3D Sign Language Translator Using Optical Flow

Riyanto Sigit, Setiawardhana, Dyah Rahma Kartika
Department of Informatics and Computer Engineering
Electronics Engineering Polytechnic Institute of Surabaya
E-mail : riyanto@pens.ac.id, setia@pens.ac.id, dyahrk@ce.student.pens.ac.id

***Abstract* – Communication is required to interact and socialize. Generally, communication used verbal communication in spoken or written words. However several people are unable to communicate verbally, since they were deaf or mute. Non-verbal communication is the most effective communication used by the deaf and mute. This communication uses sign language in the form of hand and body gestures. Hand gestures identification is the method to identify the position pattern of the hand and the body in order to translate the meaning and the purpose. This communication is intended to make normal person understand regarding the sign language that conveyed by the deaf. Many difficiluties faced by normal person to understand the sign language. Therefore interpreter tool for sign language is required to translate into easy-understood verbal language. This research only includes sign language from hand and body gestures, in order to translates into characters. The configuration of hand gestures detection must be done to describes Indonesian Sign Language (BISINDO). The optical flow method has been used in this research in detecting hand and body movements. By using this method, 65% from 20 different word has been succeeded to identified.**

Keywords; sign language, image processing, bisindo, optical flow, 3d, motion detection, hand and body movements

I. Introduction

Communication is needed to interact and to socialize. In general, communication is used verbally, that is to use spoken or written words. But in everyday life, some of people are not able to communicate verbally. Some of them have a problem in communication, such as deaf and mute. The most effective communication used by deaf and mute are through non-verbal communication. This communication uses sign language in the form of hand and body gestures.

Biometric technology is highly developed, for example, a technology that study and analyze the physical and behavior of human for authentication, then the application of technology in facilitating the use of sign language can be done.

Many studies that discuss the hand motion and hand gesture detection. As the study conducted by Chris Putu Wijayanto [1] in 2009 on his research entitled "Building Applications Training Sign Language Computer Based On People Deaf" and research conducted by Ubaidullah Umar [6] in 2011 on his research entitled "Tracking Directions Finger Gestures Webcam based using Optical Flow method", which in these two studies may be associated with our research. The study conducted by Ubaidullah Umar [6] is the development of video surveillance applications using a finger as a medium that utilizes an Optical Flow method for detecting and tracking the movement of a finger through the image processing that has been stored by the server computer. The next step is to calculate the distance of the screen left and right to determine the finger position as a mouse.

This research is the development of the finger detection by utilizing the movement of Indonesian Sign Language (BISINDO) as the main object of the research and Optical Flow method for detecting and tracking the movement of the hand through image processing using databases

A. Research Purpose

Based on the background described, the aim of this research are to make software applications as a sign language interpreter, to detect hand movements using PC camera (webcam) and to produce a communication tool for the deaf and mute.

II. Theory Support

A. Indonesia Sign Language

According to the Dictionary of Indonesian Language, Sign Language is a language that is often used by the deaf. Sign language is a language that does not use the sound of human speech or writing [4]. Deaf people use this language by combining shapes, orientation and movement of the hands and body, as well as expressions to express their thoughts.

Meanwhile, according to common dictionary of sign language Indonesia, Sign Language Indonesia is a signal that was originally taken from the signals used deaf as words or vocabulary in Indonesian, including American Sign Language (ASL) [4][5]. And from these two sources then according to research conducted by Chris Putu Wijayanto, in the manufacture of sign language training applications using BISINDO would be more useful and easier for Indonesian Sign Language is used more often when compared with SIBI [1]. Here's an example image taken cues from General Dictionary Indonesian Sign Language in Fig 1.

(a) (b)

` Fig. 1. Indonesian Sign Language of "Alhamdulillah" (a) capture from camera 1 (b) capture from camera 2

B. *Object Tracking*

Definition Object Tracking in the image is a way to follow the image elements that have been localized into moving objects over time automatically in moving image / video. Tracking has an important role in the processing of the video as it can be a benchmark to extract some estimates the property of the moving object in the video [27][28].

C. *Optical Flow*

In research conducted by Ubaidullah Umar and colleagues describe the concept of Optical Flow, Optical Flow which is forecast movement of a part of an image based on the derivative of the light intensity in a sequence of images. In 2D space, it means how much an image pixel moves between two frame images sequentially [6][7]. While in 3D space this means how far a pixel volume (voxel) moves on two successive volumes. Calculation of the derivative is based on changes in light intensity in the second frame image and volume. Changes in light intensity of a portion of the image can be caused by movements.

1.1. Algorima Lucas-Kanade

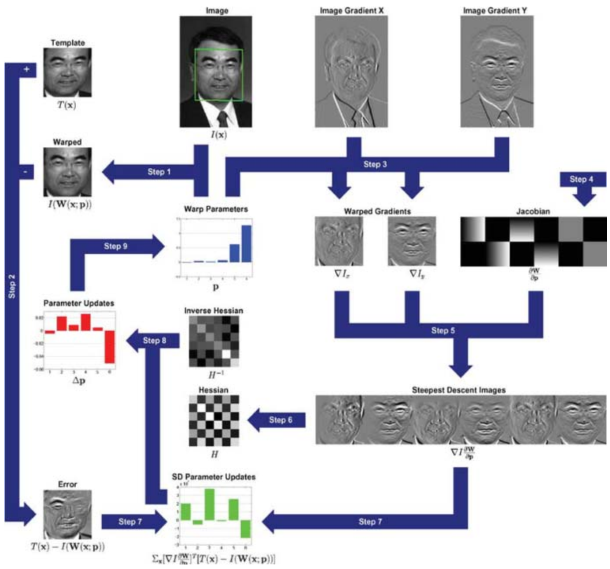

Fig. 2. *Optical Flow* Lucas-Kanade Algorithm (Lucas and Kanade, 1981)

A schematic overview of the Lucas-Kanade algorithm (Lucas and Kanade, 1981). The image I is warped with the current estimate of the warp in Step 1 and the result subtracted from the template in Step 2 to yield the error image. The gradient of I is warped in Step 3, the Jacobian is computed in Step 4, and the two combined in Step 5 to give the steepest descent images. In Step 6 the Hessian is computed from the steepest descent images. In Step 7 the steepest descent parameter updates are computed by dot producting the error image with the steepest descent images. In Step 8 the Hessian is inverted and multiplied by the steepest descent parameter updates to get the final parameter updates p which are then added to the parameters p in Step 9.

Optical Flow algorithms Lucas-Kanade in Fig. 2 can be described as follows :

1. Locate the template with expectations at the first image by using W(x,p) to calculate $I(W(x,p))$.
2. Calculate the value of the error $T(x) - I(W(x,p))$.
3. Determine the value of the gradient VI with *W(x.p).*
4. Calculate the value of Jacobian $\frac{\partial y}{\partial x}$ at *(x,p).*
5. Calculate the *steepest descent* $\nabla I \frac{\partial y}{\partial x}$.
6. Calculate the Hessian matrix.
7. Calculate the equation $\sum_x \left[\nabla I \frac{\partial W}{\partial p}\right]\left[T(x) - I(W(x;p))\right]$
8. Calculate the value Δp.
9. Update the parameter p.
10. Repeat until the algorithm $|\Delta p| \leq \varepsilon$.

D. *In Pyramidal Lucas - Kanade*

The Lucas–Kanade is a widely used differential method for optical flow estimation developed by Bruce D. Lucas and Takeo Kanade. It assumes that the flow is essentially constant in a local neighbourhood of the pixel under consideration, and solves the basic optical flow equations for all the pixels in that neighbourhood, by the least squares criterion. By combining information from several nearby pixels, the Lucas–Kanade method can often resolve the inherent ambiguity of the optical flow equation. It is also less sensitive to image noise than point-wise methods.

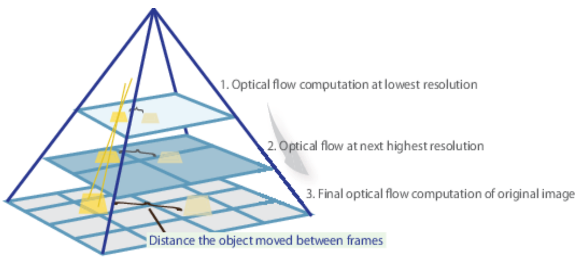

Fig. 3. *Pyramidal Optical* Algorithm (http://www.mathworks.com/help/vision/ref/vision.pointtracker-class.html?requestedDomain=www.mathworks.com)

E. *Euclidean Distance*

Euclidean distance or Euclidean metric is the "ordinary" distance between two points in Euclidean space. With this distance, Euclidean space becomes a metric space. The Euclidean distance between points x and y is the length of theline segment connecting them. In Cartesian coordinates, if x = (x1, x2,..., xn) and y = (y1, y2,..., yn) are two points in Euclidean nspace, then the distance (d) from x to y is given by the Pythagorean formula :

$$d(\overline{A}, \overline{B}) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

Where :

d = *distance*
x_1 = x coordinate 1
x_2 = x coordinate 2
y_1 = y coordinate 1
y_2 = y coordinate 2
z_1 = z coordinate 1
z_2 = z coordinate 2

III. ESIGNING AND MANUFACTURING SYSTEMS

A. *Design Systems*

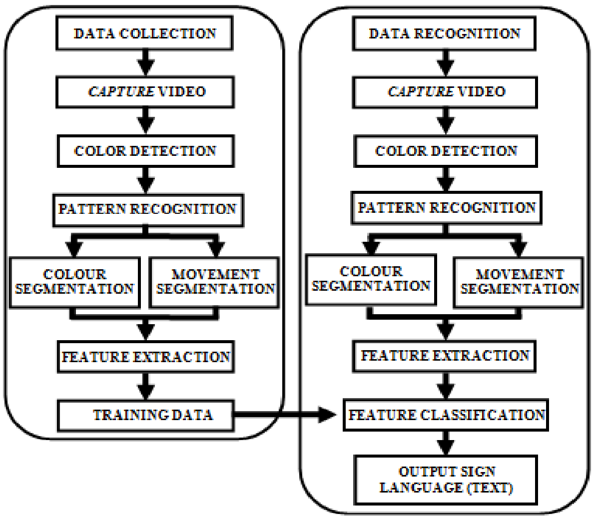

Fig. 4. Design Systems of Application

The steps to measure and to recognize the data are as follows:

1. Video capturing which is done in real time is displayed on GUI application which has been made.
2. After that, retrieve 2 data segmentations, which is color and movement segmentation on each captured image to differentiate the color and movement data using Optical Flow. Color and movement segmentation can be obtained through the object detection and pattern recognition methods as a pre-processing processes. The object detection method is obtained through the moment and the color histogram of the image, while the pattern recognition process is done by detecting the point on the hand with the unit of time and direction of movement of Optical Flow in sequence.
3. The next step is to perform feature extraction for data separation. In this process, color and movement features will be separated in sequence on different frames by calculating the position of the x, y and z at each point earned.
4. The processed data is stored in the certain folder. This data was taken as the data for training and as a comparison value in the process of recognition of sign language gestures.
5. The next step is to do the Optical Flow detection with database matching to determine the suitability of the direction of hands and body movement with words that have been saved.
6. After going through the training process and the classification by using optical flow, then the identification process and the results of language recognition hand signals are converted into text can be obtained.

B. *Object Recognition*

The process of object recognition using a web camera that has been set manually will detect the red gloves that are used by the user. The process of object recognition is through detection of the position of the object using the value of the initial position of the object and the final value of the object's position after moving. Fig. 5 illustrates the process of moving the position of an object.

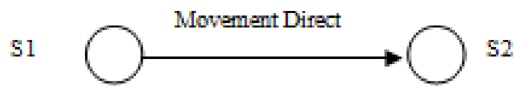

Fig 5. Illustration of Object Movement

Fig. 5 shows that the S1 is the initial position of the object and S2 is the object position after movement

IV. EXPERIMENTS AND DISCUSSIONS

The output of the data collection process is the position x, y and z as the main feature object. Data is collected in front of the camera using hand gestures as a sign language.

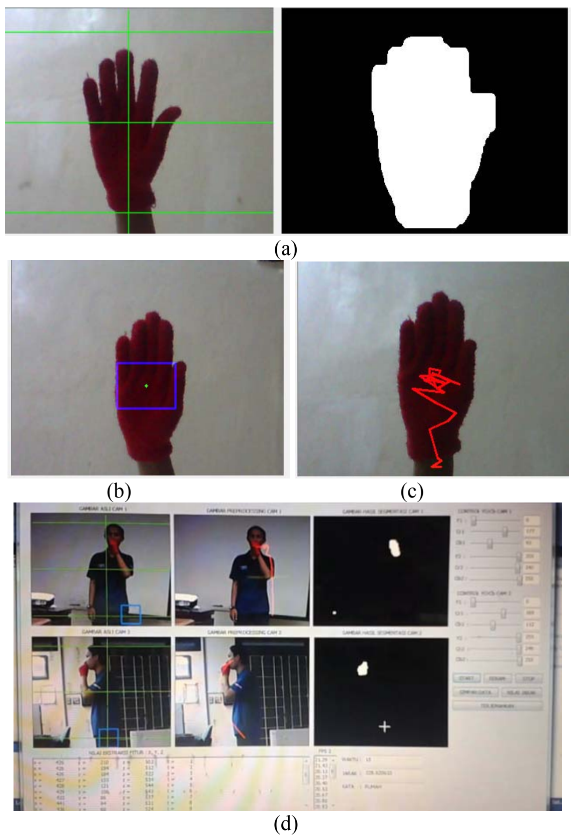

Fig. 6. Data Collection Process (a) color segmentation (b) central moment process (c) tracking image process (d) application running

Detection of the position x, y and z are used as input the classification process using Euclidean Distance.

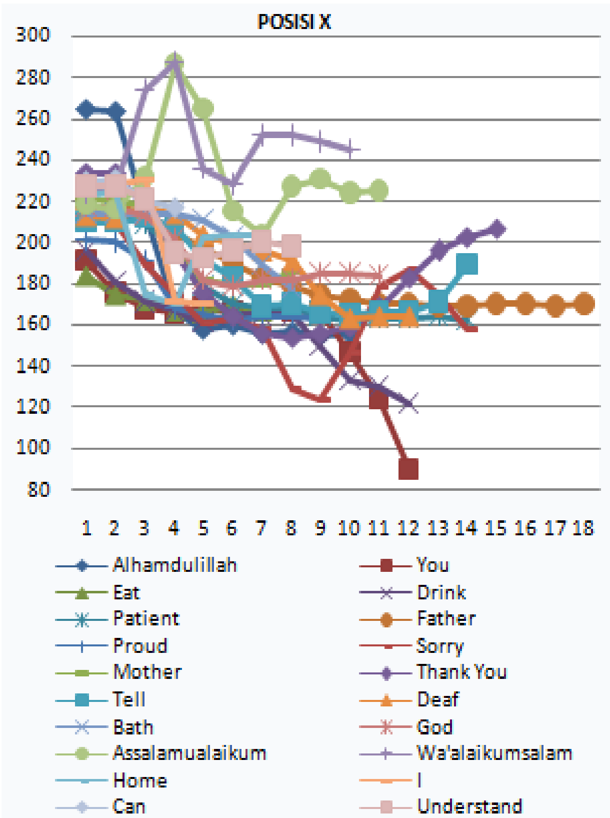

Fig. 7. Chart of X Position on each Word

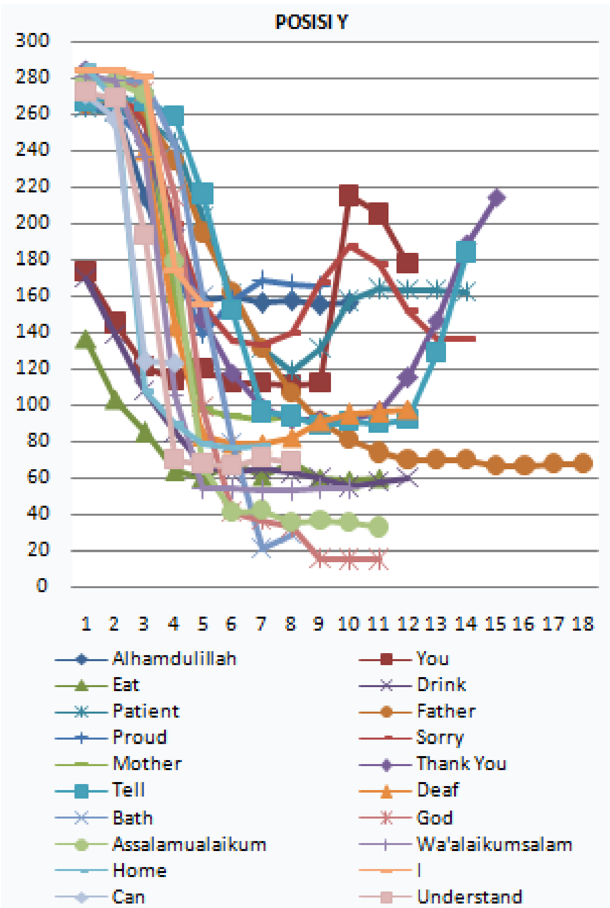

Fig. 8. Chart of Y Position on each Word

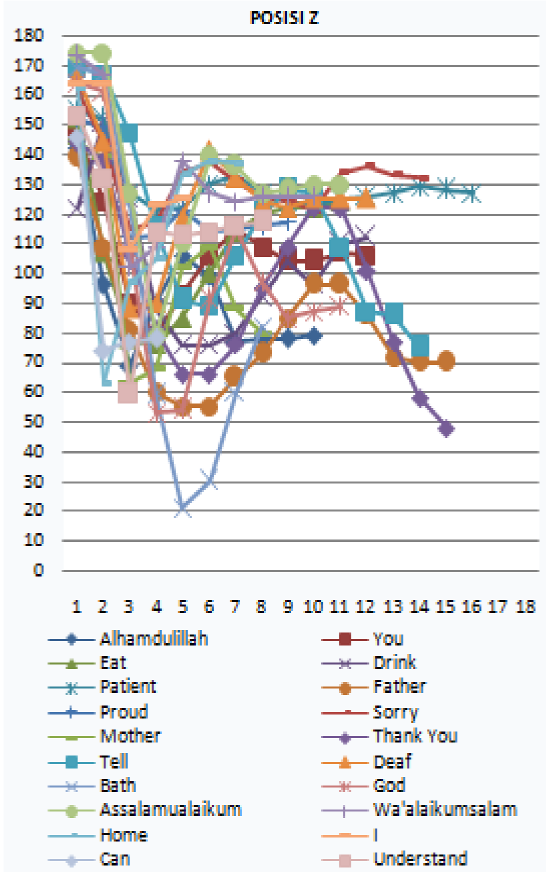

Fig. 9. Chart of Z Position on each Word

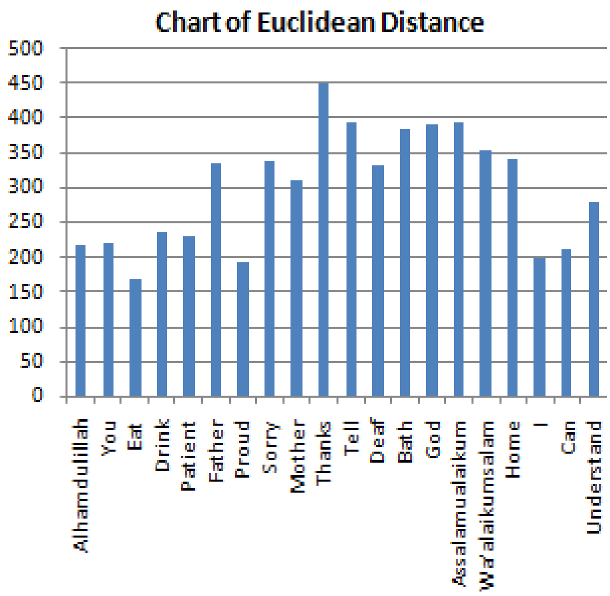

Fig. 10 Chart of Distance Value on each Word

In Fig. 10 we can see that every word has a different word patterns. The pattern obtained from the calculation of Euclidean Distance value of X and Y at each point. Pattern similarity between the words is not visible, so that each word has the potential similarity detection of different translations. It is influenced by different time taken in movement between words and movements differ between one word with another word. So from the data is allowed to do the reading sign language Indonesia.

TABLE I. WORD DETECTION

WORD DETECTION	NUMBER OF DETECTED	NUMBER OF UNDETECTED	PROCEN -TAGE
Alhamdulillah	6	4	60%
You	7	3	70%
Eat	5	5	50%
Drink	8	2	80%
Patient	7	3	70%
Father	8	2	80%
Proud	6	4	60%
Sorry	5	5	50%
Mother	7	3	70%
Thanks	7	3	70%
Tell	7	3	70%
Deaf	6	4	60%
Bath	7	3	70%
God	5	5	50%
Assalamualaikum	7	3	70%
Wa'alaikumsalam	6	4	60%
Home	6	4	60%
I	9	1	90%
Can	6	4	60%
Understand	6	4	60%

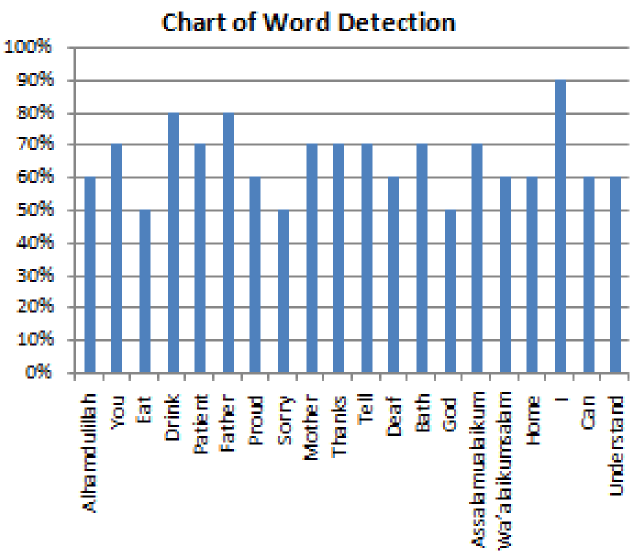

Fig. 11. Chart of Word Detection

V. CONCLUSIONS AND FUTURE WORK

Based on the experimental results, we can conclude that the process of translating Indonesian sign language into text can use the optical flow method, through the detection of x and y position as euclidean distance calculation feature. In the future, we would like to develop this application to be easy to use and add more words of sign language.

REFERENCES

[1] Chris Putut Wijayanto, —Membangun Aplikasi Pelatihan Bahasa Isyarat Berbasis Komputer Pada Orang Tuna Rungu‖, Sekolah Tinggi Manajemen Informatika Dan Komputer, Amikom, Yogyakarta

[2] Trong-Nguyen Nguyen, Huu-Hung Huynh, —Static Hand Gesture Recognition Using Artificial Neural Network‖, *Journal Of Image And Graphics, Volume 1, No.1, March, 2013*

[3] Ammar Anuar, Khairul Muzzammil Saipullah, Nurul Atiqah Ismail, Yewguan Soo, —Opencv Based Real-Time Video Processing Using Android Smartphone‖, *International Journal Of Computer Technology And Electronics Engineering (Ijctee) Volume 1, Issue 3*

[4] Pusat Bahasa. (1988). —Kamus Besar Bahasa Indonesia‖. Jakarta : Balai Pustaka.

[5] Imas A. R. Gunawan 1996 —Kamus Umum Bahasa Isyarat Indonesia‖, Jakarta : Lembaga Komunikasi Total Indonesia

[6] Ubaidillah Umar, Reni Soelistijorini, Haryadi Amran Darwito, —Tracking Arah Gerakan Telunjuk Jari Berbasis Webcam Menggunakan Metode *Optical Flow*‖, Jurusan Teknik Telekomunkasi - Politeknik Elektronika Negeri Surabaya, Surabaya

[7] Affan Mahtarami, Moch. Hariadi, —Tracking Gerak Tangan Berbasis Pyramidal Lucas-Kanade‖, Jurusan Teknik Elektro, Fakultas Teknologi Industri, Its – Surabaya

[8] Dimas Arioputra, —Analisa Perbandingan Akurasi Metode *Optical Flow* Dan *Gaussian Mixture Model* Untuk Sistem Pemantau Lalu Lintas Berbasis Computer Vision‖, Fakultas Teknik, Program Studi Teknik Komputer, Universitas Indonesia.

[9] Benedictus Yoga Budi Putranto, Widi Hapsari, Katon Wijana, —Segmentasi Warna Citra Dengan Deteksi Warna Hsv Untuk Mendeteksi Objek‖, Fakultas Teknik Program Studi Teknik Informatika, Universitas Kristen Duta Wacana Yogyakarta

[10] Kurniawan Dwi Irianto, Gunawan Ariyanto, Dedi Ary P, —Motion Detection Using Opencv With Background Subtraction And Frame Differencing Technique‖, Jurusan Teknik Informatika, Fakultas Komunikasi Dan Informatika, Universitas Muhammadiyah Surakarta.

[11] Gary Bradsky & Adrian Kaehler. *Learning Opencv : "Computer Vision With The Opencv Library"* O'reilly.

[12] Berthold K. P. Horn Brian G. Schunck, Determining *Optical Flow*, April 1980, Massachussets Institute Of Technologi Artificial Intelligence Laboratory