Solving Large-Scale Submodular Labeling Problems

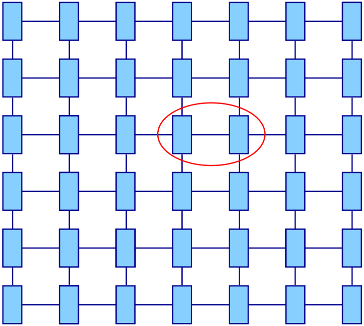
Overview of I.Kovtun's work

Bogdan Savchynskyy

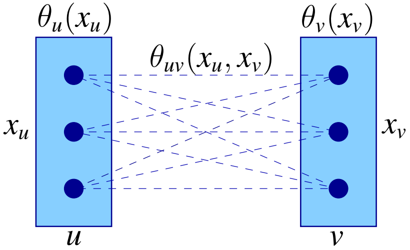
Heidelberg University,
Germany

MRF Energy Minimization

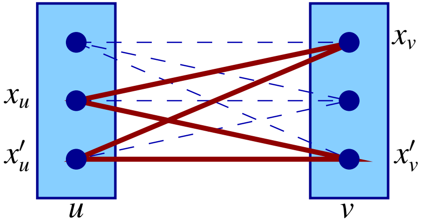
$$\min_{x \in \mathcal{X}} \theta(x) := \min_{x \in \mathcal{X}} \sum_{v \in \mathcal{V}} \theta_v(x_v) + \sum_{uv \in \mathcal{E}} \theta_{uv}(x_u, x_v)$$



graph $(\mathcal{V}, \mathcal{E})$



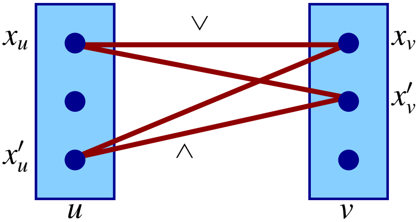
Submodular Problems



- Ordering of $\mathcal{X}_v$: $x_v \geqslant x'_v$
- $\theta_{uv}(x_u, x_v) + \theta_{uv}(x'_u, x'_v) \leqslant \theta_{uv}(x_u, x'_v) + \theta_{uv}(x'_u, x_v)$
- Example: $\theta_{uv} = \phi(x_u - x_v)$ and ϕ-convex ($|\cdot|$, $(\cdot)^2$)

$\Rightarrow \mathcal{X}^4$ operations to check submodularity in pairwise case.
(can be done in with $\mathcal{X}^2$ operations, indeed)

Submodular Problems: Generalization



- Ordering of $\mathcal{X}_v$: $x_v \geqslant x'_v$
- $\vee$ - nodewise maximum
- $\wedge$ - nodewise minimum
- Works for any $\mathcal{A} \subseteq \mathcal{V}$: $x_\mathcal{A} \vee x'_\mathcal{A}$ or $x_\mathcal{A} \wedge x'_\mathcal{A}$
- Submodularity: $\theta_\mathcal{A}(x_\mathcal{A} \vee x'_\mathcal{A}) + \theta_\mathcal{A}(x_\mathcal{A} \wedge x'_\mathcal{A}) \leqslant \theta_\mathcal{A}(x_\mathcal{A}) + \theta_\mathcal{A}(x'_\mathcal{A})$
- Sufficient condition: $\mathcal{A}$ - factors
- Necessary and suffcient $\mathcal{A} = \mathcal{V}$.

Submodular Problems: Property Of Solutions

Theorem (Known fact in submodular optimization)

Let x^* and x' be any two minimizers of $\theta(x)$. Then $x^* \vee x'$ and $x^* \wedge x'$ are minimizers as well.

Proof.

By construction:

$\theta(x^*) \leqslant \theta(x^* \vee x')$ and (1)

$\theta(x') \leqslant \theta(x^* \wedge x')$. (2)

From (1)+(2): $\theta(x^*) + \theta(x') \leqslant \theta(x^* \vee x') + \theta(x^* \wedge x')$ (3)

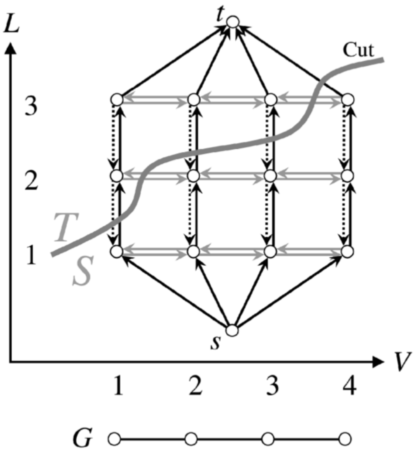
From submodularity: $\theta(x^*) + \theta(x') \geqslant \theta(x^* \vee x') + \theta(x^* \wedge x')$ (4)

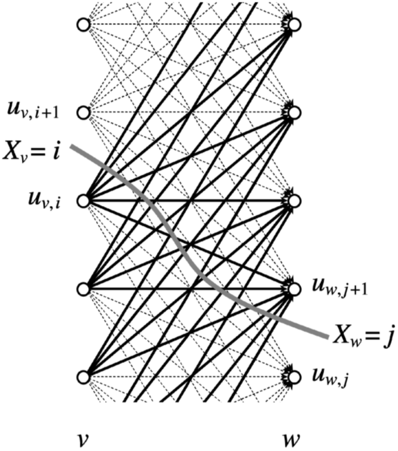
Comparing (3), (4) and using optimality of x^* and x' we get

$\theta(x^* \vee x') = \theta(x^* \wedge x') = \theta(x^*) = \theta(x')$

□

Solving with MinCut/MaxFlow





Number of edges grows as $|V||\mathcal{X}_v|^2$. For ℓ_1 norm only $|V||\mathcal{X}_v|$

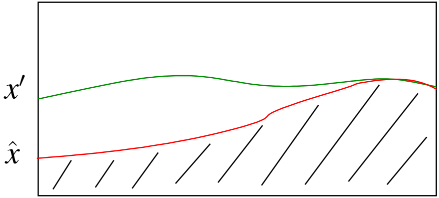
Ishikawa. Exact Optimization for Markov Random Fields with Convex Priors 2003

Schlesinger, Flach. Transforming an arbitrary MinSum problem into a binary one. 2006

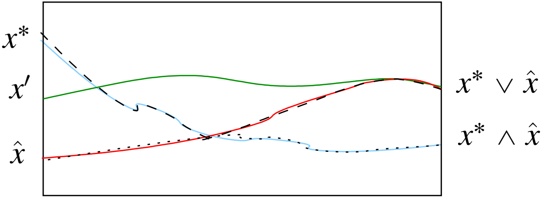
Partial Optimality

Theorem (Kovtun 2005)

Let $x' \in \mathcal{X}$, $\mathcal{X}^*(x') = \operatorname{Arg\,min}_{\substack{x \in \mathcal{X} \\ x \leqslant x'}} \theta(x)$ *and* $\hat{x} = \wedge_{x \in \mathcal{X}^*(x')} x$.
Then for all $x^* = \min_{x \in \mathcal{X}} \theta(x)$ *holds* $x^* \geqslant \hat{x}$.



Partial Optimality: Proof



Proof.

Let $x^* \not\geqslant \hat{x}$. Then $x^* \wedge \hat{x} \leqslant \hat{x}$ (1)

From submodularity: $\theta(x^*) + \theta(\hat{x}) \geqslant \theta(x^* \wedge \hat{x}) + \theta(x^* \vee \hat{x})$ (2)

From optimality of x^*: $\theta(x^*) \leqslant \theta(x^* \vee \hat{x})$ (3)

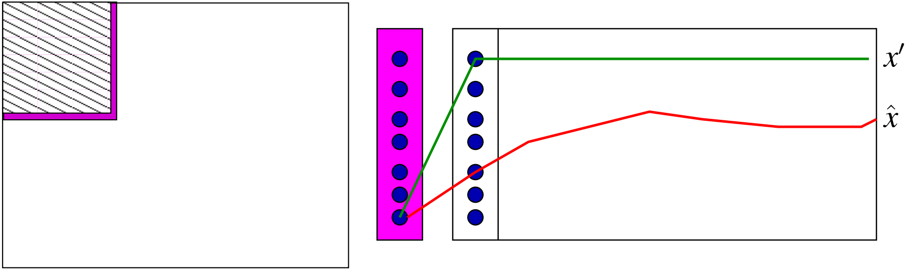
By construction and from (1) $\theta(\hat{x}) < \theta(x^* \wedge \hat{x})$. (4)

From (3)+(4): $\theta(x^*) + \theta(\hat{x}) < \theta(x^* \wedge \hat{x}) + \theta(x^* \vee \hat{x})$ - contradicts to (2)

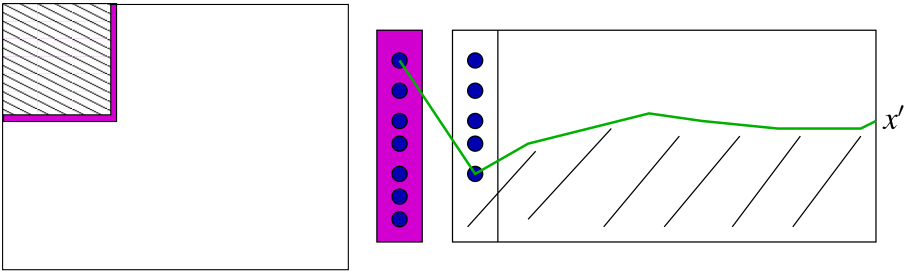
□

A symmetric theorem $\vee \leftrightarrow \wedge$ holds also

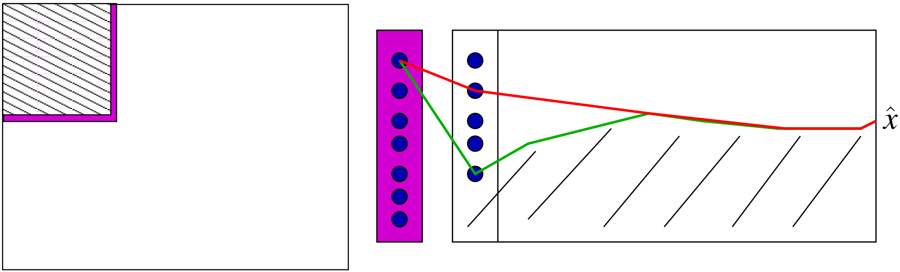
Partial Optimality: Practical Use



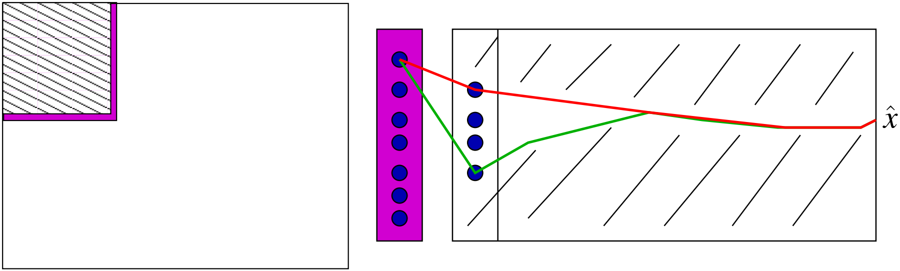
Partial Optimality: Practical Use



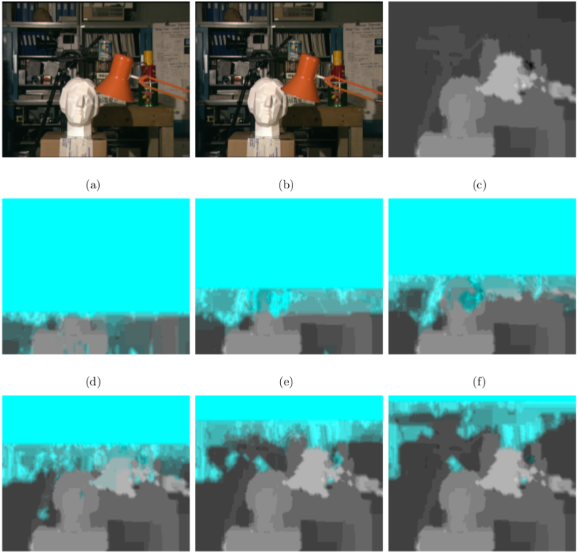
Partial Optimality: Practical Use



Partial Optimality: Practical Use



Partial Optimality: Experiment



Size $359 \times 253 \times 21$, ℓ_1 regularization. Needed: 80 Mb, Solved with 10 Mb.

Partial Optimality: Experiment

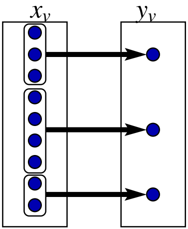
(a)

(b)

Size $996 \times 1478 \times 149$, ℓ_1 regularization. Needed: 8 Gb, Solved with 400 Mb.

Coarse-to-Fine Approach

- Max-flow problem size grows as $|\mathcal{V}||\mathcal{X}_v^2|$
- We addressed $|V|$
- Let us address $\mathcal{X}_v$
- What if $\mathcal{X}_v$ continuous (e.g. depth)?
- $\tilde{\theta}_w(y_w) = \min_{x_w \in y_w} \theta_w(x_w), \;\; x_w \in \mathcal{X}_w, \; w \in \mathcal{V} \cup \mathcal{E}$



Theorem (Raphael C. Coarse-to-fine dynamic programming 2001)

If $|y_v| = 1$ for all $v \in \mathcal{V}$ then $y_v* = \arg\min_{y \in \mathcal{Y}} \tilde{\theta}(y)$ is optimal for θ.

Theorem (Kovtun 2005)

If θ is submodular and $\mathcal{Y}$ is convex coarsening *then $\tilde{\theta}$ is submodular.*

[Zach *A Principled Approach for Coarse-to-Fine MAP Inference* CVPR 2014]

Coarse-to-Fine: Algorithm

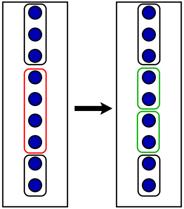
Theorem (Raphael C. Coarse-to-fine dynamic programming 2001)

If $|y_v| = 1$ *for all* $v \in \mathcal{V}$ *then* $y_{v*} = \arg\min_{y \in \mathcal{Y}} \tilde{\theta}(y)$ *is optimal for* θ.

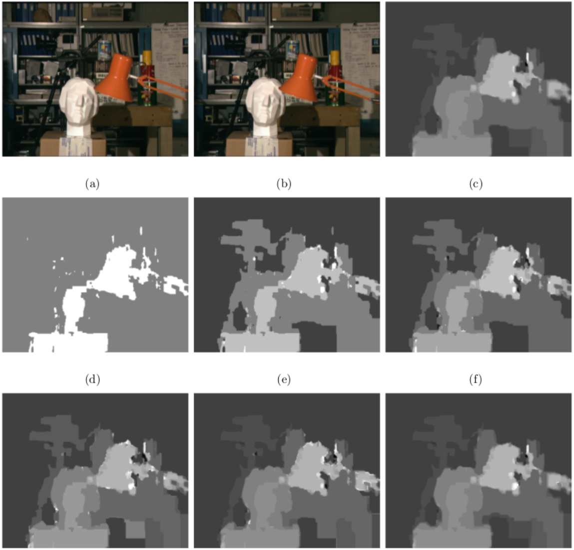
Theorem (Kovtun 2005)

If θ *is submodular and* $\mathcal{Y}$ *is* convex coarsening *then* $\tilde{\theta}$ *is submodular.*

- solve with max-flow;
- split y_v^* into 2 equal parts
- repeat until $|y_v^*| = 1$ for all $v \in \mathcal{V}$



Coarse-to-Fine: Experiment

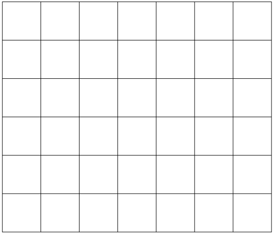


Size $359 \times 253 \times 21$, ℓ_2 regularization. Needed: 400 Mb, Solved with 50 Mb.

Generalizations

Everything is directly generalizable to higher order case.

What about Finite Algorithm for Partial Optimality Method?



- Partial Optimality Method works not for all Memory/Problem Connectivity ratio
- No guarantee for a fixed memory size.

BVZ-sawtooth(20)	80.0%	LB07-bunny-sml	15.6%	bone.n26c100	6.9%	bone_subxyz.n6c100	6.6%
BVZ-tsukuba(16)	72.8%	liver.n26c10	7.1%	bone.n6c10	8.8%	bone_subxyz_subx.n26c10	7.9%
BVZ-venus(22)	70.2%	liver.n26c100	5.3%	bone.n6c100	7.0%	bone_subxyz_subx.n26c100	6.6%
KZ2-sawtooth(20)	85.0%	liver.n6c10	7.2%	bone_subx.n26c10	6.6%	bone_subxyz_subx.n6c10	8.2%
KZ2-tsukuba(16)	69.9%	liver.n6c100	5.3%	bone_subx.n26c100	6.6%	bone_subxyz_subx.n6c100	6.6%
KZ2-venus(22)	75.8%	babyface.n26c10	29.3%	bone_subx.n6c10	6.3%	bone_subxyz_subxy.n26c10	11.3%
BL06-camel-lrg	2.0%	babyface.n26c100	30.9%	bone_subx.n6c100	6.3%	bone_subxyz_subxy.n26c100	9.5%
BL06-camel-med	2.3%	babyface.n6c10	35.4%	bone_subxy.n26c10	6.6%	bone_subxyz_subxy.n6c10	12.7%
BL06-camel-sml	4.6%	babyface.n6c100	33.7%	bone_subxy.n26c100	6.6%	bone_subxyz_subxy.n6c100	9.3%
BL06-gargoyle-lrg	6.0%	adhead.n26c10	0.3%	bone_subxy.n6c10	6.4%	abdomen_long.n6c10	1.7%
BL06-gargoyle-med	2.4%	adhead.n26c100	0.3%	bone_subxy.n6c100	6.3%	abdomen_short.n6c10	6.3%
BL06-gargoyle-sml	9.8%	adhead.n6c10	0.2%	bone_subxyz.n26c10	6.6%		
LB07-bunny-lrg	11.4%	adhead.n6c100	0.1%	bone_subxyz.n26c100	6.6%		
LB07-bunny-med	13.1%	bone.n26c10	8.7%	bone_subxyz.n6c10	6.6%		

[Shekhovtsov. *A distributed mincut/maxflow algorithm combining path augmentation and push-relabel*. IJCV 2012]

Distributed Max-Flow

[Shekhovtsov. *A distributed mincut/maxflow algorithm combining path augmentation and push-relabel*. IJCV 2012]

1. two operating modes:
   - sequential with restricted memory;
   - parallel distributed.
2. Kovtun's partial optimality implicitly included.
3. Code is available at http://cmp.felk.cvut.cz/~shekhovt/

References

- Ivan Kovtun. Phd Thesis: *Image segmentation based on sufficient conditions for optimality in NP-complete classes of structural labeling problems*, In Ukrainian. http://irtc.org.ua/image/people/kovtun
- Ivan Kovtun *Sufficient condition for partial optimality for (max,+) labeling problems and its usage*, Control Systems and Computers 2011
- Alexander Shekhovtsov. *A distributed mincut/maxflow algorithm combining path augmentation and push-relabel*. IJCV 2012
- Alexander Shekhovtsov. PhD Thesis *Exact and Partial Energy Minimization in Computer Vision*, 2013, http://cmp.felk.cvut.cz/~shekhovt/
- Victor Lempitsky and Yuri Boykov *Global Optimization for Shape Fitting* CVPR 2007