AJMAS

https://journal.utripoli.edu.ly/index.php/Alqalam eISSN 2707-7179

Original article

The System for Estimating the Number of People in Digital Images Based on Skin Color Face Detection Algorithm

Samar Ittahir, Tarik Idbeaa*, Hisham Ogorban

Department of Computer Science, Faculty of Science, Garyan University, Garyan, Libya,

ARTICLE INFO

Corresponding Email. *dr.Tarik.idbeaa@gu.ed.ly*

***Received**: 12-03-2022* ***Accepted**: 22-04-2022* ***Published**: 25-04-2022*

***Keywords**: Face Detection, Skin Color, Face Recognition, Pattern Recognition.*

ABSTRACT

Counting the number of people in many estimation systems, such as still images or video frames, is a buoyant research area that is challenging in the field of computer vision. It plays a considerable role in a variety of applications, such as security, management, education, and commerce. The purpose of this paper is to suggest a system to estimate the number of people in digital still images based on the Face Detection method. This system composed of two parts: face detection and counting of detected faces. In the detection step, the Skin Color Face Detection method was applied on the input of a digital still image. In the counting part, the obtained detected faces by the Skin Color Face Detection method have counted to estimate the number of people in an input color image with simple software and simple low-cost hardware. The skin color face detection algorithm was tested using 133 images from the People Image Groups dataset, which contains about 2573 color images of people, to test the proposed system. based on the obtained results, the best precision achieved of the proposed Skin Color face detection algorithm was 85%

Cite this article: *Ittahir S, Idbeaa T, Ogorban H. The System for Estimating the Number of People in Digital Images Based on Skin Color Face Detection Algorithm. Alq J Med App Sci. 2022;5(1):215-225.*
https://doi.org/10.5281/zenodo.6483654

INTRODUCTION

People counting is a process used in many computers vision and pattern recognition systems to estimate the number of people in a digital image or video. Such information can be used for further analysis in a wide range of practical applications related to education, commerce, security, management etc. For instance, people counting systems used for security provide information about the total number of people in the building and number of people on each floor in order to control the number of visitors. People counting can also be used for pedestrian traffic management [1-4] and for fire management, where it is considered one of the most important tools in the event of a fire. Additionally, people counting applications are widely used in commerce domains to measure marketing effectiveness. As the given examples show, people counting research is one of the most important active areas in the field of computer vision which can be applied in many areas of daily life. The output of people counting systems may be passed to other management systems used in different areas such as public transport, industry, stations, airports, etc. Many examples illustrate the importance of people counting systems. For instance, in the commerce domain, shopping centers, hypermarkets and others depend on visitor statistics to measure the marketing effectiveness and attractiveness of a merchant site, and to make new decisions. There are different people counting techniques, which are broadly classified into two main categories: vision-based counting systems and non-vision-based counting systems. The vision-based techniques can again be categorized into tracking-based and non-tracking-based counting systems (Figure 1). Each class has specific strengths and weaknesses. In general, the class of non-vision-based people counting systems includes all systems that are not based on video or image data from a camera, but on devices such as a heat sensor, light from an infrared beam, or pressure sensors [5, 6].
In recent years, vision-based systems have become more popular in different scenarios like hotspots, streets, and buildings because vision-based systems make excellent use of the infrastructure provided by security networks. In addition, vision-based people counting relies on information provided by a security camera's video feed to count people crossing.

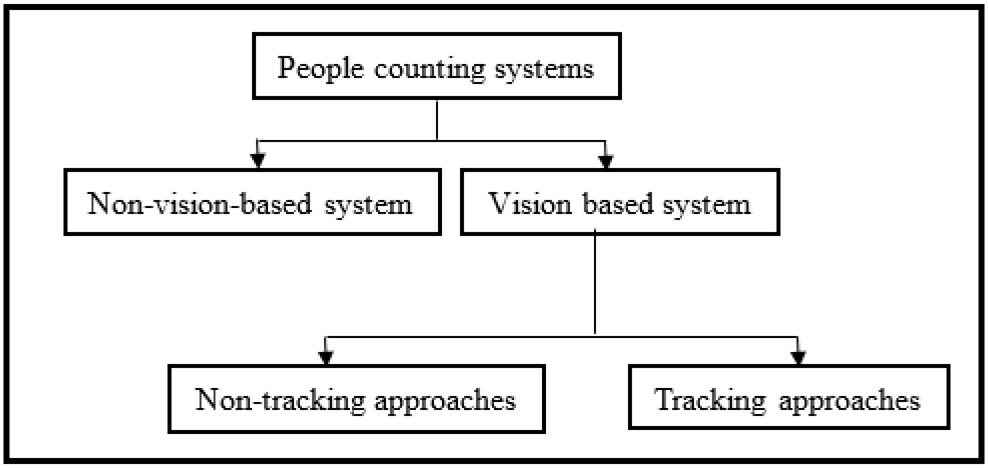

Figure 1. Classes of people counting system

The main strength of these systems is that they are not limited to narrow corridors or doorways, but the disadvantage is that they are limited by the coverage area of the cameras. Additional challenges for vision-based systems include differences in environmental conditions and occlusions between people. Some systems solve this problem by installing cameras looking directly down to the floor [9]. The non-tracking vision-based approach counts the number of people without any human tracking of groups or individuals. Thus, this approach is suitable for a large number of people in a frame, unlike the tracking approach, which is not well suited for a growing number of people. Due to their ability to count large numbers of people, these systems are generally known as crowd surveillance systems. The main weakness of the non-tracking approach is that it provides the number of people that are included in a frame, but it does not provide information on the directions of people's movements [10-12]. In this paper, the number of people in a static digital image based on skin color face detection technique were counted and estimated; It plays a very important role in the field of computer vision, which is used in various varieties of applications, such as facial recognition, photography, and marketing. Face detection identifies human faces in an input digital image with an emphasis on frontal human face detection. Therefore, face detection is an important part and the first step of automatic face recognition.

Additionally, people can be counted based on the number of faces detected in a digital image. The main problem was discussed in this study is how to count the number of people in different images. People may look different in pictures. For instance, counting the number of people in an image with people with light faces is different from counting the number of people in a crowd. Besides that, there can be a lot of different challenges in the pictures. For instance, occlusion between people and varying environmental conditions are some of the challenges of people counting systems. Input digital images have a broad degree of variation in unrestricted environments, illumination level, image dimensions, skin color, similarity between background color and clothing of people, the quality of the image, people's closeness to each other, which makes it difficult to detect people. Therefore, the aim of this study is to overcome these challenges. To achieve this, the process of the introduced system works on the skin color face detection algorithm to determine the number of people in an image. In addition, the proposed method requires hardware components of low cost and complexity. People counting systems based on the skin color face detection method provide more information, such as the location and appearance of people, that is not detected by visionless infrared beam systems, which counts the number of people without further information.

In other words, its information is limited to people counting data only. In addition, the performance of people counting systems based on an infrared counter is less accurate than digital image and video systems. The information obtained from a people counting system based on digital images can provide useful information for further analysis in different practical applications, for example, during a fire, the approximate number and the location of people in a building can be obtained. This paper has been divided into five main sections. We begin by providing a general background on people counting systems in the field of computer vision with a summary of potential issues that can arise when working on a people counting system. In section 2, we provide a review of the literature related to the field of people counting and its methods. Next, in Section 3, we explain the people counting method in detail step by step. Additionally, we present an overview of the datasets used in this study. In section 4, we discuss all the results obtained from our experiments. Finally, in section 5, we summarize the conclusions. Additionally, we list future work to give readers ideas on how to improve people counting systems.

LITERATURE REVIEW

Counting the number of people in images is an interesting study area and has been given significantly attention in recent years. In addition, it is perceived as a substantial module for many computer vision applications and provides useful information that can be used for further analysis in a wide range of practical applications related to education, commerce, security, and management. Counting the number of people in videos and digital images are types of the vision-based classes. Counting the number of people from images counts the number of people without any humans tracking of groups or individuals. Moreover, this approach is suitable for very large number of people in a frame, unlike tracking approach that does not scale well with a growing number of people. Nevertheless, it does not give information about the movement directions of people. Although counting the number of people from images and counting the number of people from videos have many mutual ingredients, they might be diverse in the core approaches of classification. Therefore, in following sections, different methods and approaches were presented that have been proposed in order to estimate the number of people.

There are many different People Counting Systems of Vision-based systems, which are the most commonly used class People Counting Systems, which requires multiple frames for processing. In addition, the counting process of vision-based systems is extended with the direction of blobs to enrich provided information. Vision-based systems are the most widely used class of people counting systems, requiring multiple frames for processing. Also, the counting process of vision-based systems is extended with the blob direction to enrich the provided information.

Much research has been done in this area. For instance, Zhao et al. [13] developed a people counting approach based on face detection, trajectory tracking and classification. A standard face detection method was used in addition to a face tracking method by combining a kernel-based tracking algorithm with a Kalman filter. They extracted the histogram of angles of the neighboring points of each potential face trajectory. K-NN was then used to classify the extracted data. As a result, by applying this approach to a dataset with more than 160 possible person trajectories, the approach achieved an accuracy rate of up to 93%.

Lempitsky and Zisserman [14] proposed a very flexible MESA-based distance learning framework for counting people in images. They confirmed that using a limited amount of training data can lead to much higher accuracy than regression counting, which directly optimizes accuracy over the image. The advantages of their system are its accurate counting and fast processing time. Using 2,000 frames of video from a camera overlooking a busy pedestrian street, they achieved a detection accuracy of 96.5%.

Liu et al. [15] present a surveillance system composed of crowd segmentation, visual tracking, counting recognition module and automatic calibration. The system was able to count the number of people leaving or entering a specific location by segmenting groups of people into individuals. To evaluate this system, they recorded a 10-minute video taken from a fixed camera about 6 meters above the ground, and the system worked effectively.

Bansal and Venkatesh [16] presented a method that uses SIFT, Fourier analysis, wavelet decomposition, GLCM features, and low-confidence head detections as multiple sources to estimate the number of people in high-density crowds from images. fixed. Using a dataset with 100 images, the method reported an accuracy of 99.4%.

Subburaman et al. [17] estimated the number of people in a crowded scene as a function of the detection region of the head using a state-of-the-art cascade of boosted integral features. Two different databases were used to assess their approach: the PETS 2012 database and the Turin Metro Station database. Therefore, an accuracy of 95% was achieved.

Rahman and Islam [18] identified the object boundaries of the entire image by converting it to a grayscale image and finding an appropriate threshold value there. Based on these limits, they counted the number of objects in the image. Additionally, the method used a watershed segmentation technique to overcome the problem of image noise, which results in counting fake objects. The method gives an accuracy of up to 96% with high definition images.

Topkaya et al. [19] presented a people counting system based on the use of different characteristics such as time, place and color. In their system, they used the HOG feature and the nonparametric nature of Dirichlet Process Mixture Models (DPMMs) to detect the number of groups and count the number of people.

The system was evaluated on the PETS [20], Peds2 [21], and BEHAVE [22] data sets and achieved an accuracy of 95.1%.

Wang et al. [23] attempted to present a simple method to count the number of people in an image using feature extraction and pattern recognition techniques. First, a Gaussian masking method and morphological background modeling were used to more efficiently detect image targets. then the HOG feature was used to extract meaningful features from the shape and appearance of people. In addition to the edge functions and texture functions, the HOG features were used to train an 11-support vector regression machine and present the number of people in the image. Thus, the method was able to achieve a

counting performance of 97.12% by applying the method to the UCSD dataset, which was partitioned into 600 images as the training set and 1,200 images as the training set.
Chan et al. [24] showed how to estimate the size of inhomogeneous crowds without using tracking techniques. Its privacy-preserving system relies on the use of the dynamic texture motion model to segment the crowd into homogeneous motion groups. then the features from each segmented region was extracted., the Gaussian regression process was then used to evaluate the consistency between the number of people and the characteristics per segment. Finally, the system was validated on a large pedestrian dataset containing 2,000 images. The result of this method was a success rate of 98%.
Dan et al. [25] present an accurate people counting system based on sensor fusion, which is robust to crowding conditions and lighting variations. The proposed system is composed of person detection based on human models, a lost depth data recovery algorithm, and individual tracking using color and depth data. First, a morphological operator processes the depth image to relieve depth artifacts such as data loss and optical noise. The human model then extracts the object from the pre-processed depth image. Finally, by applying the bidirectional matching algorithm, the track of the detected object is established. Therefore, experimental results in various test environments show that the algorithm achieves an accuracy of more than 98%. Teixeira and Savvides [26] used a lightweight method on indoor camera sensor networks to count and locate people. The algorithm uses a motion histogram to detect people based on size and motion standards. Their algorithm was tested and implemented on a network of iMote2 sensor nodes. As a result, the method was able to achieve a counting accuracy of up to 88.6%.

SYSTEM OVERVIEW

In this paper, skin color-based face detection was tested to detect multiple faces in a digital still image and then those faces were counted to estimate the number of people in an input image. People's skin color was used as an effective means of feature extraction. The skin color method used in several existing face detection applications. The skin color face detection workflow used to find the region of the face in an input image based on the color model. It is a simple and popular approach. Skin color pixels are represented by different color models such as RGB [27-29], Hue Saturation Intensity (HSI), or Hue Saturation Intensity Value (HSV) [30, 31, 32, 33], YES [34], CIE LUV [35], YCbCr [36,37], CIE XYZ [38] and YIQ [39,40], normalized RGB [41-47], etc. Detect a human face the skin color feature is not enough to achieve sufficient results. If only skin color information is used, it may cause an increase in the rate of incorrect face detection. For instance, when human faces are close to each other, it is difficult to distinguish them. Therefore, many modular systems use a combination of various features, such as size, shape, skin color characteristics, and color segmentation to detect heads and faces [32, 35, 48, 49]. The overview of the proposed system is illustrated in Figure 2.The general algorithm of this method can be divided into seven steps, which are:

i. Input a digital image.
ii. Alter the color space.
iii. Detect face region.
iv. Extract face model.
v. Reduce noise.
vi. Detect face.
vii. Count people

To improve skin detection performance in color images the combination of HSV and YCbCr color spaces was used. then the faces and count each detected face in an input digital image was perceived.

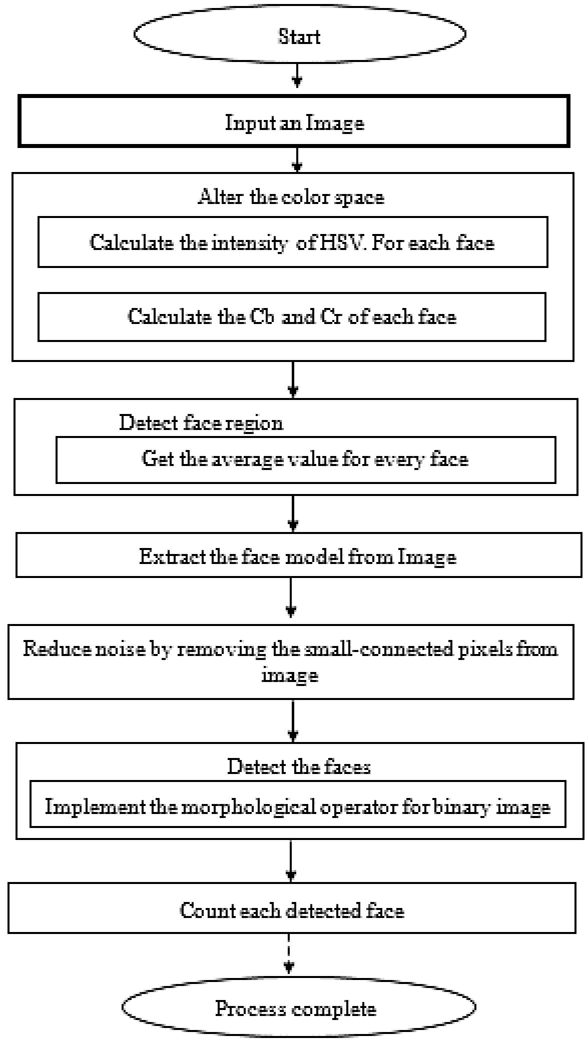

Figure 2. Steps of people counting based on Skin Color face detection.

Color Space Alteration

There are many types of color space that can be used to label the pixels in an image as skin color such as RGB, HSV or HIS, normalized RGB and YCbCr, etc. The skin color model requires selecting a suitable color space first before locating candidate face regions. Therefore, in order to improve performance of skin detection, this method combines two types of color space that are HSV and YCbCr. Thus, in the first step the input image is transformed from RGB color space to HSV and YCbCr color space individually (as shown in Figure 3) because the variance in illumination environments cannot be described by using RGB. Equation (1) shows the conversion from RGB color space to HSV color space, equations (2) and (3) show the conversion from RGB color space to YCbCr color space.

$$\begin{cases} H = arccos\dfrac{\frac{1}{2}\big((R-G)+(R-B)\big)}{\sqrt{\big((R-G)^2+(R-B)(G-B)\big)}} \\ S = 1 - 3\dfrac{min(R,G,B)}{R+G+B} \\ V = \dfrac{1}{3}(R+G+B) \end{cases} \qquad (1)$$

$$\begin{aligned} Y' &= 16 + (65.481R' + 128.553G' + 24.966B') \\ C_B &= 128 + (-37.797R' - 74.203G' + 112.0B') \\ C_R &= 128 + (112.0R' - 93.786G' - 18.214B') \end{aligned} \qquad (2)$$

Or

$$(Y', C_B, C_{BR} = (16,128,128) + (219Y, 224P_B, 224)P_R \qquad (3)$$

(a) (b) (c)

Figure 3. The result of the color space conversions: (a) RGB color image (b) YCbCr image (c) HSV image

Detection of face region

To extract skin color, the input color image (see Figure 4) is segmented based on a combination of two types of color models: HSV color space and subcolor space. YCbCr, to increase the accuracy of classifying an image region as skin. . or non-skin. Therefore, the intensity of HSV color space was calculated for each face in an input image. And then calculated the Cb and Cr color space intensity for each face in an input image to get the average value for each face in the image to obtain an adaptive threshold value for that feature. A threshold value must be applied to decide whether it is skin-related or not. Next, the image was segmented based on HSV, Cb, Cr components to detect facial skin regions and used thresholds for each pixel to estimate face area. All pixels are classified as face region if Cr_{new}, Cb_{new}, HSV_{new}values fall within the following ranges:

$$\begin{aligned} Cr - 1 < Cr_{new} < Cr + 1 \\ Cb - 1 < Cb_{new} < Cb + 1 \\ HSV - 1 < HSV_{new} < HSV + 1 \end{aligned} \qquad (4)$$

(a) (b)

Figure 4. The result of face region detection: (a) Original image (b) Segmented skin regions image

Extract face model
This step uses morphological reconstruction from morphological mathematics and robust image processing transformation to extract the shape of the facial model as a human face. Morphological reconstruction has two unique processing properties; they are a structuring element and two images, which are a marker image and a mask image, instead of a structuring element and an image. The structuring element is commonly used to specify connectivity. Also, the marker image includes start points that can be used for transformation and other images, but the transformation is limited by the mask. However, the determination of the marker and the structuring element plays an important role in the morphological reconstruction. Therefore, this method uses the open morphological reconstruction operation as a creator image with an 8*7 pixel structuring element to reconstruct the shape of detected faces, which can almost describe the shape of human faces. After extracting candidate face shape information, the skin region is classified as a facial region if:

- The height of the skin area "h" is greater than the width of the skin area "w".
- The "h/w" ratio is less than 2. In other cases, the region is classified as a faceless zone.

Reduce noises
After finding the area that contains the face pixels, some pixels may have noise and be erroneously determined to be the face. Therefore, these pixels must be removed by discarding the connected region, which is smaller than a typical face region. The proposed method used a simple binary morphology erosion operation to reduce the noise by filling the voids or gaps as shown in Figure 5.

(a) (b) (c)

Figure 5. The result of noise reduction: (a) Remove a small-connected image (b) Filled-hole image (c) Color image after hole-filling

Face detection
After implementing the above steps, people's faces are detected with a bounding box drawn around each face in an image, as shown in Figure 6.

(a) (b)

Figure 6. The result of face detection based on Skin color: (a) Original image (b) Final-result for face detected.

Count people

In the last step, the number of people are counted based on the number of faces detected in an input image, as shown in Figure 7.

(a) (b)

Figure 7. The final-result of people counting based on Skin Color Face Detection: (a) Original image (b) Final-result for counting people.

RESULTS

This part discusses and presents the results of experiments counting the number of people in an image based on the skin color face detection method; several results have obtained which are presented in the following sections. The performance of the experimental results has evaluated using the following parameters:

- True Positive (TP): Represents the number of correctly detected faces.
- False Negative (FN): Represents the number of lost faces.
- False Positive (FP): Represents the number of non-facial features detected.
- Total Heads (P): Represents the sum of true positives and false negatives.
- Correct Detection Rate (CDR) = Recall: represents the true positive divided by the total number of faces.
- False detection rate (FPR): Represents the false positive divided by the total number of faces.
- Missing Data Rate (MR): Represents the number of false negatives divided by the total number of faces.
- Precision: Represents the True Positive divided by the sum of the True Positive and the False Negative.
- Measure F: Represents ((Precision * Recall) / (Precision + Recall))* 2.

DISCUSSION

Count people

To count the number of people using the skin color face detection experiment, 133 color images have used as the test dataset, which were selected from the Groups of Images of People dataset. There were 2,573 people in the images in the dataset. When the Skin color Face Detection method was used, the people counting system correctly detected 709 people with a recall value of 79%. The system failed to detect 185 people and had 124 error detections. Table 1 summarizes the results of this method.

Table 1. The final result of people counting system based on Skin Color Detection method

P	894
TP	709
FN	184
FP	124
CDR	79%
FPR	13.8%
MR	20.5%
Precision	85.1%
F-Measure	82%

CONCLUSION

This paper introduced a system for estimating the number of people "how many faces of people are there and where they are in a digital still image" which is based on the skin color face detection method. The face detection method was tested by skin color using the people image cluster dataset. There were 2,573 people in the images in the dataset. According to the results, the conclusions have obtained are as follows:

- People counting based on skin color face detection method achieves 79% recovery.
- People counting based on skin color face detection method provides accuracy values of 85.1%.
- People counting based on the skin color face detection method results in an F-measurement of 82%.

Based on these obtained results, it shown that the proposed method works at more than 10 fps and its accuracy is greater than 80%.

Disclaimer

The article has not been previously presented or published, and is not part of a thesis project.

Conflict of Interest

There are no financial, personal, or professional conflicts of interest to declare.

REFERENCES

1. Cope A, Doxford D, Probert C. The approaches of land and recreation resource management organisations to visitor monitoring. Land Use Policy Monitoring visitors to UK country side resources.2000;17(1):59-66.
2. Wren CR, Azarbayejani A, Darrell T, Pentland AP. Pfinder: Real-time tracking of the human body. Pattern Analysis and Machine Intelligence, IEEE Transactions on.1997; 19(7): 780-785.
3. Seriani S, Fernandez R. Pedestrian traffic management of boarding and alighting in metro stations. Transportation research part C: emerging technologies. 2015; 53: 76-92.
4. Krajzewicz D, Erdmann J, Härri J, Spyropoulos T. Including Pedestrian and Bicycle Traffic in the Traffic Simulation SUMO. In 10th ITS European Congress. Helsinki. 2014 (p. 10).
5. Hashimoto K, Morinaka K, Yoshiike N, Kawaguchi C, Matsueda S., June). People count system using multi-sensing application. In Solid State Sensors and Actuators, 1997. TRANSDUCERS'97 Chicago., 1997 International Conference on 1997; (2):1291-1294

6. Son BR, Shin SC, Kim JG, Her YS. Implementation of the real-time people counting system using wireless sensor networks.International Journal of Multimedia and Ubiquitous Engineering.2007; 2(2): 63-79.
7. Noone DR, Adam SB, Robert KL. Method and system for people counting using passive infrared detectors U.S. Patent No. 9,183,686. Washington, DC: U.S. Patent and Trademark Office.2015.
8. Kim JW, Choi KS, Choi BD, Ko SJ. Real-time visionbased people counting system for the security door. In International Technical Conference on Circuits/Systems Computers and Communications.2002; 1416- 1419.
9. Septian H, Tao J, Tan YP. People counting by video segmentation and tracking. In: 2006 9th International Conference on Control, Automation, Robotics and Vision. IEEE. 2006; 1-4.
10. Chen L, Tao J, Tan YP, Chan KL. People counting using iterative mean-shift fitting with symmetry measure. In: 2007 6th International Conference on Information, Communications & Signal Processing. IEEE. 2007; 1-4.
11. Hu Y, Chang H, Nian F, Wang Y, Li T. Dense crowd counting from still images with convolutional neural networks. Journal of Visual Communication and Image Representation. 2016; 38: 530-539.
12. Zhao T, Nevatia, R., Ramakant. Bayesian human segmentation in crowded situations. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings. IEEE. 2003; II-459.
13. Zhao X, Delleandrea E, Chen L. A people counting system based on face detection and tracking in a video. In: 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance. IEEE. 2009; 67-72.
14. Lempitsky V, Zisserman A. Learning to count objects in images. In Advances in Neural Information Processing Systems. 2010; 1324-1332.
15. Liu X, Tu PH., Rittscher J, Perera A, Krahnstoever N. Detecting and counting people in surveillance applications. In: IEEE Conference on Advanced Video and Signal Based Surveillance, IEEE, 2005; 306-311.
16. Bansal A, Venkatesh KS. People counting in high density crowds from still images. arXiv preprint arXiv:1507.08445, 2015.
17. Subburaman VB, Descamps A, Carincotte C. Counting people in the crowd using a generic head detector. In: 2012 IEEE ninth international conference on advanced video and signal-based surveillance. IEEE. 2012;470-475.
18. Rahman MS, Islam MR. Counting objects in an image by marker controlled watershed segmentation and thresholding. In: 2013 3rd IEEE International Advance Computing Conference (IACC). IEEE. 2013; 1251-1256.
19. Topkaya, I. S., Erdogan, H., & Porikli, F. (2014, August). Counting people by clustering person detector outputs. In: 2014 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE. 2014;313-318.
20. PETS2009.Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. [updated Oct 2011; cited 2015 November]. Available from: ftp://ftp.pets.rdg.ac.uk/pub/PETS2009/.
21. Mahadevan V, Li W, Bhalodia V, Vasconcelos N. Anomaly detection in crowded scenes. In: 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE. 2010; 1975-1981.
22. Blunsden S, Fisher RB. The BEHAVE video dataset: ground truthed video for multi-person behavior classification. Annals of the BMVA. 2010; 4: 4.1-12.
23. Wang Y, Lian H, Chen P, Lu Z. (2014, August). Counting people with support vector regression. In: 2014 10th International Conference on Natural Computation (ICNC). IEEE. 2014;139-143.
24. Chan AB., Liang ZSJ, Vasconcelos N. (2008, June). Privacy preserving crowd monitoring: Counting people without people models or tracking. In: 2008 IEEE conference on computer vision and pattern recognition. IEEE. 2008;1-7.
25. Dan BK., Kim YS, Jung JY, Ko S J. Robust people counting system based on sensor fusion. Consumer Electronics, IEEE Transactions on. 2012; 58(3): 1013-1021.
26. Teixeira T, Savvides A. Lightweight people counting and localizing for easily deployable indoors wsns. Selected Topics in Signal Processing, IEEE Journal of. 2008; 2(4): 493-502.
27. Jebara TS, Pentland A. Parametrized structure from motion for 3d adaptive feedback tracking of faces. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE. 1997;144-150.
28. Jebara T, Russell K, Pentland A. Mixtures of eigenfeatures for real-time structure from texture. In: Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271). IEEE. 1998;128-135.
29. Satoh SI, Nakamura Y, Kanade T. Name-it: Naming and detecting faces in news videos. IEEE MultiMedia, 1999;(1): 22-35.
30. Saxe D, Foulds R. Toward robust skin identification in video images. In: Proceedings of the Second International Conference on Automatic Face and Gesture Recognition. IEEE. 1996; 379-384.
31. Kjeldsen R., Kender J. Finding skin in color images. In: Proceedings of the Second International Conference on Automatic Face and Gesture Recognition. IEEE. 1996; 312-317.
32. Sobottka K, Pitas I. Face localization and facial feature extraction based on shape and color information. In: Proceedings of 3rd IEEE international conference on image processing. IEEE. 1996; 483-486.
33. Sobottka K, Pitas I. Segmentation and tracking of faces in color images. In: Proceedings of the Second International Conference on Automatic Face and Gesture Recognition. IEEE. 1996;236-241.

34. Saber E, Tekalp AM.. Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions. Pattern Recognition Letters.1998; 19(8): 669-680.
35. Cai J, Goshtasby A. Detecting human faces in color images.Image and Vision Computing.1999; 18(1): 63-75.
36. Wang H, Chang SF. A highly efficient system for automatic face region detection in MPEG video. Circuits and Systems for Video Technology, IEEE Transactions on.1997; 7(4): 615-628.
37. Chai D, Ngan KN. Locating facial region of a head-and-shoulders color image. In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. IEEE. 1998; 124-129.
38. Chen Q, Wu H, Yachida M.. Face detection by fuzzy pattern matching. In: Proceedings of IEEE International Conference on Computer Vision. IEEE. 1995;.591-596.
39. Dai Y, Nakano Y, Miyao H. (1994, October). Extraction of facial images from a complex background using SGLD matrices. In: Proceedings of 12th International Conference on Pattern Recognition. IEEE. 1994; 137-141.
40. Dai Y, Nakano Y. Face-texture model based on SGLD and its application in face detection in a color scene. Pattern recognition. 1996; 29.6: 1007-1017.
41. Miyake Y, Saitoh H., Yaguchi H, Tsukada N. Facial Pattern Detection and Color Correction of Television Pictures for Newspaper Printing. J. Imaging Technology. 1990; 16: 165-169.
42. Crowley JL, Bedrune JM., Bekker M, Schneider M. (1994). Integration and control of reactive visual processes. In: European Conference on Computer Vision. Springer, Berlin, Heidelberg. 1994; 47-58.
43. Starner T, Pentland A. Real-time asl recognition from video using hmm's. Media Lab, Massachusetts Institute of Technology, Tech. Rep, 1996, 375.
44. Yang J, Waibel A. A real-time face tracker. In: Proceedings Third IEEE Workshop on Applications of Computer Vision. WACV'96. IEEE. 1996; 142-147.
45. Crowley JL, Berard F. Multi-modal tracking of faces for video communications. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE. 1997; 640-645.
46. Oliver N, Pentland A P, Berard F. Lips and face real time tracker. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE. 1997; 123-129.
47. Yang J, Stiefelhagen R, Meier U, Waibel A. Visual tracking for multimodal human computer interaction. In: Proceedings of the SIGCHI conference on Human factors in computing systems. 1998; 140-147.
48. Graf HP, Cosatto E, Gibbon D, Kocheisen M, Petajan E. Multi-modal system for locating heads and faces. In: Proceedings of the Second International Conference on Automatic Face and Gesture Recognition. IEEE.1996; 88-93.
49. Azad R, Shayegh HR. Novel and tuneable method for skin detection based on hybrid color space and color statistical features. arXiv preprint arXiv. 2014;1407.6506.