Energy Minimization under Constraints on Label Counts

Yongsub Lim (KAIST), Kyomin Jung (KAIST), Pushmeet Kohli (Microsoft Research)

ERC, August 26 2010

Results

1. We propose a novel approximation algorithm for energy minimization under all label count constraints.
2. We develop a fast heuristic algorithm to compute solutions under almost all label count, which we show to work well empirically.

Label Counts in Computer Vision

Constraints on label count is useful for many problems like the image segmentation, 3D image reconstruction, image denoising etc. One of concrete examples is that we can identify objects each of which has pre-defined size from images.

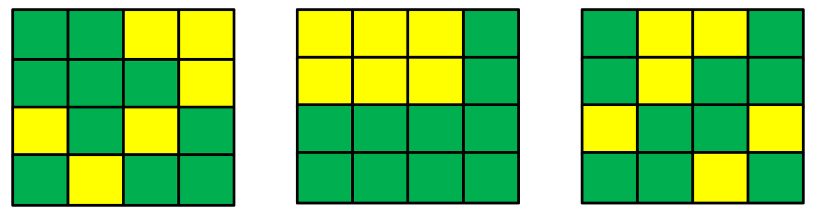
Setup

Minimizing energy functions defined on pixel-grid graphs.

$$H(\mathbf{x}) = \sum_{v \in V} \pi_v(x_v) + \sum_{(v,w) \in E} \pi_{vw}(x_v, x_w)$$

Label Count Constraint (LCC). Consider solutions so that the number of pixels having each label is fixed.

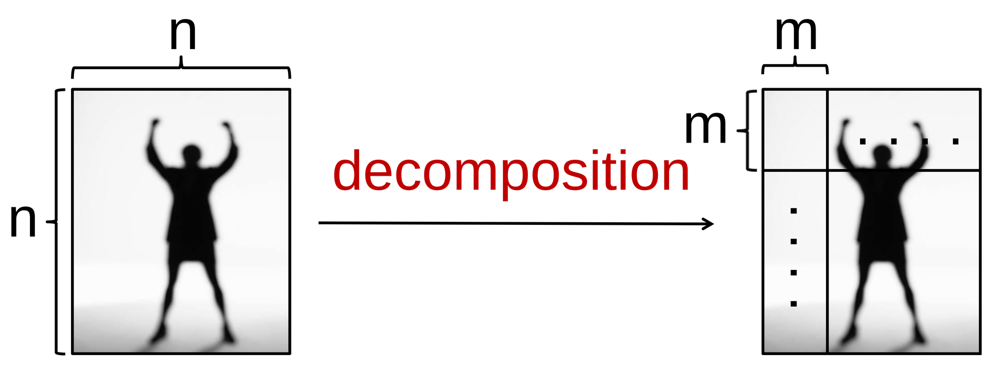
Some feasible solutions under LCC (6,10) in binary labeling.



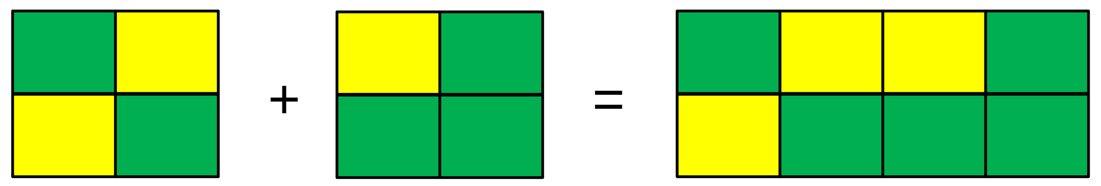
Decomposed Dynamic (DD)

Procedure

1. Decompose an image
2. For each decomposed, compute solutions for all label counts using dynamic programming.
3. Merge results of 2.



Merging Example:



(2,2) (1,3) (3,5)

Theorem. DD outputs ε-approximate solutions for all label count constraints. **DD** runs in time $O(Nk^{(1/\varepsilon)} (1/\varepsilon)^{2k+2} + N^k (1/\varepsilon)^2)$.

But DD is slow to apply to applications in practice.

Label Counts for Submodular MRF

Parametric maxflow algorithm (PP).

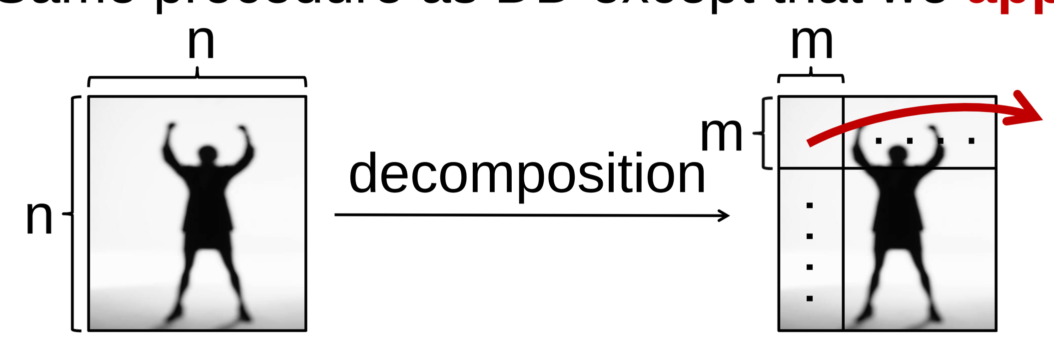
It deals with a parameterized energy function $H^\lambda(\mathbf{x}) = H(\mathbf{x}) + \lambda \sum_{v \in V} x_v$

Label counts at which the algorithm produces solutions are not controllable. Moreover, the number of those LCC is small.

Lemma [Kolmogorov, V., Boykov, Y., Rother, C, 2007].
For submodular pseudoboolean functions, the parametric maxflow algorithm produces exact solutions under some LCCs.

Decomposed Parametric (DP)

Same procedure as DD except that we apply PP to each subimage.



DD: dynamic programming
DP : parametric maxflow

DP runs fast and works well empirically.

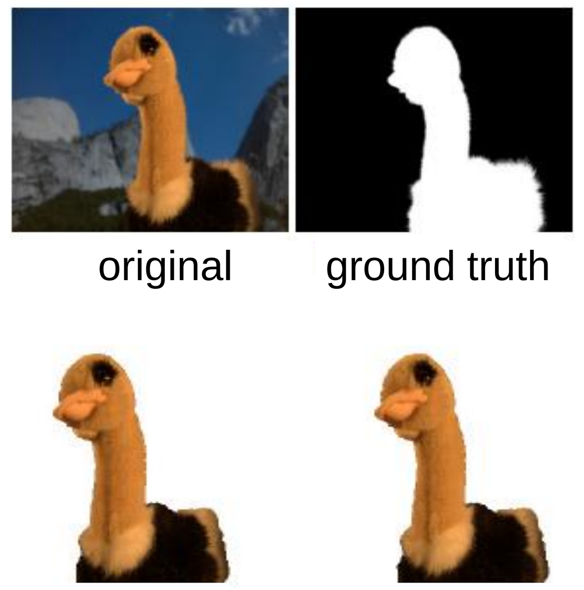
Experiments on Segmentation

Our Model:
Unary potentials are obtained by user given hints about the appearance of foreground and background segments. Pairwise potentials are defined as

$$\pi_{uv}(x_u, x_v) = |x_u - x_v| (\lambda_1 + \lambda_2 g(u,v))$$

where $g(u,v)$ is proportional to the distance of u and v's RGB colors, and λ_1, λ_2 are parameters of the model. We denote $\mu = \lambda_2 / \lambda_1$

DP produces results for almost all LCC, while PP is not.



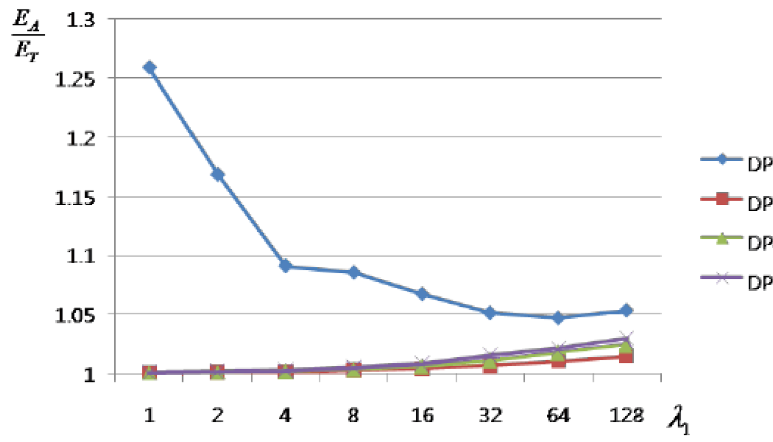
original ground truth

PP DP3

	$\mu = 20$		
λ_1	PP	DP_3	DP_5
1	0.2786	0.9828	0.9998
2	0.2552	0.9819	0.9997
4	0.2231	0.9795	0.9995
8	0.1862	0.9767	0.9994
16	0.1498	0.9736	0.9986
32	0.1164	0.9698	0.9970
64	0.0875	0.9650	0.9951
128	0.0642	0.9544	0.9925

(# of output LCCs) / (# of total LCCs)

DP has small error, even slightly smaller error than DD.



	$\mu = 10$		$\mu = 20$	
λ_1	DP_{20}	DD_{20}	DP_{20}	DD_{20}
1	1.0061	1.0079	1.0080	1.0104
2	1.0086	1.0109	1.0111	1.0140
4	1.0129	1.0157	1.0163	1.0199
8	1.0197	1.0232	1.0249	1.0295
16	1.0308	1.0360	1.0388	1.0456
32	1.0481	1.0492	1.0592	1.0611
64	1.0706	1.0713	1.0862	1.0877
128	1.1008	1.1021	1.1228	1.1253

The numbers are
(average energy of solutions by our algorithms) / (average energy of optimal solutions)

DP is much faster than DD, and reasonably slower than PP.

	Time	
	DD_{20}	DD_{25}
IM1	24m 41s	17m
IM2	32m 33s	29m 21s
IM3	24m 16s	16m 21s
IM4	25m 17s	17m 43s
IM5	27m 39s	21m 14s
IM6	26m 26s	19m 14s
IM7	33m 26s	25m 15s
IM8	27m 12s	22m 6s

	Time(seconds)				
	PP	DP_2	DP_3	DP_4	DP_5
IM1	8	10	18	24	27
IM2	11	23	31	42	53
IM3	51	22	28	35	42
IM4	4	11	15	18	19
IM5	23	42	48	63	75
IM6	12	38	42	52	62
IM7	26	44	55	76	94
IM8	17	42	46	58	70

DP performs well even on the binary image denoising problem.

original D_3 Werner's method [Werner et al. 2008]