





Performance analysis of a fuzzy disparity selector for stereo matching of image segments under radiometric variations

Akhil Appu SHETTY¹ , Vadakekara Itty GEORGE¹ , Chempi Gurudas NAYAK^{1,*} ,
Raviraj SHETTY² 

¹Department of Instrumentation and Control Engineering, Manipal Institute of Technology,
Manipal Academy of Higher Education, Manipal, India

²Department of Mechanical and Manufacturing Engineering, Manipal Institute of Technology,
Manipal Academy of Higher Education, Manipal, India

Received: 02.09.2019

Accepted/Published Online: 10.03.2020

Final Version: ..2020

Abstract: Stereo matching algorithms generate disparity maps, which contain the depth information of the environment, from two or more images of a scene taken from different viewpoints. The process of obtaining dense disparity maps is a problem which is still being actively researched. The presence of radiometric differences in the images only further complicates the stereo matching problem. In the present research work, the images are initially split into small patches of pixels, such that pixels in each patch have similar intensities. The authors attempt to study the effect of the parameters, namely, tuning parameter ‘ α ’ and the number of segments, while the images are subjected to variations in exposure and illumination. The value ‘ α ’ performs the function of a weight signifying the contribution of each data cost, when the two data costs are combined in a linear fashion. Lastly, the results of this methodology are compared with other methods that try to tackle the problem of stereo matching under radiometric variations.

Key words: Stereo matching, fuzzy logic, Middlebury stereo dataset, radiometric variations, multiple data costs.

1. Introduction

The area of dense stereo matching is a research field in computer vision that is still being extensively worked upon. In applications which require visual cues, for example, virtual reality and autonomous driving, stereo matching is beneficial, especially when it can generate dense depth information about the surroundings [1]. Even for the navigation of unmanned vehicle in a real environment, an accurate and dense disparity map can be of tremendous benefit. A smoothness assumption is used in most dense stereo matching methods. This assumption states that the pixels that correspond have similar colour values or intensities. However, in a real environment, this particular assumption might not necessarily hold, and there would be differences in the radiometric conditions in the input images. These differences in radiometric conditions, which cause the most complications in performing a dense stereo match between the input images, are the variations in illumination and exposure conditions [2]. These factors severely degrade the performance of most of the dense stereo matching algorithms. Hence, to prevent significant errors in the matching process, these stereo matching methods are applied to images taken under the same lighting conditions and with the same camera settings.

The process of stereo matching methods to produce dense results usually includes the following four stages: (1) matching cost computation, (2) cost aggregation, (3) disparity computation/optimization, and

*Correspondence: cg.nayak@manipal.edu

(4) disparity refinement. The abovementioned stereo matching methods can be segregated into two general categories, namely, global and local methods [3], based on whether the second step is included or not. In the case of local algorithms, aggregation of the matching costs are performed over varying or fixed support windows, depending on certain factors such as the intensities of the pixels in the image or colour values. They usually make assumptions about implicit smoothness through aggregation. A winner takes all strategy (WTA) is used to select the best disparity. In the case of global methods, assumptions are based on explicit smoothness. Complex global functions are used, which utilize the disparity estimated at the remaining other pixels, to determine the disparity at one pixel [4]. The main distinction between local and global methods is the usage of an energy minimization function. Typically, local methods are comparatively faster than global methods but do not usually produce reliable matches in areas which lack texture or in an area where discontinuities in depth could be found. Their counterpart, the global methods, are more expensive from a computational point of view and require much larger memory, but they can generate highly accurate results. The methodology proposed in this research work falls into the category of local algorithms.

The authors have used multiple data costs, followed by a fuzzy logic-based disparity selection method to generate the disparity map. The main contribution of this research work is the fuzzy-based disparity selection method, the data cost metric developed by combining multiple common data costs, and the study of the effect of tuning of parameters used in the data cost metrics on different radiometric conditions in the images. The authors propose a feature-based matching methodology as opposed to a deep learning-based approach. The main reason for this decision is the fact that most deep learning methods demand a large amount of processing power [5–7]. This will be an extremely limiting factor if environments with restricted resources are considered where the resources are not abundant. In such situations of limited processing power, it is desirable to opt for a more traditional methodology for stereo matching. Hence, with the intention that the proposed work should be implementable even in such demanding environments, the authors followed a traditional feature-based methodology.

In the rest of the paper, Section 2 reviews the literature in this area of research, followed by the proposed methodology in Section 3. The discussions on the results obtained and the conclusion is presented in Sections 4–6.

2. Literature review

The first significant hurdle any dense stereo matching algorithm has to overcome is to determine the most appropriate matching cost [8]. Conventionally, the methods used for the computation of matching costs are based on individual pixels, for instance, absolute differences (AD) of intensity [9] and squared difference of intensities [10]. Authors have also adopted window-based methodologies, for instance, sum of absolute differences (SAD), normalized cross-correlation (NCC), sum of squared differences [11], or common filter-based, for example, mean filtering, bilateral filtering and Laplacian of Gaussian filtering [12]. Some authors have even made use of Rank and Census, which are nonparametric based matching methods [13].

Due to their advantages, stereo matching algorithms have found increased use in real-world scenarios. Hence, the attention of algorithm development has transformed towards matching cost methods that are insensitive to and able to overcome the differences in radiometric conditions between the input images. Even though several such diverse matching cost approaches have been worked upon and developed, most of these methods can only achieve reasonable results when the difference in radiometric conditions between the right and left images is trivial. A combination of AD and census transform-based matching cost methods were also

considered to overcome this problem [14]. However, because colour information was utilized in the matching cost step, these methodologies, again, only work when the input images have minor radiometric differences.

Another instance of using multiple data costs led to the combination of census transform, gradient matching, and AD for matching cost computation [15]. The disparity selection was made on a WTA scheme. Further refinement of the disparity map was carried out through the left-right consistency check and hole filling as postprocessing. Authors claim that since they use multiple cost functions, their method overcomes the problems faced by a single method. Multiple data costs involving truncated AD of colour and gradient, mutual information (MI), census transform, and zero mean normalized cross-correlation produced reliable results [16, 17]. An adaptive support weight was used for cost aggregation and input images are segmented at different levels using a graph-based segmentation method. The results of segmentation further aids the cost aggregation method to generate an even better disparity map.

Methodologies involving preprocessing of input images have also been followed where the input images were segmented [18] into superpixels. The proposed algorithm's idea is that the pixels belonging to each segment belong to that particular object of the segment. This infers that the disparities inside each segment are similar. Each superpixel centers are matched using a census-based measure. A colour cross algorithm is proposed in this paper. The results show that using a segmentation method (not very precise), like simple linear iterative clustering (SLIC) superpixels, it is possible to get good results.

Authors also proposed a matching cost method, which is invariant to radiometric changes, named adaptive NCC (ANCC) [19]. The methodology adopted here extracts colour invariant information and utilizes the colour formation model from the input images taken under varying radiometric settings. Even though the results of the matching through this method are robust and dense, the window size used for aggregation in the major factor that determines the speed and performance of this method. It was further indicated that MI-based matching cost methods could be made use of to dampen the effect of radiometric differences between images [20].

Mutual information (MI) has been observed to handle radiometric differences well [21], but its performance depends heavily on the size of the window. It is a parametric matching method that uses the intensity of the pixels for processing and matching as compared to nonparametric methods, which use the local ordering of the pixel intensities for processing/matching. The authors further claim that if two signals are similar, then the joint entropy of the two signals will be minimum as the uncertainty between two signals is low. When joint entropy is minimized, the MI will increase. Hence for similar signals, the MI is higher. In the case of images, two constant regions will also have low joint entropy. In such cases, to avoid spurious matches, it becomes necessary to maximize the individual entropies of the samples to be compared. Hence in such cases, MI will be useful. Further work on MI [22, 23] stresses on the fact that MI is a very good candidate for stereo matching even in multimodal matching. This data cost can also easily be combined with other parameters such as gradients to produce an even more enhanced result.

Research towards the use of a maximum a posteriori estimation [24] based stereo matching to compute the disparity of the pixels in different radiometric changes was also presented. Initially, a prior on the disparities of a few pixels is built. The points selected for this purpose needs to be such that they can be matched with high certainty. Such points are selected using DAISY descriptor as this was proven to produce good results even in the presence of radiometric variations. Following this, a disparity plane was constructed using three such points whose disparities are accurately determined. Windows of 17×17 was considered during this research work. A combination of census transform and hamming distance was used as an initial cost value. The cost aggregation was carried out using a self-adaptive window strategy. The window consists of binary weights

wherein the weights are assigned to pixels based on the mean of the pixels in the window. The final cost aggregation is the convolution between this weights window and the census cost value calculated previously. A full conditional probability for a certain disparity value is applied where the prior values calculated previously are utilized and the WTA scheme of disparity selection is used to find the appropriate disparity value of the pixel in consideration. The obtained disparity was further refined using a left-right consistency check, a median filter and interpolating unknown missing disparities using piecewise constant function on the smallest valid nearest neighbor in the same image line.

3. Methodology

Given a pair of stereo images, many traditional stereo matching methods use a single data cost metric to determine the disparity map such as absolute intensity difference and squared intensity differences, sum of AD, sum of squared differences and normalized cross-correlation (NCC). The main problem area the authors are working on is the situation of varying radiometric conditions between the two stereo images under consideration. In such conditions, the performance of the traditional stereo matching algorithms tends to deteriorate.

Over the years, it has been observed from literature [25–28] that using multiple data costs provides a better solution to the problem of computing accurate disparities. Inspired by these findings, the authors propose a stereo matching metric comprising of multiple data costs to overcome the abovementioned problem. Research work done towards comparison between MI and NCC [29] claims that MI is a very useful metric that can be used for stereo matching. It is noted that MI and NCC can handle radiometric differences well, but its performance depends heavily on the size of the window. Another data cost, SAD, is also mentioned, but it has shown to produce a larger error in situations of radiometric changes.

Even though SAD, when used without any modifications, produces a large error, it had been observed through experiments that using SAD on the gradient of the images produces a better disparity map. Moreover, when compared to other methods, SAD is not as computationally complex, which is an advantage from the point of view of resources and execution time. Hence in this work, these three data costs have been considered [namely, normalized MI (NMI), NCC, and SAD of gradients of the images] and created a new metric comprising of a combination of the above-mentioned methods.

Once the matching cost metric has been decided, the next step in dense stereo matching is cost aggregation. Most of the methods use a fixed size square window for this step. It has been observed through experiments and in the literature that using a fixed window size is not a good choice as the window tends to include regions with depth discontinuities. Within a window, the depth of the objects are assumed to be smooth, i.e., all the objects inside a window have the same depth. The calculation of disparity of a pixel within a window would be correct if the abovementioned condition is met. The presence of different depths in the same window gives rise to erroneous disparities as the calculations would take into consideration even those pixels which have different depth and are still included in the window. An alternate method would be to use a window whose size is not fixed. Such a window should be able to change its shape depending on the properties of the pixels under observation. One such method is based on segmentation. Here, the reference image is initially segmented and the segmented regions are used as blocks for matching in the target image. The assumption with this methodology is that the segmented blocks belong to the same region and pixels in the same region have similar depth.

In the proposed method, the authors have segmented the input images into patches [30] such that each patch would consist of pixels belonging to the same object(region). To overcome the problem of windows having

pixels of different depths, these segmented regions are considered as windows and stereo matching is performed using these newly obtained variable sized windows. The result of the segmentation of one of the images in the dataset is shown in Figure 1. As can be observed, pixels in each segment belong to the same region and would hence have similar depth. After the calculation of cost values for each disparity, the disparity location which has the best cost is found.

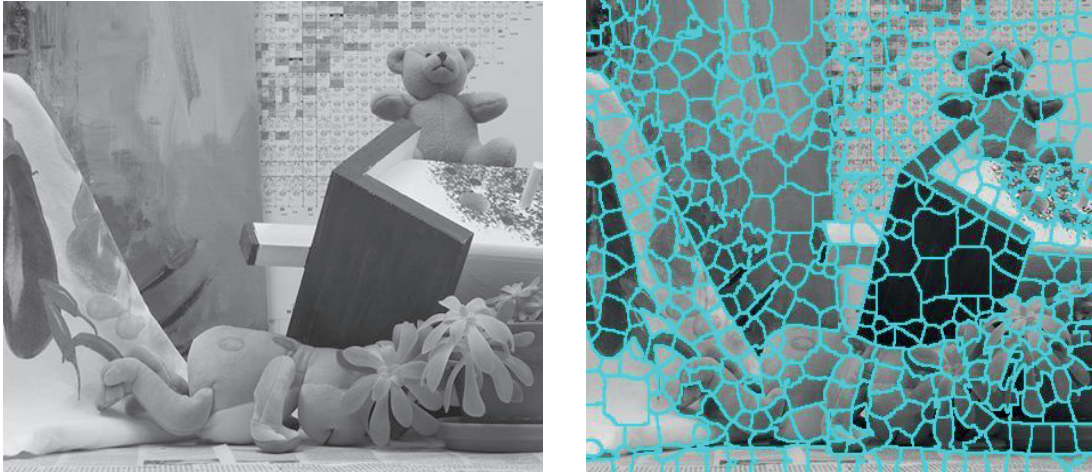


Figure 1. Original image (left) and result of segmenting image into patches (right).

This selection of disparity is usually performed by a WTA scheme. Among all the disparity positions, the location with the best cost is selected as the true disparity value. There is always a possibility that a cost function might produce multiple maximums/minimums while calculating the cost. This leads to much ambiguity. When a WTA based method is used in the presence of such ambiguities, the results tend to be erroneous. Keeping this problem in mind, an “intelligence” based disparity selection method is used wherein, instead of directly applying WTA on the results of the datacosts, the datacosts are fed into the fuzzy logic module to generate a value which is obtained through the combination of the two mentioned datacosts and apply the WTA on the result of this fuzzy logic module. The proposed methodology is represented as shown in Figure 2. The present work was conducted using a fuzzy inference system (FIS) mentioned in Figure 3. The main attribute of the FIS is the rule base. The output is decided based on these rules formulated in an “if-then” setup [31].

The first input to the FIS is cost function 1 and the second input is cost function 2. The data cost metrics use for this research are NMI, NCC and SAD of gradients as described in [32]. Cost function 1 and cost function 2 were defined as per Eq. (1) and Eq. (2).

$$\text{Cost function 1} = \alpha * \text{NMI} + (1 - \alpha) * \text{SAD of Image Gradients} \quad (1)$$

$$\text{Cost function 2} = \text{NCC} \quad (2)$$

The selection of which data costs should be used with which input to the fuzzy system was done based on experimentation. The positions of the three data costs (namely NMI, NCC, and SAD of image gradient) were interchanged in the Eq. (1) & Eq. (2) and the results were compared. Based on these results, it was concluded that the mentioned combinations gave the best results. A similar methodology was followed for the selection of

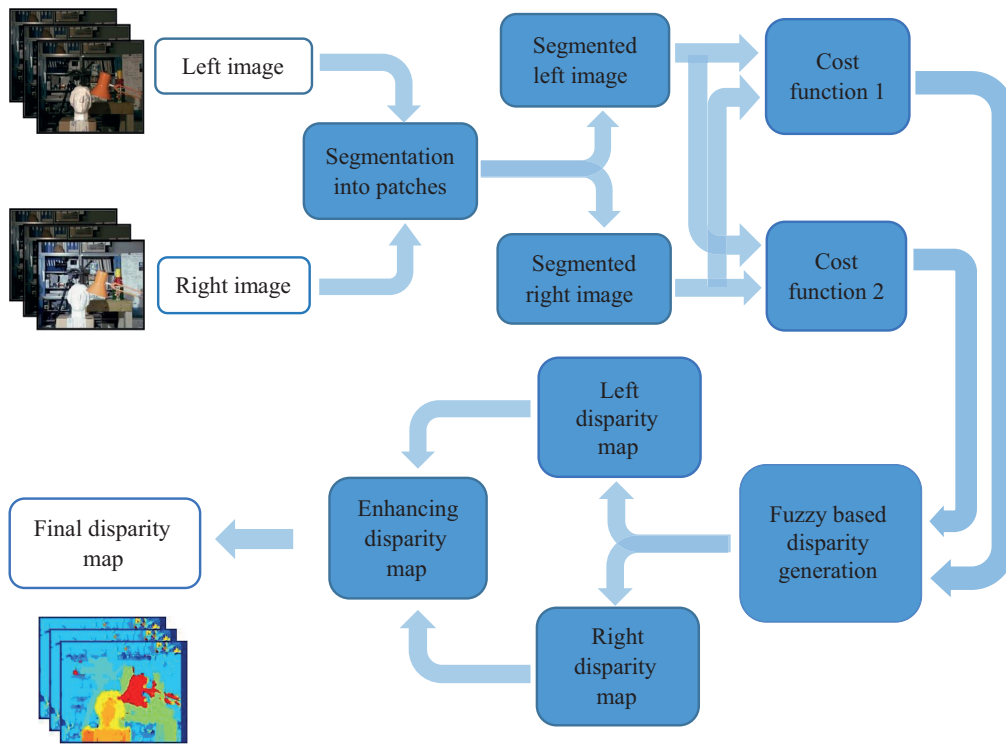


Figure 2. Proposed algorithm.

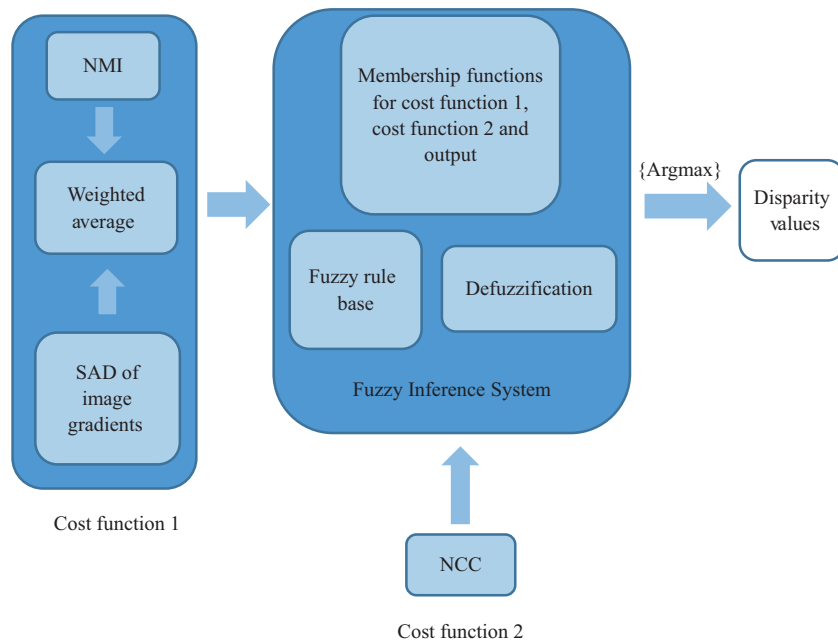


Figure 3. Fuzzy logic based disparity generation module.

the number of membership functions for each input and the number of fuzzy rules to be used. The categories of the membership functions for cost function 1, cost function 2, and the output were defined as per Eq. (3),

Eq. (4), and Eq. (5) respectively.

$$Cost\ function1 : \{Poor, Very\ bad, Bad, Average, Good, Very\ good\} \tag{3}$$

$$Cost\ function2 : \{Poor, Very\ bad, Bad, Average, Good, Very\ good\} \tag{4}$$

$$Output : \{Worst, Poor, Very\ bad, Bad, Below\ average, Acceptable, Average, Above\ average, Good, Very\ good, Extremely\ good, Great, Superb, Excellent, Best\} \tag{5}$$

The rules of the FIS are mentioned in Table 1. The membership functions of input 1 and input 2 are shown in Figures 4(a) and 4(b) respectively while the membership function of output is shown in Figure 5.

Table 1. Rules of the fuzzy inference system.

Rule number	Cost function 1	Cost function 2	Output	Rule number	Cost function 1	Cost function 2	Output
1	Poor	Poor	Worst	19	Average	Poor	Very bad
2	Poor	Very bad	Worst	20	Average	Very bad	Very bad
3	Poor	Bad	Very bad	21	Average	Bad	Below average
4	Poor	Average	Bad	22	Average	Average	Average
5	Poor	Good	Below average	23	Average	Good	Good
6	Poor	Very good	Below average	24	Average	Very good	Very good
7	Very bad	Poor	Poor	25	Good	Poor	Bad
8	Very bad	Very bad	Very bad	26	Good	Very bad	Below average
9	Very bad	Bad	Very bad	27	Good	Bad	Acceptable
10	Very bad	Average	Bad	28	Good	Average	Above average
11	Very bad	Good	Acceptable	29	Good	Good	Extremely good
12	Very bad	Very good	Above average	30	Good	Very good	Superb
13	Bad	Poor	Poor	31	Very good	Poor	Below average
14	Bad	Very bad	Very bad	32	Very good	Very bad	Acceptable
15	Bad	Bad	Bad	33	Very good	Bad	Acceptable
16	Bad	Average	Below average	34	Very good	Average	Great
17	Bad	Good	Above average	35	Very good	Good	Excellent
18	Bad	Very good	Above average	36	Very good	Very good	Best

The block diagram of the fuzzy-based disparity selection method is shown in Figure 3. As discussed previously, both left and right disparity maps are subjected to the mentioned fuzzy-based methodology. Hence both the images are subjected to segmentation. The abovementioned cost functions are applied to both the preprocessed images and the data is fed into the fuzzy-based disparity selector. This resulting left and right disparity maps are used for left-right consistency checking to obtain the final refined disparity map. For every segment in the reference image, the data cost for all the considered disparities are calculated using the cost functions mentioned above and are fed into the FIS. The output of the FIS will be a scalar value for each pair

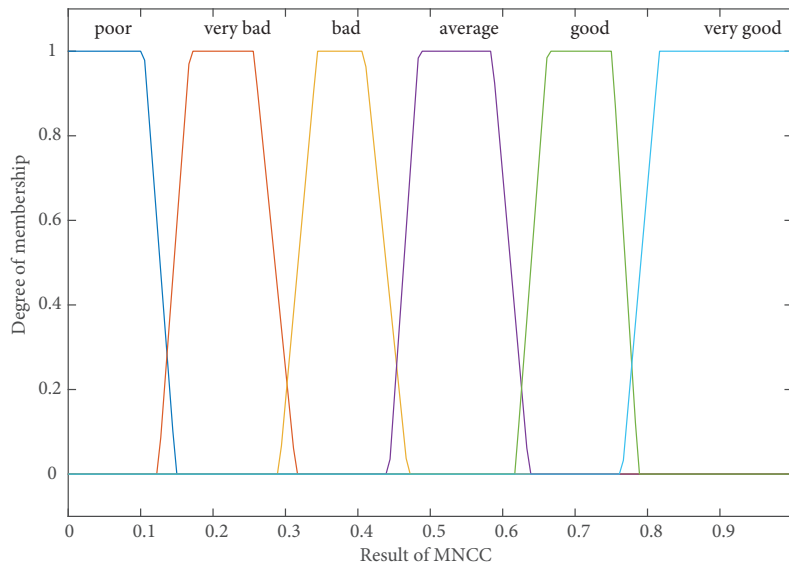


Figure 4. Membership functions of input 1 (left) and input 2 (right).

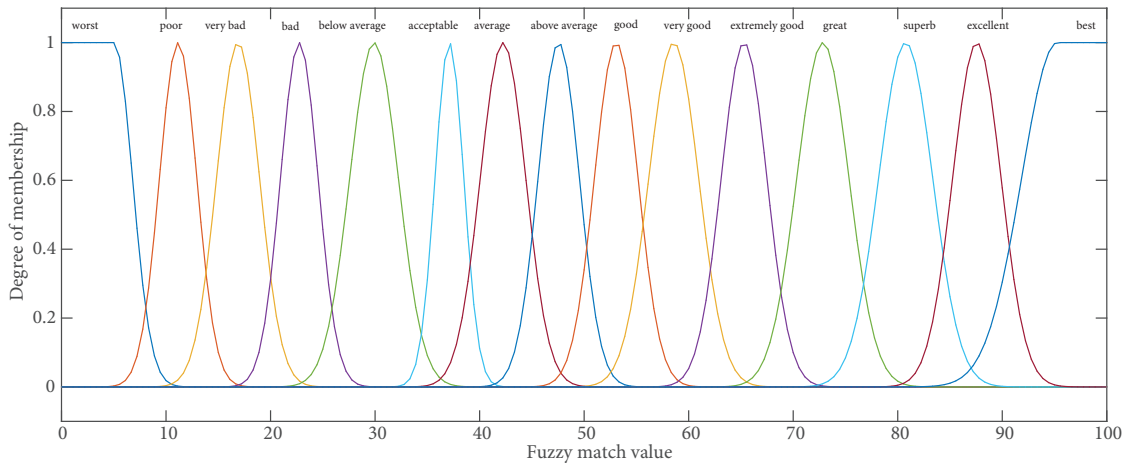


Figure 5. Membership function of output.

of inputs fed into it. Hence, for N disparity values, N scalar-valued outputs would be obtained. The WTA can now be applied to these set of values to obtain a more accurate value of disparity. The procedure is repeated for both the left and the right images to obtain two disparity maps corresponding to the left and right images. These images are then subjected to a left-right consistency check to improve the result further. The concept of left-right consistency check [33] requires both the left reference disparity map and the right reference map. It utilizes the following idea: if d_l and d_r are the left and right disparity maps, then a point (x, y) is said to be consistent only if the condition in Eq. (6) is met:

$$d_l(x, y) = d_r(x, y - d_l(x, y)) \tag{6}$$

4. Results

The images used for this experiment were obtained from the Middlebury stereo vision benchmark, which is a standard platform for researchers working in this area. The number of segments considered for this paper were 500, 600, 700 and 800. It was found experimentally that as the number of segments was reduced below 500 or increased beyond 800, the errors kept increasing. Most fluctuations were observed in the interval between 500–800 segments.

The experiments were conducted under similar conditions as observed in literature. For the case of varying exposure conditions, the light source illumination was set to *Illum1* and the exposure conditions were changed. For the case of varying illumination conditions, the exposure of the light source was set to *exp2* and the illumination was varied. The effect the number of segments and the value of ' α ' (varied from 0.1 to 0.9) have on the various illumination and exposure conditions are discussed in this section. The error parameter used here is the percentage of bad pixels. A pixel in the obtained disparity is considered to be a bad pixel if the difference between this pixel and the corresponding pixel in the ground truth is greater than a set threshold (in the experiments, the threshold is set to 1). The horizontal axis describes the radiometric conditions of the images used for our experiments. For example, the exposure condition 1/2 represents that for this particular situation, the exposure of left image is set to condition 1 and the exposure of right image is set to condition 2. The same holds true for illumination variation conditions. The image pairs as well as the levels of illumination and exposure conditions were considered as mentioned in the Middlebury dataset used for this research work.

4.1. Varying exposure conditions

From Figure 6(a), it is evident that errors for each category of segments are somewhat similar. This particular image does not have much depth discontinuities. Hence the variations in the number of segments do not have a significant effect on the results. In Figure 6(b), the tuning factor ' α ' has a significant effect in situations of lower exposure conditions, but when exposure is increased, the effect is somewhat constant. This shows that for conditions involving low exposure, a higher value of ' α ' is better. Since ' α ' gives the contribution of NMI in the data cost, thus indicating that NMI had better performance. But in cases of higher exposure situations, lower ' α ' is better, indicating SAD of gradients is a better choice in higher exposure conditions. Figure 7(a) again explains the effect of the number of segments appears to be similar at almost every radiometric condition. Similar to the Aloe dataset, Baby2 also has a comparatively lower depth discontinuity and hence the effect of the number of segments is also similar.

In Figure 7(b), it is observed that for higher exposure situations (1/1 and beyond) as long as there is a significant contribution of NMI, the errors would be more or less similar. The depth discontinuities in this image is slightly higher than Aloe. NMI has a stronger effect even in both lower and higher exposure conditions. Figure 8(a) displays the results on the 'Dolls' image pair. There are many objects in the image. This presence of multiple objects results in large depth discontinuities. Because of the presence of significant depth discontinuities in the image, in general, better results are obtained for a larger number of segments as increasing the number of segments means the pixels in the segments will have a larger chance of belonging to the same object. In Figure 8(b), as mentioned previously, in the presence of depth discontinuities, a higher value of ' α ' produces a better result there by indicating that NMI is a good cost metric while dealing with depth discontinuities. The tuning parameter ' α ' functions as a weighing parameter that indicates the contribution of SAD of gradients and the NMI datacosts when they are combined in the form of a linear equation. Higher value of ' α ' indicates a larger contribution of NMI in the overall final data cost, while its lower value tells us that the contribution of SAD of

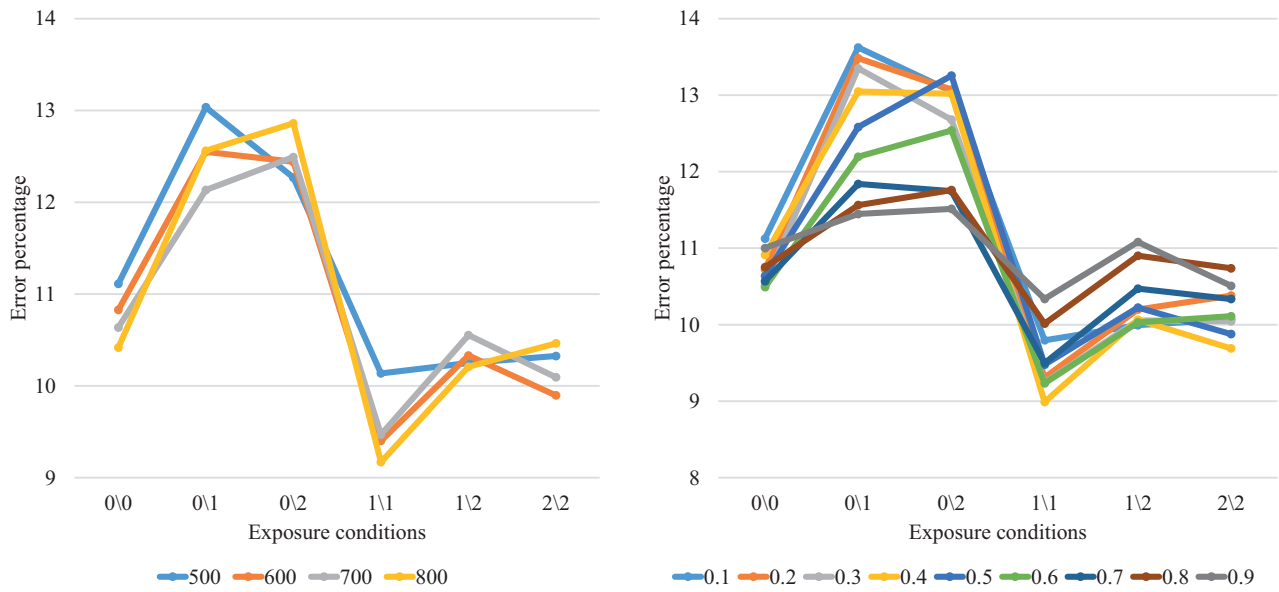


Figure 6. Effects of number of segments (left) and values of α (right) on different exposure conditions for “Aloe” dataset.

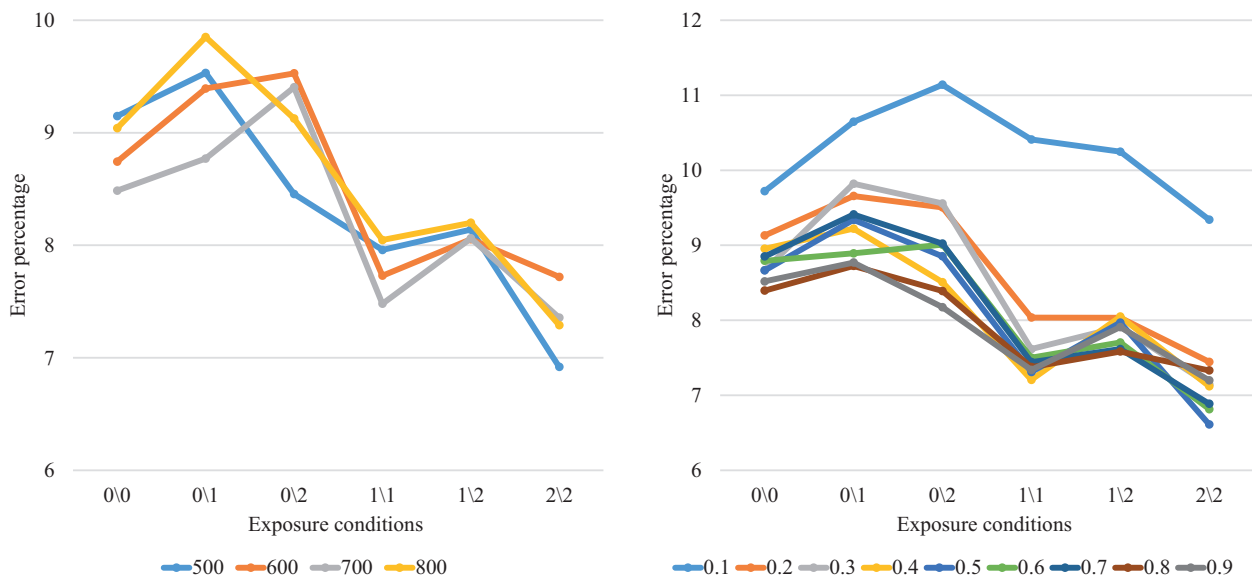


Figure 7. Effects of number of segments (left) and values of α (right) on different exposure conditions for “Baby2” dataset.

gradients data cost is larger when compared to NMI. When the depth discontinuities in the image is large, SAD of image gradients generates higher errors. For this reason, when the value of ‘ α ’ is lower (indicating higher contribution of SAD of image gradients), the errors tend to be larger.

In Figure 9(a), similar to Figure 8(a) there are large depth discontinuities. Better results are obtained for a larger number of segments as increasing the number of segments means the pixels in the segments will have a higher probability of belonging to the same object. From Figure 9(b), it can be deduced that in the presence

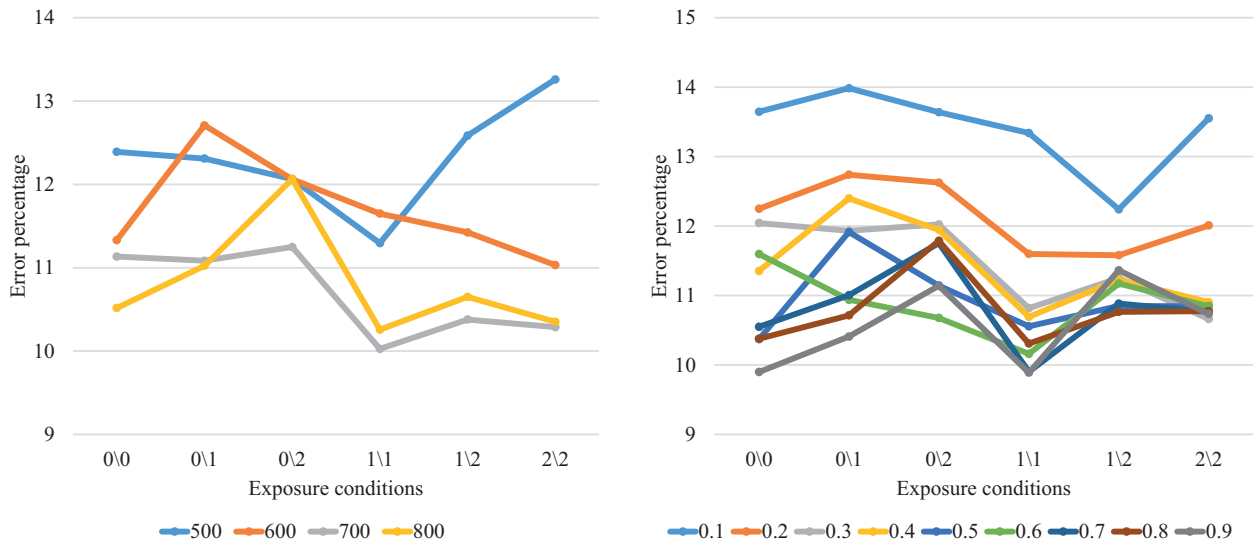


Figure 8. Effects of number of segments (left) and values of α (right) on different exposure conditions for “Dolls” dataset.

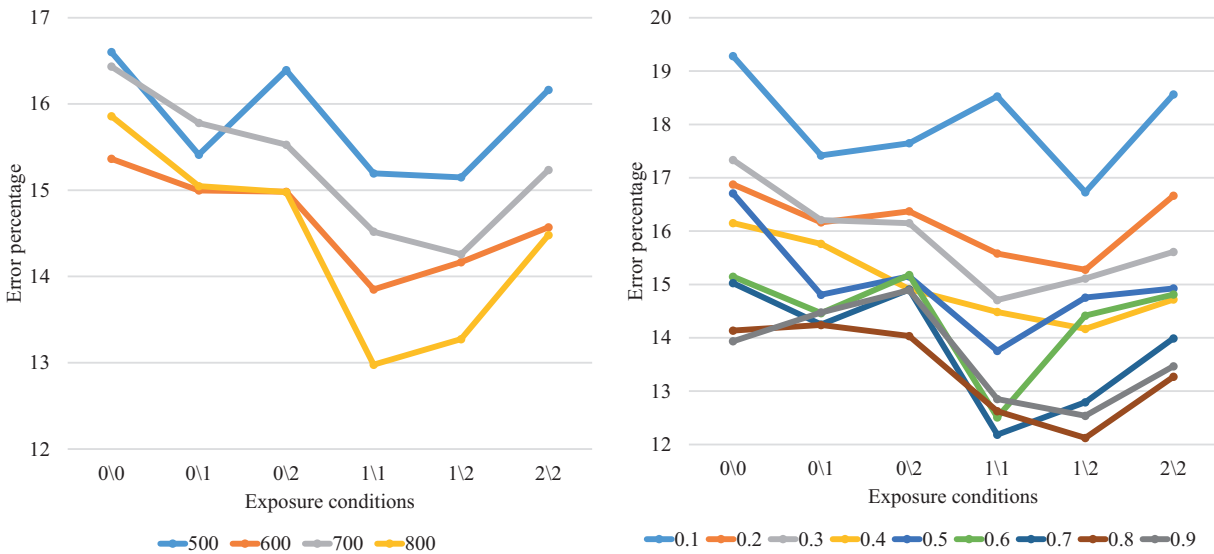


Figure 9. Effects of number of segments (left) and values of α (right) on different exposure conditions for “Reindeer” dataset.

of depth discontinuities, again, a higher value of ‘ α ’ produces a better result thereby indicating that NMI is a good cost metric while dealing with depth discontinuities. Lower ‘ α ’ would indicate that ‘ $(1-\alpha)$ ’ will have higher weightage, hence SAD of gradients would have a higher contribution to the final data cost as compared to NMI. In situations where the depth discontinuities are extremely prominent, performance of SAD of gradients tends to degrade, hence larger error is observed for lower values of ‘ α ’ in the presence of depth discontinuities.

4.2. Varying illumination conditions

Similar to the case of varying exposure conditions, we observe from Figure 10(a) that the variation in the number of segments does not have much of an effect on the performance of the algorithm. In other words, it can be assumed that for images which are similar to Aloe (lower depth discontinuities), the number of segments does not play a significant role in deciding the quality of the result. The tuning parameter ‘ α ’, on the other hand, plays a major role in the generation of disparity maps. As presented in Figure 10(b), similar to Figure 6(b), it is observed that the lower values of tuning factor ‘ α ’ produce better results. Lower values of ‘ α ’ signify that the SAD of image gradients has a larger role to play when compared to NMI.

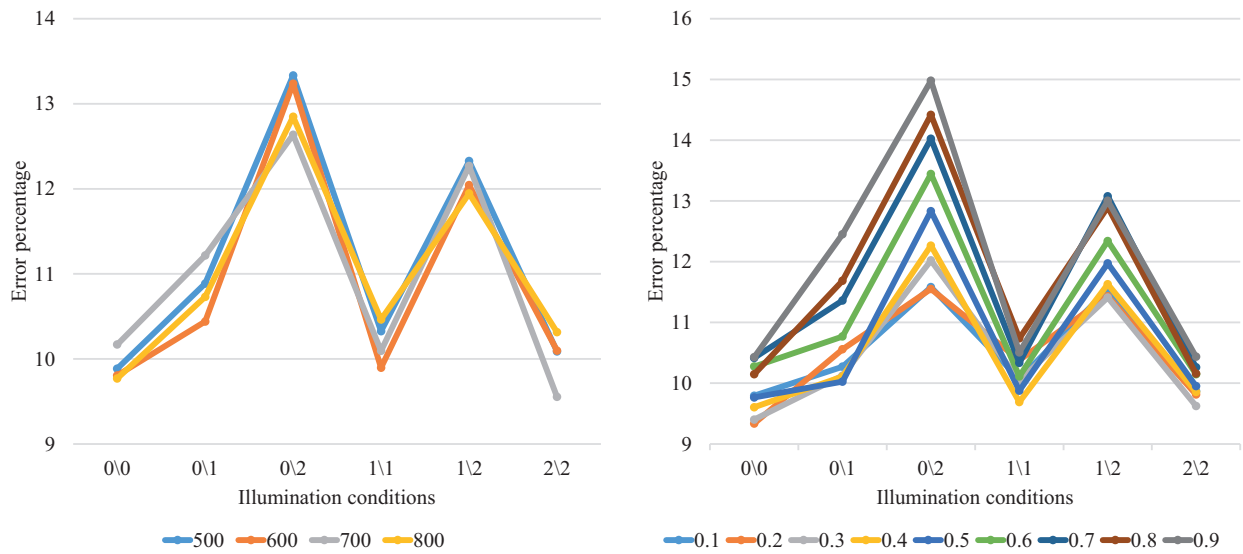


Figure 10. Effects of number of segments (left) and values of α (right) on different illumination conditions for “Aloe” dataset.

From Figure 11(a), it is observed that 500 segments produce best results. In Figure 11(b), it is evident from the graph that the mid-range of the tuning factor ‘ α ’ produces better results. This, in turn, concludes that in the case of ‘Baby2’ dataset, better results are obtained for lower values of segments and ‘ α ’ values, which force somewhat similar contribution of NMI and SAD of image gradients in the datacost. It has been noticed from Figures 12(a) and 12(b) that higher values of segments and larger values of tuning factor ‘ α ’ produces overall better results. The reason for this might be the fact that this particular image pair, namely, ‘Dolls’, from the dataset possesses higher depth discontinuities as compared to ‘Aloe’. The only exceptional cases being (0/1 and 0/2), wherein the authors noticed that at extreme variation in illumination, lower values of ‘ α ’ showed better performance, while in the majority of the cases, higher values of ‘ α ’ produced desirable results.

Observing Figures 13(a) and 13(b), similar to the previous case, better results (i.e., lower error values) were obtained for higher values of segments as well as the tuning factor ‘ α ’. Similar to the ‘Dolls’ dataset, this pair of image ‘Reindeer’, also has large depth discontinuities and also has multiple objects in them. The effect of the number of segments becomes particularly prominent in the latter half of the graph.

A visual comparison of the proposed method with other competing methods, under different radiometric conditions, is shown in the Figures 14 and 15.

Sample images of the Middlebury stereo dataset with varying exposure conditions and the disparity maps generated through the proposed methodology are presented in Figure 16.

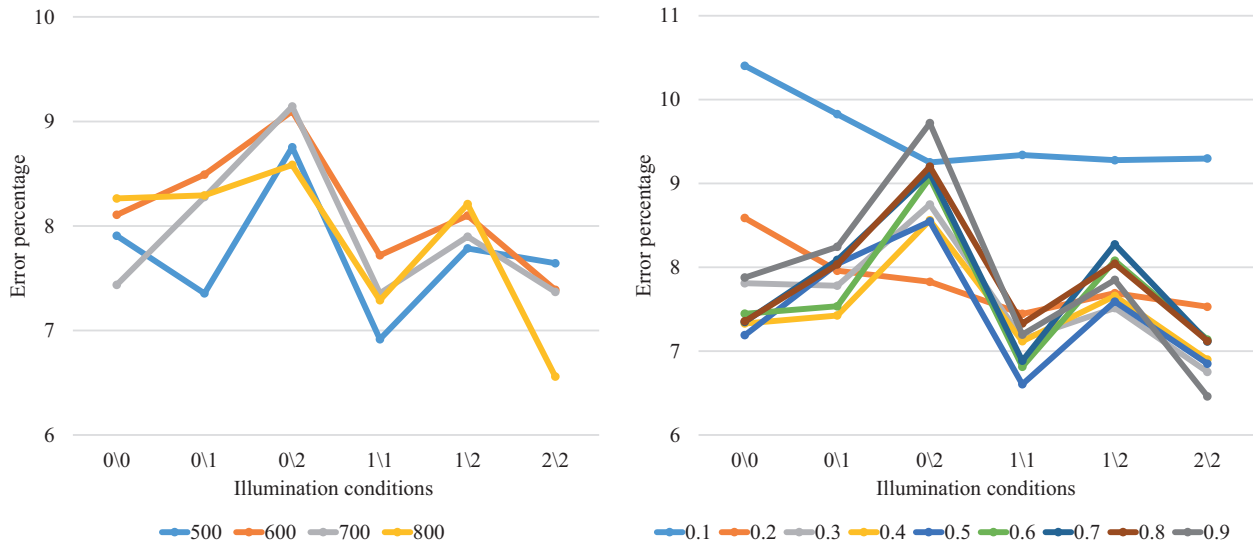


Figure 11. Effects of number of segments (left) and values of α (right) on different illumination conditions for “Baby2” dataset.

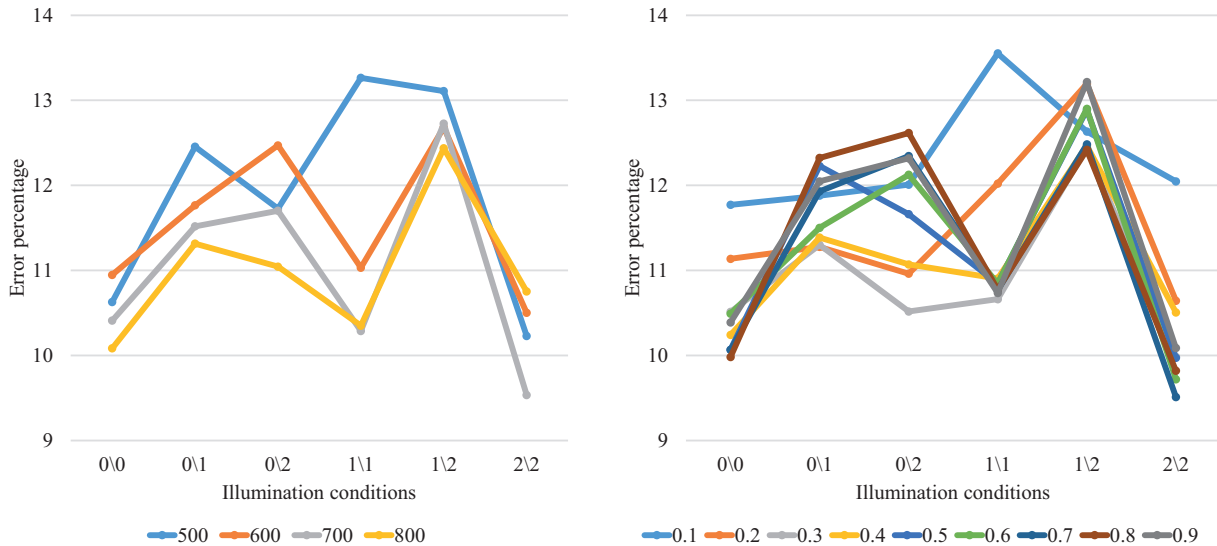


Figure 12. Effects of number of segments (left) and values of α (right) on different illumination conditions for “Dolls” dataset.

5. Discussions

We had initially assumed that the use of segments for stereo matching is beneficial in the presence of depth discontinuities. Even in the presence of radiometric variations, the use of segments was thought to improve the results as the inputs would be separated into clusters of similarly valued pixels, which was expected to reduce the chances of erroneous matches. Based on the results of the experiments displayed in the last section, it is observed that the number of segments and the tuning factor ‘ α ’ do play a significant role in determining the quality of the disparity map obtained. Figures 13(a) and 14(a) prove the previously mentioned statement regarding the influence of the number of segments on images with depth discontinuities. Even though increasing

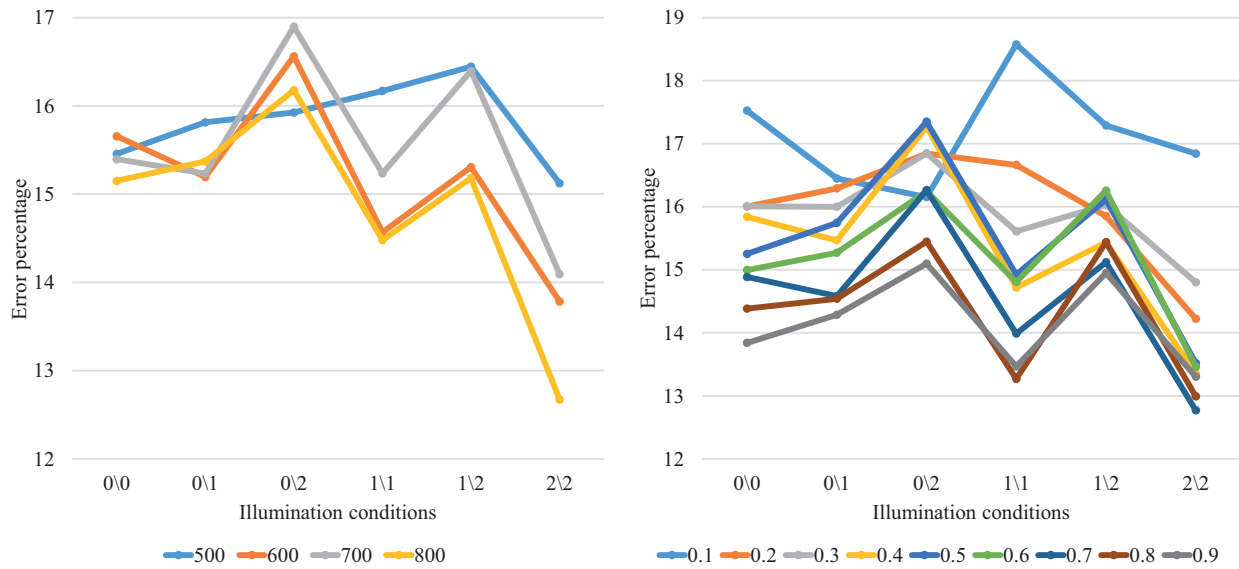


Figure 13. Effects of number of segments (left) and values of α (right) on different illumination conditions for “Reindeer” dataset.

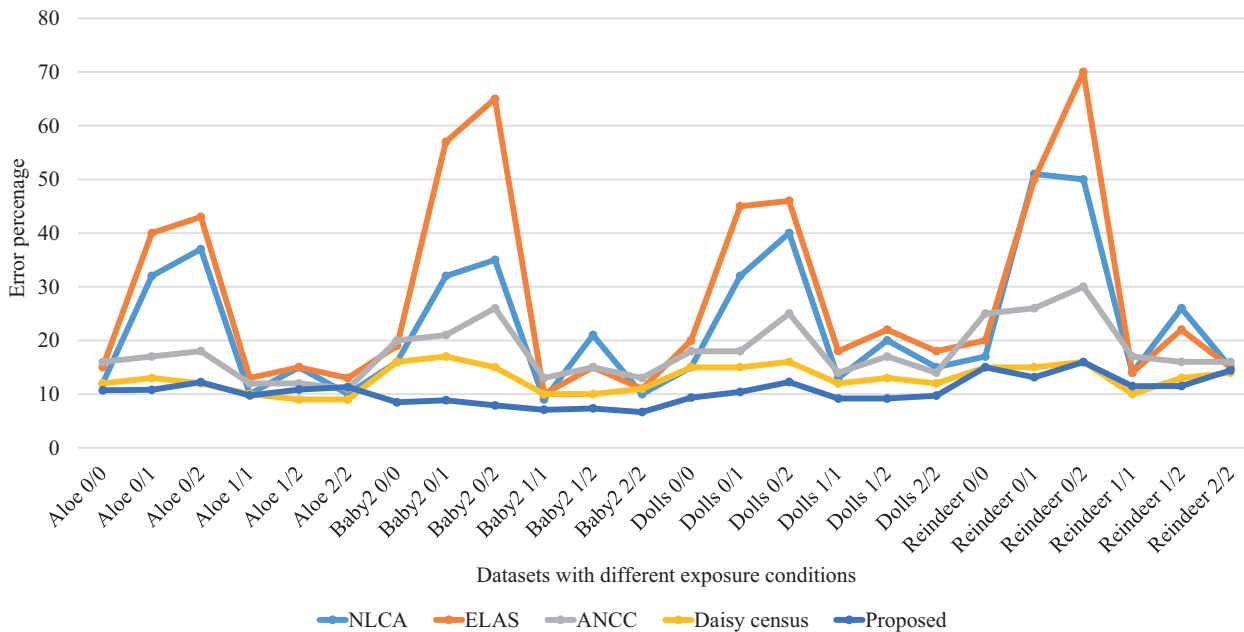


Figure 14. Comparison of proposed method with other methods under different exposure conditions.

the number of segments helps to reduce the error in such conditions, it was observed that in cases of extreme difference in depth in adjacent pixels, the proposed method is not able to perform as well as it does in other cases. A similar inference can be drawn when there is a large difference in radiometric conditions (ex. with 0 and 2 conditions). Figures 11(a) and 12(a) are a good example of this situation. Baby2 has comparatively lower depth discontinuity. According to the author’s assumptions, lower values of segments should have produced

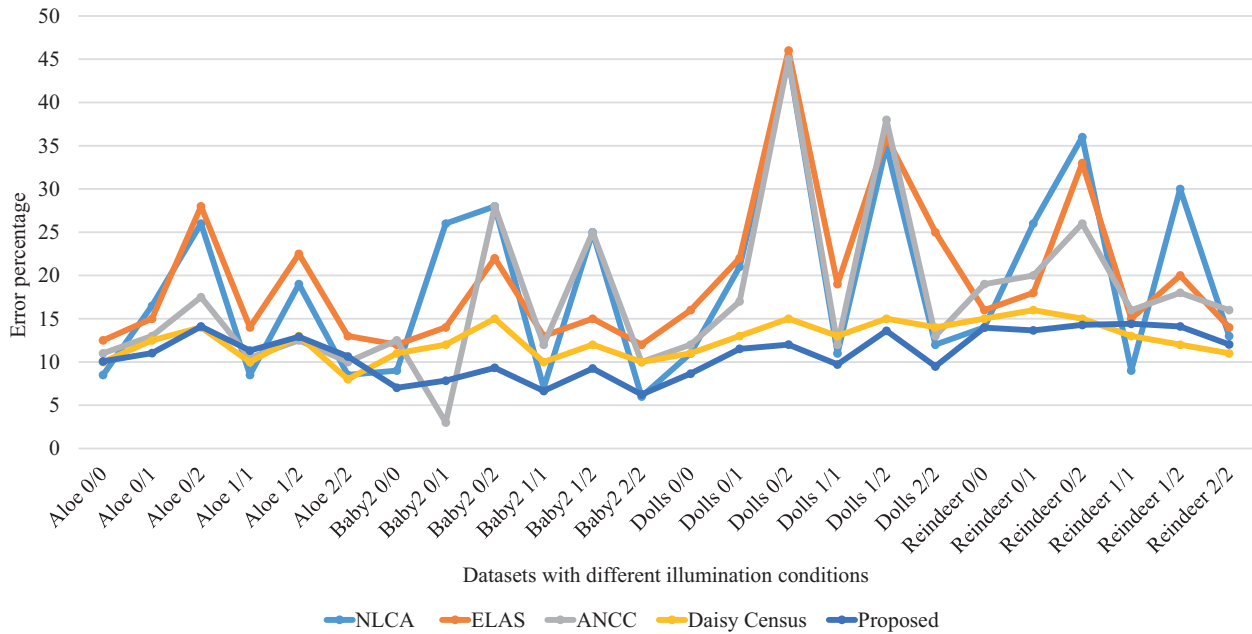


Figure 15. Comparison of proposed method with other methods under different illumination conditions.

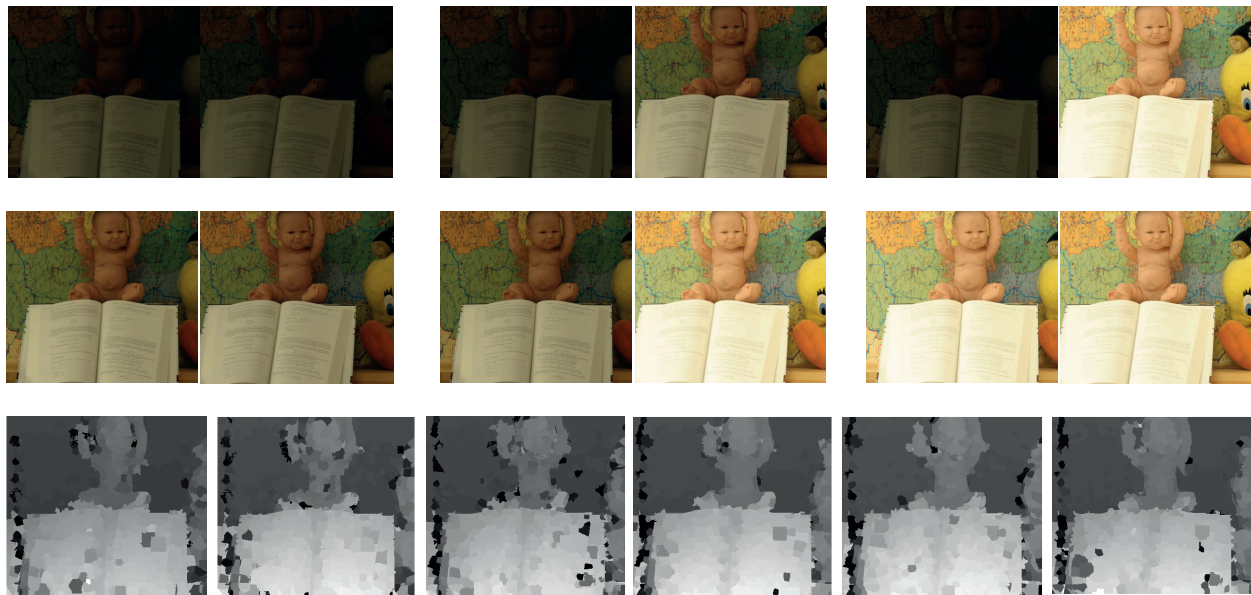


Figure 16. First and second row describe sample image pairs with different exposure conditions i.e., 0/0, 0/1, 0/2, 1/1, 1/2 and 2/2, respectively. Third row presents the results of the proposed algorithm for each of the above conditions.

better results. As observed from the graphs, this assumption does hold true for almost all the datasets and the radiometric conditions used in this experiment. A comparison of the errors of the disparity maps generated from the proposed method with that of other methods is presented in Table 2 for different illumination conditions and Table 3 for varying exposure conditions.

Table 2. Comparison of error values of disparity maps generated by the proposed method with other methods (different illumination conditions)[24].

Dataset	NLCA	ELAS	ANCC	Daisy census	Proposed
Aloe 0/0	8.5	12.5	11	10	10.07
Aloe 0/1	16.5	15	13	12.5	11.02
Aloe 0/2	26	28	17.5	14	14.1
Aloe 1/1	8.5	14	10.5	10	11.32
Aloe 1/2	19	22.5	12.5	13	12.92
Aloe 2/2	8.5	13	10	8	10.64
Baby2 0/0	9	12	12.5	11	7.02
Baby2 0/1	26	14	3	12	7.84
Baby2 0/2	28	22	28	15	9.32
Baby2 1/1	7	13	12	10	6.65
Baby2 1/2	25	15	25	12	9.24
Baby2 2/2	6	12	10	10	6.25
Dolls 0/0	11	16	12	11	8.6
Dolls 0/1	21	22	17	13	11.5
Dolls 0/2	45	46	45	15	12.01
Dolls 1/1	11	19	12	13	9.7
Dolls 1/2	35	36	38	15	13.6
Dolls 2/2	12	25	13	14	9.49
Reindeer 0/0	14	16	19	15	13.97
Reindeer 0/1	26	18	20	16	13.6
Reindeer 0/2	36	33	26	15	14.28
Reindeer 1/1	9	15	16	13	14.4
Reindeer 1/2	30	20	18	12	14.11
Reindeer 2/2	13	14	16	11	12.03

6. Conclusion

The behaviour of a dense local stereo matching algorithm is studied on images with radiometric variation wherein the disparity selection is conducted by a fuzzy logic module. The effect of the number of segments and the tuning factor ‘ α ’ on the stereo matching process is studied. It was observed that the assumption of the result improving by increasing the number of segments, in the case of depth discontinuities, is verified in almost all of the datasets and radiometric conditions, and the algorithm is in par with other methodologies which are used to overcome the problem of generating dense disparity maps in the presence of radiometric variations in the images. It was noticed that for images with large depth discontinuities, NMI with large number of segments produces a comparatively better result while in the case of images with lower depth discontinuities, SAD of image gradients performs better. Based on these observations, it can be concluded that the proposed combination of data costs, along with the fuzzy disparity selector explained in this work, are a reliable tool to generate dense disparity maps even in the presence of radiometric changes.

Table 3. Comparison of error values of disparity maps generated by the proposed method with other methods (different exposure conditions) [24].

Dataset	NLCA	ELAS	ANCC	Daisy census	Proposed
Aloe 0/0	12	15	16	12	10.71
Aloe 0/1	32	40	17	13	10.81
Aloe 0/2	37	43	18	12	12.23
Aloe 1/1	10	13	12	10	9.75
Aloe 1/2	15	15	12	9	10.85
Aloe 2/2	10	13	11	9	11.32
Baby2 0/0	16	19	20	16	8.48
Baby2 0/1	32	57	21	17	8.86
Baby2 0/2	35	65	26	15	7.9
Baby2 1/1	9	10	13	10	7.08
Baby2 1/2	21	15	15	10	7.33
Baby2 2/2	10	11	13	11	6.65
Dolls 0/0	15	20	18	15	9.36
Dolls 0/1	32	45	18	15	10.41
Dolls 0/2	40	46	25	16	12.25
Dolls 1/1	13	18	14	12	9.18
Dolls 1/2	20	22	17	13	9.18
Dolls 2/2	15	18	14	12	9.71
Reindeer 0/0	17	20	25	15	15.01
Reindeer 0/1	51	50	26	15	13.15
Reindeer 0/2	50	70	30	16	15.97
Reindeer 1/1	14	14	17	10	11.48
Reindeer 1/2	26	22	16	13	11.51
Reindeer 2/2	15	15	16	14	14.42

Acknowledgment

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

References

- [1] Yoon KJ. Stereo matching based on nonlinear diffusion with disparity-dependent support weights. *IET Computer Vision* 2012; 6 (4): 306-313.
- [2] Heo YS, Lee KM, Lee SU. Illumination and camera invariant stereo matching. In: *IEEE Conference on Computer Vision and Pattern Recognition*; Anchorage, AK, USA;2008 pp. 1-8.
- [3] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 2002; 47: 7-42.

- [4] Yang Q. A non-local cost aggregation method for stereo matching. In: IEEE Conference on Computer Vision and Pattern Recognition; Providence, RI, USA; 2012 pp. 1402-1409.
- [5] Yang G, Zhao H, Shi J, Deng Z, Jia J. SegStereo: Exploiting Semantic Information for Disparity Estimation. In: Proceedings of the European Conference on Computer Vision; Beijing, China; 2018. pp. 660-676.
- [6] Cheng X, Wang P, Yang R. Learning depth with convolutional spatial propagation network. arXiv preprint arXiv:1810.02695; 2018.
- [7] Song X, Zhao X, Fang L, Hu H. EdgeStereo: An Effective Multi-Task Learning Network for Stereo Matching and Edge Detection. arXiv preprint arXiv:1903.01700; 2019.
- [8] Szeliski R. Computer vision: algorithms and applications. Berlin, Germany: Springer Press, 2010.
- [9] Kanade T, Kano H, Kimura S, Yoshida A, Od K. Development of a video-rate stereo machine. In: IEEE Int. Conference on Intelligent Robots and Systems;Pittsburgh, PA, USA; 1995. pp. 95-100.
- [10] Simoncelli EP, Adelson EH, Heeger DJ. Probability distributions of optical flow. In: IEEE Conference on Computer Vision and Pattern Recognition; Maui, HI, USA; 1991. pp. 310-315.
- [11] Mattoccia S, Tombari F, Di Stefano L. Fast full-search equivalent template matching by enhanced bounded correlation. IEEE Transactions on Image Processing 2008; 4 : 528-538.
- [12] Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: IEEE International Conference on Computer Vision; Bombay, India; 1998. pp. 839-846.
- [13] Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence. In: European Conference on Computer Vision; Berlin, Heidelberg, Germany; 1994. pp. 151-158.
- [14] Mei X, Sun X, Zhou M, Jiao S, Wang H, Zhang X. Building an accurate stereo matching system on graphics hardware. In: IEEE International Conference on Computer Vision; Barcelona, Spain; 2011. pp. 467-474.
- [15] Hamzah RA, Ibrahim H, Hassan AHA. Stereo matching algorithm based on per pixel difference adjustment, iterative guided filter and graph segmentation. urnal of Visual Communication and Image Representation 2017; 42: 145-160.
- [16] Zhu H, Yin J, Yuan D. SVCV: segmentation volume combined with cost volume for stereo matching. IET Computer Vision 2017; 11 (8): 733-743.
- [17] Lee DH, Chang JY, Heo YS. Stereo matching using cost volume fusion for high dynamic range scenes. Electronics Letters 2017; 53 (23): 1522-1524.
- [18] Roszkowski M. Stereo matching with superpixels. In : Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments; Wilga, Poland; 2012. pp. 1-9.
- [19] Heo YS, Lee KM, Lee SU. Robust stereo matching using adaptive normalized cross-correlation. IEEE Transactions on Pattern Analysis and Machine Intelligence 2011; 33 (4): 807-822.
- [20] Hirschmuller H. Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence 2008; 30 (2): 328-341.
- [21] Hirschmuller H, Scharstein D. Evaluation of stereo matching costs on images with radiometric differences. IEEE Transactions on Pattern Analysis and Machine Intelligence 2008; 29: 1582-1599.
- [22] Yaman M, Kalkan S. Performance evaluation of similarity measures for dense multimodal stereo vision. Journal of Electronic Imaging 2016; 25: 1-31.
- [23] Pluim JP, Maintz JA, Viergever MA. Image registration by maximization of combined mutual information and gradient information. IEEE Transactions on Medical Imaging 2000; 19 (8): 809-814.
- [24] Qu Y, Jiang J, Deng X, Zheng Y. Robust local stereo matching under varying radiometric conditions. IET Computer Vision 2014; 8 (4): 263-276.
- [25] Baydoun M, Al-Alaoui MA. Enhancing stereo matching with varying illumination through histogram information and normalized cross correlation. In: IEEE International Conference on Systems, Signals and Image Processing ;Bucharest, Romania; 2013. pp. 5-9.

- [26] Jiao J, Wang R, Wang W, Dong S, Wang Z, Gao W. Local stereo matching with improved matching cost and disparity refinement. *IEEE MultiMedia* 2014; 21: 16-27.
- [27] Tombari F, Mattoccia S, Di Stefano L, Addimanda E. Near real-time stereo based on effective cost aggregation. In: *IEEE International Conference on Pattern Recognition*; Tampa, FL, USA; 2008. pp. 1-4.
- [28] Xu Y, Zhao Y, Ji M. Local stereo matching with adaptive shape support window based cost aggregation. *Applied Optics* 2014; 53 (29): 6885-6892.
- [29] Egnal G. Mutual information as a stereo correspondence measure. University of Pennsylvania, Department of Computer and Information Science.
- [30] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2014; 34: 2274-2282.
- [31] Ross TJ. *Fuzzy logic with engineering applications*. 3rd ed. University of New Mexico, New Jersey, USA: John Wiley & Sons, 2009.
- [32] Shetty AA, George VI, Nayak CG, Shetty R. Fuzzy logic based disparity selection using multiple data costs for stereo correspondence. *Turkish Journal of Electrical Engineering & Computer Sciences* 2019; 27: 377-391.
- [33] Mroz F, Breckon TP. An empirical comparison of real-time dense stereo approaches for use in the automotive environment. *EURASIP Journal on Image and Video Processing* 2012; 13: 1-19.