An Automated Paintball Turret Target Tracking Algorithm

Stéphane Franiatte, Alain April
Software Engineering Department
École de Technologie Supérieure,
Montréal, Québec, Canada
stephane.franiatte.1@ens.etsmtl.ca, alain.april@etsmtl.ca

***Abstract*-Video target tracking is a difficult problem to solve. Difficulties particularly arise, first, during sudden changes in the trajectory of a tracked object, then if changes occur in the appearance of the object or scene and also during occlusions or camera movements. The proposed algorithm has been designed to resolve the problems that occur when tracking a player during a paintball battle. The solutions we applied to this particular problem segment potential targets effectively regardless of lighting conditions, decide which target to follow and, finally, anticipate its trajectory. Techniques that we have implemented successfully include an adaptive Gaussian mixture model for background subtraction. Image subtraction is an effective segmentation solution to the problem of light variation since the images captures correspond to a very short interval of time. Mathematical morphology is then used with two objectives: 1) to eliminate the residual noise, and 2) to assemble blobs belonging to the same object. After the segmentation step, a two-step blob analysis is performed. At first, the blobs that have an area outside preset limits are discarded. Then the blobs whose aspect (bounding box aspect ratio) does not match the shape of a standing human are also discarded. The remaining blobs are tracked frame by frame. Since, by design, the turret can only shoot at one target at a time, the blob having the maximum area is selected as a target. Its position in x and y coordinates (camera coordinates) is then provided to a Kalman filter, which anticipates the target's trajectory, allowing the needed time for the turret to adjust itself before shooting.**

***Keywords*—Video target tracking; MoG; Kalman Filter; Blob detection**

I. INTRODUCTION

Integral part of the field of computer vision, intelligent video surveillance (IVS) is a very active research topic. The goal of IVS is to process video streams in real time for the purpose of detection, tracking, anticipation of trajectory and even behavior understanding. It has many applications such as site security, public transportation surveillance and traffic monitoring.

A good video tracking algorithm must resolve many issues and problems. First of all, the environment where the video is captured is often uncontrolled. Therefore, one has to have very strong a priori expectations about the environment in which the system will operate, to achieve a set level of efficiency.

Secondly, objects captured in a video may be hard to follow, because of occlusions, overlaps or sudden changes in the trajectory of the object. However, in the particular task of following targets during a paintball game, a few assumptions can be made, making the problem of IVS slightly easier to resolve.

A first assumption is that the lighting conditions are expected to vary widely from day to day and from playing fields locations. This led us to a first choice of trying a Gaussian mixture model image subtraction (MoG), for the image segmentation task. This technique does not depend upon any kind of initialization and is insensitive to lighting variation due to its very short time span criteria.

A second assumption, in our case study since a paintball field is expected to host only humans participating in a game, each of the detected objects is assumed to be a potential target. Hence, there is little decision to take about whether or not an object that is moving is a potential target, making the whole process faster.

The third and last assumption is that the tracked objects are expected to follow a fast, linear path on uncovered part of the field. Because of their high speed, the target path needs to be anticipated, allowing time to the turret to adjust itself before shooting. These kinds of linear path are quite easy to predict using a Kalman filtering (KF) technique, which is the choice we made for the target path prediction part of our algorithm.

Our proposed algorithm is divided in three successive steps: 1) preprocessing; 2) blob detection and tracking; 3) target position prediction (see figure 1).

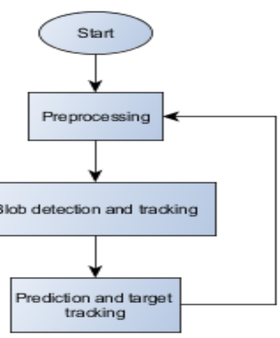

Fig. 1. Algorithm flowchart

II. Preprocessing

Many segmentation algorithms have been proposed for object detection in IVS applications. They rely on different assumptions : statistical models of the background, minimum and maximum values or adaptivity [1]. The choice we made is a combination of frame differences and statistical background models, an algorithm developed by Zoran Zivkovic [2]. His mixture of Gaussian (MoG) is particularly well suited to address the lighting variation on the paintball field; the models are constantly updated for each pixel over a reasonable time period, and new objects entering the scene are considered to be part of the background only if they remain static long enough, a situation that rarely happen in a paintball game.

At the end of the preprocessing step of our algorithm, two basic mathematical morphology operators are applied to the image. The first one is the erosion operator; the kernel used is of dimension3×3. The goal here is to get rid of the residual noise still present in the image. Then a dilation operator is applied. This time, the goal is to assemble separated blobs which are probably part of the same object; thus, the kernel is this time of dimension7×3. The kernel height / width ratio is chosen to be equivalent to the one of a standing human, guaranteeing that the dilated shape will still be roughly the same as the target's true shape.

III. Blob detection and Tracking

After the segmentation step, a component-labeling algorithm is performed. Its results are then provided to a blob tracking algorithm based on the work of Senior et al [3]. According to the authors, this algorithm is able to track objects which are partially or even fully occluded, using appearance models. These models are used to localize objects during partial occlusions and detect complete occlusions. The algorithm assigns an ID to each blob, and then try to follow them, frame after frame.

The target is chosen among the tracked blobs. Once a tracked blob lifetime is longer than twenty frames in a row, its ID is memorized and assigned to the target (see figure 2).

IV. target tracking and Next position prediction

The target path is followed and predicted by a Kalman Filter. This algorithm uses a series of measurements, sampled over time, which will constitute the history of the system. The algorithm maximizes the a posteriori probability of those previous measurements. It will then be possible to build a model to predict the next state of the system [4].

First introduced in 1960, the Kalman filter has many applications in a wide variety of signal processing contexts, from spacecraft guidance to econometrics. We use it in the context of the paintball turret to predict a physical object's future location. Because of its underlying linear system model, to use the Kalman filter, the object must be moving at constant velocity or constant acceleration [5]. In our case, this holds true most of the time.

The Kalman filter algorithm involves two steps, prediction and correction/update (see figure 2). The first step uses previous states to predict the current state, while the second step uses the current measurement of the target location to correct the state [4]

In case the target is lost for more than 20 frames in a row, it is considered to be definitively lost. A new target is then chosen among the tracked blobs (see figure 2). The actual and predicted target locations are provided to a state-machine, which takes decisions about shooting.

V. Conclusion and Future work

The whole algorithm is yet to be evaluated in real conditions. So far it has only been tested on recorded video, giving promising results in the first performance tests inspired by the work of Nascimento and Marques [1].

Future work may include machine learning to the project. So far the algorithm holds on the assumption that every detected blob is a potential target. This might not always be true; for instance, users of the turret may want it to shoot only on players from the opponent team. It would also be interesting to do gesture recognition, which would act upon the turret behavior; for instance, holding hands up, as a surrender, would make the turret to cease fire.

Since Kalman filter's underlying model of the probability distribution for its hypothesis is unimodal Gaussian, it is not possible to represent multiple hypotheses simultaneously [5] and thus to predict the path of a target that behaves erratically or that is occluded for a certain amount of time. It could then be interesting to replace the Kalman filter used in our algorithm by a Condensation algorithm, which can make multiple-hypotheses about a target's movements [6].

Acknowledgment

This work was supported by the Software engineering and Information technologies Department, École de technologie supérieure, Montréal, Canada.

References

[1] J. Nascimento and J. Marques, "Performance evaluation of object detection algorithms for video surveillance," *IEEE Transactions on Multimedia,* vol. 8, p. 13, 2006.

[2] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *IEEE 17th International Conference on Pattern Recognition*, 2004, pp. 28-31.

[3] A. Senior, A. Hampapur, Y.-L. Tian, L. Brown, S. Pankanti, and R. Bolle, "Appearance models for occlusion handling," *Elsevier Image and Vision Computing,* vol. 24, p. 10, 2006.

[4] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," University of North Carolina at Chapel Hill1995.

[5] G. Bradski and A. Kaehler, *Learning OpenCV*, 1 ed., 2008.

[6] M. Isard and A. Blake, "CONDENSATION - Conditional Density Propagation for Visual Tracking," *International Journal of Computer Vision,* vol. 29, pp. 5-28, 1998.

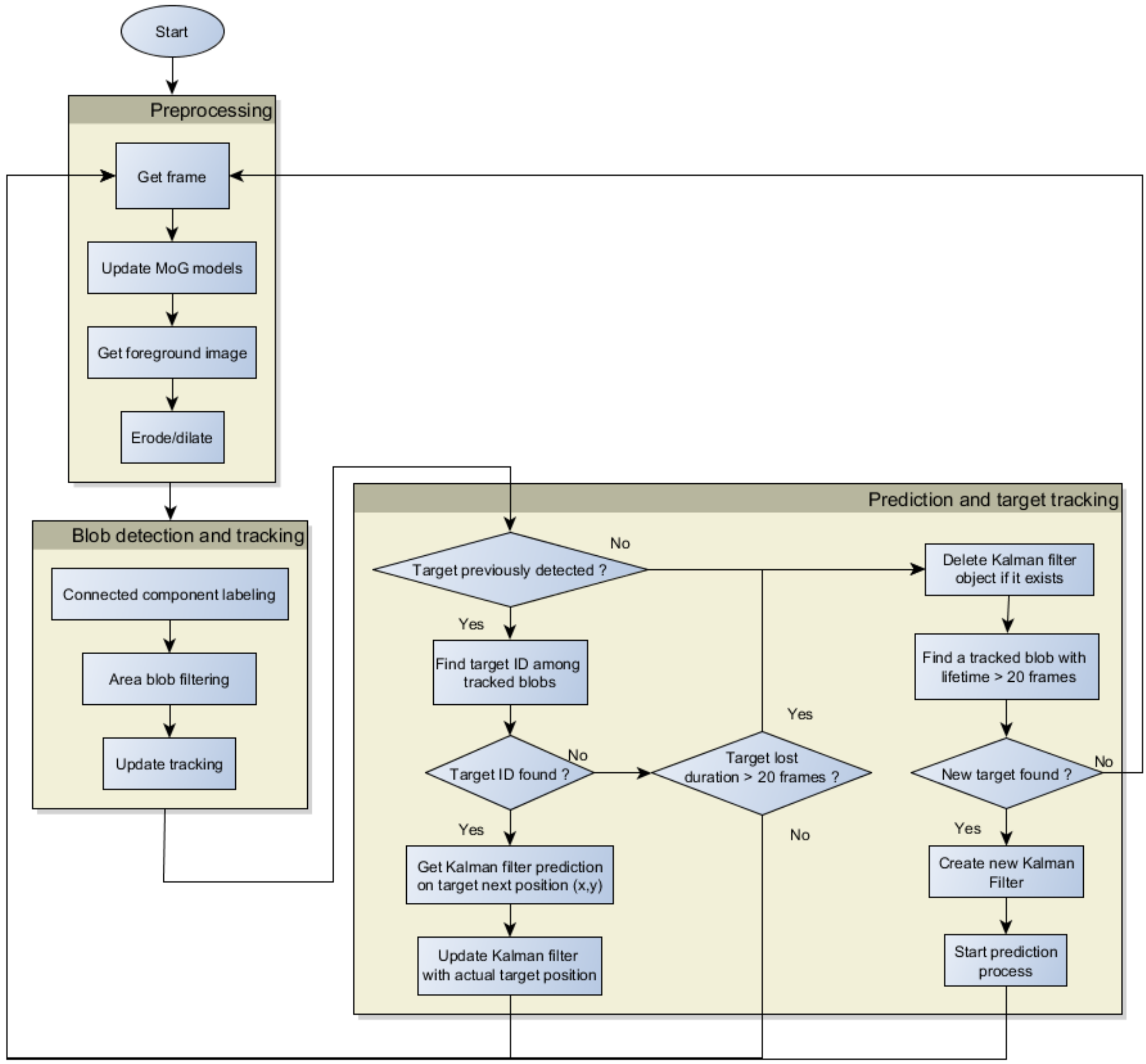

Fig. 2 The complete algorithm