Development and research of algorithms of recognition and localization of moving objects in the video stream

Maxim Ikonnikov
Peter the Great St.Petersburg Polytechnic University
Polytechnicheskaya, 29
St.Petersburg, Russia
denimqa@gmail.com

ABSTRACT

This paper describes how to create a software application for the detection of moving objects in the video stream. The paper discusses the various background subtraction algorithms and morphological operations, the comparative analysis of the methods and the choice of suitable means to implement the program. The developed software module can be used in systems of intelligent analysis of the video stream. Development performed using C ++ language and OpenCV library.

Keywords

video analysis; background subtraction; morphological operations; object detection; opencv library

1. INTRODUCTION

Background subtraction (BS) is a crucial step in many video analysis systems. BS has been widely studied since the 1990s. Many algorithms have been designed to segment the foreground objects from the background.

All of BS methods try to effectively estimate the background model from the temporal sequence of the frames.

This article compares some of the most implemented background subtraction methods according to several criteria: performance, quality and number of false detections. We use the OpenCV library to implement BS methods in our software application.

2. VIDEO ANALYSIS AND BACKGROUND SUBTRACTION

2.1 Motion

Motion is a changing position of one object relative to another objects over time.

There are 3 main different cases of motion in computer vision tasks:

1. static camera, static background
2. moving camera, static background
3. moving camera, dynamic background

2.2 Background subtraction methods

Background subtraction methods are usually used for motion detection at the first place. Some types of BS methods:

Non-Recursive methods:

1. subtracting the current and previous frames
2. averaging the determined number of previous frames
3. calculating the median of the determined number of previous frames

Recursive methods:

1. histogram method
2. Gaussian-Mixture method
3. codebook
4. visual background extraction method

We use a sequence of frames $I_1, I_2, ...I_N$ as an entry data:

$$I_k = I_k(x, y), 0 \leq x < width, 0 \leq y < height, k = \overline{1, N} \quad (1)$$

The result of background subtraction is a set of binary images where white pixels correspond to the pixels of moving objects and black pixels correspond to background.

$$M_k(x, y) = \begin{cases} 255, & (x, y) - pixel\ of\ an\ object \\ 0, & (x, y) - pixel\ of\ background \end{cases} \quad (2)$$

2.3 Morphological operations

Morphology is a broad set of image processing operations that process images based on shapes. Morphological operations apply a structuring element to an input image, creating an output image of the same size. In a morphological operation, the value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. By choosing the size and shape of the neighborhood, you can construct a morphological operation that is sensitive to specific shapes in the input image.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

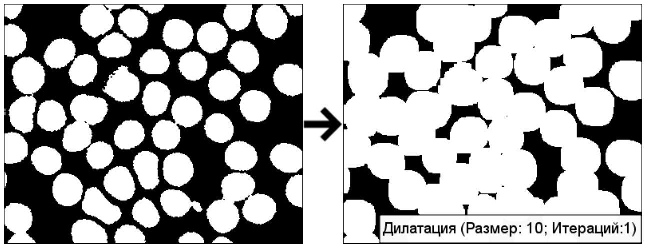

Figure 1: The result of dilation

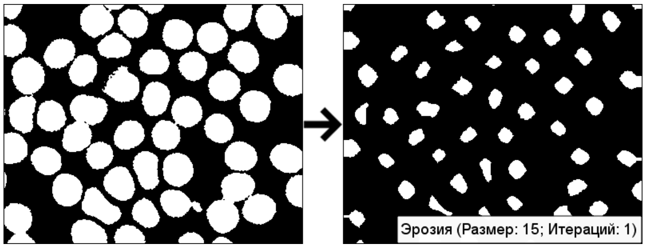

Figure 2: The result of erosion

2.3.1 Dilation

Dilation is one of the two basic operators in the area of mathematical morphology, the other being erosion. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (i.e. white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller (Fig.1).

2.3.2 Erosion

Erosion is one of the two basic operators in the area of mathematical morphology, the other being dilation. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (i.e. white pixels, typically). Thus areas of foreground pixels shrink in size, and holes within those areas become larger (Fig. 2).

2.4 OpenCV methods

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Being a BSD-licensed product, OpenCV makes it easy for businesses to utilize and modify the code.

The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc. OpenCV has more than 47 thousand people of user community and estimated number of downloads exceeding 7 million. The library is used extensively in companies, research groups and by governmental bodies.

Figure 3: Original frame

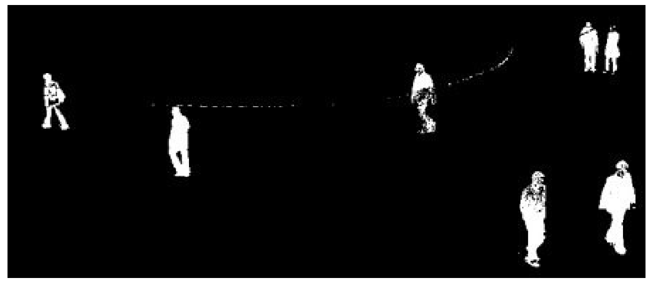

Figure 4: The result of BackgroundSubtractorMOG

2.4.1 BackgroundSubtractorMOG

It is a Gaussian Mixture-based Background/Foreground Segmentation Algorithm. It was introduced in the paper "An improved adaptive background mixture model for real-time tracking with shadow detection" by P. KadewTraKuPong and R. Bowden in 2001. It uses a method to model each background pixel by a mixture of K Gaussian distributions (K = 3 to 5). The weights of the mixture represent the time proportions that those colours stay in the scene. The probable background colours are the ones which stay longer and more static (Fig. 4).

2.4.2 BackgroundSubtractorMOG2

It is also a Gaussian Mixture-based Background/Foreground Segmentation Algorithm. It is based on two papers by Z.Zivkovic, "Improved adaptive Gausian mixture model for background subtraction" in 2004 and "Efficient Adaptive Density Estimation per Image Pixel for the Task of Background Subtraction" in 2006. One important feature of this algorithm is that it selects the appropriate number of gaussian distribution for each pixel. (Remember, in last case, we took a K gaussian distributions throughout the algorithm). It provides better adaptibility to varying scenes due illumination changes etc (Fig. 5).

2.4.3 BackgroundSubtractorGMG

This algorithm combines statistical background image estimation and per-pixel Bayesian segmentation. It was introduced by Andrew B. Godbehere, Akihiro Matsukawa, Ken Goldberg in their paper "Visual Tracking of Human Visitors under Variable-Lighting Conditions for a Responsive Audio

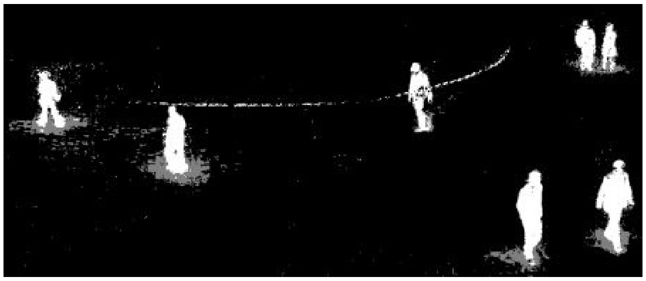

Figure 5: The result of BackgroundSubtractor-MOG2

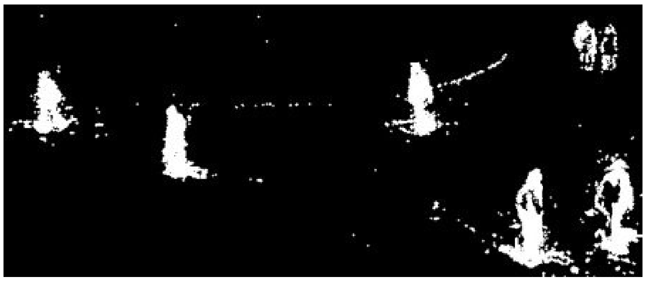

Figure 6: The result of BackgroundSubtractorGMG

Art Installation" in 2012. As per the paper, the system ran a successful interactive audio art installation called "Are We There Yet?" from March 31 - July 31 2011 at the Contemporary Jewish Museum in San Francisco, California (Fig. 6).

2.5 Analysis of some methods

The comparing results of several background subtraction methods are presented in Table 1.

Testing infrastructure:

- CPU: Intel Core i7-5700HQ 2.70 GHz
- RAM: 8 GB
- OS: Windows 8.1 Enterprise
- Frame resolution: 720x404
- FPS: 25

Table 1: The result of comparing

Method	Processing of frame	Shadow detection	Noise	Staying objects
BSMOG	100 ms	no	no	yes
BSMOG2	3 ms	yes	low	yes
BSKNN	30 ms	yes	low	yes
BSGMG	24 ms	no	high	yes
AdaptiveSelective	2 ms	no	low	no
DPMeanBGS	6 ms	no	no	yes
SJN MultiCueBGS	33 ms	no	no	yes

2.6 Steps of detecting moving objects algorithm

Step 1: Original frame
Step 2: Background Subtraction
Step 3: Erosion
Step 4: Dilation
Step 5: Finding contours
Step 6: Creating bounding rects

Figure 7: Original frame

Figure 8: Background Subtraction

3. CONCLUSIONS

In this article, we have compared several background subtraction methods from the literature, thanks to OpenCV and BGSLibrary. The algorithm for the detection of moving objects in the video stream has been developed.

Figure 9: Erosion

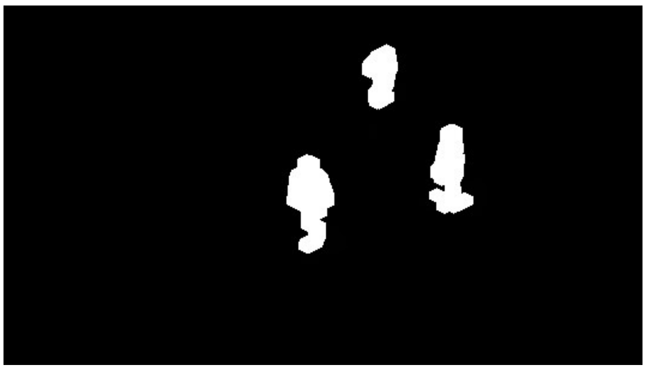

Figure 10: Dilation

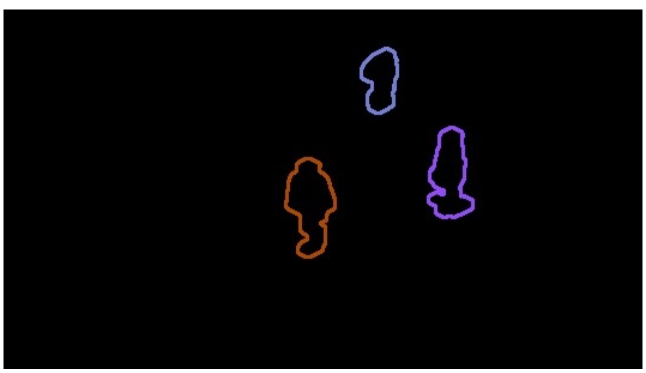

Figure 11: Finding contours

Figure 12: Bounding rects