A Robust Stereo Matching Algorithm Based on Disparity Growing of Non-Horizontal Edge Features

S.Shirazi¹, P.Moallem², M. Ashourian³

¹Electrical Engineering, Tiran Branch, Islamic Azad University, Isfahan, Iran; Email: fars423@gmail.com
²Electrical Engineering Department, Faculty of Engineering, University of Isfahan, Isfahan, Iran Email: p_moallem@eng.ui.ac.ir
³Electrical Engineering Department, Majlesi Branch, Islamic Azad University, Isfahan, Iran Email: ashourian@iaumajlesi.ac.ir

ABSTRACT

In this paper, a robust stereo matching technique using growing component seeds has been proposed in order to find a semi-dense disparity map. In the proposed algorithm only a fraction of the disparity space is checked to find a semi-dense disparity map. The method is a combination of both the feature based and the area based matching algorithms. The features are the non-horizontal edges extracted by the proposed feature detection algorithm. In order to compute the disparity values in non-feature points, a feature growing algorithm has been applied. The quality of correspondence seeds influences the computing time and the quality of the final disparity map. We show that the proposed algorithm achieves similar results as the ground truth disparity map. The proposed algorithm has been further compared with box filtering, belief propagation, random Growing-Component-Seeds and Canny-Growing-Component-Seeds. According to the obtained results, the proposed method increases the speed of computation ratio of the random GCS algorithm by 31% and the ratio of canny GCS algorithm by 13%. The experimental results show the proposed method achieves higher speed, more accurate disparity map, and the lowest RMS errors.

Keywords: stereo matching, feature based matching, area based matching, fast matching,

Mathematics Subject Classification: Image Analysais 62H35

Computing Classification System:

1. INTRODUCTION

Stereo vision is an area of study that attempts to recreate the human vision system by using two or more views of the same scene to derive 3D depth information about the scene. As an emerging technology, stereo vision algorithms are constantly being revised and developed, and many alternative approaches exist for implementing a stereo vision system (Munro, 2009), (Foggia 2007).

Stereo vision based obstacle detection is an algorithm that aims to detect and compute obstacle depth using stereo matching and disparity map. The disparity map or motion field obtained from the
matching stage may then be used to compute the 3D positions of the scene points given the imaging geometry. Matching techniques can be divided broadly into pixel-based, area-based and feature-based image matching, or a combination of them. Other types of stereo matching methods such as diffusion-based (Scharstein 1998), wavelet-based (Kim 1997), phase-based (Porr 1998), and filter-based (Jones 2000) have also been developed. Some other less common techniques for estimating image motion or optical flow are reported in (Limongiello 2007) or for medical analysis in (Jainy Sachdeva et. al. 2012) and Deddy Kurniadi (2010).

In our developed system, the pre-supposition of a stereo system are considered which are: (1) Epipolar constraint; (2) Uniqueness constraint; (3) Smoothness constraint; and (4) Ordering constraint or monotonicity constraint.

In all matching algorithms, a trade off between the accuracy and the speed exist. The smaller the field of searching and calculation, the faster the comparison phase and the less calculation cost will be. In this paper, the proposed algorithm is compared with box filtering, belief propagation, random Growing-Component-Seeds and Canny- Growing-Component-Seeds.

The next presents basics of stereo matching. We used segment-based stereo matching using belief propagation for comparison with the proposed method. The proposed method combining area based and feature based matching algorithms using growing component seeds on the features is explained in Section 3. Section 4 provides some experimental results on well-known databases. Finally, the conclusion is given in the last section.

2. MATCHING ALGORITHMS

The Different comparison criterions have been suggested for the analogy of areas and disparity computation (Haralick 1993), which most of them are the statistic equations between brightness and light intensity of the left and right window (Sun 2000). There is some kind of function matching for stereo matching such as: Sum of Squared Difference (SSD), Sum of Absolute Difference (SAD), Normalized Cross Correlation (NCC) and Sum of Hamming Distance (SHD) (Moallem 2005), (Fazli 2009),(Yoon 2012). In this paper the applied technique for matching is as follows:

\[
NCC(f, g) = \frac{(f - \bar{f})(g - \bar{g})}{|f - \bar{f}| |g - \bar{g}|}
\]  

(1)

Where \(f\) and \(g\) are the intensity values the left and right images within a box. The stereo problem is solved based on an optimization algorithm. Different optimization techniques, including dynamic programming (DP) (Mohammadi 2010), graph cuts (Jones 2000), (Limongiello 2007), (Kim 1997), and genetic algorithm (Kim 1996) have been applied under this framework. These techniques try to find a global optimized solution under some given conditions defined by limiting some parameters. In order to avoid searching the entire disparity space, a number of algorithms were proposed which greedily
grow corresponding patches from a given set of reliable seed correspondences. Such algorithms assume that the neighboring pixels have similar disparity, not exceeding the disparity gradient limit or other constraints (Haralick 1985). Some of the benchmarking algorithms which we compare our result with them are: Disparity Component Growth (Radim 2006), (Radim 2002), (Sharstein 2002); Box filtering (Xiaoyan 2012), (Sun 1997), and Segment-Based Stereo Matching Using Belief Propagation (Sharstein, 2002) (Lankton 2009) realization of this work supposes the availability of a great number of repetitions of samples responding to the same known theoretical model. In practice, as the theoretical model is unknown, we use the Monte-Carlo method based on the generation of the data by computer according to a fixed theoretical model.

3. MATCHING ALGORITHMS

In this paper, we propose a novel stereo matching technique combining the feature based and the area based algorithms together. Fig. 1 shows an overall view of the proposed algorithm.
This algorithm consists of two phases; the first is feature extraction and matching and the second one is disparity growing on a special neighborhood around the extracted features. In the following sections each part is explained in detail.

3.1. Phase I

In this paper, the extracted features are edges. This consists of the implementation of various image processing algorithms like edge detection using Sobels, Prewitt, Canny and Laplacian, and so on. A different technique is reported to increase the performance of the edge detection. We first use the ancillary conditions taken by the Canny Filter. The Canny operator is used because Canny has shown that the ideal operator that maximizes the conventional signal-to-noise ratio in detecting a particular edge correlates with the same edge model itself. However, this detection is not well localized and requires an additional localization criterion (Moallem 2006), (Shriram 2010).

As mentioned before, the comparison of non-horizontal edges has difficulties in stereo vision and therefore we choose and use these non-horizontal edges as characteristic points for comparison. The second suggested algorithm makes it possible by the derivation of non-horizontal edges, to reduce a huge amount of pixels and fulfill the comparison based on these pixels, which mainly play a more important role in the comparison. In this algorithm, the volume and capacity of the calculation is not high and the time of the progress of identifying edges’ points is relatively short. If the noise of images is so that there is no need of filtering to eliminate them, the following actions may be taken to identify the non-horizontal edges with the breadth of one pixel: First we use a 3×3 Sobel filter which estimates the gradient parallel to x direction on both left and right images.

\[
S_x = \begin{bmatrix}
1 & 0 & -1 \\
2 & 0 & -2 \\
1 & 0 & -1
\end{bmatrix}
\]

(2)

To identify the real edges which have noise, the output of the filter has to be greater than a minimum threshold. Otherwise there will be many edges which most of them are appeared because of the noise.

To solve this problem, we go on by identifying the non-horizontal edges with the breadth of one pixel (Moallem 2005), (Sadeghi, 2008) (Jahn 2000), (Pollard 1985)

The edges’ points derived based on reduced light intensity on their sides are divided into two groups of positive and negative ones. To reach a diameter of one pixel, we need to calculate the maximum and minimum of the positive and minimum edge points. A non-horizontal thinned positive edge in a left image is localized in relation to a pixel when the filter response exceeds a positive threshold \(p_0^+\) and then a local maximum in the horizontal direction is obtained, therefore:

\[
\begin{align*}
\rho_l(x, y) &> \rho_0^+ & \text{Threshold} \\
\rho_l(x, y) &> \rho_l(x-1, y) \\
\rho_l(x, y) &> \rho_l(x+1, y)
\end{align*}
\]

Local maximum

(3)
Assume that $\rho_0^+$ is the mean positive value of the filter response. The extraction of non-horizontal positive thinned edge points is the same as the negative ones:

$$\begin{align*}
\rho_i(x, y) &< \rho_0^- \\
\rho_i(x, y) &< \rho_i(x-1, y) \\
\rho_i(x, y) &< \rho_i(x+1, y)
\end{align*}$$

Local maximum

Threshold

(4)

Now we can extract feature chains (Sadeghi 2008) with lengths more than 3. Each two sequence feature points in the same chain should be in the same sequence scan lines, and the disparity in the horizontal direction should be less than 2. Chains with length less than 3 are ignored. In the feature extraction stage, for both of the left and right images non-horizontal thinned edge extraction is performed. In the matching stage only similar features are to be compared. To detect non-horizontal edge points, the left and right images are convolved with a proper gradient mask. A non-horizontal thinned positive edge is localized in relation to a pixel where the filter response exceeds a positive threshold and obtains a local maximum in the direction. The positive threshold is selected dynamically based on the mean positive value of the filter response; which is highly important in wide baseline stereo. Then the matching of the founded seeds is examined and the matching seeds are found and saved in Matrix S.

![Figure 2. (a) Non-horizontal positive and negative edge (b) canny edge](image)

### 3.2. Phase II

The second phase consists of growing extracted features around a special neighborhood. Suppose an unsorted list of disparity seeds $S$ is given. Each seed is a point in the disparity space, $s = (x, x', y)$ defined by selecting a certain neighborhood $N(s)$ around this point formed with 16 points constructed from four sub-sets $N_1(s) \cup N_2(s) \cup N_3(s) \cup N_4(s)$, see Fig. 3 where we use the same colors for each:
Figure 3. Disparity space neighborhood used in this paper.

\[
N_1(s) = \{(x - 1, x' - 1, y), (x - 2, x' - 1, y), (x - 1, x' - 2, y)\},
\]

\[
N_2(s) = \{(x + 1, x' + 1, y), (x + 2, x' + 1, y), (x + 1, x' + 2, y)\},
\]

\[
N_3(s) = \{(x', y - 1), (x \pm 1, x', y - 1), (x, x' \pm 1, y - 1)\},
\]

\[
N_4(s) = \{(x', y + 1), (x \pm 1, x', y + 1), (x, x' \pm 1, y + 1)\}.
\]

An empty matching table \(\tau\) is defined and the disparity components are grown by drawing an arbitrary seed \(s\) from \(S\). Then the similarity is computed from small image windows around pixels \((u, v)\) and \((u', v)\) by for example the normalized cross-correlation (NCC) and adding it to \(\tau\), or by individually selecting the best-similarity neighbors \(q_i\) over its four sub-neighborhoods \(N_i(s)\):

\[
q_i = (u, u', v) = \arg \max_{(x, x', y) \in N_i(s)} \text{simil}(x, x', y)
\]

(5)

And placing each of these neighbors \((q_i)\) to the seed list if their inter-image similarity exceeds a threshold. Hence, up to four new seeds are created. If a seed from the list \(S\) is already a member of the matching table, then it will be discarded. If seeds have been already selected, the matching operation will be stopped for this seed. The development will be stopped by finishing all of the seeds of \(S\). The output of this phase - here on called Matching on the Special Neighborhood around Features is a partially filled matching table which its connected regions in 3D represent disparity components grown around the initial seeds.

The proposed method uses Growing Correspondence Seeds (GCS) for matching. A fast stereo matching algorithm is proposed that visits only a small fraction of the disparity space in order to find a semi-dense disparity map. It works by growing from a set of corresponding seeds. Unlike other seed-growing algorithms, this algorithm guarantees matching accuracy and correctness, even in the presence of repetitive patterns. In this paper the NCC function matching has been used on a \(5 \times 5\) window for comparison of the image similarity in all experimental results. We prepare an empty matching table \(\tau\) and start growing disparity components by drawing a seed \(s\) from \(S\), computed by using the proposed edge detection and adding it to \(\tau\). In this way higher performance can be achieved in compare to previous works such as (Cech, 2007). In contrast to the method proposed in (Cech, 2007) which uses random seeds, here we use an \(S\) matrix that contains the edges of left and right images (Cech, 2007). In this way we combine area based and feature based matching. This will improve the accuracy for real images.
where \( y \) is an \( n \times 1 \) vector observations of the dependent variables, \( X \) is the matrix \( n \times k \) of \( k \) explanatory variable, \( \varepsilon \) the vector of \( n \) theoretical residuals and \( \beta \) the vector of the theoretical regression coefficients. It is supposed that the residuals are independent random variable of the same normal distribution of null mean and constant variance \( \sigma^2 \). The parameters to be simulated are \( X \), \( \beta \) and \( \varepsilon \), while the vector \( y \) is calculated by the model.

### 4. EXPERIMENTAL RESULTS

In this section some experimental results is provided to evaluate the robustness of the proposed method. We have used Middlebury, CMU database, the database used in (Kim 1996), and real images captured by un-calibrated cameras. The program runs on PC with these features: AMD Athlon 64x2 Dual core processor 4800+2.5 Ghz, 896 MB of RAM. For simplicity, the following abbreviation is used: Belief propagation algorithm as “BP”; The box filtering as “BF”; Random Growing Component Seeds as “GCS”; Canny Growing Component Seeds as “Canny+GCS”; non-horizontal edge Growing Component Seeds as “NHE+GCS”.

Figures 4 to 8 show the original images and the ground truth of them. Figs. 10 to 12 show the left and right original images. Figs. 4 and 6 are images with rich texture content, and Figs. 8 and 10 are the relative images without texture. Figs. 5, 7, 9, 11, and 13 show the results of all the explained algorithms. Tables 1, 2, and 3 show the RMS error, time taken, and the percent of fault matching, respectively. The RMS error is calculates based on Equation 3.

\[
RMSE_{\text{Error}} = \sqrt{\frac{\sum (D(i,j) - G(i,j))^2}{n}}
\]  

(6)

\( D(i,j) \) is the disparity value in \((i,j)\) coordinates; and \( G(i,j) \) is the disparity value of the ground truth at \((i,j)\) coordinates. Tables 1, 2 and 3 compare the results. Figures 5, 7, 9, 11, and 13 show the qualitative evaluation.
Figure 4. Original image and Ground truth of Bowling2

Figure 5. Results of all explained algorithms of Bowling

Figure 6. Original image and Ground truth of Aloe
Box Filtering (BF)  Algorithm1 (BP)  GCS algorithm

Canny+GCS  NHE+GCS

Figure 7. Results of all explained algorithms of Aloe

Original image (corner)  Ground truth

Figure 8. Original image and Ground truth of corner
Box Filtering (BF)  Algorithm1 (BP)  GCS algorithm

Canny+GCS  NHE+GCS

Figure 9. Results of all explained algorithms of corner

Original image (Head left)  Original image (Head right)

Figure 10. Original image and Ground truth of Head
Figure 11. Results of all explained algorithms of Head

Figure 12. Original image and Ground truth of real image
Box Filtering (BF)  Algorithm1(BP)  GCS algorithm

Canny+GCS  NHE+GCS

Figure 13. Results of all explained algorithms of real image

Table 1. RMSError for different algorithms

<table>
<thead>
<tr>
<th></th>
<th>Bowling2</th>
<th>Flower pots</th>
<th>Baby3</th>
<th>Aloe</th>
<th>Cones</th>
<th>Poster</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
<td>300.2</td>
<td>412.3</td>
<td>201.3</td>
<td>150.3</td>
<td>212.1</td>
<td>100.3</td>
</tr>
<tr>
<td>GCS</td>
<td>233.2</td>
<td>351.0</td>
<td>193.1</td>
<td>125.4</td>
<td>147.1</td>
<td>91.9</td>
</tr>
<tr>
<td>Canny+GCS</td>
<td>228.5</td>
<td>214.8</td>
<td>200.5</td>
<td>124.5</td>
<td>142.3</td>
<td>75.3</td>
</tr>
<tr>
<td>NHE+GCS</td>
<td>225.5</td>
<td>203.4</td>
<td>193.1</td>
<td>109.2</td>
<td>129.2</td>
<td>66.9</td>
</tr>
</tbody>
</table>
The obtained results prove that the proposed method is fairly fast and robust against shadows. As it is shown in Figures 5 to 13, BF and BP algorithms are area based algorithm, thus they operate soundly in highly textured images as well as low textured images. The GCS algorithm achieves an accurate disparity map, but due to the fact that the chosen growing points are randomly selected, longer running time is elapsed in compare to the Canny+GCS algorithm and the NHE+GCS algorithm. The Canny+GCS uses “canny” edge detector so that the pixels of horizontal edges are also grown, thus it produces a disparity map with lower disparity than the NHE+GCS algorithm. According to Table 2, the NHE+GCS algorithm is almost 31% faster than basic GCS algorithm, and the Canny+GCS is 13% faster than basic GCS algorithm.

The NHE+GCS algorithm achieves higher speed and a more accurate disparity map with the lowest RMS errors. As shown in Figure 13, the NHE+GCS algorithm works well for un-calibrated (uncertified) real images as well. According to experimental results the NHE+GCS algorithm improves the speed, accuracy, and robustness, simultaneously.
5. CONCLUSION

In this paper a novel stereo matching algorithm has been proposed, namely the NHE+GCS algorithm. Our algorithm consists of two phases; the first phase is feature extraction and feature matching and the second one is disparity growing on a special neighborhood around features extracted. The method combines feature and area based matching. The proposed algorithm has been compared with box filtering, belief propagation, random Growing-Component-Seeds and canny Growing-Component-Seeds. The box filtering algorithm based on area increases the accuracy while the belief propagation increases the speed. The random and canny GCS algorithms increase both the speed and accuracy but the proposed method increases the speed even more and meanwhile preserves the accuracy. Thus, this method achieves a high quality disparity map in lower operation time, suitable for obstacle detection which requires high speed and accuracy in computing the disparity map.

6. REFERENCES


Kim T. and Muller, J. (1996). Automated urban area building extraction from high resolution stereo imagery. IVC, 4:115–130,


Yoon, K.-J. (2012) .Stereo matching based on nonlinear diffusion with disparity-dependent support weights, IET Computer Vision, 6 , Issue: 4