CVPR2009: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami Beach, Florida. June 20-25, 2009.

Higher-Order Clique Reduction in Binary Graph Cut

Hiroshi Ishikawa

Nagoya City University
Department of Information and Biological Sciences
1 Yamanohata Mizuho, Nagoya 467-8501, Japan
hfs@ieee.org

Abstract

We introduce a new technique that can reduce any higher-order Markov random field with binary labels into a first-order one that has the same minima as the original. Moreover, we combine the reduction with the fusion-move and QPBO algorithms to optimize higher-order multi-label problems. While many vision problems today are formulated as energy minimization problems, they have mostly been limited to using first-order energies, which consist of unary and pairwise clique potentials, with a few exceptions that consider triples. This is because of the lack of efficient algorithms to optimize energies with higher-order interactions. Our algorithm challenges this restriction that limits the representational power of the models, so that higher-order energies can be used to capture the rich statistics of natural scenes. To demonstrate the algorithm, we minimize a third-order energy, which allows clique potentials with up to four pixels, in an image restoration problem. The problem uses the Fields of Experts model, a learned spatial prior of natural images that has been used to test two belief propagation algorithms capable of optimizing higher-order energies. The results show that the algorithm exceeds the BP algorithms in both optimization performance and speed.

1. Introduction

Many problems in computer vision such as segmentation, stereo, and image restoration are formulated as optimization problems involving inference of the maximum a posteriori (MAP) solution of a Markov Random Field (MRF). Such optimization schemes have become quite popular, largely owing to the success of optimization techniques such as graph cuts[5, 10, 14], belief propagation[7, 21], and tree-reweighted message passing[13]. However, because of the lack of efficient algorithms to optimize energies with higher-order interactions, most are represented in terms of unary and pairwise clique potentials, with a few exceptions that consider triples[6, 14, 29]. This limitation severely restricts the representational power of the models: the rich statistics of natural scenes cannot be captured by such limited potentials[21]. Higher order cliques can model more complex interactions and reflect the natural statistics better.

Table 1. Graph cut applicability

Order	Binary labels	Multiple labels
First	Mincut[14], QPBO[9]	$\rightarrow$ binary (α-exp. [5]), [10]
Second	$\rightarrow$ 1st order [14, 8]	$\rightarrow$ binary (fusion [29])
Higher	$\rightarrow$ 1st order (*this paper*)	$\rightarrow$ binary (fusion, *this paper*)

This has long been realized[11, 22, 25], but with the recent success of the new energy optimization methods, there is a renewed emphasis on the effort to find an efficient way to optimize MRFs of higher order. For instance, belief propagation variants[17, 23] have been introduced to do inference based on higher-order clique potentials. In graph cut, Kolmogorov and Zabih[14] found a reduction that can reduce second-order binary-label potentials into pairwise ones, followed by an algebraic simplification by Freedman and Drineas[8] (Table 1.) Kohli et al.[12] extend the class of energies for which the optimal α-expansion and $\alpha\beta$-swap moves can be computed in polynomial time. Most recently, there are at least three papers just in this CVPR, including this one, addressing the problem of minimizing higher-order Markov random fields: Komodakis and Paragios[16] employ a master-slave decomposition framework to solve a dual relaxation to the MRF problem; Rother et al.[27] use a soft-pattern-based representation of higher-order functions that may for some energies lead to very compact first-order functions with small number of non-submodular terms, as well as addressing the problem of transforming general multi-label functions into quadratic ones.

In our approach, higher-order energies are optimized by "move-making," in which higher-order energies of binary variables are reduced to first-order ones and minimized iteratively. For our approach, two recent advances in graph cut are essential. First, there is a recent generalization called "fusion move"[18, 19] of the α-expansion algorithm[5]. The energy is minimized in α-expansion by starting from an

initial labeling and iteratively making a series of moves that decrease the energy. In each iteration, the move is selected by solving a binary-label optimization problem with an s-t mincut algorithm, which can globally optimize a class of first-order binary-label energy potentials called submodular functions[14]. The move at each pixel can be i) changing the label to α or ii) keeping the current label. The fusion move generalizes this by allowing a "proposal" and the current labeling to be merged by arbitrarily choosing one of the two at each pixel to generate a new map of labels.

Second, a recent innovation allows optimization of first-order non-submodular functions. This method by Boros, Hammer, and their co-workers [2, 4, 9] is variously called QPBO[15] or roof-duality[26]. If the function is submodular, QPBO is guaranteed to find the global minimum. Even if it is not submodular, QPBO returns a solution assigning either 0, 1, or -1 to each pixel, with the guarantee that at pixels where 0 or 1 is assigned, the value is what it would be with a globally minimum labeling. Pixels that are assigned -1 have "unlabeled" values. This has a crucial impact on the move-making algorithms since the choice of the move at each iteration depends on binary-label optimization. In particular, QPBO has an "autarky" property[15] that lets us ensure that energy does not increase in move-making: we just leave the label unchanged at those pixels that are given the value -1 by the QPBO algorithm. In the context of optimizing higher-order potentials, it means that some limitations that prevented the use of these algorithms for higher-order functions can possibly be overcome.

As we mentioned, second-order potentials on binary variables can be reduced into pairwise ones[14]. However, the requirement that the result of reduction must be submodular made its actual use quite rare, if not nonexistent. Thanks to the QPBO technique, now we can think of reducing higher-order potentials into pairwise ones with a hope that at least part of the solution can be found. Although it is not a solution to every problem, the QPBO technique often allows solving a non-submodular problem approximately by giving a large enough part of the globally optimal move to iteratively improve the solution. Woodword et al.[29] recently used this strategy very successfully.

So far, the second-order case has remained the only case that could be solved using this group of techniques, because the reduction we mention above is only applicable in that case (Table 1). To be sure, a totally different reduction technique that can reduce energies of any order has been known for a long time[24]; however, to our knowledge it has never been used successfully in practice for orders higher than two. This seems to be because, even though it can reduce any function into pairwise clique potential, the result is *always* non-submodular. We discuss this in §2.2 and §2.3.

In this paper, we introduce a new reduction technique along the line of the Kolmogorov-Zabih reduction that can reduce any higher-order optimization problem on binary variables into an equivalent first-order problem. Then we use it with fusion move and QPBO in move-making algorithms to optimize higher-order multi-label energies.

We demonstrate its effectiveness by testing it on a third-order potential, an image restoration problem that has been used to test two BP algorithms[17, 23] capable of optimizing higher-order energies. We show that the fusion-move algorithm using our reduction outperforms the BP algorithms in both quality and speed.

In the next section, we describe briefly the two known reductions and discuss their limitations. In section 3, we introduce the new reduction. We describe the higher-order fusion-move algorithm using the reduction in section 4 and its experimental validation in section 5.

2. Known Reductions

There are a couple of known methods to reduce a higher-order function of binary variables to first-order one so that the minima of the reduced function can be translated easily to those for the original function. Here, we outline the known reduction methods and then discuss their limitations.

We denote the set $\{0, 1\}$ by $\mathbb{B}$ and the set of all real numbers by $\mathbb{R}$. We consider functions in binary variables, i.e., a map $f : \mathbb{B}^n \to \mathbb{R}$, where n is the number of variables. Such functions are called *pseudo-Boolean functions*. Any pseudo-Boolean function can be uniquely represented as a polynomial of the form

$$f(x_1, \ldots, x_n) = \sum_{S \subset V} c_S \prod_{i \in S} x_i, \tag{1}$$

where $V = \{1, \ldots, n\}$ and $c_S \in \mathbb{R}$ [3].

2.1. Reduction by Minimum Selection

Kolmogorov and Zabih[14] first proposed this reduction in the context of graph-cut optimization. Later, Freedman and Drineas[8] recast it into an algebraic formula.

Consider a cubic pseudo-Boolean function of $x, y, z \in \mathbb{B}$

$$f(x, y, z) = axyz. \tag{2}$$

The reduction is based on the following:

$$xyz = \max_{w \in \mathbb{B}} w(x + y + z - 2). \tag{3}$$

Suppose $\mathbb{R} \ni a < 0$. Then

$$axyz = \min_{w \in \mathbb{B}} aw(x + y + z - 2). \tag{4}$$

Thus, whenever $axyz$ appears in a minimization problem with $a < 0$, it can be replaced by $aw(x + y + z - 2)$.

What if $a > 0$? Then we flip the variables (i.e., replace x by $1 - x$, y by $1 - y$, and z by $1 - z$) of (3) and consider

$$(1 - x)(1 - y)(1 - z) = \max_{w \in \mathbb{B}} w(1 - x + 1 - y + 1 - z - 2). \tag{5}$$

This is simplified to

$$xyz = \min_{w\in\mathbb{B}} w(x+y+z-1) + (xy+yz+zx) - (x+y+z) + 1. \quad (6)$$

Therefore, if $axyz$ appears in a minimization problem with $a > 0$, it can be replaced by

$$a\{w(x+y+z-1) + (xy+yz+zx) - (x+y+z) + 1\}. \quad (7)$$

Thus, either case, the cubic term can be replaced by quadratic terms.

This reduction only works with cubic terms. For quartic term $axyzt$, the same trick works if $a < 0$:

$$xyzt = \max_{w\in\mathbb{B}} w(x+y+z+t-3), \quad (8)$$

$$axyzt = \min_{w\in\mathbb{B}} aw(x+y+z+t-3). \quad (9)$$

However, if $a > 0$,

$$(1-x)(1-y)(1-z)(1-t) = \max_{w\in\mathbb{B}} w(-x-y-z-t+1) \quad (10)$$

becomes

$$xyzt = \max_{w\in\mathbb{B}} w(-x-y-z-t+1) + (xyz+xyt + xzt+yzt) - (xy+yz+zx+xt+yt+zt) + (x+y+z+t) - 1. \quad (11)$$

Unlike the cubic case, the maximization problem is not turned into a minimization. Similarly, this does not work with any term of even degree. This poses a severe restriction for which function this reduction can be used in the case of degrees higher than 3.

Note that this has nothing to do with the submodularity of the function. The reduction is valid in the cubic case whether the function is submodular or not.

2.2. Reduction by Substitution

However, it has long since been known that the optimization of pseudo-Boolean function of *any* degree can always be reduced to an equivalent problem for quadratic pseudo-Boolean function. The method was proposed by Rosenberg[24] more than 30 years ago; it has since been recalled by Boros and Hammer[3] and, more recently, by Ali et al.[1].

In this reduction, the product xy of two variables x, y in the function is replaced by a new variable z, which is forced to have the same value as xy at any minimum of the function by adding penalty terms that would have a very large value if they don't have the same value.

More concretely, assume that $x, y, z \in \mathbb{B}$ and define

$$D(x,y,z) = xy - 2xz - 2yz + 3z. \quad (12)$$

Then it is easy to check, by trying all eight possibilities, that $D(x,y,z) = 0$ if $xy = z$ and $D(x,y,z) > 0$ if $xy \neq z$. Consider an example pseudo-Boolean function

$$f(x,y,w) = xyw + xy + y. \quad (13)$$

The reduction replaces xy by z and add $MD(x,y,z)$:

$$\tilde{f}(x,y,w,z) = zw + z + y + MD(x,y,z), \quad (14)$$

which has one more variable and is of one less degree than the original function f. Here, M is chosen to be a large positive number so that, whenever $xy \neq z$ and thus $D(x,y,z) > 0$, it is impossible for $\tilde{f}$ to take the minimum.

By repeating the above reduction, any higher-order function can be reduced to a quadratic function with additional variables; for any minimum-energy value-assignment for the new function, the same assignment of values to the original variables gives the minimum energy to the original function.

2.3. The Problem with Reduction by Substitution

Reduction by substitution has not been used very often in practice, because it is difficult to make it work. Although Ali et al.[1] used it with the help of QPBO, they only experimented with second-order potentials.

Note that, according to (12),

$$MD(x,y,z) = Mxy - 2Mxz - 2Myz + 3Mz \quad (15)$$

in (14). The first term Mxy is a quadratic term with a very large positive coefficient. This makes *in all cases* the result of reduction non-submodular, according to the following theorem:

Theorem. Let $E(x_1,\ldots,x_n) = \sum_{i,j} a_{ij}x_ix_j + \sum_i a_ix_i + const.$ and let $x_i \in \mathbb{B}$. Then E is submodular if and only if $a_{ij} \leq 0$ for all i, j. [8]

Thus, the term Mxy in (15) makes the function non-submodular. Though in the case of QPBO submodularity is not the only factor, it seems such an energy cannot be minimized very well even with the QPBO. In our experiments (§5.3) on a third-order energy with this reduction, the QPBO algorithm assigned -1 to most variables, leaving the move-making almost completely stalled.

3. The New Reduction

Our main contribution in this paper is a new reduction of higher-order pseudo-Boolean functions into quadratic ones. It is an expansion of the reduction by minimum selection we described in §2.1.

3.1. Quartic Case

Let us look at the quartic case. We would like to generalize the formulae (3) and (6). Looking at them, one notices

that the whole equation is symmetric in the three variables x, y, and z. Similarly, it stands to reason that if there exists a generalization with $xyzt$ on the LHS, the RHS should also be symmetric in the four variables, i.e.:

$$xyzt = \min_w w(1^{\text{st}} \text{ order sym.}) + (2^{\text{nd}} \text{ order sym.}). \quad (16)$$

Now, it is known that any symmetric polynomial can be written as a polynomial expression in the *elementary symmetric polynomials*. There is only one elementary symmetric polynomial of each degree; the ones we need are:

$$s_1 = x + y + z + t, \quad s_2 = xy + yz + zx + tx + ty + tz. \quad (17)$$

Also, when the variables only take values in $\mathbb{B}$, the square of a variable is the same as itself. Thus, we have ${s_1}^2 = s_1 + 2s_2$, meaning that any symmetric polynomial of second degree or less can be written as a linear combination of s_1, s_2, and 1. Thus the formula should be of the form:

$$xyzt = \min_{w \in \mathbb{B}} w(as_1 + b) + cs_2 + ds_1 + e. \quad (18)$$

We exhaustively searched for integers a, b, c, d, and e that makes the RHS positive only when $x = y = z = t = 1$ and 0 otherwise, to find

$$xyzt = \min_{w \in \mathbb{B}} w(-2s_1 + 3) + s_2 \quad (19)$$

Similarly, we also looked for quintic formula. Failing this, we increased the number of auxiliary variables and found

$$xyztu = \min_{(v,w) \in \mathbb{B}^2} \{v(-2r_1 + 3) + w(-r_1 + 3)\} + r_2, \quad (20)$$

where r_1 and r_2 are the first- and second-degree elementary symmetric polynomials in x, y, z, t, and u. We went on and found similar formulae for degrees six and seven. Then we guessed the general formula given in the following.

3.2. General Case

Now, we introduce similar reductions for general degree. Consider a term $ax_1 \dots x_d$ of degree d.

Again, it is simple if $a < 0$:

Case: $a < 0$

$$ax_1 \dots x_d = \min_{w \in \mathbb{B}} aw\{x_1 + \dots + x_d - (d-1)\}, \quad (21)$$

as suggested by Freedman and Drineas[8].

Case: $a > 0$ No similar reduction has been known in the case $a > 0$. Now, we give a new formula for such a case. This time, we need more than one auxiliary variables if the degree is larger than 4. Specifically, to reduce a term of degree d, we need the following number of auxiliary variables.

$$n_d = \left\lfloor \frac{d-1}{2} \right\rfloor \quad (22)$$

Let us denote $w = (w_1, \dots, w_{n_d}) \in \mathbb{B}^{n_d}$ and

$$S_1 = \sum_{i=1}^{d} x_i, \qquad S_2 = \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} x_i x_j. \quad (23)$$

Then, the reduction is as follows.

Even degree

$$ax_1 \dots x_d = \min_w a \left\{ \sum_{i=1}^{n_d} w_i(-2S_1 + 4i - 1) \right\} + aS_2. \quad (24)$$

Odd degree

$$ax_1 \dots x_d = \min_w a \left\{ \sum_{i=1}^{n_d - 1} w_i(-2S_1 + 4i - 1) + w_{n_d}(-S_1 + 2n_d - 1) \right\} + aS_2. \quad (25)$$

A proof of the correctness of the formulae is given in the Appendix. Note that the cubic case of (25) is different from (6) and simpler.

As we mentioned above, it is known that any pseudo-Boolean function can be written uniquely as a polynomial of binary variables. Since each term in the polynomial can be reduced to a quadratic polynomial using (21), (24), or (25) depending on the sign of the coefficient and the degree of the term, the whole function can be reduced to a quadratic polynomial that is equivalent to the original function in the sense that, if any assignment of values to the variables in the reduced polynomial achieves its minimum, the assignment restricted to the original variables achieves a minimum of the original function.

The number of additional variables per clique in the worst case is exponential in d. This is because there are not only the highest degree term but also lower-degree terms, each of which needs its own new variables. For instance, with a clique of size 5, there can be up to 1 quintic, 5 quartic, and 10 cubic terms, and 17 new variables could be needed. We would guess that six or seven variables in a clique would be about the maximum for vision problems with many such cliques, at least for the time being. Still, it is a significant improvement; even triple cliques can make a big difference relative to pairs, as [29] recently demonstrated. Also, many lower-degree terms would be shared by neighboring cliques.

4. Higher-Order Multi-Label Optimization

In this section, we describe a fusion move algorithm using the reduction in the previous section to optimize higher-order MRF energies with more than two labels. Let V be

the set of pixels and L the set of labels. We consider the following energy on a labeling $I \in L^V$ assigning a label $I_v \in L$ to each pixel $v \in V$:

$$E(I) = \sum_{C \in \mathscr{C}} f_C(I_C), \tag{26}$$

where $\mathscr{C}$ is the set of cliques and $f_C(I_C)$ denotes the local energy depending on the labels $I_C \in L^C$ assigned by the labeling I to the pixels in clique C.

The algorithm maintains the current labeling I. In each iteration, the algorithm fuses I and a proposed labeling $P \in L^V$ by minimizing a pseudo-Boolean energy. For instance, in the α-expansion algorithm, the proposal is a constant labeling with label α everywhere. Here, it can be any labeling and how it is prepared is problem specific, as is how I is initialized at the beginning.

The pseudo-Boolean energy minimized every iteration is as follows. We consider a binary labeling $X \in \mathbb{B}^V$. It consists of a binary variable $X_v \in \mathbb{B}$ for each pixel $v \in V$ that indicates the choice of the value that I_v will have at the end of the iteration. That is, $X_v = 0$ if I_v is to remain the same and $X_v = 1$ if I_v is to change to the proposed label P_v.

Let us denote by $F_C^{I,P}(\beta) \in L^C$ the labeling on clique C that I will have if the value of X on C is $X_C = \beta \in \mathbb{B}^C$:

$$\left(F_C^{I,P}(\beta)\right)_v = (1-\beta_v)I_v + \beta_v P_v. \quad (v \in C) \tag{27}$$

With this notation, we define a pseudo-Boolean function

$$\mathscr{E}(X) = \sum_{C \in \mathscr{C}} \sum_{\beta \in \mathbb{B}^C} f_C\left(F_C^{I,P}(\beta)\right) \theta_C^\beta(X_C), \tag{28}$$

where $\theta_C^\beta(X_C)$ is a polynomial of degree $|C|$ defined by

$$\theta_C^\beta(X_C) = \prod_{v \in C} \left\{\beta_v X_v + (1-\beta_v)(1-X_v)\right\}, \tag{29}$$

which is 1 if $X_C = \beta$ and 0 otherwise.

The polynomial $\mathscr{E}(X)$ is then reduced into a quadratic one using the technique described in the previous section, after which we use the QPBO algorithm to minimize it. We obtain an assignment of $0, 1$, or -1 to each pixel v and update I_v to P_v if 1 is assigned, leaving it unchanged otherwise. We iterate the process until some convergence criterion is met.

5. Experiments

The higher-order BP variants by Lan et al.[17] and Potetz[23] were both tested using a particular higher-order image restoration problem. We also use this problem to compare with them the effectiveness of our fusion algorithm using the reduction of higher-order cliques.

5.1. Image Denoising with Fields of Experts

The image restoration formulation uses the recent image statistic model called Fields of Experts (FoE)[25], which captures complex natural image statistics beyond pairwise interactions by providing a way to learn an image model from natural scenes. FoE has been shown to be highly effective, performing well at image denoising and image inpainting using a gradient descent algorithm. The FoE model represents the prior probability of an image I as the product of several Student's t-distributions:

$$p(I) \propto \prod_C \prod_{i=1}^K \left(1 + \frac{1}{2}(J_i \cdot I_C)^2\right)^{-\alpha_i}, \tag{30}$$

where C runs over the set of all $n \times n$ patches in the image, and J_i is an $n \times n$ filter. The parameters J_i and α_i are learned from a database of natural images. Both in [17] and [23], 2×2 patches were used to show that 2×2 FoE improves over pairwise models significantly.

We are given the noisy image N and find the maximum a posteriori solution given the prior model (30). The prior gives rise to a third-order MRF, with clique potentials that depend on up to four pixels. Since the purpose of the experiments is not the image restoration *per se*, but a comparison of the optimization algorithms, we use exactly the same simple model as in the two predecessors. It is an inference problem with a simple likelihood term: image denoising with a known additive noise. We assume that the images have been contaminated with an i. i. d. Gaussian noise that has a known standard deviation σ. The likelihood of noisy image N given the true image I is assumed to satisfy

$$p(N|I) \propto \prod_{v \in V} \exp\left(-\frac{(N_v - I_v)^2}{2\sigma^2}\right). \tag{31}$$

5.2. Optimization Algorithm

We use the algorithm described in section 4, initializing I by N and iterating until the energy change over 20 iterations drops below a convergence threshold θ_c. For proposal P, we use the following two in alternating iterations: i) a uniform random image created each iteration, and ii) a blurred image, which is made every 30 iterations by blurring the current image I with a Gaussian kernel ($\sigma = 0.5625$). For comparison, we also test the α-expansion proposal, i.e., a constant image that gives the same value α everywhere.

We consider the set $\mathscr{C}$ of cliques consisting of the set of singleton cliques $\mathscr{C}_1 = \{\{v\} \mid v \in V\}$ and the set $\mathscr{C}_4$ of 2×2 patches. Then the local energy is defined by

$$f_{\{v\}}(I_{\{v\}}) = \frac{(N_v - I_v)^2}{2\sigma^2}, \quad (\{v\} \in \mathscr{C}_1) \tag{32}$$

$$f_C(I_C) = \sum_{i=1}^K \alpha_i \log\left(1 + \frac{1}{2}(J_i \cdot I_C)^2\right). \quad (C \in \mathscr{C}_4) \tag{33}$$

(a) (b) (c) (d) (e)

Figure 1. Qualitative difference of denoising using cliques of different orders. (a) Original image, (b) Noise-added image ($\sigma = 20$), (c)(d) images denoised using first-order Potts model and α-expansion with two smoothing factors, (e) denoised using third-order FoE model.

Table 2. PSNR and Energy, average over 10 images using the same FoE model and different optimization.

Noise level	Lan et al.[17]	Potetz[23]	Our result
$\sigma = 10$	30.36 / 40236	31.54 / 36765	31.44 / 35896
$\sigma = 20$	27.05 / 33053	27.25 / 31801	27.43 / 30858

For the QPBO algorithm, we used the C++ code made publicly available by Vladimir Kolmogorov.

5.3. Results

First, Figure 1 shows the qualitative difference between the denoising results using first-order and third-order energies. Some more denoising examples are shown in Figure 2.

For quantitative comparison, we measured the mean energy and the mean peak signal-to-noise ratio (PSNR) for denoising results over the same set of 10 images that were also used in both [17] and [23]. The images are from the Berkeley segmentation database[20], grayscaled and reduced in size, as well as added the Gaussian noise with $\sigma = 10, 20$. The test images and the FoE model was kindly provided by Stefan Roth, one of the authors of [17, 25]. We minimized with $\theta_c = 8$ ($\sigma = 10$) and $\theta_c = 100$ ($\sigma = 20$). The PSNR and energy numbers are listed in Table 2. The energy results show that our fusion move algorithm using the quadratic energy reduced from the third-order energy outperforms the both BP variants optimizing the same energy. The PSNR is also comparable to [23] and better than [17]. Our algorithm takes 8 to 12 minutes ($250 \sim 280$ iterations) to converge on a 2.33GHz Xeon E5345 processor. By comparison, according to [23], it took 30 to 60 minutes on a 2.2GHz Opteron 275, while the algorithm in [17] takes 8 hours on a 3GHz Xeon. Thus, our algorithm outperforms the two predecessors in quality and speed, though BP is considered to be substantially more parallelizable than graph cuts.

Figure 3 shows the behavior of some numbers during the optimization when $\sigma = 20$. It suggests that for higher-order energies, α-expansion does not work very well. The energy never reached the same level as the result using the blur & random proposal. In the case of $\sigma = 10$, α-expansion did go down to about the same energy level, but took significantly longer. In the experiments using the blur & random proposal, the average percentage of the pixels labeled by the QPBO algorithm over two consecutive steps starts around 50% and almost steadily goes up to about 80% when $\sigma = 20$, and from 80% to almost 100% when $\sigma = 10$. Using the α-expansion proposal, the number is always less than 20%. When we use the "reduction by substitution" method explained in §2.2, the percentage stays at almost 0%, with averages 0.00018% and 0.091% over 100 iterations for $\sigma = 20$ and 10, with almost no energy reduction.

6. Conclusions

In this paper, we have introduced a new reduction of any pseudo-Boolean function into an equivalent quadratic one, as well as its combination with the fusion-move and QPBO algorithms for optimization of higher-order multi-label problems. We have validated the technique by minimizing a third-order potential for image denoising. The results show that the algorithm exceeds the preceding BP algorithms in both optimization capability and speed.

Although we focused on the use of this reduction in move-making algorithms, it can of course be used directly for binary-label problems. It can also be used to optimize the binary-label function converted from a multi-label one using the techniques in [28]. It can also be used with other algorithms that minimizes quadratic pseudo-Boolean optimization problems, such as BP and TRW.

We plan to make our research code for the reduction publicly available at our webpage.

Figure 2. More image restoration results. (Left column) Original images. The size is 240 × 160 pixels. (Middle column) Noisy images (σ = 20). PSNR=22.63 and 22.09 from top to bottom. (Right column) Restored images. PSNR=30.92 and 28.29 from top to bottom.

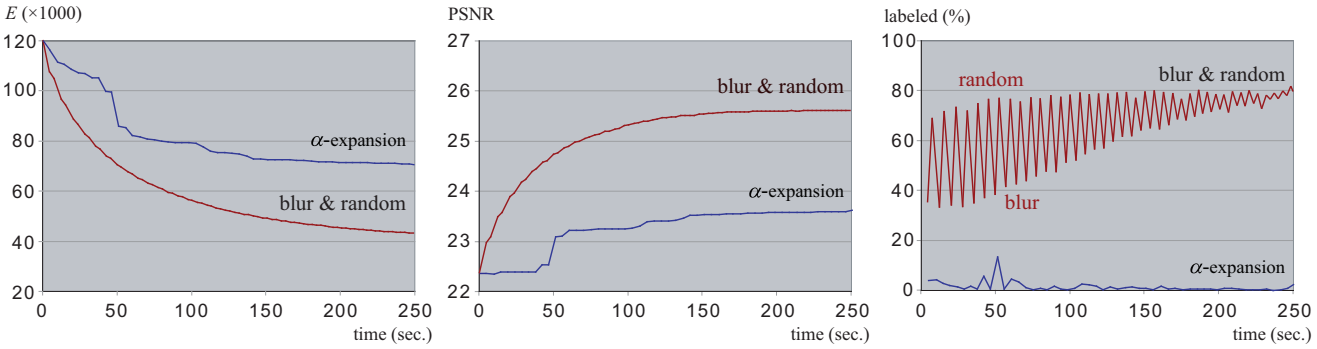

Figure 3. Some plots during the restoration of the example image in Figure 1.

Acknowledgments

We thank Stefan Roth for providing the test images and the FoE models, Brian Potetz for providing his results, and Vladimir Kolmogorov for making his QPBO code publicly available. This work was partially supported by the Kayamori Foundation and the Grant-in-Aid for Scientific Research 19650065 from the Japan Society for the Promotion of Science.

References

[1] A. M. Ali, A. A. Farag, and G. L. Gimel'farb. "Optimizing Binary MRFs with Higher Order Cliques." In *ECCV2008*, III:98–111.

[2] E. Boros, P. L. Hammer, and X. Sun. "Network Flows and Minimization of Quadratic Pseudo-Boolean Functions." RUTCOR Research Report, RRR 17-1991, May 1991.

[3] E. Boros and P. L. Hammer. "Pseudo-Boolean Optimization." *Discrete Appl. Math.* **123**(1-3):155–225, 2002.

[4] E. Boros, P. L. Hammer, and G. Tavares. "Preprocessing of Unconstrained Quadratic Binary Optimization." RUTCOR Research Report, RRR 10-2006, April 2006.

[5] Y. Boykov, O. Veksler, and R. Zabih. "Fast Approximate Energy Minimization via Graph Cuts." *IEEE TPAMI* **23**:1222–1239, 2001.

[6] D. Cremers and L. Grady. "Statistical Priors for Efficient Combinatorial Optimization Via Graph Cuts." In *ECCV2006*, III:263–274.

[7] P. Felzenszwalb and D. Huttenlocher. "Efficient Belief Propagation for Early Vision." *Int. J. Comp. Vis.* **70**:41–54, 2006.

[8] D. Freedman and P. Drineas. "Energy minimization via Graph Cuts: Settling What is Possible." In *CVPR2005* II:939–946.

[9] P. L. Hammer, P. Hansen, and B. Simeone. "Roof Duality, Complementation and Persistency in Quadratic 0-1 Optimization." *Mathematical Programming* **28**:121–155, 1984.

[10] H. Ishikawa. "Exact Optimization for Markov Random Fields with Convex Priors." *IEEE TPAMI* **25**(10):1333–1336, 2003.

[11] H. Ishikawa and D. Geiger. "Rethinking the Prior Model for Stereo." In *ECCV2006*, III:526–537.

[12] P. Kohli, M. P. Kumar, and P. H. S. Torr. "$\mathcal{P}^3$ & Beyond: Move Making Algorithms for Solving Higher Order Functions." *IEEE TPAMI*, to appear, 2008.

[13] V. Kolmogorov. "Convergent Tree-Reweighted Message Passing for Energy Minimization." *IEEE TPAMI* **28** (10): 1568–1583, 2006.

[14] V. Kolmogorov and R. Zabih. "What Energy Functions Can Be Minimized via Graph Cuts?" *IEEE TPAMI* **26**(2):147–159, 2004.

[15] V. Kolmogorov and C. Rother. "Minimizing Non-submodular Functions with Graph Cuts—A Review." *IEEE TPAMI* **29**(7):1274–1279, 2007.

[16] N. Komodakis and N. Paragios "Beyond Pairwise Energies: Efficient Optimization for Higher-order MRFs." In *CVPR2009*.

[17] X. Lan, S. Roth, D. P. Huttenlocher, and M. J. Black. "Efficient Belief Propagation with Learned Higher-Order Markov Random Fields." In *ECCV2006*, II:269–282.

[18] V. Lempitsky, C. Rother, and A. Blake. "LogCut - Efficient Graph Cut Optimization for Markov Random Fields." In *ICCV2007*.

[19] V. Lempitsky, S. Roth, and C. Rother. "FusionFlow: Discrete-Continuous Optimization for Optical Flow Estimation." in *CVPR2008*.

[20] D. Martin, C. Fowlkes, D. Tal, and J. Malik. "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics." In *ICCV2001*, pp. 416–423.

[21] T. Meltzer, C. Yanover, and Y. Weiss. "Globally Optimal Solutions for Energy Minimization in Stereo Vision Using Reweighted Belief Propagation." In *ICCV2005*, pp. 428–435.

[22] R. Paget and I. D. Longstaff. "Texture Synthesis via a Noncausal Nonparametric Multiscale Markov Random Field." *IEEE Trans. on Image Processing*, **7**(6):925–931, 1998.

[23] B. Potetz. "Efficient Belief Propagation for Vision Using Linear Constraint Nodes." In *CVPR2007*.

[24] I. G. Rosenberg. "Reduction of Bivalent Maximization to the Quadratic Case." *Cahiers du Centre d'Etudes de Recherche Operationnelle* 17. pp. 71–74, 1975.

[25] S. Roth and M. J. Black. "Fields of Experts: A Framework for Learning Image Priors." In *CVPR2005*, II:860–867.

[26] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer. "Optimizing Binary MRFs via Extended Roof Duality." In *CVPR2007*.

[27] C. Rother, P. Kohli, W. Feng, and J. Jia. "Minimizing Sparse Higher Order Energy Functions of Discrete Variables." In *CVPR2009*.

[28] S. Ramalingam, P. Kohli, K. Alahari, and P. H. S. Torr. "Exact Inference in Multi-label CRFs with Higher Order Cliques." In *CVPR2008*.

[29] O. J. Woodford, P. H. S. Torr, I. D. Reid, and A. W. Fitzgibbon. "Global Stereo Reconstruction under Second Order Smoothness Priors." In *CVPR2008*.

Appendix: Correctness Proof of the Formulae

Let us suppose that k of the d variables $x_1, \ldots, x_d$, are 1, and the rest are 0. Then, it follows from (23) that

$$S_1 = k, \qquad S_2 = \frac{k(k-1)}{2}. \tag{34}$$

Let us also define

$$m_d = \left\lfloor \frac{d-2}{2} \right\rfloor, \; A = \min_w \sum_{i=1}^{m_d} w_i(-2S_1 + 4i - 1) + S_2. \tag{35}$$

Since the variables w_i can take values independently,

$$A = \sum_{i=1}^{m_d} \min(0, -2k + 4i - 1) + \frac{k(k-1)}{2} \tag{36}$$

If we define $l = \lfloor k/2 \rfloor$, checking the cases where k is even and odd, we have $-2k + 4i - 1 < 0$ if and only if $i \le l$. Thus

$$A = \sum_{i=1}^{\min(l, m_d)} (-2k + 4i - 1) + \frac{k(k-1)}{2} \tag{37}$$

In particular, if $k \le d - 2$, it follows $l \le m_d$, which in turn means that the sum in (37) is up to l, and checking the cases where k is even and odd, we obtain $A = 0$.

Now, first we consider the even degree case (24), of which the RHS is aA. Thus, both sides are 0 if $k \le d - 2$. If $k = d - 1$, $A = 0$ similarly follows from $l = m_d$. If $k = d$, it follows from $m_d = l - 1, k = 2l$, and (37) that

$$A = \sum_{i=1}^{l-1} (-4l + 4i - 1) + l(2l - 1) = 1, \tag{38}$$

which completes the proof of (24).

As for the odd degree case (25), we have $m_d = n_d - 1$ and that the RHS is $a\,(A + \min(0, -S_1 + 2n_d - 1))$. Since $d = 2n_d + 1$, it follows $-S_1 + 2n_d - 1 \ge 0$ if $k \le d - 2$, showing that the both sides of (25) are 0. If $k = d - 1$, $-S_1 + 2n_d - 1 = -k + k - 1 = -1$. It also follows from $m_d = l - 1, k = 2l$, and (37) that $A = 1$, showing that the both sides of (25) are 0. Finally, if $k = d$, we have $-S_1 + 2n_d - 1 = -k + k - 1 - 1 = -2$ and, from $m_d = l - 1, k = 2l + 1$, and (37),

$$A = \sum_{i=1}^{l-1} (-4l + 4i - 3) + l(2l + 1) = 3, \tag{39}$$

which shows that both sides of (25) are a. □