Intelligent Video Surveillance System based on Wavelet Transform and Support Vector Machine

Meghana M. Deshpande
Student
Department of Electronics & Telecommunication
MGM's Jawaharlal Nehru Engineering College,
Aurangabad-431003, Maharashtra, India

Jaideep G. Rana
Professor
Department of Electronics & Telecommunication
MGM's Jawaharlal Nehru Engineering College,
Aurangabad-431003, Maharashtra, India

ABSTRACT
Automatic human detection is an important application where security is the main concern. As human detection problem involves classification of two objects as humans and others, a human detection using intelligent video surveillance system is presented using support vector machine to detect human in surveillance field. In this paper, in order to improve the efficiency of the machine learning 2D Wavelet transform based features are used. It consists of wavelet statistical features and wavelet co-occurrence features which are obtained from red, green and blue layers of sample images. The sample images are obtained through the video which forms training input to SVM. The experimental results demonstrate that the proposed system achieves good success rate for wavelet co-occurrence features.

Keywords
Human detection, Support Vector machine, Video surveillance, Wavelet transform.

1. INTRODUCTION
In the last few years, visual surveillance has become one of the most active research areas in computer vision, especially due to the growing importance of visual surveillance for security purposes. The traditional video surveillance needs a complete surveillance which wastes a mass of manpower and financial. With the aim to recognize moving human figure through video monitoring system, it is able to detect moving human figure in a surveillance area. Several interesting approaches for human detection have been reported till date. Ye and Wen [1] proposed the feature extraction method using DWT with support vector machine. Xinnan, Lizhong, Xuewu, and Lei [2] presented a grid and center radiating vector features with SVM classifier. Mostayed, Mazumder, Kim and Park [3] attempted classification of human using K-means clustering based on DFT and DWT features. Sudha and Bhavani [4] segmented individual frames using background subtraction algorithm from which wavelet features are extracted and used with SVM. El-Yacoubi, Shaiek and Dorizzi[5] Proposed a model-free approach and dynamics of human figure are modeled using Hidden Markov Models. Li [6] presented a novel gait recognition based on Haar wavelet and fused Hidden Markov Models. Libin and Wenxin [7] presented a algorithm in gait identification based on Haar wavelet and support vector machine.

In this article intelligent video surveillance system is developed for human detection. 2D wavelet transform based statistical and co-occurrence feature extraction algorithm developed for obtaining eighteen features from red, green and blue layers of human and non-human images. Further, support vector machine classification algorithm is trained using the wavelet based features and demonstrated its working for various possibilities of data so it can be used for human detection.

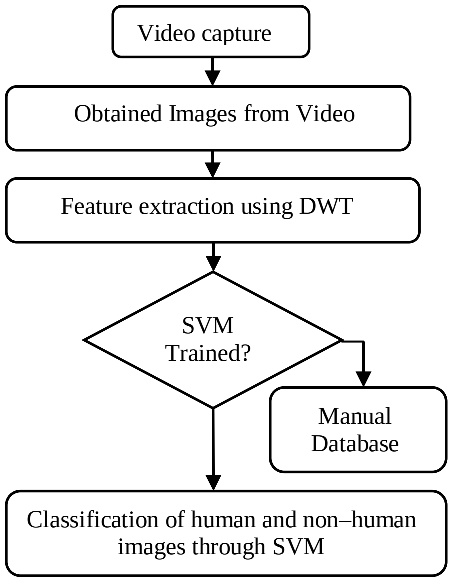

Fig 1: Process overview

A SVM based intelligent video surveillance system is developed by acquiring video of the surveillance field. Figure 1 shows process of human detection in this we have captured the video using web camera under natural lightning conditions. Then this video is converted into image sequences which are used for feature extraction process. Each obtained image stored as 24 bit color image in computer memory. Feature extracted from these images are then used for developing a SVM classification system which can be useful to detect human in surveillance field.

2. FEATURE EXTRACTION PROCESS
Wavelet analysis is an advanced feature extraction algorithm which is based on windowing technique with variable sized regions. The window size can be kept wide for low frequencies and narrow for high frequencies which lead to an optimum time frequency resolution for complete frequency range Mukane, Gengaje, and Bormane [8]. A discrete wavelet transform (DWT) is the wavelet transform process in which the wavelets in numerical analysis and functional analysis are discreetly sampled. Temporal resolution is a key advantage of wavelet transform over Fourier transform in which it captures both frequency and location information. In this study single level 2-D wavelet transform decomposition is used for features extraction. The 2-D wavelet transform will give four matrices the approximation coefficients matrix and detailed

coefficients matrices horizontal, vertical, and diagonal, respectively. The approximation matrix CA is obtained for the three different layers of red, green and blue which is converted to two types of features such as wavelet statistical features and wavelet Co-occurrence features. Statistical features are mean, standard deviation, skewness and energy and co-occurrence features are cluster shade and cluster prominence. Thus total 18 features characterize one human image from three different layers of image that is red, green and blue.

2.1 Wavelet Statistical Features

Based on available wavelet coefficients mean(1),standard deviation(2), skewness(3) and energy(4) of all sub bands up to second level decomposition are calculated as features by using following equations

Mean:

$$Mk = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} x(i, j) \qquad (1)$$

Standard deviation:

$$\sigma_k = \left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(x_k(i,j) - \mu_k(i,j)^2)\right]^{\frac{1}{2}} \qquad (2)$$

Skewness:

$$S_k = \left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{E(x_k(i,j) - \mu_k(i,j)^3)}{\sigma(i,j)^3}\right] \qquad (3)$$

Energy:

$$E_k = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|x_k(i, j)\right| \qquad (4)$$

2.2 Wavelet Co-occurrence Features

The co-occurrence matrix features are obtained from whole sample image and two levels of DWT decomposed sub-bands coefficients of sample image. Co-occurrence matrix is derived for distance vector d(i,j) i.e offset is taken as d(1,1) from co-occurrence matrix. Features namely cluster shade and cluster prominence are obtained by equation (5,6) respectively.

Cluster shade:

$$\sum_{i,j=1}^{N}\left(i - Px + j - Py\right)^{3} Co(i, j) \qquad (5)$$

Cluster prominence:

$$\sum_{i,j=1}^{n}\left(i - P_x + j - P_y\right)^{4} Co(i, j) \qquad (6)$$

3. SUPPORT VECTOR MACHINE

SVM maps input vectors to a higher dimensional space where a maximal separating hyperplane is constructed [9]. SVM maps the input patterns into a higher dimensional feature space through some nonlinear mapping chosen a priori. A linear decision surface is then constructed in this high dimensional feature space. Thus, SVM is a linear classifier in the parameter space, but it becomes a nonlinear classifier as a result of the nonlinear mapping of the space of the input patterns into the high dimensional feature space. The mathematical background of SVM formulation is explained briefly in this section. Detailed discussion on SVM is available in references[10,11].

For linearly separable data set of {(xi , ci)}, the hyperplane obtained by SVM for classifying the data into two classes can be written as:

$$w.x - b = 0$$

Which implies

$$c_i(w.x_i - b) \geq 1 \qquad 1 \leq i \leq n$$

Where, xi is input sample and ci output having either 1 or -1, a constant denoting the class. The vector w is perpendicular to the separating hyperplane and the offset parameter b allows in increasing the margin. For the linearly separable training data, these hyperplanes can be considered to maximize the distance between the extreme points of each class. The distance between the hyperplanes can be given as $2/|w|$. Therefore, maximization of the distance between the hyperplanes becomes a problem of minimization of $|w|$. The primal form of optimization problem becomes a quadratic programming optimization which can be written as,

Minimize $(1/2)\|w\|^2$;

subject to $c_i(w.x_i - b) \geq 1 \qquad 1 \leq i \leq n$

The factor of 1/2 is used for mathematical convenience. The dual form of this optimization problem leads to a classification problem which is only a function of the support vectors, i.e., the training data that lie on the margin. Originally, the SVM algorithm was developed for as a linear classifier. However, it was further extended [12] to create a nonlinear classifier by applying the kernel trick. Some common kernels include, Polynomial (homogeneous): $k(x,x) = (x,x)^d$.

Polynomial (inhomogeneous): $k(x,x) = (x, x+1)^d$;

Radial Basis Function: $k(x,x) = \exp(\gamma\|x - x\|)^2$, for $\gamma > 0$; Gaussian Radial Basis Function:

$$k(x,x) = \exp\left(-\frac{\|x - x\|^2}{2\sigma^2}\right)$$

3.1. Formulation of SVM Based Human Detection System

The training set is the foundation in constructing SVM classifier and the capability of the training set affects the performance of the SVM classifier directly. The sample set should include human sample and non-human sample. We can get the human sample through the following two methods extracting the sample from the images as indicated earlier; selecting images including human and segmenting human sample manually. The non-human sample includes all kinds of images including all kinds of animals and vehicles. The combination of human sample and part non-human sample forms the training input to SVM and an initial SVM classifier. The final sample set we obtained includes 1500 human samples and 1500 non-human samples. The number of samples required for training are selected based on plot of coefficient of variation for different number of sample shown in Figure 2 from the figure it is observed that for 1500 number of samples Coefficient of variation become constant .hence in this actually we have used 1500 samples for human and non-human images.

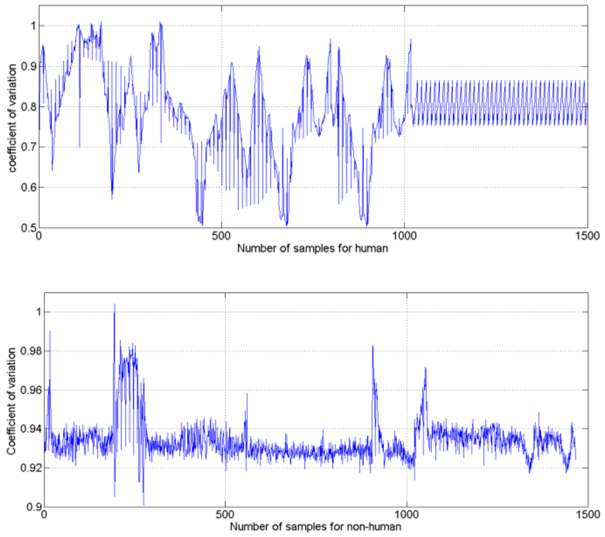

Fig 2: Sample selection method

We test our algorithm based on the mixed database we have taken video consist of human walk through surveillance region with different gestures. This video is then converted in image sequences. The pixel of the image is 320x240 PC saves the 15 images per second. Application program does feature extraction at last, the trained SVM detector is used for decision making.

4. RESULTS AND DISCUSSION

In the current study, the SVM is used for nonlinear classification of human and non-human image features using the radial basis function kernel. Fig. 3 and Fig.4 shows the classification of human and non-human sample for 320x240 pixel images. SVM classifies human and non-human data.

The performance of two features classifications is shown in Figure 3 and Figure 4 for wavelet statistical features and wavelet co-occurrence features, respectively. Wavelet statistical features include mean (MG-MB), standard deviation (StdG-StdB), Skewness (SkG-SkB) and Energy (ER-EB) and Wavelet Co-occurrence features include cluster shade (CsR-CsB) and cluster prominence (CpR-CpB). Figures shows (x) 'Human' class data, (□) 'Non-human' class data, dashed black line (-.) Shows decision boundary, and thick lines separates hyper planes. The vectors (points) that constrain the width of the margin are the support vectors dashed line is the decision boundary (Classification boundary) makes separation between two classes SVM handles this by using kernel function; It transforms the data into higher dimensional space to make it possible to perform the separation. Separation of data at wavelet decomposition level 1 is complicated so to increase the classification rate we used decomposition level 2. From Figure 3 and Figure 4, it is observed that data is well separable into two classes.

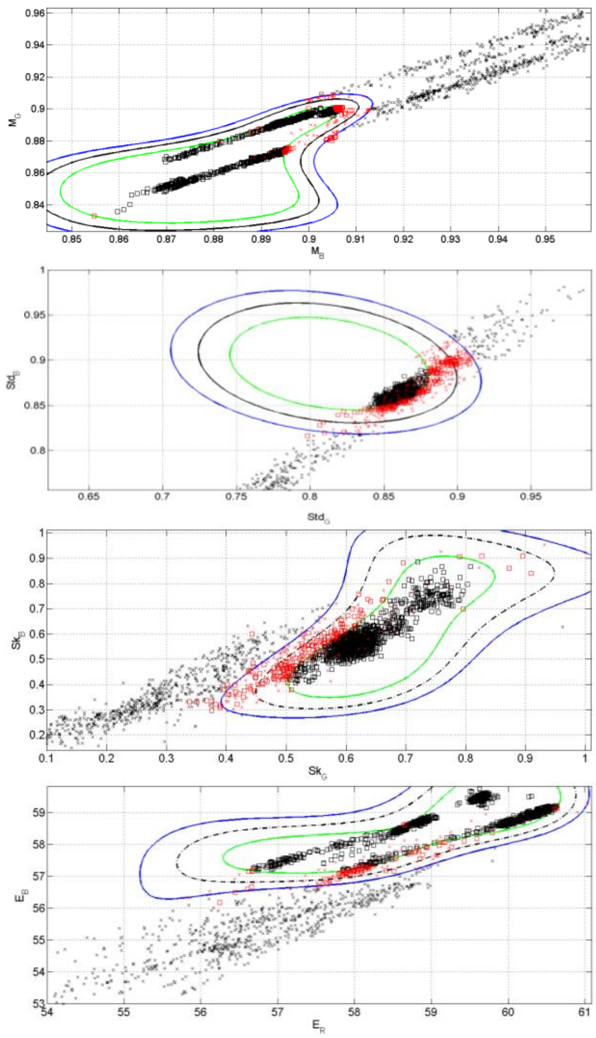

Fig 3: Two feature classification of Wavelet statistical features

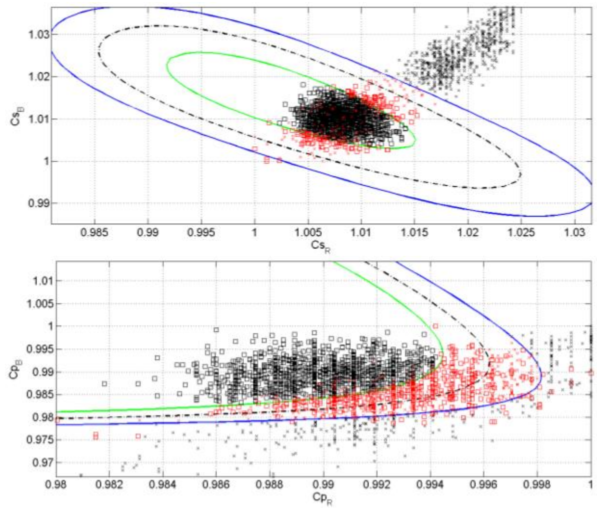

Fig 4: Two Feature classification of Wavelet Co-occurrence Features

The performance of the SVM classification is measured in terms of success rate which is the ratio of number correctly classified sample against the total number of samples used for classification. The performance of SVM is estimated for different groups of features as shown in Figure 5 Figure shows that SVM is tested for six set of features viz. mean, standard deviation, skewness, energy, cluster shade and cluster prominence. Figures shows that the cluster shade and cluster prominence features shows success rate of almost 100% performance. The mean and standard deviation features show lowest performance of about 85%, where skewness and energy based features show the success rate of 95% and 97%, respectively.

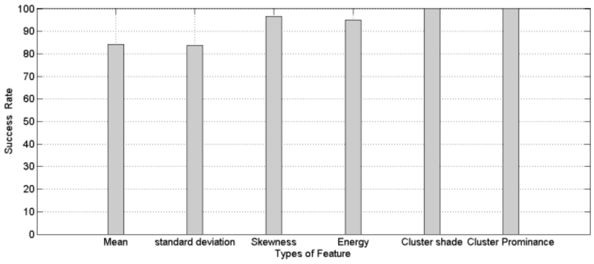

Fig 5:Success rate for different features

5. CONCLUSION

In this paper, the application of nonlinear feature extraction for human detection is presented. The features are obtained using wavelet transforms and SVM is used for classification of the data. The result of human detection using SVM shows that co-occurrence feature viz. cluster shade, cluster prominence gives success rate is of almost 100% this gives the confidence that this algorithm can be used to reduce the efforts of surveillance personnel. Further this algorithm will be learned for non ideal condition, such as variation in light, background effect etc.

6. REFERENCES

[1] Ye, B. and Y, Wen. 2007 Gait Recognition based on DWT and SVM, International Conference on Wavelet analysis and Pattern recognition, 1382-1387.

[2] Xinnan, F. Lizhong, X. Xuewu, Z. and Lei, C. 2008 The Research and Application of Human Detection Based on Support Vector Machine using in Intelligent Video Surveillance System, Forth International Conference on Natural Computation,139-143.

[3] Mostayed, A. Mazumder, M. Kim, S. and Park, S. 2008 Abnormal Gait Detection Using Discrete Wavelet Transform in Domain, 4th Kuala Lumpur International Conference On Biomedical Engineering,383-386.

[4] Sudha, L. and Bhavani, R. 2011 SVM based biometric authorization system by video analysis of human gait, International Conference on Electronics Computer Technology, 301-304.

[5] El-Yacoubi, M. Shaiek, A. and Dorizzi, B. 2011 HMM-based gait modeling and recognition under different walking scenarios, International Conference of Multimedia Computing and Systems,1-5.

[6] Li, P. 2011 Gait Recognition via Fused Hidden Markov Models, Multimedia Information Networking and Security, 187-190.

[7] Libin, D. and Wenxin, S. 2011 An Algorithm of Gait Recognition Based on Support Vector Machine, Journal of Computational Information Systems, vol. 13, 4710-4715 .

[8] Mukane, M. Gengaje, S. and Bormane, S. 2011On Scale Invariance Texture Image Retrival using Fuzzy Logic and Wavelet Co-occurrence based Features, International journal of computer applications, Vol.18(3),10-17.

[9] Pawar, p. and Jung, s. 2007 Support Vector Machine based Online Composite Helicopter Rotor Blade Damage Detection System, Journal Of Intelligent Material Systems And Structures, vol. 19, 1217-1228.

[10] Cortes, C. and Vapnik, V. 1995 Support Vector Networks, Machine Learning, 273-297.

[11] Vapnik, V. 1995 The Nature of Statistical Learning Theory, Springer-Verlag, New York.

[12] Boser, B. Guyon, I. and Vapnik, V. 1992 A Training Algorithm for Optimal Margin Classifiers, Proceedings of the Fifth Annual workshop on Computational learning Theory Pittsburgh Pennsyl vania USA, 144-152.

7. AUTHOR'S PROFILE

Ms.Meghana M . Deshpande is a student of MGM's Jawaharlal Nehru Engineering College, Aurangabad, Maharashtra, India and pursuing her M.E. Electronics.

Prof. Jaideep G. Rana is working as HOD & Professor in E & TC Department in MGM's Jawaharlal Nehru Engineering College, Aurangabad, Maharashtra, India.