Pseudo-Boolean Optimization: Theory and Applications in Vision

Strandmark, Petter; Kahl, Fredrik

2012

Link to publication

Citation for published version (APA):
Pseudo-Boolean Optimization: Theory and Applications in Vision

Petter Strandmark, Fredrik Kahl
Centre for Mathematical Sciences, Lund University
Email: {petter, fredrik}@maths.lth.se

Abstract—Many problems in computer vision, such as stereo, segmentation and denoising can be formulated as pseudo-boolean optimization problems. Over the last decade, graph cuts have become a standard tool for solving such problems. The last couple of years have seen a great advancement in the methods used to minimize pseudo-boolean functions of higher order than quadratic. In this paper, we give an overview of how one can optimize higher-order functions via generalized roof duality and how it can be applied to problems in image analysis and vision.

I. PSEUDO-BOOLEAN OPTIMIZATION

A pseudo-boolean function is a function from the set of boolean vectors of dimension $n$, denoted $\mathbb{B}^n = \{0, 1\}^n$, to the reals. Any pseudo-boolean $f$ can be uniquely represented by a multilinear polynomial of the form

$$f(x) = \sum_i a_i x_i + \sum_{i<j} a_{ij} x_i x_j + \sum_{i<j<k} a_{ijk} x_i x_j x_k + \ldots$$

The degree $m$ of $f$ is equal to the degree of the corresponding polynomial. In this paper, we are interested in efficient ways of solving the following optimization problem:

$$\min_{x \in \mathbb{B}^n} f(x), \quad (1)$$

in order to improve performance for many basic applications in computer vision.

Application problems that can be turned into such an optimization problem abound. For example, state-of-the-art methods for stereo, segmentation and image denoising are often formulated as the inference of the maximum a posteriori estimate in a Markov Random Field (MRF) and such problems can be formulated as energy minimization problems where the energy function is given by a pseudo-boolean function.

In general, the minimization problem in (1) is NP-hard so approximation algorithms are necessary. For the quadratic case ($m = 2$), one of the most popular and successful approaches is based on the roof duality bound [1], [2]. We generalize the roof duality framework for higher-order pseudo-boolean functions. Our main contributions are (i) how one can define a general bound for any order (for which the quadratic case is a special case), (ii) how one can efficiently compute solutions that attain this bound in polynomial time and (iii) to give example applications in vision for which performance is improved by this technique. Our results rely on many previous significant contributions.

A. Related work and Applications

Graph cuts is by now a standard tool for many vision problems, in particular, for the minimization of quadratic and cubic submodular pseudo-boolean functions [3], [4]. The same technique can be used for non-submodular functions in order to compute a lower bound [1].

In recent years, there has been an increasing interest in higher-order models and approaches for minimizing the corresponding energies. For example, in [5], approximate belief propagation is used with a learned higher-order MRF model for image denoising. Similarly, in [6], an MRF model is learned for texture restoration, but the model is restricted to submodular energies which can be optimized exactly with graph cuts. Curvature regularization requires higher-order models [7], [8]. Even global potentials defined over all variables in the MRF have been considered, e.g., in [9] for ensuring connectedness, in [10] to model co-occurrence statistics of objects. Another state-of-the-art example is [11] where second-order surface priors are used for stereo reconstruction. The optimization strategies rely on dual decomposition [12], [13], move-making algorithms [14], [15], linear programming [16], belief propagation [5] and, of course, graph cuts.

B. Submodularity

There is a large subset of pseudo-boolean which contains functions which are easy to minimize: the set of submodular functions. If we let $x \land y$ and $x \lor y$ mean element-wise min and max, respectively, the set can be defined as the set of all functions $f$ satisfying

$$f(x) + f(y) \geq f(x \land y) + f(x \lor y) \quad (2)$$

for all points $x, y \in \mathbb{B}^n$. In the quadratic ($m = 2$) and cubic ($m = 3$) cases, a submodular function $f$ can be minimized very efficiently via graph cuts and in polynomial time for any degree. However, for $m \geq 4$, it is a very difficult problem (co-NP-complete) to determine whether a given function is submodular [17].

As an example, the function

$$f(x) = x_1 - 4x_2 + 2x_3 - 5x_1 x_2 \quad (3)$$

is submodular but the function

$$f(x) = x_1 - 4x_2 + 2x_3 + 5x_1 x_2 \quad (4)$$

is not.
C. Roof Duality

The roof duality bound is an efficiently computable lower bound to the minimum of \( f \) in the quadratic case. While a good lower bound can be interesting in many applications, e.g., when using branch and bound, another property of roof duality is arguably even more useful: persistency. Each variable in the solution obtained via roof duality is equal to one of three possibilities: \( \{0, 1, ?\} \). It is guaranteed that a global solution exists corresponding to the parts of the solution not equal to ‘?’. Therefore, the roof duality solution is said to be partially optimal.

The fastest method of computing the roof dual for quadratic pseudo-boolean polynomials is by using graph cuts. Reductions have been explored for higher-order polynomials \( (m > 2) \), e.g., \([18]–[21]\). These methods all convert a higher-order function to a quadratic one.

Our framework builds on \([22]\) using submodular relaxations directly on higher-order terms. We define optimal relaxations to be those that give the tightest lower bound. As an example, consider the problem of minimizing the following cubic polynomial \( f \) over \( \mathbb{B}^3 \):

\[
\begin{align*}
f(x) = & -2x_1 + x_2 - x_3 \\
& + 4x_1x_2 + 4x_1x_3 - 2x_2x_3 - 2x_1x_2x_3.
\end{align*}
\] (5)

The standard reduction scheme \([21]\) would use the identity \(-x_1x_2x_3 = \min_{z \in \mathbb{B}} z(2 - x_1 - x_2 - x_3)\) to obtain a quadratic minimization problem with one auxiliary variable \( z \). Roof duality gives a lower bound of \( f_{\text{min}} \geq -3 \), but it does not reveal how to assign any of the variables in \( x \). However, there are many possible reduction schemes from which one can choose. Another possibility is \(-x_1x_2x_3 = \min_{z \in \mathbb{B}} z(-x_1 + x_2 + x_3 - x_1x_2 - x_1x_3 + x_1)\). For this reduction, the roof duality bound is tight and the optimal solution \( x^* = (0, 1, 1) \) is obtained. This simple example illustrates two facts: (i) different reductions lead to different lower bounds and (ii) it is not an obvious matter how to choose the optimal reduction.

II. GENERALIZED ROOF DUALITY

A. Problem Formulation

Consider the optimization problem in (1) where \( f \) has degree \( m \). In this paper, we will investigate the cases when \( m = 3 \) and \( m = 4 \). By enlarging the domain, we will relax the problem and look at the following tractable problem:

\[
\min_{(x, y) \in \mathbb{B}^{2n}} g(x, y),
\] (6)

where \( g : \mathbb{B}^{2n} \mapsto \mathbb{R} \) is a pseudo-boolean function that satisfies the three conditions

\[
\begin{align*}
g(x, \bar{x}) &= f(x), \quad \forall x \in \mathbb{B}^n, \quad (A) \\
g \text{ submodular}, \quad (B) \\
g(x, y) &= g(\bar{y}, \bar{x}), \quad \forall (x, y) \in \mathbb{B}^{2n} \text{ (symmetry).} \quad (C)
\end{align*}
\]

For a point \( x = (x_1, x_2, \ldots, x_n) \in \mathbb{B}^n \), we denote \( \bar{x} = (x_1, x_2, \ldots, x_n) = (1 - x_1, 1 - x_2, \ldots, 1 - x_n) \). The reason for requirement (A) is that if the range of \( f \) is included in the range of \( g \) then the minimum of \( g \) is a lower bound to the minimum of \( f \). If the computed minimizer \((x^*, y^*)\) of the relaxation \( g \) happens to fulfill \( x^* = \bar{y}^* \) then, of course, \( x^* \) is a minimizer of \( f \) as well. Even if it is not the case that \( x^* = \bar{y}^* \), we still obtain a lower bound on \( f \) and as we shall see, it is possible to extract a partial solution for a minimizer of \( f \).

Requirement (B) is also fairly obvious. Since we must be able to minimize \( g \), requiring that \( g \) is submodular is natural. The last requirement is more technical and is required in order to prove persistency. It also turns out that including (C) does not give worse lower bounds.

Let \( f_{\text{min}} \) denote the unknown minimum value of \( f \), that is, \( f_{\text{min}} = \min f(x) \). Ideally, we would like \( g(x, y) \geq f_{\text{min}} \) for all points \((x, y) \in \mathbb{B}^{2n}\). This is evidently not possible in general. However, one could try to maximize the lower bound of \( g \), \( \max_{g} \min_{x,y} g(x, y) \), that is,

\[
\max_{g} \ell \quad \text{such that} \quad g(x, y) \geq \ell, \quad \forall (x, y) \in \mathbb{B}^{2n}, \quad (7)
\]

\( g \) satisfies (A)-(C).

A relaxation \( g \) that provides the maximum lower bound will be called optimal. As we have proved in \([23]\), when \( m = 2 \), the lower bound coincides with the roof duality bound \([1]\) and therefore this maximum lower bound will be referred to as generalized roof duality.

B. Properties of the Solution

By construction, the solution to (7) gives a lower bound to the optimal value of \( f \). Another important property is that it satisfies persistency in the following way: Let \( g^* \) be the solution to (7) and \((x^*, y^*) \in \arg \min g(x, y) \). If we then let \( I \) be the set of all \( i \) such that \( x_i^* \neq y_i^* \), then there exists a global minimum \( \bar{x} \) to our original function \( f \) such that \( \bar{x}_i = x_i^* \) for all \( i \in I \). See \([23]\) for a proof.

In other words, the minimizer to \( g^* \) sometimes gives us partial information about a global minimizer of \( f \), just as roof duality does.

C. Solving the Problem

While problem (7) is a linear program in the coefficients of \( g \), the number of constraints is exponential in the number of variables \( n \). Nevertheless, it is possible to attain the roof duality bound in polynomial time \([23]\).

Requirements (A) and (C) never give any problems. Requirement (B) (submodularity) can be represented with polynomially many constraints only when \( m \leq 3 \). Recall that determining whether a function \( f \) is submodular is co-NP-complete. For \( m = 4 \) we therefore have to look at smaller sets than the entire set of submodular functions, as no reasonably sized set of linear constraint can describe all submodular functions.

The final hurdle is the exponentially many constraints involving \( \ell \) in (7). We solved this in \([23]\) using an iterative approach. Instead of maximizing \( \min_{x,y} g(x, y) \) it is enough to maximize \( g(0, 0) \), which eliminates all constraints. Then, when a solution is obtained, \( g \) can be minimized. If any persistencies are available, they can be used to simplify \( f \) and the process starts over. Otherwise, the process terminates. This will give a bound which is at least as good as the bound from (7).
III. EXPERIMENTS

We have performed experiments comparing the proposed generalized roof duality to the current state of the art. The methods we are considering are:

- **GRD** Generalized roof duality over all submodular functions for $m = 3$ and a subset for $m = 4$.
- **GRD-gen** Same as GRD, except over a larger subset for $m = 4$ [24] (which results in a larger linear program).
- **GRD-heuristic** Instead of solving a linear program (which can be expensive), a very fast heuristic is used to give an approximate solution.

The reductions proposed in [25].

**HOCR** The reductions proposed in [21].

A. Random Polynomials

In the first experiment, we generated polynomials with random coefficients:

$$f(x) = \sum_{(i,j,k) \in T} f_{ijk}(x_i, x_j, x_k),$$

where $T \subseteq \{1 \ldots n\}^3$ is a random set of triplets and each $f_{ijk}$ is a cubic polynomial in $x_i, x_j, x_k$ with all its coefficients picked uniformly in $\{-100, \ldots, 100\}$. The results for all methods can be seen in Fig. 2.

B. Image Denoising

Ishikawa [21] used image denoising as a benchmark problem for higher-order pseudo-boolean minimization. In each iteration proposals are generated in two possible ways which are alternated: by blurring the current image and picking all pixels at random. Each pixel then has a choice of staying the same or switching to the proposal, resulting in a pseudo-boolean optimization problem. The smoothness term consists of a Fields of Experts (FoE) model using patches of size $2 \times 2$. Thus, quartic polynomials are needed to formulate the image restoration task as a pseudo-boolean minimization problem.

Figure 3 shows a comparison between the different methods for this problem. Generalized roof duality performed very well, often labeling very close to 100% of the problem variables. Figure 1 shows the results for a $160 \times 240$ image, where all three methods used have comparable running time.

REFERENCES

Fig. 2: Number of persistencies, relative bounds and running time for 100 random polynomials. The set of coefficients \( T \) was drawn uniformly after making sure that all variables were used once.

(a) Cubic polynomials with \( n = 1000 \) and \( |T| = 1000 \). GRD-heuristic is not shown in the histogram because it is almost indistinguishable from GRD.

(b) Quartic polynomials with \( n = 1000 \) and \( |T| = 300 \).

Fig. 3: Restoring a small image. In each iteration a proposal is generated and each pixel can either stay the same or switch to the proposal. A quartic smoothness function is used.