A Comparative Study of Modern Inference Techniques for Discrete Energy Minimization Problems

Yufeng Jiang

1. Related Inference Studies

Since 2006 when the study was published, the field has made rapid progress. Modern vision problems involve more complex models, involve larger datasets and use machine learning techniques to train the model parameters and energies.

Taken together, these changes give rise to hard energy minimization problems that are fundamentally different than the ones considered by Szeliski *et al.* In particular, in [5] the models were restricted to 4-connected grid graphs with unary and pairwise factors only.

Apart from the study [5], there are many recent articles in computer vision which compare inference techniques for a small specialized class of models, such as [2, 4, 3]. Unfortunately, the models and inference techniques are often not publicly available. Even if they were available, the lack of a flexible software-framework which includes these models and optimization techniques makes a fair comparison difficult. Closely related is the smaller study that uses the first and now deprecated version of OpenGM. It compares several variants of message passing and move making algorithms for higher order models. In contrast to this work, polyhedral inference methods are not included and only a small number of synthetic models are used.

Outside computer vision, the Probabilistic Inference Challenge (PIC) [1] covers a broad class of models used in machine learning. They include the leading optimization techniques of PIC in their study.

Figure 1. Example for a pixel based model

2. Second-Order Models

Table 1 summarizes the results for the stereo labeling problems from [5]. For instances of this problem, methods like ogm-BUNDLE-A that solve the LP-relaxation perform best. The mrf-TRWS method from [5] produces the best solution on average, but does not solve the LP-relaxation optimally, since the average lower bound is not the tightest. Applying Lazy Flipper on top (TRWS-LF2) improves the solution further.

Pixel-Based Models For many low-level vision problems it is desirable to make each pixel a variable in the model. For 2D iamges, where variables are associated to pixels in a 2D lattice, a simple form of a factor graph model connects each pixel with its four nearest neighbors using a pairwise energy. This simple form is popular and was the sole subject of the study [5]. In their study they incorporated the models $mrf - sterro$, $mrf - inpainting$, and $mrf - photomontage$ from [5]. Additionally they used three models which have the same 4-connected structure, inpainting $inpainting - N4$, color segmentation $color - seg - N4$[1] and object segmentation $object - seg$. All sets have a small number of labels and use Potts regularizers. In $inpainting - N4$ and $color - seg - N4$ this regularizer is fixed for all factors. In $object - seg$, it depends on the image-gradient. The unary terms measure the similarity to predefined class-specific color models.

They also consider the task of in-painting in binary images of Chinese characters, $dtf - chinesechar$. The factors are learned potentials from a decision tree field. Although each variable has only two labels, their regional sparse neighborhood structure makes the resulting inference problem challenging, *c.f.* Fig.1.

In contrast to the previous case, primal move-making methods outperform LP-relaxation methods for the $mrf - photomaontage$ model, also from [5]. Table 2 provides the details. Authors believe this is because some unary terms have infinite values. As observed in [5], mrf-EXPANSION is both the fastest and achieves the best results on average. The same behavior can be seen for the synthetic inpainting instances $inpainting - N4/8$. The instance "inverse" is
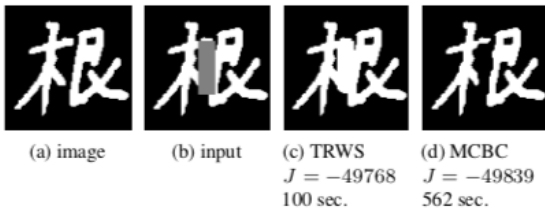
---

[1] The $inpainting - N4/8$ and $color - seg - N4/8$-models were originally used in variational approaches together with total variation regularisers. A comparison with variational models is beyond the scope of this study.

algorithm	mean run time	mean value	mean bound	best	ver. opt
FastPD	**4.47 sec**	1614255.00	-∞	0.00	0.00
FastPD-LF2	296.38 sec	1611484.33	-∞	0.00	0.00
mrf-EXPANSION	13.04 sec	1614353,00	-∞	**33.33**	0.00
mrf-TRWS	296.22 sec	1587928.67	1584746.53	0.00	0.00
ogm-BUNDLE-A	8458.49 sec	1610985.00	1584776.15	**33.33**	**33.33**
ogm-SUBGRAD-A	8533.41 sec	1752958.00	1578421.00	**33.33**	**33.33**
TRWS-LF2	668.47 sec	**1587040.00**	1584746.53	**33.33**	0.00

Table 1. mrf-stereo (3 instances)

algorithm	mean run time	mean value	mean bound	best	ver. opt
mrf-EXPANSION	**8.54 sec**	**168220.00**	-∞	**100.00**	0.00
mrf-SWAP	**11.35 sec**	**180345.00**	-∞	0.00	0.00
mrf-TRWS	253.20 sec	1243144.00	166827.07	0.00	0.00
ogm-BUNDLE-H	4771.35 sec	599206.00	111100.35	0.00	0.00
ogm-SUBGRAD-A	4734.20sec	3846787.00	26005.24	0.00	0.00
TRWS-LF2	431.34 sec	735193.00	166827.12	0.00	0.00

Table 2. mrf-photomontage (2 instances)

constructed such that the LP-relaxation over the local polytope is not tight. In this case pure primal move-making methods are superior, since they do not face the rounding problem.

References

[1] PIC2011. http://www.cs.huji.ac.il/project/PASCAL/. 1

[2] K. Alahari, P. Kohli, and P. H. S. Torr. Dynamic hybrid algorithms for MAP inference in discrete MRFs. *IEEE TPAMI*, 32(10):1846–1857, 2010. 1

[3] V. Kolmogorov and C. Rother. Comparison of energy minimization algorithms for highly connected graphs. In *ECCV*, pages 1–15, 2006. 1

[4] N. Komodakis and N. Paragios. Beyond loose LP-relaxations: Optimizing MRFs by repairing cycles. In *ECCV*, pages 806–820, 2008. 1

[5] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. *IEEE TPAMI*, 30(6):1068–1080, 2008. 1