IJID International Journal on Informatics for Development, *e-ISSN :2549-7448*
Vol. 5, No. 1, 2016

Comparison Analysis between Kinect Grayscale and Depth Image Method

Fajria Antoni
Department of Informatics Engineering
Islamic State University of Sunan Kalijaga
Yogyakarta
fajriaantoni@gmail.com

Naufal Asyhab
Department of Informatics Engineering
UIN Sunan Kalijaga
Yogyakarta
naufalasyhab@gmail.com

Ahmad Subhan Yazid
Department of Informatics Engineering
UIN Sunan Kalijaga
Yogyakarta
yazid.anfalah@gmail.com

Aulia Faqih Rifa'i
Department of Informatics
Faculty of Science and Technology
UIN Sunan Kalijaga
aulia.faqih@uin-suka.ac.id

***Abstract*—The purpose of this research is to know which *method* is more appropriate in the application of *user interfaces*. A data collection method used in this research is the literature study. This literature study method is done by reading books, scientific journals, papers, and using *internet media* as a source of information. Based on the research conducted it can be concluded that from both methods it has advantages and disadvantages. Grayscale method is good for fast processing and image processing. While the depth method is good for applications that require a high level of accuracy and require high resources.**

Keywords-Kinect; Grayscale; Depth Image; User Interface.

I. INTRODUCTION

At this time the interface that is most often used in human interaction with computers is the mouse. The mouse was originally still shaped trackball a very large discovered by Tom Cranston and Fred Longstaff, a Canadian navy in 1952 and used ball bowling. The development of the mouse was so rapid that in 1963 a new mouse method replaces the method trackball was found invented by Douglas Engelbart in 1964 [1]. But the mouse itself still lacked a lot of development, for example, the use of a mouse that was less effective in long-term activities such as design and digital painting. The use of a mouse for a long time causes repetitive stress injuries (RSI) which results in several diseases and abnormalities in the hand [2].

The development of the mouse was so rapid that in 1963 a new mouse method replace the method trackball was found to invented by Douglas Engelbart in 1964. But the mouse itself still lacked a lot of development, for example the use of a mouse that was less effective in long-term activities such as design and digital painting. The use of a mouse for a long time causes RSI which results in several diseases and abnormalities in the hand [2].

The development of technology, especially computers and electronics, triggered the development of more natural input devices. Douglas Engelbart (1964) created a mouse prototype that was first used in the Windows GUI. In 1968 it was created touch screen as an air traffic control device by EA Johnson. In 2006 Nintendo Wii as an evolutionary game controller who previously had joysticks became remote systems that could detect movement. On November 4, 2010 the Xbox Kinect by Primesense was designed for the Xbox console which immediately sold 10 million units in the first month [3].

Kinect is a Depth camera. Unlike ordinary camera, Kinect functions to capture the distance from an object in a room. Kinect uses infrared light to create a depth image. The beam is IR shot at an object and its reflection is captured again by the IR camera. This method is referred to as stereo camera triangulation. The result of an IR camera is a grayscale image. A grayscale image is an image that has a color level. This color level shows how far the object is towards the camera. The whiter the object gets closer, and vice versa [3].

Some devices that have abased stereo cam such as Kinect, for example, PS 3 move, perceptual computing intel. PS 3 move is a device that has a working system combination of stereo cam and accelerometer. Intel Perceptual computing has a system of cooperation with Kinect, namely stereo cam, the difference is that intel perceptual computing focuses on gestures finger and face, while Kinect can detect the entire body and can include 6 users.

Kinect has two methods in getting data depth, namely with color grayscale analysis and depth image analysis. Data depth is the value of the results of the second process of analysis. Data depth is data that contains the value of the distance of objects on each pixel. Grayscale analysis is a method for obtaining distance data of objects in front of Kinect by doing image processing. The brighter a pixel in the image indicates the closer it is to Kinect. analysis depth image to obtain distance data objects that are in front of Kinect by calculating the travel time of infrared light emitted by Kinect which is then stored in the data array [3].

These two methods can be used for applications interface but need to be known aspects needed in application interfaces, namely in the form of object recognition capability, geometry, tracking distance measurement objects, and efficiency. The use and selection of this method is very digital in making software, because errors in the selection of methods result in a system that is heavy, and inaccurate. Previously there had never been a study that discussed this matter, therefore an experiment was needed to be able to compare the two methods.

II. RESEARCH PURPOSES

The purpose of this study is to compare grayscale and depth image method on Kinect so that more of the precise method can be found for making applications and user interfaces.

III. RESEARCH METHOD

The steps used are prelimilinary studies, identification of system requirements, and design experiments.

A. Premilinary Studies

Preliminary studies can be interpreted as a step to obtain information from previous research that must be done, regardless of whether a study uses primary data or secondary data. In the preliminary study, the main target of this study was to develop two methods of image processing of depth Kinect and obtain data on deficiencies and advantages in several aspects.

The data collection method used is the study of literature. Data collection using the literature study method is done by reading books, scientific journals, papers, and using internet media as a source of information. Literature studies are conducted to obtain theories and methods that can be used in this study. Some theories related to this research are stereotypes of triangulation, Depth image, Grayscale image, Open NI, NITE, and Processing Programming.

B. System Requirement Identification

The hardware and software used in this study are:

1) Hardware

Hardware included are Processor AMD Phenom II X2, Memory 4096 MB, VGA ATI Radeon 5770 1029 MB 128 Bit and Xbox Kinect.

2) Software

Software included are Processing 1.5.1, Visual Studio C++ Express Edition, Open NI NITE installer and Library simple open NI.

C. Design of Experiments

In this study we will examine aspects of kinect using two methods. The aspects to be tested are the introduction of geometry, distance testing, testing tracking and system usage.

1) Geometry Shape Identification Experiments, the purpose of these experiments are to measure Kinect ability to identify

geometry shape, which divided into 3 shape, star, cube and circle shape. Star patterns are intended to determine the ability of Kinect to capture angular objects such as in joints of the human body. Data can show the level of accuracy when used to detect angular objects such as elbow, knee folds and several joints other. The cube pattern is aimed at knowing the ability of Kinect to capture objects perpendicular to the raster pixel system. The circle pattern is aimed at knowing ability to Kinect's capture curved objects like those of the human body. Data can show the level of accuracy when used to detect curved objects such as the human body.

2) Distance Experiments, purpose of experiments is to find out which method is more accurate, besides that we can find out the effective distance range for each method. By knowing the level of distance measurement and effective distance, it is expected to be a guideline in the application user interface. An experiment is done by measuring 3 objects, each of which has a different distance. The object used is a beam-shaped object. The shape of the beam is aimed so that the calculation process is easy because the surface is flat. The 3 objects were placed at different distances of 60 cm, 90 cm, and 180 cm. each distance will be tested 5 times and will get an average distance.

3) Tracking Experiments, these experiments aim to determine the level of accuracy of Kinect in the function of reading and tracking movements gesture in the user interface. An experiment is done by tracking or locking 1 criterion of the object caught on the camera and then following the coordinates of the object. The closest object is represented by the pixel brightest in the method grayscale, while the method is depth represented by the data value array smallest. The closest object to this test is the human fingertip. The exhibitor will stand at 3 different distance positions, 60cm, 90cm 120cm.

4) Resource Usage Experiments, the experiments aim to find out which method is more efficient to calculate from memory and processor used so that it is good to be used in an application. An experiment is done by calculating the difference between memory usage and the CPU usage of the idle system until the application runs stable. The idle system is when the system is running stable. Experiment uses program grayscale test, PoC test, grayscale tracking, depth tracking, grayscale range, depth range. Benchmarking resource that is needed at runtime by the system is idle. The higher the value of RAM and needs processor means the worse the results.

IV. IMPLEMENTATION AND RESULT

A. *Research Result*

Data obtained in this study were divided into two, first data in the form of images in testing pattern recognition geometric and tracking. second, data in the form of numbers obtained from distance testing, and system usage.

1) Description of geometric pattern recognition test data

In the study of geometric pattern recognition, we will conduct a study with visual observations, observations are made by comparing data with pattern samples. The final value obtained is whether the pattern is detected or not. If the pattern is seen it will be symbolized by the letter Y, whereas if it is not visible it will be symbolized by the letter N. It can be seen in Table 1-5.

TABLE I. CUBE PATTERN EXPERIMENT RESULT DATA WITH GRAYSCALE METHOD

Distance (cm)	Diameter (mm)							
	50	25	10	5	4	3	2	1
60	Y	Y	Y	Y	N	N	N	N
90	Y	Y	Y	N	N	N	N	N
120	Y	N	N	N	N	N	N	N

TABLE II. STAR PATTERN EXPERIMENT RESULT DATA WITH GRAYSCALE METHOD

Distance (cm)	Diameter (mm)							
	50	25	10	5	4	3	2	1
60	Y	Y	N	N	N	N	N	N
90	Y	N	N	N	N	N	N	N
120	Y	Y	N	N	N	N	N	N

TABLE III. CIRCLE PATTERN EXPERIMENT RESULT DATA WITH DEPTH METHOD

Distance (cm)	Diameter (mm)							
	50	25	10	5	4	3	2	1
60	Y	Y	Y	N	N	N	N	N
90	Y	Y	N	N	N	N	N	N
120	Y	N	N	N	N	N	N	N

TABLE IV. CUBE PATTERN EXPERIMENT RESULT DATA WITH DEPTH METHOD

Distance (cm)	Diameter (mm)							
	50	25	10	5	4	3	2	1
60	Y	Y	Y	N	N	N	N	N
90	Y	Y	N	N	N	N	N	N
120	Y	N	N	N	N	N	N	N

TABLE V. STAR PATTERN EXPERIMENT RESULT DATA WITH DEPTH METHOD

Distance (cm)	Diameter (mm)							
	50	25	10	5	4	3	2	1
60	Y	Y	N	N	N	N	N	N
90	Y	Y	N	N	N	N	N	N

120	Y	Y	N	N	N	N	N	N

2) Description of distance measurement data

From the experiment the distance measurement with gray scale method obtained the following data. It is shown in Table 6-7.

TABLE VI. GRAYSCALE METHOD DISTANCE MEASUREMENT EXPERIMENT RESULT DATA

Grayscale grade (0-255)	Real distance (cm)		
	60	**90**	**180**
Experiment 1	254	233	210
Experiment 2	254	230	210
Experiment 3	253	229	210
Experiment 4	254	228	211
Experiment 5	255	225	211
Average	**254**	**229**	**210**

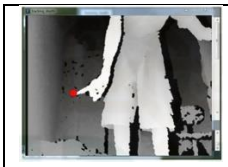
TABLE VII. DEPTH METHOD DISTANCE MEASUREMENT EXPERIMENT RESULT DATA

Grayscale grade (0-255)	Real distance (cm)		
	60	**90**	**180**
Experiment 1	604	902	1800
Experiment 2	609	908	1801
Experiment 3	613	919	1805
Experiment 4	615	923	1807
Experiment 5	628	928	1809
Average	**613**	**916**	**1804**

3) Description of tracking experiment data

Data obtained in this study are images grayscale which are processed into data Boolean which can be seen in Table 8.

TABLE VIII. TRACKING EXPERIMENT RESULT DATA

Image	Method	Distance
8 	*Grayscale*	60 cm
	*Grayscale*	90 cm
	*Grayscale*	150 cm
	*Depth*	60 cm
	*Depth*	90 cm
	*Depth*	150 cm

In that table, Kinect able to do tracking. You can see the tracking area in yellow parking in the method grayscale while the method depth is red.

4) Description of system usage experiment data

Data obtained in this research is numerical data obtained by observing the graph in monitor resources. The applications running on this computer are shown in Fig. 1-3.

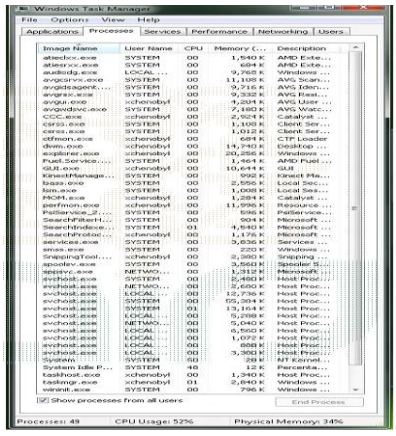


Figure 1. Running Application

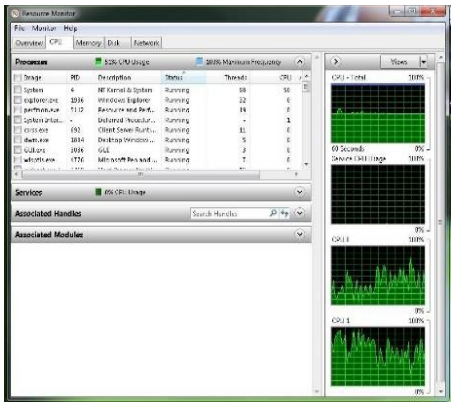


Figure 2. CPU Idle

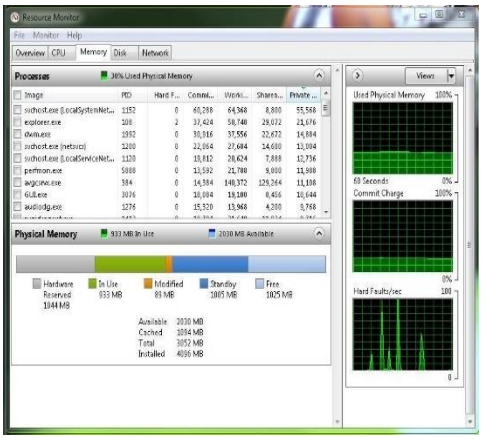


Figure 3. Memory Idle

From 3 software that tested with grayscale method, we obtain following data that are shown in Table 9.

TABLE IX. EFFICIENCY COMPARISON BETWEEN GRAYSCALE AND DEPTH METHOD

Testing	Grayscale	Depth
Pattern	CPU 100%	CPU 100%
	RAM 1093 MB	RAM 1158 MB
Distance	CPU 100%	CPU 92%
	RAM 1130 MB	RAM 1134 MB
Tracking	CPU 100%	CPU 100%
	RAM 1142 MB	RAM 1134 MB

B. *Result & Discussion*

This study was conducted to determine the comparison of Kinect methods grayscale and Kinect depth in several aspects. The first aspect is the aspect of Kinect's ability to recognize patterns. The second aspect is the reliability and accuracy of Kinect in measuring the distance of objects, the third aspect is the ability of Kinect in tracking the last aspect object is to measure the level of efficiency of the Kinect method in its use.

1) Geometric pattern recognition ability aspect

Aspect of Kinect's ability to recognize patterns is intended for user interface applications. For example, to recognize the geometry of the human body or even recognize human finger gestures. From the test results data has been determined which pattern is included is detected and which pattern is not by comparing with the pattern that has been determined. Detected will be represented by the letter Y, while undetectable is symbolized by letter N. They are shown in Table 10-16.

TABLE X. CIRCLE PATTERN EXPERIMENT DATA IN 60CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	Y	N	N	N	N	N
Depth	Y	Y	Y	N	N	N	N	N

In testing the circle pattern at a distance of 60cm in getting both methods have the same level of accuracy

TABLE XI. CIRCLE PATTERN EXPERIMENT DATA IN 90CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	N	N	N	N	N	N
Depth	Y	Y	N	N	N	N	N	N

In testing the circle pattern at a distance of 90cm in getting both methods have the same level of accuracy.

TABLE XII. CIRCLE PATTERN EXPERIMENT DATA IN 120CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	N	N	N	N	N	N
Depth	Y	Y	N	N	N	N	N	N

In testing the circle pattern at a distance of 120cm in getting both methods have the same level of accuracy.

TABLE XIII. SQUARE PATTERN EXPERIMENT DATA IN 60CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	Y	Y	N	N	N	N
Depth	Y	Y	Y	N	N	N	N	N

In testing the square pattern at a distance of 60cm, the method grayscale is better because it can detect up to 5mm diameter accuracy, while the method depth can only detect 10mm diameter.

TABLE XIV. SQUARE PATTERN EXPERIMENT DATA IN 90CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	Y	Y	N	N	N	N
Depth	Y	Y	Y	N	N	N	N	N

In testing square patterns at a distance of 90cm, the method grayscale is better because it can detect up to 10mm diameter accuracy, while the method is depth only able to detect 25mm diameter.

TABLE XV. SQUARE PATTERN EXPERIMENT DATA IN 120CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	Y	Y	N	N	N	N
Depth	Y	Y	Y	N	N	N	N	N

In testing the circle pattern at a distance of 120cm in getting both methods to have the same level of accuracy.

TABLE XVI. STAR PATTERN EXPERIMENT DATA IN 60CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	N	N	N	N	N	N
Depth	Y	Y	N	N	N	N	N	N

In testing the star pattern at a distance of 60cm in getting both methods have the same level of accuracy.

TABLE XVII. STAR PATTERN EXPERIMENT DATA IN 90CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	N	N	N	N	N	N	N
Depth	Y	Y	N	N	N	N	N	N

In testing the star pattern at a distance of 90cm, the method depth is better because it can detect up to 25 mm diameter accuracy, while themethod50 mm grayscale can only detect diameter.

TABLE XVIII. STAR PATTERN EXPERIMENT DATA IN 120CM DISTANCE

Method	50mm	25mm	10mm	5mm	4mm	3mm	2mm	1mm
Grayscale	Y	Y	N	N	N	N	N	N
Depth	Y	Y	N	N	N	N	N	N

In testing the star pattern at a distance of 120cm in getting both methods have the same level of accuracy.

2) Distance measuring aspect

Ability aspects of Kinect capability in measuring distance is intended to determine the level of accuracy of Kinect to distance or depth level. The level of depth is an important matter in depth cam because this function determines the level of good or not the camera in capturing 3d objects. From the test results data has got the following results.

In a grayscale method, Kinect will continue to display images with dynamic brightness levels based on the benchmark of the closest object. In the first case, we placed an object 60 cm in front of the Kinect that we measured with a ruler. The result is a system that displays RGB 249 colors, which means it approaches the maximum white color, namely RGB 254 as shown in Fig. 4. In the second case, we put an object at a distance of 30 cm that we measured with a ruler. The result is that Kinect detects the object with a value of 255 as shown in Fig. 5. The object in the first case changes its value to 226.

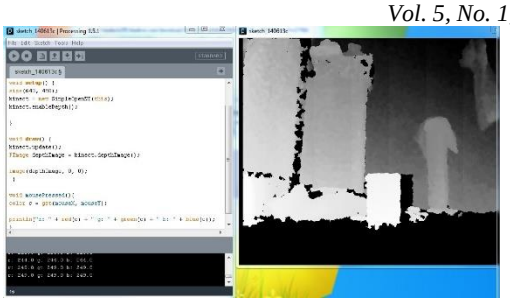


Figure 4. A grayscale method first case experiment

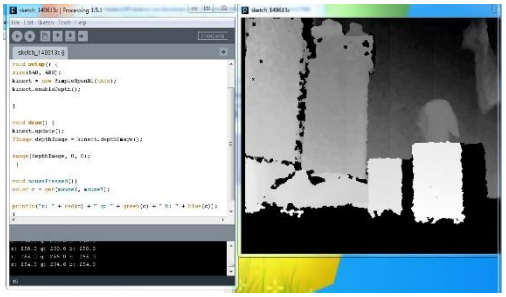


Figure 5. A grayscale method second case experiment

In depth method, the distance of the object in front of the Kinect is measured using infrared rays. The measurement system is done by calculating the time taken by infrared signal particles from the Kinect to reach the object, then captured by the IR cam. The results of this process are stored in data in the form of 2-dimensional array with a length of 435200 and an update rate of 30 times every second and have a variety of values of 11bit. By processing data in this array we can find out the depth or distance of objects from Kinect. The following is the result of a comparison of the calculated data with the grayscale method and the method depth for cube objects in several variations of distance.

In the third case, we accidentally made a random click on a dark object, the upper wall. We obtained the RGB value is 1. So far we have tried, we will never get an RGB 0 value on an object outside the blank object or object that is not detected. With this, we conclude death value grayscale kinetics as much as 255 levels with 1 additional level to represent a blank object.

Figure 6. A grayscale detection area 0 - 255

Figure 7. A blank object detection area

In the Fig. 6-7, it is explained that RGB's value 0 does not necessarily represent the farthest object, but it can also be the closest object and can appear in the middle between RGB's 255 and RGB's 1 in the real world. Whereas technically, RGB's 0 only occurs if IR rays cannot be reflected back and if there is an object exceeding the minimum limit of + - 30cm and a maximum limit of 800cm.

In the grayscale RGB's 255 – RGB's 1 area all objects will be displayed with dynamic brightness levels according to the benchmark of the closest object and the farthest object, in this case the closest object is 60 cm and the farthest object is 3m + - . so it can be concluded that in an area of 60 cm to 300 cm it will be represented by 255 brightness levels. So that every brightness level is 0.9411764705882353 cm. or has a distance accuracy of 0.9411764705882353 cm. or you can count using calculations.

In depth method, the distance of the object in front of the Kinect is measured using infrared rays. The measurement system is done by calculating the time taken by infrared signal particles from the Kinect to reach the object, then captured by the IR cam. The results of this process are stored in data in the form of array 2-dimensional with a length of 435200 and an update rate of 30 times every second and have a variety of values of 11bit. By processing data in the array this we can find out the depth or distance of objects from Kinect. Tables 17-18 are the result of a comparison of the calculated data with the grayscale method and the method depth for cube objects in several variations of distance.

TABLE XIX. A CALCULATION RESULT WITH GRAYSCALE METHOD

Real distance (cm)	Distance in grayscale (cm)	Subtracted result(cm)
60	60.94	+ 0.94
90	84.47	- 5.53
180	102.35	-77.65

TABLE XX. A CALCULATION RESULT WITH DEPTH METHOD

Real distance (cm)	Distance in grayscale (cm)	Subtracted result(cm)
60	61.3 cm	+ 1.3
90	91.6 cm	+1.6
180	180.4 cm	+0.4

From the data and both tables, we get the difference between direct measurement and measurement using both methods. The smaller the difference in the value obtained, the more accurate the method is. From the two methods, it can be concluded that grayscale is better at a distance of 60 cm with an accuracy value of 0.94 compared to 1.3 at depth, at a distance of 90cm depth is more accurate with an accuracy value of 1.6 versus 5.53 in grayscale, at 120cm depth is more accurate with an accuracy value of 0.4 compared to 77.65 on grayscale.

3) Tracking capabilities aspect

The aspect of Kinect capability in tracking is intended to determine the level of ability to track and follow the tracked object. In user, interface capabilities are tracking very important to capture movement and user gesture. From the test results data has got the following results.

In the grayscale method, accuracy is tracking much lower compared to depth because grayscale tracking done by analyzing colors, while depth is done by reading data values. But the very high level of accuracy makes tracking not smooth. In the method depth, the circle point often moves or it looks vibrating and often jumps to other objects that have similar depth levels.

Tracking jumps can be caused by other factors, namely the update rate factor. Kinect has an update rate of 30x in 1 second. But these two methods have different characteristics in each update. In the grayscale update, 30 frames of images are carried out every second. So that every 0.0334 seconds Kinect will give 1 whole frame to disperse. because the frame that is received in full will not be distorted.

In the depth method, the system works by reading data in an array. The data array contains the depth value for each pixel, if the data is calculated array based on Kinect resolution, which is 640 * 480, it will get length an array of 307,200. as explained previously, Kinect updates the rate 30 times in 1 second, or it can be interpreted Kinect will fill the data array as much as 307200 x 30 = 9,216,000 in 1 second.

From the explanation above, we get the basic difference, namely the depth of processing 1 frame intact, while the depth data that is processed is every 1 data array. Table 19 shows the test results table of tracking data.

TABLE XXI. A TRACKING EXPERIMENT RESULT

No.	Experiment	Accuracy	Vibration	Jump
1	*Grayscale* 60 cm	Less	None	None
2	*Grayscale* 90 cm	Good	None	None
3	*Grayscale* 150 cm	Good	None	None
4	*Depth* 60 cm	Good	Yes	None
5	*Depth* 90 cm	Good	Yes	Yes
6	*Depth* 150 cm	Good	Yes	Yes

From the data and tables above, it can be concluded that the method *depth* is better at the level of accuracy. The grayscale method is more stable in following both stationary and moving objects.

4) Efficiency level aspect

Aspect of Kinect ability in the level of system efficiency is intended to find out which method is suitable to be used in making the application as needed. From the test results data has obtained the following results.

The data in Table 20 is obtained by reducing the runtime value with the idle value. The smaller the CPU and ram value the better.

TABLE XXII. A USAGE EXPERIMENT RESULT

Experiment	**Grayscale**		**Depth**	
	***CPU* (%)**	**RAM (MB)**	***CPU* (%)**	**RAM (MB)**
Pattern	49	100	49	165
Distance	49	137	41	141
Tracking	49	194	49	141

From Table 20, it can be concluded that grayscale method good in the level of RAM efficiency in testing patterns and distances, while depth method more RAM efficient in tracking testing. Depth method is more efficient CPU on distance testing.

ACKNOWLEDGEMENTS

Thank you and thank you for the presence of Allah Almighty, who always bestows His mercy, guidance and blessings on us. Thanks also to parents, people closest to us who always provide moral and non-moral support, so we were able to complete this paper.

REFERENCES

[1] D. Engelbart, "Doug Engelbart Institute," 2008. [Online]. Available: http://dougengelbart.org/.

[2] F. Dernoncourt, "Replacing the computer mouse," Oct. 2014.

[3] G. Borenstein, *Making Things See*. Cambridge: O'reilly, 2012.