GENERIC CUTS: AN EFFICIENT ALGORITHM FOR OPTIMAL INFERENCE IN HIGHER ORDER MRF-MAP

Chetan Arora, Subhashis Banerjee, Prem Kalra, S.N. Maheshwari

Indian Institute of Technology Delhi, INDIA

MRF-MAP INFERENCE

Cost of assigning labeling $l_{\mathcal{P}}$ at set of pixels $\mathcal{P}$

Cost of assigning label l_p at pixel p

Cost of assigning labeling $l_{\boldsymbol{c}}$ at clique $\boldsymbol{c}$

- $$l_{\mathcal{P}}^{*} = argmin_{l_{\mathcal{P}}} E(l_{\mathcal{P}}) = \sum_{p} D_p(l_p) + \sum_{\boldsymbol{c}} W_{\boldsymbol{c}}(l_{\boldsymbol{c}})$$

MRF-MAP INFERENCE

Can be solved optimally in polynomial time, when $|\mathcal{L}| = 2$, $|c| = 2$ and W_c is submodular

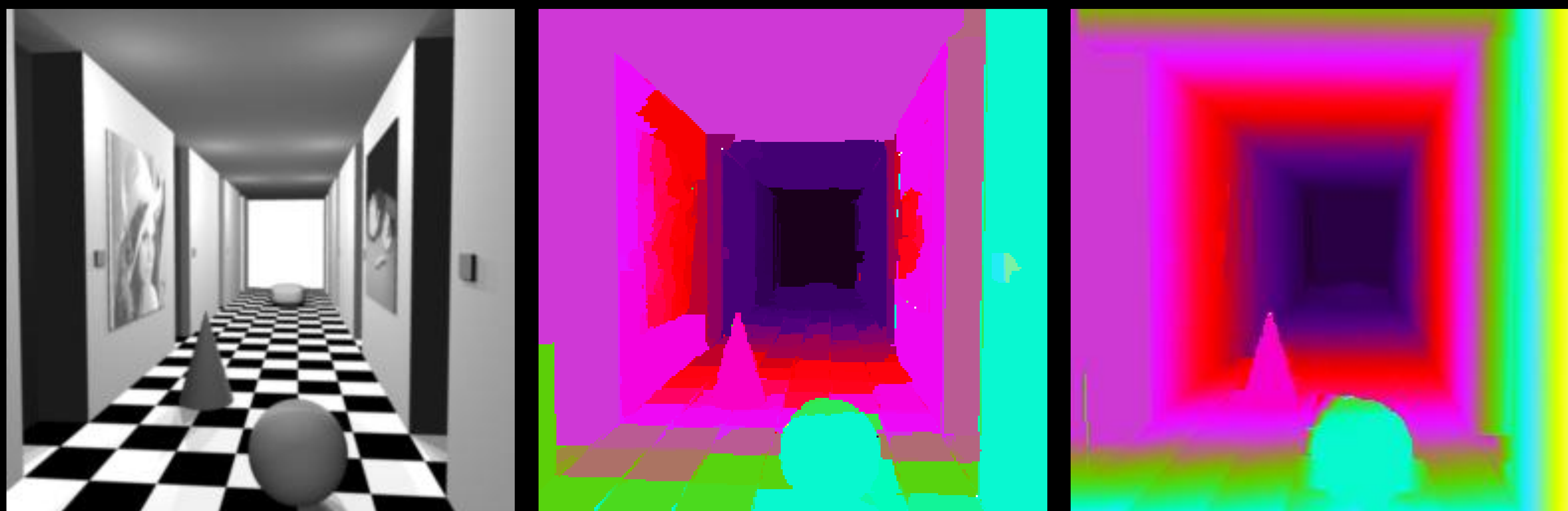
- $$l_{\mathcal{P}}^* = argmin_{l_{\mathcal{P}}} E(l_{\mathcal{P}}) = \sum_p D_p(l_p) + \sum_{\boldsymbol{c}} W_{\boldsymbol{c}}(l_{\boldsymbol{c}})$$

HIGHER ORDER MRF – WHY?

Encoding complex prior knowledge

Smooth gradients ⇒ $W_c = W_{var}(l_p, l_a, l_r) = l_p - 2l_a + l_r$

Input | Disparity – 2 Clique | Disparity – Smooth Gradient



Woodford et al., PAMI 2009

POSSIBLE INFERENCE METHODS

- Reduction to 2-clique, followed by QPBO (Ishikawa, CVPR 2009; Rother et al., CVPR 2009)
- Problem decomposition + subgradient (Komodakis and Paragios, CVPR 2009)
- LP-relaxation: e.g. Cutting-plane (Sontag et al., NIPS 2007)
- Iterated Conditional Modes (ICM)
- . . .

Rother, INRIA Summer School 2010

- New Reduction Techniques - Gruber, Boros and Zabih, ICCV 2011; Kahl and Strandmark, ICCV 2011)

LIMITATIONS

- Reduction
  - Do not preserve submodularity. Inference using QPBO leaves many nodes unlabeled even when the original higher order function is submodular
  - Increase in the size of created graph.

- BP/Dual Decomposition/ICM
  - Convergence only in the limit. No guarantee on number of steps
  - Observed slow convergence with increase in image size.

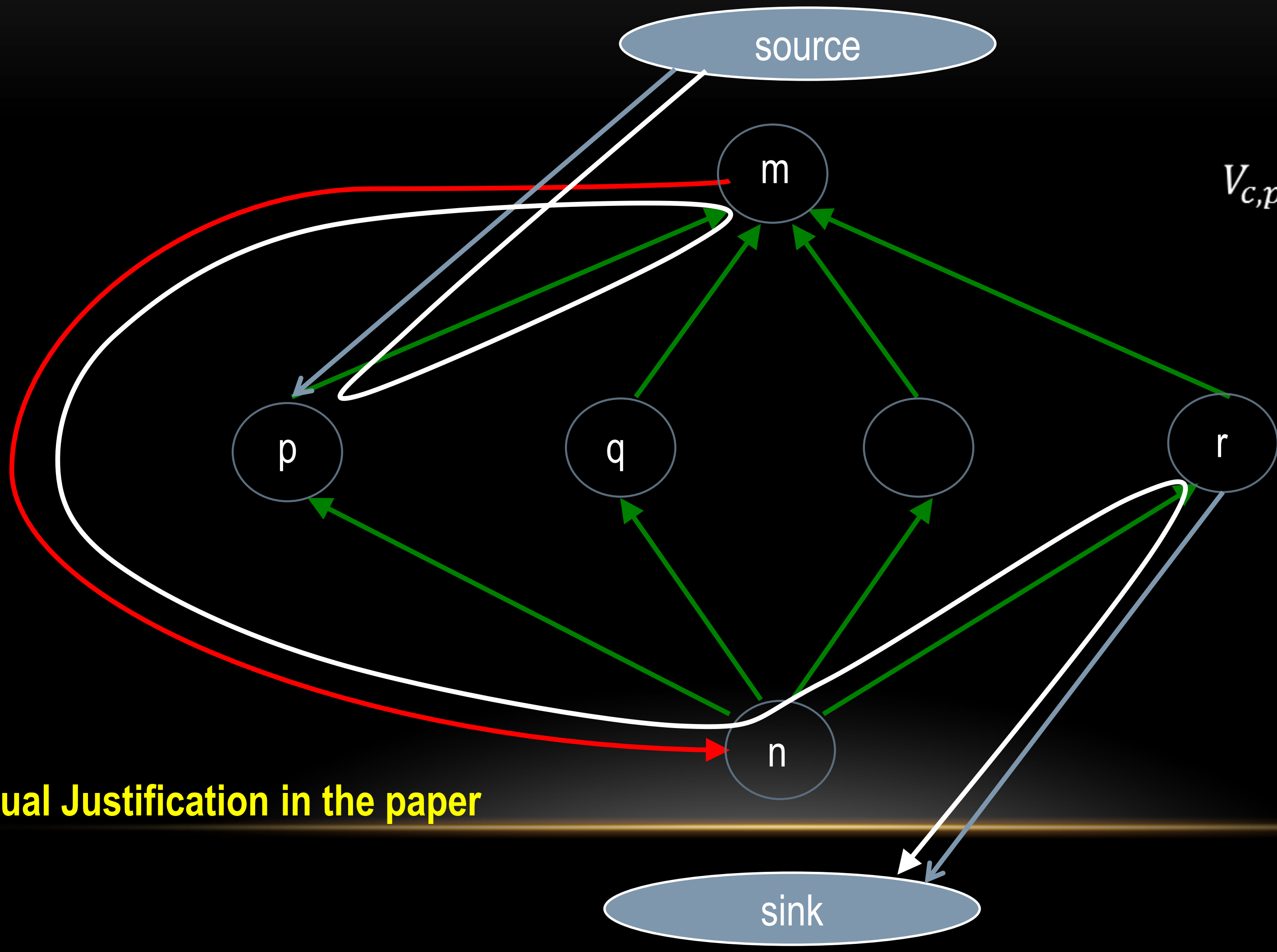
MAIN CONTRIBUTION

Need for development of direct algorithms for handling higher order clique problems

We show 2-label higher order MRF-MAP can be formulated as maxflow problems

When the clique potentials are submodular: Can be solved optimally and efficiently.

GADGET

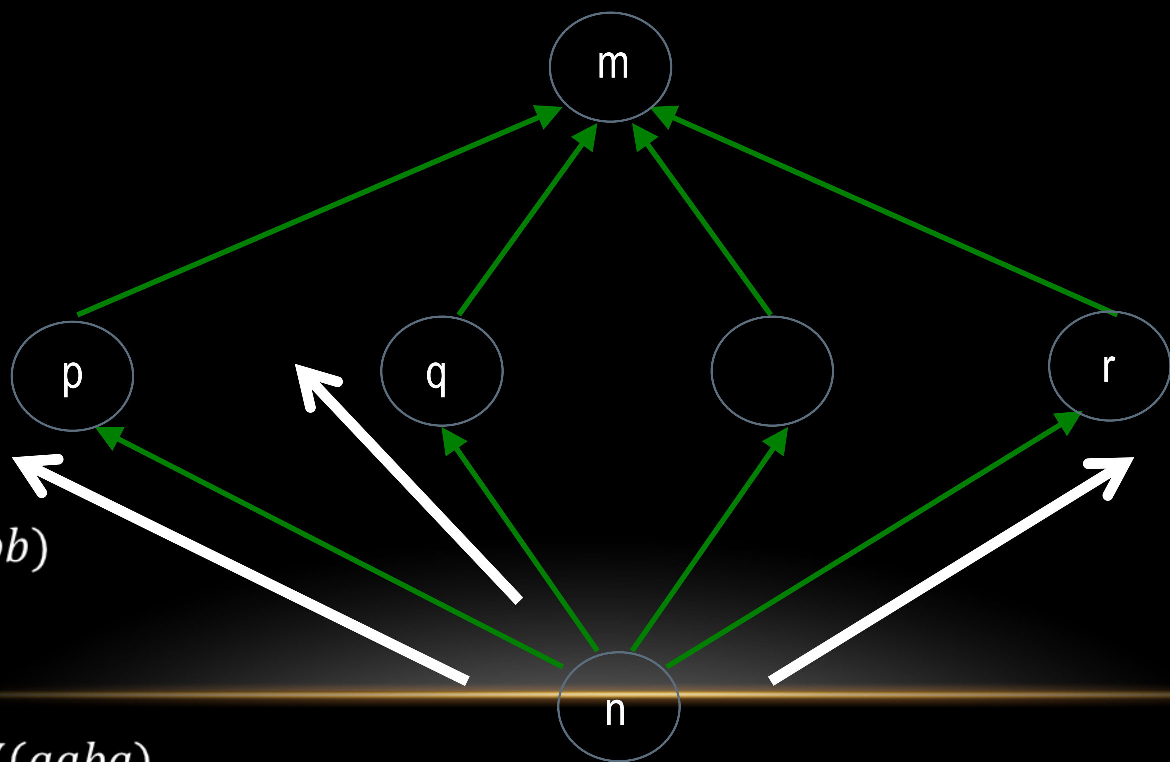


$$V_{c,p,a} = f^a_{np} - f^a_{pm}$$

Primal Dual Justification in the paper

EDGE CAPACITY – DUAL FEASIBILITY CONSTRAINT

$$\sum_{p \in \mathrm{P}: l_c^p = a} V_{c,p,a} \leq W_c(l_c) \quad \Rightarrow \quad \sum_{p \in \mathrm{P}: l_c^p = a} f_{np}^a - f_{pm}^a \leq W_c(l_c)$$



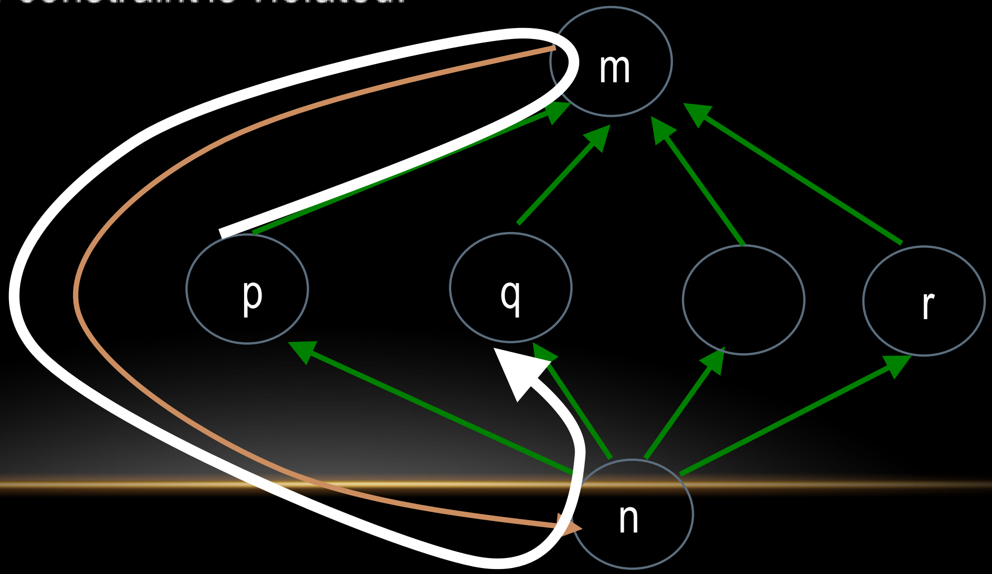
$f_{np} + f_{nq} \leq W(aabb)$

$f_{np} + f_{nq} + f_{nr} \leq W(aaba)$

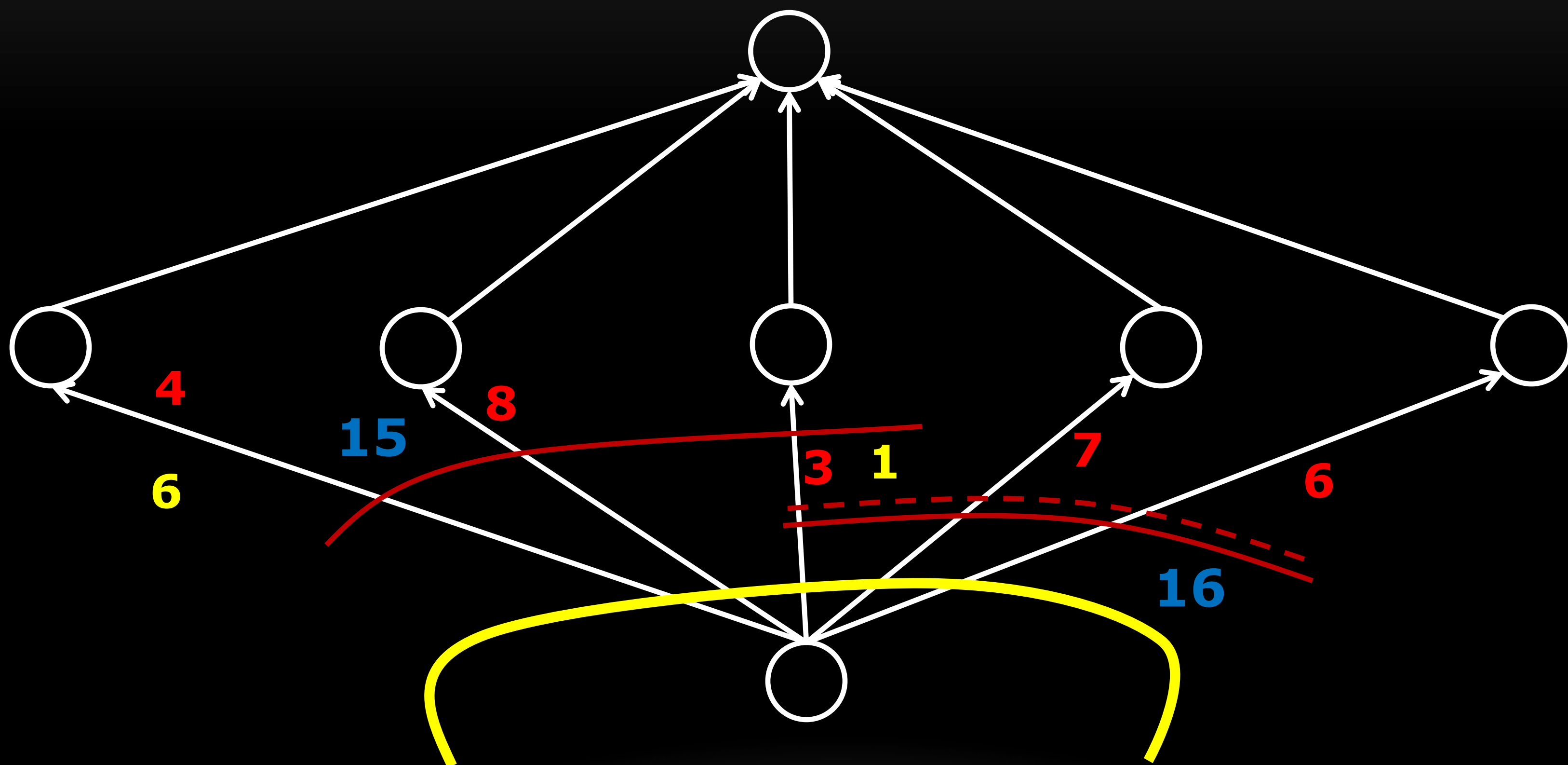
EDGE CAPACITY – DUAL FEASIBILITY CONSTRAINT

- $$Slack = W_c(l_c) - \sum_{p \in P : l_c^p = a} (f_{np}^a - f_{pm}^a)$$
- **Residual capacity** of pair of conjugate edges is equal to the ***minimum of slacks of all dual feasibility constraints*** in which it participates.
- Conditions Apply ! - Flow can be sent from node p to q, even if the residual capacity is zero, as long as no dual feasibility constraint is violated.

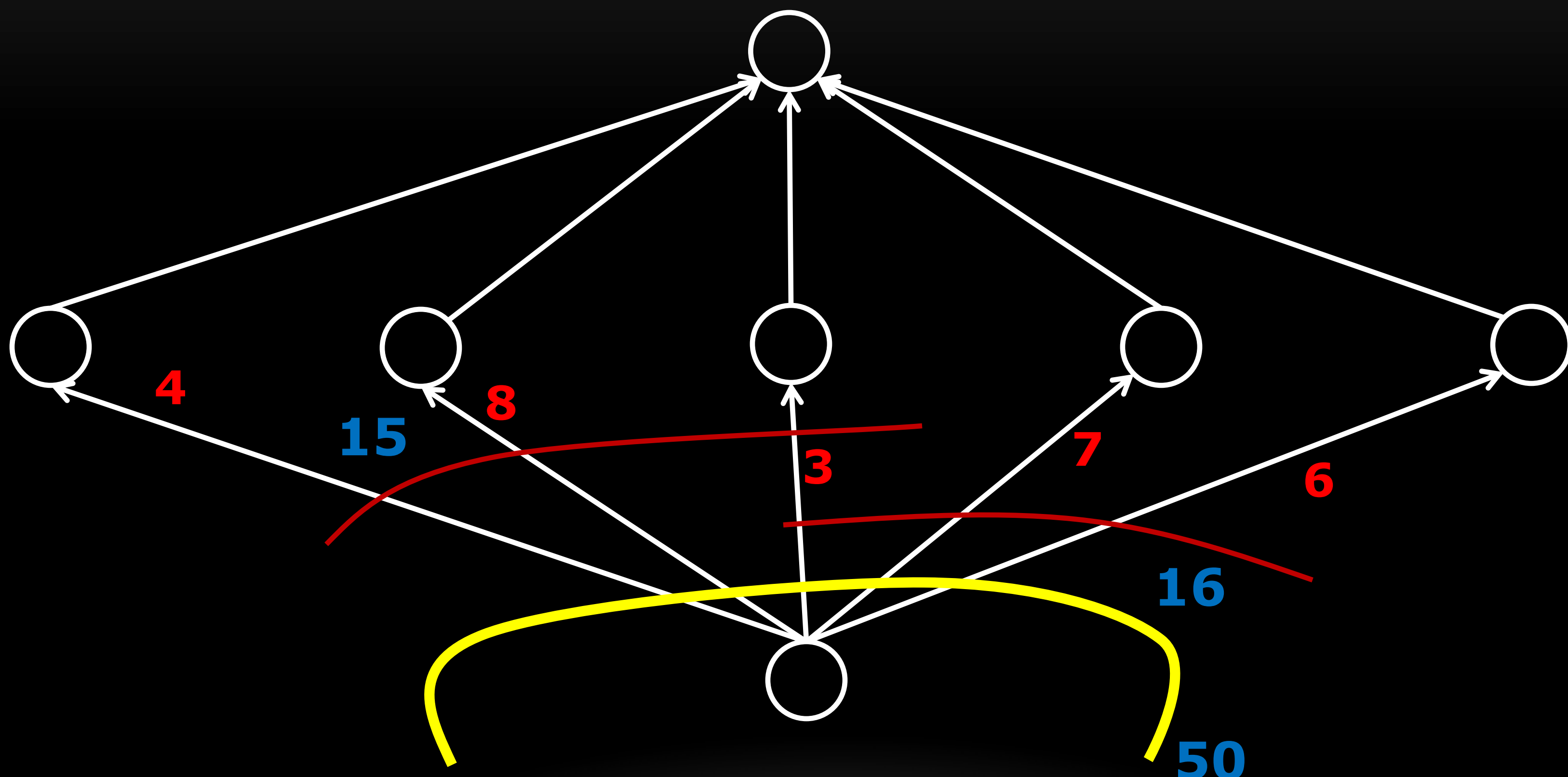
$$f_{np} + f_{nq} = W(aabb)$$



HOW DOES SUBMODULARITY EFFECTS US - MAXFLOW?



HOW DOES SUBMODULARITY EFFECTS US - MINCUT?



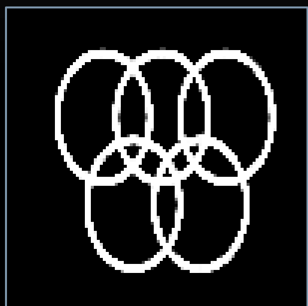
COMBINATORIAL PROPERTIES OF THE FRAMEWORK

- Complexity of the max-flow algorithm comes out to be $O(2^k k^2 n^3)$.

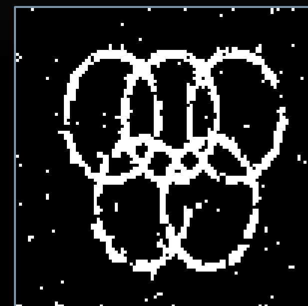
- All combinatorial results for maxflow carry over, e.g., the path length after each augmenting flow is non-decreasing.

- We can show, when costs are submodular, in the flow graph corresponding to the dual optimization problem, max flow is equal to min cut and corresponding primal and dual solutions are optimal

COMPARISON


Ground Truth

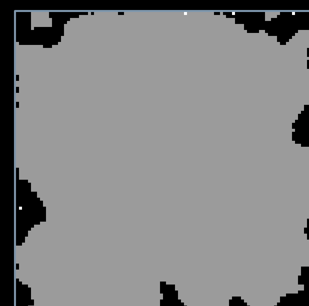

Noisy Input

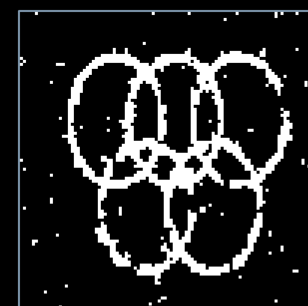

DD


MPI


ICM


TRWS


IQ

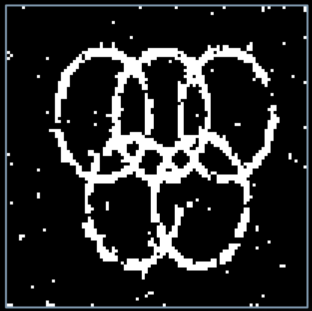

GC

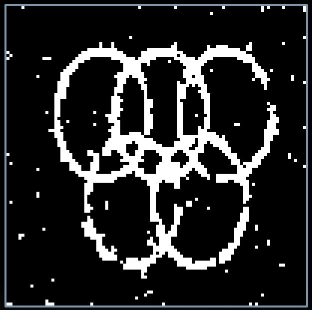
DD / MPI / ICM / TRWS: drwn.anu.edu.au/

IQ: www.f.waseda.jp/hfs/software.html

COMPARISON -ENERGY , CLIQUE SIZE=4

Image Size	DD	MPI	TRWS	IQ	GC
50X50	219496	267018	221815	232197	219161
100X100	845883	1056459	848071	919249	843056

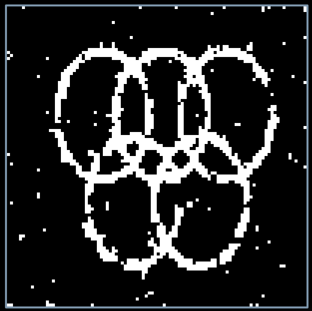

DD


GC

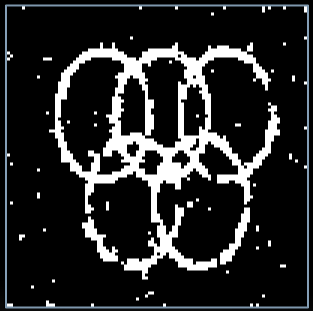

MPI

COMPARISON -TIME (MS) , CLIQUE SIZE=4

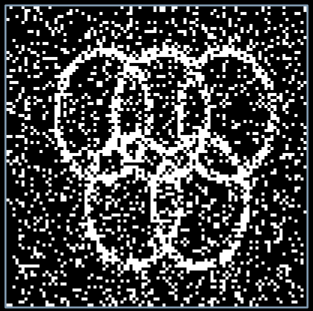
Image Size	DD	MPI	TRWS	IQ	GC
50X50	2084	8	7769	86	4
100X100	14202	41	107873	493	14



DD



GC



MPI

COMPARISON – CLIQUE SIZE V/S TIME (MS)

Image Size	Clique Size	DD	GC	IQ
100×100	4	15081	14	519
100×100	6	36683	103	7398
50×50	8	21290	159	32604
50×50	9	44872	435	206756
50×50	10	88577	1102	DNR
50×50	12	400543	7125	DNR

GENERALIZATIONS

- Can be extended directly to find approximate solutions for non-submodular functions with uniform labeling costs zero.

- Can be extended by duplicating nodes to handle general non-submodular (similar to what is done in QPBO for 2-clique non-submodular problems)

- Can be extended to handle sparse clique potential efficiently.

SUMMARY

- Proposes a new gadget to formulate 2-label multi-clique MRF-MAP as generalized flow problem.
- Guarantees **optimal** solution when the clique potential is **submodular**. Most current methods give only approximate solutions.
- Guaranteed to converge in $O(2^k k^2 n|c|^2)$ in worst case, where k is size of clique, n is number of pixels and c number of cliques. Exponentially faster than many current inference methods.