

# RELIABILITY MEASURE FOR PROPAGATION-BASED STEREO MATCHING

Guillaume Gales\*, Sylvie Chambon\*\*, Alain Crouzil\*\*, John McDonald\*

\*National University of Ireland Maynooth, Ireland  
{ggales,johnmcd}@cs.nuim.ie

\*\*IRIT Toulouse, France  
{chambon,crouzil}@irit.fr

## ABSTRACT

Seed propagation-based stereo matching can help to reduce ambiguity occurring when a pixel from one image has different putative correspondents in the other one due to difficult areas (repetitive patterns, homogeneous areas, occlusions and depth discontinuities). They rely on previously computed matches (*seeds*) to reduce the size of the search area, and thus the number of candidates. One approach of these iterative methods selects the “best” seed at each iteration to prevent the propagation of errors. However, little attention has been brought to this best-first selection criterion for which a correlation score is usually employed. This value itself does not consider any ambiguity and is not well-suited to select the most reliable seed. Therefore, in this paper we introduce a reliability measure. It has the advantage of taking into account information from the other candidates, and leads, according to the provided experimental evaluation, to better results than the correlation score alone.

## 1. INTRODUCTION

**Stereo matching** – Widely studied problem in computer vision, it consists in finding, from two images (*left* and *right*) of a same scene, which pairs of pixels correspond to the projection of a same element. A taxonomy of different stereo correspondence algorithms can be found in [1]. These algorithms are classified into two main categories: local (window-based) and global methods. The methods from the former category search for the pixels whose neighbourhoods are the most similar based on correlation measures. The methods from latter category search for the correspondences that minimize a cost function of the overall matching error. Local methods are fast and simple to implement but they are more prone to errors in difficult areas (near occlusions and depth discontinuities) than global ones, which are more time consuming. However, few papers focus on seed propagation-based technique which has the advantage of reducing errors. This is a special case of local method. As illustrated by Figure 1(a), in the basic local approach, the correspondent of one pixel is sought through a large search area (9 pixels in the given example). This search area is usually reduced to a line using epipolar geometry. Nevertheless, the size of such a search area is still large enough to introduce ambiguity as shown by Figure 1(b).

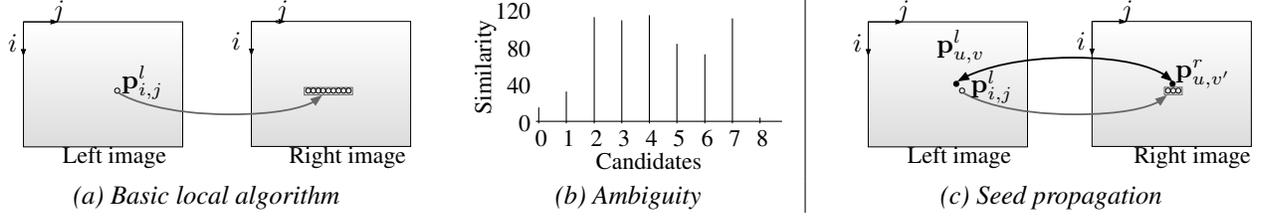
Indeed, one pixel from one image may have many different look-alikes in the other image. The idea behind seed propagation is to reduce this search area, and thus ambiguity, to the neighbourhood of reliable matches (the seeds), see Figure 1(c) where  $(\mathbf{p}_{u,v}^l, \mathbf{p}_{u,v'}^r)$  is a seed and the search area for the pixel  $\mathbf{p}_{i,j}^l$  is reduced to few pixels. Then, the newly found correspondences are added to the set of seeds and the process is reiterated until no more matches can be found.

Besides increasing the reliability of the computed matches, these methods are also suitable for small and wide baseline configurations and the results can be used by global methods which usually require an initialization step. For these reasons, in this paper, we focus on the seed propagation-based techniques.

**Motivation** – We focus on the “best-first strategy” to propagate from the “best” seed at each iteration in order to prevent the propagation of errors. A correlation score is usually employed, however, this criteria does not provide a measure of the reliability of a match since it does not reflect any ambiguity. Therefore, we propose to use a *reliability measure* instead. This measure takes into account an ambiguity term, a continuity term and a colour consistency term. Furthermore, we propose an experimental evaluation in a small baseline stereo context and compare the results with propagation using correlation as best-first selection criterion and the basic local algorithm.

**Propagation-based methods** – Seed propagation-based stereo matching relies on an initial set of trustworthy matches. Feature points can be extracted and matched using correlation or distance measures between descriptors [2, 3, 4, 5]. An alternative method consists in applying a basic local matching algorithm and selecting only the most reliable matches, satisfying a set of strong constraints [6].

Once the initial seed set is computed, the actual propagation can start. The propagation methods differ according to the way the seeds are used. In the *simultaneous approach*, at each iteration, all the seeds included within the set of seeds are propagated (*i.e.* a correspondent is sought for the neighbours of the left pixel of a seed) at the same time [6]. In the *sequential approach*, at each iteration, one seed is selected (according to its correlation score) to be propagated. This strategy is also called “best-first strategy” [2, 3, 7, 8, 4]. When no new matches satisfying a set of constraints can be found,



**Fig. 1.** (a) Basic local algorithm. (b) In this example, there is a high ambiguity *i.e.* we cannot distinguish the real correspondent between the candidates # 2, 3, 4 and 7. (c) The search area for the pixel  $\mathbf{p}_{i,j}^l$  is reduced to the neighbourhood of the correspondent of the seed ( $\mathbf{p}_{u,v}^l, \mathbf{p}_{u,v'}^r$ ) assumed reliable.

the set of seeds empties itself and the propagation stops. A threshold on the correlation score is usually used. A constraint is also used in [3] to avoid propagations in homogeneous areas.

**Sequential approach** – It is summarized by the algorithm 1 (notations are given in Table 1.), since we propagate at each iteration from the “best” seed, erroneous seeds are more likely to be selected towards the latest iterations. Therefore, some constraints can be introduced to stop the propagation before reaching these errors while having a quasi-dense result (*i.e.* a result for almost all the pixels of the image). For these reasons, in this paper, we focus on the sequential approach and more especially on the selection of the most reliable seeds.

$\mathbf{p}_{i,j}^k$	Pixel of coordinates $(i, j)$ within the image $k \in \{l, r\}$ ( $l$ for left, $r$ for right)
$d_m$	disparity value of a match $m$
$m = (\mathbf{p}_{i,j}^l, \mathbf{p}_{i,j'}^r)$	Candidate match
$s = (\mathbf{p}_{u,v}^l, \mathbf{p}_{u,v'}^r)$	Seed

**Table 1.** Notations.

**Algorithm 1** Sequential propagation-based stereo matching.

- 1: **Input:**  $\mathcal{S} \leftarrow$  initial set of seeds
- 2: **Output:** quasi dense disparity map
- 3: **repeat**
- 4:   select and remove  $(\mathbf{p}_{k,l}^l, \mathbf{p}_{k,l'}^r)$  the “best” seed from  $\mathcal{S}$
- 5:   **for** each unmatched pixel  $\mathbf{p}_{i,j}^l$  in the 8-neighbourhood of  $\mathbf{p}_{k,l}^l$  **do**
- 6:     find  $\mathbf{p}_{i,j'}^r$  the best candidate within the search area induced by the seed (see Figure 1(c))
- 7:     **if** the found match  $(\mathbf{p}_{i,j}^l, \mathbf{p}_{i,j'}^r)$  satisfies a set of constraints **then**
- 8:       accept this match
- 9:        $\mathcal{S} \leftarrow \mathcal{S} \cup (\mathbf{p}_{i,j}^l, \mathbf{p}_{i,j'}^r)$
- 10:     **end if**
- 11:   **end for**
- 12: **until**  $\mathcal{S} = \emptyset$

## 2. RELIABILITY MEASURE

Instead of a correlation score which does not reflect the ambiguity, we propose to use a reliability measure to select the

“best” seed at each iteration. This value is precomputed for the initial set of seeds, otherwise, it is computed during the matching step (step 6 of algo. 1). We define the “best” seed as a reliable seed *i.e.* a match without ambiguity. This is integrated in our reliability measure by an *ambiguity term*. When a new match computed from one seed is accepted, it is added to the set of seeds. The farther this match is from its seed, the less reliable this new seed is. This is also taken into account into our reliability measure as a *continuity term*. We also use the hypothesis that neighbouring pixels with similar colours are more likely to belong to the projections of points from a same surface, and thus to have close disparities. This is taken into account by a *colour consistency term*.

**Ambiguity term** – The correlation score itself does not take into account the ambiguity. We propose to use instead the probability induced by the correlation scores among all the candidates. We also believe that a small correlation window ( $3 \times 3$ ) helps to propagate into small details of the image. These small windows usually introduce ambiguity (because they are not as discriminant as bigger ones) but we believe that the ambiguity is reduced with a propagation algorithm (since we also rely on proximity of the pixels from a seed). We propose to use GC (Gradient Field Correlation), which gives the best results with a small window according to [9]. The ambiguity term is given by:

$$P_{GC}(m, s) = \frac{GC(\mathbf{p}_{i,j}^l, \mathbf{p}_{i,j'}^r)}{\sum_{\delta=0}^n GC(\mathbf{p}_{i,j}^l, \mathbf{p}_{i,d_s-\delta}^r)} \quad (1)$$

with  $d_s = (v' - v)$  and where  $GC(\mathbf{p}_{i,j}^l, \mathbf{p}_{i,j'}^r)$  is the correlation score<sup>1</sup> between the neighbourhoods of the candidate match,  $d_s$  is the disparity induced by the seed and  $n$  is the size of the search area. For the initial set of seeds, this is pre-computed with the largest possible search area. Thanks to the normalization, this term will be low for a highly correlated candidate if there is ambiguity (*i.e.* if other candidates are also highly correlated).

<sup>1</sup>The original  $GC_0$  is a dissimilarity measure and its formula can be found in [10]. The result varies into  $[0; 1]$  and we actually use the similarity value  $GC = 1 - GC_0$ .

**Continuity term** – The reliability score for each candidate is weighted according to the continuity hypothesis, stating that, neighbouring pixels should have close disparities. The continuity is usually ensured by a propagation algorithm when a small reduced search area is chosen. However, on one hand, if the chosen size is too small, we increase the risk of making mistakes in areas where the pixels are the projections of points from non-fronto-parallel surfaces. On the other hand, if the size is too large, we increase the risk of propagating errors in areas where the ambiguity is high. Thus, we propose to use a “large” search area and to lower the weight of candidates too far away from the seed. We used the Gaussian function to express this term:

$$P_s(m, s) = \mathcal{N}((j - j') - d_s \mid 0, \sigma_s). \quad (2)$$

For the initial set of seeds, this term is set to its maximum value  $\mathcal{N}(0 \mid 0, \sigma_s)$ . A small value of  $\sigma_s$  gives more chances to the candidates in the close neighbourhood of the right pixel of the seed to be selected<sup>2</sup>. A higher value gives more flexibility.

**Colour consistency term** – We assume that pixels within an homogeneous colour region are the projections of pixels from a same surface. We propose to use this idea in the propagation by using the Gaussian function of the distance between  $I(\mathbf{p}_{i,j}^l)$ , the color vector of the candidate  $\mathbf{p}_{i,j}^l$ , and  $I(\mathbf{p}_{u,v}^l)$ , the color vector of the left pixel of the seed  $\mathbf{p}_{u,v}^l$ :

$$P_c(\mathbf{p}_{i,j}^l, \mathbf{p}_{u,v}^l) = \mathcal{N}(\|I(\mathbf{p}_{i,j}^l) - I(\mathbf{p}_{u,v}^l)\| \mid 0, \sigma_c). \quad (3)$$

In order to use the Euclidean distance between two colours, we work in the Lab color space which is almost perceptually uniform. For the initial set of seeds, this term is set to its maximum value  $\mathcal{N}(0 \mid 0, \sigma_c)$ . Like this, we consider that a match is more reliable when the colour of its left pixel is close to the colour of the one from the left pixel of its seed.

**Reliability measure** – Finally, the reliability measure is given by:

$$P_m(m, s) = P_{GC}(m, s)P_s(m, s)P_c(\mathbf{p}_{i,j}^l, \mathbf{p}_{u,v}^l). \quad (4)$$

Although the proposed measure is intended to be used as the best-first selection criterion, since it is computed during the matching step, this term is also used instead of the correlation score to find out the best match among the candidates. A threshold  $t$  on this measure is used to accept a match.

### 3. EVALUATION AND RESULTS

**Compared methods** – We compare the following methods:

1. *Sequential propagation with the reliability measure* ( $P_{RM}$ ) Different values of  $\sigma_s \in [0.75; 30]$  and  $\sigma_c \in [0; 1 \times 10^5]$  are used to cover the set of all possible configurations. The seeds are computed by matching complementary feature points [11].

<sup>2</sup>The actual size of the reduced search area depends on the chosen  $\sigma_s$ . We can take  $4\sigma_s$  pixels since  $\mathcal{N}(x \mid 0, \sigma_s) \approx 0$  for  $x \geq 4\sigma_s$ .

A threshold  $t \in [0; 1]$  on the reliability measure is used to stop the propagation.

2. *Sequential propagation with correlation only* ( $P_{corr}$ ) We use GC with a  $3 \times 3$  window to be consistent with the proposed approach. Different sizes for the reduced search area are tested:  $s \in [1; 10]$ . A threshold  $t \in [0; 1]$  on the correlation score is used to stop the propagation.

3. *Basic local algorithm* (B.L.A.) We use GC with a  $3 \times 3$  window, to be consistent with the size chosen for the propagations, and with a  $9 \times 9$  window, more suited for this algorithm [9]. A threshold  $t_{3 \times 3 | 9 \times 9} \in [0; 1]$  on the correlation score is applied to keep the highest correlated matches giving different densities.

**Image dataset** – For our experimentation, we used stereo image pairs provided by Middlebury<sup>3</sup> [1, 12], see Figure 2. The images are epipolar rectified and the ground truth is given. It is obtained with a technique of projection of structured light (except for the *Tsukuba* pair, for which the ground truth is obtained manually, and for the *Venus* pair, where geometric constraints are used to obtain the ground truth).

**Evaluation criteria** – Unlike the methods compared by the Middlebury protocol, the outputs of the compared methods are not expected to be dense. Therefore, we propose to measure the percentage of correct matches  $C$  and the percentage of computed disparities (density)  $D$ . Let  $d_{\mathbf{p}}^{\text{th}}$  be the theoretical disparity value of a pixel  $\mathbf{p}$ . Since the ground truth is given at sub-pixel precision, and since our matching is done at pixel precision, we suppose an error when:  $|d_{\mathbf{p}}^{\text{th}} - d_{\mathbf{p}}| > 1$ . Then, our evaluation criteria are:

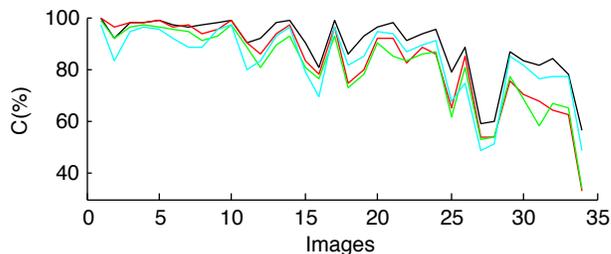
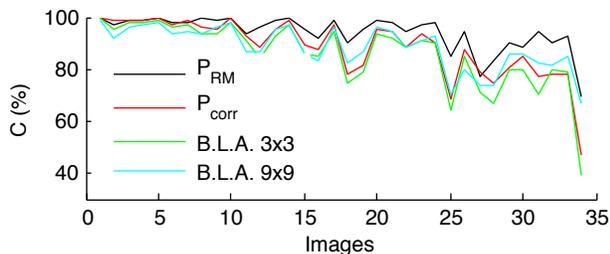
$$C = \frac{\# \text{ correct matches}}{\# \text{ matches}} \quad D = \frac{\# \text{ computed disparities}}{\# \text{ pixels}}$$

**Results** – Figure 2 presents the results on the 34 Middlebury pairs sorted according to a measure of “textureness”<sup>4</sup> from the most to the least textured image. For each method, only the configuration giving the best result is kept. Table 2 gives the values of the parameters that lead most of the time to the best result. For densities  $50\% \leq D < 75\%$ ,  $P_{RM}$  gives better results than the other methods except for the pairs #1, 3, 5, 7. These images are well-textured and are not very difficult to match, thus all the tested methods perform well. There is also an exception on the pair #27. For densities  $D \geq 75\%$ ,  $P_{RM}$  gives better results than the other methods except for the well-textured pairs #2 and 7. In both cases, the reliability measure leads to better improvement than the other methods with images showing large homogeneous areas. Moreover, B.L.A.<sub>3×3</sub> performs better than B.L.A.<sub>9×9</sub> on well-textured images (and vice-versa).

Figure 3 presents disparity maps obtained with the evaluated methods ( $D \geq 75\%$ ). Furthermore, we try a stereo pair of a building. This pair is particularly difficult to match due to large occlusions and repetitive texture patterns. Though the

<sup>3</sup>vision.middlebury.edu/stereo

<sup>4</sup>Percentage of pixels for which the squared horizontal intensity gradient is greater than some threshold [1].



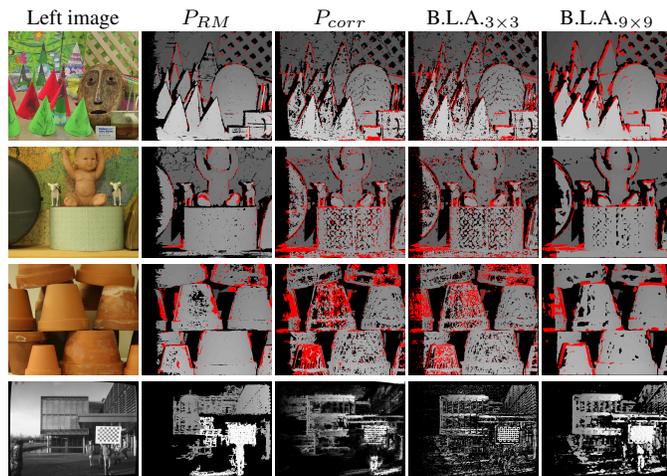
1.Cloth1 2.Aloe 3.Cloth3 4.Rocks2 5.Cloth4 6.Rocks1 7.Cloth2 8.Cones 9.Sawtooth 10.Barn1 11.Dolls 12.Teddy 13.Poster 14.Barn2 15.Moebius 16.Art 17.Bull 18.Books 19.Tsukuba 20.Baby1 21.Venus 22.Bowling2 23-24.Baby2-3 25.Laundry 26.Reindeer 27-28.Midd1-2 29.Flowerpots 30.Bowling1 31.Monopoly 32-33.Lampshade1-2 34.Plastic

**Fig. 2.** Comparison of different methods of matching. These graphs show the percentages of correct matches for the best results obtained with a density  $50\% \leq D < 75\%$  (left) and the ones obtained with a density  $D \geq 75\%$  (right).

METHOD	$50\% \leq D < 75\%$	$D \geq 75\%$
$P_{RM}(\sigma_s, \sigma_c, t)$	(1, 750, $4.7 \times 10^{-6}$ )	(1, 1000, $2.8 \times 10^{-6}$ )
$P_{corr}(s, t)$	(2, 0.8)	(2, 0.65)
B.L.A. ( $t_{3 \times 3}; t_{9 \times 9}$ )	(0.825; 0.825)	(0.725; 0.65)

**Table 2.** Modes of the values giving the best results.

result with  $P_{RM}$  is not perfect, we notice that the computed disparities are better than the ones obtained with the others methods on the facades of the building.



**Fig. 3.** Examples of results with *Cones* (#8), *Baby3* (#24) and *Flowerpots* (#29). Errors are shown in red. The last row shows the results on an image of a building.

#### 4. CONCLUSION

In this paper we used a reliability measure instead of a correlation score for propagation-based stereo matching. This measure is based on three terms which take into account ambiguity induced by other putative matches, as well as disparity continuity and color consistency hypothesis. The most reliable seeds, according to this measure, are propagated first. This helps to prevent the propagation of errors. Our evaluation showed that the proposed measure leads to better results than the correlation approaches on 85% of the dataset

( $50\% \leq D < 75\%$ ) and on 94% of the dataset ( $D \geq 75\%$ ). Further work will deal with the contribution of each term independently, as well as different methods of combining them. The result obtained on the building image is promising and this work will be extended to achieve more accurate results by using local and global constraints.

*Acknowledgement* Research presented in this paper was funded by a Strategic Research Cluster grant (07/SRC/I1169) by Science Foundation Ireland under the National Development Plan. The authors gratefully acknowledge this support.

#### 5. REFERENCES

- [1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV*, 2002.
- [2] G. P. Otto and T. K. W. Chau, "Region-growing algorithm for matching of terrain images," *IVC*, 1989.
- [3] M. Lhuillier and L. Quan, "Match propagation for image-based modeling and rendering," *PAMI*, 2002.
- [4] J. Kannala and S. S. Brandt, "Quasi-dense wide baseline matching using match propagation," in *CVPR*, 2007.
- [5] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: A survey," *Foundations and Trends® in Computer Graphics and Vision*, 2008.
- [6] Q. Chen and G. Medioni, "A volumetric stereo matching method: Application to image-based modeling," in *CVPR*, 1999.
- [7] J. Čech, *Accurate and Robust Stereoscopic Matching in Efficient Algorithms*, PhD thesis, U. of Prague, 2009.
- [8] Y. Z. Ma and W. J. Liu, "Progressive matching based on segmentation for 3D reconstruction," in *ICCV*, 2005.
- [9] S. Chambon and A. Crouzil, "Similarity measures for image matching despite occlusions in stereo vision," *PR*, 2011.
- [10] A. Crouzil, L. Massip-Pailhes, and S. Castan, "A new correlation criterion based on gradient fields similarity," in *ICPR*, 1996.
- [11] G. Gales, A. Crouzil, and S. Chambon, "Complementarity of feature point detectors," in *VISAPP*, 2010.
- [12] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *CVPR*, 2003.