Optimization of pixel labeling problems with max-norm objective functions – a summary of recent results

Filip Malmberg[1] and Krzysztof Chris Ciesielski[2]

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

April 1, 2021

An elevator pitch

"Many important optimization problems arising in image processing and computer vision, that we normally consider computationally infeasible to solve, can be solved efficiently if we use the L_∞-norm instead of other commonly used norms."

Optimization problems 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 problems 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:

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

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 $\{0, 1 \ldots, K-1\}$.

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.

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

This is a hard problem!

- Images of interest have thousands or millions of elements.
- If we only have a unary term, we can trivially assign the best label to each pixel independently.
- Through the pairwise terms, the labels assigned to these pixels become interdependent. This makes the problem much more interesting, but also *much* harder.
- In fact, finding a globally optimal solution to the labeling problem described above is NP-hard in the general case.

Optimization by minimal graph cuts

- In the general case, global optimization of this labeling problem is NP-hard, but 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* ($\approx$ convex).
- A pairwise term ϕ_{ij} is said to be submodular if

$$\phi_{ij}(0,0) + \phi_{ij}(1,1) \leq \phi_{ij}(0,1) + \phi_{ij}(1,0)\ . \qquad (3)$$

Multi-label problems

- 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.
- These approaches have been very succesful, and have made graph cuts a standard tool for solving general optimization problem in image processing.

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.

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

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:

$$E_p(\mathrm{x}) = \left(\sum_{i \in \mathcal{V}} \phi_i^p(x_i) + \sum_{i,j \in \mathcal{E}} \phi_{ij}^p(x_i, x_j) \right)^{1/p} \tag{5}$$

Minimizing L_p-norm objective functions via grah cuts

- It is straightfoward to show that similar submodularity requirements hold also for the generalized objective functions E_p for any finite p.

$$(\phi_{st}^p(0,0) + \phi_{st}^p(1,1))^{1/p} \leq (\phi_{st}^p(0,1) + \phi_{st}^p(1,0))^{1/p}. \quad (6)$$

- To use the graph cut approach, we must first verify that all pairwise terms satisfy the appropriate submodularity conditions. Otherwise, we have to resort to approximate methods.

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:

$$E_\infty(\mathrm{x}) = \max\Big\{\max_{i \in V} \phi_i(x_i), \max_{\{i,j\} \in \mathcal{E}} \phi_{ij}(x_i, x_j)\Big\}. \tag{7}$$

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.

Letting p go to ∞ – a toy example

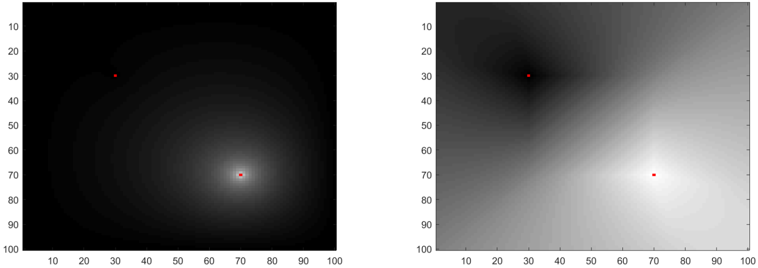


Figure 1: Left: L_2.norm, Right: L_∞-norm.

A quite remarkable result

- We have shown that in the limit as p goes to infinity, *the requirement for submodularity of the pairwise terms disappears*!
- Thus, even when the local costs are such that the problem of minimizing E_p is NP-hard for some or all finite p, a labeling minimizing E_∞ can be found in low order polynomial time! (In practice: linearithmic)

Direct optimization of max-norm problems

- In two recent papers, we present two different algorithms for optimizing binary labeling problems with the max-norm E_∞ objective function:
  - A linearithmic time algorithm for optimizing E_∞ under the condition that all pairwise terms are ∞-*submodular*.
  - An algorithm for optimizing *any* function E_∞, submodular or not. The theoretical runtime for this algorithm is quadratic, but empirically it is also linearithmic.
- A pairwise term is said to be ∞-*submodular* if:

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

Outline of our proposed algorithms

- To describe the optimization methods, 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 $S \setminus \{A\}$ is consistent, remove A from S.

A set of atoms is said to be *consistent* if it is possible to construct at least one valid labeling from the atoms in the set.
At the termination of this algorithm, the atoms remaining in S define a unique labeling. This labeling is globally optimal according to the objective function E_∞.

Checking consistency

The key issue is to determine, at each step of the algorithm, whether the remaining set of atoms is consistent.

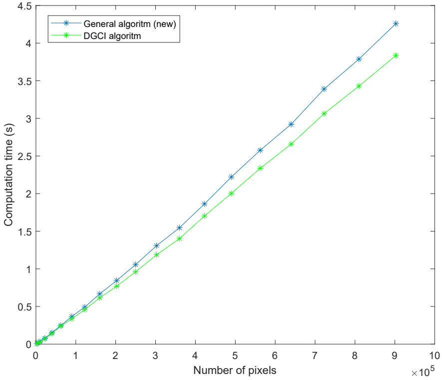
- When the all pairwise terms are ∞-submodular, we show that this check can be performed efficiently via "local" conditions. This leads to the pseudo-linear algorithm.
- In the general case, we show that the problem of determining the consistency can be phrased as a *boolean 2-satisfiability problem*, solvable in linear time. This leads to the quadratic algorithm.

The 2-SAT problem

- Consider a set of boolean variables (*true* or *false*) and a set of constraints on these variables, such that each constraint involves at most two variables. The *2-SAT problem* consists of answering the question: *Is there an assignment of truth values (i.e.,*0 *or* 1*) to the variables that satisfies all given constraints?*
- Solvable in linear time using e.g., Aspvall's algorithm.

An efficient version of the general algorithm

- Running Aspvall's algorithm for every atom we want to remove is inefficient.
- Each satisfiability problem, however, is very similar to the previous one. We have found a way to utilize this redundancy to formulate a practically efficient algorithm!



Strict/lexicographical optimization

- A potential drawback of the L_∞-norm is that it does not distinguish between solutions with high or low errors below the maximum error. This may be resolved by the notion of *stricit optimizers*.
- Two solutions are compared by ordering all elements non-increasingly by their local error value, and then performing a lexicographical comparison. A solution is a strict optimizer if it is better than or equal to any other solution according to this comparison.
- Any strict minimizer is also an L_∞-optimal solution. The limit, as $p \to \infty$, of L_p-norm minimizers is not only an L_∞-minimizer but also a strict minimizer.
- We have shown that, under certain conditions, our algorithms produce strictly optimal solutions.

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.
- Many important optimization problems that are NP-hard under other p-norms can be solved very efficiently under the max-norm!

Thank you for your attention!

Relevant publications:

- Malmberg, F, Strand, R.: When Can l_p-norm Objective Functions Be Minimized via Graph Cuts?
  International Workshop on Combinatorial Image Analysis, Springer, (2018)
- Malmberg, F., Ciesielski, K.C., Strand, R.: Optimization of max-norm objective functions in image processing and computer vision.
  International Conference on Discrete Geometry for Computer Imagery, pp. 206–218. Springer (2019)
- Malmberg, F., Ciesielski, K.C.: Two polynomial time graph labeling algorithms optimizing max-norm based objective functions.
  Journal of Mathematical Imaging and Vision, (2020).