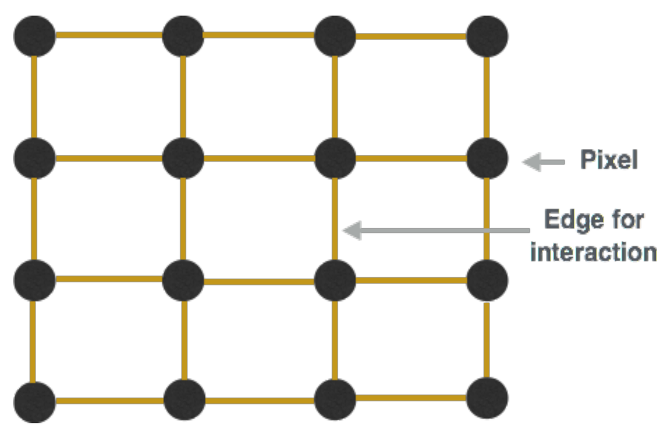
Truncated Max-of-Convex Models

Pankaj Pansari, Pawan Mudigonda
University of Oxford, The Alan Turing Institute

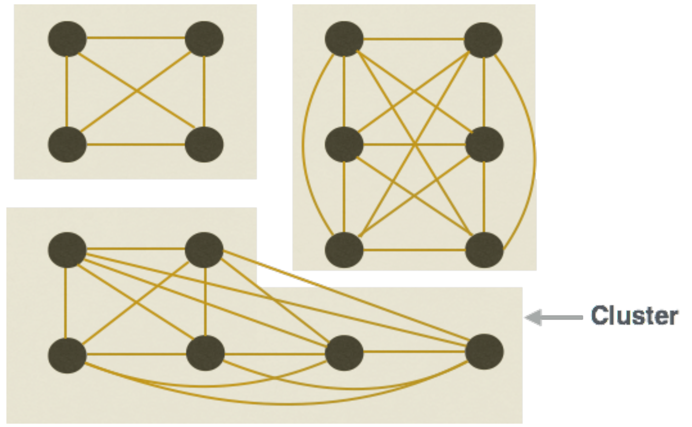
IEEE 2017 Conference on Computer Vision and Pattern Recognition

Contribution

We present a new class of higher-order Markov Random Fields (MRFs). We also design an efficient algorithm for maximum probability (MAP) inference in such models.



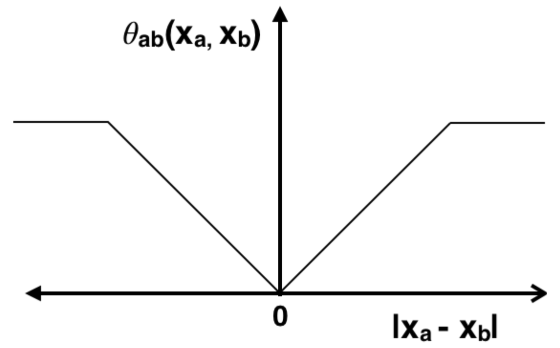
Pairwise MRF



Higher-order MRF

Introduction

Truncated convex models (TCMs) have been widely used in computer vision. In TCMs, the interaction potentials are truncated convex distances on pairs of variables:



Truncated linear distance

However, TCMs are often unable to capture useful image statistics because of the limited interactions they can represent. We present a higher-order model called truncated max-of-convex models (TMCMs) which are generalization of TCMs.

Model

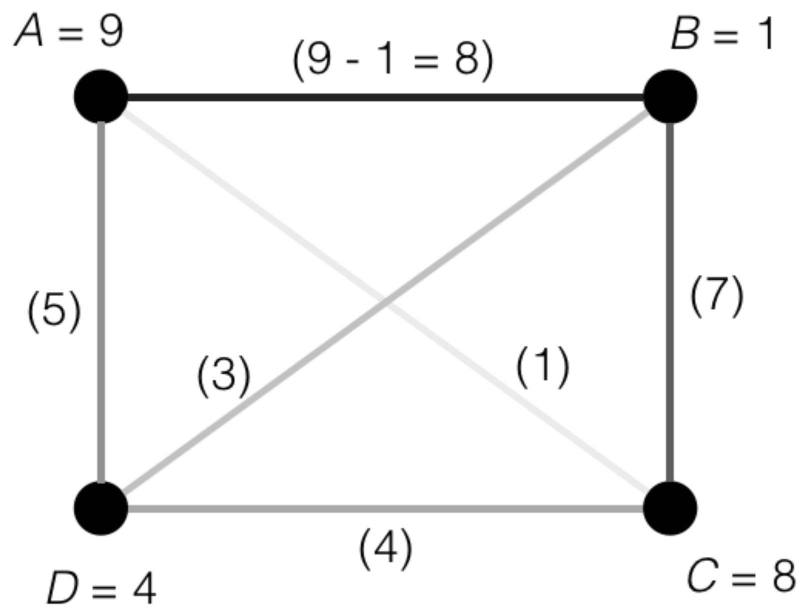
Let $\mathbf{X}$ represent the set of discrete random variables and C be the set of all cliques. We denote by $\mathbf{x_c}$ an assignment to all variables of a clique $\mathbf{x_c}$. Also, let $\theta_a(x_a)$ represent the unary potential for assigning $X_a = x_a$ and let $\theta_c(\mathbf{x_c})$ be the clique potential.

A TMCM specifies an energy function over the labelings x as:

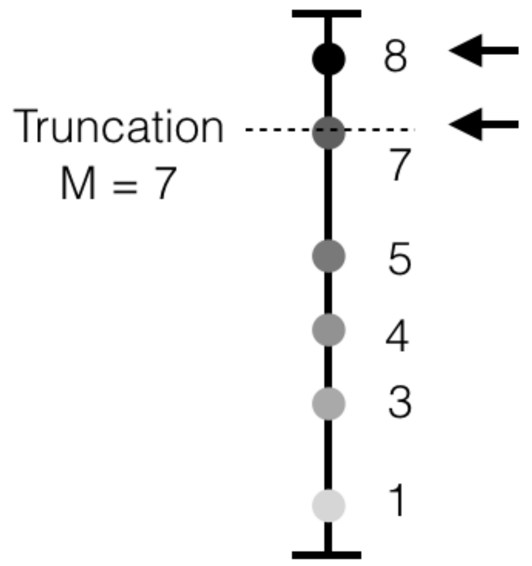
$$E(\mathbf{x}) = \sum_{X_a \in \mathbf{X}} \theta_a(x_a) + \sum_{c \in C} \theta_c(\mathbf{x_c})$$

Let $p(\mathbf{x_c})$ be an ascending-order sorted list of labels in $\mathbf{x_c}$, and let $p_i(\mathbf{x_c})$ by its i-th element. Given a convex distance function $d(.)$, a truncation factor M and an integer m, the clique potential $\theta_c(\mathbf{x_c})$ is defined as:

$$\theta_{\mathbf{c}}(\mathbf{x_c}) = \omega_{\mathbf{c}} \sum_{i=1}^{m} \min\{d(p_i(\mathbf{x_c}) - p_{c-i+1}(\mathbf{x_c})), M\}$$



Variables: *A, B, C, D*
Labels: 0-9



Let $\omega_c = 1$, $m = 2$, d = linear

$\theta(\mathbf{x_c})$ = 1. (max(8, 7) + max(7, 7))

= 14

Optimization

The MAP inference problem for TMCM is:

$$\min_{\mathbf{x} \in \mathbf{L}^N} E(\mathbf{x})$$

where $\mathbf{L}$ is the set of labels and N is the number of variables. This problem is *NP*-hard to solve exactly. We design a fast approximate inference algorithm based on range expansion and graph-cuts.

Results

Image Inpainting & Denoising

Input Image | Baseline 1 | Baseline 2 | **TMCM**

Stereo Correspondence

Ground Truth | Baseline 1 | Baseline 2 | **TMCM**

Project page: http://www.robots.ox.ac.uk/~pankaj/tmcm/
Email: pankaj@robots.ox.ac.uk