Motion Estimation using Mobile Camera From Video

Sarita P. Shinde
Department of Electronics and Telecommunication Engineering
MIT College of Engineering, Pune, Maharashtra, India

Prof. A. R. Askhedkar
Department of Electronics and Telecommunication Engineering
MIT College of Engineering, Pune, Maharashtra, India

Abstract- The typical first step in visual surveillance is capability of extracting moving objects from a video sequence captured using a static camera. For automatic initialization, tracking, pose estimation, and movement recognition the numbers of significant research advances are identified together with novel methodologies. There are mainly three algorithms for the motion detection i.e voting based motion estimation, temporal difference, and background subtraction. Out of these algorithms here I use the voting based motion estimation algorithm. This algorithm is to correctly estimate the motion of moving object. To estimate the camera movement the shifting information of edges of static background utilized by this algorithm without knowing the prior knowledge of camera motion. On the Several representative region of interest the shifting information can be identified using voting decision method. Here we implement the voting based motion estimation algorithm by estimation, compensation, moving edge correction & enhancement and moving object detection..

Keywords- spatial coding, temporal sampling, Motion detection Visual surveillance system, voting based motion estimation.

I. INTRODUCTION

For improving public safety and security, the demands for surveillance-related information are growing. In indoor and/or outdoor monitoring the video-based surveillance plays an important role for guaranteeing the quality and security of human life. Cameras deployed in the area to sense and monitor possible targets are present in the video based surveillance system, for the purpose of recording and/or analyzing. In numerous areas which need to be monitored, including transportation, industrial, commercial and areas the networked video system can be applied. Mobile camera system is an emerging trend for obtaining wider and more accurate visual monitoring data, using multi-camera. Depending on the characteristics of the moving objects existing in surveillance video the coding and compression of the surveillance video are performed. A down-sampling mechanism depends on the image contents characteristics and is then performed on image sequences for efficiently discarding the similarity within the video data. In the video-based surveillance system spatial coding, the detection of moving objects must be accurate to have desirable result.

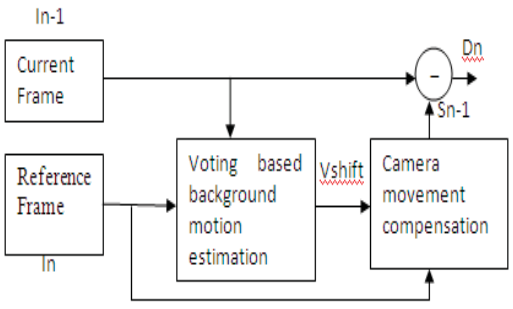

Fig.1. Voting based motion estimation and compensation

In Fig. 1. In-1, In, Sn-1 and Vshift denote the reference frame, the current frame, the frame after the compensation of motion, and the movement of camera, respectively. The motion vectors i.e moving vectors of edges of static background to estimate the motion of camera are utilized by the voting-based background motion estimation algorithm. There is no necessary to estimate the parameters like motion of camera model, such as rotation angles, inclination. Using optical flow the computation complexity and time are lesser. Various sampling techniques in temporal and spatial domains are implemented which depends on the characteristics of the moving objects and the available bandwidth of transmission network. On the quantity of moving objects the temporal sampling techniques mainly depends. Finally, a priority-based spatial coding technique is used to encode the visual information of the temporally chosen frames. The primary goal of the overall technique is to remove the redundant information from one image frame and to efficiently allocate the available network bandwidth for progressively transmitting the most important part of the image frame

II.CHALLENGES IN MOTION DETECTION WITH MOBILE CAMERA

The major challenge in video surveillance is detection of moving objects in the presence of motion of camera. It is very easy to find out the moving objects from static image or video but it is very difficult to find out moving objects with its moving direction from dynamic video or moving video.

III.CAMERA MOTION ESTIMATION AND DATA REDUCTION

There are some main processes involved in camera motion estimation and data reduction Camera Motion Estimation and Compensation, Content-Based Temporal Sampling, Priority-Based Spatial Coding. Each process poses many challenges in camera motion detection after reference image has shot, the algorithm runs until the detection of the motion. Algorithm has two images to understand whether there is considerable change that can be considered as a motion in the projected vision or not. The calculation is done in RGB format data of the images in byte arrays. Two images firstly processed to see the general change in the projected vision. This general change must not effect the image difference calculation since this change is not from a motion. This difference is stored to use in pixel difference calculation part of the motion detection algorithm. This general difference will be called correction after here. Two byte arrays are compared to see if there is considerable change in the pixels. Considerable change is determined by a threshold value which is called pixel threshold. If the difference between two pixels is greater than the pixel threshold it is then compared with the correction value. Then according to the result the pixel is labeled as black if it has changed or labeled as white if it has not changed. Now it can be seen the different pixels in the real-time image.

The new image which keeps an image of difference is our reference now and then. This image is processed to catch a motion of an entire body rather than individual minor pixels. This process of the motion detection algorithm is called blob calculation. Blob calculation is calculates the entire body of black blob size. This calculation is done by finding the radius of a blob. This radius of the blob is compared with another threshold value which is called blob threshold. This value is determined according to what sensitivity do you want to achieve on the motion detection system. In our motion detection system, this sensitivity is defined by the user of the mobile phone and passed as an argument to the algorithm. Whenever a blob with a greater radius than blob threshold, the algorithm finishes the process telling the upper layer there is a motion in the projected area. We briefly review and discuss these challenging research issues as mentioned below.

A. Camera Motion Estimation and Compensation

In this subsection, the concept and implementation of the proposed voting mechanism for estimating camera motion is given. Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence. It is an ill-posed problem as the motion is in three dimensions but the images are a projection of the 3D scene onto a 2D plane. The motion vectors may relate to the whole image or specific parts, such as rectangular blocks, arbitrary shaped patches or even per pixel. The motion vectors may be represented by a translational model or many other models that can approximate the motion of a real video camera, such as rotation and translation in all three dimensions and zoom. There are mainly two algorithms one is direct method and second is indirect method. On the outcome of the standard edge detection algorithm along with several fundamental morphological operations like Erosion and dilation, for correctly identifying the edges of moving objects the voting mechanism is based. Static objects as well as

background look like moving objects in the captured images due to the camera motion. Therefore, subtracting two frames might simultaneously obtain the edge information of both static objects and moving objects. The camera motion should be first estimated and the estimated motion is then used to compensate for the result of edge detection in order to correctly identify the visual information of moving objects. To enhance the edge detection of moving objects and to filter out that of static/background objects therefore morphological operations can also be used. The camera motion is decided based on the decision of voting of the motion vectors of pre-specified blocks in order to increase the estimation robustness.

Fig.2 shows the overall framework of the algorithm along with the moving edge correction and enhancement. The background motion of previous frame alignment can be done by the voting-based motion estimation and compensation In-1 with the current frame in. The background with the current frame has the same position information of its result frame Sn-1. Next, subtracting Sn-1 and In can obtain the frame difference Dn result. Dn could still contain slightly different information from the correct edge information because the background position in current frame in and compensated output Sn-1 might not be completely identical Therefore, the morphological operations erosion and dilation are used to filter out the differences and to emphasize the result of Dn. Dn is then turned into i.e. converted to a binary frame Rn . Then, a correct moving edge frame Mn can be obtained, by taking a logic AND operation between the current edge frame En and Rn.

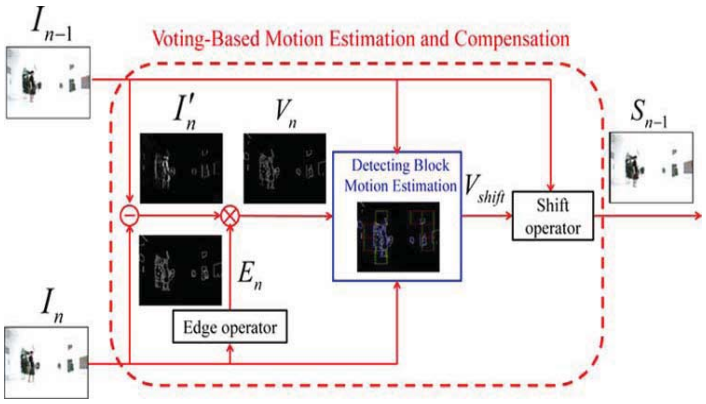

Fig.2. The flowchart of the voting-based motion estimation and compensation

B. content based temporal sampling

The motion of moving camera is relatively small, the successive captured images are almost identical When the frame rate is high enough. Hence, transmitting visual frames with temporally similar image content over a bandwidth-limited network is not efficient for bandwidth utilization. To transmit frames with the most important information about the monitoring area this is one way to tackle the issue. i.e., says N by selecting one frame from certain sequential frames. On the image content obtained from the voting-based motion estimation and compensation process the selection could based. When moving objects appear in the captured frames, a smaller value of might have a smoother video performance but lower visual quality under the requirement of the motion continuity of moving objects. A larger value of N might sacrifice the continuity of video performance for better video quality Most blurred frames might only have a small amount of edges. so, the frame with the largest number of edges among the successive frames is considered as the clearest one and the one with the most important information about the monitoring area. The frame selection can be done as follows. First of all, the standard Canny edge detection is performed to identify the location of the edge pixels. Then, count the number of pixels of changing edges on each frame and choose the largest number from sequential N frames

C. Priority-Based Spatial Coding

In spatial domain an image frame can be divided into two parts unimportant and important regions. On the result of the moving edge detection the importance can be decided. we can consider the most important information to the moving objects in the comparison with other stationary objects and background. Hence, in the result we can differentiate the regions with or without moving objects. Thus, to encode the region which have important information into a frame of higher visual quality and the region which don't have important information into a frame of lower visual quality the spatial coding algorithm can be used. so, an embedded coding algorithm, such as the set partitioning in hierarchical tree or SPIHT can be used to encode the visual quality based on the currently

determined importance and available bandwidth. The proposed spatial coding algorithm includes two parts: (1) set priority to regions, and (2) generate image code based on the priority.

The voting based motion estimation algorithm uses the following steps first the camera initialization we have to do then register the video for selecting the frames from n no. of frames. The reference frame will be the selected frame then the current frame will be taken. The comparison between the both frames will done. The name is voting based algorithm so the majority voting will done then direction offset detection and then translate and subtraction between the pixel values will occur. This result frame then applied to Object detection and tracking. Result can be shown by contours.

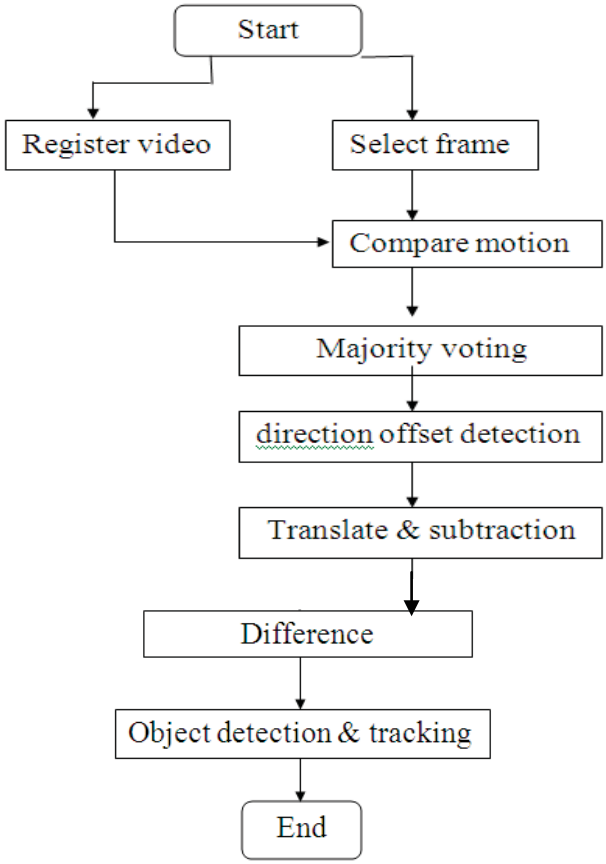

Fig. 3 Flow chart of voting based motion estimation

IV. RELATED WORK

In this Section, we describe various algorithms in the area of video segmentation especially gradual transition detection.

The moving object detection using mobile camera has been proposed by Feng-Li Lian, Yi-Chun Lin, Chien-Ting Kuo, and Jong-Hann Jean [1]. In this method, using moving camera the video is captured. The reference frame is taken first when the camera starts then current frame and the previous frame both are compared. Utilizing the frame differencing technique first the change detection from the captured images can be performed so the binary image can be generated by this technique that can be used to identify the area with significant difference between two frames or from the current frame to the background frame. Before performing the differencing technique, due to these methods it aligns the previous frames. A typically alignment process consists of two parts, background motion

estimation and motion compensation. For indoor surveillance the powerful and speedy VBME algorithm is proposed to estimate the camera motion without tracking features and knowing explicit knowledge of camera motion. Due to mobile camera the motion estimation utilizes the voting decision from the set of motion vectors which are determined by the edge feature of static objects and/or background to accurately estimate and compensate for the shifting motion. The voting-based motion estimation algorithm can decrease the computational complexity and avoid incorrect feature tracking due to only the edge characteristics are employed.

C. S´anchez-Ferreira, J. Y. Mori, C. H. Llanos proposed the algorithm for "background subtraction"[2]. The basic idea of background subtraction method is to initialize a background firstly, and then by subtracting current frame in which the moving object is present that current frame is subtracted with background frame to detect moving object. This method is simple and easy to realize, and accurately extracts the characteristics of target data, but it is sensitive to the change of external environment.

Kelson R. T. Aires, Andre M. Santana, Adelardo A. D. Medeiros proposed the optical flow algorithm for motion estimation [6]. By using color information the improvement in the optical flow estimation is done. Tasks like obstacle and movement detection in real-time and using optical flow can done. Optical flow is concerned to the movement in the eye that specifies the direction of locomotion. As the only low level data in Computational Vision applications, even with color information is available Luminance information is widely used. Using color prevents loss of information due to be luminance. Secondly in synchronization a bigger amount of data must be processed. A color image corresponds to color information can be used in optical flow estimation. In contrast to gray images case, the using of additional constraints is not necessary [6, 16, and 12]. It is not necessary to utilize all channels of a tri-chromatic image depending on the chosen color model. With the proposed method only using two color channels: YU, YV or UV. Tests have been made. Between the three options of choice the results can consists differences of 5% maximum in the number of valid measures. The optical flow method uses the motion target of the vector characteristics which changed with time to detect motion area in image sequences. It gives better performance under the moving camera, but this algorithm is very complex and complicated computation and also it needs special hardware support, so it is difficult to meet the requirements of real-time video processing

Y.-N. Li et al. have proposed a fast shot detection framework employing pre-processing techniques including thresholding and bisection-based comparisons to eliminate non-boundary regions[8]. The factors that lead to high detection speed in the proposed framework are three folds. Firstly, a large number of non-boundary frames are discarded before hard cut and GT detections. Secondly, the types of shot boundaries can be predicted in preprocessing stage, so that it can detect the preserved segments directly using the right detector. Thirdly, the burden of GT detection is eased, because the duration of each possible GT is available in advance and it is not necessary to calculate frame distances for multiple times. On Simulations and comparisons, significant speed up is achieved in the proposed framework, while the precision and recall rates can get in a satisfactory level.

Taeho Kim and Kang-Hyun Jo proposed how to generate background model and detect moving objects based on multiple background model [10]. The multiple background models is effectively estimate background scene for each frame, it has weakness when a background is temporally changed. use of temporal median filter. The result of detecting moving objects to overcome this problem, this method is good to detect moving object even though camera will move. However, this approach has disadvantages of camera shaking and moving objects. To overcome these disadvantages, find the 3-dimensional ratio for camera height vs. other objects. Camera rotation has been measuring to increase the performance of proposed algorithm. In spite of this limitation, proposed algorithm successfully generates multiple background models and detect moving object with low cost.

A fast coarse-to-fine video shot segmentation algorithm has been proposed by Liu and Jian-Xun Li [12]. The camera motion, object motion and gradual shot transition can be differentiated through this method. Based on the improved information entropy theory, the differences between the shots of a video sequence are calculated. The adaptive thresholds are implemented to select sequences of candidate shots from the video sequence. Because the camera/object motion and the gradual shot transition present similar characteristics, they are all selected as the candidate shots. Then a fast motion-edge detection algorithm is implemented to distinguish the gradual shot transition. The proposed algorithm is based on the statistical properties of the characteristics, hence compared to the single characteristic detection algorithm; the computational complexity is reduced effectively. This algorithm is reduces both the computational complexity and error detections caused by the camera/object motion effectively.

Yang Xu et al. have proposed 3- DWT based motion suppression for video shot boundary detection [13]. In this method adaptive threshold is selected, so that gradual transition and motion are adequately discriminated. Under the framework of 3-DWT framework dramatic motion are characterized. In the proposed method, motion intensity is extracted and motion suppression value(MSV) is defined, which is integrated into histogram based on edges based

methods for video shot boundary detection. This method is more efficient to detect the gradual transition and also solves the problem of motion information which suffers from noise and illumination.

Block matching algorithm is proposed by the Aroh Barjatya. This algorithm is one of those algorithms which are used to detect the motion estimation in video compression [14]. This implements and compares 7 different types of block matching algorithms which have range from the very basic Exhaustive finding to the recent advance adaptive algorithms like Adaptive Rood Pattern Search. The algorithm in this has been used in implementing various standards, ranging from MPEG1 / H.261 to MPEG4 / H.263. While ISO MPEG sets the standard. Motion estimation is the most computationally expensive and time-consuming process in the entire motion based video compression process. On reducing both of these side effects of motion estimation the research in the earliest decade has focused. The most popular and efficient of the various motion estimation techniques block matching techniques.

S. Araki, T. Matsuoka, N. Yokoya, and H. Takemura proposed the algorithm for Real-time tracking of multiple moving object contours in moving camera image sequences [16]. In this the new method is used for detection and tracking of moving objects from a moving camera image sequence using robust statistics and to find out active contour models for showing the moving object by contours. In the method, we assume that the apparent motion of the static background is dominant in an image sequence. The motion between two consecutive image frames can be represented by a transformation such as affine transformation for registering the static background. The kind of robust estimator useful for a data set including outliers we estimate affine parameters using LMedS (Least Median of Squares) method. by subtracting the previous frame transformed with estimated affine parameters from the current frame moving object regions can be detected. For tracking moving objects active contour models are also promising.

V. RESULTS

Results are obtained using MATLAB and video tool kit. Here camera is also used. When camera captures video it can be move in both direction. Such ad either from left to right or from right to left. Camera can be moved from in any one direction. From this moving video and camera, only the moving objects are identified. The edges of moving objects are identified. These moving objects are shown by contours in the results

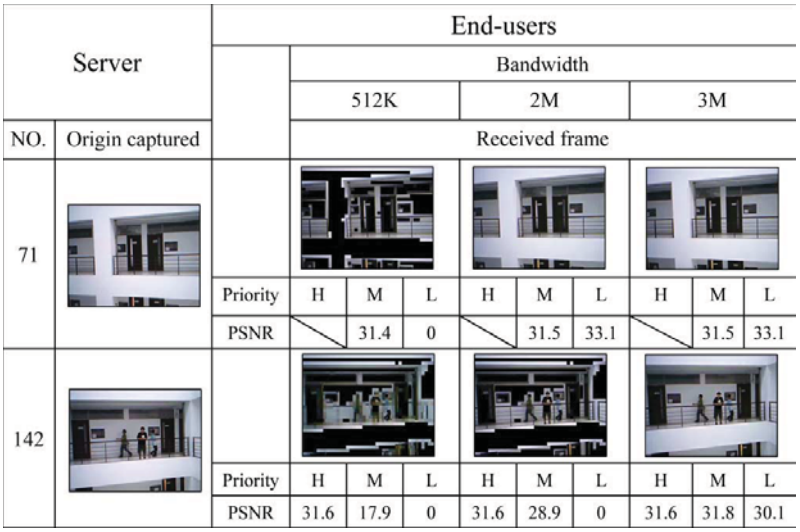

Fig. 4 Results

IV.CONCLUSION

The content-based video transmission mechanism is proposed to achieve video surveillance applications. The decision of voting over a set of detected motion vectors of edges is primarily used to estimate the shifting of static background induced by camera motion to correctly detect moving objects on the image frames captured from mobile camera. The image changes occurred by motion of camera can be compensated due to the shifting between current frame and reference frame. Two morphological operations, dilation and erosion, are used to enhance the

result of the moving edge detection. The temporal sampling criterion is based and the spatial coding process is based on the identified importance of image content on the detection of moving objects in the captured image frames. These two integrated mechanism of temporal sampling and spatial coding can simultaneously provide acceptable visual quality image frame.

REFERENCES

[1] Feng-Li Lian, Yi-Chun Lin, Chien-Ting Kuo, and Jong-Hann Jean" Voting-Based Motion Estimation for Real-Time Video Transmission in Networked Mobile Camera Systems" *IEEE transactions on industrial informatics*, vol. 9, no. february 2013

[2] C. S´anchez-Ferreira, J. Y. Mori, C. H. Llanos "Background Subtraction Algorithm for Moving Object Detection on FPGA" Department Mechanical Engineering University of Brasilia 2012

[3] S. Chen, J. Zhang, Y. Li, and J. Zhang, "A hierarchical model incorporating segmented regions and pixel descriptors for video background subtraction," IEEE Trans. Ind. Inf., vol. 8, no. 1, pp. 118–127, 2012.

[4] A. P. Dani, N. R. Fischer, and W. E. Dixon, "Single camera structure and motion," *IEEE Trans. Autom. Contr*., vol. 57, no. 1, pp. 241–246, Jan 2012.

[5] J.-H. Jean and F.-L. Lian, "Robust visual servo control of a mobile robot for object tracking using shape parameters," *IEEE Trans. Contr. Syst. Technol*., vol. 20, no. 6, pp. 1461–1472, Nov. 2012, doi 10.1109/ TCST.2011.2170573.

[6] Kelson R. T. Aires, Andre M. Santana, Adelardo A. D. Medeiros" OPTICAL FLOW USING COLOR INFORMATION PRELIMINARY RESULTS"

[7] P. Drews, R. de Bem, and A. de Melo, "Analyzing and exploring feature detectors in images," in Proc. 9th *IEEE Int. Conf. Industrial Informatics(INDIN)*, Lisbon, Portugal, Jul. 26–29, 2011, pp. 305–310

[8] Y.-N, Li, Z.-M, Lu, X.-M, Niu, "Fast video shot boundary detection framework employing pre-processing techniques" IET, Image Process, vol. 3, iss. 3, pp. 121–134, 2009.

[9] Jun Li, Youdong Ding, Yunyu Shi, Qingyue Zeng, "DWT-based shot boundary detection using support vector machine", Information Assurance and Security, vol. 1, pp. 435 – 438, 18-20 Aug 2009

[10] T. Kim and K.-H. Jo, "Segmentation of moving objects using multiple background model for industrial mobile robots," in Proc. 6th IEEEInt. Conf. Industrial Informatics (INDIN), Daejeon, Korea, Jul. 13–16, 2008, pp. 1492–1497

[11] Kelson R. T. Aires, Andre M. Santana, Adelardo A. D. Medeiros (2008). Optical Flow Using Color Information. ACM New York, NY, USA. ISBN 978-1-59593-753-7

[12] Naveen Aggarwal, Nupur Prakash, Sanjeev Sofat, "Gradual transition detection in digital videos using area correlation", *IEEE Region 10 Conference TENCON*, pp. 1-4, 14-17 Nov 2006.

[13] S. Lawrence, D. Ziou, and M.-F. Auclair-Fortier, "Motion insensitive detection of cuts and gradual transitions in digital video", Pattern ecognition and Image Analysis, Vol.14, iss. 2, pp 109–119, 2004.

[14] Aroh Barjatya, Student Member, IEEE" Block Matching Algorithms For Motion Estimation" DIP 6620 Spring 2004 Final Project Paper

[15] T. LU tong, P.N. Suganthan, "An accumulation algorithm for video shot boundary detection", Multimedia Tools and Applications, vol. 22, iss. 1, pp. 89–106, Jan 2004.

[16] S. Araki, T. Matsuoka, N. Yokoya, and H. Takemura, "Real-time tracking of multiple moving object contours in a moving camera image sequences," IEICE Trans. Inf. Syst., vol. E83-D, no. 7, pp. 1583–1591, Jul. 2000.