

Using Face Detection and Face Recognition Technology to Develop a Feature in Video Cameras for Crowd Counting

Jane Doe
Researcher

ABSTRACT

Having a digital camera with a feature to count people in a gathering without recounting a face will help to fortify modes of acquiring attendance figures for large gathering and provide exact data for records. There have been conventional methods used to know attendance in a gathering such as ticket sales. For some other gathering it is manually determined by dividing an area occupied by a crowd into sections, determining the average number of people in each section and multiplying the number of sections occupied. Aerial photography and satellites are also used for crowd counting. These methods provide close estimation and conflict sometimes when two or more methods are used. Digital camera technology develops yearly; the technology proposed in this paper suggests adding a new feature to digital cameras, such that it can be used to count people in a gathering aside covering the event normally. To have this feature, face detection and face recognition technology will be used. Face detection to detect human faces and count them; such that faces counted will be saved temporarily and the face recognition technology will ensure that faces counted at the time will not be counted again. This paper presents crowd count feature in details

General Terms

Digital Camera, Computer Vision

Keywords

Crowd Counting, Face Detection, Face Recognition

1. INTRODUCTION

Face detection is a computer technology that determines the locations and sizes of human faces in arbitrary images, it simply detect faces and ignores everything else. This technology is used for applications like video surveillance and image retrieval. Face detection technology is also used for computer-human interaction. [1] This also makes it measure up for use for the crowd count feature proposed in this paper. Face recognition is a biometric identification by scanning a person's face and matching it against a library of known faces. It is used for applications such as security systems and psychological processes. [2]

The face recognition technology will complement the face detection technology to accomplish the crowd count feature proposed here. To achieve this crowd count feature; face detection technology will detect faces; such faces are counted and saved temporarily in a library. Since the camera will still be used for its conventional purpose, it will return to positions where it has counted faces. The face recognition system will help compare faces with those in the library to ensure that no

face is recounted. Facial images will be stored temporarily to avoid repetition; saved images are automatically deleted after that viewing session especially when figures have been computed. This feature is expected to have high accuracy since it is concerned with detecting human faces. This technology will give correct figures since its accuracy will exceed other methods of crowd counting. The interesting part of this technology is that it will be a feature that can be enabled or disabled in coming models of video cameras. Meaning that along with a camera's high resolution, high zoom lens magnification and picture quality; digital video cameras will also be used to count people in any gathering so long it views the area covered by the crowd. This technology is much different from the people counter technology which is used to know the number of times people enter and leave a particular place. Entry and exit count is more important than individual count. [3] The crowd count feature is concerned with individual count especially as their face stays same during that period. The need for the crowd count feature is to have a trusted mode of getting exact figures in large gatherings. Since this feature will be available in coming cameras more than one can be used in a particular event. The feature will be useful for press reports, to determine the number of people in entertainment shows, number of students at a gathering, those present at a sporting event, and many other crowd intensive occasions. Table 1 simply presents the process of operation of the proposed Digital Camera using an individual's face as a reference. The camera counts a face and does not repeat the count since the face is temporarily saved in the library.

Table 1. Crowd Count Process with reference to a face

Process →	→	→	→
Camera View	A face is detected	The face is counted	The face is temporarily saved in the Library
Camera returns to previous position	The Face is Recognized	The face is not recounted	Since it is the same face previously saved in the Library

2. THEORETICAL DEVELOPMENT

Face detection and face recognition technology have been developed through the years; research has also made them useful for more applications. For systems where they are conventionally used, their delivery is usually increased. The novel feature presented in this paper will have face detection and face recognition technology work together; this section presents their peculiarities.

2.1 Face Detection Technology

This is a computer technology that detects facial images and ignores anything else; it determines the location and sizes of human faces in digital images. There are feature based and image based algorithm approaches for face detection. Feature based uses edge detection, skin color and symmetry analysis. Image based algorithm uses neural networks. [1] The skin color processing of the feature based algorithm is faster than any other facial feature. To obtain this, it is necessary to identify those pixels which fall within a certain range of RGB (Red Green Blue Color Model) values, and categorize them as skin pixels. [4] Skin color segmentation helps to reject non skin color region from the input image and morphological operations helps to clean up that image and remove noise. Connecting analysis is usually done on the image to obtain the various connected regions; these regions will be separated until a single region is split further. [4][5] Edge detection and Symmetry transform are used to further separate the region. Retinal connected neural network examines small window of an image and decides whether each window contains a face. This system acts between multiple networks to improve performance over a single network. [6] The system applies a set of neural network-based filters to an image then uses a sub-system to combine the outputs. The filters examine each location in the image at different scale, to get locations that might contain a face. The sub-system then merges detections from individual filters and eliminates overlapping detections. [6]

The retinal connected neural network algorithm can be used to achieve the crowd count feature, since the camera will be used to count faces in a crowd, it is strictly based on facial images

so issue of false detections (non-facial images) does not directly apply. Since this camera will be carried on or around the podium, digital zoom will be useful in viewing faces at the back. Some faces will not be seen in totality so the face can be viewed in a 20×20 window, starting from a little above the eyelid and ends below the lower lip. [7] It can be increased to view the whole face in a 30×30 pixel for those in front. Recent technology presents multi-view face detection meaning that even if a person's face is rotated along the vertical or left axis (out-of-plane rotation) or both; at the time the face is to be counted, the person will be counted once without repetition.

2.2 Face Recognition Technology

As a computer application, this system verifies or identifies a person's face from a digital image or video frame from a video source. It compares selected facial images in a facial database from the image. [2] Some facial recognition algorithm identifies faces by extracting landmarks or features from an image of the person's face. An algorithm may analyze the position, size, and/or shape of the eyes, nose, cheekbone and jaw. These features are then used to search for other images with matching features. Other algorithm normalize a gallery of face images and then compress the face data, only saving the data in the image that is useful for face detection. [2] The Facial recognition system compares face detected to faces in the database; crowd count will be a feature on a digital camera which will not be connected to any facial database. Libraries will temporarily store the faces detected; therefore a library or libraries will act as the facial database.

Three dimensional (3D) face recognition as part of this crowd count technology can help increase the efficiency for face recognition purposes. The 3D technique uses 3D sensors to capture information about the shape of a face. This feature identifies distinctive features of a face such as the contour of the eye sockets, nose and chin. This technique is not affected by changes in lighting. [8][9] Recent technologies show that high resolution images consist of facial images with an average of 250 pixels between the centers of the eyes, making the 3D efficient with this also. The count feature presented in this research will be operated with a button on the digital camera, when activated; it will detect faces, counts and save them simultaneously.

(REI QA Testing)

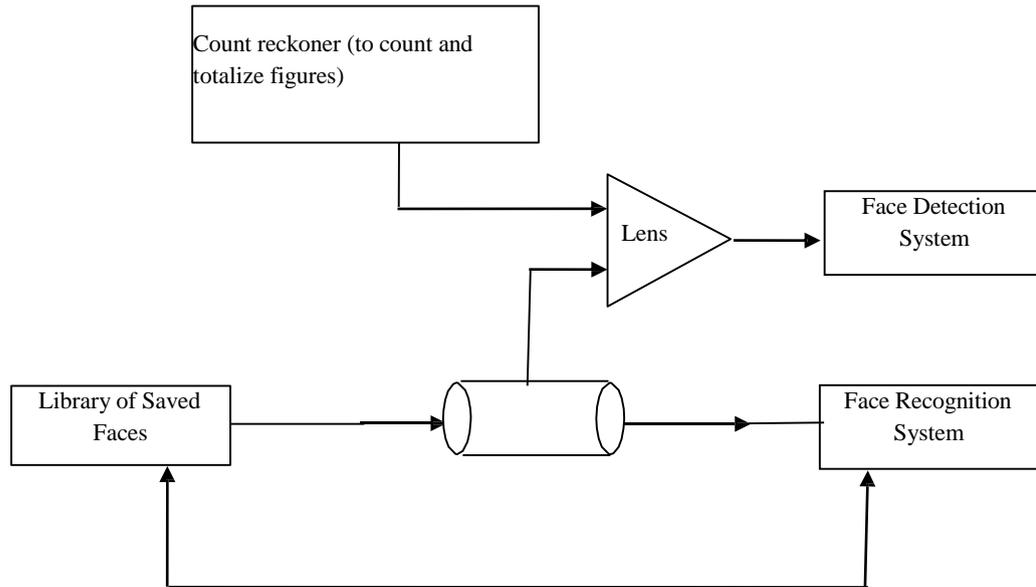


Fig 1: Schematic Diagram Showing the Crowd Counting Digital Camera Setup.

3. EXPERIMENTAL PROCEDURE

The facial recognition, detection and digital camera technology will be combined in one device to achieve this objective; normal camera size will be maintained since these technologies will be made to fit into it. Digital cameras have some features that support the crowd count feature. For example, the optical disc standard storage system provides enough digital storage to store hours of video content meaning that it can also be used to temporarily save some or all of the faces counted; the numbers of lines in the vertical display resolution, the scanning system and the number of frames or fields per second help define images clearly will aid counting and saving without distortion. [10] Figure 1 above, shows a skeletal makeup of the Crowd Counting Digital Camera, the diagram shows the Face Detection system close to the lens such that faces are viewed and identified; and saved in the library connected to the face recognition system to avoid faces being recounted. The Count reckoner counts faces and totalizes the number of faces counted. The face recognition algorithm is divided in two modules: a face image detector that finds human faces and a face recognizer determines who the person is. Both technologies allow the same framework; they both have a feature extractor that transforms the pixels of the facial image into a useful vector representation and a pattern recognizer that classifies the feature vector and searches the memory [11] to ensure that the incoming face has not been counted previously. There are various algorithms developed for face detection and face recognition technology, usually based on models like skin color and neural network. It has been observed that different human skin color give rise to compact clusters in color space such as normalized RGB (red, green, blue), YCbCr and HSI Color spaces amongst others.[4] For skin color based face detection in RGB Color space; the

pixels for skin region can be detected using a normalized color histogram that can be further normalized for changes in intensity on dividing by luminance. The pixels can then convert an [R,G,B] vector to an [r, g] vector of normalized color which provides a fast means of skin detection. [4] Some face recognition algorithm uses nodal points and inter-nodal distances to create a value unique to each facial photograph. This algorithm can be programmed with Java Swing components to generate Graphic User Interface (GUI) software. After the values are saved to the database, the captured face unique value will be matched to the closest values in the database. The will help determine that a match will be found within a reasonable margin of error. To determine the probability and accuracy of the GUI-based face recognition program, an in-depth statistical analysis can be run on the data. [12]. Many more algorithms are available for face detection and recognition system, meaning that certain subsystems may act as arbitrators between each system. It therefore implies that further developments on exiting models will leverage the crowd count feature.

4. RESULT AND DISCUSSION

Options to come with the crowd count feature include: enable/disable, reset count, count again to confirm, notification (if a face is not seen clearly to be counted) the camera will seek to count such face when in that direction again. Some of the options will work automatically as default and others will depend on the command given. Count feature maybe enabled and disabled by the user as desired, count can be repeated as desired by the user or depending on the number of hours the event will last, recount by default can be disabled / enabled to ensure further accuracy of figures. Once the count feature is enabled, the user is expected to move slowly with the camera to enable it detect, count and save faces temporarily. After counting some parts, the user is expected to

(REI QA Testing)

zoom to other parts. Due to certain obstruction or space between individuals in the gathering, the camera may not see some faces; it is expected to come back to that direction. Hair pattern recognition may also help out of such situation when developed. No two human hair patterns are same, even if they are similar in appearance; height, arrangement (head shape) and growth direction differs. This simply means that from the top view, the camera may recognize the hair of an individual, count and save so when it is returned to that direction, it will not count the person. This may also be necessary as people are tightly packed in some gathering and heads rather than faces are seen from the podium. Hair pattern recognition algorithm for crowd count feature will certainly increase the efficacy of attendance figures from aerial photography. Digital cameras will be useful to count people in any kind of setting when this technology is available. This research presents details on how to create a crowd count feature for digital cameras; a number of algorithms will be written to have the crowd count feature while research into hair pattern recognition is advanced to ensure more options for accuracy for this feature.

4.1 Conclusion

Technology to enable a digital video camera count people in a gathering is some algorithm away. This feature will be developed overtime such that all digital cameras will come with it. This will increase versatility of digital cameras with some other features that will be added to them. Increased market share, technology solutions, accurate estimations are some of the bring-along of this technology.

5. GLOSSARY

Accuracy:

A catch-all phrase for describing how well a biometric system performs; simply put is the quality of being correct, true or exact, with little or no error.

Aerial photography:

Refers to images not supported by ground based structure such that photographs of the ground are taken from elevated position(s).

Algorithm:

Is a limited sequence of instructions or steps that tells a computer system how to solve a particular problem. A biometric system will have multiple algorithms, For example: image processing, template generation, comparisons, etc.

Biometric Identification:

Is an automatic identification of living individuals by using their physiological and behavioral characteristics usually called biometrics. Biometrics can be used to describe a characteristic or a process. As a characteristic: is a measurable biological (anatomical and physiological) and behavioral characteristic that can be used for automated recognition. As a process: are automated methods of recognizing an individual based on measurable biological (anatomical and physiological) behavioral characteristics.

Biometric Systems:

Are multiple Individual components (such as sensor, matching algorithm, and result display) that combine to make a fully operational system. A biometric system maybe a component of a larger system; It is an automated system capable of:

1. Capturing a biometric sample from an end user
2. Extracting and processing the biometric data from that sample.
3. Storing the extracted information in a database.
4. Comparing the biometric data with data contained in one reference or more.
5. Deciding how well they match and indicating whether or not an identification or verification of identity has been achieved.

Capture:

Or to capture is a process of collecting a biometric sample from an individual via a sensor.

Crowd counting:

Is a technique or set of methods used to count or estimate the number of people in a crowd.

Database:

A collection of one or more computer files. For biometric systems, these files could consist of biometric sensor readings, templates, match results, related end user information, etc

Edge detection:

Is used to identify points in a digital image at which the image brightness changes sharply or more formally has discontinuities.

Feature Extraction:

Is the process of converting a captured biometric sample into biometric data so that it can be compared to a reference.

Gallery:

Is the biometric system's database, or set of known individuals, for a specific implementation or evaluation experiment.

Graphical User Interface (GUI):

Is an object-oriented display format that allows the user to select from menus and icons, using either a mouse or keystroke commands.

HSI Color Space:

HSI is the corresponding color model used to describe three major color properties which are Hue, Saturation and Intensity.

Image retrieval:

Is a computer system for browsing, searching and retrieving images from a large database of digital images.

Magnification:

Is the act of expanding something in apparent size.

Neural Network:

Is a computer architecture in which processors are connected in a manner suggestive of connections between neurons; has the ability to learn by trial and error.

Noise:

Unwanted components in a signal that degrade the quality of data or interfere with the desired signals processed by a system

Pixel:

Is a picture element, usually the smallest element of a display that can be assigned a color value.

Resolution:

Is the number of pixels per unit distance in the image. It usually describe the sharpness and clarity of an image.

Satellite:

Is anything that orbits something else, usually used to describe man made equipment that orbits around the earth or the moon.

Symmetry analysis:

Is the degree of symmetry in a Three Dimensional shape, under some class of transformations.

Video surveillance:

Is the use of video cameras to transmit a signal to a specific place on a limited set of monitors.

YCbCr Color Space:

It belongs to a family of television transmission color spaces and was developed due to increase demands for digital algorithm in handling video information. It is used as a part of color image pipeline in video systems. [13]

6. ACKNOWLEDGMENTS

I am grateful to organizations, institutions and individuals who have written papers, articles and presentations in related topics on Face Detection and Face Recognition Technology which are keywords of this paper; efforts and results of scientists across the globe towards making the world a better place will remain the determining factor for present and future decisions of change, Thank You All.

7. REFERENCES

- [1] Jesorsky, O., Kirchberg, K.J., and Frischholz, R.W, "Robust Face Detection Using the Hausdorff Distance", Third International Conference on audio and video based Biometric Authentication, Springer, Lecture Notes in Computer Science, LNCS-2091, pp 90-95, Halmstad, 6-8 June 2001.
- [2] Swarupa, N.V.S.L., and Supriya D, "Face Recognition Systems International Journal of Computer Applications", (0975 - 8887), Volume 1 – No. 29, 2010.
- [3] Aik, L.E., and Zainuddin Z, "Real-Time People Counting System using Curve Analysis Method" International Journal of Computer and Electrical Engineering, Vol 1, No.1 (1793-8198), pp77, 2009
- [4] Singh, S.Kr., Chauhan, D.S., Vatsa, M., and Singh A, "Robust Skin Color Based Face Detection Algorithm", Tamkang Journal of Science and Engineering, Vol. 6, No 4, pp 227-234, 2003.
- [5] Kovac, J., Peer, P., and Solina F, "Human Skin Color Clustering for Face Detection" pp 1-5 Eurocon 22-24 September 2003.
- [6] Hannuksela J, "Facial feature based head tracking and pose estimation" Diploma Thesis for Department of Electrical and Information Engineering, University of Oulu, Oulu, Finland. 2003.
- [7] Do, T.T., and Le, T.H, "Facial Feature Extraction Using Geometric Feature and Independent Component Analysis Department of Computer Sciences, University of Natural Sciences, HCMC, Vietnam, pp 4, 2009.
- [8] Heseltine T., Pears N., and Austin J, "Three Dimensional Face Recognition Using Surface Space Combinations" Department of Computer Science, The University Of York, 2008.
- [9] Kakadiaris I., Passalis G., Toderici G., Murtuza N., and Theoharis T, "3D Face Recognition" Encyclopedia of Biometrics, S.Z. Li, Ed. Springer, pp. 329 -338, 2009.
- [10] Andor Technology, "Digital Camera Fundamentals" Andor Publications, 2006.
- [11] Huang Y. H., and Fuh C.S, "Face Detection and Smile Detection" Proceedings of IPPR Conference on Computer Vision, Graphics and Image Processing, Shitou, Taiwan, A5-6, p. 108, 2009.
- [12] Rao V, "Face Recognition: Is It a Match" Oklahoma Academy of Science Publication, 2009.
- [13] National Science and Technology Council's (NTSC) Subcommittee on Biometrics, Biometrics Glossary, 9-14-2006.