Institut National des Sciences Appliquées de Rouen

Laboratoire d'Informatique de Traitement de l'Information et des Systèmes

Universitatea "Babeş-Bolyai"

Facultatea de Matematică şi Informatică, Departamentul de Informatică

Extended abstract for thesis defence

PHD THESIS

Speciality : Computer Science

Defended by

Miron Alina Dana

to obtain the title of

PhD of Science of INSA de ROUEN

and "Babeş-Bolyai" University

Multi-modal, Multi-Domain Pedestrian Detection and Classification: Proposals and Explorations in Visible over StereoVision, FIR and SWIR

PhD Directors:

Horia F. Pop - *Professor* - "Babeş-Bolyai" University

Abdelaziz Bensrhair - *Professor* - INSA de Rouen

Contents of the thesis

Keywords: Intelligent Vehicles, Pedestrian Detection, Far-Infrared, Short-Wave Infrared, StereoVision

Introduction

Intelligent autonomous vehicles have long surpassed the stage of a Sci-Fi idea, and have become a reality [62],[1]. The main motivation behind this technology is to increase the safety of both driver and other traffic participants. In this context, pedestrian protection systems have become a necessity. But merely passive components like airbags are not enough: active safety, technology assisting in the prevention of a crash, is vital. For this, a system of pedestrian detection and classification plays a fundamental role.

Challenges

Pedestrian detection and classification in the context of intelligent vehicles in an urban environment poses a lot of challenges:

Pedestrian Appearance and Shape. By nature, the humans have different heights and body shapes. But this variability in appearance is further increased by different cloth types. Moreover, human shape can change a lot in a short period of time (for example a person that bends to tie its shoes). Also the appearance depends on the point of view of the camera, as well as the distance between the camera and the pedestrian. Close pedestrians can bear little resemblance with the ones situated far away.

Occlusion. Occlusions represents an important challenge for the detection of any type of object, and in the case of pedestrians they can be divided into: self and external occlusions. Self-occlusion are cause especially by the pose of the object, in the case of a pedestrian that has a side-way position in relation with the point of view of the camera will certainly exhibit occlusion of some body-parts. Moreover different objects carried by the pedestrians might have the same effect (for example hats, bags, umbrellas). In the external occlusions category we include other pedestrians (especially in an urban situation), poles, other cars, as well as the situation in which the pedestrian is too close to the camera leading certain body-parts exit the field of view.

Environmental conditions. Although some meteorological circumstances might not have a direct impact on the quality of images (for example light rain), they can influence the appearance of pedestrians for cameras (for example a passer-by can open an umbrella which might lead to occlusion of the head region). Other conditions might lead to situations where the quality of retrieved images is altered (for example situations of haze, fog, snow, heavy rain etc.). Another factor that should be taken into consideration is the time of day, that has a direct impact over the amount of ambient light available - usually, during daytime the problem of pedestrian detection and classification poses less problems than during night.

Sensor choice. Each existing sensor has certain disadvantages and advantages, depending on the situation. For example, passive sensors like visible cameras can be affected by low light conditions, giving poor images with low variation in intensity across objects and background, while thermal cameras might experience the same problems when the environment has a similar temperature with the pedestrians. Active sensors, like LIDAR, have the advantage of providing distance to all objects in a scene, but they have as output a large datasets that might be difficult to interpret.

Other objects. Distinction between non-pedestrians and pedestrians might not be always simple, being difficult to construct a model that differentiates between pedestrians and any other existing objects.

Main Research Contributions

Motivated by the importance of pedestrian detection, there exist an extensive amount of work done in connection with this field. *Our objective is to study the problem across different light spectrum and modalities, with an emphasis on disparity map.*

Our main contributions can be summarised as follows:

- Creation and annotation of two databases for benchmarking of pedestrian classification, one for Far-Infrared (Thermal) and the other one in Short-Wave Infrared (SWIR).
- In the context of Thermal images, we have proposed a new feature, Intensity Self Similarity (ISS). The performance of ISS was compared on three different datasets with state of the art features.

- As a novelty, we have studied the SWIR spectrum for the task of pedestrian classification, and we have performed a comparison with the Visible domain.
- As a low cost solution, we believe that Stereo Vision is a promising solution. In this context, we have also focused on improving Stereo Matching algorithm by proposing new cost functions.
- We have studied the performance of different features across different domains (Visible, FIR) and across multiple modalities (Intensity, Motion, Disparity map)

Thesis Overview

This thesis is organized as follows (see also figure 1):

Chapter 1 presents an in-depth analysis for the motivation of a pedestrian detection system, along with an overview of existing types of sensors. Our sensor of choice is passive sensors represented by cameras sensitive to different light spectrums: Visible, Far Infrared and Short Wave Infrared. We present also a short review of the steps employed in the task of pedestrian classification and detection with an emphasise on the step of feature computation.

In **Chapter 2** we study the problematic of pedestrian classification in Thermal images (Far-Infrared Spectrum). After overviewing existing datasets of Thermal images, we have reached the conclusion that they all have important disadvantages: either the quality of the thermal images is poor and there is not possibility of direct comparison with the Visible spectrum; or the datasets are not publicly available. In this context, we have acquired and annotated a new dataset. Moreover we have proposed a feature adapted for pedestrian classification in Far-Infrared images and compared it with other state of the art features, in different conditions.

A new spectrum that can be interesting for the task of pedestrian detection and classification is the Short-Wave Infrared (SWIR). An analysis of this light spectrum is made in **Chapter 3**. After having performed some preliminaries experiments on a restricted dataset, we have acquired and annotated a dataset of SWIR images, along with the Visible correspondent. On this later dataset, we have compared the two spectrums from the perspective of different features.

Infrared cameras represent an interesting alternative to Visible cameras, and in general with better results, but remains an expensive one. In this context, StereoVision could improve the results obtained by just the employment of Visible cameras. **Chapter 4** deals with the algorithms of Stereo Matching. We propose several improvements for this algorithm, that mostly focus on the employed cost function.

Chapter 5 treats the problem of multi-modality pedestrian classification (Intensity, Depth

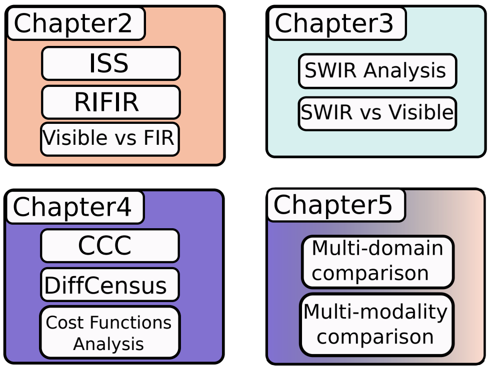

Figure 1: Thesis structure

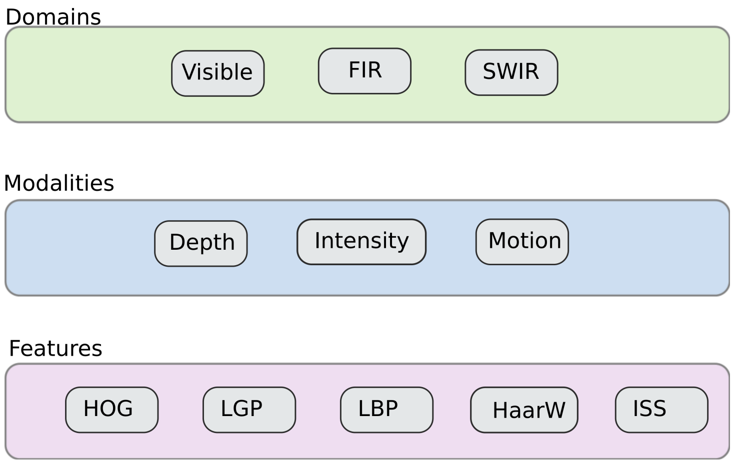

Figure 2: Domain-modality-feature relationship

and Optical Flow) in both Visible and FIR spectrum. In figure 2 is presented the difference between the domains and modalities employed. Moreover we show a preliminary analysis of the impact of the quality of the Disparity Map over the results of classification. Finally, conclusions and future work are presented in **Chapter 6**.

List of articles

Journal Papers

- **Alina Miron**, Samia Ainouz, Alexandrina Rogozan, Abdelaziz Bensrhair, *"A robust cost function for stereo matching of road scenes"*, Pattern Recognition Letters, No. 38, (2014): 70-77.

- **Alina Miron**, *"Post Processing Voting Techniques for Local Stereo Matching"*, Studia Univ. Babes-Bolyai, Informatica, Volume LIX, Number 1, (2014): 106-115

- **Alina Miron**, Samia Ainouz, Alexandrina Rogozan, Abdelaziz Bensrhair, *"Cross-comparison census for colour stereo matching applied to intelligent vehicle."*, Electronics

Letters 48.24 (2012): 1530-1532.

- **Alina Miron**, Samia Ainouz, Alexandrina Rogozan, Abdelaziz Bensrhair, Horia F. Pop,*"Stereo Matching Using radiometric Invariant measures"*, Studia Univ. Babes-Bolyai, Informatica, Volume LVI, No.3, (2011): 91-96.

Conferences

- Fan Wang, **Alina Miron**, Samia Ainouz, Abdelaziz Bensrhair, *Post-Aggregation Stereo Matching Method using Dempster-Shafer Theory*, IEEE International Conference on Image Processing 2014 (accepted)

- **Alina Miron**, Rean Isabella Fedriga, Abdelaziz Bensrhair, and Alberto Broggi, *SWIR Images Evaluation for Pedestrian Detection in Clear Visibility Conditions*, Proceedings of IEEE ITSC (2013): 354-359

- Massimo Bertozzi, Rean Isabella Fedriga, **Alina Miron**, and Jean-Luc Reverchon, *Pedestrian Detection in Poor Visibility Conditions: Would SWIR Help?*, IEEE ICIAP (2013): 229-238

- **Alina Miron**, Bassem Besbes, Alexandrina Rogozan, Samia Ainouz, Abdelaziz Bensrhair, *Intensity Self Similarity Features for Pedestrian Detection in Far-Infrared Images*, IEEE Intelligent Vehicle Symposium (2012): 1120-1125

- **Alina Miron**, Samia Ainouz, Alexandrina Rogozan, Abdelaziz Bensrhair, *Towards a robust and fast color stereo matching for intelligent vehicle application*, IEEE International Conference on Image Processing (2012): 465-468

Presentations

- One Day BMVA Symposium at the British Computer Society: *"Stereo Matching using invariant radiometric features"*, London, May 18th 2011

- Journee GdR ISIS, Analyse de scenes urbaines en image et vision, *"Stereo-vision for urban scenes."*, Nov. 8th 2012, Paris

Summary

The main purpose of constructing Intelligent Vehicles is to increase the safety for all traffic participants. The detection of pedestrians, as one of the most vulnerable category of road users, is paramount for any Advance Driver Assistance System (ADAS). Although this topic has been studied for almost fifty years, a perfect solution does not exist yet. This thesis focuses on several aspects regarding pedestrian classification and detection, and has the objective of exploring and comparing multiple light spectrums (Visible, ShortWave Infrared, Far Infrared) and modalities (Intensity, Depth by Stereo Vision, Motion).

From the variety of images, the Far Infrared cameras (FIR), capable of measuring the temperature of the scene, are particular interesting for detecting pedestrians. These will usually have higher temperature than the surroundings. Due to the lack of suitable public datasets containing Thermal images, we have acquired and annotated a database, that we will name RIFIR, containing both Visible and Far-Infrared Images. This dataset has allowed us to compare the performance of different state of the art features in the two domains. Moreover, we have proposed a new feature adapted for FIR images, called Intensity Self Similarity (ISS). The ISS representation is based on the relative intensity similarity between different sub-blocks within a pedestrian region of interest. The experiments performed on different image sequences have showed that, in general, FIR spectrum has a better performance than the Visible domain. Nevertheless, the fusion of the two domains provides the best results.

The second domain that we have studied is the Short Wave Infrared (SWIR), a light spectrum that was never used before for the task of pedestrian classification and detection. Unlike FIR cameras, SWIR cameras can image through the windshield, and thus be mounted in the vehicle's cabin. In addition, SWIR imagers can have the ability to see clear at long distances, making it suitable for vehicle applications. We have acquired and annotated a database, that we will name RISWIR, containing both Visible and SWIR images. This dataset has allowed us to compare the

performance of different pedestrian classification algorithms, along with a comparison between Visible and SWIR. Our tests have showed that SWIR might be promising for ADAS applications, performing better than the Visible domain on the considered dataset.

Even if FIR and SWIR have provided promising results, Visible domain is still widely used due to the low cost of the cameras. The classical monocular imagers used for object detection and classification can lead to a computational time well beyond real-time. Stereo Vision provides a way of reducing the hypothesis search space through the use of depth information contained in the disparity map. Therefore, a robust disparity map is essential in order to have good hypothesis over the location of pedestrians. In this context, in order to compute the disparity map, we have proposed different cost functions robust to radiometric distortions. Moreover, we have showed that some simple post-processing techniques can have a great impact over the quality of the obtained depth images.

The use of the disparity map is not strictly limited to the generation of hypothesis, and could be used for some feature computation by providing complementary information to color images. We have studied and compared the performance of features computed from different modalities (Intensity, Depth and Flow) and in two domains (Visible and FIR). The results have showed that the most robust systems are the ones that take into consideration all three modalities, especially when dealing with occlusions.

Conclusion

In this thesis we have focused on the problem of pedestrian detection and classification using different domains (FIR, SWIR, Visible) and different modalities (Intensity, Motion, Depth Map), with a particular emphasis on the Disparity map modality.

FIR. We have started by analysing Far-Infrared Spectrum. For this, we have annotated a large dataset, ParmaTetravision. Because this dataset is not publicly available, we have also acquired a new dataset called RIFIR. This has allowed us to construct a benchmark in order to analyse the performance of different features, and in the same time to compare FIR and Visible spectrums. Moreover, we have proposed a feature adapted for thermal images, called ISS. Altough ISS has a similar performance with that of HOG in the far infrared spectrum, local-binary features like LBP or LGP proved to be more robust. Moreover, in our tests, FIR consistently proved to be superior to Visible domain. Nevertheless, the fusion between Visible and FIR gave the best results, lowering the false positive rate with factor of ten in comparison with just using the FIR domain.

Since one of the main advantages of thermal images is the fact that the search space for possible pedestrians can be reduced to hot regions in the image, future work should include a benchmark of ROI extraction algorithms. Moreover, we can extend the feature comparison by testing different fusion techniques in order to find the most appropriate configuration.

SWIR With the advent of new camera sensors, a promising new domain is represented by Short-Wave Infrared (SWIR). In this context, we have experimented with two types of cameras. The preliminary experiments that were performed on a dataset that we have annotated, ParmaSWIR. This contains images taken using different filters with the purpose of isolation of different bandwidths. Since the results were promising, we have acquired another dataset, RISWIR, this time using both a SWIR and a Visible camera. On RISWIR, the short-wave infrared provided better results than the Visible one. In our opinion, this is due to the fact that acquired images in SWIR spectrum are sharper, having well-defined edges.

Further tests in SWIR domain should include different meteorological conditions, along with an evaluation during night conditions. Moreover, we believe for the results to be conclusive, SWIR cameras should be compared against several Visible cameras.

StereoVision Since Visible domain represents a low cost alternative to other spectrums, we give a special attention to Depth modality obtained by constructing the disparity map using different stereo matching algorithms. In this context, we have worked to improve existing stereo matching algorithms by proposing new cost function robust to radiometric distortions. As future work we plan on analysing the impact that post-processing algorithms have over the disparity map. In addition, in order to incorporate the findings of chapter **??**, we should improve the information contained in the areas subject to occlusions.

Multi-domain, multi-modality. In a similar manner with the way human perception uses clues given by depth and motion, a new direction of research is the combination of different modalities and features. A lot of articles tacked this problem from different features point of view for the Visible domain. Daimler Multi-cue dataset provides a way to centralize this analysis. In this context we have extended the number of features compared on the dataset with different modalities, along with several fusion scenarios. The best results were always obtained by fusing different modalities. Moreover, we extended the analysis multi-modality to a multi-domain approach, comparing Visible and FIR on ParmaTetravision dataset. Even if the FIR spectrum continues to give the best results, the fusion between Visible and Depth manages to perform close to the results given by FIR. Moreover, the fusion between Visible, Depth and FIR lowers the false positive rate by a factor for *thirty*, than just the use of FIR information.

As future work, we want to extend the analysis to include more datasets (like ETH [43]), along with a comparison of different new features. Moreover, in the multi-modalities experiments we have only treated the problem of pedestrian classification, but we plan of extending the analysis in a pedestrian detection framework.

There exist various approaches used for the task of pedestrian detection and classification task. In this thesis, we have showed that a multi-modality, multi-domain approach, and furthermore multi-feature, is essential for a good pedestrian classification system.

Bibliography

[1] The vislab intercontinental autonomous challenge. *http://viac.vislab.it/*, 2010.

[2] L. Andreone, F. Bellotti, A. De Gloria, and R. Lauletta. Svm-based pedestrian recognition on near-infrared images. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, pages 274–278. IEEE, 2005.

[3] A. Apatean, A. Rogozan, and A. Bensrhair. Objects recognition in visible and infrared images from the road scene. In *IEEE International Conference on Automation, Quality and Testing, Robotics, 2008*, volume 3, pages 327–332, 2008.

[4] Gregory P Asner and David B Lobell. A biogeophysical approach for automated swir unmixing of soils and vegetation. *Remote Sensing of Environment*, 74(1):99–112, 2000.

[5] Max Bajracharya, Baback Moghaddam, Andrew Howard, Shane Brennan, and Larry H Matthies. A fast stereo-based system for detecting and tracking pedestrians from a moving vehicle. *The International Journal of Robotics Research*, 28(11-12):1466–1485, 2009.

[6] Emmanuel P Baltsavias and Dirk Stallmann. *SPOT stereo matching for Digital Terrain Model generation.* Citeseer, 1993.

[7] Jasmine Banks and Peter Corke. Quantitative evaluation of matching methods and validity measures for stereo vision. *The International Journal of Robotics Research*, 20(7):512–532, 2001.

[8] Rodrigo Benenson, Radu Timofte, and Luc Van Gool. Stixels estimation without depth map computation. In *IEEE Conference on Computer Vision Workshops (ICCV Workshops)*, pages 2010–2017, 2011.

[9] Rodrigo Benenson, Markus Mathias, Radu Timofte, and Luc Van Gool. Pedestrian detection at 100 frames per second. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2903–2910, 2012.

[10] M. Bertozzi, A. Broggi, A. Fascioli, T. Graf, and M.M. Meinecke. Pedestrian detection for driver assistance using multiresolution infrared vision. *IEEE Transactions on Vehicular Technology*, 53(6):1666–1678, 2004.

[11] M Bertozzi, A Broggi, A Lasagni, and MD Rose. Infrared stereo vision-based pedestrian detection. In *Intelligent Vehicles Symposium*, pages 24–29. IEEE, 2005.

[12] M Bertozzi, A Broggi, M Felisa, G Vezzoni, and M Del Rose. Low-level pedestrian detection by means of visible and far infra-red tetra-vision. In *Intelligent Vehicles Symposium*, pages 231–236. IEEE, 2006.

[13] M Bertozzi, A Broggi, C Hilario Gomez, RI Fedriga, G Vezzoni, and M Del Rose. Pedestrian detection in far infrared images based on the use of probabilistic templates. In *Intelligent Vehicles Symposium*, pages 327–332. IEEE, 2007.

[14] Massimo Bertozzi, Emanuele Binelli, Alberto Broggi, and MD Rose. Stereo vision-based approaches for pedestrian detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pages 16–16. IEEE, 2005.

[15] Bassem Besbes, Alexandrina Rogozan, and Abdelaziz Bensrhair. Pedestrian recognition based on hierarchical codebook of surf features in visible and infrared images. In *IEEE Intelligent Vehicles Symposium (IV)*, pages 156–161, 2010.

[16] Bassem Besbes, Sonda Ammar, Yousri Kessentini, Alexandrina Rogozan, and Abdelaziz Bensrhair. Evidential combination of svm road obstacle classifiers in visible and far infrared images. In *Intelligent Vehicles Symposium*, pages 1074–1079. IEEE, 2011.

[17] Dinkar N. Bhat and Shree K. Nayar. Ordinal measures for image correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):415–423, 1998.

[18] E. Binelli, A. Broggi, A. Fascioli, S. Ghidoni, P. Grisleri, T. Graf, and M. Meinecke. A modular tracking system for far infrared pedestrian recognition. In *Intelligent Vehicles Symposium*, pages 759–764. IEEE, 2005.

[19] M. Bleyer and S. Chambon. Does Color Really Help in Dense Stereo Matching? In *Proceedings of the International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*. Citeseer, 2010.

[20] Michael Bleyer and Margrit Gelautz. Simple but effective tree structures for dynamic programming-based stereo matching. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 415–422, 2008.

[21] Michael Bleyer, Sylvie Chambon, Uta Poppe, and Margrit Gelautz. Evaluation of different methods for using colour information in global stereo matching approaches. *Int. Society for Photogrammetry and Remote Sensing*, pages 63–68, 2008.

[22] Alberto Broggi, Massimo Bertozzi, Alessandra Fascioli, and Massimiliano Sechi. Shape-based pedestrian detection. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 215–220. Citeseer, 2000.

[23] Alberto Broggi, Massimo Bertozzi, and Alessandra Fascioli. Self-calibration of a stereo vision system for automotive applications. In *IEEE International Conference on Robotics and Automation*, volume 4, pages 3698–3703, 2001.

[24] Matthew Brown and Sabine Susstrunk. Multi-spectral sift for scene category recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 177–184. IEEE, 2011.

[25] Alan Brunton, Chang Shu, and Gerhard Roth. Belief propagation on the gpu for stereo vision. In *The 3rd Canadian Conference on Computer and Robot Vision*, pages 76–76. IEEE, 2006.

[26] I. Cabani, G. Toulminet, and A. Bensrhair. A Fast and Self-adaptive Color Stereo Vision Matching; a first step for Road Obstacle Detection. In *Intelligent Vehicles Symposium*, pages 58–63. IEEE, 2006. ISBN 490112286X.

[27] Pietro Cerri, Luca Gatti, Luca Mazzei, Fabio Pigoni, and Ho Gi Jung. Day and night pedestrian detection using cascade adaboost system. In *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1843–1848. IEEE, 2010.

[28] S. Chambon and A. Crouzil. Colour correlation-based matching. *International Journal of Robotics and Automation*, 20(2):78–85, 2005. ISSN 0826-8185.

[29] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3): 273–297, 1995.

[30] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893, 2005.

[31] Navneet Dalal, Bill Triggs, and Cordelia Schmid. Human detection using oriented histograms of flow and appearance. In *European Conference on Computer Vision*, pages 428–441. Springer, 2006.

[32] James W Davis and Mark A Keck. A two-stage template approach to person detection in thermal imagery. In *WACV/MOTION*, pages 364–369. Citeseer, 2005.

[33] James W Davis and Vinay Sharma. Background-subtraction using contour-based fusion of thermal and visible imagery. *Computer Vision and Image Understanding*, 106(2):162–182, 2007.

[34] P. Dollár, Z. Tu, P. Perona, and S. Belongie. Integral channel features. *BMVC 2009, London, England*, 2009.

[35] Piotr Dollár, Serge Belongie, and Pietro Perona. The fastest pedestrian detector in the west. In *British Machine Vision Conference*, volume 55, 2010.

[36] Piotr Dollar, Christian Wojek, Bernt Schiele, and Pietro Perona. Pedestrian detection: An evaluation of the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(4):743–761, 2012.

[37] Nick A Drake, Steve Mackin, and Jeff J Settle. Mapping vegetation, soils, and geology in semiarid shrublands using spectral matching and mixture modeling of swir aviris imagery. *Remote Sensing of Environment*, 68(1):12–25, 1999.

[38] M Enzweiler, P Kanter, and DM Gavrila. Monocular pedestrian recognition using motion parallax. In *Intelligent Vehicles Symposium*, pages 792–797. IEEE, 2008.

[39] Markus Enzweiler and Dariu M Gavrila. Monocular pedestrian detection: Survey and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(12): 2179–2195, 2009.

[40] Markus Enzweiler and Dariu M Gavrila. A multilevel mixture-of-experts framework for pedestrian classification. *IEEE Transactions on Image Processing*, 20(10):2967–2979, 2011.

[41] Markus Enzweiler, Angela Eigenstetter, Bernt Schiele, and Dariu M Gavrila. Multi-cue pedestrian classification with partial occlusion handling. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 990–997. IEEE, 2010.

[42] S Erturk. Region of interest extraction in infrared images using one-bit transform. *Signal Processing Letters*, 2013.

[43] Andreas Ess, Bastian Leibe, and Luc Van Gool. Depth and appearance for mobile scene analysis. In *IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007.

[44] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. Liblinear: A library for large linear classification. *The Journal of Machine Learning Research*, 9: 1871–1874, 2008.

[45] Yajun Fang, Keiichi Yamada, Yoshiki Ninomiya, Berthold Horn, and Ichiro Masaki. Comparison between infrared-image-based and visible-image-based approaches for pedestrian detection. In *Intelligent Vehicles Symposium*, pages 505–510. IEEE, 2003.

[46] Pedro F Felzenszwalb and Daniel P Huttenlocher. Efficient belief propagation for early vision. *International journal of computer vision*, 70(1):41–54, 2006.

[47] Pedro F. Felzenszwalb, Ross B. Girshick, and David McAllester. Cascade object detection with deformable part models. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2241–2248, 2010.

[48] Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. *IEEE Transanctions on Pattern Analysis and Machine Intelligence*, 32(9):1627–1645, 2010.

[49] Clinton Fookes, A Maeder, Sridha Sridharan, and Jamie Cook. Multi-spectral stereo image matching using mutual information. In *2nd International Symposium on 3D Data Processing, Visualization and Transmission*, pages 961–968. IEEE, 2004.

[50] The Royal Society for the Prevention of Accidents. What are the most common causes of road accidents? `http://www.rospa.com/faqs/detail.aspx?faq=298`, 2013. Accessed: 2013-12-03.

[51] D.A. Forsyth and J. Ponce. *Computer vision: a modern approach*. Prentice Hall Professional Technical Reference, 2002.

[52] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

[53] Andrea Fusiello, Vito Roberto, and Emanuele Trucco. Efficient stereo with multiple windowing. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 858–863, 1997.

[54] Tarak Gandhi and Mohan M Trivedi. Pedestrian protection systems: Issues, survey, and challenges. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):413–430, 2007.

[55] Dariu M Gavrila and Stefan Munder. Multi-cue pedestrian detection and tracking from a moving vehicle. *International journal of computer vision*, 73(1):41–59, 2007.

[56] Andreas Geiger, Martin Roser, and Raquel Urtasun. Efficient large-scale stereo matching. In *Asian Conference of Computer Vision*, pages 25–38. Springer, 2011.

[57] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Computer Vision and Pattern Recognition (CVPR)*, Providence, USA, June 2012.

[58] David Geronimo, Antonio M Lopez, Angel Domingo Sappa, and Thorsten Graf. Survey of pedestrian detection for advanced driver assistance systems. *Transactions on Pattern Analysis and Machine Intelligence*, 32(7):1239–1258, 2010.

[59] Ross B Girshick, Pedro F Felzenszwalb, and David A Mcallester. Object detection with grammar models. In *Advances in Neural Information Processing Systems*, pages 442–450, 2011.

[60] Scott Grauer-Gray and Chandra Kambhamettu. Hierarchical belief propagation to reduce search space using cuda for stereo and motion estimation. In *Workshop on Applications of Computer Vision (WACV)*, pages 1–8. IEEE, 2009.

[61] Scott Grauer-Gray, Chandra Kambhamettu, and Kannappan Palaniappan. Gpu implementation of belief propagation using cuda for cloud tracking and reconstruction. In *IAPR Workshop on Pattern Recognition in Remote Sensing*, pages 1–4. IEEE, 2008.

[62] Erico Guizzo. How Google's Self-Driving Car Works. *http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works*, Retrieved 18 October 1011.

[63] Marc P Hansen and Douglas S Malchow. Overview of swir detectors, cameras, and applications. In *SPIE Defense and Security Symposium*, pages 69390I–69390I. International Society for Optics and Photonics, 2008.

[64] Simon Hermann and Reinhard Klette. Iterative semi-global matching for robust driver assistance systems. In *Proc. Asian Conf. Computer Vision, LNCS*, 2012.

[65] H. Hirschmuller and D. Scharstein. Evaluation of cost functions for stereo matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007. ISBN 1424411807.

[66] Heiko Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):328–341, 2008.

[67] Heiko Hirschmuller and Daniel Scharstein. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9):1582–1599, 2009.

[68] Heiko Hirschmüller, Peter R Innocent, and Jon Garibaldi. Real-time correlation-based stereo vision with reduced border errors. *International Journal of Computer Vision*, 47 (1-3):229–246, 2002.

[69] Asmaa Hosni, Michael Bleyer, Margrit Gelautz, and Christoph Rhemann. Local stereo matching using geodesic support weights. In *16th IEEE International Conference on Image Processing (ICIP)*, pages 2093–2096, 2009.

[70] M. Humenberger, C. Zinner, M. Weber, W. Kubinger, and M. Vincze. A fast stereo matching algorithm suitable for embedded real-time systems. *Computer Vision and Image Understanding*, 2010. ISSN 1077-3142.

[71] Sensors Inc. Why swir? what is the value of shorwave infrared? `http://www.sensorsinc.com/whyswir.html`, 2013. Accessed: 2013-10-12.

[72] Bongjin Jun, Inho Choi, and Daijin Kim. Local transform features and hybridization for accurate face and human detection. *Transactions on Pattern Analysis and Machine Intelligence*, pages 1423–1436, 2013.

[73] DS Kim and KH Lee. Segment-based region of interest generation for pedestrian detection in far-infrared images. *Infrared Physics & Technology*, 61:120–128, 2013.

[74] DS Kim, M Kim, BS Kim, and KH Lee. Histograms of local intensity differences for pedestrian classification in far-infrared images. *Electronics Letters*, 49(4):258–260, 2013.

[75] Andreas Klaus, Mario Sormann, and Konrad Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *Proceedings of the 18th International Conference on Pattern Recognition - Volume 03*, ICPR '06, pages 15–18. IEEE Computer Society, 2006. ISBN 0-7695-2521-0.

[76] Vladimir Kolmogorov and Ramin Zabih. Computing visual correspondence with occlusions using graph cuts. In *8th IEEE International Conference on Computer Vision*, volume 2, pages 508–515, 2001.

[77] J.Z. Kolter, Y. Kim, and A.Y. Ng. Stereo vision and terrain modeling for quadruped robots. In *IEEE International Conference on Robotics and Automation, 2009*, pages 1557–1564. IEEE, 2009.

[78] Kurt Konolige. Small vision systems: Hardware and implementation. In *ROBOTICS RESEARCH-INTERNATIONAL SYMPOSIUM-*, volume 8, pages 203–212. MIT PRESS, 1998.

[79] Sebastien Kramm and Abdelaziz Bensrhair. Obstacle detection using sparse stereovision and clustering techniques. In *Intelligent Vehicles Symposium (IV)*, pages 760–765. IEEE, 2012.

[80] Stephen J Krotosky and Mohan M Trivedi. On color-, infrared-, and multimodal-stereo approaches to pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems*, 8(4):619–629, 2007.

[81] Stephen J Krotosky and Mohan M Trivedi. A comparison of color and infrared stereo approaches to pedestrian detection. In *Intelligent Vehicles Symposium*, pages 81–86. IEEE, 2007.

[82] Raphael Labayrade, Didier Aubert, and J-P Tarel. Real time obstacle detection in stereovision on non flat road geometry through" v-disparity" representation. In *Intelligent Vehicle Symposium, 2002. IEEE*, volume 2, pages 646–651. IEEE, 2002.

[83] Lubor Ladickỳ, Paul Sturgess, Chris Russell, Sunando Sengupta, Yalin Bastanlar, William Clocksin, and Philip HS Torr. Joint optimization for object class segmentation and dense stereo reconstruction. *International Journal of Computer Vision*, pages 1–12, 2012.

[84] Bastian Leibe, Edgar Seemann, and Bernt Schiele. Pedestrian detection in crowded scenes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 878–885, 2005.

[85] Guoliang Li, Yong Zhao, Daimeng Wei, and Ruzhong Cheng. Nighttime pedestrian detection using local oriented shape context descriptor. In *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering*. Atlantis Press, 2013.

[86] J. Li, W. Gong, W. Li, and X. Liu. Robust pedestrian detection in thermal infrared imagery using the wavelet transform. *Infrared Physics & Technology*, 53(4):267–273, 2010.

[87] Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In *International Conference on Image Processing*, volume 1, pages I–900. IEEE, 2002.

[88] Qiong Liu, Jiajun Zhuang, and Jun Ma. Robust and fast pedestrian detection method for far-infrared automotive driving assistance systems. *Infrared Physics & Technology*, 2013.

[89] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[90] Jiangbo Lu, Gauthier Lafruit, and Francky Catthoor. Anisotropic local high-confidence voting for accurate stereo correspondence. In *Proc. SPIE-IS&T Electronic Imaging*, volume 6812, pages 605822–1, 2008.

[91] Mirko Mählisch, Matthias Oberländer, Otto Löhlein, Dariu Gavrila, and Werner Ritter. A multiple detector approach to low-resolution fir pedestrian recognition. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV2005), Las Vegas, NV, USA*, 2005.

[92] David Marr, Tomaso Poggio, Ellen C Hildreth, and W Eric L Grimson. A computational theory of human stereo vision. In *From the Retina to the Neocortex*, pages 263–295. Springer, 1991.

[93] Xing Mei, Xun Sun, Mingcai Zhou, Shaohui Jiao, Haitao Wang, and Xiaopeng Zhang. On building an accurate stereo matching system on graphics hardware. In *International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 467–474. IEEE, 2011.

[94] U. Meis, M. Oberlander, and W. Ritter. Reinforcing the reliability of pedestrian detection in far-infrared sensing. In *Intelligent Vehicles Symposium*, pages 779–783. IEEE, 2004.

[95] S. Meister, B. Jähne, and D. Kondermann. Outdoor stereo camera system for the generation of real-world benchmark data sets. *Optical Engineering*, 51(02):021107, 2012.

[96] A Miron, S Ainouz, A Rogozan, and A Bensrhair. Cross-comparison census for colour stereo matching applied to intelligent vehicle. *Electronics Letters*, 48(24):1530–1532, 2012.

[97] Sandino Morales and Reinhard Klette. Ground truth evaluation of stereo algorithms for real world applications. In *Computer Vision–ACCV 2010 Workshops*, pages 152–162. Springer, 2011.

[98] Karsten Mühlmann, Dennis Maier, Jürgen Hesser, and Reinhard Männer. Calculating dense disparity maps from color stereo images, an efficient implementation. *International Journal of Computer Vision*, 47(1-3):79–88, 2002.

[99] Harsh Nanda and Larry Davis. Probabilistic template based pedestrian detection in infrared videos. In *Intelligent Vehicle Symposium*, volume 1, pages 15–20. IEEE, 2002.

[100] Sergiu Nedevschi, Silviu Bota, and Corneliu Tomiuc. Stereo-based pedestrian detection for collision-avoidance applications. *IEEE Transactions on Intelligent Transportation Systems*, 10(3):380–391, 2009.

[101] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1): 51–59, 1996.

[102] Masatoshi Okutomi, Osamu Yoshizaki, and Goji Tomita. Color stereo matching and its application to 3-d measurement of optic nerve head. In *11th IAPR International Conference on Pattern Recognition*, pages 509–513. IEEE, 1992.

[103] Daniel Olmeda, Jose Maria Armingol, and Arturo de la Escalera. Discrete features for rapid pedestrian detection in infrared images. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3067–3072. IEEE, 2012.

[104] Daniel Olmeda, Cristiano Premebida, Urbano Nunes, Jose Maria Armingol, and Arturo de la Escalera. Pedestrian detection in far infrared images. *Integrated Computer-Aided Engineering*, 20(4):347–360, 2013.

[105] World Health Organization et al. Global status report on road safety 2013: supporting a decade of action. 2013.

[106] Gary Overett, Lars Petersson, Nathan Brewer, Lars Andersson, and Niklas Pettersson. A new pedestrian dataset for supervised learning. In *Intelligent Vehicles Symposium*, pages 373–378. IEEE, 2008.

[107] Constantine Papageorgiou and Tomaso Poggio. A trainable system for object detection. *International Journal of Computer Vision*, 38(1):15–33, 2000.

[108] Tomaso Poggio, Vincent Torre, and Christof Koch. Computational vision and regularization theory. *Image understanding*, 3(1-18):111, 1989.

[109] W. Richards. Stereopsis and stereoblindness. *Experimental Brain Research*, 10(4):380–388, 1970.

[110] Marcus Rohrbach, Markus Enzweiler, and Dariu M Gavrila. High-level fusion of depth and intensity for pedestrian classification. In *Pattern Recognition*, pages 101–110. Springer, 2009.

[111] Indranil Sarkar and Manu Bansal. A wavelet-based multiresolution approach to solve the stereo correspondence problem using mutual information. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 37(4):1009–1014, 2007.

[112] Ashutosh Saxena, Jamie Schulte, and Andrew Y Ng. Depth estimation using monocular and stereo cues. IJCAI, 2007.

[113] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 47(1):7–42, 2002.

[114] Jan Portmann Simon Lynen. Ethz thermal infrared dataset. *http://projects.asl.ethz.ch/datasets/doku.php?id=ir:iricra2014*, 2014.

[115] C Stentoumis, L Grammatikopoulos, I Kalisperakis, E Petsa, and G Karras. A local adaptive approach for dense stereo matching in architectural scene reconstruction. XL: XL–5/W1,219–226, 2013.

[116] F. Suard, A. Rakotomamonjy, A. Bensrhair, and A. Broggi. Pedestrian detection using infrared images and histograms of oriented gradients. In *Intelligent Vehicles Symposium*, pages 206–212. Ieee, 2006.

[117] Deqing Sun, Stefan Roth, and Michael J Black. A quantitative analysis of current practices in optical flow estimation and the principles behind them. *International Journal of Computer Vision*, 106(2):115–137, 2014.

[118] H. Sun, C. Hua, and Y. Luo. A multi-stage classifier based algorithm of pedestrian detection in night with a near infrared camera in a moving car. In *Proceedings Third International Conference on Image and Graphics*, pages 120–123. IEEE, 2004.

[119] Hao Sun, Cheng Wang, and Boliang Wang. Night vision pedestrian detection using a forward-looking infrared camera. In *International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM)*, pages 1–4. IEEE, 2011.

[120] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7):787–800, 2003.

[121] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, Aseem Agarwala, Marshall Tappen, and Carsten Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(6):1068–1080, 2008.

[122] Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision*, pages 839–846. IEEE, 1998.

[123] Jürgen Valldorf and Wolfgang Gessner. *Advanced microsystems for automotive applications 2005*. Springer Verlag, Berlin, June 2005. ISBN: 3540334092.

[124] W. van der Mark and D.M. Gavrila. Real-time dense stereo for intelligent vehicles. *Transactions on Intelligent Transportation Systems*, 7(1):38–50, 2006. ISSN 1524-9050.

[125] R.F. van der Willigen, W.M. Harmening, S. Vossen, and H. Wagner. Disparity sensitivity in man and owl: Psychophysical evidence for equivalent perception of shape-from-stereo. *Journal of vision*, 10(1), 2010.

[126] O. Veksler. Stereo correspondence by dynamic programming on a tree. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 384–390. IEEE, 2005. ISBN 0769523722.

[127] Paul Viola, Michael J Jones, and Daniel Snow. Detecting pedestrians using patterns of motion and appearance. In *9th IEEE International Conference on Computer Vision*, pages 734–741. IEEE, 2003.

[128] S. Walk, N. Majer, K. Schindler, and B. Schiele. New features and insights for pedestrian detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1030–1037. IEEE, 2010.

[129] Stefan Walk, Konrad Schindler, and Bernt Schiele. Disparity statistics for pedestrian detection: Combining appearance, motion and stereo. In *European Conference on Computer Vision*, pages 182–195. Springer, 2010.

[130] Jiabao Wang, Yafei Zhang, Jianjiang Lu, and Yang Li. Target detection and pedestrian recognition in infrared images. *Journal of Computers*, 8(4), 2013.

[131] Menghua Wang and Wei Shi. The nir-swir combined atmospheric correction approach for modis ocean color data processing. *Optics Express*, 15(24):15722–15733, 2007.

[132] Xiaoyu Wang, Tony X Han, and Shuicheng Yan. An hog-lbp human detector with partial occlusion handling. In *12th International Conference on Computer Vision*, pages 32–39. IEEE, 2009.

[133] Christian Wojek, Stefan Walk, and Bernt Schiele. Multi-cue onboard pedestrian detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 794–801, 2009.

[134] Q. Yang, L. Wang, R. Yang, S. Wang, M. Liao, and D. Nister. Real-time global stereo matching using hierarchical belief propagation. In *The British Machine Vision Conference*, pages 989–998, 2006.

[135] M. Yasuno, S. Ryousuke, N. Yasuda, and M. Aoki. Pedestrian detection and tracking in far infrared images. In *Proceedings Intelligent Transportation Systems*, pages 182–187. IEEE, 2005.

[136] Kuk-Jin Yoon and In-So Kweon. Locally adaptive support-weight approach for visual correspondence search. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 924–931. IEEE, 2005.

[137] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. *Computer Vision NECCV'94*, pages 151–158, 1994.

[138] Ke Zhang, Jiangbo Lu, and Gauthier Lafruit. Cross-based local stereo matching using orthogonal integral images. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(7):1073–1079, 2009.

[139] Li Zhang, Bo Wu, and Ram Nevatia. Pedestrian detection in infrared images based on local shape features. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[140] Liang Zhao and Charles E Thorpe. Stereo-and neural network-based pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems*, 1(3):148–154, 2000.