A Literature Survey on Removal of Ambiguities in Stereo Images

[1]Simi Thakur, [2]Gianetan S. Sekhon
[1,2]Dept. of Computer Engineering, YCOE, GuruKashi Campus, Punjabi University
Talwandi Sabo, Punjab, India

Abstract
This paper presents the study of ambiguities presents in Stereo Images. Stereo vision (stereo means "Solid" or "three dimensional" and vision means "appearance" or "sight") is the impression of depth that is perceived when a scene is viewed with both eyes. There are several ambiguities are there, when any algorithm matches the two images present in image set, these ambiguities are related to either noise, speed, efficiency or reliability. The study of various techniques of stereo matching is considered in this work. To resolve the ambiguity problem in stereo matching many methods also have been studied in this work. Define different ambiguities present in stereo images and comparative analysis of different algorithms is the major contribution of this survey paper.

Keywords
Stereo Image Set, Filter, Disparity Map

I. Introduction

Artificial intelligence is defined as "the art creating machine that perform functions that require intelligence when performed by people."(Kurzweil, 1990) Or in other words we can say that "it is a field of study that seeks to explain and emulate intelligent behavior in terms of computational processes."(Schalkoff, 1990) [12]. Pattern recognition and intelligent system is the major part of the artificial intelligence. Pattern recognition is concerned with estimating density function in a high-dimensional space and dividing the space into regions of categories and classes. Computer vision is the application of Pattern recognition; it is the combination of concepts, technique and ideas from: digital image processing, pattern recognitions, artificial intelligence, and computer graphics.

Now a day's stereo matching is one of the most active areas in computer vision because of the accurate acquiring of the image depth that is related to many applications such as robotics, autonomous system, human computer interaction and scene reconstruction etc. Although many stereo matching algorithms have been proposed, especially for generating disparity maps or depth maps, stereovision is the one of methods that can yield depth information of the scene. It uses stereo image pairs from two cameras to produce disparity maps that can be easily turn into depth maps. But still many challenging works needed to be done which caused by textureless, noise, occlusions and discontinuity. Stereo matching calculates disparity or depth with image set using pixel correspondence. Local and global are the categories in which generally any stereo matching method is divided.

There are many ambiguities problems are present while performing the stereo matching, for this purpose many methods have been proposed from decades. Early approaches use local and window-based methods, and employed a local "winner-takes-all" (WTA) strategy in depth estimation at each pixel. Later on, several global methods are proposed, which formulate the depth estimation as an energy-minimization problem. The local matching algorithm has some advantages such as low computation cost and high efficiency. However, its high error rate and low resolution makes it difficult to meet the requirement of practical applications. In the local algorithm, the disparity at a given pixel depends only on color values within a window [8]. Area based stereo matching approaches are applied that solves correspondence problem on the basis of area correlation. But this approach is completely local, to overcome this feature based approach is applied that is better and more accurate than the area based stereo matching but it can produce global matching results and thus efficiently avoid searching blindly in a wide range and computational complexity is another disadvantage for this method. Then the deficiencies of area and feature based methods are overcome by Hierarchical hybrid method. It produces better disparity maps on the edge of the scenes as well as on the occluded areas, and generates robust and dense disparity fields. The hybrid matching reduces the dependency on the quantity and quality of features so that it obtains still good results in sparse feature fields. It is computational efficient and the complexity of the HHM algorithm is acceptable [1]. The graph cut based algorithm which combines window-based local matching into a global optimization framework, that it preserves discontinuities both during the local as well as global matching phase [2]. The region based progressive algorithm is the one in which previous region-based and progressive approaches are combined together and this algorithm is among the state-of-the-art both in accuracy and efficiency.

Recently, segment-based methods have attracted attention due to their good performance. They are based on the assumption that the scene structure can be approximated by a set of non-overlapping planes in the disparity space and that each plane is coincident with at least one homogeneous color segment in the reference image. The Distinctive Similarity Measure (DSM) is essentially based on the distinctiveness of image points and the dissimilarity between them, which are both closely related to the local appearances of image points; the distinctiveness of an image point is related to the probability of a mismatch while the dissimilarity is related to the probability of a good match [6] and it also verifies that this algorithm is very efficient under point ambiguity. The algorithm disparity distribution models (DDMs) is also based on segment-to-plane. The key difference from existing segment-based methods is that no global optimal is used [9]. In this a correct disparity plane assign to a segment directly. A region adaptive local stereo matching algorithm is the one which focuses on two methods that are census transform and a fast segmentation based method that is for depth discontinuity and other low texture area respectively.

In this paper we are not exactly concern with converting 2-D object into 3-D, we can only compare the different algorithms that obtaining the exact depth of the object. Removal of ambiguities, it can be of any type either related to hardware or software and comparative analysis with different algorithms are key issues.

II. Related Work

Guoping li et al. [1] presented algorithm based on the combination of feature and area-based matching, called Hierarchical Hybrid Matching Algorithm (HHM). It provides the efficient disparity

estimation and it also overcome the deficiencies of the area and feature based algorithms. To simplify the hierarchical framework, the HHM is implemented in two different resolution layers, which are: Segmentation and feature extraction, mapping line segments into blocks, Feature matching and global block matching, local block matching.

Motilal Agrawal et al. [2] presented an algorithm which combines window-based local matching into a global optimization framework. The local matching algorithm assumes that local windows can have at most two disparities. Under this assumption, the local matching can be performed very efficiently using graph cuts. The global matching is formulated as minimization of an energy term that takes into account the matching constraints induced by the local stereo algorithm. Fast, approximate minimization of this energy is achieved through graph cuts.

Yichen WEI et al. [3] presented region based progressive stereo matching algorithm in which reliable regions are firstly identified and matched using GCPs. Remaining regions are matched progressively in a growing-like process using a global best first strategy based on a cost function that integrates disparity smoothness and visibility constraints and an ambiguity measure that is defined to be the ratio of the best and second best costs. Generally, matched regions propagate from textured areas to homogeneous and occluded areas. The progressive matching process stops if dense disparity map is obtained or a prespecified reliability threshold is reached.

This algorithm runs as-

1. Arrange all UNMATCHED regions in descending order of confidence. Set A_{amb} empty.
2. For each region R with confidence(R) $\neq 0$
(a) If ambiguity(R) $\leq \lambda$amb, label R as
MATCHED and assign it the disparity d^{best}_{R}. Update disparity functions D, D1 accordingly and repeat Step 2.
(b) If R contains GCPs with more than one disparity,
split R, label new regions as UNMATCHED and repeat Step 2.
(c) If there is an entry in A_{amb} empty or with a value larger than ambiguity(R), set the value of the entry to ambiguity(R).
3. If there are new regions labeled as MATCHED, repeat Step 2.
4. If there are UNMATCHED regions, set λamb to the maximum values in Aamb, go to Step 1; or exit when a pre-defined condition is met in case that dense matching is not desired.

Andreas Klaus et al. [5] presented an algorithm that utilizes color segmentation on the reference image and a self-adapting matching score that maximizes the number of reliable correspondences. The scene structure is modeled by a set of planar surface patches which are estimated using a new technique that is more robust to outliers. Instead of assigning a disparity value to each pixel, a disparity plane is assigned to each segment. The optimal disparity plane labeling is approximated by applying belief propagation. The resulting dissimilarity measure is given by:

$C(x, y, d) = (1-\omega) * CSAD(x, y, d) + \omega * CGRAD(x, y, d)$

$C_{SAD}(x, y, d) = \sum_{(i,j)\in N(x,y)} I1(i, j) - I_2(i + d, j)$

And

$C_{GRAD}(x,y,d) = \sum_{(i,j)\in Nx(x,y)} |\nabla xI1(i,j) - \nabla xI2(i+d,j)| + \sum_{(i,j)\in Ny(x,y)} |\nabla yI1(i,j) - \nabla yI2i + d, j)|,$

Where N(x, y) is a 3 × 3 surrounding window at position (x, y), Nx(x, y) a surrounding window without the rightmost column, NY(x, y) a surrounding window without the lowest row, ∇x the forward gradient to the right and ∇y the forward gradient to the bottom. Color images are taken into account by summing up the dissimilarity measures for all channels.

Kuk-Jin Yoon et al. [6] presented a method named Distinctive Similarity Measure (DSM) is essentially based on the distinctiveness of image points and the dissimilarity between them, which are both closely related to the local appearances of image points; the distinctiveness of an image point is related to the probability of a mismatch while the dissimilarity is related to the probability of a good match. BarısBaykant [7] presented an algorithm named region based stereo matching under this there are two techniques that are Global Error Energy Minimization by Smoothing Functions and Line Growing Based Stereo Matching, both are together used for extraction of depth information for generating 3D image by eliminating noise with the help of averaging filter to generate reliable disparity maps. Error energy is expressed as-

$$e(i,j,d) = \frac{1}{3.n.m} \cdot \sum_{x=i}^{i+n} \sum_{y=m}^{j+m} \sum_{k=1}^{3} (L(x, y+d, k) - R(x, y, k))2 \qquad [7]$$

Where e (i, j, d) is error energy, n×m window size of block matching, L (i, j, c) is left image in RGB format and R (i, j, c) is right image in RGB format

Zhihua Liu et al. [8] proposed a progressive region merging algorithm for stereo matching. Firstly, the reference image is segmented by mean shift algorithm and the initial disparity maps are calculated using loopy belief propagation method. Secondly, we derive the reliable pixels through left-right consistent constraint. In addition, we present a new approach to exclude mismatch pixels in textureless regions, which will also reduce errors in the following region merging process. Finally, a progressive region-merging algorithm based on hierarchical clustering algorithm is put forward to merge neighboring regions and refine the plane parameters. Angle and distance are used together to measure whether the two planes are the same or not, which will reduce the calculation complexity.

Lin Chen et al. [10] presented a Region- Adaptive stereo matching algorithm which focuses on different cost aggregation strategies for different regions that can apply by census transform method along depth discontinuities and a fast segmentation-based method for other low texture areas.

Total aggregation cost function is defined as:

$$C_{TOT}(p; q; d) = \omega * \frac{Cseg(p,q,d)}{n(Bp)+n(Bq)} + (1-\omega) * \frac{Cw(p,q,d)}{I^2}$$

It contains two parts: segmentation part and window part. In segmentation part, we adopt both the block images of reference and target image. CSEG is defined as:

$C_{SEG}(p; q; d) = C_{SL}(p; q; d) + C_{SR}(p; q; d)$

Where

$C_{SL}(p; q; d) = \sum_{pi\in Bp} \min(\delta(pi; qi; d); Tr)$

And

$C_{SR}(p; q; d) = \sum_{qi\in Bq} \min(\delta(pi; d; qi); Tr)$

Bp is the segment block of p and Bq is the segment block of q.

Wei-qun Li et al. [9] presents first, several disparity distribution models on segments under the piecewise smoothness assumption are investigated for high quality dense disparity maps without global optimization; second, a novel plane fitting strategy based on the investigation is developed for accurately estimating disparity planes. The processing steps are as follows:

1. Choose GCPs (Ground Control Points) which have the smallest

disparity value of the CDI (Continuous disparity interval) to fit the disparity plane and denote the plane as P.
2. A matching cost for P is calculated by

$$C(S, P) = \frac{1}{Ns} \sum_{(x,y) \in S} C(x, y, d)$$

Where d is the disparity value defined by P at the pixel (x, y) in the segment S, Ns is the number of the pixels in S.
3. Add GCPs corresponding to the next disparity value to the old ones, fit the disparity plane once more and calculate the matching cost of the new plane.
4. If the GCPs that have the largest disparity value of the CDI are added, go to step 5. Otherwise, repeat step 3.
5. Choose the plane that has the minimum matching cost as the disparity plane of the segment.

Zi-wei Zhou et al. [11] presents another algorithm based on graph cuts stereo matching but the difference is the previous methods are not used for the real time application, but this can increase the accelerate the matching speed. Firstly, the segments of the image are calculated using the Gaussians pyramid (Dogs) computing method, secondly calculate the disparity range of two corresponding image segments in image pair using the color feature, and every segment pair is matched based on graph cuts algorithm according to the offset of disparity range in parallel time, finally all the segments are assembled into a whole disparity image.

III. Comparative Analysis

[1]	Guoping Li, Yun He	2002	A HIERARCHICAL COMBINED FEATURE- AND AREA- BASED STEREO MATCHING ALGORITHM	Hierarchical hybrid Matching algorithm (HHM).	Precise disparity estimates on the edge and occluded area. Good results in sparse feature fields.	Computational cost
[2]	Motilal Agrawal, Larry S. Davis	2004	WINDOW-BASED, DISCONTINUITY PRESERVING STEREO	Graph cut based algorithm for combining window-based local stereo into a global optimization framework	Preserves discontinuities both during the local as well as global matching phase.	Computational speedup, Selection of window size, Incorporating occlusion & uniqueness constraint.
Ref No.	Author(s)	Year	Title	Technique	Results	Limitation(s)
[3]	Yichen WEI, Long QUAN	2004	REGION-BASED PROGRESSIVE STEREO MATCHING	Region-Based Progressive algorithm	Low computational complexity, capability of occlusion handling around region boundaries and disparity regularization inside regions.	Early wrong decisions, not obtained good results for much textured maps.
[4]	Li Hong, George Chen	2004	SEGMENT-BASED STEREO MATCHING USING GRAPH CUTS	Segment-based stereo matching algorithm using graph cuts	Good results in difficult areas such as textureless regions, disparity Discontinuous boundaries and occluded portions.	Not be able to handle the situation if there are disparity boundaries appearing inside the initial color segments
[5]	Andreas Klaus, Mario Sormann, KonradKarner	2006	SEGMENT-BASED STEREO MATCHING USING BELIEF PROPAGATION AND A SELF-ADAPTING DISSIMILARITY MEASURE	Segment based matching algorithm using BP and Self-Adapting	Give good results under fixed parameters and less time consuming.	Over segmentation is performed

[6]	Kuk-Jin Yoon, In So Kweon	2007	STEREO MATCHING WITH THE DISTINCTIVE SIMILARITY MEASURE	Distinctive Similarity Measure(DSM)	Better results than local methods, very efficient for the reliable local feature selection and matching and resolve the point ambiguity problem in stereo matching, focusing on the distinctiveness.	Do not give satisfied results under ideal similarity measures model
[7]	BarısBaykant ALAGÖZ	2008	OBTAINING DEPTH MAPS FROM COLOR IMAGES BY REGION BASEDSTEREO MATCHING ALGORITHMS	Region based stereo matching using Global error energy minimization and point matching.	Increase reliability of disparity maps by using filters.	Time consuming, not considered occluded area and better filter can apply to make it more robust.
[8]	Zhihua Liu, Qixiang Ye, Lu Ke, Jianbin Jiao	2009	A PROGRESSIVE REGION-MERGING ALGORITHM FOR STEREO MATCHING	Progressive region-merging algorithm	Reliable merging with its neighboring regions, which will reduce error rate and the calculation complexity.	Depth Discontinuities only occur on region boundaries, disparity is reserved.
RefNo.	Author(s)	Year	Title	Technique	Results	Limitation(s)
[9]	Wei-qun Li, Xian-ming Chen	2010	A FAST STEREO MATCHING USING IMAGE SEGMENTATION FOR HIGH QUALITY DENSEDISPARITY MAPS	Disparity distribution models (DDMs)	Performs robustly both in textureless Segments and trivial segments and the algorithm's efficiency and accuracy make it easily used in a real-time system.	If the segment is not occluded seriously, the result might be wrong.
[10]	Lin Chen, MinGang Chen, Xiao Lin, LiZhuang Ma	2012	AN EFFICIENT REGION-ADAPTIVE STEREO MATCHING ALGORITHM	Region adaptive local stereo matching algorithm	Different kinds of methods into different types of regions, fast speed performance along with better compared to other real time methods.	Segmentation is not properly defined and mean shift algorithm is time consuming And do not have high accuracy
[11]	Zi-weiZhou,Ge Li, Ji-zhuang Fan and Jie Zhao, XinYuOyang	2012	A NEW STEREO MATCHING ALGORITHM BASED ON IMAGE SEGMENTATION	Gaussians pyramid (DoGs) computing method along with graph cuts algorithm.	High matching speed with high disparity image quality.	Not mentioned.

IV. Conclusion

In this paper, different techniques related to removal of ambiguities in stereo images have been studied. Each technique has its own benefits and limitations and each technique has its own application area in computer vision. Some of the algorithms are related to local methods and some are related to global ones. Some of the algorithms are very good in the state-of-the-art both in efficiency and accuracy. Other are performed very well in the difficult areas like textureless regions, disparity discontinuous boundaries and occluded portions. But still some research is required especially in segmentation quality, reduced computational time, high accuracy, textureless regions, disparity discontinuous boundaries and occluded areas.

References

[1] Guoping Li, Yun He,"A hierarchical combined feature- and area- based stereo matching algorithm", IEEE, pp. 277-280, 2002.

[2] Motilal Agrawal, Larry S. Davis,"Window-based, discontinuity preserving stereo", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004.

[3] Yichen WEI, Long QUAN,"Region-based progressive stereo matching", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004.
[4] Li Hong, George Chen,"Segment-based stereo matching using graph cuts", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004.
[5] Andreas Klaus, Mario Sormann, Konrad Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure", IEEE, the 18th International Conference on Pattern Recognition (ICPR'06), 2006.
[6] Kuk-Jin Yoon, In So Kweon,"Stereo matching with the distinctive similarity measure", IEEE 2007.
[7] Barıs Baykant ALAGÖZ,"Obtaining Depth Maps From Color Images By Region Based Stereo Matching Algorithms", OncuBilim Algorithm and Systems Labs. Vol. 08, No. 04, 2008, pp. 1-13.
[8] Zhihua Liu, Qixiang Ye, Lu Ke, Jianbin Jiao,"A progressive region-merging algorithm for stereo matching", IEEE, pp. 142-145, 2009.
[9] Wei-Qun Li, Xian-Ming Chen,"A fast stereo matching using image segmentation for high quality", IEEE, pp. 239-243, 2010.
[10] Lin Chen, MinGang Chen, Xiao Lin, LiZhuang Ma, "An efficient region-adaptive stereo matching algorithm", IEEEICALIP, pp. 1093-1097, 2012.
[11] Zi-wei Zhou, Ge Li, Ji-zhuang Fan, Jie Zhao, XinYuOyang, "A new stereo matching algorithm based on image segmentation", IEEE International Conference on Information and Automation, pp. 861-866, June 2012.
[12] RC Chakraboorty,"Introduction to Artificial Intelligence", lec notes, June 01, 2010
[13] D. Scharstein, 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.

Simi Thakur received her B.Tech degree in computer science engineering from Adesh Institute of Engineering and Technology (Faridkot) under Punjab Technical University in 2011 and pursuing M Tech. (Regular) degree in computer engineering from Yadavindra College of Engineering Punjabi University Guru Kashi Campus Talwandi Sabo (Bathinda), batch 2011-2013. Her research interests include computer vision. At present she has guided many students regarding the research study in computer vision.