Privacy Protection in an Electronic Chronicle System

Simon Greiner[1,2] and Jie Yang[2]
[1] Department of Computer Science, University of Karlsruhe(TH), Germany
[2] School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
ulbdc@stud.uni-karlsruhe.de, jie.yang@cs.cmu.edu

***Abstract*—This paper describes our preliminary work on privacy protection in video for supporting automatic dietary assessment in obesity studies. We present an approach to protect people's identities and contents on computer screens using object detection techniques. We use the Adaboost algorithm implementation from the OpenCV framework to build a system that allows detection of faces and screens in order to obscure them and make them unrecognizable in the recorded images. We also use some post processing methods to improve the detection accuracy. Our preliminary results show encouraging results. We are currently improving the accuracy and robustness of the system towards a working system.**

I. Introduction

The advance in hardware and multimedia technologies has made it possible to build up a system that can deal with capturing, representation, organization, analysis, and presentation of temporal streams of data, captured by a patient with body mounted devices in "free-living" conditions for more accurate medical studies. For example, we are developing a unified sensor and electronic chronicle (e-chronicle) system for ubiquitous "free-living" data acquisition and management in the study of obesity. The device will integrate multiple sensors such as a microscopic video camera, an accelerometer, an oximeter, a semiconductor thermistor, a microphone, etc [1]. The device can be used to capture eating/drinking activities as well as physical activities for obesity patients. Because the video camera is configured to record the same scene as the patient's perspective, we have to address problems of privacy protection for both the patient and other people. For example, the device may capture other people in the scene or the patient's own computer screen when he/she is using a computer. In this paper, we present our ongoing research of privacy protection in videos to deal with these problems.

The problem of privacy protection in videos is poorly defined. Privacy could mean different things to different people. In this research, we adopt the common approach in practice that protects the identity of people and the content of objects from being recognized during the video playback. More specifically, we develop techniques to remove people's identities by face obscuring, as commonly seen in television, and to protect the content of computer screens by masking the screens when the screen is close enough (e.g., with a certain size). Fig. 1 illustrates the proposed concept of our system. Fig. 1(a) shows the results of privacy protection for people and Fig. 1(b) displays the results of content protection for computer screens. In order to obscure human faces and to mask computer screen, we have to robustly detect human faces and computer screens in images. Thus, the problem becomes an object detection problem in images/video. In the rest of the paper, we present the algorithm for object detection (Section II), describe implementation of the system and report the preliminary results of the experiments (Section III), and discuss future work (Section IV).

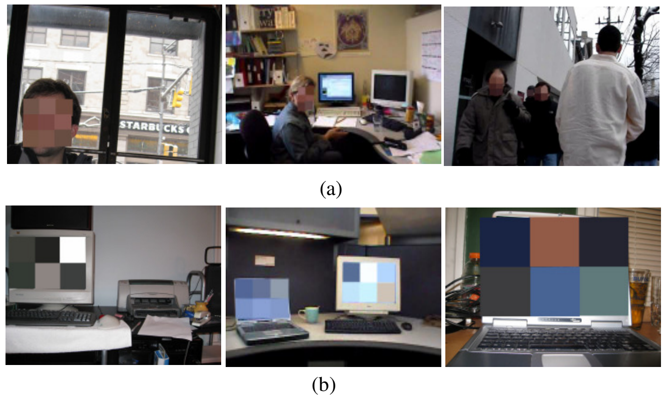

Fig 1 An illustration of privacy protection for people (a) and screens (b)

II. Object Detection

Object detection is a fundamental problem in computer vision and pattern recognition: Given an image, to determine whether or not the object is present, and, if present, determine the location and size of each object. A well studied and common used approach for object detection is the Adaboost algorithm introduced by Viola-Jones [2] and extended by [3]. The basic idea of the Adaboost algorithm is to build a series of so called weak classifiers that perform slightly better than guessing in classifying. They are represented by connecting rectangles, which define a set of pixels. These pixels' values are summed up. The difference of the sums of the rectangles is then used to separate positive from negative samples, e.g., faces from non faces. By "boosting" a weak classifier as described in [2], a strong classifier is calculated. This strong classifier can be trained using arbitrarily good false alarm rate and hit rate on training data.

We then can use the trained classifier for an object detection task. To detect objects, the input image and its sub images are forwarded to this classifier, which decides whether it contains the trained object or not. According to the position of the sub image, the position of a detected object within the image is found.

III. System Implementation

We use the OpenCV framework, an open source library maintained by the Intel Corporation, as a tool to implement our system. The OpenCV provides essential data structures and algorithms for computer vision applications. Among these algorithms, Viola-Jones algorithm is implemented and provides training and detection tools such as cvhaartraining.exe and performance.exe. We used these libraries and tools to train the classifier, detect objects in images and test our results.

A. Face Detection

For training a face detector, we used an optimized face database from other researchers [4]. To speed up the training process, small subsets of this large collection were used. Three classifiers will be compared later. Class2000, Class5000 and Class10k were trained with 2000, 5000, 10.000 face images and 2000, 5000, 5000 images without faces respectively.

For evaluation, we used the three collections with upright faces from the CMU+MIT test set for testing the classifiers quality, as most other researchers do. The evaluation dataset contains 191 images with 517 faces.

The result of these tests is displayed as a roc curve. It describes the amount of falsely positive labelled objects on the x-axis and the percentage of correctly detected objects on the y-axis. The more the curve orients to the top left corner, the better is the classifier. Fig 2 shows the promising results for our classifiers, as they become better, the more samples are used and as [4] reports very good results using this set for training.

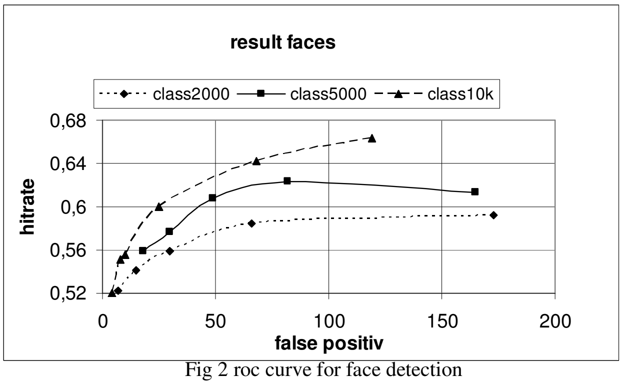

Fig 2 roc curve for face detection

B. Screen Detection

While face detection is a very well researched field, screen detection is nearly not mentioned in literature. Our first attempt is to use the same procedure as we used for faces. Therefore, we collected 1000 screen images and 2800 no-screen images on the internet. A preliminary classifier was trained using this collection. The result is a classifier that found a high number of false positives when detecting a reasonable number of objects correctly.

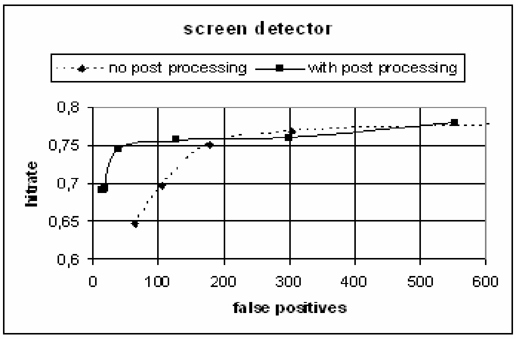

Fig 3 roc curve for screen detection

To improve this result displayed in Fig 3, the detected objects are filtered by post processing. Only those objects are considered as correct, which are also found in two of the three pervious or following frames.

For testing a short video recorded in an office environment was used. This sequence consists of 239 frames containing 314 screens. Fig 3 shows that most of the correctly detected screens were kept while dismissing many false positives during post processing.

The post processing allows a hit rate (correctly found screens divided by all found screens) nearly as well as the classifier alone, but has far less false positives.

IV. Future Work

Our preliminary work shows that object detection can be used for protecting privacy in videos. Although the described approach can provide robust detection of objects concerning privacy, it needs further improvements for real working systems. The classifiers need to be improved, especially for occluded and rotated objects.

For more complete privacy protection, we will work on detecting other objects that may reveal privacy of people, such as text in the scene, and mask them accordingly.

Acknowledgements

The research was supported by NIH under Genes, Environment and Health Initiative (GEI) (Gant Number: 1U01 HL091736 01).

The first author was partly founded by the Landesstiftung Baden-Wuertemberg and an InterAct Scholarship. The work was performed at Carnegie Mellon University.

References

[1] N. Yao, R. Sclabassi, Q. Liu, J. Yang, J. Fernstrom, M. Fernstrom, and et al., “A Video Processing Approach to the Study of Obesity,” Proceedings of IEEE ICME 2007, pp. 1727-1730, 2007.

[2] P. Viola, M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features”, CVPR, 2001

[3] R. Lienhart, J. Maydt, „An extended set of haar-like features for rapid object detection“, IEEE ICIP, pp. 900~903, 2002

[4] J. Chen, X. Chen and W. Gao. “Expand Training Set for Face Detection by GA Resampling,” 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 73-79, 2004.