Improved Stereo Matching Algorithm based on Census Transform and Dynamic Histogram Cost Computation

A. F. Kadmin[1], R. A. Hamzah[2], M. N. Abd Manap[3], M. S. Hamid[4], S. F. Abd Gani[5]

[1,3]*Fakulti Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*
[2,4,5]*Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik, Universiti Teknikal Malaysia Melaka,Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*

Abstract— **Stereo matching is a significant subject in the stereo vision algorithm. Traditional taxonomy composition consists of several issues in the stereo correspondences process such as radiometric distortion, discontinuity, and low accuracy at the low texture regions. This new taxonomy improves the local method of stereo matching algorithm based on the dynamic cost computation for disparity map measurement. This method utilised modified dynamic cost computation in the matching cost stage. A modified Census Transform with dynamic histogram is used to provide the cost volume. An adaptive bilateral filtering is applied to retain the image depth and edge information in the cost aggregation stage. A Winner Takes All (WTA) optimisation is applied in the disparity selection and a left-right check with an adaptive bilateral median filtering are employed for final refinement. Based on the dataset of standard Middlebury, the taxonomy has better accuracy and outperformed several other state-of-the-art algorithms.**

Keywords—**Stereo matching, disparity map, dynamic cost, census transform, local method**

I. Introduction

Stereo matching is one of the most promising subjects in machine vision studied by many researchers. The stereo vision system application was increasing and applied in several applications such as image processing in medical, virtual reality, autonomous navigation and many more [1]. The disparity estimation accuracy evaluation is quite crucial since small inaccuracies lead to no little impact on the result of the 3D application. Thus, it has been a targeted subject by many researchers [2]. The disparity accuracy assessment of the stereo matching algorithm can be compared with other state-of-the-art algorithms using the online evaluation platform such as Middlebury Dataset [3].

Most of the algorithms established are based on the traditional taxonomy of Schartein and Szeliski, comprised of four stages; 1. Matching cost computation, 2. Aggregation of cost, 3. Disparity selection, and 4. Final disparity refinement [4]. The taxonomy used an essential cost function of matching to measure the stereo images' corresponding points from two or multiple perspectives [5]. This taxonomy aims to acquire a disparity map; thus, it can be translated into depth assessment for depth-based processing and communications. The disparity estimation accuracy evaluation is quite crucial since small inaccuracies lead to no little impact on the result of the 3D application. The disparity measurement accuracy will be the quantitative assessment and can be compared with other algorithms [3].

It is difficult to acquire accurate correspondences due to several factors, especially from low texture regions, radiometric differences from environment illumination variations, and blurry boundaries displayed due to poor segmentation methods. Several factors contributed to these problems, such as specular reflection and non-Lambertian surfaces [6]. The stereo camera light sensor also produced illumination differences, resulting in dissimilar intensity levels corresponding to the same point in 3D space. Besides that, the irregularity of the image captured from the stereo camera also caused radiometric differences. The properties from inconsistent multiple stereo cameras like Gaussian noise, salt and pepper noise, gain setting, vignetting, etc., are the main factors for radiometric differences.

An improved taxonomy of the local stereo matching algorithm approach is developed to increase the disparity map accuracy based on modified dynamic cost computation.

This work used a modified Census Transform (CT) with a dynamic histogram to provide the cost computation in the matching cost stage, while in the aggregated cost stage, it utilised a fixed-window strategy with bilateral filtering that efficiently retains image depth information and edge in the image. In the disparity selection and refinement stage, a Winner-Takes-All (WTA) selection and a consistency check of left and right with adaptive bilateral median filtering are employed. The rest of this article is organised as follows. Section 2 explained the illumination constraints. The methodology of the taxonomy is shown in section 3. Section 4 will explain the experiment result and the discussion, while the conclusion is summarised in section 5.

II. Related Works

Most of the recent algorithms have been developed widely by several researchers based on the taxonomy consists of four stages, as shown in Figure 1. The input will be a stereo image pair and he output of this taxonomy will be the smooth, dense, or sparse disparity map [4]. Each block in the taxonomy is the area of focus by many researchers, consisting of one or several algorithms to process the image to enhance the overall performance of disparity accuracy level or time execution.

Computation of matching cost. The earliest fundamental in the stereo matching algorithm design is the matching cost, which obtains the value of disparity map. [7] proposed essential matching cost absolute difference (AD) cost initialization for real-time high-quality system. The advanced AD formula sum of absolute differences (SAD) used by [8] to calculate for each pixel based on the disparity under consideration while zero-mean SSD (ZSSD) used by [9] which eliminates each patch of average intensity and used for the comparison between mean intensity of independent and each pixels..

Another familiar matching cost computation formula used is Normalized Cross Correlation (NCC) technique for generating matching costs.

It is a similarity measure because it has a robust invariant property for global bias and gains changes [10]. [11] adopted a an improvement of Census transform consists of local texture metric used to calculate the initial cost. [12] proposed a different and new taxonomy called PatchMatch-based using superpixel cut and utilized the 3D labels of an image as matching cost computation accurately. More advanced algorithms utilized CNN-based matching cost function as segment of binary abilities in the energy function's smoothness term employed by [13].

Cost Aggregation. Traditional approach for this stage is to apply a squared shape fixed-sized window due to its easy low computational complexity and implementation [14]. Also, the fixed-size windows performance is varying according to the different sizes of images or other datasets are used. [15] adopted a cost aggregation using cross-scale technique by incorporating weighted least square with intrascale of smoothness constraint. A coarse-to-fine multi-window algorithm can be implemented by forming multiple smaller windows deals with slanted surfaces and discontinuities [16]. Various approaches have been applied, such as shiftable window and support window [17]–[19]. Still, the support window's rectangular constrained shape having the difficulty at the near the depth discontinuity regions to measure disparity values of the pixels [20].

A better solution is to implement adaptive support weight (ASW) by adjusting each point's cost weight, such as similarity and proximity between adjacent pixels and target points in a fixed size window [21]. [22] proposed a filter-Based Guidance Image as cost aggregation to decrease the complexity for computational and make it resilient by using the popular filter such as edge-aware filters; guided filter and bilateral filter. [1] produced an improvement in disparity accuracy for large area of textureless regions by implementing an Adaptive Guided Filter (AGF) over the support region. [23] introduced an improvement algorithm of stereo matching using Adaptive Weighted Bilateral Filter (AWBF) as the primary filter for the aggregation process.
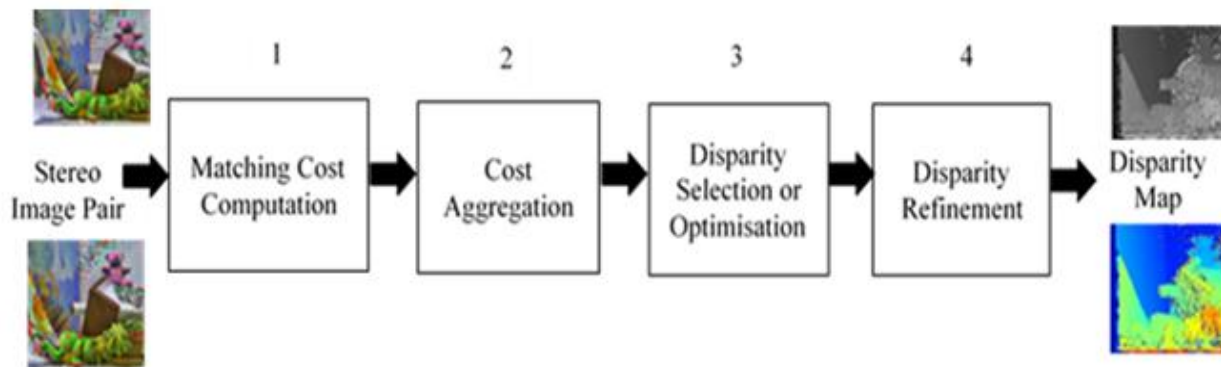
Figure 1: Traditional stereo matching taxonomy by Scharstein and Szeliski

A combination of minimum spanning tree based on support region and joining object flow in cost aggregation is adopted by [24]–[27] rather than using fixed-sized windows. An enhanced MST called 3DMST-CM was proposed by [28] handling a cases based on ambiguity of image pixel to achieve a high level accuracy in disparity map. Another variant of enhanced cost aggregation is the simplified independent component correlation algorithm (SICA) to solve mismatching at boundaries, non-texture regions and tiny details in the map [29]. Also, an attention of cost aggregation applied to the encoder-decoder network taxonomy in deep learning recently applied by [30] comprises loss of calculations and multiple outputs, which efficiently increases the accuracy of matching however it contributed to increase of training time and memory loss. Additionally, an ASW also been proposed as cost aggregation which, [31] combined ASW with CNN to end network as cost aggregation to get more support for thin structures.

Disparity Optimization. The fundamental disparity optimization is utilizing the Winner Takes All (WTA) strategy. The final disparity value of that particular pixel point is selected from the disparity range with the minimum cost value after cost aggregation. [32] adopted cross-based window voting in WTA called Range-WTA to estimate the disparity to fill white holes that represent non-consecutive disparities with the same minimum cost.

In contrast, global optimization commonly produces higher accuracy of disparity maps compare local methods at computational cost disadvantage. [33]–[36] introduced a global approach on of WTA optimization with dynamic programming and [15] Least Square optimization to find the final disparity value.

This global approach produces less errors caused by textureless, occlusion, and discontinuity regions.

Disparity Refinement. Post-processing in the stereo matching taxonomy is called disparity refinement to remove any outliers, uncertainties, and noise from the map for disparity to achieve a greater level of accuracy. These noises only can be detected by performing a left-right consistency checking, bi-modality, match goodness jumps, and occlusion constraint [27]. Additionally, [32] proposed the technique using cross voting of image-based, and a median filter is applied to perfect the depth estimation cost using the triple image approach to identify textureless regions and false matches. [37]–[39] introduced a coarse-to-fine (CTF) in the image segmentation with image labelling addition to histogram to correct the occlusion and error pixels. Besides that, various approaches have been proposed in the disparity refinement, such as the probabilistic approach, Order-Based and segment-tree structure [40].

III. METHOD

Fundamentally, the proposed stereo matching algorithm has been developed is illustrated in Figure 2. This taxonomy utilises the disparity estimation of modified matching cost with dynamic histogram and census signature in the matching cost computation to produce a 2D depth map compared with a traditional taxonomy by [41]. To accomplish the objective of this work, some processes will be conducted. The proposed work methodology consists of five main necessary stages, pre-processing, matching cost, aggregation of cost, disparity optimisation and final disparity refinement.
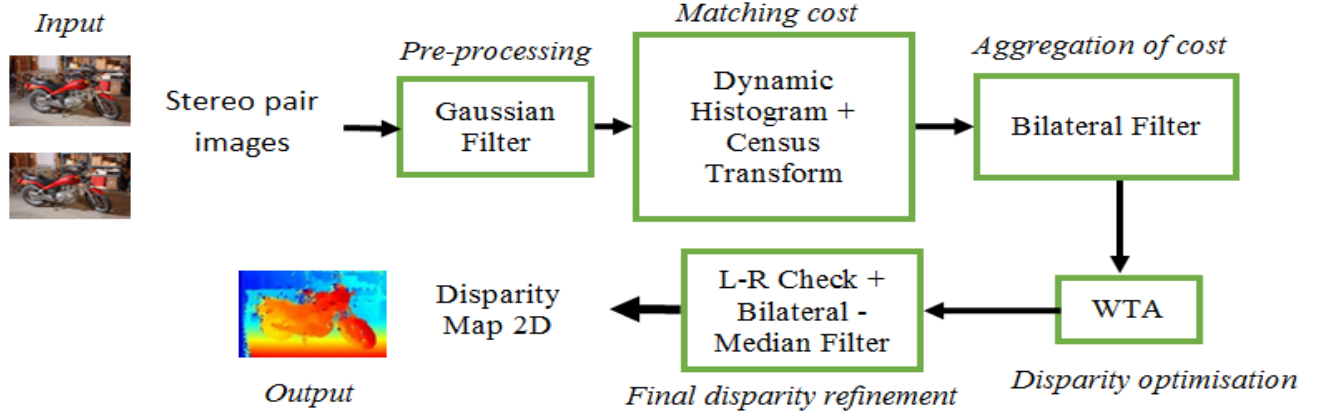
Figure 2: The taxonomy blocks of the developed algorithm

This taxonomy of stereo images is pre-processed with the 1D Gaussian filter size 1 x 3 to smooth the original images as in Eq. 1:

$$G_f(x) = \exp \frac{-x^2}{2\sigma^2} \qquad (1)$$

where x is the relative coordinate to the center of kernel and σ is the filter standard deviation. The algorithm generates sub-histogram partition, the gray level allocation for dividing each sub-histogram and employing equalisation [42] in the matching cost computation. The partitions of the sub-histogram are based on local minima parameters. Again, a 1D Gaussian filter size 1 x 3 is utilised on the main histogram to eliminate unnecessary minima. Afterwards, it creates partitions of sub-histogram based on the points that fall between two local minima, which by default the non-zero histogram of first and last are measured as minima [43]. The partitions will shun domination by some part of the histogram to others. Each sub-histogram is allocated with a specific range of gray level values that spanned an output image's histogram. The gray level allocation is determined based on the gray level span ratio in each occupied sub-histogram. This will avoid over or under enhancement to the stereo image because equalisation works separately on each sub-histogram, as in Eq. 2.

$$Y(x,y) = \sum_{k=0}^{i-1} range_{k+1} + (\sum_{k=i}^{i} range_k - \sum_{k=i}^{i-1} range_{k+1}) \; x \; \sum_{k=1}^{i} \frac{n_k}{M} \qquad (2)$$

where the Y (x, y) the output dynamic histogram. The *range* presents the dynamic range for the histogram while n_k is the quantity of pixels with the intensities of k, and M is the total of pixels.

The approach avoids both gray levels of dissimilar sub-histograms mapped to the identical value of gray level in the image output mainly due to the sequential, non-overlapping and specific grey level ranges. So, there is no loss of important information in image details [44]. CT is a non-parametric local transform formula with a local intensity relationship between center pixel and near pixels inside a matching window [45]. The matching window is defined as (2u + 1) x (2v + 1), so the equation can be presented as Eq. 3 and Eq. 4 from the Y (x, y) for both reference and target image:

$$T(x,y) = \bigotimes_{i=-u}^{u} \bigotimes_{i=-v}^{v} \xi(I(Y(x,y)), I(Y(x+i,y+i))) \quad (3)$$

$$\xi(q_1,q_2) = \begin{cases} 1, if \; q_1 > q_2 \\ 0, if \; q_1 \leq q_2 \end{cases} \qquad (4)$$

where the intensity from the gray level of the histogram for the pixel in Y (x, y) represents by I(Y(x,y)) while the $\bigotimes$ designates a bitwise catenation, the ξ is the relationship function and can be well-defined where q_1 is the centre pixel while q_2 is the closest neighbour pixel. Then, the Hamming Distance is calculated using the Hamming Function to determine the differences between both transform vectors that expressed as:

$$C_d(x,y) = Hamm(T_l(x,y), T_r(x-d,y)) \qquad (5)$$

where d is the disparity and T_l, T_r represents the left and right transform image vectors. A pixel's correspondence can be computed by measuring the cost for entire candidates in a window between the target image and reference image.

The target image provides the best cost of the window compared with the corresponding reference image. A fixed window applied using the box filter method [46]. The size of the window is established to $(2z + 1)$ x $(2z + 1)$ so that the value of cost aggregation can be determined using:

$$T^r(x,y,d) = \frac{1}{2z+1}\sum_{u=-z}^{z} C_d(x+u,y,d) \qquad (6)$$

$$A_{square}^r(x,y,d) = \frac{1}{2z+1}\sum_{v=-z}^{z} T^r(x,y+v,d) \qquad (7)$$

where the cost matrix is T^r works to store the intermediate outcomes. It assumes the stereo pair images are well rectified, showing the corresponding epipolar lines are on the same height and horizontal. The cost value will be aggregated with the bilateral smoothing filter as in Eq. 8:

$$CA(p,d) = \sum_{q \in W_p|d(q)=d_z} A_{square}^T(x,y,d) \qquad (8)$$

where d_z indicates the disparity range and wp represents the size of the matching window, including z x z radius at a centred of p pixel. This work determines the disparity selection after cost aggregation to minimise each pixel of aggregated corresponding value utilising the WTA approach to obtain the precise disparity map. The WTA approach for local methods can improve the computational cost as applied by [47]. Based on their works, the disparity maps output during this stage still producing errors in the textureless and occluded regions. The expression of WTA is delivered by Eq. 9:

$$d = \arg\min_{d \in d_z} CA(p,d) \qquad (9)$$

where the minimum cost of d from cost aggregation is selected, the cost aggregation signifies the range of allowable disparity values. The final stage of this work comprises the post-processing step defined as final disparity refinement. This approach is worked from the reference disparity map of the image, which corresponds to the right disparity map's target image. Thus, utilising the similar strategy employed by [48], the location of point p for disparity validation map is expressed by Eq. 10:

$$d_0(p) = \begin{cases} 1, & if\ |d_{LR}(p) - d_{RL}(p - d_{LR}(p))| \leq \tau_{LR} \\ 0, & otherwise \end{cases}$$

$$(10)$$

where d_{RL} and d_{LR} signify reference and target of the disparity maps. The map comprises the disparity location of valid (i.e., 0=inlier) and invalid (i.e., 1=outlier). In this work, the τ_{LR} is fixed to zero that have a similar setting with [49]. The aim is to produce the minimum error in the final map. Unnecessary noises can be removed using the Adaptive Bilateral Median Filter (ABMF) of B(p;q). The ABMF was employed as in [40] as in Eq. 8. Their work attained a great accuracy level on the removal of noise. Corresponding to Eq. 8 and Eq. 10, the value of weighted B(p, q) is adjusted to a summation of the histogram (p, d_z) that contributes to Eq. 11:

$$h(p,d_z) = \sum_{q \in W_p|d(q)=d_z} B(p,q) \qquad (11)$$

where d_z indicates disparity range and w_p represents the size of matching window, including the z x z radius at the centred of p pixel. The median of h(p, d_z) valued finalised the disparity value d′ given by Eq. 12:

$$d' = med\{d|h(p,d_z)\} \qquad (12)$$

The algorithm's quantitative and qualitative performance analysis will be employed using the Middlebury Vision Benchmark Dataset by computing the bad pixel error as in Eq. 13. Middlebury dataset was established by [4], that comprises 15 training set images. The level assessment of accuracy for each image depends on two parameters (i.e., bad pixels and average errors among all pixels in non-occluded areas (nonocc) and all pixels detected as valid pixels (all). The lower percentage for the stereo matching algorithm will be the better disparity map accuracy, allowing the algorithm's accuracy to be assessed objectively [50].

$$B = \frac{1}{N}\sum_{(x,y)}(|d_c(x,y) - d_T(x,y)| > \delta_d) \qquad (13)$$

where δ_d (eval bad thresh) is a disparity error tolerance, for this work, we used $\delta_d = 1.0$. N is the total of pixels.

IV. EXPERIMENTAL RESULT

All experiments for this work are accomplished using the Window 10 platform on desktop PC with Intel Xeon 2.6GHz processor with 64GB memory. The Middlebury dataset was standard from natural surroundings of indoor and outdoor from the stereo camera system.

Therefore, it contains very complicated images and interesting for taxonomy evaluation. The algorithm executed all the 15 stereo pair images in about 398.1 seconds with an average of about 26.54 seconds per image using the Matlab environment. The taxonomy consists of the matching cost computation algorithm utilising dynamic cost volume to improve the radiometric differences and discontinuity and disparity accuracy. Figure 3 shows the results of comparison disparity map matching on the Middlebury images 'Adirondack', 'Playtable' and 'Teddy' with the original reference images from the training dataset.

Reference Image	Ground Truth	Disparity Map

(a) Adirondack

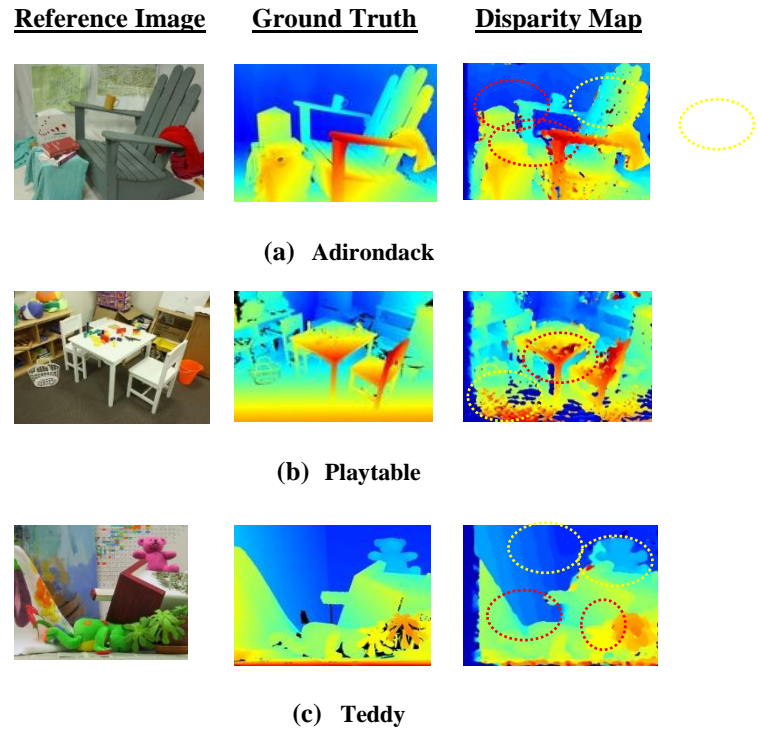**(b) Playtable**

(c) Teddy

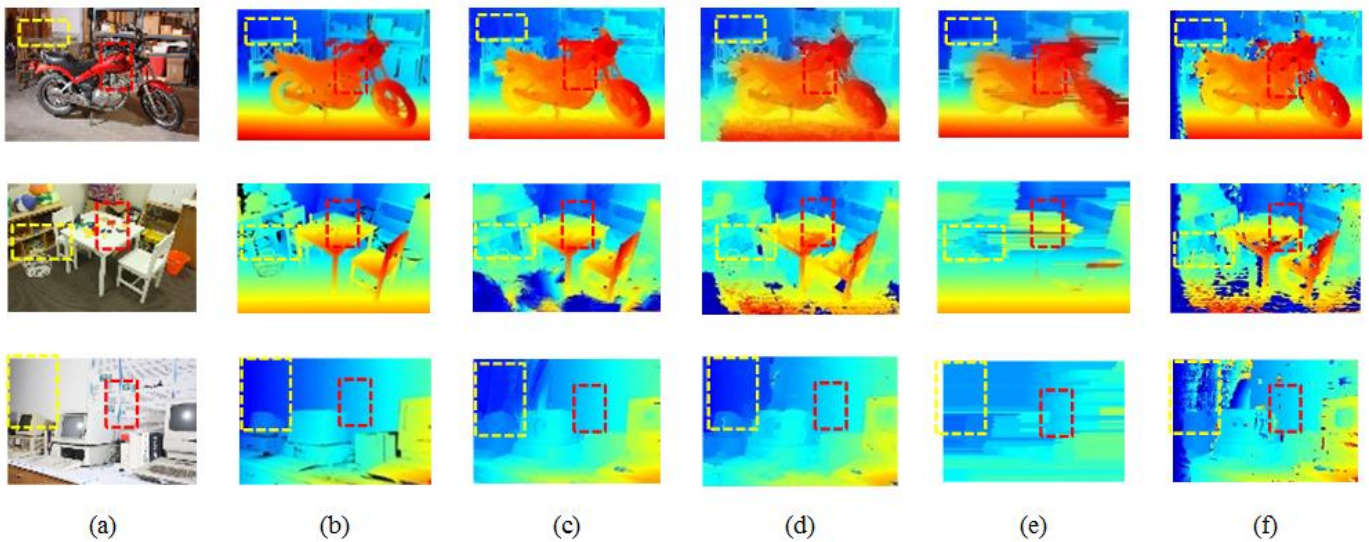Figure 3. Reference image versus the disparity map (a) Adirondack (b) PlayTable (c) Teddy

Figure 4. Disparity map comparison for Middlebury dataset 'Motorcycle,' Playtable' and 'Vintage' images (a) Reference image (b) Ground-truth (c) ISM (d) ADSR_GIF (e) R-NCC (f) Proposed taxonomy

It indicates that the taxonomy performance around the uniform and border surface is worked well, as marked in the red circle. The disparity map also shows minimum streak artefacts, which usually occur in the local method stereo vision system. Though the disparity map for the developed taxonomy is not entirely smoothed at the low texture regions, the area of illumination variations and discontinuities pro-duce better matching detection. The taxonomy performed poorly at low texture regions and occlusion areas as an around yellow circle. It can be shown in the figure as a streak at the cup in the Adirondack image and Playtable floor. The background pattern for Adirondack and Teddy images also disappears due to the taxonomy smoothed small object that is recognised as a uniform surface. The difference between uniform region to each other also very smoothed as expected that contribute the disparity map for the object is fatter. The fatter effect can be avoided by introducing sub-pixel interpolation techniques.

Figure 4 shows the comparison disparity map matching result on the Middlebury images 'Motorcycle', 'Playtable' and 'Vintage' between several established methods such as Gradient Matching with Iterative Guided Filter (ISM) [51], Adaptive Support Weight with Guided Filter (ADSR_GIF [52] and Modified NCC (R-NCC) [53]. The object scenes located at the depth are designated step by step, increasing the disparity values based on the final value from closer to farther according to colours assignment.

Although the disparity map for the developed taxonomy is not entirely smoothed than ISM and ADSR_GIF, the uniform and border surface area produce better matching detection. The low texture that contributes mismatching process caused by the plain colour and textureless surface regions and sometimes also produced a constant luminance in large areas is marked in the red box compared with other methods shown in Figure 4. The taxonomy performed poorly at low texture regions and occlusion areas, as shown around the yellow box marking. Consequently, it is more difficult and very tricky to develop the algorithm in larger low texture regions due to pixel intensities' likeness to each other. By visually evaluating the developed algorithm's disparity map, the taxonomy able to improve the edge-preserving properties in the disparity map and robust against the illumination variation and discontinuities areas. The disparity maps result has proved that the algorithm taxonomy increases the accuracy of the disparity map with other stereo vision algorithms.

Table 1 and Table 2 reports the Middlebury quantitative evaluation of the developed taxonomy for bad pixel errors and average errors assessed compared with several other state-of-the-art methods. The accomplishment of the proposed taxonomy from the experiment has been measured and compared with other local methods in the Middlebury Dataset. The outcomes show the taxonomy is among the lowest average errors that signify the proposed accuracy accomplishment competitiveness. All the compared algorithm except the proposed algorithm developed without the pre-processing stage and dynamic histogram matching cost. It indicates that the developed taxonomy is placed at the top of the comparison table, producing 41.0% and 33.0% of all and nonocc errors—the second method contributed by ADSR_GIF and trailed by ISM, and R-NCC. The developed taxonomy decreases in all errors with 5% ~ 10% and nonocc errors with 3% ~ 5%, respectively, compared to the other methods.

Table 1
Comparison Of The Results Of The Proposed Taxonomy With Other Published Taxonomys Using The Middlebury Training Database For Bad Pixel Error 2.0% Performance

Method	ISM		ADSR_GIF		R-NCC		Proposed	
	Non %	All %	Non %	All %	Non %	All %	Non %	All %
Adiron	37.3	40.0	43.6	40.0	26.2	54.3	**29.7**	**35.2**
ArtL	28.8	37.1	18.6	31.1	14.8	35.3	**22.8**	**38.5**
Jadepl	45.5	54.4	36.7	49.1	30.2	61.9	**38.2**	**52.0**
Motor	32.1	36.3	24.6	39.3	30.9	25.7	**27.4**	**34.9**
MotorE	34.3	38.5	58.6	45.1	72.9	30.2	**26.1**	**33.8**
Piano	44.7	48.1	22.8	42.3	41.6	46.2	**35.6**	**40.5**
PianoL	53.1	56.1	56.3	62.2	77.7	61.2	**47.6**	**51.5**
Pipes	34.2	43.1	49.7	35.1	64.1	46.7	**26.7**	**38.7**
Playrm	44.6	50.9	18.7	48.1	27.4	67.4	**43.3**	**50.9**
Playt	59.8	62.3	56.0	55.6	59.1	45.0	**52.0**	**56.9**
PlaytP	44.5	48.5	48.5	42.1	71.9	36.8	**34.1**	**40.9**
Recyc	38.1	40.5	32.2	39.1	50.9	47.6	**30.7**	**35.7**
Shelvs	51.8	53.0	24.5	55.2	33.9	53.4	**57.1**	**59.7**
Teddy	22.0	27.3	36.3	17.9	78.2	24.7	**15.5**	**24.3**
Vintge	55.0	57.7	79.1	61.8	80.8	81.5	**52.4**	**56.2**
Avg	39.5	44.3	37.1	41.8	48.4	45.1	**33.0**	**41.0**

It shows that the established technique is produced average results in the Table 2, producing 20.1% and 10.6% of all and nonocc errors for bad pixel errors. While the performance in average error is very good, all and nonocc errors difference with 5 ~ 10%% respectively compared to the ISM (9.65%, 6.45%), ADSR_GIF (19.8%, 11.3%) and R-NCC (22.9%, 19.8%). The taxonomy also shows a promising result of bad pixel error for images containing illumination variations and radiometric differences such as 'Recycle', 'Teddy' and 'Vintage'.

Table 2
Comparison Of The Results Of The Proposed Taxonomy With Other Published Taxonomys Using The Middlebury Training Database For Average Error Performance

Method	ISM		ADSR_GIF		R-NCC		Proposed	
	Non %	All %	Non %	All %	Non %	All %	Non %	All %
Adiron	3.63	4.64	4.84	6.40	20.5	21.2	**5.49**	**10.4**
ArtL	4.89	7.56	4.62	9.00	10.0	12.5	**7.17**	**23.8**
Jadepl	13.2	30.5	16.1	26.1	67.2	91.0	**26.1**	**53.3**
Motor	3.37	6.0	4.58	8.11	9.59	11.5	**4.37**	**12.1**
MotorE	3.35	6.01	7.22	11.4	10.6	12.7	**4.12**	**11.8**
Piano	5.22	6.20	5.20	6.15	9.12	9.59	**9.53**	**13.8**
PianoL	11.5	12.3	34.4	34.0	15.8	15.8	**22.6**	**26.0**
Pipes	6.31	11.7	7.53	14.9	21.8	27.9	**8.47**	**21.9**
Playrm	4.93	7.44	5.05	10.5	29.0	30.0	**10.7**	**26.0**
Playt	25.3	26.8	13.0	16.7	18.0	17.5	**22.9**	**29.1**
PlaytP	5.73	8.74	5.67	10.0	13.1	13.0	**7.31**	**14.4**
Recyc	3.45	3.96	3.37	4.20	22.3	22.2	**5.1**	**9.28**
Shelvs	8.40	8.71	9.49	9.97	11.5	11.7	**17**	**19.7**
Teddy	2.53	3.38	2.15	3.35	4.13	4.81	**4.04**	**13.2**
Vintge	7.83	8.81	9.64	10.9	44.3	45.1	**27.6**	**33.7**
Avg	6.45	9.65	7.81	11.3	19.8	22.9	**10.6**	**20.1**

The pre-processing stage with the matching costs in the proposed taxonomy is robust against the illumination variations, recognizing the regions with different intensity and brightness. The modified dynamic cost of matching cost computation presents a fine contour of disparity levels with the lowest noise displayed on the image. Additionally, the bilateral and median filter adoption also increased the efficiency and preserved the edges of an object.

V. CONCLUSION

An improved stereo matching taxonomy is formulated in this paper to improve disparity map accuracy, especially for radiometric differences and occlusion issues. The taxonomy achieved an impressive level of disparity accuracy with acceptable execution time. The algorithm comprises pre-processing using a Gaussian filter and a modified dynamic histogram with a census signature of matching cost with adaptive bilateral filter as cost aggregation, WTA for disparity selection and finalised using left-right consistency check with combination bilateral and median filter as the disparity refinement. The accurate disparity map obtained and competitive in the experiment compared with several other stereo matching algorithms in the Middlebury Evaluation Dataset. In the future, we will adopt a local texture histogram taxonomy to analyse the texture histogram, which improves the low texture region effectively.

REFERENCES

[1] H. Liu, R. Wang, Y. Xia, and X. Zhang. 2020. Improved cost computation and adaptive shape guided filter for local stereo matching of low texture stereo images, Appl. Sci., vol. 10, no. 5, pp. 1869(1)-1869(17).

[2] R. A. Hamzah and H. Ibrahim. 2016. Literature survey on stereo vision disparity map algorithms, J. Sensors, vol. 2016, pp. l–22.

[3] I. Cabezas, V. Padilla, and M. Trujillo. 2011. A measure for accuracy disparity maps evaluation, in Iberoamerican Congress on Pattern Recognition, pp. 223–231.

[4] D. Scharstein et al. 2014. High-resolution stereo datasets with subpixel-accurate ground truth, in German conference on pattern recognition, pp. 31–42.

[5] B. Jia, S. Liu, and Z. Du. 2016. A progressive framework for dense stereo matching, Pattern Recognit. Image Anal., vol. 26, no. 2, pp. 294–301.

[6] R. R. Basaru, C. Child, E. Alonso, and G. Slabaugh. 2015. Quantized Census for Stereoscopic Image Matching, in Proceedings - 2014 International Conference on 3D Vision Workshops, 3DV 2014, pp. 22–29.

[7] R. A. Hamzah, and H. Ibrahim. 2020. Improvement of disparity map refinement stage using adaptive least square plane fitting technique. Electronics Letters, vol. 56, no. 18, pp. 918-920.

[8] M. Dumont, P. Goorts, S. Maesen, G. Lafruit, and P. Bekaert. 2015. Real-time edge-sensitive local stereo matching with iterative disparity refinement, in Communications in Computer and Information Science, pp. 435–456.

[9] A. Buades and G. Facciolo. 2015. Reliable multiscale and multiwindow stereo matching, SIAM J. Imaging Sci., vol. 8, no. 2, pp. 888–915.

[10] U. Raghavendra, K. Makkithaya, and A. K. Karunakar. 2015. Anchor-diagonal-based shape adaptive local support region for efficient stereo matching, Signal, Image Video Process., vol. 9, no. 4, pp. 893–901.

[11] S. Zhu and L. Yan. 2017. Local stereo matching algorithm with efficient matching cost and adaptive guided image filter, Vis. Comput., vol. 33, no. 9, pp. 1087–1102.

[12] L. Li, S. Zhang, X. Yu, and L. Zhang. 2018. PMSC: PatchMatch-Based Superpixel Cut for Accurate Stereo Matching, IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 3, pp. 679–692.

[13] M. G. Mozerov and J. Van De Weijer. 2019. One-View Occlusion Detection for Stereo Matching with a Fully Connected CRF Model, IEEE Trans. Image Process.

[14] T. Matsuo, S. Fujita, N. Fukushima, and Y. Ishibashi. 2015. Efficient edge-awareness propagation via single-map filtering for edge-preserving stereo matching, Three-Dimensional Image Process. Meas. (3DIPM), Appl. 2015, vol. 9393, p. 9393OS.

[15] H. Ma, S. Zheng, C. Li, Y. Li, L. Gui, and R. Huang. 2017. Cross-scale cost aggregation integrating intrascale smoothness constraint with weighted least squares in stereo matching, J. Opt. Soc. Am. A, vol. 34, no. 4, pp. 648–656.

[16] A. Buades and G. Facciolo. 2015. Reliable multiscale and multiwindow stereo matching, SIAM J. Imaging Sci., vol. 8, no. 2, pp. 888–915.

[17] O. Zeglazi, M. Rziza, A. Amine, and C. Demonceaux. 2018. A hierarchical stereo matching algorithm based on adaptive support region aggregation method, Pattern Recognit. Lett., vol. 112, pp. 205–211.

[18] J. Lu, Y. Li, H. Yang, D. Min, W. Eng, and M. N. Do. 2017. PatchMatch Filter: Edge-Aware Filtering Meets Randomized Search for Visual Correspondence, IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 9, pp. 1866–1879.

[19] H. Shi, H. Zhu, J. Wang, S. Y. Yu, and Z. F. Fu. 2016. Segment-based adaptive window and multi-feature fusion for stereo matching, J. Algorithms Comput. Technol., vol. 10, no. 1, pp. 3–11.

[20] C. Stentoumis, L. Grammatikopoulos, I. Kalisperakis, G. Karras, and E. Petsa. 2015. Stereo matching based on census transformation of image gradients.

[21] C. H. Lin and C. W. Liu. 2015. Accurate stereo matching algorithm based on cost aggregation with adaptive support weight, Imaging Sci. J., vol. 63, no. 8, pp. 423–432.

[22] Y. Zhan, Y. Gu, K. Huang, C. Zhang, and K. Hu. 2016. Accurate Image-Guided Stereo Matching with Efficient Matching Cost and Disparity Refinement, IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 9, pp. 1632–1645.

[23] S. S. A. Razak, M. A. Othman, and A. F. Kadmin. 2019. The effect of adaptive weighted bilateral filter on stereo matching algorithm, Int. J. Eng. Adv. Technol.

[24] R. A. Hamzah, M. G. Y. Wei and N. S. N. Anwar. 2019. Stereo matching based on absolute differences for multiple objects detection. Telkomnika, vol. 17, no. 1, pp. 261-267.

[25] X. Huang and Y. J. Zhang. 2016. An O(1) disparity refinement method for stereo matching, Pattern Recognit., vol. 55, pp. 198–206.

[26] C. Zhang, C. He, Z. Chen, W. Liu, M. Li, and J. Wu. 2019. Edge-Preserving Stereo Matching Using Minimum Spanning Tree, IEEE Access, vol. 7, pp. 177909–177921.

[27] J. Zhang, Z. Liu, J. F. Nezan, and G. Zhang. 2018. Correspondence matching among stereo images with object flow and minimum spanning tree aggregation, Int. J. Adv. Robot. Syst., vol. 15, no. 2, p. 1729881418760986.

[28] Y. Xiao et al. 2020. Confidence Map Based 3D Cost Aggregation with Multiple Minimum Spanning Trees for Stereo Matching, in Asian Conference on Pattern Recognition, pp. 355–365.

[29] S. Chen, J. Zhang, and M. Jin. 2020. A simplified ICA-based local similarity stereo matching, Vis. Comput., vol. 1, pp. 1–9.

[30] Y. Zhang, Y. Li, Y. Kong, and B. Liu. 2020. Attention Aggregation Encoder-Decoder Network Framework for Stereo Matching, IEEE Signal Process. Lett., vol. 27, pp. 760–764.

[31] Y. Zhang et al. 2018. Activestereonet: End-to-end self-supervised learning for active stereo systems, in Proceedings of the European Conference on Computer Vision (ECCV), pp. 784–801.

[32] T. A. Chang, X. Lu, and J. F. Yang. 2017. Robust stereo matching with trinary cross color census and triple image-based refinements, EURASIP J. Adv. Signal Process., vol. 2017(1), no. 27.

[33] Y. J. Jeong and C. C. Jay Kuo. 2019. Stereo Matching with Confidence-Region Decomposition and Processing, J. Electr. Eng. Technol., vol. 14, no. 1, pp. 463–469.

[34] S. Zhu, D. Cao, Y. Wu, and S. Jiang. 2015. Edge-aware dynamic programming-based cost aggregation for robust stereo matching, J. Electron. Imaging.

[35] Z. Zhou, D. Wu, and Z. Zhu. 2016. Stereo matching using dynamic programming based on differential smoothing, Optik (Stuttg)., vol. 127, no. 4, pp. 2287–2293.

[36] R. H. Zhu, G. Y. Ge, G. S. Zhang, Z. Shen, and Q. Sun. 2018. Semi-global stereo matching algorithm based on AD-Census transform and multi-scan line optimization, Guangdianzi Jiguang/Journal Optoelectron. Laser.

[37] J. Jang, D. W. Shin, and Y. S. Ho. 2018. Disparity map refinement method using coarse-to-fine image segmentation, in Proceedings - 9th Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) 2017.

[38] H. Li, L. Chen, and F. Li. 2019. An Efficient Dense Stereo Matching Method for Planetary Rover, IEEE Access, vol. 7, pp. 48551–48564.

[39] M. Tian, B. Yang, C. Chen, R. Huang, and L. Huo. 2019. HPM-TDP: An efficient hierarchical PatchMatch depth estimation approach using tree dynamic programming, ISPRS J. Photogramm. Remote Sens., vol. 155, pp. 37–57.

[40] W. Wu, L. Li, and W. Jin. 2016. Disparity refinement based on segment-tree and fast weighted median filter.

[41] I. Gallo and E. Binaghi. 2005. A new algorithm for dense two-frame stereo correspondence, Int. Conf. Comput. Vis. Theory Appl.

[42] M. A. Al Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae. 2007. A dynamic histogram equalization for image contrast enhancement, IEEE Trans. Consum. Electron.

[43] B. Subramani and M. Veluchamy. 2020. Quadrant dynamic clipped histogram equalization with gamma correction for color image enhancement, Color Res. Appl.

[44] S. K. P. Kaur. 2016. A Novel Brightness Preserving Histogram Equalization Technique for Image Contrast Enhancement, Int. J. Sci. Res.

[45] J. Lim, Y. Kim, and S. Lee. 2017. A census transform-based robust stereo matching under radiometric changes.

[46] C. Ttofis, C. Kyrkou, and T. Theocharides. 2016. A Low-Cost Real-Time Embedded Stereo Vision System for Accurate Disparity Estimation Based on Guided Image Filtering, IEEE Trans. Comput.

[47] Y. Gan, R. A. Hamzah, and N. S. N. Anwar. 2018. Local Stereo Matching Algorithm Based on Pixel Difference Adjustment, Minimum Spanning Tree and Weighted Median Filter.

[48] Y. Du and K. Jia. 2019. Neighborhood correlation and window adaptive stereo matching algorithm, J. Inf. Hiding Multimed. Signal Process., vol. 10, no. 4, pp. 509–516.

[49] G. A. Kordelas, D. S. Alexiadis, P. Daras, and E. Izquierdo. 2015. Enhanced disparity estimation in stereo images, Image Vis. Comput., vol. 35, pp. 31–49.

[50] D. S. and R. Szeliski, Middlebury Stereo Evaluation - Version 3, 2020. https://vision.middlebury.edu/stereo/eval3/.

[51] R. A. Hamzah, A. F. Kadmin, M. S. Hamid, S. F. A. Ghani, and H. Ibrahim. 2018. Improvement of stereo matching algorithm for 3D surface reconstruction, Signal Process. Image Commun., vol. 65, pp. 165–172.

[52] L. Kong, J. Zhu, and S. Ying. 2020. Stereo Matching Based on Guidance Image and Adaptive Support Region, Guangxue Xuebao/Acta Opt. Sin.

[53] Y. Li and S. Fang. 2019. Removal-based multi-view stereo using a window-based matching method, Optik (Stuttg)., vol. 178, pp. 1318–1336.