

隣接要素関係性の自動重み付けによる 最適輸送シーケンスマッチング

堀江 光彦^{1,a)} 笠井 裕之^{1,2,b)}

概要: 本稿では映像や音声などのシーケンスデータに対する最適輸送問題を用いたマッチングの新たな手法を提案する。提案手法ではシーケンス構造の重要な特徴である隣接要素との関係性を考慮したマッチングの導出を目標とする。マッチング対象となるシーケンスデータに対し、シーケンスの各要素における隣接関係を表す変化量シーケンスを定義し、シーケンスデータ間の輸送距離損失に加え、変化量シーケンス間の輸送距離損失を新たに考慮することで、隣接要素間関係性を考慮したマッチングを実現する。さらに提案手法によるマッチングの有効性を確認するための実験として、実際にマッチングを行い類似した隣接関係を持つ要素同士がマッチングするかを確認し、またマッチングから得られる距離によるデータの分類を行い、既存手法との精度比較を行う。

1. はじめに

シーケンスマッチングとは2つのシーケンスデータの各部分同士を対応づけることであり、さらに求められた対応関係から2つのシーケンスデータ間の類似度を反映した距離を計算することである。シーケンスマッチングから求められるシーケンス間の対応関係および距離は動作分類、音声認識、映像検索、時系列データの分類など幅広い分野に利用される。

シーケンスデータ間のマッチングを導出する際に考慮すべきシーケンスデータの特徴として、まずシーケンスを構成するベクトルデータの独立要素としての類似度がある。ユークリッド距離や l_p ノルムといったベクトル間の距離を基に、類似した要素同士を対応づけていくことがシーケンスマッチングにおける基本となる。しかし、シーケンスを構成する要素は同じだがその順序が異なるといったシーケンスのペアに対し、その順序の違いを認識することができない。そのため、要素の類似度のみを考慮するだけでは不十分である。そこでマッチングにおいて考慮すべき第二の特徴として、各ベクトルデータの順序が挙げられる。順序

の近い要素同士をマッチングさせる制約を課すことで、同じデータで構成されるシーケンスであってもその順序の違いから必ずしも同じデータ同士がマッチングするとは限らなくなり、順序の違いが距離に反映されるようになる。さらに、第三に考慮すべき特徴として隣接要素間関係性が挙げられる。シーケンスデータにおいて隣接する要素は関連していることが多く、また各要素がシーケンス全体の中で持つ意味は隣接要素との関係性によって変わってくる。例えば、ある2つのシーケンスにおいて同じ値100を持つデータが同じ位置にあった場合でも、一方のシーケンスでは値が増加中のデータ、もう一方のシーケンスでは減少中のデータであればそのデータが持つ意味は全く異なることから、マッチングするべきではない。シーケンスマッチングにおいてはこれらのシーケンスデータの特徴を適切なバランスで考慮する必要がある。

シーケンスの各要素を確率分布におけるサンプルとみなし、最適輸送問題 (Optimal Transport:OT) [1][2] により導出される輸送計画をシーケンスマッチングとみなす手法として近年 OPW[3] が提案されている。本稿では OPW を拡張し、第三の特徴である要素の隣接関係を明示的に考慮してマッチングを導出する新たな手法を提案する。さらに提案手法では、マッチングの導出において、各要素の独立データとしての類似度と隣接関係の類似度をそれぞれの程度考慮するかの重みづけを、ハイパーパラメータによる調整ではなく、マルチビュークラスタリング (MVC) の技術を用いて自動で計算する手法についても提案する。

評価実験では、提案手法によるマッチングが隣接要素間

¹ 早稲田大学大学院 基幹理工学研究所 情報理工・通信専攻
Department of Computer Science and Communications Engineering, Graduate School of Fundamental Science and Engineering, Waseda University

² 早稲田大学 基幹理工学部 情報通信学科
Department of Communications and Computer Engineering, School of Fundamental Science and Engineering, Waseda University

a) kidsgoldppp@akane.waseda.jp

b) hiroyuki.kasai@waseda.jp

の関係性を考慮したものであるかについて、シーケンスデータとそのマッチングを可視化し視覚的に評価する。また、提案手法から計算される距離についても、最近傍法によるデータセットの分類を行い、既存手法および自動重み付けを行わない場合との数値比較から評価する。

2. 関連研究

シーケンスマッチングにおける代表的な手法として Dynamic Time Warping (DTW) [4] が提案されている。DTW では各要素間のユークリッド距離を基に、境界条件や単調性などいくつかの順序制約下でマッチング後の距離が最小となるようなマッチングを導出する手法であり、関連研究も数多く存在する [5] [6]。DTW では各要素間のユークリッド距離のみを用いてマッチングを算出していることから、第 1 節の例で挙げたように、独立したデータとしては類似しているが異なる隣接構造を持つような要素を誤ってマッチングしてしまうという問題がある。この問題に対処する関連手法も数多く存在し [7][8]、マッチング対象のシーケンスの各要素における微分値を用いて DTW を行う Derivative DTW [9] や、シーケンスの各要素を形状を表す特徴量に変換して DTW を行う shape DTW [10] などがある。

Order-preserving Wasserstein Distance (OPW)[3] は OT を用いてシーケンスマッチングを導出する手法である。DTW よりも柔軟な順序制約を持っており、局所的な順序の入れ替わりや開始位置の異なる周期的なパターンなどのデータに対しても正確にマッチングを行うことができる。本稿で提案する手法では OPW を拡張するため、以下にその概要を説明する。2つのシーケンス $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$, $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M]$ の OPW 距離 d_{OPW} は、各要素間の距離を表す距離行列 $\mathbf{D}_1 \in \mathbb{R}^{N \times M}$ を用いて以下で定義される。

$$d_{OPW}(\mathbf{X}, \mathbf{Y}) = \langle \mathbf{T}^*, \mathbf{D}_1 \rangle.$$

ここで、輸送行列 $\mathbf{T}^* \in \mathbb{R}^{N \times M}$ は以下に定義される。

$$\mathbf{T}^* = \arg \min_{\mathbf{T} \in \mathbf{U}(\mathbf{a}, \mathbf{b})} \langle \mathbf{T}, \mathbf{D}_1 \rangle - \lambda_1 \cdot I(\mathbf{T}) + \lambda_2 \cdot KL(\mathbf{T} || \mathbf{P}) \quad (1)$$

上式において、 $\lambda_1, \lambda_2 > 0$ は正則化パラメータ、 $I(\mathbf{T})$ は \mathbf{T} の逆差分モーメント、 KL はカルバック・ライブラー情報量を示す。ここで \mathbf{T} は OT における輸送行列であり、 (i, j) 要素 t_{ij} は $\mathbf{x}_i, \mathbf{y}_j$ 間の輸送量を表す。OT においては一般的に $\mathbf{x}_i, \mathbf{y}_j$ 間の輸送コスト d_{ij} が小さい値であるほど輸送量 t_{ij} は大きい値として導出される。シーケンスマッチングにおいては、輸送コストと輸送量をそれぞれ各要素間の類似度、マッチング確率に対応させることができ、 \mathbf{X}, \mathbf{Y} 間の OT を求めることは、類似した要素のマッチング確率が大きくなる最適なマッチングを求めることとみなすこ

とができる。そこで本稿では \mathbf{T} をマッチング行列と呼ぶ。さらに OPW では式 (1) の第 2 項、第 3 項に示される 2 つの順序の正則化項を付与することで、近い順序にある要素同士のマッチングを促している。

3. 提案手法

3.1 課題とアプローチ

OPW によるマッチングの導出では各要素間の距離行列 \mathbf{D}_1 との輸送損失を用いており、DTW によるマッチングの導出と同様に隣接要素間の関係性を明示的に考慮していないという課題がある。そこで、本稿では OPW における要素間の類似度に対する輸送損失に加え、隣接要素間の関係性の類似度に対する輸送損失を新たに加えることを考える。具体的には Derivative DTW[9] で導入された各要素の微分値の距離行列を \mathbf{D}_2 として新たに求め、 \mathbf{D}_2 との損失を新たに導入する。微分値は隣接要素との変化量を表すため、その距離行列を考えることで類似した隣接関係を持つ要素同士をマッチングさせることができる。さらに本稿では、2つの損失に対する重みについて、マルチビュークラスタリングの重み自動計算 [11][12] を適用することを試みる。重みをパラメータとしてデータセット全体に対して固定値に設定するのではなく、データごとに最適な値を用いることで、 k 近傍法によるデータの分類等における距離性能を向上させることを目的とする。

3.2 提案手法

OPW では \mathbf{X}, \mathbf{Y} 間の距離行列 $\mathbf{D}_1 \in \mathbb{R}^{N \times M}$ を用いた。 \mathbf{D}_1 の各成分は \mathbf{X}, \mathbf{Y} のそれぞれの要素間の距離を表す。

$$\mathbf{D}_1 := [d(\mathbf{x}_i, \mathbf{y}_j)]_{i,j} \quad (2)$$

次に、各要素の隣接関係の類似度を測るために、対象のシーケンス \mathbf{X}, \mathbf{Y} に対して微分値シーケンス \mathbf{X}', \mathbf{Y}' を導入する。ここでは \mathbf{X} に対する微分値シーケンス \mathbf{X}' とする。

$$\mathbf{X}' := [\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_N]$$

ここで \mathbf{x}'_i は以下に定義する [9]。

$$\mathbf{x}'_i := \frac{(\mathbf{x}_i - \mathbf{x}_{i-1}) + (\mathbf{x}_{i+1} - \mathbf{x}_{i-1})/2}{2} \quad (3)$$

ただし、最初（最後）の要素に対しては 1 つ後（前）の要素との差分とする。上記の定義により、 \mathbf{x}'_i は \mathbf{x}_i における微分値、すなわち各要素の隣接関係を表す。次に、 \mathbf{X}', \mathbf{Y}' 間の距離行列を $\mathbf{D}_2 \in \mathbb{R}^{N \times M}$ とする。

$$\mathbf{D}_2 := [d(\mathbf{x}'_i, \mathbf{y}'_j)]_{i,j} \quad (4)$$

\mathbf{D}_2 の (i, j) 成分 $d_{2(i,j)}$ は \mathbf{x}'_i と \mathbf{y}'_j の距離を表すため、 \mathbf{X}, \mathbf{Y} の各要素における隣接関係の類似度を表すことができる。 \mathbf{D}_2 に対する \mathbf{T} との損失を加え、OPW を基にさらに

隣接関係を考慮した新たなマッチング導出式を以下に定義する。

$$\min_{\mathbf{T} \in \mathbf{U}(\alpha, \beta)} \sum_{v=1}^2 \langle \mathbf{D}_v, \mathbf{T} \rangle - \lambda_1 I(\mathbf{T}) + \lambda_2 KL(\mathbf{T} || \mathbf{P}) \quad (5)$$

以下では、この2つの損失に対しさらに重みを付与することを考える。各距離行列 \mathbf{D}_v に対する重みを w_v とすると、重み付与後のマッチングおよび距離の導出式は以下のようになる。

$$d_{\text{提案}}(\mathbf{X}, \mathbf{Y}) = \sum_{v=1}^2 w_v \langle \mathbf{D}_v, \mathbf{T}^* \rangle$$

$$\mathbf{T}^* = \arg \min_{\mathbf{T} \in \mathbf{U}(\alpha, \beta)} \sum_{v=1}^2 w_v \langle \mathbf{D}_v, \mathbf{T} \rangle - \lambda_1 I(\mathbf{T}) + \lambda_2 KL(\mathbf{T} || \mathbf{P}) \quad (6)$$

ここで、マルチビュークラスタリングにおける重みの自動計算手法に則ると、式(6)における最適な重み w_v は以下に定義される [11]。

$$w_v = \frac{1}{2\sqrt{\langle \mathbf{D}_v, \mathbf{T} \rangle}} \quad (7)$$

この定義では w_v は最小化問題(6)の変数 \mathbf{T} についての変数でもある。そこで、 \mathbf{T} と w_v について、一方を固定しもう一方を式(6)または(7)から更新する交互最適化を行う。この交互最適化により \mathbf{T} は収束する。また、そのときの w_v の値は定義から $\langle \mathbf{D}_v, \mathbf{T} \rangle$ の値が小さい \mathbf{D}_v に対して大きい値を取る。輸送損失が小さいような \mathbf{D}_v はそのデータのマッチングにおいて考慮されるべき距離行列であり、そのような \mathbf{D}_v に対して w_v は大きい値を取るため、この重み付けは考慮すべき距離行列に対してより大きな重みを付与する最適な重み付けといえる [11]。よってこの一連の交互最適化は、重要な距離行列に大きい重みを付与しつつ、式(6)を解くことができている。

Algorithm 1 提案手法

Require: two sequences \mathbf{X}, \mathbf{Y} , parameters $\lambda_1, \lambda_2, \sigma$

Ensure: $d_{\text{Prop}}(\mathbf{X}, \mathbf{Y})$

- 1: Calculate derivative sequence \mathbf{X}' and \mathbf{Y}' by Eq.(3)
- 2: Calculate \mathbf{D}_1 and \mathbf{D}_2 by Eq.(2) and Eq.(4)
- 3: Initialize $w_1 = 1, w_2 = 1$
- 4: **while** w_v has not converged or the iteration has not reached to the maximum number **do**
- 5: $\mathbf{D} = w_1 \cdot \mathbf{D}_1 + w_2 \cdot \mathbf{D}_2$
- 6: Calculate OPW matching matrix \mathbf{T} with \mathbf{D} by Eq.(1)
- 7: Update $w_v \leftarrow 1/2\sqrt{\langle \mathbf{D}_v, \mathbf{T} \rangle} \quad v = 1, 2$
- 8: **end while**
- 9: $d_{\text{提案}}(\mathbf{X}, \mathbf{Y}) = w_1 \langle \mathbf{D}_1, \mathbf{T} \rangle + w_2 \langle \mathbf{D}_2, \mathbf{T} \rangle$

4. 実験

本節では、提案手法の評価実験を行なう。評価は、シーケンスデータおよびそのマッチングの可視化による視覚的評価と、マッチングから計算される距離を用いた最近傍法による分類実験の精度比較による評価の2つを行う。実験においては各距離行列は各要素をその最大値で割ることにより0~1の範囲に正規化している。実験に使用したデータセットは、Spoken Arabic Digit (SAD)[13]^{*1}, High-Quality Recordings of Australian Sign Language Signs (HAS)[13]^{*2}, UCR Time Series Classification Archive[14],^{*3} より FacesAll および SwedishLeaf である。

表 1 データセットの概要

データセット	クラス数	test	train	次元	長さ
SAD	10	2200	6600	13	不定
HAS	95	1330	1235	22	不定
FaceAll	7	1690	560	1	131
SwedishLeaf	15	625	500	1	128

4.1 マッチングの視覚的評価

はじめに、マッチングの可視化を通じて提案手法が OPW と比較してよりデータの隣接関係を考慮したマッチングであるかを視覚的に評価する。実験に際して、各要素の最も強いマッチングのみを可視化している。マッチングを行うデータとして SwedishLeaf データセットの train-1 と train-4 を採用した。まず実験に使用したデータを図1に示す。2つのシーケンスは同様の概形(凹凸の数や増加・減少傾向など)を持つが、その値や位置は異なる。このデータに対する OPW によるマッチングを図2に示す。枠で囲われている部分では、一方は上昇、もう一方は下降中の要素であるにも関わらず、値や順序の近さにつられてマッチングが生じてしまっている。これらに対して、図3に示す提案手法のマッチングでは隣接関係の考慮により値の増加中同士、減少中同士のマッチングを行うことができている。

4.2 最近傍法によるデータの分類実験

次に、距離の性能について評価する。マッチングから計算される距離を用いて最近傍法によるデータの分類を行い、その精度を DTW, OPW と比較する。分類精度を示す指標として Accuracy および mean Average Precision (mAP) を採用し、OPW と提案手法におけるパラメータは $\lambda_1 = 0.01, \lambda_2 = 0.01, \delta = 1$ とした。また、重みの自動計算についての評価のため、 $w_1 = 1, w_2 = 1$ と固定した場合の提案手法による分類も行った。実験結果を表2に示す。提案(固定)が重みを固定した提案手法を示す。FaceAll,

*1 <https://archive.ics.uci.edu/ml/datasets/Spoken+Arabic+Digit>

*2 [https://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+signs+\(High+Quality\)](https://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+signs+(High+Quality))

*3 https://www.cs.ucr.edu/~eamonn/time_series_data_2018/

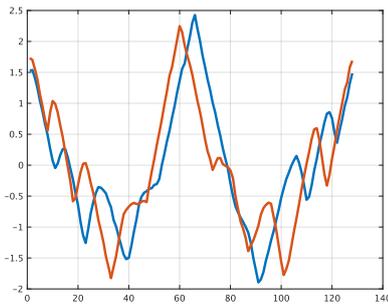


図 1 マッチング対象のシーケンスデータ

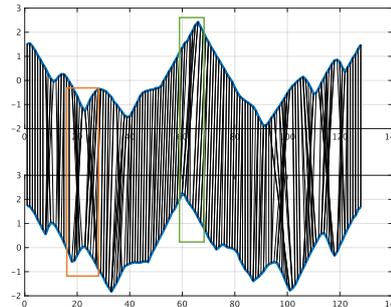


図 2 OPW

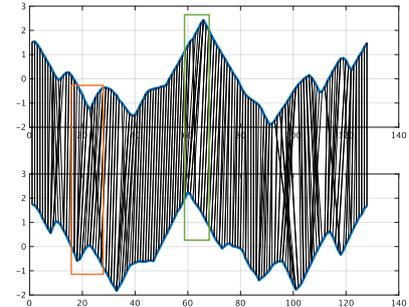


図 3 提案手法 ($w_1 = 1.07$, $w_2 = 2.69$)

表 2 最近傍法による分類精度

データセット		DTW	OPW	提案	提案 (固定)
SAD	Acc	0.9636	0.9432	0.9232	0.9327
	mAP	0.5658	0.5415	0.5228	0.5570
HAS	Acc	0.8135	0.7992	0.7820	0.7714
	mAP	0.4461	0.4257	0.4475	0.4386
FaceAll	Acc	0.8077	0.8183	0.9249	0.8959
	mAP	0.5483	0.5484	0.5999	0.5825
Swedish	Acc	0.7920	0.8800	0.9424	0.9456
Leaf	mAP	0.4928	0.5183	0.7201	0.7121

SwedishLeaf, および HAS の mAP において提案手法は既存手法よりも高い数値を記録している。また、重みを固定した場合との比較では、SAD を除く 3 つのデータセットにおいて、提案手法がより優れた性能を示している。

5. まとめ

本稿では最適輸送問題を用いたシーケンスマッチングにおける新手法の提案を行なった。提案手法ではシーケンスの各要素における隣接関係を考慮したマッチングを導出することを目的とし、シーケンスの微分値の距離行列との輸送距離損失を新たに最適化目的関数に追加した。さらに、マルチビュークラスタリングにおける重みの自動計算を用いて、シーケンスとその微分値の 2 つの輸送距離損失に対し、データごとに自動重み付けを行なった。評価実験では、提案手法が隣接関係を考慮したマッチングを実現し、既存手法と同等以上のデータの分類性能を記録することを確認した。また重みを固定した場合との精度比較から、重みの自動計算の有効性も示された。今後は、現状ではハイパーパラメータ付きの正則化項である順序に対する制約も新たに距離行列の形で表現し輸送損失とする手法の検討を行う。

参考文献

[1] G. Peyré and M. Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*,

11(5-6):355–607, 2019.
 [2] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.
 [3] B. Su and G. Hua. Order-preserving optimal transport for distances between sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):2961–2974, 2019.
 [4] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, 26(1):43–49, 1978.
 [5] Young-Seon Jeong, Myong K Jeong, and Olufemi A Omitaomu. Weighted dynamic time warping for time series classification. *Pattern recognition*, 44(9):2231–2240, 2011.
 [6] Dave Deriso and Stephen P. Boyd. A general optimization framework for dynamic time warping. *CoRR*, abs/1905.12893, 2019.
 [7] Jidong Yuan, Qianhong Lin, Wei Zhang, and Zhihai Wang. Locally slope-based dynamic time warping for time series classification. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 1713–1722, 2019.
 [8] Jae Yeol Hong, Seung Hwan Park, and Jun-Geol Baek. Ssdwtw: Shape segment dynamic time warping. *Expert Systems with Applications*, 150:113291, 2020.
 [9] Eamonn J Keogh and Michael J Pazzani. Derivative dynamic time warping. In *Proceedings of the 2001 SIAM international conference on data mining*, pages 1–11. SIAM, 2001.
 [10] Jiaping Zhao and Laurent Itti. shapedtw: Shape dynamic time warping. *Pattern Recognition*, 74:171–184, 2018.
 [11] Feiping Nie, Jing Li, Xuelong Li, et al. Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification. In *IJCAI*, pages 1881–1887, 2016.
 [12] Feiping Nie, Jing Li, and Xuelong Li. Self-weighted multiview clustering with multiple graphs. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 2564–2570, 2017.
 [13] Arthur Asuncion and David Newman. Uci machine learning repository, 2007.
 [14] Hoang Anh Dau, Anthony Bagnall, Kaveh Kamgar, Chin-Chia Michael Yeh, Yan Zhu, Shaghayegh Gharghabi, Chotirat Annh Ratanamahatana, and Eamonn Keogh. The ucr time series archive. *IEEE/CAA Journal of Automatica Sinica*, 6(6):1293–1305, 2019.