

Graph-representable Submodular Functions

- What are submodular energy functions?
- Why are they important?
- Why do we want to represent submodular energy functions by graphs?
- How does one go about finding the class of graph-representable energy functions?
- What application will the solution of this problem have?

[1]

Many early vision problems can be naturally formulated in terms of energy minimization where the energy function has the following form:

$$E(x|\theta) = \theta_{const} + \sum_{s \in v} \theta_s(x_s) + \sum_{(s,t) \in \epsilon \theta_{st}(x_s, x_t)} \quad (1)$$

Set V usually corresponds to pixels; x_s denotes the label of pixel $s \in V$ which must belong to some finite set. For motion or stereo, the labels are disparities, while for image restoration they represent intensities. θ defines parameters of the energy: $\theta_s(\cdot)$ is a unary data penalty function, and $\theta_{st}(\cdot, \cdot)$ is a pairwise interaction potential. This energy is often derived in the context of Markov Random Fields: a minimum of E corresponds to a *maximum a-posteriori* (MAP) labeling \mathbf{x} .

In general, minimizing E is an NP-hard problem, so researchers have focused on approximate minimization algorithms. The most well-known techniques are graph cuts and belief propagation. Graph cuts was introduced in the 90's and showed a major improvement over previously used simulated annealing. To my knowledge, graph cuts are currently considered to be the most accurate minimization algorithm for energy functions arising in many vision applications, e.g stereo, image restoration, image segmentation, texture synthesis. In fact, for some functions it finds a global minimum.

However, graph cuts can be applied only to a limited class of energy functions. If a function falls outside this class then one has to use other techniques such as max-product belief propagation (BP). BP can be applied to any function of the form as in eqn. 1, but it has some drawbacks. First, it usually finds a solution with higher energy than graph cuts (in cases when graph cuts can be applied). Second, BP does not always converge - it often goes into a loop. [2]

Submodularity is a property of set functions. It relates to both convexity and concavity in differing ways. An excellent example of deploying sensors in a drinking water distribution network is provided. Submodularity is a property of set functions, i.e, functions $f : 2^V \rightarrow R$ that assign each subset $S \subseteq V$ a value of $f(S)$. Here V is a finite set, commonly called the *ground set*. Also, $f(\phi) = 0$, i.e, the empty set carries no value. One definition of submodularity, based on the notion of discrete derivative, is as follows:

Defintion 1.1 (Discrete derivative): For a set function $f : 2^V \rightarrow R$, $S \subseteq V$, and $e \in V$, let $\Delta_f(e|S) := f(S \cup \{e\}) - f(S)$ be the *discrete derivative* of f at S with respect to e .

Where the function f is clear from the context, we drop the subscript and simply write $\Delta(e|S)$.

Defintion 1.2 (Submodularity): A function $f : 2^V \rightarrow R$ is *submodular* if for every $A \subseteq B \subseteq V$ and $e \in V \setminus B$ it holds that

$$\Delta(e|A) \geq \Delta(e|B). \quad (2)$$

Equivalently, a function $f : 2^V \rightarrow R$ is *submodular* if for every $A, B \subseteq V$,

$$f(A \cap B) + f(A \cup B) \leq f(A) + f(B) \quad (3)$$

Submodular functions are functions with specific properties. We attempt to find the class of submodular energy functions which can be represented in terms of a graph. This will help us in knowing the class of Markov Random Field models which can be efficiently solved using graph cuts.

References

- [1] Vladimir Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(10):1568–1583, 2006.
- [2] Andreas Krause and Daniel Golovin. Submodular function maximization. *Tractability: Practical Approaches to Hard Problems*, 3:19, 2012.