Alphabet SOUP: A Framework for Approximate Energy Minimization

Stephen Gould, Fernando Amat, Daphne Koller

1 Overview

- **Alphabet SOUP provides a general purpose framework for approximate energy minimization of arbitrary energies that scales to accommodate available resources and problem complexity.**

- Many problems in computer vision can be modeled using conditional Markov random fields.

$$\begin{array}{ll} \text{minimize} & E(\mathbf{x}) = \sum_c \theta_c(\mathbf{x}_c) \\ \text{subject to} & x_i \in \mathbf{dom}(X_i) \qquad \forall X_i \in \mathcal{X} \end{array}$$

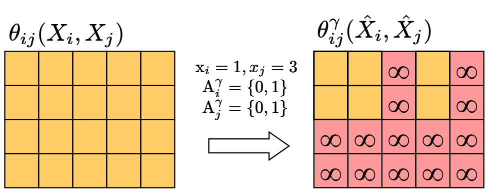
- Finding the MAP solution is NP-hard, so people resort to approximate techniques.
- Very efficient methods exist for problems with certain structure (e.g., α-expansion (Veksler, 1999) for submodular energies).
- Message passing algorithms (e.g., max-product or its convex variants) are the only general purpose approach (for problems with arbitrary energies and non-homogeneous domains).
- Static and dynamic theoretical guarantees when the inner loop of our method is exact.

2 γ-Expansion Moves

- For each i, let A_i^γ be a subset of the domain for variable X_i. Then a γ-expansion move maps each X_i from its current value x_i to a value in $\{x_i\} \cup A_i^\gamma$. This generalizes other approaches:
  - **α-expansion:** set $A_i = \{\alpha\}$ for all i.
  - **α-β swap:** set $A_i = \{\alpha\}$ for all $x_i = \beta$, $A_i = \{\beta\}$ for all $x_i = \alpha$, and empty otherwise.
  - **ICM:** set $A_i = \mathbf{dom}(X_i)$ for $i = k$ and empty otherwise.
  - **Fusion move:** set $x_i = x_i^0$ and $A_i = \{x_i^1\}$.

- We can find the γ-expansion move with minimum energy by minimizing $E^\gamma(\hat{x}; x) = \sum_c \theta_c^\gamma(\hat{x}_c)$ where

$$\theta_c^\gamma(\hat{x}_c) = \begin{cases} \theta_c(\hat{x}_c) & \text{if } \forall X_i \in \mathbf{X}_c : \hat{x}_i \in A_i^\gamma \cup x_i \\ \infty & \text{otherwise} \end{cases}$$



3 Alphabet SOUP Algorithm

start with arbitrary assignment $\mathbf{x}$
repeat
  for $k = 1, \ldots, K$
  find $\hat{\mathbf{x}}$ one γ_k-expansion-move away from $\mathbf{x}$
  if $E(\hat{\mathbf{x}}) < E(\mathbf{x})$ then $\mathbf{x} \leftarrow \hat{\mathbf{x}}$

4 Theoretical Guarantees

- **Definition (covering set):** A set of moves $\gamma_1, \ldots, \gamma_K$ is covering if for every $x_i \in \mathbf{dom}(X_i)$, there exists a γ_k such that $x_i \in A_i^k$.
- **Theorem (static):** Let $\gamma_1, \ldots, \gamma_K$ be a covering set of moves.* If $\mathbf{x}$ is a local optimum relative to $\gamma_1, \ldots, \gamma_K$, then $E(\mathbf{x})$ is within a constant factor of the optimal energy.
- **Theorem (dynamic):** By using the LP-dual message passing methods for the inner loop we can provide a global bound on the optimality gap. E.g., for the method of Globerson et al. (2007) the current assignment is within

$$\Delta = E(\mathbf{x}) - \sum_s \min_{x_s} \sum_{c \in N(s)} \delta_c^s(x_s)$$

of the optimal, where $\delta_c^s(x_s) = \min_{x_{c \setminus s}} \left\{ \{\beta_c^s(x_c)\}_{x_c \in \gamma}, \left\{ \frac{1}{|S(c)|} \theta_c(x_c) \right\}_{x_c \notin \gamma} \right\}$

*Assume $\theta_c(x_c) \geq 0$ for all cliques c, with equality only if there exists some γ_k such that $x_i \in A_i^k$ for all variables X_i in the clique.

5 Choosing the γ-expansion Moves

- Our method allows the selection of γ-expansion moves to be tailored to the problem at hand, e.g., **(static)** overlapping ordinal ranges, grouped low-energy configurations, or **(dynamic)** coordinate descent on vector-valued variables

6 Object Detection and Outlining

- **Task:** Detect object and localize using landmark based outline.
- **Model:** LOOPS (Heitz et al. (2008)).
- **γ-Expansion Moves**: Overlapping candidate landmark locations, sorted by unary potential (highest to lowest).

7 Image Completion and In-painting

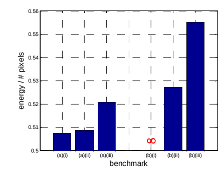
- **Task:** Fill in missing part of image using image patches.
- **Model:** Pairwise MRF of Komodakis and Tziritas (2006).
- **γ-Expansion Moves:** Non-overlapping sets of 250 patches. Patches for adjacent variables taken from adjacent locations in the image. This provides low energy pairs.

8 Cell Membrane Surface Reconstruction

- **Task:** Reconstruct surface of bacteria from tomography scans.
- **Model:** Discretized mesh: 3D orientation and 1D radial offset. Here the energy is regular, but α-expansion does not scale.
- **γ-Expansion Moves:** $(\theta_X, \theta_Y, \theta_Z, R)$ coordinates.

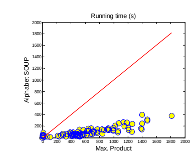
Problem	Energy per pixel (i)	Energy per pixel (ii)	Energy per pixel (iii)	Running time (s) (i)	Running time (s) (ii)	Running time (s) (iii)
40x43	**0.507**	0.509	0.521	39	1	**<1**
112x116	n/ a	**0.527**	0.555	∞	23	**5**

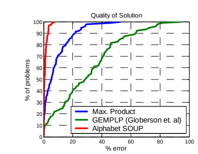
- (i) simultaneous optimization of all coordinates
- (ii) Alphabet SOUP over coordinate pairs
- (iii) Alphabet SOUP over individual coordinates
- ∞ indicates that problem could not fit into memory



9 Rosetta Protein Design

- **Task:** Find most stable configuration of amino-acids that give rise to a given 3D structure.
- **Model:** Protein design dataset from Yanover et al. (2007).
- **γ-Expansion Moves:** Subset of 50 values (modulo domain size)





10 Conclusion

- **Alphabet SOUP** is a flexible method for finding approximate solutions to the MAP inference problem for arbitrary energy functions. Our method is faster than standard max-product belief propagation, requires significantly less memory, and often produces lower energy solutions.
- Most interesting direction for further study is in providing more formal foundations for the choice of subsets used for the different variables.