The Iraqi Journal For Mechanical And Material Engineering, Vol.17, No4, Dec. 2017

VISION-BASED OBJECT DISTANCE MEASUREMENT USING MONO AND STEREO VISION

Ahmed Shany Khusheef
ahmed_shany@yahoo.com
Department of Electric, Institute of Technology-Kut/ Middle Technical University

ABSTRACT

Visual servoing (VS) is an essential research field because it is included in several robots' applications, such as "navigation and visual surveillance". The VS comprises three stages: taking an image by a visual sensor, processing the image to detect a target object, and guiding the robot to navigate the object. In this work, this issue will be studied to find the object's position in the frame of the image; and then to estimate its distance from the robot frame. The paper will provide the fundamental information of using two techniques: mono and stereo vision. C++ programming in addition to OpenCV libraries were used to develop and implement all codes. The experimental results demonstrated that the distances measured by stereo-vision were relatively accurate with errors ranged from -2.0 to 2.6 % in comparison with mono-vision's measurement errors that were fluctuated between -6.4 and 2.5%. It was also verified that the better results were achieved by using the larger distance between the cameras within the stereo-vision system.

KEYWORDS : Visual servoing, Mono-vision, Stereo-vision, OpenCV, Triangulation, Distance measurement, Image processing

قياس مسافة جسم معين عن الكاميرا باستخدام الرؤيا الاحادية وباستخدام رؤية الاستيريو

احمد شاني خشيف

قسم التقنيات الكهربائية/ المعهد التقني كوت/ الجامعة التقنية الوسطى

الخلاصة

التحكم البصري (Visual servoing) هو أحد المجالات البحثية المهمة وذلك لأستخدمه في تطبيقات عديدة من تطبيقات الروبوتات، مثل "الملاحة والمراقبة البصرية". هذا العملية تشمل ثلاث مراحل اساسية هي: اولا أخذ صورة عن طريق المتحسس البصري، ثانيا تحليل الصورة من اجل الكشف وايجاد الجسم المستهدف داخلها، واخيرا التحكم بالروبوت ليتبع ذلك الجسم. في هذا البحث، سيتم دراسة ايجاد موقع الجسم في الصورة ومن ثم ايجاد المسافة التي تفصله عن الروبوت (او الكاميرا). حيث سيتم تقديم المعلومات الأساسية لتقنيتين مستخدمتين لهذا الغرض هما: الرؤيا الاحادية (تتم باستخدام كاميرا واحدة) ورؤية الستيريو (تتم باستخدام كاميرتين). في هذا البحث تم استخدام لغة البرمجة (C++) بالإضافة إلى المكتبة المفتوحة لمعالجة الصور (OpenCV). أظهرت النتائج التجريبية أن المسافات التي تقاس بنظام الستيريو اعلى دقة بنسبة خطأ تتراوح بين (2- و 2.6%) بالمقارنة مع نتائج القياسات المستحصلة من استخدام كاميرا واحدة والتي تراوحت نسبة خطأها بين 6.4- و 2.5%بالاضافة الى ذلك قد تم اثبات انه افضل النتائج تتحقق باستخدام مسافة اكبر بين الكامرتين .

Received : 2017 - 4 - 10
Accepted : 2017 - 6 - 6

INTRODUCTION

Visual servoing (VS), which is a process of controlling the motion of a robot using the feedback visual information that is sent by a camera, is a significant robotics application **[Alizadeh, 2015].** In VS navigation, it will be difficult for the robot to follow the moving objects if their distances' data are not accessible [Goto, A. and Fujimoto, H., 2008]. In such a case, it is important to calculate the objects' location and orientation with respect to the robot's camera. The extracted information is then fed to the robot's controller that dictates the motion of the robot **[Khusheef et al., 2011A and 2013B].** There are three vision-based techniques used to measure the object distance from the camera **[Alizadeh, 2015].** The first method is called the stereovision in which two cameras are used to estimate the object's distance by measuring the relative position of object from both cameras **[Mrovlje et al. 2008], [Mahammed et al. 2013].** The second technique utilizes a single camera and it is named a mono-vision **[Alizadeh, 2015], [Tao et al. 2007].** The third method is a time-of-flight camera in which the distance of object is calculated by computing the time that the light needs to transfer to and reflect from the target object **[Drayton, 2013]**. The main drawback of third method is that the incoming signal is difficult to be separated because it is dependent on several factors such as the strength of the reflected light and the sensor's dynamic range **. Many** researchers have investigated the field of the vision-based object distance measurement. For instance, **Alizadeh [2015]** developed the single-camera algorithms that can measure the object's distance from the camera based on the image processing techniques. The distance and dimension of the object were first found using the existing techniques of image processing; however, the results were not precise due to the noise and variable light intensity. Therefore, the obtained results were then modified in "the X- and Y-directions" and for the object's orientation around the Z-axis by using the identification methods such as the least square method. The **Alizadeh's [2015]** experimental results showed that the distance measurement accuracy improved to be used within real-time robotic applications. Similarly, **Tao et al. [2007]** presented a mono-vision technique to measure the position of objects using translation and rotation matrices that utilizes the information that is extracted from the image. The near-infrared light was used as a light source to decrease the light variations of the environment and to reach a better contrast between the objects and background. **Krishnan et al. [2010]** demonstrated a complex-log-mapping technique to find the object distance from the camera. This method uses two images acquired by a single camera at two different locations in order to compute the ratio of the object’s sizes on those images **. Some** researchers have studied of using the stereovision to measure the object's distance from the camera. **Mrovlje and Vranci [2008]** proposed an algorithm for the distance measurement based on the images taken by two cameras. The target object should be firstly chosen on the left camera and then it would be detected on the right camera. The **Mrovlje's [2008]** algorithm that used the triangulation technique measured the distance of the object dependent on its displacement in both images. The authors confirmed that the better results were achieved by using longer base distance between the cameras. **Mahammed et al. [2013]** also used the triangulation technique within two cameras to measure the object's distance from the cameras. The authors in that research compared their results with the distance measurements that were done by using a digital laser range finder. The results of **Mahammed et al. [2013]** presented the good agreements with comparative small errors that were between 1.13% and -2.05% **. The** main objectives of this paper is to provide the fundamental information about finding of the object's distance based on the principles of mono and stereo vision techniques; it also offers a comparison between those two methods. It also shows implementation of such techniques by using openCV and C++ programming. The paper firstly gives the fundamental information of mono and stereo vision techniques. This is followed by analyses and discussions of the experimental results; and finally the conclusion of the work is presented.

MONO-VISION MEASUREMENT METHOD

Fig. 1 shows the schematic diagram of the mono-vision measurement technique. The distance (D) of an object from a camera can be measured using the triangle similarity technique **[Rosebrock, 2015]** as following

$$\frac{\mathrm{P\ (pixels)}}{\mathrm{f\ (pixels)}} = \tan\theta = \frac{\mathrm{h\ (cm)}}{\mathrm{D\ (cm)}} \tag{1}$$

in which P is the projection of the object in the image; f denotes the focal length of the camera; and h is the object's real height as shown in Fig. 1. The target object that has a known height h is firstly placed with a known distance D from the camera. The object's apparent height P in the image is then calculated in pixels by using the existing methods of image processing. This allows finding out the perceived camera's focal length f by using the following formula

$$\mathrm{f} = \frac{\mathrm{p} * \mathrm{D}}{\mathrm{h}} \tag{2}$$

The measured focal length f can be then used to determine any distance between the target object and the camera by solving Equation (2) to yield

$$\mathrm{D} = \frac{\mathrm{f} * \mathrm{h}}{\mathrm{p}} \tag{3}$$

STEREVISION MEASUREMENT METHOD

The method in **[Mrovlje and Vranci, 2008] and [Mahammed et al., 2013]** is adapted and used to find the object distance from the cameras; and it can be explained as following. Fig. 2 shows the geometry of two cameras that are aligned to make their optical axes parallel. Those cameras have the same characteristics (i.e. they have the identical focal length and the same view angle). The object's distance (D) can be found by performing some geometrical derivations. Note that, b indicates the base displacement between the cameras and it can be expressed as a summation of distances b_1 and b_2

$$b = b_1 + b_2 = D\tan\theta_1 + D\tan\theta_2 \tag{4}$$

As mentioned above, both camera lenses' optical axes are parallel, which means that θ_1 and θ_2 angles are between those axes and the target object. Using Fig. 2 and applying the basic trigonometry, one can find that $\tan\theta_1 = \frac{x_L}{f}$ and $\tan\theta_2 = -\frac{x_R}{f}$; where x_L and x_R are the horizontal position of the target object in the left and right cameras' images, respectively. In this case Equation (4) will be

$$b = D\left(\frac{x_L}{f} - \frac{x_R}{f}\right) = \frac{D}{f}(x_L - x_R) \tag{5}$$

Solving Equation (5) for distance D, one can yield

$$D = \frac{bf}{x_L - x_R} \tag{6}$$

Using the trigonometry again, it will lead to

$$\tan(\frac{\theta_0}{2}) = \frac{X_0}{2D} = \frac{x_0}{2f} \tag{7}$$

$$f = \frac{x_0}{2\tan(\frac{\theta_0}{2})} \tag{8}$$

in which θ_0 represents the view angle of the camera; X_0 is the width of the camera's view area; and x_0 points to the image's width in pixels. Substituting Equation (8) into Equation (6) yields

$$D = \frac{b\,x_0}{2\tan\left(\frac{\theta_0}{2}\right)(x_L - x_R)} \tag{9}$$

In Equation (9), the cameras' base displacement (the distance between cameras) (b), the width of the image in pixels (x_0), and the camera's viewing angle (θ_0) are constant and known. The difference (disparity) between the distances of projections of the target object in both cameras' images $(x_L - x_R)$ can be also calculated. Therefore, the distance of the object (D) from the camera can be determined. The cameras must be adjusted and calibrated to compensate alignment errors **[Mrovlje and Vranci, 2008], [Mahammed et al., 2013]** and therefore Equation (9) can be expressed as power relation as following

$$D = KX^Z \tag{10}$$

where $K = \frac{bx_0}{2\tan\left(\frac{\theta_0}{2}+\varphi\right)}$ in which φ is an alignment correction parameter that should be measured experimentally; X is Disparity and it equals to $(x_L - x_R)$; and Z represents a constant that will be also determined experimentally.

OBJECT RECOGNITION

In this paper, traffic Signs are used to be recognized in the image and then their distances from the camera are measured. Some image processing techniques are performed to detect and extract the traffic sign information. The Image is first acquired by camera. The noises in the image usually appear due to irregular lighting and shadows; therefore, Gaussian kernel filter is firstly applied on the original image in order to decrease or remove those noises. Then, the acquired image that typically has 3 channels (RGB color space) is changed into grey-scale that consists of only one channel. A threshold process is then applied on the grey-scale image to obtain better results. The thresholded picture is next used to identify contour and the target object later on. Finally, the x-coordinates of the target's center in the image are determined. Fig. 3 demonstrates the steps of the object (traffic sign) detection .

RESULTS AND DISCUSSION

The experiments were carried out to evaluate the accuracy of the object distance measurement techniques that were mentioned above.

Mono Vision Method

The target object (traffic sign) that has 14cm height was placed within 50cm distance in front of the camera. An image was then captured and processed to measure the height of the traffic sign in the image. The focal length of the camera used was then determined using Equation (2). In this case the focal length was 871 pixels and this value was used to measure any distance by using Equation (3). Table 1 demonstrates the experimental results of fifteen tests while Figs. 4A and B show some of the target's images at two positions during the experiments. Table 1 is also showing the real distance of the target from the camera that is premeasured by a measurement tape. It is also presenting the percentage errors between the measured distances. The experimental results demonstrated that the distances measured by mono-vision were fluctuated between -6.4 and 2.5%. The errors were small within the large distances while they raised slightly as distances decrease. However, the errors were not enormous and therefore the measured distances could be considered precise.

Stereo vision

Two cameras' settings were used in the experiments as shown in Fig. 5. The base displacement between the cameras (b) was first taken to be 25cm while the view angle of camera (θ_0) and the width of the image (x_0) were 37.5° and 640 pixels, respectively. The experiments were carried out to determine the parameters of Equation 10. The target's x_L and x_R coordinates values in both cameras were firstly found for a various distances D. The disparity x ($x_L - x_R$) for every distance were then calculated. Note that the real distances were premeasured by the measurement tape. The relationship between the disparity and the real distance of object from the cameras was then plotted by using Excel program as shown in Fig. 6. The equation of relationship curve could be found by using Excel's trend/regression analysis and it was deduced to be

$$D = 8270.2X^{-0.838} \qquad (11)$$

From Equations 10 and 11, one can find that K and Z equal to (8270.2) and (-0.838), respectively. In such a case, the alignment correction parameter (φ) can be found to be 26°. Equation 11 was then implemented by using open cv and C++ to measure the target's distance from the cameras for any disparity value extracted from the image processing. In order to check the written program, fourteen experiments were executed to measure the object's distance from the cameras. Table 1 demonstrates the experimental results that show the measured distances were relatively accurate with errors fluctuated between -2.0 and 2.6 % in comparison with mono-vision's errors that were ranged from -6.4 to 2.5%. The errors were small for short distances and they raised a little when the distances increase. Figs. 7A and B demonstrate two positions of the target in the images of left and right cameras during the experiments.

The same procedure mentioned above was repeated when the base displacement (b) (the distance between cameras) was changed to be 6cm (see Fig. 5B). In such a case, the experimental results show that the errors fluctuated between -10.0 and 10.4% and they were larger than those occurred with the wider base displacement (see Tables 2 and 3 for comparison). This makes agreement with **Mrovlje and Vranci [2008]** who confirmed that "better results are obtained with wider base (distance between the cameras)".

CONCLUSIONS

This paper provides the fundamental information about measuring of the object's distance based on the principles of mono and stereo vision methods. The object detection was first done in stages using image processing techniques: such as Gaussian kernel filter, edge detection, and contour. The distance of the object from the camera was then determined using its information that was extracted from the image. From the experimental results, it was concluded that:

1- The distances measured by the stereo-vision system were more precise than those measured by the mono-vision procedure.
2- The accuracy of the distances measured by stereo-vision system was dependent on the base displacement (the distance between cameras). Better results were achieved by using wider base displacement .

Table (1) Object distance measurement by mono vision method

Test No.	The target's real height (h) (cm)	The projection height (p) (pixel)	The target's center in the image (x, y)	Real distance (cm)	Measured distance (D) (cm)	Error (%)
1	14	244	(360, 232)	50	49	-2.0%
2	14	171	(323, 292)	75	71	-5.3%
3	14	129	(305, 267)	100	94	-6.0%
4	14	104	(297, 349)	125	117	-6.4%
5	14	86	(316, 296)	150	141	-6.0%
6	14	72	(299, 364)	175	169	-3.4%
7	14	62	(266, 372)	200	196	-2.0%
8	14	57	(256, 378)	225	215	-4.4%
9	14	49	(332, 321)	250	248	-0.8%
10	14	45	(329, 318)	275	270	-1.8%
11	14	42	(343, 326)	300	290	-3.3%
12	14	39	(242, 328)	325	316	-2.8%
13	14	34	(269, 388)	350	359	2.5%
14	14	32	(271, 390)	375	381	1.6%
15	14	30	(351, 330)	400	407	1.8%

Table (2): Object distance measurement by stereo vision (the base displacement = 25cm)

No	Real distance D(cm)	X_L (pixel)	X_R (pixel)	Disparity x (pixel)	Measured distance D(cm)	Error (%)
1	75	403	132	271	76	0.8%
2	100	316	124	192	101	1.0%
3	125	305	158	147	126	1.0%
4	150	301	179	122	148	-1.6%
5	175	220	118	102	172	-2.0%
6	200	238	152	86	198	-1.1%
7	225	244	169	75	222	-1.4%
8	250	235	169	66	247	-1.2%
9	275	234	175	59	271	-1.3%
10	300	261	209	52	302	0.6%
11	325	237	190	47	328	1.0%
12	350	243	200	43	354	1.1%
13	375	261	222	39	384	2.4%
14	400	257	221	36	411	2.6%

Table (3): Object distance measurement by stereo vision (base displacement = 6cm)

No	Real distance D(cm)	X_L (pixel)	X_R (pixel)	Disparity x (pixel)	Measured distance D(cm)	Error (%)
1	75	272	188	84	76	0.9%
2	100	304	226	78	90	-10.0%
3	125	309	244	65	138	10.4%
4	150	326	263	63	148	-1.1%
5	175	330	272	58	180	2.9%
6	200	320	265	55	204	2.0%
7	225	326	273	53	222	-1.2%
8	250	329	278	51	243	-2.7%
9	275	336	286	50	255	-7.3%
10	300	347	299	48	280	-6.5%
11	325	348	301	47	295	-9.4%
12	350	350	306	44	344	-1.8%
13	375	341	299	42	383	2.2%
14	400	352	312	40	430	7.4%

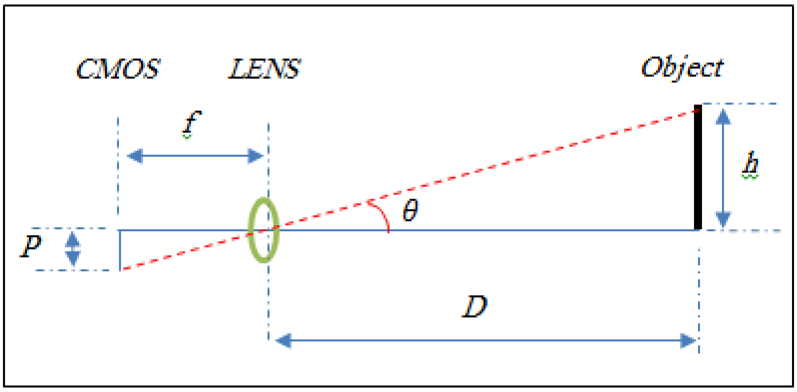

Fig. 1: The geometry of object distance from a single camera

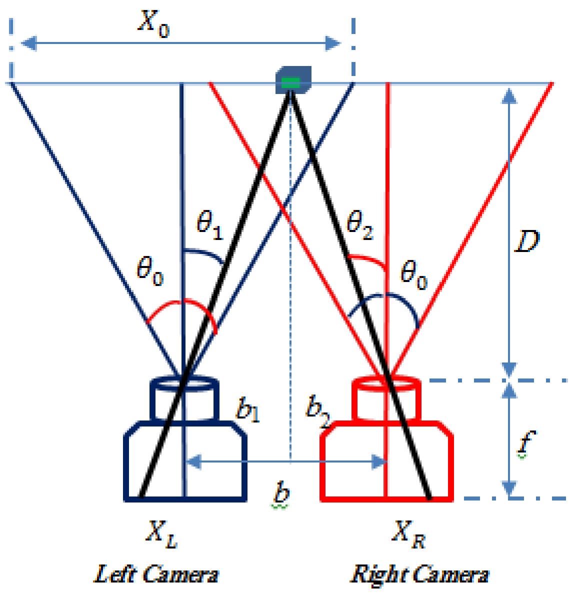

Fig. 2: The geometry of using two cameras

Fig. 3 : (A) RGB color space image, (B) Smoothing image using Gaussians Kernel filter, (C) Gray space images, (D) Binary image after using threshold technique, and (E) Traffic sign (Give way sign) detection using contour technique

(A) 1.5 m object distance from the camera

(B) 4m object distance from the camera

Fig. 4: Examples of the target's position: the original image is on the left side while the processed image is on the right side

(A)

(B)

Fig. 5: The cameras' setting: A) separated by 25cm, and B) separated by 6cm

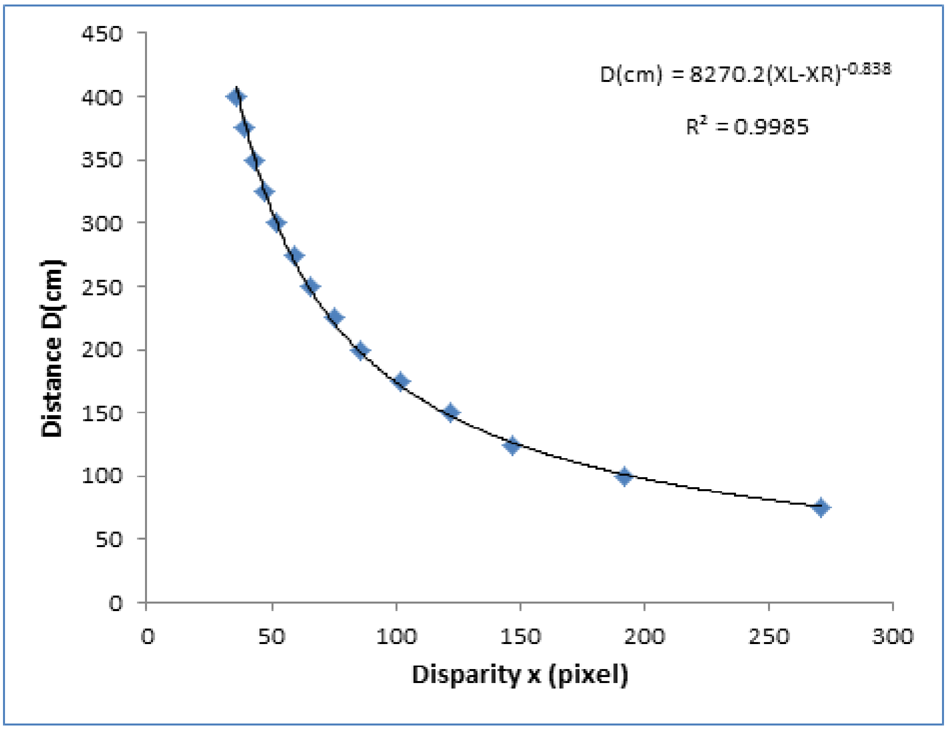

Fig. 6: The relationship between the disparity and the distance of object from the cameras

(A) 1.5m object distance from the cameras

(B) 4m object distance from the cameras

Fig. 7: Examples of the target's position shown in the images of the left and right cameras

REFERENCES

Alizadeh, P., "Object distance measurement using a single camera for robotic applications," Master's thesis, Laurentian University, Sudbury, Ontario, Canada, 2015.

Drayton, B., Algorithm and design improvements for indirect time of flight range imaging cameras. PhD thesis, Victoria University of Wellington, NZ, 2013.

Goto, A. and Fujimoto, H., "Proposal of 6-DOF visual servoing for moving object based on real-time distance identification," SICE. Annual Conference, Japan, pp. 3208-3213, 20-22 August, 2008.

Khusheef, A. S., Kothapalli, G., and Tolouei-Rad, M., "An approach for integration of industrial robot with vision system and simulation software," Journal of World Academy of Science, Engineering and Technology, vol. 58, pp. 18-23, 2011A.

Khusheef, A. S., Kothapalli, G., and Tolouei-Rad, M., "Vision-based relocation of object using a mobile robot," Emirates Journal for Engineering Research, 18(2), p.p1-16, 2013B.

Krishnan, J. V. G. Manoharan, N. and Rani, B. S., "Estimation of distance to texture surface using complex log mapping," Journal of Computer Application, Vol. 3, No. 3, 2010.

Mahammed, M. A., Melhum, A. I., and F. A. Kochery, "Object distance measurement by stereo vision," International Journal of Science and Applied Information Technology (IJSAIT), Vol. 2, No. 2, pp. 05-08, 2013.

Mrovlje, J. and Vranci, D., "Distance measuring based on stereoscopic pictures," 9th International PhD Workshop on Systems and Control: Young Generation Viewpoint, Izola, Slovenia, Oct.1- 3, 2008.

Rosebrock, A., (2015), Available: http://www.pyimagesearch.com/2015/01/19/find-distance-camera-objectmarker-using-python-opencv/

Tao, Z. Changku, S. and Shan, C., "Monocular vision measurement system for the position and orientation of remote object," International Symposium on Photo electronic Detection and Imaging, vol. 6623, 2007.