International Journal For Technological Research In Engineering
Volume 1, Issue 8, April-2014 *ISSN (Online): 2347 - 4718*

VIDEO IMAGE ENHANCEMENT AND OBJECT TRACKING

Mr. Abhilash Ghatol[1], Prof. Sourabh Sharma[2]
[1]M.Tech Student
Department of Electronics Engineering
Trinity Institute of technology and Research,
Bhopal, India.

Abstract: - To develop the real world computer vision system, detection of moving objects in video images is very important. The automatic detection of moving objects in video images is very important. The automatic detection of moving objects in monitoring system needs efficient algorithms. The common method is simple background subtraction i.e. to subtract current image from background. But it can't detect the difference when brightness difference between moving objects and background is small. The other approach is to use some algorithms such as color based subtraction technique but the costs are very high and have problem in stability. Here a method is proposed to detect moving objects using difference of two consecutive frames. The objective is to provide software that can be used on a pc for performing tracking long with video enhancement using bilinear interpolation. This project is able to track moving objects and it is structured as different blocks working together. Initially the spatial resolution and the contrast of the extracted frames of the video sequence are enhanced. The position of the object is now marked manually so as to obtain the "Region of Interest". The algorithm is implemented in MATLAB and the results demonstrate that both the accuracy and processing speed are very promising.

I. INTRODUCTION

Detection of moving objects in video images is one of the most important and fundamental technologies to develop the real world computer vision systems, such as video monitoring system, intelligent-highway system, intrusion surveillance, etc. Traditionally, the most important task of monitoring safety is based on human visual observation, which is a hard work for watchmen. Therefore, the automatic detection of moving objects is required in the monitoring system that can help a human operator, even if it cannot completely replace the human's presence. To facilitate a monitoring system, efficient algorithms for detecting moving objects in video images need to be used. The usual method for detecting moving objects is simple background subtraction that is to subtract current image from background image. However, there exist gradual illumination changes, sudden changes in illumination and other scene parameters alter the appearance of the background. Simple background subtraction is susceptible to these changes. And when the brightness difference between moving objects and the background is small, it cannot detect the difference. In order to resolve these problems, some algorithms such as color based subtraction technique and the technique based on optical flows have been proposed. But the computational costs of these methods are very high and have problem in stability. In our method Moving objects are detected from the difference of two consecutive frames. This approach uses the motion to distinguish moving objects from the background. So it is more efficient than the previous approaches. In our system, images are captured with a stationary camera. The experiment results demonstrate that both the accuracy and processing speed are very promising. Furthermore, the algorithm is robust to the changes of lighting condition and camera noise method and algorithm. Section III describes the video surveillance and its importance with the use of tracking. Section IV describes how tracking is useful for video surveillance and its application. Finally Section V presents our conclusions and future work.

II. RELATED WORK

There's a standard audio / video file format under Windows called AVI (Audio Video Interleave)."AVI" file is just a wrapper, a package that contains some audio / visual stuff,
But with no guarantees about what's inside. Microsoft created the AVI format for packaging A/V data, but it's just a specification for sticking A/V data in a file, along with some control information about what's inside.[5] It's sort of like having rules for putting boxes together ("insert flap A in slot B"), and rules for labeling the box, but anybody can put any sort of stuff in a box, in any language. You can make sense of the contents of a box only if you have the right translation kit. Each developer of a new A/V format is responsible for writing the translation kits that permit your Windows system to understand that flavor of AVI.

These kits are called "codecs", for compressor decompresses or, because the video and audio formats usually perform some form of compression to reduce the amount of data in the file. Video makes large files. A single image file can get big by itself, but video consists of frame after frame of images, which keep piling on one after another. To make video useful and shareable on PC's, the amount of data must be reduced to reasonable sizes. There are three basic approaches to reducing the amount of data in a video clip: reducing the resolution and length of the clip, trimming the individual pixels and color data, and compressing the frames. Each clever new approach to squeezing down the size of video data creates yet another incompatible format, and requires a new codec to play the resulting files.

Table 1. AVI Video Formats

Reduction	Video Format	Size (MB)	Reduce
Resolution	640x480, 30 fps, 24-bit	26.37	100%
	320x240, 30 fps, 24-bit	6.59	25%
	160x120, 30 fps, 24-bit	1.65	6%
Frame Rate	320x240, 15 fps, 24-bit	3.30	13%
	320x240, 10 fps, 24-bit	2.20	8%
Pixel Size	320x240, 10 fps, 16-bit	1.46	6%
Color Format	320x240, 10 fps, 9-bit	0.82	3%
	320x240, 10 fps, 12-bit	1.10	4%
Compression	320x240, 10 fps, JPEG 90%	0.29	1%
	320x240, 10 fps, JPEG 75%	0.18	1%

Various approaches to reducing the amount of data include:
- Resolution: Dropping the frame size from 640 x 480 to 320 x 240 to 160 x 120
- Frame rate: Lowering the frame rate from 30 to 10 frames per second
- Pixel size: Trimming the pixel depth from 32 to 16 bits per pixel
- Color format: Subsampling the color to 9- or 12-bit
- Compression: JPEG compression at different ratios

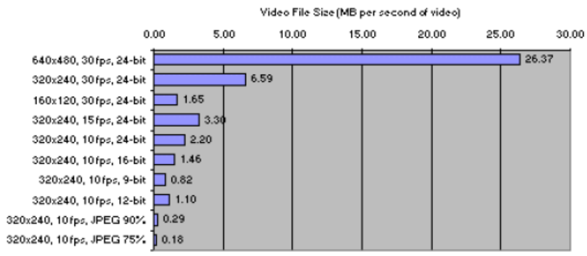


Fig. 1. Video File sizes: Chart of AVI video file sizes for a one-second video clip

But reducing the amount of data adversely affects the quality of the video. Keeping this in mind we have tried to obtain a better quality video by just increasing the spatial resolution of the individual frames.

I. *Resolution Enhancement using Bilinear Interpolation:*

Interpolation is a bit like gambling...guessing what something is going to be or to put it more mathematically, "calculating values between known pixels." An image consists of lots of pixels. For simplicity, let's stick to gray scale (black-and-white) images for now, where each pixel is simply one value - usually ranging from 0 to 255, where a higher value means more brightness, 0 being black and 255 being white. When the image is 400 x 300 pixels, and we want to enlarge it to 400%, it becomes 1600 x 1200 pixels. This means that from a mere 120,000 pixels in the original, we go to almost 2 million pixels in the enlargement. So our computer software has to "guess" what all those new pixels are going to be, based only on those 120,000 original values. The values in the original picture appear in the enlargement at every 4^{th} pixel. The newly introduced pixels, the ones in between, we have to fill in ourselves. This can be done in many ways. For example we can simply repeat each original pixel 4 times, in both horizontal and vertical directions; this is called "nearest neighbor”. This is of course a very simple and fast method, but also yields poor results (the typical jagged mosaic effect is glaringly obvious).

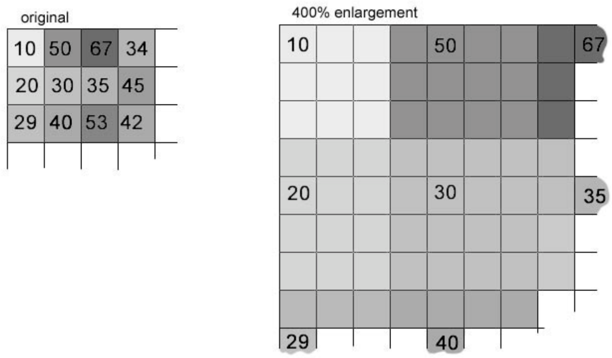


Fig. 2: 400% enlargement using interpolation

We can also gradually change the values from one to another. In the top row of the example image, the three unknown pixels between 10 and 50 would become (from left to right) 20, 30 and 40. Doing the same in vertical direction, we can fill up all the empty places. This is called "bilinear" interpolation. However, this technique gives better results but still it lacks preciseness, especially because the averaging of information results in adulates and loss of local contrast. Thus in order to be more precise we implement the bilinear interpolation along with minimization of the bending energy. Although this will lead to an increase in the complexity of the program but at the same time we will get better results with high quality.

A bilinear interpolation function can be written as

$$F(x,y) = ao + axx + ayy + axyxy$$

Now for each 2X2 block in the original image we will get a 4X4 block in the new image. We have to predict the values of the missing pixels, which can be done by matrix manipulation as shown.

$$\begin{pmatrix} F(x_1,y_1) \\ F(x_2,y_2) \\ F(x_3,y_3) \\ F(x_4,y_4) \end{pmatrix} = \begin{pmatrix} 1 & x_1 & y_1 & x_1y_1 \\ 1 & x_2 & y_2 & x_2y_2 \\ 1 & x_3 & y_3 & x_3y_3 \\ 1 & x_4 & y_4 & x_4y_4 \end{pmatrix} X \begin{pmatrix} a_o \\ a_x \\ a_y \\ a_{xy} \end{pmatrix}$$

From above equation we can calculate the coefficients { ao , ax , ay , axy } which are used to predict the missing values as shown below.

1	**3**
4	**2**

1.0000	1.6667	2.3333	**3.0000**
2.0000	2.2222	2.4444	2.6667
3.0000	2.7778	2.5556	2.3333
4.0000	3.3333	2.6667	**2.0000**

Fig. 3: Implementing bilinear interpolation

The results of the two interpolation methods are shown below.

Fig. 4: Nearest neighbor: Result not Satisfactory

Fig. 5: Bilinear: Result obtained is much better

III. OBJECT DETECTION

Detection of moving objects in video images is one of the most important and fundamental technologies to develop the real world computer vision systems, such as video monitoring system, intelligent-highway system, intrusion surveillance, etc. Traditionally, the most important task of monitoring safety is based on human visual observation, which is a hard work for watchmen. Therefore, the automatic detection of moving objects is required in the monitoring system that can help a human operator, even if it cannot completely replace the human's presence. To facilitate a monitoring system, efficient algorithms for detecting moving objects in video images need to be used.

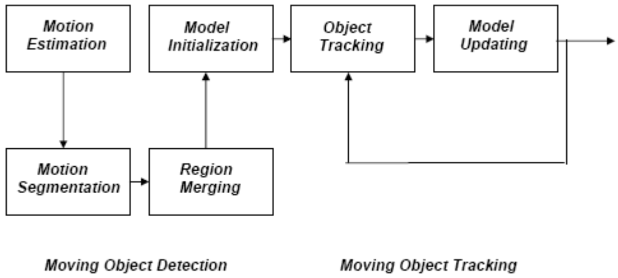


Fig 6. General Model for object detection and tracking

The usual method for detecting moving objects is simple background subtraction that is to subtract current image from background image. However, there exist gradual illumination changes, sudden changes in illumination and other scene parameters alter the appearance of the background. Simple background subtraction is susceptible to these changes. And when the brightness difference between moving objects and the background is small, it cannot detect the difference. In order to resolve these problems, some algorithms such as color based subtraction technique and the technique based on optical flows have been proposed.

But the computational costs of these methods are very high and have problem in stability. The other commonly used method for moving objects detection is frame difference. Moving objects are detected from the difference of two consecutive frames. This approach uses the motion to distinguish moving objects from the background. So it is more efficient than the previous approaches. Frame difference approach is robust to environmental changes, however, unable to detect motionless objects .Currently, all the systems using frame difference approach to detect moving objects, in which, the subtraction is done pixel-wise in luminance on the whole image, even though the object moves in relatively local range. In fact, instead of processing a whole image, only the neighborhood area around the moving object needs to be processed. So computation time can be reduced.

In this, a real-time algorithm for detecting moving objects in the image sequence is proposed, which integrates the region-based frame difference with adjusted background Subtraction. In our system, images are captured with a stationary camera. The region-based frame difference is used to extract the moving objects and the adjusted background subtraction is used to get the motionless foreground objects in the scene. The experiment results demonstrate that both the accuracy and processing speed are very promising. Furthermore, the algorithm is robust to the changes of lighting condition and camera noise.

The process sequence of the algorithm is shown in Fig. The algorithm is based on a region-based frame difference motion detection technique and adjusted background subtraction.

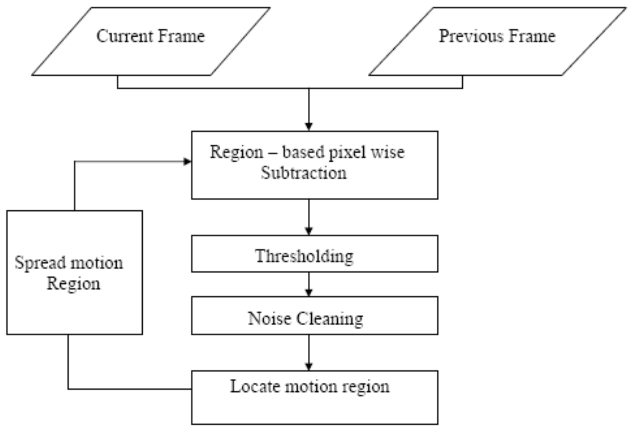


Fig 7. Block Diagram showing Object Tracking Algorithm

The purpose is to indicate the position of moving object in a video frame. By utilizing the region-based frame difference, it detects motion in the scene, and furthermore, it determines the position of moving regions. A good detection and location algorithm should be able to detect and track the suspicious objects even when they stop. So by utilizing the adjusted background subtraction, the algorithm can also detect motionless foreground objects in the scene. By using the algorithm, the moving and still objects in video images can all he detected and located.

A. *Region-based Frame DifferenceThe*

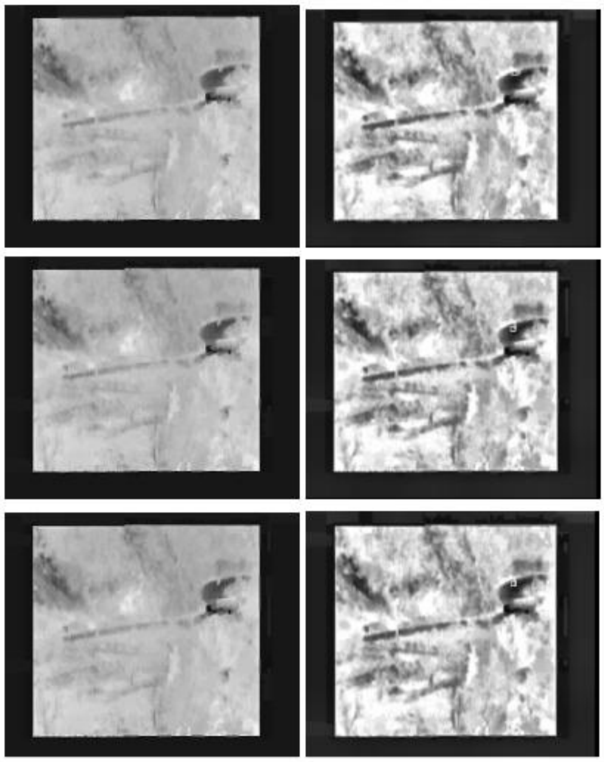
The first process of the proposed algorithm is the detection of moving objects. Frame difference technique is used, which is simple and yet powerful enough, to discriminate between moving objects and non-moving ones. The frame difference method is simply finding the absolute difference between two consecutive frames. In stream video sequences, the frame rate is more than one shot per second. Thus, if there is any object that is in motion, it will have a slight position change and the maximum change will occur at the edges of the image since the discontinuity points are there. Supposing that the intensity of a pixel at location (x, y) and time f is represented by f (x, y, t). Then the difference of two consecutive frames can be represented as.

$$D(x,y,t) = f(x,y,t) – f(x,y,t+1)$$

The noise occurring in D (x,y,t) is removed by convolving it with gaussian low pass filter. Since noise is made up of high frequency components, so most of noise is removed. After that thresholding is done i.e. selecting maximum value pixel out of all pixels and making it 1, rest are made 0. In this way coordinates of point that is one are found out and hence object is located. After that region of interest is marked around that point and in next iteration it is expected that object will be located within that region (assuming motion of object is smooth and not abrupt). Therefore the coordinates of object within region of interest are taken as center for marking region of interest for next iteration. This procedure continues and with each iteration the object is always tracked. This is shown in the sequence of frames below.

IV. EXPERIMENTS AND RESULTS

Fig. 8. Results and analysis for Video clip



V. APPLICATION DOMAINS OF VIDEO SURVEILLANCE

Video surveillance usage range in many areas starting from Security in public and commercial domains, Smart video data mining, Law enforcement and Military purposes etc. Usage in public and commercial security can be further explained as monitoring of banks, department stores, airports, museums, stations, private properties and parking lots for crime prevention and detection. Some other key areas are patrolling of highways and railways for accident detection. Surveillance of properties and forests for fire detection. Observation of the activities of elderly and infirm people for early alarms and measuring effectiveness of medical treatments. There are number of applications in the field of smart video data mining like measuring traffic flow, pedestrian congestion and athletic performance, compiling consumer demographics in shopping centers and amusement parks, extracting statistics from sport activities, counting endangered species, logging routine maintenance tasks at nuclear and industrial facilities, artistic performance evaluation and self-learning etc. Today, traffic monitoring and control is most widely used application of object detection and tracking using video surveillance. The system works on the principle of measuring the speed of vehicles and detecting red light crossings and unnecessary lane occupation. The law enforcement agencies use this system on a very high scale worldwide. The other use might be in military security for patrolling national borders, measuring flow of refugees, monitoring peace treaties, providing secure regions around bases, assisting battlefield command and control etc. The use of smart object detection, tracking and classification algorithms are not limited to video surveillance only. Other application domains also benefit from the advances in the research on these algorithms. Some examples are virtual reality, video compression, human machine interface, augmented reality, video editing and multimedia databases.

VI. CONCLUSION AND FUTURE WORK

In this paper, Moving object tracking is a key task in video monitoring applications. Object detecting and tracking has a wide variety of applications in computer vision such as video compression, video surveillance, vision-based control, human-computer interfaces, medical imaging, augmented reality, and robotics. Additionally, it provides input to higher level vision tasks, such as 3D reconstruction and 3D representation. It also plays an important role in video database such as content-based indexing and retrieval.

REFERENCES

[1] Moving object perception and tracking by use of DSP Shirai, Y.; Miura, J.; Mae, Y.; Shiohara, M.; Egawa, H.; Sasaki, S.;Computer Architectures for Machine Perception, 1993. Proceedings 15-17 Dec. 1993 Page(s):251 – 256

[2] Tracking moving objects using adaptive resolution Born, C.;Neural Networks, 1996, IEEE International Conference on Volume 3, 3-6 June 1996 Page(s):1687 – 1692 vol.3

[3] Kalman filter incorporated model updating for real-time tracking Dae-Sik Jang; Gye-Young Kim; Hyung-Il Choi; TENCON '96. Proceedings. 1996 IEEE TENCON. Digital Signal Processing Applications Volume 2, 26-29 Nov. 1996 Page(s):878 - 882 vol.2

[4] Moving object tracking by optimizing active models Daesik Jang; Hyung-Il Choi; Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on Volume 1, 16-20 Aug. 1998 Page(s):738 - 740 vol.1

[5] Moving object tracking using local windows Celenk, M.; Reza, H. Intelligent Control, 1988. Proceedings., IEEE International Symposium on 24-26 Aug. 1988 Page(s):180 – 185

[6] Object Tracking Using Block Matching. H.Kartik, D. Schonfeld, P. Raffy, F. YassaIEEE conference on Image Processing, Vol. 3, pp. 945 - 948, Jul 2003.

[7] Adaptive Filtering for Image Enhancement, T. Peli, J.S. Lim, Proc. ICASSP81, Atlanta, pp. 1117-1120, Mar.1981.

[8] Paolo Spagnolo, Marco Leo, Giovanni Attolico, Arcangelo Distante, " A Supervised Approach in Background Modelling for Visual Surveillance" in AVBPA 2003: 592-599.

[9] C. Stauffer, W.E.L. Grimson. "Adaptive Background Mixture Models for Real-Time Tracking," in Proc. Computer Vision and Pattern Recognition Conf., vol. 2, Fort Collins, CO. USA, June 1999,pp.246-252.

[10] A. D. Bue, D. Comaniciu, V. Ramesh, and C. Regazzoni. Smart cameras with real-time video object generation. In Proceedings of the IEEE International Conference on Image Processing, volume 3, pages 429–432, June 2002

[11] F. Oberti, G. Ferrari, and C. S. Regazzoni. A Comparison between Continuous and Burst, Recognition Driven Transmission Policies in Distributed 3GSS, chapter 22, pages 267–278. Video-Based Surveillance Systems. Kluwer Academic Publishers, Boston, 2002.

[12] Bhavana C. Bendale, Prof. Anil R. Karwankar. Moving Object Tracking in Video Using MATLAB, International Journal of Electronics, Communication & Soft Computing Science and Engineering.

[13] Qing Wanga, Feng Chena. An Experimental Comparison of Online Object Tracking Algorithms, aTsinghua University, Beijing, China

[14] Evandro Alves da Silva, Adilson Gonzaga. On-Road Automotive Vehicle Detection using Gaussian Mixture Model and NNs, University of São Paulo.

[15] A. Yilmaz, O. Javed and M. Shah,"Object Tracking: A Survey" in ACM Journal of Computing Surveys, Vol. 38, No. 4, 2006.

