A Combined Approach for Stereoscopic 3D Reconstruction Model based on Improved Semi Global Matching

Rajeshkannan Sundararajan and Reeba Korah
Department of Electronics Communication and Engineering, St. Joseph's College of Engineering, India

Abstract: *The effective recovery of the 3D structure of a scene using two or more 2D images of the scene, each acquired from a different viewpoint is a challenging task of stereovision. Defining pixel correspondence in stereo pairs is a fundamental process for automated image based effective 3D reconstruction. This paper presents modified Census based approach for local cost optimization where local matching cost is combined with Sum of Squared Absolute Differences (SSAD) of the image color values and then aggregated. From the aggregated cost, effective disparity map is obtained using Semi Global Matching (SGM) which improves the quality of the matches. This proposed approach represents a fusion of state of the art algorithms to improve the matching quality with reduced number of bad pixels. Finally, a stereoscopic 3D view will be obtained by merging triangulation algorithm in the realistic manner. Because of more realistic depth perception, our proposed three dimensional stereo models finds application in medical field where it improves surgical success with shorter operation time and research in space where effective analysis can be made by using calibrated photo realistic 3D model of the space structure.*

Keywords: *Cost function, disparity map, modified census transform, SAD, SGM, SSAD, triangulation.*

Received August 25, 2013; accepted March 14, 2014; published online October 29, 2015

1. Introduction

Stereovision is the extraction of 3D information from two or more digital images obtained from camera. By comparing information about a scene from two vantage points and by examining the relative positions of objects in the two panels, 3D information can be extracted. The evolution of vision based autonomous behaviours of robotics has raised new challenges which need to be tackled. Some of the practical issues of robotics based stereo vision models are handling of uneven lighting conditions, handling of multi-view stereo systems, the use of miscalibrated image sensors and the introduction of novel biological methods to robotic vision. In traditional stereo vision, two cameras displaced horizontally from one another are used to obtain two differing views on a scene similar to human binocular vision [4] which is shown in Figure 1. By comparing those two images, the relative depth information can be obtained in the form of disparities which are inversely proportional to the differences in distance to the objects. The term binocular vision can be used when two cameras are required.

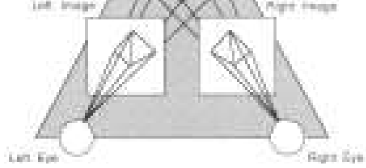
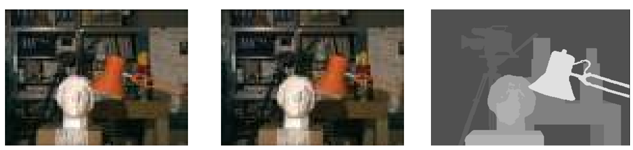


Figure 1. Object perception.

The recovery of an accurate disparity map is still challenging task due to the following reasons:

1. Pixels of half occluded regions [24] do not have correspondence which leads to incorrect matches.
2. Due to sensor noise, images get disturbed. This becomes problematic in poorly textured regions.
3. The constant brightness or color constraint is only satisfied under ideal conditions that can only roughly be met in practice.

The primary focus of the research in stereo vision is correspondence and reconstruction. The correspondence problem [5] consists of determining the locations in each camera image that are the projection of the same physical point in space. i.e., finding pairs of matched points such that each point in the pair is the projection of the same 3D point. No general solution to the correspondence problem exists due to ambiguous matches. From the corresponding points, the disparity map of the scene can be generated. The depth map is simply the reciprocal of the disparity map [6]. The effective disparity map can be obtained by matching corresponding pixels in the left and right image as shown in Figure 2.



a) Reference. b) Target. c) Disparity map.

Figure 2. Effective disparity map.

From the disparity map, 3D depth view of the scene can be extracted (i.e., recover the 3D structure) if the stereo geometry [12] is known. The reconstruction problem consists of determining three dimensional structure of the scene from a disparity map based on camera geometry. The depth of a point in space *P* imaged by two cameras with optical centers O_L and O_R are defined by intersecting the rays from the optical centers through their respective images of *P*, *p* and *p'* which is depicted in Figure 3.

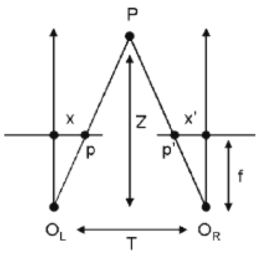


Figure 3. Disparity.

Given the distance between OL and OR called the baseline *T* and the focal length *f* of the cameras, depth at a given point may be computed by similar triangles as:

$$Z = f\,\frac{T}{d} \tag{1}$$

Where *d* is the disparity of that point, $d=x-x'$. This process is called triangulation [10] but in practice, it is difficult to build stereo systems with non-verged geometry. A role of stereo imaging technology [14] is inevitable in medical field since it improves accuracy and reliability of operation of a surgical robot system. Moreover, detailed diagnosis of small biological organs may also be carried out. The Field Programmable Gate Array (FPGA) implementation of stereo matching problem [19] has been a promising alternative towards real time speed. The architecture and the design methodologies are unique for FPGA implementation of stereo matching problems. The great computational capabilities can be achieved by using parallel pipelined processing [15], providing with proper scalability opposed to the serial behaviour of most software based techniques.

The rest of the paper is organized as follows: Section 2 gives an idea about related survey. Section 3 explains about proposed model and section 4 discusses and evaluates the simulation results. Section 5 draws the conclusion.

2. Related Works

In early days, most of the analysis in stereovision was concerned with the processing of aerial photographs for depth estimation. Baha *et al.* [1] presented a new neural network based approach for disparity computation which is the combination of area and feature based matching method. Initial disparity computation is done using Back-Propagation followed by accurate disparity refinement by using image segmentation method. For matching process, feature points are detected using Harris detector and valid match is obtained by ZNCC. The experimental validation shows that the efficiency and the quality of the disparity map are greatly improved and the processing time is also reduced. Bleyer and Gelautz [4] discussed about the role of color segmentation to allow the handling of large un-textured regions and precise localization of depth boundaries.

Stentoumis *et al*. [24] attempted for dense matching based on local optimization. Here, the matching cost is combined with the Absolute Difference (AD) of image color values and a new cost volume had been computed by aggregating over cross-window support regions with a linearly defined threshold on cross-window expansion. Then, aggregated costs had been refined using a scan-line optimization technique. Hu and Huang [12] proposed a new stereo matching algorithm based on square and gradient for binocular vision. Here, an image line is divided into a series of ranges by comparing the gradients of the points in left and right image lines. Sobel operator was used to calculate the gradients of the images. The best matching in each range was found based on the summary of squared differences. Hirschmiller and Scharstein [10] had evaluated the insensitivity of different matching costs with respect to radiometric variations of the input images. Both pixel-based and window-based variants had been taken in to account and their performances in the presence of global intensity changes and local intensity changes were measured.

Zabih and Woodfill [29] dealt with the performance of non parametric local transforms for computing visual correspondence. Factionalism produced outliers with consistent distributions. The motivation of this approach was to obtain better results near the edges of the object. Yang and Wang [27] attempted to address the problem of redundant computation in depth computation. He presented a belief propagation based global algorithm that generates high quality results while maintaining real-time performance. In order to improve the number of reliable correspondences with edge information, Klaus *et al.* [17] proposed an algorithm which was based on color segmentation and self-adapting dissimilarity measure (combination of Sum of Absolute intensity Differences and a gradient based measure). The conjunction of color segmentation, a self-adapting matching score, a robust plane fitting technique as well as BP-optimization yields superior matching quality.

Humenberger *et al.* [13] suggested a new segmentation-based approach for disparity optimization in stereo vision. The local cost calculation was done with a census-based correlation method and it was compared with standard sum of ADs. The confidence of a match was measured which was used to estimate non-confident or non-textured pixels.

Birchfield and Tomasi [3] proposed a new form of dynamic programming to match epipolar scan lines independently and propagates information between the scan lines to refine the disparity map and the depth discontinuities. Fife and Archibald [7] proposed a new sparse census transform which reduces resource utilization with increased flexibility and the same was implemented on FPGA resulting throughput of 500Mbps with good correlation accuracy. Heo *et al.* [9] attempted to develop an iterative framework which constructs joint probability density function of log-chromaticity color space images to give a linear estimate of corresponding pixels. They devised a new matching cost by combining Mutual Information (MI), SIFT descriptor with the use of stereo color histogram equalization to produce accurate depth maps and color consistent images. Although, many area based stereo matching algorithm have been proposed, the selection of proper shape and size of the matching window is still a challenging issue in real time vision system.

Gupta and Cho [8] proposed a new approach that utilizes two correlation windows to improve the performance of the algorithm with real-time suitability. This method shows good result with sharp disparity map at boundaries and in low textured images. Ttofis *et al*. [26] designed a new disparity map computation architecture that integrates Sobel edge detection unit which reduces search space and dense map is produced by optimizing a global energy function. Finally, this prototype model is implemented on the Xilinx ML505 FPGA evaluation Platform, achieving 50 fps for 1,280×1,024 image. The increased bandwidth requirement due to Census transform based Semi Global Matching (SGM) is solved by a novel sub pixel interpolation methodology which is proposed by Pantilie and Nedevschi [21] and it has been implemented on CUDA. Perri *et al.* [23] proposed a new hardware oriented approach that utilize adaptive support weights during Census transformation and weighted sum of SAD is used as the dissimilarity metric. The above approach is implemented on Virtex 6 FPGA with 68 fps for 640×480 stereo images using 80,000 slices.

Thomas *et al.* [25] described the design and implementation of 3D stereo vision on Virtex-5 FPGA. Here, the researchers discussed about a lot of ongoing effort to support real time processing of 3D stereo vision system in which they proved a rate of 30 fps is desirable for the human eye. The implementation of different technologies with maximum resolution and processing rate is presented in Table 1.

Table 1. Implementation of 3D stereo vision in different technologies.

Resolution	Disparity	fps	Technology	Year
160×120	32	30	DSP	1997
320×240	20	150	Spartan-3	2007
320×240	16	75	Virtex-II	2010
640×480	128	30 4	Stratix S80	2006
640×480	64	230	Virtex-4	2010

From the above comparison, the maximum performance was obtained with the Virtex-4 FPGA implementation.

3. Proposed Methodology

Over the years, a wide variety of measures for the implementation of stereo matching approaches has been discussed in literature. In general, area-based matching is used to calculate the costs for each matching candidate and then optimize them for correct matches. Once the local costs are computed, a minimum search (winner takes all) is used to find the best match. Another strategy is global optimization [28] to enhance the probability of correct matching. In our proposed work, the whole scan line or the whole image will be used to calculate the cost. Using this technique, better results can be achieved on texture less area. The block diagram of our proposed model is shown in Figure 4. For this vision analysis, we assume that images have been rectified so that epipolar [12] lines corresponding to scan lines simplify the algorithm and improve its efficiency.

Using stereo pair, the cost function can be computed by combining modified census transform [23] and the Sum of Squared Absolute Difference (SSAD). From the minimum path cost which is computed by SGM technique, effective disparity will be derived.

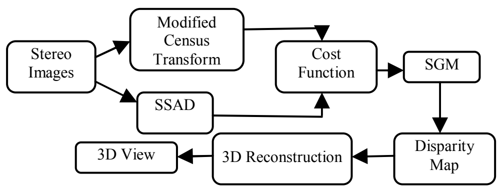


Figure 4. Block diagram of proposed method.

3.1. Modified Census Transform

The census transform [2] is in the class of area based non parametric transform which is well suited for implementation in FPGA as the majority of the operations are performed using bit manipulation and simple arithmetic. The census transform also, exhibits relative insensitivity to ADs in imaging devices and operates well on images with high dynamic range. An example of a window with radius “1” is the following centered on a pixel with intensity 130.

$$\begin{pmatrix} 127 & 128 & 131 \\ 126 & 130/129 & 129 \\ 126 & 131 & 133 \end{pmatrix}$$

But census transform is not suitable for low texture images due to loss of information associated with each pixel. Hence, we use a modified one which compares the mean of the 3×3block pixel intensities [11] rather than the center value. For this modified census, instead of taking centre value as 130, we take average value of intensity 129 as centre. Pixels with lower intensities are given a value of 1 and pixels with higher intensities are given a value of 0. Pixels with intensity values equal to the intensity value of the center pixel can be treated as either a 1 or 0, but must remain consistent. The transformed matrix would be:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & X & 1/0 \\ 1 & 0 & 0 \end{pmatrix}$$

Then, this matrix is rearranged into a bit string to capture the entire window resulting in the eight-symbol pattern "11010100" of primary image. Using this pattern, we find the closest matching pixel in the secondary image by comparing bit strings using Hamming distance [18]. Pixel windows with similar intensities relative to the center pixel will give smaller Hamming distances. The same pixel position in the secondary image yields a disparity of zero and an infinite range estimate. This process continues until the predetermined maximum disparity is reached.

3.2. Sum of Squared AD

The AD [29] is a simple measure which is widely used in matching. Sum of ADs is an algorithm which takes the AD between each pixel in the original block and the corresponding pixel in the block being used for comparison of measuring similarity. These differences are summed to create a simple metric of block similarity. It is calculated by subtracting pixels within a square neighbourhood between the reference image I_1 and the target image I_2 followed by the aggregation of ADs within the square window. If the left and right images match exactly, then the resultant image will be zero.

$$\sum_{(i,j)\in W} \left| I_1(i,j)^2 - I_2(x+i,\ y+j)^2 \right| \tag{2}$$

For example, the template image is 3 by 3 pixels in size while the search image is 3 by 5 pixels in size. Each pixel can be represented by a single integer from 0 to 9.

$$\underset{Template}{\begin{pmatrix} 2 & 5 & 5 \\ 4 & 0 & 7 \\ 7 & 5 & 9 \end{pmatrix}} \qquad \underset{Search\ Image}{\begin{pmatrix} 2 & 7 & 5 & 8 & 6 \\ 1 & 7 & 4 & 2 & 7 \\ 8 & 4 & 6 & 8 & 5 \end{pmatrix}}$$

There are exactly three unique locations available within the search image where the template may fit the left side of the image, the centre of the image and the right side of the image and AD values for each location is given below:

$$\underset{Left}{\begin{pmatrix} 0 & 2 & 0 \\ 3 & 7 & 3 \\ 1 & 1 & 3 \end{pmatrix}} \qquad \underset{Centre}{\begin{pmatrix} 5 & 0 & 3 \\ 3 & 4 & 5 \\ 3 & 1 & 1 \end{pmatrix}} \qquad \underset{Right}{\begin{pmatrix} 3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & 3 & 4 \end{pmatrix}}$$

The ADs are added together to give a SAD value of 20, 25 and 17 corresponding to each of the three image patches. From these SAD values, it is clearly understood that the right side of the search image is the most similar to the template image because it has the least difference as compared to the other locations. However, this is not suitable method for low light intensity images where noisy disparity maps occur. As an improvement over the Sum of ADs, we have used the SSAD to rectify the above weakness of SAD. This SSAD algorithm is used to compute the corresponding point between images for finding disparity maps. We conclude that for each pixel in IMG0, finding corresponding points in IMG1 and then the AD of corresponding pair is calculated. Finally, ADs with same disparity ae accumulated to get SAD. From SAD, all the values are accumulated for obtaining SSAD.

3.3. Semi Global Matching

The use of row-wise 1Dconstraints in dynamic programming results in depth maps with "streaking effect" [27] as shown in Figure 5.

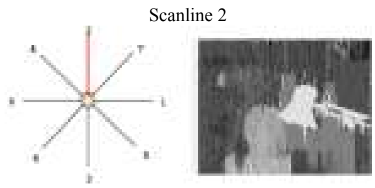


Figure 5. Streaking effect.

To overcome streaking effects, we have proposed a new method which combines SGM, modified census transform and SSAD. SGM [17] is a scalable method which approximates global optimization by combining multiple 1D optimization criteria to improve dense matching results. In SGM, a cost array [20] is formed over the costs of all disparities from the minimum disparity to the maximum disparity for all the pixels and the minimum path cost is calculated from the pixel energy Equation 3 which aims to determine the disparity image D such that the global energy $E(D)$ [21] is a minimum.

$$\begin{aligned} E(D) = \sum_p \Big(C(p, Dp) &+ \sum_{q\in Np} P1T\left[\left| Dp - Dq \right| = 1 \right] \\ &+ \sum_{q\in Np} P2T\left[\left| Dp - Dq \right| > 1 \right] \Big) \end{aligned} \tag{3}$$

The first term of equation calculates the sum of a pixel wise matching cost $C(p, Dp)$ for all pixels p at their disparities Dp. The function T [ ] is defined to return 1 if its argument is true and 0 otherwise. The second term of the energy function penalizes small disparity differences of neighbouring pixels Np of p with the cost $P1$. Similarly, the third term penalizes larger disparity steps with a higher penalty $P2$.The value of $P2$ does not depend on the size of the disparity, which preserves discontinuities. The costs along the paths from all directions r are summed, we get:

$$s(p,\ d) = \sum_r Lr(p,\ d) \tag{4}$$

For each pixel p, the disparity d is chosen that corresponding to the minimum cost and is given below:

$$Dp = arg\ min\ S(p, d) \tag{5}$$

SGM has very good trade off between runtime and accuracy. It is also, robust against radiometric differences and is well suited for remote sensing and mobile robotics. However, in practice, the pixel wise matching cost purely depends on texture of the image. Therefore, global matching finds difficulties on untextured area. Those areas are interpolated smoothly with the support of neighbouring better textured areas to produce accurate results at depth borders [13] and uniform regions.

3.4. 3D Reconstruction Model

Stereo reconstruction is based on the concepts of epipolar geometry [12]. An overview of a stereo 3D reconstruction system is shown in Figure 6. The effective stereo reconstruction algorithm [3] needs to solve two basic problems: One is correspondence and the other is reconstruction. In order to locate a common point of reference in the two images, a small window (called correlation window) from both images is being evaluated by comparing the window from the left image with the window from the right image. Once, its correspondence is located in both images, an effective disparity map is derived using its pixel coordinates. Then, reconstructed 3D structure of the scene is formed by projecting images onto a common image plane using triangulation.

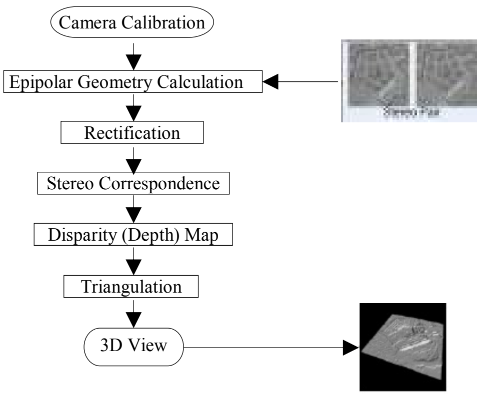


Figure 6. 3D reconstruction system.

The following steps are involved in reconstruction process:

- *Step* 1: Disparity Map obtained from SGM is taken as input for the reconstruction.
- *Step* 2: Disparity Map is then converted to gray scale and concatenated along three dimensions.
- *Step* 3: A composite image is obtained from the concatenated image and right image. If the two images are of different sizes, the smaller dimension images are padded with zeros so that both images are of the same size before creating the composite image.
- *Step* 4: Then, the above composite view is fused with the left image to obtain the final reconstructed image.

3.5. Complexity Analysis

The computational complexity of MCT and SSAD is quadratically related to the window size used to aggregate the matching costs and its implementation is linear with the number of image pixels. The truncated squared AD in range and the gradient between corresponding pixels are used to calculate the matching costs. The FPGA implementation of MCT and SSAD are ideal because of their high accuracy though the complexity of algorithm is three times with high demand on computation time. Due to squared AD computation, the computation time of SSAD is slightly greater than SAD and SSD methods. But its accuracy of dense disparity map is very much better and there must be a trade off between computation time for real time implementation and accuracy. The time analysis in sec of SAD, SSD, NCC and SSAD [22] with window size 9×9 is tabulated in Table 2 and its graphical representation is shown in Figure 7.

Table 2. Comparison of computation time.

Matching Method	Computation Time (in sec)
SSAD	0.19320
SAD	0.14264
SSD	0.14596
Histogram based Matching	0.25
NCC	0.27938

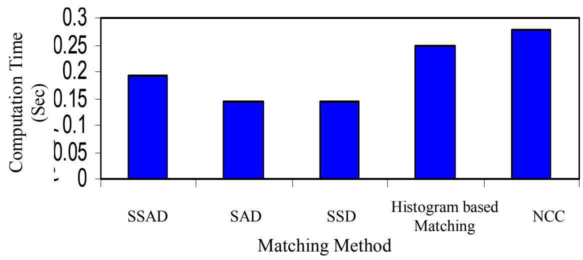


Figure 7. Matching method vs computation time for tsukuba stereo pair (window size 9×9).

The pixel wise matching of all pixels is another reason of complexity for SGM. The complexity can be reduced by considering edges of stereo image which is obtained from the use of edge detector unit matching and optimal disparity range selection. Due to low computation complexity and storage requirement, real time implementation of local methods is quite adoptable. By selecting larger supporting window, the incorrect disparity estimation of local methods at texture less, repeating pattern regions can be minimized and the overhead has been dealt with about 10% of their runtime at full resolution.

3.6. Applications

A lot of research work has been carried out in the area of stereoscopic medical imaging and surveillance system [16]. Many practical devices have been realized and applied in the medical and surveillance field. Some of the recent applications are discussed below:

1. Accuracy of the surgery can be improved by providing clear three dimensional endoscopic

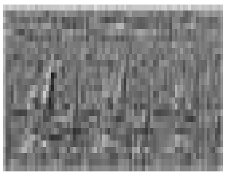
images of interior of organ or a cavity of the body to the surgeon.
2. Investigation of 3D structure of complex biological specimens will be made by providing realistic 3D morphology of the object being examined. This can be done by using stereo microscope which produces a three dimensional visualization of the object.
3. By using vision based motion tracking technology, 3D position parameters of fingers are calculated to detect breast cancer.
4. Finally, unauthorized entry can be detected in hazardous and military areas by using contour based stereo vision system.

4. Results and Discussions

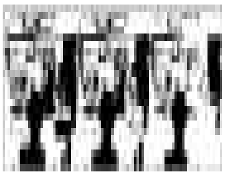
The work presented in this paper is a new hybrid census based vision approach and is tested with different Middlebury datasets. Figure 8 shows different stages of our proposed method using one of the Middlebury input stereo data set.

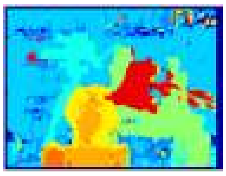
a) Original tsukuba stereo pair (left image and right image).



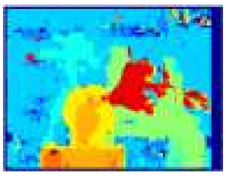
b) Census Transformed image.



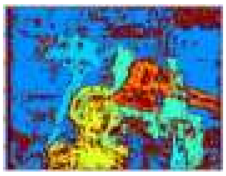
c) Modified Census output.



d) SAD correlated output.



e) SSAD correlated output.



f) Final proposed SGM output.

Figure 8. Results of proposed SGM output on stereo pair.

In Census, the centre pixel intensity value is replaced by the bit string such that if current pixel intensity is less than that of the centre, assign bit to 0 else 1. In the modified Census, instead of the centre pixel value, the mean value of the entire pixel block is assigned to centre pixel and compared with neighbourhood pixels. Some invisible parts of stereo pair in Census as shown in Figure 8-b are present in modified Census transform as shown in Figure 8-c. Similarity metrics will be done by using suitable correlation method SSAD which is applied on the transformed vector of left and right stereo pairs. Compared with existing SAD method, effective recovery is achieved by SSAD which is shown in Figures 8-d and e. Objects closer to the camera have larger disparity and appear brighter in the disparity map. The matching results which are obtained from the above proposed approach are further improved by applying SGM which finds optimum disparity of the image is shown in Figure 8-f. From the above result, we conclude that improved SGM has shown effective matching with minimum number of bad matches. Finally, reconstructed view is partially obtained from the disparity of proposed model by the above suitable reconstruction approach as shown in Figure 9. Our proposed matching algorithm gives better result with considerable reduction of bad matches which is listed in Table 3.

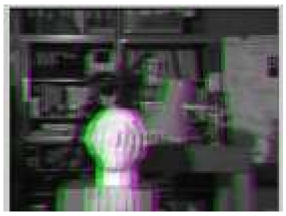


Figure 9. Reconstructed output.

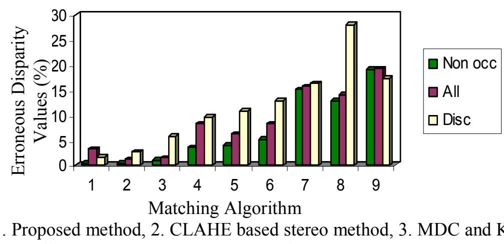
Table 3. Detail of bad matches.

Matching Algorithm	No. of Bad Pixels	% of Bad Matches
Proposed Method	288	6.12
Histogram based Matching	844	7.57
Rank+SSD	1200	10.34
Census+SAD	2952	12.35

Based on erroneous disparity values of pixel, the improved performance of our proposed method with other existing algorithm is tabulated in Table 4 and its graphical representation is shown in Figure 10.

Table 4. Erroneous disparity values.

Algorithm	Non occ	All	Disc
Proposed Method	0.4168	3.3178	1.6699
CLAHE based Stereo Method	0.4135	1.266	2.755
MDC and KP	1.08	1.59	5.8
DoubleBP	3.53	8.3	9.63
ADCensus	4.1	6.22	10.9
CoopRegion	5.16	8.31	13.0
HistAggr2	15.2	15.7	16.4
RTCensus	12.9	14.1	28.1
AdaptingBP	19.1	19.3	17.4



1. Proposed method, 2. CLAHE based stereo method, 3. MDC and KP
4. DoubleBP 5. ADCensus 6. CoopRegion, 7. HistAggr2
8. RTCensus, 9. AdaptingBP

Figure 10. Matching algorithm Vs

Where Non occ represents the subset of the non occluded pixels, Disc represents subset of the pixels near the occluded areas and All represents subset of the pixels being either non occluded or half-occluded. The above listed results show that our proposed method performs at par with other state-of-the-art algorithms by the reduced erroneous disparity values.

5. Conclusions

The stereo correspondence problem remains an active area for research in computer vision. One of the major problems in stereovision is occlusion which gives more number of mismatched pixels with inferior quality of disparity. This above quality of results is improved considerably by our combined MCT and SSAD method with less number of mismatches and it has been tested on different stereo image pairs from Middlebury benchmark. Further, improvement is achieved by using SGM technique and finally, partial implementation of reconstruction is done to get three dimensional views. More and more recent applications demand not only accuracy but real time operation as well. The real time implementation of highly scalable systolic based architecture will be proposed to achieve high processing speed while maintaining the accuracy of 3D image with less number of resources. By using double buffering technique combined with pipelined processing to integrate sub pixel for high resolution images, the required reconfigurable hardware can be designed in future and the implementation will be done on a Xilinx Virtex-5 to generate 150 fps of dense disparity map.

We hope to enhance our proposed method in a forthcoming work for improving the effectiveness of the vision by including edge detection process over the input images which will further reduce computational space and complexity.

References

[1] Baha N. and Larabi S., "Neural Disparity Map Estimation from Stereo Image," *the International Arab Journal of Information Technology*," vol. 9, no. 3, pp. 217-224, 2012.

[2] Ben-Ari R. and Sochen N., "Variational Stereo Vision with Sharp Discontinuities and Occlusion Handling," available at: https://pdfs.semanticscholar.org/11eb/df7cd46cf3b5b2ff500ae7762248931b1ce3.pdf, last visited 2003

[3] Birchfield S. and Tomasi C., "Depth Discontinuities by Pixel-to-Pixel Stereo," *in Proceedings of the 6th International Conference on Computer Vision*, Bombay, India, pp. 1073-1080 1998.

[4] Bleyer M. and Gelautz M., "A Layered Stereo Matching Algorithm using Image Segmentation and Global Visibility Constraints," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 59, no. 3, pp. 128-150, 2005.

[5] Brown M., Burschka D., and Hagar G., "Advances in Computational Stereo," *IEEE Transaction on Pattern and Machine Intelligence*, vol. 25, no. 8, pp. 993-1008, 2003.

[6] Delaunoy A., Prados E., and Belhumeur P., "Towards Full 3D Helmholtz Stereovision Algorithms," *in Proceedings of* the *10th Asian Conference on Computer Vision*, New Zealand, pp 39-52, 2010.

[7] Fife W. and Archibald J., "Improved Census Transforms for Resource-Optimized Stereo Vision," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 1, pp. 60-73, 2013.

[8] Gupta R. and Cho S., "Window-based Approach for Fast Stereo Correspondence," *IET Computer Vision*, vol. 7, no. 2, pp. 123-134, 2013.

[9] Heo Y., Lee K., and Lee S., "Joint Depth Map and Color Consistency Estimation for Stereo Images with Different Illuminations and Cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1094-1106, 2013.

[10] Hirschmuller H. and Scharstein D., "Evaluation of Cost Functions for Stereo Matching," *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, pp.1-8, 2007.

[11] Hirschmuller H., "Stereo Vision in Structured Environments by Consistent Semi Global Matching," *in Proceedings of IEEE* Conference *on Computer Vision and Pattern Recognition*, pp. 2386-2393, 2006.

[12] Hu T. and Huang M., "A New Stereo Matching Algorithm for Binocular Vision," *International Journal of Hybrid Information Technology*, vol. 3, no. 1, pp. 1-8, 2010.

[13] Humenberger M., Engelke T., and Kubinger W., "A Census based Stereo Vision Algorithm using Modified Semi Global Matching and Plane Fitting to Improve Matching Quality," *in Proceedings of IEEE Conference on Computer Society*, San Francisco, pp. 77-84 2010.

[14] Humenberger M., Zinner C., Weber M., Kubinger W., and Vincze M., "A Fast Stereo Matching Algorithm Suitable for Embedded Real-Time Systems," *Computer Vision and Image Understanding*, vol. 114, no. 11, pp. 1180-1202, 2010.

[15] Jia Y., Zhang X., Li M., and An L., "A Miniature Stereo Vision Machine (MSVM-III) for Dense Disparity Mapping," *in Proceedings of the 17th International Conference on Pattern Recognition*, pp. 728-731 2004.

[16] Kanadi T., Kano H., and Kimura S., "Development of a Video Rate Stereo Machine," *in Proceedings of International Robotics and Systems Conference*, pp. 95-100, 1995.

[17] Klaus A., Sormann M., and Karner K., "Segment based Stereo Matching using Belief Propagation and a Self Adapting Dissimilarity Measure," *in Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, pp. 15-18 2006.

[18] Kolmogorow V. and Zabih R., "Computing Visual Correspondence with Occlusions via Graph Cuts," available at: http://www.cs.cornell.edu/rdz/Papers/KZ-occ-pami-submit.pdf, last visited 2001.

[19] Longfield S. and Chang M., "A Parameterised Stereo Vision Core for FPGAs," *in Proceedings of the 17th IEEE Symposium on Field Programmable Custom Computing Machine*, Napa, pp. 263-266 2009.

[20] Min D., Lu J., and Do M., "Joint Histogram based Cost Aggregation for Stereo Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, pp. 2539-2545, 2013.

[21] Pantilie C. and Nedevschi S., "SORT-SGM: Subpixel Optimized Real-Time Semiglobal Matching for Intelligent Vehicles," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 3, pp. 1032 - 1042, 2012

[22] Patil S., Nadar J., Gada J., Motghare S., and Nair S., "Comparison of Various Stereo Vision Cost Aggregation Methods," *International Journal of Engineering and Innovative Technology*, vol. 2, no. 8, pp. 222-226, 2013.

[23] Perri S., Corsonello P., and Cocorullo G., "Adaptive Census Transform: A Novel Hardware-Oriented Stereovision Algorithm," *Computer Vision and Image Understanding*, vol. 117, no. 1, pp. 29-41, 2013.

[24] Stentoumis C., Grammatikopoulos L., Kalisperakis I., and Karras G., "Implementing an Adaptive Approach for Dense Stereo Matching," available at: http://portal.survey.ntua.gr/main/labs/photo/staff/gkarras/Karras_ISPRS_2012a.pdf, last visited 2012.

[25] Thomas S., Papadimitriou K., and Dollas A., "Architecture and Implementation of Real Time 3D Stereo Vision on a Xilinx FPGA," *in proceedings of the 21st International Conference on Very Large Scale Integration*, Istanbul, pp. 186-191, 2013.

[26] Ttofis C., Hadjitheophanous S., Georghiades A., and Theocharides T., "Edge-Directed Hardware Architecture for Real-Time Disparity Map Computation," *IEEE Transactions on Computers*, vol. 62, no. 4, pp. 690-704, 2013.

[27] Yang Q., Wang L., Ruigang Yang R., Wang S., Liao M., and Nisťer D "Real Time Global Stereo Matching using Hierarchial Belief Propagation," available at: http://vis.uky.edu/~liaomiao/publication_files/realtimebp_bmvc2006.pdf, last visited 2006.

[28] Yang Q., Engels C., and Akbarzadeh A., "Near Real-time Stereo for Weakly-Textured Scenes," available at: http://www.cs.cityu.edu.hk/~qiyang/publications/bmvc-08-qingxiong-yang.pdf, last visited 2008.

[29] Zabih R. and Woodfill J., "non parametric local transforms for computing visual correspondence," *in proceedings of the 3rd european conference on computer vision*, sweden, pp. 151-158, 1994.

Rajeshkannan Sundararajan received his BE degree in Electronics and Communication Engineering from Government College of Engineering, Thirunelveli, Manonmaniyam Sundaranar Unversity, India in 1998 and ME in Applied Electronics from Sathyabama University, chennai, India in 2004. He is pursuing PhD in the field of Computer Vision Systems, Anna University, Chennai, India. Currently, he is working as associate professor in the Department of Electronics and Communication, St. Joseph's College of Engineering, Chennai, India. He has presented a paper in an ieee conference held at chennai and published a paper in international journal. His research interest includes image processing and vlsi design.

Reeba Korah received her BE degree in Electronics Engineering from Marathwada University, India in 1992 and ME degree in Mechatronics from Anna University, India in 2000. She secured her PhD from Anna University in 2009. Currently, she is working as professor and head of Department of Electronics and Communication of St. Joseph's College of Engineering, Chennai, India. She has published six research papers in international journals and many in national and international conferences. Her research interests are in VLSI design and signal processing.