Distance and Orientation Estimations based on Variable Focal Length Determination and Fisheye Image Modeling

Tewodros A. Aregawi [O*], Oh-Yeol Kwon [*], Soon-Yong Park [**], Sung-Il Chien [*]
* School of Electrical Engineering, Kyungpook National University
** School of Computer Science and Engineering, Kyungpook National University
tewodros@ee.knu.ac.kr, koy0606@ee.knu.ac.kr, sypark@knu.ac.kr, sichien@ee.knu.ac.kr

Abstract

Our method estimates the focal length of a fisheye lens camera in order to adjust the image distortions that occur in the fisheye lens image. We extract matching feature points by using SIFT feature space and further reduce their number so that they become more stable and robust. Using these selected matching points, we estimate the distance moved by the camera based on the pixel movements on the feature space with the help of the estimated focal length. Besides we estimate the orientation angle based on the correlation method between the reference and testing images in which we minimized the correlation error by removing high intensity areas and ambiguous image center regions.

1. Introduction

In our everyday life, mobile robots can be used in many applications, like in surveillance, navigation, services in military operations, and so on. Self-localization of robots is therefore one of the mandatory tasks for a mobile robot to estimate its location and perform subsequent tasks. Localization [1] is the problem of determining the position of a mobile robot from sensor data. In other words, it is the problem of estimating robots' position (location and orientation) relative to its environment [2]. Our work can be applied in the area of mobile robots for navigation and self-localizations because we can estimate the distance and orientation using relatively simple approaches.

In this paper, we present distance and orientation estimation of a camera based on images obtained by the fisheye lens camera. Even though fisheye lens has a wide field of view, it creates distortions [3] [4]. Therefore, we propose the concept of the normalized focal length at different positions of the image to compensate for this problem. This variable focal length is used to calculate the real distance moved by the camera. In estimating the distance, we have used matching feature points from SIFT algorithm [5]. Orientation of the camera relative to its reference position can be determined using the correlation method. The obtained results show that our proposed method can determine the distance moved by the robot and its orientation.

2. Determination of Focal Length Variations on Fisheye Lens System

It is known that fisheye lens images suffer from distortions, i.e., the size or shape of an object becomes smaller when we traverse from the center of the image to the boundary. If we assume that two points on a scene are fixed, the distance between the two points projected around the center of the fisheye lens image is bigger than distance between the two points projected near the boundary. This nonlinearity in the image can be interpreted by the variation of the focal length. In other words the focal length is bigger around the center of the image and it gets smaller around the boundary. To correct these distortions, we have devised an approach of calculating focal lengths at different positions of the fisheye lens image. The approach we have adopted is to find the relationship between distances moved by the camera to the moved image points as in Fig. 1. When the camera is moved as much as *D*, the image points are moved *d* in the image plane from the center. *H* is the height of the camera to the roof and *f* is the camera focal length which is variable in our case depending on its distance from the center of the image. If distance *d* in the image plane is divided by the pixel size of the CMOS sensor, we can convert it to the corresponding

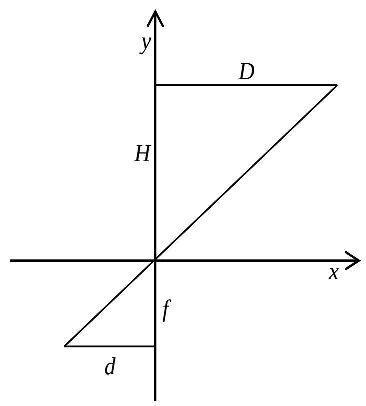


Fig. 1. Image model for estimating variable focal length of fisheye lens

pixel values. Therefore,

$$p = \frac{d}{\mu_s} \quad (1)$$

where μ_s is the pixel size of the sensor and d is the corresponding actual distance travelled on the image plane. We can then have the following simple relation for Fig. 1.

$$\frac{D}{H} = \frac{d}{f} \quad (2)$$

$$f_n = \frac{f}{\mu_s} = \frac{pH}{D} \quad (3)$$

where f_n is the normalized focal length of the fisheye lens camera we have used. Another reason to use the normalized focal length is that even though the dimension of a CMOS sensor is not often known, we still can calculate f_n.

Based on this relation, we have estimated the normalized focal length at different positions experimentally producing a lookup table relating f_n and *p*. These focal lengths will be used to calculate the actual distance moved by the camera at the last stage.

In order to calculate the focal length of the fisheye lens camera, we put a marker object right ahead of the camera. The object then is adjusted to be located at the center of the image sensor. We moved the camera relative to the object marker and we calculate the normalized focal length accordingly. We set up *H* to 140*cm* from the roof to the camera, and *D* varies at intervals of 25*cm* for a total distance of 600*cm*. The pixel distances (*p*) are obtained from the images that are captured at the moved distance (*D*). The estimated results are obtained from Eq. (3), and can be shown in Fig. 2. We can observe from Fig. 2 that the normalized focal length drops as the marker moves from the center of the image to the outer boundary.

3. SIFT Feature Space, and Distance and Orientation Estimation

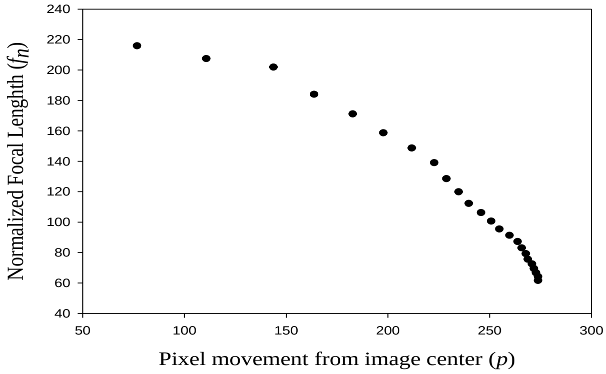


Fig. 2. Graph of normalized focal length vs. pixel movement from center

The captured images by a fisheye camera are used as the input images to estimate the distance and orientation of the camera position. From these images one is a reference image fixed in a single position. The test images obtained from the marker object moved some distance and/or that are rotated with respect to the reference image. In distance estimation, the feature points of the reference and the test images are extracted and are matched with each other. After the matching procedure is done, we calculate how far the camera has moved with respect to the reference image. We use the correlation method to estimate the orientation angle.

3.1 SIFT Feature Space

In many cases, effective matching of feature points between two images is the first step to solving various problems in computer vision. Therefore extracting image features that are stable and robust against noise is quite preferable for suitable matching. SIFT algorithm [5] is one of the popular methods and is used widely for this purpose. SIFT is an algorithm of extracting and describing features in an image. These feature points are highly stable, robust against noise, and also invariant to scaling, rotation, and illumination change. First SIFT features are extracted from the reference image and these features are matched with the test images based on the minimum Euclidean distance measure. Even though the SIFT matching feature points are stable and robust against noises, they still can make wrong matches.

In order to minimize the problems of the mismatching of feature point matches, we have further selected some matching feature points from the SIFT feature space by considering the magnitudes of the matching feature points. The higher the magnitude is, the better the matching will be, because feature points with higher magnitudes are more important and stable points than the others. But ambiguous feature points like points that are affected by high intensity or points that are located around the outer boundary of the fisheye lens camera tend to lead to wrong matches. Based on this observation, we have selected 40 key points from the matching feature points to estimate the distance of the camera movement.

3.2 Fisheye Image Modeling for Distance Estimation

In real world environments, a camera is fixed at a mobile robot and the robot can move in an arbitrary direction. Therefore, for the camera to localize itself effectively, we have come up with a model that holds this fact. We have used the relationship between the actual moved distance of a camera and the corresponding movement in the image plane as shown in Fig. 3. In this figure the camera points to the roof of the environment and it is moved as much as *D*. Then the original scene point at (0,0,*H*) moves to point $A(x_s, y_s, H)$, whereas its corresponding point moves from point (0,0,-*f*) to point $a(x_s', y_s', -f)$ in the image plane.

Thus, the arbitrary movement of the robot is described

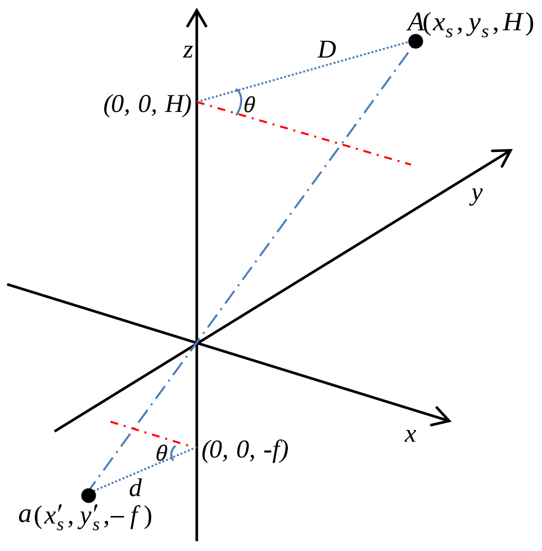


Fig. 3. Pinhole image modeling

in terms of point *a* and *f* which is the focal length of the fisheye lens. From point *A* we can define the distance (*D*) moved by the camera in the real world while image point movement corresponding to point *a* is represented by *d* or *p* which is measured in pixel unit. In the image plane, however, image point *a* corresponding to point *A* is located at *z*=-*f*. For the camera movement in Fig. 4, Eq. (2) still holds true so that we can estimate the distance (*D*) from that equation. But now *D* is $\sqrt{x_s^2 + y_s^2}$, the distance moved by the camera. The movement in the image plane is given by:

$$d = \sqrt{x_s'^2 + y_s'^2} \tag{4}$$

d is the actual distance moved in the image plane as given in Eq. (1). If d is divided by μ_s, we can obtain eq. (1). Assume that a point in a testing image is matched with a point in a reference image and that *p* values of the two points are p_1 and p_2 in the image plane. We can show that robot movement *D* can be obtained by p_1 and p_2. And we have:

$$D = \frac{\Delta pH}{\overline{f_n}}, \tag{5}$$

where $\Delta p = |p_1 - p_2|$, $\overline{f_n} = (f_{n1} + f_{n2})/2$ in which f_{n1} and f_{n2} are the normalized focal lengths corresponding to p_1 and p_2. $\overline{f_n}$ is their average.

For estimating *D* of the camera, f_n must be given in advance when *p* obtained in the test image. This step needs the data of the graph of Fig. 2. We can think of interpolating equation for two variables, *p* and f_n. For simplicity, we use the nearest neighbor method, that is, if *p* is given, we find the nearest *p* in the graph and obtain the corresponding f_n. This is because f_n changes very slightly for most of *p* data.

3.3 Evaluation Methods for Distance Estimation

Using SIFT and the matching algorithm, we have obtained 40 matched feature points linking the reference and testing images. From a pair of matching points, we can estimate the distance (*D*) and this means that we have 40 estimated distances. The first evaluation method is the average method in which the 40 distances are simply averaged. However, to increase estimation accuracy, it is found that a more elaborate method is needed.

We observed that wrong matches, if they happen, provide either smaller or larger distances because these points are located around the center or the outer rim of boundary of the image. Based on our analysis, the distances of some error-prone matching pairs are usually less than the average of the moved points and therefore, we set this average value as the lower threshold. To remove large distances prone to wrong matching, we first calculate the standard deviation (σ) as well as the average (μ). The upper threshold is determined as μ+2.5σ. Thus, we only accept the distances between these two thresholds, as is called selected range method.

3.4 Orientation Estimation

A camera can be rotated in any angle so that its orientation is determined by the image at its current position with respect to the image at its original position. For estimating the angle of rotation of the camera, we have used the correlation between the reference and the testing images. In general correlation is used to compare the similarity of two signals, resulting in a signal that shows this similarity and reaches its maximum at the case that the two signals match best [6].

We then use this correlation method to estimate the rotation angle of the robot as to the reference image. The reference image is rotated by 1° interval and correlated with the candidate image. Therefore, the peak of the correlation of the two images will be estimated as the angle of the orientation of the camera.

As for the Site 3 images (Fig 4. (c)), the symmetry is found for an image and its 180° rotated image and sometimes they may look quite similar. Therefore to overcome this problem and others, we have reduced the high intensity areas to reasonable value with the help of Otsu thresholding method [7]. And we have cropped the center of the image at some radius so that uncharacteristic regions on the image will be removed.

4. Experimental Results

4.1 Distance Determination

Our experiment is performed with four data sets (Site 1, Site 2, Site 3, and Site 4) in which all the heights are the

Table 1. Distance estimation Results with average method and selected range method

Distance (D)	Average method			Selected range method		
	average value	absolute error	Standard deviation	average value	absolute error	Standard deviation
15	11.82	3.5	8.68	15.08	2.61	4.57
30	22.37	7.87	16.45	30.92	6.46	9.95
45	29.8	15.2	19.03	40	9.42	9.13
60	38.06	21.94	24.13	50.61	14.77	10.84
Average		12.13	17.07		8.32	8.62

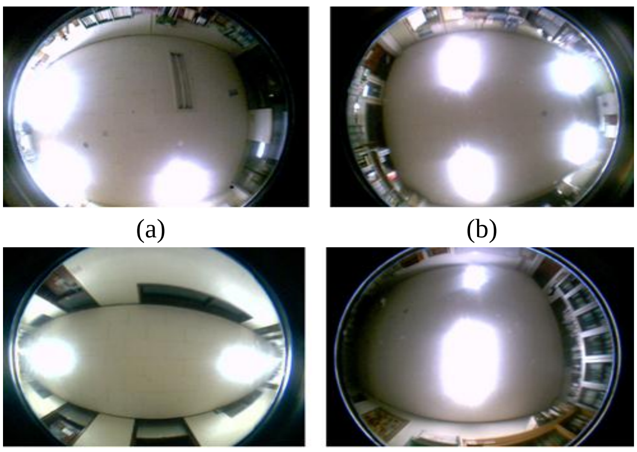
same and fixed to 140*cm*. For estimating the moved distance of the camera, we have used 40 selected matching feature points from the image. The camera is moved 15*cm*, 30*cm*, 45*cm*, and 60*cm* in each path but each data set has different number of objects and different lighting conditions. The sample images of the data sets are shown in Fig. 4.

The distance estimation results obtained by both of our methods are given in Table 1 for all of the four data sets consisting of four different paths. The measurement unit is centimeter. Table 1 contains the numerical values measured in centimeter for the averages of estimated distances, their average absolute distance error and their standard deviation of the four data sets.

From Table 1, we can observe that the average absolute distance error of the first method is around 1.5 times larger than the selected range method, of which the standard deviations are very small. The important point to be noted here is that as a test image is located farther away from the reference image, the accuracy of the distance estimation gets smaller because of severe distortion introduced. Thus, the possibility of mismatching also increases.

4.2 Orientation Angle Determination

The experiment is performed in four different locations where the height of the camera is fixed to 140*cm* from the roof to the camera. The camera is rotated about the center of a reference image with intervals of 30° in a full circle. We use the correlation method discussed in the orientation estimation section. As already discussed, the maximum



(a) (b) (c) (d)

Fig. 4. Sample images taken by fisheye lens camera at different locations: (a) Site 1 (b) Site 2 (c) Site 3 (d) Site 4

Table 2. Orientation estimation results using correlation method for in-plane rotation

Angle (θ)	Average estimated angle	Average error of estimated angle
0°	0°	0°
30°	27.5°	4°
60°	58.75°	4.75°
90°	91.5°	3.5°
120°	125°	6.5°
150°	150.25°	4.25°
180°	178.75°	1.75°
210°	205.75°	4.75°
240°	237.25°	3.75°
270°	272°	2°
300°	304.25°	4.25°
330°	332.75°	2.75°
Average of errors		3.52°

peak of the correlation result is the estimated angle. We have overcome the high intensity or highlight problem by changing the signals to suitably lower intensity values by using Otsu theresholding [6], and by cropping the central part of the image with a radius of 140 pixels, since the high intensity areas contribute to big errors during calculating the correlation result of two images. Therefore, we determine a threshold to isolate the high intensity areas so that we can reduce only the high intensity. The estimation results are shown in Table 2 in which the estimated angle and its error are shown. From Table 2 we can see that the orientation estimations show stable results. When the camera has moved 15*cm*, 30*cm* from the reference image with the 30° interval angles in the full circle rotation, we still have good estimation results. Yet the accuracy of the orientation estimation is not guaranteed when camera is moved beyond 30cm.

5. Conclusions

We have proposed a method of determining the normalized focal length of a fisheye lens system to compensate for modeling lens distortions, which in turn is needed for estimating the distance moved by a camera. Feature points are matched with each other based on the noise tolerant SIFT feature space. Besides we have calculated the orientation of the camera about a fixed reference point using the correlation method. Our results confirm that the results obtained using our proposed methods for estimating the distance and orientation are quite useful for mobile robot navigation.

References

[1] D. Fox, W. Burgard, and S. Thrum, “Active markov localization for mobile robots,” Robotics and Autonomous Systems, vol. 25, pp. 195-207, 1998.

[2] S. Thrun, D. Fox, W. Burgard, and F.Dellaert, “Robust Monte Carlo localization for mobile robots,” int. Proc. of National conference on Artificial Intelligence, vol. 128, pp. 1-49, 2001.

[3] X. Xiong and B. Choi, “Position estimation algorithm based on natural landmark and fish-eyes’ lens for indoor mobile robot,” Communication Software and Networks (ICCSN), pp. 1-5, 2011.

[4] S. Li, “Full-view spherical image camera,” IEEE International Conference on Pattern Recognition, pp. 1-5, 2006.

[5] D. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, pp. 147-151. 1988.

[6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Pearson Prentice Hall, 2008.

[7] N. Otsu, “A threshold selection method from gray-level histogram”, IEEE Trans. On Systems, vol. 9, pp. 62-66, 1979.