Globally optimal solutions for energy minimization in stereo vision using reweighted belief propagation

Talya Meltzer, Chen Yanover and Yair Weiss
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Jerusalem, Israel
{*talyam,cheny,yweiss*}@*cs.huji.ac.il*

Abstract

A wide range of low level vision problems have been formulated in terms of finding the most probable assignment of a Markov Random Field (or equivalently the lowest energy configuration). Perhaps the most successful example is stereo vision. For the stereo problem, it has been shown that finding the global optimum is NP hard but good results have been obtained using a number of approximate optimization algorithms.

In this paper we show that for standard benchmark stereo pairs, the *global* optimum can be found in about 30 minutes using a variant of the belief propagation (BP) algorithm. We extend previous theoretical results on reweighted belief propagation to account for possible ties in the beliefs and using these results we obtain easily checkable conditions that guarantee that the BP disparities are the global optima. We verify experimentally that these conditions are typically met for the standard benchmark stereo pairs and discuss the implications of our results for further progress in stereo.

1 Introduction

Considerable progress in stereo vision has been achieved by formulating the problem in terms of energy minimization [3, 6, 11, 12, 16]. To illustrate the power of energy based methods, Figure 1(a) shows the "tsukuba" image, a benchmark image for stereo vision. Figure 1(b) shows the output of a standard SSD based algorithm followed by a "min" filter [14]. As can be seen, the use of a single window size is problematic — windows that are too small may not have enough structure in them to resolve the correspondence and windows that are too large cause noticeable artifacts at the occlusion boundaries and completely miss thin structures.

The energy based methods, in contrast, do not use a window of analysis. Rather, an energy function is defined which has a local term that meaures the goodness of a correspondence at a single pixel and a pairwise term that penalizes differences in disparity between neighboring pixels. The disparity is then found by running an algorithm that attempts to minimize this energy function. Perhaps the two most successful energy minimizers for this problem are Graph Cuts [3, 11] and belief propagation [6, 12, 16]. Figure 1(c) shows the output of the Graph Cuts algorithm as implemented in [12] (visually similar results are also obtained using belief propagation). Unlike the normalized correlation output, the energy minimization approach is able to provide sharp boundaries and preserves thin structures.

Despite the success of these methods, their output is still not perfect. For example, in Figure 1(c), the video camera is chopped in half. One can think of two different approaches for improving the output: (1) changing the energy function and (2) finding a different minimizer for the same energy function. Deciding between these two approaches is currently difficult because both belief propagation and Graph Cuts are only guaranteed to find *local* minima of the energy function. We do not know if a better optimizer would find a better solution.

Obviously, if we had a method that is capable of finding the *global* optimum of the energy function we would have a better idea of how to proceed. Unfortunately, it has been shown that for energy functions typically used in stereo, finding the global optimum is NP complete [3]. While this makes it extremely unlikely that we will be able to find the global optimum for *all* images in polynomial time, it leaves open the option for finding the global optimum for *some* images. In this paper, we show that a modification of belief propagation provides such an algorithm. In particular, we show that for the standard stereo benchmark images, the global minimum can be found in about 30 minutes per image.

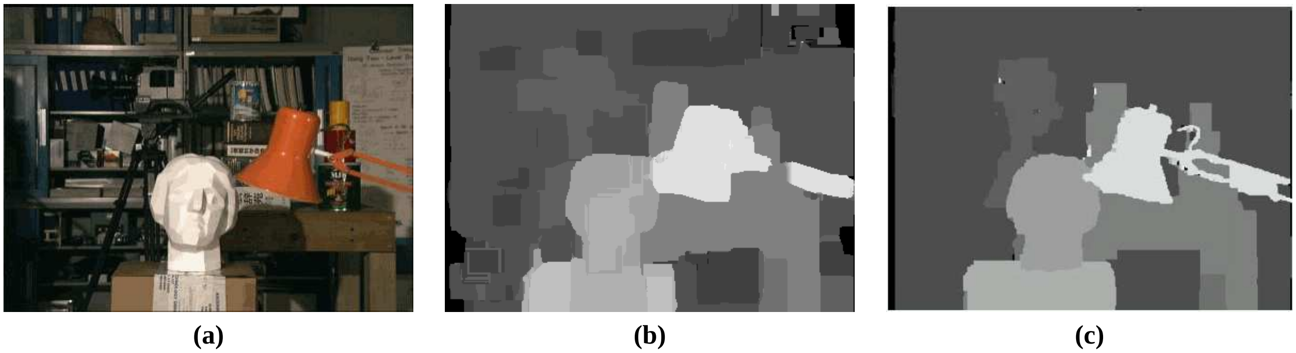


Figure 1: **(a)** A single frame from the "tsukuba" benchmark image. **(b)** The output of a standard SSD based algorithm, and **(c)** the output of an energy based method. The energy based method performs noticeably better but still misses some structure. Is the problem due to poor optimization or a bad energy function?

1.1 Linear Programming Relaxation and Reweighted Belief Propagation

The algorithm we will use for finding the global optimum is called "Tree Reweighted Belief Propagation" (TRBP) [8,9,18,19]. The algorithm is closely related to *Linear Programming Relaxations*, a standard approach in computer science for approximating combinatorial problems [1]. In this section, we give a very brief introduction to this connection and refer the reader to [8,19] for more details.

Denote by x the disparity image, $E(x_i, x_j)$, the pairwise compatability cost and $E(x_i)$ is the local data cost. Then the global minimum of the energy, x^* is:

$$x^* = \arg\min_x \sum_{i,j} E_{ij}(x_i, x_j) + \sum_i E_i(x_i) \qquad (1)$$

By introducing indicator variables, $q_i(x_i), q_{ij}(x_i, x_j)$ we can reformulate this as an integer program:

$$\text{minimize: } J(\{q\}) \quad = \qquad (2)$$
$$= \sum_{i,j} \sum_{x_i,x_j} q_{ij}(x_i, x_j) E_{ij}(x_i, x_j) \quad + \quad \sum_i \sum_{x_i} q_i(x_i) E_i(x_i)$$

subject to:

$$q_{ij}(x_i, x_j) \in \{0, 1\} \qquad (3)$$
$$\sum_{x_i,x_j} q_{ij}(x_i, x_j) = 1 \qquad (4)$$
$$\sum_{x_i} q_{ij}(x_i, x_j) = q_j(x_j) \qquad (5)$$

The *Linear Programming (LP) Relaxation* involves replacing the hard constraint $q_{ij}(x_i, x_j) \in \{0, 1\}$ with the relaxed constraint $q_{ij}(x_i, x_j) \in [0, 1]$. This relaxed problem can now be solved in polynomial time using an LP solver, and if all the relaxed variables $q_{ij}(x_i, x_j)$ in the LP solution happen to be nonfractional (i.e. they satisfy the hard constraint $q_{ij}(x_i, x_j) \in \{0, 1\}$) then we have found the global optimum of the energy.

Note that there is usually no way to know in advance whether the LP solution will be fractional or not — we just have to run it on a particular problem and check. Unfortunately, as [4,8] have observed, running standard LP solvers on the stereo problem is not practical with today's computing hardware. This is simply due to the large number of variables and constraints. In most stereo problems, the disparities are discretized to 30 possible states, so that the pairwise indicator $q_{ij}(x_i, x_j)$ is 30×30 matrix and for a 200×200 image there are $2 \times 200 \times 200$ of these matrices. We have found that the largest image we can solve with a standard, interior-point LP solver [17] is 39×39.

Wainwright and colleagues suggested a different way of solving LP relaxations, which is related to the Lagrangian dual of the LP problem. Their algorithm, TRBP, is a variant of BP that differs slightly in the message update equations. Define pairwise potentials $\Psi_{ij}(x_i, x_j) = \exp(-E_{ij}(x_i, x_j))$ and singleton potentials $\Psi_i(x_i) = \exp(-E_i(x_i))$, the TRBP algorithm iterates the following equation:

$$m_{ij}(x_j) \leftarrow \alpha \max_{x_i} \Psi_{ij}^{1/\rho_{ij}}(x_i, x_j) \Psi_i(x_i) \frac{\prod_{k \in N_i \setminus j} m_{ki}^{\rho_{ki}}(x_i)}{m_{ji}^{1-\rho_{ji}}(x_i)} \qquad (6)$$

where α is a normalization constant. After one has found a fixed point of these message update equations, the pairwise and singleton beliefs are defined as:

$$b_i(x_i) = \alpha \Psi_i(x_i) \prod_{j \in N_i} m_{ji}^{\rho_{ji}}(x_i) \qquad (7)$$
$$b_{ij}(x_i, x_j) = \alpha \Psi_i(x_i) \Psi_j(x_j) \Psi_{ij}^{1/\rho_{ij}}(x_i, x_j) \cdot \qquad (8)$$

$$\cdot \quad \frac{\prod_{k \in N_i \backslash j} m_{ki}^{\rho_{ki}}(x_i)}{m_{ji}^{1-\rho_{ji}}(x_i)} \frac{\prod_{k \in N_j \backslash i} m_{kj}^{\rho_{kj}}(x_j)}{m_{ij}^{1-\rho_{ij}}(x_j)}$$

The edge weights ρ_{ij} depend on the graph topology and for a grid graph they need to be strictly less than one. Specifically, we used $\rho_{ij} = 0.5$ for all edges. Note that for $\rho_{ij} = 1$, TRBP reduces to simple BP. Note also that the memory requirements of TRBP are similar to those of BP and are far less than second order LP solvers.

A useful property of TRBP beliefs is that they satisfy **max-marginalization**: $\max_{x_j} b_{ij}(x_i, x_j) = b_i(x_i)$.

Wainwright et al. showed that under certain conditions it is possible to transform the pairwise beliefs and singleton beliefs $b_{ij}(x_i, x_j), b_i(x_i)$ into indicator variables $q_{ij}(x_i, x_j), q_i(x_i)$ so that these indicator variables will be a solution to the LP problem. Perhaps the most important case is when the beliefs have no ties.

TRBP=MAP Theorem: Let $b_i(x_i)$ be beliefs calculated from fixed-point messages of TRBP. If there are no ties in these beliefs — for every i the maximum of $b_i(x_i)$ is attained at a unique value x_i^* — then x^* is the global minimum of the energy function.

The proof is given in [18, 19].

Kolmogorov [8] applied TRBP to the stereo problem and found that the conditions of the TRBP=MAP theorem *do not* hold for the standard benchmark images. That is, he found that many of the nodes had "ties" in them. He suggested a heuristic for breaking these ties and obtained lower energy than Graph Cuts using this heuristic, but emphasized that these disparities *are not* necessarily the global minimum of the energy function.

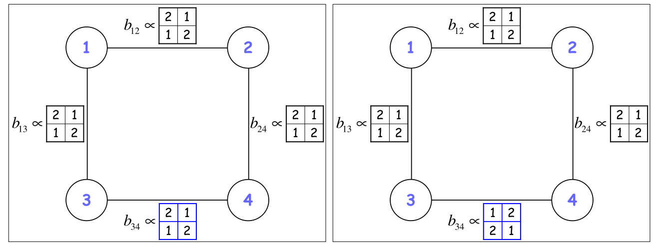
2 Theory

We now derive results and algorithms to obtain MAP solutions even when the TRBP beliefs have ties.

Our first results are based on a recent result by Wainwright et al. [19]:

TRBP with ties Theorem: Let $b_i(x_i)$ be beliefs calculated from fixed-point messages of TRBP (possibly with ties). If there exists an assignment x^* such that for every connected pair of pixels ij, $b_{ij}(x_i^*, x_j^*)$ maximizes the pairwise belief $b_{ij}(x_i^*, x_j^*) = \max b_{ij}(x_i, x_j)$ and for every pixel i, $b_i(x_i^*)$ maximizes the singleton belief, then x^* is the global minimum of the energy function.

Figure 2 illustrates the theorem. In figure 2(a), there are four pixels for which ties exist, and it is easy to see that choosing $x^* = 0$ will maximize the local singleton and pairwise beliefs. On the other hand, in figure 2(b) it is easy to show that there exists no assignment x^* which will maximize all the pairwise beliefs. This is because the pairwise beliefs create a *frustrated cycle*. To satisfy the top three



(a) **(b)**

Figure 2: An illustration of the TRBP with ties theorem. For the four nodes on the left, it is possible to find an assignment x^* that maximizes the pairwise and singleton beliefs. Our theorem proves that this means that x^* is the global optimum. For the four nodes on the right it is impossible to find such an assignment.

beliefs in the figure requires $x_1 = x_2 = x_3 = x_4$ while satisfying the fourth pairwise beliefs requires $x_4 \neq x_3$.

A naive way of searching for x^* would be to perform an exhaustive enumeration over all pixels in which there are ties in the beliefs. This is of course exponential in the number of pixels with ties. Fortunately, one can do this much more efficiently by taking advantage of the structure of the graph. In fact, as the following observation proves, when the set of pixels in which there are ties forms a singly connected set, there is no need to perform any search.

Observation (1): Let X_T denote the set of pixels for which there are ties in the TRBP beliefs and X_{NT} the set of pixels for which the TRBP beliefs are uniquely maximized by x_{NT}^*. If the graph of pixels X_T contains no cycles, then there exists an extension of x_{NT}^* to x^* that satisfies the conditions of the TRBP with ties theorem and hence x^* is the global minimum of the energy.

Proof: The proof is based on a single pass construction for the extension. For simplicity of exposition we describe the construction in the case where the set of pixels with ties X_T is a chain, and we label these nodes $x_1, x_2, x_3 \cdots x_n$. We choose an assignment for x_1^* as one of the maximizing values of $b_1(x_1)$. We then choose an assignment for x_2 as $x_2^* = \arg\max_{x_2} b_{12}(x_1^*, x_2)$ (when $\arg\max_{x_2} b_{12}(x_1^*, x_2)$ is not unique, arbitrarily set x_2^* to one of these maximizing states). The fact that the beliefs satisfy max-marginalization guarantees that x_2^* also maximizes $b_2(x_2)$ and that $b_{12}(x_1^*, x_2^*)$ maximizes $b_{12}(x_1, x_2)$. We now continue and choose $x_3^* = \arg\max_{x_3} b_{23}(x_2^*, x_3)$ and continue in this fashion until we have extended x^*. By construction, for any pair of nodes ij in the chain x_i^*, x_j^* maximize the pairwise beliefs and the singleton beliefs. Now consider a pair of nodes ij for which i is in the chain and j is not. Due to the max-marginalization property, x_i^*, x_j^* must also maximize b_{ij}.

Observation (2): Let X_T denote the set of pixels for

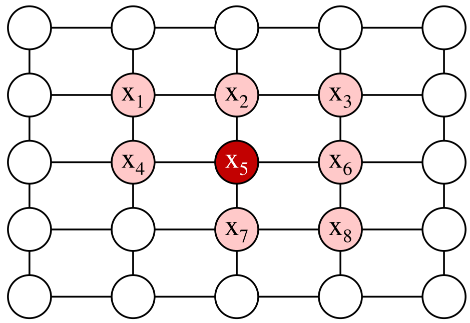


Figure 3: An illustration of the extended TRBP with ties theorem. The red region, T, contains tied pixels. We define $b_T(x_1, ..., x_8) = \prod_{<ij>\in T} b_{ij}^{\rho_{ij}} \cdot b_5^{c_5}$, with $c_5 = 1 - (\rho_{25} + \rho_{45} + \rho_{56} + \rho_{57})$. If there exists an assignment over the region x_T^* that maximizes b_T and for each $i \in \partial T$, x_i^* maximizes b_i, then we have the global optimum.

which there are ties in the TRBP beliefs and X_{NT} the set of pixels for which the TRBP beliefs are uniquely maximized by x_{NT}^*. Define a new cost function $C(x_T) = \sum_{i,j\in T} C_{ij}(x_i, x_j)$, with $C(i,j) = 0$ if x_i, x_j maximize b_{ij} and ϵ otherwise. If $\min_{x_T} C(x_T) = 0$, then there exists an extension of x_{NT}^* to x^* that satisfies the conditions of the TRBP with ties theorem and hence x^* is the global minimum of the energy.

Note that $C(x_T)$ is a cost involving only pairwise costs among the tied pixels and we can minimize it by finding the MAP assignment in an undirected graphical model. The complexity of minimizing $C(x_T)$ will typically be far less than a full exponential enumeration over all tied pixels and depends on the clique size in the junction tree [5].

3 Resolving Frustration

We have derived a novel theorem that significantly extends the class of TRBP beliefs from which one can provably extract the MAP solution.

Strong TRBP with ties Theorem: Let $b_i(x_i), b_{ij}(x_i, x_j)$ be beliefs calculated from fixed-point messages of TRBP (possibly with ties) with edge appearance probabilties ρ_{ij}. Let T be the set of tied pixels as before and let the boundary of T, ∂T, be the subset of nodes in T which have at least one neighbor in NT. Define $b_T(x_T) = \prod_{<ij>\in T} b_{ij}^{\rho_{ij}}(x_i, x_j) \prod_{i\in T\setminus\partial T} b_i^{c_i}(x_i)$ with $c_i = 1 - \sum_j \rho_{ij}$. If there exists an assignment x_T^* that maximizes $b_T(x_T)$ and for all nodes on the boundary of T, $i \in \partial T$, x_i^* maximizes $b_i(x_i)$, then (x_T^*, x_{NT}^*) is the MAP configuration.

Figure 3 illustrates the theorem. The red region, T, contains tied pixels. The boundary of T, ∂T, contains the nodes x_1 through x_7 eaceot x_5. Assuming $\rho_{ij} = \frac{1}{2}$ we would define $b_T(x_1, ..., x_8) = (b_{12}b_{23}b_{14}b_{25}b_{36}b_{45}b_{56}b_{57}b_{68}b_{78})^{-1/2} \cdot b_5^{-1}$, If there exists an assignment over the region x_T^* that maximizes b_T and for each $i \in \partial T$, x_i^* maximizes b_i, then we have the global optimum.

Before sketching the proof of the theorem, we provide some immediate corollaries:

Corollary 1: If all nodes on the boundary of the tied pixels ∂T have uniform beliefs, then the non-tied beliefs are optimal $x_{NT}^* = x_{NT}^{MAP}$

This is because for uniform beliefs on the boundary, any assignment x_T maximizes the beliefs on the boundary.

Corollary 2: If all nodes in the graph are binary, then the non-tied beliefs are optimal $x_{NT}^* = x_{NT}^{MAP}$

This is because for binary nodes, tied beliefs must be uniform. Corollary 2 has also been proven using a different technique by Kolmogorov and Wainwright [10].

The proof of the strong TRBP=MAP theorem is based on the following theorem, which is proven in [20]:

Convex GBP=MAP Theorem Let b_α, b_i be beliefs formed from messages of Generalized Belief Propagation with a convex free energy. Suppose there exists x^* such that for each α, i x_α^* maximizes b_α and x_i^* maximizes $b_i(x_i)$. Then that x^* is the MAP configuration.

This theorem shows that tree reweighted BP is a special case of a general class of algorithms that can be used to find MAP configurations. Tree reweighted BP is a special case of GBP where all the regions only include pairs of nodes and the free energy is indeed convex. Hyper-tree reweighted BP [19] is another special case of GBP where some regions may be larger than pairs of nodes, but there are many more convex free energies that are not in the class of tree reweighted or hypertree reweighted algorithms [7, 20].

The Convex GBP=MAP theorem suggests a strategy for dealing with cases where TRBP beliefs lead to frustration - to run convex GBP with regions that contain large subsets of the tied pixels. In fact, if we could run convex GBP in which all the tied pixels x_T form a single region, we would expect no frustrations in the output. Unfortunately, the complexity of GBP algorithms grow exponentially with the size of the largest region so this strategy is not practical. However, we can show that by using the beliefs b_i, b_{ij} obtained from running TRBP and defining b_T as in the theorem, *we do not have to actually run GBP*. The beliefs b_i, b_{ij}, b_T can be shown to be beliefs that arise from messages of GBP with a convex free energy and all we have to check is whether the conditions of the convex GBP=MAP theorem hold. If there exists an assignment x_T^* that maximizes $b_T(x_T)$ and for all nodes on the boundary of T, $i \in \partial T$, x_i^* maximizes $b_i(x_i)$, this is enough to guarantee that the conditions of the convex GBP=MAP theorem hold hence we have a global optimum.

We note that the conditions of the strong TRBP with ties theorem are sufficient conditions for optimality of x_{NT}, but they are *not* necessary. We have found cases where the conditions of the theorem do not hold but nevertheless the assignments of the non-tied pixels are indeed optimal. We also note that the theorem corresponds to running GBP with a particlar region graph - with a large region containing all tied pixels. It is straightforward to extend the result to other region graphs (e.g. where the large region contains a subset of the tied pixels). As in all other applications of GBP, the question of "how to choose the correct regions" is an open one.

An additional algorithm for resolving frustrations is to condition on one of the nodes involved in a frustrated cycle. Suppose, node x_1 is involved in a frustrated cycle and it has two maximizing disparities 1 and 2. We now run TRBP three times - once when x_1 is constrained to have disparity 1, once when x_1 is constrained to have disparity 2 and once when it is constrained not to have disparity 1 or 2. If we can find the MAP for each of these constrained problems, we can take the best configuration out of the three and that will be the MAP of the unconstrained problem.

In summary we have the following algorithm.

1. Run TRBP until convergence and identify the tied pixels x_T. If this set is empty, we have the MAP.

2. Construct a new graphical model that includes only the tied pixels and the possible states are only those that maximize the beliefs. The pairwise potentials are 1 if the pair maximizes the pairwise belief and ϵ otherwise. Use the Junction Tree algorithm to find x_T^*, the MAP in this new graphical model. If x_T^* has energy equal to zero then we have the MAP.

3. Construct a new graphical model that includes only the tied pixels and the original possible states. The pairwise potentials are $b_{ij}^{\rho_{ij}}(x_i, x_j)$ and the singleton beliefs are $b_i^{c_i}(x_i)$ for all nodes not on the boundary and uniform potentials for all nodes on the boundary. Use the junction tree algorithm to find all MAP configurations. If one of the MAP configurations also maximizes the local beliefs for all boundary nodes, then we have the MAP.

4. Choose a node x_i involved in a frustrated cycle (i.e. part of a pair of nodes i, j for which $E(x_i^*, x_j^*) \neq 0$ in the reduced graphical model of step 2). Run TRBP $k+1$ times where k is the number of disparities that maximize $b_i(x_i)$, and in each run the disparity at x_i is constrained as above. For each constrained problem, run TRBP and use steps 2,3 to find the MAP. If we can find the MAP for all constrained problems, then the best configuration of these $k+1$ is the MAP of the unconstrained problem.

4 Experiments

We ran TRBP on the images in the Middlebury stereo benchmark set [15]. We used the same energy function used by Tappen and Freeman [12]. The local cost is based on the Birchfield-Tomasi matching cost [2] and the pairwise energy penalizes for neighboring pixels having different disparities. The amount of penalty depends on the intensity difference between the two pixels — the smaller the intensity difference the larger the penalty. We used the BP code provided by Marshall Tappen and modified the code to perform TRBP rather than BP. We found that each iteration of TRBP took less than a second and using a very harsh convergence threshold (10^{-8}) convergence was typically obtained in about 2000 iterations, thus yielding run times of approximately half an hour per image.

For the second and third stages in the algorithm we used junction tree code provided by Kevin Murphy [13]. Even though the junction tree code is in Matlab, the run time is negligible compared to the TRBP run time (typically around 30 seconds for the junction tree).

Figure 4 shows the results for the four benchmark images with the standard parameters used in [12]. The top row shows the input images and the bottom rows show the TRBP results. Red pixels are those for which the TRBP belief had a tie. In all cases the ties occur at depth edges and the beliefs are undecided between the disparities at the two sides of the edge. As observed by [8], for each image there are some red pixels and hence the TRBP=MAP therorem does not hold. The second from bottom row shows the global optimum computed using our algorithm followed by the results of graph cuts and BP (replotted from [12]).

The global minimum produces smoother solutions but does not solve many of the problems present in the approximate solutions. For example, in the tsukuba image, the global minimum also "misses" half of the video camera and erroneously fills in the thin structure holding up the lamp.

Table 1 compares the global minimum of the energies for the four images to those found using Graph Cuts and belief propagation. Following [12] we also varied the parameters of the energy function to see how the difficulty of the problem varies. An indication of the hardness of the problem is provided by the number of tied pixels in the beliefs achieved using TRBP, and also by the size of the maximal clique in the junction tree. The table also shows what fraction of the problems each steps of our algorithm can solve. For about 50% of the problems we looked at, step 2 of our algorithm (which is cheapest) found the MAP. Step 3 of the algorithm is more expensive and in some cases even intractable (when the junction tree is too large). We are currently exploring running TRBP on the small graphical model in step 3 for these cases. Step 4 of the algorithm is the most powerful but also the most expensive.

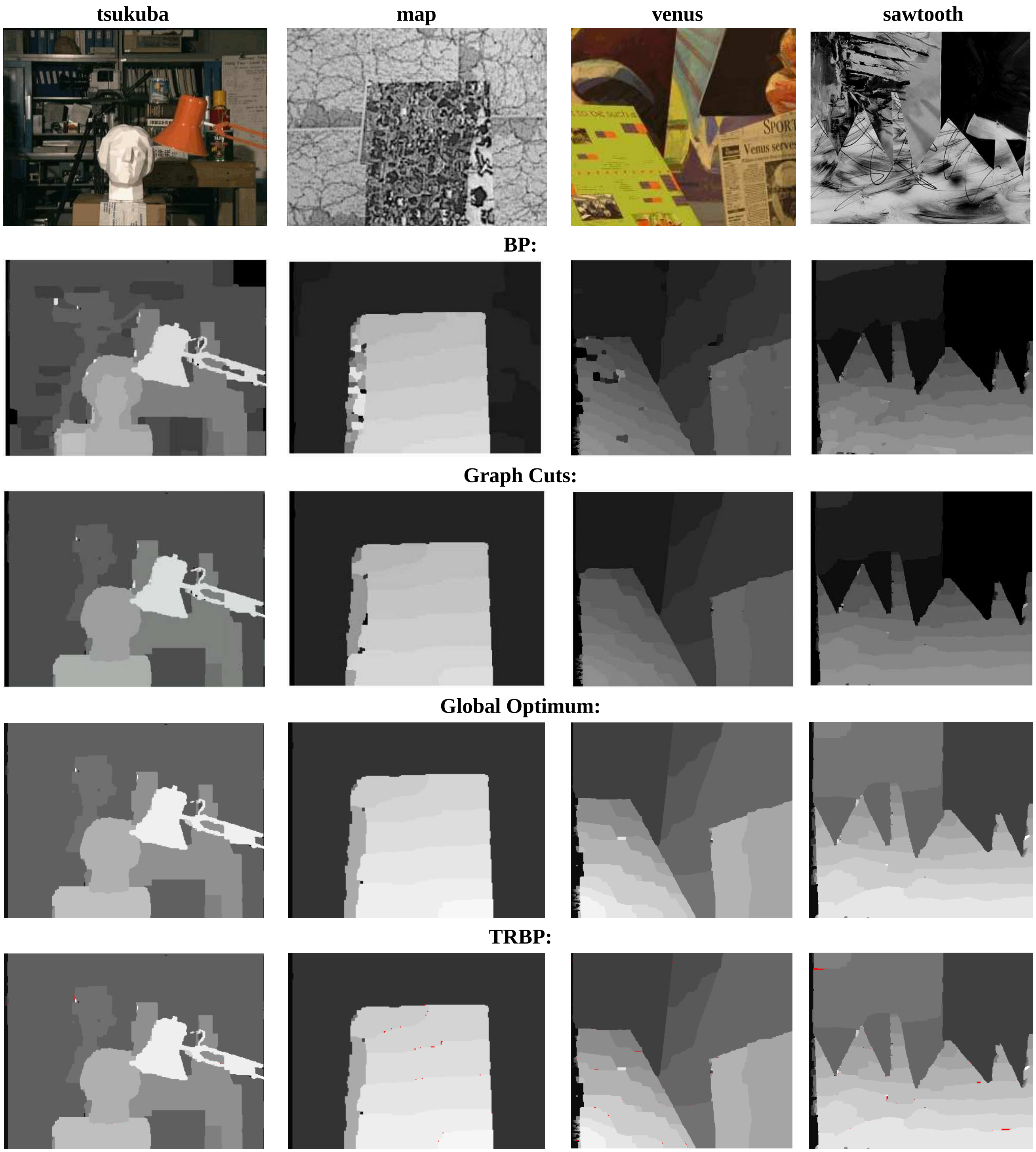


Figure 4: Comparison of the solutions obtained by BP, Graph Cuts and the global minimum using identical energy functions. The BP and Graph Cuts solutions are replotted from [12]. The global minimum produces smoother solutions but does not solve many of the problems present in the approximate solutions.

	MRF params			Energy			Algorithm			
	T	s	P	*Graph Cuts*	*Belief Propagation*	*Global minimum*	#Ties	Max. cliq.	MAP in step 2	MAP in step 3
tsukuba	0	20	1	461041	475829	459993	209	5	-	-
	4	20	2	489159	526006	487932	134	4	+	+
	4	20	4		589790	520097	53	3	-	+
	8	20	2	503751	538787	502206	140	6	-	-
	8	20	4	553215	620444	552302	55	5	-	+
	0	50	1	602355	633092	600086	56	3	+	+
	4	50	2	645865	758117	643946	20	2	+	+
	4	50	4	696251	899215	695236	85	6	-	?
	8	50	2	663845	775085	662543	15	2	+	+
	8	50	4	739000	941129	737699	10	1	+	+
map	0	20	1	309574	315455		429	16	-	?
	4	20	2	313123	322334		470	16	-	?
	4	20	4	315770	335110	314738	71	4	+	+
	8	20	2	320677	336503		355	20	-	?
	8	20	4	331302	366189	330150	85	7	-	?
	0	50	1	366289	385652	365501	32	3	+	+
	4	50	2	372693	405933	371993	29	2	+	+
	4	50	4	375979	427580	375734	71	4	+	+
	8	50	2	384342	442665	382160	37	4	+	+
	8	50	4	399455	518615	399383	19	3	+	+
venus	4	50	2	$1442 \cdot 10^3$	$1501 \cdot 10^3$	$1399 \cdot 10^3$	80	3	-	-
sawtooth	4	50	2	$1652 \cdot 10^3$	$1713 \cdot 10^3$	$1595 \cdot 10^3$	182	4	+	+

Table 1: The energies obtained by *Graph Cuts* and *BP* compared to the *global minimum energy* found using *TRBP* on the sequences with different parameters. The energy gives a cost sP for a disparity change if the intensity difference between neighboring pixels is less than T and s otherwise. The number of tied pixels and the maximal clique size found while processing the junction tree give an indication of the hardness of the problem. The second column from the right indicates whether observation 2 (used in step 2 of our algorithm) succeeded in provably finding the MAP. The last column indicates whether the strong TRBP theorem (used in step 3) succeeded in provably finding the MAP. '+' denotes success in finding the MAP, '-' failure, while '?' means that the step was too expensive computationally.

5 Discussion

Finding the global minimum of the stereo energy function for any image is NP-complete. Nevertheless, in this paper we have shown that using tree reweighted belief propagation, it is possible to find the global minimum for several standard benchmark images. TRBP is similar to linear programming relaxations in that it is guaranteed to find the global optimum when the beliefs are non-fractional, but unlike general purpose LP solvers, it can be applied to these large problems in a matter of minutes. We have extended the theory of TRBP to allow for the case of ties in the beliefs and derived easily verified conditions for the TRBP beliefs with ties to give the global minimum. We have verified experimentally that these conditions indeed hold on the standard benchmark images.

As can be seen, the global minimum of the energy function does not solve many of the problems in the BP or Graph Cuts solutions. This suggests that the problem is not in the optimization algorithm but rather in the energy function. A promising problem for future research is to learn better energy functions from ground truth data.

References

[1] D. Bertismas and J. Ttsitskikilis. *Introduction to linear optimization*. Athena Scientific, 1997.

[2] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):401–406, 1998.

[3] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. In *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, 1999.

[4] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001.

[5] R. Cowell. Advanced inference in Bayesian networks. In M.I. Jordan, editor, *Learning in Graphical Models*. MIT Press, 1998.

[6] P. Felzenszwalb and D. Huttenlocher. Efficient belief propagation for early vision. In *Proceedings of IEEE CVPR*, pages 261–268, 2004.

[7] T. Heskes. On the uniqueness of loopy belief propagation fixed points. *Neural Computation*, 2004.

[8] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. Technical Report 2005-38, Microsoft Research, 2005.

[9] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. In *Proceedings AI Stats*, 2005.

[10] V. Kolmogorov and M. J. Wainwright. On the optimality of tree-reweighted max-product message passing. In *Uncertainty in Artificial Intelligence (UAI 2005)*, 2005.

[11] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004.

[12] M.Tappen and W.T. Freeman. Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. In *ICCV*, 2003.

[13] Kevin Murphy. The Bayes net toolbox for matlab. *Computing Science and Statistics*, 33, 2001.

[14] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Compter Vision*, 2002.

[15] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003)*, volume 1, pages 195–202, 2003.

[16] J. Sun, H.-Y. Shum, and N.-N. Zheng. Stereo matching using belief propagation. In *ECCV*, volume 2, pages 510–524, 2002.

[17] R. Vanderbei. LOQO. `www.princeton.edu/˜rvdb`.

[18] M. J. Wainwright, T. Jaakkola, and A. S. Willsky. Exact MAP estimates by (hyper)tree agreement. In *Advances in Neural Processing Systems (NIPS)*. MIT Press, 2003.

[19] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky. MAP estimation via agreement on (hyper)trees: Message-passing and linear-programming approaches. Technical Report UCB/CSD-03-1269, UC Berkeley CS Division, 2003.

[20] Y. Weiss, C. Yanover, and T. Meltzer. Linear programming and belief propagation with convex free energies. Technical report, Leibniz Center, Hebrew University of Jerusalem, 2005.