Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 13, Number 9 (2017), pp. 6355-6366

http://www.ripublication.com

A Survey and Analysis of Mathematical Algorithms on Stereo Images and Its Techniques

[1]Mrs. V. Kavitha and [2]Mr. G. Balakrishnan

[1]Research Scholar, Bharathiyar University, Coimbatore, India.
[2]ME.., PH.D, Director, Indira Ganesan College of Engineering, Trichy, India.

Abstract

Stereo Images are using many conventional algorithms, which are very expensive and inconvenient in estimating disparity in the occlusion, discontinuities and texture less regions in the images. In order to overcome these issues, in this paper, we propose some methods and techniques. .In this paper going to see about different types of methods, techniques, algorithms are used into stereo images. Advanced Elevation Models delivered by computerized photogrammetry workstations are regularly utilized as a part in complex Geographic Information Systems displaying. Since the exactness of GIS databases must be inside a predefined extend for fitting investigation of the data and resulting basic leadership, a precise is required. Customary picture coordinating systems might be delegated either region based or include based techniques. These picture coordinating strategies couldn't conquer the divergence discontinuities issue and just supply a Digital Surface Model. The range based picture coordinating technique is utilized to supply thick differences. Picture edge recognition and surface examination procedures are utilized to discover houses and tree territories. Both these parts are robustified with a specific end goal to maintain a strategic distance from out layers.

Keywords: Disparity Map, Multiple Windows, Adaptive Windows, Mean Shift Algorithm

1. INTRODUCTION

Computer vision is currently an important field of research. It includes methods such as image acquisition, processing, analysis, and understanding [1]. Computer vision techniques attempt to model a complex visual environment using various mathematical methods. The main purpose of computer vision is to define the world that we see based on one or more images and to restructure its properties, such as its illumination, shape, and colour distributions. Stereo vision is an area within the field of computer vision that addresses an important research problem: which is the reconstruction of the three-dimensional coordinates of points for depth estimation. A stereo vision system is consists of a stereo camera. Namely, two cameras placed horizontally. The two images captured simultaneously by these cameras are then processed for the recovery of visual depth information [2]. The challenge is to determine the best method of approximating the differences between the views shown in the two images to map (i.e., plot) the correspondence (i.e., disparity) of the environment. Initially, a disparity map represents corresponding pixels that are horizontally shifted between the left image and right image. New methods and techniques for solving this problem are developed every year and exhibit a trend toward improvement in accuracy and time consumption.

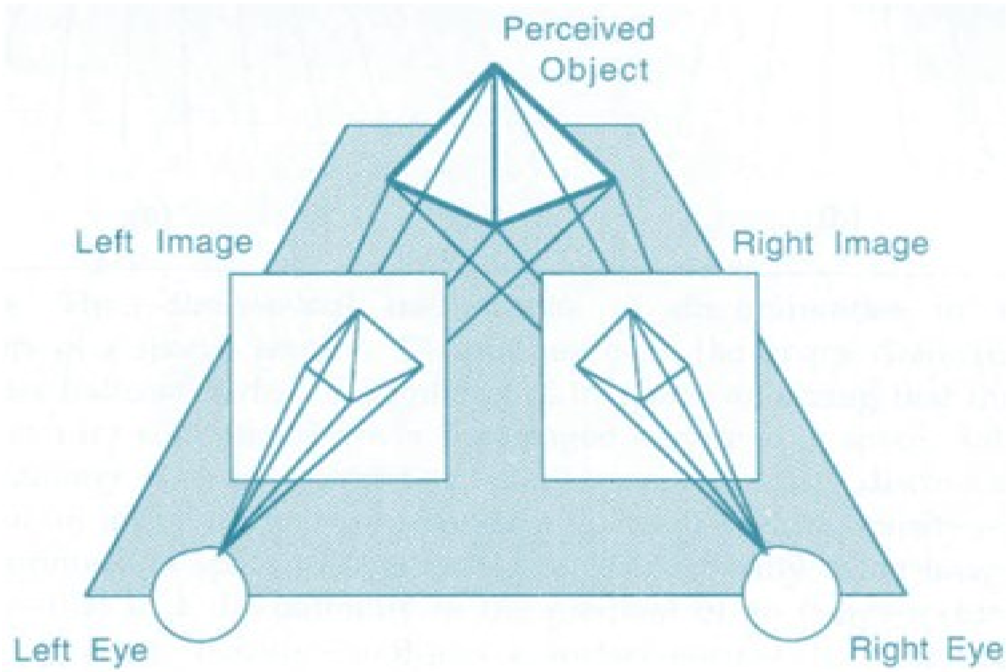

Figure 1: Goal of Stereo Vision

Another device that is used to acquire depth information is a time-of-flight (ToF) or structured light sensor. Such a device is a type of active sensor, unlike a classic stereo vision camera. Devices of this type such as the Microsoft Kinect are cheap and have led to increased interest in computer vision applications. However, these active sensors suffer from certain characteristic problems [3]. First, they are subject to systematic errors such as noise and ambiguity, which are related to the particular sensor that is used. Second, they are subject to non-systematic errors such as scattering and motion blur. According to the comparative analyses performed by Foix et al. [4], Kim et al. [5], and Zhang et al. [6], To F devices perform satisfactorily only up to a maximum distance of approximately 5–7 meters and are too sensitive to be used in outdoor environments, especially in very bright areas. Because of these limitations of ToF sensors, stereo vision sensors (i.e., passive sensors) are more reliable and robust; they are able to produce high-resolution disparity maps and are suitable for both indoor and outdoor environments [7]. The recovery of the 3D structure of a scene using two or more images of the 3D scene, each acquired from a different viewpoint in space and The term binocular vision is used when two cameras are employed (Figure.1).

In stereo vision uniqueness outline figures required to builds number of pixels per picture. This wonder makes the coordinating issue be computationally mind boggling [8]. The enhancements to and decrease in computational many-sided quality that have been accomplished with late advances in equipment innovation have been gainful for the headway of research in the stereo vision field. In this manner, the principle inspiration for equipment based usage is to accomplish constant handling [9]. Continuously stereo vision applications are characterizing self-governing driving, 3D gaming, and self-sufficient mechanical route, quick however exact profundity estimations are required [10]. Extra preparing equipment is along these lines important to enhance the handling speed.

A review on stereo vision difference outline would be profitable to the individuals who are keen on this exploration region. The catchphrases are for the most part utilized into stereo vision/stereo vision calculation and their segments that were looked were the title, dynamic, and watchwords/record terms of the papers in the databases. These papers may speak to commitments to principal calculation advancement, investigation, or utilization of stereo vision calculations. Their pattern lines are expanded and demonstrated that the field of stereo vision stays dynamic in innovative work and it's turned into an intriguing and testing zone of research. This paper gives a short prologue to the best in class advancements fulfilled with regards to such calculations. This all works are surveyed to most recent distributed stereo vision calculations and classifies them into various phases of preparing, which depend on the scientific classification proposed by Scharstein and Szeliski [11]. This paper additionally talks about two sorts of usage stages for these calculations (i.e., programming based and equipment based). In programming based stages, the strategies are actualized just on a standard CPU, with no other extra handling equipment. Rather than equipment based stages, the calculations are executed on a CPU, with a GPU or FPGA as an independent framework.

2. LITERATURE REVIEW

Year	Author	Proposed Methodology
2002	Scharstein and Szeliski [11]	Proposed taxonomy for vision algorithms and provided a quality metric to compare and evaluate multiple blocks of algorithms and also provided a test bed for measurable evaluation of stereo depth map algorithms. The test bed or benchmarking dataset consists of four images (Tsukuba, Venus, Teddy, and Cones) which are available at http://www.middlebury.edu/stereo.
2003	Brown et al. [3]	Reviewed in stereo vision disparity map algorithms regarding correspondence methods and occlusion handling methods for real time implementations.
2008	Tombari et al. [20]	Survey and compared to different methods of cost aggregation for stereo correspondence through accuracy and computational requirements.
2008	Lazaros et al. [12]	Reviewed developments in stereo vision algorithms are implemented via software and hardware categorized in terms of major attributes. The comparison of local and global methods provided by previously developed algorithms implemented on software and hardware based platforms was presented in this work.
2011	Tombari et al. [21]	Contribution and evaluation of stereo vision depth map algorithms in terms of their 3D object recognition ability.
2013	Tippetts et al. [8]	Reviewed stereo vision algorithms and their suitability for resource-limited systems. They have compiled and presented an accuracy and runtime performance data for all stereo vision disparity map algorithms in the past decade with an emphasis on real time performance.

3. PROCESSING STAGES OF STEREO VISION ALGORITHMS

3.1 Disparity Map Algorithms

Most stereo vision disparity map algorithms have been implemented using multistage techniques. These techniques, as codified by Scharstein and Szeliski, consist of four main steps as shown in Figure 2[11]. The input images are obtained from stereo vision sensors (i.e., from at least two cameras). Commonly, these cameras are arranged horizontally and set up which produce two or more corresponding images. For the explanation or the process as described by adopted taxonomy and their input images are assumed to be rectified images.

Next, the image pair to be analysed will pass through all of the blocks, in sequence, beginning with Step and ending with Step. The output of this process is a smooth disparity map. In essence, each block represents one or more algorithms whose performance can be measured based on the expected output. This taxonomy is widely used by many current developers of stereo vision disparity map algorithms [8, 12].

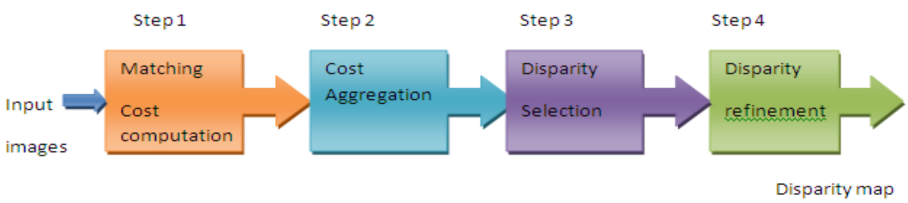

Figure 2: Simple framework for stereo vision Algorithm

In general, stereo vision disparity map algorithms can be classified into local or global approaches. A local approach is also known as area based or window based approach. This is because the disparity computation at a given point (or pixel) depends only on the intensity values within a predefined support window. Thus, such method considers only local information and therefore has a low computational complexity and a short run time. Local methods include all four steps of the taxonomy. Examples of implementation of such methods are provided by the work of Mattoccia et al. [13], Arranz et al.[14], and Xu et al. [15]. The disparity map value assignment is achieved through winner take all (WTA) optimization. For each pixel, the corresponding disparity value with the minimum cost is assigned to that pixel. The matching cost is aggregated via a sum or an average over the support window.

A global method treats disparity assignment as a problem of minimizing a global energy function for all disparity values. Such a method is formulated as an energy minimization process with two terms in the objective function (i.e., a data term, which penalizes solutions that are inconsistent with the target data and a smoothness term, which enforces the piecewise smoothing assumption with neighbouring pixels). The smoothness term is designed to retain smoothness in disparity among pixels in the same region. The disparity map is produced by assigning similar depth values to neighbouring pixels. Global methods produce good results but are computationally expensive. Therefore, they are impractical for use in real time systems. Global methods typically skip Step of the taxonomy depicted in Figure 2 (i.e., they do not perform cost aggregation and therefore contain only three steps) [16-18]. Markov random field (MRF) modelling is the approach that is most common approach used in global methods. This type of modelling uses an iterative framework to ensure smooth disparity maps and high similarity between matching pixels.

3.1.1 Cost Aggregation

Cost aggregation is the most important stage and general performance of a stereo vision disparity map algorithm but especially for local methods. The purpose of cost aggregation is to minimize matching uncertainties. Cost aggregation is need to the information obtained for a single pixel upon calculating the matching cost is not sufficient for precise matching. Local methods aggregate the matching cost by summing them over a support region [11]. This support region is typically defined by a square window centred on the current pixel of interest, as shown in Figure 3(a). The most straightforward aggregation method is to apply a simple low-pass filter in the square support window. The fixed-size window (FW) technique (e.g., binomial or Gaussian, uniform (box filters)) suffers an increased error rate when the size of the support window is increased over a certain threshold. Moreover, this method requires the parameters to be set to values suitable for the particular input dataset. Otherwise, it tends to blur object boundaries [19]. To avoid fattening artifacts near depth discontinuities, methods using shifting window or multiple windows (MW) as well as methods using adaptive windows (AW), windows with adaptive sizes, or adaptive support weights (ASW) have been developed.

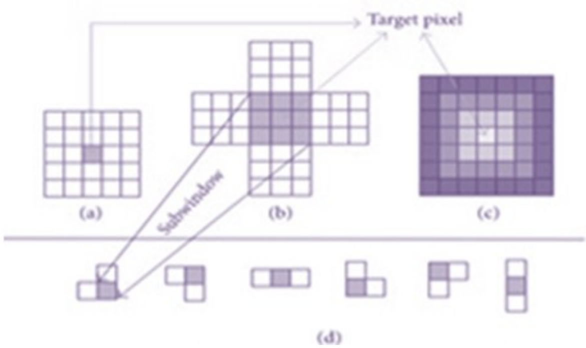

Figure 3: Cost aggregation Algorithm

In the MW technique, multiple windows are selected from among a number of candidates based on the support windows that produce smaller matching costs. This method was implemented by Hirschmüller et al. [20] and Veksler [21] in their previous studies of real time stereo vision disparity map algorithms. Their experimental results reveal difficulties in preserving dedicated pixel arrangements in disparity maps, especially at object boundaries. This occurs because of the shape of the support windows. This approach is imperfect for a small number of candidates. To resolve this problem, the AW technique was developed to reduce the errors in the disparity map caused by boundary problems. In this method, the support regions are constructed as approximations to the local image structures. Figure 3(b) illustrates the application of this method with five sub-windows with dimensions of 3 × 3. These sub-windows must be located near the target pixel as shown in Figure 3. The cost aggregation with the minimum matching cost value for this pixel is calculated. For example, the cost can be calculated as the summation over the target pixel sub-window and any two other adjacent sub-windows. The chosen shape of the valid matching windows for aggregation can therefore be any of the shapes shown in Figure 4(d). In practice, the shape of the adaptive window is adaptively varied to reflect the local image content, such as corners and edges.

The AW technique was implemented by Lu et al. [22] who achieved high quality results both near depth discontinuities and in homogenous regions. Lu's work was improved upon by Zhang et al. [23] through a modification to the concept of adaptive support regions. They developed and support regions with arbitrarily adaptive shapes and implemented the algorithm on a GPU for real time applications. The shapes of these support regions are more flexible and are not restricted to be rectangles. These authors achieved high matching accuracy with real time implementation. AW techniques and the algorithm attempts to find support windows that fit the shape or size of each region, while preventing them from crossing object boundaries. Furthermore, this technique is able to reduce computational costs as discussed by Chen and Su [24]. These authors proposed a shape adaptive low complexity technique for eliminating computational redundancy between stereo image pairs for pixels matching. They grouped pixels with the same depth value to reduce the number of computations.

A comparative study of the use of different support region techniques in the cost aggregation stage was performed by Fang et al. [25]. This study addressed the FW, AW, and ASW approaches. The authors concluded that the most advantageous technique for cost aggregation is the ASW approach. In this techniques are described each pixel in the support region is assigned a support weight, which depends on its intensity dissimilarity and spatial distance from the anchor pixel as shown in Figure 3(c). The target pixel which is located at the center is assigned different weight depending on distance as indicated by the different tone of colors.

Generally, for typical ASW a technique, (11) is used to aggregate the matching costs at pixel and disparity where a square support window is centered on pixel. The window size is a user defined parameter. The value of the function represents the possibility that a pixel will possess a disparity value similar to that of the window's center pixel represents a target pixel with a disparity value. Ideally, should return a value of "1" if pixels and have equal disparity values and "0" otherwise. Chen et al. [26] developed a trilateral filter based on the ASW approach with using a bilateral filter. They also added a new weighted term to increase the robustness against object boundaries.

Essentially, in ASW application, a higher weight will be allocated to a pixel if its intensity is more similar to that of the anchor pixel and if it is located at a smaller distance from the anchor pixel, as implemented by Zhang et al. [27]. This method is produce a disparity map in which the object boundaries are well preserved and their accuracy is very high compared with the previous methods reported in their literature. Hosni et al. [28] presented an extensive evaluation of ASW regions. They performed their test on a GPU to evaluate whether the speed and computational efficiency were sufficient for real time responses. Their evaluations indicated that the ASW approach produces outstanding results in terms of both computational efficiency and the quality of the generated disparity maps. Nalpantidis and Gasteratos [29] developed a new approach based on the ASW technique. They combined it with the quantified gestalt law to calculate a weighting factor. In general, a correlation weight reflects the proximity, similarity, and continuity between both input images (i.e., left and right images).

3.2 Mean Shift Algorithm

Segmentation is mainly used into subdividing an image into its constituent regions or object. The level up to which the subdivision is carried out depends on the problem being solved. [30]. The mean shift algorithm is a nonparametric clustering technique which does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters. Given n data points xi, i = 1, ..., n on a d-dimensional space Rd, the multivariate kernel density estimate obtained with kernel K(x) and window radius h is:

$$f(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \tag{1}$$

For radically symmetric kernels, it suffices to define the profile of the kernel k(x) satisfying:

$$K(\mathbf{x}) = c_{k,d}\, k(\|\mathbf{x}\|^2) \tag{2}$$

Where ck,d is a normalization constant which assures K(x) integrates to 1. The modes of the density function are located at the zeros of the gradient function f(x) = 0. The gradient of the density estimator (1) is:

$$\begin{aligned}\nabla f(\mathbf{x}) &= \frac{2c_{k,d}}{nh^{d+2}}\sum_{i=1}^{n}(\mathbf{x}_i-\mathbf{x})g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)\\ &= \frac{2c_{k,d}}{nh^{d+2}}\left[\sum_{i=1}^{n}g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)\right]\left[\frac{\sum_{i=1}^{n}\mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n}g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}-\mathbf{x}\right]\end{aligned} \tag{3}$$

Where g(s) = −k _ (s).

The first term is proportional to the density estimate at x computed with kernel G(x) = cg,dg(x 2) and the second term:

$$\mathbf{m}_h(\mathbf{x}) = \frac{\sum_{i=1}^{n}\mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n}g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}-\mathbf{x} \tag{4}$$

is the mean shift. The mean shift vector always points toward the direction of the maximum increase in the density. The mean shift procedure, obtained by successive.

• Computation of the mean shift vector mh(xt),

• Translation of the window xt+1 = xt +mh(xt) is guaranteed to converge to a point where the gradient of density function is zero.

Using Mean-Shift on Colour Models Two approaches:

1) Create a colour —likelihood image, with pixels.

2) Weighted by similarity to the desired colour (best for un-coloured objects) Represent colour distribution with a histogram. Use mean-shift to find region that has most similar distribution of colours.

4. CONCLUSION AND FUTURE WORKS

In this paper analysed and some algorithms, techniques, methods for enhance the Image Processing. Each one method or algorithm have some performance ratio not only the advantages and also have some drawbacks within that. In future work will choose any one algorithm which is most efficient and suitable to do better accuracy for image process and then apply some enhancement within that to proof much better than the old performance.

REFERENCES

[1] R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2010.

[2] B. H. Bodkin, Real-time mobile stereo vision [M.S. thesis], University of

Tennessee, 2012.

[3] M. Hansard, S. Lee, O. Choi, and R. P. Horaud, Time-of-Flight Cameras: Principles, Methods and Applications, Springer, 2012.

[4] S. Foix, G. Alenyà, and C. Torras, "Lock-in time-of-flight (ToF) cameras: a survey," IEEE Sensors Journal, vol. 11, no. 9, pp. 1917–1926, 2011.

[5] M. Y. Kim, S. M. Ayaz, J. Park, and Y. Roh, "Adaptive 3D sensing system based on variable magnification using stereo vision and structured light," Optics and Lasers in Engineering, vol. 55, pp. 113–127, 2014.

[6] S. Zhang, C. Wang, and S. C. Chan, "A new high resolution depth map estimation system using stereo vision and depth sensing device," in Proceedings of the IEEE 9th International Colloquium on Signal Processing and Its Applications (CSPA '13), pp. 49–53, IEEE, Kuala Lumpur, Malaysia, March 2013.

[7] W. Kazmi, S. Foix, G. Alenyà, and H. J. Andersen, "Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: analysis and comparison," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 88, pp. 128–146, 2014.

[8] B. Tippetts, D. J. Lee, K. Lillywhite, and J. Archibald, "Review of stereo vision algorithms and their suitability for resource-limited systems," Journal of Real-Time Image Processing, pp. 1–21, 2013.

[9] X. Xiang, M. Zhang, G. Li, Y. He, and Z. Pan, "Real-time stereo matching based on fast belief propagation," Machine Vision and Applications, vol. 23, no. 6, pp. 1219–1227, 2012.

[10] L. Nalpantidis and A. Gasteratos, "Stereovision-based fuzzy obstacle avoidance method," International Journal of Humanoid Robotics, vol. 8, no. 1, pp. 169–183, 2011.

[11] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International Journal of Computer Vision, vol. 47, no. 1–3, pp. 7–42, 2002.

[12] N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, "Review of stereo vision algorithms: from software to hardware," International Journal of Optomechatronics, vol. 2, no. 4, pp. 435–462, 2008.

[13] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 8, pp. 993–1008, 2003.

[14] Á. Arranz, Á. Sánchez, and M. Alvar, "Multiresolution energy minimisation framework for stereo matching," IET Computer Vision, vol. 6, no. 5, pp. 425–

434, 2012.

[15] L. Xu, O. C. Au, W. Sun et al., "Stereo matching by adaptive weighting selection based cost aggregation," in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '13), pp. 1420–1423, Beijing, China, May 2013.

[16] J. Salmen, M. Schlipsing, J. Edelbrunner, S. Hegemann, and S. Luke, "Real-time stereo vision: making more out of dynamic programming," in Computer Analysis of Images and Patterns, pp. 1096–1103, Springer, Berlin, Germany, 2009.

[17] J. M. Pérez and P. Sánchez, "Real-time stereo matching using memory-efficient Belief Propagation for high-definition 3D telepresence systems," Pattern Recognition Letters, vol. 32, no. 16, pp. 2250–2253, 2011.

[18] Y.-C. Wang, C.-P. Tung, and P.-C. Chung, "Efficient disparity estimation using hierarchical bilateral disparity structure based graph cut algorithm with a foreground boundary refinement mechanism," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 5, pp. 784–801, 2013.

[19] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 8, pp. 993–1008, 2003.

[20] F. Tombari, S. Mattoccia, L. D. Stefano, and E. Addimanda, "Classification and evaluation of cost aggregation methods for stereo correspondence," in Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08), pp. 1–8, Anchorage, Alaska, USA, June 2008.

[21] F. Tombari, F. Gori, and L. Di Stefano, "Evaluation of stereo algorithms for 3D object recognition," in Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV '11), pp. 990–997, IEEE, Barcelona, Spain, November 2011.

[22] F. Ekstrand, C. Ahlberg, M. Ekström, and G. Spampinato, "High-speed segmentation-driven high-resolution matching," in Seventh International Conference on Machine Vision (ICMV 2014), p. 94451Y, International Society for Optics and Photonics, 2015.

[23] B. J. Tippetts, D.-J. Lee, J. K. Archibald, and K. D. Lillywhite, "Dense disparity real-time stereo vision algorithm for resource-limited systems," IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 10, pp. 1547–1555, 2011.

[24] S. H. Lee and S. Sharma, "Real-time disparity estimation algorithm for stereo camera systems," IEEE Transactions on Consumer Electronics, vol. 57, no. 3, pp. 1018–1026, 2011.

[25] R. K. Gupta and S.-Y. Cho, "Window-based approach for fast stereo correspondence," IET Computer Vision, vol. 7, no. 2, pp. 123–134, 2013.

[26] D. Scharstein and R. Szeliski, Middlebury Stereo Evaluation—Version 2, 2015.

[27] A. Fusiello, U. Castellani, and V. Murino, "Relaxing symmetric multiple windows stereo using Markov Random Fields," in Computer Vision and Pattern Recognition, vol. 2134 of Lecture Notes in Computer Science, pp. 91–105, Springer, 2001.

[28] R. Yang and M. Pollefeys, "Multi-resolution real-time stereo on commodity graphics hardware," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. I/211–I/217, Madison, Wis, USA., June 2003.

[29] H. Hirschmüller and D. Scharstein, "Evaluation of cost functions for stereo matching," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07), pp. 1–8, Minneapolis, Minn, USA, June 2007.

[30] Dr. Ali Hassan Al-Fayadh, Hind Rostom Mohamed, Raghad Saaheb Al-Shimsah, "CT Angiography Image Segmentation by Mean Shift Algorithm and Contour with Connected Components Image", International Journal of Scientific & Engineering Research Volume 3, Issue 8, August-2012.