A New Multi-label Energy & Algorithm

Input: Set of observations \( \mathcal{P} \) (pixels, features, matches, ...)  
Set of candidate labels \( \mathcal{L} \) (objects, geometric models, ...)  
Output: Find a labeling \( f: \mathcal{P} \rightarrow \mathcal{L} \) that minimizes energy \( E(f) \).

We focus on two useful energies:

\[
E(f) = \sum_{p \in \mathcal{P}} D_p(f_p) + \beta \sum_{l \in \mathcal{L}} \mathbb{I}[f_p = l] 
\]

(label count)

\[
E(f) = \sum_{p \in \mathcal{P}} D_p(f_p) + \sum_{pq \in \mathcal{N}} V_{pq}(f_p, f_q) + \sum_{L \in \mathcal{L}} h_L \mathbb{I}[\mathbb{1} f_p \in L] 
\]

(smooth costs)

(label costs)

We extend \( \alpha \)-expansion algorithm [1] to handle label costs (2).
- Approximation guarantees
- C++/MATLAB code: http://vision.csd.uwo.ca/code/
- Faster code for special case (1) provided.

Energy (2) better than (1) for vision:
- smooth costs — because vision data is spatially correlated
- label subset costs — because labels may naturally form groups

We demonstrate many multi-model fitting applications in vision.

Why Label Costs?

Label costs are a regularizer that encourages ‘simple’ labelings.  
Energy (1) useful for geometric models like motion [4,5,6].

Has many interpretations:
- minimum description length (MDL) [8]
- information criteria [6]
- planning / operations research [3]
- sparse coding

Why label subset costs?

Encourages labelings to use fewer groups.

Special case of co-occurrence / context
- penalizes each ‘category’ of labels
- labels may form meaningful hierarchy

Relation to EM and \( K \)-means

\( EM = 'soft' \) labeling; better for data with high overlap.  
\( K \)-means = unregularized discrete energy (1) where \( \beta = 0 \).

Optimizing (2): \( \alpha \)-expansion w/ label costs

Suppose label set \( \mathcal{L} = \{ \alpha, \beta, \gamma \} \). Consider a single \( \alpha \)-expansion:

current labeling

expanding \( \alpha \) ...

optimal expansion

new labeling

Graph cut for standard expansion step:

Expansion step w/ label costs:
- optimal expansion changes! 
- encode costs \( h_\beta \) as high-order binary potentials [2], added in each expansion step.

Applications: multi-motion [4,5,6], multi-homography [6], auto segmentation [7,8], context-based segmentation, ...

example input

label costs only

smooth costs only

label + smooth

auto segmentation examples