# A Distance Measurement Method Using Single Camera For Indoor Environments

Zhaleh Sadreddini, Tuğrul Çavdar, and Hossein Barghi Jond
Department of Computer Engineering
Karadeniz Technical University
Trabzon, Turkey
Email: {zh.sadreddini, ulduz, barghi}@ktu.edu.tr

***Abstract*—A distance measurement method for a single-camera condition is proposed. The aim is to implement a way to find the distance of an object in front of the imaging system in indoor environment. In this method, the first step is to extract the appropriate vertical floor line from the snapshots sequence. Moreover, the line passing through the base side of the object is found simultaneously with the floor line. Then, the part of the floor line starting from this point is measured in pixels, by taking the intersection point of these lines, and converted to distance between the camera and the object. The proposed method is applicable for flat indoor buildings with regular lines on the floor, while many buildings are in such a situation. Experimental results are carried out by using MATLAB to estimate the measured distance by the proposed method.**

***Keywords—distance measurement; single camera; Hough transform; indoor environment; line detection***

## I. Introduction

Distance measurement is a key role for an intelligent robot or an agent to understand its workspace. This feature has been embedded in many of Driver Assistance Systems [1, 2] and mobile robot applications [3]. There are many tools and methods used to determine the distance to objects or obstacles for various operating environments. Some tools send signals to object such as laser or ultrasonic rangefinder. However, some tools only receive information about the object position such as measurements based on vision algorithms.

Vision algorithms have been used to extract lines and points to determine the location of robots in [3]. In vision methods, the standard form is to use the stereo systems which require at least two cameras or changing camera position and parameters. Moreover, this requirement increases system cost and power consumption. Therefore, providing effective methods that can be used in single-camera systems are required. Holzmann and Hochgatterer built a prototype mobile phone for indoor and outdoor distance measurement [4]. Lamza, Wrobel and Dziech measured distance with different focus settings [5]. However, they need either modification in the camera or the parameters of the lens. Di, Shang and Wang measured distance with a single rotating camera by using Perspective Projection Model with a relative accuracy about 3% [6]. Sugimoto, Kanie, Nakamura and Hashizume used single camera and ultrasound for 3D localization [7]. Ali, Kurokawa and Uesugi developed a method having a single camera in order to improve the precision measurement in machining technology [8]. Zhang, Stahle, Gaschler, Buckl and Knoll presented an approach based on Random Finite Set (RFS) Statistics to estimate a vehicle's trajectory in complex urban environments by using a fixed single camera [9]. Dong offered a vision measurement method for estimating the distance of the object in static scene, which requires single camera with 3-axis accelerometer sensor and rotated around a fixed axis [10].

In this paper, a method based on a single-camera is proposed where vision techniques is used to measure the distance of objects by assuming camera is stationary. The proposed method can be employed especially for flat indoor buildings with regular lines on the floor. This type of floors can be seen in many buildings such as schools, shopping centers, museums, hospitals and airports that have regular and striated surfaces.

A filter could be applied to reduce noise from the imaging system as in [11]. In this study, Median filter has been employed to improve the image quality. To obtain a binary image, an optimal automatic threshold detection method based on the image histogram is used (the Otsu's method [12]). Also, the morphological operations, which are useful for the separation and characterization of shapes [13], have been applied to the extracted components from binary images.

In this work, it is assumed that the floor lines are regular and easy detectable to reduce the complexity of the distance measurement problem In that case the extraction operation always becomes successful. However, different algorithms are available to extract lines, which are based on a vision-system as Hough Transform (HT) [14], Mei transforms [15] and K-means based method [16]. In this work, Hough Transform approach is used for floor lines extraction. It should be mentioned that HT is known as a powerful tool for graphic element extraction from images because of its robustness for noisy images.

Experiments are carried out in MATLAB platform to observe how much the measured distance by the proposed method converges to the actual distance of the object.

The rest of this paper is organized as follow. In the next section, a distance measurement system based on the proposed method is presented. In the third section, the results of the experiments are demonstrated and discussed. The last section is the conclusion.

## II. Measurement System Implementation

The overview of the proposed distance measurement system is shown in Fig. 1. The implementation of this system can be divided into three main stages: pre-processing, establishment of correspondence, and calculating the distance.

### A. *Pre-processing Stage*

At this stage, the required features of each image are extracted in order to use in subsequent processes. Firstly, the snapshots that taken from the camera are considered in both color and grayscale forms in order to perform next operations in parallel on the both image forms. The color image is used for detection and extraction of the objects, while the grayscale one is used to extract floor lines. To detect the object from the background, the color characteristic is used. Here, the object's color is red. In this stage, the color image is subtracted from the corresponding grayscale one. Then, the object is detected and extracted from the background. Detection of the object with red color is shown in Fig. 2.

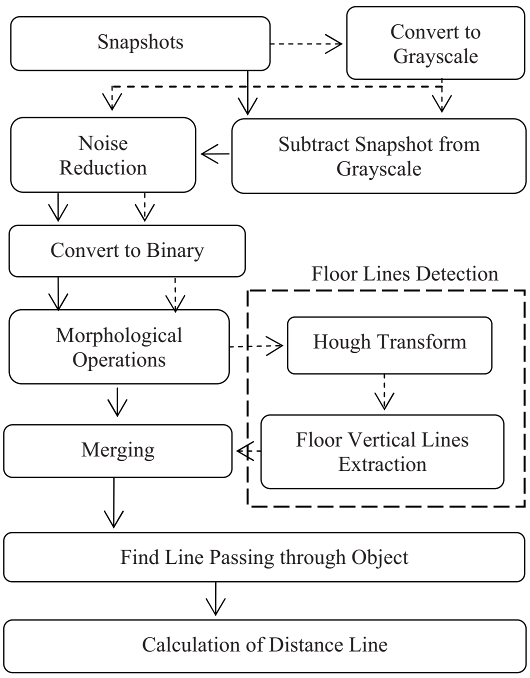
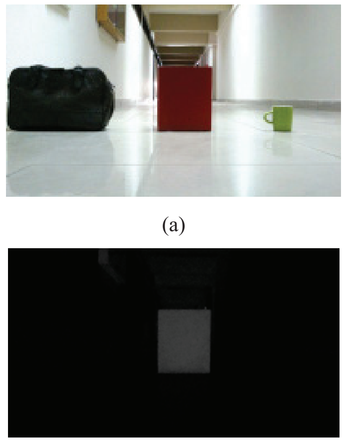



Fig. 1. Flow diagram of the system implementation

Afterwards, Median filtering operation is used to reduce noise in color and grayscale images so that the objects and floor lines can be detectable better. In this filter, the noisy values are replaced with the median value of the neighborhood by the mask.

After filtering operation, the next step is to convert the color and grayscale images to binary images using Otsu's threshold method. This method assumes that the pixels are the components of foreground or background. The threshold value is found as so it minimizes the sum of the foreground and the background distribution. Details are mentioned as follow.



(a)

(b)

Fig. 2. (a) Color snapshot, (b) subtracting color image from corresponding grayscale to detect the red object.

An image can be considered as a two-dimensional function of the intensity of the grayscale levels. This function contains $N$ pixels with grayscale levels from $1$ to $L$. The number of pixels in $i^{th}$ grayscale level is denoted by $f_i$, and the probability of $i^{th}$ gray level in the image will be equal:

$$P_i = \frac{f_i}{N} \tag{1}$$

To find the threshold between two groups of the background and the foreground pixels, the pixels are classified into two classes, $C_1$ with gray levels in the range $[1,\ldots,t]$ and $C_2$ with gray levels in the range $[t{+}1,\ldots,L]$. Then, the probability distributions of gray levels for both classes are calculated as follow respectively:

$$w_1 = \sum_{i=1}^{t} P_i, \quad w_2 = \sum_{i=t+1}^{L} P_i \tag{2}$$

Also, means for the $C_1$ and $C_2$ classes are obtained as follows respectively:

$$m_1 = \frac{1}{w_1}\sum_{i=1}^{t}(i)(P_i), \quad m_2 = \frac{1}{w_2}\sum_{i=t+1}^{L}(i)(P_i) \tag{3}$$

Based on Otsu approach, the intra-class variance of the threshold image is obtained as follows:

$$v_B^2 = w_1(m_1 - m)^2 + w_2(m_2 - m)^2 \tag{4}$$

Where $v_B$ is the intra-class variance and $m$ is the mean intensity for the whole image. The optimal threshold is chosen so that the intra-class variance $v_B^2$ is minimized.

The next step operates on binary images. Binary morphology consists of basic operations: dilation and erosion, and several combinations of them. This operation uses structural elements as 4− or 8−neighbor. However, the structural element values can be 1 or 0. Here, these values are considered as 1.

Dilation causes to expand the connected sets of 1 in a binary image. It can be used for expanding the shapes and filling small holes, gaps and empty spaces. Assuming the mask B is the 3x3 matrix whose all cells are 1, the area of A under B becomes 1 if any cell of B intersects with 1-valued pixel of A. Erosion causes to shrink the connected sets of 1 in a binary image. If all cells of B intersect with 1-valued pixel of A, the area of A under B becomes 1, else all pixels become 0. Thus, erosion can be used for shrinking shapes and removing small protrusions.

Opening and closing operations are useful combination of dilation and erosion. Either closing or conditional dilation operations are used to fill small holes on a binary image. Opening operation can be used to remove the slim connections. In the proposed system, noise has been partially reduced by using the combination of the above operators. The result of these operations is shown in Fig. 3 for an example image.

Fig. 3. Color snapshot after getting binary form with Otsu and applying morphological operations.

### B. Establishing Correspondence

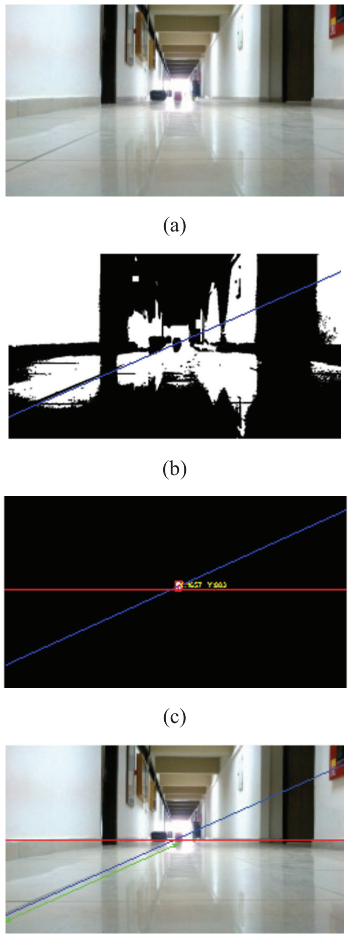
In order to extract the floor lines, the binary form of the grayscale image is sent to the Hough algorithm. To understand the approach, suppose a pixel ($x_p$, $y_p$) in an image that is part of the floor line. It is clear that any straight-line including this point could be shown with the slope-intercept form as below:

$$y_p = m_1 x_p + b_1 \tag{5}$$

Where $m_1$ denotes the slope or gradient and $b_1$ represents the y-intercept of the line. For any second pixel ($x_q$, $y_q$), if the ($m_2$, $b_2$) tuple of the ($x_q$, $y_q$) is the same with the ($m_1$, $b_1$) then the line is passed through both ($x_p$, $y_p$) and ($x_q$, $y_q$). Hough transform uses a selection tool to choose value ($m$, $b$) which gives the line including the most of pixels in the image. The result of the Hough transform on one of the binary grayscale images (Fig. 4 (a)) is shown in Fig. 4 (b).

### C. Distance Calculation

As shown in the flow chart in Fig. 1, the found lines on both color and grayscale images are merged in one image. The aim is to find the distance line in the merged image. Then, the length of the distance line to outer surface will be measured as distance in pixels. The results of mentioned steps are shown in Fig. 4 (c). Also, the desired distance (green line) and measured distance (blue line) are illustrated in Fig. 4 (d).



(a)

(b)

(c)

(d)

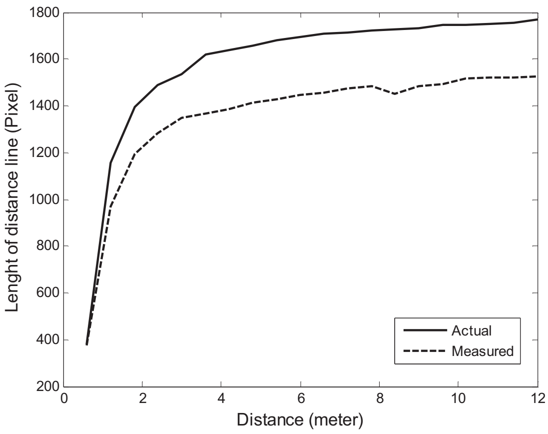
Fig. 4. (a) Input color snapshot, (b) HT results, (c) Finding the line passing through the base side of the object, (d) Illustration of the found lines in the color snapshot.

## III. DISCUSSION

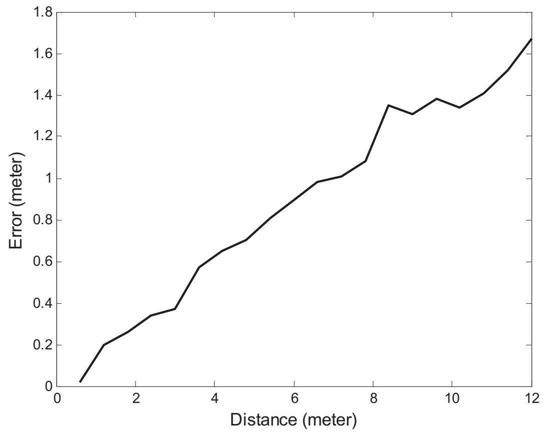
The length of the distance line could be mapped to real distances (for example, in the metric system). The performance of the method is directly related to the HT performance by detecting accurate floor lines. The proposed method is implemented in environmental conditions that HT works well. Fig. 5 (a) shows the real distance compared to the measured distance. Fig. 5 (b) shows the error of the measured distance versus the real distance. Also, this figure proves that the proposed method has a good performance even in the long distance between the camera and the object. As a result, the proposed method is a good candidate to be applied for real applications. It is obvious that the measurement precision has a direct relation with the floor line density.

## IV. CONCLUSION

This paper presents a successful implementation of a distance measurement method based on vision algorithms. The main idea is quite new and it can be implemented with simple image processing techniques. The proposed method is




(a)




(b)

Fig. 5. (a) The real distance versus the measured in pixels, (b) Error between the real and the measured distances in meter

designed for a single-camera condition. This method is applicable for flat indoor environments that have lines on the floor, as many buildings. However, line extraction is the main stage of the method. The distance is calculated in pixels. Experimental results are carried out to observe the relationship between real and measured distances. Results of experiments prove that the proposed method has a good performance and it can be applied to real applications.

## REFERENCES


[1] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, R. Schmidt, T. Graf, High Accuracy Stereo Vision System for Far Distance Obstacle Detection, IEEE Intelligent Vehicles Symposium, 292 – 297, 2004.

[2] D. Seo, H. Park, K. Jo, K. Eom, Omnidirectional stereo vision based vehicle detection and distance measurement for driver assistance system, 39th Annual the IEEE Industrial Electronics Society, 5507 – 5511, 2013.

[3] M. Milkhalsky and J. Sitte, Active motion vision for distance estimation in autonomous mobile robots, IEEE Conference on Robotics, Automation and Mechatronics, 1060 – 1065, 2004.

[4] C. Holzmann, M. Hochgatterer, Measuring Distance with Mobile Phones Using Single-Camera Stereo Vision, 32nd International Conference on Distributed Computing Systems Workshops, 88 – 93, 2012.

[5] A. Lamża, Z. Wróbel, A. Dziech, Depth Estimation in Image Sequences in Single-Camera Video Surveillance Systems, Communications in Computer and Information Science, 368 , 121-129, 2013.

[6] H. Di, Q. Shang and S. Wang, A Virtual Binocular Vision Range Finding Method of Remote Object Based on Single Rotating Angle Indexing Camera, 9th International Conference on Electronic Measurement & Instruments ICEMI'2009.

[7] M. Sugimoto, N. Kanie, S. Nakamura and H. Hashizume, An Accurate 3D Localization Technique using a Single Camera and Ultrasound, International Conference on Indoor Positioning and Indoor Navigation, November 13-15 2012.

[8] H. Ali, S. Kurokawa and K. Uesugi, Camera based precision measurement in improving measurement accuracy, J. Measurement, 49, 138-147, 2014.

[9] F. Zhang, H. Stahle, A. Gaschler, C. Buckl and A. Knoll, Single Camera Visual Odometry Based on Random Finite Set Statistics, 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal, October 7-12 2012.

[10] S. Dong, Vision Measurement Method with Single Camera and 3-Axis Accelerometer Sensor, Fourth International Conference on Intelligent Control and Information Processing (ICICIP) , Beijing, China, June 9 – 11, 2013.

[11] Y. Zhu, Ch. Huang, An Improved Median Filtering Algorithm for Image Noise Reduction, International Conference on Solid State Devices and Materials Science (2012) 609 - 616.

[12] P.-S. Liao, T.-S. Chen, P.-C. Chung, A Fast Algorithm for Multilevel Thresholding, Journal of Information Science and Engineering 17 (2001) 713 -727.

[13] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis, and Machine Vision, PWS, 1998.

[14] J.-Q. Song, M. R. Lyu, A Hough transform based line recognition method utilizing both parameter space and image space, Pattern Recognition 38 (2005) 539 - 552.

[15] T.-W. Mei, Intelligent vision based vehicle side collision warning system on highway, Dissertation, National Chiao Tung University 2006.

[16] H.-S. Lho, J.-W. Kwon, D.-K. Chwa, S.-K. Hong, Vision-based Corridor Line Detection Using K-means Algorithm, International Conference on Control, Automation and Systems (2007) 1052 - 1056.