Multi algorithm object tracking for effective surveillance camera operation.

Sander Tiganik
Institute of Computer Science, University of Tartu
Supervisor: Artjom Lind

April 25, 2016

Abstract

This article gives insights into how one might create their own surveillance camera program, which uses storage space efficiently by only recording events that are of interest to the user (I.e movement), thereby reducing storage space required.

1 Introduction

These days we are surrounded by cameras. Whether they are in out pockets, backpacks or mounted on a house wall, we never seem to be able to evade the electronic eyes. As time passes cameras also not only get smaller and better they also get smarter. But all this progress comes with a downside. As cameras get better and the video feeds they create get clearer and sharper, so does the requirement for storage space to save the captured images. That is why when you look at security camera footage it usually look very low detail and has very low FPS (frames per second). That is because surveillance camera owners usually don't want to spend an enormous amount of money on storage space and also sometimes have legal obligations to store the recordings for a certain amount of time (I.e shops and stores may be required to store recordings for a certain amount of days by law). Now if those recording would be encoded in regular video encodings, have a full high-definition resolution and a 25 FPS framerate, then we are talking about potentially Terabytes of information from a one week period. This article proposes a way to decrease that amount drastically, by only taking high-definition video if there is something of interest on the footage (I.e movement) and take regular surveillance camera video footage the rest of the time, thus both increasing the quality and usefulness of the recording and also drastically decreasing the amount of space required to store the recordings.

2 Idea

When it comes to object tracking, there are a lot of different algorithms for achieving it. Unfortunately each of the algorithms has certain strengths and weaknesses. Be their weakness changing light conditions, moving camera or only high contrast detection, we can not allow these faults to create a situation where important video segments are not recorded because the algorithm had faults.

In this article we propose that when we layer multiple algorithms on top of each other and then compare the results of the algorithms to determine moving objects, we get a much more accurate result than one might get using only a single movement detection algorithm. We are going to concentrate on a two algorithm system that uses background subtraction and a gaussian mixture model to determine movement on each frame of the video feed. However this approach is not limited to just two algorithms. One can use an arbitary amount of algorithms to find the optimal setup for detecting movement in sequential frames.

3 Background subtraction

Background subtraction is the most basic of object tracking techniques. Although primitive it still yields satisfactory results as long as certain criteria about the environment don't change, such as lighting conditions and shadows.
It works by taking the first frame of the video feed and converting it to grayscale (only gray colors) as a reference. It can also be any other frame, as long as there is no movement in the image. As time progresses and more frames are processed, each of those new frames is converted to grayscale and subtracted as absolute difference from the original image. All the pixels that are not 0 (0 means that there is no difference between the two pictures) are objects that have moved into the line of sight of the camera and need to be tracked. [1] That's how it would work in a perfect world.

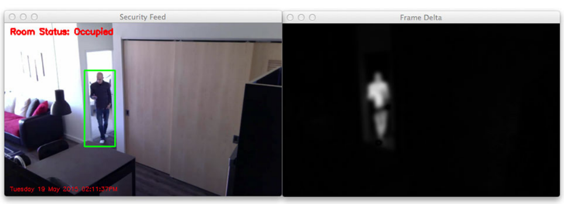

Figure 1: Video frame - left. Absolute difference of video frame and reference frame - right. [2]

Unfortunately background subtraction has a big flaw when light conditions and shadows change, because they will also show up as movement. This flaw can be somewhat repaired by periodically taking new reference pictures as time progresses, making the background subtraction algorithm more stable. Still it suffers from false positives and object shadow highlighting. This means that background subtraction will often highlight areas of the picture where there might not actually be anything (light condition glitch) or highlight shadows of objects passing by (since the area of the shadow also changes in color value). All these flaws combined make the background subtraction algorithm unreliable when it comes to security footage.

4 Mixture of Gaussian values (MOG)

The second algorithm in our system is another background subtraction algorithm, but this one is a bit more complicated. The algorithm is called background subtraction using mixture of gaussian distribution values. It's base principle is the same as with the background subtraction, but it has a little added on to it. It works by using Gaussian distribution (mathematical statistics) to calculate the probability of a pixel being observed. The higher the probability, the higher is the chance that the pixel is a background pixel. If the probability of the pixel is low, it is most likely a new pixel. [1] This approach has the positive side of correcting itself and adapting to different environments, since it uses new frames to update and fix the current distribution models. The algorithm has many parameters that can be tweaked to achieve maximum efficieny in detecting objects.

Figure 2: Original frame. [3]

Although this approach also has it's flaws. The algorithm uses complicated math to calculate the objects that are moving across the screen and though it is very accurate in recognizing objects it may not always highlight the entire object on the screen. This is why the last and final step of unification of found objects is required for the proper functioning of our system.

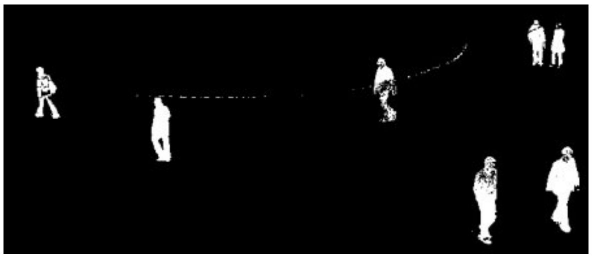

Figure 3: Background substraction using mixture of gaussian distribution values. [3]

5 Final Result

The final result is comprised of both object tracking methods. The way to do it, is to run both algorithms separately on the input frame being processed and detect all the objects possible. After that comes the unifying process where one takes objects from both algorithms and compares them to one another to find objects that overlap. If there are overlapping objects there is a good chance that they are part of the same object, so we use the contour found by the background subtraction (it is more sensitive and has a bigger object tracking contour) and draw it on to the frame. Also while there are objects detected we enable high-definition video capturing.

6 Possible future advancements

As computing power of computer systems increase, we will be able to use ever more complicated algorithms and also combinations of algorithms, in order to decrease the amount of faults created by the system. Eventually we will see fault free systems, that make mistakes so rarely that it is considered an abnormality rather than a bug. When object tracking gets better we can combine it with object classification and data mining to gather information about objects while the footage is still being recorded for even more valuable and useful results for interested parties. At one moment it will most likely be impossible to escape the electronic eyes surveilling our every day lives and while for some this might be a scary notion the author finds that the benefits of this kind of information recording could largely outweigh the small sacrifices we would have to make in privacy.

7 Conclusion

By combining multiple object tracking algorithms we lower the amount of false positives the system gives us, therefore allowing us to rely on the system to record all important occurences in high-definition video while recording all unimportant background footage in low quality and low FPS. This means that the amount of storage space required to make surveillance recordings is drastically smaller than recording the entire footage in high-definition, and that the footage is still very information rich and useful to anyone who needs it (E.g security guards, law enforcement).

References

[1] M. Piccardi ; Comput. Vision Group, Univ. of Technol., Sydney, NSW, Australia, Background subtraction techniques: a review, Systems, Man and Cybernetics, 2004 IEEE International Conference on, 3099 - 3104 vol.4 , DOI: 10.1109/ICSMC.2004.1400815

[2] Adrian Rosenbrock, Basic motion detection and tracking with Python and OpenCV, last visited: 25.04.2016, address: `http://www.pyimagesearch.com/2015/05/25/basic-motion-detection-and-tracking-with-python-and-`

[3] OpenCV documentation: Background Subtraction, last visited: 25.04.2016, address: `http://docs.opencv.org/3.1.0/db/d5c/tutorial_py_bg_subtraction.html#gsc.tab=0`