International Journal of Engineering & Technology, 7(4.16) (2018) 124-127

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Development of Hand Tracking and Hand Recognition for Virtual Manufacturing

Gandjar Kiswanto*, Dody Rakhmat Ramadhan, Adjeng Ayu Setiani, Ferdiansyah Zhultriza, Muhammad Fathin Juzar, Rachmad Muhammad Suryantoro

Department of Mechanical Engineering, Universitas Indonesia, Depok 16424, Indonesia
**Corresponding author: gandjar_kiswanto@eng.ui.ac.id*

Abstract

Development of information and technology very rapidly encourage people to create a simulation system. This simulation system can be used to simulate the manufacturing field. To make this simulation become more naturally it is used hand tracking and hand recognition as input device. This paper discusses the implementation background subtraction, CAMSHIFT and convex hull algorithms for Hand Tracking and Hand Recognition. Input given in the form of palm coordinates then the system will do the tracking by creating boundaries rectangle. Then separated between the hand image and the background on the boundaries. After that done calculate image processing to determine the pose of the hand. The result of this research is a system that will perform tracking and recognition and then the result next stage to be simulated

Keywords: *Simulation, Background Subtraction, CAMSHIFT, convex hull, virtual manufacturing*

1. Introduction

By communicating people can present their ideas. In the world of mechanical engineering visual communication is used to convey ideas such as when conveying ideas about the design of an engine block. One of the communication media can use simulation. Simulations are used when systems in the real world cannot be used because of a limitation or when the system has not been built but we want to get predicted results.

In the world of manufacturing simulation in the planning process becomes a very important aspect, with the simulation can be known the advantages and disadvantages of system performance that will be made so that it can be improved on the system. Simulation can help identify and display problems that cause the system not working properly. Simulation can be used to analyze small changes that could actually have a big impact. Simulation can provide alternative solutions in case of problems in the system. Simulation also can be used to observe critical points and can also be used to provide recommendation to the management, and to reduce unnecessary cost

A good simulation is a simulation that can describe the actual event. To make the simulation more naturally then use the movement and gesture of the hand as input directly not through other intermediaries. So, it is necessary to develop a reliable method of hand tracking and hand recognition.

The outline of the method used to perform hand recognition and hand tracking is to do a background subtraction to separate the background by hand, then do the tracking on the palm using CAMSHIFT algorithm developed by Gary Bradsky in his paper "Computer Vision Face Tracking for Use in a Perceptual User Interface" [1], then perform recognition using convex hull algorithm developed by Amiraj Dhawan in his paper" Implementation of Hand Detection based Techniques for Human Computer Interaction ". [2]

2. Image Processing

2.1. Background Subtraction

Background subtraction is used to detect moving objects. this algorithm is very useful for static background or static camera. the basic idea to do a background subtraction is to use the background image, the background is used as a reference. background image must be an image without moving objects, it would be better if using a background that has a homogeneous colour. there are various background subtraction algorithms but used for this research is the frame algorithm "frame different theoretical techniques". this algorithm is the easiest algorithm to do. The current frame is simply the subtracted from the previous frame, and if the difference in pixel values for a given pixel is greater than a threshold, then the pixel is considered part of the foreground [3].

$$|(\text{frame}_1\text{-frame}_{i-1})| > T_h$$

Advantages of this method are it is a lowest computation speed of all the methods, background model is highly adaptive. Which means, this method adapts changes in the background faster than any other methods, since background is just the previous frame (precisely 1/fps). [4]

2.2. Thresholding

Thresholding can be used to create binary images of grayscale images by changing the value of all pixels below the threshold to zero and changing all values above the threshold into one.

To get a good global thresholding is to minimize misclassification on the pixel. to minimize the error must make the histogram to be bimodal. One method that can be used is Otsu method. This algorithm assumes there are two tops of background and foreground.

Copyright © 2018Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The purpose of the Otsu method is to divide the grayscale image histogram into two different regions automatically without requiring the user's help to enter the threshold value. [5]

2.3. CAMSHIFT

If described briefly the CHAMSHIFT algorithm has the following explanation [6]
1. Selecting the desired starting part that contains the object you want to track, this part of the square coordinate of the object
2. The next step creates a colour histogram of the selected object
3. Perform a back-projection histogram on the frame
4. Find the centre of mass in search widows using the MeanSHIFT method based on the probability distribution image
5. Placing search windows to the point that has been obtained using step four and do the literacy until you find the most central point
6. Process the next frame with the search window position from step five

2.4. Contour

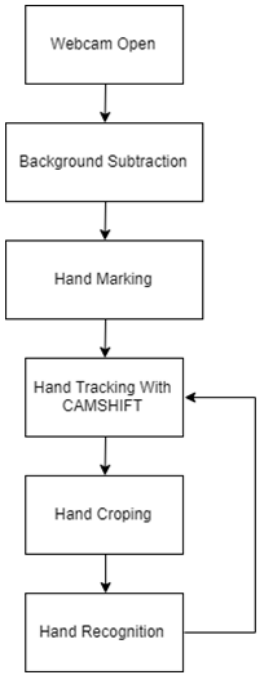

Fig. 1: Architecture of hand tracking and hand recognition

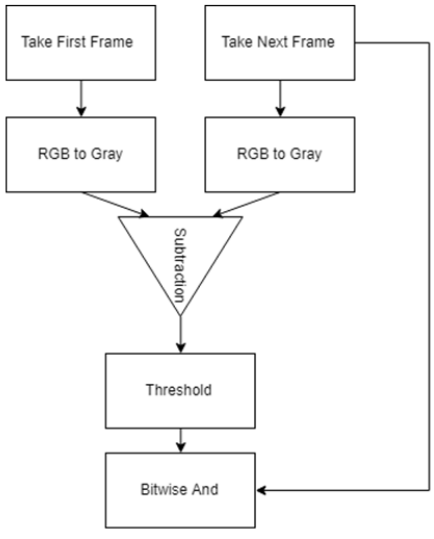

Fig. 2: Algorithm background subtraction

Binary images that have been obtained by using thresholding is just a view that has no meaning. To be able to retrieve the information on the image it is necessary to have advanced processing. One of the processing is to use contour. Chain code can be used to represent contour by means of connected circuit from straight line segment with certain direction and length. The representation used is based on eight connectivity that resembles the direction of the wind with a targeted direction numbering [7].

3. System Modeling

Hand recognition and hand tracing systems that have been made have several stages, each stage will be explained, in order to facilitate the understanding can see the figure 1. In Figure 1 shows the flow diagram of hand tracking algorithm and hand recognition that has been made. explanation of each step in the flowchart image 1 as follows

3.1. Webcam Open

Webcam is used to get video. Because the input used in the form of video then the camera needed to turn on.

3.2. Background Subtraction

For more detail can see figure 2. In figure 2 describes the stages of the background subtraction. first need to acquisition background, then change the image RGB into Grayscale image. the next step to do subtract with the next frame that has been changed to grayscale. Thresholding is then done to separate background and foreground. background and foreground can be separated by changing the colour, the background is black while the foreground is white. then the last thing calculates with logical bitwise and. the results obtained can be seen in Figure 3.

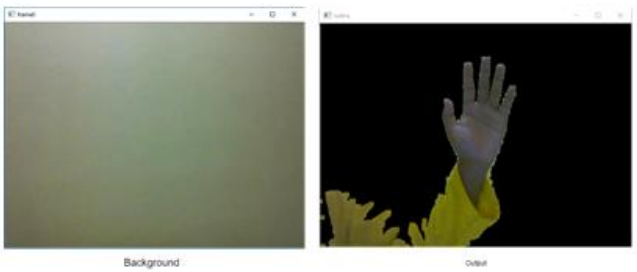

Fig. 3: Output from background subtraction

3.3. Hand Marking

To make a rectangular boundary, and save coordinates every corner from the boundary. the colour in the boundary will be used as a reference.

3.4. Hand Tracking with CAMSHIFT

At this stage hand tracking is done by using CAMSHIFT algorithm. The CAMSHIFT algorithm consists of 3 stages. The stages consist of back projection, MeanSHIFT, and tracking. for Back Projection and MeanSHIFT will be explained using the flow diagram, as in Figures 4 and 5. In figure 4 start with changing the hand image became HSV image. Then take the Hue value of the HSV image. The value of Hue has a range between 0-180. Then turn it into a grayscale image. From this grayscale image is made into a histogram so it can be changed to the image back projection. Back projection is to calculate the weight of each colour that is on the histogram and then change each pixel in the grayscale image according to the weigh colour of the pixel. Suppose yellow colour has percentage of 30% then if found yellow pixels will be done

multiplication 0.3 x 255 so get result 76.5. The second stage is doing MeanSHIFT algorithm. In this algorithm the first thing to do is determine the size of the diameter of the circle. Then do the calculations to determine the centre of its mass. After that move the centre of the circle to the centre of mass. The last stage is to do tracking. tracking is done by performing back projection and MeanSHIFT algorithms on each frame.

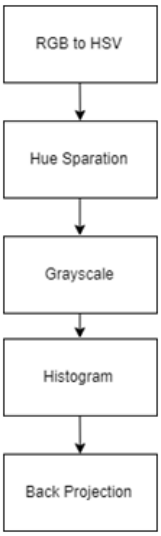

Fig. 4: Algorithm back projection

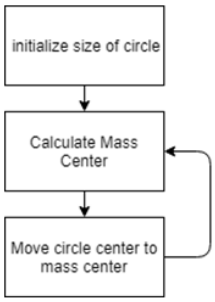

Fig. 5: Algorithm MeanSHIFT

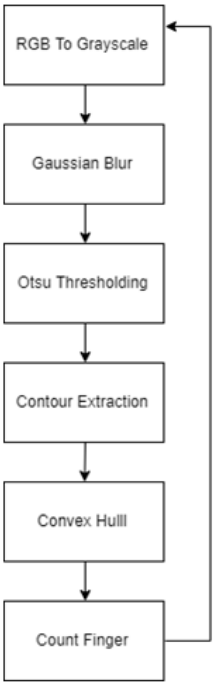

Fig. 6: Algorithm Hand Recognition

3.5. Hand Cropping

Cropping is done by removing outside pixels from boundary. the purpose of cropping is to eliminate noise that is outside the boundary. Because the noise will affect the results of recognition. Noise is usually in the form of faces or objects that have the same color as skin.

3.6. Hand Recognition

For recognition given the only limit to identify six kinds of palm poses. Holding pose, pose one, pose two, pose three, pose four, pose five. The algorithm used as in figure 6.

1. RGB to grayscale: The RGB image is then co verted to gray-scale image
2. Gaussian Blur: Gaussian blur is used to eliminate angular angles that are too sharp,
3. Otsu Thresholding: Otsu Thresholding researchers use to separate between the hand object with the background that is in the rear.
4. Contours extraction: making contours on the largest threshold image
5. Convex Hull: This stage is the core of this recognition algorithm, in Figure 7 there is a line that has a black colour indicates convex hull, red dots at the end of the black line is called the defect, while the green line is the contours. To determine the hand poses to form poses 2 to pose 5 simply use the following algorithm. When the hand image has been obtained and has convex hull, each convex hull value is formed from the convection in the hand image but not necessarily every convexity is the convention of the finger. See the segments formed by the ABC line, to ensure that the convexity comes from the fingers, it is necessary to calculate the angular widths formed by sides b and c, if the angle has a value below 900 then it is considered a convexity derived from the fingers. So if we get an angle whose value is below Ninety then the palm pose shows the number two and so on. However, this algorithm does not apply to determine hand positions and number one. To determine the hand poses the researcher makes a new algorithm which to determine the hand pose shows the pose one or grasp quite simply by comparing the lines D and E, if the line D has a long over the line E then the palm pose shows the number one and if the line D is not exceeding the E line then the hand pose shows a holding pose, for more details can see figure 8.



Fig. 7: Palm hand

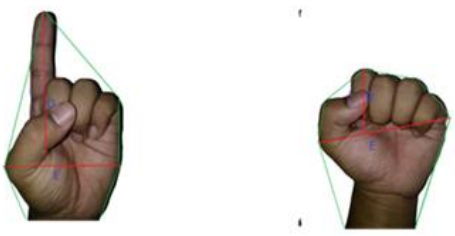

Fig. 8: One Pose and Holding Pose

4. Result

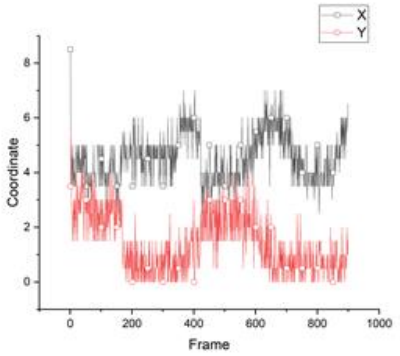

Fig. 9: Deviation (x,y) on Translation Motion

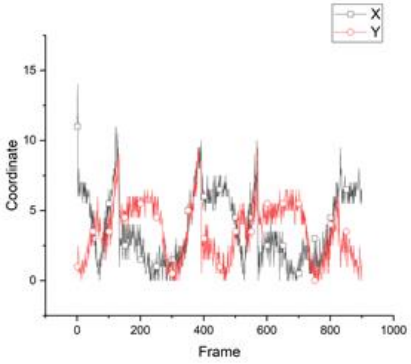

Fig. 10: Deviation (x,y) on Rotation Motion

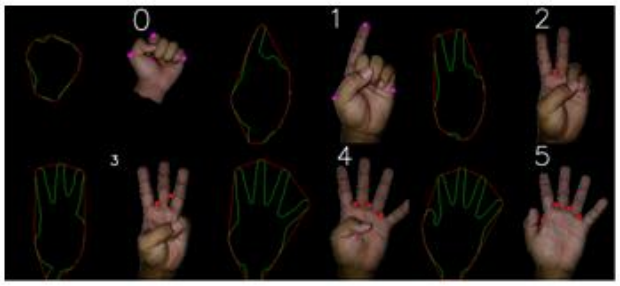

Fig. 11: Successful Recognition

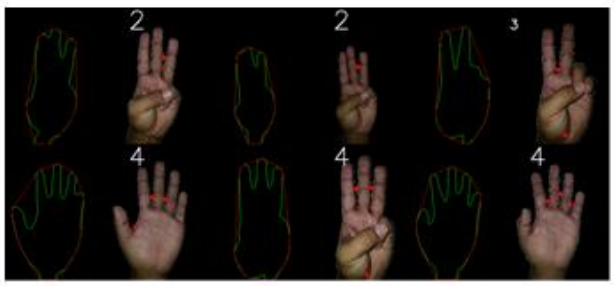

Fig. 12: Fail Recognition

Conducted measurements to see the performance of tracking. measurements are made on hand images that move translations and rotations. the result of the measurement of translational motion is in figure 9 and the rotational motion in figure 10. This measurement is done by comparing the coordinates of the tracking result with the actual coordinates. The biggest deviation occurring in translational motion is 8 pixels. The biggest deviation in rotation motion is 7 pixels. For the measurement of recognition there will be 120 hand images with different poses, for each pose there are 20 images. from the measurement results there are 20 incorrect image recognitions. sample results of recognition can be seen in Figure 11 and 12.

5. Conclusion

The CAMSHIFT algorithm has great performance for tracking object translation and object rotation. the resulting deviation is small enough that it will not be visible or affect the virtual manufacturing. the disadvantage of the CAMSHIFT algorithm is that it will interfere with a background or object that has the same color as the hand image. In addition, this algorithm will fail to track if the background used has the same colour as the hand image. To eliminate errors that occur because the background color should be done background subtraction. Background subtraction will remove the entire background leaving only the hand image.
Recognition will work if there is no noise image around the hand. other than that the hand should be open clearly, the hand is not open clearly will cause the wrong recognition. The direction of coming light is also very important because it causes a shadow that can cause recognition errors. The direction of the light comes good when it comes from the top of the hand, the shadow will be below and not disturbing.

Acknowledgement

The authors are thankful to the Department of Mechanical Engineering, Engineering Faculty, Universitas Indonesia for supporting this research

References

[1] G. Bradski, “Computer Vision Face Tracking For Use in a Perceptual User Interface,” Intel Technology Journal, 1998.
[2] V. H. Amiraj Dhawan, “mplementation of Hand Detection based Techniques for,” International Journal of Computer Applications, vol. 72, 2013.
[3] R. C. R. E. W. a. S. L. E. Gonzalez, “Digital image processing using,” Gatesmark Publishing, vol. 2, 2009.
[4] S. S. Deepjoy Das, “Implementation And Performance Evaluation Of Background Subtraction Algorithms,” International Journal on Computational Sciences & Applications , vol. 2, 2014.
[5] P. Hidayatullah, Pengolahan Citra Digital, Bandung: INFORMATIKA, 2017.
[6] “OpenCV,” OpenCV, 23 February 2018. [Online]. Available: https://docs.opencv.org/3.4.1/d7/d4d/tutorial_py_thresholding.html. [Accessed 14 July 2018].
[7] M. J. Y. a. H. S. Ren, “Tracing boundary contours in a binary image,” Image and vision computing, 2002.