

Radiometric Invariant Dense Disparity Estimation for Real Time Stereo Correspondence

^{1,2}Deepambika V.A. and ^{3,4}M. Abdul Rahiman

¹Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India

²Department of ECE, LBS Institute of Technology for Women, Trivandrum, Kerala, India

³Managing Director C-Apt Kerala, India

⁴Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India

Article history

Received: 29-10-2018

Revised: 13-02-2019

Accepted: 11-04-2019

Corresponding Author:

Deepambika V.A.

Department of ECE, LBS

Institute of Technology for

Women, Trivandrum, Kerala,

India

Email: deepambika.va@gmail.com

Abstract: Computer stereo vision tries to mimic human vision by grabbing multiple views of the same scene and cognizing it. The stereo correspondence will find out the matching pixels between the two views based on the Lambertian criteria, which results in disparity. The distance of the objects from the camera can be calculated using this disparity. But in the real world scenario, this Lambertian assumption may not work always due to the radiometric variation between the image pairs and the conventional approaches results in erroneous disparity. In this work, for doing the radiometric invariant stereo matching, the simple local binary pattern is used. The correspondence is done by using semi global block matching method, which can handle the depth changes of curved surfaces and slanting surfaces by adding suitable penalty terms. The performance evaluation of the proposed shows lesser error rate in the range of 0.14% - 0.3883% and run time requirement of 0.20 milliseconds only. This radiometric invariant stereo correspondence attains accuracy as that of global method with run time speed as that of local method and is suitable for most of the real time stereo vision applications.

Keywords: Radiometric Invariant, Stereo Matching, Local Binary Pattern, Semi Global Block Matching, Dense Disparity

Introduction

In recent years, binocular stereo vision has been playing a major role in computer vision applications such as robotic vision, medical imaging, autonomous vehicles and augmented reality. Stereo vision tries to achieve the abilities of human vision by electronically grabbing and cognizing the images (Szeliski, 2010). By measuring the difference in relative positions between the matching pixels, the stereo correspondence can infer depth of the objects from two views of a scene. The matching is done based on a cost function that gives the similarity measure between the conjugate pairs. The aggregation of this cost results in the Disparity Space Image (DSI). Using this DSI, the depth of objects from the camera can be calculated with the help of triangulation. The correspondence can be of either local methods or global methods (Scharstein and Szeliski, 2002). Often, the real time applications in computer vision faces the trade-off between speed and accuracy, hence they usually rely on local area based

correspondence. The crux of stereo correspondence always lies in finding the similarity between two images under challenging conditions such as high dynamic range imaging (Park *et al.*, 2017), radiometric variations, occlusions, textureless regions and repetitive patterns. The radiometric variations can be due to the change in illumination conditions or the camera exposure time variations. In such real world scenario, the traditional approaches for the stereo correspondence will result in increased error. The common non-parametric approaches to tackle these radiometric variations are Census and rank transforms (Ramin and Woodfill, 1994) followed by a window based cost aggregation. Most of the approaches suggested by the stereo vision researchers to handle these kinds of challenges are done at high computational expense. In single view still image applications, a number of binary operators are used to handle non-uniform texture regions that arises from the sudden illumination changes. Of these the Local Binary Pattern (LBP) operator transforms these images captured under radiometric variations to an illumination invariant one in a better

manner (Ojala *et al.*, 2001). This work incorporates simpler and accurate LBP method for the radiometric invariant conversion of stereo images.

The cost aggregation strategies of local stereo correspondence are faster than the global approaches. Hence, local approaches are used for real time applications where the speed of the algorithm is the main concern. Common cost functions used for the correspondence search are Sum of Absolute Differences (SAD) (Hamzah and Hamzah, 2016), Sum of Squared Differences (SSD) (Marghany *et al.*, 2011) and Normalized Cross Correlation (NCC) (Shen, 2011) or their adaptive versions (Yoon and Kweon, 2006) on both gray scale and RGB images. Of these, the correlation window approach, also known as area based approach assign the same disparity value for all the pixels within a window. This approach is based on the planar surface assumption, i.e., all the pixels within a window are assumed to be at same distances from the camera. They can provide dense disparity map but fail to handle small depth changes such as slanted and curved surfaces. Semi-Global Methods (SGM) takes the advantages of global and local correspondences (Hirschmuller, 2005), which use the pixel wise matching as in global methods and can provide faster results as in local methods. Therefore the planar surface assumption used by the local correspondence search may not always works. Compared to global methods, the semi global matching can give faster results, which can meet the real-time requirements. Also semi global approaches always outperform the local methods in terms of its accuracy. By incorporating the window based approach to SGM method, the speed of the algorithm can be further increased. This work uses the faster window based approach of semi-global matching algorithm, termed as Semi-global block matching (SGBM) method. The algorithm can attain accuracy as in global methods without sacrificing the run time speed of local methods. Often, the real world scenes contains curved surfaces and slanting surfaces such as the edge of a bottle and the roof of a building. To handle the small and large depth discontinuities of such surfaces, smaller and larger penalty terms have been used to get accurate depth results. In this system, the LBP will transform the input stereo pair to radiometric invariant form. SGBM can handle curved and slanting surfaces and can provide accurate dense disparity map at a higher frame rate. Thus a new LBP-SGBM method is presented here for the accurate radiometric invariant stereo correspondence.

This work targets in the formulation of stereo vision system for autonomous robots, which is suitable for various real-time indoor and outdoor applications. This passive range finding system can handle curved surfaces such as a bottle and can be used for the pick and place robots. It can also be used as the sensor for the navigation of autonomous robots in unknown environments.

Related Work

Based on the application requirements, a large number of stereo correspondence approaches have been developed in the recent years. But most of the local stereo correspondence methods are done based on the assumption that the pair of input images are taken under same radiometric conditions. But in real world scenario, this may not happen. There can be radiometric variations due to illumination changes and camera exposure time variations. In such conditions, the conventional stereo correspondence approaches can give only inaccurate disparity results. To handle these challenges, a couple of pre-processing steps have been proposed by some researchers. Two non-parametric transforms are proposed for visual correspondence in (Ramin and Woodfill, 1994). Mutual information have been used for the radiometric invariant visual correspondence in (Mustafa and Kalkan, 2015), (Guanying *et al.*, 2018). A contrast invariant local stereo correspondence employs different spatial frequency channels in (Ogale and Aloimonos, 2005). Binary descriptor-based line scan stereo correspondence is presented in (Valentín *et al.*, 2017). ANCC have been proposed for radiometric invariant stereo matching (Heo *et al.*, 2011). Combined census and adaptive window for handling radiometric changes is presented in (San and War, 2017). Census filtering and hamming distance based cost aggregation is proposed in (Sarika *et al.*, 2015). Evaluation of stereo matching cost for varying radiometric images is introduced in (Hirschmuller and Scharstein, 2009; Stolc *et al.*, 2016) proposes stochastic binary local descriptor for the stereo correspondence. LBP is used for vehicle detection in (Neumann *et al.*, 2017). Often the real time applications rely on dense local correspondence methods as they can provide decent disparity results with higher frame rate. But they fail to handle small depth discontinuities as in curved surface and in slanting surface. Semi global methods use penalty terms for dealing with these depth discontinuities (Hirschmüller *et al.*, 2012). Khomutenko *et al.* (2016) used direct fish eye stereo matching with semi-global approach. Disparity steps changes are handled by using penalty terms in (Taimoor and Afanasyev, 2018). The iterative SGM method uses neighbourhood disparity values to get the distance maps (Hermann and Klette, 2012). SGM with surface orientation priories is used for stereo matching (Daniel *et al.*, Scharstein, 2017). Eunah *et al.* (2017) proposes a hierarchical stereo matching approach for low resolution images. Andreas *et al.* (2010) uses triangulation on a set of support points for robust stereo matching. Qingxiong (2012) presents an adaptive non-local cost aggregation approach for the stereo matching. Mei *et al.* (2013) proposes tree-based cost aggregation with segment tree construction algorithm (ST-1) and aggregation with enhanced segment tree algorithm (ST-2). Shenyong *et al.* (2017) presents a non-local stereo matching method with an initial cost and multiple weight.

Most of these approaches are memory intensive and their complexity will depend on the complexity of the scene. The main objective of this proposed system is the development of an accurate and radiometric invariant stereo correspondence that can give precise disparity maps at depth discontinuities such as curved surfaces and slanting surfaces without sacrificing the run time speed of local methods.

Features and Contributions

This is the prime stereo matching approach that exploits the advantages of LBP, local correspondence and global method. Contributions of this work include:

- This method incorporates computationally simpler yet efficient local binary pattern for getting the radiometric invariant stereo matching. This highly discriminative operator will convert the input image to an illumination invariant one
- For getting the accuracy and run time speed, this framework uses semi global block matching stereo correspondence. This SGBM approach is not based on planar surface assumption as in conventional window based approaches. It uses the smaller and larger penalty terms for handling the depth discontinuities in the curved surfaces and slanting surfaces

On error evaluation with different input images taken under different radiometric conditions, this fast and accurate stereo matching approach shows less error rate in the range of 0.14%-0.3883% and run time requirement of 0.20 milliseconds only. The performance analysis

shows that, the algorithm gives stable accurate results. It is insensitive to various parameters such as selection of window size, choice of penalty functions and radiometric changes. This method is suitable for most of the range finding real time applications.

Radiometric Invariant Stereo Matching

The proposed radiometric invariant stereo matching consist of transformation of input image pair to LBP followed by SGBM method. Figure 1 shows the system overview.

Local Binary Patterns

The common form of Local binary pattern assigns values for the image pixels with in a 3x3 window, based on a threshold level. This operator can be used for different sizes (Ojala *et al.*, 2002) by using circular neighbourhood of any radius. The LBP operator can be derived as follows.

Assume the gray scale image $I_G(X, Y)$ and let the g_p represent the intensity value of a pixel at position (x, y) . If g_s represents the intensity value of sample points in equal radius r from the centre point C , then:

$$g_s = I_G(x_s, y_s) \quad s = 0, 1, \dots, C - 1 \quad (1)$$

$$X_s = X + r \cos (2 \pi s / C) \quad (2)$$

$$Y_s = Y - r \sin (2 \pi s / C) \quad (3)$$

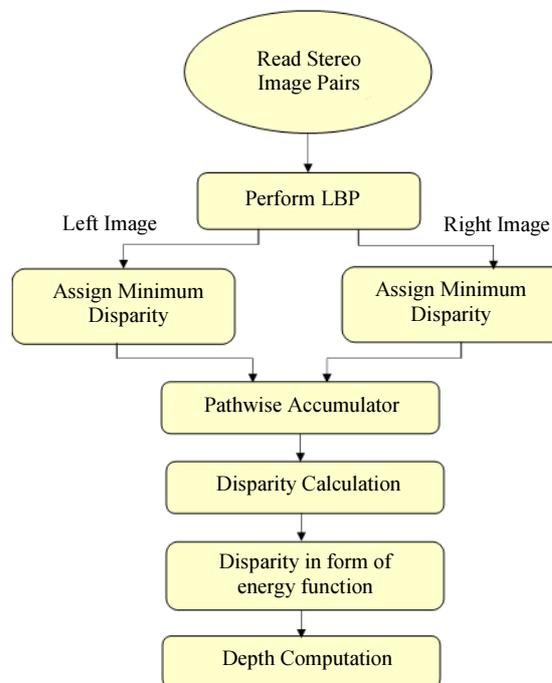


Fig. 1: System overview of LBP-SGBM algorithm

Let us assume that the texture of the image $I_G(X, Y)$ with in an area is described by intensity distribution of $C+I_G(C > 0)$:

$$T = t(g_p, g_0, g_1, \dots, g_{s-1}) \quad (4)$$

If the difference between the gray level value of centre pixel and the neighbouring pixels are taken, then the gradient information will be retained:

$$T = t(g_p, g_0 - g_p, g_1 - g_p, \dots, g_{s-1} - g_p) \quad (5)$$

Consider that centre pixel is independent of these differences, then the Equation (5) becomes:

$$T \approx t(g_p) t(g_0 - g_p, g_1 - g_p, \dots, g_{s-1} - g_p) \quad (6)$$

The term $t(g_p)$ is the intensity value at the centre point of $I_G(x, y)$ and it does not contain any information regarding the texture variations. But the second factor in the Equation (6) gives the texture variations in that area. But the extraction of this texture details from the image is a difficult task. To meet this, vector quantization was used by (Li *et al.*, 2012). In this the features points are quantized to reduce the dimensionality. But in this type of quantization, though $g_s - g_p$ are invariant to mean gray level of image, this will not hold for other gray value changes. To achieve the invariance, only the signs of the differences are taken:

$$t(s(g_0 - g_p), s(g_1 - g_p), \dots, s(g_{s-1} - g_p)) \quad (7)$$

The thresholding function:

$$H(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases} \quad (8)$$

Taking the weighed sum of threshold differences, the modified LBP operator can be written as:

$$LBP_{S,R}(X_s, Y_s) = \sum_{s=0}^{s-1} H(g_s - g_p) 2^s \quad (9)$$

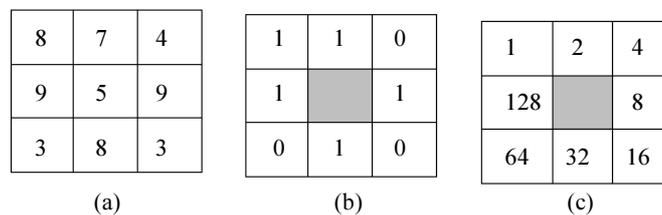
The local binary pattern operator transforms an image into an image of integer labels, describing small-scale appearance of the image.

Figure 2 shows the formation of LBP codes centered on a pixel for a 3x3 fragment of an image. Here the intensity level of centre pixel is compared with its eight neighbourhood ones, in clockwise direction from top-left corner of the block. The encoding is done based on the Equation (8). The resultant pattern is shown in Fig. 2b, in which the smaller and higher values compared to the centre pixels are encoded as 0 and 1 respectively. The weight assigned to each pixel in the manner of 2^s is shown in Fig. 2c. This block level operation can be completed for entire image using sliding window method.

Semi-Global Block Matching Algorithm

SGBM is the faster window based approach of SGM. It is performed by taking the mean data from each block as pixel energy. In this method the cost arrays are formed by arranging the cost values corresponding to minimum disparity value to the maximum disparity value for the entire pixels. The matching cost is based on (Hernandez-Juarez, 2016; Birchfield and Tomasi, 1999). A search is carried out in five directions for the minimum path cost based on the pixel energy. Penalties are added for the disparity differences with the neighbouring pixels along the path. The cost minimum corresponds to the best matching pixels and the disparity value is computed for this minimum cost value.

The algorithm is insensitive to various parameters such as selection of window size, choice of penalty terms and radiometric changes. The streaking effect that happens in global method has been eliminated in SGBM by symmetrically computing the matching cost through multiple paths instead of single scan line as shown in Fig. 3.



$$\begin{aligned} \text{Pattern} &= 10101001 \\ \text{LBP} &= 1+2+8+32+128 = 171 \\ C &= (8+7+9+8+9)/5 - (4+3+3)/4 = 5.7 \end{aligned}$$

Fig. 2: Calculation of LBP codes (a) Image block (b) Pattern ($S(g_s - g_p)$) (c) Weights (2^s)

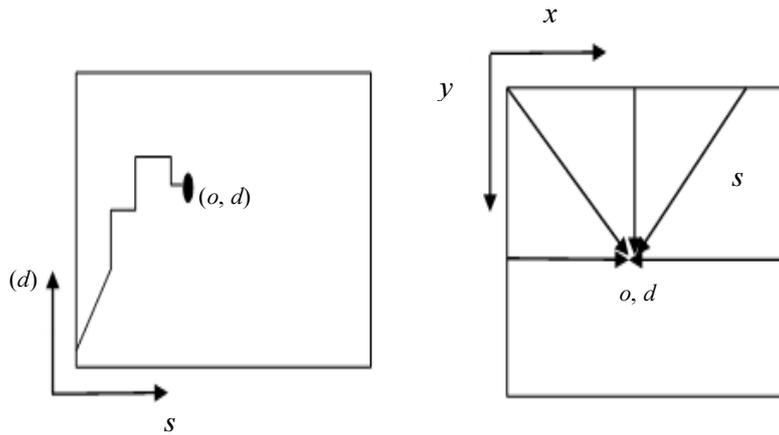


Fig. 3: Optimisation paths for the cost aggregation

The path accumulation along the direction ‘s’ is given by $L'_s(o, d)$:

$$\begin{aligned}
 L'_s(o, d) &= C(o, d) + \min(L_s(o-s, d), \\
 &L'_s(o-s, d-1) + P_1, \\
 &L'_s(o-s, d+1) + P_1, \\
 &\min_i L'_s(o-s, i) + P_2
 \end{aligned} \tag{10}$$

The first term, $C(s, d)$ is the cost function for the similarity measure, which can be calculated as:

$$C(o, d) = \min(d(o, o-d, I_L I_R), d(o-d, o, I_R I_L)) \tag{11}$$

Here the I_L and I_R are the left and right input images. The difference in intensity level between the two neighbouring pixels of the left image is given as:

$$d(o, o-d, I_L I_R) = \min_{o-d-0.5 \leq o-d+0.5} |I_L(o) - I_L| \tag{12}$$

The second term in the Equation (10) gives the regularity of the disparity field. The cost value of current pixel, $C(o, d)$ is added to the previous pixel cost, along the direction ‘s’ will give the final cost for finding the disparity value. The following cost values are used for this:

- The cost at the pixel with disparity values $d-1$ and $d+1$ with smaller penalty term P_1 (for handling smaller disparity steps) gets added
- The cost at the previous pixel with disparity values less than $d-1$ and greater than $d+1$ with higher penalty term P_2 (for handling larger disparity steps) gets added
- (iii) The minimum value of the previous pixel is subtracted in order to limit the increasing value of

$L'_s(p, d)$ along the path. The maximum value of $L'_s(o, d)$ is limited by $C_{max} + P_2$, where C_{max} is the maximum cost

The penalty term P_1 is added for dealing with small disparity changes such as curved surfaces and slanted surfaces. Term P_2 is added for higher disparity changes such as depth discontinuities at object boundaries. On aggregation of the cost value in all 5 directions will give the final disparity map. The aggregated cost for the pixel p with disparity d in the direction of s is given as:

$$C_T(o, d) = \sum_s L'_s \tag{13}$$

Semi global method targets to minimize the global energy function E_G , for the disparity space image, D :

$$E_G(D) = \sum_p \left(C_T(o, D_o) + \sum_{q \in N_o} P_1 I[|D_o - D_q| = 1] + \sum_{q \in N_o} P_2 I[|D_o - D_q| > 1] \right) \tag{14}$$

Here o and q are pixel indices in the image and N_o is the number of pixels in the neighbourhood of the pixel o . $C_T(o, D_o)$ is the pixel-wise cost over the entire image with disparity D_o . P_1 is the penalty for small changes in disparity value, where the change in disparity to the neighboring pixel is not more than one pixel. P_2 is the penalty term for higher disparity changes to neighbouring pixels.

In block based stereo matching, a window of desired size centred at a pixel of interest is taken from the reference image. The similarity check is done using this window to find its conjugate pair in the target image. Small correlation windows are preferred for getting faster results and can preserve the sharp object boundaries. But they lack in accuracy. On the other hand

large windows give more accurate disparity results. Here the correspondence search will be much slower because of their iterative nature. Often there exists a trade-off between accuracy and run time.

Parameters used for SGBM Implementation

- (a) minDisparity – It is the minimum value assigned to the disparity levels. Normally it is taken as zero. If there is any pixel shift occurring after the rectification, a minimum disparity is to be added to the obtained disparity in order to compensate for this shift
- (b) numDisparity – Numdisparity shows the number of disparity levels in the obtained disparity map. The difference between maximum and minimum disparity value gives the Numdisparity
- (c) Block Size – It is the size of the window used for stereo matching. It should be of sizes 3, 5, 7... to 15
- (d) P_1 – To handle small depth discontinuities such as slanted surfaces and curved surfaces, small penalty term P_1 is used
- (e) P_2 – To handle higher depth discontinuities such as object boundaries, higher penalty term P_2 is used. The value of P_2 will be always larger than the value of P_1
- (f) disp12MaxDiff – Maximum permissible difference in the left-right disparity check. To disable check, this value is set as a non-positive one
- (g) preFilterCap – It is the clipping level of the pre-filtered image. The algorithm first takes the x-derivative of each pixel value and clips off it within the interval $[-preFilterCap, +preFilterCap]$. Cost computation is done for these filtered values
- (h) uniquenessRatio-It is the percentage margin allotted for the cost minimum. Cost minimum is taken based on “winner-takes-all” optimisation strategy for a block of preferred size. The margin for the cost variation is in between 5-15
- (i) speckleWindowSize -It is the maximum size of smooth disparity regions, while considering the speckle noise in that region. To disable speckle filtering, set it as 0. Otherwise, set the size in 50-200 range
- (j) speckleRange – It gives the maximum variation of disparity levels within each connected components

Result and Analysis

The algorithm is implemented and tested in Open CV 2.1.12 on Intel Core i3 processor with 2.10 GHz clock frequency and 4GB of RAM. Rectified stereo images from Middlebury data sets are used as input. (<http://vision.middlebury.edu/stereo/data/>).

The algorithm has been tested for different stereo input image pairs, with variable window sizes varying from 3,5,7 up to 15.

Performance Evaluation

For assessing the performance of the proposed method, it has been evaluated based on the accuracy, amount of bad matching pixel (Scharstein and Szeliski, 2002) and the execution time. Analysis is done based on ground truth disparity map. The Root Mean Square Error (RMSE) in the obtained disparity can be calculated using the ground truth disparity map (Scharstein and Szeliski, 2003) based on the Equation (15):

$$E = \left(\frac{1}{N} \sum_{(i,j)} |d_c(i,j) - d_t(i,j)|^2 \right)^{\frac{1}{2}} \quad (15)$$

Here ‘N’ is the number of pixels in disparity maps.

The number of bad matching pixels in the computed disparity result can be calculated using the Equation (16):

$$B = \frac{1}{N} \sum_{(i,j)} |d_c(i,j) - d_t(i,j)| > \delta_d \quad (16)$$

where, δ_d is the threshold level of the bad matching pixels. Here it is taken as one.

The evaluation of the proposed method is carried by using different stereo image pairs taken with varying illumination condition and exposure time.

Figure 4a-d shows the input image, ground truth disparity, LBP image and the disparity map results obtained for LBP-SGBM. Figure 5 shows the results of LBP-SGBM stereo correspondence on image pair Art with varying (left illumination/right illumination). Illumination combinations used for the correspondence search are (1/1), (1/2), (1/3), (2/2), (2/3) and (3/3). Figure 6a-i shows the percentage RMS error in the obtained disparity map results of images- Art [463×370], Cones [450×375] and Reindeer [447×370], taken under Illumination 1, 2 and 3 for a window size of 5×5.

The error evaluation shows a minimum error of 0.175167% and a maximum error of 0.3083% in the result of image Art taken under illumination 1 and Reindeer taken under illumination 3. Table 1 shows the percentage of bad matching pixels. Figure 7 shows the RMS Error in the obtained disparity map of different input images, Aloe, Baby Cones, Bowling, Wood and Tsukuba with varying window sizes.

Table 2 and 3 shows the performance evaluation based on RMS error and bad matching pixels for image pairs taken with different exposure time. In this the image Art shows a minimum error of 0.198067% under illumination 2 and image Reindeer shows a maximum error of 0.3883% under illumination 3.

Table 4 shows the amount of bad matching pixels in image pairs Aloe, Art, Mobies and Dolls with varying (left illumination/right illumination). Illumination combinations used are (1/1), (1/2), (1/3), (2/2), (2/3) and (3/3).

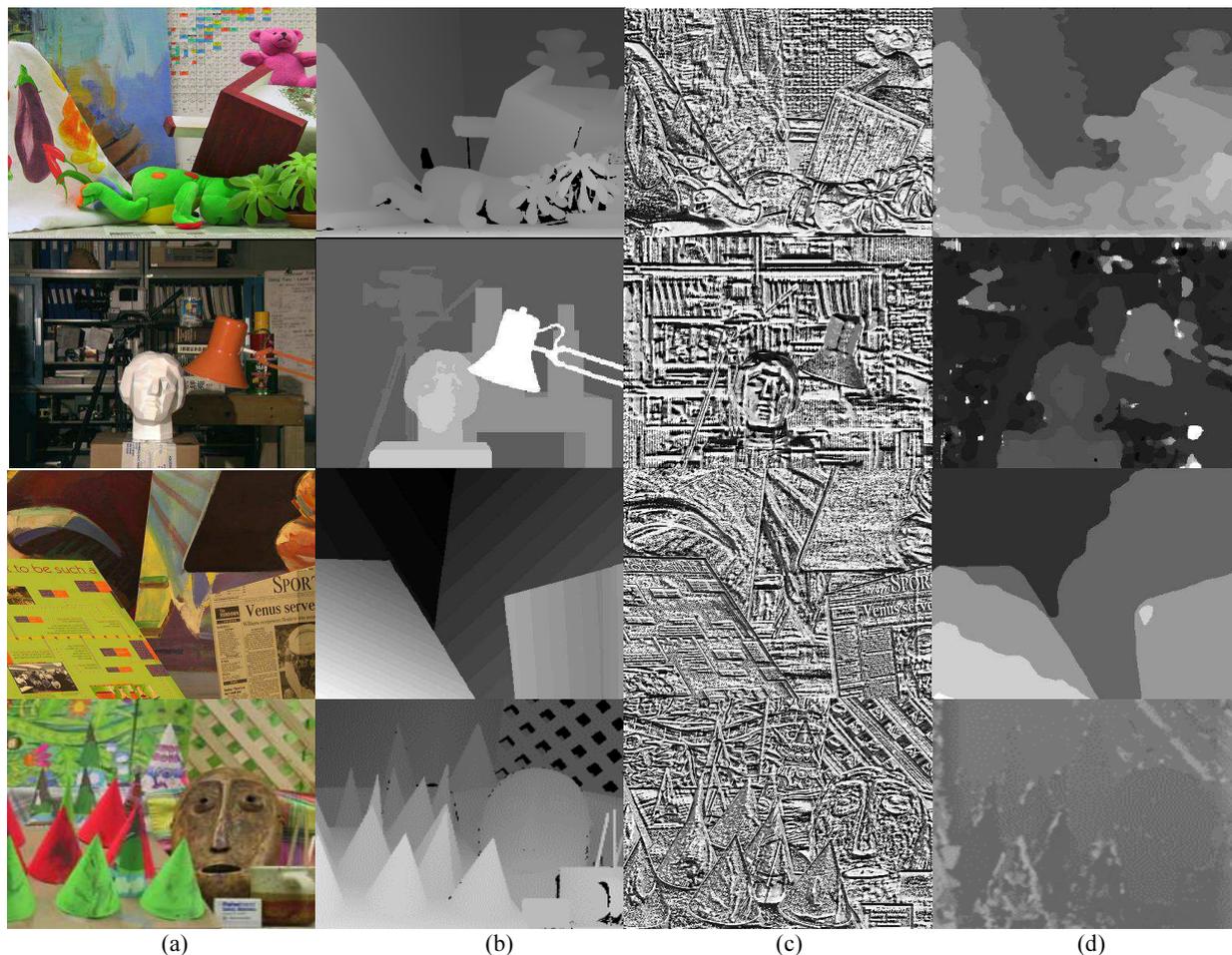


Fig. 4: Results of LBP-SGBM (a) Input images –Left view (b) Ground truth disparity (c) LBP image (d) Result of LBP SGBM

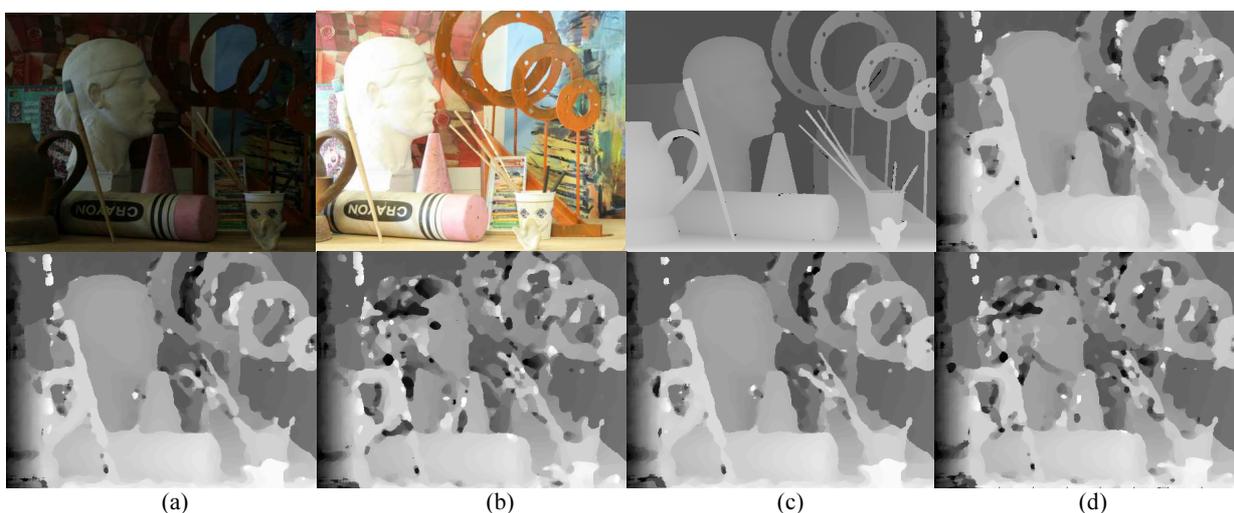


Fig. 5:Results of LBP-SGBM stereo correspondence on image pair Art with varying illumination. (a) The left image with illumination (1) - exposure (0) (b) The right image with illumination (3) - exposure (0), (c) Ground truth disparity. (d) to (i) obtained disparity map with varying (left illumination/right illumination). Illumination combinations are (1/1), (1/2), (1/3), (2/2) and (2/3) respectively

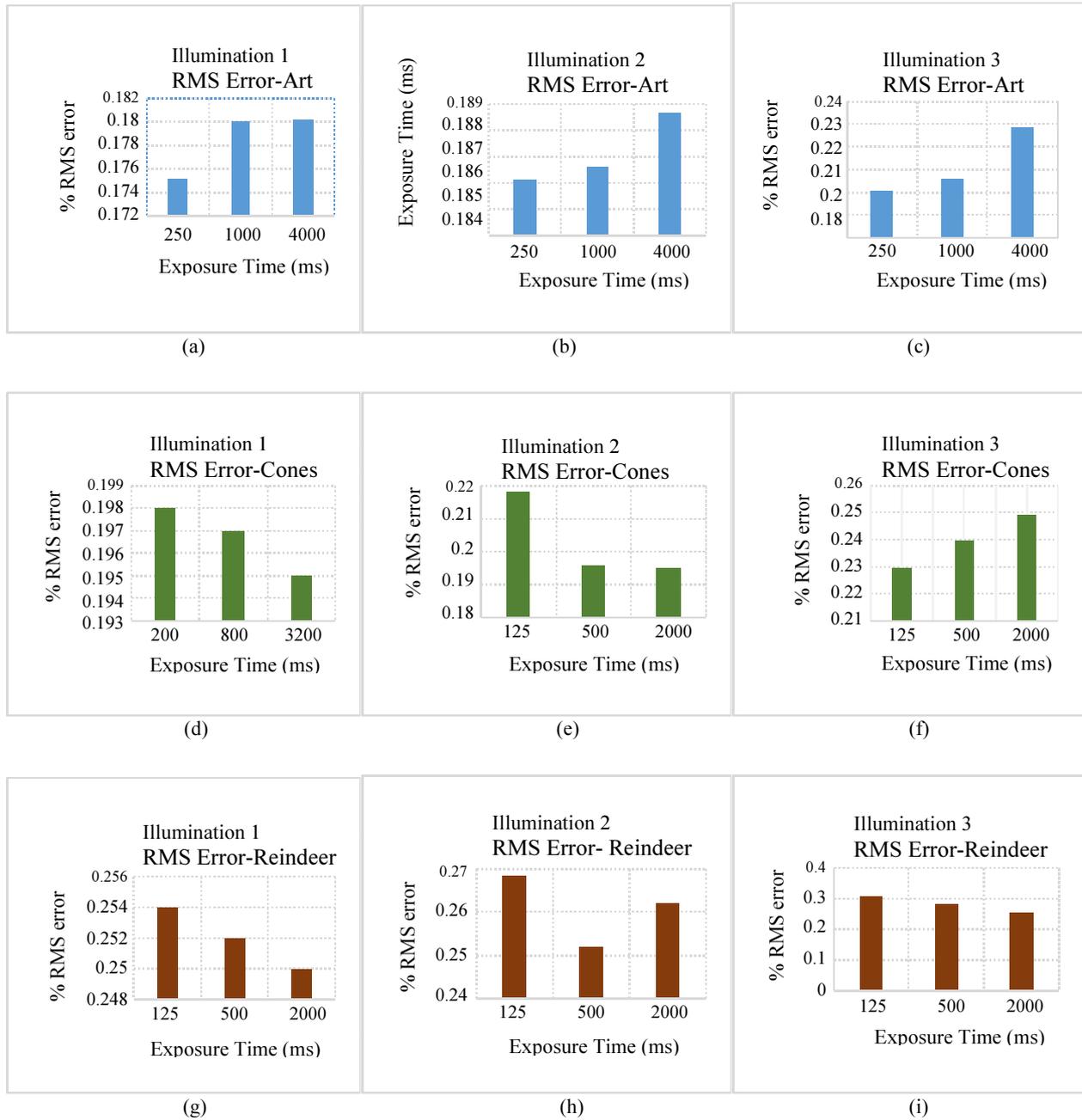


Fig. 6: RMS Error in the obtained disparity map of different input images. (a), (b) and (c) using the image Art. (d), (e) and (f) using the image Cones (g), (h) and (i) using the image Reindeer, under Illumination 1, 2 and 3

Table 1: Percentage of bad matching pixels under varying illumination conditions and exposure time

Input Image	Illumination 1 Exposure Time (ms)			Illumination 2 Exposure Time (ms)			Illumination 3 Exposure Time (ms)		
	250	1000	4000	250	1000	4000	250	1000	4000
Art	8.32	8.43	8.41	8.76	8.77	8.78	9.73	9.86	11.04
Cones	9.33	9.31	9.23	10.52	9.54	9.43	11.82	12.54	12.43
Reindeer	12.43	12.42	12.24	13.52	12.53	12.63	14.52	13.53	12.43

Table 2: Percentage of bad matching pixels in the disparity results with varying left illumination/right illumination combinations.

Input Image	Aloe						Art					
Left/right illumination	1/1	1/2	1/3	2/2	2/3	3/3	1/1	1/2	1/3	2/2	2/3	3/3
% of bad matching pixels	6.41	7.78	10.58	5.2	9.68	5.51	5.93	5.482	19.69	4.30	16.195	3.02
Input Image	Mobies						Dolls					
Left/right illumination	1/1	1/2	1/3	2/2	2/3	3/3	1/1	1/2	1/3	2/2	2/3	3/3
% of bad matching pixels	5.97	6.31	9.62	6.35	8.03	6.31	18	18	28	18	20	16

Table 3: Percentage of RMS error for varying Left exposure time/right exposure time

Input image	Illumination 1 Exposure time (ms)		Illumination 2 Exposure time (ms)		Illumination 3 Exposure time (ms)	
	Left	Right	Left	Right	Left	Right
Art	250	4000	250	4000	250	4000
	0.19317		0.19807		0.19617	
Cones	200	3200	125	2000	250	2000
	0.278		0.277		0.375	
Reindeer	200	3200	125	2000	125	2000
	0.334		0.3483		0.3883	

Table 4: Percentage of bad matching pixels for varying left exposure time/right exposure time

Input image	Illumination I Exposure time (ms)		Illumination II Exposure time (ms)		Illumination III Exposure time (ms)	
	Left	Right	Left	Right	Left	Right
Art	250	4000	250	4000	250	4000
	9.125		9.332		9.423	
Cones	200	3200	125	2000	125	2000
	13.015		13.334		15.89	
Reindeer	200	3200	125	2000	125	2000
	14.985		15.235		16.845	

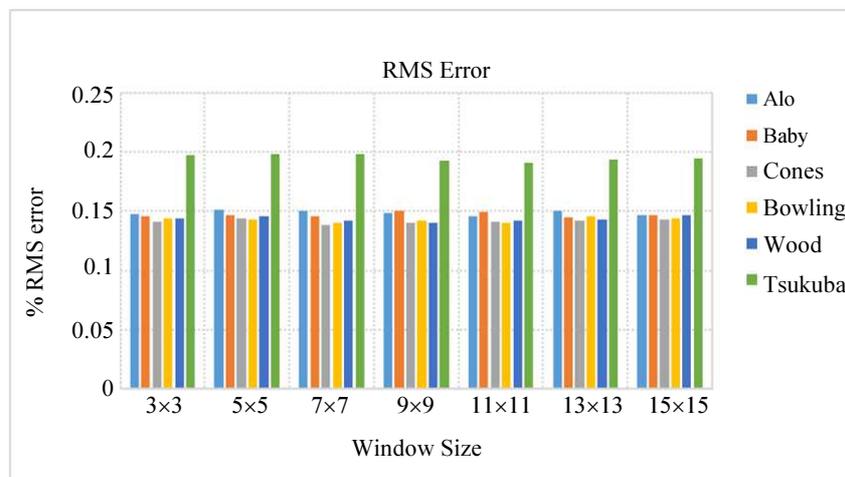


Fig. 7: RMS Error in the obtained disparity map of different input images Aloe, Baby Cones, Bowling, Wood and Tsukuba using varying window sizes

Comparison with the Existing Methods

Table 5 and 6 shows the comparison of proposed radiometric invariant LBP SGBM stereo matching approach based on amount of bad matching pixels with the existing methods of its similar kind.

Hirschmuller *et al.* (2002) proposed correlation based stereo matching. This method uses LoG filter for reduced boarder errors. Kai and Hanebeck (2001) used normalized correlation for stereo correspondence. Hirschmuller (2008) done stereo matching using semiglobal and mutual information.

Table 5: Comparison of LBP-SGBM with the existing methods based on amount of bad matching pixels

Methods	Cones	Art	Teddy
Eunah <i>et al.</i> (2017)	4.80	10.8	8.700
Andreas <i>et al.</i> (2010)	5.00	13.3	11.500
Qingxiong (2012)	8.68	5.02	10.890
Mei <i>et al.</i> (2013) ST-1	7.11	9.22	10.880
Mei <i>et al.</i> (2013) ST-2	6.62	5.27	9.950
Shenyong <i>et al.</i> (2017)	7.24	4.49	9.570
Proposed method	4.51	3.02	4.026

Table 6: Comparison of LBP- SGBM method with the existing method based on amount of bad matching pixels

Method	Left/right Illumination	Rank/BT (Ramin, Woodfill 1994)	MI (Hirschmuller 2008)	NCC (Kai and Hanebeck, 2001)	Log/BT (Hirschmuller <i>et al.</i> , 2002)	Proposed Method
Input Image	1/1	7	5	9	14	6.4100
Aloe	1/2	9	18	11	15	7.7800
	1/3	19	50	21	26	10.5800
	2/2	6	3	10	18	5.2000
	2/3	18	50	19	20	9.6800
	3/3	7	6	10	12	5.5100
	Art	1/1	15	15	17	29
1/2		15	32	21	29	5.4818
1/3		39	60	41	40	19.6900
2/2		15	14	20	30	4.3000
2/3		30	42	39	40	16.1950
3/3		13	14	19	28	3.0200
Mobies	1/1	9	10	11	15	5.9700
	1/2	22	11	15	16	5.9700
	1/3	20	38	22	21	9.6200
	3/2	11	11	12	14	6.3500
	2/3	18	30	19	20	8.0300
	3/3	10	10	11	12	5.4400
Dolls	1/1	7	11	10	18	4.0190
	1/2	9	19	11	18	5.0790
	1/3	33	38	34	28	20.2200
	2/2	7	11	10	18	5.8600
	2/3	24	30	29	20	17.9300
	3/3	7	9	9	16	4.0260

Ramin and Woodfill (1994) used non-parametric approach for stereo matching. Eunah *et al.* (2017) proposes a hierarchical stereo matching approach in two-scale space for low resolution images. Andreas *et al.* (2010) uses triangulation on a set of support points for robust stereo matching. (Qingxiong, 2012) presents an adaptive Non-Local cost aggregation approach on a tree structure for the stereo matching. Mei *et al.* (2013) proposes a tree-based cost aggregation with segment tree construction algorithm (ST-1) and aggregation with enhanced Segment tree algorithm (ST-2). Shenyong presents a nonlocal stereo matching method with an initial cost and multiple weight.

This proposed method uses local binary pattern for radiometric invariance and semiglobal block matching for handling the curved surfaces and slating regions with the real time speed. From the table, it can be seen that this approach gives accurate results compared to the other ones.

Conclusion and Future Scope

In this work we combined the run time efficiency of local stereo correspondence and the accuracy of global correspondence along with the radiometric invariant performance of local binary pattern. Other than the existing window based methods that are based on planar surface assumption, this method can handle the curved and slanting surfaces. Here smaller and higher penalty terms are added to handle small and larger disparity changes. The performance evaluation shows lesser error rate in the range of 0.14%-0.3883% and the run time requirement of 0.20 milliseconds only.

This work targets in the formulation of stereo vision system for autonomous robots, which is suitable for various real-time indoor and outdoor applications. This passive range finding system can handle curved surfaces such as a bottle and can be used for the pick and place robots. It can also be used as the sensor for the navigation of autonomous robots in unknown environments.

Also this system can be used as robotic vision system for precise assembly in automotive industry. This binocular stereo vision can be extended to multi view stereo for view synthesis and panoramic vision. Another application of our stereo matching system is in aerial imaging for finding the range of buildings from an air vehicle.

Most of the real time applications in the area of stereo vision demand algorithms with higher frame rate without compromising its accuracy. As the robots started to take over hard and risky tasks in many aspects of our modern life, incorporation of the latest outbreaks in neuroscience findings in computer vision will help in solving the open issues in intelligent robot vision. Future scope of this work targets on the incorporation of deep learning in our system for getting the cognitive capabilities of human vision to deal with the occluded areas and high dynamic range views. Such an intelligent vision system will be of great use in the indoor and outdoor applications such as rescue robots and industrial robots.

Acknowledgment

We thank Prof. S. Bhavani, Head of the Department of ECE, Karpagam University, Coimbatore, Tamilnadu, India, for sharing her valuable knowledge and suggestions that significantly improved this article on its earlier version, even though any errors if there, are of our own.

Author Contributions

Deepambika V.A.: Has done the algorithm implementation, analysis and preparation of the manuscript.

M. Abdul Rahiman: Contributed in the design of the concept, experimental analysis and article revision.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

- Andreas, G., M. Roser and R. Urtasun, 2010. Efficient large-scale stereo matching. Proceedings of the Computer Vision (ACCV'10), Springer, Berlin, Heidelberg, pp: 25-38.
DOI: 10.1007/978-3-642-19315-6_3
- Birchfield, S. and C. Tomasi, 1999. Depth discontinuities by pixel-to-pixel stereo. Int. J. Computer Vision, 35: 269-293.
DOI: 10.1109/ICCV.1998.710850
- Daniel, S., T. Taniyai and S.N. Sinha, 2017. Semi-global stereo matching with surface orientation priors. Proceedings of the International Conference on 3D Vision, Oct. 10-12, IEEE Xplore press, Qingdao, China. DOI: 10.1109/3DV.2017.00033
- Eunah, C., S. Lee and H. Hong, 2017. Hierarchical stereo matching in two-scale space for cyber-physical system. Sensors, 17: 1680.
DOI: 10.3390/s17071680
- Guanying, H., Z. Wu, J. Li and S. Li, 2018. Underwater target detection and 3D reconstruction system based on binocular vision. Sensors 18, 10: 3570.
DOI: 10.3390/s18103570
- Hamzah, R.A. and H. Ibrahim, 2016. Literature survey on stereo vision disparity map algorithms. J. Sensors, 1: 01-23. DOI: 10.1155/2016/8742920
- Heo, Y.S., K.M. Lee and S.U. Lee, 2011. Robust stereo matching using adaptive normalized cross-correlation. IEEE Trans. Pattern Analysis Machine Intelligence, 33: 807-822.
DOI: 10.1109/TPAMI.2010.136
- Hermann, S. and R. Klette, 2012. Iterative semi-global matching for robust driver assistance systems. Computer Vision-ACCV, 7726: 465-478.
DOI: 10.1007/978-3-642-37431-9_36
- Hernandez-Juarez, D., A. Chacón, A. Espinosa, D. Vázquez and J.C. Moure *et al.*, 2016. Embedded real-time stereo estimation via semi-global matching on the GPU. Procedia Comput. Sci., 80: 143-153.
DOI: 10.1016/j.procs.2016.05.305
- Hirschmuller, H. and D. Scharstein, 2009. Evaluation of stereo matching costs on images with radiometric differences. IEEE Trans. Pattern Analysis Machine Intelligence, 31: 1582-1599.
DOI: 10.1109/TPAMI.2008.221
- Hirschmuller, H., 2005. Accurate and efficient stereo processing by semi-global matching and mutual information. IEEE Computer Society Conference Computer Vision Pattern Recognition, 2: 807-814.
DOI: 10.1109/CVPR.2005.56
- Hirschmuller, H., 2008. Stereo processing by semiglobal matching and mutual information. IEEE Trans. Pattern Analysis Machine Intelligence, 30: 328-341.
DOI: 10.1109/TPAMI.2007.1166
- Hirschmüller, H., M. Buder and I. Ernst, 2012. Memory efficient semi-global matching. SPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., I-3: 371-376. DOI: 10.5194/isprsannals-I-3-371-2012
- Hirschmuller, H., P. Innocent and J. Garibaldi, 2002. Real-time correlation-based stereo vision with border reduced errors. Int. J. Computer Vision, 47: 229-246. DOI: 10.1023/A:1014554110407
- Kai, B. and U.D. Hanebeck, 2001. Template matching using fast normalized cross correlation. Proc. Optical Pattern Recognition XII, 4387: 95-103.
DOI: 10.1117/12.421129

- Khomutenko, B., G. Garcia and P. Martinet, 2016. Direct fisheye stereo correspondence using enhanced unified camera model and semi-global matching algorithm. Proceedings of the 14th International Conference on In Control, Automation, Robotics and Vision, Nov. 13-15, IEEE Xplore press, Phuket, Thailandpp, pp: 1-6.
DOI: 10.1109/ICARCV.2016.7838761
- Li, L., L. Zhao, Y. Long, G. Kuang and P. Fieguth, 2012. Extended local binary patterns for texture classification. *Image Vision Comput.*, 30: 86-99.
DOI: 10.1016/j.imavis.2012.01.001
- Marghany, M., M. Razlan, B. Tahar and M. Hashim, 2011. 3D stereo reconstruction using sum square of difference matching algorithm. *Scientific Res. Essays*, 6: 6404-6411. DOI: 10.5897/SRE11.1661
- Mei, X., X. Sun, W. Dong, H. Wang and X. Zhang, 2013. Segment-tree based cost aggregation for stereo matching. Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, Jun. 23-28, IEEE Xplore press, Portland, OR, USA, pp: 313-320.
DOI: 10.1109/CVPR.2013.47
- Mustafa, Y. and S. Kalkan, 2015. An iterative adaptive multi-modal stereo-vision method using mutual information. *J. Visual Communication Image Representation*, 26: 115-131.
DOI: 10.1016/j.jvcir.2014.11.010
- Neumann, D., T. Langner, F. Ulbrich, D. Spitta and D. Goehring, 2017. Online vehicle detection using Haar-like, LBP and HOG feature based image classifiers with stereo vision preselection. *Intelligent Vehicles Symposium (IV)*, IEEE.
DOI: 10.1109/IVS.2017.7995810
- Ogale, A.S. and Y. Aloimonos, 2005. Robust contrast invariant stereo correspondence. Proceedings of the IEEE International Conference on Robotics and Automation, April, 18-22, IEEE Xplore press, Spain. DOI: 10.1109/ROBOT.2005.1570218
- Ojala, T., K.O. Valkealahti and E. Pietikäinen, 2001. Texture discrimination with multidimensional distributions of signed gray-level differences. *Pattern Recognition*, 34: 727-739.
DOI: 10.1016/S0031-3203(00)00010-8
- Ojala, T., M. Pietikainen and T. Maenpaa, 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Analysis Machine Intelligence*, 24: 971-987. DOI: 10.1109/TPAMI.2002.1017623
- Park, W.J., J. Seo-Won, K. Seok-Jae, J. Seung-Won and K. Sung-Jea, 2017. Stereo vision-based high dynamic range imaging using differently-exposed image pair. *Sensors*, 17: 1473.
DOI: 10.3390/s17071473
- Qingxiong, Y., 2012. A non-local cost aggregation method for stereo matching. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Jun. 16-21, IEEE Xplore press, Providence, RI, USA, pp: 1402-1409.
DOI: 10.1109/CVPR.2012.6247827
- Ramin, Z. and J. Woodfill, 1994. Non-parametric local transforms for computing visual correspondence. Proceedings of the European Conference on Computer Vision, Berlin, Heidelberg, Springer.
DOI: 10.1007/BFb0028345
- San, T. and N. War, 2017. Local stereo matching under radiometric variations. Proceedings of the 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Jun. 26-28, IEEE Xplore press, Kanazawa, Japan, pp: 245-249.
DOI: 10.1109/SNPD.2017.8022728
- Sarika, S., V.A. Deepambika and M. Abdul Rahman, 2015. Census filtering based stereomatching under varying radiometric conditions. *Procedia Computer Science*, 58: 315-320.
DOI: 10.1016/j.procs.2015.08.026
- Scharstein, D. and R. Szeliski, 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Computer Vision*, 47: 7-42. DOI: 10.1109/SMBV.2001.988771
- Scharstein, D. and R. Szeliski, 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Computer Vision*, 47: 7-42. DOI: 10.1023/A:1014573219977
- Scharstein, D. and R. Szeliski, 2003. High-accuracy stereo depth maps using structured light. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Jun. 18-20, IEEE Xplore press, Madison, WI, USA.
DOI: 10.1109/CVPR.2003.1211354
- Shen, Y., 2011. Efficient normalized cross correlation calculation method for stereo vision based robot navigation. *Frontiers Computer Science China*, 5: 227-235. DOI: 10.1007/s11704-011-9190-2
- Shenyong, G.H.G., H. Zhang and Y. Zhang, 2017. A nonlocal method with modified initial cost and multiple weight for stereo matching. *J. Sensors*, 2017: 1-12. DOI: 10.1155/2017/9374870
- Stolc, S., K. Valentin and R. Huber-Mörk, 2016. STABLE: Stochastic binary local descriptor for high-performance dense stereo matching. *Electronic Imaging*, 14: 1-7.
DOI: 10.2352/ISSN.2470-1173.2016.14

- Szeliski R., 2010. Computer vision: Algorithms and Applications. Springer Science and Business Media, ISBN-10: 1848829345.
- Taimoor, S.S. and I.M. Afanasyev, 2018. Stereo vision-based optimal path planning with stochastic maps for mobile robot navigation. Proceedings of the International Conference on Intelligent Autonomous Systems Springer. Baden-Baden, Germany, pp: 40-55.
- Valentín, K., R. Huber-Mörk and S. Štolc, 2017. Binary descriptor-based dense line-scan stereo matching. J. Electronic Imaging, 26: 1-13.
DOI: 10.1117/1.JEI.26.1.013004
- Yoon, K.J. and I.S. Kweon, 2006. Adaptive support-weight approach for correspondence search. IEEE Trans. Pattern Analysis Machine Intelligence, 28: 650-656. DOI: 10.1109/TPAMI.2006.70