UDC 629.7.014-519:778.54 (045), DOI: 10.18372/1990-5548.51.11715

[1]V. M. Sineglazov,
[2]J. V. Trach

IDENTIFICATION OF STATIC AND MOVING OBJECTS BY UAVs

Educational & Scientific Institute of Information and Diagnostic Systems, National Aviation University,
Aviation Computer-Integrated Complexes Department, Kyiv, Ukraine
E-mails: [1]svm@nau.edu.ua, [2]jultrach@rambler.ru

Abstract—*The real-time method to detect and track moving objects from unmanned aerial vehicles using a single camera.*

Index Terms—Unmanned aerial vehicle; image processing; tracking.

I. INTRODUCTION

Modern military technology is impossible without effective automated monitoring systems and escorting of targets. One of the ways, that significantly increase the combat capability of modern machines, are high-quality cameras day and night surveillance and devices to automatically track visually selected targets, which greatly facilitate the operator's work and increase its effectiveness.

Computers provide ample opportunities for automatic image processing, which are still only used to a small extent. Such simple tasks as recording only movements have rather universal solutions. However, more complex tasks, especially from the field of computer vision, still do not have common solutions, and the proposed algorithms, as a rule, are applicable only under certain conditions and limitations. In connection with this, to date, this field of research is of great interest both in scientific and applied terms. One important but completely unsolved sub-task is the automatic allocation of moving objects, which is a necessary preliminary step in solving such surveillance tasks as access control systems, background removal, robot orientation in space, driver assistance systems, face recognition, facilitation of scene perception the human eye and many others.

Army aviation today is one of basic constituents of Ground forces of the Armed Forces of Ukraine (GF). It presence considerably promotes battle and mobilization possibilities of Ground Forces. Unmanned aerial vehicles can operate both: in proximity to the battle area, and over enemy territory.

One potential benefit of unmanned aerial vehicles (UAVs) is that they could fill a gap in current border surveillance. In particular, technical capabilities of UAVs could improve coverage along remote sections of the borders.

This article is devoted to the problem of detection in real-time foreground objects in a video taken from UAV with a non-static camera. One of the most common applications of such a task is tracking algorithms for people who use background subtraction to find objects. Such algorithms are used in security video surveillance, coverage of sports broadcasts. Also, object tracking algorithms are used to create additional effects during video conferencing, for example, to replace the background.

Detection and tracking of dynamic objects has become an important field for the correct development of many multidisciplinary applications, such as traffic supervision [1], autonomous robot navigation [2], [3], and surveillance of large facilities [4]. In this article we primarily focused on detection of moving objects from aerial vehicles (UAVs) for surveillance, although other potential applications could also benefit from the results.

II. PROBLEM STATEMENT

Detecting objects with the help of computers is effective and modern method of video processing. High accuracy has made this method popular and widespread.

The problem statement of object detection in image available in several versions:

- the object of observation moves, drone with a camera fixed,
- immovable object of observation, drone with a camera moves;
- observation object moving, a drone with a camera is also moving.

We need to obtain the coordinates of the observed object relative to the location and orientation of the observer or the direction of the object relative to the orientation of the observer and therefore determine in what direction and how fast observed objects are moving. In general, the problem is solved by complex object recognition in obtained images from video sequence according the most common "visual characteristics" in tracking objects, such as movement, shape, color and intensity of light.

The main goal of UAV patrolling is right capturing the object and tracking it from frame to frame. While detecting the objects the hardest task is to identify weather it is the person, animal, car, tree, or building. There are lots of methods for objects detection, recognition, classification, and then tracking. All this method have some special features that they extract in the video, the main disadvantage of these methods is – all of them aren't universal, it means, that if this method works for this particular video and area, it would not work so good if we'll change the altitude of flight(consequence: size of objects will change). So, in this paper, we tried to investigate several methods of objects detection and tracking, to identify the objects in a given video. The desired output is a video in which the objects are marked in every frame of the video.

III. Review of Methods

There are three key steps in video analysis.

1. Detection of interesting objects.
2. Tracking of such objects from frame to frame.
3. Analysis of object tracks to recognize their behavior. The common methods of object tracking and object classification are presented in the Fig. 1.

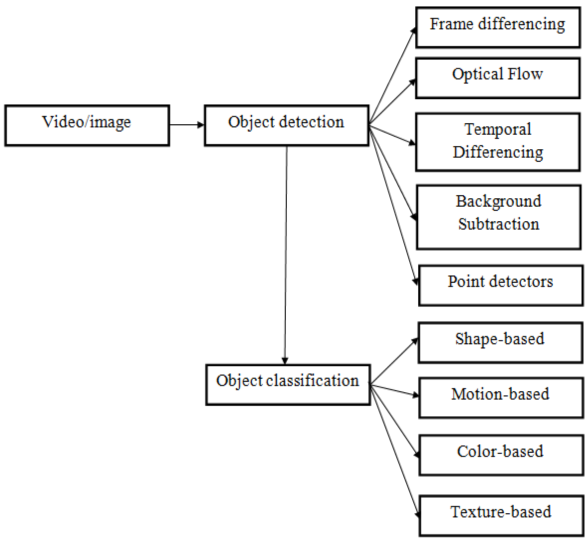

Fig. 1. General steps in object detection and classification

For the analysis was chosen such most commonly used algorithms for capturing and tracking of objects:

- algorithm of patterns of motion (Motion Templates) is based on finding the boundaries of objects in each frame of video. The shift of boundary in a new frame relatively to the previous frame defines motion vector of object;
- mean-Shift is based on a mathematical model, which calculate local extreme of density distribution of a set of characteristic points, that is, the algorithm tracks the displacement of the center of mass points that define the object tracking, receiving the output vector of the object;
- viola-Jones is based on the detection of sets pixels in the frame that match the pre-selection patterns composed of black and white rectangles. To recognize different objects it is required a unique set of templates, created by training the algorithm on a particular subject.

To start any tracking algorithm it is necessary, some way, to initialize the initial tracking area or array of specific points.

IV. Problems Solution

The motion detection algorithm begins with the segmentation part where foreground or moving objects are segmented from the background. The simplest way to implement this is to take an image as background and take the frames obtained at the time *t*, denoted by $I(t)$ to compare with the background image denoted by B. So, we can segment out the objects using image subtraction technique of computer vision meaning for each pixel in $I(t)$, take the pixel value denoted by $P[I(t)]$ and subtract it with the corresponding pixels at the same position on the background image denoted as $P[B]$.

In mathematical equation, it looks like:

$$P[F(t)] = P[I(t)] - P[B]. \qquad (1)$$

The background is assumed to be the frame at time *t*. This difference image would show some intensity for the pixel locations which have changed in the two frames. This approach will only work for cases where all foreground pixels are moving and all background pixels are static. A threshold is put on this difference image to improve the subtraction.

$$\left|P[F(t)] - P[F(t+1)]\right| > \text{Threshold}. \qquad (2)$$

Running Gaussian average

Every pixel is characterized by mean μ_t and variance σ_t^2 , let us initialize that every pixel is background.

$$\mu_0 = I_0, \quad \sigma_0^2 = \text{some default value}. \qquad (3)$$

Here, I_t is the value of the pixel's intensity at time *t*.

As the background may change over the time (due to illumination changes, change in altitude of UAV flight, or non-static background objects). To cope with that change, at every frame t, every pixel's mean and variance must be updated, as follows:

$$\mu_t = \rho I_t + (1-\rho)\mu_{t-1},$$
$$\sigma_t^2 = d^2\rho + (1-\rho)\sigma_{t-1}^2, \quad (4)$$
$$d = |(I_t - \mu_t)|,$$

Here, ρ determines the size of the temporal window that is used to fit the probabilistic density function (usually $\rho = 0.01$) and d is the Euclidean distance between the mean and the value of the pixel.

Now we can classify a pixel as background if its intensity lies within some confidence interval of its distribution's mean:

$$\frac{|(I_t - \mu_t)|}{\sigma_t} > k \rightarrow \text{Foreground},$$
$$\frac{|(I_t - \mu_t)|}{\sigma_t} \le k \rightarrow \text{Background}, \quad (5)$$

where k is a free threshold (usually $k = 2.5$).

So, the equation for the mean is changed accordingly:

$$\mu_t = M\mu_{t-1} + (1-M)(I_t\rho + (1-\rho)\mu_{t-1}), \quad (6)$$

where $M = 1$ when I_t is considered foreground and $M = 0$ otherwise. So when $M = 1$, that is, when the pixel is detected as foreground, the mean will stay the same. As a result, a pixel, once it has become foreground, can only become background again when the intensity value gets close to what it was before turning foreground.

To detect object in the video first we should analyze the image.

If each pixel resulted from a particular surface under particular lighting, a single Gaussian would be sufficient to model the pixel value while accounting for acquisition noise. If only lighting changed over time, a single, adaptive Gaussian per pixel would be sufficient. In practice, multiple surfaces often appear in the view frustum of a particular pixel and the lighting conditions change. Thus, multiple, adaptive Gaussians are necessary. We use a mixture of adaptive Gaussians to approximate this process.

Each time the parameters of the Gaussians are updated, the Gaussians are evaluated using a simple heuristic to hypothesize which are most likely to be part of the "background process." Pixel values that do not match one of the pixel's "background" Gaussians are grouped using connected components. Finally, the connected components are tracked from frame.

Firstly we grab 2 or 3 successive frames (Fig. 2). Secondly, we need to divide every image according to the RGB (to get red, blue and green parts of our image) we do this for the better segmentation of image, as it is easier to get more details in that way. We obtained 3 images from each frame; next our step will be receiving the grayscale image of them. Summing them up we get (Fig. 3).

Using the segmentation algorithm we obtained (Fig. 4).

Subtraction of the background image from the current video frame. This step includes pixel-by-pixel subtraction of the frame rate of the video and the background image.

Fig. 2 The screenshot of input video taken from the UAV

Fig. 3 Grayscale image

The selection of pixels belonging to the background and the object is the construction of a binary image. It is considered that the pixel belongs to the object and has a white color in the binary image if the difference in the intensity of the background and the current frame for a given pixel exceeds a certain threshold value.

Fig. 4 Result of segmentation algorithm

If the subject is darker than the background on the current frame, subtraction of the frame from the background will not lead to correct results, and vice versa, if the object of the current frame is lighter than the same area of the background frame, subtraction of the background from the current frame will give incorrect results. Because the luminance component is used, the result of the image difference will be white (for moving objects) and gray (for shadows) pixels if you set thresholds for them, otherwise the shadow will merge with the object and the area it occupies will be much more. At this stage, it is also necessary to filter (using median filtering) to remove noise – single white pixels. Since only objects are informative for the problems described above, further cut-off transformations cut off the shadow (Fig. 5).

a b

Fig. 5. The results of object detection: (a) is the initial image; (b) is the cut-off image

Algorithm of the object detection and tracking:

1) camera calibration;
2) video capture;
3) taking several successive frames;
4) noise reduction of any frame with the help of Gausses filter;
5) detection of moving targets by the method of background subtraction;
6) contours highlighting (Scanning algorithm with the assignment of numerical values to the scanned pixels);
7) objects classification (Method of histograms of directed gradients and linear classifiers);
8) tracking (Search for the center of the object in the neighborhood of its position on the previous frame);
9) determination of the motion parameters (determination of the speed taking into account the camera parameters and the motion of the UAV).

These algorithms provide the classification of objects and determine the speed of their movement under various environmental conditions and minimum equipment costs. The algorithms for automatic object tracking set an automatic image algorithms including background substitution, which also requires a certain image filtering of noise. It was also considered improving search and object recognition partly distorted form dots tracking moving objects.

V. Experiments and Results

In our experiments, we used two different image sequences coming from the same UAV sensor with an image resolution of 1280×720 pixels. The first scene has a GSD of 0.345 m pixel and consists of 1201 single images with 33270 moving target detections in summary. Also the format of color image frame is 24-bit in a RGB system. We manually labeled 5067 detections distributed across the whole sequence either for being building or vehicle. The second image sequence has 733 images with approximately 4461 detections and a GSD of about 0.225 m pixel. This scene was manually.

Among the detections, for example, in shown frame, there are 40 objects. All not detected objects are moving only very slowly or not at all. There are several wrong detections (it caused by the pixels similarities in that area).

The two example sequences differ in UAV altitude, object shadow, and number of objects. Furthermore, it is most likely, that no object is appearing in both image sequences.

In the case of stable operation, when there is no strong change in lighting or camera settings, the accuracy of the algorithm based on the Gaussian mixture exceeds the accuracy of other methods. However, once the lighting changes significantly, the number of errors increases dramatically. Many systems embed a separate step to handle such situations. For example, by sharp changes in the histogram, the brightness of the pixels in the frame is determined by the instant of light changes, after which changes are made to the background model (or vice versa, the colors of the pixels of the next video frame are changed). We used this algorithm from Matlab without changes, without embedding such processing. With its embedding, we admit that an algorithm based on a mixture of Gaussians would be able to show the greatest accuracy on test video clips. However, usually such treatments are heuristics using a number of manually selected parameters. The proposed algorithm copes with this problem without any additional modifications (Table I).

TABLE I

CLASSIFICATION RATES

Correct classification rate	89.68%
False detected	6.02%
Not detected objects	5.33%

IV. Conclusion

In this paper we have shown a probabilistic method for background subtraction. It involves modeling each pixel as a separate mixture model also it needs some image segmentation. We implemented a real-time approximate method which is stable and robust. The method requires several parameters which are robust to different cameras and different scenes.

This method deals with lighting changes by adapting the values of the Gaussians. It also deals with multi-modal distributions caused by shadows, secularities, swaying branches, computer monitors, and other troublesome features of the real world which are not often mentioned in computer vision. It recovers quickly when background reappears and has an automatic pixel-wise threshold. All these factors have made this tracker an essential part of our activity and object detection and classification research. This system has been successfully used to track people, and cars, remote control vehicles in a lab setting. All these situations involved different objects being tracked. This system achieves our goals of real-time performance over extended periods of time without human intervention.

References

[1] T. Gao, Z. Liu, and S. Yue, “Traffic video-based moving vehicle detection and tracking in the complex environment.” *Cyber. Syst.* vol. 40, pp. 569–588, 2009.

[2] Heinze, N., Esswein, M., Kr¨uger, W., and Saur, G., “Image exploitation algorithms for reconnaissance and surveillance with UAV,” *in Proceedings of SPIE vol. 7668, Airborne intelligence, surveillance, reconnaissance (ISR) systems and applications VII*, April 2010.

[3] https://alchetron.com/Background-subtraction-1816821-W

Received December 18, 2016

Sineglazov Viktor. Doctor of Engineering Science. Professor.

Educational & Scientific Institute of Information and Diagnostic Systems, Aviation Computer-Integrated Complexes Department, National Aviation University, Kyiv, Ukraine.

Education: Kiev Polytechnic Institute. Kiev, Ukraine (1973).

Research interests: Air Navigation, Air Traffic Control, Identification of Complex Systems, Wind/solar power plant.

Publications: more than 600 papers.

E-mail: svm@nau.edu.ua

Julia Trach. Student.
Educational & Scientific Institute of Information and Diagnostic Systems, Aviation Computer-Integrated Complexes Department, National Aviation University, Kyiv, Ukraine.
Research interests: image processing
E-mail: jultrach@rambler.ru

В. М. Синєглазов, Ю. В. Трач. Ідентифікація статичних і рухомих об'єктів, за допомогою БПЛА
Розглянуто метод виявлення та відстеження рухомих та статичних об’єктів безпілотними літальними апаратами.
Ключові слова: безпілотний літальний апарат; обробка зображень; відстеження.

Синєглазов Віктор Михайлович. Доктор технічних наук. Професор.
Навчально-науковий інститут інформаційно-діагностичних систем, кафедра авіаційних комп’ютерно-інтегрованих комплексів, Національний авіаційний університет, Київ, Україна.
Освіта: Київський політехнічний інститут. Київ, Україна (1973).
Напрям наукової діяльності: аеронавігація, управління повітряним рухом, ідентифікація складних систем, вітроенергетичні установки.
Кількість публікацій: більше 600 наукових робіт.
E-mail: svm@nau.edu.ua

Трач Юлія Володимирівна. Студентка.
Навчально-науковий інститут інформаційно-діагностичних систем, кафедра авіаційних комп’ютерно-інтегрованих комплексів, Національний авіаційний університет, Київ, Україна.
Напрям наукової діяльності: обробка зображень
E-mail: jultrach@rambler.ru

В. М. Синеглазов, Ю. В. Трач. Идентификация статических и подвижных объектов, с помощью БПЛА
Рассмотрен метод обнаружения и отслеживания подвижных и статических объектов беспилотных летательных аппаратов.
Ключевые слова: беспилотный летательный аппарат; обработка изображений; слежение.

Синеглазов Виктор Михайлович. Доктор технических наук. Профессор.
Учебно-научный институт информационно-диагностических систем, кафедра авиационных компьютерно-интегрированных комплексов, Национальный авиационный университет, Киев, Украина.
Образование: Киевский политехнический институт. Киев, Украина (1973).
Направление научной деятельности: аэронавигация, управление воздушным движением, идентификация сложных систем, ветроэнергетические установки.
Количество публикаций: больше 600 научных робот.
E-mail: svm@nau.edu.ua

Трач Юлия Владимировна. Студентка.
Учебно-научный институт информационно-диагностических систем, кафедра авиационных компьютерно-интегрированных комплексов, Национальный авиационный университет, Киев, Украина.
Направление научной деятельности: обработка изображений.
E-mail: jultrach@rambler.ru