I J C T A, 9(4), 2016, pp. 1809-1816

Reliable Stereo Matching with Depth Measurement in Videos

T. Disilva[1], R. Vedhapriya Vadhana[2] and Ruba Soundar[3]

Abstract

The stereo matching algorithm analyses the accuracy of disparity map by using simple local window based method. The proposed algorithms are developed for depth quality in the video processing. The traditional approach is not suitable for smaller window size. But, the proposed technique is perfect suitable for smaller window sizes. It is dominated by the aggregation window size, the resulting depth quality with respect to various window sizes and demonstrated. Where, the depth has calculated from the disparity result. The proposed method to determine the disparity estimate of each image pixel and a simple Sum of Absolute Differences (SAD) matching technique was adopted. This method improves the accuracy of stereo matching. The experimental results verify the efficiency and reliability of proposed method.

Index terms: Disparity, Local window based system, stereo matching algorithm, dilation, Sum of absolute difference.

1. INTRODUCTION

The success of new 3D services is a reality due to the increase of quality of experience. Although there are some factors and initial measurement devices in this field, there are still no common way and procedure to compare 3D video contents and integrated solutions and obtain an evaluation of quality. Considering the Qualinet White Paper, quality of experience is defined as the degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state. Quality is influenced by different factors such as content, network, device, application, user expectations and context, and more when 3D video is considered. While evaluating visual quality assessment over standard 2D video, there are several factors that are typically taken into account, such as sharpness, blocking effect, or blurring. The difference in perception between regular 2D video and 3D stereoscopic video extends the list of factors that will affect its visualization, thus, emerging the need to study them in order to offer a more complex 3D quality assessment. Stereoscopic 3D video perception is based on the fact that two different video signals are captured in order to feed each of the viewer's eyes. There is a signal aimed to be received by the left eye and another one aimed to be received by the right eye. This system tries to recreate the experience of watching a real world scene, where two different images are captured by each eye and the difference between them depends on the position of the elements in the world related to the viewer's position. This means that the system is feeding the observer with a disparity depth cue. But the experience of watching 3D TV is significantly different from a natural view, as the point of view is prefixed by the fixed point of view of the camera lenses that have captured the scene and so is the focus. Furthermore, in natural viewing, the eyes focus and converge to the same distance, but when looking at a 3D object displayed on a screen, a viewer's eyes must focus on the screen while, at the same time, they converge on a point in space that may be located beyond the screen, on the screen, or in front of the screen. This is known as the vergence accommodation conflict. This conflict

[1,2,3] Electronics and communication engineering Francis Xavier Engineering College Tirunelveli, Tamilnadu, India, *Emails: disjesus134@gmail.com, vedhapriya.vadhana@gmail.com, rubasoundar@gmail.com*

limits the amount of parallax that a viewer can tolerate avoiding discomfort, also known as the zone of comfort. This paper aims to study the effects of stereoscopic disparity in quality assessment through the analysis of depth maps of a sequence and its temporal evolution. We try to quantify objectively the effects of parallax, depth and motion, exporting the common situations in which discomfort is substantial, from opinions of observers derived from empirical and subjective tests.

2. EXISTING SYSTEM

In traditional algorithms can be divided into global optimization approaches and local window based approaches. An extensive complexity and quality comparisons of global and local optimization algorithms were reported. Global optimization provides high depth quality because of its energy optimization functions, but its computation load may be impractical for hardware implementation. Many local based algorithms use either a variable window size or an irregular window shape to deal with depth discontinuity regions. Different cost aggregation, post processing and pre-processing techniques are also developed to enhance the depth quality of local window-based methods. Generally the main hardware cost of local window based algorithms is dominated by the aggregation window size investigated the resulting depth quality with respect to various window sizes and demonstrated that the quality is unacceptable if the window size is smaller than 13 × 13. A relatively large window size is usually chosen for obtaining reliable matches, which leads to high complexity and memory cost for hardware design especially for irregular window shapes. Stereo-matching algorithms can be divided into two categories: global optimization approaches and local window based approaches.

2.1. Background of global based approaches

The Global optimization approaches such as dynamic programming graph cuts and belief propagation obtain good depth accuracy because of their energy optimization functions. The smoothing term often measures the differences between neighboring pixels' disparities where ñ is a monotonically increasing function of the disparity difference. Global optimization often achieves better depth quality than that obtained with local window-based approaches at the cost of high computational complexity and storage requirements, which may be impractical for VLSI implementation. In contrast local stereo-matching algorithms are more suitable for hardware realization and are more likely to meet the demands of real-time computation, though they usually sacrifice depth quality.

2.2. Background of local based approaches

Title The development of local window-based algorithms focus on the matching cost computation and aggregation. Specifically such algorithms choose their local window sizes and shapes, aggregate the matching costs according to the defined cost function and determine the final disparity using a local winner take all optimization at each point. The simplest local window-based method matches the corresponding pixels with a fixed square window. The method assumes that the pixels in the window have similar depth; therefore, estimation error occurs when the window shifts to a depth discontinuity. The algorithm in uses a shifting window to deal with depth discontinuity. When the window goes through a depth discontinuity the location of the window that minimize the matching cost is selected. Several studies have proposed modifying the shape of the window for properly handling depth discontinuity, such as those using a compact window and an adaptive binary window. The compact window algorithm chooses a window size and shape by optimizing over a large class of compact windows via a minimum ratio cycle. The minimum ratio cycle is used to find the window with the lowest matching cost. The boundaries of compact windows are located at edges of the image content. The adaptive binary window is used to segment regions with a given depth, and thus the window shape is irregular. To construct the support matching window the color similarity of pixels in a region.

3. MOTIVATION

To develop a Stereo-Matching algorithm and analyse the accuracy of disparity by using simple local window-based methods. Thus the preserving reliable information only.

4. PROPOSED SYSTEM

To The development of local window-based method focus on the matching cost computation and aggregation. The simplest local window-based method matches the corresponding pixels with a fixed square window. Stereo matching algorithm to generate the initial depth map so further raising the depth quality is possible extension.

4.1. Block diagram

In this Figure 4.1 a depth-reliability-based stereo-matching algorithm is proposed to generate a good depth result even for difficult cases. Note that the required matching window size is smaller than those of existing window-based approaches with similar depth quality. Therefore, the proposed algorithm consumes very limited resources and is thus suitable for VLSI realization. The algorithm can be divided into three stages: the SAD phase, the depth reliability computation phase and the reliable depth propagation phase, The SAD phase consists of a simple small window-based cost computation performed using a WTA decision unit. The results, containing many mismatches, are examined in the reliability computation phase. Poor decisions are removed and reliable depth values are extended to fill unreliable regions in the depth propagation stage.

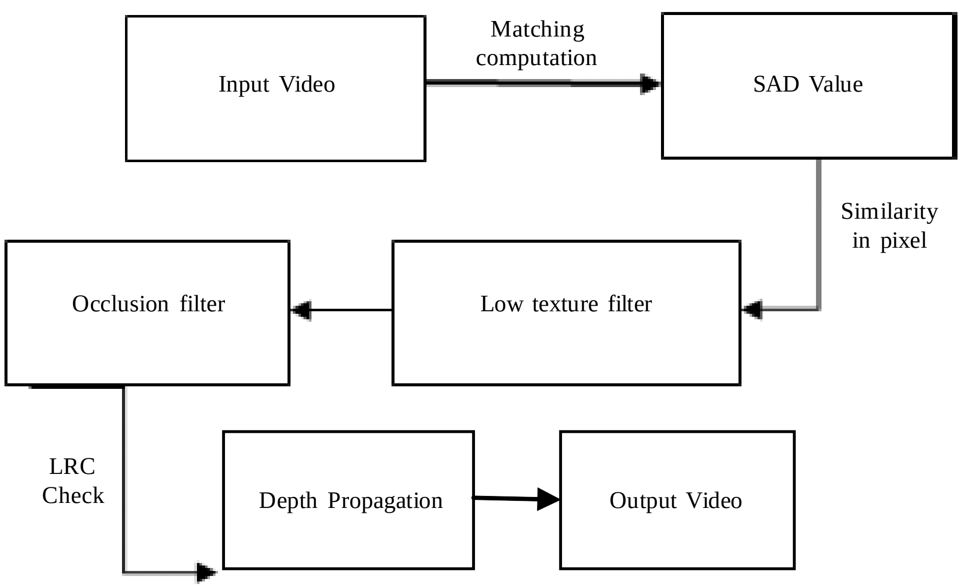

Figure 1: Depth measurement in the video processing

4.2. Matching Cost Computation

Matching cost computation is an initial cost computation of the stereo matching algorithm. Methods for matching cost computation can be classified into parametric and non-parametric methods. Parametric methods use pixel intensity Absolute Difference (AD) but are difficult to implement in the real-world environment because of their sensitivity to parameters of the cameras used as well as external conditions such as radial difference over exposure etc. On the other hand non-parametric measures such as rank census transform or gradient based measures are more robust than parametric methods. They transform intensity data into feature data and thus exhibit high robustness to illumination variations and exposure differences.

4.3. Disparity Computation

The process of disparity computation where input images are segmented first and then the same matching points in the left and right images are found. This procedure plays a very important role in our proposed stereo system. This idea is illustrated for an arbitrarily located 3D point P. Let a distant object is viewed by two cameras positioned in the same orientation but separated by a distance known as the baseline. Then the object will appear in a similar position in both stereo images. The distance between the objects in left and right images is known as disparity d defined by 8, where xL and xR are x coordinates of the projected 3D coordinate onto the left and right image planes IL and IR.

4.4. Stereo Matching Based on SAD

The basic geometry of the parallel camera at which rectified images come out. The baseline (T) denotes a line connecting the centers of two lenses and the focal length (f) means a distance from a center of lens to image plane. When one point P on 3-D object is projected onto two image planes by perspective projection two points indicate A and B respectively. The pixel B onto which P is projected in right image is called the corresponding pixel of A. A plane through the baseline and point P is termed an epipolar plane and the two straight-line intersections of an epipolar plane with the two image planes are called epipolar lines.

Due to $y1 = y2$ each point in one image must be observed in the other image on a known epipolar line which is the epipolar constraint. Suppose that the window size on each image is $M \times N$ every number in the matrix indicates a pixel of the window. The disparity denotes a distance between the centers of two windows. The SAD algorithm is one of the area-based approaches.

4.5. Depth measurement

High level computer vision tasks, such as robot navigation and collision avoidance, require 3-D depth information about the surrounding environment at video rate. Current general purpose microprocessors are too slow to perform stereo vision at video rate. For example, it takes several seconds to execute a medium-sized stereo vision algorithm for a single pair of images on a 2 GHz general-purpose microprocessor. To overcome this limitation, designers in the last decade have built custom designed hardware systems to accelerate the performance of the vision systems.

Software implementation allows one to exploit the parallelism that usually exists in image processing and vision algorithms and to build systems to perform specific calculations very quickly compared to software. By processing several parts of the data in parallel, we can speed up the overall functioning and achieve video-rate performance

Table 1
Background Dominated

S. no	Depth value	Number of accumulated
1	50	800
2	100	1500
3	150	500
4	200	0

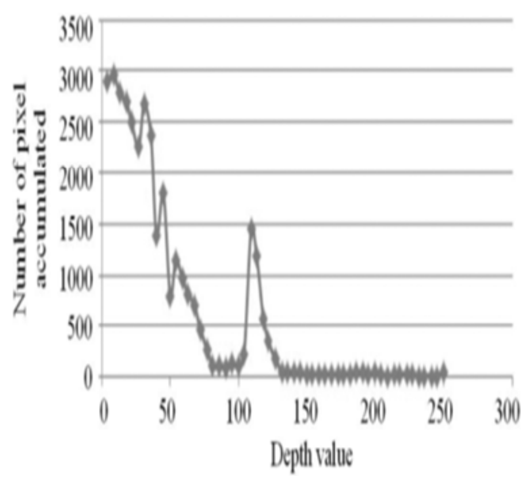

Figure 2: Background dominated

4.6. Accuracy

Here the accuracy has been restricted to analyse visual characteristics in terms of results of implementations achieved and categorize each approach as providing limited intermediate or high (L, M, H)accuracy in terms of quality or visibility of object detection. The quality of results was good. Object shadows of the object can be seen. It is been clearly seen occlusion in an image and accuracy is high.

4.7. Compressed Domain Video Analysis

Video compression techniques such as DCT coding, Quantization, Entropy coding, Motion estimation are widely used in video compression techniques. The main focus of this paper is to analyse video compression techniques required for video processing especially to discover how much amount of data to compressed, which techniques is faster and visual quality better and so on. We evaluate the video compression techniques for finding compression ratio in terms of performance speed and accuracy.

5. RESULT AND DISCUSSION

This chapter explains the results obtained in the proposed system using the Xilinx and shows the simulated MATLAB window for the proposed scheme. The MATLAB window at various simulation times are taken and given below:

5.1. Traffic video

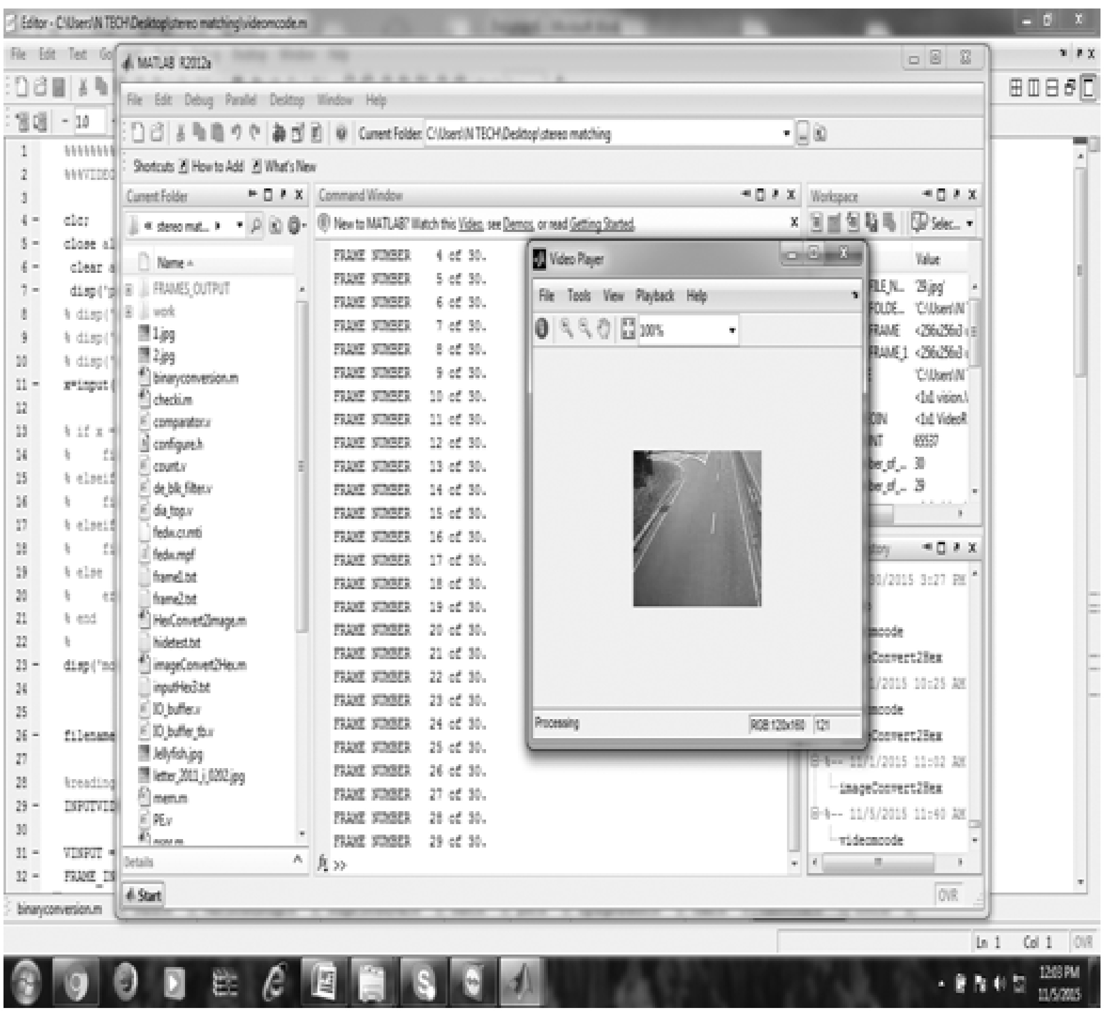

Figure 3: Input video by stereo matching flow

5.2. Disparity Dilation

In this figure 4 the traffic video have good quality of results. So, it has clearly seen in the traffic video and accuracy is high.

5.3. SAD Calculation

In this figure 5 is Mvx1, Mvy1 denotes the pixel variation of two frames. If both the output values are similar, it denotes no variation in pixel; if they are different, it denotes the pixel variation.

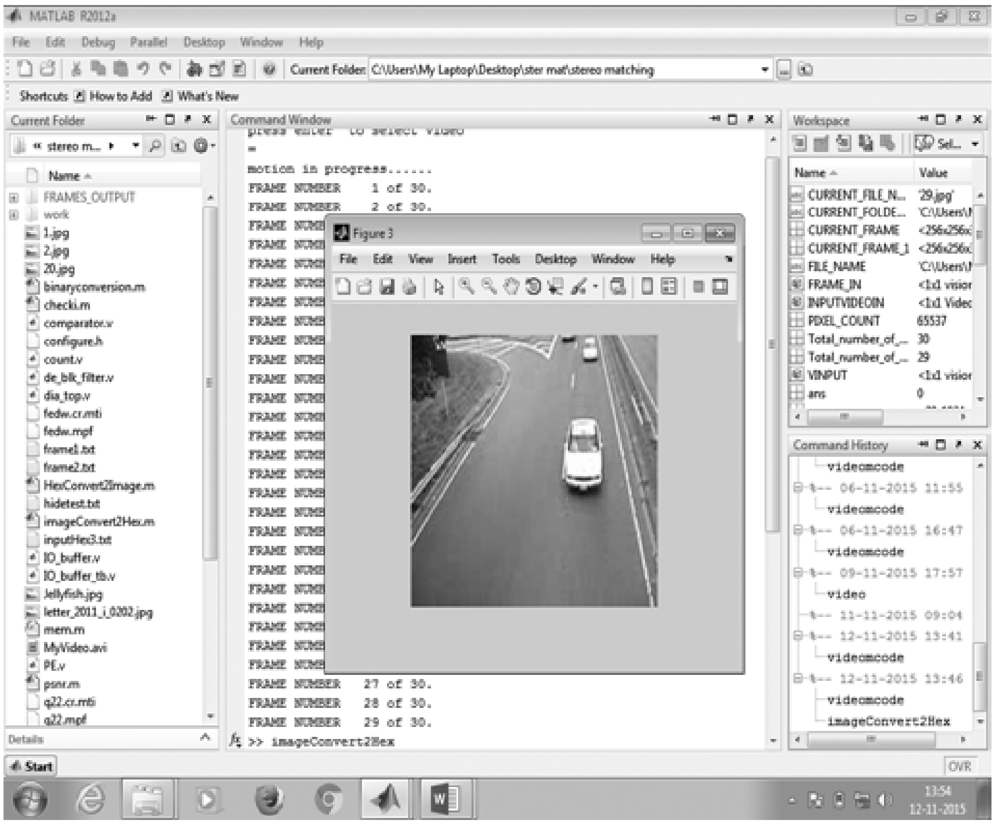

Figure 4: Disparity dilation in the traffic video

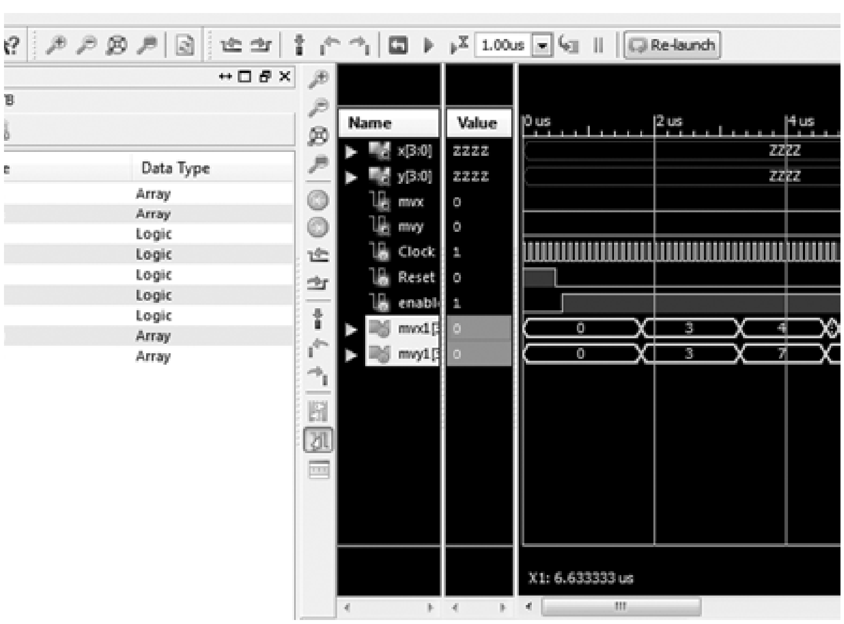

Figure 5: SAD calculation using Xilinx

5.4. Parameter Calculation

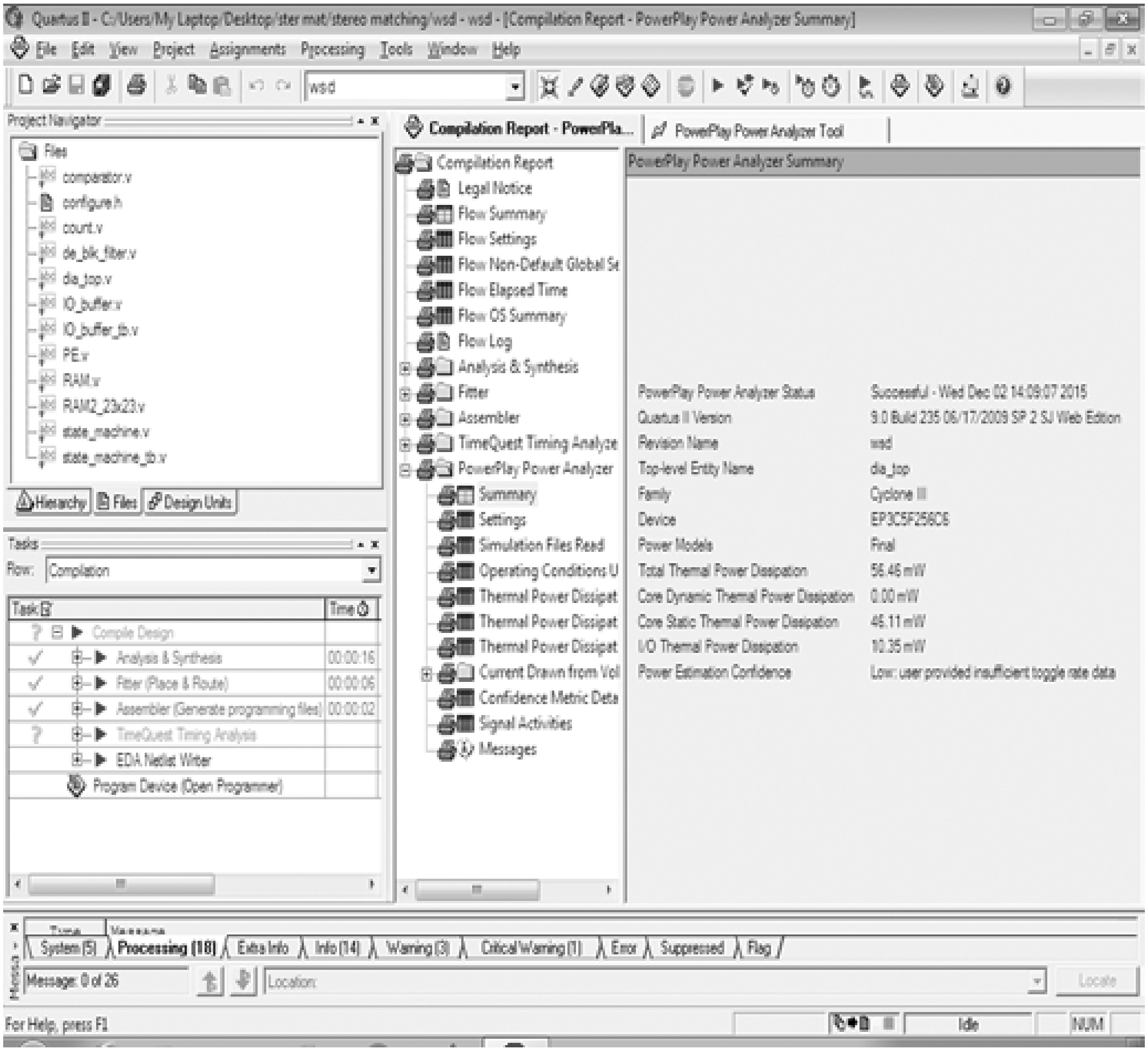

Figure 6: Parameter calculation

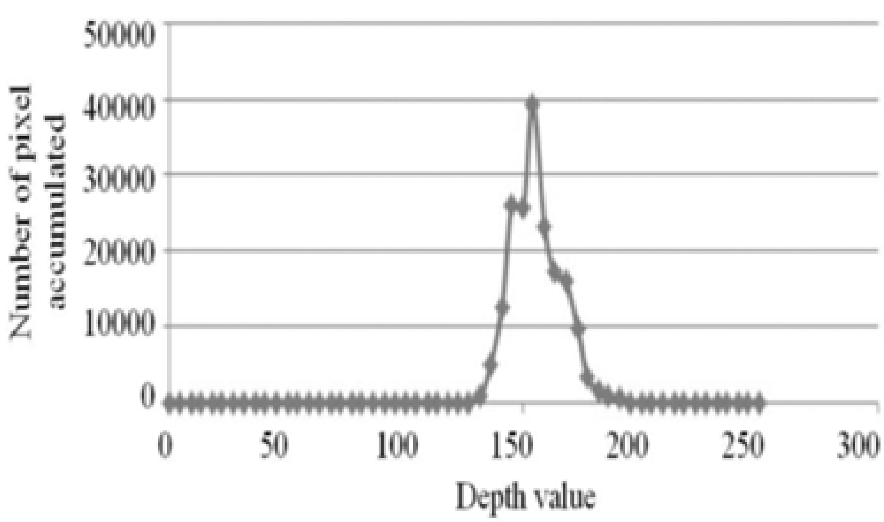

Figure 7: Normal distribution

6. CONCLUSION

In this paper, we proposed that can be applied of the stereo-matching techniques in the video processing. To introducing more aggressive stereo-matching algorithms to generate the initial depth map for further raising the depth quality is the possible extension of this paper. We analyze the performance of the proposed scheme through simulations. The simulation results showed that our scheme works better even in mobile environments.

In future work will be focus on, to motion estimation detection with high accuracy in the video processing by using hybrid Stereo matching algorithm.

REFERENCES

[1] Anhong Wang Mei Zhao, Bing Zen, Lei Liu4 and Huihui Bai, July (2014), "Depth Coding Based on Compressed Sensing with Optimized Measurement and Quantization" Journal of Information Hiding and Multimedia Signal Processing ISSN 2073-4212 Ubiquitous International Volume 5, Number 3.

[2] Chavan, Kapil, Halkarnikar IOSR Journal of Computer Engineering (IOSR-JCE), (Sep-Oct. 2014), "Occlusion detection in video sequences" e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 16, Issue 5, Ver. V, pp. 01-08.

[3] Frank Perbet, Björn Stenger, Atsuto Maki, (2009), "Random Forest Clustering and Application to Video Segmentation" Toshiba Research Europe Cambridge Research Laboratory 208 Science Park.

[4] Hubert, B. Stabernack, and F. Zilly, (May 2013), "Architecture of a low latency image rectification engine for stereoscopic 3-D HDTV processing," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 5, pp. 813–822.

[5] Kang, and Y. Ho, (2012), "Depth video coding using adaptive geometry based intra prediction for 3D video system" IEEE Trans. Multimedia, vol. 14, no. 1, pp. 121-128.

[6] Kamencay P., Breznan M. (2011), In Proceedings of 21st International Conference Radioelektronika "A stereo depth recovery method using belief propagation" Brno Czech Republic, pp. 383-386.

[7] Liu, A. Wang, Z. Li, and K. Zhu (2011), "An improved distributed compressive video sensing based on adaptive sparse basis", Proc. of the IEEE International Conference on Robot, Vision and Signal Processing, pp. 137-140.

[8] Nelson Yen-Chung Chang, Tsung-Hsien Tsai, Bo-Hsiung Hsu, Yi-Chun Chen, and Tian-Sheuan Chang (JUNE 2010), "Algorithm and Architecture of Disparity Estimation With Mini-Census Adaptive Support Weight" IEEE Transactions on circuits and systems for Video technology, Vol. 20, No. 6.

[9] Patrik Kamencay, Martin Breznan, Roman Jarina, Peter Lukac, Martina Zachariasova (April 2012), "Improved Depth Map Estimation from Stereo Images Based on Hybrid Method" Radio Engineering, Vol. 21, No. 1.

[10] Shen, W. S. Kim, S. K. Narang, A. Ortega, and Wey (2010), "Edge-adaptive transforms for efficient depth-map coding", Proc. of the 28th Picture Coding Symposium, pp. 566-569.

[11] Stein and M. Hebert, (Sep. 2007), "Combining local appearance video in Motion cues for Occlusion boundary detection" in Proc. Brit. Mach no. 6.

[12] Zhang, A. Wang, B. Zeng, and L. Liu, (2014), "Adaptive distributed and compressed video sensing", Journal of Information computer science in Multimedia Signal Processing, vol. 5, no. 1, pp. 98-106.

[13] Zhao, C. Zhu, Z. Chen, and L. Yu, (2011), "Depth no synthesis error in Model for view synthesis" in 3-D video, IEEE Trans. Image Processing.