

DOI: 10.17725/rensit.2020.12.407

Image Processing Techniques Based Crowd Size Estimation

Mohanad Abdulhamid

AL-Hikma University, <http://alhikma.edu.iq>.

P.O. Box 10069, Baghdad, Iraq

E-mail: mob1hamid@yahoo.com

Lwanga Wanjira

University of Nairobi, <http://uonbi.ac.ke>

P.O. Box 30197-00100, Nairobi, Kenya

E-mail: researcher12018@yahoo.com

Received February 26, 2020; peer reviewed March 30, 2020; accepted April 06, 2020

Abstract: Image processing algorithms are the basis for image computer analysis and machine Vision. Employing a theoretical foundation, image algebra, and powerful development tools, Visual C++, Visual Fortran, Visual Basic, and Visual Java, high-level and efficient computer vision techniques have been developed. This paper analyzes different image processing algorithms by classifying them in logical groups. In addition, specific methods are presented illustrating the application of such techniques to the real world images. In most cases more than one method is used. This allows a basis for comparison of different methods as advantageous features as well as negative characteristics of each technique is delineated. The main objective of this paper is to use image processing techniques to estimate the size of a crowd from a still photograph. The simulation results show that the different images have different efficiencies.

Keywords: Image processing; crowd size estimation

UDC 004.932.2

For citation: Mohanad Abdulhamid, Lwanga Wanjira. Image Processing Techniques Based Crowd Size Estimation. *RENSIT*, 2020, 12(3):407-414. DOI: 10.17725/rensit.2020.12.407.

CONTENTS

1. INTRODUCTION (407)
 2. DESIGN STAGES (408)
 3. IMPLEMENTATION AND RESULTS (410)
 4. CONCLUSION (413)
- REFERENCES (413)

1. INTRODUCTION

A crowd is something beyond a simple sum of individuals. It has collective characteristics which could be described in general terms such as "angry crowd", and "peaceful crowd". A crowd can assume different and complex behaviors as those expected by their individuals.

Understanding crowd behavior helps in designing pedestrian facilities, for major layout modifications to existing areas and for the daily management of sites subject to crowd traffic. Conventional manual measurement techniques are not suitable for comprehensive data collection of patterns of site occupation and movement. Real time monitoring is tedious and tiring but safety-critical.

When congestion (crowd density) exceeds a certain level, this being dependent upon the collective objective of the crowd and the environment, danger may occur for a variety of reasons. Physical pressure may result directly in injury to individual or to the collapse of parts of the physical environment.

Crowd density analysis could be used to measure the comfort level in public spaces or detect potentially dangerous situations. There are real models developed to estimate the number of people in crowded scenarios using computer vision techniques such as; pixel-based analysis, texture-based analysis, and object-level analysis.

An important and challenging problem related to the crowd phenomenon is crowd simulation which relates to the reproduction of realistic crowds based on computer graphics techniques. Animations of crowds find applications in, for instance, evaluation of crowd management techniques where for instance simulation of the flow of people leaving a football stadium after a match.

Study of methods used by human observers may help in the choice of image processing algorithms likely to be useful in automatic assessment of crowd behavior. Some works related to the topic of this paper can be found in literatures[1-5].

2. DESIGN STAGES

The development of the crowd size estimation program algorithm is divided into four main stages according to our computation strategy block diagram shown in **Fig. 1**. Then each main stage is broken into sub-steps according to the algorithm use. Every sub-step is planned as a distinct algorithm, which could be written and tested separately, then incorporated into the main stage.

2.1. IMAGE ACQUISITION

At this stage we read the image from a digital camera or phone camera. The angle of capture could be aerial, at an angle or directly frontal view.

2.2. IMAGE ENHANCEMENT

We carry out image processing techniques at this stage. They include the following:

- Segmentation
- Morphological Operations
- Edge Detection

2.2.1. SEGMENTATION

This technique segments an image into various components for object recognition. The read image is first converted to grayscale. The objects pixels are separated within an image from background pixels. Thresholding graylevel techniques separate an object from the background based upon the graylevel histogram of an image. We use the graylevel discontinuities within the image. These discontinuities are then used to separate objects within an image from the background.

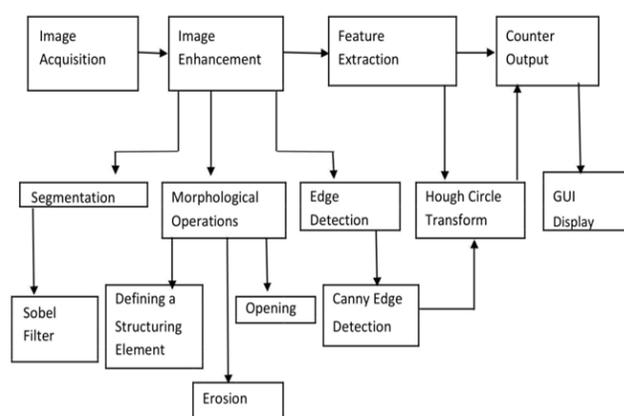


Fig. 1. Computation strategy block daigram.

We use the gradient magnitude operation as the segmentation function. A popular one is the sobel operator which creates an image which emphasizes edge and transitions. Thus, we use the sobel edge masks, imfilter and some simple arithmetic to compute the gradient magnitude. The gradient is high at the borders of the objects and low mostly inside the objects.

2.2.2. MORPHOLOGICAL OPERATIONS

We perform the following morphological operations on the segmented image.

- Define a circular structuring element 'disk'
- Erosion
- Opening

2.2.2.1. DEFINING STRUCTURE ELEMENT

A structuring element is a second set of pixels with peculiar shape that acts on the pixels of the image to produce an expected result. Thus we choose a structuring element the same size and shape as the objects we want to process in the input image, in our case we are using a circular structuring element to enable us identify the circular shape of the heads.

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. The morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

Morphological techniques probe an image with a small shape or template called a structuring element. The structuring element is positioned at all possible locations in the image and it is compared with the corresponding neighborhood of pixels. Some operations test whether the element "fits" within the neighborhood, while others test whether it "hits" or intersects the neighborhood.

It is this that determines the precise details of the effect of the sobel operator on the image. The structuring element is sometimes called the kernel, but we reserve that term for the similar objects used in convolutions. The structuring element consists of a pattern specified as the coordinates of a number

of discrete points relative to some origin.

2.2.2.2. EROSION

If a binary image is eroded, the resultant image is one where there is a foreground pixel for every origin pixel where its surrounding structuring element sized, fit within the object. We combine two sets using vector subtraction of set elements and is the dual operator of dilation. It's used to shrink the size of the heads for easier identification.

The basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (i.e. white pixels, typically). Thus areas of foreground pixels shrink in size, and holes within those areas become larger. This therefore emphasizes the circular form of the heads for easy identification and analyzing.

2.2.2.3. OPENING

Application of an erosion immediately followed by dilation of the eroded object using same structural element. It closes or fills the gaps between objects used for smoothing outline of objects after a digitization followed. It eliminates thin protrusion.

We can use morphological opening to remove small objects from an image while preserving the shape and size of larger objects in the image. Thus we choose the operations to improve the quality of the eroded image.

2.2.3. EDGE DETECTION

Edge is the boundary between an object and its background. If edges of images can be identified with precision, all the objects can be identified and their area, perimeter, shape etcetera can be calculated. We decided to use the Canny Edge detection because it detects strong edges plus it will find weak edges that are associated with strong edges. As edge detection is a fundamental step in computer vision, it is necessary to point out the true edges to get the best results from the matching process. That is why it is important to choose edge detectors.

The smoothing concept has been applied in this Gaussian operation, so the finding of errors is effective by using the probability. The next advantage is improving the signal with respect to the noise ratio and this is established by Nonmaxima suppression method as it results in one pixel wide ridges as the output. The third advantage is better detection of edges especially in noise state with the help of thresholding method. The major disadvantage is the

time consumption because of complex computation of Gradient calculation for generating the angle of suppression.

2.3. FEATURE EXTRACTION

Features are inherent properties of data, independent of coordinate frames. In this context we are looking to extract heads for counting in estimation of a crowd. Therefore, we incorporate the use of Hough Transform to identify and mark the heads in form of circles. Hough Transform is a mapping algorithm that processes data from a Cartesian coordinate space into a polar parameter space.

It is most useful for finding geometric lines and shapes in binary images. The Hough transform is a technique which can be used to isolate features of a particular shape within an image. Because it requires that the desired features be specified in some parametric form, the classical Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses, etc. A generalized Hough transform can be employed in applications where a simple analytic description of a feature is not possible. Due to the computational complexity of the generalized Hough algorithm, we restrict the main focus of this discussion to the classical Hough transform. Despite its domain restrictions, the classical Hough transform retains many applications, as most manufactured parts (and many anatomical parts investigated in medical imagery) contain feature boundaries which can be described by regular curves. The main advantage of the Hough transform technique is that it is tolerant of gaps in feature boundary descriptions and is relatively unaffected by image noise.

2.4. COUNTER OUTPUT

We display the total circle detected to be through a graphic user interface. **Fig. 2** shows the flowchart of the design of the graphic user interface.

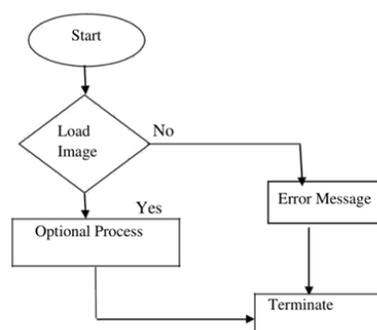


Fig. 2. Flowchart.

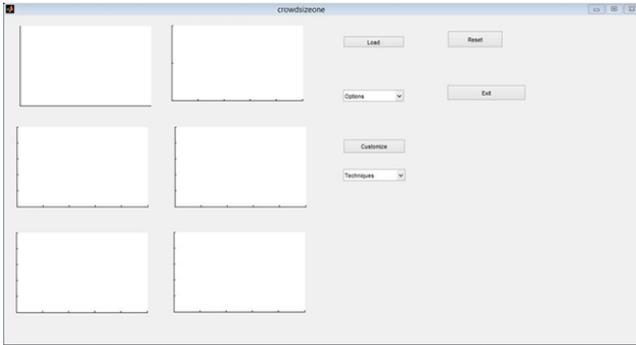


Fig. 3. Graphic user interface-1.

Graphic user interface-1 in **Fig. 3** shows the blank user interface before loading the image. Also graphical user interface in operation-1 in **Fig. 4** shows the user interface after loading the image and the total estimated crowd size displayed in a separate window.

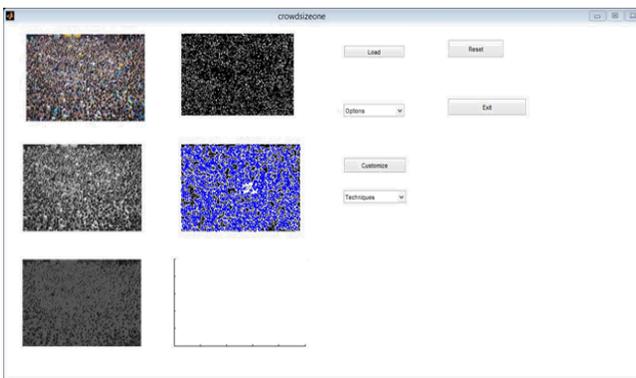


Fig. 4. Graphic user interface in operation-1.

Fig. 5 shows the design stage of the Graphic user interface with the different axes for displaying results of various image processing techniques on the image. Moreover, we incorporate the use of pushbuttons and pop-up menus to provide a wide variety of user options.

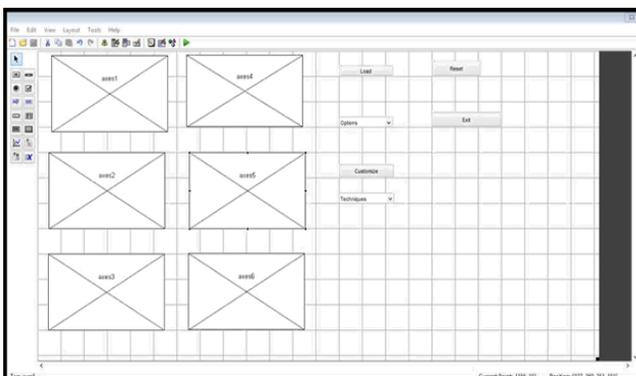


Fig. 5. Design stage of graphic user interface.



Fig. 6. Original image.

3. IMPLEMENTATION AND RESULTS

Having Designed the program as well as the graphic user interface, we implement it on 14 crowd pictures and tabulate the results of the different behaviors to various image processing techniques as well as the results of accuracy. First, taking the below choice photo (**Fig. 6**) for illustration of our design implementation.

The image in **Fig. 6** is what we use to put through the various image processing techniques designed in the design stage. It is a picture of a crowd at a concert taken from an angle aerial point of view. As depicted in the picture, the "heads" in the foreground are bigger than those at the background part of the image. We estimate a radius range in the Hough transform to accommodate the variant sizes. However still some "heads" are not counted and the program counts some hats as heads. Nevertheless we are able to achieve a good level of accuracy as indicated later.

Below is the systematic hierarchy of output pictures (**Figs. 7-10**) as the program is running. We indicate the various processes below each image for easier understanding.

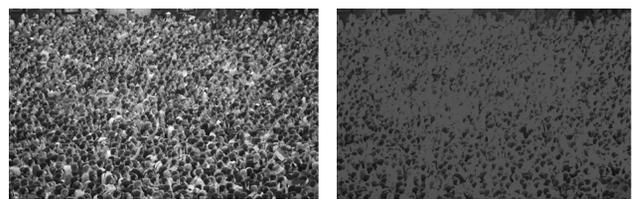


Fig. 7. Grayscale image. Fig. 8. Opening-reconstructed image.

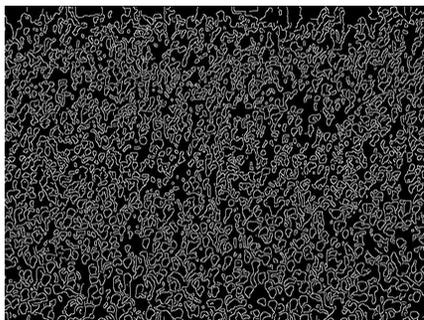


Fig. 9. Edge detected image.

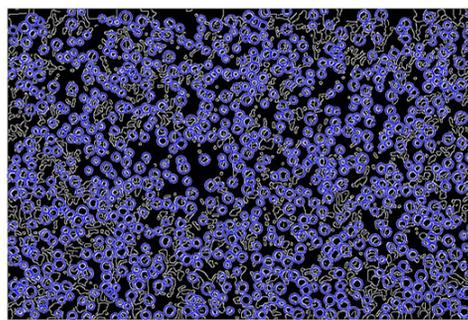


Fig. 10. Hough circle transformed image.

Fig. 11 indicates the total number of heads counted by the program is 776. Manually, we count the number of people in the picture and came up with an estimate of 1,100 people. Hence the percentage accuracy is; $(776/1100) * 100 = 70.545454\%$

However not all the images could achieve this level of accuracy due to their characteristics. For instance some are taken from a side-view point of view and their arms and legs are counted too thus the total achieved exceeding total number of people in the picture. Moreover, we eliminate some image processing techniques while adding others to certain images to establish which combination of processes works for each individual image.

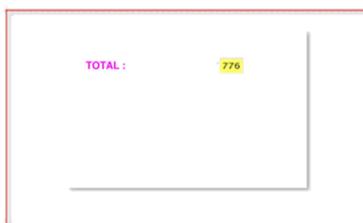


Fig. 11. Total of circles.

Figs. 12-25 show the 14 different images of crowds we work with.



Fig. 12. Crowd 1.



Fig. 13. Crowd 2.



Fig. 14. Crowd 3.



Fig. 15. Crowd 4.



Fig. 16. Crowd 5.



Fig. 17. Crowd 6.



Fig. 18. Crowd 7.



Fig. 19. Crowd 8.



Fig. 20. Crowd 9.



Fig. 21. Crowd 10.



Fig. 22. Crowd 11.



Fig. 23. Crowd 12.



Fig. 24. Crowd 13.



Fig. 25. Crowd 14.

Table 1 shows the behavior of the above images to different combinations of image processing techniques. Number represents the images, and (x) marks processes undertaken. Table 1 shows the program total, manual total, and the efficiency for each image.

Table 1

| Behavior of images | | | | | | | | | | | | | | | |
|--------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|----------------|
| Techniques | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | Original Image |
| Morphology | x | x | x | | x | | x | x | | x | x | x | x | x | x |
| Segmentatiob | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| Erosion | x | x | x | x | x | x | | x | x | x | x | x | x | | x |
| Dilatation | | | | | | | | | | | | x | | x | |
| Opening | x | x | x | | x | x | x | x | | | | | x | | x |
| Closing | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| Edge Detection | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| Manual total | 1200 | 57 | 84 | 168 | 20 | 170 | 96 | 1500 | 80 | 100 | 40 | 110 | 65 | 50 | 1100 |
| Program total | 782 | 43 | 62 | 83 | 17 | 85 | 42 | 632 | 15 | 52 | 12 | 48 | 28 | 23 | 776 |
| Accuracy (%) | 65.2 | 75.4 | 73.8 | 49.4 | 85.0 | 50.0 | 43.8 | 42.1 | 18.8 | 52.0 | 30.0 | 43.6 | 43.0 | 46.0 | 70.5 |

4. CONCLUSION

In this paper, there was an understanding of what a crowd is and how to incorporate various image processing techniques using algorithms to estimate a crowd size on a photograph. Exploration and familiarization with the different image processing techniques used to analyze the images was carried out and tested. A software was designed to be used for analysis and extraction of features of a photograph that most likely represent a person, which in our case was heads. The software was designed in such a fashion to take the image and determine what the system needs to do to get the total number of people and display sequential effects on an image through a

graphic user interface as well as total. Though tedious, total number of people were counted manually so as to be compared and contrasted with the software output.

REFERENCES

1. M. Jiang, J. Huang, X. Wang, J. Tang, and C. Wu. An approach for crowd density and crowd size estimation. *Journal of Software*, 2014, 9(3):757-762.
2. Lwanga Wanjira. *Crowd size estimator using image processing techniques*. Graduation Project, University of Nairobi, Kenya, 2014.
3. N. Kulkarni, A. Rana, and A. Patre. Crowd analysis and density estimation using

- surveillance cameras. *International Journal of Computer Science and Information Technologies*, 2015, 6(6):5044-5047..
4. Rai, R. Meshram. Automatic estimation of crowd size and target detection using Image processing. *Indian Journal of Computer Science and Engineerin*, 2017, 8(3):358-362, 2017.
 5. M. Aziz, F. Naeem, M. Alizai, and K. Khan. Automated solutions for crowd size estimation. *Social Science Computer Review*, 2017, 36(5):610-631.

Моханад Абдулхамид

Ph.D., Assistant Prof.

Университет Альхикма

Р.О. Вох 10069, Багдад, Ирак

moh1hamid@yahoo.com

Лванга Ванджира

B.Sc., Assistant lecturer

Университет Найроби

Р.О. Вох 30197-00100, Найроби, Кения

researcher12018@yahoo.com.