International Journal of Pure and Applied Mathematics
Volume 116 No. 22 2017, 467-477
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

An Efficient Method to Detect Moving Object with Features by Using GLCM and ANN

[1]N. Venkatesvara Rao, [2]D. Venkatavara Prasad and [3]M. Sugumaran

[1]Information Technology,

Kings Engineering College, Chennai.

JNTUK, Kakinada.

raosusi1969@yahoo.co.in

[2]JNTUK, Kakinada.

CSE, SSN College of Engineering, Chennai.

dvvprasad@ssn.edu.in

[3]JNTUK, Kakinada.

CSE, Pondicherry Engineering College, Pondicherry.

sugu@pec.edu

Abstract

Object tracking monitors the spatial and temporal changes in the video sequence. Object detecting and tracking has wide applications in medical imaging, vision-based control, video surveillance and robotics. Detection of moving object is a challenging task and tracking it requires the location and shape of object in every frame. Background subtractions are very useful, simple and efficient technique which reduces the computational time. In background subtraction, video frames are compared with the reference background model and the deviation of the image is identified. Feature is extracted using Gray Level Co-occurrence Matrix and is classified using Artificial Neural Network. The main goal is to detect the number of moving objects in a video sequence in a robust manner.

Key Words: Background subtraction, bilateral filter, Gray Level Co-occurrence Matrix (GLCM) and Artificial Neural (ANN).

1. Introduction

The world is in the middle of an exciting period where new technologies breakthrough in areas of digital computation and telecommunication. Video and image processing is getting interesting due to the diversity of applications. Many emerging technologies are available for digital image processing. An image is an imitation of an object or thing, a vivid or a graphical description. An improved version of an image can be formed using different image processing techniques by a series of actions or operations. The need for the improvement of pictorial information for human interpretation has paved a way to the stem of image processing. Compared to image analysis, video introduce many additional challenges [1].

In video analysis, real time moving object detection is an important application. Moving object detection is used to find the changes in the images and if there is any changes, then it extracts the moving object. Moving object detection is done by separating the foreground object i.e., moving object from the static information known as the background. [2] The background subtraction is the key step used in moving object detection to detect persons, vehicles, etc.

While processing the video, first video is converted into frames of images. [3]Preprocessing is an initial step which has to be done in order to remove the noise in the images. Enhancement of the image will also be achieved by preprocessing techniques.

Feature extraction is a method for capturing visual content of images for indexing and retrieval. Feature extraction is a form of dimensionality reduction technique. The input data transformation to form a set of features is called feature extraction. The feature selected should contain enough information about the image and no other specific knowledge should be needed for their extraction [4].The selected feature should be easy to compute. The selected feature can be of any type like color, shape, etc. There are different feature extraction techniques. Principal component analysis, Gabor filter, Chain Code Histogram of Character Contour, GLCM, etc. are some of the feature extraction techniques.

The extracted features are classified according to the needs. There are different classification techniques [5]. Among those artificial neural network is a classifier which can be trained according to the needs. These neurons can be trained according to the inputs and it can be easily modified according to the criteria chosen for classification.

2. Literature Review

Zhang et al., (2016), proposed a new algorithm introduced for the background subtraction to detect moving objects in a video sequence. The method uses a spatiotemporal-based algorithm to find precise background. By background

subtraction method, the region of foreground is found by comparing the input video with the reference image. Instead of fixed threshold, this method adapts a no-fixed threshold mechanism. Fixed threshold to find the similarity of the pixel will lead to error while determining the background when the area contains high background dynamics. Due to the pixel based detection, this method outperforms many state-of-the-art methods.

Lee et al., (2016), suggested a method that focused on background subtraction for multimedia applications. Hash-table-based function is used for background subtraction method. The cost is low because of its simple algorithm. Background sets are defined on reduced image and color spaces. Due to color-space reduction, color noises can be avoided. The proposed method gives high speed results and also the complexity is less due to hash-table-based functions. This method is mostly used for dynamic backgrounds. It is mainly used in mobile or embedded environments.

Lloyd et al., (2017), proposed a method to detect the events automatically from videos to minimize violence. The abnormal crowd detection is done by utilizing computer vision technique. GLCM texture feature analysis is used for crowd density estimation. Temporal coding is used to describe the crowd dynamics. Parameter values are selected with an in-depth evaluation. Haralick texture features are selected with GLCM. The method is highly effective at discriminating normal and abnormal behavior and can operate in real time.

Devi et al., (2016) proposed a method to detect features using GLCM. Iris recognition system is used for identification. Spatial relation between varying intensity values is represented with non-filtered GLCM. Circular Hough transform is used to find the center co-ordinates and radius of the pupil in iris. Edges within the image is detected using canny edge detector.After segmenting, normalization is done to get the polar co-ordinates. The features are extracted with GLCM and is mainly used for recognizing individuals.

Gaba et al., (2016), proposed a method aimed at motion detection for automatic video analysis. Complete detection of moving object can be achieved even when there is changes in the brightness, variations in the surrounding environment and even when there is noise in the background. The method is pixel dependent and is based on non-parameterized approach to build the model. A unique methodology is also used to identify and to eliminate the ghost object. The proposed technique gives both quantitative and qualitative results with reduced processing time. Abedin et al., (2017) proposed a method to detect and recognize using fuzzy and neural network. Fuzzy rule based color system was used to detect the traffic sign. Speeded Up Robust Features descriptor is used for training the artificial neural network. Recognition of the traffic signal is done with ANN by extracting the features from segmented sign area in gray scale. The performance of ANN is compared with other classifier techniques where ANN outperforms other methods.

Abdelhedi et al., (2016) proposed a method to recognize human actions using artificial neural network. Due to abundant visual spatio-temporal information, features are hardly generalized. Hu moment and Zernike moment are used to find feature vector of an action. Feed forward neural network is used as a classifier. The negative effect due to the complex actions can be eliminated by the proposed method. The recognition rate is increased for complex actions and the method is effective and robust.

3. Proposed Methodology

The proposed method was used for detecting the moving object video sequence with high accuracy. This was done by using GLCM as feature extractor and the classification is done by training the artificial neural network (ANN) with the feature metrics.

Block Diagram

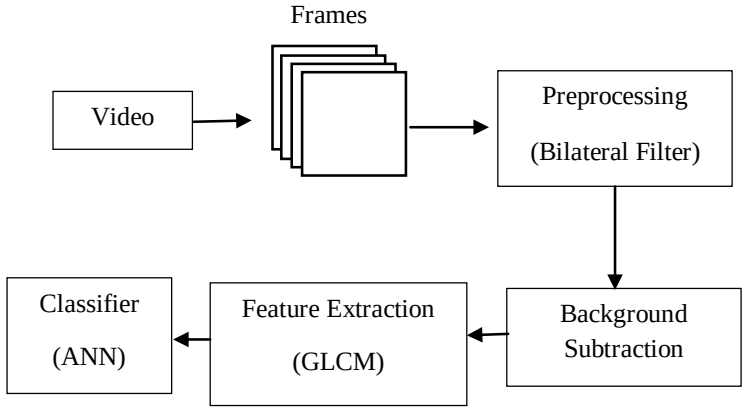


Figure 1: Block Diagram of the Proposed System

Background Subtraction

Arithmetic operations on images like addition, subtraction, mean, etc., are commonly used to combine and transform two or more images to form a new image with better display or to highlight certain features in the image. Also with single image as input, arithmetic operations can be done on all pixels in order to modify brightness and to enhance contrast.

Image subtraction operator takes the differences between the two input images. Here the absolute difference between the pixel values are considered. Background subtraction is mostly used to detect changes in a sequence of images of a continuous scene or to diagnose a moving object. By using the subtraction technique, the conspicuity of an object of interest can be increased by eliminating the background surrounding the object. (Brandan, M.E et al., 2016). For any two images f_r and f_s the image difference is given by the following equation.

$$f_r - f_s \qquad (1)$$

4. Feature Extraction

Feature extraction concerns on finding shapes in an image. First a feature has to be selected. Feature selection is the process where a subset of a feature is selected from a group of original features. By choosing particular features by certain criteria, the feature space can be optimally reduced. Feature extraction is the process in which a new set of features will be created. Different kinds of features can be extracted like shapes, histogram, color, spectral, etc. There are different feature extraction techniques like histogram leveling, fast fourier transform, discrete wavelet transform, gray level co-occurrence matrix, etc.

5. Gray Level Co-Occurrence Matrix

The GLCM technique is a statistical method for feature extraction. Features are important piece of information. By extracting the feature, the processing time and the complexity will be reduced. The spatial relationship of the pixels is taken into an account for extracting features.

The co-occurrence matrix of the image can be determined using GLCM. Pixel relationship with other pixels in terms of distance and angle can be determined. The number of rows and columns in a matrix is equal to the number of gray levels in the GLCM matrix.The formula for GLCM of an image with size $N \times N$ is given by

$$p(i,j) = \sum_{i=1}^{N} \sum_{j=1}^{N} \begin{matrix} 1, I(x,y)=i \text{ and } I(x+\Delta_x, y+\Delta_y)=j \\ 0, otherwise \end{matrix} \qquad (2)$$

$p(i,j)$ refers the joint probability of co-occurrence at an offset (Δ_x, Δ_y) with spatial positions $x \& y$ and having an intensities $i \& j$. The distance d and the angle θ between the pixels to its neighborhood are specified by the offset (Δ_x, Δ_y).

The number of distinct gray levels in the quantized image is represented by G. $p_i(i)$ and $p_j(j)$ represent the marginal probability matrix. The various features are represented by the following notations.

$$p_i(i) = \sum_{j=1}^{G} p(i,j) \qquad (3)$$

$$p_j(j) = \sum_{i=1}^{G} p(i,j) \qquad (4)$$

Mean

The mean values are denoted by μ_i and μ_j. It gives the contribution of individual pixel intensity for the entire image.

$$\mu_i = \sum_{i=1}^{G} p_i(i) \quad (5)$$

$$\mu_j = \sum_{j=1}^{G} p_j(j) \quad (6)$$

Standard Deviation

The standard deviation is denoted by σ_i and σ_j

$$\sigma_i^2 = \sum_{i=1}^{G} (i - \mu_i)^2 p_i(i) \quad (7)$$

$$\sigma_j^2 = \sum_{j=1}^{G} (j - \mu_j)^2 p_j(j) \quad (8)$$

Variance

Variance is used to find how each pixel varies from the neighboring pixel and is used in classifying different regions.

$$Variance = \sum_{i,j} \left(p(i,j) \frac{1}{N} \right)^2 \quad (9)$$

Entropy

Entropy: Entropy is the measure of spatial disorders within the image. For textually uniform image, entropy will be small.

$$Entropy = -\sum_{i,j} p(i,j) \log(p(i,j)) \quad (10)$$

Energy

Energy: Energy is used to measure the local homogeneity and it is opposite to entropy. Energy will measure the textural uniformity.

$$Energy = \sum_{i,j} p(i,j)^2 \quad (11)$$

Homogeneity

It is also called as Inverse Difference Moment. It measures image homogeneity and assumes larger values for smaller gray tone differences in pair elements. It is more sensitive to the presence of near diagonal elements in the GLCM. It has maximum value when all elements in the image are same. When energy is constant, homogeneity decreases if contrast is increased.

$$Homogeneity = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i-j)^2} p(i,j) \quad (12)$$

Correlation

The correlation feature is a measure of gray tone linear dependency in the image.

$$Correlation = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} p(i,j) \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} \quad (13)$$

Contrast

Contrast: Contrast is defined as the local gray level variation in pixel intensity value. It is the difference between the highest and the lowest values of a contiguous set of pixels. It measures the amount of local variations present in the image.

$$Contrast = \sum_{i=0}^{G-1} n^2 \left\{ \sum_{i=1}^{G} \sum_{j=1}^{G} p(i,j) \right\}, |i-j| = n \quad (14)$$

Classification

Once the feature has been extracted, it is classified using artificial neural network.

6. Artificial Neural Network

An artificial neural network is the mathematical model of the biological neural network. This network is made by the interconnection of artificial neurons together. Each neuron is a computational unit. Inputs are received by these neurons and are processed to produce the output. The information flow depends on the connection between the neurons. Neurons consist of inputs that is multiplied by the weights and computed with a mathematical function to find the activation characteristics of the neuron.

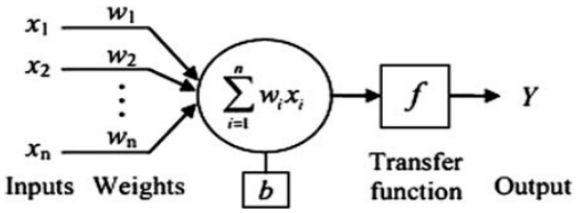


Figure 2: Structure of Neuron

Neural network is an adaptive system whose structure changes during the learning phase. There are three different learning paradigms in ANN. They are supervised learning, unsupervised learning and reinforcement learning. In supervised learning, the network is trained by giving both inputs and outputs. In unsupervised learning, only the inputs are given according to which the weights of the neurons are changed and the network is trained accordingly. Reinforcement learning is a special type of supervised learning. Here instead of output, this network provides only whether the output is good or bad.

7. Results and Discussion

This section explains how the moving object is detected with GLCM and ANN.

First the videos are converted into frames. Then the noises in the frames are removed using bilateral filter. It is a non-linear filter which preserves the edges of the images.

Background Subtraction

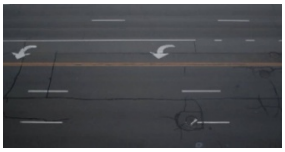
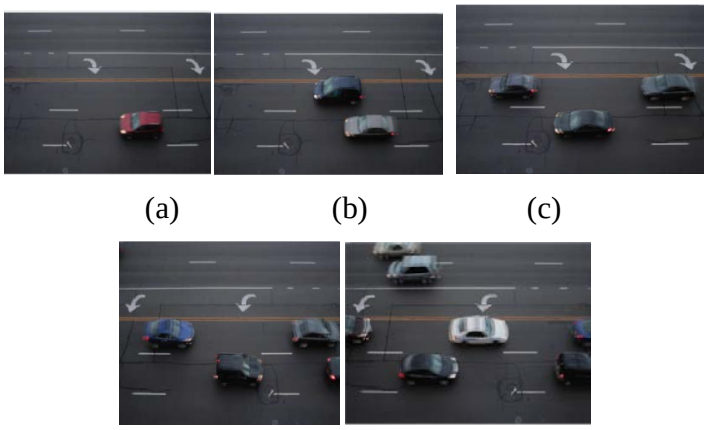


Figure 3: Background Model

In the background subtraction technique, the deviation in the pixels of the frame is identified which will be the pixels of the moving object in the video sequence. First a background model is considered and each frame is compared with this background model and the difference of the two images will be detected. The background model is shown in the figure 3.

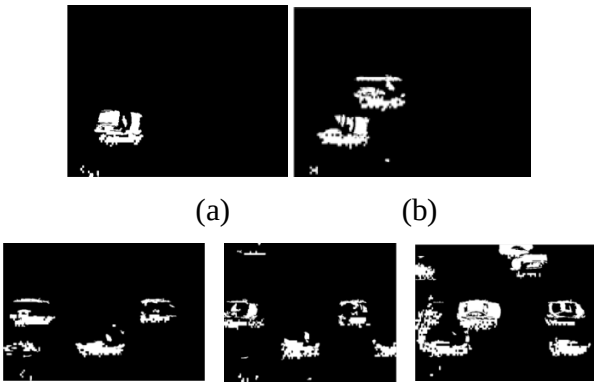
The current frames are shown in the figure 4. Each current frame is compared with the background model and the foreground mask is detected which gives the binary image of the moving object in the video and it is shown in figure 5.



(a) (b) (c)

(d) (e)

Figure 4: Current Frames



(a) (b)

(c) (d) (e)

Figure 5: Foreground Mask after Background Subtraction

Features are the base by which the objects dimensions like shape and size are detected. Gray Level Co-occurrence Matrix (GLCM) technique is used to find the different feature metrics like entropy, energy, homogeneity, contrast, correlation, mean, standard deviation, etc. are identified. Table 1 gives the extracted feature values of each frames. Then the extracted featured are classified using artificial neural network technique. In the proposed method, a supervised learning technique is considered and trained according to the collected database. The neural network classifies the frame by counting the number of moving objects in the video sequence.

Table I: Feature Values Extracted Using GLCM

Number of moving objects	Contrast	Correlation	Energy	Entropy	Homogeneity	Mean	RMS	Standard deviation	Variance
1	0.0461	0.9821	0.9466	0.1788	0.9992	0.0816	0.2372	0.4959	0.2184
2	0.0723	0.9823	0.9154	0.2577	0.9987	0.1209	0.3699	0.6663	0.4035
3	0.1016	0.9781	0.9031	0.2860	0.9982	0.3375	1.2906	1.7407	2.9136
4	0.1326	0.9777	0.8759	0.3465	0.9976	0.9063	3.3267	3.9628	15.1660
More than four	0.2511	0.9796	0.7432	0.6038	0.9955	1.8197	4.7372	5.2411	25.8694

8. Conclusion

Moving object detection has wide real time application. In this paper the moving object is detected from the video using GLCM and ANN. There are many existing techniques for background subtraction, feature extraction and for feature classification. For processing of images, first the noise in the video is removed with bilateral filter. Background subtraction method is applied to detect the object and to reduce the processing time.We use a normal subtraction technique to find the foreground objects from the background image. GLCM based feature extractor is used for extracting the features of the object to be detected. Then artificial neural network is used for feature classification and is trained to detect the number of objects which are in motion.

References

[1] Laptev I., Ramanan D., Sivic J., Guest Editorial Video Recognition, International Journal of Computer Vision 119 (3) (2016), 217-218.

[2] Chen S., Xu T., Li D., Zhang J., Jiang S., Moving object detection using scanning camera on a high-precision intelligent holder, Sensors 16 (10) (2016).

[3] Qian X., Liu X., Ma X., Lu D., Xu C., What Is Happening in the Video?—Annotate Video by Sentence, IEEE Transactions on Circuits and Systems for Video Technology 26 (9) (2016), 1746-1757.

[4] Saputra W.A., Herumurti D., Integration GLCM and geometric feature extraction of region of interest for classifying tuna,

International Conference on Information & Communication Technology and Systems (2016), 75-79.

[5] Curtis P., Harb M., Abielmona R., Petriu E., Feature selection and neural network architecture evaluation for real-time video object classification,IEEE Congress on Evolutionary Computation (CEC) (2016), 1038-1045.

[6] Brandan M.E., Contrast-enhanced digital mammography and other techniques based on image subtraction, AIP Conference Proceedings on AIP Publishing 1747 (1) (2016).

[7] Zhang Y., He Q., Wang H., Guan G., Xu T., Chen H., Background subtraction based on pixel clustering, IEEE International Conference on Information and Automation (2016), 1496-1500.

[8] Lee H., Kim H., Kim J.I., Background Subtraction Using Background Sets With Image-and Color-Space Reduction, IEEE Transactions on Multimedia 18 (10) (2016), 2093-2103.

[9] Lloyd K., Rosin P.L., Marshall D., Moore S.C., Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (GLCM)-based texture measures, Machine Vision and Applications 28 (3-4) (2017), 361-371.

[10] Devi K., Gupta P., Grover D., Dhindsa A., An effective texture feature extraction approach for iris recognition system, International Conference on Advances in Computing, Communication, & Automation (ICACCA) (2016), 1-5.

[11] Gaba N., Barak N., Aggarwal, S., Motion detection, tracking and classification for automated Video Surveillance, IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (2016), 1-5.

[12] Abedin Z., Dhar P., Hossenand M.K., Deb K., Traffic sign detection and recognition using fuzzy segmentation approach and artificial neural network classifier respectively, International Conference on Electrical, Computer and Communication Engineering (2017), 518-523.

[13] Abdelhedi S., Wali A., Alimi A.M., Human activity recognition based on mid-level representations in video surveillance applications, International Joint Conference on Neural Networks (IJCNN) (2016).