

# 非マルコフ的制約を導入した最適弾性マッチング

福富 正弘<sup>†</sup> 小川原光一<sup>††</sup> フォンヤオカイ<sup>†††</sup> 内田 誠一<sup>†††</sup>

<sup>†</sup>九州大学大学院システム情報科学府 〒 819-0395 福岡市西区元岡 744

<sup>††</sup>和歌山大学 システム工学部 〒 640-8510 和歌山市栄谷 930

<sup>†††</sup>九州大学大学院システム情報科学研究科 〒 819-0395 福岡市西区元岡 744

あらまし 時系列パターン間の最適弾性マッチングの一手法を提案する．弾性マッチングとは，2 パターン間の各時刻の対応関係を定める方法であり，例えばパターンに生じた時間的非線形伸縮を吸収する目的で利用される．本手法の最大の特徴は，離れた時刻の対応関係の直接制約，すなわち非マルコフ的制約を実現している点である．動的計画法 (DP) やその確率的拡張である HMM に基づいた従来一般的な弾性マッチングでは，原理的に隣接時刻間の制約条件すなわち単純マルコフ的制約しか扱えない．非マルコフ的制約は，弾性マッチングに求められる新機能として，今後様々な展開できる可能性を持っている．非マルコフ的制約を扱えるようにするために，DP ではなくグラフカットを用いて弾性マッチング問題を解く．グラフカットは DP とは全く異なる最適化原理であるため，非マルコフ的制約付きの最適化問題を扱える．本論文では本手法の原理を説明するとともに，その有効性をオンライン文字データを対象とした実験を通して検証する．

キーワード 弾性マッチング，非マルコフ的制約，グラフカット

## 1. はじめに

本論文では，時間的非線形伸縮を伴う時系列パターンの弾性マッチングの一般化について検討する．動作や音声などの時系列パターンには，一般に時間的非線形伸縮が伴う．この非線形伸縮により，ユークリッド距離や内積計算をそのまま行っても，2 パターン間の相違度や類似度を適切に評価できない．特に時間長まで変わってしまうと，パターンの次元が変わるため，距離計算そのものが不可能になる．このため，時系列パターンを扱う場合には，DP [1] や HMM [2] といった，2 パターンの時刻間の非線形な対応付け (写像) を行う手法，いわゆる弾性マッチングが利用されている<sup>(注1)</sup>

弾性マッチングを行う際には，一般に何らかの制約条件が課せられる．すなわちパターン間の完全自由な対応付けを許容しない．むしろ，対象となるパターンの性質に着目した制約を積極的に導入することで，不自然な対応関係の発生を抑制する．

制約条件の代表例に，単調性や連続性がある．単調性とは，時間的に逆転する対応付けを許さない制約である．連続性とは，時間的に大きなジャンプを含むような対応付けを許さない制約である．単純化した例として系列「こんにちは」を考えると，それを「こんちは」と対応付けるのは許容しても，「こちんには」(単調性違反) や「こんは」(連続性違反) と対応付けるのは許さない．単調性と連続性は，位相保存写像を近似実現するものであり，不自然な対応関係の抑制に妥当な制約と言える．

これら単調連続性制約をはじめとして，DP や HMM に基づく一般の弾性マッチングにおける制約条件は，全て単純マルコフ的 (以下，マルコフ的) である．マルコフ的制約とは，隣接する時刻間の対応関係に関する制約を指す．例えば，時系列パターン  $X = x_1, \dots, x_t, \dots, x_T$  上の隣接 2 時刻  $t, t+1$  が，パターン  $Y = y_1, \dots, y_{\tau}, \dots, y_{\tau'}$  上の時刻  $\tau, \tau'$  にそれぞれ対応する場合を考える．このとき，単調性は  $\tau \leq \tau'$ ，連続性は  $\tau - \varepsilon \leq \tau'$  なる制約条件で表される．いずれも隣接 2 時刻  $t, t+1$  に関する制約なので，これらはマルコフ的である．

これに対し，本論文では，非マルコフ的制約を導入した弾性マッチング法を提案する．以下に述べるように非マルコフ的制約は非常に有用でありながら，これまでそれを導入した手法は，筆者らの知る限り存在しない．従って本手法は，非マルコフ的制約という弾性マッチングに有用な新機能を初めて実現したものである．そして本手法を出発点として今後様々な方向に展開できる可能性を持っている．

非マルコフ的性質を直感的に理解できる例として手書き数字の「0」を考える．この「0」の筆記を表す時系列パターンは，上部の始端点から下降を開始し，下まで降り切ってから再び上昇し，始端点と同じ位置を終端点として終了する．この単純な時系列の生成過程を考えると，実はマルコフ性の下では表現できない．マルコフ性の下では，隣接時刻間のみが直接の依存関係がある．このため，単純に言えば，始端点から離れると，すぐに始端点のことを忘れてしまう．終端点付近では始端点位置との依存関係はほとんど残っていない．よって，始端点終端点を一致させて閉曲線にしたいとすることも不可能になる．このように，始端点と終端点という時間的に離れた箇所につ

(注1): DNA 塩基配列のような記号列パターンも，非線形伸縮が起こるという点で，DP と HMM が広く用いられている [3]．本論文で提案する手法は，こうした記号列パターンにも問題なく利用できる．

いて、それらの依存関係を直接制御するような仕組み、すなわち非マルコフ性が必要になる。

この「0」の例は、生成過程における非マルコフ性であり、弾性マッチングにおける時刻対応関係における非マルコフ性とは多少異なる<sup>(注2)</sup>。それでもこの例を出して主張したかったのは、非マルコフ性が決して特異なものではなく、我々の身の回りのパターンにおいて頻出しているという事実である。例えば、全体的な形のバランスを取りながら字を書いていくというのは「過去すべての筆記結果を見ながら」次の筆記方向を定めるという意味において、非マルコフ性の現れである。すなわち、次の筆記方向は現時点の筆の位置だけでは定まらない。筆記に限らずとも、我々は往々にして直前よりさらに前の行動履歴を参照しながら次時刻の行動を計画する。周期運動などの複雑なパターンのモデル化も、マルコフ性の下では限界がある。

従来の弾性マッチング法で非マルコフ性が扱われなかった理由は、従来法が最適化法としてDP(HMM内部のViterbiアルゴリズムもDP)を常套的に用いていたことが大きいと考えられる。DPは逐次決定過程上での最適化法であり、非マルコフ性を扱えない。理論的には、 $N(> 1)$ 時刻前に依存するような非マルコフ的問題( $N$ 重マルコフ的問題)も、マルコフ的問題に等価変換することができるので、DPでも解けないことはない。しかし、計算量が $N$ に関する指数オーダーで増加するため、現実的ではない。古くは如何にDP計算に有利なように等価変換するかについての研究もあったが[4]、抜本的な解決には至っていない。このように常套手段であるDPが非マルコフ性を扱えないために、自ずと問題側もマルコフ的なものに限定されてきたと考えられる。

本手法は、DPではなくグラフカットを最適化法として用いることで、弾性マッチング問題への非マルコフ的制約の導入を実現している。単純に言えば、グラフカットはDPとは全く異なった原理に基づく最適化法であるため、マルコフ性の縛りが無く、比較的自由に非マルコフ的制約を導入できる。DPと同じく大局的最適解を保証できる方法であり、画像修復[5],[6]、物体認識[7]、輪郭マッチング[8],[9]、領域分割[9]~[12]等、広く利用されている。後に詳述するように、非マルコフ的制約を導入してもアルゴリズムがさほど複雑化しない点も利点である。実際、本手法の計算量は多項式時間の範囲で収まる。

グラフカットを用いた弾性マッチング自体は、1990年代後半に例えばIshikawaら[13]により提案されている。同文献では、ステレオマッチングにおけるエッジ線素上の画素対応付け問題にグラフカットが利用されている。

(注2): ただし、それはあくまで表層的な違いであり、本手法を多少の拡張すればこうした問題をも統一的に扱えることを注意しておく。具体的な拡張法については、付録で若干触れるに留め、詳細は稿を改めて論じる予定である。

この問題自体は時系列マッチングではないが、弾性マッチング問題の一種であり、やはり数多くの手法がDPを用いてきた。同文献は、DPの代わりにグラフカットが利用できることを示したのである。

本手法は、Ishikawaらの手法[13]を拡張して、非マルコフ的制約を積極的に組み込んだものである。この組み込みは極めて容易で、大幅な原理的拡張や解法の変更は不要である。ただし、上でも強調したように、非マルコフ的制約の導入効果もたらず意義は非常に大きいと期待され、この嚆矢たることこそが本論文の価値である。また様々な応用分野では弾性マッチングの最適化法として依然としてDPが多用されている現状で、グラフカットをより積極的に利用することの動機づけにもなると考えている。

本論文の構成は以下の通りである。2.では、従来のDPによる弾性マッチング、すなわちマルコフ制約下で弾性マッチングについて述べる。次に、3.では、DPの代わりにグラフカットを用いた弾性マッチング[13]について詳述する。そして4.では、本手法すなわち非マルコフ的弾性マッチングについて述べる。5.では時系列パターンの一種であるオンライン文字パターンを対象とした実験結果およびその考察を与える。最後に6.では結論を述べる。

## 2. DPによる弾性マッチング

### 2.1 定式化とマルコフ的制約

2つの時系列パターン  $X = x_1, \dots, x_t, \dots, x_T, Y = y_1, \dots, y_\tau, \dots, y_T$  の弾性マッチングを考える。ここで、 $x_t$  および  $y_\tau$  は各時刻の特徴ベクトルである。弾性マッチングとは、 $X$  と  $Y$  間の時間的対応付け  $\tau = u_t$  を最適化する問題である。ここで、制御変数  $u_t (t = 1, \dots, T)$  は  $x_t$  を  $y_{u_t}$  に対応付けることを意味する。この対応付けは図1(a)中に示すように、 $t-\tau$  平面中の曲線(マッチングパス)として表現される。

マルコフ的制約すなわち隣接時刻における対応関係( $u_t$  と  $u_{t+1}$ )の間に範囲制約を考える場合、弾性マッチングは、次の目的関数  $F$  の  $U = u_1, \dots, u_T$  に関する最小化問題として定式化される。

$$\min F = \min_{\substack{u_1, \dots, u_T \\ \alpha_{t,t+1} \leq u_{t+1} - u_t \leq \beta_{t,t+1}}} \sum_{t=1}^T d_t(u_t) \quad (1)$$

ここで、 $d_t(u_t)$  は  $x_t$  と  $y_{u_t}$  を対応付ける場合のコスト(局所コスト)を表し、例えば  $d_t(u_t) = \|x_t - y_{u_t}\|$  と定義される。

条件  $\alpha_{t,t+1} \leq u_{t+1} - u_t \leq \beta_{t,t+1}$  はマルコフ的制約である。この制約は、隣接2時刻の対応先での差異  $u_{t+1} - u_t$  を特定の範囲内  $[\alpha_{t,t+1}, \beta_{t,t+1}]$  に抑える効果がある。なお、この  $u_{t+1} - u_t$  はマッチングパスの勾配であるため、この条件は傾斜制限[1]とも呼ばれる。1.で例示した単

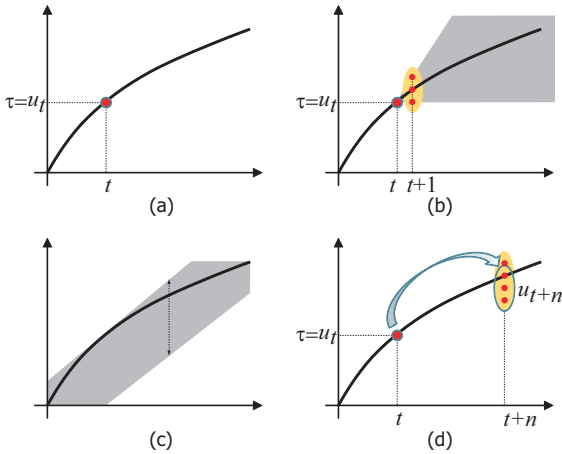


図1 (a) 対応付け  $\tau = u_t$  による弾性マッチングの表現, (b) マルコフ的制約 (単調連続性制約), (c) 整合窓制約, (d) 非マルコフ的制約

調連続性制約ならば, 図1 (b) に示すように, 全ての  $t$  において  $\alpha_{t,t+1} = \alpha = 0$ ,  $\beta_{t,t+1} = \beta = \epsilon$  とすればよい (同図では  $\epsilon = 2$ ). なお, 多くの場合, この条件に加え, 始末端条件  $1 = u_1, T = u_T$  を課する. これは  $X, Y$  の始末端点を共に対応付ける条件である.

## 2.2 DP による解法

式 (1) は, 動的計画法 (DP) により解くことができる [1]. 具体的には,

$$f_t(u_t) = d_t(u_t) + \min_{\alpha_{t-1,t} \leq u_t - u_{t-1} \leq \beta_{t-1,t}} f_{t-1}(u_{t-1}) \quad (2)$$

とおくと,  $\min F = f_T(T)$  の形に変形できることを利用する. 式 (2) を DP 漸化式という. この DP 漸化式を,  $t = 2$  から  $T$  まで順に全ての  $\tau$  について計算すると大局的最適解が得られる.

再度の説明になるが, マルコフ的制約は, 隣接時刻の  $u_t, u_{t+1}$  間の関係を直接制限する. その結果, 例えば隣接間の傾斜制限を  $[\alpha, \beta] = [0, 2]$  とすると,  $u_{t+n} - u_t$  の取り得る範囲は  $[0, 2n]$  となる. このように, マルコフ的制約では, ある程度時刻が進んだ際の取り得る範囲が大きくなる. 別の言い方をすれば, 間隔  $n$  が離れば離れるほど,  $u_t$  が  $u_{t+n}$  に及ぼす影響が小さくなる. この様子を図1 (b) に示す. このため, 離れた2時刻の対応付け  $u_t, u_{t+n}$  の間に何かしらの制約を導入する手段として, マルコフ的制約は不十分であると言える.

## 2.3 整合窓

DP マッチングでは, 単調連続性制約すなわちマルコフ的制約に加え, 整合窓制約 [1] を導入することが多い. これは, 各時刻  $t$  で独立に課せられる制約で, 単に時刻  $t$  での  $u_t$  の範囲に上限下限を与えるものである. 図1 (c) にその様子を示す. 0次マルコフ的制約とも呼べるこの制約は, マルコフ的制約では抑えきれない極端な対応付けをある程度排除することもできる. ただし, 単純なだけに窓幅を変えるだけの単純な制御しかできない. 例え

ば, 窓幅を狭くして極端な対応付けを排除しようとするれば, 弾性マッチングの効果そのものも抑えてしまう事態も起こりうる.

## 3. グラフカットによる弾性マッチング

### 3.1 概要

式 (1) の最適弾性マッチング問題の大局的最適解は, グラフカットでも求まる [13]. DP では図2 (a) に示す  $t-\tau$  平面上で系列  $U = u_1, \dots, u_T$  を最適化していた. これに対し, グラフカットに基づく手法では,  $U$  がそのグラフのカット (切断) となるような有向グラフを準備する. 同図 (b) はこのグラフの例である. 詳細は後述するが, (a) (b) を見比べると類似した構造が背後に見て取れる. そして同図 (a) 内でパスとして表されている  $U$  は, (b) ではカット (図中点線) として表される. この準備の上で, DP において様々なパスを探索しながら最適な  $U$  を導出するのと同様, グラフカットでは様々なカットのうち最小カットを導出する.

図2 (b) の有向グラフ  $G = (V, E)$  は一見複雑である. これは単調連続性条件や始末端条件などを満たしたカットが求まるようにするための「仕掛け」である. グラフ中のエッジの多くは無量大の重みを持っている. すなわちそのエッジを切断するようなカットは許されない. これがまさに都合の悪いカット, すなわち制約を満たさない対応付けを除外するための「仕掛け」である. 以下では, これを機能ごとに分解して示した同図 (c)-(h) を用いて説明する.

なお, 以下で説明するものは, DP との対比の都合上, 文献 [13] の手法における制約をより厳しくしたものとなっている. すなわち, 同文献では, 隣接時刻間の関係を, 単調連続性のような不等式制約ではなく, ペナルティ  $\lambda |u_{t+1} - u_t - 1|$  により緩く制御していた. こうしたペナルティは, 以下で述べるエッジの無量大の重みを, 有限値に置き換えることで容易に実現する. いずれにせよ, この違いは本質的なものではない. 本手法と同文献の差異は, あくまで非マルコフ的制約の導入の有無にある.

### 3.2 グラフのノード

有向  $G$  のノード集合  $V$  には, 2つの特殊なノード source, sink がある. 有向グラフ上のフローはこの source から出て sink に流れ込む. それ以外のノードは, 図2 (a) の格子点  $(t, \tau)$  それぞれに対し2個ずつ準備される  $v_{t,\tau}^0$  および  $v_{t,\tau}^1$  である. (c) ではこのノード対を小楕円で囲み, 対であることを明示している.

### 3.3 局所コスト

このノード対  $v_{t,\tau}^0$  および  $v_{t,\tau}^1$  の間にはエッジが存在し, そこには局所コスト  $d_t(\tau)$  が重みとして付与される. このため, もし  $v_{t,\tau}^0$  から  $v_{t,\tau}^1$  へのエッジが切断 (カット) されれば, そのカットには  $d_t(\tau)$  のコストが印加される.

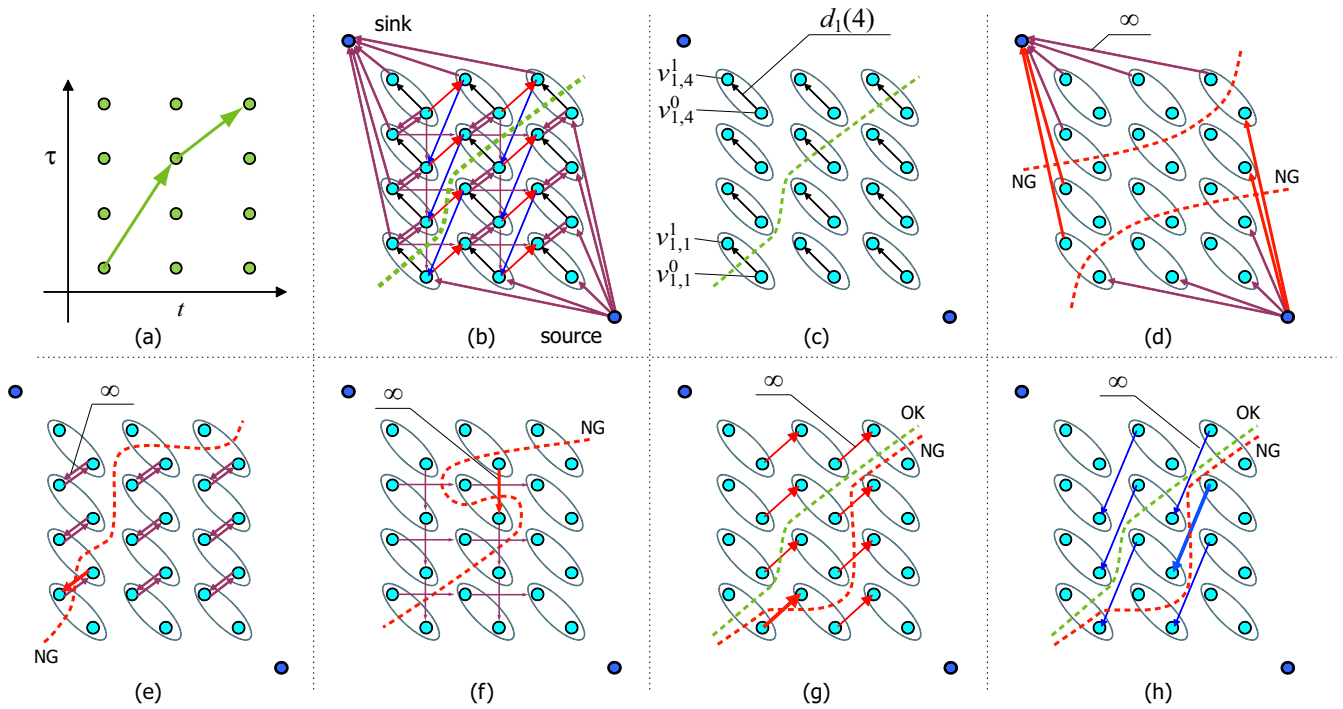


図2 グラフカットによる最適弾性マッチングのためのグラフ。(a) DP で用いる  $t$ - $\tau$  平面およびマッチングパスの例。(b) グラフカットによる弾性マッチングのためのグラフの全体像。(c)-(h) グラフ (b) を構成するエッジの個別表示。(c) 局所コスト  $d_i(u_i)$  を表すエッジ。(d) 始末端条件を表すエッジ。(e)(f) 一価性条件を表すエッジ。前者は  $x_i$  に複数の  $y_\tau, y'_\tau$  が対応する状況を回避し、後者はパスの逆戻りを防ぐ。(g) 下限値  $\alpha_{t,t+1} = 1$  を実現するためのエッジ。(h) 上限値  $\beta_{t,t+1} = 2$  を実現するためのエッジ。

これはすなわちそのカットにより  $x_t$  と  $y_\tau$  を対応づけたことを意味している。このことから、(a) 上のマッチングパスのコストは、(c) 上のカット(点線)のコストに等しいことがまずわかる。

### 3.4 始末端条件

同図 (d) は始末端条件すなわち  $1 = u_1, T = u_T$  を表現するエッジである。いずれも無限大の重みを持っており、カットすることは許されない。このため許されるカットは、左下から入って(すなわち  $v_{1,1}^0$  と  $v_{1,1}^1$  の間のエッジをカット)、右上へ抜ける(すなわち  $v_{T,T}^0$  と  $v_{T,T}^1$  の間のエッジをカット)のみに限定される。このように始末端条件が実現していることがわかる。

### 3.5 一価性条件

図 (e) および (f) はグラフカットによる解法に独特の条件である。ここでは両者を併せて一価性条件と呼ぶことにする。具体的には、(e) は  $x_i$  に複数の  $y_\tau, y'_\tau$  が対応する状況を回避するためのエッジ、(f) はパスの逆戻りを防ぐためのエッジである。すなわち  $t$  が  $t+1$  に進行するに従って、マッチングパスすなわちカットも(停留や後戻りせずに)必ず右方向に進行していくための条件である。別の言い方をすれば、対応付け  $\tau = u_t$  により  $t$  に対応する  $\tau$  が一意に定まるための条件である。この意

味によりここでは一価性条件と呼んでいる。

ここで初歩的な補足を与える。図 (f) の“NG”のカットにおいて無限大のコストを与えているのは一か所のエッジ(太矢印で表示)だけであることに注意されたい。他にも多くのエッジをカットしているように見えるが、フローの流れ(source→sink)からみて、それらはカットにならない。

### 3.6 単調連続性条件

最後に単調連続性条件  $\alpha_{t,t+1} \leq u_{t+1} - u_t \leq \beta_{t,t+1}$  をグラフ上に設定する。これは下限  $\alpha_{t,t+1}$  と上限  $\beta_{t,t+1}$  をそれぞれ設定することで実現する。なお、図2を含め本節では簡単のために、 $\alpha_{t,t+1}$  および  $\beta_{t,t+1}$  は  $t$  に依らずそれぞれ 1, 2 であるとする。

図 (g) は下限値を設定したものである。いずれも重み無限大を持つエッジにより設定される。今  $\alpha_{t,t+1} = 1$  を想定しているので、 $t$  が増えるごとに、パスすなわちカットの  $\tau$  値は少なくとも一つ増加する必要がある。もし増加しない場合、同図中“NG”を付されたカットのように、この無限大のエッジ(太矢印で表示)を切断してしまう。従って必ず増加する必要がある。増加する分については幾ら増加しても問題ない(同図“OK”のカットも同種のエッジを切断しているように見えるが、一価性条件のところ述べたのと同様、このエッジは切断されて

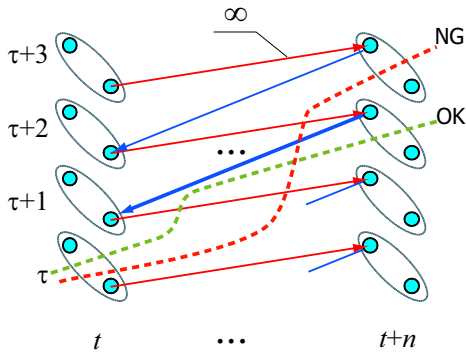


図3 非マルコフ的制約の導入． $\alpha_{t,t+n} = 1, \beta_{t,t+n} = 3$  の例

いない.)

図(h)は上限値  $\beta_{t,t+1} = 2$  を設定したものである．やはり重み無限大を持つエッジにより設定される．“NG”をつけたカットは， $d_2(1)$ の次に $d_3(4)$ を通過，すなわち3つ分増加しており，上限値  $\beta = 2$  を超えた増加をしている．このカットを見ると，このエッジの一つ(太矢印で表示)を切断しており，従って実際に選択されることはないことがわかる．

以上により単調連続性条件も導入できた．再三述べたように，この条件はマルコフ的制約である．なお，以上の議論からわかるように，パラメータ  $\alpha$  や  $\beta$  を変えると，このエッジの傾きが変わる(別のノード間を結ぶ)ことがわかる．時刻  $t$  毎にこれらパラメータの値を変えることも可能である(例えば  $\alpha_{1,2} = 1$  だが  $\alpha_{2,3} = 0$  とするなど)．この場合  $t$  と  $t+1$  の間毎に同図(g)(h)のエッジの傾きを変えればよい．

## 4. 非マルコフ的弾性マッチング

### 4.1 非マルコフ的制約

非マルコフ的弾性マッチングでは，パターン  $X$  と  $Y$  のマッチングにおいて，隣接時刻間のみではなく，非隣接時刻間すなわち  $u_t$  と  $u_{t+n}$  間 ( $n \geq 1$ ) にも直接制約を導入する．DPで非マルコフ的問題を効率的に扱うことは原理的に不可能である．このため，その重要性に反して，これまで最適弾性マッチング問題で導入されたことはなかった．

本論文では，DPではなく，前節のグラフカットによる最適弾性マッチングに対して，非マルコフ的制約を導入する．グラフカットの利用により，非マルコフ的制約下での大局的に最適なマッチング結果が極めて効率的に求まる．さらに制約の導入は極めて容易であり，以上の議論を踏まえれば，その原理も明確である．すなわち実装上は小さな変更で済む．ただし，それが意味するところは，今後の発展可能性を考えると非常に大きい．

非マルコフ的弾性マッチングは，(1)と同じ目的関数  $F$  を，(1)とは異なる制約条件のもとで最小化する問題として定式化される．

$$\min F = \min_{u_1, \dots, u_T} \sum_{t=1}^T d_t(u_t) \quad (3)$$

$$\alpha_{t,t+n} \leq u_{t+n} - u_t \leq \beta_{t,t+n}$$

最小化の際の制約条件が  $t+1$  から  $t+n$  になっていることに注意されたい．このように非マルコフ的制約を導入することにより， $n$  時刻進んだ先の対応付けを直接制限することができ，従来以上に細かな制御が可能になる．ここで  $t$  と  $t+n$  間の制約条件は， $u_t$  と  $u_{t+n}$  へのみ直接課せられている点に注意されたい．すなわち，図1(d)に示すように，本論文で扱う非マルコフ的制約は， $u_{t+1}, \dots, u_{t+n-1}$  には直接影響しない．

すべての  $t$  と  $t+n$  の組について制約を導入する必要はない．さらに制約毎に  $n$  を変えてもよい．例えば，時刻4と9，7と11の間には  $\alpha, \beta$  による制約があるが，5と10の間にはない，といった状況も許される．逆に言えば，どの時刻とどの時刻の間にもどのようなパラメータ  $\alpha, \beta$  で非マルコフ的制約を導入するかについては，事前に何らかの方法で決定しておく必要がある．4.3ではその方法一例を示す．

### 4.2 グラフカットによる解法

グラフカットに基づく最適弾性マッチングへの，非マルコフ的制約の導入は極めて容易である．前節では，単調連続性条件を表現するために，隣接時刻のノード間のみエッジを挿入していた．これに対し，離れた時刻のノード間 ( $v_{t,\tau}^0$  と  $v_{t+n,\tau'}^1$ ) にエッジを挿入するだけで，非マルコフ的制約を導入できる．

図3は非マルコフ的制約を表したエッジである．重みは無量大である．これは時刻  $t$  と  $t+n$  間に  $1 = \alpha_{t,t+n} \leq u_{t+n} - u_t \leq \beta_{t,t+n} = 2$  なる非マルコフ的制約を入れた場合に相当する．すなわち，時刻  $t$  から  $n$  時刻進んだ後に，最小で1，最大で2だけ対応する  $\tau$  が進んでほしい，という制約である．同図の  $t+n$  を  $t+1$  と読み替えれば，図2における単調連続性条件の実現と全く同様のエッジが利用されていることがわかる．

非マルコフ的制約を導入した場合，DPで無理に解こうとすれば計算量爆発を起こすのに対し，グラフカットならば計算量の増加を最小限に抑えられる．実際，グラフカットの計算量はグラフの規模に依るので，マルコフ的制約も非マルコフ的制約も計算時間への影響は同じである．従って，グラフカットの一般的なアルゴリズムが多項式時間であることを考えれば，非マルコフ的弾性マッチング問題もやはり多項式時間で解ける．

付録1.に示すように，グラフカットの枠組みではさらに拡張された非マルコフ的制約を用いることも可能である．具体的には， $\tau$  に応じて  $\alpha_{t,t+n}, \beta_{t,t+n}$  を変化させることができる．さらに付録2.に示すように，1次元時系列パターン間だけでなく，1次元パターンと2次元パターンのマッチングにも本手法は拡張可能である．ただし，注意すべきは，そうして拡張された非マルコフ的制

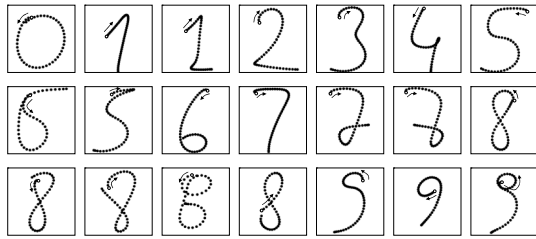


図 4 標準パターン Y

約については、任意のものがグラフカットの枠組みで扱えるわけではない。こうしたグラフカットによる非マルコフ的弾性マッチングの限界についても、付録 3. で触れる。

### 4.3 学習による非マルコフ的制約の自動設定

非マルコフ的制約をどこの時刻対にどのようなパラメータ値 ( $\alpha_{t,t+n}, \beta_{t,t+n}$ ) で導入するかは、非マルコフ的制約を活用する上で重要な検討項目である。以下では、非マルコフ的弾性マッチングを認識タスクに利用する場合について、非マルコフ的制約を学習により自動的に設定する簡便な方法を示す。

認識タスクにおける非マルコフ性制約導入の目的は、以下の 2 点に配慮しながらマッチングを制約することにある：(i) 同一カテゴリ内に生じる非線形伸縮だけを正しく吸収、(ii) 異なるカテゴリのパターンについては不必要な非線形伸縮により合わせ過ぎを起こさせない。

そこで、同一カテゴリのパターン間において、従来型のマルコフ的弾性マッチングを行い、カテゴリ内に実際に生じる非線形伸縮  $U$  を多数収集する。次に、収集された非線形伸縮について、各時刻対  $t, t+n$  において対応する時刻の変化  $u_{t+n} - u_t$  の最大値と最小値を求める。この最小値および最大値を、時刻対  $t, t+n$  の上下限パラメータ  $\alpha_{t,t+n}, \beta_{t,t+n}$  として設定する。これを全ての時刻対  $t, t+n$  について行うことで、そのカテゴリ内に起こる非線形伸縮の範囲を、非マルコフ的制約として設定できる。

## 5. 実験

### 5.1 実験条件

以上で述べた非マルコフ的弾性マッチングの効果を検証するために、オンライン文字データを対象として実験を行った。以下の通り、本実験は、制約条件に係る部分を除き、シンプルに実施された。これは非マルコフ的制約の効果を吟味しやすくするためである。すなわち認識率を追及することは本実験の目的ではない。

実験に使用したデータは Ethem Alpaydin Digit のオンライン手書き数字データ (“0”–“9”) である。非マルコフ的制約の学習用 (4.3) に各数字 700 パターン、認識時のテストパターン  $X$  用に各数字 300 パターンを用いた。標準パターン  $Y$  には図 4 に示す 21 種を用いた。各時刻の特徴量はペン先位置を表す 2 次元座標特徴を用い

表 1 認識率 (%)

	マルコフ	マルコフ+整合窓	非マルコフ
全体平均	86.6	88.7	89.4
クラス “6”	75.7	83.3	86.3
クラス “1”	67.0	73.3	74.7

た。局所コスト  $d_t(u_t)$  には、各時刻  $t$  の平均・共分散によるマハラノビス距離を用いた。各テストパターン  $X$  について、全 21 種の標準パターン  $Y$  と、非マルコフ的弾性マッチングを行い、得たコスト  $\min F$  に基づく最小距離識別則で、認識結果を定めた。

比較対象として、(i) 隣接時刻間のみを制約を導入した手法 (マルコフ的弾性マッチング)、および (ii) (i) に整合窓を導入した手法、を用いた。整合窓を導入したのは、それと非マルコフ的制約のどちらがより有効な制約であるかを見る為である。整合窓については、最大限の工夫を施した。すなわち学習パターンにより各時刻  $t$  における  $u_t$  の最大値および最小値を求め、それに基づいて  $t$  毎に異なる整合窓幅を設定した。すなわち図 1(c) の状況よりも詳細に設定したものである。

### 5.2 実験結果

まず認識に要した時間を示す。300(パターン)  $\times$  10(クラス)  $\times$  21(カテゴリ) =  $6.3 \times 10^5$  回の非マルコフ的弾性マッチング処理に要した時間は、民生用 PC を用いて 300s 以下であった。従ってマッチング 1 回あたりはおおよそ 5ms である。DP で非マルコフ的制約を導入すればすぐに実行不能になることを考慮すれば、これは非常に高速であると言える。

次に認識率を表 1 に示す。同表より、非マルコフ的制約を導入することにより、わずかながらではあるが認識率が向上していることがわかる。向上した理由などについては、次節で考察する。

### 5.3 考察

非マルコフ的制約により認識率が向上した理由について、“6” および “1” の二つのテストパターンを例に挙げて考察する。図 5 (a) はそれらテストパターンである。マルコフ的制約のみ (整合窓あり) の場合、これら “6” および “1” は、それぞれ “8” と “9” に誤認識されていた。それに対し、非マルコフ的制約を導入した場合は、いずれも正しく認識された。

図 (b) は従来のマルコフ的弾性マッチング (整合窓あり) により誤認識された結果である。左側に最短距離を与えた標準パターン、右側に対応付け結果  $U$  を示している。後者について、横軸はテストパターン  $X$  の時刻  $t$  を、縦軸は標準パターン  $Y$  の時刻  $\tau$  をそれぞれ示している。テストパターン “6” と標準パターン “8” のマッチングパスを見ると、前半で急激に変化している。これは、“6” と “8” の後半部分が非常に類似しているため、前半部分は粗に、後半部分は密に対応付けすれば、 $\min F$  を

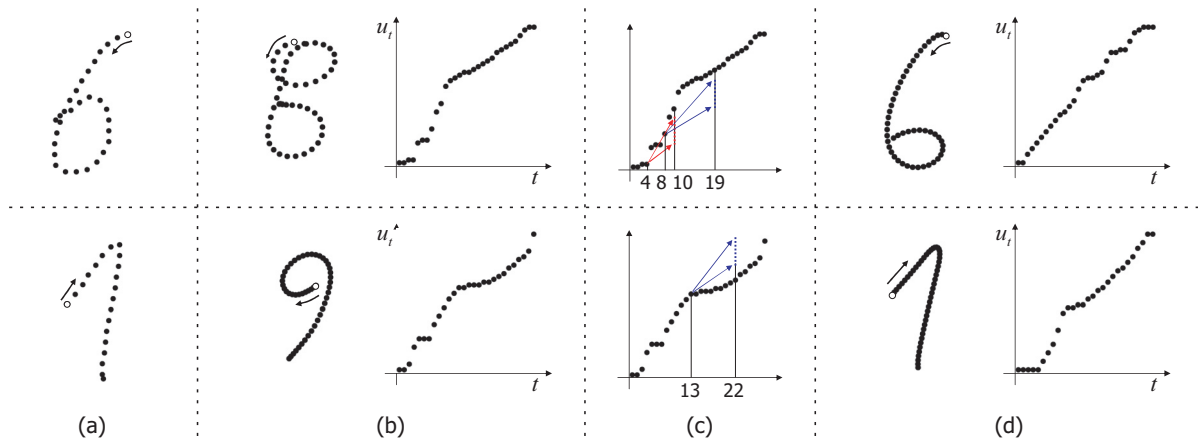


図5 入力パターンとその認識結果。(a) 入力パターン (b) マルコフ的弾性マッチング結果。左：対応付けられたカテゴリの標準パターン。右：対応付けの様子。(c) 非マルコフ的制約の影響。(d) 非マルコフ的弾性マッチングによる結果。

極小化できるためである。このようなマッチングパスの急激な伸縮が、マルコフ制約下では許容される。そしてそれが原因で今回の誤認識“6”→“8”が起っていた。

図(c)の上の図はこれら“6”と“8”の非マルコフ的弾性マッチングの際に利用された、時刻(4,10)および(8,19)における非マルコフ的制約を、先のマッチングパス上に重畳表示したものである。この制約により、2時刻 $t, t+n$ において、 $u_{t+n}$ が点線で示した範囲内でなければならない。同図に示した2時刻では、時刻が離れているにもかかわらず、非マルコフ的な相関があり、制約が存在していた。結局、この非マルコフ的制約下では、(b)のようなマッチングパスは許されず、結果的に、極端な伸縮吸収が抑制される。そして最終的には正しいカテゴリとのマッチング(d)のほうが小さな $\min F$ を与え、正しい認識結果が得られた。下段の文字対“1”と“9”についても同様の考察が可能である。

## 6. まとめと今後の課題

時間的非線形伸縮を伴う時系列パターンの認識のために、非マルコフ的弾性マッチングを提案した。これは隣接時刻間のみではなく、非隣接時刻間にも制約を導入する手法である。従来、こうした非マルコフ的制約はあまり顧みられることがなかった。しかし、実世界には非マルコフ的制約を受けるようなパターンが多く見られる。このため、本手法の開発は、今後そうしたパターンを扱う問題に道を開くことと期待している。

非マルコフ的制約の導入は、マッチングの最適化のために、旧来用いらてきたDPの代わりに、グラフカットを採用したことで実現できた。計算量も多項式時間で済むため、実用的価値も高い。オンライン文字パターンを対象とした実験を行い、本手法の有効性を確認した。

非マルコフ的制約が実用的に扱えるようになったために、今後は多様な方向での展開を期待している。それら方向性について以下、具体例を挙げる。

- 第一に、非マルコフ的制約の学習法がある。今回

は単純な学習法を利用したが、マルコフ的弾性マッチングの結果を利用した手法であるために、その精度には限界があると思われる。むしろ対象に潜む非マルコフ性を積極的に抽出できるような学習方式が望ましい。

- 第二に、非マルコフ的制約の拡張が挙げられる。これは付録1で触れた $\tau$ に応じて変化する制約である。利用価値や拡張限界を吟味したい。

- 第三には、マッチング問題自体の拡張である。今回は時刻 $t$ と $\tau$ を対応付ける問題を解いていたが、付録2に示すように $t$ と $x, y$ 座標を対応付けるようなマッチング問題も原理的に扱える(例えば、文字「画像」に時系列文字データをフィッティングする場合がこれに相当する。冒頭で述べた「0」の筆記もこの類である。)

謝辞 本研究の一部は、JST 戦略的創造研究推進事業チーム型研究(CREST)「共生社会に向けた人間調和型情報技術の構築」に依った。

## 文 献

- [1] 内田誠一, “DP マッチング概説 ~ 基本と様々な拡張 ~ ” 信学技報, PRMU2006-166, 2006.
- [2] L. R. Rabiner and B. H. Juang, “An Introduction to Hidden Markov Models,” IEEE ASSP Mag., 3(1), 1986.
- [3] R. Eddy, A. Krogh, G. Mitchison, R. Durbin, *Biological Sequence Analysis*, Cambridge University Press, 1999.
- [4] U. Bertele and F. Brioschi, *Nonserial Dynamic Programming*, Academic Press, 1972.
- [5] S. Kumar and M. Hebert, “Discriminative Fields for Modeling Spatial Dependencies in Natural Images,” Proc. NIPS, 2003.
- [6] D. Cremers and L. Grady, “Statistical Priors for Efficient Combinatorial Optimization via Graph Cuts,” Proc. ECCV, 2006.
- [7] D. Anguelov, B. Taskar, V. Chatalbashev, D. Koller, D. Gupta, G. Heitz, and A. Ng, “Discriminative Learning of Markov Random Fields for Segmentation of 3D Scan Data,” Proc. CVPR, 2005.
- [8] F.R. Schmidt, E. Toppe, D. Cremers, and Y. Boykov, “Efficient Shape Matching via Graph Cuts,” Proc. CVPR, 2007.
- [9] F.R. Schmidt, E. Toppe, and D. Cremers, “Efficient Planar Graph Cuts with Applications in Computer

- Vision,” Proc. CVPR, 2009.
- [10] Y. Boykov and M. P. Jolly, “Interactive Graph Cuts for Optimal Boundary Region Segmentation of Objects in N-D Images,” Proc. ICCV, 2001.
- [11] B. Price, B. Morse, and S. Cohen, “Geodesic Graph Cut for Interactive Image Segmentation,” Proc. CVPR, 2010.
- [12] T. Schoenemann, F. Kahl, and D. Cremers, “Curvature Regularity for Region-based Image Segmentation and Inpainting: a Linear Programming Relaxation,” Proc. ICCV, 2009.
- [13] H. Ishikawa and D. Geiger, “Occlusions, Discontinuities, and Epipolar Lines in Stereo,” Proc. ECCV, 1998.
- [14] 内田誠一, 迫江博昭, “動的計画法に基づく単調連続 2 次元ワーブ法の検討,” 信学論, J81-D-II(6), 1998.
- [15] D. Keysers and W. Unger, “Elastic Image Matching Is NP-Complete,” Pattern Recog. Lett., 24(1-3), 2003.

## 付 録

### 1. 非マルコフ的制約の拡張

本論文で用いた非マルコフ的制約では, ある 2 時刻  $t, t+n$  間の制約を  $\tau = u_t$  に依らず一定としていた. すなわち  $\alpha_{t,t+n}$  も  $\beta_{t,t+n}$  も  $\tau$  には依存せず一定であった. グラフで考えれば, これは時刻  $t, t+n$  間のエッジは,  $\tau$  方向で一定していることを意味する.

これに対し, 実際には,  $\tau$  に応じて  $\alpha_{t,t+n}, \beta_{t,t+n}$  を変化させることも可能である. すなわち, 時刻  $(t, t+n)$  間の制約に関するパラメータを  $\alpha_{t,t+n}^{\tau}, \beta_{t,t+n}^{\tau}$  のように,  $\tau$  に応じて可変にできる. この実現するには, 図 A.1 (a) のようにエッジを作ればよい. 同図は, 例として,  $\alpha_{t,t+n}^{\tau} = 1, \beta_{t,t+n}^{\tau} = 3$  および  $\alpha_{t,t+n}^{\tau+1} = 1, \beta_{t,t+n}^{\tau+1} = 2$  と設定している.

### 2. 1 次元-2 次元マッチング問題への拡張

本手法の考え方は, 時系列すなわち 1 次元パターン間のマッチング問題だけでなく, より次元の高いパターンのマッチング問題にも適用可能である. すなわち, 次節 3. で述べるように限界はあるにしろ, 問題を適切に設定すれば, 様々なマッチング問題がグラフカットで解け, そして非マルコフ的制約を導入できる. 非マルコフ的制約の旨みは, 1. の「0」筆記始終端問題のように, むしろこうした問題においてさらに活きるだろう.

実証例として図 A.2 のような 1 次元-2 次元マッチング問題を考える. これは同図 (a) のようにグラフ表現された 2 次元パターン上に 1 次元パターンをマッチングする問題である. (要するにグラフ (a) 上の最適経路問題と思えばよい.) 文字を例にすれば, タブレットから入力された時系列パターンとしてのペン先座標列を, その文字の標準画像にフィッティングする問題に相当する. 逆に, 画像として入力された文字をその文字の標準筆跡時系列モデルと照合する問題と解釈してもよい. いずれにせよ, この場合は  $x-y$  平面という 2 次元空間を扱っている点が重要である. なお同図では, 簡単のために, 2 列で進行するような幅の細いグラフとなっているが, これは図示のわかりやすさのためであり, 実際にはさらに列数が多くても構わない.

この問題がグラフカットで解けることを図 A.2 で示す. DP でこの問題を解く場合, 同図 (b) に示す平面上のパス最適化問題として扱う. そして, 本論文で述べた議論により, これと等価な結果を与えるグラフカット用のグラフが同図 (c) のように得られる. なお同図では簡単のため, 一価性および始終端条件のためのエッジは表示していない. このグラフは, 図 2 で述べたものと同形式であり, グラフカットで最適解が求まることがわかる. そして, 図 3 と同形式のエッジをさらに追加すれば, 非マルコフ的制約を容易に導入できることもわかる. なおここで次節で述べるような制約違反は生じていないことを断っておく. ページ数の都合上省略するが, この方法により 1. で述べた「0」筆記の始終端問題も解決できる.

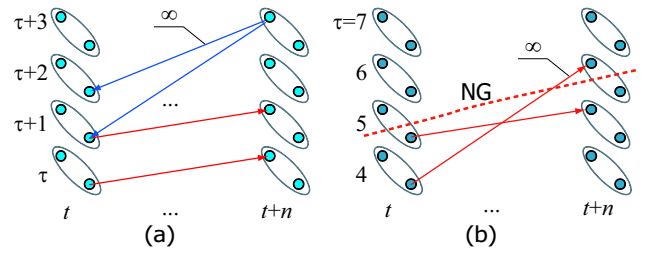


図 A.1 (a) 非マルコフ的制約の拡張. (b) グラフカットでは扱えない非マルコフ的制約

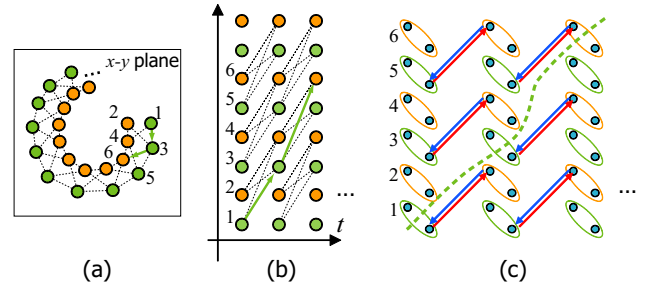


図 A.2 1 次元-2 次元マッチング問題への拡張

### 3. 導入できない非マルコフ的制約

グラフ上で任意の  $(t, \tau)$  と  $(t+n, \tau')$  の間にエッジを結ぶことができ, 結果的に任意の非マルコフ的制約を導入できそうに見えるが, 実はそうではない. すなわち, グラフカットで任意の非マルコフ的制約を扱えるわけではない. 実際には, 次式を  $\forall \tau < \tau'$  について満たす非マルコフ的制約のみが導入できる.

$$\tau + \alpha_{t,t+n}^{\tau} \leq \tau' + \alpha_{t,t+n}^{\tau'} \quad (\text{A}\cdot\text{1})$$

$$\tau + \beta_{t,t+n}^{\tau} \leq \tau' + \beta_{t,t+n}^{\tau'} \quad (\text{A}\cdot\text{2})$$

逆にこの条件式を満たさない  $\alpha_{t,t+n}^{\tau}, \beta_{t,t+n}^{\tau}$  については, グラフカットで解を求めることはできない.

この限界を見るために, 式 (A.1) を満たさない場合を考える. この場合, 図 A.1 (b) に示すような, 交差するエッジが発生する. 同図は,  $u_t$  が 4 ならば  $u_{t+n}$  は 7 以上,  $u_t$  が 5 ならば  $u_{t+n}$  は 6 以上という 2 条件を想定している. ここで図中のカットを考えると, これは  $u_t = 5$  の時であるから, 後者の条件を満たせばよく, そして確かに  $u_{t+n} = 6$  であるから, 条件を満たしている. しかしながら, このカットをよく見ると (このカットにとって無関係な) 前者の条件に由来したエッジを切断してしまっている. 結果的に, 条件を満たしながら, このエッジが解として求まることはない.

この制約から, グラフカットによる弾性マッチングを画像の単調連続弾性マッチング [14] に適用できない. これは以下のように説明される. まず 2 画像をそれぞれラスタスキャンし, 1 次元化したものを  $X, Y$  とする. この時, 画像の弾性マッチングで用いられる単調連続性は, 1 次元化された  $X, Y$  上ではラスタスキャンの性質により離れた地点間の制約すなわち非マルコフ的制約となる. 特に,  $X(t)$  だけでなく,  $Y(\tau)$  にも不連続的に制約が入ることになり, 結果的に図 A.1 (b) のような交差が多発したグラフになる. すなわち, 画像の最適弾性マッチングは (A.1) や (A.2) を満たさない状況であり, グラフカットでは解けないことが証明される.

実は, 画像の単調連続弾性マッチング問題は, その NP-Hard 性が証明されている [15]. もしグラフカットで任意の非マルコフ的制約を扱えるのであれば, 画像の弾性マッチングも多項式時間で解けることになり, 上記証明との矛盾が発生する. この事実も, グラフカットで扱える非マルコフ的制約には限界があることを例証している. ただし, 付録 2. のように, グラフの工夫により解ける問題も多いことを再度強調しておく.