

Real Time Object Tracking based on Inter-Frame Coding: A Review

Shraddha Mehta¹, Vaishali Kalariya²

RK. University, Rajkot

Abstract

Inter-frame Coding plays significant role for video Compression and Computer Vision. Computer vision systems have been incorporated in many real life applications (e.g. surveillance systems, medical imaging, robot navigation and identity verification systems). Object tracking is a key computer vision topic, which aims at detecting the position of a moving object from a video sequence. The application of Inter-frame Coding for low frame rate video, as well as for low resolution video. Various methods based on Top-down approach just like kernel based or mean shift technique are used to track the object for video. So, Inter-frame Coding algorithms are widely adopted by video coding standards, mainly due to their simplicity and good distortion performance for object tracking.

Keyword

Inter-frame Coding, Computer Vision, Low Frame Video, Resolution.

1. INTRODUCTION

Object tracking is a key computer vision topic, which aims at detecting the position of a moving object from a video sequence. Motion detection was initially developed for video encoding. Motion vectors can be used to predict changes in the scene between two or more video frames. The size of video data is reduced by encoding only the current frame and the motion vectors, from which several future frames can be recovered. Motion detection for object tracking has distinctly different requirements. In object tracking, there is a need to interpret the information given by the motion vectors. In Video compression algorithms all background changes, moving objects, and camera motion need to be encoded, but the meaning of the motion vectors is not important. In Object tracking have to detect only the moving objects and filter out the noise, small camera motions, and insignificant other motions (e.g. rain drops or falling leaves). Grouping or some other kind of association of the motion vectors into a meaningful scene representation is the primary goal of motion detection for object tracking.

2.MODEL OF INTERFRAME CODING FOR OBJECT TRACKING

2.1 Motion Segmentation:

The first task is to distinguish the target objects in the video frame. Most pixels in the frame belong to the background and static regions, and suitable algorithms are needed to detect individual targets in the scene.

2.2 Noise removal:

Segmentation of the image frame into moving and non-moving pixels, region growing is used to locate and identify each moving object. Each group of pixels that are connected to each other is labeled as one blob. An envelope is established around each object of interest by use of morphological erosion and dilation filters. The erosion-dilation operation also removes blobs that are very small and thus likely to be noise. Additional noise is filtered by ignoring blobs less than S pixels in size.

2.3 Representation of a blob:

The variables and their description in the data structure used for representing an individual blob Blob-ID is the number that is given to a blob. The track number of the blob, as assigned by the tracking algorithm, is stored in this variable. The meanR, meanG and meanB variables store the average intensity values of red, green, and blue colors of all pixels belonging to the blob.

2.4 Representation of a track:

Track-ID stores the unique track number for the track. Starting-Frame and ending-Frame variables store the first and last frame in which the track was seen, respectively. Lost Count Variable keeps count of how many consecutive frames for which a track has been 'lost'. This is kept so that a track that is lost for more than 5 frames can be deleted and closed.

2.5 Object Tracking:

It performs object tracking by finding correspondences between tracks in the previous frame and blobs in the current frame.

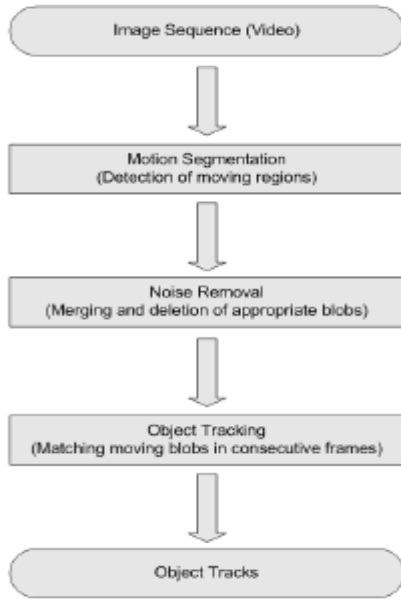


Figure 1. Model Of Inter-frame Coding For Object Tracking

3. SEGMENTATION

Segmentation defines the divide the image into smaller parts. There is mainly 2 types of segmentation that is Foreground Segmentation and Background Segmentation.

3.1 Foreground Segmentation

Detection of object to be tracked is the primary step in object tracking process. The object can be detected either once in the first frame or in every frame in the video. The goal of segmentation is to find out the semantically meaningful regions of an image and cluster the pixels belonging to these regions. It is very expensive to segment all static objects in an image. However, it is more practical to segment only the moving objects from video using spatial-temporal information in sequence of images (frames). A segmentation method should be generic and should not depend on other factors like color, shape and motion. Also, segmentation method should not computationally intensive and should require less memory.

3.2 Background Subtraction

The background subtraction method is one of the very simple and promising approaches for extracting moving objects from video sequences. In background subtraction approach, we compare current frame with a reference frame known as background image. A significant difference indicates the presence of moving objects.

4. TOP-DOWN APPROACH

Top down Approach is goal oriented and find the positions of the object in the current frame using a hypothesis generated at the start of the tracking based on parametric representation of the target.

The top-down approach uses different block sizes in a manner similar to image pyramids. The basic inter-frame coding is applied using a large block size. The goal is to have the entire object covered only with few blocks. Then the blocks size is reduced by half in each dimension and the block matching is applied again but only for those blocks for which motion was detected during the first step. This can be repeated as many times as desired or until the block size becomes 1*1 pixels.

So, In Short in this approach object has Large Block just like 1*1 pixels then it divided into 2*2 pixel size of blocks and then divided into 4*4 pixel size of blocks in that entire Object is covered.

So, for using the Top Down Approach we can use the formula for n number of levels is : $2^n \times 2^n$. where $n=1,2,3,4,\dots$ Depends on the how much the object is large for calculate the motion.

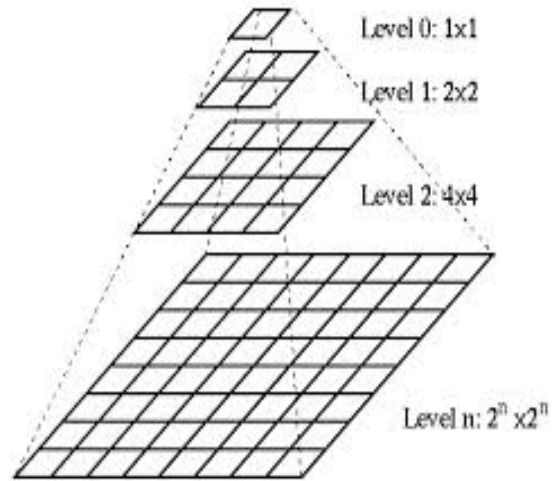


Figure 2 : Top-Down Approach

The top down approaches often rely on external input to initialize the tracking process. These tracking methods use different object characteristics, such as color, texture, shape and motion. In which we only focus on Motion in this paper. One of the popular method in this category is mean-shift based tracking or kernel based tracking. A regular pyramid is a hierarchy. Each level λ contains an array of cells. A cell of a regular pyramid is determined by its position (i, j, λ) in the hierarchy, (i, j) are its coordinates within the level λ . The cells on $\lambda=0$ (base level) are either directly the pixels of the input image or the result of any

local computation, like filters, on the image. We obtain each pyramid level recursively by processing the level below.

The reduction window gives us the children-parent relationships. Each cell in level $\lambda + 1$ has a reduction window of $N \times N$ children at level λ .

5. MEAN-SHIFT TECHNIQUE

While mean-shift based technique we use number of points around of one dot detect any possible translation and uniform scaling transformations. And one point inside alerts us when we lose the object which detects and estimates the changes in the object movement. We assume a textured object with continuous and smooth motion. We consider translation, rotation and uniform scaling transformations.

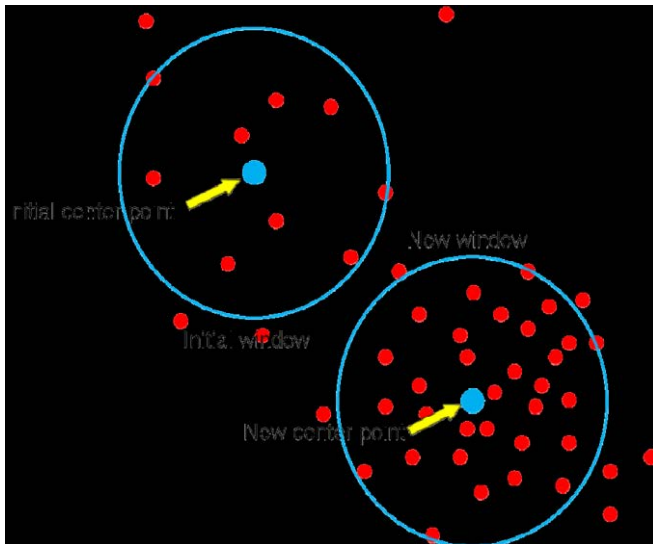


Figure 3 : Mean Shift Technique

6. PROCESS OF MEAN SHIFT TRACKING

Start from the position of the model in the current frame
Search in the model's neighborhood in next frame
Find best candidate by maximizing a similarity function
Report the same process in the next pair of frames

7. EXAMPLE OF OBJECT TRACKING

In this Example Object is considered as a man which is moving while the ball is coming in front of him. So in this only man's face is tracked for 58 Seconds. So, this example is useful to understand the object tracking using mean-shift Technique.

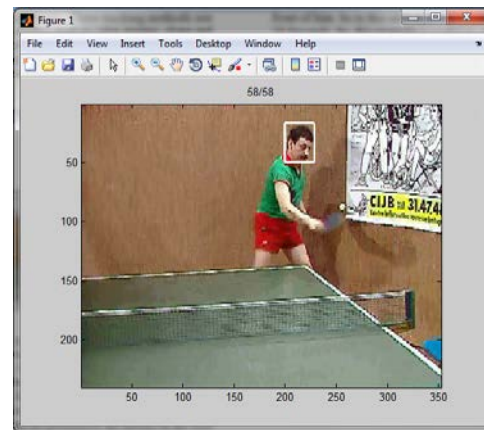
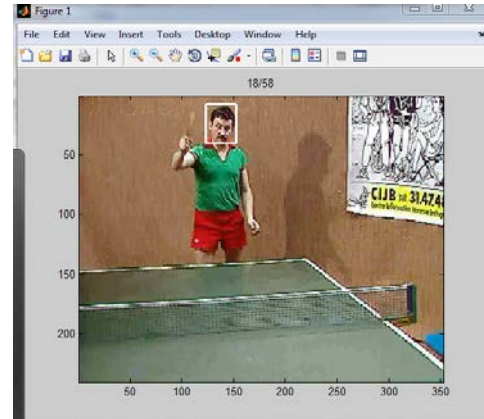


Figure 4. Tracking results: Video sequences - Tracks 18 and 58 through a set of frames

8. RESULT OF MEAN SHIFT

Experiment was performed to assess the tracking performance of Mean Shift Tracking. The mean shift based tracker proved to be robust to partial clutter and camera motion. Since no motion model has been assumed, the tracker adapted well to the non-stationary character of the pedestrian's movements. It starts from the position of the model in the first frame and then searches in the model's neighborhood in next frame, followed by finding best candidate by maximizing a similarity function. And then repeats the same process in the next pair of frames.

Table 1. Extracted Features For Tennis Video Sequences

(Frame, Object)	Area	Width	Height
(1,1)	250*350	6	160
(18,1)	250*350	4	170
(40,1)	250*350	10	190
(47,1)	250*350	15	210
(58,1)	250*350	20	240

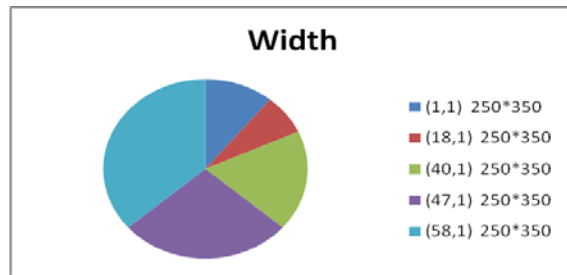


Figure 5. Visual Tracking Representation

9. CONCLUSION

Significant progress has been made in object tracking during the last few years. Several robust trackers have been developed which can track objects in real time in simple scenarios. However, it is clear from the papers reviewed in this survey that the assumptions used to make the tracking problem tractable, for example, smoothness of motion, high contrast with respect to background, etc., are violated in many realistic scenarios and therefore limit a tracker's usefulness in applications like automated surveillance, human computer interaction, video retrieval, traffic monitoring, and vehicle navigation.

10. FUTURE WORK

While the Texture of Object and Background have Uniform Color Then it will difficult to track that object with motion estimation So, It will be try to solve using top-down approach.

REFERENCES

- [1] cigdem Ero,glu Erdem and Bülent San - Video Object Tracking With Feedback of Performance Measures - IEEE Transactions on circuits and systems for video technology, vol. 13, no. 4, April 2013.
- [2] ElAzim, S.A., "Efficient Object Tracking Technique Using Block-Matching Algorithm" IEEE – 2012.
- [3] Mahbub Murshed ,MD.Hasanul Kabir ,Oksam Chae , "Moving Object Tracking-An Edge Segment Approach" International Journal Of Innovative Computing Information And Control-2011.
- [4] Manjunath,Narayana and Donna Haverkamp. A Mean-shift algorithm for tracking multiple moving objects in outdoor surveillance video. In Fourth Joint IEEE International Workshop on Object Tracking and Classification in and Beyond the Visual Spectrum, In conjunction with IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2011, June 2011.
- [5] Karthik Hariharakrishnan, Dan Schonfeld, "Fast Object Tracking Using Adaptive Block Matching" IEEE – 2010.
- [6] Chandrika Kamath ,Sen-ching Cheung, "Block Matching for Object Tracking", Lawrence National Laboratory Livermore, 2010.

- [7] Y. Fu, T. Erdem, and A. M. Tekalp, "Tracking visible boundary of objects using occlusion adaptive motion snake," IEEE Trans. Image Process, vol. 9, no. 12, pp. 2051–2060, Dec. 2009.
- [8] D. Schonfeld and D. Lelescu, "VORTEX: video retrieval and tracking from compressed multimedia databases multiple object tracking from MPEG-2 bitstream," J. Vis. Commun. Image Represent. (Special Issue on Multimedia Database Management), vol. 11, pp. 154–182, 2009.