World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:12, No:11, 2018

Real Time Object Tracking in H.264/ AVC Using Polar Vector Median and Block Coding Modes

T. Kusuma, K. Ashwini

Open Science Index, Computer and Information Engineering Vol:12, No:11, 2018 publications.waset.org/10009772.pdf

***Abstract*—**This paper presents a real time video surveillance system which is capable of tracking multiple real time objects using Polar Vector Median (PVM) and Block Coding Modes (BCM) with Global Motion Compensation (GMC). This strategy works in the packed area and furthermore utilizes the movement vectors and BCM from the compressed bit stream to perform real time object tracking. We propose to do this in view of the neighboring Motion Vectors (MVs) using a method called PVM. Since GM adds to the object's native motion, for accurate tracking, it is important to remove GM from the MV field prior to further processing. The proposed method is tested on a number of standard sequences and the results show its advantages over some of the current modern methods.

***Keywords*—**Block coding mode, global motion compensation, object tracking, polar vector median, video surveillance.

I. INTRODUCTION

REAL time object tracking has significant role in many applications nowadays such as video communication, human computer interaction, video compression, video surveillance, monitoring and security etc. Especially with increasing demands in surveillance applications, tracking in compressed video streams is essential for memory administration and time and speed necessities over the system for the constant situations. Adaptable video coding (AVC), an augmentation of H.264/AVC, empowers the video administrations with low resolutions yet with high reconstruction quality for decoding [1]. Our proposed tracking algorithm makes use of BCM information and MVs information from the H.264/AVC compressed bitstream, which could be used to enhance region analysis and scene interpretation in the compressed domain [13]. In the literature, however, even systems that fully decode the bit stream are sometimes referred to as compressed domain approach [14], [15], on condition that pixel values are not recovered for all frames. The pixel information leads to lower accuracy, but the advantage of compressed domain methods in practical applications is their generally lower computational cost. This is due to the fact that part of decoding can be avoided. In video, the viewer's attention usually remains on a salient region for several consecutive frames.

The proposed method uses only the MVs and BCM from the compressed bit stream to perform tracking. In addition, since most video contents nowadays are only accessible in the compacted shape, unraveling is required with the end goal to produce pixel space data. Then again, "compacted area" approaches make utilization of the information from the packed video bit stream, for example, MVs square coding modes, movement repaid forecast residuals or their change coefficients [2]-[6]. It is outstanding that movement strongly affects saliency in unique scenes. Consequently, various saliency estimation strategies have been recommended that depend on transient changes [16]. The resultant data streams still compose a valid data stream but with comparatively low frame resolution or data rate.

This paper presents real time object tracking method that uses only "MVs and BCM" information to perform fast and fairly accurate tracking. The method assumes a video bitstream coded using H.264/AVC [14].

II. MV-BASED MOTION

MV is a segment in the movement estimation process in half and half video encoding framework to show the situation of current Macro Block (MB) in the reference outline. Consequently, by encoding this pointer and some level of distinction information to compensate the mistakes, the encoder can acquire well pressure productivity. MV comprises of two sections: flat and vertical part. MBs estimate is regularly 16 by 16 pixels and can be separated into segments to expand pressure effectiveness. H.264/AVC defines 4 segments (16x16, 16x8, 8x16, and 8x8) i.e. basic MBs and 3 sub-segments for 8x8 segments (8x4, 4x8, and 4x4). We identify movement event in video scene by utilizing a technique that arranges MBs in a casing dependent on the estimation of MVs of every MB as proposed [3].

At the point when a moving object goes through a specific locale in the scene, it will produce unique MVs in the comparing spatio-worldly neighborhood, a few MVs will relate to the foundation, others to the objection itself, and Mvs themselves might be extremely not the same as each other, particularly if the dispute is adaptable. Then again, a region of the scene secured altogether by the foundation will have a tendency to have steady MVs, caused for the most part by the camera movement. MVs in a given spatiotemporal neighborhood could be utilized as marker of the distance of moving objects, which thus indicates potential for pulling in consideration.

The aim of this paper is to provide a framework for tracking moving objects in compressed domain based on MVs and associated BCM alone. The object of interest is selected by the user in the first frame, and then tracked through subsequent frames [12]. The standard vector median would be the vector

---

Kusuma T. Research scholar, and Dr Ashwini K. Research Supervisor are with the Department of Computer Science is with the Global Academy of Technology, Bangalore, Karnataka, India (e-mail: kusuma_t@ yahoo.com, dr.ashwini.k@ gat.ac.in).

from the list with minimum total distance and the existence of outliers, however, has an adverse effect on the vector median. We offer to do this based on the adjacent MVs using a method called PVM.

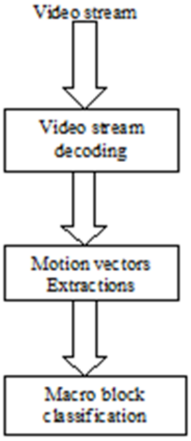

Fig. 1 MV-based Motion Detections

III. PVM FOR INTRACODED BLOCKS

Intra-coded blocks have no related MVs. For this purpose, we employ MVs of the first-order neighboring MBs (North, West, South, and East) that are not intra-coded. Fig. 2 shows a sample intra-coded MB along with its first-order neighboring MBs. In this example, the goal is to find **v***(***b***)* for all blocks **b** in the intra-coded MB [4]. We collect MVs of the 4 ×4 blocks from the neighboring MBs that are closest to the current intracoded MB and store them in the list *V*. For this example, the list of MVs is

V = *(***v**1, **v**2, **v**3, **v**3, **v**4, **v**4, **v**5, **v**5, **v**6, **v**6, **v**6, **v**6*)*

In PVM, the representative vector is computed in polar coordinates as follows: Let V= $((v_i)_{i=1:n}$ be the list of input vectors, sorted according to their angle from –π +π. Then a collection of m= [(n+1)/2] vector is selected as $\widehat{\mathrm{V}} = (\mathrm{v_i})_{\mathrm{i=I:I+m-1}}$, where index I is found as

$$\mathrm{I}= j\sum_{i=j}^{j+m-2}\theta_i \tag{1}$$

and θ_i denotes the angle between vectors v_i and v_{i+1}

The next stage is to allot a negotiator vector from this addition of vectors. The standard vector middle would be from the bad condition with the base total separation to different vectors in the in bad condition. The presence of exceptions, be that as it may, has a surly impact on the vector middle.

The PVM $\hat{V}$ contains approximately half the original number of vectors in V chosen such that the sum of angles between them is minimum. Hence, they are clustered in a narrow beam. The PVM $\hat{v}$ is constructed as follows: its angle is chosen to be the median of angles of the vectors in $\hat{V}$, while its magnitude is set to the median of magnitudes of the vectors in V, that is

$$\llcorner\hat{v} = median(\llcorner v_i)_{i=I:I+m-1} \tag{2}$$

$$\parallel \hat{v} \parallel_2 = \mathrm{median}(\parallel v_i \parallel_2)_{i=1:} \tag{3}$$

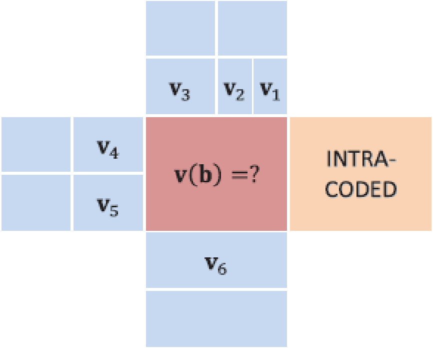

Fig. 2 One of the first-order neighboring MBs is also intracoded, and the remaining neighbors have MVs assigned to variably sized blocks

Algorithm 1:

```
Input  x: (array) x-component of a set of vectors
  y: (array) y-component of a set of vectors
Output
  med_x: (real value) x-component of output vector
  med_y: (real value) y-component of output vector
Step1: define constants
Step 2: compute the length of polar vector median
Step 3: compute the angle of polar vector median
Step 4: discard vectors with zero (or close to zero) length
Step 5: calculate difference angles in clockwise direction
Step 6: calculate polar median filtered vector
```

Fig. 3 shows the effect of assigning PVM of neighboring blocks to intra-coded blocks on frame of the Hall Monitor sequence [7]-[11].

Fig. 3 Intracoded blocks indicated by yellow squares

The intra-coded blocks are indicated in Fig. 3, the labeling with zero MV assignment to intra-coded MBs is shown in Fig. 4, while the labeling with PVM assignment to intra-coded MBs is shown in Fig. 5. All pixels in a given block are assigned the label of that block. By comparing with the manually segmented ground truth, one can identify correctly labeled object pixels (true positives - TP), non-object pixels incorrectly labeled as object pixels (false positives - FP), and missed object pixels that are labeled as non-object pixels (false negatives - FN). It is easy to see that detection is improved by

PVM assignment around the man's feet in the bottom parts of Figs. 4, 5

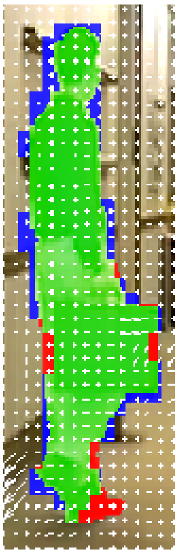

Fig. 4 Tracking result without PVM assignment

MVs are commonly created around regions of movement. In any case, MVs, some of the time, exist in territories in which there are no moving items. MVs do not generally speak to genuine movement since they are created to enhance the pressure proportion. We need to pick the right MV. Fig. 4 indicates macroblocks where MVs exist. The filled region speaks to where MVs are beyond threshold. At the point when the objective individual is strolling, MB around the individual has an extensive MV. On the correct side, there is no moving object yet rather there are huge MVs, or MV noise.

Fig. 5 Tracking result with PVM assignment. True positives (TPs) are shown as green, false positives (FPs) as blue and false negatives (FNs) as red

TABLE I
TOTAL AVERAGE OF PRECISION, RECALL, AND F-MEASURE (IN PERCENTAGE) OF PROPOSED METHOD

Precision	Recall	F -Measure
73.8	86.9	79.3
49.9	92.6	63.3
82.5	66.4	72.5
82.1	67.3	73.1
71.5	82.1	76.4
89.3	88.7	88.8
90.2	95.9	92.9
86.3	90.5	88.2

Average: Precision = 76.9, Recall = 85.4, F measure = 79.

VI. GMC

Global Motion (GM), caused by camera movement, affects all pixels in the frame. Since GM adds to the object's native motion, for accurate tracking, it is important to remove GM from the MV field prior to further processing. Given the parameters of the affine model, $[m_1 \ldots \ldots m_6]$, a block centered at (x, y) in the current frame will be transformed to a quadrangle centered at (x', y') in the reference frame, where (x', y') is given by

$$\begin{aligned} x^1 &= m_{1+}m_2x + m_3y \\ y^1 &= m_{4+}m_5 + m_6y \end{aligned} \tag{4}$$

The MV due to affine motion is given by

$$\mathrm{V(x, y)} = (x' - x, y' - y) \tag{5}$$

This method reduces the influence of outliers in a reweighted iteration procedure based on the estimation error obtained using least squares estimation. This iterative procedure contains until convergence. GMC is based on perspective transformation Algorithm [9]-[11].

Algorithm 2.

```
Step1. define constants (smallest block size in H.264)
Step2. center coordinates in image size only keep large size blocks
       and discard small size blocks and skip/intra coded blocks
Step3. remove the region of object (imagine the object in the
       previous frame has the same position in the current frame)
Step4. reduce weights for those blocks which are different from their
       neighbors
Step5. if there is not enough blocks to estimate GME then use all
       blocks
Step6. discard those blocks which we ignored it before
Step7. only use those blocks which might have background motion
Step8. if there is not enough blocks to estimate GME stop iterative
       estimation
```

In our approach, for the purpose of GME, we also discard the MVs [8] from the region that was occupied by the object in the previous frame. To get fast convergence to a stable solution and escape from being trapped in many local minima, we initialize the weighing factor for i-th MV based on its dissimilarity to neighboring MVs.

Fig. 7 Target at frame of *Coastguard* superimposed by scaled MV field after GMC

V CONCLUSION

We have presented an approach to track moving objects in a H.264/AVC-compressed video. BCM and MVs are the proposed method, the H.264/AVC compressed stream is used to perform tracking process. As a result, the proposed method has a fairly low processing time and provides high accuracy

ACKNOWLEDGMENT

The authors are grateful for having the opportunity to experience a wonderful time at Global Academy of Technology, establishing research facility, and an exceptionally friendly environment.

REFERENCES

[1] A. Borji and L. Itti. State-of-the-art in visual attention modeling. IEEE Trans.Pattern Anal. Mach. Intell., 35(1):185-207, 2013. 2, 5, 9, 24, 25
[2] V. Mahadevan and N. Vasconcelos. Spatiotemporal saliency in dynamic scenes. IEEE Trans. Pattern Anal. Mach. Intell., 32(1):171-177, 2010. 35, 41, 61, 93, 88.
[3] Z. Liu, H. Yan, L. Shen, Y.Wang, and Z. Zhang. A motion attention model based rate control algorithm for H. 264/AVC. In The 8th IEEE/ACIS International Conference on Computer and Information Science (ICIS'09).
[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the H. 264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol.
[5] S. H. Khatoonabadi and I. V. Bajic. Still visualization of object motion in compressed video. In Proc. IEEE ICME'13 Workshop: MMIX, 2013.
[6] G. Abdollahian, Z. Pizlo, and E.J. Delp. A study on the effect of camera motion on human visual attention. In Proc. IEEE ICIP'08, pages 693-696, 2008.
[7] Tracking database for standard video sequences. http://www.sfu.ca/~ibajic/ datasets.html. [Online]..
[8] Y. M. Chen, I. V. Baji´c, and P. Saeedi. Motion segmentation in compressed video using Markov random fields. In Proc. IEEE ICME'10, pages 760–765, 2010. 96.
[9] A. Smolic, M. Hoeynck, and J. R. Ohm. Low-complexity global motion estimation from P-frame motion vectors for MPEG-7 applications. In Proc. IEEE ICIP'00,volume 2, pages 271-274, 2000..
[10] W. Zeng, J. Du, W. Gao, and Q. Huang. Robust moving object segmentation on H.264/AVC compressed video using the block-based MRF model. Real-Time Imaging, (11):290–299, 2005.
[11] M. G. Arvanitidou, A. Glantz, A. Krutz, T. Sikora, M. Mrak, and A. Kondoz. Global motion estimation using variable block sizes and its application to object segmentation. In Proc. IEEE WIAMIS'09, pages 173{176, 2009. 67, 104
[12] J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. Proceedings of the IEEE,78(4).
[13] 264/AVC reference software. http://iphome.hhi.de/suehring/tml/. (Online).