

変数フリップによる高階グラフカットの拡張

石川 博

名古屋市立大学 大学院システム自然科学研究科

〒 467-8501 名古屋市瑞穂区瑞穂町字山の畑 1

E-mail: hfs@ieee.org

あらまし 高階グラフカットは、ブール値変数についての任意階のエネルギーを同値な 1 階のエネルギーに変換し QPBO などのグラフカット手法を使って最小化することを繰り返すことで、任意階の多値エネルギーをも最小化する。本稿では、任意階ブール値変数エネルギーの 1 階への変換アルゴリズムを、変数の 0 と 1 を入れ替えるという単純な操作(フリップ)を変換の前後に行うことによって、エネルギー中に現れる変数の数について指数的な数の新しい変換を与える。また、可能な変換の中から、特定の場合を実験により検証し、従来の変換より性能が改善されることを示す。さらに、知られている高階エネルギー最小化法のいくつかは、この拡張アルゴリズムの特殊な場合として統一的に理解できることも示す。

キーワード グラフカット, マルコフ確率場, 最適化, 高階, 高次, フリップ

Flipping Variables to Extend Higher-Order Graph Cuts

Hiroshi Ishikawa

Department of Information and Biological Sciences, Nagoya City University

1 Yamanohata, Mizuho, Nagoya 467-8501, Japan.

E-mail: hfs@ieee.org

Abstract Higher-order graph cuts minimize multi-label higher-order energies by repeatedly minimizing a binary higher-order energy, converting it into a first-order one and then using graph-cut techniques such as QPBO. In this paper, we give an exponential number of different higher-to-first-order conversions in terms of the number of occurrences of the variables in the energy by flipping variables before and after the existing conversion. We experimentally show that a certain specific conversion outperforms the existing conversion. We also show that some known higher-order energy minimization methods can be understood as a special case of this generalized formulation.

Key words Graph cut, Markov Random Field, Optimization, Higher order, Flipping variables

1. はじめに

領域分割 [14], [17], ステレオ [18], 画像復元 [16] などの問題を, MRF などの確率モデルとして定式化し, エネルギー最小化問題として解く方法, 特にグラフカット [3][6][12], 信念伝播法 (BP) [4][13], またツリー重み再配分メッセージ伝達法 (TRW) [10] などの高効率なエネルギー最小化アルゴリズムが最近一般化している。

エネルギーが最高 k 個の点に依存する関数の和で書けると, $k-1$ 階のエネルギーと呼ぶが, [7] ではそれまで最大 2 階に限られ, また実際にはほとんど 1 階, つまり各画素と隣接 2 画素のみに依存する項の和として表されたエネルギーしか使われてこなかったグラフカットによるエネルギー最小化を, 任意階に適用可能な方法に一般化した。その核心は, 任意階のブール値変数エネルギーを同値な 1 階エネルギーに変換する方法であり, そ

れを QPBO 及び融合移動と組み合わせ, 高階の多値エネルギーのグラフカットによる最適化を可能にした。

本稿では, この任意階ブール値変数エネルギーの 1 階への変換アルゴリズムを, 変数の 0 と 1 を入れ替えるという単純な操作(フリップ)を変換の前後に行うことによって一般化する。これにより, 原理的にエネルギー中に現れる変数の数について指数的な数だけの異なる変換を得るが, 本稿ではそのうち特定の場合を実験により検証し, 既存の変換より性能が改善されることを示す。また, 特殊な形の高階エネルギーをグラフカットによって最小化する方法がこの数年の間に提案されているが, そのいくつかは本稿で一般化されたアルゴリズムの特殊な場合と考えられることも示す。

2. 背景

本節ではまず高階グラフカット [7] について説明する。

2.1 エネルギー最小化

例えば領域分割は、領域を表すラベルを各画素に一つずつ割り振る問題と考えることができる。ラベル付け X に対して実数値 $E(X)$ を与える関数をエネルギーと呼ぶ。エネルギーが小さいほど X が表す領域分割がより望ましいように $E(X)$ を定義し、それを最小化する X を探すのが、エネルギー最小化による定式化である。ラベル付け X が画素 v に与えるラベルを X_v とする。よく使われるのは、エネルギーが各画素及び隣接する2画素の組のラベルのみに依存する関数の和として書き表される次式の形で、これを1階のエネルギーと呼ぶ：

$$E(X) = \sum_{v \in V} f_v(X_v) + \sum_{\{u,v\} \in \mathcal{E}} f_{uv}(X_u, X_v).$$

ここで V は画素の、 \mathcal{E} は隣接画素の組の集合である。また3画素までの組のある集合を \mathcal{C} とし、右辺を

$$\sum_{\{v\} \in \mathcal{C}} f_v(X_v) + \sum_{\{u,v\} \in \mathcal{C}} f_{uv}(X_u, X_v) + \sum_{\{u,v,w\} \in \mathcal{C}} f_{uvw}(X_u, X_v, X_w)$$

の形としたものは、2階のエネルギーと呼ばれる。

以下、これを任意階で定式化する。画素などに代表される「位置」の抽象化をサイトと呼び、その集合を V とする。いくつかのサイトの組、つまり V の部分集合をクリークと呼ぶ。クリークの集合 \mathcal{C} 、ラベルの集合 L を固定する。ラベル付け、つまり V から L への写像を以下では配置と呼び、その集合を L^V で表す。エネルギーとは配置 X に対して実数値 $E(X)$ を与える関数である。任意のクリーク C に対して、配置 X は C 上の配置 $X_C = (X_v)_{v \in C} \in L^C$ 、つまり C の各サイトに X により与えられたラベルの組を決める。エネルギーは次のような形に分解できると仮定する：

$$E(X) = \sum_{C \in \mathcal{C}} f_C(X_C), \quad (1)$$

ただし $f_C(X_C)$ は X_C のみ^(注1)に依存する関数である。 \mathcal{C} 内最大のクリークのサイト数が k のとき、 $k-1$ 階のエネルギーと呼ぶ。特に、ラベルの集合 L がブール値の集合 $\mathbb{B} = \{0, 1\}$ のとき、ブール値変数のエネルギーと呼ぶ。以下、 \mathbb{R} は実数の集合を表す。

2.2 高階グラフカット

高階グラフカットアルゴリズムは、クリークの集合 \mathcal{C} 、式(1)の形のエネルギー $E(X)$ 、配置 X の初期値、提案配置の生成法、及び繰り返しの終了条件を与えられて、 $E(X)$ を近似的に最小化する X を求める。

配置 X を変数として、提案配置 P を与えられた生成法により繰り返し生成し、 X と融合して $E(X)$ を最小化する X を探す。提案配置の生成方法は任意だが、その選択

(注1)：もちろん他にも、与えられたデータなど、最適化の過程で変化しないものには依存する。

にアルゴリズムの性能が依存する。

融合は、サイト v 毎に現在値 X_v が提案値 P_v を選ぶことで決まる。その選択を与えるブール値の配置 Y を考える。つまり、 Y に従って v 毎に $Y_v = 0$ または 1 だが、

$$X'_v \leftarrow \begin{cases} X_v & (Y_v = 0 \text{ のとき}) \\ P_v & (Y_v = 1 \text{ のとき}) \end{cases} \quad (2)$$

によって X を更新した配置 X' を得る。そのエネルギー $E(X')$ は Y の関数 $\mathcal{E}(Y)$ として以下のように与えられる。

クリーク C における X' 、すなわち X'_C は、 Y の与える C 上の配置 Y_C によって決まるが、 $Y_C = \gamma \in \mathbb{B}^C$ であるときその成分は次式で与えられる：

$$(x_C^{X,P}(\gamma))_v = \begin{cases} X_v & (\gamma_v = 0 \text{ のとき}) \\ P_v & (\gamma_v = 1 \text{ のとき}) \end{cases} \quad (3)$$

すると Y_C の変数 $Y_v (v \in C)$ についての $|C|$ 次の多項式

$$Y_C^{(\gamma)} = \prod_{v \in C} \{\gamma_v Y_v + (1 - \gamma_v)(1 - Y_v)\},$$

は $Y_C = \gamma$ ならば 1 、それ以外の場合 0 である。これらの記号を使って Y の変数 $Y_v (v \in V)$ についての次の多項式を定義する：

$$\mathcal{E}(Y) = \sum_{C \in \mathcal{C}} \mathcal{E}_C(Y_C), \quad (4)$$

$$\mathcal{E}_C(Y_C) = \sum_{\gamma \in \mathbb{B}^C} f_C(x_C^{X,P}(\gamma)) Y_C^{(\gamma)} \quad (5)$$

すると(5)右辺の和は、 Y_C の取りうる各ベクトル値 γ につき、 $Y_C = \gamma$ のときのみ $f_C(x_C^{X,P}(\gamma))$ になり、それ以外のは 0 になる項の和である。この値は、 $Y_C = \gamma$ から(2)によって決まる X'_C の局所エネルギー $f_C(X'_C)$ の値である。従って多項式 $\mathcal{E}(Y)$ は Y に従って(2)によって変化した X' のエネルギー $E(X')$ を与える。

このブール値変数エネルギー $\mathcal{E}(Y)$ を §2.4 で述べる方法により2次に変換し、グラフカットアルゴリズムの一種であるQPBOアルゴリズム[2][11]に与えると、最小解の一部が得られるので、 $Y_v = 1$ が与えられれば X_v を P_v に変え、それ以外なら変えない。QPBOアルゴリズムの性質から、この操作によってエネルギーは増加しないことが保障される。以上を終了条件が満たされるまで繰り返す。

2.3 擬ブール関数

一般に、ブール値をとる変数の実数値関数 $f: \mathbb{B}^n \rightarrow \mathbb{R}$ を擬ブール関数という。擬ブール関数の性質として以下が知られている([5]参照)。まず、変数の番号の集合を $U = \{1, \dots, n\}$ とすると、任意の擬ブール関数は多項式の形に一意に表される：

$$f(x_1, \dots, x_n) = \sum_{S \subset U} c_S \prod_{i \in S} x_i$$

ここで c_S は実数係数である。つまり、ブール値変数エネルギーは常に多項式に変換できる。また、 $k-1$ 階のエネルギーは k 次の多項式で表される。2 次の擬ブール関数は、展開したとき 2 次の項の係数がすべて負であるとき劣モジユラであると言われる。このとき、その関数はグラフカットによって大域的に最小化することができる。

2.4 階数変換

上の $\mathcal{E}(Y)$ のようなブール値変数についての一般の関数を 1 階の関数に変換、つまり最大 2 つの変数に依存する項の和の形に書き表す変換が、高階グラフカット [7] の核心である。本稿ではこの変換を拡張するのであるが、まず [7] における変換を説明する。

ここでいう階数変換とは、任意次数の多項式に新たに付加変数を加えて同値な 2 次の多項式に変換することで、階数を 1 階に下げる変換である。ここで同値とは、変換後の多項式が最小値をとる変数への値の与え方があれば、その値を元からある変数に与えることによって元の多項式の最小値を得ることができるという意味である。

そのため [7] では、多項式を展開して、各単項式 $ax_1 \cdots x_d$ を 2 次の多項式に置き換える次の変換を与えた。

$a < 0$ のとき

$$ax_1 \cdots x_d = \min_{w \in \mathbb{B}} aw\{S_1 - (d-1)\} \quad (6)$$

$a > 0$ のとき

$$ax_1 \cdots x_d = \min_{w_1, \dots, w_d \in \mathbb{B}} a \sum_{i=1}^{n_d} w_i (c_{i,d}(-S_1 + 2i) - 1) + S_2 \quad (7)$$

ただし、

$$n_d = \left\lfloor \frac{d-1}{2} \right\rfloor, \quad c_{i,d} = \begin{cases} 1 & d \text{ が偶数で } i = n_d, \\ 2 & \text{それ以外.} \end{cases}$$

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

である。この公式によれば、例えば $xyzt$ は

$$\min_{w \in \mathbb{B}} w(-2(x+y+z+t)+3)+xy+yz+zx+tx+ty+tz \quad (8)$$

に変換される。

3. 階数変換の拡張

本節では、変換の前後に変数フリップをすることにより、上に述べた階数変換を拡張する。この拡張によって、非常に多数の異なる変換を得ることができる。

3.1 変数フリップ

ブール値 $b \in \mathbb{B}$ について、 b の否定、つまり $1-b$ を \bar{b} で表すことにする。変数の否定を混ぜることによって、

同じ擬ブール関数を様々な書き方で表すことができる：

$$\begin{aligned} xyzt &= (1-\bar{x})yzt = yzt - \bar{x}yzt \\ &= yzt - \bar{x}zt + \bar{x}\bar{y}zt = zt - \bar{y}zt - \bar{x}zt + \bar{x}\bar{y}zt. \end{aligned}$$

このような操作を変数フリップと呼ぶことにする。

本稿の主な貢献は、任意階の擬ブール関数を 1 階の変数に変換する前に、変数の一部をフリップし、変換後に再びフリップして元の変数に戻すことによって、新しい変換公式を得るという単純なアイデアである。

ここで注意する必要があるのは、当然ながら、ある変数 x とそれをフリップした \bar{x} は独立でないことである。一方、QPBO などの最小化アルゴリズムは、全ての変数に自由に値を与えられることを前提としている。従って、関数を最小化する前に、各変数 x について、 x か \bar{x} のどちらかしか関数中に現れないようにする必要がある。

公式 (6) および (7) の左辺の単項式 $ax_1 \cdots x_d$ は、変数 x_1, \dots, x_d の全てが 1 であるときだけ 0 でない値をとる関数である。これを 0 と 1 の任意の組み合わせに拡張する。

任意のブール値ベクトル $\gamma = (\gamma_1, \dots, \gamma_d) \in \mathbb{B}^d$ について、次のように $\{1, \dots, d\}$ の部分集合 I_γ^0, I_γ^1 を定義する。

$$I_\gamma^0 = \{i \mid \gamma_i = 0\}, \quad I_\gamma^1 = \{i \mid \gamma_i = 1\}. \quad (9)$$

次に、 $x = (x_1, \dots, x_d)$ をブール値変数のベクトルとし、 $x^{(\gamma)} = (x_1^{(\gamma)}, \dots, x_d^{(\gamma)})$ を x と γ から次のように定義されるブール値変数のベクトルとする。

$$x_i^{(\gamma)} = \begin{cases} x_i & \gamma_i = 1 \text{ のとき} \\ \bar{x}_i & \gamma_i = 0 \text{ のとき} \end{cases}$$

この $x^{(\gamma)}$ を、 γ -フリップした変数ベクトル x と呼ぶ。すると、単項式 $ax_1^{(\gamma)} \cdots x_d^{(\gamma)}$ は、 $x = \gamma$ のとき値 a を、それ以外のとき値 0 をとる：

$$ax_1^{(\gamma)} \cdots x_d^{(\gamma)} = \begin{cases} a & x = \gamma \text{ のとき,} \\ 0 & \text{それ以外} \end{cases}$$

すると、任意の擬ブール関数 $f: \mathbb{B}^d \rightarrow \mathbb{R}$ について、

$$f(x) = \sum_{\gamma \in \mathbb{B}^d} f(\gamma) x_1^{(\gamma)} \cdots x_d^{(\gamma)}$$

が成り立つ。さらに、任意の $\lambda \in \mathbb{R}$ について

$$f(x) = \lambda + \sum_{\gamma \in \mathbb{B}^d} (f(\gamma) - \lambda) x_1^{(\gamma)} \cdots x_d^{(\gamma)} \quad (10)$$

も成り立つ。特に、

$$\lambda = \max_{\gamma \in \mathbb{B}^d} f(\gamma) \quad (11)$$

と取ることによって、(10) の係数を全て 0 以下にすることができる。こうすることにより、任意の擬ブール関数を変換 (6) のみを使って (7) を使わずに 2 次擬ブール関数に変換することができる。次節で示すように、Rother [15] の変換は実はこれである。

表 1 階数変換に必要な付加変数の数.

次数	d	3	4	5	6	7	8	9
多項式として変換 (最悪のケース)	$N^{\text{all}^+}(d)$	1	5	17	49	129	321	769
多項式として変換 (平均)	$(N^{\text{all}^+}(d) + N^{\text{all}^-}(d))/2$	1	5	16.5	45.5	114	270	617.5
全て負の係数((10)式)としてから変換	$2^d - 1$	7	15	31	63	127	255	511

3.2 拡張公式

次式で γ -フリップした S_1, S_2 を定義する:

$$S_1^{(\gamma)} = \sum_{i=1}^d x_i^{(\gamma)} = \sum_{i \in I_\gamma^1} x_i + \sum_{i \in I_\gamma^0} \bar{x}_i = - \sum_{i=1}^d (-1)^{\gamma_i} x_i + |I_\gamma^0|$$

$$S_2^{(\gamma)} = \sum_{i=1}^{d-1} \sum_{j=i+1}^d x_i^{(\gamma)} x_j^{(\gamma)} = \frac{S_1^{(\gamma)}(S_1^{(\gamma)} - 1)}{2}$$

すると, 拡張された階数変換公式は次のようになる:

$a < 0$ のとき

$$ax_1^{(\gamma)} \cdots x_d^{(\gamma)} = \min_{w \in \mathbb{B}} aw \{S_1^{(\gamma)} - (d-1)\} \quad (12)$$

$$= \min_{w \in \mathbb{B}} aw \left\{ - \sum_{i=1}^d (-1)^{\gamma_i} x_i - |I_\gamma^1| + 1 \right\}$$

$a > 0$ のとき

$$ax_1^{(\gamma)} \cdots x_d^{(\gamma)} = \min_{w \in \mathbb{B}^{n_d}} a \sum_{i=1}^{n_d} w_i (c_{i,d}(-S_1^{(\gamma)} + 2i) - 1) + aS_2^{(\gamma)} \quad (13)$$

これは, フリップした変数も 0 か 1 の値をとることは同じであり, 普通の変数として公式 (6) および (7) を適用できることからわかる.

3.3 付加変数の数

次に, これらの変換によって d 次の多項式を階数変換するときに必要な付加変数の数を考える. d 変数 d 次の多項式を展開したときの単項式の数, i 次の項は $\binom{d}{i}$ 個であり, 定数項まで数えて合計 2^d 個である. (ブール値のため同じ変数の冪乗は等しいことに注意する.) 3 次以上の項の数は $2^d - 1 - \frac{d(d+1)}{2}$ 個であり, 次数 d とともに指数的に増える. 3 次以上の項を 2 次に変換するには少なくとも一つの新しい変数を増やす必要があるため, 必要な付加変数の数も次数について指数的に増加する.

上述のように, (10) 式において λ を (11) によって定めれば, 任意の擬ブール関数を係数が全て 0 以下のフリップされた単項式の和として表すことができ, これを変換 (12) のみを使って 2 次擬ブール関数に変換することができる. その場合, 3 次以上の項それぞれに一つの付加変数で済むので, 係数のうち一つは常に 0 になることから, 必要な付加変数の数は $2^d - 1$ である.

一方, 擬ブール関数を多項式として展開して [7] の変換 (6), (7) を使った場合, 最悪の場合, 全ての係数が正で

あるとすると, 変換 (7) に必要な新しい変数の数は

$$N^{\text{all}^+}(d) = \sum_{i=3}^d \binom{d}{i} \left\lfloor \frac{i-1}{2} \right\rfloor = 2^{d-2}(d-3) + 1$$

であり, 逆に全ての係数が負であれば

$$N^{\text{all}^-}(d) = \sum_{i=3}^d \binom{d}{i} = 2^d - 1 - \frac{d(d+1)}{2}$$

である. 平均的なケースでは, 係数の半分は正, 半分は負であるとする, 表 1 に示すように, $d \leq 7$ の場合には, 多項式として表して変換する方が, (10) のように表すより平均して少ない数の付加変数ですむ.

3.4 変換の多様性

任意のブール値変数エネルギーを多項式として表し, その中に現れる変数 x それぞれについて, $1 - \bar{x}$ に置き換えることができる. それを (12) と (13) によって 1 階に変換し, 最後にフリップされた変数 \bar{x} を $1 - x$ で置き換えれば, 元のエネルギーの 1 階への変換ができる.

例えば, $xyzt$ は次のようにも変換できる.

$$xyzt = -\bar{x}yzt - \bar{y}zt + zt$$

$$= \min_{u, w \in \mathbb{B}} \{-u(\bar{x} + y + z + t - 3) - w(\bar{y} + z + t - 2) + zt\}$$

$$= \min_{u, w \in \mathbb{B}} \{-u(-x + y + z + t - 2) - w(-y + z + t - 1) + zt\}.$$

これは従来の変換 (8) とは明らかに異なる.

同じ変数でも, 出てくる場所ごとにフリップしてもしなくてもよい:

$$xyz - 2xzt = (1 - \bar{x})yz - 2xzt$$

$$= \min_{v, w \in \mathbb{B}} \{yz - v(\bar{x} + y + z - 2) - 2w(x + z + t - 2)\}$$

$$= \min_{v, w \in \mathbb{B}} \{yz - v(-x + y + z - 1) - 2w(x + z + t - 2)\}$$

このように, エネルギー中に現れる各変数について, フリップするかしないかを選択できるため, 変数の現れる数について指数的な数だけの異なる変換を得ることができる.

階数変換に求められることは, できるだけ最適化が容易な関数に変換することと, できるだけ変数の増加をしないで変換することである. つまり, それらの意味で最適な変換を見つけないという, 新たな問題が現れたことになる. そのような最適な変換を見つけないアルゴリズム,

あるいは少なくともヒューリスティックが望まれるところだが、その解決は将来の研究が待たれる。本稿では代わりに、次節では知られているいくつかのアルゴリズムとの関係を明らかにし、また次々節では特定の場合の性能を実験的に評価する。

4. 既存手法との関係

ここ数年の間に、限られた形の高階エネルギーについていくつかの最小化法が提案されてきた。本節では、それらのうち、前節で定義した変換方法によって拡張された高階グラフカットの特殊な場合として記述できるものについて関係を述べる。すなわち、Rotherら [15] によるブール値変数エネルギー変換と、 \mathcal{P}^n Potts モデルと呼ばれる Kohliら [8] による方法である。

一方、Rotherら [15] の選択関数を使った多値最小化および Kohliら [9] のロバスト \mathcal{P}^n Potts モデルについては、同様の記述はできないようである。

4.1 Rotherらによる Type-I および Type-II 変換

§3.1 で述べたように、 λ を (11) によって定めれば、和 (10) 中の係数を全て 0 以下にすることができる。従って、任意の擬ブール関数を変換 (12) のみで 2 次に変換することができる。実はこれが Rotherら [15] による Type-I および Type-II 変換と同一であることを以下に示す。

実数 $\theta > 0$ 、ブール値ベクトル $\gamma \in \mathbb{B}^d$ について、擬ブール関数 $\psi(x) : \mathbb{B}^d \rightarrow \mathbb{R}$ を次のように定義する。

$$\psi(x) = \theta - \theta x_1^{(\gamma)} \cdots x_d^{(\gamma)} = \begin{cases} 0 & x = \gamma \text{ のとき} \\ \theta & \text{それ以外} \end{cases}$$

4.1.1 Type-I 変換

本稿で提案した変換 (12) を使って、 $\psi(x)$ を次のように変換できる。

$$\begin{aligned} \psi(x) &= \theta + \theta \min_{w \in \mathbb{B}} w \left\{ \sum_{i=1}^d (-1)^{\gamma_i} x_i + |I_\gamma^1| - 1 \right\} \\ &= \theta + \theta \min_{w \in \mathbb{B}} w \left\{ \sum_{i \in I_\gamma^0} x_i + \sum_{i \in I_\gamma^1} (1 - x_i) - 1 \right\} \\ &= \theta \min_{w \in \mathbb{B}} \left\{ 1 + vw + (1 - v) \sum_{i \in I_\gamma^0} x_i + w \left(\sum_{i \in I_\gamma^1} (1 - x_i) - 1 \right) \right\} \end{aligned}$$

ここで $v = \bar{w} = 1 - w$ と定義し、定義 (9) と $vw = 0$ であることを注意する。

ブール値ベクトル $x \in \mathbb{B}^d$ とブール値 $v, w \in \mathbb{B}$ について、

$$\varphi(x, v, w) = 1 + vw + (1 - v) \sum_{i \in I_\gamma^0} x_i + w \left(\sum_{i \in I_\gamma^1} (1 - x_i) - 1 \right)$$

と定義すると

$$\varphi(x, 0, 0) = 1 + \sum_{i \in I_\gamma^0} x_i \geq 1 = \varphi(x, 1, 0)$$

$$\varphi(x, 1, 1) = 1 + \sum_{i \in I_\gamma^1} (1 - x_i) \geq 1 = \varphi(x, 1, 0)$$

から、 $\varphi(x, v, w)$ の最小値が $v = w$ で得られるときは $\varphi(x, 1, 0)$ も最小値をとることがわかる。つまり、 $\varphi(x, v, w)$ の最小値は v と w に常に逆の値を持たせても得られ、2 つの変数 v, w を持つ自由度は最小値をより小さくする効果を持たないことがわかる。従って、 $v = \bar{w}$ と固定しても最小値は変わらないので

$$\psi(x) = \theta \min_{w \in \mathbb{B}} \varphi(x, \bar{w}, w) = \theta \min_{v, w \in \mathbb{B}} \varphi(x, v, w)$$

を得る。これは [15] で Type-I 変換と呼ばれている

$$\psi(x) = \min_{v, w \in \mathbb{B}} \theta \left\{ v + (1 - w) - v(1 - w) + \sum_{i \in I_\gamma^0} (1 - v)x_i + \sum_{i \in I_\gamma^1} w(1 - x_i) \right\}$$

と一致する。

4.1.2 Type-II 変換

一方、[15] において Type-II 変換は次で定義される：

$$\begin{aligned} \psi(x) &= \theta + \frac{\theta}{2} \min_{w \in \mathbb{B}} \left\{ w(d - 2) + w \left(\sum_{i \in I_\gamma^1} (1 - x_i) - \sum_{i \in I_\gamma^0} (1 - x_i) \right) \right. \\ &\quad \left. + (1 - w) \left(\sum_{i \in I_\gamma^1} x_i - \sum_{i \in I_\gamma^0} x_i \right) + \sum_{i \in I_\gamma^0} x_i - \sum_{i \in I_\gamma^1} x_i \right\} \end{aligned}$$

これは次のように書き直すことができる。

$$\psi(x) = \theta + \frac{\theta}{2} \min_{w \in \mathbb{B}} w \left\{ d - 2 + \sum_{i=1}^d (-1)^{\gamma_i} (2x_i - 1) \right\} \quad (14)$$

拡張された変換 (12) を使って、 $\psi(x) = \theta - \theta x_1^{(\gamma)} \cdots x_d^{(\gamma)}$ は次のように書きなおすことができる。

$$\begin{aligned} \psi(x) &= \theta + \theta \min_{w \in \mathbb{B}} w \left\{ \sum_{i=1}^d (-1)^{\gamma_i} x_i + |I_\gamma^1| - 1 \right\} \\ &= \theta + \frac{\theta}{2} \min_{w \in \mathbb{B}} w \left\{ \sum_{i=1}^d (-1)^{\gamma_i} \cdot 2x_i + |I_\gamma^1| + (d - |I_\gamma^0|) - 2 \right\} \\ &= \theta + \frac{\theta}{2} \min_{w \in \mathbb{B}} w \left\{ \sum_{i=1}^d (-1)^{\gamma_i} (2x_i - 1) + d - 2 \right\} \end{aligned}$$

これは (14) と一致する。

以上のように、[15] の Type-I および Type-II 変換はともに変換 (12) と同一である。

4.2 Kohliらによる \mathcal{P}^n Potts モデル

Kohliら [8] は

$$f_C(X_C) = \begin{cases} \mu_l & \text{全ての } v \in C \text{ で } X_v = l \text{ のとき} \\ \mu_{\max} & \text{それ以外} \end{cases} \quad (15)$$

の形の高階エネルギーを定義し \mathcal{P}^n Potts モデルと呼び、この形のエネルギーは α - β 交換および α 拡張で効率よく最小化できることを示した。ここで $\mu_l \in \mathbb{R}$ は各ラベル $l \in L$ について $\mu_{\max} \geq \mu_l$ を満たすように定義される。

エネルギー (1) が (15) で定義されるときに高階グラフカットで最小化することを考える。またそのとき、提案配置は α - β 交換および α 拡張 [3] を「エミュレート」することにする。以下では、定数および 1 次の項は省略する。

4.2.1 α - β 交換

α - β 交換アルゴリズムは、高階グラフカットアルゴリズムによって次のように「エミュレート」することができる。現在の配置が $X \in L^V$ であるとして、二つのラベル $\alpha, \beta \in L, \alpha \neq \beta$ について提案配置を各 $v \in V$ に

$$P_v = \begin{cases} \alpha & X_v = \beta \text{ のとき} \\ \beta & X_v = \alpha \text{ のとき} \\ X_v & \text{それ以外} \end{cases}$$

を与える配置として定義する。

クリーク C 上のブール値の配置を $\gamma \in \mathbb{B}^C$ として C の部分集合

$$C_l = \{v \in C \mid X_v = l\} \quad (l \in L), \\ C_\gamma^0 = \{v \in C \mid \gamma_v = 0\}, \quad C_\gamma^1 = \{v \in C \mid \gamma_v = 1\}$$

を定義する。

γ に従った C 上の融合後の配置 $x_C^{X,P}(\gamma)$ は (3) によって定義される。クリーク C に対するエネルギーの値は

i) $C_\alpha \cup C_\beta = C$ ならば

$$f_C(x_C^{X,P}(\gamma)) = \begin{cases} \mu_\alpha & C_\gamma^0 = C_\alpha \text{ かつ } C_\gamma^1 = C_\beta \text{ のとき} \\ \mu_\beta & C_\gamma^1 = C_\alpha \text{ かつ } C_\gamma^0 = C_\beta \text{ のとき} \\ \mu_{\max} & \text{それ以外} \end{cases}$$

ii) $C = C_l, l \neq \alpha, l \neq \beta$ ならば、全ての $v \in C$ について $P_v = X_v = l$ だから

$$f_C(x_C^{X,P}(\gamma)) = \mu_l$$

iii) それ以外のとき、 α - β 交換後にクリーク全体が同じラベルになることはないから

$$f_C(x_C^{X,P}(\gamma)) = \mu_{\max}$$

定数でない i) の場合を考えると、高階グラフカットのブール値変数エネルギー (4) で

$$\mathcal{E}_C(Y_C) = \mu_{\max} + \delta_\alpha \prod_{v \in C_\alpha} \bar{Y}_v \prod_{v \in C_\beta} Y_v + \delta_\beta \prod_{v \in C_\alpha} Y_v \prod_{v \in C_\beta} \bar{Y}_v$$

とすれば上のエネルギーが表されることがわかる。ただし $\delta_\alpha = \mu_\alpha - \mu_{\max}, \delta_\beta = \mu_\beta - \mu_{\max}$ である。

$\delta_\alpha \leq 0, \delta_\beta \leq 0$ だから、このエネルギーは拡張された階数変換 (12) によって

$$\delta_\alpha \prod_{v \in C_\alpha} \bar{Y}_v \prod_{v \in C_\beta} Y_v = \min_{u_C \in \mathbb{B}} \delta_\alpha u_C \left(1 - |C_\beta| - \sum_{v \in C_\alpha} Y_v + \sum_{v \in C_\beta} Y_v \right) \\ \delta_\beta \prod_{v \in C_\alpha} Y_v \prod_{v \in C_\beta} \bar{Y}_v = \min_{w_C \in \mathbb{B}} \delta_\beta w_C \left(1 - |C_\alpha| + \sum_{v \in C_\alpha} Y_v - \sum_{v \in C_\beta} Y_v \right)$$

と変換できる。この変換によって最大でクリーク数の 2 倍の変数が増えられるが、実際には多くのクリークは上の ii) または iii) に該当するため、それよりは付加変数の数は少なくなる。

上の形のエネルギーは劣モジュラではないが、 $X_v = \beta$ である v についてのみ Y_v をフリップしてエネルギーを

$$\mathcal{E}_C(Y_C) = \mu_{\max} + \min_{u_C \in \mathbb{B}} \delta_\alpha u_C \left(1 - \sum_{v \in C_\alpha} Y_v - \sum_{v \in C_\beta} \bar{Y}_v \right) \\ + \min_{w_C \in \mathbb{B}} \delta_\beta w_C \left(1 - |C| + \sum_{v \in C_\alpha} Y_v + \sum_{v \in C_\beta} \bar{Y}_v \right)$$

とすることで、大域的最適化可能になる。

4.2.2 α 拡張

次に α 拡張を「エミュレート」する提案配置を考える。つまり、ある $\alpha \in L$ と全ての $v \in V$ について

$$P_v = \alpha$$

とする。現在の配置が $X \in L^V$ であるとして、

$$C_{\bar{\alpha}} = \{v \in C \mid X_v \neq \alpha\}$$

とする。(3) によって定義される $\gamma \in \mathbb{B}^C$ に従ったクリーク C 上の融合後の配置 $x_C^{X,P}(\gamma)$ に対するエネルギーの値は

$$f_C(x_C^{X,P}(\gamma)) = \begin{cases} \mu & \text{全ての } v \in C \text{ で } \gamma_v = 0 \text{ のとき} \\ \mu_\alpha & \text{全ての } v \in C_{\bar{\alpha}} \text{ で } \gamma_v = 1 \text{ のとき} \\ \mu_{\max} & \text{それ以外} \end{cases}$$

となる。ただし、 μ はクリーク C 上の現在の配置 X_C に従って次のように定義される。

$$\mu = \begin{cases} \mu_l & \text{全ての } v \in C \text{ で } X_v = l \text{ のとき} \\ \mu_{\max} & \text{それ以外} \end{cases}$$

ここで $\delta = \mu - \mu_{\max}$ および $\delta_\alpha = \mu_\alpha - \mu_{\max}$ と定義すると、高階グラフカットのブール値変数エネルギー (4) で

$$\mathcal{E}_C(Y_C) = \mu_{\max} + \delta \prod_{v \in C} \bar{Y}_v + \delta_\alpha \prod_{v \in C_{\bar{\alpha}}} Y_v,$$

と定義すればよいことがわかる。従って、 $\delta \leq 0$ かつ $\delta_\alpha \leq 0$ よりこのエネルギーは (12) によって

$$\delta \prod_{v \in C} \bar{Y}_v = \min_{u_C \in \mathbb{B}} \delta u_C \left(1 - \sum_{v \in C} Y_v \right) \quad (16)$$

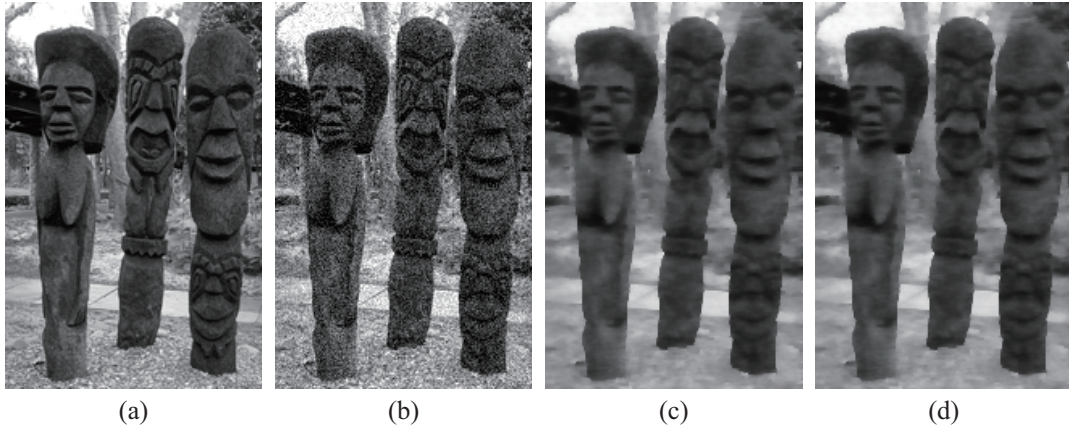


図1 エネルギー最小化の例として使った画像復元. (a) 元の画像 (b) ノイズを加えた画像 (c) 高階グラフカット [7] による結果 (d) 本稿提案の変数フリップ高階グラフカットによる結果.

$$\delta_\alpha \prod_{v \in C} Y_v = \min_{w_C \in \mathbb{B}} \delta_\alpha w_C \left(\sum_{v \in C_\alpha} Y_v + 1 - |C_\alpha| \right) \quad (17)$$

と2次に変換される.

再び, この変換によって最大でクリーク数の2倍の変数が加えられるが, 実際には現在の配置がクリークの全ての画素に同じラベルを与える場合以外には (16) の部分は0になる. また, (17) は劣モジュラであり, (16) は u_C をフリップすることで劣モジュラになるため, QPBO を使えば大域的に最小化できる.

以上により, Kohli ら [8] による \mathcal{P}^n Potts モデル (15) の α - β スワップ及び α 拡張による最適化は, 本稿で拡張された高階グラフカットの特殊な場合であることが示された.

5. 実験

本稿の主な貢献は前節までの理論的なものである. §3.4 で述べたように, 本稿で提案した階数変換の拡張によれば, エネルギー中に現れる変数の現れる数について指数的な数だけの異なる変換を得ることができる. 従って最適変換を見つけるアルゴリズム, あるいは少なくともヒューリスティックが望まれるところだが, その解決は今後の研究が待たれる. 本稿では代わりに, 特定の方針で変換を選び, その性能を実験的に [7] の結果と比較評価する.

5.1 変換方針

この特定の変換方針とは, 各単項式について, 必要ならば1つの変数をフリップして係数を負に変え, 変換公式 (12) のみを使うものである. つまり,

1. 与えられたエネルギーを多項式として表す.
2. 3次以上の単項式全てについて, 次数の高いものから順に以下の変換をする. 単項式を $ax_1 \cdots x_d$ として
 - a) $a < 0$ ならば (12) によって2次に変換する.

- b) $a > 0$ ならば $ax_1 \cdots x_d$ を $a(1 - \bar{x}_1)x_2 \cdots x_d$ に置き換え, $-a\bar{x}_1 \cdots x_d$ を (12) によって2次に変換する.

3. フリップされた変数 \bar{x} を $1 - x$ に戻す.

例えば $xyzt$ は次のように変換される.

$$\begin{aligned} xyzt &= (1 - \bar{x})yzt = yzt - \bar{x}yzt \\ &= (1 - \bar{y})zt - \bar{x}yzt = zt - \bar{y}zt - \bar{x}yzt \\ &= zt - \min_{u \in \mathbb{B}} u(\bar{y} + z + t - 2) - \min_{w \in \mathbb{B}} w(\bar{x} + y + z + t - 3) \\ &= zt - \min_{u, w \in \mathbb{B}} \{u(-y + z + t - 1) - w(-x + y + z + t - 2)\} \end{aligned}$$

5.2 エネルギー最小化問題

実験には公開されている [7] のコードとデータを使用して, 階数変換部分のみ置き換えて性能を比較した. エネルギー最小化問題は画像復元問題で, (1) で次のように定めたエネルギーを使用する:

$$\begin{aligned} f_{\{v\}}(X_{\{v\}}) &= \frac{(N_v - X_v)^2}{2\sigma^2}, \quad (\{v\} \in \mathcal{C}_1) \\ f_C(X_C) &= \sum_{i=1}^4 \alpha_i \log \left(1 + \frac{1}{2} (J_i \cdot X_C)^2 \right). \quad (C \in \mathcal{C}_4) \end{aligned}$$

ただし, N_v は与えられた画像 N の画素 v における値, クリークの集合は $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_4$ と表され, \mathcal{C}_1 は画像の画素全体, \mathcal{C}_4 は画像中の全ての 2×2 パッチからなる. また J_i は 2×2 行列, α_i は実数で両者はモデルのパラメータである. ここでは [7] と同じものを用いた. 「内積」 $J_i \cdot X_C$ は対応する位置の値を掛け合わせて加えたものである.

最小化アルゴリズムは §2.2 で説明した高階グラフカットアルゴリズムである. 階数変換部分のみ置き換えて性能を比較した. 提案配置 P は [7] で使われているもののうち, 次の2つを交互に提案するものを用いた.

- a) 毎回画素毎に生成される一様乱数画像.
- b) 繰り返し30回毎に X をガウス核 ($\sigma = 0.5625$) でぼかした画像.

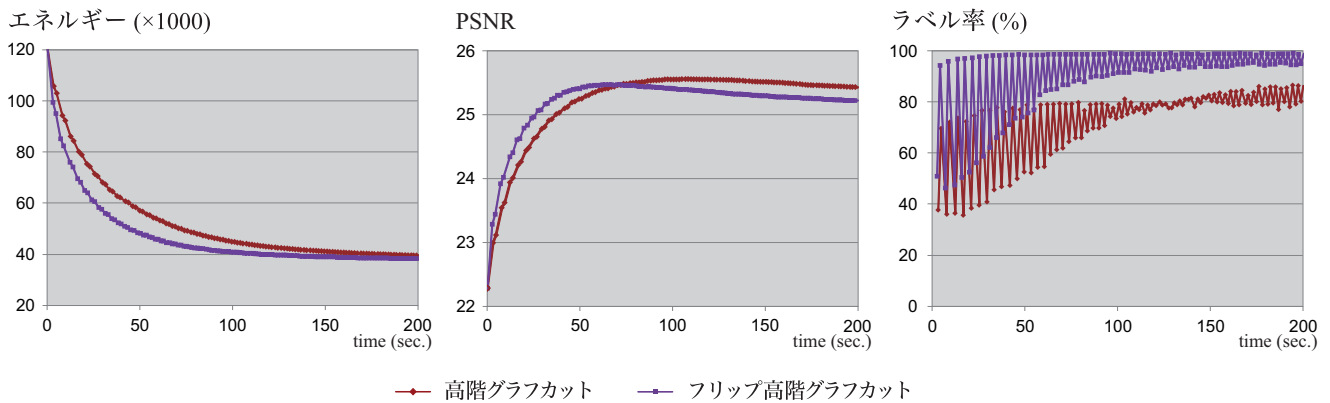


図2 エネルギー最小化の過程におけるエネルギー減少, ノイズ除去の効果(PSNR), および QPBO でラベル付けされた画素の比率についての, 高階グラフカット [7] と本稿提案の変数フリップ高階グラフカットとの比較.

5.3 実験結果

図1に高階グラフカット [7] と本稿提案の変数フリップ高階グラフカットによる画像復元の例を示す. 視覚的には両者の結果はほぼ見分けがつかない. 図2はエネルギー最小化の過程におけるエネルギー減少, ノイズ除去の効果(PSNR), および QPBO でラベル付けされた画素の比率を表す. 上述のように, 実験は [7] と階数変換部分のみ置き換えて比較した. 4次の式の場合, [7] の変換では (8) のようになるから, §5.1 で説明した今回の方法の方が階数変換で加えられる変数の数は多くなる. それにもかかわらず, エネルギーの減り方は今回の方法の方が若干速いことがわかる. これは, 付加変数が多いため一回のQPBO アルゴリズムにかかる時間は長くても, 右端のグラフからわかるように毎回ラベル付けされる画素の比率が高いためということがわかる. エネルギーがより低いのに PSNR は必ずしも高くないが, これは最小化アルゴリズムの問題ではなく, 本来, より低いエネルギーがよりよい結果を表すべきである事前分布モデルに問題があることを示している.

6. むすび

本稿では, 任意階ブール値変数エネルギーの1階への変換アルゴリズムを, 変数の0と1を入れ替えるという単純な操作(フリップ)を変換の前後に行うことによって拡張した. これにより, エネルギー中に現れる変数の数について指数的な数の異なる変換を得るが, 本稿では特別な場合を実験により検証し, 既存の変換より性能が改善されることを示した. また, ここ数年に発表された高階エネルギー最小化法のいくつかは, この一般化されたアルゴリズムの特殊な場合と考えられることも示した.

謝辞 本研究は栢森情報科学振興財団研究助成及び科学研究費補助金萌芽研究 19650065 の補助を受けた.

文 献

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

[2] E. Boros, P. L. Hammer, R. Sun, and G. Tavares, "A max-flow approach to improved lower bounds for quadratic unconstrained binary optimization (QUBO)," *Discrete Optimization* **5**(2):501–529, 2008.

[3] Y. Boykov, O. Veksler, R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts." *IEEE T. PAMI* **23**(11):1222–1239, 2001.

[4] P. Felzenszwalb and D. Huttenlocher, "Efficient Belief Propagation for Early Vision." *Int. J. Comp. Vis.* **70**(1):41–54, 2006.

[5] P. L. Hammer and S. Rudeanu, *Boolean Methods in Operations Research and Related Areas*, Springer-Verlag, Berlin, Heidelberg, New York, 1968.

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

[7] 石川 博. "高階グラフカット." *MIRU2009*, pp. 7–14.

[8] P. Kohli, M. P. Kumar, and P. H. S. Torr, "P³ & Beyond: Move Making Algorithms for Solving Higher Order Functions," *IEEE T. PAMI* **31**(9):1645–1656, 2009.

[9] P. Kohli, L. Ladicky, and P. H. S. Torr, "Robust Higher Order Potentials for Enforcing Label Consistency," *Int. J. Comp. Vis.* **82**(3):302–324, 2009.

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

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

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

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

[14] 永橋知行, 藤吉弘亘, 金出武雄. "平滑化処理の繰り返しによる画像セグメンテーションのためのグラフカット." *MIRU2007*, pp. 241–248.

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

[16] ハリム サンディ, 和田俊和. "グラフカットを用いた画像の粒状ノイズ抑制." *MIRU2008*, pp. 109–116.

[17] 高谷優樹, ザカリヤレット, 斎藤英雄. "スポーツフィールドの平面性を考慮したグラフカット法による選手領域のセグメンテーション." *MIRU2009*, pp. 953–959.

[18] O. J. Woodford, P. H. S. Torr, I. D. Reid, and A. W. Fitzgibbon, "Global stereo reconstruction under second order smoothness priors." *CVPR2008*.