International Journal of Electronics Communication and Computer Engineering
Volume 5, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Design and Implementation of a Human Detector and Counting System using MATLAB

P. O. Otasowie
Electrical/Electronic Engineering Department,
University of Benin, Benin City, Nigeria
Email: potasowie@yahoo.co.uk

I. A. Edeoghon
Computer Engineering Department,
University of Benin, Benin City, Nigeria
Email: isi.edeoghon@uniben.edu

***Abstract* – This research work is aimed at design and implementation of a human detector and counting system using MATLAB. Human detection systems are used to give a measure of the number of people in a given area in a given time. The resolution of this measurement is entirely dependent upon the sophistication of the technology used.**

In this project computer vision technology was employed with the Viola-Jones algorithm using MATLAB Release 2012 (A) platform to design and implement such a system. Efforts were made to analyze its viability, improve on its accuracy as well as outline the challenges faced in its implementation.

The work was deemed satisfactory as it was able to detect and count some numbers of humans in a specific environment as presented by images.

***Keyword* – MATLAB, Viola-Jones, Human Detector, Counting Systems, Computer Vision.**

I. Introduction

A human counter is a device used to measure the number and direction of people traversing a certain passage or entrance per unit time.Many technologies are used in counting persons; infra-red beams, computer vision, thermal imaging and pressure sensitive mats. [1]

Each technology has its merits and demerits although no system has a one hundred per-cent accuracy factor.Of all the afore-mentioned counter technologies the computer vision system is employed in this work.

Computer vision is a field that includes methods for acquiring, processing analyzing and understanding images and, in general high dimensional data from the real world in order to produce numerical or symbolic information e.g. in the form of decisions.

A theme in the development of this field has been to duplicate the abilities of human vision by electronically perceiving and understanding an image.

This image understanding can be seen as disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics and learning theory. [2]

Advantages of Computer vision technologies

- Relatively high accuracy(95%)
- Flexible to customize.
- Highly scalable
- Easy to integrate with other systems

Disadvantages

- High Cost
- High power consumption
- Accuracy can be affected by shadow, floor background, poor image quality, differing light levels.

Uses

- *Retail*
  Necessary to calculate the percentage of stores visitors that make purchase of goods.
- *Optimizing staff shifts*
  Staff requirements are related to the density of visitor traffic and services such as cleaning and maintenance are done typically when traffic is at its lowest.
- *Queue management and customer tracking*
  Enables management of queue in offices, banks, schools.
  Visitor statistics can be monitored to measure marketing effectiveness.
- *Security*
  People counting/detection system can be used to detect an intruder which can set of an alarm
- *Future technologies*
  It is a major step in the future of artificial intelligence developing systems that can accurately tell items apart from each other and properly class and identify them, useful in robotic engineering and cybernetics.

II. Statement of Problem

Human detection/counting systems are very necessary in the modern world. The need to develop intelligent security systems that can monitor footage and 'understand' what they see is highly important. Rather than have human personnel sit all day observing CCTV cameras an intelligent system would be able to detect an object (in this case human) of maybe a specific number and raise an alarm under any of the above conditions according to some timed program. Retail shops might also need to have information on the number of people going in or out of it's stores in order to monitor effectiveness and impact of its market and make surveys and evaluations. Organisations/hobbyists might want to use human traffic information to: (i) limit crowds (crowd control) (ii) monitor building or seating capacity (iii) Management of queues.

III. Review of Related Work

- *Tally counter*

A hand-held tally-counter, sometimes called a clicker-counter, would be used; one press per person. To reset the counter, one would have to turn a knob, resetting most counters' display to "0000".

- *Infrared beams*

The simplest form of counter is a single, horizontal infrared beam across an entrance which is typically linked to a small LCD display unit at the side of the doorway or can also be linked to a PC or send data via wireless links and GPRS. Such a beam counts a 'tick' when the beam is broken, therefore it is normal to divide the 'ticks' by two to get visitor numbers. Dual beam units are also available from some suppliers and can provide low cost directional flow 'in' and 'out' data. Accuracy depends highly on the width of the entrance monitored and the volume of traffic.Horizontal Beam Counters usually require a receiver or a reflector mounted opposite the unit with a typical range up to 6 metres (20 ft), although range finding beam counters which do not require a reflector or receiver usually have a shorter range of around 2.5 metres (8 ft 2 in).Vertical beams are somewhat more accurate than horizontal, with accuracies of over 90% possible if the beams are very carefully placed. Typically they do not give 'in and out' information, although some directional beams.[2]

- *Computer vision*

Computer vision systems typically use either a closed-circuit television camera or IP camera to feed a signal into a computer or embedded device. Some computer vision systems have been embedded directly into standard IP network cameras. This allows for distributed, cost efficient and highly scalable systems where all image processing is done on the camera using the standard built in CPU. This also dramatically reduces bandwidth requirements as only the counting data has to be sent over the Ethernet.Accuracy varies between systems and installations as background information needs to be digitally removed from the scene in order to recognize, track and count people. This means that CCTV based counters can be vulnerable to light level changes and shadows, which can lead to inaccurate counting. Lately, robust and adaptive algorithms have been developed that can compensate for this behavior and excellent counting accuracy can today be obtained for both outdoor and indoor counting using computer vision. [2]

IV. Design Considerations

A number of things were factored into the building of this program.

1. The use of images instead of videos to enable a wide variety of objects to be analyzed in a short time and to test the robustness of the system.
2. The ability of the software to use other features apart from face detection that are as well human defining.
3. Modular, simple structuring and explicit comments on the program so that improvements and other additions can be easily made.
4. Converting the images to grayscale before detection in other to reduce processing time.

V. Design Specifications

1. The appropriate dimensions of test image should be less than or equal to 640 X 480 pixels (about 60kB).
2. Images should not be blurred or underexposed.
3. Images can be black or white or coloured.

VI. Design Procedure

The project was designed using MATLAB 2012 Release A using the Viola-Jones object detector tool = the vision. Cascade Object Detector whose operation is based on the Viola-Jones algorithm[3][4][5]. This tool analyses an image and returns a matrix whose dimensions are specific to a feature or features detected. Each feature detected returns a matrix with one row.The object detector therefore returns an M-by-4 matrix defining M bounding boxes containing the detected objects. This method performs multiscale object detection on the input image, I. The detector detects objects defined by an input string, the input string defines the type of object to be detected. The detector can detect a variety of objects, including faces and a person's upper body. The default is configured to detect faces.

The M by 4 matrix returned by the detector is *the key* because M returns a value which is equal to the number of items of a specified object that has been detected. In this project the term ‘profiler’ has been coined to refer to the type of detectors used. Following from this, in this work five (5) ‘profilers’ were used, they are:

1) Face
2) Upper body
3) Nose
4) Mouth
5) Profile Face

These model detectors or ‘profilers’ have been trained previously by default integration in matlab. The idea is to get inputs from the returned values of each M by 4 matrix which will be indicate how many of the afore mentioned ‘profilers’ have been detected by using matlab coding.

The image is imported into the code, resizing is applied and each image is subjected to a grueling test algorithm involving each of all five profilers to determine all aspects of upper body human features. Profilers return a matrix whose rows (M) indicate the number of features detected. These profiler returned row values are examined; based on the occurrence of the mode, or the equality between the face detector, nose detector or body detector and the best fit for each of these conditions is chosen.

VII. Analysis of Program

In order to detect faces the following code is used:

Matlab code:

```
% Detect faces
bbox = step(faceDetector, I);
```

This detects a face that has already been called with the imread() command and stored as a variable I.

Matlab code:

```
A=bbox;
```

This instruction assigns a variable A to the matrix bbox which is a face detector object.

For instance the value of A after processing an image ‘2face and annie.jpg’ is:

A = 92 11 29 29
65 22 31 31

A is a 2 by 4 matrix and in this case M(row) is 2 and hence the face detector has returned M as 2 meaning in simple terms two faces have been detected.

The result of M can be obtained thus:

[x y]=size(A);

Hence x=M=2.

To further improve on its accuracy and detecting abilities other profilers are introduced.

Matlab code:

```
% Create a cascade detector object.
bodyDetector = vision.CascadeObjectDetector();
```

To detect upper bodies the code above is used to declare the object detectorfunction.

Matlab code:

```
bbox = step(bodyDetector, I);
```

This detects an upper body that has already been called with the imread() command and stored as a variable I.

Matlab code:

```
B=bbox;
```

This instruction assigns a variable B to the matrix bbox which is a body detector object.

For instance the value of B after processing an image '2face and annie.jpg' is

B = 78 3 68 62
38 8 85 77

B is a 2 by 4 matrix and in this case M(row) is 2 and hence the body detector has returned M as 2 meaning in simple terms two upper bodies have been detected.

The result of M can be obtained thus:

[x y]=size(B);

Hence x=M=2.

In this same manner the other 'profilers' ;i) Nose ii) Mouth iii) Profile-Face were tested and all the results examined by the algorithm developed in order to pick the best fit and the highest occurrence to accurately pinpoint the number of people in the image selected. See figure 1.

VIII. Result Analysis

Testing of the system was performed using a test set of thirty-five (35) different images. Each test was performed under the following conditions:

1. Image size (Resolution).
2. Profilers test.
3. Algorithm test.

These were done to determine the optimum conditions that would give the highest accuracy in order to improve on the system and set operational conditions for implementation to ensure best results. The end result would thus enable development of fully robust software. From figure 2 and 3, y represents actual values, x detected. The equation describes the model of the system.

From testing both models it (the system) was found to have an accuracy of between 84.9-94.8%. This shows that the algorithm developed for assessing the various values returned by the profilers is suitable enough to make the system viable for implementation.

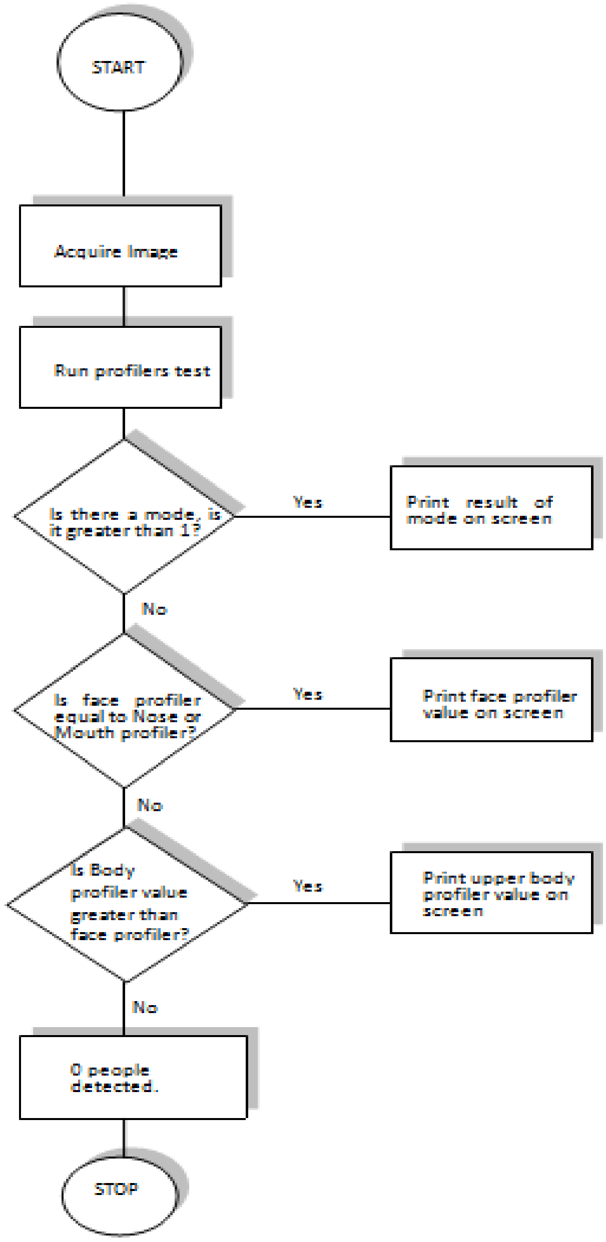

Fig.1. Flow Chart for developed algorithm

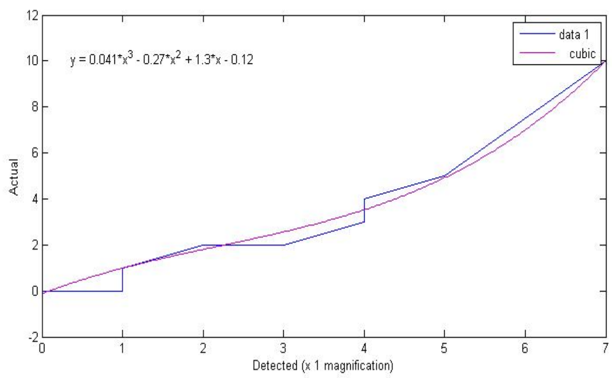

Fig.2. Plot of Actual number of people against detected (x1 magnification)

International Journal of Electronics Communication and Computer Engineering
Volume 5, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

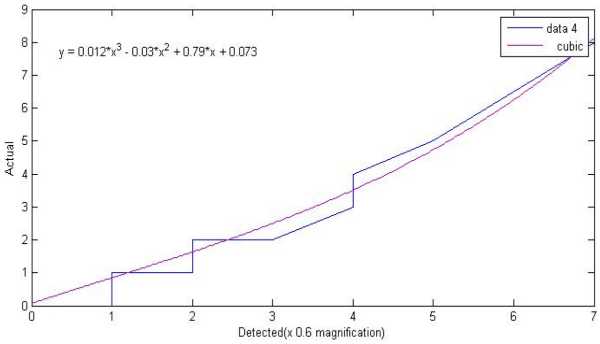

Fig.3. Plot of Actual number of people against detected (x0.6 magnification)

IX. Conclusion

In this project we have been able to build a human detection/counting system using real life still images. This was made possible by using the Viola-Jones detection algorithm as its backbone. We have been able to use a series of detectors based on the above algorithm to enable refine the recognition and detection process. The system has been tested to determine its accuracy, its limitations and the optimal operating conditions. It has been found to be at least 84% accurate under poor visual conditions and more than 90% under optimum visual conditions. With some minor embellishments in the area of aesthetics (graphical user-interface) the program can be ported to other programming languages and be employed for use in the outside world.

References

[1] Digital Image, http://en.wikipedia.org/wiki/Digital_image, site accessed 7th June 2014
[2] Digital Imaging, http://en.wikipedia.org/wiki/Digital_imaging, site accessed 7th June 2014
[3] Paul Viola & Michael j. Jones, International Journal of Computer Vision 57(2), 2004, U.S.A, "Robust Real-Time Face Detection." Pdf Pages 139-140, 143-145.
[4] Viola-Jones algorithm, http://en.wikipedia.org/wiki/Viola Jones_object_detection_framework, site accessed 7th June 2014
[5] Object detection, http://www.mathworks.com/discovery/object-detection.html, site accessed 10th June 2014
[6] Applications of object detection,
[7] http://www.di.ens.fr/~laptev/objectdetection.html , site accessed 10th June 2014
[8] Rafael.C.Gonzalez , Richard E. Woods , Steven L. Eddins , 2009, "Digital image processing using Matlab. "Gatesmark Publishing® A Division of Gatesmark,® LLC, United States of America, Second edition. Pdf, Pages 13-15
[9] Daniel L. Swets&Bill Punch, Year unknown, ,"Genetic Algorithms for Object Localization in a Complex Scene." Publisher unknown, U.S.A, Pdf Pages 2-6.
[10] E.R. Davies, 2012, "Computer and Machine Vision", Fourth Edition. Elsevier Inc., pages 459-466.
[11] B Bhanu and O. Faugeras, "Segmentation of images having unimodal Distributions." 1982, IEEE Trans Pattern Anal. Mach Intell., vol PAMI-4,Pages 408-419.
[12] S.Boukharouba, J. M. Rebordao, and P.L. Wendel, 1985, "An amplitude segmentation method based on the distribution function of an image, "Comput. Vision Graphics Image Process., vol, pages 47-59.
[13] W.Doyle, 1962. "Operation useful for similarity-invariant pattern recognition,"J. Assoc. Comput. Mach., vol 9, pages 259-267.
[14] Rowley, H., Baluja, S., and Kanade, T. 1998. Neural network-based face detection. IEEE Patt. Anal. Mach. Intell., pages 20:22-38
[15] Sung, K. and Poggio, T. 1998. Example-based learning for based face detection. IEEE Patt. Anal. Mach. Intell, 20:39-51
[16] Tieu, K. and Viola, P.2000, Boosting image retrieval. In Proceedings of the IEEE conference on computer Vision and Pattern Recognition.
[17] Tieu, K. and Viola, P. 2000. Boosting image retrieval.In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Author's Profile

Isidaehomen. A. EDEOGHON
is a lecturer in the department of Computer Engineering of the University of Benin, Benin City, Nigeria. He holds a B.Eng. and M.Eng. degree in Computer Engineering and Electrical/Electronic Engineering in 2005 and 2014 respectively. His area of research includes software engineering, image processing and telecommunication systems.

OTASOWIE O. Paul
received a Ph.D (2008) in Electronics and Telecommunication Engineering from the University of Benin, Benin City Nigeria .
He is now an Associate Professor at the Department of Electrical and Electronic Engineering, University of Benin, Benin City, Nigeria. He is author and co author of more than fourty Scientific Papers and three Books Chapters.
He has supervised over twenty masters degree candidates and currently supervising four Ph.D. Students.
His research interest includes micro wave Signal Propagation, Mobile Communication, Image Processing, Energy Efficiency, Solar Energy for Home Lighting and BTS among others.