

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

# Annotation in Various Images: A Survey

Abhijeet A. Chincholkar<sup>1</sup>, Yash A. Jaiswal<sup>2</sup>, Ahilya A. Pardhi<sup>2</sup>, Rupa S. Khandarkar<sup>2</sup>

Assistant Professor, Dept. of EXTC Engineering, JCOET, Yavatmal, Maharashtra, India<sup>1</sup>

UG Student, Dept. of EXTC Engineering, JCOET, Yavatmal, Maharashtra, India<sup>2</sup>

**ABSTRACT:** In this paper we are depicting various methods in order to annotate an image or a video sequence. In this paper various methods are sequentially addressed in an approach to mark an object from various data inputs like ultrasound images, human samples in video sequence. Stationary and dynamic camera's inputs and multi sequential camera inputs. In the mentioned approaches the object is recognized and method optically separated from its source for a wider analysis. Which certainly provides all detailed and rich information about the object and its source? Also Earth Observation images are been tracked in order of velocity observations. Also by annotating an image or video, the image or video damage is found and is either repaired or eliminated there by again enhancing the original video quality accordingly. In today's widespread of Annotation techniques, Ultimately, this would be a proof of concept of a new framework for obtaining low cost precise annotations of objects. The task of target image annotating is the key component of this paper. It provides the input to the high level image processing such as recognition or the re-identification which basically initialize the research. The techniques listed below are in demands in the recent times, especially in the mass events, human activities, which ultimately makes annotation a critical requirement where most of the pictorial evidences are concerned.

**KEYWORDS:** TSI(Time spatial image detection), Subtraction of background in video surveillance, Gesture algorithm based sensors, Ultrasound images detectors.

### I. INTRODUCTION

A background subtraction technique is presented in this paper that learns the variation of each pixel throughout the video in terms of differences of averages. And also the pixels colors of current frame with previous frame are learnt with average differences. Background subtraction differentiating foreground objects from the stationary parts of the scene is one of the modules of annotation. Valuable low level cue information is obtained by the module for performing high-level tasks of motion analysis, like motion estimation, tracking, etc. In recent research, 3D scene reconstruction and geometric model matching are the basis of perception where a 3D sample point cloud are used to match with the trained features. A semantic perception method is proposed in this work which is based on spatiosemantic features. The defining of this features are natural, symbolic way, such as geometry and spatial relation.

### II. PRESENT THEORY AND PRACTICES

ANUVA CHOWDHURY proposes 'A Background Subtraction Method Using Color Information in the Frame Averaging Process', accurate and reliable detection of moving object from a video sequence is often hard to do in real case. The most common approach is image segmentation which identifies the object from the video sequence that differs significantly from a background model. Among the other segmentation methods background subtraction is the effective and efficient one. A background subtraction method must gratify some challenges, first of all the method must diminish the noise from the sensor that results alteration in pixel color in the probable background from the original in the current video sequence. Secondly the method must automatically fiddle with the changing scene conditions. All these challenges must satisfy in real time with lower computational complexity. In this paper we represent a background subtraction technique that learns the variation of each pixel throughout the video in terms of average and average differences of pixels color of current frame with the previous one. The BGR (Blue, Green and Red) image is first converted to HSV (Hue, Saturation and Value) color space. And then the algorithm is applied to the H plane only as H plane is invariant to illumination changes. Morphological operations are done for removing image noises. In the

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

last stage edges are detected to find out whether there is any false detection or not. Our experimental result shows that this method is very much reliable in changing scene conditions and noise elimination.

EUGEN GANEA proposes 'A System For Image Processing To Automatic Annotation', a system for image automatic annotation considering the information extracted from the image through the segmentation and the boundary detection process. Using domain ontology, we add semantic information to regions from the images which belong of to the domain. The technique proposed in the article uses a flexible hexagonal-grid structure which is over-lap on the image pixels and we extract for each segmented region the dominant color and prominent geometric properties. These characteristics are used in an inference algorithm to determine a decision tree. The instances of the decision tree together with the objects, corresponding to the low level features extracted from the images, are used for image annotation. The graph representation of images is the output of all the phases of system processing: the pre-processing phase, the segmentation step and finally, the automatic annotation. Our technique, which combines the visual feature descriptors with inference methods, has a good time complexity because it associates the semantic labels to regions by using the decision tree model.

MARKUS EICH proposes 'An Application Of Fuzzy DL-Based Semantic Perception To Soil Container Classification' semantic perception and object labeling are key requirements for robots interacting with objects on a higher level. Symbolic annotation of objects allows the usage of planning algorithms for object interaction, for instance in a typical fetch and- carry scenario. In current research, perception is usually based on 3D scene reconstruction and geometric model matching, where trained features are matched with a 3D sample point cloud. In this work we propose a semantic perception method which is based on spatiosemantic features. These features are defined in a natural, symbolic way, such as geometry and spatial relation. In contrast to point-based model matching methods, a spatial ontology is used where objects are rather described how they "look like", similar to how a human would described unknown objects to another person. A fuzzy based reasoning approach matches perceivable features with a spatial ontology of the objects. The approach provides a method which is able to deal with sensor noise and occlusions. Another advantage is that no training phase is needed in order to learn object features. The use-case of the proposed method is the detection of soil sample containers in an outdoor environment which have to be collected by a mobile robot. The approach is verified using real world experiments.

YOGINEE B. BRAMHE(PETHE) proposes 'An Implementation Of Moving Object Detection, Tracking And Counting Objects For Traffic Surveillance System' moving vehicle detection in digital image sequences is one of the key technologies of Intelligent Transportation Systems (ITS). Traffic Surveillance System is being more and important with the enlarging of urban scale and increasing number of vehicles. This Paper presents an intelligent vehicle counting method based on blob analysis in traffic surveillance. The algorithm is composed of moving object segmentation, blob analysis, and tracking. By analyzing the blob of vehicles, the meaningful features are extracted. In addition, the speed of each vehicle and the vehicle flow through a predefined area can be calculated by analyzing blobs of vehicles. The experimental results show that the proposed system can provide useful information for traffic surveillance. We analyze the procedure of video-based traffic congestion system and divide it into graying, binarization, denosing and moving target detection. The system first reads video and converts them into grayscale images. We also put forward a Boundary block detection algorithm with noise reduction to identify the moving objects.

PAUL SWOBODA proposes 'Anatomical Landmark Detection Using Multiple Instance Boosting With Spatial Regularization' a novel multiple instance boosting approach with spatial regularization for detecting anatomical landmark to alleviate the manual annotation burden and to address imprecise annotations. It features three contributions. The first is the introduction of soft max cost function for better handling the practical situation in object detection that most positive bags only contain very few true positives while including the ISR rule and AdaBoost as special examples. The second is to exploit for better detection the spatial context embedded in a medical image, specifically the grid arrangement of the training instances with strong correlation. This is in contrast with conventional methods that treat instances in a bag independently. The third is to encourage a concentrated detection response map so that the final detection result can be derived with more confidence. The latter two contributions are realized using total

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

variation regularization. Experimentally the proposed approach achieves significantly better detection performance than state-of-the-art detection methods in detecting anatomical landmarks with few or even no annotations.

C. CANTON-FERRER proposes ‘Audiovisual Event Detection Towards Scene Understanding’ acoustic events produced in meeting environments may contain useful information for perceptually aware interfaces and multimodal behavior analysis. In this paper, a system to detect and recognize these events from a multimodal perspective is presented combining information from multiple cameras and microphones. First, spectral and temporal features are extracted from a single audio channel and spatial localization is achieved by exploiting cross correlation among microphone arrays. Second, several video cues obtained from multi-person tracking, motion analysis, face recognition, and object detection provide the visual counterpart of the acoustic events to be detected. A multimodal data fusion at score level is carried out using two approaches: weighted mean average and fuzzy integral. Finally, a multimodal database containing a rich variety of acoustic events has been recorded including manual annotations of the data. A set of metrics allow assessing the performance of the presented algorithms. This dataset is made publicly available for research purposes.

KUNLIN CAO proposes ‘Automated Catheter Detection In Volumetric Ultrasound’ ultrasound guided catheter insertion is a common procedure in current clinical practice, but it requires a skilled ultrasound practitioner to correctly acquire the images so that the catheter and tip are well visualized. Automated detection of the catheter location in ultrasound images can help in procedural guidance as well as surveillance of catheter position post placement, especially when the ultrasound imaging is being performed by a non-sonographer. Accurate and fast localization of the catheter is a very challenging task because of the poor observation ability of the catheter in ultrasound images. In this paper, we present a novel algorithm for fast catheter detection in 3D ultrasound images. We begin by generating a catheter likelihood map using a physical model of the catheter. For fast detection, we perform a greedy optimization, where the likelihood map is projected onto a single image plane using a variation on the maximum-intensity-projection approach. The highest likelihood curves on this plane then help us to determine the curved planes on which the catheter may be located. The potential catheter locations are then detected on these curved planes. We finally apply some post-processing to enable robust detection of the catheter. We illustrate the performance of our proposed method through quantitative comparisons with expert annotations as well as qualitative results on 26 images obtained from 7 subjects (of which 15 are with and 11 are without a catheter).

CHENG-LAIING LAI proposes ‘Automatic Path Modeling By Image Processing Techniques’ intelligent video surveillance system, including camera calibration, foreground region detection, moving object detection, moving object tracking and path modeling. This study used the data of the moving trajectory of a moving object as the path modeling data. However, the data may contain incorrect trajectory data, such as wrong foreground region detection data, wrong moving object tracking data, or moving object trajectory not on the normal path, thus resulting in incorrect path. This study first used Background Subtraction to capture moving objects, such as pedestrians or vehicles from the video, and then applied Morphology Operation and Connected Components to eliminate noise and label every individual moving object. Finally, gravity center of each moving object was calculated to obtain the path modeling data. Different from previous path modeling, this study used reward and punishment mechanism to automatically adjust path modeling weight, thereby reducing the impact of inferior trajectory on path, and improving the path model performance with the new path.

ALEXANDER R.T proposes ‘Co-Training Of Context Models For Real-Time Vehicle Detection’ a simple way to reduce the amount of required training data in context-based models of realtime object detection. We demonstrate the feasibility of our approach in a very challenging vehicle detection scenario comprising multiple weather, environment and light conditions such as rain, snow and darkness (night). The investigation is based on a real-time detection system effectively composed of two trainable components: an exhaustive multiscale object detector (“signal-driven detection”), as well as a module for generating object-specific visual attention (“context models”) controlling the signal-driven detection process. Both parts of the system require a significant amount of ground-truth data which need to be generated by human annotation in a time-consuming and costly process. Assuming sufficient training examples for signal-based detection, we demonstrate that a co-training step can eliminate the need for separate ground-truth data to

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

train context models. This is achieved by directly training context models with the results of signal-driven detection. We show that this process is feasible for different qualities of signal-driven detection, and maintains the performance gains from context models. As it is by now widely accepted that signal-driven object detection can be significantly improved by context models, our method allows training strongly improved detection systems without additional labor, and above all, costing.

NAFI UR RASHID proposes 'Detection And Classification Of Vehicles From A Video Using Time-Spatial Image' detection and classification of vehicles are the most challenging tasks of a video-based intelligent transportation system. Traditional detection and classification methods are based on subtraction of estimated still backgrounds from a video to find out the moving objects. In general, these methods are computationally highly expensive, and in many cases show poor detection and classification performance, especially when differences between pixel intensities of vehicles and backgrounds are small. In this paper, we present a novel detection and classification method that employs an analysis of time-spatial image (TSI) obtained from a virtual line on the frames of a video so that the dependencies of pixel intensities of still and moving objects of the video may be reduced. First, the TSI is segmented to count the number of vehicles those cross the virtual line. Then, a feature-based classification scheme is proposed to classify these vehicles. The classification scheme utilizes the shape of the segmented regions of the TSI as well as that of appropriate frames of a video to extract the certain features of the moving objects. Experimental results on a number of real video sequences demonstrate that the proposed method provides higher accuracy in counting and classifying the vehicles as compared to that of the conventional background subtraction-based methods.

MEENAKSHI PANWAR proposes 'Hand Gesture Recognition Based On Shape Parameters', pattern recognition and Gesture recognition is the growing fields of research. Being a significant part in non verbal communication hand gestures are playing vital role in our daily life. Hand Gesture recognition system provides us an innovative, natural, user friendly way of interaction with the computer which is more familiar to the human beings. Gesture Recognition has a wide area of application including human machine interaction, sign language, immersive game technology etc. By keeping in mind the similarities of human hand shape with four fingers and one thumb, this paper aims to present a real time system for hand gesture recognition on the basis of detection of some meaningful shape based features like orientation, centre of mass, status of fingers, thumb in terms of raised or folded fingers of hand and their respective location in image. The approach introduced in this paper is totally depending on the shape parameters of the hand gesture. It does not consider any other mean of hand gesture recognition like skin color, texture because these image based features are extremely variant to different light conditions and other influences. To implement this approach we have utilized a simple web cam which is working on 20 fps with 7 mega pixel intensity. On having the input sequence of images through web cam it uses some pre-processing steps for removal of background noise and employs K-means clustering for segmenting the hand object from rest of the background, so that only segmented significant cluster or hand object is to be processed in order to calculate shape based features. This simple shape based approach to hand gesture recognition can identify around 45 different gestures on the bases of 5 bit binary string resulted as the output of this algorithm. This proposed implemented algorithm has been tested over 450 images and it gives approximate recognition rate of 94%.

SHIVA SUNDARAM proposes 'Hotflashes: Thumbnailing Videos Of Social Gatherings By Detecting Camera Flash Illuminated Frames', Automatic annotation of video clips is desired to efficiently thumbnail user generated content available in the Internet. Automatic techniques typically focus on a selected set of inherent video features (such as scene-cut or shot boundaries) that are deemed to be salient. Along this line, the presence of camera flash light illumination is a feature of interest. This event is usually triggered manually (by a photographer) within a scene and the selected frame(s) they occur in are deemed to be interesting in the recording: for instance, the appearance of a celebrity in a party. In this paper we present a method to detect video frames that contain flashes originating from still cameras in user generated video clips. We focus on designing features for flash illumination detection using various measures of luminance change within video sequences. Using the proposed method, we obtain detection performance of approximately 89% which is 8% absolute improvement over the baseline method that uses only change in average illumination. We also illustrate a case where flashes are automatically detected in video clips of social gatherings that can be used for thumb nailing and browsing.

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

TIMM SCHOENING proposes ‘Investigation Of Hidden Parameters Influencing The Automated Object Detection In Images From The Deep Seafloor Of The HAUSGARTEN Observatory’, Detecting objects in underwater image sequences and video frames automatically, requires the application of selected algorithms in consecutive steps. Most of these algorithms are controlled by a set of parameters, which need to be calibrated for an optimal detection result. Those parameters determine the effectivity and efficiency of an algorithm and their impact is usually well known. There are however further non-algorithmic impact factors (or hidden parameters), which bias the training of a machine learning system as well as the subsequent detection process and thus need to be well understood and taken into account. In the context of mega fauna detection in benthic images, we investigate the effects of some of these parameters on our machine learning based detection system iSIS. The images to be analyzed were taken at the deep-sea, long-term observatory HAUSGARTEN in which five experts labeled seven distinct object classes as an annotation gold standard. We found, that the hidden parameters from imaging as well as the fusion of expert knowledge could partly be compensated and were able to achieve detection performances of 0.67 precision and 0.87 recalls. Despite the efforts to compensate the hidden parameters, the detection performance was still varying across the image transect. This poses the potential occurrence of further hidden parameters not taken into account so far

TOM S. H. LEE proposed ‘Learning Categorical Shape From Captioned Images’, Given a set of captioned images of cluttered scenes containing various objects in different positions and scales, we learn named contour models of object categories without relying on bounding box annotation. We extend a recent language-vision integration framework that finds spatial configurations of image features that co-occur with words in image captions. By substituting appearance features with local contour features, object categories are recognized by a contour model that grows along the object’s boundary. Experiments on ETHZ are presented to show that 1) the extended framework is better able to learn named visual categories whose within class variation is better captured by a shape model than an appearance model; and 2) typical object recognition methods fail when manually annotated bounding boxes are unavailable.

LEONID PISHCHULIN proposes ‘Learning People Detection Models From Few Training Samples’, People detection is an important task for a wide range of applications in computer vision. State-of-the-art methods learn appearance based models requiring tedious collection and annotation of large data corpora. Also, obtaining data sets representing all relevant variations with sufficient accuracy for the intended application domain at hand is often a non-trivial task. Therefore this paper investigates how 3D shape models from computer graphics can be leveraged to ease training data generation. In particular we employ a rendering-based reshaping method in order to generate thousands of synthetic training samples from only a few persons and views. We evaluate our data generation method for two different people detection models. Our experiments on a challenging multi-view dataset indicate that the data from as few as eleven persons suffices to achieve good performance. When we additionally combine our synthetic training samples with real data we even outperform existing state-of-the-art methods.

AUSTIN ABRAMS proposes ‘LOST: Long-Term Observation Of Scenes (With Tracks)’, We introduce the Long-term Observation of Scenes (with Tracks) dataset. This dataset comprises videos taken from streaming outdoor webcams, capturing the same half hour, each day, for over a year. LOST contains rich metadata, including geolocation, day-by-day weather annotation, object detections, and tracking results. We believe that sharing this dataset opens opportunities for computer vision research involving very long-term outdoor surveillance, robust anomaly detection, and scene analysis methods based on trajectories. Efficient analysis of changes in behavior in a scene at very long time scale requires features that summarize large amounts of trajectory data in an economical way. We describe a trajectory clustering algorithm and aggregate statistics about these exemplars through time and show that these statistics exhibit strong correlations with external meta-data, such as weather signals and day of the Week.

PIERRE BLANCHART proposes ‘Mining Large Satellite Image Repositories Using Semi-Supervised Methods’, The increasing number and resolution of earth observation (EO) imaging sensors has had a significant impact on both the acquired image data volume and the information content in images. There is consequently a strong need for highly efficient search tools for EO image databases and for search methods to automatically identify and recognize structures within EO images. In this paper, we present a concept for an earth observation image data mining system mixing an auto-annotation component with a category search engine which combines a generic image class search and an object

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

detection feature. The proposed concept relies thus on three distinct components which are detailed successively: in the first part, we describe the auto-annotation component, in the second part, the generic category search engine and in the third part, the object detection tool. In the concluding part of the paper, we provide an insight into how these three components can be related to each other and used in a complementary way to arrive at a system which combines the advantages of both the auto-annotation systems and the category search engines.

YOUSUN KANG proposes 'Multiband Image Segmentation And Object Recognition For Understanding Road Scenes', a novel method for semantic segmentation and object recognition in a road scene using a hierarchical bag-of-textons method. Current driving-assistance systems rely on multiple vehicle-mounted cameras to perceive the road environment. The proposed method relies on integrated color and near-infrared images and uses the hierarchical bag-of-textons method to recognize the spatial configuration of objects and extract contextual information from the background. The histogram of the hierarchical bag-of-textons is concatenated to textons extracted from a multiscale grid window to automatically learn the spatial context for semantic segmentation. Experimental results show that the proposed method has better segmentation accuracy than the conventional bag-of-textons method. By integrating it with other scene interpretation systems, the proposed system can be used to understand road scenes for vehicle environment perception.

K SUSHEEL KUMAR proposes 'Multiple Cameras Using Real Time Object Tracking For Surveillance And Security System', multiple cameras using real time tracking for surveillance and security system. It is extensively used in the research field of computer vision applications, like that video surveillance, authentication systems, robotics, pre-stage of MPEG4 image compression and user inter faces by gestures. The key components of tracking for surveillance system are extracting the feature, background subtraction and identification of extracted object. Video surveillance, object detection and tracking have drawn a successful increased interest in recent years. An object tracking can be understood as the problem of finding the path (i.e. trajectory) and it can be defined as a procedure to identify the different positions of the object in each frame of a video. Based on the previous work on single detection using single stationary camera, we extend the concept to enable the tracking of multiple object detection under multiple camera and also maintain a security based system by multiple camera to track person in indoor environment, to identify by my proposal system which consist of multiple camera to monitor a person. Present study mainly aims to provide security and detect the moving object in real time video sequences and live video streaming. Based on a robust algorithm for human body detection and tracking in videos created with support of multiple cameras.

### III. OBSERVATIONAL TABLE OF DIFFERENT OBJECT TRACKING TECHNIQUES AND VARIOUS ASPECTS

Sr. No.	Author	Year	Proposed method	Dataset used	Parameter used for evaluation	Results
1.	Anuva Chowdhury, Sang-jin Cho,	2011	A background subtraction Technique that learns the variation of each pixel.	Video sequence from static cameras	Background subtraction; frame averaging process; edge;Color; brightness.	Method is very Much reliable in changingscene conditions and noise elimination
2.	Eugen Ganea, Dumitru Dan Burdescu, Marius Brezovan,	2010	Flexible hexagonal-grid Structure	Segmented video sequence.	Image processing; graph representation; decisionTree; inference algorithm; automatic annotation	The extracted syntactic features give the signaturefor the real objects from image and will be used in theshape recognition process
3.	Markus Eich	2013	A semantic perception method which is based on spatio-semantic features.	Real life images of Soil	Point-based model matching, spatial ontology. Laser	This approachcan be used to detect objects in a cluttered environment

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

				container.	Generated point cloud,	based only on the semantic description of the features
4.	Yoginee B. Bramhe (Pethe , P.S. Kulkarni	2011	Based on blob analysis in traffic surveillance.	Stationary cameras.	Binarization; background subtraction; boundary Detection.	It is detecting the moving object according to its boundary fulfilling the objectives and requirements.
5.	Paul Swoboda, David Liu, S. Kevin Zhou	2013	Spatial Regularization for detecting anatomical landmark	Milboost s Images.	Anatomical landmark detection, multiple instance Learning & boosting, spatial regularization	The proposed spatial regularization helps to incorporate the needed spatial context when dealing with inaccurately labeled data.
6.	C. Canton-Ferrer, T. Butko, C. Segura, X.	2009	Weighted mean average and fuzzy integral	Multiple cameras and microphones.	A set of metrics allow assessing the performance of the presented algorithms.	Multimodality yields A noticeable increase of the detection rate
7.	Kunlin Cao, David Mills, Kedar A. Patwardhan	2013	Generating a catheter likelihood Map using a physical model of the catheter.	Ultrasound images.	Ultrasound, 3D, catheter, detection	Method is able to locate catheter trajectory with satisfactory accuracy and robustness in real time.
8.	CHENG-LAING LAI, KAI-WEI LIN	2010	A Background Subtraction method to capture Moving objects	Sequences from moving cameras.	Path modeling; Foreground detection; Video surveillance	Automatically adjust path Modeling weight reducing the error of path model
9.	Alexander R.T. Gepperth	2012	An exhaustive multiscale object detectors well as a module for Generating object-specific visual attention	Video sequence from traffic areas.	Sliding-window type Object detector	For vehicles, co-training of context models is a feasible option which will allow a huge reduction of cost and human effort when constructing powerful Hybrid detection architectures.
10.	Nafi Ur Rashid, Niluthpol Chowdhury Mithun	2010	Time spatial image (TSI)	Video sequence.	Detection and classification of vehicles, Time spatial Images, Video surveillance.	Method provides higher accuracy in counting and classifying the vehicles as compared to that of the conventional background subtraction-based methods.
11.	Meenakshi Panwar		Algorithms for hand gesture interpretation.	Web cam which is working on	Image processing, hand gesture recognition, Pattern	This simple shape based approach to hand gesture recognition can

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

				20 fps with 7 mega pixel intensity.	recognition, Human computer interaction, K-meansClustering.	identify Around 45 different gestures on the bases of 5 bit binary stringresult.
12.	Shiva Sundaram, Vladan Velisavljevic and Yujie Qin	2011	Flash illumination detection using various measures of luminancechange within video sequences.	Video clips.	Flash detection, feature extraction, contentbasedVideo analysis, video summarization, video thumbnailing, video ranking, user generated content	Method illustrate a case where flashes are Automatically detected in video clips of social gatherings that can be Used for thumbnailing and browsing.
13.	Timm Schoening_, Melanie Bergmanny,	2012	Machine learning based detection system isis, illumination Correction algorithm	Underwater image sequences And video frames	Pre-processing Should prevent any correlation between the f1-rate and the Image area.	Consideration of hidden parameters has shown to beimportant in underwater image analysis.
14.	Tom S. H. Lee, SanjaFidler, Alex Levinshtein, Sven Dickinson	2012	Language-vision integration framework that finds spatialconfigurations of image features that co-occur with words in Image captions.	Captioned images.	Vision integration, Image annotation, Perceptual grouping, Object categorization, Semi-supervised Shape learning	Method is able to handle large variationsin scale and minor changes in viewpoint, deformation And articulation also strong dependence on manual annotation.
15.	Leonid Pishchulin, Arjun Jain ,Christian Wojek, MykhayloAnd riluka,		Rendering-based reshaping method in order to Generate thousands of synthetic training samples from only A few persons and views.	3D shaped models from computer graphics.	State-of-the-art systems , the HOG-detector.	Using recordings from several hundreds of peopleshould allow to reach performance levels that are beyondwhat can be reached with today's manually and tediously Labeled data.
16.	Austin Abrams, Jim Tucek, Joshua Little, Nathan Jacobs, Robert Pless		Robust Anomaly detection, and scene analysis methods based On trajectories.	Videos from outdoor webcam.	Rich metadata, outdoor surveillance.	Method show that the simpleclusteringtechniq ue presented in the previous section providesa way to create summaries of behavioral patterns that varyin different scenes for weather and day of the week.
17.	Pierre Blanchart, Marin Ferecatuand MihaiDatcu	2011	Auto-annotation component with a category search engine which Combines a generic image class search and an object detection feature.	Earth Obseravtio n images.	Satellite imagery; Auto-annotation; Contentbasedimage retrieval (CBIR) systems; Object detection;Active learning; Semi-	Method limits the number of computations When it comes to evaluating the decision function of the Used classifier.



# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

					supervised methods;	
18.	Yousun Kang, Koichiro Yamaguchi, Takashi Naito, and Yoshiki Ninomiya	2011	Hierarchical bag-of-textons method.	Multiple vehicle mounted cameras.	Bag-of-textons, object recognition, road scene, semantic segmentation, vehicle environment perception.	Method has better segmentation accuracy Than the conventional bag-of-textons method, also method can be used to understand roadside perception.
19.	K Susheel Kumar, Shitala Prasad, Pradeep K. Saroj, R.C. Tripathi.	2010	Robust algorithm for human body Detection and tracking in videos created with support of Multiple cameras.	Multi-cameras sequence.	Real time objects tracking; Real time object detection; object identification; surveillance system and background subtraction.	The algorithm is tested for different video data set. The detected object is represented by its centroid and the rectangular shape around The object boundary. This would be helpful in surveillance systems.

## IV. CONCLUSION

We presented a framework for the challenging task of training an object detector from weakly labelled data. Our framework includes a novel initialization routine that is able to provide a good initial annotation of the training set to start iteratively training an object detector. Following such a framework we have tested state-of-the-art pedestrian descriptors. Also segmentation algorithm based on edge detection method, which combines aspects of mathematical morphology, filtration and gradient operations are taken into consideration here. Also the overlapping of images which is disadvantage that occurred is dealt here in order to get riddance from the sharper edges of the images which make it harder to merge.

## REFERENCES

- [1] Anuva Chowdhury, Sang Jin Cho, Vi-Pil Chong, "A Background Subtraction Method using Color Information in the Frame Averaging Process", The 6th International Forum on Strategic Technology, Vol. 2, pp. 1275-1279, 2011.
- [2] Eugen Ganea, Dumitru Dan Burdescu, Marius Brezovan, Liana Stansky, Cosmin Stoica, "A System for Image Processing to Automatic Annotation", Fifth International Multi-conference on Computing in the Global Information Technology, pp. 87-92, 2010.
- [3] Markus Eich, "An Application of Fuzzy DL-Based Semantic Perception to Soil Container Classification", IEEE, ISSN: 2325-0526, pp. 1-8, 2013.
- [4] Yoginee B. Bramhe (Pethe), P.S. Kulkarni, "An implementation of Moving Object Detection, Tracking and Counting Objects for Traffic Surveillance System", International Conference on Computational Intelligence and Communication Systems, ISBN: 978-0-7695-4587-5, pp. 143-148, 2011.
- [5] Paul Swoboda, David Liu, S. Kevin Zhou, "Anatomical landmark detection using multiple instance boosting with spatial regularization", IEEE 10th International Symposium on Biomedical Imaging: From Nano to Macro San Francisco, CA, USA, ISSN: 1945-7928, pp. 218-221, 2013.
- [6] C. Canton-Ferrer, T. Butko, C. Segura, X. Gir'o, C. Nadeu, J. Hernando, J.R. Casas, "Audiovisual Event Detection Towards Scene Understanding", IEEE, ISSN: 2160-7508, pp. 81-88, 2009.
- [7] Kunlin Cao, David Mills, Kedar A. Patwardhan, "Automated catheter detection in volume ultrasound" 2013 IEEE 10th International Symposium on Biomedical Imaging: From Nano to Macro San Francisco, CA, USA, ISSN: 1945-7928, pp. 37-40, 2013.
- [8] Cheng-Laiing Lai, Kai-Wei Lin, "Automatic Path Modeling By Image Processing Techniques", Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, Vol.1, pp. 2589-2594, 2010.
- [9] Alexander R.T. Gepperth, "Co-training of context models for real-time vehicle detection", Intelligent Vehicles Symposium Alcalá de Henares, Spain, Vol. 3 Issue 1, pp. 146-166, 2012.
- [10] Nafi Ur Rashid, Niluthpol Chowdhury Mithun, Bhadhan Roy Joy, S. M. Mahbur Rahman, "Detection And Classification of Vehicles from A Video Using Time-Spatial Image", Vol 13, Issue 3, pp. 1215-1225, 2012.
- [11] Shiva Sundaram, Vladan Velisavljevic and Yujie Qin, "Hotflashes: Thumbnailing Videos Of Social Gatherings By Detecting Camera Flash Illuminated Frames", Vol 2, pp. 1-4, 2009.
- [12] Timm Schoening, Melanie Bergmann, Tim W. Nattkemper, "Investigation of hidden parameters influencing the automated object detection in images from the deep seafloor of the HAUSGARTEN observatory", ISBN 978-1-4673-0829-8, pp. 1-5, 2012.

## International Journal of Innovative Research in Science, Engineering and Technology

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 9, September 2015**

- [13] Tom S. H. Lee, SanjaFidler, Alex Levinshtein, "Sven Dickinson Learning Categorical Shapefrom Captioned Images", Ninth Conference on Computer and Robot Vision, pp. 228-235, 2012.
- [14] Leonid Pishchulin, Arjun Jain, Christian Wojek, Mykhaylo Andriluka, Thorsten Thorm"ahlen,BerntSchiele, "LearningPeopleDetectionModelsfromFewTraining Samples", ISSN 1063-6919, pp. 1473-1480, 2011.
- [15] Austin Abrams, Jim Tucek, Joshua Little, Nathan Jacobs, Robert Pless, "LOST: Longterm Observation of Scenes(with Tracks)", ISBN 978-1-4673-0233-3, pp. 297-304.
- [16] Pierre Blanchart, Marin Ferecatu and MihaiDatu, "Mining large satellite imager positories using semi Supervised Methods", IEEE, pp. 1585-1588, 2011.
- [17] Yousun Kang, Koichiro Yamaguchi, Takashi Naito, and Yoshiki Ninomiya, "Multiband Image Segmentation and Object Recognition for Understanding Road Scenes" IEEE Transactions On Intelligent Transportation Systems, Vol. 12, No. 4, December 2011.
- [18] K Susheel Kumar, Shitala Prasad, Pradeep K. Saroj, R.C. Tripathi, "Multiple cameras using real time object tracking for surveillance and security system", IEEE, ISBN 978-6-7695-4246-1, pp. 213-218, 2010.