# An Approach for Features Matching Between Bilateral Images of Streo Vision System Applied for Automated Heterogeneous Platoon

**Mohammad Alfraheed**

**Department of Computer Science and Information Technology, Faculty of Science , Tafila Technical University, Jordan**

**E-mail: alfraheed@ttu.edu.jo**

**[ABSTRACT]**

Recently, the stereo vision system (SVS) has been developed in measuring tasks. Using SVS in measuring tasks plays an important role in automated highway system (AHS) because the SVS can be used instead of high cost distance sensors. The AHS is being developed to be run in different environments (i.e. unstructured and dynamic environment) and to form different shapes of platoon (i.e. heterogeneous platoon). In this work, a proposed method has been developed to improve the performance of the SVS in terms of automated heterogeneous platoon. The first stage of improvement has been here introduced by proposing a method for matching the left and right image of SVS (i.e. bilateral image). The significant contribution of the proposed method is to localize the points of interest in matching considering the shape of context assigned to the localized object of interest. The idea behind this development is to localize alternative connected points whenever the back view of the preceding vehicle BVPV (i.e. reference object) is influenced by the environment, including but not limited to sunshine reflection. In order to eliminate the surrounding objects of the BVPV, the Histogram of Oriented Gradient (HoG) has been developed by a proposed enhanced procedure. The latter depends on passing a precise knowledge about the position of edges and enhancing the gradient of intensity values. As for feature extraction, the proposed method has been developed to use the smoothed image generated by DoG instead of using the original image.

**[KEYWORDS]** *Stereo Vision System, Automated Heterogeneous Platoon, Features Matching*

# 1. INTRODUCTION

Recently, the Stereo Vision System (SVS) has been developed not only to monitor the wide range of the surrounding but also to measure the distance about the object of interest [1]. Today, the automated highway system (AHS) is being extended to be robust to withstand the influences of dynamic and unstructured environments (e.g. desert). Moreover, the AHS is being developed to adapt the heterogeneous vehicles (i.e. different shapes of vehicles) [2].

The SVS plays an important role in AHS through detecting, tracking, locating, and recognizing the heterogeneous. Measuring distance about the object of interest enables AHS to reduce the cost because the SVS could be used instead of other measuring devices (i.e. RIDAR and LIDAR).

However, the need for high accuracy prevents the SVS to be applied in automated heterogeneous platoon [1].

In SVS, the disparity image has to be accurate enough to generate the depth map associated with the object of interest. Alfraheed et al [3] [4] have introduced the back view of the proceeding vehicle (BVPV) as a reference object for the automated heterogeneous platoon. The problem arisen in their suggestion is that the number of the corresponding points between the left and right images of the SVS (i.e. bilateral images) is not high enough to generate an accurate depth map and disparity images. Starting from this challenge, the extracted features are used to increase the number of corresponding points by using them to match bilateral images. Thus, the corresponding points generated by SURF [5] have to be increased for calibrating SVS successfully.

Furthermore, these points have to be extracted in dynamic environments (e.g. hazy weather, dusty weather or etc ). Our significant contribution is to develop a robust and effective approach that is able to extract these points for Automated Highway Systems or Automated Vehicles in these environments.

The novelty introduced here is to improve the current work of the SVS in terms of the automated heterogeneous platoon. The SVS has to be more accurate once it is used to measure the distance of object of interest. In AHS, the reference object is located at different distances which vary from 3 meters to 10 meters. Therefore, the SVS can be installed in AHS instead of high cost distance measurement sensors. In this work, the first stage of improvement introduced aims to efficiently extract the corresponding points in context of feature matching instead of point matching. The extraction has to be also compatible with different distances of the object of interest.

The rest of the paper is organized as follows; after showing the introduction in section 1, the related work is given in section 2. The third section shows how the proposed method has been developed whilst the fourth section provides the results and discussions of the experiment. Finally, the conclusion is given in the last section.

# 2. RELATED WORK

In this section, the methods applied in this research paper were state of the art in feature extraction. Moreover, the literature review has been evaluated by highlighting the drawbacks of these researches in context of automated highway systems or automated heterogeneous platoon systems.

The feature extraction is one of the significant steps which enables the vision system to calculate the depth information and construct the object of interest in three-dimensional images (3D). The most of vision system are usually calibrated using chessboard (which is placed in front of the SVS) [6]. Some FPGA-based vision system deal with the embedded FPGA algorithms to reconfigure and calibrate the stereo images [7]. In automated heterogeneous platoon, the SVS is fixed on the front of the following vehicle in order to capture the BVPV. Once the platoon is electronically coupled, the SVS has to be calibrated based on the features of the BVPV. Practically, the platoon's system has to configure the SVS based on BVPV [1].

Instead of chessboard, Kim and Park [8] have presented a new stereo matching algorithm based on feature link matching. The latter utilizes the length and color

information of these features in stereo images. Their method is so effective to decide correct correspondence and to increase the accuracy of stereo matching. Moreover, the point of features are determined by FAST (Feature from Accelerate Segment Test) extractor. Despite of the successfully achieved results, the FAST extractor is not tested against the external effects (i.e. sunshine).

Therefore, Lin et al [9] have proposed a stereo matching algorithm based on the dynamic programming to reduce the noise which may lead to incorrect results during stereo matching. In addition, they have an assumption that the disparity of the scene is always smooth. However, this assumption is not always guaranteed in the dynamic environment.

In contrast, Wang Xiaoli et al [10] have proposed a binocular stereo matching algorithm which uses SURF (Speed Up Robust Features) to construct SURF detector pyramid [11]. The latter enables the proposed algorithm to improve the matching speed of the binocular stereo vision. In addition, their algorithm searches for a specific feature point on a specific scale layer generated by multi-scale analysis. As a consequence, the search range is greatly reduced. Since there are five constraints (i.e. threshold and symmetric constraints) to run successfully the algorithm, the latter is not applicable in automate and real time application.

Furthermore, the Genetic Algorithm (GA) has been developed in [12] for solving the stereo matching problem. The proposed approach has used the scan-line algorithm for the stereo matching problem. As a result, the accuracy has been increased approximately 20% compared to results without using GA. The problem arisen by GA is generated from the various components of the GA framework (i.e. crossover operation). The GA usually requires a crossover mechanism which often degrades the performance of the algorithm either by ignoring relevant data or by increasing the algorithm's time complexity significantly.

Having a point of interest is a significant step in this work. Several of methods have been presented to detect points for stereo vision [13] [14] [15]. The SIFT (the Save Invariant Feature Transform) [16] is generally used

to find the matching points in stereo vision. Here, the SIFT is used to localize the point of interest in the PVPV. Alhwarin et al [16] have presented a novel method to speed up SIFT feature matching. They have extended SIFT feature by a few pair wise independent angles, which are invariant to rotation, scale and illumination changes. Practically, when taking a picture from two different angles, the color of a certain pixel is changed due to lighting conditions. The dynamic environment of heterogeneous automated platoon generally causes the color changing problem. Thus, the aim of this work is not only to speed up the feature matching but also to increase the accuracy of matching.

Recently, Zhang et al [11] have proposed an efficient image matching algorithm based on SURF. This approach has enormous advantages of less computations and short time-consuming. Moreover, the Random Sample Consensus algorithm is used to eliminate the false match and wrong match points. In this work, their proposed approach is developed in terms of the automated heterogeneous platoon and in dynamic environment.

## 3. PROPOSED METHOD

Within this work, a proposed approach is introduced to improve the performance of stereo image matching in context of automated heterogeneous platoon. Generally, the image matching has been divided into two significant tracks, which are image matching based on image value and feature based image matching [17]. Since the application of this work relates to the real time applications, the second track of the image matching (based on the features) is chosen to develop the proposed approach.

The popular approach feature based image matching has been introduced in [18] as a SURF (Speed Up Robust Features) approach. Within the later, the points of interest are extracted based on the scale-and rotation invariant [11]. In other words, the FAST-Hessian detector is used to localize the points of interest. The Haar wavelet is used to generate a descriptor of the localized points and to extract features associated with them. The idea to extract the features in SURF is similar to SIFT (Scale Invariant Features Transform). In SIFT,

the different scale image is applied using the Gaussian function to localize the point of interest [16]. Although both approaches are looking for the strongest points of interest, they do not take the shape context in their account. Some points are discarded because they are not available in the lower scale of the scale invariant.

Recently, Zhang et al [11] have proposed an efficient image matching algorithm based on SURF. That approach has enormous advantages of

less computations and short time-consuming. Moreover, the Random Sample Consensus algorithm is used to eliminate the false matches and wrong match points. Although the successful results of their proposed approach due to eliminating the false matching, the approach has not been applied in dynamic environment and it does not take the shape of interest in its account.

Since having points of interest is a significant step in stereo image matching, the proposed approach is developed to extract points associated with edges of BVPV. The latter is considered here as a shape of interest in order to improve the efficiency of the proposed matching approach. Localizing the BVPV is necessary to keep the proposed approach running. Therefore, the BVPV has been detected and tracked in a video stream [4] [3].

The idea behind this development is first to use BVPV as a reference point in bilateral images of SVS. This would be especially beneficial since the BVPV enables the SVS to restrict its functions to a particular part of frame instead of processing the whole frame. Therefore, the running time of the image matching is reduced [1]. Secondly, the proposed approach is required to increase the number of points used in the stereo image matching. Increasing the number of points enables the approach to be robust and effective against the influences of the dynamic environment of the automated heterogeneous platoon (e.g. shadow, dust, sunshine reflection, hazy weather ..., etc.). Once some points of interest are eliminated in the stereo image matching due to either false matching or color changing problem, other new points are undertaken to image matching instead of those eliminated points. Since the extracted points are

associated with the edges of BVPV, they are clearer than the eliminated points and close to them.

The proposed approach is required to match the BVPV located on the left and right images of SVS. The approach is supplemented by the detection and tracking agents [4] [3] to extract key points required to match the stereo images in video stream. Therefore, the proposed approach has to be also applicable under real time constraints. Additionally, the dynamic environment represents a challenge for localizing key points because it often changes the appearance of BVPV. Therefore, the agent has to be adaptable in dynamic environment.

The significant contribution of this work is to localize the points of interest in the image matching considering the shape of context assigned to the localized object of interest. This would be especially beneficial since it enables the SVS to recognize its reference object (i.e. the Back View of Preceding Vehicle BVPV) for calibration process. Meaning that, the number of the points of interests is increased whenever the BVPV is moved away from the stereo cameras (i.e. the distance between BVPV and SVS is approximately 10 – 12 meters).

Furthermore, the shape of context enables the proposed approach to localize the points of interest in the normal scale of the Gaussian function (i.e. $\sigma = 1$) instead of looking for the point of interest at the lower scale (i.e. $\sigma = 6$). Although the localized points at lower scale are stronger than those at the normal scale, the shape of context offers the connected points of interest which enables the proposed approach to localize alternative points whenever the reference object is exposed to environmental influences (e.g. sunshine reflection, shadow ... etc). In the Following, the main steps of the proposed approach are presented.

## 4. CONCLUSION

Within this work, the mismatching of the corresponding points of the Stereo Vision System (SVS) has been addressed in terms of the automated highway heterogeneous platoon. The proposed method has been developed for improving the matching results of the

SVS. Firstly, the proposed method depends on localizing the object of interest. In this work, the Back View of the Preceding Vehicle (BVPV) is taken as a reference object for SVS. The BVPV is then processed to extract its edges based on the Difference-of- Gaussian (DoG). The extracted edges of BVPV are enhanced to ensure that their points are connected. In order to eliminate the surrounding objects of the BVPV, the Histogram of Oriented Gradient (HoG) has been developed by a proposed enhanced procedure. The latter depends on passing a precise knowledge about the edges position and enhancing the gradient of intensity values. Furthermore, the step of the feature extraction has been developed to use the smoothed image generated by DoG instead of using the original image. As a consequence, the effects of the illumination, the distraction of the gradient and distraction of the local intensity are reduced. Once the features are extracted, the matching step is applied. The similarity value of each enhanced block is calculated based on the Euclidean Distance. The latter has been used because it is efficiently applied not only in image analysis but also in pattern recognition. For each enhanced block, if the similarity value is greater than 99%, the block is a perfect match. If the similarity value is greater than 99% and there are more than two corresponding enhanced blocks, a shifting procedure has been applied to align those blocks. The proposed method distinguishes itself from other methods through its ability to increase the corresponding points between bilateral images of SVS. In addition, it can be adaptable with the dynamic environment and heterogeneous platoon. The limitation of this study is to run the proposed method in the dynamic environment. However, the proposed method is not tested in other environments (e.g. hazy weather).

**REFERENCES:**

[1]  M. Alfraheed, A. Dröge, M. Klingender, D. Schilberg, and S. Jeschke, "A mechanism to improve Stereo Vision Systems in automated heterogeneous platoons," in 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2012, pp. 425–432.

[2]  M. Alfraheed, A. Dröge, M. Klingender, D. Schilberg, and S. Jeschke, "Longitudinal and Lateral Control in Automated Highway Systems: Their Past, Present and Future," in Intelligent Robotics and Applications, 2011, pp. 589–598.

[3] M. Alfraheed, A. Dröge, D. Schilberg, and S. Jeschke, "Automated heterogeneous platoons in unstructured environment: Real time tracking of a preceding vehicle using video stream," in 2014 5th International Conference on Information and Communication Systems (ICICS), 2014, pp. 1–6.

[4] M. Alfraheed, A. Dröge, R. Kunze, M. Klingender, D. Schilberg, and S. Jeschke, "Real time detection of the back view of a preceding vehicle for automated heterogeneous platoons in unstructured environment using video," in 2011 IEEE International Conference on Systems, Man, and Cybernetics, 2011, pp. 549–555.

[5] Y. Bai, L. Zhuo, B. Cheng, and Y. F. Peng, "Surf feature extraction in encrypted domain," in 2014 IEEE International Conference on Multimedia and Expo (ICME), 2014, pp. 1–6.

[6] S. Arya, D. M. Mount, "Approximate Nearest Neighbor Queries in Fixed Dimensions", Open Journal, No. 1, May. 2018.

[7] S. Jin et al., "FPGA Design and Implementation of a Real-Time Stereo Vision System," IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 1, pp. 15–26, Jan. 2010.

[8] C. I. Kim and S. Y. Park, "Fast Stereo Matching of Feature Links," in Visualization and Transmission 2011 International Conference on 3D Imaging, Modeling, Processing, 2011, pp. 268–274.

[9] J. Lin, D. Yan, X. Hu, Q. Xing, and B. Yang, "Dynamic programming algorithm for stereo correspondence of contour," in 2012 5th International Congress on Image and Signal Processing, 2012, pp. 866–870.

[10] W. Xiaoli, Y. Lei, W. Lirong, and X. Jing, "Characteristic Point Match Algorithm Based on the SURF in Binocular Stereo Vision," in 2012 Fifth International Conference on Intelligent Networks and Intelligent Systems, 2012, pp. 302–305.

[11] B. Zhang, Y. Jiao, Z. Ma, Y. Li, and J. Zhu, "An efficient image matching method using Speed Up Robust Features," in 2014 IEEE International Conference on Mechatronics and Automation, 2014, pp. 553–558.

[12] E. Kiperwasser, O. David, and N. S. Netanyahu, "A Hybrid Genetic Approach for Stereo Matching," in Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation, New York, NY, USA, 2013, pp. 1325–1332.

[13] L. Trujillo and G. Olague, "Automated Design of Image Operators that Detect Interest Points," Evol. Comput., vol. 16, no. 4, pp. 483–507, Dec. 2008.

[14] T. Khan, M. Biglari-Abhari, G. Gimel'farb, and J. Morris, "Fast Point-of-interest Detection from Real-time Stereo," in Proceedings of the 27th Conference on Image and Vision Computing New Zealand, New York, NY, USA, 2012, pp. 79–84.

[15] L. Trujillo, G. Olague, E. Lutton, and F. Fernández de Vega, "Multiobjective Design of Operators That Detect Points of Interest in Images," in Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, New York, NY, USA, 2008, pp. 1299–1306.

[16] F. Alhwarin, D. Ristić–Durrant, and A. Gräser, "VF-SIFT: Very Fast SIFT Feature Matching," in Pattern Recognition, 2010, pp. 222–231.

[17] B. Zhang, Y. Jiao, Z. Ma, Y. Li, and J. Zhu, "An efficient image matching method using Speed Up Robust Features," in 2014 IEEE International Conference on Mechatronics and Automation (ICMA), 2014, pp. 553–558.

[18] B. Zhang, Y. Jiao, Z. Ma, Y. Li, and J. Zhu, "An efficient image matching method using Speed Up Robust Features," in 2014 IEEE International Conference on Mechatronics and Automation (ICMA), 2014, pp. 553–558.

[19] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.

[20] D. Liu and J. Yu, "Otsu Method and K-means," in Proceedings of the 2009 Ninth International Conference on Hybrid Intelligent Systems - Volume 01, Washington, DC, USA, 2009, pp. 344–349.

[21] M. M. H. Daisy, S. TamilSelvi, and L. Prinza, "Gray scale morphological operations for image retrieval," in 2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET), 2012, pp. 571–575.

[22]  GONZALEZ, DIGITAL IMAGE PROCESSING USING MATLAB 2E. Tata McGraw-Hill Education, 2009.

[23]  Z. Huijuan and H. Qiong, "Fast image matching based-on improved SURF algorithm," in 2011 International Conference on Electronics, Communications and Control (ICECC), 2011, pp. 1460–1463.

[24]  N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, vol. 1, pp. 886–893 vol. 1.

[25]S. Chen, J. Li, and X. Wang, "A Fast Exact Euclidean Distance Transform Algorithm," in 2011 Sixth International Conference on Image and Graphics, 2011, pp. 45–49.

[26]  C. Pornpanomchai and A. Phaisitkulwiwat, "Fingerprint Recognition by Euclidean Distance," in 2010 Second International Conference on Computer and Network Technology, 2010, pp. 437–441.

[27]  VisLab - Dipartimento di Ingegneria dell'Informazione - Parco Area delle Scienze - Università di Parma, "VisLab | Extend Your Vision," VisLab, 17-Oct-2017. [Online]. Available: http://vislab.it/. [Accessed: 17-Oct- 2017].

**CONTRIBUTORS:**
[1] 4, 구민준, 국민대학교

.