Towards Automatic Detection of Child Pornography

Napa Sae-Bae[1], Xiaoxi Sun[1], Husrev T. Sencar[2,3], and Nasir D. Memon[1,2]

[1] CSE Department
NYU, Polytechnic School of Engineering
Brooklyn, NY 11201 USA

[2] Research Institute
NYU Abu Dhabi
Abu Dhabi, UAE

[3] Department of Computer Engineering
TOBB ETU
Ankara 06560, Turkey

***Abstract*—This paper presents a child pornographic image detection system that identifies human skin tones in digital images, extracts features to detect explicit images and performs facial image based age classification. The novelty of the technique relies on the use of a robust and very fast skin color filter and a new set of facial features for improved identification of child faces. Tests on a dataset containing explicit images taken under different illuminations and reflecting a diversity of human skin tones, show that explicit images can be differentiated from benign images with around 90% accuracy. Similarly, tests performed on adult and child facial images yielded an accuracy of 80% in detecting child faces. Test conducted on 105 images involving semi-naked children (with no sexual context) revealed that the system has true positive rates of 83% in detecting explicit-like images and 96.5% in detecting child faces.**

I. INTRODUCTION

The ease of access, relative anonymity, and the borderless nature of the Internet has made child pornography a problem for governments and law enforcement officials around the world. Despite the existence of a great many laws worldwide that prohibit both possession and distribution of such content, it has proven to be an extremely difficult problem to curtail online storage and distribution of child pornography. From a technical stand point, there have been two thrusts in combatting child pornography. The first focusses on distribution by monitoring the Internet to identify sites, determining offenders and restricting access to such content. The second focusses on detecting possession of child pornography when conducting forensic examinations of suspected computers, cameras and data storage devices and examining active or deleted files [1], [2]. In both, the ability to automatically identify child pornography is crucial.

In this work, we propose an automated method for identifying child pornographic images. However, the performance of the proposed technique is evaluated only on explicit-like child images since experimentation on actual child pornographic images is not permissible by law. In order to evade definitional ambiguities and limit the scope of the problem, we define child pornography as an image that depicts clearly prepubescent children in a sexually explicit manner. This definition is also in line with the findings of investigators that most arrested child pornography possessors had images of children younger than the age of 12, and most of the images were explicitly showing sexual acts [3]. Therefore, the ability to detect nudity and detect children in images is crucial for identifying child pornography. Inspired by prior work in these areas, this paper presents a system designed to detect explicit images and child faces. These two are then combined to detect images with children with large parts of their skin exposed. Our hypotheses is that such a system will be able to detect child pornography images, albeit with some false positives. In fact a version of the proposed system has been implemented as part of the Adroit Photo Forensics Tool [4], which is being used by law enforcement officers around the world.

The next two subsections provide an overview of related work and a give a sketch of the proposed method and its key contributions. Section II, provides specifics of the proposed explicit image detection method and gives performance results. Details on child face detection and results showing its efficacy are given in Section III. We provide experimental results for the overall system in Section IV.

A. Related Work

From a content analysis point of view, there have only been a few works proposed to detect child pornography. One of the most comprehensive studies in this area was performed by Ulges *et al.* [5] who adapted a well-established visual recognition method, *i.e.,* bag-of-visual words feature representation, to design a solution to this problem. This technique is based on clustering a collection of regions of interest in images into visually coherent patches, called visual words. Each word is described in terms of the low-frequency DCT transform coefficients of the identified patches in the YUV color space, and a classification model is built based on the presence of those visual words. Tests performed on child pornographic images showed that their technique achieves an equal error rate of 11-23% depending on the dataset used to represent the benign images.

In [6], Tanner devised a method to complement an existing pornography scanner, *i.e.* RedLight software, with child pornography detection capability. The method automatically extracts features from facial landmarks, like eyes, nose and mouth, to build a classification model for age estimation, and is reported to have an accuracy of 70% in identifying children aged 12 years or younger. In [7], Islam *et al.* proposed the use of a skin tone model explicitly for children and affine-invariant geometric descriptors to distinguish between child pornographic and benign images. However, no results on the performance of the method were provided.

Numerous techniques have been proposed for detection of explicit images [8]–[13]. Since detection of skin tones is an important indication of the presence of humans, skin tone detection lies at the core of all these techniques. Various color spaces have been considered to obtain a better representation of skin tones and lessen the influence of the variable illumination conditions. Skin tone detection is typically followed by the extraction of a vector of features that capture low-level characteristics such as texture, or high-level features that require complex computations to interpret the content of the image. The resulting features are then integrated with a statistical classifier to determine if the image in question is explicit.

Similarly, many techniques have been proposed for facial image based age-group classification and age prediction. The earliest work on age-group classification is by Kwon *et al.* [14]. This work classified a facial image of an individual into three groups, namely, baby, young adult and senior, using a set of features that describe cranio-facial development and wrinkle evidence. Horng *et al.* [15] later proposed use of geometric features in addition to wrinkle features for classification. Another direction in child face recognition research has been in the development of age prediction models [16]–[19]. However, many of these works require some prior information to estimate age (like images collected over time), and it has been reported

978-1-4799-5751-4/14/$31.00 ©2014 IEEE

by Ramanathan *et al.* that estimation of the age range, as opposed to actual age, would be more feasible [20].

B. Key Features of Proposed System

From a practical standpoint, the two most important qualities of a child pornography detection system are accuracy and speed. Our approach aims at achieving a good trade-off between these two requirements. For this purpose a skin tone filter that improves upon existing skin tone filters is proposed. Low- and high-level visual features extracted from skin tone regions are further augmented by facial features. These features are then used for designing classification systems that not only discriminate between explicit and non-explicit images but also attempt to perform age classification of human faces.

Our contributions mainly lie in two areas. The first is in the detection of skin regions more accurately and effectively. For this, a dataset of more than 5K images that reflects diversity of human skin colors under different illuminations was created. A set of skin color classification models from the literature that attempt to represent the variation in human skin colors in a variety of color spaces were tested against this dataset to identify the best-performing and redundant ones. The selected filters were consolidated and modeled together under a binary classification system to create a hybrid skin tone filter. Simulation results show that the hybrid skin tone classifier performs $15-20\%$ better than the individual skin tone filters. Compounding the color information by a number of features that utilize smoothness and connectedness of skin regions, a classification system is built that is both efficient and effective in quickly examining a large storage device and can be applied to online monitoring of Internet traffic.

The second contribution is an improved version of a face detection based on OpenCV face algorithms in conjunction with the proposed skin tone filter and a novel face classification model based on simple facial component detection algorithm. Specifically, the face classification model utilizes geometric information as well as characteristics of eyes, nose and mouth regions, in order to verify whether a given face image is indeed a child face or an adult face. The proposed method is robust to different types of facial images, *e.g.*, images with various facial expressions, poses, angle or facial orientations. Our results show that our age classification system can achieve an accuracy of around 80%.

The proposed system involves two core modules: 1) explicit image detection (EID), and 2) face detection and classification. Each of these are further detailed in the following sections.

II. Explicit Image Detection

The proposed Explicit Image Detection (EID) module consists of four main components: (1) preprocessing; (2) skin tone filter; (3) visual feature extraction; and (4) classification.

A. Preprocessing

The color representation of a skin obtained by a camera is influenced by many factors like illumination, surface angle to the light and the proprietary color processing algorithm (such as color demosaicing, white balance adjustment and illumination color correction) employed by a digital camera. When combined with the variety of different skin colors, the ability to represent color in a way that is invariant to changes in illumination becomes an important problem in skin tone detection. To address this problem we utilize a dynamic white balance adjustment method [21]. All images are preprocessed with this technique to decrease sensitivity to lighting variations prior to any other processing.

B. Skin tone filter

Skin tone detection is typically performed in two steps. First, a given pixel is transformed into an appropriate color space that allows accurate and compact modeling of skin-color distribution. Second a decision rule is applied to identify skin-color pixels. A skin color model (filter) can never be wholly perfect as each filter reacts differently to variations in illumination and skin colors. To locate the connected pixels that correspond to areas where human-skin appears in the image, we considered 13 proposed skin models that define explicitly (through a number of rules) the boundaries for how skin tones cluster in various color spaces and compared their performances in terms of their detection accuracy. Tests were performed on skin and non-skin patches extracted from more than 10K images representing a diverse range of explicit images. Based on the results, poorly performing and redundant skin filters were eliminated. The remaining seven filters, which operate in RGB [22], YUV [23], [24], YIQ [24], [25], YCbCr [26], [27], HSV [28], [29] and custom color spaces [30], although exhibiting varying performance, had different types of errors indicating different expertise in identifying colors of skin and sensitivity to variations in lighting.

Preliminary simulation results revealed that five out of the seven selected color filters consistently misidentified certain shades of red color as a skin color. All these skin filters were modified by updating the decision parameters used when deciding whether a pixel is skin color or not. Then, we designed a hybrid skin tone filter by combining the seven filters in a single filter using a 5-out-of-7 voting scheme. The results indicated that the hybrid skin filter performs better than any of the individual skin filters by at least 10% in accuracy.A non-linear classification system (based on SVM) was also used to represent skin colors in each of the corresponding color space. This was realized by treating representations in each color space as part of a feature vector and creating a binary classifier from the obtained feature vectors. The performance of the resulting classifier was found to be 8% better than that of the hybrid skin filter. To further increase the speed, classification results for all the RGB color space values was created and stored as a look-up table for fast access.

C. Visual feature extraction

Since skin colors are very common in nature, the output of the skin tone filter will include many objects that indeed have skin colors but are not part of a human skin. To further analyze the images, we deploy a region growing method to remove noise and connect broken regions and extract a 11-dimensional feature vector from the detected skin-color image regions to capture the notion of explicitness in an image. These features include both low-level visual features such as, skin texture, shape, ratios, and high-level ones such as faces and bodily forms. Human skin exhibits relatively smooth patterns; therefore, textural properties can be used to eliminate most non-skin blobs with skin-tone colors. To quantify the degree of smoothness of human skin regions, the energy of DCT coefficients in different bands was measured and used as features. This was done by computing the energy distribution along horizontal, vertical and diagonal directions in DCT domain from all (8×8) image blocks in skin-color image regions to obtain three features [31] [32]. The other visual features used included the following.

- **Pixel ratios:** The number of skin color pixels in an image is a good indicator of explicit images. Moreover, humans in explicit images usually lie in the focal plane and are typically centered. Hence, features based on the ratio of skin pixels and the number of pixels in the largest skin blob to total number of pixels and to pixels in the center area of an image were measured (4 features).

- **Distribution of skin blobs:** Many explicit images include nakedness, which essentially translates to the detection of comparatively large sized skin regions. For this purpose the number of connected skin sub-regions, the distances between them, and their relative sizes were used as features (2 features).
- **Face features/indicators:** In most explicit images, one or more faces would be present. Furthermore, these faces will not typically dominate the image in size(like face portraits). To capture this property face detection was applied to image and the size and connectedness of the face to other detected images was measured (2 features).

D. *Classification*

In the final step, the features extracted from the skin color filtered images were combined through a binary classification system (a support vector machine) to better model the relation between these features and to automatically discriminate explicit images from non-explicit images. Experiments were performed on 10K manually labeled images, half of which were explicit and half were non-explicit. Explicit images were obtained from multiple web sites that host adult images and contained a sample of all skin colors and categories listed on those sites. Non-explicit images included a variety of images (including landscape, seashore, cityscape, and pictures of painting, *etc.*) obtained from web sites and public datasets, and it also included bikini body images (1000 images) and images of people and portraits (1000 images). We utilized half of the dataset for training by randomly selecting half of each sub-category and the rest for testing. The detection accuracy on this image set was measured to be 92%. Since many studies in EID report performance results on custom datasets, a comparison has not been possible. However, given the composition of the dataset, the achieved accuracy can be considered among the highest reported in the literature.

III. CHILD FACE DETECTION

Once an image is classified as explicit, the next step is to verify whether any child face appears in the image. To detect a child face in a given image, we propose a two step process. The first is to detect all possible faces in the image using face detection algorithm. The second is to classify those faces into two groups: child or adult faces.

A. *Face Detection*

In our face detection algorithm, we take all the potential face objects detected earlier using OpenCV in the EID module as an input. One problem associated with face detection is that low false-negative rates can be achieved only at the expense of relatively high false-positive rates [33]. In other words, when the goal is to detect all the faces in an image, many non-face objects will also be detected as faces. Consequently, a child face classifier will be forced to decide on the age-group of those non-face objects which could, in turn, increase false alarm rate of the system. To address this issue, we eliminate all detected face objects that have no eyes and if the majority of their pixels have non-skin colors. For this, we deployed the OpenCV eye detection and the skin-color filter described in the previous section.

To evaluate performance of the proposed face detection technique, we ran OpenCV face detection on 319 people images collected from internet and found 389 objects. Then, each of these 389 objects are verified whether they were a real face object using the proposed method. The verification result in terms of confusion matrix is presented in Table I. It is seen that the proposed model can reject many false detection objects detected by OpenCV face detection resulting in much higher precision from 74.81% to 96.01%. In other words, detected objects can be recognized with 99.23% accuracy.

TABLE I: Confusion matrix of the face detection model when applying OpenCV eye detection and skin tone filter detection on 389 faces and 131 non-face that are detected by OpenCV face detection

Classified Label	Actual Label	
	Face	Non-Face
Face	385	16
Non-Face	4	115

B. *Child Face Classification*

In this section, details of the proposed child face classifier are described. Given a face image, first the positions of the two eyes in the face are located. Second, the face image is rotated according to the angle of the line drawn from the left to the right eye. The third and fourth steps involve searching for horizontal lines that indicate mouth and nose positions, respectively. In the fifth step, positions of these facial components are used to segment the face image into 9 regions before extracting descriptive features from each of the regions. In addition, distances between facial components are used as part of the feature set. Lastly, an SVM model is used to classify the image represented by this feature set.

1) Eye Localization: While several eye localization techniques have been proposed, many rely on a classification model requiring a large number of training samples [34], [35]. Consequently, classification performance depends on the variation of eye images and non-eye images in the training set. In this work, we propose a simple eye localization technique that does not require any training samples. Our proposed eye localization technique utilizes two characteristic of the pupil assuming the eyes are properly illuminated. The first characteristic is due to the fact that there will be a reflection point located near the pupil center. The second concerns the shape of the pupil, *i.e.*, a small bright spot in a dark region. Our eye localization method is summarized as follows.

First, an eye reflection coefficient $E(i,j)$ is computed by applying an average filter with a window of size 7×7 on the Sobel edge gradient magnitude of a given face image. Next, a pupil pattern matching coefficient $P(i,j)$ is computed by applying the following filter on the original face image. Note that this mask is designed to highlight the spot, or location, that match a pupil pattern.

$$Mask = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 3 & 3 & 3 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 3 & 3 & 3 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 3 & 3 & 3 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

Finally an eye candidate coefficient, $Eye(i,j) = -\frac{E(i,j)}{P(i,j)}$ is computed to locate eye positions. That is, we search for $\underset{(i,j)}{\operatorname{argmax}}\, Eye\,(i,j)$ on both left and right sides of the image.

2) Face rotation: Once both eyes are detected, the orientation of the face image is normalized according to the position of the two eyes in order to simplify other facial components detection steps. That is, the face image is rotated such that the positions of the left and right eyes are on the same horizontal line by rotating with the angle $\theta = tan^{-1}((y_2 - y_1)\,/\,(x_2 - x_1))$.

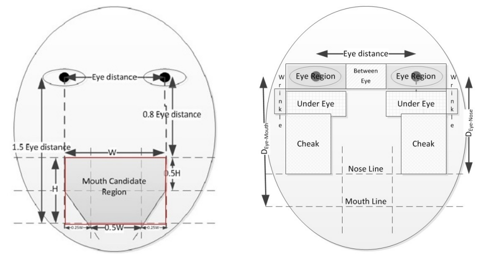
3) Mouth detection: Given the face image with normalized orientation, we search for a horizontal line that indicates the mouth's position. Previous work to locate mouth and nose positions has used a geometric face model, i.e., the ratios of distances between each pair of components [36]. However, these positions are just a rough estimation as geometric face models differ from person to person. In addition, the component distances derived from face images of the same person at different angles are different. S. Baskan *et al.* [37] addressed this issue by employing two models to detect the position of closed mouth and open mount to reflect intensity difference between teeth and lips. To simplify their model, we used relative intensity difference in comparison to skin pixel intensity instead of absolute intensity in conjunction with the following observations. First, intensity of pixels in mouth region is far different from those of skin pixels. Second, the largest dimension of mouth shape typically lies on horizontal line.

We first define a mouth candidate region as the pentagon shape shown in Fig. 1a. Then, let C be the image mask indicating this region such that $C(i,j)$ is set to 1 if the coordinate (i,j) is within the limits of the mask and to 0 otherwise. Let $I(i,j)$ be intensity of each pixel. Then, the skin intensity I_{Skin} is computed as the median intensity of all pixels within the region, $I_{Skin} = \text{median}\{I(i,j) \times C(i,j)\}$, and the mouth intensity $I_{Mouth}(i,j)$ is computed as $I_{Mouth}(i,j) = \{I(i,j) - I_{Skin}\} \times C(i,j)$. Next, we define the mouth matching coefficient, $M(i,j)$ as $M(i,j) = \frac{1}{3}\sum_{k=i-1}^{i+1} |I_{Mouth}(k,j)|$. Finally, the horizontal position of the mouth is determined as $\underset{i}{\text{argmax}} \sum_j M(i,j)$.

4) Nose detection: Once the mouth is detected, subsequently, we search for a horizontal line that identifies nose position. An earlier work by S. Baskan *et al.* [37] has proposed to use projections on x and y coordinates to search for a nose location. In particular, they utilize the characteristic of nostrils in frontal view face images. That is, there are two dark regions areas, nostrils, on the left and on the right part of the face image. As a result, the method is not robust to certain face angles where these two regions could be hidden. In this work, we generalize the definition of nose position as the line that provides a clear separation between low and high contrast areas. In addition, the fact that nose is an object found between eyes and mouth can be used to reduce search space for nose detection. Our nose detection method works as follows:

Let D be the distance between the two eyes. A candidate position of nose is scanned on the vertical line across the middle between the two eyes beginning from the point that is $0.45 \times D$ lower than the eyes to the mouth's position. For each position, we define a region of interest as a square box with the width D and the height $0.15 \times D$ centering at the candidate position. Then, the box is divided equally into upper and lower part, and a ratio between variances of intensity of pixels in those two parts is computed. Lastly, the vertical position of nose is defined as the one with the largest ratio.

5) Feature extraction: In general, human faces age in two ways: structural and textural development [38]. From a baby to a teenager, the size of the face increases as well as distances between facial components. From teenager to adult, size and position of facial components remain constant. Instead, facial texture is changed due to development of muscles, size of pores, stress, *etc.*, which causes wrinkles, moles, and aging spots. Using this fact, we extract two types of features, namely, distance and textures features. For distance features, three distances between facial components are derived: a distance between the two eyes and two vertical distances between the eyes and nose and between the eyes and mouth, respectively.



(a) Mouth candidate region **(b)** Region segmentation

Fig. 1: Facial segmentation

Fig. 2: Samples from the 105 explicit-like images subjected to blurring.

For textures features, we use seven co-occurrence matrix properties extracted from each of the nine specific areas, as depicted in Fig. 1b.

6) Classification: Finally, a face image represented by the 66 features described above is classified using an SVM classifier. Particularly we used Libsvm to conduct experiments. To test the performance of child face classification system on real world images, we collected around 2,250 images from internet and combined them with 750 images from FG-NET Database in order to sufficiently train the SVM model. The model was validated using 20-fold cross validation where 50% of samples are used for training. Our results indicate that, with 95% confidence level, the system could achieve at least 81.63% accuracy in identifying child faces.

IV. Validation of Overall System

We created a dataset to estimate the overall performance of the proposed system. Note that, in the previous experiment, we have used the actual explicit images of adults in validating EID modules. However, due to infeasibility of working on actual child pornographic images, we limited ourselves to explicit-like child images. For this, we collected 105 child images from internet that involved semi-naked children in non-sexual contexts. Some sample images are presented in Fig. 2. All these images were analyzed by the EID and child face detection modules. Out of the 105 images, 87 were classified by the EID as explicit. In these 84 images, a total of 93 objects were detected, where 86 of them were indeed child faces and the rest were false detection. Out of these 86 child faces, 77 were correctly classified. Three out of the six non-face images, however, were also incorrectly classified as child faces. This yields a true positive rate of 89.53% in child face detection. With this result, a 74.19% detection of child pornography images could be expected given that all the faces are identified by the OpenCV face detection algorithm. Although we have no way of verifying accuracy on real child pornography images, it should be noted that the presented techniques have been incorporated in the Adroit Photo Forensics Tools used by law enforcement officials around the world.

References

[1] Anandabrata Pal and Nasir Memon, "The evolution of file carving," *Signal Processing Magazine, IEEE*, vol. 26, no. 2, pp. 59–71, 2009.

[2] Nasir Memon and Anindrabatha Pal, "Automated reassembly of file fragmented images using greedy algorithms," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 385–393, 2006.

[3] Janis Wolak, David Finkelhor, and Kimberly J. Mitchell, "Child-pornography possessors arrested in internet-related crimes: Findings from the national juvenile online victimization study," *National Center for Missing & Exploited Children Report*, 2005.

[4] "Adroit photo forensics," *http://digital-assembly.com*, 2013.

[5] Adrian Ulges and Armin Stahl, "Automatic detection of child pornography using color visual words," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–6.

[6] Kylie Tanner, "Modeling automated detection of children in images," 2011.

[7] Mofakharul Islam, Paul A. Watters, and John Yearwood, "Real-time detection of children's skin on social networking sites using markov random field modelling," *Inf. Secur. Tech. Rep.*, vol. 16, no. 2, pp. 51–58, May 2011.

[8] Michael J Jones and James M Rehg, "Statistical color models with application to skin detection," *International Journal of Computer Vision*, vol. 46, no. 1, pp. 81–96, 2002.

[9] David A Forsyth and Margaret M Fleck, "Automatic detection of human nudes," *International Journal of Computer Vision*, vol. 32, no. 1, pp. 63–77, 1999.

[10] Jiann-Shu Lee, Yung-Ming Kuo, Pau-Choo Chung, E Chen, et al., "Naked image detection based on adaptive and extensible skin color model," *Pattern Recognition*, vol. 40, no. 8, pp. 2261–2270, 2007.

[11] Weiming Hu, Ou Wu, Zhouyao Chen, Zhouyu Fu, and Steve Maybank, "Recognition of pornographic web pages by classifying texts and images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 6, pp. 1019–1034, 2007.

[12] Bruno Jedynak, "Blocking adult images based on statistical skin detection," *Electronic Letters on Computer Vision and Image Analysis*, vol. 4, no. 2, pp. 1–14, 2004.

[13] Thomas Deselaers, Lexi Pimenidis, and Hermann Ney, "Bag-of-visual-words models for adult image classification and filtering," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008, pp. 1–4.

[14] Young Ho Kwon and N. da Vitoria Lobo, "Age classification from facial images," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, jun 1994, pp. 762 –767.

[15] Wen-Bing Horng, Cheng-Ping Lee, and Chun-Wen Chen, "Classification of age groups based on facial features," *Tamkang Journal of Science and Engineering*, vol. 4, no. 3, pp. 183–192, 2001.

[16] Xin Geng, Zhi-Hua Zhou, and Kate Smith-Miles, "Automatic age estimation based on facial aging patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 12, pp. 2234–2240, 2007.

[17] Narayanan Ramanathan and Rama Chellappa, "Modeling age progression in young faces," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. IEEE, 2006, vol. 1, pp. 387–394.

[18] Yun Fu and Thomas S Huang, "Human age estimation with regression on discriminative aging manifold," *Multimedia, IEEE Transactions on*, vol. 10, no. 4, pp. 578–584, 2008.

[19] Xiaodan Zhuang, Xi Zhou, Mark Hasegawa-Johnson, and Thomas Huang, "Face age estimation using patch-based hidden markov model supervectors," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008, pp. 1–4.

[20] Ramanathan N., Chellappa R., and Biswas S., "Age progression in human faces : A survey.," in *Journal of Visual Languages and Computing*, 2009, Advances in Multimodal Biometric Systems'09.

[21] Ching-Chih Weng, Homer Chen, and Chiou-Shann Fuh, "A novel automatic white balance method for digital still cameras," in *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*. IEEE, 2005, pp. 3801–3804.

[22] Abhishek Choudhury, Marcus Rogers, Blair Gillam, and Keith Watson, "A novel skin tone detection algorithm for contraband image analysis," in *Systematic Approaches to Digital Forensic Engineering, 2008. SADFE'08. Third International Workshop on*. IEEE, 2008, pp. 3–9.

[23] Dong Wang, Jinchang Ren, Jianmin Jiang, and Stan S Ipson, "Skin detection from different color spaces for model-based face detection," pp. 487–494, 2008.

[24] Hong Zhu, Shuming Zhou, Jianying Wang, and Zhongke Yin, "An algorithm of pornographic image detection," in *Image and Graphics, 2007. ICIG 2007. Fourth International Conference on*. IEEE, 2007, pp. 801–804.

[25] Lijuan Duan, Guoqin Cui, Wen Gao, and Hongming Zhang, "Adult image detection method base-on skin color model and support vector machine," in *Asian conference on computer vision*, 2002, pp. 797–800.

[26] Adrian Ulges and Armin Stahl, "Automatic detection of child pornography using color visual words," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, july 2011, pp. 1 –6.

[27] Jau-Ling Shih, Chang-Hsing Lee, and Chang-Shen Yang, "An adult image identification system employing image retrieval technique," *Pattern Recognition Letters*, vol. 28, no. 16, pp. 2367–2374, 2007.

[28] Son Lam Phung, A Bouzerdoum Sr, and D Chai Sr, "Skin segmentation using color pixel classification: analysis and comparison," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 1, pp. 148–154, 2005.

[29] Francesca Gasparini and Raimondo Schettini, "Skin segmentation using multiple thresholding," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 60610F–60610F.

[30] Abbas Cheddad, Joan Condell, Kevin Curran, and Paul Mc Kevitt, "A skin tone detection algorithm for an adaptive approach to steganography," *Signal Processing*, vol. 89, no. 12, pp. 2465–2478, 2009.

[31] KW Chun, KW Lim, HD Cho, and JB Ra, "An adaptive perceptual quantization algorithm for video coding," *Consumer Electronics, IEEE Transactions on*, vol. 39, no. 3, pp. 555–558, 1993.

[32] Golam Sorwar, Ajith Abraham, and Laurence S Dooley, "Texture classification based on dct and soft computing," in *Fuzzy Systems, 2001. The 10th IEEE International Conference on*. IEEE, 2001, vol. 2, pp. 545–548.

[33] Paul Viola and Michael Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 4, 2001.

[34] Peng Wang, Matthew B Green, Qiang Ji, and James Wayman, "Automatic eye detection and its validation," in *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*. IEEE, 2005, pp. 164–164.

[35] Paola Campadelli, Raffaella Lanzarotti, and Giuseppe Lipori, "Precise eye localization through a general-to-specific model definition.," in *BMVC*. Citeseer, 2006, pp. 187–196.

[36] Frank Y Shih and Chao-Fa Chuang, "Automatic extraction of head and face boundaries and facial features," *Information Sciences*, vol. 158, pp. 117–130, 2004.

[37] Selin Baskan, M Mete Bulut, and Volkan Atalay, "Projection based method for segmentation of human face and its evaluation," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1623–1629, 2002.

[38] Narayanan Ramanathan, Rama Chellappa, Soma Biswas, et al., "Age progression in human faces: A survey," *Visual Languages and Computing*, vol. 15, pp. 3349–3361, 2009.