International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Review on Energy Minimization via Fast Approximation Algorithms

Harsimran Kaur[1], Neetu Gupta[2]

[1]M.Tech, ECE Department, GIMET

[2]Assistant Professor, ECE Department, GIMET

Abstract: ***Label cost optimization proposes a new improvement in label cost function, improving existing moves of α-expansion algorithm and introducing some new moves for this algorithm. Instead of k-mean optimization algorithm new technique i.e. fast approximation can be used in α-expansion algorithm to optimize label cost function. In order to perform comparison, different metrics of energy minimization will be considered. An appropriate comparison will be drawn among proposed technique and previous well known techniques. The objective is to effectively optimize energies so that satisfactory image segmentation can be obtained (represented with different labels respective to different objects).***

Keywords: Energy minimization, labels, α-expansion, segmentation, NP-hard, local minimum, non-parametric histogram.

1. Introduction

Energy minimization is of strong practical importance to computer vision. Energy expresses our criteria for a good solution i.e. low energies are good, high energies are bad, independent of any algorithm. Even for low level vision problems we are confronted by energies that are computationally hard to minimize. Computer vision is full of 'labelling' problems cast as energy minimization. For example, the data to be labelled could be pixels, interest points, point correspondences, or mesh data such as from a range scanner. Depending on the application, the labels could be either semantic (object classes, types of tissue) or describe geometry/appearance (depth, orientation, shape, texture).

1.1 Label Cost Energy Minimization

There are many labelling problems for which the labels naturally form groups. In computer vision, a recent trend is the use of 'context' to resolve ambiguities in object recognition. The idea is that certain groups of labels are self-consistent because they tend to appear together, e.g. the {car, road, sky} labels all belong to the "outdoors" context, while {table, chair, wall} all belong to the "indoors" context. In computer graphics, one may wish to automatically classify the faces of a 3D mesh into semantic parts for the benefit of artists and animators. The part labels arm, tail, and wheel naturally belong to different groups based on their context (humanoid, quadruped, and vehicle). In operations research, facility location can be cast as a labelling difficulty, and hierarchical variants have been studied. All of these disparate labelling problems are similar from an optimization point of view. When labels are clearly grouped in a hierarchy, the costs in the energy are naturally structured.

1.2 How α-Expansion Works

The α-expansion algorithm performs local search using a powerful class of 'moves'. Given an initial labelling ^f and some particular label $\alpha \in L$, a α-expansion move gives each variable the following binary choice: either keep the current label ^fp, or switch to label α. Let Mα(^f) denote the set of all moves (labellings) that can be generated this way, in other words $M\alpha(\hat{f}) = \{f : fp \in \{\hat{f}p\} \cup \{\alpha\}\}$.

All variables are simultaneously allowed to keep their current label or to switch, so there are an exponential number of possible moves. For each choice of α, we must efficiently find the best possible move. In practice, this sub-problem is solved by casting it as a graph cut and using combinatorial algorithms to compute the optimal binary configuration [8].Because a graph cut finds the best move from an exponential number of possibilities, the α-expansion algorithm is a very large-scale neighbourhood search (VLSN) technique and is very competitive in practice [9].

With respect to some current labelling ^f, the full set of possible expansion moves is $M(\hat{f}) = \cup_{\alpha \in L} M\alpha(\hat{f})$. The α-expansion algorithm simply performs local search over the full search neighbourhood M (^f). Perhaps surprisingly, local search with expansion moves will terminate with a labelling ^f that is within a constant factor from the globally optimal labelling f∗.

The α-expansion algorithm is generally implemented as shown below.

α-EXPANSION (E) — local search with expansion moves
1 f^:= arbitrary labelling
2 repeat
3 for each $\alpha \in L$
4 f := $\text{argmin}_{f \in M\alpha(\hat{f})} E(f)$ // solve via graph cut
5 if E(f) < E(f^)
6 f^:= f
7 until converged
8 return f^

Licensed Under Creative Commons Attribution CC BY

1. Energy Minimization Algorithms

2.1 Iterated Conditional Modes (ICM)

Iiterated conditional modes [1] uses a deterministic "greedy" strategy to find a local minimum. It starts with an estimate of the labelling, and then for each pixel it chooses the label giving the largest decrease of the energy function. This process is repeated until convergence, which is guaranteed to occur, and in practice is very rapid. Unfortunately, the results are extremely sensitive to the initial estimate, especially in high-dimensional spaces with non-convex energies (such as arise in vision) due to the huge number of local minima. ICM is assign each pixel the label with the lowest data cost. This resulted in significantly better performance.

2.2 Graph Cuts

The two most popular graph cuts algorithms [4], called the *swap move* algorithm and the *expansion move* algorithm were introduced in [7]. These algorithms rapidly compute a strong local minimum, in the sense that no "permitted move" will produce a labelling with lower energy. For a pair of labels α, β, a swap move takes some subset of the pixels currently given the label α and assigns them the label β, and vice-versa. The swap move algorithm finds a local minimum such that there is no swap move, for any pair of labels α, β that will produce a lower energy labelling. Analogously, we define an expansion move for a label α to increase the set of pixels that are given this label. The expansion move algorithm finds a local minimum such that no expansion move, for any label α, yields a labelling with lower energy. The criteria for a local minimum with respect to expansion moves (swap moves) are so strong that there are many fewer minima in high dimensional spaces compared to standard moves. In the original work of [7] the swap move algorithm was shown to be applicable to any energy where Vpq is a semi-metric and the expansion move algorithm to any energy where Vpq is a metric. The results of [9] imply that the expansion move algorithm can be used if for all labels α, β, and γ, $Vpq(\alpha, \alpha) + Vpq(\beta, \gamma) \leq Vpq(\alpha, \gamma) + Vpq(\beta,\alpha)$. The swap move algorithm can be used if for all labels α,β $Vpq(\alpha, \alpha) + Vpq(\beta, \beta) \leq Vpq(\alpha, \beta) + Vpq(\beta,\alpha)$. (This constraint comes from the notion of regular, i.e. sub modular, binary energy functions, which are closely related to graph cuts.) If the energy does not obey these constraints, graph cut algorithms can still be applied by "truncating" the violating terms [10].

2.3Max-product loopy belief propagation (LBP)

To evaluate the performance of LBP, we can implement the max-product LBP version, which is designed to find the lowest energy solution. The other main variant of LBP, the sum-product algorithm, does not directly search for a minimum energy solution, but instead computes the marginal probability distribution of each node in the graph. In general, LPB is not guaranteed to converge, and may go into an infinite loop switching between two labelling present a number of ways to speed up the basic algorithm. In particular, LBP implementation uses the distance transform method which significantly reduces the running time of the algorithm [2].

2.4 Tree-reweighted message passing (TRW)

It is a message-passing algorithm similar, on the surface, to LBP. An interesting feature of the TRW algorithm is that it computes a lower bound on the energy. The original TRW algorithm does not necessarily converge, and does not, in fact, guarantee that the lower bound always increases with time. An improved version of TRW was used, which is called sequential TRW, or TRW-S. In this version, the lower bound estimate is guaranteed not to decrease, which results in certain convergence properties. In TRW-S we first select an arbitrary pixel ordering function $S(p)$. The messages are updated in order of increasing $S(p)$ and at the next iteration in the reverse order. Trees are constrained to be chains that are monotonic with respect to $S(p)$.

This Introduction covers the terminology and techniques used for the cost labeling approach. Future work will be focused around improvement in label cost function, improving existing moves of α-expansion algorithm and introducing some new moves for this algorithm. Some, new technique will be used in α-expansion algorithm to optimize label cost function and utilize it for better results. Gaussian mixture models are formed by combining multivariate normal density components. Gaussian mixture models are often used for data clustering [11]. Clusters are assigned by selecting the component that maximizes the posterior probability. Like k-means clustering, Gaussian mixture modeling uses an iterative algorithm that converges to a local optimum. Gaussian mixture modeling may be more appropriate than k-means clustering when clusters have different sizes and correlation within them.

2. Literature Survey

The objective of the literature review is to find and explore the benefits of Energy Minimization algorithms and also what are the different problems in existing algorithms and techniques. The main goal of this literature review is to find the gaps in existing research and methods and also what will be the possible solutions to overcome these holes.

Andrew Delong et.al [5] In this paper author describes the α-expansion algorithm has had a significant impact in computer vision due to its generality, effectiveness, and speed. It is commonly used to minimize energies that involve unary, pair wise, and specialized higher-order terms. Their main algorithmic contribution is an extension of α-expansion that also optimizes "label costs" with well characterized optimality bounds. Label costs penalize a solution based on the set of labels that appear in it, for example by simply penalizing the number of labels in the solution. As energy has a natural interpretation as minimizing description length (MDL) and sheds light on classical algorithms like K-means and expectation-maximization (EM).Label costs are useful for multi-model fitting and several such applications: homography detection, motion segmentation, image segmentation, and compression.

Licensed Under Creative Commons Attribution CC BY

Lena Gorelick et.al [6] In this paper author describes computers vision problems elegantly in terms of energy minimization by characterizing a class of energies with hierarchical costs and proposing a novel hierarchical fusion algorithm. In semantic segmentation one could rule out unlikely object combinations via hierarchical context. In geometric model estimation, one could penalize the number of unique model families in a solution, not just the number of models—a kind of hierarchical MDL criterion. Hierarchical fusion uses the well-known α-expansion algorithm as a subroutine, and offers a much better approximation bound in important cases.

Olga Veksler et.al [7] In this paper author addresses the problem of minimizing a large class of energy functions that occur in early vision. The major restriction is that the energy function's smoothness term must only involve pairs of pixels. Two algorithms are proposed that use graph cuts to compute a local minimum even when very large moves are allowed. The first move considered is an α-β swap: for a pair of labels α β; this move exchanges the labels between an arbitrary set of pixels labeled and another arbitrary set labeled β. The first algorithm generates a labeling such that there is no swap move that decreases the energy. The second move considered is a α-expansion: for a label α, this move assigns an arbitrary set of pixels the label α. The second algorithm, which requires the smoothness term to be a metric, generates a labeling such that there is no expansion move that decreases the energy

Vladimir Kolmogorov[8] Minimum cut/maximum flow algorithms on graphs emerged as an increasingly useful tool for exact or approximate energy minimization in low-level vision. As the combinatorial optimization literature provides many min-cut/max-flow algorithms with different polynomial time complexity. The algorithms include both Goldberg-Tarjan style "push-relabel" methods and algorithms based on Ford-Fulkerson style "augmenting paths". These algorithms have been benchmarked on a number of typical graphs in the contexts of image restoration, stereo, and segmentation.

Yuri Boykov et.al [3] In the last few years, several new algorithms based on graph cuts have been developed to solve energy minimization problems in computer vision. Each of these techniques constructs a graph such that the minimum cut on the graph also minimizes the energy. Yet, because these graph constructions are complex and highly specific to a particular energy function, graph cuts have seen limited application to date. In this paper, a characterization of the energy functions that can be minimized by graph cuts has been given. However the results are restricted to functions of binary variables. Energy functions can be minimized using graph cuts, among the energy functions that can be written as a sum of terms containing three or fewer binary variables. A general-purpose construction to minimize such an energy function using a necessary condition for any energy function of binary variables to be minimized by graph cuts i.e.concerning energies of the form $E = D + V$ [2].By making the strong assumption that *both* D and V are *tree metrics*, and can compute a global optimum. However, most applications do not satisfy the metric assumption on data costs D.

Richard Szeliski et.al [9] Among the most exciting advances in early vision has been the development of efficient energy minimization algorithms for pixel-labeling tasks such as depth or texture computation. It has been known for decades that such problems can be elegantly expressed as Markov random fields, yet the resulting energy minimization problems have been widely viewed as intractable. Algorithms such as graph cuts and loopy belief propagation (LBP) have proven to be very powerful. However, the trade-offs among different energy minimization algorithms are still not well understood.

3. Discussion

Image can be segmented by assigning different labels (represented by different colors) to different objects. Label costs penalize a solution based on the set of labels that appear in it, for example by simply penalizing the number of labels in the solution. There should be sufficient number of labels; too many labels do not represent good segmentation as multiple labels may represent subpart of single object. On the other hand, in case of too less number of labels, a single label may represent multiple objects. Label cost can be associated with energy terms (combination of various energies associated with images e.g. Smoothing Energy, Bending Energy, Elastic energy etc.). Most labeling problems in computer vision and machine learning are ill-posed and in need of regularization, but the most useful regularization algorithms often make the problem NP-hard. The objective is to effectively optimize energies so that satisfactory image segmentation can be obtained (represented with different labels respective to different objects). In order to meet the objective, first task will be to define some label cost function in terms of energies. Unsupervised segmentation will be performed to assign labels by clustering simultaneously over pixels and color space using Gaussian Mixtures (for color images) and nonparametric histograms (for gray-scale images). Then based upon fast approximation based combinatorial optimization algorithm is implemented to minimize label cost function and redefine labels. α-expansion algorithm is already available for this purpose. The work will be focused around improvement in label cost function, and incorporating elastic energy for this algorithm.

4. Conclusion and Future Scope

Different metrics of energy minimization will be considered for performance comparison. An appropriate comparison will be drawn among proposed technique and previous well known techniques. The objective is to effectively optimize energies so that satisfactory image segmentation can be obtained (represented with different labels respective to different objects). New combinatorial optimization algorithm will be proposed to show promising experimental results with the new moves, which we believe could be used in any context where α -expansions are currently employed.

Licensed Under Creative Commons Attribution CC BY

References

[1] Besag, J.: On the statistical analysis of dirty pictures (with discussion). Journal of the Royal Statistical Society, Series B 48 (1986) 259–302.
[2] Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient belief propagation for early vision. In: CVPR. (2004) 261–268.
[3] Yuri Boykov and Olga Veksler. Graph Cuts in Vision and Graphics: Theories and Applications. In Nikos Paragios, Yunmei Chen, and Olivier Faugeras, editors, *Handbook of Mathematical Models in Computer Vision*, pages 79–96. Springer US, 2006.
[4] Energy Minimization using Graph Cuts 2011 http://rise.cse.iitm.ac.in/wiki/ images /0/07/Deepak.pdf
[5] Andrew Delong · Anton Osokin "Fast Approximate Energy Minimization with Label Costs" International Journal of Computer Vision, vol. 96, no. 1, pp. 1–27, January 2012.
[6] Lena Gorelick, Olga Veksler "Minimizing Energies with Hierarchical Costs" International Journal of Computer Vision, vol. 100, no. 1, pp. 38–58, October 2012.
[7] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(11):1222–1239, November 2001.
[8] Vladimir Kolmogorov "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", In IEEE Transactions on PAMI, Vol. 26, No. 9, pp. 1124-1137, Sept. 2004.
[9] Richard Szeliski, Ramin Zabih, "A Comparitive Study of Energy Minimization Methods for Markov Random Fields" Pattern Analysis and Machine Intelligence, IEEE Transactions on (Volume:30 , Issue: 6) 2008.
[10] Wainwright, M.J., Jaakkola, T.S., Willsky, A.S.: MAP estimation via agreement on (hyper)trees: Message-passing and linear-programming approaches. IEEE Trans Info Theory 51 (2005).
[11] Image Segmentation using Gaussian Mixture Models by Rahman Farnoosh, Gholamhossein Yari and Behnam Zarpak,Department of Applied Mathematics, University of Science and Technology, 16844,Narmak,Tehran, Iran.

Licensed Under Creative Commons Attribution CC BY