Adult Image Content Classification Using Global Features and Skin Region Detection

Hakan Sevimli[1,2], Ersin Esen[1,3], Tuğrul K. Ateş[1,3], Ezgi C. Ozan[1,3],
Mashar Tekin[1,2], K. Berker Loğoğlu[1,4], Ayça Müge Sevinç[1,4], Ahmet Saracoğlu[1,3],
Adnan Yazıcı[2] and A. Aydın Alatan[3]

[1] TÜBİTAK Space Technologies Research Institute
[2] Department of Computer Engineering, M.E.T.U.
[3] Department of Electrical and Electronics Engineering, M.E.T.U.
[4] Graduate School of Informatics, M.E.T.U.

{hakan.sevimli, ersin.esen, tugrul.ates, ezgican.ozan, mashar.tekin, berker.logoglu, muge.sevinc, ahmet.saracoglu}@uzay.tubitak.gov.tr
yazici@ceng.metu.edu.tr, alatan@eee.metu.edu.tr

Abstract. A method for adult content classification and nudity detection is presented. Objective of this method is to classify images into different classes, varying on the degree of adult content. We utilize MPEG-7 descriptors to represent visual information. Skin regions are detected to model adult content more precisely, as well as to eliminate false-positives. Proposed method is tested with conventional image sets. Experimental results indicate that the algorithm has an acceptable detection performance.

Keywords: adult image classification, nudity detection, MPEG-7 visual descriptors, close-up face detection, skin color detection

1 Introduction

The amount of visual content available on the internet is well beyond manual analysis. Adult classification of images is one of the major tasks for semantic analysis of visual content. Modern approach to this problem is introducing rating mechanisms to prevent unsolicited access to this type of content. This prevention is especially critical for children [1]. It is an obvious fact that such a rating mechanism should make use of automatic analysis. One desired property of such a system is the opportunity to dynamically adjust the content severity level. For instance, different clients may require different restriction levels at different times [2].

In the literature, different adult image filtering methods are presented. The detection of skin areas is investigated in [3] where skin color is used in combination with other features such as texture and color histograms. Most of these systems build on neural networks or Support Vector Machines [4] as classifiers. One of the pioneering works is done by Forsyth *et al.* [5] where they combine tightly tuned skin filter with smooth texture analysis. After skin detection, geometric analysis is applied to detect of bare parts in human body figures. Another work is conducted by Duan *et* *al.* [6]. Their study is based purely on skin color detection and SVM. The images are first filtered by skin model and outputs are classified. Rowley *et al*. [7] propose a system that includes both skin color and face detection where they utilize a face detector to eliminate the effects of skin regions that belong to face area on the skin map. Yoo [8] suggests retrieving labeled images from a database where an image is labeled as adult content if most of the similar images are labeled that way.

Previous studies on this topic show that images which contain close-up face(s) mislead adult content classification systems [9], since these images have the same characteristics with ordinary nude images regarding skin color features. These studies suggest the use of face detection systems to overcome this problem.

Our aim is to rate any image as being in one of the classes defined in [10]: *normal*, *swimming suit*, *topless*, *nude* and *sexual content*. We present a system with four components. First component is skin color detection. Second component is close-up face detection. Third component is feature extraction from images by using shape, color and texture descriptors. The fourth component is to classification.

2 Proposed Algorithm

The proposed algorithm has four fundamental components: close-up face detection, skin color detection, feature extraction and classification. The general system structure is shown in Fig. 1.

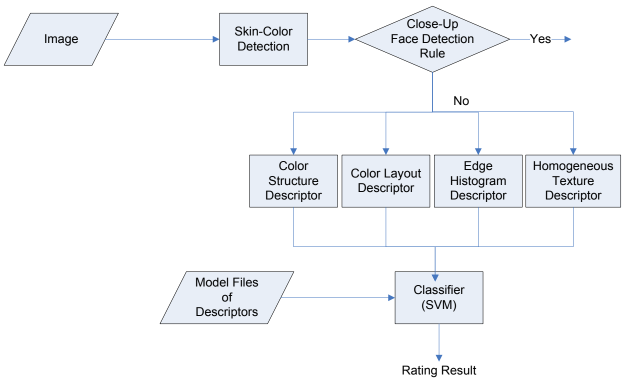

Fig. 1. Overall flow of adult rating classification algorithm.

2.1 Skin Region Detection

For the skin region detection step of this study, the method proposed by Jones and Rehg [3] is utilized. This method is based on inferring pixels on statistical skin and non-skin models, which are represented and trained with GMM. Jones and Rehg utilize millions of skin pixels for training. Mean and covariance results of this study are directly used in our method. Skin masks are obtained from skin detection and connected component labeling algorithms as shown in Fig. 2-a and 2-b. The skin mask is used to determine the regions that contain human bodies or body parts in the original image. These regions define the bounding box, which encapsulates all the parts that shows the skin color characteristics. Also skin detection is used for face elimination phase which will be described in the next section.

(a) **(b)** **(c)**

Fig. 2. (a) Input image, (b) its skin mask and (c) An example for close-up face elimination. Rectangular box indicates the face box given by the face detector.

2.2 Face Detection and Close-up Face Rule

Images containing many faces or a close-up face can be classified as nude images [9] causing false alarms. To solve this problem, we should be able to detect and eliminate close-up faces. For face detection, the method of Viola and Jones [13] is utilized.
By combining the face areas with skin masks, the ratio of total skin pixels that lie on the face area over the total skin pixel count of skin region can be calculated. If this ratio is large enough, it indicates that body parts other than the face may be found in the image. This means corresponding input images may contain nudity. Otherwise the image should be labeled as non-nude. The ratio can also be interpreted as a measure to determine the scale of human bodies in the image. Fig. 2-c shows an example of a query image, where total area of skin regions are large compared to image proportions. The test rule is $T = \tilde{A} - A / \tilde{A}$. If T is below a certain threshold, then the image is labeled as non-nude.

2.3 Feature Extraction

In feature extraction step, visual descriptors are extracted by using four different methods. These four low-level feature extraction methods are chosen from MPEG-7 descriptors [11, 12]. These are *Color Structure Descriptor* (CSD), *Color Layout Descriptor* (CLD), *Edge Histogram Descriptor* (EHD) and *Homogeneous Texture Descriptor* (HTD). These four different descriptors capture the visual information from different perspectives and provide compact representations.

2.4 Classification

We employ pattern classification on extracted features with SVM's due to their well reported potential in the literature [4]. We use multi-category SVM with a radial basis function (RBF) kernel. SVM are trained with the features extracted from the images by using the color and texture descriptors mentioned before. The OpenCV implementation of SVM is used [14].

3 Experiments

3.1 Experimental Setup

The system extracts four MPEG-7 descriptors for each image. The data set [10] that is used in the experiments has five different classes, which are *normal* images (class 1), *swimming suit* images (class 2), *topless* images (class 3), *nude* images (class 4) and *sexual activity* images (class 5).

The experimental image set consists of 1702 images for each class [10]. For each of the five classes, 1000 images are used for training and 500 images are used for testing. The image dataset, which is used to measure the success rates of close-up face elimination rule, consists of 799 frontal female faces of the FERET dataset [15]. If face elimination labels as non-nude, the classification is accomplished successfully.

3.2 Results

The classification results of the proposed system are shown in Table 1. It should be emphasized that whole image area is utilized for inactive skin detection case during training and test phases. These results indicate that color descriptors are more suitable than texture descriptors to adult image classification.

Comparing the results in Table 1 according to the use of skin color detection reveals that skin color detection slightly increases the performance. Compared to a very similar study [10], success rates of the proposed system are relatively higher, since Kim *et al.* [10] have not utilized any skin detection mechanism. Besides rating adult images with respect to their content, the results give an idea about nudity detection. It is obvious that rating images is a much more complex task than nudity detection. The results of nudity detection, where the classes 2, 3, 4 and 5 are combined into a single nude class, are represented in Table 2. In the second stage of the experiments, the effectiveness of the close-up face elimination rule is observed. Exact counts of eliminated images with respect to adult classes and the female set are shown in the Table 3. False eliminations on classes 1 to 5 are rare.

The close-up face check eliminates the images that are not likely to contain nudity. Thus they are classified as class 1 (normal images). Remaining images are classified with four feature descriptors and obtained results given in Table 4. Texture descriptors do not contribute much to the close-up face elimination process but color descriptors are more successful.

Table 1. Classification confusion matrices of experiments (with/without skin detection).

Descriptor	***Query Class***	***Predicted Class*** None	Swimsuit	Topless	Nude	Sexual
Color Structure	None	**92.4/92.8**	2.8/3.8	2.6/2	1.6/1	0.6/0.4
	Swimsuit	6.8/8	**51.8/51.8**	7.6/7.2	27.6/27.2	6.2/5.8
	Topless	7.4/10	18/20.6	**55.4/50.2**	17.8/17.8	1.4/1.4
	Nude	13.615.4	22.2/29	6.4/3.8	**56.8/50.8**	1/1
	Sexual	1/1	39/34	3/2.6	11.4/14.8	**45.6/47.6**
Color Layout	None	**80.2/75.4**	5.4/9.4	7.6/9	1.2/0	5.6/6.2
	Swimsuit	8/8.2	**22.6/27.4**	15.6/22.6	12.2/2.8	41.6/39
	Topless	5.2/6.4	20.4/23	**42/44.8**	4.6/2.4	27.8/23.4
	Nude	14/10.6	17.2/25	18.8/23.2	**12.6/10**	37.4/31.2
	Sexual	3/3.6	20/16.4	19.6/30.8	5.6/2.2	**51.8/47**
Edge Histogram	None	**82.2/80.8**	4.8/5.6	5.5/4.6	2.2/2.6	5.6/6.4
	Swimsuit	7.8/8	**41.4/34**	19.8/24.4	18.2/16.8	16/16.8
	Topless	15/15.4	25.4/30.8	**25.4/23.4**	18.2/13.4	16/17
	Nude	12.8/14.4	20/15.4	8.4/15.2	**47.4/43.8**	11.4/11.2
	Sexual	4.2/4.4	13/14.2	9/7.6	8.6/7.4	**65.2/66.4**
Homogeneous Texture	None	**74.4/74.2**	3.4/3.8	12.8/13	4.6/4.8	4.8/4.2
	Swimsuit	10.8/11	**28.8/25.2**	16.2/17.8	25/28.6	19.2/17.4
	Topless	7.4/7.2	16.6/15.4	**54.8/54.2**	8.2/8.4	13/14.8
	Nude	4.8/4.8	19/14.4	15/15	**31.6/36.8**	29.6/29
	Sexual	2.4/2.6	18/16.2	7.6/7.8	22/22.4	**50/51**

Table 2. Success rates of nudity detection when classes 2, 3, 4 and 5 are combined

Desc	***w. Skin Detection***	***wo. Skin Detection***
CSD	92.80	91.40
CLD	92.45	92.80
EHD	90.05	89.45
HTD	93.65	93.60

Table 3. Number of images that are labeled non-nude by face elimination rule

Image Set	***Eliminated***	***Percentages***
None	1	0.06(incorrect)
Swimsuit	28	1.64(incorrect)
Topless	33	1.94(incorrect)
Nude	7	0.41(incorrect)
Sexual	25	1.47(incorrect)
FERET Female	576	72.1(correct)

Table 4. Confusion values for FERET frontal female image set.

Descriptor	***Resulting Class*** None	Swimsuit	Topless	Nude	Sexual
Color Structure	**91.61**	2.13	0.13	6.13	0.0
Color Layout	**83.60**	5.89	1.00	3.13	6.38
Edge Histogram	**74.34**	7.13	5.63	2.38	10.51
Homogeneous Texture	**72.47**	7.88	5.51	4.38	9.76

4 Conclusion

In this study, a method that makes use of visual descriptors with skin detection is proposed for adult image classification. Additionally, we incorporate a close-up face elimination mechanism to get rid of false alarms. Experiments indicate that the proposed system can distinguish nude and non-nude cases. An important observation is that descriptors and decision of using skin detection or not is very crucial.

Acknowledgements. This work is funded by TÜBİTAK (The Scientific and Technological Research Council of Turkey) KAMAG TARAL 1007 (Support Programme for Research Projects of Public Institutions 1007) Research Programme, Project no. 107G229.

References

1. Zheng Q.F., Zeng W., Wen G., Wang W.Q., Shape-Based Adult Image Detection. In 3rd International Conference on Image and Graphics, pp. 150—153 (2004).
2. Deselaers T., Pimenidis L., Ney H., Bag-of-Visual-Words Models for Adult Image Classification and Filtering. In 19th International Conference on Pattern Recognition, pp. 1--4 (2008)
3. Jones M.J., Rehg J.M., Statistical Color Models with Application to Skin Detection. In International Journal of Computer Vision, Vol 46, No. 1, pp. 81--96 (2002).
4. Duda R.O., Hart P.E., Stork D.G., Pattern Classification. John Wiley & Sons, USA (2001)
5. Fleck, M., Forsyth, D.A., Bregler, C., Finding Naked People. In 4th European Conference on Computer Vision, pp. 593--602. Springer (1996).
6. Duan L., Cui G., Gao W., and Zhang H., Adult Image Detection Method Base-on Skin Color Model and Support Vector Machine. In 5th Asian Conference on Computer Vision. pp. 780--797 (2002).
7. Rowley H. A., Jing Y., Baluja S., Large Scale Image-Based Adult-Content Filtering. In 1st International Conference on Computer Vision Theory. pp. 290--296 (2006).
8. Yoo S.J., Intelligent Multimedia Information Retrieval for Identifying and Rating Adult Images. In 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems. vol. 1of Lecture Notes in Artificial Intelligence 3213, pp. 164--170 (2004).
9. Jeong C.Y., Kim J.S., Hong K.S., Appearance-Based Nude Image Detection. In 17th International Conference on Pattern Recognition, pp. 467--470 (2004).
10. Kim W., Yoo S.J., Kim J.S., Nam T.Y., Yoon K., Detecting Adult Images Using Seven MPEG-7 Visual Descriptors. In Human. Society@ Internet, pp. 336-339 (2005).
11. Manjunath B.S., Introduction to MPEG-7: Multimedia Content Description Interface. John Wiley & Sons (2002).
12. Manjunath B.S., Ohm J.S., Vasudevan V.V., Yamada A., Color and Texture Descriptors. In IEEE Transactions on Circuits and Systems for Video Technology. pp. 703--715 (2001).
13. Viola P., Jones M., Rapid Object Detection using a Boosted Cascade of Simple Features. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition. vol. 1, pp. 511 (2001).
14. Open Computer Vision Library, http://sourceforge.net/projects/opencvlibrary
15. Phillips P.J., Moon H., Rizvi S.A., Rauss P.J., The FERET Evaluation Methodology for Face Recognition Algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 22, pp. 1090--1104 (2000).