# Face Detection for Surveillance Systems Application

Tanko Danial Salka, Marsyita Binti Hanafi, Syamsiah Mashohor, and Sharifah Mumtazah Syed Ahmad

***Abstract*—** Face detection in digital image has been widely explored over the past few decades. Despite the significant progress, detecting human faces in unconstrained and complex images remains a challenging problem in computer vision especially for the images captured by surveillance cameras. This paper proposed a face detection method, which is robust enough to handle various imaging conditions such as illumination, size, pose, and orientation. To achieve this goal, our method incorporates skin color and shape to identify the regions which are most likely to contain a face. The system applied two different types of filters, namely RGB and Dynamic Chrominance filters for the skin detection and the resulting image is merged with an inverted Canny edge image to aid the separation of the skin and non-skin regions. A morphological operation is applied to remove noises. Finally, Euler's number that works based on the geometrical properties of human faces is applied to locate faces. Experimented on a Large Variability Surveillance Camera face database that consists of 1200 image, the proposed face detection system produced 98.4% of true detection rate.

***Keywords*—** Face detection, RGB filter and Dynamic Chrominance filter

## I. INTRODUCTION

IN the past few years, video surveillance systems have received great attention and are widely used due to the need for safety in the society [1]. Many cities around the world have installed Close Circuit Television (CCTV) cameras in public places such as banks, airports, schools, market places, and etc. The use of the CCTV system has contributed to the reduction of crimes. However, detecting human faces from the images captured by a CCTV cameras is a challenging task due to the low quality of images, which are mostly low in resolution, variation in illumination and occlusion. In recent years, many techniques have been developed to tackle these problems. In [2], the authors proposed a face detection system which is based on skin color. They performed image enhancement to the RGB image followed by skin region segmentation using RGB and YCbCr color spaces and edge detection to classify the skin and non-skin region. The detected skin region was identified as face or non-face using the geometrical properties of human face, such as area, bounding box proportions, centroid and extent. The proposed algorithm was able to detect faces in different illumination conditions, sizes and expressions. However, the algorithm produced high number of false detection for the images with a complex background and the algorithm worked well for the frontal upright faces only.

In [3], the authors presented a simple and an effective method for detecting human faces. This method used Cb and Cr chrominance components to detect the skin color. The reason of using the chrominance component is to remove the effect of illumination. Euler's number was used to calculate the number of holes in the skin region that determined the location of the face region. Experimented on images with single and multiple persons, which were captured at different background, the system achieved 96% and 92.67% accuracy, respectively.

In [4], the authors applied YCbCr color space to model human skin and Gray level co-occurrence matrix (GLCM) to extract features that best represent skin. Then, Tamura texture was used to remove the non-skin region recognized by GLCM. The proposed method produces high detection rate, which is more than 99%. However, the authors failed to mention the total number of images used in the experiment and the proposed algorithm did not consider other ethnic groups.

The authors in [5] and [6] proposed an algorithm that combined HSI and YCbCr color spaces alongside with Prewitt and Canny edge detection techniques. Other methods involved in their work are connected component analysis, morphological operation, region labeling, Euler's test and template matching. The authors showed that this method outperformed the methods that worked based on the color of the skin only. Nevertheless, the system produced high percentage of false detection rate.

In this paper, we present a face detection algorithm for the images captured by CCTV systems, which is able to tackle the problems of detecting faces due to the variation in illumination, ethnicity and occlusion. The algorithm consists of a combination of the skin and edge detection methods, and Euler number method. The rest of the paper is organized as follows: Section 2 describes the methods used to detect faces in images captured by CCTV system and the experimental results are presented in section 3. Finally, conclusions and future works are discussed in section 4

## II. METHODOLOGY

The methods used to detect faces in the images captured by CCTV system is illustrated in Fig.1. The images are under various imaging conditions such as variation in illumination,

Tanko Danial Salka, Marsyita Binti Hanafi, Syamsiah Mashohor, Sharifah Mumtazah Syed Ahmad are with Department of Computer and Communication Systems, Faculty of Engineering, UPM, Serdang, Malaysia.

size, pose, and orientation. The first stage in the algorithm is to detect the skin region and the edges, separately. The skin is detected using two types of filters, namely, RGB and Dynamic Chrominance filters. Meanwhile, the edges are detected using Canny edge detector. Then, the results from both methods are combined based on the concept of logical ADD. In addition, morphological operation is applied to remove noises. Finally, the face region is located using Euler's number that works based on the geometrical properties of human faces. Each process is discussed in detail in the following sections.

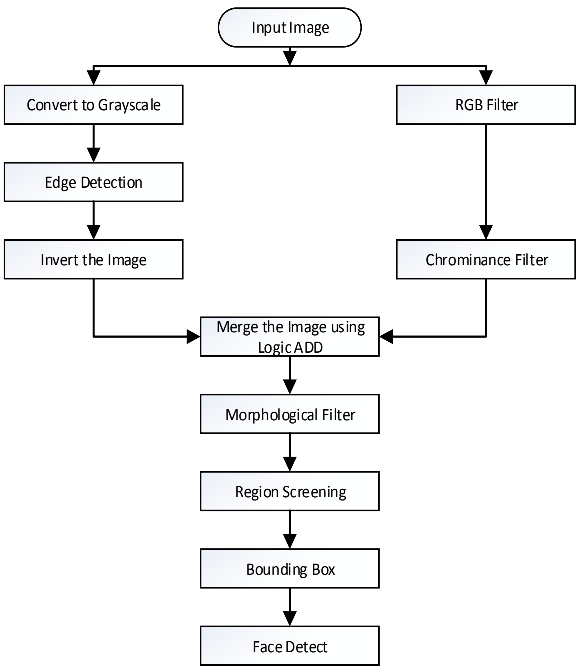


Fig. 1. Flow Chart of the Proposed Method

### A. *RGB Filter*

The RGB filter is used to eliminate most of the pixels that are non-skin pixels so that in the next stage, the classification will be executed on the pixels that are most probably belong to the skin. The filter applies a set of rules, which is based on the work published in [8], where the range of pixel values in red (R), green (G) and blue (B) channels are identified from developed histograms. The example of images used to develop the histograms are shown in Fig. 2

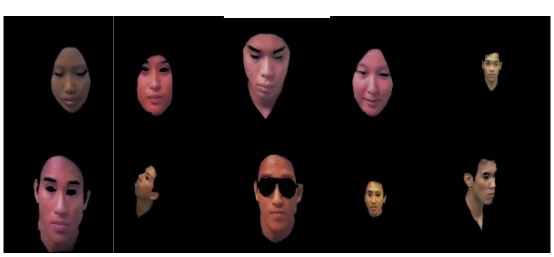

Fig 2. Examples of skin mask

Based on the generated histograms, we found that the range of the pixel values that represent skin for each channel are as follows.

$$skin = \begin{cases} 58 < R < 210 \\ 30 < G < 165 \\ 30 < B < 180 \end{cases} \quad (1)$$

However, the inter-channel relationship is not considered in equation (1). We observed that the R channel (red) has a higher value than the G (green) and B (blue) channels in the RGB color model, regardless of different ethnic groups. It was reported in [9] that the R channel has a greater influence on the skin and this channel should be adjusted so that the range of values for R is larger than B and G channels. The reason is that the skin color is mostly consists of a combination of red, yellow and brown colors, where yellow and brown can be produced using R, therefore, the value of R must be greater than G and B. In terms of inter-channel relationship, we found that the difference between R and G (R minus G), and R and B (R minus B) that best represented the skin color across different ethnic must be set larger than 15. Thus, a rule as follows is identified

$$RGB\ Filter = \begin{cases} R > B \ \cup R > G \\ R - G > 15 \ \cup R - B > 15 \end{cases} \quad (2)$$

### B. *Chrominance filter*

The filter consists of the optimal combination of color components, which are the combination of $H_1$, $H_2$, S and I components. The optimal combination that best represent the skin areas across different ethnics is searched by performing experiments on several color spaces, namely *CIEL a* b**, *CIEL UV*, *HSV*, new *HSL* and normalized *RGB* (*rgb*), which is following the work discussed in [8]. The filter is functioned to classify a pixel belongs to class "skin", according to its value, *z*, of a single component of a certain color system, as:

$$X_{(z;\,a,b)} = \begin{cases} 0 & if\ z \le a\ or\ z\ \ge b \\ 1 & if\ a < z < b \end{cases} \quad (3)$$

Parameter $a$ and $b$, where $a < b$, are calculated from the histograms that are developed using training dataset so that $a = X_{min}$ and $b = X_{max}$. In our approach, we have three sets of $X_{min}$ and $X_{max}$ for each selected component, where each set represents the range of the skin values for African, Chinese, Malay and others. The "others" is defined as the skin of respondents who are originated from the Middle East. Based on our observation, we found that the range of values for each component that represent "others" is approximately similar to the range of skin values for Malays and this is the main reason that Malays and others were grouped in one class. Table I shows the values of the $X_{min}$ and $X_{max}$ for each component that best represent African, Chinese, and, Malay and others.

TABLE I
THRESHOLD VALUES FOR THE $H_1$, $H_2$, S AND I COMPONENTS.

| Ethnic Group | Component 1 | Component 2 | Component 3 | Component 4 |
|---|---|---|---|---|
| African | $H_{1\_min} = 0$ and $H_{1\_max} = 0.14$ | $H_{2\_min}=0.9$ and $H_{2\_max}= 1$ | $S_{min}=0.05$ and $S_{max} = 0.53$ | $I_{min}= 0.01$ and $I_{max} = 0.18$ |
| Chinese | $H_{1\_min} = 0$ and $H_{1\_max} = 0.24$ | $H_{2\_min}=0.9$ and $H_{2\_max}=1$ | $S_{min} = 0.2$ and $S_{max} = 0.6$ | $I_{min}= 0.02$ and $I_{max} = 0.23$ |
| Malay and other | $H_{1\_min} = 0$ and $H_{1\_max} = 0.2$ | $H_{2\_min}=0.85$ and $H_{2\_max}= 1$ | $S_{min}=0.25$ and $S_{max} = 0.62$ | $I_{min}= 0.02$ and $I_{max} = 0.23$ |

The first set that consists of the range of values for each component that represents the skin of the African is labelled as set A, whereas for "Malay and others" and "Chinese" are labelled as set B and C, respectively. The filter works dynamically by choosing only one set at a time (A, B or C) and the selection is based on the total number of pixels detected as skin by the RGB filter and the three RGB regions, which are defined as follows.

$$0 < r_1 < 85$$
$$86 < r_2 < 170$$
$$171 < r_3 < 255$$

Set A is chosen by the chrominance filter when most of the pixels that are detected by the RGB filter as skin have the values that are in the range of the first region, $r_1$, whereas set B is chosen when most of the pixels that are detected by the RGB filter have the values, which are in the second region, $r_2$. If most of the pixels have values in the third region, $r_3$, then, set C is chosen. The process is illustrated in Fig. 3.

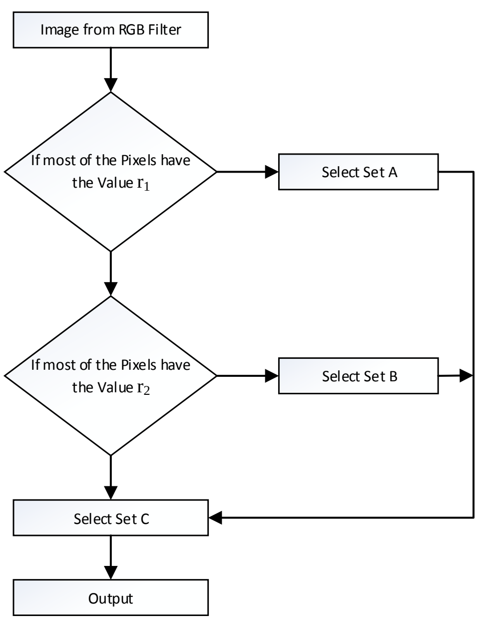


Fig. 3. Illustrate the flow chart of the process.

*C. Edge Detection*

In this work, the edges are detected using a Canny edge detector. The detector finds the edges by looking for local maxima of the gradient of image. The gradient is calculated using the derivative of a Gaussian filter. This method achieve three useful objectives namely; good localization, good detection, and good representation for locating an object contour. The Canny method applies two thresholds to the gradient: a low threshold for high edge sensitivity and a high threshold for low edge sensitivity. In the experiment, the low threshold is set to 0.05 and the high threshold is set to 0.2. These values are identified as the values that best produce the edges.

*D. Morphological Filter*

The filter consists of morphological closing and opening, which are used to reduce the number of pixels of true skin detected as non-skin and true non-skin detected as skin. The main function of the filter is to remove salt and pepper noises in the images. The mathematical representation for opening is

$$A \circ B = (A \ominus B) \oplus B \quad (4)$$

and for closing is

$$A \cdot B = (A \oplus B) \ominus B \quad (5)$$

Where $\ominus$,$\oplus$,$\circ$ and $\cdot$ denotes erosion, dilation, opening and closing, respectively.

*D. Regions Screening*

After detecting the skin color region, the next task is to locate the position of a face in an image by considering facial features such as eyes, mouth and eyebrows. The output of the morphological filter, which is the image that has been removed the noises is then converted into a binary image. The facial features in the binary image are considered as holes. As we know, a face must consists of facial features. Based on this knowledge, we know that the region in the binary image without any holes in it is considered as non-face. Hence, region without holes is rejected. The method used to reject the region without holes is Euler's number method. The Euler's number, E is calculated by deducting the number of connected component C, from the number of holes, H, as shown as follows.

$$E = C - H \quad (6)$$

In our work, the Euler's number is calculated using a built in function that is available in Matlab. The region is considered as face if the Euler's number produces a negative value and this is because, C is considered as the largest connect skin region and therefore, it must equal to 1. However, the total number of holes inside the region is limited to 6 and this is based on the knowledge that there are 6 facial features on faces, namely eyes, eyebrows, nose and mouth. Hence, the range of Euler's number used in this work is as follows.

$$0 < E < -5 \quad (7)$$

## III EXPERIMENTAL RESULTS

The experiments are conducted on 1200 images, which are obtained from a Large Variability Surveillance Camera Face Database [10]. The images in the database are captured using an indoor PTZ camera under controlled condition with a variation in illumination, poses, scales and occlusions across different ethnics (Malays, Indians, Chinese, Africans and others). In the experiments, the total number of images used in the evaluation and testing is 500 and 700 images, respectively. The results of various stages in the proposed method are revealed in Fig. 4. Fig. 4a shows the example of original images used in the experiment and Fig. 4b shows the results produced by the proposed skin detection method that classifies the skin and non-skin pixels using the RGB filter and Dynamic Chrominance filters. The detected edges are shown in Fig. 4c. Fig. 4d shows the output produces by the morphological filter and Fig. 4e shows the location of the face region that has been identified by the Euler's number method. The performance of the proposed algorithm is observed using true and false detection rates, as shown in Table II. In addition, a comparison with other methods [3, 5 and 11] that worked based on skin detection and Euler's number methods is shown in Table II as well. The result shows that the proposed method outperforms the other methods.

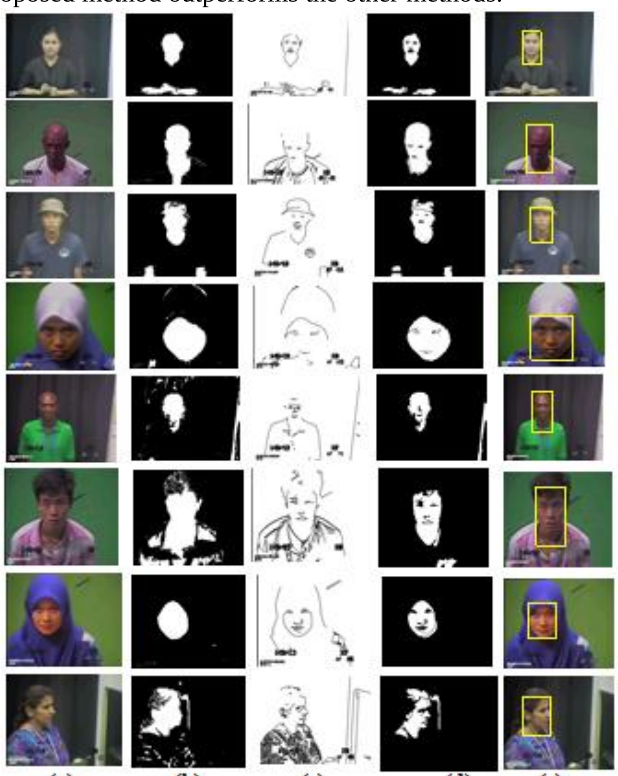


Fig. 4 Illustrate the various stages in the proposed method: (a) Original Image (b) Skin detection (c) Edge detection (d) Noise reduction (e) Final result

TABLE II
PERFORMANCE EVALUATION AND A COMPARISON WITH SEVERAL METHODS

| Method | True Detection Rate (TDR) % | False Detection Rate (FDR) % |
|---|---|---|
| Reema A. et al (2013) [5] | 84.60 | 7.69 |
| Samir El K. (2014) et al [11] | 98.02 | 1.98 |
| Pallabi Saikia et al 2012 [3] | 96 | |
| | 92.67 | |
| **Our method** | **98.40** | **1.60** |

## VI. CONCLUSIONS

In conclusion, this work proposes a method that detects human faces from images captured by surveillance camera that works based on a combination of skin color segmentation; edge detection and geometrical properties of a face. The technique for the skin detection uses two different types of filters, namely, the RGB and Dynamic Chrominance filters. The resulting skin image and the inverted canny edge detected image were compared in order to produce a finalized skin region and the noises were removed using a morphological filter. The final step is to locate a face in the images based on the knowledge that there are six facial features in each face. This step was executed using Euler's number method. The results show that the proposed method outperforms several methods, as shown in Table 2, by producing high accuracy, which is 98.4%. However, in some cases, missed detection occurs and in future, this problem will be tackled by introducing a texture-based analysis in the algorithm.

## REFERENCES

[1] M. Valera and S. A Velastin Intelligent distributed surveillance systems: a review". Vision, Image and Signal Processing, IEE Proceedings. 192-204, (2005)

[2] D. Ghimire & J. A. Lee. Robust Face Detection Method Based on Skin Color and Edges. *Jips*, *9*(1), 141–156. http://doi.org/10.3745/JIPS.2013.9.1.141, (2013)
http://dx.doi.org/10.3745/JIPS.2013.9.1.141

[3] S. Pallabi, J. Gollo and K. Margaret: Face Detection using Skin Colour Model and distance between Eyes 1,2. *International Journal of Computing, Communications and Networking, Volume 1, No.3.*, *1*(3), 102–105. (2012).

[4] N. K. El-Abbadi, A. Abdul and A. Qazzaz: Detection and Segmentation of Human Face. *International Journal of Advanced Research in Computer and Communication Engineering*, *4*(2), 90–94. http://doi.org/10.17148/IJARCCE.2015.4220, (2015).

[5] A. Reema and S. Namrata: Face Detection in Digital Images Using color Spaces and Edge Detection Techniques, International Journal of Advanced Research in Computer Science and Software Engineering. Volume 3, Issue 6, Pp. 718-725, (2013)

[6] H. C. V. Lakshmi, & S. PatilKulakarni: Segmentation Algorithm for Multiple Face Detection for Color Images with Skin Tone Regions. *Signal Acquisition and Processing, 2010. ICSAP '10. International Conference on*, 162–166. http://doi.org/10.1109/ICSAP.2010.42, (2010).

[7] R. N. Bin Abdul, K. C. Wei, and J. See: RGB-H-CbCr Skin Colour Model for Human Face Detection. *Proceedings of The MMU International Symposium on Information & Communications Technologies (M2USIC 2006)*.

[8] K. Klevan, Marsyita. B. H. and R. Abdul: Optimal Combination Of Color Components For Skin Detection. *4th International Conference on Artificial Intelligence with Applications in Engineering and Technology*, (1), 171–176. http://doi.org/10.1109/ICAIET.2014.36 (2014).

[9] Z. Zhang, M. Wang, and Z. Lu: A Skin Color Model Based on Modified GLHS Space, International Journal of information hiding and multimedia signal processing, *5*(2), 144–151, (2014).

[10] D. S. Tanko, B. H. Marsyita, B. M. Syamsiah and M. S. A. Sharifah: Large Variability Surveillance Camera Face Database, Seventh International Conference on Computational Intelligence, Modelling and simulation, 86–90. http://doi.org/10.1109/CIMSim.2015.21, (2015).

[11] El K. Samir, S. Abderrahim and A. Mustapha: A New Robust Face Detection Method Based on Corner Points, International Journal of Software Engineering and Its Applications.Vol. 8, No. 11, pp. 25-40, 2014