Multi Label Generic Cuts: Optimal Inference in Multi Label Multi Clique MRF-MAP Problems

Chetan Arora

The Hebrew University of Jerusalem (now with IIIT Delhi)

S.N. Maheshwari

Indian Institute of Technology Delhi

IEEE 2014 Conference on Computer Vision and Pattern Recognition

MRF-MAP Inference Problem

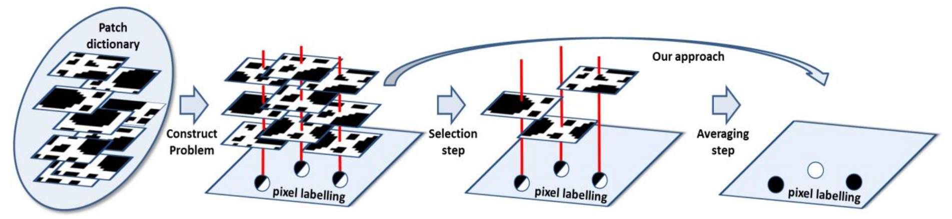
$$argmin\, E(l_P) = \sum_{p \in P} D_p(l_p) + \sum_{c \in C} W_c(l_c)$$

P is the set of pixels, C is the set of cliques.
D_p is per pixel unary/data cost.
W_c is per clique prior/clique potential.

- Inference problem is NP hard in general

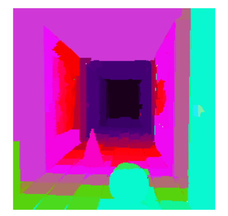
Higher Order MRF-MAP

Allow more complex clique potential based upon learnt patterns [4].

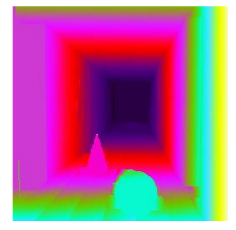


Structural constraints based upon shape and gradients can only be encoded using higher order potentials [5].

Reference Image



Disparity 2-cliquel



Disparity 3-clique

Inference Algorithms

Problem Type	Optimal Inference	Approximate Inference
2-Label First Order	Graph Cuts (max flow)	QPBO [3]
2-Label Higher Order	**Generic Cuts** [1]	**Approximate Cuts**[2], Reduction [7]
Multi-Label First Order	Ishikawa [6]	Alpha Expansion
Multi-Label Higher Order	**Proposed Algorithm**	Message Passing Variants [8-11]

Multi Label Generic Cuts (MLGC) - Encoding

1. Let $L = \{l_1, l_2, \ldots, l_m\}$ be the ordered label set.
2. For each pixel p we introduce an m-tuple $b_p = (b_1^p, b_2^p, \ldots, b_m^p)$.
3. The m possible label states of pixel p are represented by states $(0,1,1,\ldots,1), (0,0,1,\ldots,1), \ldots, (0,0,\ldots,0,0)$ of the m-tuple. That is the label at position i in L is represented by the state of b_p in which i Boolean variables from left have value 0 and remaining $(m-i)$ variables have value 1.
4. $B(l_p)$ denotes the state of b_p corresponding to the label l_p of p.

Multi Label Generic Cuts (MLGC) - Transformation

1. We explain the transformation for 2-cliques. The technique generalizes to k-cliques.
2. For the 2-ary clique potential function $W_{\{p,q\}}(l_p, l_q)$ defined on labels of pixels in clique $\{p,q\}$, we introduce a $2m$-ary function $W^b_{\{p,q\}}\left(B(l_p), B(l_q)\right): \mathfrak{B}^{2m} \rightarrow \mathfrak{R}$.
3. The domain of W^b is 2^{2m} states of the $2m$ binary variables. Of these only m^2 correspond to labels of pixels p and q. These states are referred to as *feasible* states, remaining states are called *infeasible*.
4. Value of W^b is equal to W for feasible states and set to infinity for infeasible states.
5. This transformation ensures that the minimum as well as minimum energy state for the $2m$-ary function W^b is the same as that of the original multi-label problem.

Submodularity Preservation

If the original multi-label clique potential W is submodular, then the binary label clique potential function W^b is also submodular.

Example

Step 1: Input

Labeling	Pot.	Labeling	Pot.	Labeling	Pot.	Labeling	Pot.	Labeling	Pot.	Labeling	Pot.
$W(a,a,a)$	50	$W(a,b,c)$	130	$W(b,a,b)$	120	$W(b,c,a)$	120	$W(c,a,c)$	160	$W(c,c,b)$	110
$W(a,a,b)$	110	$W(a,c,a)$	110	$W(b,a,c)$	150	$W(b,c,b)$	100	$W(c,b,a)$	140	$W(c,c,c)$	60
$W(a,a,c)$	140	$W(a,c,b)$	130	$W(b,b,a)$	90	$W(b,c,c)$	90	$W(c,b,b)$	120		
$W(a,b,a)$	80	$W(a,c,c)$	120	$W(b,b,b)$	70	$W(c,a,a)$	150	$W(c,b,c)$	110		
$W(a,b,b)$	100	$W(b,a,a)$	100	$W(b,b,c)$	100	$W(c,a,b)$	170	$W(c,c,a)$	130		

Step 2: Convert to binary function

Labeling	Pot.	Labeling	Pot.
$W^b(011011011)$	50	$W^b(001001000)$	100
$W^b(011011001)$	110	$W^b(001000011)$	120
$W^b(011011000)$	140	$W^b(001000001)$	100
$W^b(011001011)$	80	$W^b(001000000)$	90
$W^b(011001001)$	100	$W^b(000011011)$	150
$W^b(011001000)$	130	$W^b(000011001)$	170
$W^b(011000011)$	110	$W^b(000011000)$	160
$W^b(011000001)$	130	$W^b(000001011)$	140
$W^b(011000000)$	120	$W^b(000001001)$	120
$W^b(001011011)$	100	$W^b(000001000)$	110
$W^b(001011001)$	120	$W^b(000000011)$	130
$W^b(001011000)$	150	$W^b(000000001)$	110
$W^b(001001011)$	90	$W^b(000000000)$	60
$W^b(001001001)$	70		

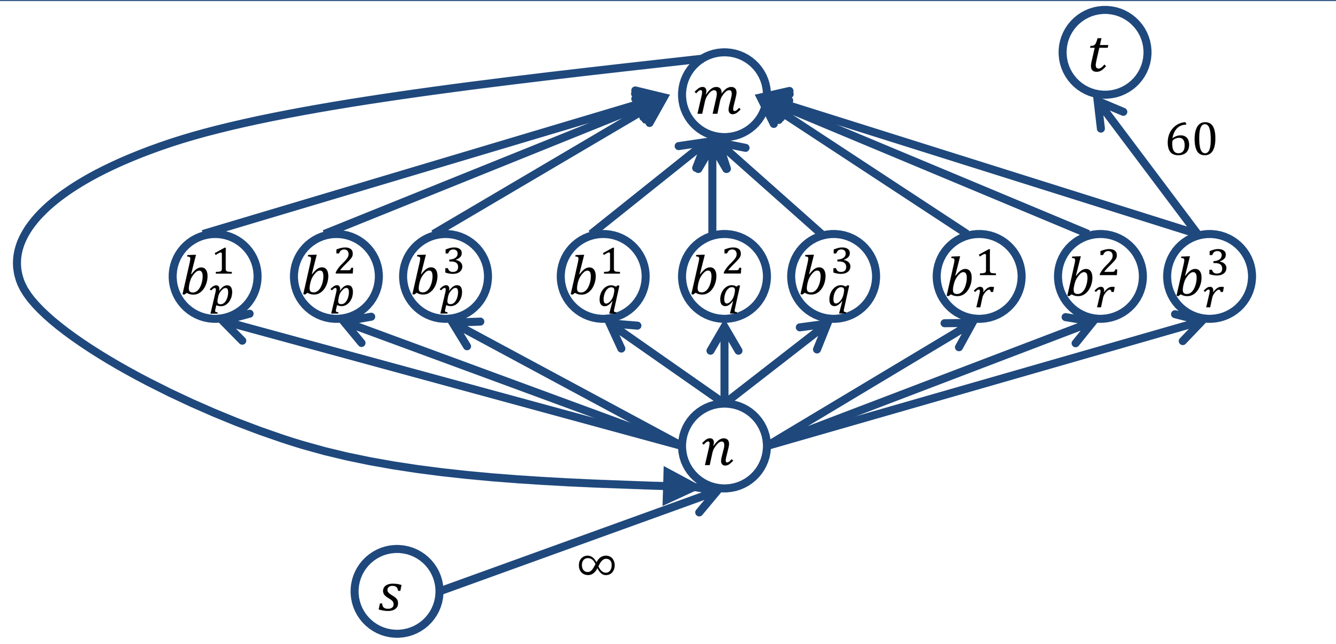
All unary costs are zero.

Step 3: Reparametrize to make $W^b(0,0,\ldots,0) = 0$

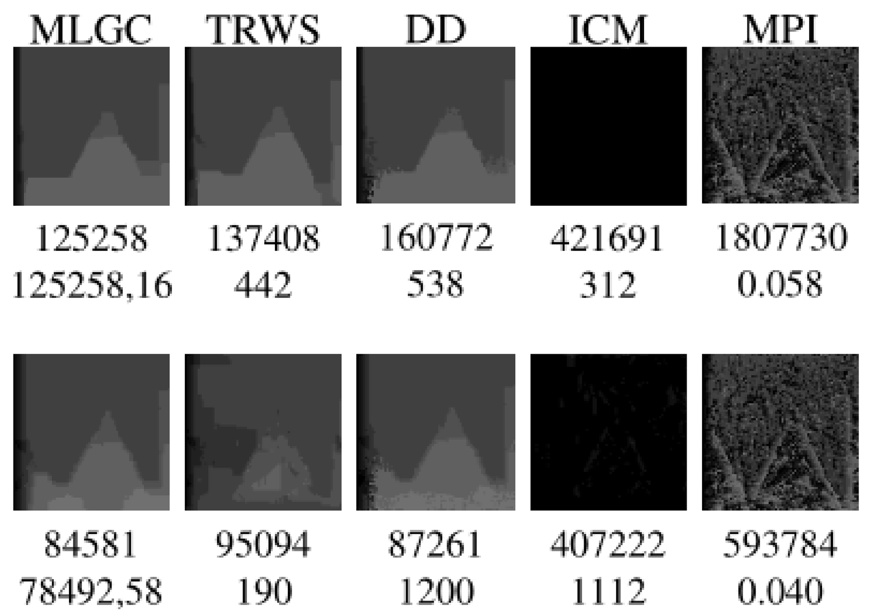
Labeling	Pot.	Labeling	Pot.
$W^b(011011011)$	50	$W^b(001001000)$	40
$W^b(011011001)$	110	$W^b(001000011)$	120
$W^b(011011000)$	80	$W^b(001000001)$	100
$W^b(011001011)$	80	$W^b(001000000)$	30
$W^b(011001001)$	100	$W^b(000011011)$	150
$W^b(011001000)$	70	$W^b(000011001)$	170
$W^b(011000011)$	110	$W^b(000011000)$	100
$W^b(011000001)$	130	$W^b(000001011)$	140
$W^b(011000000)$	60	$W^b(000001001)$	120
$W^b(001011011)$	100	$W^b(000001000)$	50
$W^b(001011001)$	120	$W^b(000000011)$	130
$W^b(001011000)$	90	$W^b(000000001)$	110
$W^b(001001011)$	90	$W^b(000000000)$	0
$W^b(001001001)$	70		

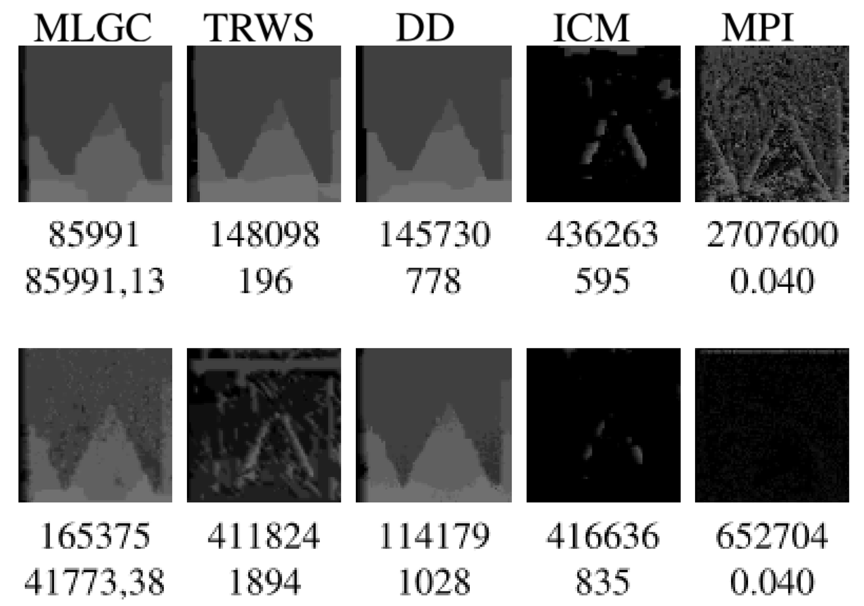
$D(b_r^3(0))$ is now 60

Flow Graph



Results





- Each row shows disparity computation using different clique potential (submodular as well as non-submodular)
- Number below each figure shows primal, dual and time taken in seconds.

MLGC runs order of magnitude faster with superior visual quality even for non-submodular clique potentials

References

1. C. Arora, S. Banerjee, P. Kalra, and S.N. Maheshwari. Generic Cuts: An efficient optimal algorithm for submodular MRF-MAP problems with higher order cliques. In ECCV, 2012.
2. C. Arora , S. Banerjee, P. Kalra, and S.N. Maheshwari. Fast Approximate Inference in Higher Order MRF-MAP Labeling Problems. In CVPR, 2014.
3. V. Kolmogorov and C. Rother. Minimizing nonsubmodular functions with graph cuts-a review. In TPAMI, 2007.
4. C. Rother, P. Kohli, W. Feng, and J. Jia. Minimizing sparse higher order energy functions of discrete variables. In CVPR, 2009.
5. Woodford, O., Torr, P., Reid, I., Fitzgibbon, A.: Global stereo reconstruction under second order smoothness priors. In CVPR, 2008.
6. H. Ishikawa. Exact Optimization for Markov Random Fields with Convex Priors. In TPAMI 2003.
7. Ishikawa, H.: Transformation of general binary MRF minimization to the first-order case. In TPAMI 2011.
8. **[DD]** N. Komodakis and N. Paragios. Beyond pairwise energies: Efficient optimization for higher-order MRFs. In CVPR, 2009.
9. **[MPI]** D. Koller and N. Friedman. Probabilistic Graphical Models: Principles and Techniques. Adaptive Computation and Machine Learning. MIT Press, 2009.
10. **[TRWS]** V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. In TPAMI, 2006.
11. **[ICM]** J. Besag. On the statistical analysis of dirty pictures. In Journal of the Royal Statistical Society, 1986.