Detecting and Tracking of Traditional Instrument with Different Camera Vision

Diyan Purwati
Departement of Informatics and Computer
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
diyanp@ce.student.pens.ac.id

Akhmad Subhan Khalilullah
Departement of Creative Multimedia
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
subhankh@pens.ac.id

Dadet Pramadihanto
Departement of Informatics and Computer
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
dadet@pens.ac.id

***Abstract* — In this paper, we present an approach for detecting, recognizing and tracking varying number of objects through both have similarity shape and colors. The objects tracking is the varying notes of Indonesian traditional instrument called "gambang". Our method build an idea how to recognize well the notes and classify each objects in real time tracking. The houghlines standard method is used to search the right lines of the shape of instrument. At the level of detecting object, the square and contours method is used. The real time tracking is based on the color of contour object using hsv color filter. We demonstrate the viability of this approach by experimenting on several videos of the instrument from different angle of view 2D camera. The instrument labeling testing with the camera view from the left side has 100% success, from the center side has 95% success, and from the right side has 90% success.**

Keywords — vision; object tracking; muti square detection; houghlines transform; hsv color filter; instrument tracking;

I. INTRODUCTION

The development of humanoid robots today is not only used to help human work, but the intelligence implanted in robots is now used to develop humanoid robots for entertainment. The development of entertainment robots attracts a lot of attention and interest researcher to create the ability of robots which resembles human ability. The ability of humanoid nao robots to play metallophone instruments [1] not only in addition to entertainment, such robots could also be employed for teaching musical instruments to children or in the treatment of autistic people [2]. From this background, the EEPIS Robotic Research Center develop the ability of humanoid robot to play Indonesian traditional music instrument named "gambang" [9] or xylophone.

This research first is build the visual process aims to recognize the notes of the instrument with moving camera and flexible position of camera. The notes recognition is processed by canny edge detection, houghlines detection, contour and moment processing. The HSV color tracking method is used to keep the notes recognize well while moving camera track the notes song. The parameter value to identify the notes is based on the pixel position and sorting method.

The camera used in this research is 2D RGB Camera with not fixed position and the size of "gambang" or xylophone musical instrument used has a length of 131 cm, width of the longest wood arrangement 58 cm, width of the smallest wooden structure 31 cm, and 29 cm of the iinstrument height. The gambang musical instrument consists of 20 pentatonic tones with different beam widths 7 cm to 3.5 cm.

Fig.1. Indonesian Traditional Instrument named Gambang

Tracking of multiple objects from 'real world' is one of the most complicated problems in computer vision since the motion of these objects is unpredictable and cannot be assumed. Tracking is also "application depended" task. There is no one general tracking methodology for solving all tracking problems [8]. Therefore, different algorithms and procedures are needed and used for different applications. However, it is possible to rank tracking applications in two general categories: the ones that need a moving camera and the ones where a stable camera is enough. The object can move rapidly to any direction, objects can also move isolated, together with other objects or been hidden for some time before they reappear in the tracking scene. In this paper, we used color filter to help tracking some of the instruments detected in the frame.

II. RELATED WORK

The previous related research is that Nao Robot Humanoid has succeeded to doing visual tracking and auditory feedback from the instrument sound. But the different is that the phisical size and characteristic of the instrument between metallophone and "gambang" or xylophone is totally different. The camera used by nao robot humanoid is 3D camera and this research used 2D camera. Based from that case, the method used in nao robot humanoid is model based object pose tracking [1] can't be implemented in this research. So the algorithm in this research is build based on the characteristic of the camera and instrument of "gambang" with 2D RGB Camera using hsv color filter to track with flexibel position of camera view to the instrument. The challenges in this research is how to keep recognize the notes while tracking with correct labeling.

The same approach which is done in the present and previous research that is at the preprocessing stage. The research on nao robots requires an edge detection process to adapt the model of the pose hypothesis that is built with the actual instrument condition. The errors between the pose model and edge detection results are calculated so that the pose model is able to adjust the position of the instrument according to edge detection. In this research, the edges detection is used to detect the lines and shapesof the instrument. There is no error calculation between the detected shape and the actual shape if the instrument.

III. METHODOLOGY

In this section, we present the procedures to recognizing and tracking the instruments of the "gambang" in Fig.2

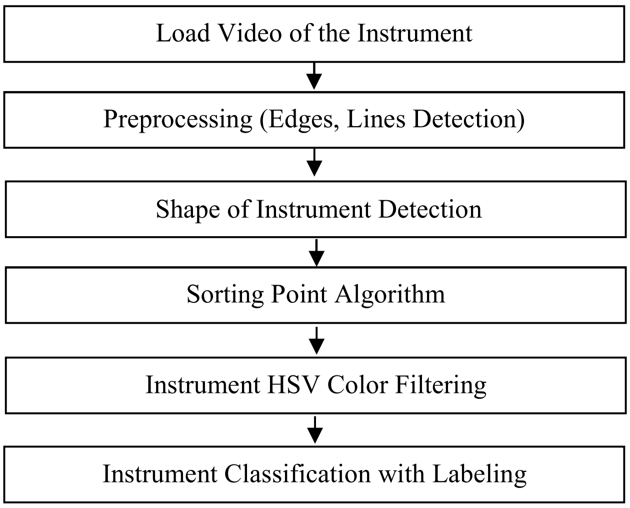

Fig.2. Diagram of the Algorithm

A. Various of Camera View of The Instrument

The field of view camera has an important role in the process of recognizing the instrument. There is no limit to the position and point of view of the camera requires that the method must be able to adjust the detection of instruments from various sides with various instrument sizes to be seen on the frame. In fig.3 to 5 are some of the camera positions for testing methods used in this research.

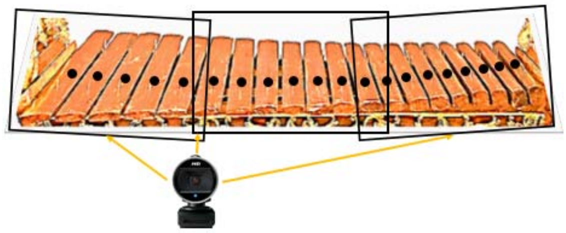

Fig.3. Camera View from the Left Side

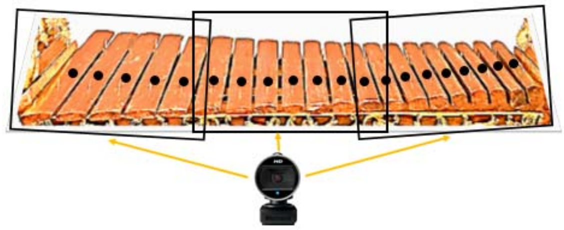

Fig.4. Camera View from the Center Side

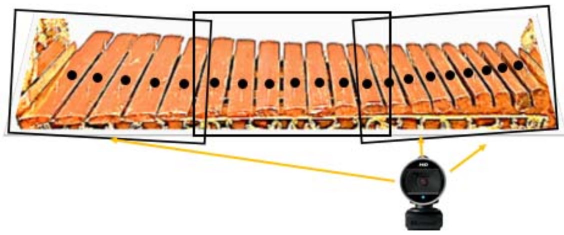

Fig.5. Camera View from the Right Side

B. Preprocessing Frame

Preprocessing step aims to recognize the instrument correctly without noise and ignore other object which potentially becomes a noise. Some of the preprocessing steps are:

1) Preprocessing

a) Canny Edges Detection

The result of canny edge detector will produce the edges of the instrument frame. The parameter on canny edge detection are threshold, so the intensity edge value that is close to black or white can be more manageable. The high threshold parameters produce a high degree of gray intensity and only maintain a high degree of grayness and removes the gray value under that parameter. Based on this process it is known that high contrast values on image produce a high degree of grayness so that it is detected as an edge. If the condition occurs then the contrast in the instrument area will be detected into edges and become a noise to the process of the shape detection. The desired edge detection result is precisely on the edges of the shape of the instrument square. Figure 6 is the result of edge detection with canny edge detection but there is noise edge detection within the instrument area.

Fig.6. Canny Edge Detection Result

b) Scaling, Brightness and Blur the Frame

The implementation of scaling is to resize the frame become smaller, it affects the extent of pixels processed in the image so that it will reduce the image processing. The implementation on brightness reduction because the instrument area reflects the light thus affecting the edges detection process. Scaling, brightness and blur process can reduce the noise around the instrument and produce edges on the instrument better.

Fig.7. Egde Detection with Scaling, Brightness and Blur Result

C. *Shape of Instrument Detection*

To detect the instrument from the preprocessing result require some process method, aims to get the correct shape of the instrument. Some of the method process are:

1) Instrument Detection

a) Standard Hough Lines Transform

Required input to do Hough Transform is the result of an edge detection. Hough Transform works by mapping the dots on the image into space parameters (hough space). Based on a function that defines the shape to be detected, So the houghlines transform contains function mappings from the points (edges) which is defined by the equation of the line so that a collection of points that resemble the line will be detected as a line. The equation of polar coordinate system is :

$$y = -\frac{cos\theta}{\sin\theta}x + \frac{r}{\sin\theta}$$

Where r = x cos θ + y sin θ, in general any point (x0, y0) can be defined by rθ = x0. Cos θ + y0. Sin θ so that each pair (rθ, θ) represents every line be passed (x0, y0).

The result of houghlines transform method is the line detection on instrument based on edge detection result. The output of this method is two points that represent the detected line from an edge detection (pt1.x, pt1.y, pt2.x, pt2.y). With this method then the point then becomes a reference to draw a line like figure 8 and 9.

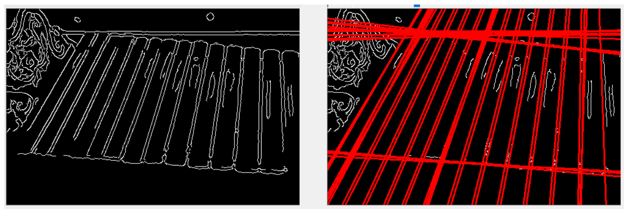

Fig.8. Hough Lines Transform Probablity Left Side

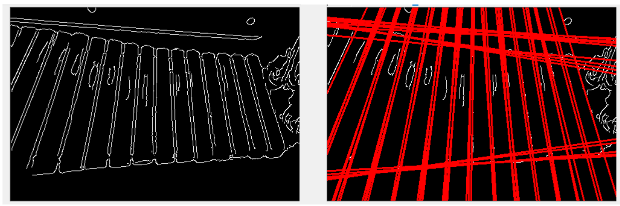

Fig.9. Hough Lines Transform Probablity Right Side

b) Find Square Shape

To recognize the instrument shape based from the lines detection, we use find squares method. The algorithm of find squares is :

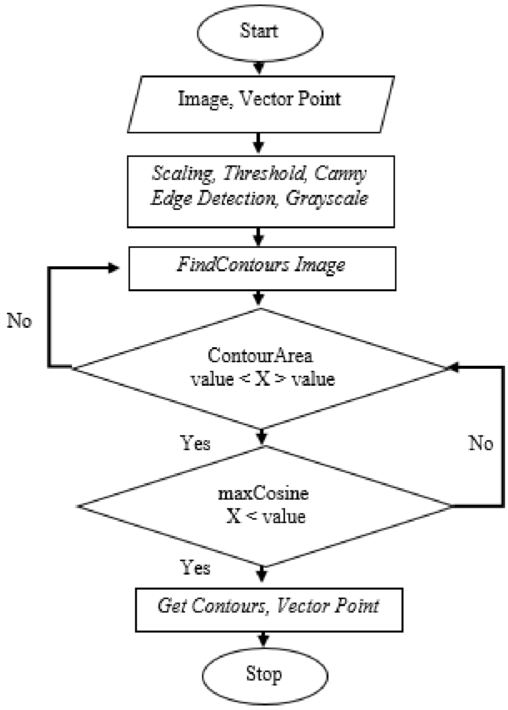

Fig.10. Find Squares Flow Chart

The algorithm of the find squares method is operate 4 vector points on the result image of the line detection by the process including threshold, canny, findcontours and contour area conditions as well as detected angle. The output of the findsquares process is a 4 point vector point which can be seen by drawing polylines or drawcontours. The intruments has different angle which need to detect, so we use a function to find a degree of angle of 4 point points (pt1.x, pt1.y, pt2.x, pt2.y) and look for the length or distance from the point forming a vector line. Then look for the angle formed from the two lines.

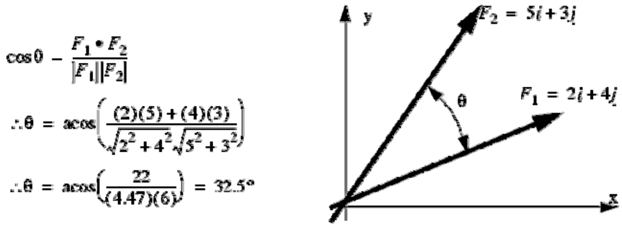

Fig.11. The Equation of Vector Angle

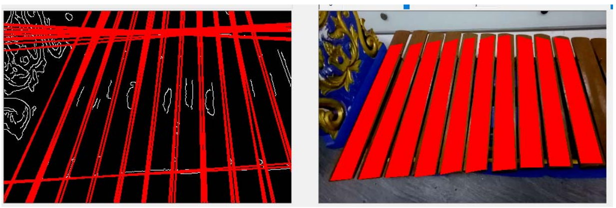

Fig.12. Find Squares Result Left Side

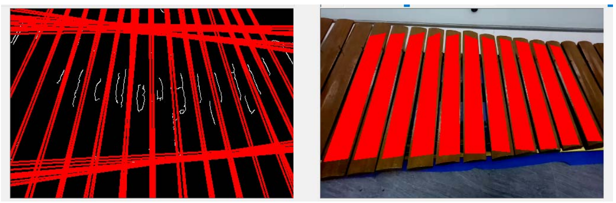

Fig.13. Find Squares Result Center Side

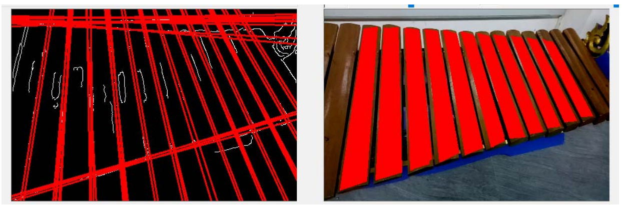

Fig.14. Find Squares Result Right Side

We used contour moment to get the centroid of the contour instrument. The equation of contour moment on gray value function:

$$m_{p,q} = \sum_{i=1}^{n} I(x,y)x^p y^q$$

$$\mu_{p,q} = \sum_{i=0}^{n} I(x,y)(x - x_{avg})^p (y - y_{avg})^q$$

D. Sorting Point Algorithm

The results from the search for threshold contours can not be directly computed for labeling and storing the coordinates of objects or instruments. In fact, the contour moment detects randomly all the objects in the frame, so it needs an algorithm to sort the contours in the frame from the moment position. So it is necessary to obtain center points every moment especially for the current case required sorting according to the position of x on the pixel camera. The required moment parameter value is the midpoint of each instrument contour sorted from the position of small x is moment 0 and so on it increases the value of the center point of the contour of the instrument.

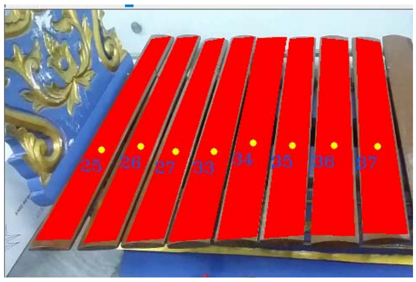

Fig.15. Sorting of Contour Moment Based Pixels Position

E. Instrument HSV Color Filtering

The color moments aid in representing the object and also used in image retrieval applications [4]. In this work, the red coloring of the instrument is performed to detect the object with a hsv range filter of red color. The color values used as filters are RGB of red RGB (255,0,0) converted to HSV value (0,130,255). This color giving is given manually in the process of contour depiction, so the color can be changed according to the best color conditions to be done HSV color filter.

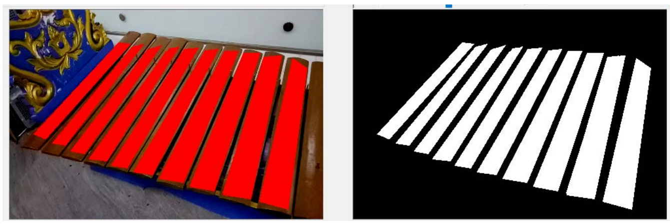

Fig.16. Coloring and HSV Color Filtering Left Side

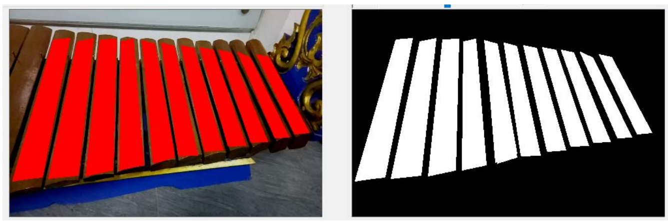

Fig.17. Coloring and HSV Color Filtering Right Side

Tracking of multiple objects means associating each of the segmented objects in the previous frame with the corresponding object in the next consecutive frame of a video sequence. The association is achieved by correlating the features of moving objects in the successive frames. In this work, Chi-Square dissimilarity measure [5] is used to measure the degree of association of the objects features in the successive frames.

F. Classification with Labeling Algorithm

To make labeling adjustments when the camera is in motion, a counting algorithm for incoming and outgoing instances is required. An easy way to implement is to check the existence of line intersection. Due to the characteristics of the instrument to be detected when the square conditions are formed intact and located at the edge area of the frame, a line between frames has a possibility that the object will be formed and lost. From the specified line, we need to know the characteristics of the lines formed from each instrument by studying how the contours are formed. So from the drawcontour function can also be draw rectangle that maintains a straight line from the outer point of the contour of the instrument.

The result of making of rotated rectangle of each instrument is got the characteristic of point formed and got angle value from slope of instrument. Figure 18 is the result of the representation of the rotated rectangle point points with the color characteristic at this point point corresponding to the angle formed from the instrument square.

Fig.18. Drawing Rotated Rectangle

From the characteristic obtained the instrument angle has a small value from the left position, the center instrument of the camera position will be worth 90 and the right to the greater the value of the square angle formed. When the angle condition is <90 then the blue square point point is at the lowest of pixels (x, y). For an angle condition> = 90, the square points of red and white are parallel to form a straight line of the bottom of the square. Then the result of the characteristics of the instrument points formed in the computation to be detected whether there is intersection with the created line. By creating a boolean condition program then when intersection means that there is an instrument out of the frame. The boolean condition program is created from some form condition of the instrument outgoing in the left side frame.

The algorithm created is able to adjust the labeling of the instrument with the movement of the camera from left to right. Figure 19 shows a yellow line created as an instrument condition exits the frame and activates the boolean program. The success obtained up to this classification process requires the calibration of parameter values to detect edges (canny edge detection) and detect the line (houghlines transform). so the precise calibration value determines the success of the labeling method.

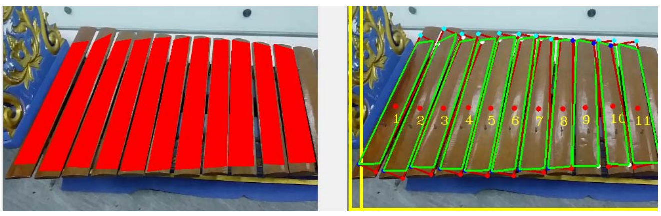

Fig.19. Labeling Left View from Left Position Camera

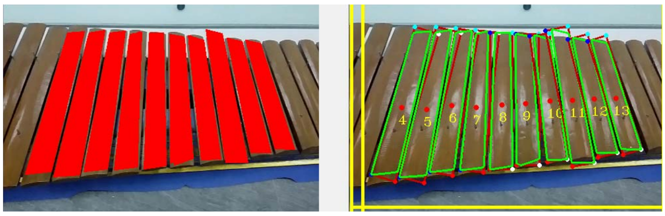

Fig.20. Labeling Center View from Left Position Camera

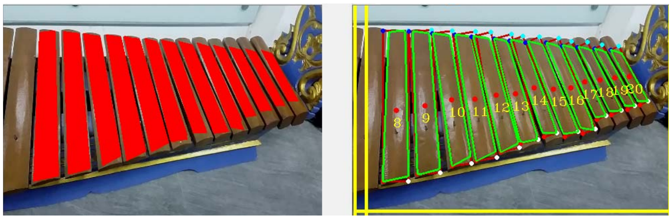

Fig.21. Labeling Right View from Left Position Camera

IV. EXPERIMENT RESULT

The experimental result obtained from testing methods against different camera positions. Table 1 is testing methods against different camera positions.

a) Camera view from the right, left, center side

Table 1. Tracking and Labeling Testing

Instrument	Left Side	Center Side	Right Side
1	Success	Success	Success
2	Success	Success	Success
3	Success	Success	Success
4	Success	Failed	Success
5	Success	Success	Success
6	Success	Success	Failed
7	Success	Success	Failed
8	Success	Success	Success
9	Success	Success	Success
10	Success	Success	Success
11	Success	Success	Success
12	Success	Success	Success
13	Success	Success	Success
14	Success	Success	Success
15	Success	Success	Success
16	Success	Success	Success
17	Success	Success	Success
18	Success	Success	Success
19	Success	Success	Success
20	Success	Success	Success
Error	0%	5%	10%

From the test results and observations are known that the characteristics of the gambang instruments have different characteristics on each side and point of view of the camera. The diversity of parameters to be considered in image processing in this instrument strongly influenced by light and the number of carvings on the right and left side adds edge detection which makes the line detection process disturbed and produces noise failure detection of the shape. Another parameter is the success of the line that determines the success of the form also influenced by the area between the line as the instrument area. Beside that, the failure rate on the identification failed to occur due to the poor detection process of the instrument resulting in an unstable object for tracking and identification. The best data retrieval position is from left-sided camera point of view.

V. CONCLUSION

The proposed of this paper is built an automatic system to detect, recognize, tracking, identifiy the instrument. The system for pre-processing as a process to detect the shape of the instrument successfully performed through the process of brightness, scaling, blur, edge detection, line detection, square detection, contour, moment. HSV filter color method used works well and can form according to the results of the detection of the form as a multi-tracking method. The success of the tracking labeling algorithm for instrument classification has not been fully successful. There is still a fault of up to 10% with experiments and observations to obtain appropriate calibration values and is appropriate for each side of the camera's point of view. The success of the classification and labeling process is influenced by the results of line detection. Stable line detection results will be easier to classify.

The vision process must be capable of producing maximum, precise, fast results with little error or noise. It is worth studying the characteristics of the processed object and searching for as many parameters as can be calculated into an algorithm for vision computing as needed. It takes a lot of experiments and observations to know the characteristics of the image processing results on each side rather than the camera's point of view. So the next plan for this research is how to make an algorithm to do a calibration adjustment automatically based on the instrument characteristic to get the high success rate of labeling. Another plan is make an algorithm to adjust the labeling method from rigth to left camera moving. The last plan is, implementation in humanoid robot to playing traditional instrument from Indonesia.

VI. REFERENCES

[1] D. Maier, R. Zohouri, M. Bennewitz, Using Visual and Auditory Feedback for Instrument-Playing Humanoids, In Proc. of the IEEE RAS International Conference on Humanoid Robots. 2014.

[2] I. Fujimoto, T. Matsumoto, P. R. S. De Silva, M. Kobayashi, and M. Higashi. Study on an assistive robot for improving imitation skill of children with autism. In Proc. of the Int. Conf. on Social Robotics (ICSR), 2010.

[3] Nidhal El Abbadi, Lamis Al Saadi. Automatic Detection and Recognize Different Shapes in an Image. ICJSI International Journal of Computer Science Issues, Volume 1D Issues 6 No 1, 2013.

[4] Chandrajit M, Girisha R, Vasudev T, Motion Segmentation from Surveillance Video Sequence using Chi-Square Statistics, Proceedings of ELSEVIER International Conference on Emerging Research in Computing, Information, Communication and Applications. (ERCICA-14).

[5] O. Pele and M. Werman, The Quadratic-Chi Histogram Distance Family, in K. Daniilidis, P. Maragos, N. Paragios, Eds.: Berlin, Heidelberg: ECCV 2010, Part II, LNCS 6312, 2010, pp. 749–762.

[6] W. Xiangyu, and L. Xiujuan, The study of moving target tracking based on Kalman-CAMShift in the video, 2nd International Conference on Information Science and Engineering (ICISE), pp.1-4, 2010.

[7] L. A. Sechidis, P. Patias, V. Tsioukas, Low Level Tracking of Multiple Object, Commission III, WG 5, 2008.

[8] Panji Prasetyo, Seni Gamelan Jawa Sebagai Representasi dari Tradisi Kehidupan Manusia Jawa : Suatu Telaah dari Pemikiran Collingwood, Fakultas Ilmu Pengetahuan Budaya. Universitas Indonesia, 2012.

[9] Jong Hun Park, Gang Seong Lee, Jin Soo Kim, Seuc Ho Ryu, Sang Hun Lee. (2016). *Independent Object Tracking from Video using the Contour Information in HSV Color Space.* Indian Journal of Science and Technology. Vol. 9 S(1).

[10] Ganesh S. Patel, Jaykumar S. Dhage. (2017). *Real-Time People Counting Method with Head Localization and Tracking System Consideration.* International Journal of Innovative Research in Computer and Communication Engineering. Vol. 5 Issue 4.