InFeRno - an Intelligent Platform for Recognizing Pornographic Web Pages

Sotiris Karavarsamis, Nikos Ntarmos, and Konstantinos Blekas

Computer Science Department
University of Ioannina
{sotiris, kblekas, ntarmos}@cs.uoi.gr

Abstract. We present InFeRno, an intelligent autonomous system to assess and arbitrate pornographic web page acquisition. In the processing pipeline, an adaptive Gamma correction method is applied on each assessed image in order to equalize excessive illumination, and to effectively harness skin segmentation. Based on the contour of each segmented image, a series of low-level features are extracted in the course of discriminating pornographic images. A multi-objective Support Vector Machine scheme for classifying images is proposed, which we evaluate experimentally. Our evaluation of the overall system demonstrates that the proposed system exhibits satisfactory results in discriminating pornographic web pages.

Keywords: skin segmentation, feature extraction, contour extraction, Multi-Objective Support Vector Machine

1 Introduction

Despite the usefulness and ease-of-access to a plethora of information scatted on the web, the Internet has become a hostile environment for unprotected people like children. Pornography is considered as sensitive information that is believed to be harmful for some groups of people. In the course of autonomously discriminating and blocking access to such content, Forsyth et al [1] was the first to devise a system comprising a figure grouper that inferred the existence of nude human figures. However, the proposed method required over 6 minutes to classify an image even on a workstation, thus hindering the potential of real-world applications. Wang et al [2] proposed a pornography elimination system called WIPE^{TM}, which employed Daubechies wavelet analysis and extraction of invariant central moments. Pornography filtering has also recently been a problem in web search, that is filtering out pornographic results in user search queries. For instance, Rowley et al [3] have proposed a method for identifying nude images which has been part of Google^{TM} safe search. Hu et al. [4] proposed a system to recognize pornography by classifying text and images. The system employs a combination of discrete and continuous text classifiers but also an image classifier to provide partial decisions over portions of a web page. The votes cast over parts of a web page are later fused in order to yield a page-wide content

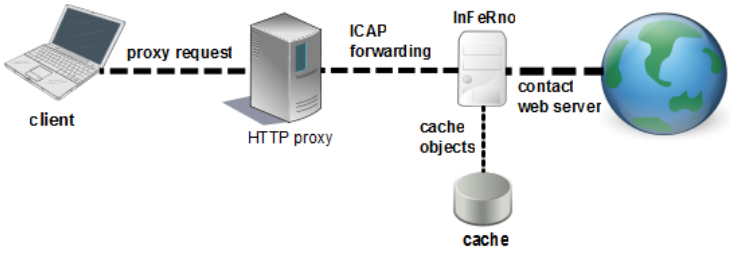

Fig. 1. InFeRno architecture

decision. In a similar sense, the POESIA project [5] is an open-source system for blocking pornography. It employs a combination of image and harmful-symbol filtering but also text classification via Natural Language Processing techniques. In this work we present InFeRno, an intelligent autonomous pornography elimination system employing extraction of low-level visual features solely from self-contained images. The contribution imposed by the proposed system is three-fold. First, an evaluation of a proposed multi-objective porn classification scheme based on the Support Vector Machine classifier is provided. We additionally evaluate a set of low-level visual features that exhibit state-of-the-art performance in the course of discriminating pornographic images when combined with our classification scheme. Our last contribution is the full-fledged implementation of our porn image recognition engine, which is integrated in an ICAP[1]-based web-proxy (see Section 2).

2 System architecture

InFeRno pornography elimination core is embedded in an ICAP server that is capable of communicating with any ICAP-enabled HTTP proxy software (see Fig 1). A request for a web page or image forwarded by a browser to the proxy infrastructure is in turn forwarded to the ICAP server in order to be assessed. InFeRno contacts the corresponding remote web server hosting the object, and fetches the source of the requested web page. Once the source is transferred, InFeRno examines the underlying Document Object Model (DOM) and fetches any image referenced in the DOM model. To reduce network overhead, InFeRno implements a caching mechanism so that assessed images and entire web pages need not be re-classified by the engine in subsequent requests. Compared to a direct connection without intermediate HTTP proxying, InFeRno is capable of serving user requests with an observed overhead of roughly 1 second per web page when moderate-sized images are present in requested web pages. A fair increase in the user's turnaround time is observed in web pages containing many

[1] ICAP stands for *Internet Content Adaptation Protocol* and is described in RFC 3507

images consisting of more than 1 million pixels due to a considerable increase of processing time in contour and feature extraction.
The proposed system is capable of performing in dual mode, that is it can either prohibit acquisition of an entire web page, or it can instead blur out only objectionable images contained in web pages. The system core can also be embedded in other interesting settings. For instance, it can be integrated in email scanner daemons, network firewalls, etc. Additionally, the system administrator can determine an acceptance threshold regarding the relative amount of pornographic content being permitted per web page (perhaps caused by image misclassification).

2.1 Image preprocessing and skin region detection

InFeRno employs a simple pixel-based technique which uses deterministic rules imposing relations between the R, G and B color channels of the visual signal, surveyed in [6]. In particular, the technique determines whether a pixel $p_i = (R_i, G_i, B_i)$ refers to skin if its color channel values satisfy the criterion's relationship rules. We have experimentally observed that the rule-based technique provides satisfactory performance over our dataset of pornographic, bikini and benign images. However, it provides many false positives when excessive illumination is present in the scene or objects with skin-like color are present. Thus, to further harness skin segmentation and improve overall performance, we pre-process each inspected image by using the adaptive Gamma correction method proposed in [7].

2.2 Contour extraction

In order to extract the contour of human nudes, we follow the methodology explained in [8]. To approximate the contour of nude entities, we employ a region splitting scheme that iteratively partitions the image plane in quadrants based on a region splitting criterion. Initially, the image plane is partitioned into four equal quadrants. Subsequently, the corresponding skin and non-skin intensity histograms of pixels present in each quadrant are computed. Given the intensity histograms, we can compute the following statistics: a) *skin to non skin ratio*, and b) *kurtosis measure*, which can effectively describe the flatness properties of the skin and non-skin pixel histograms. Based on the aforementioned metrics, we decide if a region shall be further split by introducing thresholds for each measurement. The splitting procedure can be repeated an arbitrary number of times on each quadrant. The outer corners of the obtained connected split-regions of the last iteration step are regarded as control points of the convex hull that localizes a Region of Interest (ROI) in the image.

2.3 Feature Extraction and Classification Scheme

Based on the contour of human figures described in the previous section (see Section 3), we obtain the following set of 4 features from the ROI:

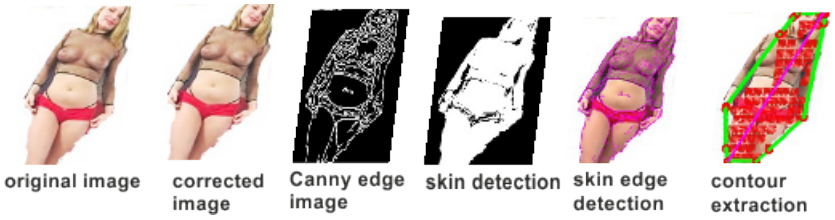

Fig. 2. the processing steps a suspected image undergoes until classification

- mean and variances of the RGB signal color channels (6 constaints in total) of all non-skin pixels in the obtained contour. These features can better describe the spatial properties of the detected non-skin pixels, which aids to reducing negativeness of falsely classified non-skin pixels.
- ratio of the total skin and non-skin pixels delimited by the contour, which is a useful feature for discriminating benign bikini images and illicit nude images
- The 7 invariant spatial Hu moments [9] of the surface delimited by the obtained contour. These features provide a strong cue for discriminating porn and face images. They can also describe the invariance of a contour with respect to standard geometric transformations like rotation, scaling and translation.
- angle of the diameter of the convex hull delimiting a ROI in the image, which is a cue that describes the posture of the obtained ROI in the image.

The feature vector constructed by the aforementioned features consists of a total of 15 numerical constraints. Based on this vector and our database of features obtained from our manually assembled dataset of 660 pornographic images, 700 bikini images and 4320 assorted benign images, we employ a combination of three Support Vector Machine classifiers [10] trained using linear kernels in a majority-vote scheme. The proposed classifier fusion scheme strengthens the confidence with regard to labeling each image. In order to evaluate the performance of the image classifier, we randomly partition our dataset in a training set comprising 75% of the total features and a testing set comprising 35% of the total features. The porn, bikini and benign SVM classifiers exhibited a cumulative false rate of 2%, 3% and 1.2% respectively. In turn, the entire image classification scheme demonstrated a cumulative false rate (considering false positives and false negatives) as low as 1%. The latter poses an indication that the multi-objective classification model is a strong model, and especially when combined with the aforementioned set of features.
In our future work one of our main concerns will be researching the effects of taking advantage of companion textual information contained in web pages as prior knowledge in the process of classifying images and subsequently web pages (see, for example, the works [5] and [4]).

References

1. M. Fleck, D. Forsyth, and C. Bregler, "Finding naked people," *Computer VisionECCV'96*, pp. 593–602, 1996.
2. J. Ze Wang, J. Li, G. Wiederhold, and O. Firschein, "System for screening objectionable images," *Computer Communications*, vol. 21, no. 15, pp. 1355–1360, 1998.
3. H. Rowley, Y. Jing, and S. Baluja, "Large scale image-based adult-content filtering," in *International Conference on Computer Vision Theory and Applications*, 2006.
4. W. Hu, O. Wu, Z. Chen, Z. Fu, and S. Maybank, "Recognition of pornographic web pages by classifying texts and images," *IEEE transactions on pattern analysis and machine intelligence*, pp. 1019–1034, 2007.
5. "Poesia content filtering." http://www.poesia-filter.org/.
6. V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques," in *Proc. Graphicon*, vol. 3, Citeseer, 2003.
7. Y. Shi, J. Yang, and R. Wu, "Reducing Illumination Based on Nonlinear Gamma Correction," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 1, IEEE, 2007.
8. J. Yang, Y. Shi, and M. Xiao, "Geometric feature-based skin image classification," in *Proceedings of the intelligent computing 3rd international conference on Advanced intelligent computing theories and applications*, pp. 1158–1169, Springer-Verlag, 2007.
9. M. Hu, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 179–187, 1962.
10. C. Bishop and SpringerLink, *Pattern recognition and machine learning*, vol. 4. Springer New York, 2006.