

高階グラフカット

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 画像分野の多くの問題がマルコフ確率場などのエネルギー最小化問題として表されるが、実際に使われるエネルギーは、少数の 2 階の例外を除いて、最大 2 点に依存する項の和からなる 1 階のものに限られてきた。自然画像の統計的性質をよりよく表現するためには、より多くの点に依存する項を持つ、高階のエネルギーの使用が望まれるが、今までそれを効率的に最小化する方法が知られていなかった。本稿では、任意階のエネルギーを同値な 1 階のエネルギーに変換することにより、QPBO などのグラフカット手法を使って高階エネルギーを最小化する方法を述べる。信念伝播法 (BP) による高階エネルギー最小化法の検証に使われた 3 階のエネルギーによる画像復元実験の結果、それらの BP 手法に比べて、最適化の度合で有意に上回り、また速度は数十倍～数百倍高速であることも報告する。

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

Higher-Order Graph Cut

Hiroshi Ishikawa

Department of Information and Biological Sciences, Nagoya City University

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

E-mail: hfs@ieee.org

Abstract Many vision problems today are formulated as energy minimization problems such as Markov Random Fields. Because of the lack of efficient algorithm to optimize energies with higher-order interactions, most such energies are of first order, with a few second-order exceptions. In this paper, we introduce a new reduction technique that reduces any higher-order binary optimization problem into an equivalent pairwise problem. We demonstrate this technique by applying it to optimize a third-order potential, which allows clique potentials that depend on four pixels.

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

1. はじめに

領域分割 [20], ステレオ [27], 画像復元 [26] などの問題を、マルコフ確率場 (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は、グラフカット [5][9][14], 信念伝播法 (BP) [6][19], またツリー重み再配分メッセージ伝達法 (TRW) [12] などの高効率なエネルギー最小化アルゴリズムの出現によって最近とみに一般化した。これらは、それ以前のアニーリング法や ICM (Iterated Conditional Mode) などに比べてより低いエネルギーを高速に得ることができるが、適用可能なエネルギーはより限られている。

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^(注1)が、グラフカッ

(注1): この定義は紛らわしく、依存する点の数とずれるだけでなく、後述の多項式による表現でも $k-1$ 階のエネルギーは k 次の多項式で表される。つまり、「階」と「次」で一つずれることに留意されたい。

表1 グラフカットによる最適化

	2 値	多値
1 階	最小切断 [14], QPBO [8]	α 拡張 [5], 凸関数 [9]
2 階	1 階に還元 [7][14]	融合移動 [7][14][27]
3 階以上	1 階に還元 (本稿)	融合移動 (本稿)

トで最小化可能なエネルギーの階数は最大 2 階に限られ、また実際にはほとんど 1 階のものしか使われてこなかった。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がほとんどであった。

これは高階のエネルギーが無用であるからではなく、効率的にそれを最小化するアルゴリズムの欠如が最大の原因である。実際、ビジョンや画像処理では、自然画像の特徴を捉えるために、より高次の統計の重要性が繰り返し指摘されてきた [10][21][24] ほか、BP [16][22] や、(特殊な場合に) グラフカット [15] によってより高階のエネルギーを最適化する試みも活発化している。

グラフカットのなかでも α 拡張 [5] が最もよく使われるが、これは 2 値の最適化問題を繰り返し解くことにより、近似的に多値の最適化問題を解くものである。2 値の最適化問題は劣モジュラ性という条件を満たすとき、 s - t 最小切断アルゴリズムを用いることで大域的に解を求めることができる。上述のように 2 階まで最適化可能なのは、2 階の 2 値問題を同値な 1 階の問題に変換する方法が知られているからである [14]。その場合でも、変換結果が劣モジュラでない場合、グラフカットは使用不能だったが、最近、劣モジュラでないエネルギーにも解の一部を与える QPBO アルゴリズム [2] [4] [8] が知られるようになり、融合移動という α 拡張の一般化と合わせて 2 階のエネルギーを最小化した例 [27] も出てきている。

本稿では、以上のように最小化可能なエネルギーが最高 2 階までであったグラフカットを、任意階に適用可能な方法に一般化する。その核心は、任意階の 2 値エネルギーを同値な 1 階のエネルギーに変換する方法である。これを 2 階の場合と同様に QPBO 及び融合移動と組み合わせることにより、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する。表 1 に、現在の主要なグラフカットの適用可能範囲と本稿の位置づけをまとめた。

高階エネルギーを最小化する BP による方法が 2 つ知られている。これらは共通の画像復元問題で性能を評価しているため、本稿の手法の評価も同じ問題を使って行った。4 点に依存する項の和で表される 3 階エネルギーの最小化実験の結果、本手法は到達エネルギーで有意に既存手法を上回り、計算速度では遥かに上回った。

2. 背景

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、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：

$$E(X) = \sum_{v \in V} f_v(X_v) + \sum_{\{u,v\} \in \mathcal{E}} f_{uv}(X_u, X_v). \quad (1)$$

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。クリークの集合を \mathcal{C} 、ラベルの集合を L とする。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ぶ。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。エネルギーは次のような和に分解できると仮定する：

$$E(X) = \sum_{C \in \mathcal{C}} f_C(X_C), \quad (2)$$

ただし $X_C = (X_v)_{v \in C} \in L^{|C|}$ はクリーク C 内のサイトに X により与えられたラベルを成分とするベクトル、 $f_C(X_C)$ は X_C のみ^(注2)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。

2.1 1 階エネルギー

これまで 1 階のエネルギー、つまり (2) の \mathcal{C} が最高 2 個のサイトの組しか含まない、(1) の形である場合について、グラフカットアルゴリズムが発達してきた [11]。

最小切断 これがグラフカットの最も基本で、ラベルが 2 値 ($|L|=2$) の場合に、エネルギー (1) が劣モジュラ条件

$$f_{uv}(0,0) + f_{uv}(1,1) \leq f_{uv}(0,1) + f_{uv}(1,0)$$

を満たせば、大域最小化することができる。ここで、また以下常に、2 値の場合には $L = \{0, 1\}$ とする。

凸関数 ラベルが 3 つ以上の数値の場合、(1) の $f_{uv}(X_u, X_v)$ が $X_u - X_v$ の凸関数であるときには大域最小化できることが知られている [9]。

α 拡張 α 拡張アルゴリズムは任意の配置 X から始めて、繰り返し移動をする。一回の移動、すなわち配置の変更は、一つのラベル α を固定して、各サイトについて現在のラベルと α のどちらを与えるかを選択することで行う。この選択は 2 値の問題を定義し、それが劣モジュラならば、最小切断によって大域的に最もエネルギーを減少させる移動を選ぶことができる。全てのラベル α について繰り返すことにより、 $E(X)$ を近似的に最小化する。

融合移動 融合移動 [17] は、この α 拡張の自然な一般化である。繰り返し毎に、任意の配置を 2 つ固定しておき、各サイトについてどちらの配置の与えるラベルをとるかという 2 値問題を解くことで新しい配置を見つける。

QPBO 融合移動は自然なアイデアであるが、最初からこれが使用されなかったのは、 α 拡張移動が一定の条件の下で劣モジュラであるのに対して、一般の融合移動は劣モジュラでないからである。ところが QPBO アルゴリズム [2] [4] [8] [13] [25] の出現により融合移動を実用的に考えることができるようになった。QPBO は 2 値の問題において必ずしも全てでない一部のサイトにラベルを与

(注2)：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。

えるが、任意の配置 Y をとり、サイト v に QPBO がラベルを与えればそれを Z_v とし、与えなければ $Z_v = Y_v$ として得られる配置 Z が $E(Z) \leq E(Y)$ を満たすことを保障する [13]. Y に大域最小解をとれば Z も大域最小解であり、 X の与えるラベルは常に大域最小解の一部になっていることが解る。またエネルギーが劣モジュラなら全てのサイトにラベルを与えて大域最小化することも保障される。

融合移動と QPBO の組み合わせ α 拡張の場合と同様に、任意の配置からはじめて、繰り返し移動をすることによって多値の最適化問題を近似的に解くことを考える。ただし、 α 拡張のように固定されたラベルへの変更のみを許すのではなく、任意の新しい提案配置を繰り返し毎に生成し、現在の配置と、この提案配置の融合移動を繰り返す。上記の QPBO の性質により、移動を選ぶ 2 値問題が劣モジュラでなくても、QPBO によってラベルが与えられないサイトは現在のラベルにとどめることによって、エネルギーが増加しないことが保障される。そのため、QPBO が十分多くのサイトに 0 か 1 を与えている限り、繰り返しによって最小化を図ることができる。

2.2 2階エネルギー

高階のエネルギーについては、2 階の 2 値問題を同値な 1 階の 2 値問題に変換する方法 [7][14] が知られている。多値の場合も、移動アルゴリズムを使うことで 2 階であれば近似的に解くことができる。変換後のエネルギーは劣モジュラでない場合が多いので、QPBO との組み合わせが重要である。例えば、Woodward ら [27] は、融合移動と QPBO の組み合わせによって曲率を考慮したエネルギーを最小化するステレオを試みている。

2.3 高階エネルギー

本稿では、これまで 2 階までしか可能でなかった 1 階エネルギーへの変換を、任意の高階 2 値エネルギーについて可能にする方法を述べ(4 節)、またこれを融合移動及び QPBO と組み合わせ、多値の場合に高階のエネルギーを最小化する方法を与える(5 節)。

次節ではまず、既存の変換法を説明する。

3. 既存の変換法

一般に、2 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。ここで、 $\mathbb{B} = \{0, 1\}$ 、 \mathbb{R} は実数の集合を表す。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 [3]:

$$f(x_1, \dots, x_n) = \sum_{S \subseteq U} c_S \prod_{i \in S} x_i. \quad (3)$$

つまり、2 値エネルギーは常に多項式に変換できる。

3.1 最小値選択への変換

まず、上述した、2 階の 2 値問題を同値な 1 階の問題に

変換する方法を紹介する。Kolmogorov と Zabih [14] がグラフカットの構成のひとつとして発見し、後に Freedman と Drineas [7] によって以下のように擬ブール関数の言葉で整理されたこの変換は、次の恒等式に拠っている。

$$xyz = \max_{w \in \mathbb{B}} w(x+y+z-2). \quad (4)$$

各変数は 0 か 1 の値しかとらないから、この式を確かめるには x, y, z の値の組み合わせ 8 通りを確かめればよい。いま、 $a < 0$ とすると、

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

となり、最小化問題に 3 次の項 $axyz$ が現れた場合、係数 a が負であれば、 $aw(x+y+z-2)$ でこれに代えても最小値とそれをとる x, y, z は変わらないことがわかる。つまり、その関数で最小値を与える x, y, z は、元の関数でも最小値を与える。一方 $a > 0$ の場合、(4) の x を $1-x$ に、 y を $1-y$ に、 z を $1-z$ に置き換えて整理すると、

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

$$xyz = \min_{w \in \mathbb{B}} w(x+y+z-1) + (xy+yz+zx) - (x+y+z) + 1 \quad (5)$$

のように \max が \min に替わるため、 $axyz$ は

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

で代えればよい。このように、 a が正負どちらでも 3 次の項を 2 次以下に変換できる。この変換は劣モジュラ性とは関係がなく、いつでも正しいことに注意する。

ただし、この変換は 3 次の場合にしか使えない。より正確には a が正の場合には 3 次までしか使えない。例えば 4 次の場合、 $a < 0$ ならば上と同様に

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

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

となるが、 $a > 0$ だと

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

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

となり、3 次の場合のように \max を \min に替えることができない。このように、偶数次で係数が正の場合にこの一般化は使えない。また 5 次以上の奇数次の場合も、展開後の式に 4 次などの偶数次で係数が正の項が現れるので、やはりそのままでは使えない。

3.2 置換による変換

実は、任意次数の擬ブール関数を 2 次のそれに変換する全く違う方法が 30 年以上前から知られている。これ

は Rosenberg [23] によって発見され、最近にも紹介されている [1][3]. この方法では、2変数の積 xy を新しい変数 z で置き換え、関数が最小値をとるときには新変数が元の積と同じ値をとらざるを得ないように、それらが異なるときに非常に大きな値をとる関数を全体に加える. 具体的には、 $x, y, z \in \mathbb{B}$ として

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

と定義すると、全ての場合を試すことにより容易に確認できるように、 $D(x, y, z) = 0$ なら $xy = z$ であり、 $xy \neq z$ なら $D(x, y, z) > 0$ である. 例として3次擬ブール関数

$$f(x, y, w) = xyw + xy + y$$

で例えば xy を z で置き換え $MD(x, y, z)$ を加えると、

$$\tilde{f}(x, y, w, z) = zw + z + y + MD(x, y, z) \quad (6)$$

と2次の関数になる. ここで M は正の定数で、 $xy \neq z$ すなわち $D(x, y, z) > 0$ のときには \tilde{f} が最小値をとることができないように十分大きくとる.

この操作を繰り返すことにより、任意の高次擬ブール関数を2次擬ブール関数にすることができる. こうして作った関数が最小値をとる変数の値を、元からある変数に与えれば、元の関数も最小値をとる.

この変換の問題点は、結果としてできた関数を最適化することが実際には困難な点にある. その原因は、(6)の

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

にある. 第一項 Mxy は非常に大きな正の係数を持つ2次の項だが、多項式で表した2次擬ブール関数は、2次の項の係数が全て負であるとき、またそのときに限り劣モジュラであることが知られている [7]. そのため、この項 Mxy のために変換後の関数は常に劣モジュラでなくなってしまう. また、同じ正の係数でも M は非常に大きいためか、この方法によって変換されたエネルギーは QPBO を使ってもほとんど全てのサイトにはラベルが与えられず、十分最小化できないことが実験的に確かめられる.

4. 高階への拡張

以上のように、今まで3階以上のエネルギーで使えるグラフカットは知られていなかった. 本節では、任意階の2値エネルギーを同値な1階の2値エネルギーに変換する新手法を述べる. これは、3.1節で説明した最小値選択の一般化であり、実験で示すように融合移動及びQPBOとの組み合わせで実用可能である.

4.1 3階の場合

まず3階、つまり多項式で表して4階の場合を考える. 3階の場合に、式(5)の両辺が x, y, z の対称式、つまり変数の置換に対して不変な形になっていることに注目す

る. 同様に4階の場合も、元の式が対称式 $xyzt$ なので、変換後も x, y, z, t の対称式にならざるを得ないと思われる. 従って、もし4階の場合に2次に変換する、(5)に類似の公式があるなら

$$xyzt = \min_{w \in \mathbb{B}} w(1 \text{ 次対称式}) + (2 \text{ 次対称式})$$

の形になるはずである.

一方、任意の対称式は基本対称式の多項式として表されることが知られている. x, y, z, t の2次以下の基本対称式は次の2つである.

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

また、変数が \mathbb{B} の値しかとらない場合には $x^2 = x, y^2 = y$ などが成り立つから、 $s_1^2 = s_1 + 2s_2$ となり、2次までの任意の対称式は基本対称式 s_1, s_2 と定数の一次結合で書ける. 従って、望まれる公式があるなら次のような形になるはずである.

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

そこで、右辺が $x = y = z = t = 1$ のときだけ正数で、それ以外は0になるような整数 a, b, c, d, e を乱暴りに探したところ、次式が発見された.

$$xyzt = \min_{w \in \mathbb{B}} w(-2s_1 + 3) + s_2 \quad (7)$$

5次についても調べたが、同様の等式は見いだされなかった. そこで外部変数を2つに増やしたところ、次式が見つかった.

$$xyztu = \min_{(v, w) \in \mathbb{B}^2} \{v(-2r_1 + 3) + w(-r_1 + 3)\} + r_2. \quad (8)$$

ここで r_1, r_2 は x, y, z, t, u の1, 2次の基本対称式である.

同様に、外部変数を増やしつつ探して6, 7次と公式を得た後、それらから一般式を推測したのが、次に述べる一般階の公式である.

4.2 一般階の公式

$a < 0$ のとき Freedman と Drineas [7] が指摘したとおり、 $a < 0$ であれば簡単で、次式が成り立つ.

$$ax_1 \dots x_d = \min_{w \in \mathbb{B}} aw\{x_1 + \dots + x_d - (d - 1)\}. \quad (9)$$

$a > 0$ のとき 本稿で報告する新公式では、 $a > 0$ で次数が4より大きい場合には、2つ以上の外部変数を必要とする. d 次の項を変換するためには次の数だけの変数を要する(ここで $\lfloor x \rfloor$ は x を超えない最大の整数を表す):

$$n_d = \left\lfloor \frac{d-1}{2} \right\rfloor.$$

これらの変数をまとめて $w = (w_1, \dots, w_{n_d}) \in \mathbb{B}^{n_d}$ と書き、

また 1 次及び 2 次の基本対称式を次のように書く：

$$S_1 = \sum_{i=1}^d x_i, \quad S_2 = \sum_{i=1}^{d-1} \sum_{j=i+1}^d x_i x_j. \quad (10)$$

すると公式は以下のようになる。

$a > 0$, 偶数次のとき

$$ax_1 \dots x_d = \min_{w \in \mathbb{B}^{n_d}} a \sum_{i=1}^{n_d} w_i (-2S_1 + 4i - 1) + aS_2. \quad (11)$$

$a > 0$, 奇数次のとき

$$ax_1 \dots x_d = \min_{w \in \mathbb{B}^{n_d}} a \left\{ \sum_{i=1}^{n_d-1} w_i (-2S_1 + 4i - 1) + w_{n_d} (-S_1 + 2n_d - 1) \right\} + aS_2. \quad (12)$$

これらの公式の証明は付録で与える。なお、3 次の場合

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

は (5) と異なるがより簡単である。また (7), (8) はそれぞれ一般の公式 (11), (12) の特殊な場合である。

上述のように任意の擬ブール関数は多項式 (3) の形に一意に表される。公式 (9), (11), (12) によって各項を同値な 2 次式に置き換えることができるから、任意の擬ブール関数を同値な 2 次擬ブール関数に変換することができる。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。

一般の d 次関数を 2 次に変換するのに必要な新たな変数の数は、最悪の場合 d について指数的である。これは、最高次だけでなくそれ以下の項もそれぞれ変換を要するからである。このためあまり高階の関数の変換は実用的でないが、5~6 階ならば十分実用可能である。今までは 2 階が最高であり、2 階でも 1 階と大きく違う [27] ことから、本変換には十分大きな意義があると考えられる。

5. 高階融合移動

本節では、前節の新公式を使って高階エネルギーを最小化するための融合移動アルゴリズムについて述べる。アルゴリズムはクリークの集合 \mathcal{C} 、式 (2) 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、 $E(X)$ を最小化する。

配置 X を変数として、提案配置 P を繰り返し生成し、 X と融合して $E(X)$ を最小化する X を探す。提案配置は任意だが、その選択にアルゴリズムの性能が依存する。

融合は、サイト v 毎の現在値 X_v か提案値 P_v の選択 $Y_v (= 0 \text{ or } 1)$ を与える 2 値の配置 Y を考え、選択の結果得られるエネルギーを Y の関数と考えた 2 値エネルギー

の最小化問題を解くことで決定する。 $Y_v = 0$ ならば X_v は変わらず、 $Y_v = 1$ ならば P_v になることとする。

従って 2 値のエネルギーは Y に対して対応した X の変化後のエネルギーを与える。クリーク C に属するサイトに対応する Y の変数のベクトルを $Y_C = (Y_v)_{v \in C}$ とする。 $Y_C = \beta \in \mathbb{B}^{|C|}$ である場合に選択の結果 X が C の各サイトに与えるラベルを成分とするベクトルを $x_C^{X,P}(\beta) \in L^{|C|}$ とすると、その成分は次式で与えられる：

$$(x_C^{X,P}(\beta))_v = (1 - \beta_v)X_v + \beta_v P_v \quad (v \in C).$$

また Y_C のとり得るベクトル値 $\beta \in \mathbb{B}^{|C|}$ について、

$$\theta_C^\beta(Y_C) = \prod_{v \in C} \{\beta_v Y_v + (1 - \beta_v)(1 - Y_v)\},$$

は Y_C の変数についての $|C|$ 次の多項式であり、 $Y_C = \beta$ 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：

$$\mathcal{E}(Y) = \sum_{C \in \mathcal{C}} \sum_{\beta \in \mathbb{B}^{|C|}} f_C(x_C^{X,P}(\beta)) \theta_C^\beta(Y_C),$$

この式の内側の和は、 Y_C の取りうる各ベクトル値 β につき一頂ずつの和だが、その項は $Y_C = \beta$ のときのみ 1 になり、それ以外ときは 0 である。またその係数は、 $Y_C = \beta$ に従った X の選択後の局所エネルギー $f_C(X_C)$ の値である。従って $\mathcal{E}(Y)$ は Y に従って変化させた X についてのエネルギー $E(X)$ を与える。

この多項式 $\mathcal{E}(Y)$ を前節の公式により 2 次に変換し QPBO アルゴリズムに与えると、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 [13] 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。

6. 実験

高階エネルギーを最小化する BP アルゴリズムが二つ知られている。Lan ら [16] と Potetz [22] によるものだが、どちらも特定の高階画像復元問題を使って検証された。本節では、前節で述べた高階融合移動アルゴリズムを同じ問題に適用し、それらのアルゴリズムと性能比較する。

6.1 Fields of Experts による画像復元

この画像復元問題は、Fields of Experts (FoE) [24] と呼ばれる最近の統計モデルを使う。FoE は 2 画素の間だけでない高階にわたる自然画像の複雑な統計的性質を、自然画像から統計モデルを学習することによって捉え、ノイズ除去や画像インペインティングにおいて非常に有効であることが示されている。FoE モデルは画像の事前確率をいくつかの Student の t 分布の積として表現する：

$$p(X) \propto \prod_C \prod_{i=1}^K \left(1 + \frac{1}{2} (J_i \cdot X_C)^2 \right)^{-\alpha_i}. \quad (13)$$

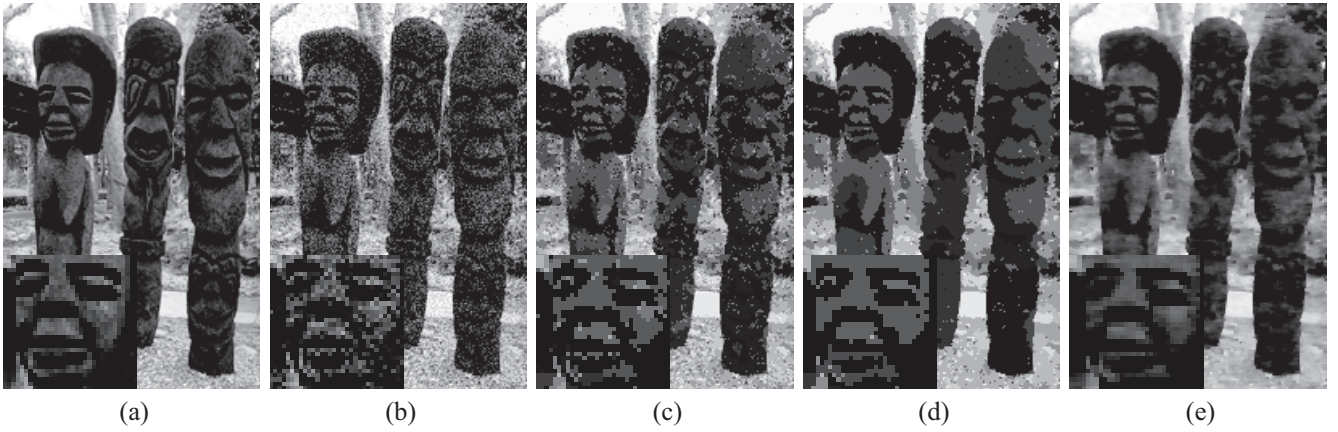


図1 1階エネルギーとの画像復元比較. 各画像の左下は左上の顔部分の拡大. (a) 元画像. 160×240 画素. (b) ノイズ入り画像 ($\sigma = 20$). (c) 1階 Potts エネルギー (15) による復元画像 ($\lambda = 8000$). (d) 同 ($\lambda = 16000$). (e) 3階 FoE エネルギーによる復元画像.

ここで C は画像中の全ての $n \times n$ 画素パッチにわたり, J_i は $n \times n$ 画素のフィルターである. モデルのパラメータは J_i と α_i で, 自然画像のデータベースから学習される. 比較対象の [16] と [22] はどちらも 2×2 画素のパッチを使い, 既存の 2 画素間統計を使った 1 階のモデルと比べてノイズ除去の性能を向上することを示した.

この画像復元問題では, ノイズ入り画像 N を与えられて, 事前確率 (13) を使った最大事後確率推定を行う. この事前確率は $2 \times 2 = 4$ 画素のクリークまで依存する 3 階の MRF を定義する. ここでも比較のために [16] [22] と全く同じモデルを使用する. ノイズ入り画像 N は, 元画像 X に既知の標準偏差 σ を持つ i.i.d. Gaussian ノイズが加えられてきたものと仮定する. 従って, X を与えたときの N の尤度が, 画素の集合を V として,

$$p(N|X) \propto \prod_{v \in V} \exp\left(-\frac{(N_v - X_v)^2}{2\sigma^2}\right) \quad (14)$$

を満たすことを仮定する.

6.2 最小化アルゴリズム

最小化には前節の融合移動アルゴリズムを使う. クリークの集合は, 全ての単一ピクセルからなる \mathcal{C}_1 と全ての 2×2 パッチからなる \mathcal{C}_4 を使い $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_4$ とする. X は与えられたノイズ入り画像 N で初期化する. 事前確率 (13) と尤度 (14) から Bayes の定理によって事後確率 $p(X|N)$ を得, それを最大化する局所エネルギーとして

$$f_{\{v\}}(X_{\{v\}}) = \frac{(N_v - X_v)^2}{2\sigma^2}, \quad (\{v\} \in \mathcal{C}_1)$$

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

を得る. 本実験では, 提案配置 P (この場合は提案画像) として次の 3 つをそれぞれ別の実験で使い比較した.

(i) **α -expansion.** α 拡張類似として, $0 \sim 255$ の定数画

像をランダムな順番で, ただし全て出るまでは同じ値は繰り返さないで提案.

(ii) **blur & random.** 次の 2 つを交互に提案:

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

(iii) **gradient.** X をエネルギーの事前分布部分の勾配 (X による微分) 方向に動かした画像を提案.

終了条件は, 過去 20 回の繰り返しにおけるエネルギー低下量が閾値 θ_c 以下になることとする.

6.3 結果

まず図 1 に 1 階エネルギーと 3 階エネルギーとの定性的な違いを示す. 1 階のエネルギーには

$$E(X) = \sum_{v \in V} (N_v - X_v)^2 + \lambda \sum_{(u,v) \in \mathcal{E}} \text{Potts}(X_u, X_v) \quad (15)$$

を使った. ここで \mathcal{E} は上下左右に隣接した画素の組の集合, $\text{Potts}(x,y)$ は $x = y$ なら 0, $x \neq y$ なら 1 を与える関数である. 最適化には α 拡張を使った. 復元画像から, 高階エネルギーは自然画像の特徴を捉え, より自然な画像を得ることができるのがわかる.

本実験の目的は画像復元そのものではなくエネルギー最適化法の評価である. 定量的比較評価のために, [16] [22] で使われたのと同じ 10 個の画像を使用した. これらは, Berkeley segmentation database [18] の画像をグレイスケール化し縮小したものに, i.i.d. Gaussian ノイズ ($\sigma = 10, 20$) を加えたものである. 画像及び FoE パラメータ J_i, α_i は, [16] [24] の著者の一人 Stefan Roth から入手した. 閾値は $\theta_c = 8$ ($\sigma = 10$), $\theta_c = 100$ ($\sigma = 20$) とした. 画像復元結果のエネルギー及び PSNR を表 2 に示す.

この表から, 本稿で与えた高階融合移動によるグラフカットアルゴリズムは, エネルギー最小化の性能で BP



図2 画像復元結果. (左) 元画像. 240 × 160 画素. (中) ノイズ入り画像 ($\sigma = 20$). (右) 復元画像.

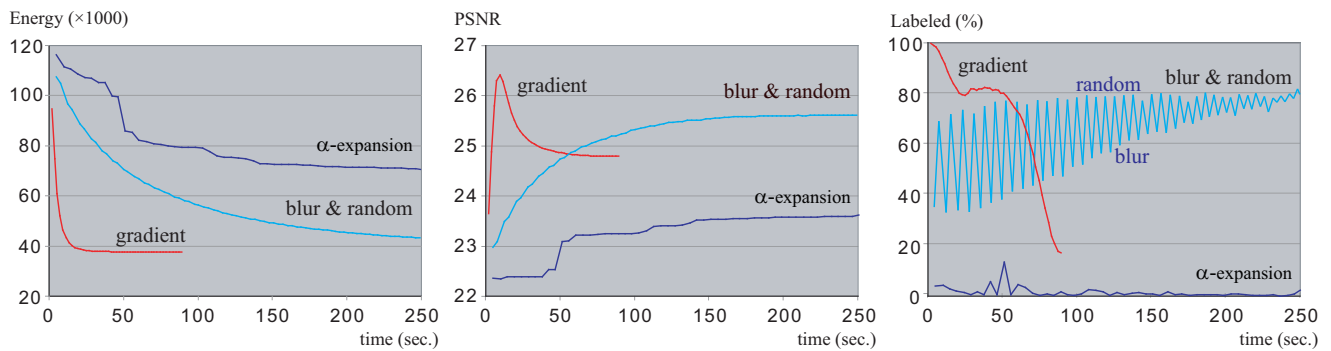


図3 融合移動の提案配置による挙動の違い. 図1のテスト画像の場合. (左) エネルギー (中) PSNR (右) QPBOによりラベル付けられた画素の比率.

表2 同一FoEモデルを使った画像復元実験による最適化法の比較. 10画像の平均PSNR/平均到達エネルギー. PSNRは大きい方が, 到達エネルギーは小さいほうがよい.

ノイズ	Lanら [16]	Potetz [22]	本稿
$\sigma = 10$	30.36 / 40236	31.54 / 36765	31.44 / 35896
$\sigma = 20$	27.05 / 33053	27.25 / 31801	27.43 / 30858

アルゴリズムによる結果 [16] [22] を上回ることがわかる. 図2にさらに画像復元の例を示す.

図3は提案配置による挙動の違いを示す. まず, エネルギーについては, blur & random と gradient は同レベルのエネルギーに到達したが, gradientの方が10倍以上速かった. また, α 拡張ではエネルギーは他の二つの場合と同レベルに達しなかった. これはQPBOによってラベルをつけられるサイトの比率が非常に低いからである. 高階エネルギーでは α 拡張は使えないことがわかる. PSNRについても, α 拡張は低いレベルにとどまった. PSNRについてはblur & randomが安定した挙動を見せたのに対して, gradientでは最初に高い値をとった後に逆に低下する現象が見られた. この間エネルギーは低下している. これは最適化よりFoEモデルの問題だ

が, このような挙動は元画像が不明な実用には考慮が必要となると思われる. 参考までに, 3.2で紹介した置換による変換を試みたところ, ラベル付けられる比率がほとんど0% ($\sigma = 20$ で100回の平均が0.00018%)にとどまり, エネルギーは減少せず, PSNRも改善しなかった.

本稿のアルゴリズムは2.33GHz Xeon E5345 プロセッサで, 提案配置がgradientの場合1分以内, blur & randomでも10分程度で収束した. 一方, [22] は2.2GHz Opteron 275で30から60分, [16] は3GHz Xeonで8時間かかる. これからの論文にある. これから, 速度については本手法は既存手法を遥かに上回るといえる.

7. むすび

本稿では, 任意階2値エネルギーの同値な1階エネルギーへの変換による, 高階グラフカットアルゴリズムを与えた. 検証のために, 既存のBPによる最適化法と同一データ及びモデルを使った画像復元実験により比較し, 最適化の度合で有意に既存手法を上回り, 速度では大幅に改善することを示した. なお, 本稿で紹介したアルゴリズムの一実装を著者のウェブページで公開している.

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。また Stefan Roth 博士に FoE モデルとテスト画像及び実験結果の、Brian Potetz 博士に実験結果の提供を受けた。また実験には Vladimir Kolmogorov 博士の公開している QPBO の C++ による実装を使用した。

文 献

- [1] A. M. Ali, A. A. Farag, and G. L. Gimel'farb. "Optimizing Binary MRFs with Higher Order Cliques." *ECCV2008*.
- [2] E. Boros, P. L. Hammer, and X. Sun. "Network flows and minimization of quadratic pseudo-boolean functions." Technical Report RRR 17-1991, RUTCOR Res. Rep., 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, R. Zabih. "Fast Approximate Energy Minimization via Graph Cuts." *IEEE T. PAMI* **23**:1222-1239, 2001.
- [6] P. Felzenszwalb and D. Huttenlocher. "Efficient Belief Propagation for Early Vision." *Int. J. Comp. Vis.* **70**:41-54, 2006.
- [7] D. Freedman and P. Drineas. "Energy Minimization via Graph Cuts: Settling What is Possible." *CVPR2005*.
- [8] P. L. Hammer, P. Hansen, and B. Simeone. "Roof Duality, Complementation and Persistence in Quadratic 0-1 Optimization." *Math. Prog.*, **28**:121-155, 1984.
- [9] H. Ishikawa. "Exact Optimization for Markov Random Fields with Convex Priors." *IEEE T. PAMI* **25**(10):1333-1336, 2003.
- [10] 石川 博. "ヒト視覚系から示唆される高階ステレオ事前分布." *MIRU2006*, pp. 128-134.
- [11] 石川 博. "グラフカット." コンピュータビジョン最先端ガイド 1, 八木康史・斎藤英雄 (編), 第 2 章, アドコム・メディア, 2008.
- [12] V. Kolmogorov. "Convergent Tree-Reweighted Message Passing for Energy Minimization." *IEEE T. PAMI* **28**(10):1568-1583, 2006.
- [13] V. Kolmogorov and C. Rother. "Minimizing Non-submodular Functions with Graph Cuts - A Review." *IEEE T. PAMI* **29**(7):1274-1279, 2007.
- [14] V. Kolmogorov and R. Zabih. "What Energy Functions Can Be Minimized via Graph Cuts?" *IEEE T. PAMI* **26**(2):147-159, 2004.
- [15] P. Kohli, M. P. Kumar and P. H. S. Torr. "P3 & Beyond: Move Making Algorithms for Solving Higher Order Functions." *IEEE T. PAMI*, to appear, 2008.
- [16] X. Lan, S. Roth, D. P. Huttenlocher, and M. J. Black. "Efficient Belief Propagation with Learned Higher-Order Markov Random Fields." *ECCV2006*, II:269-282.
- [17] V. Lempitsky, C. Rother, and A. Blake. "LogCut - Efficient Graph Cut Optimization for Markov Random Fields." *ICCV2007*.
- [18] 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." *ICCV2001*, pp. 416-423.
- [19] T. Meltzer, C. Yanover, and Y. Weiss. "Globally Optimal Solutions for Energy Minimization in Stereo Vision Using Reweighted Belief Propagation." *ICCV2005*, pp. 428-435.
- [20] 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.
- [21] R. Paget and I. D. Longstaff. "Texture Synthesis via a Non-

causal Nonparametric Multiscale Markov Random Field." *IEEE T. Image Processing*, **7**(6):925-931, 1998.

- [22] B. Potetz. "Efficient Belief Propagation for Vision Using Linear Constraint Nodes." *CVPR2007*.
- [23] I. G. Rosenberg. "Reduction of Bivalent Maximization to the Quadratic Case." *Cahiers du Centre d'Etudes de Recherche Operationnelle* 17. pp. 71-74, 1975.
- [24] S. Roth and M. J. Black. "Fields of Experts: A Framework for Learning Image Priors." *CVPR2005*, II:860-867.
- [25] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer. "Optimizing Binary MRFs via Extended Roof Duality." *CVPR2007*.
- [26] 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.
- [27] O. J. Woodford, P. H. S. Torr, I. D. Reid, and A. W. Fitzgibbon. "Global Stereo Reconstruction Under Second Order Smoothness Priors." *CVPR2008*.

付 録

公式 (11)(12) の証明

$$m_d = \left\lfloor \frac{d-2}{2} \right\rfloor, \quad A = \min_{w \in \mathbb{B}^{n_d}} \sum_{i=1}^{m_d} w_i (-2S_1 + 4i - 1) + S_2$$

と定義する. 変数 x_1, \dots, x_d のうち k 個が 1 で残りが 0 であるとする, (10) より $S_1 = k, S_2 = k(k-1)/2$. 変数 w_i がそれぞれ独立して自由にとれるから

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

そこで $l = \lfloor k/2 \rfloor$ とすると, $-2k + 4i - 1 < 0 \Leftrightarrow i \leq l$ より

$$A = \sum_{i=1}^{\min(l, m_d)} (-2k + 4i - 1) + \frac{k(k-1)}{2} \quad (16)$$

特に $k \leq d-2$ のときは $l \leq m_d$ だから (16) の和は l まで, k が偶数・奇数の場合を調べてともに $A = 0$ を得る.

さて, まず偶数次の場合の (11) を見ると, $a > 0$ だから右辺 = aA であり, $k \leq d-2$ ならば両辺が 0 で一致している. $k = d-1$ のときも $l = m_d$ だからやはり $A = 0$ となる. $k = d$ のときは $m_d = l-1, k = 2l$ と (16) より

$$A = \sum_{i=1}^{l-1} (-4l + 4i - 1) + l(2l - 1) = 1 \quad (17)$$

となり, (11) の両辺が a となり証明された.

奇数次の場合, $m_d = n_d - 1$ で

$$\text{右辺} = a(A + \min(0, -S_1 + 2n_d - 1))$$

だが, $d = 2n_d + 1$ だから $k \leq d-2$ なら $-S_1 + 2n_d - 1 \geq 0$ より両辺が 0 で一致する. また $k = d-1$ のときは $-S_1 + 2n_d - 1 = -k + k - 1 = -1$ で, また $m_d = l-1, k = 2l$ より (17) と同様 $A = 1$ だから やはり (12) の両辺が 0 で一致する. 最後に $k = d$ のときは $-S_1 + 2n_d - 1 = -k + (k-1) - 1 = -2$ で, また $m_d = l-1, k = 2l+1$ と (16) より

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

より両辺 a となり (12) が証明された. \square