Segmenting and Classifying the Moving Object from HEVC Compressed Surveillance Video

JothiKumar R, Durai Hariharan K C, Akilan U

Department of Computer Science and Engineering

R M K Engineering College, Tamil Nadu, India

Abstract - *Moving **object detection** is one of the most important tasks for intelligent surveillance, **HEVC** introduces a host of new coding features that can be further abused for moving object **segmentation** and **classification**. Therefore object segmentation and classification directly from HEVC compressed video has a new challenge. In this paper, We develope a method for object segmenting and classifying moving objects, exactly, vehicle and person in the HEVC compressed video. First, **Motion Vector** (MV) interpolation for intra-coded prediction unit and MV outlier removal are employed for preprocessing. Second, blocks with non-zero motion vectors are gathered into the connected foreground regions. Third, object region tracking based on temporal constancy is applied to connected foreground regions to remove the noise regions. Finally, a person and vehicle classification model using the spatial-temporal HEVC syntax words is trained to classify the moving objects, either person or vehicle.*

1. INTRODUCTION

Moving Object detection has one of the Challenging tasks for any video application. There most of the video content is stored in compressed format encoding with international video stranded, Such as HEVC. To obtain the original frame we have to perform decoding. In analysis video at the large scale such as content analysis an search for a large surveillance network, The complexity of video decoding become a major problem of real-time system, To overcome this compression domain approaches have been explored for direct video content analysis which extracts features directly from bitstream syntaxes, such as Motion Vector (MV) and Block Coding model, using minimum amount of decoding effort. The computational domain approach advantage is low computational complexity since full-scale decoding

Recently, a number of moving object segmentation and classification algorithms in the H.264/AVC compression domain have been reported. Temporally accrued MVs are further interpolated spatially to obtain a dense field. It should be noted that motion vectors extracted from the compressed bitstream may not represent the true object motion since they are determined to optimize the coding efficiency. To address this issue, other information, such as DCT coefficients and macroblock (MB) partition, are used to detect and track moving objects. For example, the number of bits used by each 4×4 block is used to detect moving objects in from H264 videos. This method is able to segment the object with a relatively accurate shape since it operates on 4×4 blocks. HEVC is the updateted international standard for video coding. Very little work has been done on video object analysis directly from HEVC compressed videos. In addition, new features and syntax elements in HEVC, such as coding, prediction, and transform units, can be exploited to further improve the overall performance. Compared to existing methods in the prose, the major contributions of this work lie in the following aspects. First, we have developed and evaluated new HEVC syntax features in the spatiotemporal domain to achieve efficient object classification and maintain motion coherence and spatial compactness. Second, we have successfully extended the conventional "bag of words" model in the pixel domain to the compression domain.

The rest of the paper as follows.

- Moving object segmentation
- Moving object classification
- Concludes the paper.

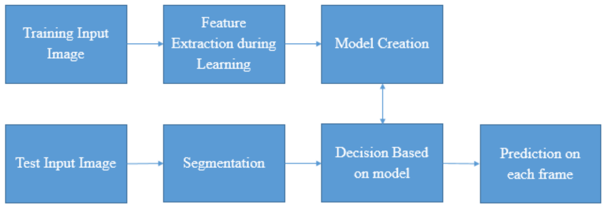
Overview

The general objective of this paper is to develop a framework for compression-domain moving object detection and classification directly from HEVC videos.

It consists of two process:

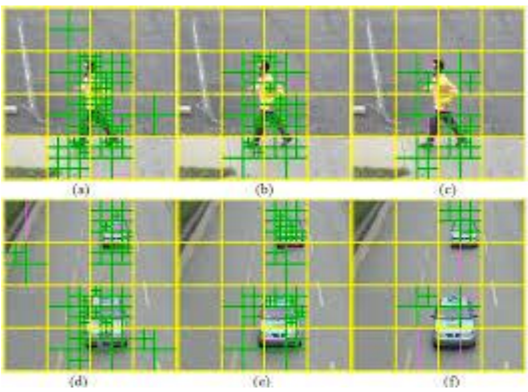
Moving object Segmentation & Person Vehicle Classification

Each stage involves a training phase to learn the model from the training data and a testing phase to apply the learned model to test videos.



2. Moving Object Segmentation

HEVC has introduced a quad-tree based coding approach where each picture is splited into square coding tree units (CTUs) . Each CTU is the root of a coding tree, which is further divided into coding units (CUs). Their sizes can be adaptively chosen by using a quad-tree based partitioning with the tree leaves representing the CUs. Each CU is a root for a prediction tree. The prediction tree has only one level and describes how a CU can be further split into so-called prediction blocks (PUs), where 8 different partition modes are used for inter-coded CUs and only two modes are used for intra-coded CUs. These are important features in HEVC which significantly improves the overall coding efficiency.



Figs a & d. show an example of block partitions for a surveillance video frame with moving persons and vehicles. Here, the largest square blocks, smaller square blocks and rectangular blocks represent CTU, CU, and PU, respectively. We observe that moving objects often have smaller CUs and Pus when compared to background image regions. In addition, object parts with non-rigid motions often have much smaller CUs and PUs. This is because they have difficulty in finding good matches from previous frames. This provides a key for the rigidness of object motion.

The squares in Figs. b & e represent blocks with non-zero motion vectors whereas the square blocks in Figs. c & f .represent intra coding modes.

We can see that some parts of persons and vehicles are coded with non-zero motion vectors. However, some other parts of persons and vehicles are coded with intra modes or zero motion vectors. In addition, some background blocks, which are supposed to be static, have non-zero motion vectors due to noise or background changes. Therefore, instead of solely relying on motion vector information for compression-domain moving object detection and segmentation, our method combines various HEVC syntax elements, including motion vector information, CU and PU sizes, and intra/inter/skip coding modes to form a comprehensive feature vector to characterize the moving object in the compression domain.

Our foreground-background classification operates on 4×4 blocks to achieve fine-grain segmentation of the moving objects. If the CU and PU have block partitions larger than 4×4, all 4×4 blocks inside are using the same motion vectors and coding modes as their covering block. Since the motion field is generally smooth, we also include the motion information of blocks in the spatiotemporal neighborhood to characterize the current block.

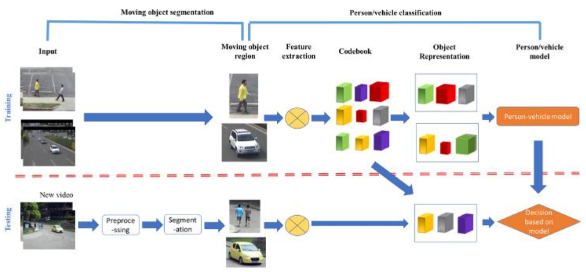
A linear SVM classifier is trained to classify all 4×4 blocks into the foreground and background blocks. During moving object detection and segmentation, the SVM classifier is applied to each 4×4 block in the current frame and then linked component analysis is used to group all foreground blocks into foreground objects.

3. Moving Object Classification

For object classification in surveillance videos, we aim to classify the detected moving objects into persons and vehicles using HEVC syntax features in the compression domain. In this work, we propose to explore an approach called a set of HEVC syntax words. Due to its histogram-type description, it is able to handle large variations in object locations, sizes, and viewing angles, as well as obstructions. This is particularly important for object detection-classification since detected object patches may not be aligned well with the object boundary.

This method has the following major steps:

- Analyzing each coding block within the moving object region using HEVC syntax features
- constructing a codebook using a clustering method
- representing each moving object using a normalized histogram of codewords from this codebook
- train a two-class classifier to classify the moving objects into persons and vehicles.
-



The main problem is to identify effective features in the compression domain which have sufficient discrimination power between persons and vehicles. In this work, through extensive experiments and performance evaluations, we have identified four types of features, namely, the absolute value of motion vectors, CU sizes, prediction modes, and motion vector difference. The absolute value of the motion vector relates to the velocity of the object, which is a simple yet important for discrimination feature for persons and vehicles. Since the motion field within the inflexible objects, such as vehicles, is often smooth, the motion vector of spatial and temporal neighborhood blocks are also used as an important feature.

Although both persons and vehicles are outlined with smaller CU sizes, we have found that the distribution of different CU sizes within the object region is one of the most characteristic features for persons and vehicle classification, since vehicles often exhibit consistent motion within the object regions. Specifically, large CUs often appear on the boundary of the moving persons. However, for vehicles, they appear both on the boundary and at the center. For person-vehicle classification in the HEVC compression domain, prediction modes and motion vector differences are also used. We observe that persons often undergo non-rigid deformations. In this case, it is harder to find a good match for the CU and PU. More blocks within the region of person are coded with intra prediction modes when compared to those of slow-moving and fast-moving cars. Because motion vectors within the moving vehicle are more consistent than those in persons, motion vector differences between neighborhood blocks of vehicles are smaller than those of persons. The motion vector differences of the x component and y component are computed separately. Specifically, the motion vector difference of the current block is computed as follows:

$$MVD_{N_{i,j}} = \begin{cases} abs\left(MV_{N_{i,j}}^{x} - MV_{C}^{x}\right) + abs\left(MV_{N_{i,j}}^{y} - MV_{C}^{y}\right), & if\ MV_{N_{i,j}} \neq 0 \\ 0, & else \end{cases} \quad (1)$$

$$MVD\_C_0^t = \frac{\sum_{i=0,j=0}^{i=2,j=2} MVD\_N_{i,j} + count/2}{count} \quad (2)$$

Where □□_Ni,j denotes the motion vector of the current block,$MV\_Ni,j\ x$ denotes the x component of MV_Ni, j, $MV\_Ni,j\ y$ denotes the y component of MV_Ni,j , MV_C denotes the motion vector of the current block, $MV\_C\ x$ is the x component of $MV\_\ C$, MV_Cy denotes the y component of MV_C, MVD_Ni,j denotes the motion vector difference between the current block and its neighborhood block Ni, j, $count$ is the number of neighborhood blocks with non-zero motion, and $MVD\_C0\ t$ is the total motion

vector difference of current block and all of its neighborhood blocks. Meanwhile, □□□_$C0\ r1$ and $MVD\_C0\ r2$ denote the motion vector difference of its temporal neighborhood blocks in reference frames $r1$ and $r2$, respectively

4. CONCLUSIONS

In this paper, we segment and classify the moving objects in an HEVC-compressed video. Firstly, the moving object region is automatically segmented by using the feature vectors extracted from the HEVC compression domain. Then, we discovered the possibility of applying the representation of temporal-spatial words to classify the moving objects in the HEVC compression domain.

The proposed method required considerably low processing time, yet still provides high accuracy.

REFERENCES

[1] M. Grundmann, V. Kwatra,M. Han, and I. Essa, “Efficient hierarchical graph-based video segmentation,” in Proc. IEEE

Conf. Computer Vision and Pattern Recognition, pp. 2141–2148, Jun. 2010.

[2] Hidetomo Sakaino, "Video-Based Tracking, Learning, and Recognition Method for Multiple Moving Objects", IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 5, pp. 1661–1674, Oct. 2013

[3] C. Poppe, S. De Bruyne, T. Paridaens, P. Lambert, and R. Van de Walle, "Moving object detection in the H.264/AVC compression domain for video surveillance applications," J. Visual Commun. Image Represent., vol. 20, no. 6, pp. 428–437, Aug. 2009.

[4] Fatih Porikli, Faisal Bashir, and Huifang Sun, "Compression domain Video Object Segmentation," IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 1, pp. 2–14, Jan. 2010.

[5] Pei Dong, Yong Xia, Li Zhuo and Dagan Feng, "Real-time moving object segmentation and tracking for H.264/AVC surveillance videos," in Proc. IEEE Int. Conf. Image Processing, Brussels, pp. 11–14, Sep. 2011.