Optimal Parameter Estimation
for MRF Stereo Matching

R. Gherardi, U. Castellani⋆, A. Fusiello, and V. Murino

Dipartimento di Informatica, Università di Verona
Strada Le Grazie 15, 37134 Verona, Italy

Abstract. This paper presents an optimisation technique to select automatically a set of control parameters for a Markov Random Field applied to stereo matching. The method is based on the Reactive Tabu Search strategy, and requires to define a suitable fitness function that measures the performance of the MRF stereo algorithm with a given parameters set. This approach have been made possible by the recent availability of ground-truth disparity maps. Experiments with synthetic and real images illustrate the approach.

1 Introduction

Three-dimensional (3D) reconstruction is a fundamental issue in Computer Vision, and in this context, structure from stereo plays a major role. The process of stereo reconstruction aims at recovering the 3D scene structure from a pair of images by searching for *conjugate points*, i.e., points in the left and right images that are projections of the same scene point. The difference between the positions of conjugate points is called *disparity*. The search is based on a matching process that estimates the "similarity" of points in the two images on the basis of local or punctual information [10]. A wide class of methods based on Markov Random Fields (MRF) models [8] has been recently introduced (see [10] for a review). Even if those methods have proved effective for the estimation of the stereo disparity, they often need a thoroughly phase of manual tuning of the free parameters that occurs in the MRF functional, using trial and errors.

In this paper we propose a technique capable of automatic selection of the "best" free parameters, based on an optimisation algorithm and a suitable fitness function that measures the performance of the MRF stereo algorithm with a given parameters set.

We consider the probabilistic stereo method R-SMW, proposed in [4], where the winner-takes-all approach of the Symmetric Multiple Window (SMW) algorithm [5] is relaxed by exploiting the non-determinism of the MRF. The MRF functional has two free parameters that in this paper are computed as the result of an optimisation based on the Reactive Tabu Search [6, 1], which mitigates the problem of local minima trapping while driving the search to unexplored regions

⋆ Corresponding author: Tel +39 045 802 7988, Fax +39 045 802 7068, E-mail `umberto.castellani@univr.it`

of the solution space. The fitness function is defined by comparing the output disparity with a known ground-truth.

Similar approaches, based on Genetic Algorithms, have been proposed in the past, focusing on different applications [2, 9].

2 The R-SMW method

The R-SMW algorithm [4] is a probabilistic relaxation of the SMW algorithm [5], using a Markov Random Filed.

In general, when an MRF model is applied to computer vision problems, the image is interpreted as a realization of a discrete stochastic process in which each pixel i is associated to a random variable D_i. By applying the Bayes rule and the Hammersley-Clifford theorem, the estimated solution (i.e., the solution that satisfies the Maximum A-Posteriori probability criterion) is obtained with the minimisation of $U(d \mid g) = U(g \mid d) + U(d)$, where $U(g \mid d)$ is the *observation model* and $U(d)$ is the *a-priori model*. In particular, the observation model describes the noise that degrade the image and the a-priori model describes the a-priori information independent from the observations, like, for instance the smoothness of the surfaces composing the scene objects. In the proposed application, in order to deal with the stereo problem [3], the scene is modelled as composed by a set of planes located at different distances to the observer, so that each disparity value corresponds to a plane in scene. Therefore, the a-priori model is *piecewise constant*.

Whilst the a-priori model impose a smoothness constraint on the solution, the observation model should describe how the observations are used to produce the solution. In the R-SMW method, the observation term encodes a multiple windows heuristic inspired from the SMW algorithm. For each site, we take into account all the disparity values in the neighbourhood weighted by the related SSD error values.

First a disparity map is computed using the simple SSD matching algorithm, taking in turn the left and the right images as the reference one. This produces two disparity map, which we will call left and right, respectively. A left-right consistency constraint is implemented by coupling the left disparity and the right disparity values.

In order to define the MRF model, we introduce two random fields D^l and D^r to estimate the left and the right disparity map, two random fields G^l and G^r to model the left and the right observed disparity map, and two random field S^l and S^r to model the SSD error. The field D^l (or equivalently D^r) will yield the output disparity.

In the following we shall describe the MRF functional, by defining the the a-priori model, the observation model, and the left-right consistency term[1].

---

[1] In the next two subsections, will shall omit superscript l and r in the field variables. It is understood that the a-priori model and the observation model applies to both left and right fields

2.1 A-priori model

With the a-priori term we encode the hypothesis that the surfaces in the scene are locally flat. Indeed, we employ a piecewise constant model, defined as:

$$U(d) = \sum_{i \in I} \sum_{j \in N_i} \delta(d_i, d_j) \tag{1}$$

where d_i and d_j are the estimate disparities value (the realization of the field D) and the function $\delta(x, y)$ is defined as:

$$\delta(x, y) = \begin{cases} 0 & if \quad x = y \\ 1 & otherwise \end{cases} \tag{2}$$

This term introduces a regularisation constraint, imposing that all pixels assume the same value in a region, thereby smoothing out isolated spikes.

2.2 Observation model

In order to mimic the behaviour of the SMW algorithm, the observation model term introduces a local non-isotropic relaxation, favouring the neighbour observations with the lowest SSD value:

$$U(g, s \mid d) = \sum_{i \in I} \sum_{j \in N_i \cup \{i\}} \delta(d_i, g_j) \cdot \left(\frac{1}{s_j}\right) \tag{3}$$

where g is the observation disparity map (the realization of the field G), s is the observed SSD values (the realization of the field S), and d is the disparity estimate (the realization of the field D).

In this term, the estimate value at site i, d_i, is compared with all its observed neighbours $\{g_j\}_{j \in N_i}$ and with g_i. When d_i takes the disparity of one (or more) of its neighbours, one (or more) term(s) in the sum vanishes. The lower is the SSD error of the chosen disparity, the higher is the cost reduction.

2.3 Left-right consistency constraint term

Let d_i^l be the left disparity (i.e., the disparity computed taking the left image as the reference) at site i, and d_i^r the right disparity at site i. The left-right consistency constraint states that: $d_i^l = -d_{i+d_i^l}^r$. The corresponding energy term is:

$$V(d^l, d^r) = \sum_{i \in I} \delta\left(d_i^l, -d_{i+d_i^l}^r\right) \tag{4}$$

In this way we introduces a payload when the left-right constraint is violated.

2.4 Summing up

The final MRF functional writes:

$$U\left(d^l, d^r \mid g^l, s^l, g^r, s^r\right) = k_1 \cdot \left[U\left(g^l, s^l \mid d^l\right) + U\left(g^r, s^r \mid d^r\right)\right] +$$
$$+ k_2 \cdot \left[U\left(d^l\right) + U\left(d^r\right)\right] + k_3 \cdot V\left(d^l, d^r\right) \quad (5)$$

where $U\left(g^l, s^l \mid d^l\right)$ and $U\left(g^r, s^r \mid d^r\right)$ are the observation model applied to the left and right disparity reconstruction, $U\left(d^l\right), U\left(d^r\right)$ are the a-priori models and $V\left(d^l, d^r\right)$ is the left-right constraint term. The positive weights k_1, k_2, k_3 are the parameters that control the performance of the algorithms. As the absolute magnitude of the functional is not important in the MRF minimisation, we can set $k_1 + k_2 + k_3 = 1$, thereby reducing the free parameters to only two.

It is customary to adjust these parameters using trial and error; in the following of this paper we will describe our strategy for automatic optimal parameters selection based on Reactive Tabu Search [1].

3 Tabu Search

Tabu Search (TS) is a meta-heuristic introduced by [6] that systematically imposes constraints on the feasible solutions to permit exploration of otherwise forbidden regions of the search space. In particular, TS will not only remember the current and best solution but it will also keep information on the itinerary through the last solutions visited. Such information will be used in order to guide the transition from the current to the next solution.

The following components defines the TS.

Fitness function: this is a scalar function defined over the solution set, that return a score for each solution.
Move: a move is a procedure by which a new (feasible) solution is generated from the current one.
Neighbourhood: a neighbourhood of a solution is the set of all the solutions that can be reached with one move.
Tabu list: this is a list of moves that are forbidden (or tabu). Its length is fixed but it is updated dynamically with the last move that was picked.
Aspiration conditions: these are rules that overrides tabu restrictions. If the aspiration condition is satisfied, a tabu move becomes allowed.

The TS algorithm can be described as follows:

1. Given a starting solution, compute its fitness.
2. Generate the neighbourhood of the current solution, or, equivalently, a set of candidate moves. A move is allowed if it is not tabu or it satisfies the aspiration condition. Pick the allowed move that get to the best neighbouring solution and consider it to be the new current solution.
3. Repeat step 2 until some termination conditions are satisfied.

At each iteration, the chosen move is put in the tabu list, thereby preventing the algorithm to go back to recently visited solutions. However, given the fixed size of the tabu list, the search might be trapped in a cycle of length greater than the size list. In order to cope with this drawback, the *Reactive Tabu Search* (RTS) [1] has been proposed, which dynamically adjusts the tabu list size. In particular, the size is increased when configurations are repeated, otherwise it is reduced. Another reactive mechanism is the *escape*, which consist of a number of random step, and it is triggered whenever too many configurations are repeated too often. The reader is refereed to the bibliography [1, 7] for more information on TS and RTS.

4 RTS applied

The RTS is used to maximise a *fitness function* that measures the performance of the R-SMW stereo algorithm as the difference between the estimated disparity and the ground truth. The independent variables of the fitness function are the weights k_1 and k_2 (k_3 is recovered as $k_3 = 1 - k_1 - k_2$.) In more details the computation of the fitness function proceeds as follows:

- given a solution (parameter set) $s^{(\ell)} = (k_1^{(i)}, k_2^{(i)})$,
- run the stereo process with $s^{(\ell)}$ and find the disparity map D^ℓ,
- compute the fitness $f(s^{(\ell)}) = -\mathrm{err}(D^\ell, D^o)$, where D^o is the ground truth disparity.

Following [10], the disparity error is given by the fraction of wrong matches in non-occluded regions:

$$\mathrm{err}(D^\ell, D^o) = \frac{1}{N} \sum_{(i,j)\in I\setminus B} \delta\left(D^\ell(i,j), D^o(i,j)\right) \qquad (6)$$

where N is the number of pixel, B is the set of occluded pixels (provided with the ground truth), and $\delta(x, y)$ has been defined in Equation (2).

A *solution* of the RTS is a point in the the region of the plane k_1, k_2 limited by the axes and by the line $k_2 = 1 - k_1$ (the search space). A *move* consists in changing the parameters value in such a way that the ratio between two of them is preserved. This gives three directions along which one can move starting from the current solution. A discrete change in the value of one parameter is obtained by flipping one bit of its binary representation. The *tabu list* is always updated with the last chosen move. The *aspiration condition* says that if a move leads to a better solution it is chosen even if is tabu.

5 Experiments

In this section experiments are reported for both synthetic and real cases.

First we estimated the optimal parameters for Random Dots Stereograms (RDS). The fitness function was computed using a square RDS and and circular
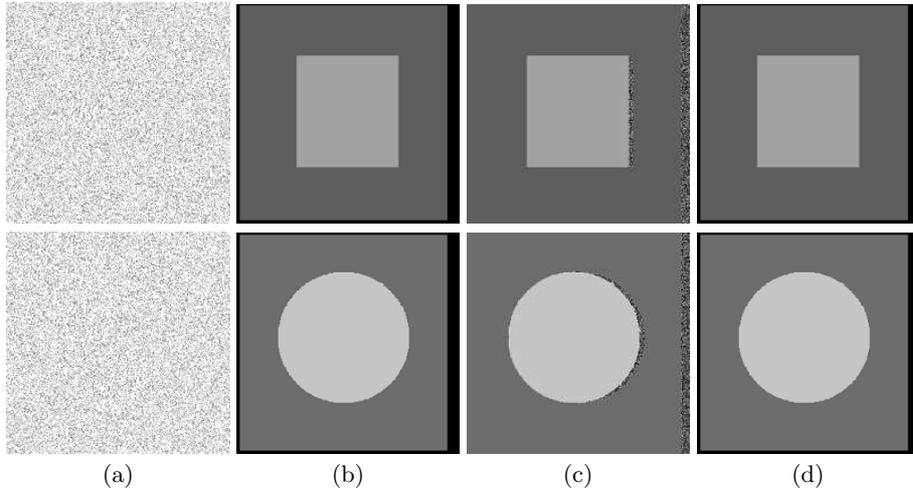
Fig. 1. Random Dots Stereograms. (a) Left image; (b) disparity obtained by SMW; (c) disparity obtained by R-SMW with optimal parameters; (d) ground truth disparity.

RDS (Figure 1). The RTS optimisation found the following values for the parameters $k_2 = 1, k_1 = k_3 = 0$, which reproduced the behaviour of the original SMW algorithm (not considering occluded areas), as can be seen in Figure 1. These values of the parameters make sense: the planar a-priori term is not needed since there is no noise; for the same reason and also because occlusions are not considered in the fitness function, the left-right consistency term is switched off.

Then we carried out experiments with the Middlebury data set [10], which is emerging as the de-facto standard data set for testing the performance of stereo algorithms. It consists of four stereo pairs: Map, Venus, Sawtooth and Tsukuba (Figure 2). The parameters estimation has been carried out using all the four stereo pairs (the global fitness function is the sum of the fitness for each set). The optimal parameters are reported in Table 1, in the column "Joint parameters". Figure 2 show the results obtained with these parameters for each stereo pair.

In order to assess the sensitivity of the parameters to the specific data set used for training, we estimated the optimal parameters separately for each stereo pair. The results are reported in Table 1, in the column "Single parameters". The error columns refer to the disparity error (i.e. the opposite of the fitness) value achieved by the given parameters set on a specific stereo pair.

It is worth noting that Sawtooth and Venus images are similar and so are the optimal parameters computed for these two stereo pairs. This seems to suggest that there are optimal parameters for classes of similar images.
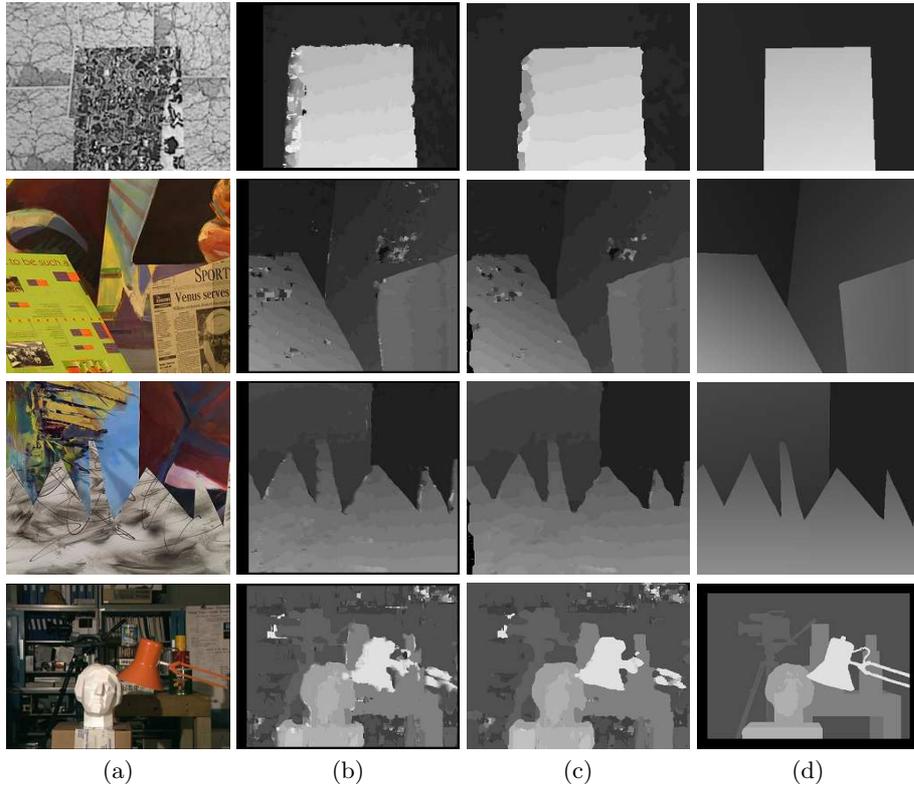
Fig. 2. Real experiments. (a) Left image of the stereo pair; (b) disparity obtained by SMW; (c) disparity obtained by R-SMW with optimal parameters; (d); ground truth disparity. Each row corresponds to a different stereo pair: (from top to bottom) Map, Venus, Sawtooth, and Tsukuba.

Table 1. Errors achieved by different parameters sets on different stereo pairs.

Stereo pair	Single parameters	Err.	Joint parameters	Err.
Map	(0.03, 0.88, 0.09)	0.26	(0.68, 0.23, 0.09)	0.55
Venus	(0.67, 0.26, 0.07)	2.92	(0.68, 0.23, 0.09)	3.16
Sawtooth	(0.69, 0.21, 0.10)	2.38	(0.68, 0.23, 0.09)	2.41
Tsukuba	(0.56, 0.12, 0.32)	4.67	(0.68, 0.23, 0.09)	4.71

6 Conclusion

The purpose of this paper has been to show that parameters tuning can be automated by using an optimisation strategy. We concentrated on stereo matching with a MRF-based algorithm (R-SMW) and used Reactive Tabu Search for parameters optimisation. The core ingredient is the fitness function, that measures

the performance of a particular parameters set. The usefulness of such an approach is based on the claim that there are optimal parameters that are valid for classes of images, instead of being image-specific. Future work will aim at substantiating this claim.

References

1. R. Battiti and G. Tecchiolli. The reactive tabu search. *ORSA Journal on Computing*, 6(2):126–140, 1994.
2. Luigi Cinque, Stefano Levialdi, Gianluca Pignalberi, Rita Cucchiara, and Stefano Martinz. Optimal range segmentation parameters through genetic algorithms. In *Proceedings of the International Conference on Pattern Recognition*, pages 1474–1477, 2000.
3. U. R. Dhond and J. K. Aggarwal. Structure from stereo – a review. *IEEE Transactions on Systems, Man and Cybernetics*, 19(6):1489–1510, November/December 1989.
4. A. Fusiello, U. Castellani, and V. Murino. Relaxing symmetric multiple windows stereo using markov random fields. In M. Figueiredo, J. Zerubia, and A.K. Jain, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, number 2134 in Lecture Notes in Computer Science, pages 91–104. Springer, 2001. Proceedings of the Third International Workshop, EMMCVPR 2001.
5. A. Fusiello, V. Roberto, and E. Trucco. Efficient stereo with multiple windowing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 858–863, Puerto Rico, June 1997. IEEE Computer Society Press.
6. F. Glover and M. Laguna. *Tabu Search*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
7. A. Hertz, E. Taillard, and D. de Werra. A tutorial on tabu search. Technical report, EPFL, Départment de Mathétiques, MA-Ecublens, CH–1015, Lausanne, 1995.
8. S. Z. Li. *Markov Random Field Modeling in Computer Vision*. Computer Science Workbench. Springer, Tokyo, 1995.
9. G. Pronzato and A.M. Wallace. Adaptive control of a boundary detection algorithm. In *Proc. of IEE Int. Conf. on Image Processing and its Applications,*, pages 356–360, Dublin, July 1997.
10. D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, May 2002.