

極値分割と Dynamic Time Warping による データ類似度評価手法

森 郁海^{1,2,a)} 中村 嘉隆² 稲村 浩²

受付日 2020年4月22日, 採録日 2020年11月5日

概要: データの発生源と消費先が近傍である場合, エッジ機器上で AI を利用するユースケースが考えられる. しかしながら, エッジ機器上で AI のための学習も行う際は, クラウドコンピューティングと異なりデータの収集範囲が限られるため, 訓練データ不足が深刻な問題となる. 本稿では, このような訓練データ不足に対し, 収集済みの類似データを学習に流用するためのデータ類似度評価手法を提案する. 提案する類似度評価手法は, データの類似度を, データの特徴量に基づき求めるデータ分布の形状に着目し計算する. データ分布の類似度計算において, データ分布の形状が平行移動 (シフト) したり, 伸縮したり, 相似形であったりする場合でも類似度が高くなるように, 類似度評価区間を極値で分割し, 区間ごとに Dynamic Time Warping 距離を求め, 各距離を結合することで類似度を得る. 画像認識でのユースケースを想定した評価環境下では, データ分布の形状にシフトや伸縮, 相似形が存在するデータ分布の比較においても, 類似画像を抽出可能であり, さらに, ユーザによる調整が必要なパラメータが不要であることを確認した.

キーワード: 類似度, 区間分割, Dynamic Time Warping, 極値

A Method of Data Similarity Evaluation Using Dynamic Time Warping with Partition by Peaks

IKUMI MORI^{1,2,a)} YOSHITAKA NAKAMURA² HIROSHI INAMURA²

Received: April 22, 2020, Accepted: November 5, 2020

Abstract: There are many use cases for using AI on edge devices in case of the data source and the consumer are very close. However, lack of training data is a serious problem because the range of data collection of edge devices is limited. In this paper, we propose a method of data similarity evaluation using Dynamic Time Warping (DTW) with partition by Peaks for diverting training data that had already collected at other domains. The proposed method partitions the data similarity evaluation interval by peaks of feature data distribution and calculates DTW distance for each partition to evaluate similarity in order to cope with peak shifts and similar in peak shape of feature data distribution. And then the similarity between data distributions is calculated by combining each distance with pre-defined weights. In evaluation assuming the use case of image recognition, we confirmed that the proposed method can extract similar images from source image set even if peak shifts and similar in peak shape exist when comparing feature data. In addition, we also confirmed that the proposed method does not require any parameters that have to be adjusted by the user under this evaluation.

Keywords: similarity, partition, Dynamic Time Warping, peaks

1. はじめに

5G, エッジ処理の実現により, AI エッジコンピューティングの研究が活発に行われている [1]. AI エッジコンピューティングのプラットフォームとして Azure IoT Edge [2] や AWS IoT Greengrass [3] などがあり, エッジ機器上で AI

¹ 三菱電機株式会社情報技術総合研究所
Information Technology R&D Center, Mitsubishi Electric Corporation, Kamakura, Kanagawa 247-8501, Japan

² 公立はこだて未来大学
Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan

^{a)} Mori.Ikumi@ab.MitsubishiElectric.co.jp

サービスを容易に実現可能な環境が整いつつある。AI エッジコンピューティングにより、従来クラウドで集中的に運用していた AI サービスをエッジ上に分散配置することができ、リアルタイム性も格段に向上する。

一方、個人の健康管理や自動車の運転支援などのユースケースでは、IoT 機器ごとの個別最適化が要求される。この場合、IoT 機器ごとに異なる AI 学習モデルを作成する必要があるが、クラウドへデータを集約する形態と比べデータの収集範囲が限られるため訓練データが不足し、AI 学習モデルに十分な精度が得られないことがある。

FedHealth [4] は、ウェアラブルヘルスケアデバイスを対象として、プライバシーの関係上ユーザデータがクラウド上で共有できない場合に、ユーザで共通に使用する AI 学習モデルと、個別のユーザ環境で取得したデータから作成した AI 学習モデルを基に、新たな AI 学習モデルを構成するフレームワークを提案している。このフレームワークは、ユーザデータから AI 学習モデルを作成するため、ある程度の訓練データ量を要求する。たとえば、身体的特徴や日常の活動パターンが似た人物 A、B の宅内カメラ画像で心拍数や行動認識による健康管理を想定する場合、A の環境下では訓練に必要な画像データが十分あるが、B の環境下では十分存在しない状況では、B の AI 学習モデルを作成することができない。このようなユースケースでは本稿で提案する、B の環境下で取得した少量のデータから A の訓練データとの類似性を判定し、A の許可を得たうえで A の訓練データを B の環境下での学習データに転用する方法の有効性が期待できる。

この方法は、転移学習（データ転移）[5]、[6] と呼ばれ、転移学習が成功するためには、他の環境で使用した訓練データ（以降、「転移元データ」と表記）と転移先の訓練データ（以降、「転移先データ」と表記）の類似度が高いことが1つの条件である。この類似度を高精度に算出することができれば、識別クラスのサンプルを偏りなく含む、良い訓練データセットを構成することができる。

仮に、識別クラスのサンプル数に偏りがある不均衡データセットで分類器を作成した場合、次のような現象が生じ、クラス識別性能が低下する。たとえば、訓練データのクラス分布が「人間」:「トラ」= 99 : 1 の場合、つねに「人間」と推論する分類器が作られると、分類器の「人間」に対する正解率は 99% となる。一方、「トラ」に対する正解率は 0% となる。推論時の入力データのクラス分布が、「人間」:「トラ」= 5 : 5 であった場合、この分類器は「人間」を過検出してしまう。このような状況が発生すると、「人間」と「トラ」で異なる処理を提供する AI サービスの場合、「トラ」に対しても「人間」の処理が割り当てられ、AI サービスの品質が低下する。

転移先データの単一サンプルと、転移元データの各サンプルとの類似度を高精度に算出することで、前述の例にお

いては、転移先データのサンプルに「トラ」が1つでもあれば、転移元データのサンプル集合から類似度が高いものを機械的に抽出し、必要に応じてクレンジングすることで、高品質な「トラ」の集合を構築することができる。このように、精度の高い類似度の算出は、不均衡データセットの生成を抑制し、AI サービスの品質低下を防ぐために重要な役割を果たす。

筆者らはこれまでに、転移元データ集合と転移先データ集合の統計的性質（データ分布）が類似しているかを調べる手法として、ピアソン相関係数や統計的仮説検定、コサイン類似度を用いる方法を検討してきた [7]。インターネット上に公開されている行動認識向けの訓練データを、アノテーション（行動ラベル）ごとに分割したものを転移元データ集合とし、それらからサンプリングした少数のデータを転移先データ集合として、前述の手法を使用して類似性の高かった転移元および転移先データ集合の組の各行動ラベルが一致する精度を評価した。その結果、精度が最も高かったのはピアソン相関係数を用いるものであった。ピアソン相関係数は、転移元データのデータ分布と転移先データのデータ分布の増減が同じような傾向を示す場合の類似度の算出に有効であるが、次に示すようなデータ分布の形状がシフトしたり、伸縮したり、相似形であったりする場合も類似度を高く算出すべきである。

たとえば、環境光などの変化に起因する照明変動によって、同一被写体の画像においてもヒストグラムなどの特徴量が増える [8]。昼の画像では、明るい画素値の出現頻度が増え、夜の画像では、暗い画素値の出現頻度が増えるのは、太陽光による照明変動が影響する。しかし、被写体が持つ本来のヒストグラム形状は、照明変動が生じてても、ヒストグラム形状が明るい方が暗い方にシフトするだけで、形状の特徴は保存されているはずである。同様に、被写体が画面に占める割合が変化する場合、背景除去などを適切に行うことで、被写体が持つ本来のヒストグラム形状は、相似形として取得できる*1。また、被写体のコントラストが低い画像は、ヒストグラムの分布が狭くなり、コントラストが高い画像は、ヒストグラムの分布が広がるため、ヒストグラム形状が伸縮することがある [9]。このように、ヒストグラム形状がシフトしたり、伸縮したり、相似形であったりするときは、転移学習が有効な場合がある。

データ分布の形状変化のうち形状のシフトについては、2つのデータを少しずつずらしながら類似度を算出するスライディングウィンドウ法を用いてもよいが、類似度算出に大きく影響するスライド幅をユーザが調整する必要がある。

本稿では、転移元データと転移先データの類似度評価を転移元データのデータ分布と転移先データのデータ分布の類似度に基づいて行う場合において、データ分布の形状が

*1 相似な平面図形の面積比は、相似比の二乗となるため。

シフトしたり、伸縮したり、相似形であったりしても類似度を高く算出できる効果が見込める Dynamic Time Warping (DTW) 距離 [10] を使用した類似度評価手法を提案する。

具体的には、転移先データから抽出した特徴量に基づきデータ分布を作成し、そのデータ分布を極値で区間分割する。この分割区間ごとに転移元データの特徴量を基に作成したデータ分布との DTW 距離を算出後、それぞれの分割区間での DTW 距離を線形結合することで類似度評価を行う。区間分割を行うことで、転移元データのデータ分布と転移先データのデータ分布の全体に対して DTW を適用したときに生じる、過度なフィッティングを防止する。

なお、筆者らの研究 [7] では訓練データのデータ分布に着目したが、本稿では画像認識での転移学習への応用を考え、画像の特徴量のデータ分布の類似度を判定する。双方ともデータ分布の類似性を評価するため、問題の本質は変わらないことに注意する。

有効性の評価では、画像認識での転移学習のユースケースを想定し、類似の特徴量を持つ画像に対して、提案手法が高い類似度を持つものとして抽出できるかを検証する。さらに、この評価環境下において、スライディングウィンドウ法のスライド幅のような、ユーザによる調整が必須のパラメータが不要であることを確認する。

2. 関連研究

データ分布の類似度を測る既存手法には、筆者らの研究 [7] で検討対象にした手法が存在する。これらの手法には、1章で述べたように、データ分布の形状がシフトしたり、伸縮したり、相似形であったりする場合に対応できないという課題がある。この課題を解決するために、本研究では、データ分布の形状変化が生じても類似度を高く算出できる効果が見込める、時系列データの類似度を測る手法に着目してデータ分布の類似度を求めることを考える。この章では、時系列データの類似度を測る既存手法を説明する。

2.1 時系列データの類似度計算手法

Aghabozorgi らのサーベイ [11] によれば、類似度計算手法は *Shape-based*, *Compression-based*, *Feature-based*, *Model-based* の 4 つの方式がある。

Shape-based は、2 組の時系列データの形状を、時間軸に対する非線形な伸縮によって可能な限り一致させようとするもので、ユークリッド距離、DTW [10]、LCSS [12]、MVM [13] などがある。これらは、短い時系列データに適用することが多い。この方式は、時間軸の影響を無視するため、データ分布の形状のシフトに対応できるが、シフト量が大きすぎると画像としての類似性が低下する恐れがある。また、データ分布の形状のうち高さを厳密に判定するため、幅広く類似画像を検索するには極値の高さの変化に

対し許容を持たせる必要がある。

Compression-based は、時系列データを何らかの形で圧縮 (変換) し、圧縮後の情報から類似性を求めるもので、CDM [14]、自己相関、カルバック・ライブラ情報量 [15]、ピアソン相関係数などがある。これらは、時系列の長さによらず使用できる。この方式は、圧縮方法に精度が左右されるため、扱うデータによって適切な圧縮方法をユーザが選択する必要がある。

Feature-based は、時系列データを低次元の特徴量に変換し、特徴量から類似性を求めるもので、統計学 (統計的仮説検定など)、係数 (Jaccard 係数 [16]、Dice 係数 [17], [18]、Simpson 係数 [19] など)、コサイン類似度などがある。これらは、長い時系列に適用することが多い。この方式も *Compression-based* と同様に、適切な次元削減方法を選択する必要がある。

Model-based は、時系列データの予測モデルがどの程度類似しているかを求めるもので、HMM [20]、ARIMA [21] などのモデルを利用する。これらは、長い時系列に適用することが多い。この方式も *Compression-based*, *Feature-based* と同様に、モデルの選定が精度に大きく影響する。

本稿では、ユーザによる調整が不要な *Shape-based* 方式を用いる。*Shape-based* 方式でよく使用されるのは、ユークリッド距離と DTW である [11]。以降で、ユークリッド距離、DTW について詳細を述べる。

2.1.1 ユークリッド距離

ユークリッド距離は、2 つの点 (ベクトル) をつなぐ線分であり、ベクトル $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$ のユークリッド距離 $d(\vec{x}, \vec{y})$ は式 (1) のように定義される。

$$d(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

ユークリッド距離は、データの差の二乗和の平方根であり、ベクトルの成分間の重みを考慮しない。本稿では、特徴量のデータ分布の比較におけるデータ分布の形状のシフトに対しては類似度への影響を少なくし、データ分布の形状が相似形である場合は、類似度を高く算出する性質を持つことが望ましい。しかしながら、ユークリッド距離は、座標系の各軸の値のずれの程度が異なる意味を持つ状況に対応することができない。

2.1.2 Dynamic Time Warping (DTW)

DTW は、2 つの時系列データの各点の距離を総当たりで比較し、2 つの時系列データの距離が最短となるパスを見つける手法である。ここでの距離指標に特に制限はないが、誤差の絶対値などを使用することが多い。基本的な概念は 1959 年に提案され [10]、その後、音声認識、筆跡照合、手話認識、データマイニング、時系列データクラスタリングなど幅広い分野に応用されている [22]。

2 つの時系列データから類似波形を探索するため、デー

タ分布の形状が異なる場合は類似度を低く算出し、波形が時間軸に対し平行移動したようなデータ分布の形状がシフトするパターンや形状が伸縮するパターンの類似度は高く算出する。このことから、2.1.1 項で述べた本稿で求める性質に近い性質を持つ。

本稿では、データ分布の形状のシフトや伸縮をどれだけ許容するかを制御できるように、転移先の極値で区間を分割し、分割区間内で DTW 距離を算出する。

3. 提案手法

転移元データと転移先データそれぞれの特徴量に基づいて求めたデータ分布の形状のシフトや伸縮、相似形に対応するために、まず、転移先データの特徴量に基づいて求めたデータ分布の極値で評価区間を分割し、分割区間ごとに転移元データの特徴量に基づいて求めたデータ分布との DTW 距離を算出後、それぞれの結果を結合し、最終的な類似度を得る方法を提案する。

提案手法では、次の手順で類似度評価を行う。

- (1) 特徴抽出とデータ分布の導出
- (2) 区間分割
- (3) 分割区間の類似度評価と結合

3.1 特徴抽出とデータ分布の導出

DTW 距離を類似度指標とし、転移元データと転移先データの類似度を求めるために、転移元データと転移先データそれぞれの特徴量ベクトルを求め、特徴量ベクトルから 2 次元データ分布を計算する (図 1)。このデータ分布どうしを DTW で比較することで、類似度を高く算出すべきデータ分布の形状変化に対応する。本稿では、プライバシーの関係上、人物ではなく動植物の画像を評価に用いる。評価に使用するデータは、柄 (模様) に特徴があるため、特徴量をグレースケール化した画素とし、画素値の出現頻度 (ヒストグラム) を 2 次元データ分布とする。正確には、グレースケール化した画像の画像値からカーネル密度推定によって画素値 x を確率変数とする確率質量関数 $p(x)$ を求め、 $(x, p(x))$ を 2 次元データ分布とする。ヒストグラムではなく確率質量関数を使用することで、画像の母集団のデータを推定するため、画像の標本に含まれるノイズの影響を抑えることができる。

ただし、 $(x, p(x))$ は転移元データと転移先データの特徴を十分反映しており、波形状に近似できればどのようなものでもかまわない。たとえば、 $(x, p(x))$ に Bag-of-Features (BoF) [23] を用いる方法がある。BoF は、訓練データから抽出した特徴ベクトル群をクラスタリングし、そのクラス ID と出現頻度を生成する手法である。よって、 x にクラス ID、 $p(x)$ にそのクラスの出現頻度を設定すればよい。

特徴抽出は、画像であれば HOG [24] や SIFT [25]、音声であれば MFCC [26] などを用いる。BoF は特徴が出現す

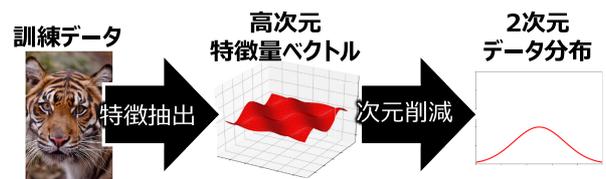


図 1 特徴抽出とデータ分布の導出の流れ

Fig. 1 Feature extraction and dimension reduction flow.

る位置情報が損失するが、特徴次元が下がり推論時の計算量を抑えられるため、計算資源の限られたエッジ環境でも利用可能と考えられる。

このように、BoF における特徴抽出方法を変更することで、動植物などの画像分類の他、スマートスピーカで使用される音声認識などのユースケースにも応用可能である。

3.2 区間分割

3.1 節で求めた転移元と転移先画像の $p(x)$ について、データ分布の形状のシフトや伸縮、相似形に対応するために、転移先の $p(x)$ の極小値で分割する。極小値から次の極小値の間には必ず 1 つの極大値が含まれる。極大値は、 $p(x)$ の形状の特徴の 1 つととらえることができ、この特徴的な形状が分割区間内、あるいはその近傍に含まれていれば、部分的な類似性があると判断する。

3.1 節で述べたとおり、画像をグレースケールに変換するため、画素の表現空間が 8 ビットである場合、画素値は 0 から 255 の間で変化する。よって本稿では、 x の区間を $[0, 255]$ とする。ただし、本項では区間を $[a, b] := \{x \in Z \mid a \leq x \leq b\}$ と定義する。転移先の $p(x)$ (以降、 $p_t(x)$ と表記) の極小値の集合が、 $\{p_t(x_1), p_t(x_2), \dots, p_t(x_n)\}$ (ただし、 $x_i < x_{i+1}$) であるとき、分割区間を式 (2) のように定義する。ただし、 $b_0 = 0$ 、 $b_i = x_i$ ($i = 1, \dots, n$)、 $b_{n+1} = 255$ である。

$$[b_0, b_1], [b_1, b_2], \dots, [b_n, b_{n+1}] \quad (2)$$

ここで、分割区間に転移元の $p(x)$ (以降、 $p_s(x)$ と表記) の極大値が入りやすくなるよう、各区間にマージン m_i を導入することを考える (式 (3))。ただし、 m_i は必須パラメータではないため、必要に応じユーザが設定する。

$$[b_0, b_1 + m_0], [b_1 - m_1, b_2 + m_1], \dots, [b_n - m_n, b_{n+1}] \quad (3)$$

本稿では、 m_i を式 (4) のように定め、パラメータ $overlap$ を使用して比率で表現する。

$$\left\{ m_i = overlap \cdot (b_{i+1} - b_i) \mid \begin{array}{l} i = 0, \dots, n \\ 0 \leq overlap \leq 1 \end{array} \right\} \quad (4)$$

そして、 m_i の導入によって x の区間が $[0, 255]$ を超えないよう、分割区間 Iv_i を式 (5) のように修正する (図 2)。

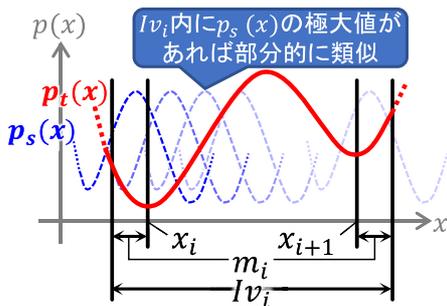


図 2 区間分割の考え方

Fig. 2 Conceptual diagram of the partition I_{v_i} .

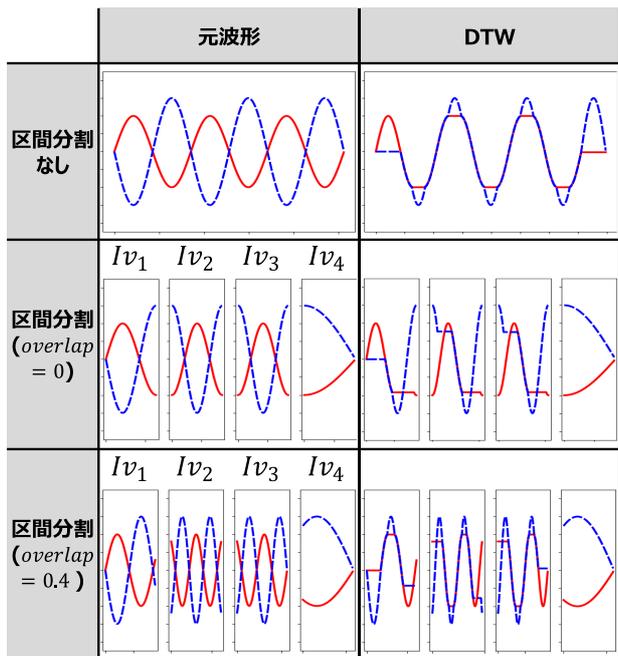


図 3 提案手法の区間分割と DTW 適用イメージ

Fig. 3 An example of applying the partitioning and DTW.

$$I_{v_i} = \begin{cases} [b_i, \min(b_{i+1} + m_i, 255)], & i = 0 \\ [\max(0, b_i - m_i), \min(b_{i+1} + m_i, 255)], & i = 1, \dots, n - 1 \\ [\max(0, b_i - m_i), b_{i+1}], & i = n \end{cases} \quad (5)$$

本稿では、 m_i を分割区間長に比例するように設定したが、原理的には固定値でもよい。 m_i を分割区間長に比例させる利点は、 $overlap$ パラメータにより分割区間長が最大でも分割区間の 3 倍となるため、過剰な形状一致を抑制できることである。単純には、スライディングウィンドウ法のように $p_s(x)$ の元波形全体をずらしながら DTW 距離を算出する方法もあるが、波形をずらす際に DTW 距離算出の対象とならない部分が発生し、類似度算出精度が低下する恐れがある。加えて、ユーザによるスライド幅の調整が必要になる。

図 3 は、提案手法の区間分割と DTW の適用イメージである。赤実線が $p_t(x)$ の波形で、青破線が $p_s(x)$ の波形である。区間分割を行わない通常の DTW では、元波形の

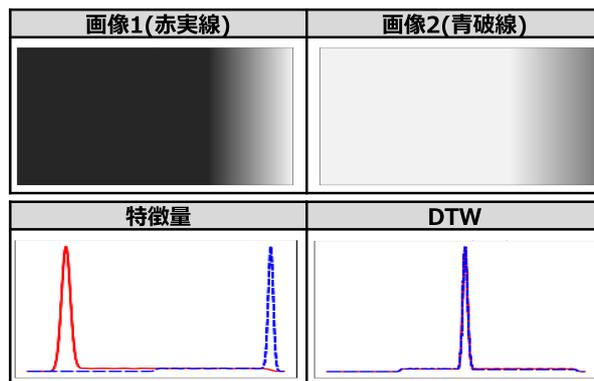


図 4 不適切な形状一致例

Fig. 4 An example of inappropriate shape matching.

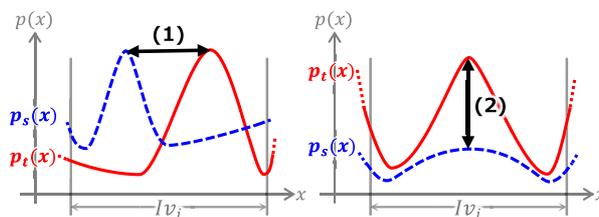


図 5 極大値の伸縮の許容範囲のイメージ

Fig. 5 An example of the tolerance range of the local maximum.

位相のずれが大きい場合でも形状を一致させている (図 3 上段右)。しかし、位相のずれが大きすぎる場合、転移元と転移先の画像の類似性が著しく低下する恐れがある。図 4 は、不適切な形状一致の例である。画像 1 と画像 2 は、見た目が異なり、特徴量のデータ分布の極大値のシフト量も大きいですが、DTW では、このような場合でも波形の大部分が一致してしまう。区間分割を行うと、 $p_t(x)$ の波形の極小値で分割され、その区間内で $p_s(x)$ の波形との形状一致判定が行われる (図 3 中段右)。 $overlap = 0$ の場合の I_{v_1} 、 I_{v_2} に着目すると、 $p_s(x)$ の波形の極大値が部分的にしか含まれないため、半分程度の一致にとどまる。一方、 $overlap = 0.4$ では、 I_{v_1} 、 I_{v_2} 内に $p_s(x)$ の波形の極大値を含むようになるため、大部分が一致するようになる (図 3 下段右)。このように、 m_i の値を変化させることで、波形の形状一致の度合いを制御することができるが、 m_i の導入によって必要以上に形状一致が許容される恐れがあり、 m_i の要否も含め調整の余地がある。

3.3 分割区間の類似度評価と結合

特徴として重視する極大値を含む区間の DTW 距離を類似度に強く反映させるために、各分割区間の DTW 距離に重み付けをして結合する。

類似度評価の前に $p_t(x)$ 、 $p_s(x)$ の値を $[0, 1]$ に正規化する。DTW は、 x 軸方向のシフトや伸縮は許容するが (図 5 (1))、 $p(x)$ 軸方向の伸縮 (極値の高さ) は許容しない。一方、転移学習では、転移元データを変換してから転

移することがある [5]. たとえば, $p_t(x)$, $p_s(x)$ の全体形状について, スケールのみが異なる場合, $p_s(x)$ を定数倍する変換を介することで転移が可能となる. そこで, $p_t(x)$, $p_s(x)$ をそれぞれ最大値と最小値を基準にスケール変換し, $p(x)$ 軸方向の伸縮 (図 5(2)) に定数倍の許容を持たせる. 正規化後の $p(x_j)$ の値 $p(x_j)_{norm}$ を, 式 (6) に示す. ここで, $p(x)_{min}$ は $p(x)$ の最小値, $p(x)_{max}$ は $p(x)$ の最大値である.

$$p(x_j)_{norm} = \frac{p(x_j) - p(x)_{min}}{p(x)_{max} - p(x)_{min}} \quad (6)$$

正規化済の $p_t(x)$ を $p_t(x)_{norm}$, 転移元データ集合 \mathbb{S} の元 $s_k \in \mathbb{S}$ の正規化済の $p_{s_k}(x)$ を $p_{s_k}(x)_{norm}$ とすると, 重みパラメータ $w_i \in [0, 1]$ を用いて, 類似度 $s(t, s_k)$ を式 (7) のように定義する. ただし, $DTW(t, s)_{Iv_i}$ は分割区間 Iv_i における時系列データ t , s の DTW 距離である.

$$s(t, s_k) = \sum_{i=0}^n w_i \cdot DTW(p_t(x)_{norm}, p_{s_k}(x)_{norm})_{Iv_i} \quad (7)$$

4. 評価

4.1 評価方法

提案手法の有効性を評価するために, AI がよく用いられる画像認識分野での転移学習のユースケースを想定して, 類似する特徴量を持つ画像に対し, 提案手法が高い類似度を持つものとして順位付けできるかを確認する.

また, 提案手法で用いている区間分割には調整の余地があり, いったんマージン m_i (*overlap* パラメータ) を導入するが, m_i の変化に対して類似度が大きく変化せず, *overlap* の調整は実用上不要であることを評価により示す.

4.1.1 前提条件

評価用の画像は, 3.1 節で述べたようにプライバシーの関係上, 人物ではなく動植物の画像を用いる (図 6). また, これらの画像は, 光源やコントラスト, 画像に対する被写体の占有率が異なることに加え, 被写体が似ているが同一の動植物でないケースを想定する. たとえば, 個人の健康管理での転移学習では, 異なる住居に同一人物がいる状況が少なく, 環境ノイズが存在する条件下で身体的特徴が似た人物を探すのが現実的である. さらに, 立つ, 座る, 転ぶなどの人物の動作認識の学習では, 完全に同一被写体でなくても, 転移学習が有効である可能性が高い.

本評価では, 転移先画像中の動物と似ている動物を「分類学的に似ている生物」と定義する. 転移元データセットには, 転移先の動物画像と似ている動物のほか, 分類系統上の距離が離れた動物や植物画像をいくつか入れる. この転移元データセットから転移先画像の動物と似ている動物を含む画像の類似度を高く算出できるかを評価する.

本稿で想定する画像認識分野での転移学習では, 類似画像の検索においてリアルタイム性を求めない夜間バッチな

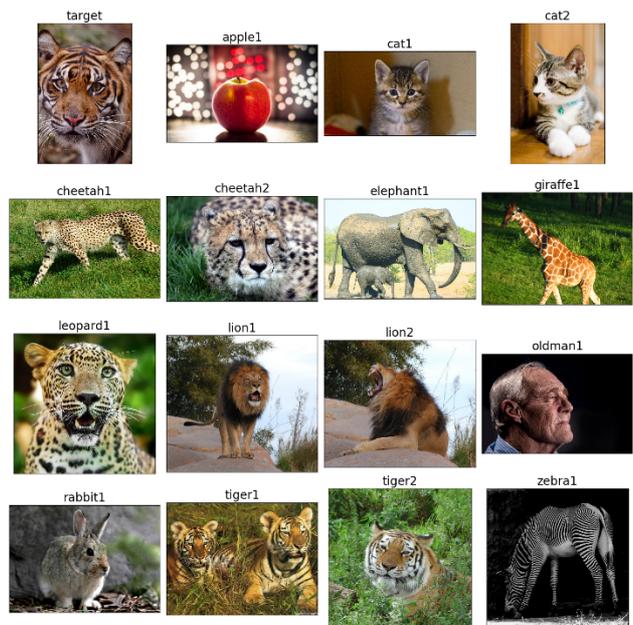


図 6 評価用画像一覧

Fig. 6 List of images for evaluation.

どでの実行を想定する. そのため, 検索時間は評価対象外とする. 加えて, 転移学習に必要な画像を収集する範囲をある程度広くする必要があるので, 転移元の画像を保存・検索する環境は, ある程度の計算能力のある PC 相当のエッジ機器, または, クラウドとする.

評価対象の画像から得られる特徴量は, 画像をグレースケール化後, カーネル密度推定によって求めた画素値の確率質量関数を $[0, 1]$ に正規化したもの ($p(x)_{norm}$) とする. $p(x)_{norm}$ は, 画像の特徴をよく抽出できているものとする. すなわち, $p_t(x)_{norm}$ と $p_s(x)_{norm}$ が類似の形状を持つ場合, 元の画像どうしの類似度も高くなる.

また, 画像は背景などの識別対象以外の領域 (ノイズ) を含むが, 本評価ではノイズ除去は実施しない. 加えて, 画像解像度や画像サイズの変更も行わない.

4.1.2 評価項目

転移元の画像集合を, 転移先の画像 (*target* 画像) との類似度が大きい順に順位付けできているかを確認する. 本評価では, 動植物の画像を用いて, 以下を評価する.

- (1) *target* 画像と似た画像ほど順位が上位であるか.
- (2) *overlap* パラメータによって, 類似度が著しく変動しないか.

提案手法と比較する方式は, 筆者らのピアソン相関係数を用いる方式 [7] (以降, 「ピアソン相関係数」と表記), 区間分割を行わず DTW を適用する方式 (以降, 「DTW」と表記) の 2 方式とする.

評価 (1) における類似度評価手法ごとの評価関数と順位付け規則を, 表 1 に示す. 各手法は, $p_t(x)_{norm}$ と $p_s(x)_{norm}$ を入力として評価関数を用いて類似度を出力し, 順位付け規則に従い順位を決定する.

表 1 評価関数と順位付け規則

Table 1 Evaluation function and rank-order rule.

類似度評価手法	評価関数	順位付け規則
提案手法	$s(t, s_k)$ (変域:= $[0, \infty)$)	昇順
ピアソン相関係数	ピアソン相関係数 (変域:= $[-1, 1]$)	降順
DTW	DTW 距離 (変域:= $[0, \infty)$)	昇順

また、提案手法の *overlap* パラメータによる順位変動とそのときの類似度の変化を確認する。理論上は、*overlap* が 0 に近いほどピアソン相関係数での類似度評価結果に近づく、*overlap* が 1 に近いほど DTW での類似度評価結果に近づくはずである。なお、評価 (1) における順位は、4.1.1 項より $p_t(x)_{norm}$ と $p_s(x)_{norm}$ の形状に近い画像ほど高くなると定義したため、背景などの環境ノイズによって、*target* 画像の動植物の分類系統と近いものが上位にこない場合があることに注意する。

評価 (2) では、 $overlap = 0$ と $overlap > 0$ の場合で、類似度が著しく変動しないことを確認し、マージン m_i の導入が実用上不要であることを示す。また、提案手法、ピアソン相関係数、DTW がどの程度の距離（類似度の差）をもって順位付けしているかを可視化し、順位付けのロバスト性を確認する。たとえば、似ている画像どうし（猫の画像どうしやチータの画像どうしなど）は、多少順位が入れ替わっても問題ないため、似ている画像どうしは、近い距離に配置されるのが望ましい。一方、明らかに似ていない画像どうしは、遠い距離に配置するのがよい。すなわち、上記の特性を持っていれば、類似画像の類似度がノイズによって多少変化したとしても、順位変動を抑えられる（ロバスト性がある）といえる。

表 1 に示したとおり、評価関数の変域および順位付け規則が類似度評価手法ごとに異なるため、転移元画像間の類似度の距離を共通の基準で確認できるように評価関数を次のように変更する。まず、ピアソン相関係数の評価関数に対してのみ出力値に -1 を乗算し、順位付け規則を昇順にする。そして、すべての手法の評価関数の出力値を、式 (6) と同様の方法で $[0, 1]$ に正規化する（以降、この値を「相対距離」と表記する）。

この相対距離を基にクラスタリングし、*overlap* パラメータを変化させても、クラスタリング結果が変わらないことを確認する。さらに、転移用の訓練データセットの構築において最も重要なのは、転移先データとの類似度が高い転移元データが検索されることである。つまり、相対距離が最も近いクラスタが変動しなければ、有効な訓練データセットの構築に必要なデータが、検索によって抽出されやすくなると考えられる。

表 2 評価条件一覧

Table 2 List of evaluation conditions.

類似度評価手法	値
提案手法、ピアソン相関係数、DTW	
<i>overlap</i>	0, 0.05, 0.1, 0.2, 0.4
w_i	1 (重み付けなし)
評価用画像	動植物画像 16 枚 (図 6) (<i>target</i> 画像 1 枚, 転移元画像集合 15 枚) (ImageNet の API から入手[27])

表 3 PC ハードウェア環境

Table 3 Evaluation hardware specifications.

	値
CPU	Intel Core i5-8500 CPU @ 3GHz (6 コア 6 スレッド)
メモリ	64GB
ディスク	Samsung SSD 860 EVO 1TB × 2 (シーケンシャル Read 550 MB/s) (シーケンシャル Write 520 MB/s)

表 4 PC ソフトウェア環境

Table 4 Evaluation software specifications.

	値
OS	Windows 10 Enterprise バージョン 1909 (64bit)
ランタイム	Python 3.6
主な使用ライブラリ	OpenCV[28] (グレースケール化) scipy.stats[29] (カーネル密度推定) scipy.signal[30] (極値算出) dtw[31] (DTW 距離算出)

4.2 評価条件

表 2 に評価条件の一覧を示す。

4.3 評価環境

PC のハードウェア環境を表 3 に、ソフトウェア環境を表 4 に示す。

4.4 評価結果

4.4.1 手法ごとの類似順位の変化

target 画像の特徴量（赤実線）と転移元画像の特徴量（青破線）、 $overlap = 0$ 時の区間分割位置（縦実線）をプロットしたグラフを図 7 に示す。また、手法ごとの類似順位の変化をプロットしたものを図 8 に示す。

動植物の表面の柄によって画素値の確率質量関数が決まると期待するため、*target* 画像（トラ）は、*tiger1*, *tiger2* や *cheetah1*, *cheetah2* などの順位が高いことが望ましい。

ピアソン相関係数は増減関係を重視するため、*target* 画像の特徴量と形状が重なるような *cheetah1* (1 位), *tiger2* (2 位), *tiger1* (4 位) は上位だが、*cheetah2* (7 位) のように極大値を引き延ばした形状の順位が下がり、*x* 軸方向の伸縮への許容性が低い。また、*giraffe1* (5 位) や *cat1* (6 位) のように、データ分布の増減関係が似ているものの順位が高く出ており、*p(x)* 軸方向の伸縮への許容性が高す

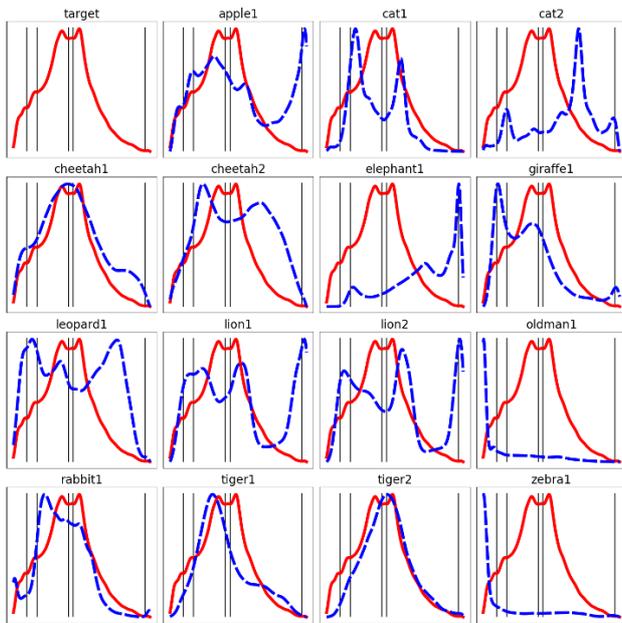


図 7 特徴量と分割区間

Fig. 7 Feature data distributions and their peaks.

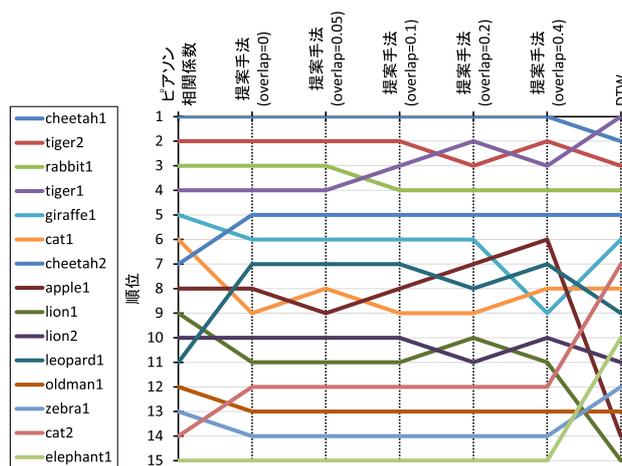


図 8 手法ごとの類似順位の変化

Fig. 8 Change in order of similarity by evaluation methods.

ざる。

提案手法 ($overlap = 0$) は, *cheetah2* (7位→5位), *leopard1* (11位→7位) の順位が上がり, *giraffe1* (5位→6位) や *cat1* (6位→9位) の順位が下がっており, 期待した結果に近づく. $overlap$ を 1 に近づけると, 極値のシフトに寛容となり, 似た特徴量を持つ *cheetah1*, *tiger2*, *rabbit1*, *tiger1* の順位が変動するようになる. 同様に, *giraffe1*, *leopard1*, *apple1*, *cat1* の間と, *lion1*, *lion2* の間の順位にも変動が生じる.

DTW は, *cat2* や *elephant1* のような極値が大きくシフトするパターンの順位を高くしている. 一方で, *apple1* や *lion1* のような極値の数が異なるパターンの順位を低くする. 提案手法の $overlap$ が 1 に近いパターンよりもかなり極値のシフトに寛容となっていることが分かる.

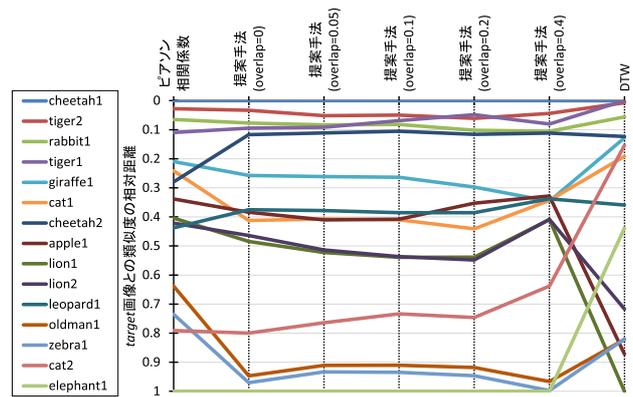


図 9 手法ごとの類似度の相対距離の変化

Fig. 9 Change in relative distance of similarity for each method.

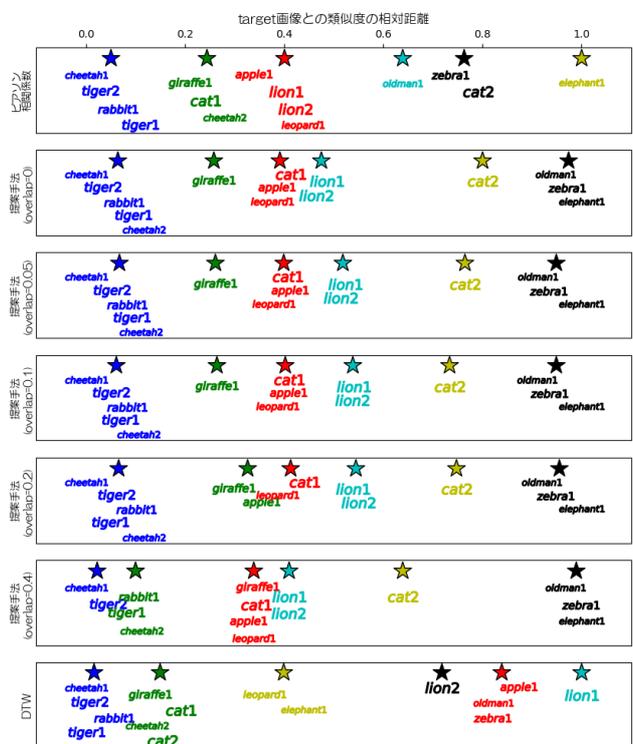


図 10 k -means ($k = 6$) クラスタリング例

Fig. 10 k -means ($k = 6$) clustering using the similarity.

4.4.2 $overlap$ パラメータによる類似度の変化

target 画像との類似度の相対距離の変化を手法ごとにプロットしたものを図 9 に示す. また, 図 10 は, *target* 画像との類似度の相対距離を用いて, k -means ($k = 6$) でクラスタリングした例である. 図中の星印は, 各クラスタの中心点である. なお, 相対距離が 0 に近いほど *target* 画像と似ていることを示している.

4.1.2 項で示したように, 重要なのは相対距離が最も近いクラスタが変動しないことである. 本評価条件での類似画像検索結果で期待する範囲を図 11 に示す. すなわち, トラ, ライオン, チータ, ネコの画像が, 相対距離の最も近いクラスタに配置されればよい. また, 分類系統において

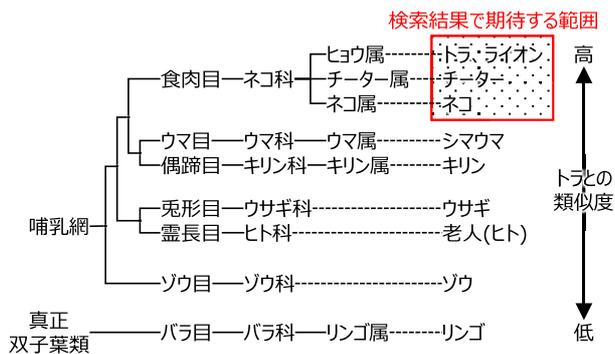


図 11 類似画像検索の結果で期待する範囲

Fig. 11 Expected range of the similar image search results.

近いものどうしが近くに配置される方がよい。

ピアソン相関係数, DTW, 提案手法とも, 相対距離の最も近いクラスにトラ, チータ, ネコが含まれ, 図 11 で示した期待する範囲内である. よって, 時系列データの類似度計算手法が有効に機能しており, 本評価条件においては類似画像検索の最も重要な要件を満たしているといえる.

しかし, ピアソン相関係数は, $\{cheetah2, giraffe1, cat1\}$ や $\{zebra1, cat2\}$ をクラスタリングしており, 分類系統においてやや離れているものが同一クラスに配置されている.

同様に, DTW でも, $lion1, lion2$ が異なるクラスに分けられ, $\{leopard1, elephant1\}$ をクラスタリングするなど, 分類系統の近いものどうしを近くに配置していない. 一方, 提案手法は $overlap = 0 \sim 0.1$ の場合, $\{cheetah1, cheetah2, tiger1, tiger2, rabbit1\}$ や $\{lion1, lion2\}$ のようにネコ科のものどうしをクラスタリングしており, $\{giraffe1\}$ や $\{oldman1, zebra1, elephant1\}$ のようなトラと分類系統上離れているものは, 異なるクラスに配置しているため, 想定する動作に近い. $rabbit1$ や $apple1$ がネコ科のクラスに入っているが, これは画像に識別対象以外の環境ノイズ (背景など) が混入して特徴量のデータ分布の形状が $target$ 画像のものと意図せず似たことが影響していると考えられる. 一方で, $overlap$ を大きくしても, おおむねクラスタリング結果は変わらないが, クラスタ間距離の近いものどうしが近づく傾向にある. ただし, 順位の変動はクラス内にとどまっている.

本評価では, 類似度を基に順位付けしたが, k -means のようなクラスタリングアルゴリズムを利用し, $target$ 画像との類似度が近い順にクラスを特定した後, 各クラス内の画像を提示すれば順位を安定させることが可能である.

5. 考察

本評価においては, 転移先と転移元データの特徴量の極値のシフトや極値の形状の相似形に対して, 提案手法がピアソン相関係数と比較し, 適切なロバスト性をもって類似度の評価が可能であることを確認した. さらに, DTW で

問題となる過剰な形状一致による類似度の誤判定を, 提案手法の極値による区間分割によって解決することができた.

また, 提案手法は, $overlap$ パラメータの導入により, 類似順位と類似度が若干変化したが, 変動の範囲は限定的であった. 実用上は, クラスタリングアルゴリズムの使用によって順位変動を吸収できる見込みで, 本評価条件下では, $overlap$ パラメータは不要である.

制約事項として, 類似性評価対象データの特徴量は, 本評価で用いた $p(x)_{norm}$ のように, x_i に対し $p(x_i)_{norm}$ がただ 1 つ定まるものを使用する必要がある. そして, 提案手法は特徴量のデータ分布の極値を重視するため, 訓練データから特徴量を抽出する段階で, 外れ値やノイズを含むとその影響を受けやすい. 本評価では, 画素値の確率質量関数を特徴量としたため, 画像全体の色彩を特徴として強く反映する. したがって, 背景画像の割合が大きいと意図しない特徴が抽出されることがある. 以上のことから, 一般的な機械学習における訓練データの前処理と同じく, 特徴抽出時に外れ値やノイズの除去を行い, 意図する特徴のみが含まれるように留意しなければならない.

6. 結論

本稿では, 転移学習のためのデータ類似度評価手法を提案した. 提案手法は, 特徴量のデータ分布の比較における極値のシフトや極値の形状の相似形に対応するために, データ類似度評価区間を極値で分割し, 区間ごとに DTW 距離を算出し, 各距離を結合して最終的な類似度を得る. 画像認識での転移学習のユースケースを想定した評価の結果, 極値のシフトや極値の形状の相似形が存在する特徴量のデータ分布の比較においても, 類似画像を抽出可能であった. また, 同評価環境下では, 提案手法は, ユーザによる調整が必要なパラメータが不要であることを確認した.

今後は実用化に向け, 類似度評価におけるクラスタリングアルゴリズムのチューニングや, 並列分散処理による検索速度の向上などを目指していく.

参考文献

- [1] 総務省: 次世代 AI × ICT データビリティ戦略, 総務省 (オンライン), 入手先 https://www.soumu.go.jp/main_content/000501300.pdf (参照 2020-03-08).
- [2] Microsoft Azure: Azure IoT Edge とは, Microsoft (オンライン), 入手先 <https://docs.microsoft.com/ja-jp/azure/iot-edge/about-iot-edge> (参照 2018-12-14).
- [3] Amazon Web Services: AWS IoT Greengrass とは, Amazon Web Services (オンライン), 入手先 <https://docs.aws.amazon.com/ja-jp/greengrass/latest/developerguide/what-is-gg.html> (参照 2018-12-14).
- [4] Chen, Y., Qin, X., Wang, J., Yu, C. and Gao, W.: Fed-Health: A Federated Transfer Learning Framework for Wearable Healthcare, *IEEE Intelligent Systems*, DOI: 10.1109/MIS.2020.2988604 (2020).
- [5] 神島敏弘: 転移学習, 人工知能学会誌, Vol.25, No.4, pp.572-580 (2010).

- [6] 中山英樹：深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習，電子情報通信学会技術研究報告，Vol.115, No.146, pp.55-59 (2015).
- [7] 森 郁海，伊藤岳広，中村嘉隆，稲村 浩：AI/IoT ソリューションへの転移学習適用における学習モデル検索システムの提案と評価，第 27 回マルチメディア通信と分散処理ワークショップ論文集，pp.141-148 (2019).
- [8] 喜多泰代：二次元濃度ヒストグラムを用いた画像間変化抽出，電子情報通信学会論文誌，Vol.J90-D, No.8, pp.1957-1965 (2007).
- [9] 大美英一，齊藤彦彦：最頻度エッジ方向に基づく画像の傾き検出と視覚的傾斜角との関係，電気学会論文誌 C (電子・情報・システム部門誌)，Vol.124, No.1, pp.112-118 (オンライン)，DOI: <https://doi.org/10.1541/ieejieiss.124.112> (2004).
- [10] Bellman, R. and Kalaba, R.E.: On adaptive control processes, *IRE Trans. Automatic Control*, Vol.4, No.2, pp.1-9 (1959).
- [11] Aghabozorgi, S., Shirkhorshidi, A.S. and Wah, T.J.: Time-series clustering: A decade review, *Information Systems*, Vol.53, pp.16-38 (online), DOI: 10.1016/j.is.2015.04.007 (2015).
- [12] Vlachos, M., Kollios, G. and Gunopulos, D.: Discovering Similar Multidimensional Trajectories, *Proc. 18th International Conference on Data Engineering*, pp.673-684 (2002).
- [13] Latecki, L.J., Megalooikonomou, V., Wang, Q., Lakaemper, R., Ratanamahatana, C. and Keogh, E.: Elastic Partial Matching of Time series, *Knowledge Discovery in Databases: PKDD2005*, pp.577-584 (2005).
- [14] Keogh, E., Lonardi, S., Ratanamahatana, C., Wei, L., Lee, S.H. and Handley, J.: Compression-based data mining of sequential data, *Data Mining and Knowledge Discovery*, Vol.14, pp.99-129 (2007).
- [15] Kullback, S. and Leibler, R.A.: On Information and Sufficiency, *Annals of Mathematical Statistics*, Vol.22, No.1, pp.79-86 (online), DOI: 10.1214/aoms/1177729694 (1951).
- [16] Jaccard, P.: Distribution de la Flore Alpine dans le Bassin des Dranses et dans quelques régions voisines, *Bulletin de la Société Vaudoise des Sciences Naturelles*, Vol.37, No.140, pp.241-272 (1901).
- [17] Sørensen, T.: A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons, *Kongelige Danske Videnskabernes Selskab*, Vol.5, No.4, pp.1-34 (1948).
- [18] Dice, Lee R.: Measures of the Amount of Ecologic Association Between Species, *Ecology*, Vol.26, No.3, pp.297-302 (online), DOI: 10.2307/1932409 (1945).
- [19] Szymkiewicz, D.: Une contribution statistique a la geographie floristique, *Acta Societatis Botanicorum Poloniae*, Vol.34, No.3, pp.249-265 (1934).
- [20] Smyth, P.: Clustering Sequences with Hidden Markov Models, *Proc. 9th International Conference on Neural Information Processing Systems (NIPS'96)*, pp.648-654 (1996).
- [21] Kalpakis, K., Gada, D. and Puttagunta, V.: Distance Measures for Effective Clustering of ARIMA Time-Series, *Proc. 2001 IEEE International Conference on Data Mining*, pp.273-280 (2001).
- [22] Senin, P.: Dynamic Time Warping Algorithm Review, *Information and Computer*, Science Department University of Hawaii at Manoa Honolulu, USA, Vol.855, pp.1-23 (2008).
- [23] Csurka, G., Dance, C.R., Fan, L., Willamowski, J. and Bray, C.: Visual Categorization with Bags of Keypoints, *ECCV International Workshop on Statistical Learning in Computer Vision*, pp.1-22 (2004).
- [24] Dalal, N. and Triggs, B.: Histograms of Oriented Gradients for Human Detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol.1, pp.886-893 (online), DOI: 10.1109/CVPR.2005.177 (2005).
- [25] Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints, *Int. J. Comput. Vision*, Vol.60, No.2, pp.91-110 (2004).
- [26] Bridle, J.S. and Brown, M.D.: An Experimental Automatic Word-Recognition System, *JSRU Report*, No.1003, Joint Speech Research Unit, Ruislip, England (1974).
- [27] ImageNet: API documentation, ImageNet (online), available from <http://image-net.org/download-API> (accessed 2020-03-24).
- [28] OpenCV: Changing Colorspaces, OpenCV (online), available from https://docs.opencv.org/4.1.1/df/d9d/tutorial_py_colorspaces.html (accessed 2020-03-24).
- [29] SciPy: Kernel density estimation, *scipy.stats* (online), available from <https://docs.scipy.org/doc/scipy/reference/tutorial/stats.html> (accessed 2020-03-24).
- [30] SciPy: Peak finding, *scipy.signal* (online), available from <https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.argrelemin.html> (accessed 2020-03-24).
- [31] Pierre Rouanet: dtw 1.4.0, pypi (online), available from <https://pypi.org/project/dtw/> (accessed 2020-03-24).



森 郁海 (正会員)

2009年公立ほこだて未来大学大学院システム情報科学研究科博士前期課程修了。同年三菱電機株式会社入社。2019年公立ほこだて未来大学大学院システム情報科学研究科博士後期課程入学。クラウドソリューションに関する

研究に従事。



中村 嘉隆 (正会員)

2007年大阪大学大学院情報科学研究科博士後期課程修了。同年奈良先端科学技術大学院大学情報科学研究科助教。2010年大阪大学大学院情報科学研究科特任助教。2011年公立ほこだて未来大学システム情報科学部助教。

2016年より同大学准教授。博士(情報科学)。情報セキュリティに関する研究に従事。電子情報通信学会，IEEE各会員。本会シニア会員。



稲村 浩 (正会員)

1990年慶應義塾大学大学院理工学研究科修士課程修了。同年日本電信電話(株)入社。1998年よりNTTドコモ。2016年より公立はこだて未来大学教授。博士(工学)。モバイルネットワーク, スマートデバイスのシステム

ソフトウェアに関する研究開発に従事。電子情報通信学会, ACM, IEEE 会員。本会フェロー, シニア会員。