A Review Paper on Video Surveillance System and Proposal to Enhance Existing Surveillance System

Mr. Amar B Kiravekar¹ Mr. Pradip C Bhaskar²
¹M. Tech. Student ²Associate Professor
¹,²Department of Technology
¹,²Shivaji University, Kolhapur, University, Kolhapur, Maharashtra, India

Abstract—In surveillance systems the videos recorded may not be clear and high quality, if they are recorded in low illumination background or in bad weather conditions like fog or rain. So there can be a fast and effective method of getting cleared and enhanced images from surveillance videos under said conditions, where we can get a good visual quality images from acquired videos with motion detection and processing these images through illumination adjustment, haze detection and haze removal algorithms in FPGA. The system can be useful in highly sensitive places like banks, shopping malls and high way transportation, where constant monitoring is required. Accordingly a rigorous literature review has been carried out and eventually it has been concluded that processing these images by image processing algorithms in FPGA increases speed of operation and is low cost in nature.

Key words: Haze Detection and Removal, Illumination Adjustment, Motion Detection, Surveillance Systems

I. INTRODUCTION

The surveillance systems are being implemented in shopping malls, high way transportation, and sensitive places where there is a need to monitor what’s happening over there. In such places a constant need to observe is required along with recording of videos in regard to that. The saved videos are reviewed as and when required in these systems if any contradictory incident has happened at that place. It helps to find out what exactly happened at that instant and the responsible person for that. The required and needful procedure is followed thereafter.

The sensitive places like shopping malls, banks and other security required places implement web cameras for constant monitoring and recording the incidents happening. These cameras give cleared view when the incident has happened in daytime when illumination is high but it fails to give a cleared video as well as image quality if the recorded videos are during the night time or in low illumination background.

However, in bad atmospheric conditions like rainfall, fog or low illumination the recorded videos cannot be clear as the pixels get blurred, so under these conditions this surveillance system can give a cleared and enhanced output image from the recorded videos that are acquired after the motion detection and by processing them under image processing algorithms and implementing them on FPGA. This will result in full recognition of images as the images would be enhanced and dehazed.

So the system can be developed for surveillance where we would extract high quality images from videos that are recorded or saved, and if motion has been detected the triggering to FPGA will be given which will result in automatic acquisition of images from videos and would get repeated if motion detection still persists by providing the time interval for that. The saved images from personal computer would be given to image processing algorithms to further enhance and dehaze the image to clarify it at greater extent such that under sensitive places it would be beneficial for exact identification and recognition. The algorithms are implemented through FPGA for faster processing and is low cost with feasible in nature.

II. LITERATURE REVIEW

A. Lightness and Retinex Theory

In 1971, Land and McCann [1] developed the Retinex theory as a simple and effective model of visual perception. The Retinex theory assumed that the intensity of visible light reaching a camera and the observers eyes depends on the product of two components one is illumination which represents the ray of light source from the sun or sky in the scene and second is reflectance which represent the illumination variation of the object surfaces. According to the assumption on observed image can be expressed as,

\[ I(x,y)=L(x,y) \times R(x,y) \]

Here, \( I(x,y) \) – Pixel of on observed image located in the \( (x,y) \) coordinate

\( L(x,y) \) - Illumination of the scene

\( R(x,y) \) - Reflectance of object

\( L'(x,y)=R'(x,y) \times G(L'(x,y)) \)

The image enhancement algorithm first estimates the illumination denoted as \( L(x,y) \) by some signal processing technique, and the estimated reflectance denoted as \( R(x,y) \) can be determined. Then, the estimated illumination \( L'(x,y) \) can be adjusted by a nonlinear operator \( G(\cdot) \) to enhance the visual quality of the image. Finally, the adjusted luminance data can be obtained as

\[ I(x,y)=R(x,y) \times G(L(x,y)) \]

The image \( I \) is an enhanced version of the input observed image. The main challenge of the Retinex method is how to estimate the illumination component \( L'(x,y) \) properly. So in order to overcome this a new technique was developed called as EMD.

B. The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis

1) N. E. Huang et al [7]

The combination of the empirical mode decomposition (EMD) method and the associated Hilbert spectral analysis has offered a powerful method for nonlinear non stationary data analysis. Central to the present approach is the sifting process to produce the IMFs, which enables complicated
data to be reduced into such a form that the instantaneous frequencies can be defined. These IMFs forming the basis of the decomposition are complete and practically orthogonal. The expansion in terms of the IMF basis has the appearance of a generalized Fourier analysis with variable amplitudes and frequencies. It is the first local and adaptive method in frequency–time analysis.

According to this theory the image $I'$ is an enhanced version of the input observed image. Recently powerful signal analyzing technique called Empirical mode Decomposition (EMD) is proposed. By extracting oscillation information of the data EMD decomposes a signal into several frequency components, which are called character intrinsic mode functions(IMFs) and a last function that can’t be decomposed into any IMFs which is called residue. The process of EMD required several iterations to obtain each IMFs. In this a signal has to be splitted into number of frequency bands and it also requires large iterations hence it is not a convenient method for use. To analyze 2-D data such as a 2-D image and to overcome the exhaustive computation a new method called as FABEMD was proposed.

C. Fast and Adaptive Bi Dimensional Empirical Mode Decomposition for the Real-time Video Fusion:

1) Maciej Wielgus, Adrian Antoniewicz, Michał Bartyś, Barbara Pütz

To analyze 2-D data such as 2-D image and to overcome the exhaustive computation a fast and adaptive bi dimensional empirical mode decomposition (FABEMD) was proposed. The Bidimensional Empirical Mode Decomposition (BEMD) method proved to be capable of producing high quality results of infrared (IR) and visible (VIS) images fusion. However, large complexity of this algorithm does not contemporarily allow for real-time implementation, necessary in many typical applications of VIS-IR fusion, e.g., in environment monitoring. In contrast, the Fast and Adaptive Bidimensional Empirical Mode Decomposition (FABEMD), the variant of BEMD, in which signal envelope is extracted by means of statistical filters rather than 2D spline interpolation, has an ability to overcome this shortcoming. We evaluate FABEMD method outputs in the context of VIS-IR fusion and present developed real-time VIS-IR video fusion system based on one chip Field Programmable Gate Array.

In image application some existing research found that each BIMF and the residue contain different information of an image such as color variety, edge, noise and illumination. The first BIMF contains the majority of color variation and the residue contains the illumination trend. With this assumption we treat the residue as the illumination plane $L^*(x,y)$ in the Retinex theory for adjusting the uneven illumination. According to FABEMD process the residue is included in the average envelop of each $I^*$th decomposition. To reduce the computational load, we only calculate the first average envelope as the illumination plane $L^*(x,y)$. Then, an adaptive gamma correction is developed to adjust $L(x,y)$ to enhance the illumination contrast without oversaturation. In this method the processed image is not that much cleared and enhanced so to adjust the illumination levels for increasing the visual quality of the image the new technique of hardware architecture has been developed called as FIEEMD.

D. A Low-Cost Hardware Architecture for Illumination Adjustment in Real-Time Applications

1) You-Horng Shiao, Pei-Yin Chen, Member, IEEE, Hung-Yu Yang, and Shang-Yuan Li

For real-time surveillance and safety applications in intelligent transportation systems, high-speed processing for image enhancement is necessary and must be considered. In this paper, we propose a fast and efficient illumination adjustment algorithm that is suitable for low-cost very large scale integration implementation. Experimental results show that the proposed method requires the least number of operations and achieves comparable visual quality as compared with previous techniques. To further meet the requirement of real-time image/video applications, the 16-stage pipelined hardware architecture of our method is implemented as an intellectual property core. This design yields a processing rate of about 200 MHz by using TSMC 0.13-µm technology. Since it can process one pixel per clock cycle, for an image with a resolution of QSXGA $(2560 \times 2048)$, it requires about 27 ms to process one frame that is suitable for real-time applications. In some low-cost intelligent imaging systems, the processing rate can be slowed down, and the hardware core can run at very low power consumption.

In this method, first color image is mapped from the RGB into the HSV color area. The HSV domain is a perfect approach for color descriptions that are natural and intuitive to humans. Second the first average envelope extracted as the illumination plane $L$ by using FIEEMD only from the intensity layer no. of the HSV domain. Third a novel gamma correction is employed to adjust the illumination component to improve the usual quality of details in low and high luminance regions. Finally, the enhanced image represented in the HSV color space is converted into the RGB domain.

The processed image from this method has good illumination levels such that the image has been enhanced at greater extent but the real time video surveillance with automatic recording of the respective images leads to bulk storage of videos on to the PC which ultimately results in wastage of huge amount of memory. However to minimize this a real time motion detection has been introduced with FPGA implementation.

E. FPGA-Based Real-Time Motion Detection for Automated Video Surveillance Systems

1) Sanjay Singh 1*, Chandra Shekhar 1† and Anil Vohra 2†

Design of automated video surveillance systems is one of the exigent missions in computer vision community because of their ability to automatically select frames of interest in incoming video streams based on motion detection. They have focused on the real-time hardware implementation of a motion detection algorithm for such vision based automated surveillance systems. A dedicated VLSI architecture has been proposed and designed for clustering-based motion detection scheme. The working prototype of a complete standalone automated video surveillance system, including input camera interface, designed motion detection VLSI architecture, and output display interface, with real-time
relevant motion detection capabilities, has been implemented on Xilinx ML510 (Virtex-5 FX130T) FPGA platform. The prototype system robustly detects the relevant motion in real-time in live PAL (720’576) resolution video streams directly coming from the camera.

The method may acquire respective images through videos but the image might not be clear if the videos have been recorded during bad weather conditions or foggy atmosphere, hence there is a need to dehaze the image with haze detection and haze removal algorithm to enhance the visual quality of the respective image.

**F. Hardware Implementation of a Fast and Efficient Haze Removal Method**

1) Yu-Horng Shiau, Hung-Yu Yang, Pei-Yin Chen, and Ya-Zhu Chuang

In this method, a fast and efficient haze removal method is presented. It employs an extremum approximate method to extract the atmospheric light and propose a contour preserving estimation to obtain the transmission by using edge- preserving and mean filters alternately. The method can efficiently avoid the halo artifact generated in the recovered image. To meet the requirement of real-time applications, an 11-stage pipelined hardware architecture for our haze removal method is presented. It can achieve 200 MHz with 12.8K gate counts by using TSMC 0.13-μm technology. Simulation results indicate that our design can obtain comparable results with the least execution time compared to previous algorithms and is suitable for low-cost high-performance hardware implementation for haze removal.

The method contains three procedures: atmospheric light estimation, transmission estimation, and scene recovery. First, an extremum approximate method with a minimum filter is presented to estimate the atmospheric light. Second, it employs an edge detection method to find the contour of objects and apply edge-preserving filter and mean filter alternately to estimate the transmission. Finally, a saturation correction method is proposed to refine the recovered scene. Even though the image has been processed with haze detection and haze removal circuit the image might be dehazed but yet it needs to be enhanced with illumination adjustment algorithm for proper illumination levels to increase its visual quality.

**G. Streaming Elements for FPGA Signal and Image Processing Accelerators**

1) Peng Wang and John McAllister

Field-programmable gate array (FPGA) devices boast abundant resources with which custom accelerator components for signal, image, and data processing may be realized, however, realizing high-performance, low-cost accelerators currently demands manual register transfer level design. Software- programmable soft processors have been proposed as a way to reduce this design burden, but they are unable to support performance and cost comparable to custom circuits. This paper proposes a new soft processing approach for FPGA that promises to overcome this barrier. A high-performance, fine- grained streaming processor, known as a streaming accelerator element, is proposed, which realizes accelerators as large-scale custom multicore networks. By adopting a streaming execution approach with advanced program control and memory addressing capabilities, typical program inefficiencies can be almost completely eliminated to enable performance and cost, which are unprecedented among software- programmable solutions. When used to realize accelerators for fast Fourier transform, motion estimation, matrix multiplication, and sobel edge detection, it is shown how the proposed architecture enables real-time performance and with performance and cost comparable with hand- crafted custom circuit accelerators and up to two orders of magnitude beyond existing soft processors.

**H. A Novel Video/Photo Recorder Using an Online Motion Sensor Triggered Embedded System**

1) Öğuz Gora

The embedded systems have gained more importance as these systems are especially dedicated to specific tasks which are handled by highly optimized solutions. One of the interesting areas of embedded systems use is multi-media. Producing, processing, streaming various multimedia types and interacting with the physical environment is very common. Similar to these studies, controlling and observing the specified area by multi-media tools are the necessities for many reasons such as security. Here they have presented a method of video and photo recording of any moving object by using open source operation system (Raspbian-a distribution of Linux) and software (Python – a high-level programming language). The system is triggered by a motion sensor and it collects visual data from a specified area for limited duration. The collected data is published on internet via dedicated web site.

The system works by itself but with a web interface many control abilities are possible. A user can for example, change operation modes, length of video recordings and access files from system memory. The operation system that works on embedded microcomputer is a distribution of Linux, Raspbian. Python programming language is used as operating system and controls the camera. The developed program in Python operates the camera. The other tasks of the program are saving the records on the system and communicating with dedicated web page. The web page interfaces the records to the users and gives opportunities to make some adjustments on the system with cooperation of the microcomputer.

**I. Adaptive Background Mixture Models for Real-Time Tracking**

1) Stauffer and Grimson

A common method for real-time segmentation of moving regions in image sequences involves “background subtraction,” or thresholding the error between an estimate of the image without moving objects and the current image. The numerous approaches to this problem differ in the type of background model used and the procedure used to update the model. This paper discusses modeling each pixel as a mixture of Gaussians and using an on-line approximation to update the model. The Gaussian distributions of the adaptive mixture model are then evaluated to determine which are most likely to result from a background process. Each pixel is classified based on whether the Gaussian distribution which represents it most effectively is considered part of the background model. This results in a stable, real-time outdoor tracker which reliably deals with lighting changes, repetitive motions from clutter, and long-term scene
changes. This system has been run almost continuously for 16 months, 24 hours a day, through rain and snow.

J. An Efficient Technique for Illumination Adjustment Using CLAHE Algorithm

1) V.Vembuselvi, T.Murugan [1]  
For real time applications in intelligent transportation system processing an image at high speed is essential and must be considered. In this paper they have proposed a low complexity efficient technique for illumination adjustment based on Retinex theory and CLAHE algorithm. From the experimental results it can be concluded that the proposed method achieved comparable good visual quality than the other enhancement methods. This method can process one pixel per clock cycle, for an image with a resolution of QSXGA (2560 × 2048), it requires about 25 ms to process one frame that is suitable for real-time applications. In some low-cost intelligent imaging systems, the processing rate can be slowed down. The procedure of this method is explained as follows.

First, an image in RGB colour space is changed into HSV domain. The reason for choosing HSV domain is, it is intuitive to humans. Next, average envelope is found as illumination plane which is denoted as L. By using CLAHE algorithm illumination component L is adjusted to get good visual quality in high and low luminance regions. At last enhanced image that is represented in HSV colour domain is again converted into RGB colour space. The procedure followed in FIEEEMD is explained as follows.

Estimate the local median and local maximum points of intensity layer of the image. In empirical mode decomposition technique local maximum and local minimum points are calculated. But in FIEEEMD technique local median point is calculated instead of local minimum points since it produces smooth illumination output and it results in an increased contrasting reflectance. By connecting all the local maximum and local median points, upper and lower envelopes can be estimated respectively. The first average envelope is calculated by averaging the lower envelope and upper envelope values that is calculated in Step 2 and it is denoted as EA. It results in estimated illumination plane. To get enhanced image, illumination adjustment must be done using CLAHE algorithm.

III. PROPOSED METHOD

In this proposed method we are going to do motion detection in a surveillance video and extract the respective images from this video by using the MATLAB tool. The extraction of these images can be single or multiple in nature, depending upon the number of detected objects. A captured image in surveillance shows an irrevocable loss of visual information. In some places under strong background illumination or in dark environments. This is mainly because the dynamic range of natural scene is far larger than the dynamic range of image captured by common digital devices, such overflow will cause a blurred image particularly in the low luminance regions and viewer can’t obtain enough information of the image. In addition to that a captured image might be corrupted due to fog or haze. The degraded image substantially loses contrast and reduces the visual quality of the object in the scene. Due to this in spite of having a video recording of any incidence the further investigation by observing the video is not becoming possible.

So in order to bring effortless investigation by observing the recorded video of any incidence at least by selecting the snaps of the video with typical interval at real time is possible. There can be two problems with snaps which need to be overcome, one is adjusting illumination levels of the image and second is to detect haze and remove it from the image.

Hence the acquired images are sent to the FPGA where the illumination adjustment, haze detection and haze removal algorithms are present. The images are further processed in this FPGA to adjust the illumination level as required by the respective image to get cleared and enhanced view, followed by processing that image for detection and removal of haze.

![Fig. 1: Expected changes to be made in current surveillance system](image)

**IV. CONCLUSION**

By studying and analyzing of the literature review it can be said that recorded videos may not be clear when recorded under low illumination and foggy atmosphere as the pixels get blurred so there is a need to developed a system which can capture only those images which are required with motion detection and enhance the visual quality of that respective image by processing it with fast image processing algorithms and implementing it over the FPGA which is expected to overcome the specified problems.

**REFERENCES**


