Map Quality Assessment

Asim Imdad Wagan[1], Afzal Godil[1], Xiaolan Li[1,2]
[1]National Institute of Standards and Technology
[2]Zhejiang Gongshang University
{wagan, godil, lixlan}@nist.gov

ABSTRACT
The maps generated by robots in real environment are usually incomplete, distorted, and noisy. The map quality is a quantitative performance measure of a robot's understanding of its environment. Map quality also helps researcher study the effects of different mapping algorithms and hardware components used. In this paper we present an algorithm to assess the quality of the map generated by the robot in terms of a ground truth map. To do that, First, localized features are calculated on the pre-evaluated map. Second, nearest neighbor of each valid local feature is searched between the map and the ground truth map. The quality of the map is defined according to the number of the features having the correspondence in the ground truth map. Three feature detectors are tested in terms of their effectiveness; these are the Harris corner detector, Hough Transform and Scale Invariant Feature Transform.

Categories and Subject Descriptors
I.2.9 [**Robotics**]: Robot Map Quality.

General Terms
Algorithms

Keywords
Harris corner detector, Scale Invariant Feature Transform, Hough Transform

1. INTRODUCTION
Assessing the map quality generated by the robots is one of the useful ways to assess the capability of a robot's understanding of its surrounding environment. Robot map quality is one of the quantitative measures which can be helpful in determining which robots will perform better in the field. When a robot moves, it generates the map which helps in its localization and planning. Normal image noise measures are not suitable to assess the map quality because the differences in maps are structural in nature. To assess these kinds of maps we had assumed that the quality is defined not in terms of overall image but on the structural detail contained inside the image. This can be observed in the ground truth image as shown in figure 1 while one test map is shown in figure 2 which was generated by the robot. So we consider a map as accurate even if there is noise and distortions present but it contains all the salient details of the ground truth.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.
PerMIS'08, August 19–21, 2008, Gaithersburg, MD, USA.
Copyright 2008 ACM 978-1-60558-293-1…$5.00.

There is not much work done in this field. Most of the work done is either in the field of image quality measure or in the range data quality measure for moving robots. The initial work was done by Chandran et al in [12] and [13]. Their method measures the quality of the map from 3D point cloud generated from the robot. This point cloud is further classified into plausible and suspicious patches using the conditional random fields. The number of suspicious patches is used to calculate the quality. Although this method seems promising but it is dependent upon the mapping where the data generated is in 3D point cloud format. The testing was done with data generated from 3D laser scanners.

In [14] the authors have used the polylines to model the shapes from the robot generated maps and utilization of these models for the solution of the Simultaneous localization and mapping (SLAM) problem. Although this paper is not directly related with the map quality problem but it provides an interesting insight into map generation which can be used to identify different parts of the map. As identification of different regions inside the map can be helpful in assessing the map quality.

Recently a manual map evaluation toolkit," Jacob's Map Analysis Toolkit" has been suggested by [16]. This application provides the manual map viewer which is designed in a way to help the quality assessor to judge the quality of the map by overlapping the maps over each other. This toolkit also provides a simple measure to assess the map quality based on evolutionary algorithms. However, using the evolutionary algorithms does not guarantee that the results for the map quality will be similar when used each time.

Another recent development is the usage of localized features for map quality assessment [17]. Authors have used the rooms as the localized features. They propose an algorithm to detect the rooms and then find the map quality using these localized features. But it is not necessary that a map will always contain features such as rooms and room like structures. Most of the time the robot generated maps do not contain lines but the collection of point clouds generated from the sensors. In that case, it will be difficult to identify rooms.

Our proposed algorithm is based on techniques which try to cover most of the short comings found in other algorithms. There are still many cases where our algorithm can fail. The reason to use multiple features is that if one kind of algorithm fails in some specific type of map, there are still two other options to judge the map quality correctly.

Some of the limitations are because of inherent nature of the algorithms used which will be discussed later in this paper.

Rest of our paper discusses our algorithm which consists of three separate sub parts. Our algorithm can be described in following steps:

1. Generation of the localized features on the map.
2. Similar features are identified in Target image and ground truth image.
3. Quality is calculated from the final number of matched features.

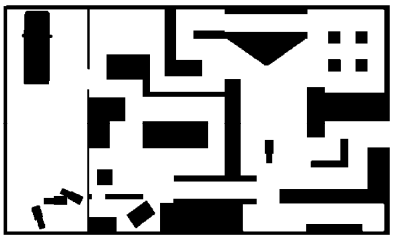

Figure 1. Ground Truth Map.

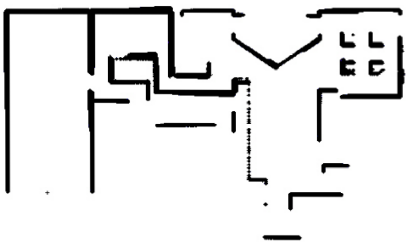

Figure 2. Robot generated Map

2. MAP QUALITY

2.1 Introduction

Measuring the map quality is a very difficult task because it is difficult to define quality in terms of the image. There can be different criteria to define the map quality. Some of them can be based on the noise generated in the maps and on the other side some can be based on the rotation and translation observed in the generated map. There can be many ways to define the map quality. But one important factor which we want to measure in the map quality for robot is the structural details in the map, so although there might be some other noise in the map it is assumed that any map is accurate if it thoroughly represents all the important structure features when compared to the ground truth. So to assess this measure we cannot use the noise to signal ratio or some other measures. We are proposing a novel method to assess the map quality based on three separate algorithms each corresponding to different type of features found in the map. These are Harris Corner Detector, Hough Transforms and Scale Invariant Feature transform. These measures will gives us three values which can be used to assess the quality of the map in three different terms.

2.2 Harris based algorithm

Our first algorithm is defined on the principals of closest point matching. Let us assume we are given two images to compare named X and Y. To compare these images we need to find interest points in these images. These images are binary so we have limited choice in selecting the interest point algorithms. Most of the interest point detectors work on gray scale or color images. The interest points should be useful with enough detail so that they can be compared with points in other image.

Corner detectors are effective in case we have binary images so we have chosen Harris corner detector [8] [9]. This algorithm is very effective in capturing corners and is effectively invariant to rotation, scale, illumination variation and image noise. This is a desirable metric which will enable us to deal with minor noise, rotation and scale problems in the map, see figure 3.

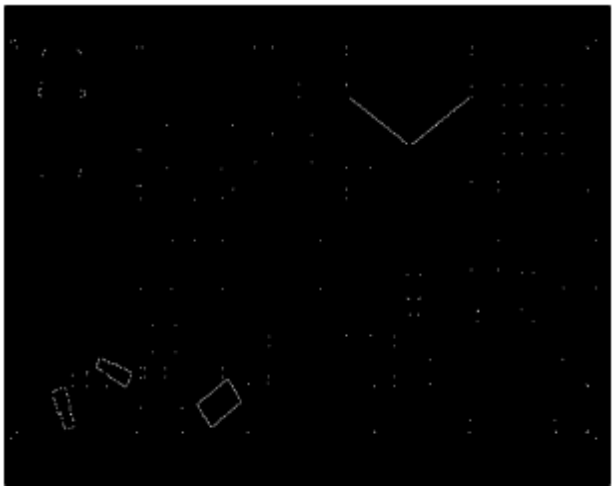

Figure 3. Harris corner detector

After calculating the interest point using the Harris corner detector, we use the closest point matching process to generate the vector maps which are later used for calculating the quality metric. To generate the vector map we find the corners which are closest to the point under consideration and then use that point in map and find its closest point in the ground truth and eliminate those points from both maps with increase in the value for true points matched counter for the map quality.

2.3 Hough based algorithm

To account for the structural detail we have used Hough transform [5] [6] to transfer the map from Euclidean space to Hough space. This has the benefit of identifying lines in the image. These lines are compared according to the position of lines as points in the Hough space. Hough space is created by exchanging the Euclidean coordinates with the parameterized values form the parametric form of the equation of the line.

$$r(\theta) = x\cos\theta + y\sin\theta \tag{1}$$

This helps in identifying lines easily as in the Hough space the points with large values will be highly likely to represent the lines. This same process can be repeated to generate the space for circle and other geometrical objects detection. A variation of the Hough transform which is known as the generalized Hough transform, can be used to detect different type of arbitrary shapes in the

image. This can be used to detect lines, squares (rooms etc), circle (roundabouts etc) in the map which will be a more generalized way to calculate the map quality.
After detection of these features the matching features can be located in the ground truth map and compared for the map quality as described in the last section.

2.4 Scale Invariant Feature Transform

Scale Invariant Feature Transform (SIFT) was introduced by the David Lowe in [15]. Since then SIFT based localized feature have gained prominence among researchers due to there invariability to rotation scale and even dynamic changes. To assess the map quality we have proposed an algorithm based the SIFT. SIFT feature are calculated from extrema detection by finding the extrema points from difference of Gaussian images as shown in equation 2, where the G_m and G_n represent the Gaussian filters at multiple scales and I is the original image. These points are further processed to find out the stable point under various conditions like edge response and low contrast point elimination.

$$DOG(I) = (G_m * I) - (G_n * I) \tag{2}$$

SIFT points detection is the first part of the process, after detection usually a descriptor is calculated and stored for each point so that it can be used to compare point from different images. The length of the SIFT detector is equal to 128 elements, which is basically the directional histogram of the local region.

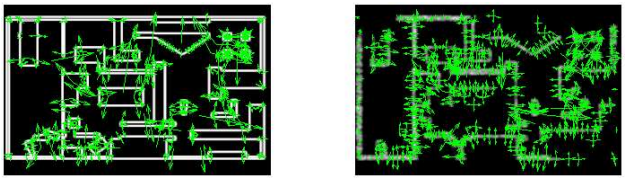

Figure 4. SIFT features on ground truth and robot generated maps.

For our algorithm we have used the following procedure:

1- First the entropy [18] of the image is calculated so that important regions with high entropy are identified. As our maps are binary images it is necessary to convert them into multiple scales with more information so that useful features are calculated.

2- This image is passed on to the SIFT for feature detection and descriptor calculation, see figure 4 for an example of SIFT features.

2.5 Closest Point Matching

Closest point matching is performed by finding the closest point to the corresponding interest points in one image to another. Each point in the ground truth is mapped in a one to one fashion between the ground truth image and the target image. To keep points from matching to a point which is extremely far, the matching is performed only for the points which exist below a specified threshold. So it generates a displacement map for each point from one image to another image. The obvious benefit is the localized identification of the object interest points.

The closest point match can be described by equation 3.

$$\text{Match} = \text{Dis}\,(\text{FV}(\text{P}(\text{x},\text{y})) - \text{FV}(\text{P}_\theta(\text{x},\text{y}))) \tag{3}$$

Where equation 3 describes that the match is the point which is equivalent to the point in one map to the corresponding region in another map under an specified threshold, where FV is the feature vector of the P(x,y) and Dis is the distance between two corresponding feature vectors. Only in the case of the SIFT features the comparing criteria is based on the calculated descriptors.

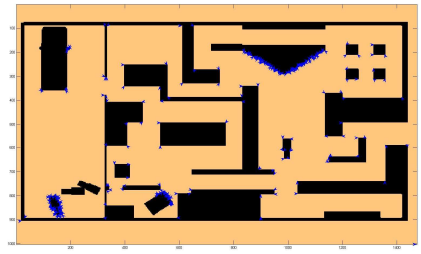

Figure 5. Map showing displacement of interest points.

Figure 6. Displacement in test image.

2.6 Vectorial Space

The displacement or vector map calculated in the last step provides much more information regarding the kind of distortion which appeared in the image. This way this vector map is a localized distortion map in the image. This can be done in both directions to identify the missing features which were not captured and extra features which don't really exist. The figure 5 shows the displacement of closest points in ground truth while the vectorial space is shown in figure 6 for the test image.

2.7 Quality Measure

The map quality measure is calculated using the ratio between the set of features. The map quality can be defined mathematically as

$$q = RMF / GTF \tag{4}$$

where RMF are the number of valid feature points found in the robot generated map while GTF are the number of feature points in the ground truth map.

The map quality obtained from the set of test images (as show in Figure 8) is shown in Table 1 and figure 7.

Image quality assessment is difficult [7] [10] because for each case there can be different criteria to define the quality. For these robot generated maps the most important quality measure is the amount of features or landmarks (points, lines, etc) which are contained in the generated map. That is why we have based our quality measure on the feature having same shapes. We have not

used the texture and color information because the maps are only binary images.

Table 1. Quality values obtained with different algorithms.

Map	Hough	SIFT	Harris
1	0.61818	0.52966	0.68981
2	0.74545	0.55085	0.81481
3	0.41818	0.47175	0.35648
4	0.58182	0.50000	0.67593
5	0.47273	0.41102	0.84722
6	0.50909	0.40395	0.71759
7	0.49091	0.39407	0.40741
8	0.30909	0.45763	0.24074

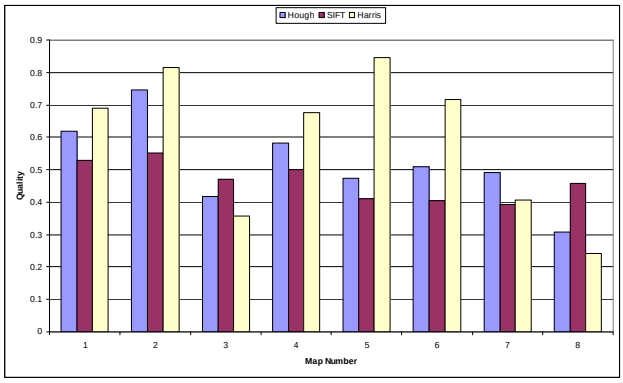

Figure 7. Comparative view of different algorithms.

A very subtle issue is with the finding of the quality of the maps when they are the subset of a larger map. The ground truth is assumed to be the superset of all the maps so it contains all the features and information. So to assess the quality of the map which is smaller than the ground truth, we have to identify the subset from ground truth for which the map was generated. This remains an issue with this algorithm although for maps which are equivalent to the ground truth the algorithm gives fairly accurate results.

Only other remaining issue is the utilization of the threshold. Utilization of threshold can be a problem because we will not be able to match features if the maps are not aligned as in the case of Harris and Hough transform but this is not the case for SIFT based detector because it can detect matches even if they are far away, independent of scale, rotation and dislocation. Although for the Harris and Hough alignment of the map remains an important point. Alignment can be achieved by a startup marker that identifies a stable point between the robot generated map and the ground truth. A map can be considered more accurate if it consistently shows good performance in all three measures.

3. LIMITATIONS

This system is only suitable for offline-measurement for the quality of the maps. As per definition the measure of quality is very difficult to define because requirements on which the map quality is based can be changed according to the need.

This algorithm measures the quality only on the basis of the information content of the image. These maps only contain bi-level images without any additional information. Map distortions and noise are not considered because the information is intact even with the added noise.

Some of the limitations which are observed are due to the type of maps used for processing. If the map has noise, such as, a jagged line or map with distortions, most likely the Harris corner detector will find lots of corners which could give erroneous results. Also Hough transform will fail for the case when point cloud data is separated quit far apart. Similarly for the SIFT case, if there is too much noise in the maps, this will introduce additional features which can cause problems during comparison of the features, because closely related features will give similar results.

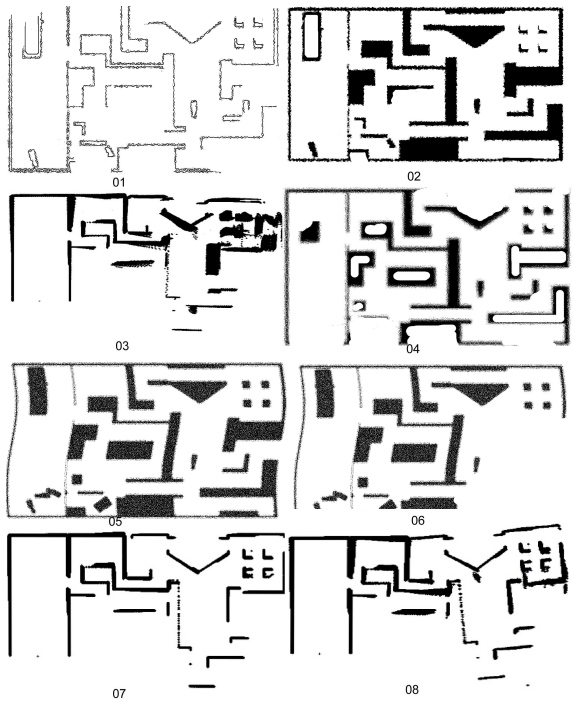

Figure 8. Maps used for the comparison.

4. CONCLUSION

We have tested our algorithm on the test map images generated by robots. The map images are also augmented with additional set of artificially created images to check the quality.

In conclusion, we have devised an automated method to calculate the quality of the maps generated by the robot. We have used the Harris corner detector to detect the interest points while we have used Hough transforms to detect lines, another important localized feature which we have used is scale invariant feature transform (SIFT). In the end we propose the three measures that can define the map quality in three separate terms. We also provide a vectorial map that basically tells us the local distortions found in the image.

ACKNOWLEDGMENTS

We would like to acknowledge Chris Scrapper and Raj Madhavan for their help and the map data. We would also like to thank the SIMA program and the IDUS program for supporting this work.

REFERENCES

[1] B. Girod, “What's wrong with mean-squared error,” Digital Images and Human Vision, A. B. Watson, Ed. Cambridge, MA: MIT Press, 1993, pp. 207-220.

[2] P. C. Teo and D. J. Heeger, “Perceptual image distortion,” in Proc. SPIE, vol. 2179, 1994, pp. 127-141.

[3] A. M. Eskicioglu and P. S. Fisher, “Image quality measures and their performance,” IEEE Trans. Commun., vol. 43, pp. 2959-2965, D

[4] M. P. Eckert and A. P. Bradley, “Perceptual quality metrics applied to still image compression,” Signal Processing, vol. 70, pp. 177-200, Nov. 1998.

[5] J. Illingworth and J. Kittler, “A survey of the Hough transform,” *Computer Vision, Graphics, and Image Processing* 44, no. 1 (1988): 87-116.

[6] D. H. Ballard, “Generalizing the Hough transform to detect arbitrary shapes,” *Pattern Recognition* 13, no. 2 (1981): 111-122.

[7] Z. Wang and A. C. Bovik, “A universal image quality index,” IEEE Signal Processing Letters, vol. 9, pp. 81-84, Mar. 2

[8] C. Harris and M. Stephens, “A combined corner and edge detector,” *Alvey Vision Conference* 15 (1988): 50.

[9] M. Trajkovic and M. Hedley, “Fast corner detection,” *Image and Vision Computing* 16, no. 2 (1998): 75-87.

[10] Z. Wang, A. C. Bovik, and L. Lu, “Why is image quality assessment so difficult,” in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing Orlando, FL, vol. 4, May 2002, pp. 3313-3316.

[11] J. L. Mannos and D. J. Sakrison, “The effects of a visual fidelity criterion on the encoding of images,” IEEE Trans. Inform. Theory, vol. IT-4, pp. 525-536, 1974.

[12] M. Chandran and P. Newman, “Motion Estimation from Map Quality with Millimeter Wave Radar,” Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on (2006): 808-813.

[13] Manjari Chandran-Ramesh and Paul Newman, Assessing Map Quality using Conditional Random Fields, International Conference on Field and Service Robotics (FSR), Chamonix, July 2007.

[14] D. Wolter and L. J. Latecki, “Shape matching for robot mapping.” Proceedings of 8th Pacific Rim International Conference on Artificial Intelligence, Auckland, New Zealand, August (2004).

[15] D. G. Lowe, "Distinctive image features from scale-invariant key points", IJCV, vol. 2, no. 60, pp. 91-110, 2004.

[16] Ioana Varsadan, Andreas Birk, and Max Pfingsthorn, "Determining Map Quality through an Image Similarity Metric", RoboCup 2008: Robot WorldCup XII, L. Iocchi, H. Matsubara, A. Weitzenfeld, C. Zhou (Eds.) Lecture Notes in Artificial Intelligence (LNAI), Springer.

[17] Johannes Pellenz,Dietrich Paulus, “Mapping and Map Scoring at the RoboCupRescue Competition “ , Robotics Science and Systems Workshop, 2008.

[18] Gonzalez, R.C., R.E. Woods, S.L. Eddins, *Digital Image Processing Using MATLAB*, New Jersey, Prentice Hall, 2003, Chapter 11.

[19] C. J. van Rijsbergen (1979) Information Retrieval (London: Butterworths)

[20] Kondrak, G., Marcu, D. and Knight, K. (2003) "Cognates Can Improve Statistical Translation Models" in Proceedings of HLT-NAACL 2003: Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, pp. 46--48