An application of mutual information based stereo matching for ADAS

Akhil Appu Shetty[1], V I George[2], C Gurudas Nayak[3], Raviraj Shetty[4]
[1,2,3,4]*Department of Mechanical and Manufacturing Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India*

Abstract: An important aspect of Advanced Driver Assistance Systems (ADAS) is the collision avoidance mechanism. The use of stereo matching techniques to attain this objective can prove to be very beneficial. The disparity maps obtained through stereo matching are a direct indication of the distance of the various objects present in the immediate vicinity. The data extracted from these disparity maps can be used for vehicular control. A data cost comprising of mutual information and a modified census transform is proposed for the generation of disparity maps. The proposed methodology involves converting the image section under consideration into a binary string which can be used for further processing. For this purpose, a new polyvalent census transform is proposed and its performance on stereo image data sets is discussed.
Keywords – Stereo Matching, Census Transform, Disparity maps, Middlebury Stereo Dataset, KITTI dataset.

I. INTRODUCTION

With the conception of smart cities, Intelligent Transport Systems (ITS) becomes an indispensable component of commute for citizens. In any city, mobility is a key concern- be it going to school, college, office or for any other purpose the citizens use the transport system. The aim of ITS is to achieve traffic efficiency through the minimization of traffic issues. Users are provided with prior information regarding the traffic, automated vehicular control in situations that demand it, real time running information regarding local convenience, availability of seat etc. which improves the safety and comfort of the passengers and also aids to reduce the time of commute.

An important component of ITS is the Advanced Driver Assistance System or ADAS. These systems utilizes the information acquired through various sensors embedded in the vehicle to alert the driver of imminent collisions or obstacles in the immediate vicinity, hence preventing hazardous situations for the driver. [1]

The research conducted in [2] ascertained that most of the major causes for on road vehicle accidents are due to the delay as well as untimely application of brakes by the driver along with inadequate level of braking torques generated in hazardous situations. The European New Car Assessment Program (Euro NCAP) has identified that the delayed braking intervention is due to the driver's lack of concentration and attention as well as general human error [3]. Research which have been conducted in the vehicle safety field to diminish these concerns suggested the introduction of Advanced Driver Assistance Systems (ADAS) [4]. The prominent areas of usage of ADAS are Collision Avoidance [5] [6], Blind Spot Monitoring [7] and Autonomous Emergency Braking (AEB) [8].

Among various methodologies used to detect the obstacles in ADAS systems, one prominent method is the use of cameras in the form of a stereo vision system embedded into the vehicle [9]. As the images are captured through the cameras, they can be processed to generate the disparity maps. These disparity maps provide us with enough data to estimate the distance of the various objects in the vicinity of the vehicle. This information can be used to generate control signals for the vehicle itself.

Stereo matching is a process which extracts depth information of a scene, in the form of a disparity map, from two or more images taken from different positions. This concept is recently being used in many research areas which include advanced driver assistance systems (ADAS), intelligent robotic systems and 3-D object recognition which are well applicable in real world scenarios. Nonetheless, stereo matching methodologies face many real world problems while extracting accurate 3D data, especially in situations common in real-world outdoor scenarios such as presence of shadows, sudden depth discontinuities and radiometric variations [10].

There are some applications of stereo matching which do not require extremely accurate results but the frame rate of the generated disparity maps is of considerable importance. One such application is to use a stereo camera setup embedded on a vehicle and use the generated disparity maps for further processing. Such applications which do not need highly accurate disparity maps, adopt a local approach of stereo matching, rather than global matching techniques [11]. The local methods consider a small area of the image, process it and then move on to the next segment. By processing the data in this manner, the execution speed of local algorithms is much faster than global methods. Taking this into consideration, the research work presented in this article will consider local matching methods for obtaining dense disparity maps rather than global methodologies.

The work presented in this research paper describes a methodology to generate disparity maps from a stereo camera setup which can be utilized for collision avoidance in ADAS systems.

II. LITERATURE REVIEW

Since the problem of stereo matching has been well worked on, many local methodologies have been previously proposed.

The sum of absolute differences (SAD), simply takes the absolute difference between the reference template and the matching target and calculates the sum [12]. Though the results are obtained much faster than other local methods, the SAD is heavily dependent on the intensities of the pixels under consideration as it directly uses them for its processing. This can lead to results that are extremely inaccurate in the presence of even the slightest difference in intensity variation.

Cost metrics like normalized cross correlation (NCC) and mutual information (MI) are more tolerant to such variations as compared to SAD. While NCC normalizes the pixel intensities by subtracting the average of the pixels from each pixel, MI calculates the appropriate match through probabilities of occurrences of the intensity values [13].

III. PROPOSED METHODOLOGY

The mutual information between two elements is presenters as:

$$mutual\ information\ (a,b) = entropy(a) + entropy(b) - joint\ entropy\ (a,b)$$

$$entropy(a) = -\sum_{i=0}^{n} p(i)\log(p(i))$$

$$joint\ entropy(a,b) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(i,j)\log(p(i,j))$$

Another metric which has been used repeatedly for stereo matching is census transform (CT). The advantage that this cost metric has over the others, is speed. CT is comparatively faster than all the metrics mentioned previously.

This method transforms the pixels within a window and then performs the required match using the hamming distance between the transformed template pixels and the transformed target pixels. The transformation equations are mentioned below [14].

$$CT(p_1,p_2) = \begin{cases} 1, & p_1 > p_2 \\ 0, & otherwise \end{cases}$$

In the above equation p_2 is the central pixel in the window and p_1 is the pixel under consideration. This would transform the template and the target in to two binary strings of equal lengths. The similarity between these strings can be calculated through hamming distance between them. The target-template pair with the least hamming distance would be the best match.

This would be able to compensate, up to a certain extent, the problems mentioned previously. To further enhance the results, the CT can be used in conjunction with other cost metrics. The target and template can be transformed through CT and the similarity can be calculated using the cost metrics instead of hamming distance. Though this might sound feasible, the binary nature of the transform restricts its usage in very useful and strong cost metrics like MI since it uses probability rather than intensities of the pixels to calculate a match. Since there are only two bins, (0 and 1) in the CT, mutual information will not have a lot of data to calculate the appropriate probabilities which would result in an erroneous match.

Hence to overcome this difficulty, a polyvariate census transform (PCT) is proposed. It is defined as

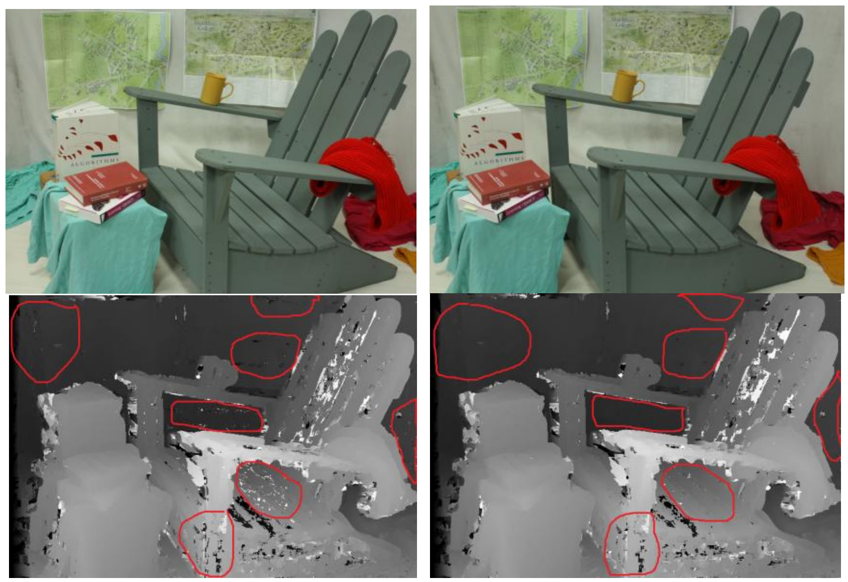
$$PCT(p_1,\bar{p}) = \begin{cases} p_1, & p_1 > \bar{p} \\ \bar{p}, & otherwise \end{cases}$$

In the above equation, p_1 is the pixel under consideration and $\bar{p}$ is the mean of all the pixels in the window being considered. All the pixels whose values are equal to or below the mean value are replaced with the mean and the remaining are unchanged.

As it can be observed, since the final transformed data is no longer a string of just binary values, MI or other probability based matching metrics can hence be used as a matching cost on this transformed data.

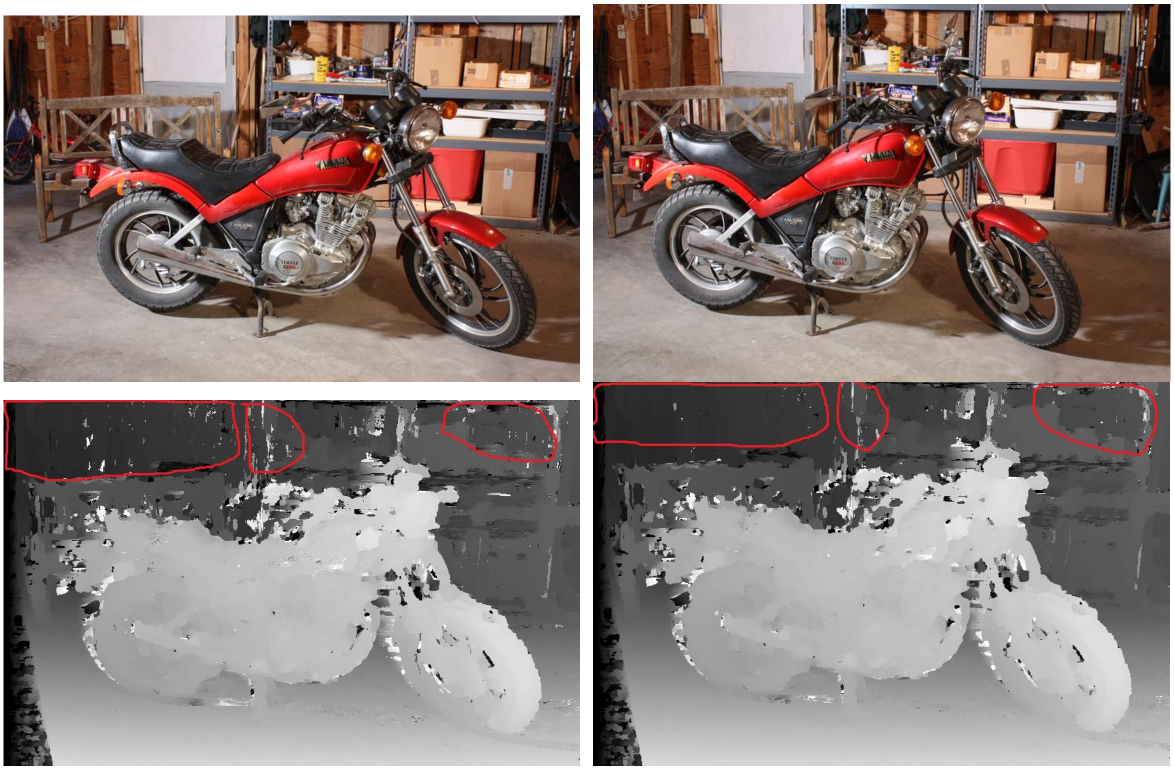
IV. RESULTS

For the purpose of testing the feasibility of the proposed algorithms, images from the Middlebury stereo dataset [6] were considered, which is one of the most frequently used data set for stereo matching researchers. Sample datasets and the results of applying the traditional census transform as well as the proposed method are presented in figure 1.

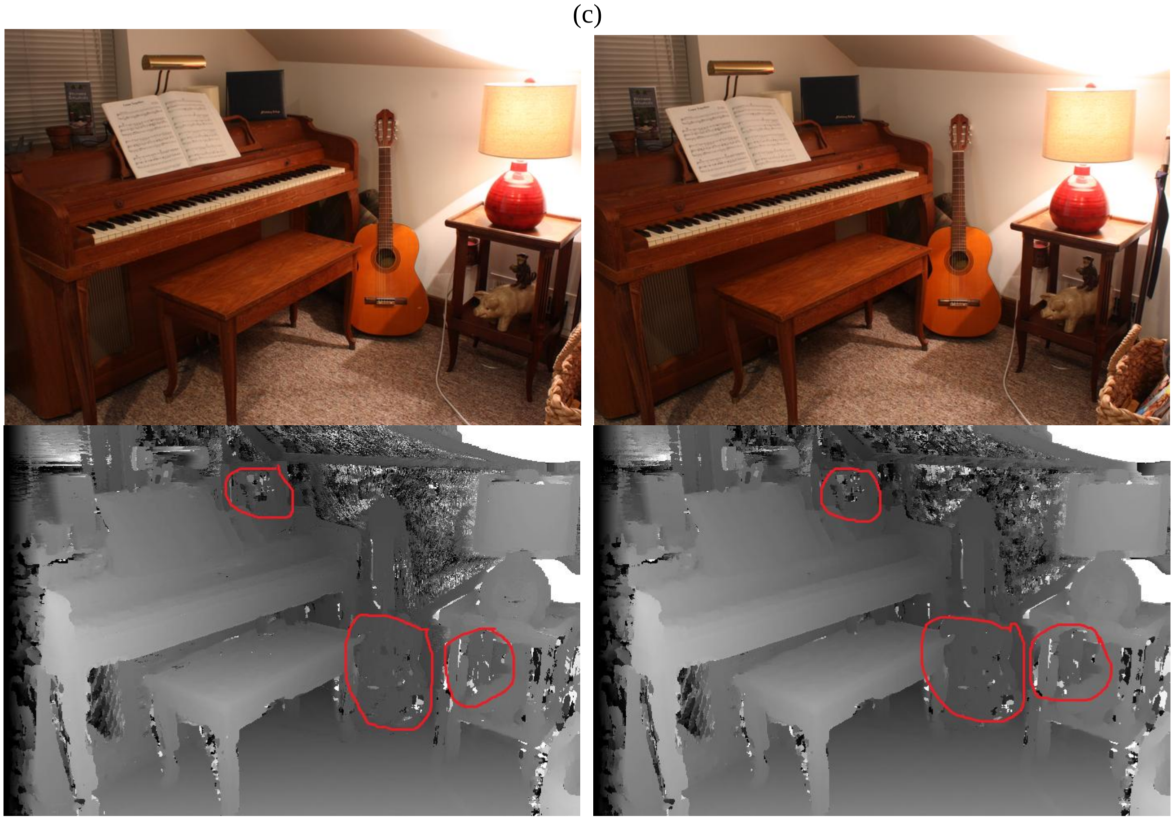


(a)

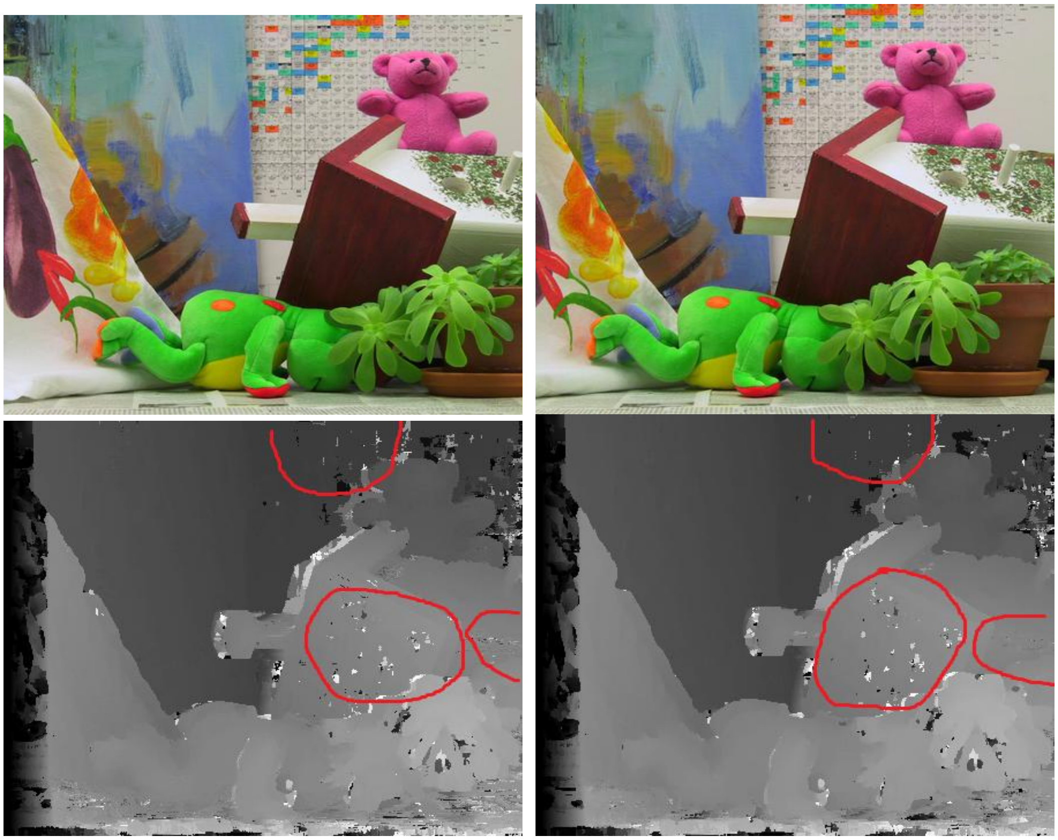
(b)



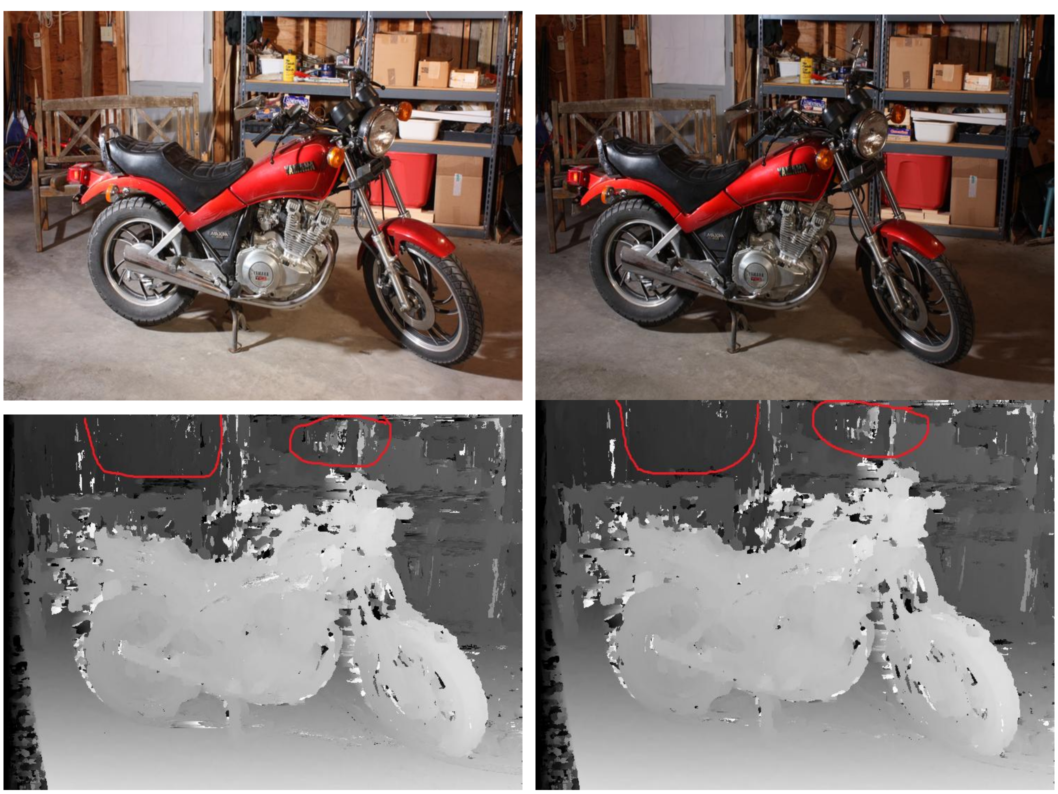
(c)



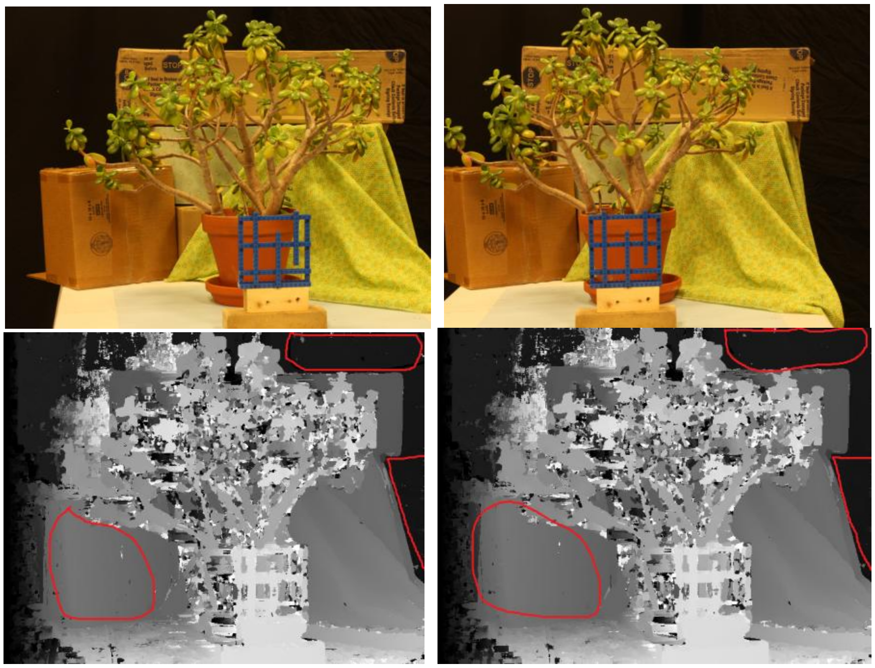
(d)



(e)



(f)



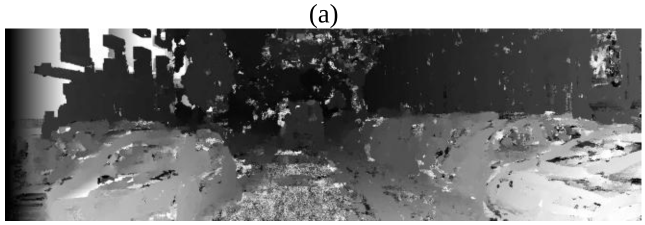
(g)

Figure 1. Top images are the original left and right image pairs with the bottom left image being the result of the original CT method and the bottom right being the result of the proposed PCT method for (a) Adirondack (b) Recycle (c) Motorcycle (d) Piano (e)Teddy (f) MotorcycleE (g) Jadelant respectively [15].

It can be observed from the results that the proposed PCT combined with MI provided better results when compared to the combination of MI and traditional CT. Hence the assumption, that increasing the number of bins would improve the results when using MI, has been proved.

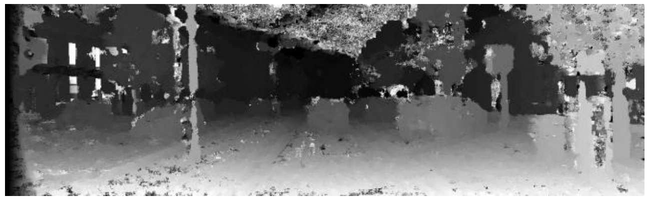
Once satisfied that the proposed algorithm functions as expected, tests on real life out door images were conducted. For this purpose, images from KITTI data set [16] were utilised. The results are presented in figure 2.

(a)



(b)

(c)



(d)

Figure 2. Left image samples from KITTI dataset [16] (a) & (c). The respective disparity mps (b) & (d).

Upon observing the results obtained by using the KITTI dataset, it is understood that the disparity maps obtained provide a considerable idea of the obstacles present in the vicinity of the cameras. These results can be further processed and the extracted information can be used to generate control signals to avoid collisions in ADAS environments.

V. CONCLUSION

In the presented paper, a methodology for generating dense disparity maps is presented. The authors have discussed the combination of MI and a modified CT as a cost metric for stereo matching. The traditional CT method produces only binary data which cannot be used with probability based matching methods. The proposed CT method generates a multi-valued data from census transform. This can be used in conjunction with probability based matching methods like mutual information, which will result in an enhanced disparity map. As observed from the results, the proposed PCT provided better result as compared to the traditional CT. These disparity maps can be further used for collision avoidance in driver assistance systems.

VI. ACKNOWLEDGMENTS

This research work was funded by the Dr. T.M.A. Pai PhD scholarship program by the Manipal Academy of Higher Education, Manipal. The authors would like to thank the department of Instrumentation & Control Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, for allowing the research to be conducted in their laboratory and providing the required facilities.

VII. REFERENCES

[1] Umar Zakir Abdul, Hamid; et al. (2017). "Autonomous Emergency Braking System with Potential Field Risk Assessment for Frontal Collision Mitigation". 2017 IEEE Conference on Systems, Process and Control (ICSPC). Retrieved 14 March 2018.

[2] J. J. Breuer, A. Faulhaber, P. Frank, and S. Gleissner, “Real world safety benefits of brake assistance systems,” in 20th International Technical Conference on the Enhanced Safety of Vehicles (ESV), no. 07-0103, 2007.

[3] M. Paine, D. Paine, C. Newland, and S. Worden, “Encouraging safer vehicles through enhancements to the ncap rating system,” in 22nd International Technical Conference on the Enhanced Safety of Vehicles (ESV), no. 11-0107, 2011.

[4] U. Z. A. Hamid, K. Pushkin, H. Zamzuri, D. Gueraiche, and M. A. A. Rahman, “Current collision mitigation technologies for advanced driver assistance systems–a survey,” PERINTIS eJournal, vol. 6, no. 2, 2016.

[5] U. Z. A. Hamid, M. H. M. Ariff, H. Zamzuri, Y. Saito, M. A. Zakaria. M. A. A. Rahman, and P. Raksincharoensak, “Piecewise trajectory replanner for highway collision avoidance systems with safe-distance based threat assessment strategy and nonlinear model predictive control,” Journal of Intelligent & Robotic Systems, pp. 1–23, 2017.

[6] U. Z. A. Hamid, H. Zamzuri, P. Raksincharoensak, and M. A. A. Rahman, “Analysis of vehicle collision avoidance using model predictive control with threat assessment.” in 23rd ITS World Congress (Melbourne). ITS World Congress, 2016.

[7] F. R. A. Zakuan, H. Zamzuri, M. A. A. Rahman, W. J. Yahya, N. Hazima, F. Ismail, M. S. Zakariya, K. A. Zulkepli, and M. ZulfaqarAzmi, “Threat assessment algorithm for active blind spot assist system using short range radar sensor,” ARPN Journal of Engineering and Applied Sciences, vol. 12, pp. 4270–4275, 2017.

[8] B. Fildes, M. Keall, N. Bos, A. Lie, Y. Page, C. Pastor, L. Pennisi, M. Rizzi, P. Thomas, and C. Tingvall, “Effectiveness of low speed autonomous emergency braking in real-world rear-end crashes,” Accident Analysis & Prevention, vol. 81, pp. 24–29, 2015.

[9] Gehrig, Stefan, and Uwe Franke. "Stereovision for ADAS." Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort (2014): 1-25.

[10] H. Hirschmüller, D. Scharstein., “Evaluation of stereo matching costs on images with radiometric differences”, IEEE Transactions on Pattern Analysis, pp. 1582-1599, 2008.

[11] A Hosni, M. Bleyer, M. Gelautz, C. Rhemann, “ Local stereo matching using geodesic support weights”, in Int. Conf. on Image Proc., IEEE. pp. 2093-2096, 2009.

[12] M. Baydoun, M.A. Al-Alaoui, “Enhancing stereo matching with varying illumination through histogram information and normalized cross correlation”, in IEEE Systems, Signals and Image Processing, IEEE. pp. 5-9, 2013.

[13] G. Egnal, “Mutual Information as a Stereo Correspondence Measure”, Technical Report, University of Pennsylvania, Philadelphia, PA, USA, 2000.

[14] H.W. Jo, B. Moon, “A modified census transform using the representative intensity values”, in IEEE SoC Design Conf., Gyungju, South Korea, pp. 309-310, 2015.

[15] http://vision.middlebury.edu/stereo/data/ (Last accessed 10/02/2019)

[16] http://www.cvlibs.net/datasets/kitti/ (Last accessed 10/02/2019)