

# Web Image based Filtering – Human Composition Matrix based on Skin Detection

Shubhangi Adate<sup>1</sup> Shankar M Patil<sup>2</sup>

<sup>1</sup>ME Student <sup>2</sup>Associate Professor

<sup>1,2</sup>Department of Computer Engineering

<sup>1,2</sup>BVCOE, Kharghar, Navi Mumbai India

**Abstract**— In this paper, we describe various skin detection methods (HCM), image filtering methods and comprehensive comparative study among these methods proposed to detect adult classified images. It is based on concept of pattern recognition techniques. First the images are changed to identify areas with low color intensity by the color model. In the next part of the proposed system specify that the image is filtered using skin detection. The main aim has to segment a person or people within the image. By counting all pixels (red and green channel) with identical skin tone, then we can treat the image with porn content. The development of Internet make changes in dramatically falling costs of data storage and improvement in coding technology are generating wide variety of wall papers, images, animation, graphics, sound and video. Now a day's it is obvious that the computers with internet connection. There are facilities in large amount of adult images for free to access. This kind of media is available for children and making problem for many parents. The searching principal of Internet browser programs to avoid adult classified content should be filtering of images. There are ways available to stop adult classified content images on computers. This process is carried out by blocking unwanted sites or identifying images that show explicit content.

**Key words:** RGB, Human Skin Filter, Image Retrieval, Color Model

## I. INTRODUCTION

Images are part of World Wide Web in this visual world. The statistics of more than 10 million HTML webpage reveal that 80% of web pages contain images and that on average there are about 30.2 images per HTML web page. These images are used to make effective and eye catching Web contents. However, images are also contributing to harmful (e.g. pornographic) or even illegal Internet content. So effective filtering of images plays vital role in a web image filtering solution. To block adult content, some representative companies as Net Nanny and Surf Watch, operate by maintaining lists of URL s and newsgroups and require constant manual updating. Now a day's tons of literature is available in the web. Detection based on image content analysis has the advantage to process equally all the images without the need for updating, so will produce more effective filtering. The fact is that there is a relation between images with maximum pixels of skin and images of adult content. So we need to develop a skin detector. Skin color offers an effective and efficient way to detect the adult image content. There are so many people already worked in this area. The WIPE system developed by Wang, Li, Wiederhold and Firschein uses a manually -specified color histogram model as a prefilter in an analysis pipeline. The images with less percentage of skin pixels are treated as non-offensive. Images that contain considerable skin pass on to a final stage of

analysis where they are classified using wavelet features. Forsyth s research group has designed and implemented an algorithm to screen images of naked people. Their algorithms involve a skin filtering method and human features matrix. The skin color model used by Fleck, Forsyth and Bregler consists of a manually specified region in a log-opponent color space. Detected regions of skin pixels from the input image to a geometric filter based on skeletal structure. Skin detection describes recognition of human skin pixels from an image. It plays an important role in various functionalities such as face detection, searching the images and filtering the content on the web. Skin color tone discrimination approaches can be defined in two basic types. They are physical-based approaches and statistical approaches. Further, Statistical approaches can be divided into two models, parametric approaches and non-parametric approaches. Parametric model approaches uses the skin color distribution in parametric form, such as Gaussian or Gaussian mixture. In nonparametric model approaches histograms, used to represent density in color space. Jones and Rehg proposed model for skin color detection by estimating the intensity of skin and non-skin color pixels in the color space using labeled training data. Both parametric and nonparametric statistical models usually perform color segmentation in color spaces that reduce the varying illuminant. The most common color spaces have been used are, normalized RGB and HSV.

## II. RELATED WORK

### A. RGB Color Model

The RGB color model is a prominent color model in which the primary colors are red, green and blue light are added together in various combinations to reproduce a broad range of colors. The name comes from the three basic colors Red, Green, and Blue. The RGB color stack is illustrated in the following Figure 1. RGB Color Model The main objective of the RGB color stack is used for various functionalities like recognizing, visualizing, sensing and display of images in electronic gadgets like televisions, PCs and high end mobiles.

The RGB color model is an exemplar in the sense that three light beams are combined together to get a final color. The formation of colors in RGB is carried out by three colored light beams (one red, one green, and one blue) should be superimposed.

A beam of each basic color is called as a component of that color. Where each can have arbitrary intensity either from fully off or fully on is in that mixture. Zero intensity for each component results as darkest color (no light in turn treated as black) and full intensity of each gives a white. A color in the RGB color model is described by indicating mixture of red, green, and blue. It is included in each component which can vary from zero to a defined maximum

value. This depends on the functionality of respective color. The component values are often stored as numerical values in the range 0 to 255.

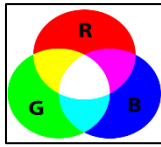


Fig. 1: RGB Color Model

### B. HSV Color Model

HSV color model is a non-linear transformation of the RGB space color. The colors are obtained by a combination of the three values: the Hue (H), Saturation (S), and itself value (V). These values are represented in a circular diagram, as illustrated in the following Figure 2.

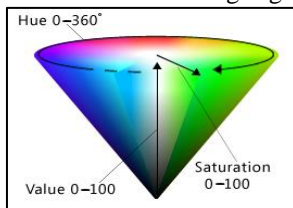


Fig. 2: HSV Color Model

HSV Color Model The three magnitudes combination can be of following values

- Hue: The kind of color (e.g. red, green, or yellow). These are represented as a degree of Angle. The possible values are ranges from 0 to 360° (although for some applications are normalized from 0 to 100%).
- Saturation: Is represented as the distance from the axis of the black-white glow. The range values may vary from 0 to 100%.
- Value: Represents the height in the blackwhite axis. The range values may vary from 0 to 100%. 0 is always considered as black. 100 could be specified white or a more or less saturated color based on saturation. By using this color model as an input image is converted using the mathematical expressions (1) to (3) that are shown below.

$$\begin{aligned}
 H &= \cos^{-1} \frac{\frac{1}{2} [(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \\
 S &= 1 - 3 \frac{\min(R, G, B)}{R + G + B} \\
 V &= \frac{1}{3}(R + G + B)
 \end{aligned}$$

### C. Gaussian Models

There are two approaches of Gaussian models that are single Gaussian model (SGM) and Gaussian mixture models (GMM). The Single Gaussian model is used to estimate skin color probability or to estimate the Gaussian density function in the two dimensional color planes. The Single Gaussian model can only simulate a human race with a skin color which is not suitable for images where there are different light. The Gaussian mixture model estimates more than a single Gaussian skin color probability density function. Because it can estimate the density distribution of arbitrary shape. The Gaussian mixture model can be used to achieve better results than the single Gaussian model when it is used to estimate the skin color space distribution. Gaussian mixture model is used frequently. For example Jeong et al used a Gaussian mixture model based on Bayesian inference to detect skin region.

Hassan pour et al proposed an adaptive skin color model based on the Gaussian mixture model to handle the changing lighting or imaging conditions in which EM algorithm is used to initially estimate the number and weights of skin color clusters. Difficulty in Gaussian mixture modeling method is the need for achieving number of single Gaussian distribution and how to determine the optimal number.

## III. METHODOLOGY

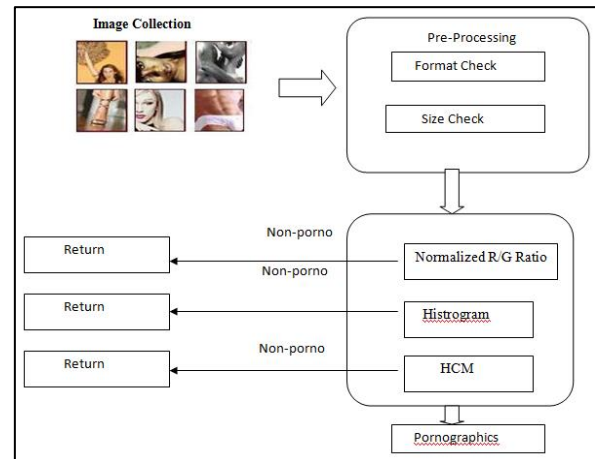


Fig. 3: Methodology

### A. Steps

#### 1) Preprocessing

A web page to be analyzed and collection of different images store in database. In preprocessing stage the image types must be JPG/JPEG, BMP, PNG, or PCX. so our project are used JPEG Format. After that check the size format so the image sizes must be in the range of 50x50 pixels and 1300x1300 pixels. After that, the web sites are passed to image-content filter. Then to remove noise and eliminate irrelevant, visually unnecessary information. Noise is unwanted information that can result from the image acquisition process.

#### 2) Processing

System consists of three main processes: normalized R/G ratio, histogram analysis, and matrix composition based on skin detection.

##### a) Normalized R/G Ratio

Normalization is any data transformation that adjusts for these effects and allows the data from two samples to be appropriately compared.

Normalized R/G ratio which is using the pixel ratios (red and green color channels).

Image processing, normalized RGB is a representation, which is easily obtained from the RGB values by a simple normalization procedure:

$$\begin{aligned}
 r &= R/R+G+B \\
 g &= G/R+G+B \\
 b &= B/R+G+B
 \end{aligned}$$

The sum of the three normalized components is known (r+g+b = 1). As the r and g on the brightness of the source, RGB color is diminished by the normalization. Certain values of R/G ratio were used as skin presence indicators. Therefore, we applied only two color tones (red and green) for image consideration in our work.

If a web site is not blocked by the normalized R/G ratio comparison phase, it is passed to the next phase, which is histogram analysis.

### B. Histogram Analysis

In image processing, a histogram is a graph demonstrating the number of pixels in an image at each different intensity value of that image HR(i) be the color histogram of image i.

An image can then be represented by a feature vector H as follows:

$$H = \{HR(1), HR(2), \dots, HR(M)\}$$

If some of images are in the range of training set results, they may be pornographic images. Afterwards, they will be processed using Image Analysis.

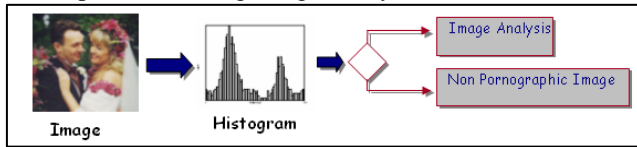


Fig.2.2Image analysis of Histogram

If the web sites cannot be blocked by histogram analysis, they will be passed to the next process, the human composition matrix based on human skin detection.

### C. Human Composition Matrix based on Human Skin Detection

We propose a new technique for pornographic web site filtering. It is called the Human Composition Matrix (HCM). If a web site contains several nude images.

The classification of skin regions into two classes: Skin or non-skin.

The HCM module is done in five steps:

1) *Step 1*

Classify the pixels of the color image into two classes (skin / non-skin) using a skin detector, the result is a binary image.

2) *Step 2*

Segmenting the binary image into a square matrix format. The results show that the 6x6 matrix size is most effective for this task.

3) *Step 3*

Choosing a suitable color space. Converting the image into the same color space that was used in the training phase.

4) *Step 4*

Classifying each pixel using the skin classifier to either a skin or non-skin.

5) *Step 5*

Elimination of non-significant regions.

### D. Skin Color based methods - Advantages

- Allows fast processing
- Robust to geometric variations of the skin patterns
- Robust under partial occlusion
- Robust to resolution changes
- Eliminate the need of cumbersome tracking devices or artificially places color cues
- Experience suggests that human skin has a characteristic color, which is easily recognized by humans.

### E. Experimental Results

1) *The Experimental Results of Filtering by Image Analysis*  
The proposed system is implemented using DOT NET (C sharp) language and tested using the database image of JPEG format. Database contains collection of 4000 pornographic images are used in our experiment. The results are shown in table 2. Sample images were downloaded from www.debonaiblog.com, www.gettyimages.com and

www.porn.com. All results are taken by using Personal computer Pentium IV 2.4 GHz. processor with 1 GB Memory.

Query Images	algorithm	websites	Accuracy	Over-blocking
Img1.jpg	R/G ratio	pornographic images	90.60	6.00
Img2.jpg	HCM	pornographic images	95.20	2.00
Img3.jpg	HCM	pornographic images	96.50	1.00
Img4.jpg	R/G ratio	pornographic images	91.30	5.00
Img5.jpg	HCM	pornographic images	98.20	1.00

Table 1: Experimental Results

### IV. CONCLUSIONS

This paper describes a comparative study of all color models to detect images with an adult classified content in color images. The experiment using the RGB, HCM for recognition of skin in the given image. These models work effectively although in some images. There could be some fault tolerance due to the image lighting conditions and type of image when taken for experiment. There is another factor that can be by a bad interpretation of the system. The importance of the comparison among the color models was done to know efficient color model to recognize skin pixels in given content. This way know whether the input image is a porn content image or not, at final could prove that the system carry out effectively. Hence, we could conclude that HCM color model has greater edge than other color model.

### REFERENCES

- [1] A. de Carvalho, A. Brayner et al., —Grand Challenges in Computer Science Research in Brazil – 2006-2016, Brazilian Computer Society, Tech. Rep., May 2006.
- [2] S. Agarwal and A. Awan, —Learning to detect objects in images via a sparse, part-based representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, pp. 1475–1490, 2004, member-Dan Roth.
- [3] L.M. Bergasa and M.Mazo and A. Gardel and M.A. Sotelo and L. Boquete, .Unsupervised and adaptive Gaussian skin-color model, Image and Vision Computing, 2000,18,987-1003
- [4] Xiaoyin Wang; Changzhen Hu; Shuping Yao; “An Adult Image Recognizing Algorithm Based on Naked Body Detection”, Computing, Communication, Control, and Management 2009. (CCCM 2009). ISECS International Colloquium on, pp. 197 – 200, 2009.
- [5] Jorge A. Marcial Basilio, Gualberto Aguilar Torres, Gabriel Sánchez Pérez, L. Karina Toscano Medina, Héctor M. Pérez Meana, “Explicit Image Detection using YCbCr Space Color Model as Skin Detection”, In Proc. Of 5<sup>th</sup>
- [6] C. Habis, and F. Krsmanovic, “Explicit Image Filter,” Stanford Univ., 2005. WSEAS Int. Conf. on Computer Engineering and Applications (CEA ‘11), pp. 123 – 128, 2011.

- [7] (2011) The Paraben's Porn Stick Detection website. [Online]. Available: <http://paraben-sticks.com/porn-detection-stick.html>.
- [8] D. A. Forsyth, M. Fleek, and C. Bregler, "Finding naked people", Proc.Forth European Conference on Computer Vision. pp 593-602, 1996.
- [9] G. Wiederhold and J. Z. Wang, "WIPE (TM): Wavelet Image Pornography Elimination; A System for Screening Objectionable Images".
- [10] Li Chen, J. Zhou, Z. Liu, W. Chen and G. Xiong, "A Skin Detector Based on Neural Network", In IEEE 2002 International Conference on Communication, Circuits and Systems and West Sino Expositions, vol. 1, 615 – 619, 2002.

