Optimization of Max-Norm Objective Functions in Image Processing and Computer Vision

Filip Malmberg[1], Krzysztof Chris Ciesielski[2], Robin Strand[1]

[1]Dept. of Information Technology, Uppsala University, Sweden
[2] Department of Mathematics, West Virginia University and MIPG, Department of Radiology, University of Pennsylvania

April 10, 2019

Optimization in image processing

- Many fundamental problems in image processing and computer vision, such as image filtering, segmentation, registration, and stereo vision, can naturally be formulated as optimization problems.
- Often, these optimization problems can be described as *labeling* problems, in which we wish to assign to each image element (pixel) an element from some finite set of labels.

Optimization in image processing

We seek a label assignment configuration $\mathbf{x}$ that minimizes a given objective function E, which in the "canonical" case can be written as follows:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) , \quad (1)$$

where:

- $\mathcal{V}$ is the set of pixels in the image.
- $\mathcal{E}$ is the set of all adjacent pairs of pixels in the image.
- x_i denotes the label of vertex i, belonging to a finite set of integers $\{0, 1 \ldots, K-1\}$.

Optimization in image processing

We seek a label assignment configuration **x** that minimizes a given objective function E, which in the "canonical" case can be written as follows:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \;, \qquad (2)$$

where:

- $\mathcal{V}$ is the set of pixels in the image.
- $\mathcal{E}$ is the set of all adjacent pairs of pixels in the image.
- x_i denotes the label of vertex i, belonging to a finite set of integers $\{0, 1 \ldots, K-1\}$.

Optimization in image processing

We seek a label assignment configuration $\mathbf{x}$ that minimizes a given objective function E, which in the "canonical" case can be written as follows:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \,, \tag{3}$$

where:

- $\mathcal{V}$ is the set of pixels in the image.
- $\mathcal{E}$ is the set of all adjacent pairs of pixels in the image.
- x_i denotes the label of vertex i, belonging to a finite set of integers $\{0, 1 \ldots, K-1\}$.

Optimization in image processing

We seek a label assignment configuration $\mathbf{x}$ that minimizes a given objective function E, which in the "canonical" case can be written as follows:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \,, \tag{4}$$

where:

- $\mathcal{V}$ is the set of pixels in the image.
- $\mathcal{E}$ is the set of all adjacent pairs of pixels in the image.
- x_i denotes the label of vertex i, belonging to a finite set of integers $\{0, 1 \ldots, K-1\}$.

Data and regularization terms

- The functions $\phi_i(\cdot)$ are referred to as *unary* terms. Each unary term depends only on the label x_i assigned to the pixel i, and they are used to indicate the preference of an individual pixel to be assigned each particular label.
- The functions $\phi_{ij}(\cdot,\cdot)$ are referred to as *pairwise* terms. Each such function depends on the labels assigned to two pixels simultaneously, and thus introduces a dependency between the labels of different pixels. Typically, these terms express that the desired solution should have some degree of smoothness, or regularity.

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \tag{5}$$

Data and regularization terms

- The functions $\phi_i(\cdot)$ are referred to as *unary* terms. Each unary term depends only on the label x_i assigned to the pixel i, and they are used to indicate the preference of an individual pixel to be assigned each particular label.
- The functions $\phi_{ij}(\cdot, \cdot)$ are referred to as *pairwise* terms. Each such function depends on the labels assigned to two pixels simultaneously, and thus introduces a dependency between the labels of different pixels. Typically, these terms express that the desired solution should have some degree of smoothness, or regularity.

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \tag{6}$$

Data and regularization terms

- The functions $\phi_i(\cdot)$ are referred to as *unary* terms. Each unary term depends only on the label x_i assigned to the pixel i, and they are used to indicate the preference of an individual pixel to be assigned each particular label.
- The functions $\phi_{ij}(\cdot, \cdot)$ are referred to as *pairwise* terms. Each such function depends on the labels assigned to two pixels simultaneously, and thus introduces a dependency between the labels of different pixels. Typically, these terms express that the desired solution should have some degree of smoothness, or regularity.

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) . \tag{7}$$

Optimization by graph cuts

- In the general case, global optimization of this labeling problem is NP-hard.
- In special cases, globally optimal solutions can be found efficiently.
- For the binary labeling problem, with $K = 2$, a globally optimal solution can be computed by solving a max-flow/min-cut problem on a suitably constructed graph. This requires all pairwise terms to be *submodular*.
- A pairwise term ϕ_{ij} is said to be submodular if

$$\phi_{ij}(0,0) + \phi_{ij}(1,1) \leq \phi_{ij}(0,1) + \phi_{ij}(1,0) \ . \qquad (8)$$

Optimization by graph cuts

- At first glance, the restriction to binary labeling may appear very limiting.
- The multi-label problem can, however, be reduced to a sequence of binary valued labeling problems using, e.g., the *expansion move* algorithm (Boykov et al. 2001, Kolmogorov et al. 2004)
- Thus, the ability to find optimal solutions for problems with two labels has high relevance also for the multi-label case.

Generalized objective functions

Looking again at the labeling problem described above, we can view the objective function E as consisting of two parts:

- A *local* error measure, in our case defined by the unary and pairwise terms.
- A *global* error measure, aggregating the local errors into a final score. In the case of E, the global error measure is obtained by summing all the local error measures.

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \tag{9}$$

Generalized objective functions

Looking again at the labeling problem described above, we can view the objective function E as consisting of two parts:

- A *local* error measure, in our case defined by the unary and pairwise terms.
- A *global* error measure, aggregating the local errors into a final score. In the case of E, the global error measure is obtained by summing all the local error measures.

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \tag{10}$$

Generalized objective functions

Looking again at the labeling problem described above, we can view the objective function E as consisting of two parts:

- A *local* error measure, in our case defined by the unary and pairwise terms.
- A *global* error measure, aggregating the local errors into a final score. In the case of E, the global error measure is obtained by summing all the local error measures.

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}(x_i, x_j) \tag{11}$$

l_p-norm objective functions

If we assume all terms to be non-negative, minimizing E can be seen as minimizing the l_1-norm of the vector containing all unary and pairwise terms. A natural generalization is to consider minimization of arbitrary l_p-norms, $p \geq 1$, i.e., minimizing:

$$E_p(\mathbf{x}) = \left(\sum_{i \in \mathcal{V}} \phi_i^p(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}^p(x_i, x_j) \right)^{1/p} \tag{12}$$

What is the effect of p?

- The value p can be seen as a parameter controlling the balance between minimizing the overall cost versus minimizing the magnitude of the individual terms.
- For $p = 1$, the optimal labeling may contain (few) arbitrarily large individual terms as long as the sum of the terms is small.
- As p increases, a larger penalty is assigned to solutions containing large individual terms. This forces local errors to be distributed more evenly across the image domain.

Letting p go to ∞

As p approaches infinity, the objective function approaches the ∞-norm, or max-norm, of the local errors:

$$E_\infty(\mathbf{x}) = \max\Big\{\max_{i \in V} \phi_i(x_i), \max_{\{i,j\} \in \mathcal{E}} \phi_{ij}(x_i, x_j)\Big\}. \qquad (13)$$

In this case, the global error is completely determined by the largest local error. Intuitively, this means that the local errors are distributed as evenly as possible across the image domain.

Why are we interested in max-norm problems?

- It is well known that special cases of the max-norm optimization problems given above can be solved very efficiently (quasi-linear time) using MSF-cuts, a.k.a. *watershed cuts*. These methods has mainly been used for image segmentation, with only a few papers considering more general optimization problems.
- Some of these solvable problems are outside of the class of problems we can solve by minimal graph cuts. (i.e., some multilabel cases).
- Yet, no systematic study has (to our knowledge) been made to determine what class of max-norm optimization problems we can actually solve!

A missing paper?

Kolmogorov and Zabih
What energy functions can be minimized via graph cuts?
IEEE PAMI, 2004

Direct optimization of max-norm problems

- In our paper, we present an efficient (quasi-linear time) algorithm for optimizing binary labeling problems with the max-norm E_∞ objective function.
- Our algorithm is structurally similar to algorithms for MSF/watershed cuts.
- We prove that the algorithm produces a globally optimal solution provided that all pairwise terms are ∞-*submodular*:

$$\max\{\phi_{ij}(0,0), \phi_{ij}(1,1)\} \leq \max\{\phi_{ij}(1,0), \phi_{ij}(0,1)\}. \quad (14)$$

Note the symmetry

- We can find globally optimal solutions for E_1 objective functions, if all pairwise terms are 1-submodular. (Using graph cuts)

$$\phi_{ij}(0,0) + \phi_{ij}(1,1) \leq \phi_{ij}(0,1) + \phi_{ij}(1,0)\ . \tag{15}$$

- We can find globally optimal solutions for E_∞ objective functions, if all pairwise terms are ∞-submodular. (Using our proposed algorithm)

$$\max\{\phi_{ij}(0,0), \phi_{ij}(1,1)\} \leq \max\{\phi_{ij}(1,0), \phi_{ij}(0,1)\}. \tag{16}$$

Outline of our proposed algorithm

- To describe the method, we introduce the notion of unary and binary solution *atoms*.
- A *unary* atom represents one possible label configuration for a single vertex.
- A *binary* atom represent a possible label configuration for a pair of adjacent vertices.
- Thus, for a binary labeling problem, there are two unary atoms associated with every pixel and four binary atoms for every pair of adjacent pixels.
- Each atom has a *weight* given by the corresponding unary or binary term of the objective function.

Outline of our proposed algorithm

The algorithm works as follows:

- Start with a set S consisting of all possible atoms.
- For each atom A, in order of decreasing weight:
  - If A is still in S, and is not the only remaining atom for that vertex/edge, remove A from S.
  - After the removal of A, S may contain incompatible atoms. Iteratively remove incompatible atoms until S contains no more incompatible atoms.

Beyond ∞-submodular functions?

- For graph cuts/E_1 problems, minimizing non-submodular functions is NP-hard in the general case.
- In the E_∞ case, we can minimize *at least* all ∞-submodular functions with two labels.
- Open question: Is the problem of minimizing E_∞ functions that are not ∞-submodular also NP-hard? (Spoiler alert: *it appears not!*)

Conclusions

- Optimization problems, specifically pixel labeling problems, are frequently occuring in image processing applications.
- We are specifically interested in problems where the objective function is given by the max-norm of the local errors.
- For many such problems, globally optimal solutions can be found very efficiently, in quasi linear time using "MSF-like algorithms". We have initiated a systematic study of these problems, to determine exactly what class of problems can be solved in this way.
- An important first result, presented here, is that the class of solvable problems includes all binary labeling problems with ∞-submodular pairwise terms.

Thank you for your attention!