A Survey on Real Time Object Detection, Tracking and Recognition in Image Processing

C. Hemalatha
Department of Electronics and
Instrumentation
Bharathiar University
Coimbatore
India

S. Muruganand
Department of Electronics and
Instrumentation
Bharathiar University
Coimbatore
India

R. Maheswaran
Department of Electronics and
Instrumentation
Bharathiar University
Coimbatore
India

ABSTRACT
Object detection, tracking and recognition in real time is a very essential task in computer vision. There are lots of research work have been done in this area. Yet it needs to be accuracy in recognizing object. The most objective of this review is to present an overview of the approaches used and also the challenges involved. In this paper we concentrate on different object detection methods, tracking and recognition methods are discussed.

Keywords
object detection, object tracking, object recognition, survey

1. INTRODUCTION
In computer vision, Real Time object recognition is one among the foremost research area. The object recognition systems are going to include detection and tracking the object is in motion and its position relevant to scene. Lots of research work has been done in this emerging field. It is very useful for the quality control and inspection task in industry. Object recognition is very successful one in controlled environment. But it is terribly tough in uncontrolled atmosphere. Several methods have been proposed to recognize objects in real time. Section two describes the review of real time object detection and tracking. The over review of object recognition is described in section three and conclusion is well delineating in section four.

2. REVIEW OF REAL TIME OBJECT DETECTION AND TRACKING
Moving object detection is a first step for object classification and object recognition. Chih-Hsien Hsia et al [1] proposed Efficient modified directional lifting-based discrete wavelet transform (MDLDWT) for moving object detection. The proposed method detects multiple moving objects very perfectly even though the object shape issues in the low resolution configuration is vague. The merit of MDLDWT is reduced the cast and protect the fine shape information in low resolution.

Victoria Yanulevskaya et al [2] proposed an object-based visual attention theory for the task of salient object detection. They assume proto-object being a unit of attention and argue that notion of an object should be taken into account while assessing object saliency. In this paper two types of object saliency is used and they combined together detects the object in the image. The two objects salient are center surround saliency and integrated saliency. Centre surround saliency detects an object different from its surrounding object how many rare details present in the objects are detected by integrated saliency.

The contribution of Pranab Kumar Dhar et al [3] proposed an efficient moving object detection method is used for video surveillance system using enhanced edge localization mechanism and gradient directional masking. This method detects moving object with better accuracy than existed methods relating to edge based methods. Because it works on current successive frames and used an edge localization method. And another advantage is it does not affected by illumination changes. Pranab Kumar Dhar et al. compared their error rate with Kim & Hwang and Dewan & Chae error rate. Comparing the error rate results, the proposed method had less error.

Table1: Comparison of error rate (in percentage) among the proposed and existing methods

Cases	Kim & Hwang method	Dewan & Chae method	Proposed method
Indoor images	2.74	5.37	1.78
Outdoor images	5.57	8.93	4.85
Foggy images	21.51	34.03	15.52
Sunny images	8.98	19.73	6.54

Nijad Al-Najdawi et al [4] proposed an Automated Real-Time People Tracking System Based on KLT Features Detection. They used Kanade-Lucas-Tomasi (KLT) technique to detect features of both continuous and discontinuous nature of an object. They proposed Kalman filter for object tracking. The main advantage of this paper is low cast automatic object tracking has proposed. Contrast analysis based algorithm is proposed for object detection in real time for night time visual surveillance [5]. The proposed method detect and tracking an object very effectively at night time. In this [6], shape based object detection has been done. They consider object detection is a matching problem between image and model segments. To rectify matching problem they find dominant sets in the correspondence graph. Multiple objects can be detected in an image when the proposed method gets all the dominant sets.

Carlos Cuevas and Narciso García [7], proposed the background modelling algorithm to detect real time moving object. The proposed method combined the background model and foreground model to detect an object from complex image very quality taken by non-completely static cameras. It finds the bandwidth matrices for the kernels which are used in background modelling. This proposed method updating the background model for reducing the misdetections. Bahadir Karasulu and Serdar Korukoglu [8], proposed Moving object detection and tracking by using annealed background subtraction method in videos. Current frame is subtracted from image and it is used to classify the pixel either foreground or background by comparing the difference with the threshold. Simulated annealing (SA) technique is used to rectify the p – median problem. The total weighted distance between demand points (nodes) and the closest facilities to demand points are used to find the p number. SA-based hybrid method is developed for performance optimization of back, which is used to detect and track object(s) in videos.

Mandellos et al [9], is used to detect and tracking vehicles in traffic surveillance using background subtraction. Scatter background information carried in a series of frames, at pixel level using Histogram-based filtering procedure, generating reliable instances of the actual background. Background instance was reconstructed on demand under any traffic conditions. In this [10], moving objects are detected and tracked using Adaptive background method. They used the median filter for removing the noise in the objects taken from the video sequence to get background image. One of the main advantage of this proposed method is it runs quickly and detect and tracking an object very effectively.

Ling Cai et al [11], proposed a stereo vision-based model for multi-object detection and tracking in surveillance. Illumination variation, shadow interference, and object occlusion problems are overcome by using stereo model. They identified the feature points after they projected into 2D ground plane. For grouping the projected points according to their height values and locations on the plane, a kernel-based clustering algorithm is used. In [12], Bangjun Lei and Li-Qun Xu proposed a detection and tracking of objects in wide range of outdoor surveillance and monitoring scenarios in real-time video analysis system. Adaptive background modelling technique is used to extract the foreground regions. A blob analysis is used for object tracking. It gives better result for non-crowded and the static state of scene. Ian Fasel, Bret Fortenberry and Javier Movellan [13], proposed to find objects. Image generation and derive optimal inference algorithms are used. It generates ratio models for object versus background and it is learned using boosting methods. The advantage of this method is very optimal for finding objects. Christian Goerick, Detlev Noll and Martin Werne [14], artificial neural network is used for real time car detection and tracking. It reduced design cost and complexity. Bijan Shoushtarian and Helmut E. Bez, [15], three different dynamic background subtraction algorithms are used. Based on the performance it is defined as Selective Updating using Temporal Averaging, Non- foreground Pixels of the input image based Selective Update and Selective Updating using Temporal Median were only different for background pixels. Among three, Selective Updating using Temporal Median gives the better results than the rest of the two. The advantage of this algorithm is dynamic reference image is created for each input image. And another advantage is it gives better result in uncontrolled indoor and outdoor scene.

Paolo Piccinini et al [16] detected and localizing the duplicate objects in pick and place applications. SIFT keypoint extraction and mean shift clustering is used to object model and the image with real time. Euclidean transform is used to delimiting points onto the current image. Gavrila, D.M. and Philomin, V [17], presented a real time vehicle detection based on shape based algorithm with Distance Transforms is used to capture the different object shapes. Stochastic optimization techniques (i.e. simulated annealing) used to distribute shape in offline. Matching between the above objects are detected. A boosted cascade of simple features based algorithm is used to detect object showed in [18], integral image is used to detect features. Selection of a small number of critical visual features from a larger set and yields extremely efficient classifiers with the help of Adaboost learning algorithm. Complex classifiers arranged in cascade for background regions of the image to be quickly discarded while spending more computation on promising object. It runs at 15 frames per second without resorting to image in real time applications. Anuj Mohan et al [19], proposed an example-based object detectors. To train human body four distinct example based detectors are used. These four detectors are trained for to detect the head, legs, left arm, and right Arm after that these components are present in configuration, second example-based classifier detects the person or non-person. Component based approach and the ACC data classification architecture are used to improve the object detection. Jianxin Wu et al [20], proposed a C4: A Real-time Framework for Object Detection. The speed of the proposed method is 20 frames per second. This is one of the main advantages. For capturing contour cues, comparison taken among the neighboring pixels. CENTRIST visual descriptor is used for contour based object detection.

3. REVIEW FOR OBJECT RECOGNITION IN REAL TIME

Woo-Han Yun, Sung Yang Bang and Daijin Kim, proposed that objects are recognized with the help of relational dependency among the objects represented by the graphical model [21]. It is designed by transition matrix. Cascaded adaboost detector is used for object detection. Using the softmax function, object candidate is estimated by a logistic regression. Christian Wohler [22] proposed pedestrian recognition in real-time vision systems using Autonomous in situ training of classification modules. Train the front and rear views of pedestrians in dark clothes using classifier gives better recognition capabilities. Supervised training algorithms are applied when pedestrians wearing both dark and bright clothes. Pedestrian recognition is done by situ training process. Markus Ulrich, Carsten Steger and Albert Baumgartner proposed modified Generalized Hough Transform (GHT) for object recognition in realtime [23]. Improving the performance of GHT, some modification done in GHT. Comparing the GHT with modified GHT, the computation time and memory are less to modified GHT. Nicoletta Noceti et al [24], presented on-line 3D object recognition in videos using Spatio-temporal constraints. Object modeling (training sequence) and the recognition phase (test sequences) are constrained by local scale-invariant features. Ad hoc matching procedure is used for recognition. It has a view-based approach to 3D object recognition in videos. The proposed method is based on the use of local scale-invariant features and captures the essence of the object appearance selecting those features that are stable in time (i.e.,across view-point changes) and meaningful in space. A compact description of the object, tolerant to scale,

illumination, and view-point changes are produced by these features

Lowe, D.G [25], objects are recognized with help of local scale-invariant features. This is not affected by image scaling, translation, and rotation, and it is partially invariant to illumination changes and affine or 3D projection. These properties of features used for object recognition. Features are detected by filtering approach. It identifies stable points in scale space. The image keys are created and used as input to a nearest neighbor indexing method which identifies candidate object matches. Lowe, D.G. [26] proposed 3D object recognition using Local feature view clustering.

Matching of invariant local image features are used for Object recognition. Combining multiple images of a 3D object into a single model is presented for object recognition. 3D objects are recognized from any viewpoint, imaging conditions provide the non-rigid changes through generalization of models, and improved robustness through the combination of features. S. Belongie et al [27], presented object recognition using shape matching by shape contexts. It finds the similarity between shapes and it is used for object recognition. Correspondences between points on the two shapes and correspondences to estimate an aligning transform are used to measure the similarity.

Mirmehdi, M. et al [28] presented a object recognition using Feedback control strategies. It is used to find instances of a generic class of objects by improving on established single-pass hypothesis generation and verification approaches. Optimal sets of low-level features are used to reduce the number of hypotheses. The feedback control directs the search for missing information using Top-down recognition. Hui Chen and Bir Bhanu, [29] presented a Highly Similar 3D Objects Recognition in Range Images using feature embedding and SVM rank learning techniques. They compared their methods with the Geometric Hashing (GH) and the performance of the proposed gives better result than the GH method.

Table:2 Comparison of the proposed approach with GH in terms of the indexing performance and the search time(in seconds) for the nearest neighbors

Indexing Performance			Time	
ς(Fraction of Dataset)	SVM rank learning	GH	SVM rank learning	GH
0.5%	(53.3%, 61.7%)	(1.65%, 32.7%)	(11.1, 1.06)	(9.6, 0.54)
5%	(82.54%, 81.72%)	(23.59%, 57.68%)		
10%	(87.74%, 87.41%)	(35.61%, 69.93%)		
15%	(91.51%, 90.69%)	(45.99%, 76.49%)		
20%	(92.92%, 91.89%)	(53.77%, 81.55%)		
25%	(94.39%, 93.34%)	(59.67%, 85.88%)		
30%	(94.81%, 94.38%)	(64.39%, 89.57%)		
35%	(95.75%, 95.35%)	(69.33%, 91.42%)		

Dirk Walther et al [30], presented recognition of multiple objects in cluttered scenes based on bottom-up visual attention. Their result is compared with of David Lowe's recognition algorithm with and without attention. It gives better result than David Lowe's method. Manuele Bicego et al [31], proposed appearance-based 3D object recognition using A Hidden Markov Model approach. For obtaining overlapped sub-images object is visited in a raster-scan fashion is used. Wavelet coefficients are employed for sub image. HMM's used to model the sequence of vectors for initialization. Classification is done by nearest neighbor rule and the distance is measured by HMM. This method recognizes an object very accurate in the case of heavily occluded or distorted objects. Kirt Lillywhite et al [32] presented an object recognition using Evolution-Constructed (ECO) features. It creates datasets based only on a set of basic image transforms. There is no need for a human expert to build feature sets or tune their parameters and ability to generate specialized feature sets for different objects and no limitations to certain types of image sources. This is the major advantage of this method. C. Wohler and J.K. Anlauf presented an adaptable time delay approach neural network for real time object recognition to complete image sequences at a time instead of single images [33].Mahmoud Meribout, Mamoru Nakanishi and Takeshi Ogura [34] presented a real-time object recognition using parallel hardware architecture. A new parallel HT algorithm is used for 2D object recognition in unknown shape. CAM as a core processor used to support HT algorithm. Young, S.S et al [35] proposed Object recognition using multilayer Hopfield neural network. It consists of several cascaded single-layer Hopfield networks.

4. CONCLUSION

Review of object detection, tracking and recognition is briefly overviewed in this paper. The motive of this paper is who are interested in this field for doing research work it will definitely help to choosing the related algorithm for their work accordingly.

5. REFERENCES

[1] Chih-Hsien Hsia and Jing-MingGuo., "Efficient modified directional lifting-based discrete wavelet transform for moving object detection", Signal Processing, article in press.

[2] Victoria Yanulevskaya, Jasper Uijlings, Jan-Mark Geusebroek, "Salient object detection: From pixels to segments", Image and Vision Computing Vol. 31, Pg. No. 31–42, 2013.

[3] Pranab Kumar Dhar et al., "An Efficient Real Time Moving Object Detection Method for Video Surveillance System", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 5, No. 3, September, 2012.

[4] Nijad Al-Najdawi et al., "An Automated Real-Time People Tracking System Based on KLT Features Detection",The International Arab Journal of Information Technology, Vol. 9, No. 1, January 2012.

[5] Kaiqi Huanga,et al., "A real-time object detecting and tracking system for outdoor night surveillance", Pattern Recognition Vol.41, Pg. No. 432 – 444, 2008.

[6] Xingwei Yang et al., "Contour-based object detection as dominant set computation", Pattern Recognition Vol. 45 Pg.No.1927–1936, 2012.

[7] Carlos Cuevas and Narciso García , "Improved background modeling for real-time spatio-temporal non-parametric moving object detection strategies", Image and Vision Computing Vol.31, Pg. No. 616–630,2013.

[8] Bahadir Karasulu and Serdar Korukoglu, "Moving object detection and tracking by using annealed background subtraction method in videos: Performance optimization", Expert Systems with Applications Vol. 39, Pg. No. 33–43, 2012

[9] Nicholas A. Mandellos et al, "A background subtraction algorithm for detecting and tracking vehicles", Expert Systems with Applications, Vol.38, Pg. No. 1619–1631, 2011

[10] Ruolin Zhang and Jian Ding, "Object Tracking and Detecting Based on Adaptive Background Subtraction", Procedia Engineering, Vol. 29 Pg. No.1351 – 1355, 2012

[11] Ling Cai et. al., "Multi-object detection and tracking by stereo vision", Pattern Recognition Vol.43, Pg. No. 4028–4041, 2010.

[12] Bangjun Lei and Li-Qun Xu, "Real-time outdoor video surveillance with robust foreground extraction and object tracking via multi-state transition management", Pattern Recognition Letters Vol.27, Pg. No. 1816–1825, 2006

[13] Ian Fasel, Bret Fortenberry and Javier Movellan, "A generative framework for real time object detection and classification'', Computer Vision and Image Understanding Vol. 98, Pg.No.182–210, 2005

[14] Christian Goerick and Detlev Noll, "Artificial neural networks in real-time car detection and tracking applications", Pattern Recognition Letters Vol.17, Pg. No. 335-343, 1996.

[15] Bijan Shoushtarian and Helmut E. Bez, "A practical adaptive approach for dynamic background subtraction using an invariant colour model and object tracking", Pattern Recognition Letters Vol. 26, Pg. No. 5–26, 2005.

[16] Paolo Piccinini et al, "Real-time object detection and localization with SIFT-based clustering", Image and Vision Computing Vol.30, Pg. No. 573–587, 2012

[17] Gavrila, D.M. and Philomin, V, "Real-time object detection for "smart" vehicles", , The Proceedings of the Seventh IEEE International Conference on Computer Vision Vol. 1, Pg. No. 87 – 93,1999.

[18] Viola, P. and Jones, M , "Rapid object detection using a boosted cascade of simple features" ,Computer Vision and Pattern Recognition, Proceedings of the IEEE Computer Society Conference, Vol.1,Pg. No. I-511 - I-518, 2001

[19] Anuj Mohan and Tomaso Poggio, "Example-Based Object Detection in Images by Components", IEEE Transactions on pattern analysis and machine intelligence, vol. 23, no. 4, april 2001

[20] Jianxin Wu et al, "C4: A Real-time Object Detection Framework", IEEE transactions on image processing, June 15, 2013

[21] Woo-Han Yun, Sung Yang Bang and Daijin Kim, "Real-time object recognition using relational dependency based on graphical model", Pattern Recognition Vol. 41, Pg .No. 742 – 753, 2008.

[22] Christian Wohler, "Autonomous in situ training of classification modules in real-time vision systems and its application to pedestrian recognition", Pattern Recognition Letters Vol.23, Pg. No. 1263–1270, 2002

[23] Markus Ulrich, Carsten Steger and Albert Baumgartner, "Real-time object recognition using a modified generalized Hough transform", Pattern Recognition Vol.36, Pg. No. 2557 – 2570, 2003.

[24] Nicoletta Noceti et al, "Spatio-temporal constraints for on-line 3D object recognition in videos", Computer Vision and Image Understanding Vol.113 Pg. No. 1198–1209, 2009.

[25] Lowe, D.G, "Object recognition from local scale-invariant features", Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 2, Pg. No. 1150 – 1157, 1999.

[26] Lowe, D.G, " Local feature view clustering for 3D object recognition", Computer Vision and Pattern Recognition Proceedings of the 2001 IEEE Conference on Computer Society , Vol. 1, Pg. No. I-682 - I-688, 2001

[27] S. Belongieet al, "Shape Matching and Object Recognition Using Shape Contexts", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 Issue 4, Pg. No. 509-522, April 2002

[28] Mirmehdi, M. et al, "Feedback control strategies for object recognition", IEEE Transactions on Image Processing, Vol. 8, Issue: 8 , Pg. No. 1084 – 1101

[29] Hui Chen and Bir Bhanu, "Efficient Recognition of Highly Similar 3D Objects in Range Images", IEEE Transactions on Pattern analysis and machine intelligence, vol. 31, no. 1, Pg. No. 172-179, January 2009.

[30] Dirk Walther et al, "Selective visual attention enables learning and recognition of multiple objects in cluttered scenes", Computer Vision and Image Understanding Vol. 100, Pg. No.41–63, 2005.

[31] Manuele Bicego et al, "A Hidden Markov Model approach for appearance-based 3D object recognition", Pattern Recognition Letters Vol.26, Pg. No. 2588–2599, 2005.

[32] Kirt Lillywhite et al, "A feature construction method for general object recognition", Pattern Recognition Vol.46, Pg. No. 3300–3314, 2013.

[33] C. Wohler and J.K. Anlauf , " real time object recognition on image sequences with adaptable time delay neural network algorithm- applications for autonomous vehicles", image and computer vision, Vol.19, Pg. No. 593-618, 2001

[34] Mahmoud Meribout, Mamoru Nakanishi and Takeshi Ogura, "A parallel algorithm for real-time object recognition", Pattern Recognition, vol. 35, Pg. No. 1917–1931, 2002

[35] Young, S.S et al, "Object recognition using multilayer Hopfield neural network", IEEE Transactions on Image Processing Vol. 6 , Issue. 3 , Pg. No. 357 - 372 ,1997

AUTHOR'S INFORMATION
[1]Research Scholar, Department of Electronics and Instrumentation, Bharathiar University, Coimbatore,Tamil Nadu. India.
[2]Associate Professor Department of Electronics and Instrumentation, Bharathiar University, Coimbatore,Tamil Nadu. India.
[3]Research Scholar, Department of Electronics and Instrumentation, Bharathiar University, Coimbatore,Tamil Nadu. India.

C.Hemalatha received M.Sc and M.Phil degree in Electronics and Instrumentation from Bharathiar University, Coimbatore, Tamil Nadu, and India in 2010 and 2011 respectively. She is pursuing Ph. D (Electronics and Instrumentation) in Bharathiar University. Her research interests include Digital Image Processing, object recognition, Neural Network and Signal Processing.

S.Muruganand received his M.Sc degree in Physics from Madras University, Chennai, Tamil Nadu, India, and the Ph.D degree from Bharathiar University in 2002. He is working as an Assistant Professor in the Department of Electronics and Instrumentation, Bharathiar University, Coimbatore, India. His area of interest is Embedded Systems, Sensors, Physics and Digital Signal Processing

R. Maheswaran received M.Sc degree in Electronics and Instrumentation from Bharathiar University, Coimbatore, Tamil Nadu, and India in 2010. He is pursuing Ph.D (Electronics) in Bharathiar University. His research interests include Embedded
System, LabVIEW, Digital Image Processing ,Wireless Sensor Network, object recognition.