

Comparative Analysis Between YCbCr Color Space and Viola-Jones Algorithm Based Method for Face Detection

¹Deepali G. Ganakwar, ²Dr. Vipulsangram K. Kadam,

¹Research Scholar, ²Professor,

¹Department of Science & Technology,

¹Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India.

Abstract : This paper provides a short insight of face detection using YCbCr and Viola jones based algorithm. Face detection is one of the most active research area in computer vision literature because of challenging nature of face as an object with countless applications. The goal of this paper is to present comparative analysis of human face detection systems. Face detection is a difficult task in image analysis which has each day more and more applications. We can define the face detection problem as a computer vision task which consists in detecting one or several human faces in an image. It is one of the first and the most important steps of Face analysis. In this paper we presented two methods of face detection. As the number of proposed techniques increases, survey and evaluation becomes important.

IndexTerms - Face detection, color space, integral image, adaboost algorithm, cascade.

I. INTRODUCTION

Face detection is becoming an active research area spanning several disciplines Such as image processing, pattern recognition, computer vision, neural networks, cognitive science, neuroscience, psychology and physiology. It is a dedicated process, not merely an application of the general object recognition process [4][5]. It is also the representation of the most splendid capacities of human vision.

Automatic face detection is the cornerstone of all applications revolving around automatic facial image analysis including face recognition and verification, face tracking for surveillance, facial behavior analysis, facial attribute recognition, gender/age recognition [1] [12] etc.

The goal of face detection is to determine whether or not there are any faces in the image and if the image is present then it return the image location and extent of each face. While this appears as a trivial task for human beings, it is an extremely tough task for computers, and has been one of the top studied research topics in the past few decades.

It is basically an image segmentation problem as the image is to be segmented into two parts: one containing faces and the other representing non-face regions. Face detection takes images/video sequences as input and locates face areas within these images. This is done by separating face areas from non-face background regions. Facial feature extraction locates important feature (eyes, mouth, nose and eye-brows) positions within a detected face [12].

In General face detection system input image is passed to the system for pre-processing. Image may vary in format, size and resolution and can include frames of video. In the next step pre-processing is done, which normalized the image and also remove noise. The classifier decides face and non-face class based on information learned from during training. Finally, the output locates face region.

II. METHODS OF FACE DETECTION

3.1 Face Detection using Viola Jones Algorithm

AdaBoost algorithm was firstly proposed by Freund and Schapire in 1995. In 2001, Viola and Jones applied this algorithm to detect faces and successfully constructed a detection system which can detect human face in real time [6] [7]. They use the integral image to solve the speed problem and uses cascade classifier to achieve high detection rates. The details of this algorithm are described as the following subsections.

The Viola-Jones object detection framework [6] is an object detection framework which provide robust and competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones. Even though it can be trained to detect a variety of object classes, it was motivated mainly by the task of face detection. This face detection framework is capable of processing images extremely rapidly and achieving high detection rates. There are three main stages of face detection framework.

3.1.1 Haar like Features:

Haar-like feature are rectangular digital image features. A rectangular Haar-like feature is defined as the difference of the sum of pixels in white rectangles and the sum of pixels in black rectangles [6] [8]. The typical rectangular features are as shown in Fig. 1:



Fig.1. Haar-like features.

Viola and Jones use integral image to calculate the value of rectangular features. The value of coordinate (x, y) in an integral image is the sum of all pixels which located on the up and left region of the original image at coordinate (x, y).

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

Where $ii(x, y)$ and $i(x', y')$ respectively are the value of integral image at coordinate (x, y) and the pixel value of original image at coordinate (x, y) . Due to the use of the integral image, the value of Haar-like feature can be calculated very fast only by using addition and subtraction operation [9].

3.1.2 Adaboost Algorithm:

There are a lot of Haar-like features in an image and the AdaBoost algorithm is used to select the best of them as weak classifiers. In this stage classifiers are constructed by selecting a small number of important features (rectangle features) using Adaboost algorithm. From vast number of features computed in stage 1 we are interested in only selected few features that would enable us to detect face with great accuracy. For this, we use Adaboost Algorithm [9] to select principal features and to train classifiers that would be using them. Aim of this algorithm is to create strong classifier from linear combination of weak classifier. AdaBoost provides an effective learning algorithm.

3.1.3 Cascading:

The third major stage of this method is a combining successively more complex classifiers in a cascade structure [8] which dramatically increases the speed of the detector by focusing attention on promising face like regions of the image. This cascade structure consists of classifiers. It works in a manner that initial classifiers are simpler and they are used to reject majority of sub-windows and at end complex classifiers are used to achieve low false positive rates. After skin color segmentation, the image is sent to the cascaded classifier to detect faces. If the first strong classifier judges the image as no face, then terminate the detection process. Otherwise the image will be sent to the next strong classifier for further verifying.

3.2 YCbCr Color space based Face detection method

For Detection of skin color in color images is a very popular and useful technique for face detection. Color is an important feature of human faces. Using skin-color as a feature for tracking a face has several advantages [13]. Color processing is much faster than processing other facial features. In the skin color detection process, each pixel was classified as skin or non-skin based on its color components.

This color space is a variant of the YCbCr model proposed by de Dios and Garcia [14]. It differs from the YCbCr model in the usage of C_g (G-Y) component instead of the C_b (B-Y) component for the second channel. By using C_g component instead of C_b this color space is reported to have better performance in skin detection than the YCbCr model [15]. Transformation from RGB to this color space is by using following calculations:

$$Y = 16 + 65.481R + 128.533G + 24.966B$$

$$C_g = 128 - 81.085R - 112G + 30.915B$$

$$C_r = 128 + 112R - 93.786G - 18.214B$$

where R, G, and B components ranges from 0 to 1.

3.3.1 Skin Detection:

The color space is converted from RGB to YCbCr so that we can segment the image into skin-like regions and non-skin regions in this space. Same skin detection technique is applied to both methods. Thresholding segmentation is the simplest method, which has better computing performance than the Gaussian skin color model. We regard a pixel as skin if it satisfies the following conditions:

Table 3.1: Threshold Values

Sr No.	YCbCr
01.	$Y > 80$
02.	$77 < C_b < 127$
0.3	$135 < C_r < 175$

The result of thresholding segmentation is a binary image where white represent the skin-like regions and black represent the non-skin regions. Then we use the median filter to remove noise. The median filter can remove noise with less attenuation of edges [6] [11]. Fig. 3 is the results of skin color segmentation.

3.2.2 Gaussian Model

Variables Skin model is considered as mathematical model that indicates skin color as an image of human face. Each of the faces demonstrates that, the image is converted from RGB space to YCbCr space. The YCbCr color space displayed $X(CbCr)_T \in R^2$ the destiny function is,

$$f(CbCr) = \frac{1}{2\pi|C|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(X - \mu)^T C^{-1}(X - \mu)\right\}$$

$$X = \begin{pmatrix} C_b \\ C_r \end{pmatrix}, \mu = \begin{pmatrix} \mu_{Cb} \\ \mu_{Cr} \end{pmatrix}, C = \begin{pmatrix} C_{CbCb} & C_{CbCr} \\ C_{CrCb} & C_{CrCr} \end{pmatrix}$$

The effectiveness of color space coordinate is more as compared to empirical threshold which identifies the coordinates. A binary image is produced after pixel based classification [10]. Morphological smoothing operations help to extract the final skin color areas [14]. It is very much effective to utilize



Fig. 2. The skin-color segmentation:(a) Original face image (b) Image in YCbCr space(c) Similarity image (d) Binary image morphological operations

morphological operations and earlier information for face detection. The morphological activity can disentangle image information while safeguarding their basic shapes characteristics and is able to eliminate any irrelevancies (Haralick and Shapiro, 1993). Thus it can determine an exact shape of the skin segment [10][13]. Face region by applying morphological dilation task with a 3-by-3 organizing components for several times (5 to 10 here) trailed by a similar number of disintegration activities utilizing the equivalent structuring element. Inside the segmented face region, there exist holes that relate to nose, eyes and mouth etc. with the help of dilation operation these holes are filled. Later, the results obtained from dilation operation are treated with erosion operation so as to restore the shape of the face [10]. The regions obtained from dilation and erosion operation is depicted in fig.5.

III. RESEARCH METHODOLOGY

4.1 Data and Sources of Data

we have used "Minear, M. & Park, D.C. (2004). A lifespan database of adult facial stimuli. Behavior Research Methods, Instruments, & Computers, 36, 630-633" database. And some other database. Minear, M. & Park, D.C database consist of more than 300 individual faces ranging from ages 18 to 93. It has faces of 218 adults age 18–29, 76 adults age 30–49, 123 adults age 50–69, and 158 adults age 70 and older [16].



Fig.3. Some examples of Database

4.2 Theoretical framework

Two methods are compared on the basis of false detection and precision. False detection can be calculated by taking ratio of total number of falsely detected faces to the total number of faces in given image (ideally it should be 0). Precision is also a ratio of total number of falsely detected faces to the total number of faces present in given image (Ideally it should be 1).

- False detection= Total Number of falsely detected faces/ Total number of faces
- Precision= total Number of correctly detected faces/Total number of faces

Following are some experimental results.



Fig.4. Some outputs of Viola-Jones Algorithm Based Face detection Method



Fig.5. Some outputs of YCbCr Color space based face detection method

IV. RESULTS AND DISCUSSION

4.1 Results of face detection for different type of database

Method	Type of Image	Total No. of Faces	Correctly Detected Faces	Falsely Detected Faces	Precision	False Detection
Viola-Jones algorithm based face detection	Front view face	12	12	0	1.0	0
	Side View Face	12	00	0	0.0	0
	Multiple Face image	54	48	11	0.8888	0.2037
YCbCr Color space based face detection method	Front view face	12	11	1	0.9166	0.0833
	Side View Face	12	10	2	0.8333	0.166
	Multiple Face image	54	46	13	0.8518	0.2407

Table 4.1 displayed correctly detected faces, falsely detected faces, false detection, and precision for both face detection methods. In this analysis, three types of databases are used like image with front view face, side view face and multiple faces. From table 4.1 it is clearly indicating that the type of face database highly affects the precision and false detection of the face detection method. Thus we can conclude that viola-jones algorithm based face detection work better with the image with a front view but absolutely fails for the image with side view.

The descriptive statistics indicated that false detection rate and precision of YCbCr color space is better for all type of database.

REFERENCES

[1] G M.H. Yang, D. J. Kriegman, N. Ahuja, "Detecting faces in images: A survey", IEEE Transaction on Pattern Analysis and Machine Vol.24 (1), p.p.34-58, 2002.

[2] S. Chitra, G. Balakrishnan , "Comparative Study for Two Color Spaces HSCbCr and YCbCr in Skin Color Detection", Applied Mathematical Sciences, Vol.6, no.85, pp.4229 - 4238, 2012.

[3] Devendra Singh Raghuvanshi, Dheeraj Agrawal, "Human Face Detection by using Skin Color Segmentation, Face Features and Regions Properties", International Journal of Computer Applications, Vol. 38, No.9, pp.14-17,2012.

[4] Smita Tripathi ,Varsha Sharma and Sanjeev Sharma, "Face Detection using Combined Skin Color Detector and Template Matching Method", International Journal of Computer Applications,Vol.26,No.7,pp.5-8,2011.

[5] Govindaraju, V., "Locating human faces in photographs", International journal in Computer Vision Vol.19 (2), pp.129-146, 1996.

[6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR , Vol.1, pp.I-511 - I-518, 2001.

[7] Binesh T.R and P. Simon, "Fast Pedestrian Detection using smart ROI separation and Integral image based Feature Extraction", International Journal on Computer Science and Engineering (IJCSSE), Vol.4, No.11, pp.1771-1779, 2012.

- [8] R. E. Schapire, Y. Freund, P. Bartlett, and W. S. Lee, "Boosting the margin: a new explanation for the effectiveness of voting methods", *Annals of Statistics*, Vol. 26, No.5, pp.1651-1686, 1998.
- [9] Zeng Hao, Qin Feng, Lin Kaidong, "An Optimized Face Detection Based on Adaboost Algorithm", 2018 International Conference on Information Systems and Computer Aided Education (ICISCAE), pp 375-378, 2018.
- [10] C. Emmy Premal, S. S. Vinsley, "Image Processing Based Forest Fire Detection using YCbCr Colour Model", 2014 International Conference on Circuit, Power and Computing Technologies [ICCPCT], pp 1229-1237, 2014.
- [11] Sukmawati N Endah, Helmie A Wibawa, Retno Kusumaningrum, "Skin Detection Based on Local Representation of YCbCr Color Moment", 2017 1st International Conference on Informatics and Computational Sciences (ICICoS), pp 65-70, 2017.
- [12] Hesham A. ALABBASI, Florica MOLDOVEANU, "Human face detection from images, based on skin color", Proceedings of the 18th International Conference on System Theory, Control and Computing, Sinaia, Romania, pp.532-537, October-2014.
- [13] Minghua ZHAO, Yonggang ZHAO, "Skin color segmentation based on improved 2D Otsu and YCgCr", International Conference on Electrical and Control Engineering, 2010.
- [14] WAN Li, "Face Detection Method Based on Skin Color and AdaBoost Algorithm", International Conference on Computational and Information Sciences, 2010.
- [15] Bashir Muhammad 1, Syed Abd Rahman Abu-Bakar, "A Hybrid Skin Color Detection using HSV and YCgCr Color Space for Face Detection", IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 2015.
- [16] Minear, M. & Park, D.C. (2004) "A lifespan database of adult facial stimuli" *Behavior Research Methods, Instruments, & Computers*, 36, 630-633.

