Robust 3D Line Extraction from Stereo Point Clouds

Zhaojin Lu, Seungmin Baek, Sukhan Lee
School of Information and Communication Engineering
SungKyunKwan University
Suwon, Korea
luzhaojin@gmail.com smbaek@ece.skku.ac.kr lsh@ece.skku.ac.kr

Abstract— The paper describes a robust method to extract 3D lines from stereo point clouds. This method combines 2D image information with 3D point clouds from a stereo camera. 2D lines are first extracted from the image in the stereo pair, followed by 3D line regression from the back-projected 3D point set of the images points in the detected 2D lines. In this paper, RANdom SAmple Consensus (RANSAC) is used to estimate 3D line from the 3D point set, the Mahalanobis distance from each 3D point to the 3D line is derived, and the statistically motivated distance measure is used to compute the support for the detected 3D line. Experimental results on real environment with high level of clutter, occlusion, and noise demonstrate the robustness of the algorithm.

***Keywords*—3D line extraction, stereo image, robust regression**

I. INTRODUCTION

Many applications in home service robotics require robust line features by stereo vision. The paper presents an algorithm for robust 3D line extraction from stereo point clouds. 3D Line feature is one of the most important features for 3D object recognition in computer vision area. 3D object recognition has been one of the major problems in robotics field and studied intensively recently. Many researchers have proposed various 3D object recognition approaches [1], among them, the model-based recognition method is the most popular one for dealing with the recognition and pose estimation, In home service robot applications, a home working environment exhibits a different spectral signatures that naturally are occurring phenomena. The fact that in-site environment and man made objects (e.g. building, furniture, refrigerator, etc.) display most straight edges bring line features to critical important features in object recognition.

Computer stereo vision is popular used in mobile robotics, and two cameras take pictures of the same scene, but they are separated by a distance - exactly like our eyes. Compared with laser rangefinder, stereo systems are inherently inexpensive, have small dimensions, and may provide 3D information at full frame rate and simultaneous acquisition of range data and images which can find visual features while measuring the distance to object. A lot of stereo devices have been commercialized, such as Videre, Bumble Bee, etc., a popular product, the Videre camera has been used in our work.

Unfortunately, the quality of stereo data is not accurate and each stereo point cloud has uncertainty, due to the modeling process from perspective images that uncertainty of an object point depends on heavily on its location in object space with respect to the cameras [6]. Correlation-based technique has been adopted by the Videre system. An analysis of errors associated with searching for disparities using a correlation window was presented in [2]. As we know that stereo system can not work well with poor textured scene, and another disadvantage of stereo system is that the disparities (differences between corresponding points in the two images) are, mainly, parallel with the baseline of the system (the line that joins the optical centers of the two cameras), the search for corresponding points thus needs to be done along lines parallel with this baseline.

Due to the poor accuracy of stereo data, using stereo range data alone would not get acceptable 3D lines. Our strategy is that to combine the information both from 2D image and stereo point clouds (indeed, this is an additional advantage of stereo system). All lines are firstly extracted from 2D image and then back-project these 2D lines to the 3D point clouds to get the subset of 3D points. RANdom SAmple Consensus (RANSAC) which is an algorithm to estimate parameters of a mathematical model from a set of observed data which contains outliers. Its principle is well explained in [3][4], RANSAC algorithm has been used in our work for line detection from point set which corresponds to the specific 2D line segment, the main advantage of RANSAC is its ability to do robust estimation of the model parameters with a high degree of accuracy even when outliers are present in the data set. RANSAC algorithm is based on sampling strategies: Hypotheses are generated and their support is measured in the point clouds, which requires computing the support of each point for a given hypothesis, i.e. the probability that the point is explained by the hypothesis. The simplest way is to assign the probability 1 to all points within a certain threshold distance and assign 0 to all other points, as we know that in stereo point clouds, the individual uncertainties of each point are different, so the distance measure must take into account the individual uncertainty, because they are highly inhomogeneous in stereo point clouds. In this paper we analyzed uncertainty of each point according to the range resolution of the Videre stereo camera, and Mahalanobis distance [5] is used here for measuring the unbiased distance from point to hypothesized 3D line, Mahalanobis distance differs from Euclidean distance in that it takes into account the individual uncertainties of each 3D point and is scale-invariant. So the statistically correct way is to use the value of probability density function(pdf) at the given Mahalanobis distance to compute the support of the hypothesis, Gaussian pdf has been adopted in our work, The main contribution of this paper is to develop a statistic approach to extract 3D line robustly from stereo point clouds.

The remainder of this paper is organized as follows: In Section 2, we briefly review related works. Section 3 presents the uncertainty analysis of stereo point clouds. In Section 4 details of robust regression of stereo point clouds are explained, which includes derivation of Mahalanobis distance from a 3D point to a 3D line, together with the implementation detail of RANSAC algorithm. Experimental results and discussions are presented in Section 5. Our conclusions follow in Section 6.

II. RELATED WORKS

Stereo vision provides real time, full-field distance information, and is useful in many applications in a wide variety of fields, including robotics, people-tracking [10], environment modeling[11], obstacle avoidance[12], mapping and navigation[8][9]. Instead of working directly with raw stereo point clouds, feature-based applications firstly transform the raw stereo points into geometric features. Among many geometric primitives, line segments is the simplest one, especially it is easy to describe most of the man-made indoor environment for home service robot. In 3D line extraction, ranges images sensed by laser range finders are frequently used, because dense and accurate 3D points can be obtained, and so many line extraction methods have been proposed based on 2D laser range finder [13]. In general, the handling and analysis of 3D data obtained from range finders require expensive computational costs, because the 3D data is a huge collection of unstructured 3D points. In addition, the range finder itself is still costly and complicated to operate and maintain compared with stereo camera. These are significant factors for using stereo vision system instead of range finder. The stereo matching is of two general types: area based and feature based. Our commercial stereo camera is area based for real time application. Two types of feature based are point-like local features such as corners, SIFT and line segments. Line segment based stereo matching is a choice for 3D line extraction directly [13]. But after extraction of line segments from the left and right images, the procedure for line matching is very complicated. Since we have stereo point clouds from commercial stereo camera, combine the 2D image information and stereo point clouds together, 3D line can be extracted in a robust and fast way. In the vision area, RANSAC algorithm is used to detect mathematical features like straight lines and circles. In the field of automatic buildings modeling based on range information, many authors suggest its use for achieving different tasks. For example, Ameri and Fritsch [7] use RANSAC algorithm for detecting the building roof planes. Forlani [15] apply RANSAC algorithm in order to correct the building roof segmentation result which are obtained using a partition in 8 classes of the gradient orientation. In our case, RANSAC algorithm is used to detect 3D lines.

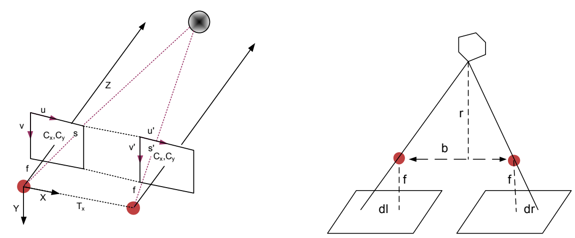
K. Schindler [6] introduced a new approach for robust regression in photogrammetric point clouds. They described the local object point precision and the Mahalanobis distance to plane is derived to allow unbiased regression statistically by taking into account the individual uncertainties of the points. Our work is developed based on Konrad and Horst's work, significant difference between the two approaches is that ours application is real 3D line extraction from real stereo point clouds instead of plane regression by synthetic data.

III. UNCERTAINTY ANALYSIS OF STEREO POINT CLOUDS

A brief introduction of stereo model and uncertainty propagation in the stereo reconstruction process will be presented in this section. Stereo analysis is the process of measuring range to an object based on a comparison of the object projection on two or more images. The fundamental problem is finding corresponding elements between the images. Once the match is made, the range to the object can be computed using the image geometry.

A. Stereo Model and Disparity

For a good understanding of stereo processing, it is necessary to understand what is the stereo geometry as illustrated in Fig.1.



(a) (b)

Figure 1. (a) Basic stereo geometry, (b) Definition of Disparity: offset of the image location of an object.

Fig.1 (a) shows the relationship of two ideal stereo cameras. The global coordinate system is centered on the focal point (camera center) of the left camera. It is a right-handed system, with positive Z in front of the camera, and positive X to the right. The camera principal ray pierces the image plane at C_x,C_y, which is the same in both cameras. The focal length is also the same. The images are lined up, with v=v' for the coordinates of any scene point projected into the images. The distance between the focal points is aligned with the X axis. Fig.1 (b) displays a simplified view of stereo disparity. Two images of the same object are taken from different viewpoints. The distance between the viewpoints is called the baseline **b**. The focal length of the lenses is **f**. The horizontal distance from the image center to the object image is **dl** for the left image, and **dr** for the right image. The relationship of above parameters is:

$$r = \frac{b \cdot f}{d} \qquad \text{where} \quad d = dl - dr \tag{1}$$

$$\frac{\partial r}{\partial d} = -\frac{b \cdot f}{d^2} = -\frac{\left(\frac{b \cdot f}{d}\right)^2}{b \cdot f} = -\frac{r^2}{b \cdot f} \tag{2}$$

$$\Rightarrow \Delta r = \frac{r^2}{b \cdot f} \cdot \Delta d$$

It is easy to see that the range resolution is a function of the range itself (1). At closer ranges, the resolution is much better than farther. Range resolution is governed by the equation (2), the range resolution **Δr**, is the smallest change in range that is discernable by the stereo geometry, given a change in disparity of **Δd**.

B. Uncertainty Propagation of Stereo Points

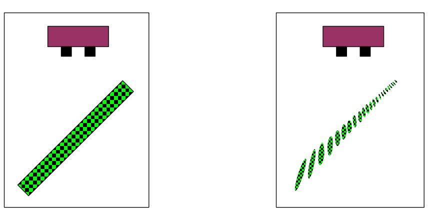
Many problems in computer vision can be couched in terms of parameter estimation from image-based measurements. Such problems arise in stereo vision, with the estimation of the fundamental matrix and in many other areas. Because such problems are typically very sensitive to noise, this uncertainty noise information is usually expressed in terms of covariance matrices. A measured image point is described by its coordinates and covariance matrix:

$$p = [u, v]^T, \qquad \Sigma_p = \begin{bmatrix} s_{uu} & s_{uv} \\ s_{uv} & s_{vv} \end{bmatrix} \tag{3}$$

A 3D point **q** is reconstructed by image points from the left image and right image of the stereo system. The 3D point coordinates and covariance matrix describe the point position in 3D space with a 3D probability distribution. The q is given as (4):

$$q = [x, y, z]^T \qquad \Sigma_q = \begin{bmatrix} s_{xx} & s_{xy} & s_{xz} \\ s_{xy} & s_{yy} & s_{yz} \\ s_{xz} & s_{yz} & s_{zz} \end{bmatrix} \tag{4}$$

In our case, stereo point clouds are generated from Videre stereo camera with short baselines in order to enable automatic matching with area-based correlation. In such a recording setup the depth information has high uncertainty.



(a) (b)

Figure 2. Uncertainty propagation (a) Top view of stereo camera and object, (b) Reconstructed 3D points, given by ellipsoid error bound.

From Fig.2 we can see that the uncertainty of point depends on the camera positions. If the point is close to the cameras, the uncertainty is low; otherwise uncertainty is high.

IV. MAHALANOBIS DISTANCE TO A 3D LINE AND RANSAC ALGORITHM FOR 3D LINE DETECTION

A. Mahalanobis Distance to A 3D Line

In statistics, Mahalanobis distance is a distance measure introduced by P. C. Mahalanobis [5]. It is based on correlations between variables by which different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample set to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant, i.e. not dependent on the scale of measurements.

The parametric equation for a 3D line is:

$$\begin{cases} x = x_0 + a \cdot t \\ y = y_0 + b \cdot t \\ z = z_0 + c \cdot t \end{cases} \tag{5}$$

Where $m_0 = (x_0, y_0, z_0)^T$ is some point on the line and $l = (a, b, c)^T$ is a vector defining the direction of the line, t is the parameter whose value is varied to define points on the line. Since the distance measure is to determine the probability that a point belong the 3D line, the distance measure must take into account the uncertainty of an each individual point. Mahalanobis distance can be used here, and the Mahalanobis distance from an uncertain 3D point x to a given point $\bar{x}$ is defined as:

$$D_M = \sqrt{(x - \bar{x})^T \Sigma_x^{-1} (x - \bar{x})} \tag{6}$$

To get the Mahalanobis distance to a 3D line, we have to apply a whitening transform and normalization to the covariance matrix Σ_q and the 3D line (geometrically means transform the ellipsoid to a unit sphere). We transform the observed point q linearly so that we obtain a new vector q' which is white, i.e., its components are uncorrelated and their variances equal unity. In other words, the covariance matrix of q' equals the identity matrix, the centering is simply a shift from q to the origin $(0,0,0)^T$ of the coordinate. Since covariance matrix Σ_q is symmetric, one popular method for whitening is to use the Singular Value Decomposition (SVD) given the rotation matrix **R** and the scale matrix **V**.

$$\Sigma_q = RVR^T, \qquad \mathrm{V} = diag(a^2, b^2, c^2) \tag{7}$$

R aligns the coordinate system with the ellipsoid axes, while the element a, b, c of V compensate the non-uniform scale along different axes. Then the whitening transform of the point m_0 on the line is (8):

$$m_0' = T(m_0) = V^{-1/2} R^T (m_0 - q) \tag{8}$$

And the whitening transform of the direction vector l of the line can be derived by (9):

$$l' = T(l) = V^{-1/2} R^T l \tag{9}$$

So the Mahalanobis distance $D_M(q)$ from the transformed line (m_0', l') to the center of the unit sphere can be computed by (10), and the center of the unit sphere lies in the origin of the new coordinate system (see Fig.3)

$$\cos\theta = \frac{m_0' \cdot l'}{|m_0'| \cdot |l'|}, \quad D_M(q') = |m_0'| \cdot \cos\theta = \frac{m_0' \cdot l'}{|l'|} \quad (10)$$

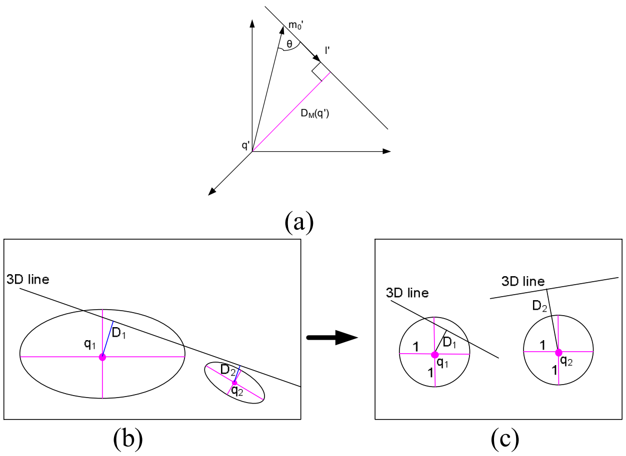


Figure 3. Mahalanobis Distance (a) The distance from a point to a 3D line, (b) ,(c) Comparing distance in the presence of uncertainty of individual point. The probability of belong to the 3D line is higher for point $q_1$than the point q_2 although D_1>D_2. After transforming the error ellipsoids to the unit sphere the distance measures the probability correctly.

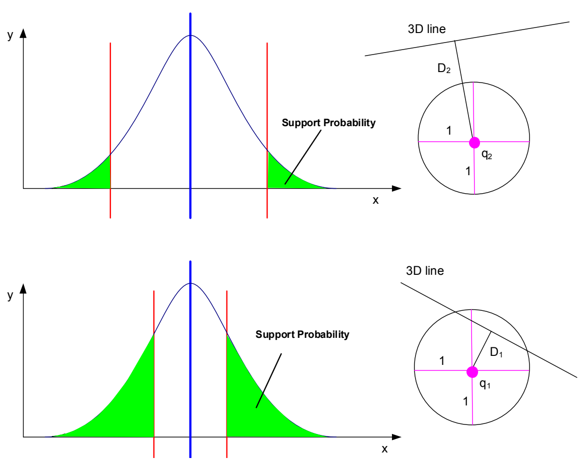


Figure 4. Probability analysis of Mahalanobis distance, below figure has higher probability than the above figure based on the Mahalanobis distance.

Gaussian probability density function (pdf) will be used here for probability analysis of the Mahalanobis distance, and a point support for the 3D line is thus determined by the percentile rank of $D_M(q')$ in the normalized Gaussian probability density function.

$$S(q) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-D_M(q)}^{+D_M(q)} e^{-\frac{t^2}{2}} dt = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{-D_M(q)} e^{-\frac{t^2}{2}} dt \quad (11)$$

The total support for a 3D line in a stereo point set $\{q_1, q_2, ... q_n\}$ is the sum of the support values $S(q_i)$

$$S = \sum_{i=1}^{n} S(q_i) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \int_{-\infty}^{-D_M(q)} e^{-\frac{t^2}{2}} dt \quad (12)$$

B. *RANSAC Algorithm for 3D Line Detection*

The principle of RANSAC algorithm is to search the best 3D Line among a 3D point cloud. It selects 2 points randomly and get the initial parameters of the corresponding 3D line. Then it detects all points of the point set support the 3D line by probability analysis through Mahalanobis distance instead of threshold in Euclidean distance. If the new result is better, then it replaces the saved result by the new one.

Algorithm: RANSAC for 3D Line Detection

Input data:

- 3D point set (*point_set*) which is a matrix of [X Y Z];
- Number of iteration *N*

```
BestSupportProbability = 0; BestLine = NULL
while( iteration <= N )
   pts = Choose 2 points randomly from (point_set)
   Line = pts2Line(pts)
   MahalanobisDistance = dist2Line(Line, point_set)
   Probability = MDist2Prob(MahalanobisDistance)
   SupportProbability = SumProb(Probability)
   if ( BestSupportProbability < SupportProbability)
     BestSupportProbability = SupportProbability
     BestLine = Line
   end
   iteration = iteration + 1;
endwhile
```

In the above pseudo code, **pts2Line** calculates the Line parameter from 2 randomly chosen points, **Dist2Line** calculates the Mahalanobis distance between point set to the given 3D line, **Mdist2Prob** converts the Mahalanobis distance to support probability for a 3D line, and **SumProb** calculates the total support for a 3D line based on (12).

V. Experimental Results

Due to the uncertainty of stereo point clouds, candidate 3D lines are selected starting from the 2D line segment which can be back-projected to the 3D point clouds. The data format of the stereo point clouds is (x, y, z, r, g, b, u, v), in which (x, y, z) is the 3D coordinate of the stereo point clouds, and (r, g, b) is the pixel value (RGB information) of the 2D image at the (u, v) coordinate in the 2D image. So from (x, y, z, r, g, b, u, v) data format, the relationship between the stereo point clouds and 2D image can be corresponded. First of all, the edgels(edge pixels) are detected by Canny algorithm, followed by a simple algorithm which inspired from David Lowe's Link.c function from the Vista image processing library [16] by linking the edge pixels together into chains, and then remerge the broken branches into line segments. Secondly, after extracting 2D line segments from the 2D image, and find the corresponding 3D point set along the same (u,v) coordinate as the 2D line segments. Then these sub-3D point sets will be regressed to a straight 3D lines based on RANSAC algorithm and probability analysis of Mahalanobis distance, which takes account of the uncertainty analysis of stereo point clouds. The Videre stereo camera is used for this experiment, together with the mobile robot (Fig.5).

(a) Videre Stereo Camera **(b)** Mobile Robot

Figure 5. Equipments for experiments.

Figure 6. Experimental results: (a) 2D Image, this image include a juice dispensor. (b) Edge image by Canny algorithm. (c) stereo point clouds . (d) stereo point clouds with the back-projected 3D point set which corresponding to the 2D lines. (e) 2D line segments. (f) 3D Lines from stereo point clouds.

Fig.6 shows that most features of the juice dispenser have been extracted. Compared with 2D lines in Fig.6 (e), some 3D lines are missed in Fig.6 (f). The main reason is that there has no enough or accurate stereo point clouds along the 2D line Segments. From Fig.6(a) we can see that the left-bottom part has no much texture information, then the reconstruction of stereo images is very difficult since correspondences between the left and right image are not good enough, as shown in Fig.6(c) no enough point clouds from the left-bottom part of the scene.

VI. Conclusions

The paper proposes a new, robust algorithm for 3D line extraction from stereo point clouds. This approach combines 2D image information with stereo point clouds together, which is a good additional advantage of stereo system. We analyzed the uncertainties of the point clouds and derived the Mahalanobis distance from a point to a 3D line, motivated from statistically distance measure, Gaussian pdf is used to compute the support probability for the 3D line of each individual point, RANSAC algorithm is used to detect the 3D line from point set. As shown in the experimental results, we extracted robust 3D lines by using the introduced approach.

Acknowledgment

This paper was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea. This work is partly supported by the Science and Technology Program of Gyeonggi province.

References

[1] S. H. Lee, E. Y. Kim, and Y. C. Park, "3D Object Recognition using Multiple Features for Robotic Manipulation," IEEE Int. Conference on Robotics and Automation (ICRA), pp. 3768-3774, May 2006.

[2] Y. Xiong and L. Matthies, "Error analysis of a real-time stereo system," IEEE Int. Conference on Computer Vision and Pattern Recognition (CVPR), San Juan, Puerto Rico, pp. 1087-1093, June 1997.

[3] M. Fischler, and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, vol. 24(6), pp. 381--395. 1981.

[4] D. A. Forsyth and J. Ponce, Computer vision, a modern approach, Prentice Hall, ISBN 0-13-085198-1, 2003.

[5] P.C. Mahalanobis, "On the generalized distance in statistics," Proceedings of the National Institute of Science of India , pp. 49-55, December., 1936.

[6] K. Schindler and H. Bischof, "On robust regression in photogrammetric point clouds," Proc. 25th DAGM Pattern Recognition Symposium (DAGM), Magdeburg, Germany, 2003.

[7] B. Ameri, and D. Fritsch, "Automatic 3D building reconstruction using plane-roof structures," Published by the *ASPRS*, Washington DC, 2000.

[8] M. Agrawal, K. Konolige, and R.C. Bolles, "Localization and Mapping for Autonomous Navigation in Outdoor Terrains: A Stereo Vision Approach," IEEE Workshop on Application of Computer Vision (WACV), Austin Texas, 2007.

[9] K. Konolige, et al, "Outdoor Mapping and Navigation using Stereo Vision," Proceedings of the International Symposium on Experimental Robotics (ISER), Brazil, July, 2006.

[10] R. Munoz-Salinas, E. Aguirre, M. Garcia-Silvente, "People detection and tracking using stereo vision and color," Image and Vision Computing, vol. 25 (6), pp. 995-1007, June, 2007.

[11] S. Y. Park, M. Subbarao, "A multiview 3D modeling system based on stereo vision techniques," Machine Vision and Application, 16 (3), pp. 148-156, May, 2005.

[12] H. G. Jung, Y. H. Lee, B. J. Kim, et al, "Stereo vision-based forward obstacle detection," International Journal of Automotive Technology, vol. 8 (4), pp. 493-504, Aug, 2007.

[13] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A comparison of line extraction algorithm using 2D laser rangefinder for indoor mobile robotics," IEEE/RSJ Proceedings. Int. conference on intelligent robotics and systems, IROS, Edmonton, Canada, Aug, 2005.

[14] G. Karimian, A. Raie, K. Faez, "On Some Geometric and Structural Constraints in Stereo Line Segment Matching," Int. Conference On Intelligent Computing (ICIC), pp. 1202-1208, Aug, 2007.

[15] G. Forlani, C. Nardinocchi, M. Scaioni, and P. Zingaretti, "Complete classification of raw Lidar data and 3D reconstruction of buildings," Pattern Analysis & Applications, vol. 8(4), pp. 357–374, 2006.

[16] A. R. Pope and D. G. Lowe, "Vista: A software environment for computer vision research," IEEE Conference on Computer Vision and Pattern Recognition, pp. 768-772, Seattle, June, 1994.