Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 36 (2014) 579 – 586

Procedia Computer Science

Complex Adaptive Systems, Publication 4
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2014-Philadelphia, PA

Utilization of Robust Video Processing Techniques to Aid Efficient Object Detection and Tracking

Anand Balasubramanian, Shreyamsh Kamate, Nuri Yilmazer

Texas A&M University-Kingsville, Kingsville, TX 78363, USA

Abstract

In this research, data acquired by Unmanned Aerial Vehicles (UAVs) are primarily used to detect and track moving objects which pose a major security threat along the United States southern border. Factors such as camera motion, poor illumination and noise make the detection and tracking of moving objects in surveillance videos a formidable task. The main objective of this research is to provide a less ambiguous image data for object detection and tracking by means of noise reduction, image enhancement, video stabilization, and illumination restoration. The improved data is later utilized to detect and track moving objects in surveillance videos. An optimization based image enhancement scheme was successfully implemented to increase edge information to facilitate object detection. Noise present in the raw video captured by the UAV was efficiently removed using search and match methodology. Undesired motion induced in the video frames was eliminated using block matching technique. Moving objects were detected and tracked by using contour information resulting from the implementation of adaptive background subtraction based detection process. Our simulation results shows the efficiency of these algorithms in processing noisy, un-stabilized raw video sequences which were utilized to detect and track moving objects in the video sequences.

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology

Keywords: *Enhancement; Optimization; Stabilization; Block Matching; Background Subtraction*

1. Introduction

Illegal trespassing and border intrusions by immigrants are a great problem posted against the U.S. border security force and the Department of Homeland Security. Various technical and scientific approaches have been proposed over the past decade to ensure protection to the citizens of U.S. The main stay of these protection processes is video surveillance across the southern border to track movement of illegal immigrants. This video surveillance process is carried out using various methods namely, by mounting cameras on watch towers or using UAVs to constantly patrol the border. The images are captured using various types of image sensors mounted on UAVs. The main area of research is in finding out a way to extract more information from the captured data

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology
doi:10.1016/j.procs.2014.09.057

without complicating the hardware involved in the process. In this research, we used a UAV called "Phantom" for capturing test videos in a south Texas university area in order to test our algorithms. Phantom is an aerial filming multi-rotor system with a remote control unit range of 300 meter and a flying speed of 6 mph. Phantom is mounted with a camera on its base. The camera used is GoPro Hero 3 Black edition, which is capable of producing videos with 4K resolution and up to 30 frames per second. In the following sections we shall discuss various video enhancement algorithms and object detection algorithms which have been proposed by various researchers that could be incorporated for the processing of videos captured by the UAV and use it for the purpose of tracking illegal movement across southern border. Image enhancement is the process of converting the obtained raw image data into a more advantageous and convenient form. Image enhancement is also used to emphasize some particular regions of the image which could be used for special purposes like object tracking and detection. Various manipulations on the pixels of the captured image help in suppressing noise and also help in extricating more data from the captured image data. Further these processed and enhanced images are aimed to extract moving object that are of interest. The raw video sequences captured by the UAV are a primary source of information perused by Homeland Security. In order to make the videos more desirable and useful for the purpose of object tracking, the video needs to be made noise-free and clearer. These requisites are accomplished by the following four stages by pixel manipulation of the images in the video sequences.

2.1. Noise Removal:

Noise present in the video is an inadmissible parameter while processing video data. In order to de-noise the video sequence we use VBM3D algorithm [1] which is based on Box Matching and 3D Filtering (BM3D) [2]. This is a de-noising algorithm which has been used to remove specific noises. In BM3D noise removal is achieved by converting similar block in the image into similar image fragments called groups. These blocks are examined by sliding window method which helps in stacking blocks with similar features in a 3D stack called 'groups' and the procedure is called 'grouping'. The similarity between the blocks in the 3D stacks result in high level correlation between the blocks. These blocks are treated with 3D de-correlating transforms which attenuates the noise present in the image. The noise attenuation is attributed to the shrinkage of transform coefficients by wiener filtering and this procedure is termed as 'collaborative filtering'. We obtain the estimates of these blocks after an inverse of the 3D transform is applied. The estimates of the blocks are combined together to obtained the noise free reconstructed image.BM3D algorithm has been extended to process noisy image sequences in videos and is called Video Block matching 3D filtering (VBM3D). In this technique the standard deviation of the noise has not been determined by calculation before processing, instead it is assumed.
In VBM3D, the noisy video captured is represented as

$$R(x,y;t) = f(x,y;t) + n(x,y;t) \quad (1)$$

Where $R(x,y;t)$ is the noisy video signal, $f(x,y;t)$ is the true signal and $n(x, y;t)$ is the noise present in the video with standard deviation σ_n.

V-BM3D algorithm

Step 1.Grouping blocks with similar features into stacks and obtaining an estimate using grouping and collaborative filtering

Step 2.Final estimates of the 3D blocks and collaborative wiener filtering that uses frequency domain mapping of the corresponding groups from the basic estimate.

Figure 1. VBM3D algorithm.

2.2. Noise Estimation:

The noise removal process carried out by the VBM3D algorithm does not take into account the exact noise standard deviation of the videos into account. Therefore, in order to overcome this issue we perform Noise Estimation using Immerkaer Method [3]. Here the basic Noise Model is made use of on each image in the video.

$$I(x,y) = f(x,y) + n(x,y) \quad (2)$$

where I is the Noisy image, f is the Noise-free image and n is additive noise. The noise is assumed to be additive white noise with zero mean. Variance of the noise is the parameter that is required. It is used along with the VM3D to remove noise from the sequences. According to the Immerkaer Method [3], a Laplacian operator of the form shown below is used to calculate the standard deviation σ_n.

N=

1	-2	1
2	-4	2
1	-2	1