the effects of the occluded regions are minimal due to the nature of the scenes and the lack of obscuration. The similarity measures are found to be as in Table 5.

These results are not sufficient to match object pairs, or the entire scene. In every row, except for 0_3, the highest similarity measure does not correspond to the true matching. Thus, it is necessary to use additional information, such as the calculated rotational differences, and the scaling ratios. Any combination of two of the three measures (similarity measures, rotational differences, and scaling ratios) is sufficient to identify the true matching, but a combination of all three yields the widest margin between the correct and next best matching. That is why the matching degree mentioned in Section 4 derives from the three measures.

As expected, the matching with the highest degree is found to be the true matching, with a value of 0.997. The next closest matching swaps the relationships of objects 0 and 3, with objects 1 and 3, and only earns a matching degree of 0.976. The high value is because objects 0 and 1 are both about the same size and the relationships between them and object 3 only differ by a few degrees. Only 8 of the 720 matching possibilities achieve a matching degree above 0.9.

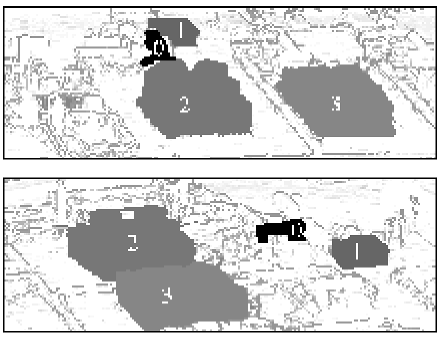

Fig. 7. These two images are from the same scene shown from two different viewpoints.

TABLE V
SIMILARITY MEASURES OF OBJECT PAIRS IN FIG. 7 USING FUZZY REGIONS.

	0_1	0_2	0_3	1_2	1_3	2_3
0_1	0.484	0.484	0.925	0.736	0.928	0.363
0_2	0.52	0.525	0.808	0.817	0.824	0.386
0_3	0.457	0.451	0.896	0.689	0.86	0.345
1_2	0.329	0.322	0.699	0.514	0.665	0.255
1_3	0.308	0.3	0.644	0.479	0.613	0.242
2_3	0.733	0.701	0.384	0.491	0.407	0.696

VI. CONCLUSIONS

We have shown that scene matching in LADAR (Laser Radar) imagery can be performed by manipulating the scene viewpoints and constructing fuzzy regions. The approach described in this paper is powerful because it combines the robustness of force histogram comparison for two-dimensional object pair identification with known information about the three-dimensional structure of the scene. These tools used with existing object recognition tools can create a powerful scene recognition system. One limitation of our approach is that it does not handle perspective distortion. If this parameter could be found in an image, then the transformed image would indeed be the scene as viewed above. The obvious next goal of this ongoing research will be to use what has been learned here to apply to scene recognition.

ACKNOWLEDGMENT

The Office of Naval Research (ONR) grant N00014-96-0439 supports this work.

REFERENCES

[1] Dongcheng Shi, Liqang Han, Ying Liu, "A Scene Matching Algorithm Based on the Knowledge of Object Edges," *IEEE International Conference on Intelligent Processing Systems*, October 1997, pp. 1442-1445.

[2] R. Venkatsan, R. Sastry, N. Ranganathan, "A VLSI Architecture for Heirarchical Scene Matching," *11th International Conference on Pattern Recognition*, 1992, pp. 214-217.

[3] J. Lemaire, O. LeMoigne, "Development of a Scene Recognition System with Imprecise Descriptions," *IEEE Procedings on the International Conference on Image Processing*, 1996, pp. 979-982.

[4] M. Celenk, "Color Scene Matching," *IEEE Proceedings of the 27th Southeastern Symposium on System Theory*, 1995, pp. 284-288.

[5] M.S. Abon-El-Ela, H. El-Amroussy, "A Machine Vision System for the Recognition and Positioning of Two-Dimensional Partially Occluded Objects," *8th Mediterranean Electrotechnical Conf.*, 1996, pp. 1087-1092.

[6] V. Sequeira, J Goncalves, M Ribeiro, "Active View Selection for Efficient 3D Scene Reconstruction," *IEEE Proceedings of the International Conference on Pattern Recognition,* 1996, pp. 815-819.

[7] P. Ballard, F. Vacherand, "The Manhattan Method: A Fast Cartesian Elevation Map Reconstruction from Range Data," *IEEE Conference on Robotics and Automation*, 1993, pp. 143-148.

[8] P. Matsakis, L. Wendling, "A New Way to Represent the Relative Position of Areal Objects," *IEEE PAMI (Pattern Analysis and Machine Intelligence)*, vol. 21, no. 7, pp. 634-643, 1999.

[9] O. Sjahputera, J. Keller, P. Matsakis, P. Gader, J. Marjamaa , "Histogram Based Spatial Relation Similarity Measure for 2D Objects," *NAFIPS'2000* (*North American Fuzzy Information Processing Society*), Atlanta, Georgia, USA, July 2000, *Proceedings*, pp. 392-396.

[10] P. Matsakis, O. Sjahputera, J. Keller, J. Marjamaa, "Matching Image Object Pairs while Retrieving the Pose Parameters," *IEEE PAMI (Pattern Analysis and Machine Intelligence)*, submitted.