Superdifferential Cuts for Binary Energies

Tatsunori Tanial1, Yasuyuki Matsushita2, Takeshi Naemura1
1The University of Tokyo, Japan. 2Osaka University, Japan.

This paper makes the following contributions:
• we propose an optimization method for broad classes of higher-order and pairwise non-submodular functions that allows arbitrary directions of optimization and is thus robust to initialization.
• our method generalizes previous optimization methods including from early [4] to state-of-the-art methods [1, 3, 5].

Many low-level vision problems such as image segmentation, binarization, denoising, and tracking are often formulated as binary energy minimization. For example, in image segmentation, the use of Markov random field formulations and graph cuts has been becoming one of primary approaches [1, 2, 3, 5]. In this approach, the energy function is typically formulated as

\[ E(S) = R(S) + Q(S), \]

where \( R(S) \) describes appearance consistencies between resulting segments \( S \) and given information about target regions, and \( Q(S) \) enforces smoothness on segment-boundaries. The form of \( R(S) \) is often restricted to simple linear (i.e., pixelwise unary) forms because graph cuts allow globally optimal inference only for unary and submodular pairwise forms of energies. However, recent studies [1, 2, 3, 5] have shown that the use of higher-order information (or non-linear terms) can yield outstanding improvements over conventional pixelwise consistency measures.

In general, higher-order terms involve difficult optimization problems. Recent promising approaches try reducing energies by iteratively minimizing either first-order approximations (gradient descent approach) [2] or upper-bounds (bound optimization approach) [1, 5] of non-linear functions using graph cuts. The bound optimization approach has some advantages over the gradient descent approach [1]: It requires no parameters (e.g., step-size) and never worsens the solutions during iterations. But we must in turn derive appropriate bounds for individual functions. A notable work is auxiliary cuts (AC) [1] by Ayed et al., where they derive general bounds for broad classes of non-linear functionals for segmentation. However, the bounds derived in [1] are formulated to successively reduce target regions; thus the resulting segments are restricted within initial segments. Such a property actually limits the applications and accuracy of the method.

In order to derive more accurate and useful bounds, we revisit a submodular supermodular procedure (SSP) [4], a general bound optimization scheme for supermodular functions. We then propose a bound optimization method as generalization of SSP. Unlike SSP, our method can be used for non-supermodular functions; and unlike AC, it allows bi-directional optimization (see Fig. 2 for an illustration in segmentation) and can produce more accurate approximation bounds. We further show that our method can be seen as generalization of AC and some state-of-the-art method [3].

Figure 1: Intuitive illustration of the proposed piecewise linear approximations in comparison with the gradient-based approximations. The blue and red lines, and orange points show the energy function \( E(S) \), its approximations, and solutions \( S^0, S^1, S^2 \) obtained at each iteration, respectively. (a) In the gradient descent approach (e.g., FTR [2] and LSA-TR [3]), the energy function is approximated by local gradients (green) with trust regions (or step-size). The overall approximation at \( S^1 \) is depicted as the red curve, and its minimization results in falling in a bad local minimum \( S^2 \). (b) In our approach, the energy function is approximated by piecewise linear upper-bound functions. Here, the approximation bounds are updated in a coarse-to-fine manner (i.e., solid to dashed red lines) in order to avoid bad local minimums.

Figure 2: Matching foreground color distribution using our SDC-GEO, pPBC, and FTR with two types of initialization. Unlike pPBC, our method allows arbitrary directions of optimization.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.