# Intelligent Video Object Classification Scheme using Offline Feature Extraction and Machine Learning based Approach

**Chandra Mani Sharma[1], Alok Kumar Singh Kushwaha[2] ,Rakesh Roshan[3] , Rabins Porwal[4] and Ashish Khare[5]**

**[1,3,4]Department of Information Technology, Institute of Technology and Science
Ghaziabad, U.P., India**

**[2] Department of Computer Engg. and Application, G.L.A. University,
Mathura, U.P., India**

**[5] Department of Electronics and Communication, University of Allahabad,
U.P., India**

## Abstract

Classification of objects in video stream is important because of its application in many emerging areas such as visual surveillance, content based video retrieval and indexing etc. The task is far more challenging because the video data is of heavy and highly variable nature. The processing of video data is required to be in real-time. This paper presents a multiclass object classification technique using machine learning approach. Haar-like features are used for training the classifier. The feature calculation is performed using Integral Image representation and we train the classifier offline using a Stage-wise Additive Modeling using a Multiclass Exponential loss function (SAMME). The validity of the method has been verified from the implementation of a real-time human-car detector. Experimental results show that the proposed method can accurately classify objects, in video, into their respective classes. The proposed object classifier works well in outdoor environment in presence of moderate lighting conditions and variable scene background. The proposed technique is compared, with other object classification techniques, based on various performance parameters.

***Keywords:*** *Object Classification, Machine Learning, Real-time Video Processing, Visual Surveillance.*

## 1. Introduction

Object classification is an important step in object detection [1], object activity recognition[2], content based video retrieval, visual surveillance[3], etc. Object classification techniques can be categorized primarily into two groups; Shape-based object classification [3-5] and Motion-based object classification [6-8]. The concept of Machine Learning can be used for classification and the training in such classification system is either based on pixels or features. The feature-based approaches are preferred over pixel-based approaches, for object classification task, because encoding of ad-hoc domain knowledge using features is easier and it is difficult to train finite quantity of data using pixels [1]. Moreover, the feature-based object classification methods are better than the pixel-based methods in terms of speed. Haar-like features have been used, in synergy with different training techniques, for creating machine learning based classifiers [10-11]. Sialat et al.[10], in their pedestrian detection system, used Haar-like features along with the *decision tree*. A versatile object recognition technique has been proposed in [11], using multi-axial Haar-like features and a compact cascaded classifier. Viola et al.[1] also used the modified reminiscent of Haar-basis functions [12], for accomplishing object detection task.

The number of available features, for training the classifier, may be considerably high. Not all the features have equal importance, therefore some sort of mechanism is used to extract those features which are more important from classification point of view. Boosting combines different weak classifiers to form highly accurate predictors [13]. According to [14] AdaBoost is the best available binary classifier. Viola and Jones [1] proposed an object detection framework based on Haar-like features and AdaBoost, and used this framework for detecting the frontal human faces in real-time. One notable contribution, of the work of Viola et al.[1], was the concept of Integral Image, for fast extraction of Haar-like features. Since they have only two classes in their face detection case: Human and Non-Human. So they used AdaBoost for fast training of features. AdaBoost has a serious limitation that it can only be used to solve binary classification problem. AdaBoost requires that the accuracy of constituent classifiers be more than 50%. This condition, in AdaBoost,

is not easy to fulfill in case if number of the classes, in classifier, exceeds two.
This paper proposes a multiclass object classification technique based on Haar-like features and uses a Stage-wise Additive Modeling using a Multiclass Exponential loss function (SAMME) [13]. SAMME was originally used to classify the chemical data into multiple classes in [13]. This paper reports the classification results, obtained by the application of SAMME on image data. We use Integral Image representation, for fast feature evaluation and then multiple weak classifiers are trained using SAMME. The observation weights, associated with the feature points, are changed depending on their accuracy of classification during training process. Finally, the classifiers are linearly combined to form a strong classifier.
The rest of the paper is organized as follows: Section 2 explains the multiclass boosting of classifier training, Section 3 describes the methodology of the proposed object classification technique, experimental results have been described in Section 4 and conclusions are given in Section 5.

## 2. Multiclass Boosting of Classifier Training

Objective of classification is to distribute the input vector (like feature points) into a set of K classes $C_k$ , where k=1,…K-1. These classes are disjoint so that ,each vector point x can be assigned to one and only one class. Boosting is an approach, used to improve the performance of classifier-training. Using boosting, the weak classifiers are trained sequentially and are finally merged, to form a classifier with adequate confidence of performance. A weight is assigned to every feature-point, to be classified, before classification. The feature points gain or lose weights during this process. Points, which are classified with error, gain weight and the points, which are correctly classified, lose weight. Points, with the higher weight value, seek more attention in the next level of classification, in order to classify them in the right classes. There are various classifiers [13- 15] available, which are created using this approach. AdaBoost [15] is a good approach for binary classification and there are two requirements associated with the classifiers in the AdaBoost- (i) the classifier must have an accuracy greater than 50% and (ii) the classifiers should be capable of representing the weighted data points. For the first condition, if the achieved accuracy is exactly 50% then distribution weights will not be updated. if the accuracy is less than 50% then the updation of distribution weights will take place in opposite direction. For binary classification, the random accuracy of classification of a feature point, in one of the two classes, should be at least 50% [15]. The second condition can be bypassed easily if the samples are taken from the training data set with replacement according to the weight distribution and then passing to the component classifier. Suppose we try to use AdaBoost with a multiclass classifier, having three classes, then random accuracy will be 33.33%, which violates the first requirement for the use of AdaBoost. This constraint can be solved using two popular approaches [16]- (i) *one-against-all* and (ii) *one-against-one*. In case of *one-against-all* approach a separate model for each class is trained to distinguish the samples of that class from the samples of the remaining classes. For a data point to be classified using *one-against-all* approach, the class which gets the highest class prediction from the probabilistic binary classifiers, is assigned for that data point. In case of *one-against-one* approach, a classification model for each pair of classes is created. If the number of classes are K then there will be K(K-1) such models. Here, the class for a data point is decided by the voting classifiers in ensemble. The performance of such model based approaches is hampered by the management of models. Multiple models cause the processing speed to slow down. Therefore these approaches are not efficient options, to be used, for real-time visual surveillance applications. The problem of multiclass boosting can be solved by transforming it into several two-class problems. The general approach based on this concept is called Error-Correcting Output Coding (ECOC) [17]. ECOC is a method of making the most of the transformation of the multiclass problem into several binary problems. A simple ECOC method is known as Hamming Encoding. The problem of multiclass boosting of classifiers can be nicely solved with the SAMME [13], without creating any extra model and experimental results show that the execution speed is at par with binary AdaBoost based methods.

## 3. The Proposed Method

In proposed technique, the *Integral Image* representation is used for fast feature evaluation and boosting of cascaded classifiers is performed using SAMME. A more robust classifier is created by linearly combining the multiple weak classifiers. For experimentation, we consider three classes: *Human, Car* and *Non-Human-Car*. After classifier training, the objects, in the video, are classified into *Human* and *Car* classes. The proposed method has following steps:

### 3.1 Sample collection

The sample images for training the classifier are collected first. We have collected images for three classes- humans, cars and images which belong neither of these two from our own captured images and images from standard datasets like CalTek101, MIT-CMU datasets. We have created our own data set which consists of 4,000 images of humans, 3,500 images of cars and 5,000 images which are neither humans nor cars. These images were resized to dimension 60x60 that they consist of only one object per image. This was performed in order to make the classifier learn more domain information from small number of images and this helps to improve the accuracy of the classifier.

### 3.2 Integral Image and Haar-basis function

We use simple Haar-like features which are reminiscent of Haar-basis functions as used in[1]. The use, of features instead of pixels, makes the classifier system work faster and helps in encoding the domain knowledge with finite quantity of data. We use three types of features- *two-rectangle feature*, *three-rectangle feature*, and *four-rectangle feature*. The Difference, between sums of pixels, gives the value of two-rectangle feature. The regions are horizontally or vertically adjacent and have the same size and shape (shown in Fig.1). A *three-rectangle feature* is used to compute the sum within two outside rectangles, subtracted from the sum in a center rectangle and a *four-rectangle feature* computes the difference between diagonal pairs of rectangles.

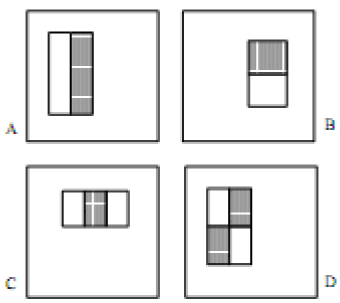



Fig1: Rectangle features

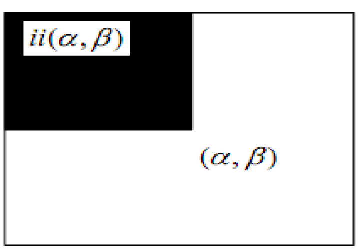



Fig2: Integral Image

The number of features, in a small detection window, can be considerably very high. For example, a base detection window of size 25x25 contains more than 200,000 such rectangle features. This huge number of features is very difficult to compute. To speed up this Computation process, we use *Integral Image* representation originally proposed by Viola et al. [1]. The Integral Image, $ii(\alpha,\beta)$, shown in Fig.2, at location $(\alpha,\beta)$ contains the sum of all pixels above and left of $(\alpha,\beta)$ and can be computed in a single pass over image using the following pair of equations.

$$\omega(\alpha,\beta)=\omega(\alpha,\beta-1)+i(\alpha,\beta) \tag{1}$$

$$ii(\alpha,\beta)=ii(\alpha-1,\beta)+\omega(\alpha,\beta) \tag{2}$$

where, $\omega(\alpha,\beta)$ is the cumulative row sum and $i(\alpha,\beta)$ is the original image.

### 3.3 Multiclass boosting with SAMME

We use Stage-wise Additive Modeling using a Multiclass Exponential loss function (SAMME) [13] for boosting the cascade of multiclass classifiers. Viola et al.[1] used binary AdaBoost for their face detection system which fails in case of multiclass classification because of certain constraints over classifier's accuracy. The proposed method has three classes of classification and a modified AdaBoost is used for this. SAMME is an extension of AdaBoost with a modification. The key for applying SAMME for multiclass problem is that the component classifiers are no longer required to achieve an accuracy greater than 50%, but instead need only it to be better than random guessing. Suppose we are given a set of training data $\{(\sigma_1,c_1),....(\sigma_n,c_n)\}$, where the input vector $\sigma_i$ € Rp, and the $c_i$ qualitatively assumes values in a finite set {1,2,…,K} where K is the number of classes. The training data are independently and identically distributed samples from an unknown probability distribution. The goal is to find a classification rule $\phi(\sigma)$ from the given training data, so that for a given new $\sigma$, we can assign it to a class label c from {1,…,K}. A complete SAMME multiclass boosting algorithm is given as below-

<u>SAMME Multiclass Boosting Algorithm</u>

- $\theta$ is vector of observation weights and $\lambda$ is the vector length
- M is the number of stages in classifier.
- $\alpha(\sigma)$ denotes a weak classifier
- $\phi(\sigma)$ denotes the final classifier, created using the linear combination of many weak classifiers
- $\eta^m$ is the score assigned to m^th classifier.

1. Initialize the observation weights $\theta_i=\frac{1}{\lambda}$, i=1,2,………, $\lambda$.
2. For $m=1$ to $M$
   (a) Fit a classifier $\alpha^{(m)}(\sigma)$ to the training data using weights $\theta_i$
   (b) Compute the misclassification error rate as,
   (c) Compute

$$\eta^m = \log \frac{1 - err^{(m)}}{err^m} + \log(K-1) \quad (3)$$

(d) Set $\theta_i \leftarrow \theta_i . \exp(\eta^{(m)} \bullet \delta(c_i \neq \alpha^m(\sigma_i)))$ (4)

for i=1,2,….., $\lambda$ .

(e) Re-normalize the value of $\theta_i$

**3.** Output

$$\phi(\sigma) = \arg\max_k \sum_{m=1}^{M} \eta^m \bullet \delta(\alpha^m(\sigma) = k) \quad (5)$$

Here, weak classifiers are linearly combined to form a strong classifier.

**4.** End

The above boosting algorithm is similar to AdaBoost but with a small but crucial difference in Eq. 3. The term log(K-1) has been added in the equation. From this term, it is evident that at K=2, SAMME reduces to AdaBoost but in multiclass case (K>2) the term log(K-1) is critical. One of the benefits is that for ηm to be positive, we just need have (1-err(m))>1/K or the accuracy of each classifier to be better than the random accuracy rather than 50%. In the present case, we have three classes- *Human*, *Car* and *Non-Human-Car*. Here, the accuracy of the weak classifiers needs to be only more than 33.33%, rather than being more than 50%. SAMME behaves as the forward stage-wise additive modeling for multiclass classification case.

## 4. Experimental results

The proposed object classification technique has been tested on several videos, captured in real outdoor environment. The results, with one such representative video, are given in Fig.3 using the proposed approach. In these results, red windows are used for classifying the human objects and blue windows are used to classify the cars. Fig.3 shows the object classification results with natural lighting conditions.

One can observe from the results presented in Fig. 3 that there are three human objects and one car in the video. The video was shot at a frame resolution of 640×480 and frame rate of 20 frames per second. The video consists of total 1600 frames and we have posted results at difference of 25 frames. In the starting of video the humans start moving towards camera and later on occupy the random motion paths. The car shown in the video has been parked and does not move in video. The proposed technique is capable of handling the partial inter-class and intra-class occlusions. Video frames no. 50, 225, 250 and 275 have the partial human to human occlusion and frames no. 1025, 1050 and 1400 show the occlusion between car and human. The proposed method classifies the objects even in these frames accurately and in case of full occlusion proposed method has the capability of resuming quickly. As it is clear from the results given in Fig.3 that the human objects have the varying poses of their bodies as they move in the video in appearing in all the possible natural poses. Moreover, they acquire various views like frontal, side and back to the camera. Also, their speed of movement varies. Unlike various existing classification methods, the proposed method classifies the various objects accurately irrespective of their views of appearance.

Many existing object classification schemes require the lighting conditions to be good in the video and can work only in presence of the statically good light. But from practical point of view it is not desirable. As in case of visual surveillance the videos are captured in the outdoor environment and fluctuations in the lighting are unavoidable. The classification results using proposed technique in poor lighting condition are given in Fig. 4. The video for this experiment was shot in evening at University of Allahabad campus. The resolution of video frames is 640×480 and was shot at frame rate of 20 frames per second. The video consists of a total 800 frames. The classification results with this video are given below starting from frame number 25 to frame number 800 at a difference of 25 frames. It can be observed that in the video three humans and one car appear. Humans are walking in the random directions. Car has been parked and does not move. The humans while walking cause occlusion within themselves and with car. Frames 250 and 675 show the partial occlusion between humans where one human is partially occluded from another. Frames 625, 650, 725, 750, 775 and 800 show the occlusion between car and human. In frame number 625 and 650, one human object occludes the car while in frame number 725, 750, 775 and 800 two humans appear before car and occlude it. It is evident from the results that the proposed technique accurately classifies the objects in all the aforementioned frames also.

We have tested the proposed method on several other realistic videos shot in outdoor environment. The average execution speed on these videos has been observed as 23 frames per second. The detection accuracy of the proposed classifier system has been estimated in our experiments between 91% and 98%. The tradeoff between the accuracy and the execution speed can be customized easily depending upon the user's requirement. The execution speed can be increased by slightly decreasing the detection accuracy and the detection accuracy can be increased by a slight decrement in execution speed.

frame 25

frame 50

frame 75

frame 100

frame 125

frame 150

frame 175

frame 200

frame 225

frame 250

frame 275

frame 300

frame 325

frame 350

frame 375

frame 400

frame 425

frame 450

frame 475

frame 500


frame 475


frame 500


frame 725


frame 750


frame 525


frame 550

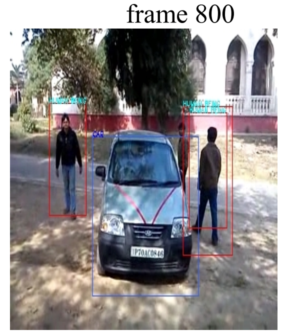

frame 775


frame 800


frame 575


frame 600


frame 825

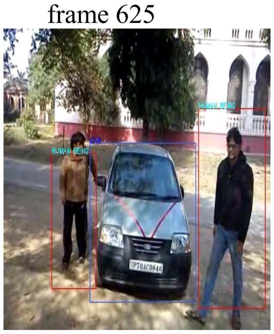

frame 850

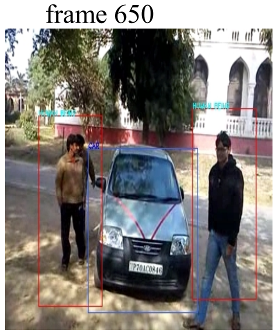

frame 625


frame 650


frame 875

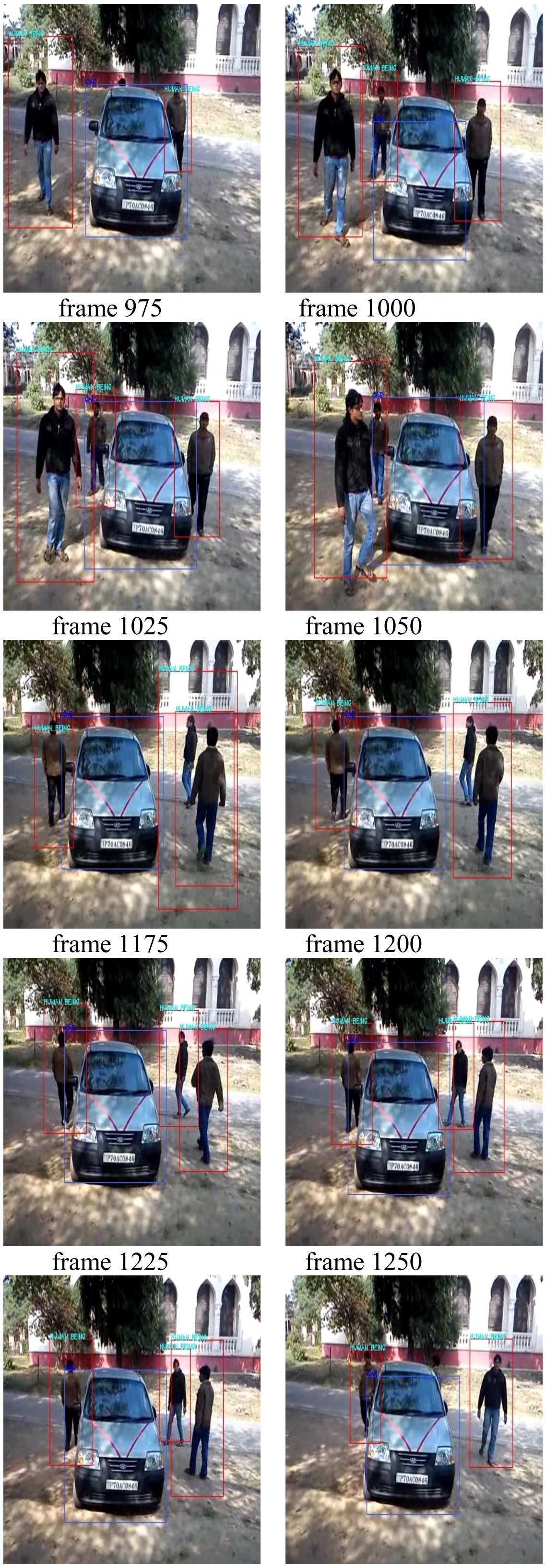

frame 900


frame 675


frame 700


frame 925


frame 950



frame 975 frame 1000
frame 1025 frame 1050
frame 1175 frame 1200
frame 1225 frame 1250
frame 1275 frame 1300

frame 1325 frame 1350
frame 1475 frame 1500
frame 1525 frame 1550
frame 1575 frame 1600

**Fig. 3:** Object classification with the proposed technique in normal lighting condition.

The operation of the proposed method has also been tested with the implementation of a real world system. The system was kept under test operation for 60 hours continuously. The system was configured with a single mounted camera. The system processed the captured video stream in real-time without any time delay. Object classification produced adequately accurate results.

### 4.1 Comparison of object classification methods

We compare the proposed method of object classification with various other existing state-of-the-art methods. This comparison is shown in table 1. Chosen parameters for comparison are: number of classes, number of training samples, object classification accuracy, processing speed, method used for feature extraction and machine learning algorithm used for classifier-training. This performance statistics has been observed on the uniform hardware set up. The performance of the proposed technique is comparable to other existing techniques. The peak detection accuracy of our method is 98% which is more than the peak accuracy of the object classification methods [19-22]. It can be observed from the table 1 that the computational speed is also comparable to the other methods. The proposed method has computational speed of 25 frames per second which is good enough for real-time processing of video frames.

Object classification is a very crucial step towards object activity recognition and behavior prediction. If done accurately then it can considerably improve the quality and accuracy of the various applications such as in monitoring and visual surveillance systems, content based video retrieval and indexing, access control to restricted areas and crowd flux analysis. Machine learning based approaches for object classification are gaining much popularity compare to the other approaches of object classification methods. This is because of that a huge amount of unannotated training data is readily available for the unsupervised training of these systems.

Also, various standard annotated training sets are available including a range of object classes in order to facilitate the supervised training of the system as well. The customizable applications for this purpose have started appearing which include training data set as well a user friendly kit to train with that data set. Generally, these types of kits are able to detect only a very few number of object classes. Majority of them can only recognize (distinguish) only one class of object. Generally these systems are known as object detectors (contrast with object classifiers) and are able to detect, generally, only a single type of objects out of various types of objects such as human detector, car detector, cow detector etc. These detector systems may also be trained to detect some special part of an object instead of recognizing the object as a whole e.g. Viola et al.[1] developed a human face detector system which was capable of tagging the human faces in an image. The system is slower and takes much time to detect the faces. Sun et al.[20] also detected the vehicles using their vehicle detector system. This system is even slower than the system of Viola et al.

**Table1**: Comparison of the proposed method with other existing methods for object classification

| Method Name | Criteria for Evaluation | | | | |
|---|---|---|---|---|---|
| | ***Number of Classes and Training Samples*** | ***Accuracy*** | ***Speed*** | ***Method used for Feature Extraction*** | ***Machine Learning Algorithm used*** |
| Kato et al. [19] | 2 Classes: 5000 Vehicle and 5000 Non-Vehicle samples to train | 89.0% to 96.0% | 0.5- 1.7 Sec per Frame | Not Specified | MC-MQDF - Linear Classifier |
| Sun et al. [20] | 2 Classes: 2500 Vehicle, 2500 Non-Vehicle Samples to train | 88.9% to 96.4% | 1.0-2.0 Sec per Frame | PCA, Wavelet, Gabor Features | Neural Networks & Support Vector Machines |
| Viola et al. [1] | 2 classes: 2000 Face and 3000. Non-Face \Samples to train | 78.3% to 98% | 0.07- 0.5 Sec per Frame | Haar-like features | AdaBoost |
| Opelt et al. [21] | 3 Classes: 450 Person, 350 Bike 250 Non Bike / Person Samples to train | 65.0 to 83.5% | 2.1-3.0 Sec per Frame | Intensity Moments, SIFTs | AdaBoost |
| **Proposed Method** | 3 classes: 4000 Human, 3500 Car and 5000 Non-Car/Human Samples to train | 91.6% - 98.3% | 0.01 - 0.05Sec per Frame | Haar-like Features | SAMME |

## 5. Conclusions

A real-time multi object classification approach has been proposed in this paper. The approach first trains a multiclass classifier using Haar-like features and SAMME boosting strategy. The concept of *Integral Image* is used for fast feature evaluation. The proposed technique is a

good substitute for AdaBoost and can work beyond the limitations of AdaBoost in multiclass environment.
The proposed method has a peak detection accuracy of 98.30% and it can process 25 frames per second. The comparison, with various other existing object classification methods, proves the novelty of the method. Experimental results show that the proposed technique works fine even in poor lighting conditions and in the presence of partial occlusion.

## References

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Int. Conf. on Computer vision and Pattern Recognition, vol. 1, pp. 83-87, 2001.

[2] C. M. Sharma, A.K.S. Kushwaha, S. Nigam and A. Khare, "Automatic human activity recognition in video using background modeling and spatio-temporal template matching based technique," in Proc. ACM Int. Conf. on Advances in Computing and Artificial Intelligence, pp.97-101, India, 2011.

[3] W. Hu and T. Tan, " A survey on visual surveillance of object motion and behaviors, " IEEE Trans. on Systems, Man, and Cybernetics, vol. 34, no. 3, pp. 334-352, 2006.

[4] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y.Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L.Wixson, "A system for video surveillance and monitoring," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep., CMU-RI-TR-00-12, 2000.
http://www.ri.cmu.edu/publication_view.html?pub_id=3325

[5] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in Proc. IEEE Workshop on Applications of Computer Vision, pp. 8–14, 1998.

[6] Y. Kuno, T. Watanabe, Y. Shimosakoda, and S. Nakagawa, "Automated detection of human for visual surveillance system," in Proc. Int. Conf. on Pattern Recognition, pp. 865–869, 1996.

[7] R. Cutler and L. S. Davis, "Robust real-time periodic motion detection, analysis, and applications," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22, pp. 781–796, 2000.

[8] A. J. Lipton, " Local application of optic flow to analyze rigid versus non-rigid motion," in Proc. *ICCV* Workshop on Frame-Rate *Vision*, Corfu, Greece, *September* 1999.
http://www.amazon.co.uk/application-non-rigid-Technical-University-Institute/dp/B0006RFKKW

[9] C. Stauffer, "Automatic hierarchical classification using time-base co-occurrences," in Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 335–339, 1999.

[10] M. Sialat, N. Khlifat, F. Bremond, and K. Hamrouni ,"People detection in complex scene using a cascade of boosted classifiers based on Haar-like Features," in Proc. IEEE Int. Symposium on Intelligent Vehicles, pp. 83-87, 2009.

[11] J. Nishimura and T. Kuroda, " Multi-axial haar-like features and compact cascaded classifier for versatile recognition, " IEEE Sensors Journal, vol. 10, no. 11, pp. 1786-1795, 2010.

[12] C. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," in Proc. IEEE Int. Conf. on Computer Vision, pp. 555-562, 1998.

[13] J. Zhu, H. Zou, S. Rosset, and T. Hastie, " Multiclass adaboost, " Int. J. of Statistics and Its Interface, vol. 2, pp. 349-360, 2009.

[14] L. Breiman, "Bagging predictors ," Int. J. of Machine Learning, vol. 24, pp. 123–140, 1996.
http://portal.acm.org/citation.cfm?id=231989

[15] Y. Freund, and R. Schapire, "A decision theoretic generalization of on-line learning and an application to boosting, " J. of Computer and System Sciences, vol. 55, 119-139, 1997.

[16] M. Rodriguez, "Multi-class boosting," 2009.
http://amachinelearningtutorial.googlecode.com/files/boosting_paper.pdf

[17] I. H. Witten, and E. Frank, "Data mining: Practical machine learning tools and techniques," 2005.
http://200.201.9.37/mestrado/docs/WittenFrank.pdf

[18] A. Laika and W. Stechele, "A review of different object recognition methods for the application in driver assistance systems," in Proc. IEEE 8th Int. conf. on Image Analysis for Multimedia Interactive Services, 2007.
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4279117

[19] T. Kato, Y. Ninomiya, and I. Masaki, " Preceding vehicle recognition based on learning from sample images," IEEE Trans. on Intelligent Transportation Systems, vol. 3, no. 4, pp. 252–260, 2002.

[20] Z. Sun, G. Bebis, and R. Miller, " Monocular pre-crash vehicle detection: features and classifiers, " IEEE Trans. on Image Processing, vol. 15, no. 7, pp. 2019-2034, 2006.

[21] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. V. Seelen, "An image processing system for driver assistance," Int. J. of Image and Vision Computing, vol. 18, no. 5, pp. 367–376, April 2000.

[21] A. Opelt, A. Pinz, M. Fussenegger, and P. Auer, "Generic object recognition with boosting," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 28, no. 3, pp. 416–431, 2006.

**Chandra Mani Sharma** received his B.Sc. from Meerut College in 2006, MCA in year 2009 from ITS Ghaziabad and M.Tech. in Computer Science from Devi Ahilya University, Indore in year 2011. He worked as a research fellow on UGC-Sponsored research project *An Intelligent Video Surveillance System for Human Behavior Analysis* at IPCV Labs, University of Allahabad from May, 2010 to June, 2011. He has obtained various awards and scholarships including ITS-meritorious scholarship, IETE scholarship and MHRD-India scholarship. Currently, he is holding the position of assistant professor in IT at ITS Ghaziabad, India.

**Alok Kumar Singh Kushwaha** received his B.Sc. and M.Sc. from University of Allahabd. He has done his M.Tech in Software Engineering from Devi Ahilya University, Indore in 2011. He is working as an assistant professor in Computer Science at GLA University Mathura.

**Rakesh Roshan** received his B.Sc. degree from Ranchi University, MCA from UPTU Lucknow, M.Tech (IT) from Mysore and he is pursuing his Ph.D from MBU, Solan. His area of interest includes artificial intelligence, data bases, and software engineering. He is IBM DB2 and RAD certified professional. Currently he is working as an assistant professor in IT at ITS Ghaziabad, India.

**Dr. Rabins Porwal** earned his B.Sc. degree from CSJMU Kanpur, M.Sc. and Ph.D from DEI Agra. He is senior fellow of IEEE. He got various awards and scholarships during his academic and research career including JRF, SRF, CSIR fellowship, gold medal in graduation, national scholarship from 10$^{th}$ standard up to graduation etc. He is treasurer of Computer Society of India (Ghaziabad section). He is life member of CSI and The Indian Science Congress Association (ISCA). He is a member of ACM and Indian Society for Technical Education (ISTE). He is currently working as an associate professor in IT at ITS Ghaziabad, India.

**Dr. Ashish Khare** received his B.Sc, M.Sc and Ph.D degrees from University of Allahabad. Dr. Khare has more than 50 research papers in international/national journals and conference proceedings. His research area includes image processing, medical imaging, video processing, soft computing, and algorithm analysis. He is currently employed as an assistant professor at Department of Electronics and Communication, University of Allahabad.