American Journal of Science and Technology
2015; 2(6): 346-353
Published online January 26, 2016 (http://www.aascit.org/journal/ajst)
ISSN: 2375-3846

American Association for Science and Technology

Stereo Vision Distance Estimation Employing Canny Edge Detector with Interpolation Algorithm

Raad Hamdan Thaher, Zaid Khedher Hussein

Electrical Engineering Department, College of Engineering, AL-Mustansiriyah Univrsity, Baghdad, Iraq

Email address

Raadthaher55@gmail.com (R. H. Thaher), zaid_khedher@yhaoo.com (Z. K. Hussein)

Citation

Raad Hamdan Thaher, Zaid Khedher Hussein. Stereo Vision Distance Estimation Employing Canny Edge Detector with Interpolation Algorithm. *American Journal of Science and Technology*. Vol. 2, No. 6, 2015, pp. 346-353.

Keywords

Stereo Vision,
Disparity,
SAD,
CBMA,
Interpolation,
Bilinear

Received: October 14, 2015
Revised: November 26, 2015
Accepted: November 28, 2015

Abstract

Stereo vision system is used to reconstruct a 3D scene from 2D images captured by a pair of optical cameras (left and right) to estimate the distance of the object. The modified version for the Sum of Absolute difference (SAD) algorithm is called the Canny Block Matching Algorithm (CBMA) which is used to find the Disparity map. The algorithm consist of two parts the Canny edge detector and Block matching technique with Sum of Absolute Difference (SAD) to determine the disparity map to reduce the execution time. The system has been implemented using two cameras arranged in a manner to enhance the detection range of objects from (30cm to 4m). The Interpolation algorithm consists of the Median filter and the Interpolation techniques, Such as: Bilinear, 1st order and 2^{nd} order polynomial, to enhance the output images from the three algorithms. The error percentage had been reduced about 2% and the execution time is reduced by the step size windows. The CBMA algorithms implemented using MATLAB (8.0) technical programing language where the Interpolation design in Microsoft Visual Basic (6.0). The practical system for stereo vision was implemented and tested.

1. Introduction

Stereoscopic vision is a technique used for inferring the 3D position of objects from two or more simultaneous images of a scene, Reconstruction of the world seen (3D information) by stereo vision cameras. [1]

The depth is obtained by computing stereo matching between two images. The technique is called triangulation, which represents the process of finding coordinates of a 3D point based on its corresponding stereo image points [2].

A number of studies have been reported on the disparity estimation problem since 1970's like Marr-Poggio In 1976 and 1979 [3, 4]. They first work in stereo vision system and used a cooperative algorithm to extract disparity information from stereo image pairs. Many researches in the literature deals with scope of this work [5-9].

There are many problems for estimating distance in stereo vision system to evaluate the performance of algorithm. Two criteria have been employed to approximate the execution time to the real time performance and to reduce the error so that the objects in the disparity map should be accurate enough to recognize.

However stereo vision system and stereo camera geometry are described by Raad [10].

2. Interpolation Techniques

Image interpolation is an important image processing operation applied in computer graphics, an image interpolation algorithm is used to convert an image from one resolution (dimension) to another resolution without losing the visual content in the picture [11]. In this work we take into account the performance of the most used interpolation techniques: bilinear, 1st order and 2nd order polynomial.

They are described below:

2.1. Bi-linear Interpolation

The commonly used linear methods, bilinear interpolation, have advantages in simplicity and fast implementation. Bilinear interpolation takes a weighted average of the four (2x2) neighborhood pixels to calculate its final interpolated value. The result is much smoother image than the original image [12, 13].

The interpolation kernel for bilinear interpolation is: [14]

$$\mathrm{u(x)} = \begin{cases} 0 & |x| > 1 \\ 1 - |x| & |x| < 1 \end{cases} \tag{1}$$

where x = distance between interpolated point and grid point.

As shown in Figure (1), the intensity value *I(x, y)* at the interpolated point *(P)* at *(x, y)* in the image can be estimated as: [13]

$$I(x,y) = \frac{(x_2-x)(y_2-y)}{(x_2-x_1)(y_2-y_1)} I(x_1,y_1) + \frac{(x-x_1)(y_2-y)}{(x_2-x_1)(y_2-y_1)} I(x_2,y_1) + \frac{(x_2-x)(y-y_1)}{(x_2-x_1)(y_2-y_1)} I(x_1,y_2) + \frac{(x-x_1)(y-y_1)}{(x_2-x_1)(y_2-y_1)} I(x_2,y_2) \tag{2}$$

$$= w_1(x_1,y_1) + w_2(x_2,y_1) + w_3(x_1,y_2) + w_4(x_2,y_2) \tag{3}$$

where *I(x1, y1)*, *I(x1, y2)*, *I(x2, y1)*, and *I(x2, y2)* are the intensity values of the four neighboring pixels.

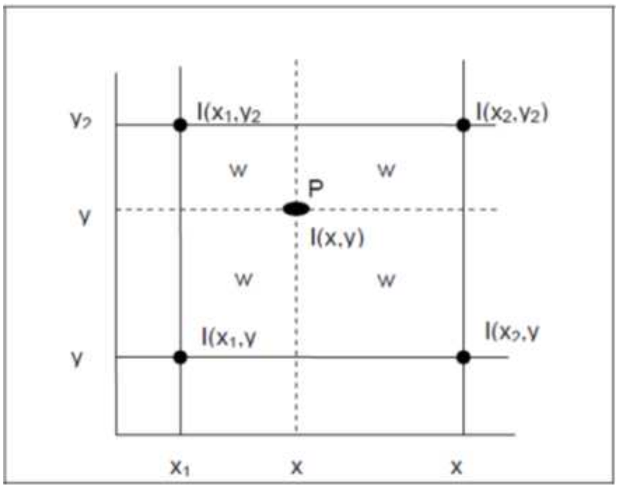

Figure 1. *Bilinear interpolation using four neighboring points.*

2.2. Polynomial Interpolation

Polynomial interpolation is the interpolation of a given data set by a polynomial: given some points, find a polynomial which goes exactly through these points. Only the first-order (linear) polynomial and the second-order polynomial were Used [15].

2.2.1. 1st Order Polynomial

The 1st order polynomial in one variable with constant coefficients and a polynomial in a single variable have the form: [16]

$$a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n \tag{4}$$

where $a_0, \dots, a_n$ are constants, or more generally elements of a ring, and x is variable. The usual (commutative, distributive) laws are valid for arithmetic operations [16].

2.2.2. 2nd Order Polynomial

The 2nd order polynomial is polynomial function have one or two variables with constant coefficients had the form: [17]

$$a_nx^ny^m + \cdots + a_{22}x^2y^2 + a_{21}x^2y + a_{12}xy^2 + a_{11}xy + a_{10} + a_{01}y + a_{00} \tag{5}$$

where $a_{00}, \dots, a_{nm}$ are constants, or more generally elements of a ring and x ,y are variables.[18]

For the matching algorithms used, the error measured between the computed disparity map $d_{C(x,y)}$ and the ground true disparity map $d_{T(x,y)}$ for total number of pixels (N) is [19]:

$$\boldsymbol{p} = \frac{1}{N} \sum_{(x,y)} | d_{C(x,y)} - d_{T(x,y)} | \tag{6}$$

For Interpolation techniques there two methods used, they are:

1. Mean Square Error (MSE)

The (MSE) of an estimator measures the average of the squares of the "errors", (i.e, the difference between the estimator and what is estimated). If $\hat{Y}$ is a vector of n predictions, and Y is the vector of the true values, then the (estimated) MSE of the predictor can be expressed by [20]:

$$MSE = \frac{1}{n} \sum_{i=0}^{n} (\hat{Y}_i - Y_i)^2 \tag{7}$$

3. Median Absolute Deviation (MAD)

The (MAD) is a robust measure of the variability of a univariate sample of quantitative data. For a univariate data set X_1, X_2, ..., X_n, the MAD is defined as the median of the absolute deviations from the data's median [21]:

$$MAD = median_i(|X_i - median_j(X_j)|) \tag{8}$$

4. Tested Stereo Vision Images

The real and rectified stereo vision image pairs, provided as a database with ground-truth (true disparity map) that is required to calculate the percentage error. One of the stereo pairs called (Tsukuba), as shown in Figure 2 [22].

5. Canny Block Matching Algorithm (CBMA)

The algorithm consists of two stages a Canny edge detector and the Block matching technique using SAD to determine the disparity map.

5.1. Canny Edge Detector

The Canny edge detector is an edge detection operator using a multi-stage algorithm to detect a wide range of edges in the images. The Canny edge detector uses the first derivative of a Gaussian; so to begin with, the raw image is convolved with a Gaussian filter and the edge direction angle is rounded to one of four angles representing vertical, horizontal and the two diagonals. Canny uses thresholding with hysteresis, and they require two thresholds (high and low) [23].

The Canny operator has three variables (Sigma standard deviation of Gaussian (σ) and two thresholds (t_{low} and t_{high})): sigma range from (0.5 - 5), t_{high}h range from (0.01-0.2), and t_{low} depends on t_{high} as $t_{low} = 0.4 * t_{high}$.

(a) Left image (b) Right image (c) true disparity map

Figure 2. *Tsukuba stereo image pair and true disparity map.*

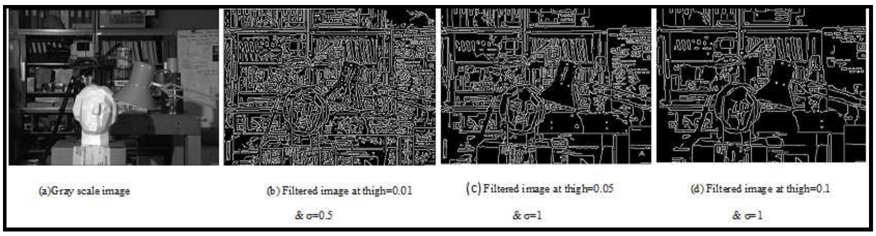

Figure 3. *Tsukuba image, (a) before applying Canny filter. And after applying Canny edge detector at different threshold and sigma values: (b) att$_{high}$=0.01 & σ=0.5, (c) at t_{high}=0.05& σ=1and (d) at t_{high}=0.1& σ=1.*

Figure 3 shows the images before and after applying Canny edge detector at different thresholds and sigma values, The optimal result for this image is obtained at threshold value as in ((c) ($t_{low} = 0.4 * t_{high}$, t_{high}0.05 & σ=1 and the time is 0.174785 seconds)), it can be considered as an optimal value since it compromises between accuracy of matching results and speed up the execution time.

5.2. Block Matching Techniques

Block matching is a technique used in stereo vision to solve the correspondence problem, each block from the left image, is matched into a block in the right image by shifting the left block over the searching area for pixels in right image, as shown in Figure 4 [24].

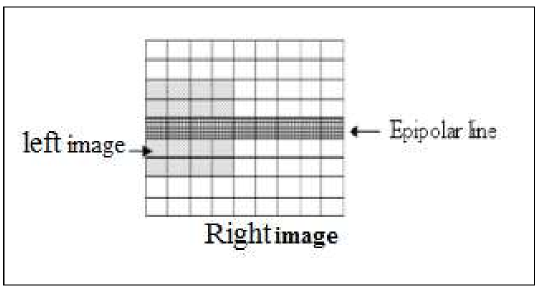

Figure 4. *SAD block matching techniques.*

In order to reduce the computational time, in which the search for the matched blocks is performed by moving the reference window of the reference image block-by-block over the target image, and the disparity value calculated for the central pixel of the window represents the same disparity value for all window's pixels neighboring the central pixel. By this method large amount of SAD computation time can be reduced.

After testing the CBMA algorithm with the pair's images for (Tsukuba) the results for the disparity map are obtained as shown in Figure (5), the execution time and error calculation according to CBMA algorithm are shown in the Table 1.

Table 1. *The execution time and error calculation according to CBMA.*

Window-size	Error percentage (%)	Execution time (Sec.)
[3×3]	23.87	0.4379
[5×5]	18.03	0.3771
[7×7]	14.91	0.3984
[9×9]	13.56	0.3657
[11×11]	12.68	0.3593
[13×13]	11.24	0.3718
[15×15]	12.93	0.3689
[17×17]	13.20	0.3611
[19×19]	13.40	0.3685
[21×21]	13.11	0.3709
[23×23]	14.22	0.3728
[25×25]	14.47	0.3639
[27×27]	13.40	0.3655

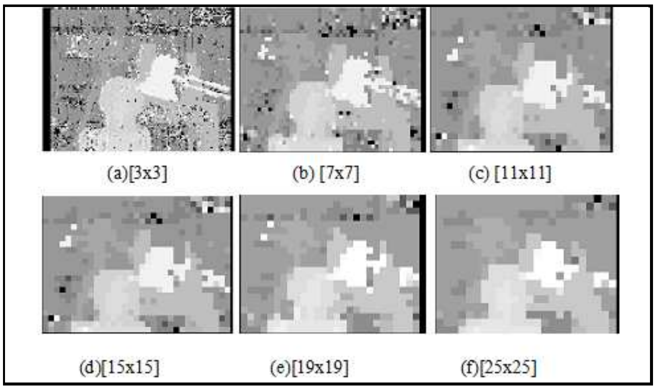

Figure 5. *Disparity maps of Tsukuba image result from CBMA for different window sizes.*

From Table 1 it can be noted that a large amount of SAD computation can be reduced, consequently the computation time is dramatically reduced the execution time between (0.4379 and 0.3593) sec is good for application to robots. The Error percentage with different windows shows the optimal window is [13x13] and gives minimum error percentage with execution time of (0.3718) sec. So, the algorithm (CBMA) has reduced the execution time, but the Error percentage is slightly more than the other algorithms. The window [13x13] is considered as a suitable (acceptable) value for disparity map computation.

6. Interpolation Algorithm

To enhance the output images of the disparity map that can be done from the matching algorithms (CBMA) which used the median filter method with some Interpolation techniques mention earlier.

6.1. Median Filter

The median filter is an effective method that can distinguish the out-of-range isolated noise from image features such as edges and lines. Specifically, the median filter replaces a pixel by the median, instead of the average, of all pixels in a neighborhood w [25]:

$$y[m,n] = median\{x[i,j], (i,j) \in w\} \quad (9)$$

where w represents a neighborhood defined by the user, centered around location $[m,n]$ in the image. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value.

In this paper we concentrate on the performance of interpolation techniques: bilinear, 1st order and 2^{nd} order polynomials.

6.2. Algorithm Steps & Flow Chart

The Interpolation algorithm can be described as follows:

Step1. Read disparity map image for (CBMA) algorithm and any window size.

Step2. Set the Block size value and in case used only 3 or 5.

Step3. Apply median filter to the input image and save the new image.

Step4. Calculate the bilinear Interpolation and save the new image and compute the mean squared error (MSE) and the median absolute deviation (MAD).

Step5. Calculate the 1^{st} order polynomial and save the new image and compute the mean squared error (MSE) and the median absolute deviation (MAD).

Step6. Calculate the 2^{nd} order polynomial and save the new image and compute the mean squared error (MSE) and the median absolute deviation (MAD). The flowchart shown in Figure 6.

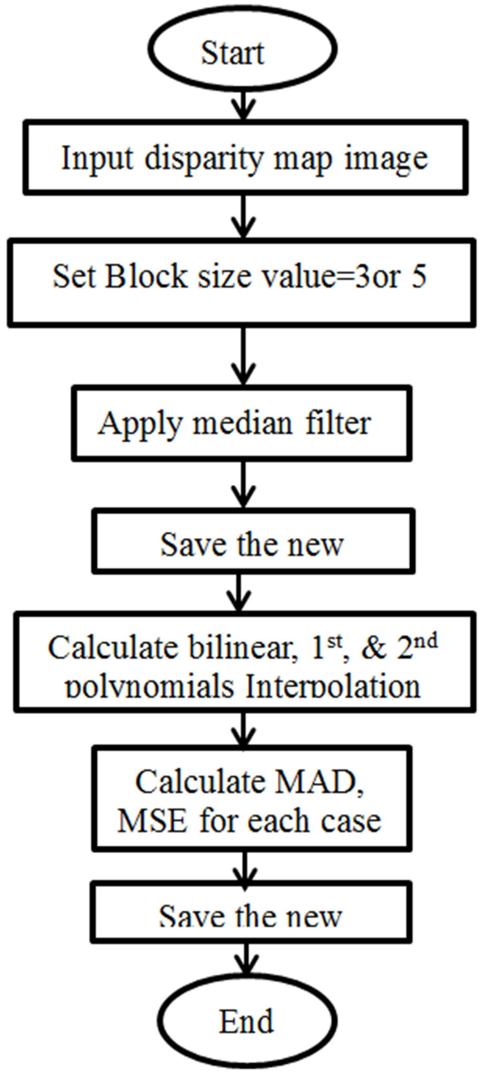

Figure 6. *The flow chart of Interpolation algorithm.*

6.3. The Execution Time with Interpolation Techniques

To compute the Execution time for the Interpolation techniques, first the actual time is registered before the of matching process starts and after it is accomplished then the time difference is divides by the step size that it used to do the interpolation image:

$$T=\frac{T_{tot}}{step\ size^2} \quad (10)$$

So to determine the time for the Interpolation process equation (10) becomes

$$T=\frac{T_{tot}}{step\ size^2}+\text{Interpolation time} \quad (11)$$

Where: T_{tot} is the time from the matching algorithms, step size is the block mask to calculate the interpolation image on the disparity map image and it is square because of using the square windows, and T is the total time after using the interpolation.

6.4. The Results

The algorithm (CBMA) outputs disparity maps for all windows sizes from (3x3) to (27x27) for the three Interpolation algorithms (Bilinear interpolation, 1st Order, and 2nd Order Polynomials) in Microsoft Visual Basic 6 and for two Block size (3*3) and (5*5) the results are shown in Table (2) which compares the Results for CBMA disparity map images from the three Interpolations (Bilinear, 1st order. & 2nd order polynomials) and (MSE, and MAD) are calculated for two block size (3*3) and (5*5) and the Error Percentage of the output disparity maps before and after the interpolation.

Table 2. *The Interpolation techniques (Bilinear interpolation, 1st Order, and 2nd Order Polynomials) and (MSE and MAD) for CBMA.*

Window Size CBMA	BS3 MAD	BS5 MAD	BS3 MSE	BS5 MSE	Error percentage (%)before Interpolation	Error percentage (%)After Interpolation
[3x3]bilinear	12	18	536.9	1018.9	21.62	19.36
1st Pol	13.3	19.2	572	1065.4		
2nd Pol	21.5	26.6	1312.9	1873.1		
[5x5]bilinear	6.3	9.8	221.5	424.5	17.04	16.99
1st Pol	6.8	10.7	231.9	455.6		
2nd Pol	14	18.5	765.3	1163.8		
[7x7]bilinear	5.4	9	206	405.3	13.73	12.37
1st Pol	5.9	9.8	216.2	423.8		
2nd Pol	12.8	17.1	654.6	1045.6		
[9x9]bilinear	4.2	6.9	130.5	248	12.79	11.21
1st Pol	4.5	7.4	135.4	258.1		
2nd Pol	10.1	13.6	493.6	725.6		
[11x11]bilinear	3.8	6.5	142.5	256.5	12.96	10.43
1st Pol	4	6.9	146.4	263.9		
2nd Pol	10	14	473.5	793.4		
[13x13]bilinear	3.3	5.5	102.8	192.2	11.34	10.54
1st Pol	3.4	5.8	104.8	197.4		
2nd Pol	8.4	12	390.8	644.3		
[15x15]bilinear	2.5	4.3	86.7	162.1	13.74	10.32
1st Pol	2.6	4.4	88.2	165.0		
2nd Pol	6.5	9.9	260.8	480.9		
[17x17]bilinear	2.5	4.3	85.2	155.9	12.34	11.73
1st Pol	2.7	4.6	87.3	159.6		
2nd Pol	7.3	10.4	330.9	580		
[19x19]bilinear	2.3	4.2	83.8	161.2	14.10	12.45
1st Pol	2.4	4.4	85.1	163.3		
2nd Pol	6.6	10.4	247.5	572.8		
[21x21]bilinear	2	3.5	63.2	123	14.23	12.65
1st Pol	2	3.6	63.9	124.3		
2nd Pol	5.4	9.1	179.7	491.1		
[23x23]bilinear	1.9	3.3	50.9	93	13.57	11.37
1st Pol	2	3.4	51.7	95		
2nd Pol	5.4	8.8	205.1	485		
[25x25]bilinear	1.5	2.3	33.9	51.7	14.25	12.22
1st Pol	1.6	2.4	34.3	53.2		
2nd Pol	4.2	8	145.5	480.3		
[27x27]bilinear	1.2	2.2	29.2	59.1	14.35	12.25
1st Pol	1.3	2.3	29.4	59.6		
2nd Pol	3.8	6.3	129.2	321.3		

Figures 7 & 8 show a comparison for the CBMA algorithm between the three Interpolation techniques (Bilinear, 1st order, & 2nd order polynomial) for both (MAD) and (MSE) respectively.

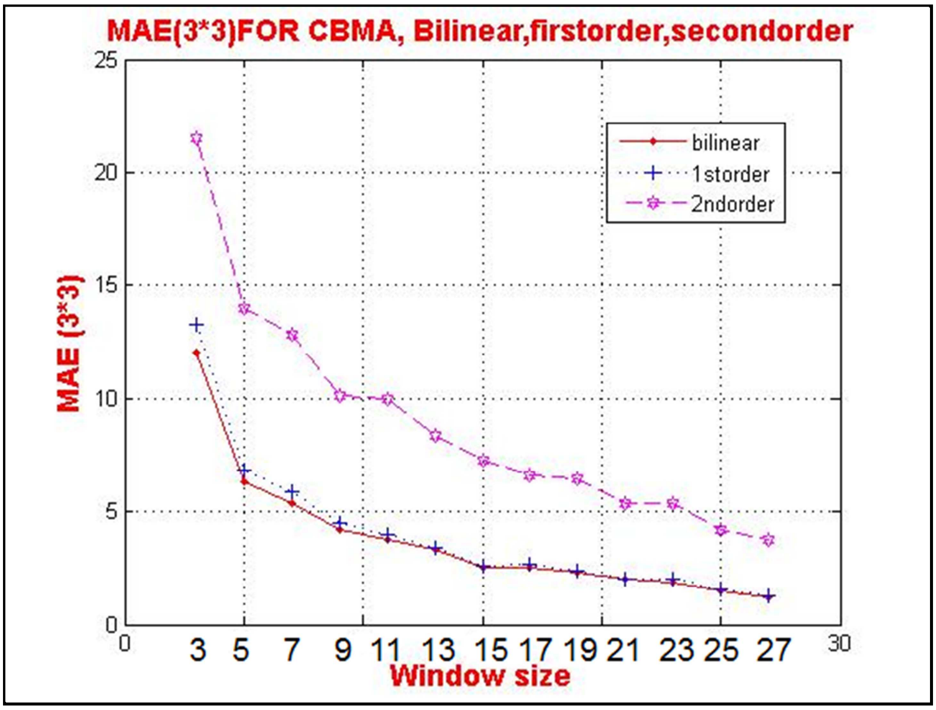

Figure 7. *Comparison of results (MAD) for CBMA between (Bilinear, 1st order, & 2nd order polynomial) block size (3*3).*

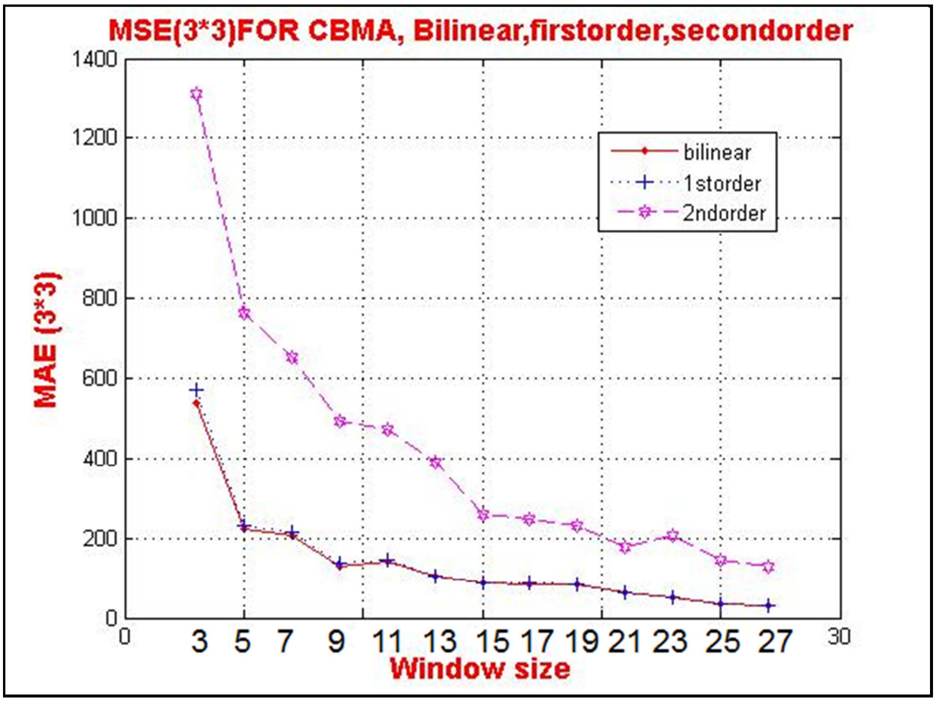

Figure 8. *Comparison of Results (MSE) for CBMA between (Bilinear, 1st order, & 2nd order polynomial) block size (3*3).*

It can be noted form the Figures 7, 8 and Table 2 that the Bilinear interpolation is better for block size (3*3) because it gives less values of the MSE and MAD and less Error percentage about (1% to 2%) which can be enhanced for real disparity map image.

Interpolation computation for block size (3*3) and (5*5) is shown in Figures (9), (10) respectively.

From Table 2 it can be shown that the Bilinear Interpolation gives the best results when used a block size (3*3) for (MAD) and (MSE) for the algorithms (CBMA) and the second nearest results is a 1st order polynomial and for the Error percentage computation for the new images (disparity map) from the Bilinear Interpolation of (3*3) block size the results is less about (1 to 2%) from the original disparity map for the (CBMA) algorithm.

It can be observed from Figures7 & 8 that the MSE (for three Interpolation techniques) give better results (less error) between the Bilinear and the 1st order polynomial than the MAD.

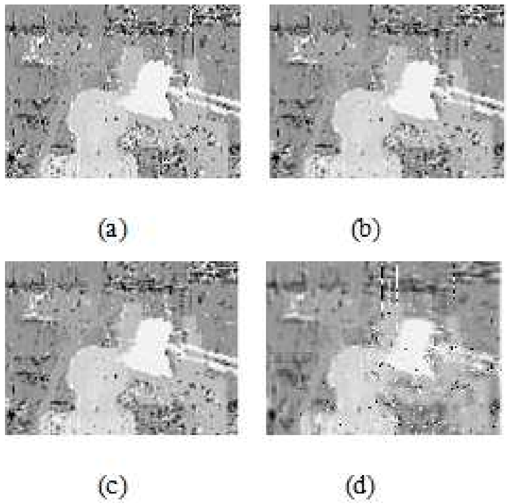

Figure 9. *Window size for CBMA is (3x3) block size (3) to compute the Interpolation (a) Origin CBMA (b) Bilinear (c) 1st order & (d) 2nd order.*

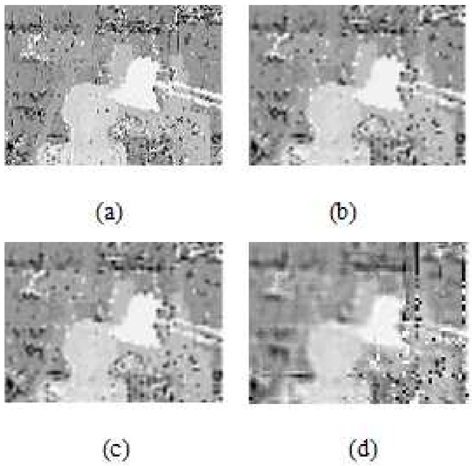

Figure 10. *Window size for CBMA is (3x3) block size (5) to compute the Interpolation (a) Origin CBMA (b) Bilinear, (c) 1st order & (d) 2nd order.*

7. Stereo Vision Hardware

Figure 11. *The implemented stereo vision system.*

The vision system hardware is implemented and of two identical models USB web cameras connected to PC with USB cable, as shown in Figure (11). The camera specification, High Resolution CMOS Color Sensor, Resolution: 1300k pixel (interpolated 5M), Interface: 2.0, Photo quality of 3 Megapixels. The system operates at a resolution of 320×240.

8. Depth Measurement

After stereo cameras have been calibrated and their parameters (intrinsic and Extrinsic) have been calculated, and rectification the images from that parameters the depth of the scene objects can be extracted. The correspondence algorithm has been applied to extract disparity map. To find the depth (z) for a stereo vision system by using the equation:

$$z = \frac{f * b}{x_L - x_R} = \frac{f * 16}{x_L - x_R} \quad (12)$$

where (b) baseline equals 16 cm,(f)focal length, $(x_L - x_R)$ is the disparity map, z is the depth.

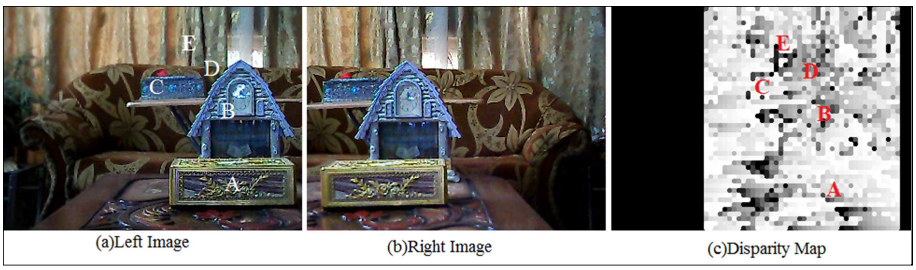

Figure 12. *Stereo images captured by the practical stereo cameras and the disparity map by selecting points at different depths.*

Table 3. *Distance measurements of the five points of Figure (12).*

Tested points	disparity value (d) in (pixels)	Real depth (cm)	Estimated depth (cm)	Depth Error percentage (%)
A	77	121	121.887	+ 0.72 %
B	59	160	159.02	- 0.6 %
C	40	236	234.59	- 0.59 %
D	31	297	302.7	+ 1.9 %
E	29	313	323.58	+ 3.38%

The estimated depths of randomly selected points in the real scene are calculated according equation (12), in which the focal length (f) that represents one of the internal parameters of the camera is extracted from camera calibration as mentioned before, where the horizontal focal length in pixels of the left camera is α=586.491 pixels.

It can be noted that the error in depth results at points A, B and C is less than that at points D and E because the disparity values obtained at points A, B and C are correct.

9. Conclusions

Stereo vision system is a very important technique and widely used in real time performance devices and the matching algorithms that are used to calculate the output suffer from the execution time and output accuracy image (Error percentage). For CBMA, some things have been done to show the results which are suitable for a real time performance to reduce the execution time of the algorithm in three major steps:

a) The Canny edge detector filter has a good effect on reducing the execution time for matching algorithm.
b) The block matching technique with the SAD ((Block-by-Block) search which provides less operation to compute the SAD when it runs in the matching algorithm).
c) The selection of the optimal window size effects on both the execution time and error percentage.

For a better accuracy, the disparity map is used to reduce the error percentage to enhance the output image using the interpolation algorithm, with median filter and the interpolation techniques, the bilinear interpolation techniques have been used to enhance the output results images from the three algorithms. The error percentage has been reduced about 2% and the execution time is reduced. Therefore, the total execution time will be divided by the step size windows (block size).

References

[1] Iocchi L. and Konolige K., "A Multiresolution Stereo Vision System for Mobile Robots", Proc. of Al Workshop on New Trends in Robotics, 1998.

[2] G. Macesanu, S. Grigorescu, T.T. Cocias, and F. Mol doveanu, "An Object Detection and 3D Reconstruction Approach For Real-Time Scene Understanding" Bulletin of the Transylvania University of Brasov, Series I: Engineering Sciences • Vol. 4 (53) No. 1. 2011.

[3] Marr D., and Poggio T., "Cooperative Computation of Stereo Disparity", Science, New Series, Vol. 194, No. 4262, pp. 283-287, Oct. 15, 1976.

[4] Marr D., and Poggio T., "A Computational Theory of Human Stereo Vision", proceedings of the Royal Society of London. Series B, Biological Sciences, Vol. 20, No. 1156, pp. 301-328, May 23, 1979.

[5] Eric W., and Grimson L., "Computational Experiments with a Feature Based Stereo Algorithm", Massachusetts Institute of Technology, Jan., 1984.

[6] Hakkarainen J., and Lee, "A 40×40 CCD/CMOS Absolute-Value-of-Difference Processor for Use in a Stereo Vision System", IEEE, Journal of Solid-State Circuits, Vol. 28, No. 7, pp. 799-807, July 1993.

[7] Labayrad R. and Aubert D. "Robust and Fast Stereovision Based Road Obstacles Detection for Driving Safety Assistance", IAPR Workshop on Machine Vision Application, pp. 624-627, Japan, Dec., 2002.

[8] Vatansever M., "3D Reconstruction Using a Spherical Spiral Scan Camera", M.Sc. Thesis, Computer Engineering, Izmir Institute of Technology, Izmir, 2006.

[9] Fengjun HU and Yanwei Zhao "Comparative Research of Matching Algorithms for Stereo Vision", Journal of Computational Information Systems 9: 13, PP 5457–5465, 2013.

[10] Raad H., Zaid K. "Stereo Vision Distance Estimation Employing SAD with Canny Edge Detector", International Journal of Computer Applications (0975 – 8887) Volume 107. No. 3, PP 38-43. December 2014.

[11] T. Acharya & P.S. Tsai "Computational Foundations of Image Interpolation Algorithms", ACM Ubiquity Vol. 8, 2007.

[12] Parth Bhatt, Sanjay Patel "Image Enhancement Using Various Interpolation Methods", International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555. Vol. 2, No. 4, PP 799-803. August 2012.

[13] Bax Smith, "Design Assignment Matlab's Implementation of Bilinear Interpolation". 2002.

[14] Vaishali Patel, Kinjal Mistree, "A Review on Different Image Interpolation Techniques for Image Enhancement", International Journal of Emerging Technology Volume 3, Issue 12, PP 129-133. December 2013.

[15] Gordon K. Smyth" Polynomial Approximation", John Wiley & Sons, Ltd, Chichester, 1998.

[16] http://en.wikipedia.org/wiki/Polynomial

[17] http://mathworld.wolfram.com/Polynomial.html

[18] http://en.wikipedia.org/wiki/Quadratic_function

[19] Scharstein D., "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms", International Journal of Computer Vision 47(1/2/3), USA: Kluwer Academic Publishers, pp. 7-42, 2002.

[20] http://en.wikipedia.org/wiki/Mean_squared_error

[21] http://en.wikipedia.org/wiki/Median_absolute_deviation

[22] Myron Z. Brown, "Advances in Computational Stereo", IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 25, NO. 8, August, 2003.

[23] http://fourier.eng.hmc.edu/e161/lectures/smooth_sharpen/node2.html

[24] http://vision.middlebury.edu/stereo/

[25] http://homepages.inf.ed.ac.uk/rbf/HIPR2/median.htm