Depth Image Extraction using Adaptive K-Means Clustering Algorithm

Assist. Prof. Dr. Hasanen S. Abdullah

University of Technology
Computer Science Department

Sanaa Sh. Ahmed

AL-Ma'moun University College
Computer Science Department

Abstract

This work presents a proposed approach to extract depth map of stereoscopic images depended on segmentation of   lightness values of pixels 'L' using adaptive K-Means cluster. The proposed approach finds the disparity map of segmentation lightness pixels and refines those segment by using morphological filtering and connected components analysis. Experimental results from Middlebury dataset show that the proposed approach performs good results in term of accurate depth and time consuming compared with classical Sum of Absolute Difference (SAD) approach and SAD with Gradient Difference (GRAD) algorithm.

Keywords-Stereoscopic Images, SAD, GRAD, Depth Map, Disparity Map.

استخلاص عمق الصورة باستخدام خوارزمية
Adaptive K-Means

أ.م.د. حسنين سمير عبدالله   سناء شهاب احمد

الجامعة التكنولوجية /قسم علوم الحاسوب   كلية المأمون الجامعة

المستخلص:-

إن هذا العمل يمثل خوارزمية مقترحة الى استخلاص عمق الصورة بالنسبة للصور Stereoscopic اعتمادا على تقطيع النقاط المضيئة باستخدام خوارزمية -Adaptive K Means . حيث تم ايجاد خريطة التباين لتقطيع النقاط المضيئة وتحسين هذه القطع باستخدام morphological filter وتحليل المكونات المترابطة .

حيث اثبتت النتائج باستخدام Middlebury Dataset ان الخوارزمية المستخدمة تعطي نتائج جيدة بمصطلح العمق الدقيق للصورة والوقت المستغرق للتنفيذ مقارنة مع طريقة SAD التقليدية وطريقة SAD with GRAD algorithm .

1. Introduction

A 3D world contain information about the third dimension. This dimension is imaginable through human vision as type of "binocular disparity". Human eyes are positioned a short distance and through them can take pictures from different views in the real world. The human brain can build depth information from these various views of the real world.

Presentation of three-dimensions takes benefit of this event by creating two slightly diverse images of each scene as well as then presen them to the human eyes. Via an proper "disparity" and "calibration" of parameters, an acceptable three dimensional view may be recognized.

An essential step in every 3-D structure is the 3-D contented creation. A number of particular cameras have been layout to create 3-D forms straightly. For instance, a stereoscopic two-cameras that makes utilize of a co-planar pattern of two independent, monoscopic cameras, every capturing one eye's

vision, and "depth information" is calculated by used binocular disparity [1].

"Stereoscopy", as well recall stereoscopic or 3-D imaging is a method for generation or enhancing the imagination of depth in an image via means of stereo psis for binocular vision. General stereoscopic techniques show set of pair images individually to the left and right eye of the observer. These 2-D images are to be then collective in the mind to offer the observation of 3-D depth. This method is recognized from 3-D views that show an image in three complete dimensions, allow the viewer to raise information about the 3-D items being showed via head and eye actions [2].The deepness is gained through calculating stereo matching among two images. This method is known as "triangulation", which performs the method of detecting coordinates of a 3-D point depend on its matching stereo image points.

There are several troubles for evaluating 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 must be precise enough to be distinguished [3].

Stereo vision system has two cameras located side by side so as to get left as well as right images. The cameras are isolated via a predefined distance (T), known as baseline, and should have the same focal length (f). For every pixel in the left image, an equivalent pixel is matched in the right image. The geometry of the camera could drastically affecting the amount of processing that demands matching strategy. It is assumed that the parallel camera configured, for this reason .The cameras are positioned on a horizontal plane and have parallel optical axes [4]. Figure(1) shows stereo vision system where O,O' are two cameras, f is the focal length, x,x' are projection point at image plane, X is the world point ,Z is the depth.
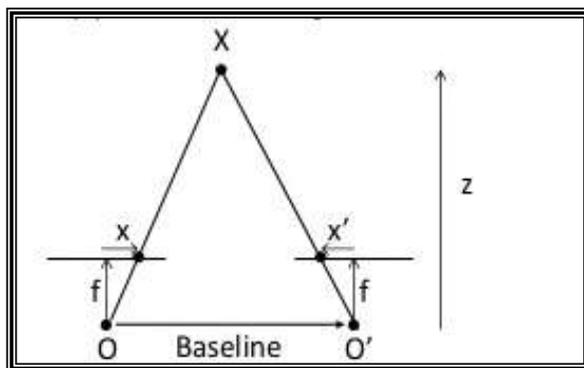
Figure (1) Stereo Vision System

Eyes of human are critical to change in brightness more than in color. For this reason left and right images which are input to algorithm convert from "RGB" to the "Lab color space" and which retains only lightness 'L' values of its pixels [5].

  In this work, a method was suggested to calculate the disparity map which aims to find depth map of stereoscopic image by using adaptive K-means cluster a lgorithm .

This algorithm is fast for computation compared with the K-means algorithm .This algorithm is instead of determined number of clusters which are initially like K-means algorithm .It randomly selects a number of clusters (k).

2. Related Work

  One of the most popular categorization of existing stereo depth estimation algorithms is depended on "global" vs. "local" methods. Local algorithms are numerical techniques which are typically depended on correspondence. The correspondence method includes finding the corresponding pixels in left and right images of a stereo pair through assembling costs [for example, "sum of absolute differences (SAD)", "sum of squared differences (SSD)"," normalized

cross correlation (NCC)"] within a region or block. In contrast, Global algorithms are depended on clear smoothness hypothesis that are resolve throughout diverse optimization methods.[6]

In 2012, Mahmoud Abdelhamid, Jeffery Beers and Mohammed Omar [7], proposed an approach to calculate the depth using a correlation matching routine which programmed by using a SSD algorithm to search for the corresponding points from the left and the right images. The suggested modifications were in two folds; reducing the search domain by establishing a ROI for the value-added data and through a rectification step that is done to reduce the 2D search domain into 1D .This means the search is done only on horizontal lines. The second fold relied on the use of detected objects' centroids rather than their pixels' intensities. The developed code results in better accuracy and repeatability, without any false positives at faster processing times.

In 2014,R.H Thaher and Z.K.Hussein [3],proposed method (CBMA)" to get the Disparity map.The algorithm contained dual parts the Canny edge detector as well as Block matching technique within SAD to conclude disparity map to decrease the implementation time. The determination of window size doesn't merely manipulate the algorithm's implementation time, however as well as the precision of disparity map. The outcome illustrate superior results with less error in comparison to depth of the real objects, which are superior outcome of disparity (d) values calculated by the CBMA.

In 2016,Juber A Sheikh and M.A.Joshi [8], proposed an approach to extraction of disparity map (Depth information) in duration of milliseconds with the use of two stereo images of that certain scene. The duration take for calculate of the map of disparity is from 25 and 30 sec utilizing combined SAD

and GRAD algorithm. For implementing system in real time, the duration of execution should be as small as possible. This algorithm for real time stereo match and depth extracting is introduced to reduce the consumption of processor cycle. In addition to minimizing the time of execution, this approach also minimizes the possibilities of false matches and therefore results avoid the disadvantages of incorrect matches with the use of bidirectional matching.

3. Proposed Depth Image Extraction Approach

This carrent proposed approach gets the depth map depends on the segmentation of the lightness pixels .Left and right images which are used in this work must be converted from RGB color space to gray scale color space and perform the stereo matching algorithm on intensity of lightness pixels only. Also, selected window size as 9*9 pixels and disparity is (0 - maximum disparity).

Lightness 'L' values of the left image pixels are segmented by the usage of adaptive k-Means clustering algorithms. The runtime of clustering must be is reduced.

Segment boundary detection achieved by comparing the cluster assignment of each pixels with that of its 8-connected component pixels. If any of these pixels is found various , the pixels are to be marked as'1'(which is on segment boundary), otherwise is '0'(which is not on a segment boundary). This step generates the boundary map later than segmentation.

Later, this step will execute a morphological operations to enhance the boundary map due to many of pixels included to the segment boundary due to boundaries of clusters which concern by lightness 'L' values only. Morphological operations using to remove like noisy pixels ,as follows:

Fill: isolated interior pixels are filled like center pixels in

                        1 1 1
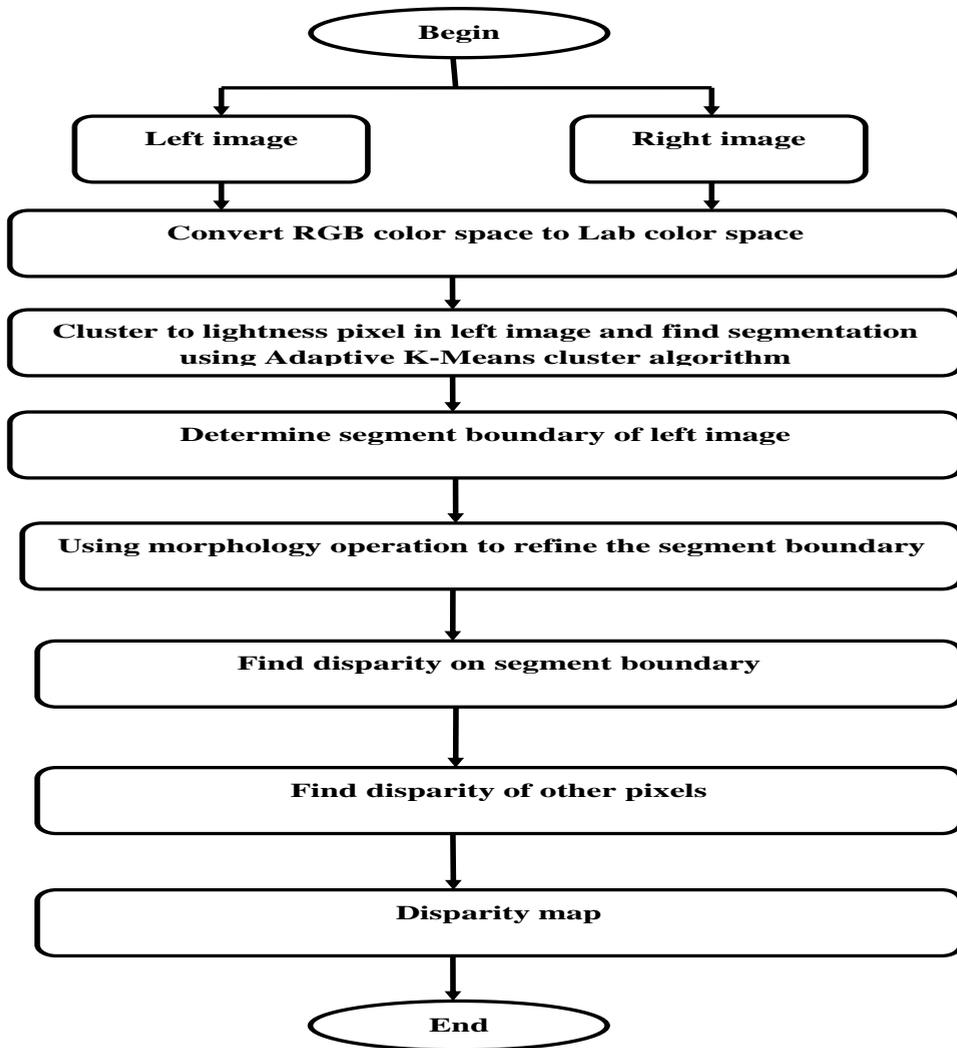   1 0 1
             1 1 1

Remove: center pixel is removed. For example ,set center pixel to '0' if all of its four-connected component are '1', therefore leave-taking just boundary pixels on. In the next step the disparity map of boundary pixels was found by using SAD cost function to matching lightness 'L' values of left as well as right image pixels.

Disparity map of the remaining pixels in each row is calculated by depended on disparities that have already been calculated. This disparity founded in two steps:
Step1: Fill the intermediary pixels in the row with the disparity value of two boundary pixels which have same disparity values. Pair of points which are calculated the disparity map it consideration belong to the same object in reference image, therefore all intermediary pixels belong to that same object.

Step2: Peeking at disparity values of neighboring pixels to estimate the disparity of pixels which have not yet been determined.

Figure(2) illustrates the flowchart of proposed approach

```
                          ┌──────────────┐
                          │    Begin     │
                          └──────────────┘
              ┌───────────────────┐   ┌───────────────────┐
              │    Left image     │   │    Right image    │
              └───────────────────┘   └───────────────────┘
          ┌──────────────────────────────────────────────────┐
          │   Convert RGB color space to Lab color space      │
          └──────────────────────────────────────────────────┘
          ┌──────────────────────────────────────────────────┐
          │ Cluster to lightness pixel in left image and find │
          │ segmentation using Adaptive K-Means cluster       │
          │ algorithm                                          │
          └──────────────────────────────────────────────────┘
          ┌──────────────────────────────────────────────────┐
          │     Determine segment boundary of left image      │
          └──────────────────────────────────────────────────┘
          ┌──────────────────────────────────────────────────┐
          │ Using morphology operation to refine the segment  │
          │ boundary                                           │
          └──────────────────────────────────────────────────┘
          ┌──────────────────────────────────────────────────┐
          │       Find disparity on segment boundary          │
          └──────────────────────────────────────────────────┘
          ┌──────────────────────────────────────────────────┐
          │          Find disparity of other pixels           │
          └──────────────────────────────────────────────────┘
          ┌──────────────────────────────────────────────────┐
          │                 Disparity map                     │
          └──────────────────────────────────────────────────┘
                          ┌──────────────┐
                          │     End      │
                          └──────────────┘
```

The execution steps of the approach on pair images it was received from Middlebury stereo vision dataset, has been selected two image pairs (left and right) of (Tsukuba, Venus) images and run the method on these images .The results compare   the two classical approach:approach1 (SAD)and

approach2(SAD with GRAD algorithm) as shown in tables (1 and 2).

To estimate the value of error of disparity map, will use the method of computing error percentage of bad matching pixels (B) of disparity map with deference to ground truth image which obtainable by the Middlebury dataset and as given by [9].

$$B = \frac{1}{N} \sum_{(x,y)} (|d_C(x, y) - d_T(x, y)| > \delta_d)$$

Where $d_T(x, y)$ are ground truth disparities $and\ d_C(x, y)$ are computed disparities , and δ_d denotes to disparity error tolerance, which is taken as (95) in this work according to the scale of taken images. Also N represent window size which is taken as 9*9 pixels in this method which givings best results .

In figure(3), left image shows the disparity map for Tuskuba image which was obtained from a proposed approach, while right image shows ground truth for this image from Middlebury dataset.

Figure (3) Comparison between Depth Map Of Tsukuba Image To Middlebury Ground Truth.

In figure(4), left image shows the disparity map for Venus image which was obtained from a proposed approach, while right image exposure ground truth for this image from Middlebury dataset.
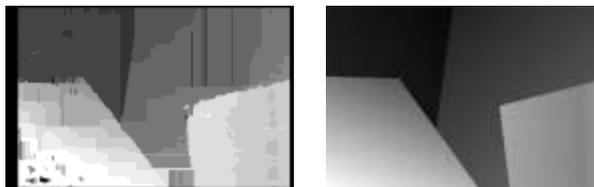

Figure (4) comparison between depth map of Venus image to Middlebury ground truth.

In table(1) and table(2)the obtained results from the proposed method are compared   with other methods used to extract depth map for Tsukuba and Venus image.

 Table (1) Performance Comparison of Proposed Approach for Tsukuba Image

	MAE	Time
Approach1 (SAD) [10]	6.33%	6.62 Sec
Approach2 (SAD with GRAD alg.)[8]	5.59%	21.16 Sec
Proposed Approach	4.92%	1.96 Sec

Table(1) shows MAE and time consuming for execution a proposed approach   for Tsukuba image. So that, MAE proposed method was 4.92% and it is better than MAE for SAD method and SAD with GRAD algorithm. In addition , run time of proposed method was 1.96 sec, and it less than two other method which maintained in this table.

Table (2) Performance comparison of proposed approach for Venus image

|	MAE	Time
Approach1 (SAD) [10]	8.34%	9.94 Sec
Approach2 (SAD with GRAD alg.)[8]	7.88%	29.95 Sec
Proposed Approach	6.70%	2.26 Sec

Table(2) show MAE and time consuming for execution a proposed approach for Venus image. So that , MAE proposed method was 6.70% and it better than MAE for SAD method and SAD with GRAD algorithm. In addition , run time of proposed method was 2.26 sec, and it is less than two other methods which are maintained in this table.

Figure(5) showing the error percentage using MAE of Tsukuba, Venus image for a proposed approach compare to SAD method and SAD & GRAD algorithm. x-axis represent used approach and y-axis represent MAE.
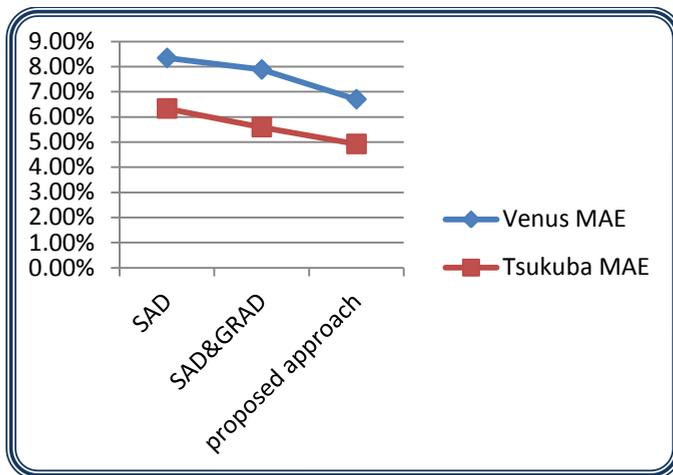
Figure(5) the error percentage using MAE of Tsukuba, Venus images

Figure(6) showing the time consuming of Tsukuba, Venus image for a proposed approach compare to SAD method and SAD & GRAD algorithm. x-axis represent used approach and y-axis represent time consuming
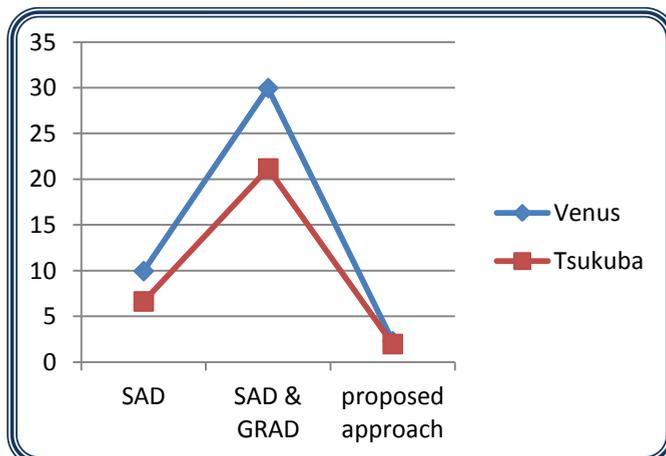
Figure(6) The time Consuming Of Tsukuba, Venus Images

5. Discussion of Results

This work extracts depth map in stereoscopic images by using two images pair (left and right) from Middlebury dataset (Tsukuba,Venus).Firstly, it compared the disparity map of proposed approach with ground truth of Middlebury dataset by using bad matching pixel measure as shown in figure(3) and figure(4). Secondly, the methad compard the results of proposed approach with two existing approaches(SAD, SAD and GRAD algorithm) according to MAE and time consuming. Where so it was observed in proposed approach that the MAE for Tsukuba and Venus images is better than SAD approach and SAD with GRAD algorithm, as shown in figure(5) .Also ,time taken to execute proposed approach is reduced than the existing approaches, as shown in figure(6).

This proposed approach used window size(block) as 9*9 pixels and maximum disparity between (0-15), and we did not take advantage of all pixels in image, but took segmentation of lightens "L" values of pixels which are reduced by processing time which means that the disparity map is more accurate than other approaches.

6. Conclusion

This paper presents a proposed approach for extraction depth map of stereo images which depend on value of lightness pixels 'L' values by using segmentation of these pixels by adaptive K-means cluster algorithm.
Adaptive K-means algorithm is considered faster algorithm because it determines randomly number of clusters .For this reason , this proposed approach executed in less time with fast processing .

As was demonstrated the value of MAE to execute this approach is reduced to 2% compared to SAD approach and 1% compare to SAD with GRAD algorithm. This means that the error percentage of this proposed approach compared to the

other two methods was better than ath. So that, run time for this approach was less than time of two other approach.
According to this result, this proposed approach given best results with accurate depth and reduces the time taken to execution method.

7.Reference

[1]   Q. Wei, "Converting 2D to 3D: A survey," International Conference *Page(s)*, vol. 7, p. 14, 2005.

[2]   J. M. Dominic, "Recent Trends in 2D to 3D Conversion: A Survey," *www.ijraset.com Issue IV*, vol. 2, no. Iv, pp. 388–395, 2014.

[3]   R. H. Thaher and Z. K. Hussein, "Stereo Vision Distance Estimation Employing SAD with Canny Edge Detector," vol. 107, no. 3, pp. 38–43, 2014.

[4]   W. Miled and J. Pesquet, "Disparity Map Estimation Using A Total Variation Bound," *3rd Can. Conf. Comput. Robot Vis.*, pp. 48–48, 2006.

[5]   P. Ganesan, V. Rajini, B. S. Sathish, and K. B. Shaik, "CIELAB Color Space based High Resolution Satellite Image Segmentation using Modified Fuzzy C-Means Clustering," *MAGNT Res. Rep.*, vol. 2, no. 6, pp. 199–210, 2014.

[6]   S. Mattoccia, "Stereo Vision Algorithms for FPGAs," *IEEE Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 636–641, 2013.

[7] Mahmoud Abdelhamid, Jeffery Beers and Mohammed Omar, "Extracting Depth   Information Using a Correlation Matching Algorithm", Journal of   Software Engineering and Applications, vol 5, no.5, pp.304, 2012.

[8]   Juber A Sheikh and M. A. Joshi, "Stereo Images for Depth Measurement," IEEE International Conference On Recent   Trends   In   Electronics   Information

Communication Technology, pp. 1267–1271, 2016.

[9]   Scharstein, D. and   Szeliski, R.," A taxonomy and evaluation of dense two-frame
stereo correspondence algorithms". International journal of computer vision,47(1-3),
 pp.7-42, 2002.

[10] C. S. Panchal and A. B. Upadhyay, "Depth Estimation Analysis Using Sum of Absolute Difference Algorithm," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 3, no. 1, pp.6761–6767,2014.