A REVIEW ON OBJECT DETECTION IN THERMAL IMAGING AND ANALYSING OBJECT AND TARGET PARAMETERS

1M. NARENDRAN, 2MANONMANI LAKSHMANAN, 3SAI MOHAN, 4LEKHITA SHARMA, 5R. PRANAY SHARMA

1,2,3,4,5SRM University, Bharathi Salai, Andavar Nagar, Ramapuram, Chennai, Tamil Nadu 600089
E-mail: vp@rmp.srmuniv.ac.in

Abstract - The use of thermal imaging has become a major application in military. Thermal images provide a clear image even during night times. Using image recognition in identifying the target and embedded system in aiming the enemy, we can attack the enemy at a faster rate than usual. The thermal image is captured, enemy is identified automatically and the enemy is targeted. This application could be utilized in war fields where the enemy is targeted with the least efforts of man. This will promote the efficiency of attack in the war fields. For image processing, we can use OpenCV or MATLAB software. First the object is going to be detected from the image and that image has to be aligned at the centre of the object. After detecting the object, the next task is to recognize the object as gun, the instruction is going to aim the gun to target the object. For object detection SVM technology is better thing to use.

Keywords - Support Vector Machine(SVM), Opencv, retinex filtering, background subtraction(BS), Gaussian mixture modules(GMM), Convolution Neural Network(CNN)

I. INTRODUCTION

Thermal cameras are used to display the difference in heat through the thermo gram which takes in the heat waves and produces output in the form of images. Object detection is a process which is used to detect object from an image. Features are taken from the image and analysed. Thus, the aim of this project is to build a system that places the gun, which acquires the image from the image recognition module, and tilts itself in a way that it targets the enemy and shoots. The objective is to build an artificially intelligent gun that requires less or no manual support except the image. This could increase the efficiency of attack when it comes to defence in the army. Recently, support vector Machine (SVM) based algorithms have been proposed to detect and track the infrared dim small targets. The SVM are based on the concept of decision planes which defines the boundaries between a set of objects belonging to different classes. The thermal image is processed:

- Applying retinex filtering to remove noise and improve filtering
- Support vector machine detects the classifier from the background using discrete wavelet transform, mean, entropy and variance.
- Using Convolution Neural Networks to increase the quality of the image
- Fusing thermal and visible imagery using Gaussian mixture models

Tracking the dim objects along with fusing layers of thermal image with image from the visible spectrum could provide a better clarity of the surroundings in the image. By combining the visible spectrum and thermal imaging as one, more information about the object being tracked can be obtained as they provide good support to each other. The segmentation of moving objects in corresponding video frames involves. ‘Background subtraction method’. Here each of visible as well as thermal is modelled as a mixture of Gaussians for background subtraction, analysis and fusion rule applied to build robust moving object detection system. The quality if the image can be increased further on by using the convolution neural network. Once the image we have formulated, we must recognize the object in the image. Intelligent video surveillance system based on image recognition is widely employed to effectively avert crimes and provide public security. We can use that model to identify our enemy. Once the enemy is identified, we are going to position the gun automatically to the target. After getting instruction from authorities, the gun is shot over the enemy. The major advantage of the machine, which cannot be reached by a normal human being.

II. LITERATURE SURVEY

Detecting dim small targets in infrared images and videos is the main concern of this paper. Support Vector Machine (SVM) algorithm is being used here which is quite efficient to detect and track the infrared dim small objects [1]. Concept used in SVM is based on decision planes that defines boundaries between two objects categories. In the starting Retinex filtering to remove the noise and improve the image quality [1]. Then in the second step, object in infrared image is detected through the support vector machine classifier from the background [1]. Four very important features used by SVM classifier are Discrete Wavelet Transform (DWT), mean, entropy and variance [1]. SVM classifier separates objects from background [1]. Identification of Moving objects is the main concern of this paper [2]. And for understanding more about
the object’s movement, special fusion of visible spectrum and Thermal imaging is used [2]. Visible spectrum and Thermal imaging are inherently complementary to each other so it gives better information [2]. The segmentation of moving objects in corresponding video frames involves ‘Background Subtraction’ method [2]. Visible and Thermal images are combined as group of Gaussians for background subtraction, and then blob analysis and fusion rule is applied [2].

Identification of Moving objects is the main concern of this paper also [3]. The technique used is background subtraction (BS) and fusion of thermal and visible imagery using Gaussian mixture models (GMM) [3]. Dynamic adaption of fusion scheme can be seen as it will work automatically for daytime/night-time and accordingly make changes [3]. There are 3 fusion techniques used, first is augmenting the GMM models with thermal information in prior to foreground segmentation [3]. The second is fusion of outputs of BS applied to each sensor separately [3]. The third is linear combinations of both images forming Hybrid image [3].

The main focus of this paper is to recognize objects in extreme conditions such as night, erratic illumination condition [4]. This need can be fulfilled using multimodal sensors [4]. This paper aims to enhance the low-resolution thermal image according to the extensive analysis of existing approaches [4]. Special thermal image enhancement using convolutional neural network (CNN) called in TEN, is used for direct learning an end-to-end mapping a single low-resolution image to the desired high-resolution image [4].

This paper suggest the first thermal image enhancement method based on RGB data [5]. Behaviour, activity, or dynamic information, of people or objects [5]. Intelligent video surveillance system (IVS) based on recognition is widely employed to effectively avert crime and provide public security [5]. High complexity is present in processing real-time data and analysis [5]. Many technologies, mostly intelligent techniques like neural systems, fuzzy logic, SVM, genetic algorithm emerged out as basis for intelligence in such systems [5]. A lossless image compression algorithm aimed at compressing thermal images captured, stored and transmitted by a Nano-satellite system [6]. This algorithm is specially applied to thermal grey scale images [6]. A balance is achieved between compression ratio and computational power consumed [6]. Method contains various stages [6]. Compression efficiency is achieved using each stage and subsequent hardware implementation to achieve so [6]. Algorithm tries to minimize the memory, which will further be really useful in real-time image compression on Nano-satellite systems [6].

The main concern of this paper is the light changes and object detection in thermal infra-red image to overcome the problems of varying lighting conditions, thermal infra-red cameras are often used [7]. The object detection approach used here is principal component analysis (PCA) based machine techniques for image classification [7]. Multiple supervised machine learning algorithms and an unsupervised machine learning algorithm are analysed, evaluated and compared to get the best outcome [7]. According to this paper, a PCA based Gaussian classifier and a Mahalanobis distance based classifier are the best choice for detection and tracking [7].

III. ANALYSIS ON

There are various processes the video must undergo after the snapshot has to be taken. The first step is to take the image and filter the image to get a clear view of it. There are various ways to filter the image. Some ways include: Linear filtering, Average filtering, Median filtering and Adaptive filtering. Median Filtering sets the value of the output pixel to the average of the pixel values in the neighbourhood around the corresponding input pixel. The value of an output pixel is determined by the median of the neighbourhood pixels, rather than the mean, according to the median filtering. The median is much less sensitive than the mean to extreme values. Median filtering is therefore better able to remove these outliers without reducing the sharpness of the image.

The basic Retinex filtering for multistate monochrome images can be expressed:

\[ R(x, y) = \sum_{k=1}^{K} W_k (\log I(x, y) - \log (F_k(x, y) * I(x, y))) \]

The next process after image filtering is the support vector machine (SVM). In machine learning, support vector machines are used with associated learning algorithms that analyse data used for classification and regression analysis. Some training examples, each marked as belonging to one or the other of two categories, an SVM algorithm builds a model that assigns new examples to one or the other, making it a classifier that is non-probabilistic. SVM model is the representation of the object as points in space, mapped in such a way that the examples of the separate categories, that are created, are divided by a clear gap as wide as possible. New samples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.
Image fusion is the fusion of two or more images of a scene to improve visual quality or feature extraction and detection. The fusion process is weighted sum of thermal and visible face information with two factors $a$ and $b$. The highest level of accuracy is achieved in image fusion that is about 98.42%.

CNNs are often used in image recognition. It has achieved least error rate. CNNs were used to assess video quality in an objective way after manual training of the images; the resulting system had a very low root mean square error.

Disorted images have an increasingly common characteristic with modern digital cameras. By contrast, those kinds of images rarely trouble humans. Various features cannot be determined that easily in normal images. But CNN does this work perfectly through CNN pooling.

Video has temporal dimension. However, some extensions of CNNs into the temporal domain. One approach is to treat space and time as equivalent dimensions of the input and perform convolutions in both time and space. Another way is to combine the features of spatial and temporal stream. Un-overlooked learning schemes for training spatial-temporal features have been introduced, based on Convolutional Gated Restricted Boltzmann Machines and Independent Subspace Analysis.

The object must be separated from the background. This is done using background separation. There are various conventional methods used in the background subtraction. They are using frame differencing, mean filter, running Gaussian average, background mixture models.

IV. EXPECTED OUTCOME

The project, as a whole, would be to monitor the situations, especially near the border. The camera input (video) is passed on to the computer where it’s processed and rectified for the background subtraction and other filters without ever losing the detail of the input signal. Next it would convert to thermal imagery and look for the plastic explosives. The process of differentiation would mainly involve the colour of the gun which is a fixed range of purple. Once any resemblance is found, a certain human around it would be detected. Once the human is detected, it will be made the center of the capture. The situation updates will be sent to the commander in charge for the location.

We are using Support Vector Machine (SVM) for differentiating the planes and proper detection and ranging of the target discovered. This is for the estimation of the target distance and direction from the camera and the gun. This would prevent any errors in the process. For the filtering of the image, we are using the Retinex filtering as it has a good standard.

In the further stages of the project, we would like to include a gun that can lock the target (identified object (man)) and wait for the instructions. This can be very helpful in keeping a watch from safer places while we track down each potential move by the enemy. This would make the tracking easy, especially in places like Kashmir, where the snow is freezing. For the demonstration of locking the target in gun, we would like to use a laser and point it at the locked target.

CONCLUSION

The biggest threat for a large democracy like ours has been invasion by enemies. The threat increases during the night times. We aim at providing a possible solution to this problem with our project. As we aim at monitoring the area, using just a camera, a gun and a person sitting in the command room, either near or far. The exciting part about the project is that we can use this for surveillance of the area even during the night times effectively. The thermal imaging and IR cameras can provide us with great inputs even during the night times. This could potentially result in us tracking down every move from the enemy/anyone for that matter. The risk of giving the control to machine is voided as we have the control of shooting with the person in the command room.

ACKNOWLEDGEMENT

This paper wouldn’t have been possible without our guide Mr. Narendran sir. We would also like to show
our deepest gratitude to our HOD and Dean.

REFERENCES

[2] Prof. Supriya Mangale and Dr. Madhuri Khambete for “Moving Object detection using Visible Spectrum Imaging and Thermal Imaging”
[4] Yukyung Choi, Nmail Kim, Soonmin Hwang and In So Kweon for “Thermal Image Enhancement using Convolution Neural Network”

★★★★