Volume 3, Issue 10, October 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Face Body Detection Method for Filtering X-Rated Content

Peddi Karthik*, Santhosh.G, B.Kranthi Kiran, P.Sammulal
Dept. of CSE& JNTUH
India

Abstract-- X-rated filtering is essential while publishing any content over Internet. This is because the x-rated has its adverse impact on children and other people who view them unintentionally. In order to filter x-rated, an efficient skin detector is required. Skin detector is used for other purposes as well including gesture recognition, face detection and so on. In the last decade many researchers were carried out on x-rated filtering. They focused on various color spaces for skin detection. Nevertheless, there was no comprehensive work on the performance evaluation of various color spaces used to detect naked human beings. In this paper we propose a novel solution to the problem of skin detection in porn images using a pixel based approach. We also built a prototype application to demonstrate the proof of concept. The empirical results reveal that the proposed technique is effective and can be used in the real world applications to filter x-rated.

Index Terms – Skin detection, x-rated, x-rated filtering, image processing

I. INTRODUCTION

Internet has become a part of everyday life for people of all walks of life. This is because it has become the electronic super highway where ones information needs are satisfied. Internet has become a global communication medium and people of all age groups use it for their needs for computing and communication. When Internet was first invented by Tim Berners Lee, the web sites were static and meant for information sharing [1], [2]. Later the web applications became dynamic with server side processing with ASP (Active Server Pages), JSP (Java Server Pages) and technologies like Servlet and EJB (Enterprise Java Beans). With the invention of AJAX (Asynchronous JavaScript and XML), the WWW has become rich in providing user experience. Now millions of people of different professions use Internet across the globe in day to day needs pertaining to their professions. Mostly web professional web portals, commercial applications like e-Commerce applications, banking applications and so on are used by people all over the world. As the content over Internet is growing day by day, malicious people started polluting the content by adding pornographic images and videos over Internet. It does mean the content over web is contaminated as harmful content such as violence, hatred and x-rated. Especially x-rated has become harmful to society in one way or other. It has created more problems than advantages. Moreover the content is published over Internet illegally [3], [4]. It is also observed that the students of all countries are addicted to Internet in terms of x-rated. More details on the x-rated and issues associated with can be found in [5], [6], [7], [8], [9], [10], and [11]. While educated adults might have control over themselves to reduce impact on the x-rated, the young ones cannot. This is the problem which has identified the need for filtering x-rated over Internet. By filtering x-rated it is possible to prevent its publications illegally. Text filtering can also be used to filter pornographic content. However, it causes problems when it also filters genuine content. The textual approaches include SurfWatch, CyberSitter, and NetNanny [12].

Content based Image Retrieval has also become a famous means of accessing images over Internet. A survey on this revealed that more than 70% web pages contain images incorporated as part of the content rendered [13]. Out of 7000 web pages studied in a survey 72% of pornographic web pages contain more than 5 images. The number of scientific publications subjected to pornographic filtering in the past is presented in figure. 1.

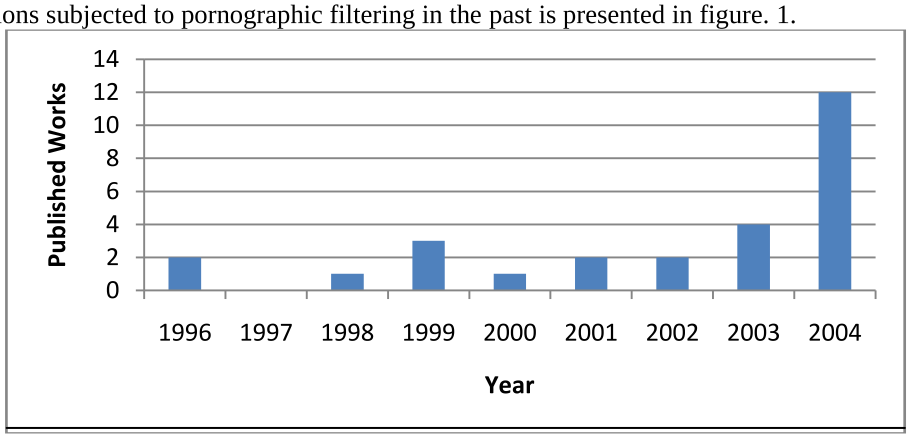

Fig.1 – Number of scientific publications subjected to pornographic filtering

As can be seen in figure 1, it is evident that the number of scientific publications that are subjected to pornographic filtering was increasing year by year. Skin detection has been used for identifying x-rated. In the literature it is found that 21 color spaces are possible to incorporate such solutions. In this paper we evaluated all possible color spaces for skin detection as part of filtering x-rated. The rest of the paper is structured as follows. Literature on x-rated filtering and skin detection techniques is reviewed in section II. Section III provides information about proposed pixel based approach for skin detection. The dataset used in this paper is described in section IV. Section IV presents experimental results while section V concludes the paper.

II. RELATED WORKS

This section reviews literature on x-rated filtering approaches including skin detection. First of all content based x-rated filtering is focused here. First of all the work related to x-rated filtering is done by Forsyth et al. [14], [15], [16]. However, their research was on filtering of nude pictures using IRgBy color space. The skin detection technique applied here is ordinary lookup table. Similar skin detection approach is followed in [17]. However, RGB is used as color space instead of IRgBy color space. Many color spaces are used for experiments on skin detection in [18], [19], [20] and [21]. The used color spaces include Nrgb, IRgBy, HSV, and RGB. The skin detection technique used is Bayesian lookup table. YUV and its variant and YIQ and its variant are used as color spaces in [22] while ordinary lookup table is used as skin detection approach. CIE-lab color space is used in [23] for experiments on skin detection. In [24] ordinary lookup table is used for skin detection while it explored color spaces such as YCb, Cr, IRgBy. In [25] only RGB color space is sued using the technique known as ordinary lookup table. Histogram based skin detection technique is used in [26] and [27] for skin detection while the experiments are made on RGB color space. In [28] the combination of RGB and HSV are considered as color spaces for skin detection. It used the skin detection technique known as ordinary lookup table.
Multiple color spaces are explored in [29] which make use of ordinary lookup table as skin detection approach for skin detection. The color spaces used include HSV, RGB, IRgBy, and YCbCr. The RGB color space along with ordinary lookup table is used by many researchers as explored in [30], [31], [32], [33], [34] and [35]. Color space such as CbCr is used along with skin detection approach named "ordinary lookup table" for skin detection. A custom color space is tried out by the authors of [36] that make use of skin detection technique such as Bayesian lookup table. In [37] two color spaces are used for experiments namely RGB, and RGB^r using a skin detection technique Bayesian lookup table. The color spaces diversity is found in literature with respect to skin detection. The color spaces explored in the literature include TSL, Nrgb, RGB^r, Nrgb, IRgBy, IRgBy+, CIE-Luv, YIQ, YUV and variations of YUV and YIQ, CIE-lab, YCbCr, CIE-XYZ, HLS, HSI, HSV, and variants of HSI and HSV. RGB is also used as one of the color spaces.

III. PROPOSED PIXEL BASED SKIN DETECTION

Pixel based skin detection is proposed in this paper. Each pixel value is considered between 0 – 255. The color spaces considered are c1, c2 and c3. The detection technique assumes that there is a function represented by P: $[0,255]^3 \rightarrow [0,1]$. According to which P(c) represents the probability of c which is related to a skin area. The function P is also known as SPM (Skin Probability Map). A lookup table (LUT) is used to achieve the desired solution. The lookup table is of two types. Ordinary Lookup Table (OLUT) and Bayesian Lookup Table (BLUT). These are the two approaches to compute P value. According to the proposed skin detection technique, the P value is computed as follows.

$$P(\vec{c}) = \frac{H_s(\vec{c})}{\max_{\vec{d}}\{H_s(\vec{d})\}}.$$

And the true and false positives are computed as follows.

$$TP = \sum_{\vec{c} \in [0,L]^p} H_s(\vec{c}) LUT(\vec{c}),$$

$$FP = \sum_{\vec{c} \in [0,L]^p} H_c(\vec{c}) LUT(\vec{c}).$$

With Bayesian lookup table, the P value is computed as follows.

$$P(skin|\vec{c}) = \frac{P(\vec{c}|skin)p(skin)}{P(\vec{c}|skin)p(skin) + P(\vec{c}|\sim skin)p(\sim skin)},$$

Algorithm Steps

1. Preprocessing an image(eliminates junk data)
2. Loading an image
3. Reading an image
4. Calculating RGB values

5. Recognition or detection algorithms (Face Body Detection Algorithms)
6. Comparing the RGB values
7. Knowing the grey value(however higher grey has a more porn content)
8. Making it as a blur image.

DATASETS

The datasets which are nothing but pornographic images were collected from Google search. The dataset contains more than 2000 porn images. All images are found in RGB color space as they are color images. They are in JPEG compression format. There was no preprocessing done on these images. There are three categories of porn images. The first category is a set of images transferred directly from web. The second category is that they are collected and edited by using editing software. The third category has images that are professionally manipulated.

IV. PROTOTYPE IMPLEMENTATION

We built a prototype application in Java platform. The application demonstrates the proof of concept which makes use of proposed skin detection technique for filtering x-rated images. The application has user-friendly interface which facilitates end users to gain access to various functionalities of the application with ease. The environment used for building prototype includes a PC with 4 GB RAM, core 2 dual processor and Windows 7 operating system.

V. EXPERIMENTAL RESULTS

Using our prototype application we made number of experiments. For instance the experiments are made with various color spaces containing perfect, partial, excessive, and irrelevant.

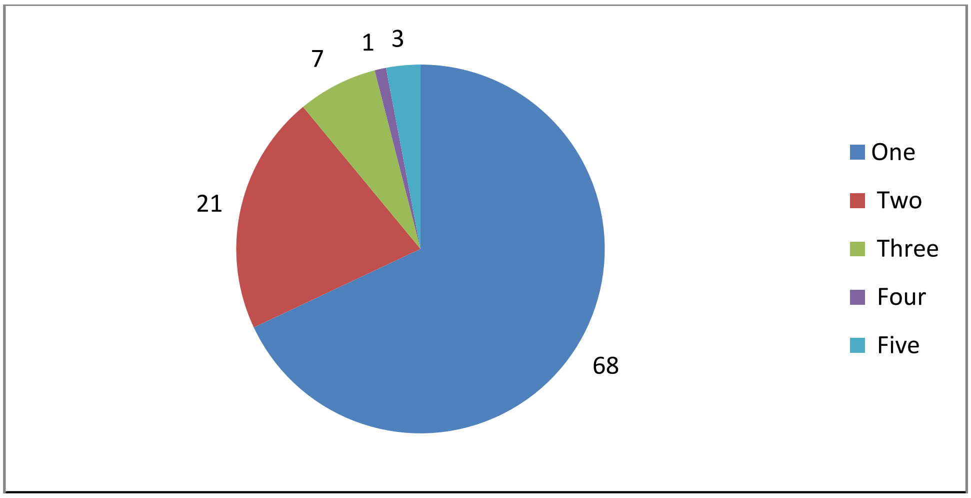

Fig. 2 – Histogram of aspect ratios

As can be seen in figure 2, statistics of images are shown in terms of number of people present in images. The web pages with up to five images are considered. The pie chart shows the number of pages with 1, 2, 3, 4, and 5 images incorporated.

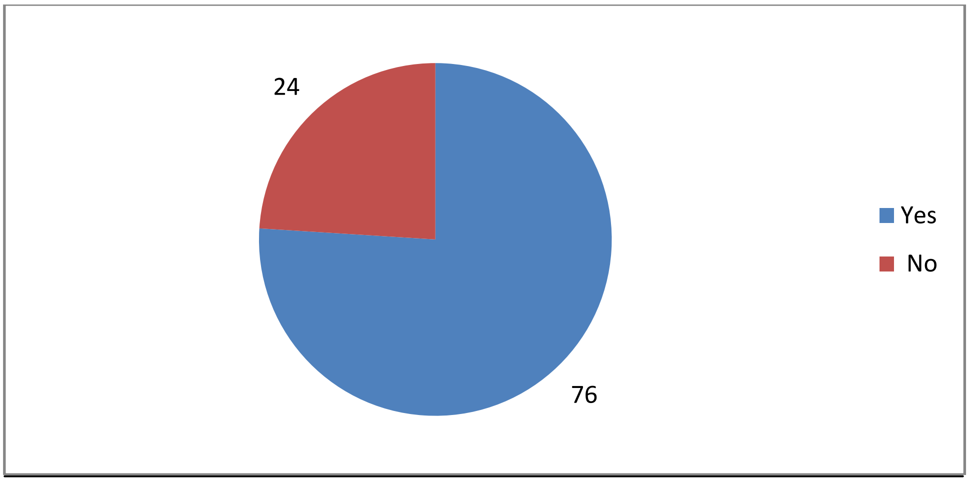

Fig. 3 – Presence of face in the image

As can be seen in figure 3, statistics are shown in terms of whether face is present in images. Majority of images have face while some images have no face involved.

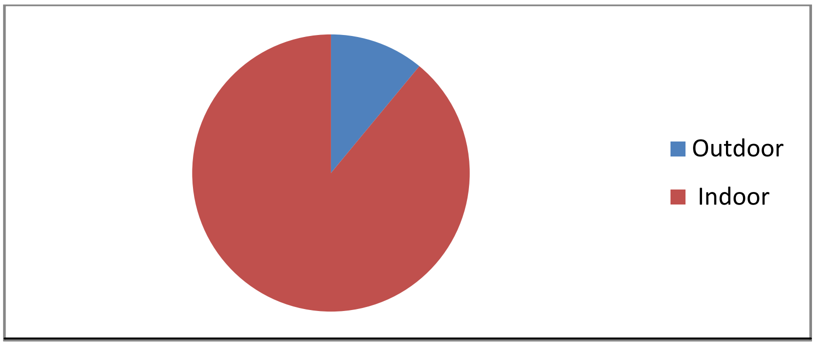

Fig. 4 – Nature of images

As can be seen in fig. 4, it is evident that the images used in the dataset are of two types namely outdoor and indoor. The indoor images are prevailing when compared with outdoor images.

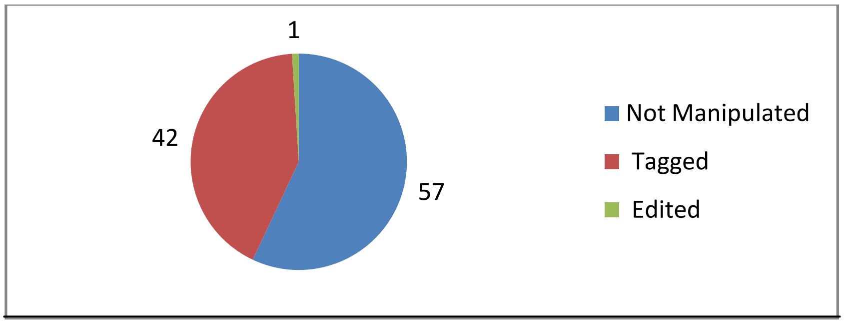

Fig. 5 – Distribution of types of images

As can be seen in fig. 5, there are three kinds of images. They are images which are not manipulated, imaged which are tagged and images that have been modified.

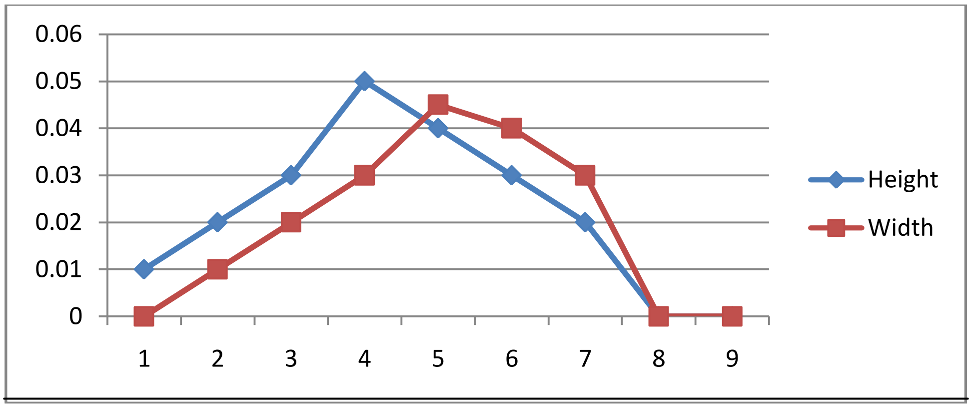

Fig. 6 – Histogram of height and width

As can be seen in figure 6, the histogram of height and width of images are presented.

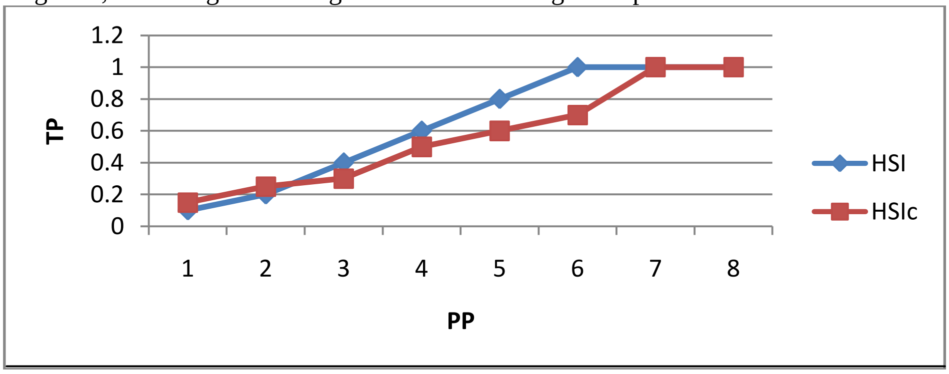

Fig. 7- Histogram of HSI and HSIc

As can be seen in figure 7, the histogram of HSI and HSIc of images are presented.

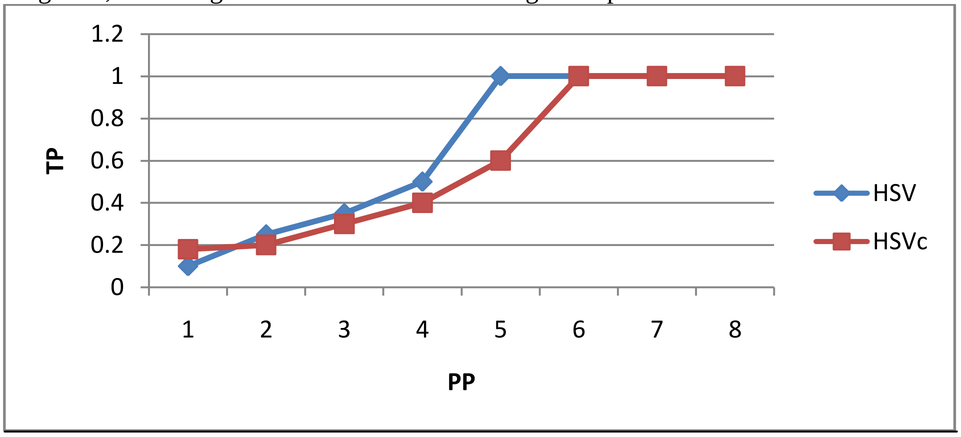

Fig. 8- Histogram of HSV and HSVc

As can be seen in figure 8, the histogram of HSV and HSVc of images are presented.

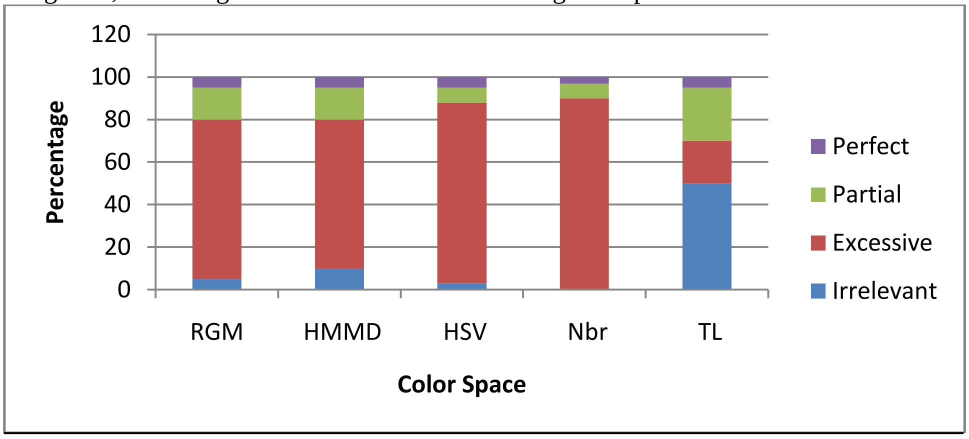

Fig. 9 – Results of computing the skin map

As can be seen in figure 9, the experimental results are presented. The rate of classification of port images if visible for various kinds of color spaces.

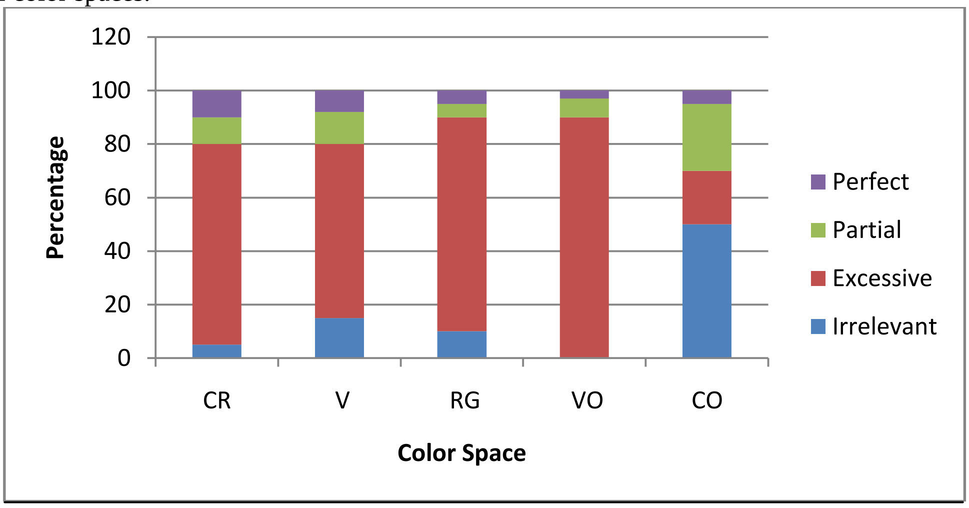

Fig. 10 – Results of computing the skin map

As can be seen in figure 10, the experimental results are presented. The rate of classification of port images if visible for various kinds of color spaces.

VI. CONCLUSIONS

In this paper we implement a novel skin detection technique which is pixel-based. We make use as many as 21 color spaces for making such experiments. Every color space is represented in seven different ways. We also evaluated the color spaces for best performance. Two approaches have been used for skin detection for filtering x-rated. Out of them the first one is ordinary while the second one is known as Bayesian LUT-based skin detection. We built a prototype application that shows the effectiveness of the proposed solution. The empirical results reveal that, in case of ordinary solution Nbr is best while 1-D color spaces are best for the other kind of skin detection. More accurate classification was possible with Bayesian approach.

REFERENCES

[1] Burke D., "Cybersmut and the first amendment: A call for a new obscenity standard," *Harvard Journal of Law & Technology*, 1996.

[2] Ding C., Chi C.-H., Deng J. and Dong C.-L., "Centralized content-based web filtering and blocking: How far can it go," in *Proceeding of IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 1999.

[3] Faucette J. E., "The freedom of speech and a frightened university's censorship of sex on the internet," *Duke Law Journal*, 1995.

[4] "Internet x-rated: Are children at risk?" http://www.nap.edu/catalog/10261.html?onpiwebextra- 050202.

[5] Krone T., "A typology of online child x-rated offending," Australian Government, Australian Institute of Criminology, Trends and Issues in Crime and Criminal Justice, No. 279, *www.ahtcc.gov.au*, 2004.

[6] Taylor M., Quayle E. and Holland G., "Child x-rated, the internet and offending," *ISUMA - Canadian Journal of Policy Research*, Vol. 2(2), 2001.

[7] Coopersmith J., "X-rated, videotape, and the internet," *IEEE Technology and Society Magazine*, Vol. Spring, pp. 27-34, 2000.

[8] Mehta M. D. and Plaza D., "Content analysis of pornographic images available on the internet," *The Information Society*, Vol. 13(2), pp. 153-161,1997.

[9] "Family safe media," http://www. familysafemedia. com/ x-rated_ statistics. html.

[10] WebsenseTM," http://www.websense.com/ products/ resources/ wp/ hr_wp.pdf, 2003.

[11] "Internet filter review," http:// www. internetfilterreview .com/ internet-x-ratedstatistics. html, 2003.

[12] Greenfield P., Rickwood P. and Tran H., "Effectiveness of internet filtering software products," CSIRO Mathematical and Information Science, 2001.

[13] Starykevitch B. and Daoudi M., "Poesia software architecture definition document," http://www.poesiafilter. org/pdf/Deliverable_3_1.pdf, 2002.

[14] Fleck M., Forsyth D. and Bregler C., "Finding naked people," in *European Conference on Computer Vision*, Buxton B. and Cipolla R., Eds. Berlin, Germany: Springer-Verlag, 1996, pp. 593- 602.

[15] Forsyth D. A. and Fleck M. M., "Identifying nude pictures," in *Third IEEE Workshop on the Applications of Computer Vision (WACV '96)*, Sarasoto, Florida, 1996, pp. 103-108.

[16] Forsyth D. A. and Fleck M., "Automatic detection of human nudes," *International Journal of Computer Vision*, Vol. 32(1), pp. 63-77, 1999.

[17] Wang J. Z., Wiederhold G. and Firschein O., "System for screening objectionable images using daubechies' wavelets and color histograms," *Computer Communications*, Vol. 21(25), pp. 1355- 1360, 1998.

[18] Bosson A., Cawley G. C., Chan Y. and Harvey R., "Non-retrieval: Blocking pornographic images," in *Proceedings of International Conference on the Challenge of Image and Video Retrieval, Lecture Notes in Computer Science*, Vol. 2383. Springer- Verlag, 2002, pp. 60-70.

[19] Chan Y., Harvey R. and Bagham J., "Using colour features to block dubious images," in *Proceedings of Eusipco 2000*, 2000.

[20] Chan Y., Harvey R. and Smith D., "Building systems to block x-rated," in *Challenge of Image Retrieval, BCS Electronic Workshops in Computing Series (CIR'99)*, Eakins J. and Harper D., Eds., Newcastle, 1999, pp. 34-40.

[21] Smith D., Harvey R., Chan Y. and Bangham J. A., "Classifying web pages by content," in *IEE European Workshop on Distributed Imaging*, Vol. 99/109, 1999, pp. 8/1-8/7.

[22] Jiao F., Gao W., Duan L. and Cui G., "Detecting adult image using multiple features," in *Proceedings of International Conferences on Infotech and Info-net ICII'01*, Beijing, China, 2001, pp. 378-383.

[23] Drimbarean A. F., Corcoran P. M., Cuic M. and Buzuloiu V., "Image processing techniques to detect and filter objectionable images based on skin tone and shape recognition," in *Proceedings of the International Conference on Consumer Electronics (ICCE)*, 2001, pp. 278-279.

[24] Cao L.-L., Li X.-L., Yu H.-H. and Liu Z.-K., "Naked people retrieval based on adaboost learning," in *Proceedings of the First International Conference on Machine Learning and Cybernetics*, Beijing, China, 2002, pp. 1133-1138.

[25] Jones M. J. and Rehg J. M., "Detecting adult images," HP Cambridge Research Laboratory, http:// crl.research.compaq.com/ projects/ vision/ adult_detection.html, 2002.

[26] Yoo S.-J., "Intelligent multimedia information retrieval for identifying and rating adult images," *Lecture Notes in Artificial Intelligence*, Vol. 3213, pp. 164-170, 2004.

[27] Yoo S.-J., Jung M. ho, Kang H. B., Won C. S. and Choi S.-M., "Composition of mpeg-7 visual descriptors for detecting adult images on the internet," *Lecture Notes in Computer Science*, Vol. 2713, pp. 682-687, 2003.

[28] Schettini R., Brambilla C., Cusano C. and Ciocca G., "On the detection of pornographic digital images," *Visual Communications and Image Processing*, Vol. 5150, pp. 2105-2113, 2003.

[29] Lin Y.-C., Tseng H.-W. and Fuh C.-S., "X-rated detection using support vector machine," in *16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP 2003)*, Kinmen, ROC, 2003, pp. 123-130.

[30] Zheng H., Daoudi M. and Jedynak B., "Adult image detection using statistical model," in *Compression et ReprÈsentation des Signaux Audiovisuels*, Lille, France, 2004.
[31] Zheng H., "Maximum entropy modelling for skin detection: with an application to internet filtering," Ph.D. dissertation, University of Lille, France, 2004.
[32] Zheng H., Daoudi M., Tombelle C. and Djeraba C., "Adult image filtering for internet safety," in *Multimedia Security Handbook*, Fuhrt B. and Kirouski D., Eds. CRC Press, 2004.
[33] Zheng H., Daoudi M. and Jedynak B., "Adult image filtering for web safety," in *Proceedings of 2nd International Symposium on Image/Video Communications Over Fixed and Mobile Networks*, Brest, France, 2004, pp. 77-80.
[34] Zheng H., Liu H. and Daoudi M., "Blocking objectionable images: adult images and harmful symbols," in *Proceedings of IEEE International Conference on Multimedia and Expo*, Taipei, Taiwan, 2004.
[35] Zheng H., Daoudi M. and Jedynak B., "Blocking adult images based on statistical skin detection," *Electronic Letters on Computer Vision and Image Analysis*, Vol. 4(2), pp. 1-14, 2004.
[36] Jinfeng Yang T. T. W. H., Fu Z., "A novel approach to detecting adult images," in *Proceedings of 17-th International Conference on Pattern Recognition (ICPR'04)*, Cambridge, UK, 2004.
[37] Zheng Q.-F., Zeng W., Wen G. and Wang W.-Q., "Shape-based adult image detection," in *Proceedings of the Third International Conference on Image and Graphics*, 2004, pp. 150-153.

AUTHORS

Peddi.Karthik, He completed M.Tech (CSE) in JNTUH College of Engineering, Jagtial, TG, INDIA. He has received B.Tech degree in Information Technology. His main research interest includes Image Procesing and Natural Language Processing.

Gajarla Santhosh, He completed M.Tech (CSE) in JNTUH College of Engineering, Jagtial, TG, INDIA. He has received B.Tech degree in Information Technology. His main research interest includes Image Processing and Cluster Computing.

Sri B.Kranthi Kiran Obtained his B.Tech(C.S.E) degree from JNTU, Hyderabad in the year 2000. M.Tech(C.S.E) From Osmania University, Hyderabad in year 2006. He is working as Officer-in-charge Academic Section(OAS) JNTUH College Of Engineering, Nachupally, Karimnagar from 1-10-2010 onwards.

Dr. P Sammulal, He did his PhD (Computer Science and Engineering) from Osmania University, Hyderabad, India. He has received the B.E degree from Osmania University in 2000 and MTech degree from JNT University in Computer Science and Engg., in 2002. He is in teaching and research since 2002. He has published 30 papers in International conferences/journals. His main research interest includes Parallel, Distributed, Cluster, Grid Computing concepts. He is working as an Assistant Professor in JNTUH College of Engineering (JNTUH University), Jagtial, INDIA.

