Moving Object Detection in Video Streams Received from a Moving Camera

Sergey Starkov, Maksim Lukyanchenko

National Research Nuclear University MEPhI, Obninsk, Russia
starkov@iate.obninsk.ru, maksim.lukyanchenko@gmail.com

Abstract. Detection of moving objects in a video stream received from a moving camera is difficult computer vision task, because the motion of the camera blends with the motion of the objects in the scene. In order to tackle this problem, we propose a method based on optical flow calculation and Delaunay triangulation. Given a sequence of frames, firstly, we extract the corner feature points using ORB algorithm and compute optical flow vectors at the extracted feature points. Secondly, we separate the optical flow vectors using K-Means clustering method. Third, we classify each cluster into camera and object motion using its mean scatter value. Finally, we represent the moving object using Delaunay triangulation.

Keywords: moving objects, moving camera, unstable background, ORB, optical flow, clustering, Delaunay triangulation

1 Introduction

Detection of moving objects of interest and tracking of such objects from frame to frame is an important tasks in systems that perfoms of video data, such as video surveillance systems, industrial robots, unmanned vehicles etc.

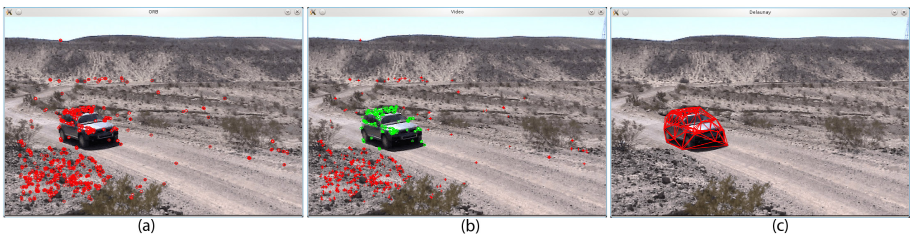

Fig. 1. Moving Object Detection in Video Streams Received from a Moving Camera (a)ORB feature points. (b)Classification of feature points, (c)Moving object in video stream

Based on the type of motion between frames, all objects can be divided into two classes: static and dynamic. Static objects maintain their position on the sequence of frames, dynamic objects change their position in space. There has been considerable research focusing on the separation of static and dynamic objects in video sequences taken from camera in stationary positions.

The classical object detection methods can not be applied directly for detecting such objects in a scenario with a moving camera because there exist multiple sources of motions from both the camera and the moving objects. In our research we focus attention on the problem of moving object detection in a video stream captured using a moving camera. To detect moving objects in a moving camera environment we need to discriminate between camera motion and object motions. Generally, there are three approaches to detect the moving object under a moving camera:

- Compensation of camera motion by ego-motion estimation [1][2];
- Separation of motions vectors in the input sequence using motion models [3][4];
- Segmentation of the camera and object motions using the graph cut algorithm [5][6].

Some of these methods need an additional algorithmic stage to select the moving object motion model, others require considerable computation time. We decided to design an improved moving-object detection method using data from a free-moving camera with a non-stationary background, which provides both high detection performance and fast processing speed. In this paper we demonstrate the proposed approach with initial results.

2 Motion estimation

For extracting structured information about the object of interest in the image, we search for feature points of the scene.

A feature point in a scene $\boldsymbol{M}$ is called a scene point if it is coplanar with other points in image neighborhood $O(M)$ that can be distinguished from all other neighborhoods $O(N)$, which in turn are composed of several points $\boldsymbol{N}$.

2.1 Feature Point Detection

In the proposed method, to find the feature points of the image, the Oriented Fast and Brief detector (ORB) is used (Fig. 1 (a)). When using this technique, it is assumed that the intensity of the corner point is offset from the center and this displacement vector can be considered as feature point direction. To calculate a descriptor of the point $\mathbf{p}(x, y)$, ORB compares brightness values of points located in its vicinity. This algorithm is invariant to image rotation, scale change, and changes in lighting level, so it satisfies the main qualities required of robust feature detectors and is suitable for a reliable estimation of moving singular points.

2.2 Optical flow computation

To determine the movement of objects in two-dimensional space using optical sensor systems, algorithms in computer vision and image processing make use of optical flow - the apparent motion of the image, which is a shift of each point between two consecutive frames.

In our approach, we compute the optical flow vectors of image points by searching for the corresponding feature points between two consecutive image frames using the pyramidal Lucas-Kanade method. This process consists of two tasks: generation of image pyramid and search for the correspondence feature points on the image pyramid.

3 Motion Clustering

Clustering is the division of the set of input vectors into groups (clusters) on the degree of "similarity" to each other.

In this paper, we cluster feature points using the length L and direction θ of optical flow vectors. The feature points are described in the optical flow coordinate (L, θ).

All optical flow coordinate (L, θ) were divided into blocks. Randomly were selected the initial points for clustering. Number of clusters is an input parameter of the method. In the present implementation of the algorithm is assumed two: background and foreground.

4 Motion classification

The clusters generated are to be separated into those that relate to the movement of the camera and those to moving objects.

In the proposed framework, we assume that the background occupies a larger area of the frame than moving objects. Thus, the points that relate to the background have a greater dispersion than singular points belonging to objects in motion (except in cases where the background has a large amount of small details). The assignment of each cluster to a background or a moving object can be done using the measure of spread of the points within the cluster. To determine the measure of the spread of points within each cluster, in the present work, we use the standard deviation s as a discriminative metric.

Cluster, which has the highest standard deviation, be deemed to apply to the background (Fig. 1 (b)).

5 Moving object detection

To select the area that relates to a movable object, in the proposed framework, Delaunay triangulation has been used. Triangulation is a planar partition of the 2D space by plane figures, one of which is an outer infinity, and the rest are triangles.

When using Delaunay triangulation, for all resulting triangles, points of the cluster except for points at vertices lie outside the circle circumscribed about the triangle.

After constructing the Delaunay triangulation, the resulting set of triangles with edges length exceeding a predetermined threshold are removed (Fig. 1 (c)).

6 Conclusion

In this work, we have developed an effective method for separation of moving objects in the scene using data from an input video stream in the presence of a non-stationary background. This method shows high frame rate performance - 20-21 fps on a computer with a processor Intel Xeon E5420 1333 MHz and 4GB RAM. However, this value does not satisfy the operation mode in real time (>24 frames per second). In order to improve the real-time performance of the algorithm, we envisage that in subsequent implementations, in addition to algorithmic optimization, implementation will be carried out on a graphics card using software optimization libraries for CUDA and OpenCL. Also planned are:

- Exploration of the possibility of introducing additional parameters to improve the quality of clustering.
- Implementation of automatic identification of cluster numbers in the step for clustering singular points.
- Implementation of automatic identification of cluster numbers in the step for clustering singular points.

References

1. Hayman E., Eklundh J.: Statistical background subtraction for a mobile observer. In: Proc. IEEE ICCV 2003. (2003)
2. Ren Y, Chua CS, Ho YK: Statistical background modeling for non-stationary camera. Pattern Recogn **24(1-3)** (2003) 183–196
3. Borshukov GD, Bozdagi G, Altunbasak Y, Tekalp AM: Motion segmentation by multistage affine classidication. IEEE Trans Image Process **6(11)** (1997) 1591–1594
4. Ke Q, Kanade T: A subspace approach to layer extraction. In: IEEE CVRP. (2001)
5. M, X.J.S.: Motion layer extraction in the presence of occlusion using graph cuts. IEEE Trans Pattern Anal Mach Intell **27(10)** (2005) 1644–1659
6. D, S.T.C.: High resolution motion layer decomposition using dualspace graph cuts. In: In: Proc. IEEE CVPR. (2008)

Выделение движущихся объектов сцены из входящего видеопотока при наличии нестационарного фона

Сергей Старков, Максим Лукьянченко

НИЯУ МИФИ, Обнинск, Россия
starkov@iate.obninsk.ru, maksim.lukyanchenko@gmail.com

Аннотация Автоматическое выделение движущихся объектов сцены из входящего видеопотока – одна из важнейших задач анализа изображений. За последнее время было предложено большое количество методов решающих данную задачу при условии неподвижности видеокамеры. В основе этих методов лежит принцип накопления кадров и выявления изменений в них. Однако при наличии подвижной камеры применение данного подхода становится невозможным. Вместе с тем развитие автономных беспилотных транспортных средств требует решения и этой задачи компьютерного зрения. В представленной работе предложен метод нацеленный на решение указанной задачи, в его основе лежит выделение ключевых точек изображения, вычисление оптического потока, сегментация изображения на фон и объекты относящиеся к переднему плана, маркирование участков изображения. Произведена оценка результатов работы предложенного метода.

Ключевые слова: движущийся объект, подвижная камера, нестабильный фон, оптический поток, кластеризация, триангуляция.