L. Sweeney and R. Gross. Mining Images in Publicly-Available Cameras for Homeland Security. AAAI Spring Symposium on AI Technologies for Homeland Security. Palo Alto, 2005.

# Mining Images in Publicly-Available Cameras for Homeland Security

**Latanya Sweeney and Ralph Gross**

Data Privacy Laboratory, School of Computer Science
Carnegie Mellon University, Pittsburgh, PA 15213-3890
{latanya, rgross}@cs.cmu.edu

**Abstract**

A dramatic increase or decrease in the number of people appearing at a location can be an indicator that something has happened that may be of interest to law-enforcement, public health, or security. This work demonstrates how low quality camera images can be used to automatically alert when an unusual number of people are absent or present at a location. We report on experiments using publicly available, inexpensive cameras already operational over the Web. A "historical database" (H) for each camera is constructed by capturing images at regular time intervals and applying a face detection algorithm to store the number of faces appearing in each image ("face count"). Later, given an image *X* having timestamp *t*, if the face count of *X* is significantly higher or lower than the expectation inferred from H for time *t*, an unusual number of people are considered to be present in the image.

## Introduction

Publicly available webcams can be used in an automated alert system to signal when an unusual number of people are absent (or present) at a location. Imagine an absence of people on Wall Street during what should be a typical work day. Or, alternatively, consider a large number of people assembling unexpectedly. Are these the results of acute events (e.g., a terrorist attack or a local accident)? If the number of people involved is very large, then the cause is easier to determine. But in cases where the number of people is statistically significant but visually subtle, then what may be an early warning sign could go undetected by law-enforcement, local security, or public health. Many authorities have an interest in unexpected changes in the numbers of people appearing in public locations. The work described herein seeks to provide an automated means of detecting such changes.

### Vision for Bioterrorism Surveillance

One motivation for early detection of people missing from public locations is bioterrorism surveillance. Aerial releases of many biological agents are of concern to the Centers for Disease Control and Prevention (CDC); among these is anthrax, which has signs and symptoms reminiscent of the flu [2004]. Interviews with the family members of those exposed to an accidental release of anthrax in Sverdlovsk in 1979 provide evidence that being absent from work and public spaces tends to be an early behavior of those exposed [Guillemin, 1999]. In the days immediately following exposure, people sought home remedies. Only in latter stages was professional assistance sought, leaving public health authorities unaware and the contaminated sites publicly accessible. Detecting whether significant numbers of people are absent from work is important because doing so can save a large number of lives in the case of biological agents [Mostashari and Hartman, 2003]. A goal in this work is to determine by automation whether a significant number of people are absent from public spaces. We provide a prototypical system that reports how many people appear at (or are absent from) a location compared to what is normal for the time of day, day of week, and season for that location.

### Publicly-Available Webcams

In 2003, Sweeney developed tools that mine the World Wide Web ("Web"). for web addresses (or "URLs") of on-line cameras. Found URLs were stored im a database named CameraWatch More than 6,000 webcams showing people in public spaces around the USA were found. These cameras primarily show sidewalks, highways and places of interest. . A sample of 200 camera URLs from the 6,000 URLs in the CameraWatch collection is available in a searchable on-line format. Because these cameras tend to view spaces where people are located, they pose a unique opportunity to identify unusual activity in the general population. Yet, the images tend to be of low quality thereby challenging established approaches to computer vision. In this work, we perform automated review of images from these low-end cameras.

These webcams were not placed on-line by us, but by a diverse collection of people, organizations and government entities for a variety of purposes. This work demonstrates how such an amorphous collection can be used to acquire knowledge about the population.

Some cameras are installed for surveillance purposes. The work reported herein also works on images captured on private surveillance camera networks, e.g. closed caption television (CCTV).

1. A "historical database" (**H**) for a camera is constructed by capturing images at regular time intervals throughout each day and applying a face detection algorithm to each image captured. The number of faces appearing in each image ("face count") along with a date and timestamp are stored.
2. Averages and standard deviations are computed for each time interval in **H** over varying classes to which the date and time stamp belongs. Date classes include current weekdays/weekend, day of week, time of month, and seasons. Holidays and events are a special class.
3. Given an image *X* having timestamp $t_X$, if the face count of *X* is more than 2 standard deviations above the average (or 2 standard deviations below the average) of any class to which $t_x$ belongs, then an unusual number of people are considered to be present in the image.

**Figure 1. Overall approach for detecting whether an unusual number of people are appearing in a camera image.**

## Methods and Results

Hutwagner et al. at CDC proposed an aberration detection model for bioterrorism surveillance of hospital data [2002]. Deviations in current data are compared to a historical average. Our approach builds on this notion. In the initial phase a historical database of images is constructed from which averages are computed (steps 1 and 2 in Figure 1). Then, in an operational phase, a decision is made about the number of people appearing in a given image based on shifts away from the mean (step 3 in Figure 1).

### Materials

In conducting experiments, we used: (1) the CameraWatch database; (2) a specific webcam; and, (3) a face detection program. Each of these are described below.

**CameraWatch Database and The Webcam.** The CameraWatch database contains more than 6,000 URLs of publicly available webcams [Sweeney, 2003]. One URL, www.earthcam.com/cams/newyork/timessquare/index.html, was selected from CameraWatch as the subject of these experiments and is termed "the Webcam." The Webcam views Time Square (1552 Broadway at 46th Street, New York, NY 10036). Figure 2 provides an example of an image from the Webcam.

**Face Detection Program.** Given an image, Schneiderman's face detection program ("Face Detection Program") reports the number of faces found by exhaustively comparing varying size model faces against all pixel positions [2004]. The Webcam was selected in part because there is sufficient light, day and night, for the Face Detection Program to operate.

**Historical Database.** Images from the Webcam were captured every 10 minutes, 24 hours a day, for 4 work days in the winter (December 6, 7, 8 and 10, 2004). It rained for 2 of the 4 days. The Face Detection Program was applied to each image and the results stored, along with the date and timestamp of the image capture.


**Figure 2. View from a publicly available webcam used to answer the question: is there an unusual number of people absent from (or present in) public spaces? Ten faces were automatically detected in this image.**

### Experiment 1: Average Face Counts

Figure 3 shows the average number of faces detected for each time interval in the Historical Database. Someone is almost always present and the most number of people are present at lunch and dinnertime. 3am has fewest people. Averages for rain and non-rain days are smoother, less noisy curves of the same shape as in Figure 3.

### Experiment 2: Detected Faces and Actual People

A manual count of the actual number of people appearing in each image was done. People whose heads (not necessarily faces) appear in the store or on the sidewalk up to the mailboxes are included. See Figure 2, which has a manual count of 11 people. A linear regression between the number of faces detected in the Historical Database and the actual number of people counted in the earlier part of the day (no rain) has $R^2$=0.76 (best is 1). Performance deteriorates with rain (perhaps due to occlusion by umbrellas and crowding under umbrellas) and other confounders. No performance difference was found between daylight and street lights.

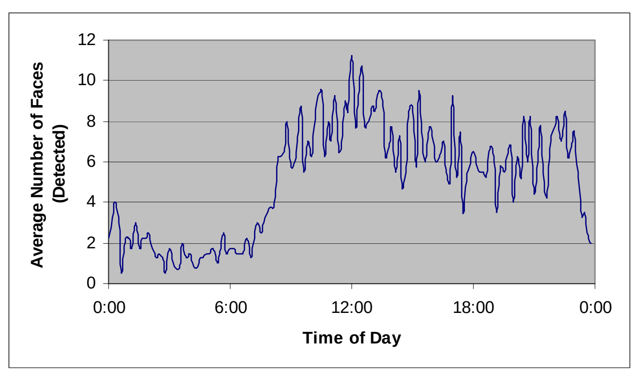


**Figure 3. Average number of faces detected in webcam images (see Figure 2) during a standard weekday from one week in the winter. Noon (12:00) is the most popular time and 3am (3:00) is least popular.**

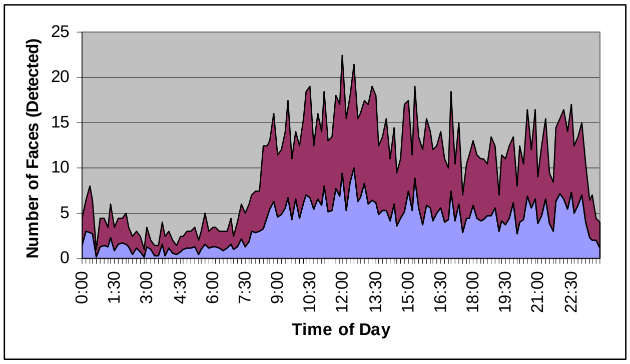


**Figure 4 Range of values considered "normal" for the time of day (middle area). Values above (top) or below the middle area (bottom area) are considered "unusual."**

### Experiment 3: Alarm Thresholds

The range of values considered "normal" for the Historical Database for each time interval is computed by taking the averages (Figure 3) and adding and subtracting 2 standard deviations from each. Figure 4 shows the resulting range of values. If the number of people typical at noon appeared in a camera image captured at 3am (or vice versa), the population would be considered "unusual" for that time period being above (or below) the range of normal values.

## Related Work in Computer Vision

One might wonder whether face recognition would improve results. Several large scale evaluations of face recognition systems have shown that the performance of face recognition systems is most affected by face pose and illumination, and to a much lesser degree by image resolution and image quality [Gross et al., 2001; Phillips et al., 2003; Blackburn et. al, 2000]. Blackburn et al. reduced the distance between the eyes from 60 pixels to 15 pixels with only minimal performance degradation —from 94% to 92% rank 1 (top choice) accuracy for the best performing system. They reduced image quality through increasing JPEG compression up to 40:1. Rank 1 accuracy decreased moderately (from 63% to 56%). Other imaging conditions such as face pose and illumination were kept constant. Images captured in unconstrained conditions with low-quality cameras are currently still beyond the capabilities of face recognition systems.

A number of research groups have addressed the problem of monitoring traffic, either on a highway or around city intersections. However, these systems require high quality image input and/or significant user interaction at system startup time. Beymer et al. [1997] proposed a system to track vehicles under congestion, but for the tracking algorithm to work the user has to manually specify a set of camera-specific parameters. Eikvil et al. [2001] proposed to use Hidden Markov Models to perform traffic surveillance. Again the system requires a supervised initialization phase.

## Discussion

While this work provides a proof of concept that webcams can report whether there is an unusual number of people at a location, much more work is needed to validate, generalize and improve these findings. For example, counting faces as a proxy for people is problematical. One can imagine a camera view in which the backs of people (not their faces) are usually visible. Constructing a "person detector" rather than using a face detector seems prudent.

In concluding, we address privacy. With almost 6,000 publicly-available web cams already showing public spaces around the USA, this approach offers a low-cost national surveillance system. But it offers it to anyone in the world. Not only is that a concern of personal privacy, but perhaps also of national security. On the other hand, there may be many worthy uses for this network of webcams. For this reason, we introduce the notion of "smart cameras" in which high-level information from webcams, and not actual images, is shared. We envision that smart cameras will include programs within, such as this work, and would provide reports, updates and alerts, and not images.

## Acknowledgements

Much gratitude goes to Sylvia Barrett and Katya Mirochnik for aid in reviewing and capturing images, and to Henry Schneiderman for the use of his face detection program.

## References

Beymer, D., McLauchlan, P., Coifman,, B. and Malik, J. A Real-Time Computer Vision System for Measuring Traffic. *IEEE Conference on Computer Vision and Pattern Recognition*, 1997.

Blackburn, D. M. and Bone, M. and Phillips, P. J. *Facial Recognition Vendor Test 2000*: evaluation report, 2000.

CDC, 2004, <www.bt.cdc.gov/agent/agentlist.asp>.

Eikvil, L. and Huseby, R. Traffic Surveillance in Real-time using HMMs. *Scandinavian Conference on Image Analysis*, 2001.

Gross, R. and Shi, J. and Cohn, J. Quo Vadis Face Recognition? *3rd Workshop on Empirical Eval Methods Computer Vision* 2001.

Hutwagner, L., Seeman, M., Thomson, W., Treadwell, T. CDC: Early Aberration Reporting. *Syndromic Surveillance Conf.* 2002.

Guillemin, J. *Anthrax: The Investigation of a Deadly Outbreak.* Dec. 1999. Univ. of California Press.

Mostashari, F. and Hartman, J. Syndromic surveillance: a local perspective. Journal of Urban Health, 80 (2) 2003.

Phillips, P.J. and Grother, P. and Ross, J. M. and Blackburn, D. and Tabassi, E. and Bone, M. *Face Recognition Vendor Test 2002*: evaluation report, 2003.

Schneiderman, H. Learning an Restricted Bayesian Network for Object Detection. *IEEE CVPR*, 2004.

Sweeney, L. *CameraWatch*. Data Privacy Lab, 2003. <privacy.cs.cmu.edu/dataprivacy/projects/camwatch/>