REDEFINING HISTOGRAMS OF ORIENTED GRADIENTS DESCRIPTORS FOR HANDLING OCCLUSION IN PEOPLE DETECTION

AHMAD SANI

UNIVERSITI TEKNOLOGI MALAYSIA

REDEFINING HISTOGRAMS OF ORIENTED GRADIENTS DESCRIPTORS FOR HANDLING OCCLUSION IN PEOPLE DETECTION

AHMAD SANI

A dissertation submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

JULY 2014

Dedicated with love to those I care,,,

ACKNOWLEDGEMENT

In the name of Allah, the most beneficent and the most merciful; all praise be to Him, the sole lord of the universe; and peace and blessing be upon the Prophet Muhammad, the seal of all prophets.

In this limited space I would like to show my gratitude to all people around; the family for their endless support, and friends for standing beside me. My thanks go to my supervisor, Assoc. Prof. Dr. Mohd Shafry Mohd Rahim, for his patience; and to Universiti Teknologi Malaysia and the Faculty of Computing as well. I was honoured to be a student at this university and under Dr. Shafry's supervision.

A very special thank you is due to my mother for everything, and to my sister, Asst. Prof. Dr. Alawiyah, as well for her special favour. God bless them all.

ABSTRACT

Object detection is a big challenge for researchers to address the issues that affect accurate detections. The Histogram of oriented gradients (HOG) descriptors have been used extensively for object detection on challenging conditions with good results. However, occlusion remains a well-known issue where some parts of an object are only visible. Conventional frameworks using HOG descriptors for detecting whole human body are prone to occluded people and thus causing errors on implementations running these frameworks. This issue was addressed in different approaches like hybrid feature extraction algorithms. Rather than following such approaches, the problem is addressed in this research by redefining HOG descriptors to work in a part object (body) detection framework that is powered by a finite-state machine. Experiments were conducted to test a new method of extracting HOG descriptors for parts of human body via slicing a window's HOG descriptor into four. The part body detection framework utilises each part's descriptor whereas three detected parts are sufficient to declare that the window has an assumed instance. Support vector machines (SVM) were used to classify the extracted parts and a finite-state machine was employed for handling the detected parts. Training and testing were made on subsets from the INRIA Person Dataset and the results favoured the part body detection framework over the conventional whole body framework. For a test set of 50 positive images of occluded people, the part body detection framework successfully detected 46 true positive while the whole body framework detected 36. Moreover, the former had less false positive detection than the latter having 80 false positive windows comparing to 289.

ABSTRAK

Pengenalpastian objek, merupakan cabaran besar kepada para pengkaji bagi mengenalpasti isu-isu yang mempengaruhi ketepatan pengesanan. Pengesan *Histogram of Oriented Gradients* (HOG) telah digunakan secara meluas dalam mengenalpasti objek, dalam keadaan yang mencabar; dengan keputusan yang baik. Namun, pertindihan kekal sebagai isu yang sedia dimaklumi di mana hanya sesetengah bahagian sesuatu objek kelihatan. Rangka kerja menggunakan pengesan HOG untuk mengesan seluruh badan manusia cenderung menghalang keterlihatan manusia; yang menimbulkan ralat dalam pelaksanaannya. Isu ini diatasi melalui pendekatan berbeza seperti menggunakan algoritma yang mengesan *feature* campuran. Berbanding menggunakan pendekatan tersebut, masalah dirungkai dalam kajian ini melalui penelitian semula terhadap pengesan HOG supaya beroperasi mengikut rangka kerja pengesanan bahagian objek(badan) yang dikuasai oleh *finite-state machine*. Kajian dijalankan bagi menguji kaedah baru untuk merungkai pengesan HOG mengikut bahagian badan manusia melalui penghirisan tetingkap pengesan HOG kepada empat(4) bahagian. Rangka kerja pengesanan bahagian badan merungkai setiap bahagian; walaupun tiga(3) bahagian yang dikesan sudah mencukupi untuk mengesahkan tetingkap mempunyai imej seperti yang diandaikan. *Support Vector Machines* (SVM) berfungsi mengelaskan bahagian yang dikesan dan *Finite-State Machine* digunakan untuk mengendalikan bahagian yang dikesan. Latihan dan ujian dilakukan ke atas subset INRIA *Person Dataset* dan keputusan memihak kepada rangka kerja pengesanan bahagian badan berbanding seluruh badan. Untuk ujian ke atas 50 set imej positif manusia yang terlindung, rangka kerja pengesanan bahagian badan berjaya mengesan 46 imej positif benar manakala hanya 36 bagi rangka kerja seluruh badan. Malahan, rangka kerja pengesanan bahagian badan mempunyai kadar pengesanan positif palsu yang lebih rendah berbanding seluruh badan dengan perbandingan 80 dan 289 tetingkap positif palsu.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Detecting people in computer vision has been an active research area for its promising applications. The research in computer vision (and computer science in general) has given computers abilities inspired by nature such as the sense of sight; and nowadays there are many implementations in road-safety and surveillance having these abilities. Many algorithms that run these implementations are not exclusive only for *people detection*; researchers have extended them to other classes like animals, cars, and generally other objects. *Detection frameworks* or *detectors* vary from using pattern recognition algorithms in a holistic approach to exhaustive scanning at every location in the image. However, from observing nowadays detectors, there are made of two main processes of *feature extraction* and *machine learning*. *Features* are used as a more convenient way to describe object, thus they are often called *descriptors*, while machine learning is used for classifying these features to answer the question whether an instance of an object exist or not in an image. Designing detectors following this way of using two processes gives better understanding of problems in computer vision and thus they are addressed properly. The problem of occlusion, for instance, is well known to affect detections where parts of instances are overlapping or hidden by other objects; thus the research community proposed detecting any visible part in an approach known by *part object detection* (or *part-based detection*) rather than the other approach of *whole object detection* (the conventional way to detect the whole object). Other problems such as view variations, poses, carrying belongings, clothes, scale, etc. were addressed as well as many features were proposed to handle such problems. However, the histograms of oriented gradients (HOG) descriptors proved to be robust to many issues and this dissertation aims to improve the use of these descriptors via redefining them into a framework of part object detection.

1.2 Research Background

Object detection is a task with different issues regarding how the objects appear in images. First, there was the question of how to find/detect instances using their shapes (Borgefors, 1988); object were matched to a template that is rigid to any object variations, thus more convenient ways to describe object were required. Oren *et al.* (1997) proposed a detection framework that has influenced many nowadays detectors; they proposed an exhaustive approach for scanning images for any instance using *sliding windows* where Haar wavelet features are extracted then a support vector machine (machine learning) classifies whether instances are there or not. Viola and Jones (2001) introduced their *face detection* (a specific domain in object detection) framework and went successful for real-time performance thanks to the fast computing in Haar wavelets and the integral image.

However, while these features have good detection performance when the variation of intensities between the object and the background is high, the detection frameworks faces troubles when the intensities variation are low due to issues related to cluttered background, clothes, belongings, etc. where texture-based features are not suitable. Gradient-based features came to be better for such situations where an object's silhouette can be captured well. Adding the statistical concepts of orientation histograms at local patches had given more robust features such as the scale-invariant feature transform (SIFT) by Lowe (1999). SIFT returns keypoints used basically for recognising objects, yet Mikolajczyk *et al.* (2004) implemented SIFT features for object detection where their detector performed well. This performance motivated further studying local histograms of oriented gradients, in terms of optimising computation time (Yan and Sukthankar, 2004) or increasing detection accuracy (Mikolajczyk and Schmid, 2005). Yet among many, the histogram of oriented gradients (HOG) descriptor (Dalal and Triggs, 2005) managed to capture the human body silhouette efficiently where experiments on HOG descriptors have showed excellent performance on the MIT pedestrian dataset and the INRIA person dataset. HOG descriptors work well in capturing the whole human body shape that may appear in different poses, views, clothes, and under different lightening conditions (the illumination factor). Yet, like SIFT, it has slower computation time comparing to fast feature extraction algorithms like Haar wavelets. There have been some algorithms for HOG-based descriptors aimed for speeding up the process by selecting likely valid features using machine learning methods, for example using AdaBoost for feature selection (Qiang *et al.*, 2006). While others aimed for increasing the accuracy and detection rate regardless of time costs (as computer hardware are advancing), such as a work by Satpathy *et al.* (2014).

In general, the amount of works increased for people detection (as well as object detection) research in the last decade; a survey by Geronimo *et al.* (2010) discussed the domain of people detection becoming a very active topic and draw the attention for building pedestrian protection systems (PPS) in advanced driver assistance systems (ADAS). Each work aimed to addressing challenges in object detection. But since HOG descriptors has done well with many issues, they were used widely for human and object detection in general (Kittipanya-ngam and Lung, 2011). Moreover, HOG descriptors are favoured over SIFT since the former does not require a license while the SIFT algorithm was patented (Lowe, 2004a) and requires fees for its implementations. However, the problem of occlusion remains a great challenge in computer vision to be covered and researchers followed different approaches to handle it. However, part object detection frameworks remain a naïve yet effective way to detect any visible part and join them using a model. Based on these frameworks, researchers introduced large but well-developed detectors employing a number of algorithms. For example, a work by Felzenszwalb *et al.* (2010) presented a part object detection framework that uses deformable part model to classify a number of objects and was able to handle simple occlusion. Wang *et al.*

(2009) presented a hybrid detection framework for handling occluded scenes using HOG descriptors and local binary pattern (LBP) features where it aims first to find occluded regions and then run the part detector where needed. This work was followed recently in similar concept by Marin *et al.* (2014). Although all these works handled varying levels of occlusion well, the framework inside remains complex with many algorithm that cost time.

Using simpler frameworks may lead to affordable answers for handling occlusion rather than consume more time on e.g. extracting different descriptors; what if the same descriptor is used *redefined* in a part object detection framework? And as a part object detector, what model will be used to join the detected parts?

1.3 Problem Statement

Although HOG descriptors have been utilised extensively for human and objects detection in general, their performance drops down when they deal with occlusion. Occlusion is a problem where parts of people are hidden by others causing more missing rates while detecting. In the available works that were reviewed, occlusion was handled in different detection frameworks using different algorithms and/or features. However, since part object detection frameworks can be used to handle occlusion and HOG has better performance on separate objects, occlusion can be handled when HOG is redefined to work in a part object detection framework that is less complex (without the need to use features other than the available HOG descriptor). This framework adds for HOG descriptors the robustness to occlusion and introduces a simple model that handles the detected part efficiently.

1.4 Research Aim

This dissertation aims to develop a variant HOG descriptor that is capable to work in a part object detection framework that is robust to occlusion. The framework should be less complex by redefining and utilising the available HOG descriptors and properly handles the detected parts.

1.5 Research Objective

From the discussion in the previous sections, the objectives for this dissertation are:

1. to redefine HOG descriptors for parts of the human body in a part object detector;
2. to develop a part object detection framework able to handle occluded people; and
3. to evaluate the performance of the new HOG descriptors and the developed detector.

1.6 Research Scope

This dissertation with its objectives is outlined to the following:

1. Input images used are still images, whether they were originated from photographs or still frames (snapshots) from videos.
2. Input images are pre-processed by balancing the illumination.
3. The detection framework is made for detecting people walking or standing; no consideration for people shown in half such as sitting on benches or on bicycles and so forth.

4. The occlusion is defined by one part of the human body, either one of the arms or of the legs, being hidden by another.
5. People in input images may wear any clothes, carry belongings, and be in different poses or from different views as long as there are visible three parts to detect.
6. The detection runs on one scale of the image and input images for testing are pre-processed by cropping them to make all people be in one defined size.

1.7 Research Significance

Object detection is a promising research area with endless implementations such as ADAS vehicle systems, robotics, and surveillance. However, more challenges occur as implementations grow. Occlusion is an everyday situation that affects any system relying on object detection frameworks. Therefore, there is a need to further the research on new HOG-based descriptors which are robust to occlusion and less complex to implement. Furthermore, the dissertation's detection framework that utilises the new descriptors can be extended to work on other object classes for general object detection. The dissertation will contribute to different research areas, i.e. occlusion in object detection, introducing redefined HOG descriptors for human parts, and a new part object detection frameworks.

1.8 Chapter Summary

This chapter discussed the dissertation topic and the problem it addresses. In a nutshell, a redefined HOG descriptor are used in a part object framework for detecting people whose parts are hidden by other objects. Furthering such research will help overcoming the occlusion problem that affects performance in object detection. The next chapter discusses some works related to the dissertation topic.

REFERENCES

Almuashi, M. A.,2013. *Recognition of Partially Occluded Objects in 2d Images.* Master of Science, Universiti Teknologi Malaysia.

Andreopoulos, A. and Tsotsos, J. K.,2013. 50 Years of Object Recognition: Directions Forward. *Computer Vision and Image Understanding.* 117(8)**,** 827-891.

Aziz, K. A. A., Ramlee, R. A., Abdullah, S. S. and Jahari, A. N.,2009. Face Detection Using Radial Basis Function Neural Networks with Variance Spread Value. *Proceedings of the 2009 International Conference of Soft Computing and Pattern Recognition*. 4-7 Dec. Malacca, Malaysia: IEEE, 399-403.

Azizpour, H. and Laptev, I.,2012. Object Detection Using Strongly-Supervised Deformable Part Models. In Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y. and Schmid, C. (Eds.). *Computer Vision.* (pp. 836-849) Berlin: Springer Berlin Heidelberg.

Bay, H., Tuytelaars, T. and Gool, L.,2006. Surf: Speeded up Robust Features. In Leonardis, A., Bischof, H. and Pinz, A. (Eds.). *Computer Vision.* (pp. 404-417) Berlin: Springer.

Bertini, M., Bimbo, A. D. and Nunziati, W.,2005. Common Visual Cues for Sports Highlights Modeling. *Multimedia Tools and Applications.* 27(2)**,** 215-218.

Borgefors, G.,1988. Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm. *Transactions on Pattern Analysis and Machine Intelligence.* 10(6)**,** 849-865.

Boser, B. E., Guyon, I. M. and Vapnik, V. N.,1992. A Training Algorithm for Optimal Margin Classifiers. *Proceedings of the 1992 5th annual workshop on Computational learning theory*. 27-29 Jul. Pittsburgh, Pennsylvania: ACM, 144-152.

Chao, Z., Bichot, C.-E. and Liming, C.,2011. Visual Object Recognition Using Daisy Descriptor. *Proceedings of the 2011 International Conference on Multimedia and Expo*. 11-15 Jul. Barcelona, Spain: IEEE, 1-6.

Chi-Hung, C., Hsieh, J.-W., Luo-Wei, T., Chen, S.-Y. and Kuo-Chin, F.,2009. Carried Object Detection Using Ratio Histogram and Its Application to Suspicious Event Analysis. *Transactions on Circuits and Systems for Video Technology*. 19(6)**,** 911-916.

Crow, F. C.,1984. Summed-Area Tables for Texture Mapping. *Computer Graphics.* 18(3)**,** 207-212.

Dalal, N.,2006. *Finding People in Images and Videos.* Doctor of Philosiphy, Institut National Polytechnique de Grenoble.

Dalal, N. and Triggs, B.,2005. Histograms of Oriented Gradients for Human Detection. *Proceedings of the 2005 Conference on Computer Vision and Pattern Recognition* 25 Jun. San Diego, CA: IEEE, 886-893.

Dalal, N., Triggs, B. and Schmid, C.,2006. Human Detection Using Oriented Histograms of Flow and Appearance. In Leonardis, A., Bischof, H. and Pinz, A. (Eds.). *Computer Vision.* (pp. 428-441) Berlin: Springer.

DTREG,2014. *Svm - Support Vector Machines*. Accessed on 30 June, 2014. http://www.dtreg.com/svm.htm

Everingham, M., Gool, L., Williams, C. I., Winn, J. and Zisserman, A.,2010. The Pascal Visual Object Classes (Voc) Challenge. *International Journal of Computer Vision.* 88(2)**,** 303-338.

Felzenszwalb, P., McAllester, D. and Ramanan, D.,2008. A Discriminatively Trained, Multiscale, Deformable Part Model. *Proceedings of the 2008 Computer Vision and Pattern Recognition*. 23-28 Jun. Anchorage, AK: IEEE, 1-8.

Felzenszwalb, P. F., Girshick, R. B., McAllester, D. and Ramanan, D.,2010. Object Detection with Discriminatively Trained Part-Based Models. *Transactions on Pattern Analysis and Machine Intelligence.* 32(9)**,** 1627-1645.

Freund, Y. and Schapire, R. E.,1997. A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences.* 55(1)**,** 119-139.

Gavrila, D. M.,1999. The Visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding.* 73(1)**,** 82-98.

Gavrila, D. M. and Philomin, V.,1999. Real-Time Object Detection for Smart Vehicles. *Proceedings of the 1999 7th International Conference on Computer Vision*. 1999. Kerkyra: IEEE, 87-93.

Geronimo, D., Lopez, A. M., Sappa, A. D. and Graf, T.,2010. Survey of Pedestrian Detection for Advanced Driver Assistance Systems. *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* 32(7)**,** 1239-1258.

Guoqing, X., Xiaocui, W., Li, L. and Zhengbin, W.,2011. Real-Time Pedestrian Detection Based on Edge Factor and Histogram of Oriented Gradient. *Proceedings of the 2011 International Conference on Information and Automation*. 6-8 Jun. Shenzhen: IEEE, 384-389.

He, X. and Niyogi, P.,2002. Locality Preserving Projections. *Advances in Neural Information Processing Systems.* 16(1)**,** 153-160.

Heewook, J., Joo Kooi, T., Ishikawa, S. and Morie, T.,2011. Applying Hog Feature to the Detection and Tracking of a Human on a Bicycle. *Proceedings of the 2011 International Conference on Control, Automation and Systems*. 26-29 Oct. Gyeonggi-do: IEEE, 1740-1743.

Hotelling, H.,1933. Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology.* 24(6)**,** 417-441.

Hui-Xing, J. and Yu-Jin, Z.,2007. Fast Human Detection by Boosting Histograms of Oriented Gradients. *Proceedings of the 2007 4th International Conference on Image and Graphics*. 22-24 Aug. Sichuan: IEEE, 683-688.

Hui, L., Lei, Y., Xiaoyu, W., Shengmiao, X. and Youwen, W.,2012. Static Hand Gesture Recognition Based on Hog with Kinect. *Proceedings of the 2012 International Conference on Intelligent Human-Machine Systems and Cybernetics*. 26-27 Aug. 2012. Nanchang, China: IEEE, 271-273.

Huttenlocher, D. P., Klanderman, G. A. and Rucklidge, W. J.,1993. Comparing Images Using the Hausdorff Distance. *Transactions on Pattern Analysis and Machine Intelligence.* 15(9)**,** 850-863.

Jae-Joon, H., Changkyu, C., ByungIn, Y., Dusik, P. and ChangYeong, K.,2013. Real-Time Hand Shape Recognition by Orientation Invariant Data Learning for Smart Tv. *Proceedings of the 2013 International Conference on Consumer Electronics*. 11-14 Jan. 2013. Las Vegas, NV: IEEE, 552-553.

Jiaolong, Y., Wei, L. and Yunde, J.,2012. Face Pose Estimation with Combined 2d and 3d Hog Features. *Proceedings of the 2012 International Conference on Pattern Recognition*. 11-15 Nov. 2012. Tsukuba, Japan: IEEE, 2492-2495.

Kah-Kay, S. and Poggio, T.,1998. Example-Based Learning for View-Based Human Face Detection. *Transactions on Pattern Analysis and Machine Intelligence.* 20(1)**,** 39-51.

Khastavaneh, H.,2013. *Moving Object Segmentation in Video Sequences Using Vector Quantization.* Master of Science, Universiti Teknologi Malaysia.

Kinh, T. and Viola, P.,2000. Boosting Image Retrieval. *Proceedings of the 2000 Conference on Computer Vision and Pattern Recognition*. 13-15 Jun. Hilton Head Island, SC: IEEE, 228-235.

Kittipanya-ngam, P. and Lung, E.,2011. Hog-Based Descriptors on Rotation Invariant Human Detection. In Koch, R. and Huang, F. (Eds.). *Computer Vision – Accv 2010 Workshops.* (pp. 143-152) Berlin: Springer Berlin Heidelberg.

Kobayashi, T., Hidaka, A. and Kurita, T.,2008. Selection of Histograms of Oriented Gradients Features for Pedestrian Detection. In Ishikawa, M., Doya, K., Miyamoto, H. and Yamakawa, T. (Eds.). *Neural Information Processing.* (pp. 598-607) Berlin: Springer.

Liang, Z. and Thorpe, C.,1999. Stereo- and Neural Network-Based Pedestrian Detection. *Proceedings of the 1999 International Conference on Intelligent Transportation Systems*. 1999. Tokyo: IEEE, 298-303.

Lienhart, R. and Maydt, J.,2002. An Extended Set of Haar-Like Features for Rapid Object Detection. *Proceedings of the 2002 International Conference on Image Processing*. 22-25 Sep. Santa Clara, CA: IEEE, 900-903.

Linh, D., Buu, B., Vo, P. D., Tran, T. N. and Le, B. H.,2011. Improved Hog Descriptors. *Proceedings of the 2011 3rd International Conference on Knowledge and Systems Engineering*. 14-17 Oct. Hanoi: IEEE, 186-189.

Loch, F. G.,2010. *Image Processing Algorithms Part 6: Gamma Correction*. Accessed on 19 Oct, 2013. http://www.dfstudios.co.uk/2010/12/image-processing-algorithms-part-6/

Lowe, D. G.,1999. Object Recognition from Local Scale-Invariant Features. *Proceedings of the 1999 7th International Conference on Computer Vision*. 20-27 Sep. Kerkyra (Corfu), Greece: IEEE, 1150-1157.

Lowe, D. G.,2004a. *US6711293 B1*. United States: The University of British Columbia, Vancouver (CA).

Lowe, D. G.,2004b. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision.* 60(2)**,** 91-110.

Maji, S., Berg, A. C. and Malik, J.,2008. Classification Using Intersection Kernel Support Vector Machines Is Efficient. *Proceedings of the 2008 Conference on Computer Vision and Pattern Recognition*. 23-28 Jun. Anchorage, AK: IEEE, 1-8.

Marin, J., Vazquez, D., Lopez, A. M., Amores, J. and Kuncheva, L. I.,2014. Occlusion Handling Via Random Subspace Classifiers for Human Detection. *Transactions on Cybernetics.* 44(3)**,** 342-354.

Menotti, D., Najman, L., Facon, J. and de Araujo, A. A.,2007. Multi-Histogram Equalization Methods for Contrast Enhancement and Brightness Preserving. *Transactions on Consumer Electronics.* 53(3)**,** 1186-1194.

Mikolajczyk, K. and Schmid, C.,2005. A Performance Evaluation of Local Descriptors. *Pattern Analysis and Machine Intelligence.* 27(10)**,** 1615-1630.

Mikolajczyk, K., Schmid, C. and Zisserman, A.,2004. Human Detection Based on a Probabilistic Assembly of Robust Part Detectors. In Pajdla, T. and Matas, J. (Eds.). *Computer Vision - Eccv 2004.* (pp. 69-82) Berlin: Springer.

Mohan, A., Papageorgiou, C. and Poggio, T.,2001. Example-Based Object Detection in Images by Components. *Transactions on Pattern Analysis and Machine Intelligence.* 23(4)**,** 349-361.

Munder, S. and Gavrila, D. M.,2006. An Experimental Study on Pedestrian Classification. *Pattern Analysis and Machine Intelligence.* 28(11)**,** 1863-1868.

Neagoe, V., Ciotec, A. and Barar, A.,2012. A Concurrent Neural Network Approach to Pedestrian Detection in Thermal Imagery. *Proceedings of the 2012 9th International Conference on Communications*. 21-23 June 2012. Bucharest, Romania: IEEE, 133-136.

Neagoe, V. and Ropot, A. D.,2002. Concurrent Self-Organizing Maps for Pattern Classification. *Proceedings of the 2002 1st International Conference on Cognitive Informatics*. 2002. Calgary, Alberta, Canada: IEEE, 304-312.

Onis, S., Sanson, H. and Garcia, C.,2008. Efficient Object Detection Robust to Rst with Minimal Set of Examples. *Proceedings of the 2008 International Joint*

Conference on Computer Vision and Computer Graphics Theory and Applications. 22-25 Jan. Funchal, Portugal: Institute for Systems and Technologies of Information, Control and Communication (INSTICC), 179-185.

Oren, M., Papageorgiou, C., Sinha, P., Osuna, E. and Poggio, T.,1997. Pedestrian Detection Using Wavelet Templates. *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*. 17-19 Jun. San Juan: IEEE, 193-199.

Osman, S. M. and Abu-Bakar, S. A. R.,2005. Classification of Moving Objects Using Active Vision. *Proceedings of the 2005 1st International Conference on Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering*. 14-16 Nov. 2005. Kuala Lumpur, Malaysia: IEEE, 112-114.

Paglieroni, D. W., Ford, G. E. and Tsujimoto, E. M.,1994. The Position-Orientation Masking Approach to Parametric Search for Template Matching. *Transactions on Pattern Analysis and Machine Intelligence.* 16(7)**,** 740-747.

Papageorgiou, C. P., Oren, M. and Poggio, T.,1998. A General Framework for Object Detection. *Proceedings of the 1998 6th International Conference on Computer Vision*. 4-7 Jan. Bombay: IEEE, 555-562.

Pearson, K.,1901. On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine.* 2(6)**,** 559-572.

Perrotton, X., Sturzel, M. and Roux, M.,2008. Automatic Object Detection on Aerial Images Using Local Descriptors and Image Synthesis. In Gasteratos, A., Vincze, M. and Tsotsos, J. (Eds.). *Computer Vision Systems.* (pp. 302-311) Berlin: Springer.

Porikli, F.,2005. Integral Histogram: A Fast Way to Extract Histograms in Cartesian Spaces. *Proceedings of the 2005 Conference on Computer Vision and Pattern Recognition*. 20-25 Jun. Cambridge, MA: IEEE, 829-836.

Qiang, Z., Mei-Chen, Y., Kwang-Ting, C. and Avidan, S.,2006. Fast Human Detection Using a Cascade of Histograms of Oriented Gradients. *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition*. 17-22 Jun. New York, NY: IEEE, 1491-1498.

Qing Jun, W. and Ru Bo, Z.,2008. Lpp-Hog: A New Local Image Descriptor for Fast Human Detection. *Proceedings of the 2008 International Symposium on*

Knowledge Acquisition and Modeling Workshop. 21-22 Dec. Wuhan: IEEE, 640-643.

Rowley, H. A., Baluja, S. and Kanade, T.,1998. Neural Network-Based Face Detection. *Transactions on Pattern Analysis and Machine Intelligence.* 20(1)**,** 23-38.

Satpathy, A., Jiang, X. and Eng, H. L.,2014. Human Detection by Quadratic Classification on Subspace of Extended Histogram of Gradients. *Transactions on Image Processing.* 23(1)**,** 287-297.

Schneiderman, H. and Kanade, T.,2004. Object Detection Using the Statistics of Parts. *International Journal of Computer Vision.* 56(3)**,** 151-177.

Sheng, W.,2011. A Review of Gradient-Based and Edge-Based Feature Extraction Methods for Object Detection. *Proceedings of the 2011 11th International Conference on Computer and Information Technology*. 31 Aug - 2 Sep. San Juan: IEEE, 277-282.

Shyang-Lih, C., Fu-Tzu, Y., Wen-Po, W., Yu-An, C. and Sei-Wang, C.,2011. Nighttime Pedestrian Detection Using Thermal Imaging Based on Hog Feature. *Proceedings of the 2011 Conference on System Science and Engineering*. 8-10 Jun. Macao: IEEE, 694-698.

Sidla, O. and Rosner, M.,2007. Hog Pedestrian Detection Applied to Scenes with Heavy Occlusion. *Proceedings of the 2007 Intelligent Robots and Computer Vision XXV: Algorithms, Techniques, and Active Vision*. 9-10 Sep. Boston, MA: SPIE, 8-19.

Sonka, M., Hlavac, V. and Boyle, R.,2007. *Image Processing, Analysis, and Machine Vision.* Stamford, Connecticut: Cengage Learning.

Sung, K.,2011. Recent Videogame Console Technologies. *Computer.* 44(2)**,** 91-93.

Szarvas, M., Yoshizawa, A., Yamamoto, M. and Ogata, J.,2005. Pedestrian Detection with Convolutional Neural Networks. *Proceedings of the 2005 Intelligent Vehicles Symposium*. 6-8 June 2005. Las Vegas, NV: IEEE, 224-229.

Tang, S., Andriluka, M. and Schiele, B.,2013. Detection and Tracking of Occluded People. *International Journal of Computer Vision.* 11263(1)**,** 1-12.

Toderici, G., Aradhye, H., Pasca, M., Sbaiz, L. and Yagnik, J.,2010. Finding Meaning on Youtube: Tag Recommendation and Category Discovery.

Proceedings of the 2010 Conference on Computer Vision and Pattern Recognition. 13-18 Jun. San Francisco, CA: IEEE, 3447-3454.

Tola, E., Lepetit, V. and Fua, P.,2010. Daisy: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo. *Transactions on Pattern Analysis and Machine Intelligence.* 32(5)**,** 815-830.

Tsotsos, J. K., Liu, Y., Martinez-Trujillo, J. C., Pomplun, M., Simine, E. and Zhou, K.,2005. Attending to Visual Motion. *Computer Vision and Image Understanding.* 100(1–2)**,** 3-40.

Viola, P. and Jones, M.,2001. Rapid Object Detection Using a Boosted Cascade of Simple Features. *Proceedings of the 2001 Conference on Computer Vision and Pattern Recognition*. 8-14 Dec. Kauai: IEEE, 511-518.

Viola, P. and Jones, M.,2004. Robust Real-Time Face Detection. *International Journal of Computer Vision.* 57(2)**,** 137-154.

Viola, P., Jones, M. J. and Snow, D.,2003. Detecting Pedestrians Using Patterns of Motion and Appearance. *Proceedings of the 2003 9th International Conference on Computer Vision*. 13-16 Oct. Nice, France: IEEE, 734-741.

Vondrick, C., Khosla, A., Malisiewicz, T. and Torralba, A.,2013. Hoggles: Visualizing Object Detection Features. *Proceedings of the 2013 International Conference on Computer Vision*. 1-8 Dec. 2013. Sydney, NSW: IEEE, 1-8.

Wang, X., Han, T. X. and Yan, S.,2009. An Hog-Lbp Human Detector with Partial Occlusion Handling. *Proceedings of the 2009 12th International Conference on Computer Vision*. 29 Sept - 2 Oct. Kyoto: IEEE, 32-39.

Xingbao, M., Jing, L. and Yingchun, D.,2012. An Extended Hog Model: Schog for Human Hand Detection. *Proceedings of the 2012 International Conference on Systems and Informatics*. 19-20 May. Yantai: IEEE, 2593-2596.

Yafei, Z., Weidong, W. and Yuehai, W.,2011. A Real-Time Hand Gesture Recognition Method. *Proceedings of the 2011 International Conference on Electronics, Communications and Control*. 9-11 Sept. 2011. Ningbo, China: IEEE, 2475-2478.

Yan, K. and Sukthankar, R.,2004. Pca-Sift: A More Distinctive Representation for Local Image Descriptors. *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition*. 27 Jun - 2 Jul. Pittsburgh, PA: IEEE, 506-513.

Yi, Y. and Ramanan, D.,2013. Articulated Human Detection with Flexible Mixtures of Parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 35(12)**,** 2878-2890.

Yihua, S., Jinfeng, Y. and Renbiao, W.,2007. Reducing Illumination Based on Nonlinear Gamma Correction. *Proceedings of the 2007 International Conference on Image Processing*. 16 Sep - 19 Oct. San Antonio, TX: IEEE, 529-532.

Zaklouta, F. and Stanciulescu, B.,2011. Warning Traffic Sign Recognition Using a Hog-Based K-D Tree. *Proceedings of the 2011 Intelligent Vehicles Symposium*. 5-9 June 2011. Baden-Baden, Germany: IEEE, 1019-1024.

Zhao, W., Han, X., Wenhao, H., Feng, W. and Kui, Y.,2013. Real-Time Sift-Based Object Recognition System. *Proceedings of the 2013 International Conference on Mechatronics and Automation*. 4-7 Aug. Takamatsu, Japan: IEEE, 1361-1366.