Journal of Applied Information Science
Volume 5 Issue 2 2017
ISSN.: 2321-6115

Studies on Various Type of Human Detection Algorithms for Multiple and Occluded Persons in Static Images

M. Shanmugasundaram[1], N. Shanmuga Vadivu[2]
[1]Dept. of CS, College of Computer Science & Information Systems, JAZAN University, Kingdom of Saudi Arabia.
[2]Dept of ECE, RVS College of Engineering and Technology, Coimbatore, Tamil Nadu, India. Email: shanmugavadivun@gmail.com

Abstract: Detecting and tracking human in still or video images provides a promising technology development and solution to many real world problems. Moreover, detecting human may be the first step to put forward the next logical steps for many applications. But, it is a challenging task due pose, dresses, color and occlusion. This paper proposes a study of human detection in static images in different view. In the literature, numerous works had been proposed to detect a single human in an image. So, the survey has been conducted for detection of multiple humans without occlusion, detection of multiple human with occlusion and human detection in fused image. Due to the difficulties found during the process of human detection such as occlusion and shadow, people in group, main focus has been given to multiple-human detection.

Keywords: Human detection, Pose, Occlusion, fusion, Machine learning, Object detection, Feature extraction

I. Introduction

The process of object detection and tracking in static and video images leads to lot of research interest in certain areas such as surveillance, medical and security. Detecting human object provides a lot of exposure in many fields. It is a next logical step after the successful development of face detection algorithms. It plays a vital role in providing solutions to many research requirements and is very useful in pedestrian detection and surveillance. Over a period of decades, algorithm developed for human detection has been getting more popular and more focus has been given to the enhancement of existing work in the literature [24][33][36]. Initially, interest was given to single human and it has been exponentially increasing to multiple human objects.

Even though most of the research is going towards detecting human, the researchers feel more difficulties to detect and recognize human object due to its wide variability in appearance. Moreover, some other criteria adds more challenges like clothing, different articulations, textures and styles, variations in illumination and clutter backgrounds [4][8]. In the literature, algorithms are developed to detect a single human [2], multiple human [35] and occluded human. Recently detecting the human object from fused image gets more popular and number of applications uses different fusion techniques [36][39][40].

Technically, algorithms for human detection are organized into two parts: 1. Feature Extraction and 2. Object Classification. Initially, the process of extracting features includes some preprocessing steps such as noise removal and segmentation. Classification mainly involves identifying the type of objects. Machine learning is one of the widely used concepts to classify the features. By using this, a detector used to recognize only human object is constructeld from a large number of training examples. Scholars have mainly been using two such a machine learning algorithms which are SVM [5][1][2] and Adaboost [11][8]. Object detection can also be done using edge detectors such as Sobel, canny etc.

Section-II gives an overview description of the proposed techniques. Analysis of multiple human detection is shown in section-III. Section-IV provides a detailed study of multiple humans with occluded condition is discussed. Finally section-V deals with detections in fused image.

II. Classification Overview

The survey conducted to this issues is classified into single human detection, multiple, occluded and fusion based human detection. Sufficient work had been carried out to detect single human object successfully. So, this paper does not focus much on this. Section-III gives the various techniques used for detecting multiple human in a single image. Several humans appeared in a single image with occluded situation is given in section-IV. Finally, recovering human objects in a fused image which is combined two images of the same scene captured by two different devices is presented in section-V. Techniques given in this paper is organized as shown in Fig. 1.

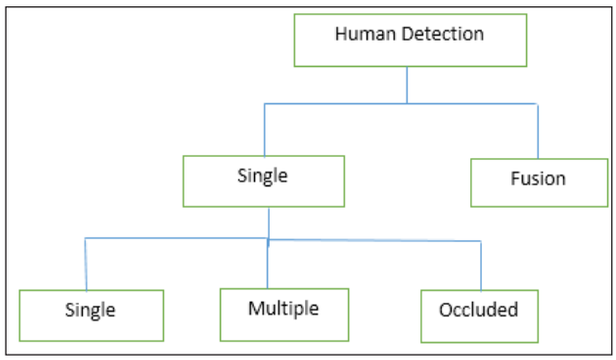

Fig. 1: Different Human Detections

III. Multiple Human Detection

Many scholars have been working towards detecting multiple humans in scenarios includes crowd in the ground, group of pedestrians, customers in shopping mall etc., Even though, it is hard task due to pose and cloths, it's results lead to a lot of benefits to the real world applications.

A. Region of Interest

A two step algorithm to detect several human in crowd was developed by Smita, *et al*. [5]. In the first step, a technique called ROI (Region of Interest) was applied to get features. Next, SVM (Support Vector Machines) was used to classify the human object from other objects. In the experiment, the algorithm achieved 90% of the success in terms of detecting multiple human in the static images. A method has been applied using HOG (Histogram of Oriented Gradients) by Qing, *et al*. [2]. In the previous approach, HOG has been applied for the entire appearance of the human. But, in this method, HOG has been applied for detecting different part of the human body includes head and shoulder based on depth map. For this, Kindest camera was used to capture the images which produces depth map information of images unlike ordinary camera which produces normal two dimensional image information. Then both depth map information and HOG features are combined to obtain the necessary features. For the purpose of classification, SVM was used. Except few false alarm rates, they have achieved almost 91% of the detection rate. Sanofer, *et al.* [16] developed a pedestrian detection system that uses a part-based approach. The methodology includes two stages to classify the features. Candidates are detected using a Haar cascade classifier and then part verification is done by using the cascade object detector. Part based is done using upper body and lower body. Finally, they are combined into one and are fed into SVM classifier.

B. Silhouette

Many researchers calculate gradients of the image to acquire the silhouette of the objects. Silhouettes are useful to detect, identify and track many objects. Dalal, *et al*. [1] designed a new method for recognizing objects based on linear SVM. This method has been proposed to detect human objects with the aid of gradients. Then, the parameter gamma can be normalized as color from the gradients. After that, the gradients are subjected to weighted vote into spatial and orientation cells. Finally, n number of overlapping spatial blocks is considered to normalize the contrast of the target object. HOG's are collected over detection window. After training with machine learning algorithm the software does person/ non- person classification. Alexandre, *et al*. [25] proposes a multiple human detection algorithm. It uses a Conditional Random Field (CRF) which is modelled using a two-hypothesis framework: a pair of detections corresponds either to the same person or not. In this approach, unsupervised learning algorithm was used for classification and sliding window detection was used for detection of humans.

C. Background Subtraction

Background subtraction is another method to separate the required objects from the unnecessary details. Based on this technique, the image signal is divided into foreground and background individual signals. By subtracting the background signal by image, the objects with their features are highlighted. To detect pedestrians, Codebook Background Subtraction (CBBS) method was proposed by Yuan, *et al*. [6] to develop pedestrian samples. They have employed Histogram of Oriented Gradients for feature extraction from both positive and negative training datasets. After that they used SVM for the purpose of classification. They have achieved almost 80% of the result in pedestrian walking on the road.

D. Motion

JongSeok *et al*. [23] propose a simple and efficient method to automatically detect multiple humans using motion information. The system developed by them takes image from both camera system and stationary system. In case of a mobile system, ROI can be found using the frame difference through the ego-motion compensation from two consecutive image frames. In case of still images, Haar features can be used to train the detector. Csaba, *et al*. [24] presents a Bayesian detection framework using shape and motion cues to obtain a maximum a posteriori (MAP) solution for human detection. It has shape templates using contour integration based on images. It also uses a non-parametric approach which generates informative object parts and judges the presence of humans.

E. Mixing Techniques

Mixing more than two techniques offer more benefits. In this way, the foremost popular HOG can be combined with some other failure techniques which provides exponent success rate in detecting the objects. Qiang, *et al*. [7] came up with a novel

approach of integrating the Cascade-of-Rejecters with the HOG features to achieve a fast and accurate human detection system. However the fixed size blocks advocated by Dalal & Triggs [1] is not informative enough to allow fast rejections in the early stages of the cascade. So they have used variable sized blocks for fast detection. Continually, Viola, *et al*. [8] took a different approach of integrating intensity information with motion information for pedestrian detection. Their proposed system works directly with images extracting short terms patterns of motion, as well as appearance information, to detect all instances of potential objects. Motion information can be extracted from pairs or sequence of images in various ways, including optical flow and block motion estimation. They encouraged AdaBoost as a classifier. Swati *et al*. [22] presents a multi-resolution grey-scale invariant approach to detect multiple human. The multi resolution is important for objects of different size and grey-scale invariance is important due to uneven illumination and within class variability. The proposed method is based on integration of central moments upon multi resolution grey-scale invariant local binary patterns operator. LBP is invariant against different resolutions of space scale and monotonic change in grey-scale. It has high computational accuracy due to use of moment operator.

Table I: Multiple Human Detection

Authors	Aim	Comparisons	Methodology	Quantitative Analysis
Smita et. al. (2014)	Fast human detection in crowd of people.	Component-based parts detectors.	Single detection window classifier.	Plot a graph of FPR vs TPR
Qing Tian et. al. (2013)	Pedestrian detection under complex backgrounds.	Sho Ikemura's algorithm, Traditional single HOG.	Kinect depth map, HOG features of head and shoulder features	DET vs FPPW.
Dalal et. al., (2005)	visual object recognition and human detection	Haar wavelets descriptors,	HOG, feature vector and SVM.	DET curves vs FPPW.
Yuan et. al., (2011)	pedestrian detection	Viola's Haar-lfeatures, SIFT descriptor , Shapelet features.	HOG- feature extraction. CBBS- a novel feature with HOG.	FPR vs TPR
Qiang et. al., (2006)	To develop a fast and accurate human detection system.	Papgeorgiou & Poggio Haar-based, Viola's Haar-like wavelets.	HOG, AdaBoost classifier	Cascade level, reject rate.
Viola P et. al., (2003)	A pedestrian detection system with motion information.	Viola Jones Haar-like wavelets	AdaBoost classifier	ROC, FPR and Detection Rate.
Sanofer et. al., (2014)	pedestrian detection system.	Viola's Haar like features	Haar cascade classifier cascade object detector	false positives per frame.

Table II: Occluded Multiple Human Detection

Author	Aim	Comparisons	Methodology	Quantitative Analysis
Bo Wu et. al., (2005)	Detecting inter-occluded humans	Papageorgiou's method, Felzenszwalb's shape models, Viola's object detection	Silhouette-based features, part detectors	ROC curves
Sitapa et. al., (2013)	multiple object detections.	PLS Sliding window method	Shape Covering, Bounding Box Filtering, Binary Optimization	Miss rate vs FPPI
Yu-Ting et. al., (2008)	To employ both intensity and gradient based eatures for occluded human detection	boosted cascade, Bayesian model	Intensity-based features, SIFT based HOG descriptors, SVM	FPPW vs Miss Rate
Guang Shu, Afshin Dehghan, Omar Oreifej, Emily Hand, Mubarak Shah (2012)	Proposing part-based tracking-by-detection framework	Trajectories, max-weight independent set algorithm	detection parts, SVM classifiers	FPPW vs Miss Rate

Hofmann, M, Kiechle, M, Rigoll, G (2011)	To detect occluded persons in multi-view camera networks	edge templates, Haar like features wavelets, Sift-like orientation features	HOG with kernel density	False Positive vs Miss Rate
Ivan et. al., (2014)	To propose people detector	spatio-temporal alignment sliding window	HOG with linear SVM	False Positive vs Miss Rate.
Xiaoyu et. al.,	To propose a partial occlusion detection	Bayesian model, LBP operator	Global detectors, Part detectors, SVM for classification	False Positive vs Miss Rate
William et. al.,	For occluded human detection	HOG, Wavelets features, PCA-SIFT and Shape contexts	HOG with PLS (Partial Least Squares)	FPPW vs Miss Rate

IV. Occluded Human Detection

After the success of detecting multiple human, researchers are concentrated towards detecting a group of people in a single image. Here, the human objects are overlapped one after another. So it is not so easy to find out most of the significant features directly. Most of the required features are hidden by another human. Hiding the body parts by one human to another body is called occlusion.

A. Part based Detection

Bo Wu, *et al*. [3] focused the problem of locating multiple occluded humans from static images. Their mechanism contains two-fold. For the first level, detectors have been used for learning from a novel set of silhouette oriented features that is called 'Edgelet'. It finds features for individual parts like head, torso, legs and hands and makes small connected chains of edges. These features are more suitable for human detection as they are relatively invariant to clothing differences. In the second level, they combine the results of various part detectors to form a joint likelihood for human. Guang, *et al*. [32] addressed a robust part-based pedestrian tracking-by-detection framework. With the use of part based detectors, detection in occlusion scenario has been solved. They used set of detection parts for detecting the parts of the human body like hands, head, legs and body. Finally, all the detected parts are composed and are fed in to the classifier. Then person specific SVM classifiers are trained using for detections.

B. HOG

Hofmann, *et al*. [31] presents a novel method to detect multiple partially occluded persons in multi-view camera networks. They have used multiple cameras to get the multi view with different angles of the scene. HOG is employed with kernel density estimation for information fusion. Xiaoyu, *et al.* [37] proposed a partial occlusion method by combining Histogram of Oriented features with Local Binary Pattern features. Here, detection in the worst condition is performed with the aid of scanning the whole body and part of the body. In the sense, the authors designed and proposed two detectors includes part and global. Global detectors perform whole scanning and Part detectors for local regions. And they have made use of SVM for classification. William, *et al.* [20] proposed a novel approach for occluded human detection by combining both Histogram of Oriented Features and Partial Least Square Analysis. The features extracted from HOG are very high dimensional so it computationally intensive. So they proposed PLS for dimensionality reduction. Ikemura *et al* [30] proposed a method for detecting humans by Relational Depth Similarity Features (RDSF). It is based on depth information obtained from a Time of flight (TOF) camera. In this method, the features obtained from a similarity of depth histograms are calculated that represents the relationship between two local regions. It added more advantage in terms of increasing speed in the process of detection by using raster scanning in a 3D space. Yu-Ting, *et al*. [10] proposed an approach to detect humans in an image using novel cascaded structure. They developed an object detection framework that is both efficient and accurate. Their approach consists two-fold. First, both intensity-based and gradient-based features are considered to develop a better model of pedestrian class. Second, a new cascaded structure is presented to exploit the stage-wise (meta stage) classification information and the interstate cross-reference information so that the detection accuracy and efficiency can be increased.

C. Miscellaneous Techniques

An interesting technique was newly invented to predict humans occluded each other above 50%. It's name is called double-person detector which uses bounding boxes of two people. Siyu *et al* [31] proposed this technique to detect and track occluded people. Later on, a technique called joint person detector was developed following the above technique. It can be trained to detect single as well as occluded people in crowd. Piotr *et al*. [28] proposes a new approach for pedestrian detection system. Caltech pedestrian database which is very unique compared to other existing datasets was utilized. It involves a couple of magnitude for training. In this method, per-image concept was used instead of per-window features. The input images fed in to the system can be returned as a bounding box and a confidence score was evaluated for every human detection operation. Jian *et al* [29] proposes a fast method to detect humans from stationary surveillance videos. It is developed on a cascade of

LogitBoost classifiers which uses covariance matrices as object descriptors. Their proposed methodology proved that human shape information contained in foreground observations can be dramatically improved when using jointly with appearance cues. In order to reduce computational cost, covariance matrices of feature subsets were used rather than the full set in boosting. Sitapa, *et al*. [4] presents a quadratic unconstrained binary optimization (QUBO) framework for multiple object detections with spatial overlaps. QUBO mainly deals with the problem to find a binary vector that maximizes the quadratic objective function. Their framework consists of two steps. Detection confidence score map can be generated by applying an existing pedestrian detector in the first phase. Then, the method samples a large but finite set of plausible candidates. A unary confidence score for each candidate and a pair wise score for overlapped candidates are computed in order to represent the quality of that proposed detection. Grouping unary and pair wise scores into a cost matrix to form the objective function for a QUBO problem was carried out in the second phase.

V. Human Detection in Fused Image

Military applications developed to detect human objects in terms of identifying terrorist, enemies from neighbor country in order to avoid penetrating them in country, face difficulties due to lot of reasons. Country like India contains small to very big hills and continuous forest belts in its border. Chance is more to penetrate the enemies. It is a big challenge to the border security force or army battalion to monitor them. But, it is possible to keep monitoring and based on that easy decision making technically. Even though, the visual camera is unable to capture humans hidden inside the forest, IR is able to capture the humans. Because of its nature of temperature based imaging, objects in the scene having high temperature can be highlighted in the image. Usually, human and animals produce more temperature than plants and trees. Based on this idea, humans hidden by the sources of forest can be recovered and found by a technique called fusion.

Emmanuel, *et al*. [29] proposed a methodology to track the pedestrians using thermal imaging. For that, visual and IR devices were used to capture the images. Then, the details taken from both are merged by pixel by pixel to provide a complement single fused image. In the experiment, medium and high level fusion were also carried out. After fusing, a technique called background subtraction normally used for segmentation was applied to the consecutive fused images to identify the changes of human objects. This is very important to track the objects. Another method was presented by Grassi, *et al*. [30] for detecting pedestrians and classifying them according to their movement patterns. For getting fused image, lidar scanner was used along with IR camera to get images for their experiment. Classifying different objects among the group of objects are done through extracting invariant features from raw data. A good solution was given with the help of multi-view image fusion to overcome the difficulties of occlusion situation. Hofmann, *et al*. [31] presents a novel method to detect multiple partially occluded persons using multi-view camera networks. Number of cameras with different angles was utilized to capture the various view of the scene. HOG is employed with kernel density estimation for segmenting and detecting the humans in fused image. Instead of merging the images directly, the details of depth and vision can be taken to fuse. Bohui, *et al*. [33] presented a novel approach to detect human based on depth and vision information fusion. Depth feature can be unique as it is powerful against illumination changes, discriminative to background clutter and stable to human pose variation.

There is a serious issue in detecting the human in visual image. One strange issue is shadow of human which imitates a real human and make the algorithm to confuse. So the net result is false-alarm. To avoid the, Thomas, *et al.* [34] came up with different definition called shape and heat flow-based technique. In addition to fuse, detection and classification of humans in unrestricted poses can also be performed. In their experiment, vision and IR fusion was used to detect and SVM was used to classify the humans from other objects. Ivan, *et al*. [35] proposed a human detection method using laser and monocular devices. Using a pixel-level fusion algorithm, the two different information of a same scene has been merged to the further steps to be carried out. A high powerful detector to detect people object was utilized which gives better result in some uncertain conditions like false positive and detection and occlusion. Hideya, *et al*. [36] offered a person detection algorithm in visual and IR fused image. The algorithm is built by two parts to detect the human. The various parts of the human body have been identified by parts detectors first and second, Gravity Center movement pattern method was used detect the parts. In body parts method, Multi-slit method for head region extraction is especially used.

Recently, gait recognition is being more popular and it gives the facial expression of human being such as happiness, sadness, laugh, angry, etc. Now a day's researchers have been trying to acquire the feeling of human being from the image. Ju Han, *et al* [38] proposed a gait system in which silhouette of the object is extracted from colour and thermal image sequences by automatic image registration to detect humans. Obviously, the initial step of most of the gait recognition approaches is human silhouette extraction. Next, a background subtraction method is applied to both colour and thermal images to extract preliminary human body silhouettes from the background. After segmentation, human genetic algorithm is used for human detection. Bir Bhanu *et al.* [39] proposed an approach for detecting humans using multi-modal measurements. The approach is based on Time Delay Neural Network to fuse the audio and video data at the feature level for detecting the walker with multiple persons in scene. Background subtraction is implemented to detect changes. For audio features spectrogram from registered audio data was computed. Then TDNN is subjected to back propagation for training to detect human.

TABLE III: HUMAN DETECTION IN FUSED IMAGE

Author	Aim	Comparisons	Methodology	Quantitative Analysis
Emmanuel et. al., (2006)	Pedestrian tracking using thermal imaging	-	medium and high level fusion, background subtraction	False & True Positives
Ana Pérez et. al., (2010)	Detecting pedestrians & classifying them	Haar like wavelets, PCA features, Invariant features	Fusion of LIDAR scanner and infrared, invariant features	False & True Positive Rates
Bohui et al., (2013)	human detection using depth & vision fusion	HOG, LBP, Edge Orientation Histogram, Silhoutte Sparse Representation	Depth feature, background subtraction	FFPW vs Miss Rate
E. Thomas et al., (2009)	Human detection system using image fusion	-	Visual & IR pixel fusion, SVM for classification	FFPW vs Miss Rate
Thi Thi Zin et al., (2011)	Person detection in visual & IR fusion	-	Body parts detection and Gravity Center movement pattern methods	False Positive vs Miss Rate

VI. CONCLUSION

Generally, the survey has been conducted to get to know the techniques and algorithms used to detect the human objects in a static image. In this paper, detecting the human is categorized in to detecting single, multiple and occluded human in a single and fused images. More work was done over a single human detection and the successful rate was 100%. Some of the techniques like single and several part- based detectors get almost 80% of the success in occluded situation. Still, there is false alarm rates found in these techniques due to shadow, tree and post lamp which imitate similar to human being. Finally it is concluded that fusion technique give better performance. Because, the shadow is not appeared in IR image. Other objects looks like human object cannot be highlighted in IR image.

REFERENCES

[1] N. Dalal, and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* vol. 1, pp. 886-893, June 2005.

[2] Q. Tian,, Zhou B, W.-H. Zhao, Y. Wei, and W.-W. Fei, "Human detection using HOG features of head and shoulder based on depth map," *JSW,* vol. 8, no. 9, pp. 2223-2230, 2013

[3] B. Wu, and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors," *IEEE International Conference on Computer Vision* (*ICCV)*, vol. 1, pp. 90-97, 2005.

[4] S. Rujikietgumjorn, and R. T. Collins, "Optimized pedestrian detection for multiple and occluded people," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 3690-3697, 2013

[5] S. Kakade, N. J. Uke, and N. Jain, "Real time human detection using hog feature as human descriptor and variable size sliding window," *Advances in Computer Science and Information Technology (ACSIT)*, vol. 1, no. 1, pp. 27-29, Oct. 2014.

[6] Y. Xu, L. Xu, D. Li, and Y. Wu, "Pedestrian detection using background subtraction assisted support vector machine," *Intelligent Systems Design and Applications (ISDA)*, 2011 11th International Conference, Cordoba, pp. 837-842, Nov. 2011.

[7] Q. Zhu, M. C. Yeh, K. T. Cheng, and S. Avidan, "Fast human detection uses a cascade of histograms of oriented gradients," In *CVPR '06: Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1491-1498. Washington, DC, USA: IEEE Computer Society, 2006.

[8] P. Viola, M. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," In *Proceedings of IEEE International Conference on Computer Vision* (*ICCV)*, vol. 2, pp. 734-741, 2003.

[9] C. Papageorgiou, and T. Poggio, "A trainable system for object detection," *Int. J. Comput. Vis*., vol. 38, no. 1, pp. 15-33, Jun. 2000.

[10] Y.-T. Chen, and C.-S. Chen, "Fast human detection using a novel boosted cascading structure with meta stages," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1452-1464.

[11] S. L. Phung, and A. Bouzerdoum, "Detecting people in images: An edge density approach," *IEEE International*

Conference on Acoustics, Speech and Signal Processing, 2007.

[12] K. M. Hofmann, and G. Rigoll, "Late fusion for person detection in camera networks," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* pp. 41-46, 2011.

[13] G. Shu, A. Dehghan, O. Oreifej, E. Hand and M. Shah, "Part-based Multiple-person tracking with partial occlusion handling," In *Conference on Computer Vision and Pattern Recognition* (*CVPR*), 2012.

[14] I. Huerta, G. Ferrer, F. Herrero, A. Prati, & A. Sanfeliu, "Multimodal feedback fusion of laser, image and temporal information," *Proceedings of the International Conference on Distributed Smart Cameras* (*ICDSC*), 2014.

[15] O. Tuzel, FPorikli F and Meer p, "Human detection via classification on riemannian manifolds", in *Conference on Computer Vision and Pattern Recognition* (*CVPR*), 2007.

[16] S. Munder, and D. Gavrila, "An experimental study on pedestrian classification," *PAMI,* vol. 28, no. 11, pp. 1863-1868, 2006

[17] C. C. Chang, and C. J. Lin, "LIBSVM: A library for support vector machines", 2001. Software Available at www.csie.ntu.edu.tw/ cjlin/libsvm.

[18] Z. Lin, and L. S. Davis, "A pose-invariant descriptor for human detection and segmentation," In *ECCV,* 2008.

[19] J. S. Lim, "Two-dimensional signal and image processing," *Communications and Radar,* p. 710, Englewood Cliffs, NJ, Prentice Hall, 1990.

[20] S. Nigam, and A. Khare "Multiresolution approach for multiple human detection using moments and local binary patterns," *International Conference on Computational Vision and Robotics (ICCVR)*, 2014.

[21] JongSeok Lim and WookHyun Kim, "Detection of Multiple Humans Using Motion Information and Adaboost Algorithm based on Harr-like Features", *International Journal of Hybrid Information Technology (IJHIT),* vol. 5, no. 2, April 2012.

[22] H. Bischof, and C. Beleznai, "Fast human detection in crowded scenes by contour integration and local shape estimation," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2246-2253, June 2009.

[23] A. Heili, C. Chen, and J.-M. Odobez, "Detection-based multi-human tracking using a CRF model," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1673-1680, Nov. 2011.

[24] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP Human Detector with Partial Occlusion Handling," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 32-39, Oct. 2009.

[25] W. R. Schwartz, A. Kembhavi, D. Harwood and L. S. Davis, "Human detection using partial least squares analysis," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 24-31, Oct. 2009.

[26] P. Doll´ar, C. Wojek, B. Schiele and P. Peron, "Pedestrian detection: A Benchmark," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 304-311, June 2009.

[27] J. Yaoa, and J.-M. Odobez, "Fast Human Detection from Joint Appearance and Foreground Feature Subset Covariances," *Computer Vision and Image Understanding (CVIU),* vol. 115, no. 10, pp. 1414-1426, Oct. 2011.

[28] S. Ikemura, and H. Fujiyoshi, "Real-Time Human Detection using Relational Depth Similarity Features," *Asian Conference on Computer Vision (ACCV)*, 2010.

[29] S. Tang, M. Andriluka, and B. Schiele, "Detection and Tracking of Occluded People," *International Journal of Computer Vision (IJCV)*, 2014.

[30] E. Goubet, J. Katz, and F. Porikli, "Pedestrian tracking using thermal infrared imaging," *Defense and Security Symposium*. *International Society for Optics and Photonics*, 2006.

[31] Grassi, Ana Pérez, Vadim Frolov and Fernando Puente León, "Information fusion to detect and classify pedestrians using invariant features," *Information Fusion,* vol. 12, no. 4, pp. 284-292, 2011.

[32] B. Zhu, Y. Ding, J. Hua, K. Hao, and L. Hong, "Fusion of depth and vision information for human detection," *Journal of Computational Information Systems,* vol. 9, no. 20, pp. 8147-8154. 2013.

[33] G. E. Thomas, P. D. Frazier, and M. F. Chouikha, "Improved human detection using image fusion," *Proceedings of the IEEE ICRA Workshop on People Detection and Tracking*, Kobe, Japan. 2009.

[34] I. Huerta, G. Ferrer, F. Herrero, A. Prati, & A. Sanfeliu, "Multimodal feedback fusion of laser, image and temporal information," In *Proceedings of the International Conference on Distributed Smart Cameras*, p. 25, ACM, 2014.

[35] T. T. Zin, H. Takahashi, H. Hama, and T. Toriu, "Fusion of infrared and visible images for robust person detection," *INTECH Open Access Publisher*, pp. 1-26, 2011.

[36] H. Ju, and B. Bhanu, "Fusion of color and infrared video for moving human detection," *Pattern Recognition,* vol. 40, no. 6, pp. 1771-1784, 2007.

[37] B. Bhanu, and X. Zou, "Moving human detection based on multi-modal sensor fusion," In *Computer Vision and Pattern Recognition Workshop*, *(CVPRW),* pp. 136-136., 2004.