

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 4 日現在

機関番号：12611

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500274

研究課題名(和文) 言葉によるコンピューティングのための時系列データの言語化基盤技術の開発

研究課題名(英文) Development of a basic technology for verbalizing time-series data for computing with words

研究代表者

小林 一郎 (KOBAYASHI, Ichiro)

お茶の水女子大学・大学院人間文化創成科学研究科・教授

研究者番号：60281440

交付決定額(研究期間全体)：(直接経費) 4,000,000円、(間接経費) 1,200,000円

研究成果の概要(和文)：複数の時系列データの間関係を言語化する手法を開発した。時系列データの相関関係を捉え、SAXにより数値情報を記号化し、データ間の距離を改良した編集距離により計測され言葉で表現する。さらに動画に映る物体と人との相互作用を言語化する枠組みを提案した。時系列データを SAXおよび動的計画法により記号化・圧縮し、対数線形モデルを用いることによりそのパターンと対応する意味ラベルとの関係を学習した。人の動作を説明する様々な文章を基にバイグラムモデルを構築し、尤もらしい文を生成した。また、転移学習を適用し、不足する言語資源を補う手法を提案した。提案手法は実験によりその有効性が検証された。

研究成果の概要(英文)：We aim to develop a method to verbalize multiple time-series data with words. We calculate the correlation coefficient between two time-series data and convert the data into symbols by SAX. The edit distance of the symbols of two time-series data is calculated and expressed with words. We also propose a method to verbalize the interaction between a human and an object. SAX with dynamic programming is applied to the time-series data to extract their patterns, and then the correspondence between the patterns and their semantic labels are learned by a log-linear model. We build a bi-gram model based on the collected linguistic resources. The most likely sentences to explain the interaction are generated by solving the bi-gram model with dynamic programming. Moreover, to enhance the linguistic resources used to generate sentences, we adopt transfer learning of n-gram language models. Through experiments, we have confirmed that our proposed framework works well.

研究分野：総合領域

科研費の分科・細目：情報学

キーワード：時系列データ 言語化 SAX 編集距離 バイグラムモデル テキスト生成 機械学習 動画

1. 研究開始当初の背景

今日、多くのセンサが使用されているが、そのようなセンサから観測されるデータのほとんどは時系列データである。時系列データは統計的に処理されることが多く、人が理解しやすいようにするためには可視化などによって、データの挙動を視覚的に捉える方法が一般的である。しかし、一方で内容をわかりやすく言葉で説明する手法も提案されている。しかし、先行研究の多くが時系列データの局所的特徴を捉えたものになっており、人が時系列データの大局的な動向を視覚で認識し説明する言語化はできていない。また、複数の時系列データの関係を要約し言語化する手法(例:最近の各国の景気動向の比較、デフレの傾向と株価の関連等)は開発されていない。また、時系列データ解釈の視点を変えた際に、いくつかの局所的な特徴を意味的に総括する語彙で表現することができない。

2. 研究の目的

時系列データの言語化において、未だできていない複数の時系列データを捉え、それらの大局的な動向を捉えることにより、時系列データを言語化する手法を提案する。また、提案する手法を実際に二つの応用システム(1)複数時系列データ間の特徴点の抽出と(2)人の動作を対象にした確率的言語生成への取り組み、として実装し、提案手法の有効性を検証する。

3. 研究の方法

(1) 複数時系列データ間の特徴点の抽出
本研究においては、SAX (Symbolic Aggregation approXimation) を使って時系列データを文字列に変換し、記号間の編集距離を使って二つの時系列データの類似度を測る。通常の編集距離は、対応する個々の記号の比較において、記号が異なる数、または、記号を一致させるのに必要なコストを2つの時系列データの距離(差異)とするが、本研究では、記号列の動向を比較し、同じ動向を持つ記号列に変更するのに要するコストを新たに編集距離として採用する(図1中、アルファベットの下の数値がそれぞれの時系列データの動向を示す)。

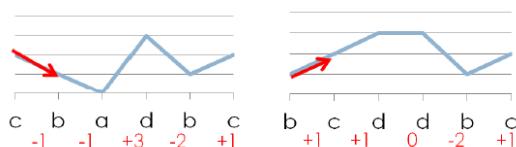


図1 編集距離の拡張

また、抽出されたこの2つの数値列をマッチングすることによって動向を比較し、2つの時系列データ間の以下に示す2つの関係を取得する。

1. 類似した動き

動きを示す値が全く同じ箇所。または、動きを示す値が正(上昇)なら「+」を、負(下降)なら「-」を、0(一定)なら「0」を当てはめ、その記号が同じ箇所。

2. 対称の動き

正負は違うものの、動きを示す値の絶対値が全く同じ箇所。または、動きを示す値が、正(上昇)なら「+」を、負(下降)なら「-」を当てはめ、その記号が全く逆の箇所。

・相関関係に基づく特徴抽出

時系列データの相関関係は、相関係数の値により次の3つのタイプに分類される。

(i) 相関係数が正に高い組, (ii) 相関係数が負に高い組, (iii) 相関係数の低い組

ここで、(i) は類似の動きをする時系列データ、(ii) は対称の動きをするデータ、(iii) は関連性が無いとされる時系列データであることを示す。本研究では(i)(ii)(iii)において、比較する時系列データから特異な特徴を抽出することにより、ユーザに時系列データの比較において新たな気づきを与える。

上記(i)(ii)(iii) に対する抽出方法をそれぞれ示す。

(i) 相関が正に高い組

おおよそ類似するデータ間において対称に類似する箇所の抽出を行う。まず SAX によるフレームの間隔を大きくとり編集距離を用いたマッチングをすることによって、大まかに見て「類似した動きをする箇所」を抽出する。その後フレームの間隔を小さくとり、また編集距離を用いたマッチングをすることによって、部分的に「対称の動きをする箇所」を抽出する。

(ii) 相関が負に高い組

おおよそ対称に類似するデータ間における類似箇所の抽出を行う。(i) 相関が正に高い組と同様にまず SAX によるフレームの間隔を大きくとり編集距離を用いたマッチングをすることによって、大まかに見て「対称の動きをする箇所」を抽出する。その後 SAX のフレーム間隔を小さくとり、編集距離を用いたマッチングをすることによって、部分的に「類似した動きをする箇所」を抽出する。

(iii) 相関が低い組

関連性の低いデータ間における類似箇所また対称に類似する箇所の抽出を行う。まず、データを時間軸上に細かく分けそれぞれ相関係数を取り、部分的に相関の高い箇所を見つける。その箇所に対し SAX によるフレームの間隔を小さくとり編集距離を用いたマッチングをすることによって、部分的に「類似した動きをする箇所」また「対称の動きをする箇所」を抽出する。

・抽出された特徴点の言語化

抽出によって得られた特徴点を、定義した編集距離を用い、あらかじめ用意したテンプレートに当てはめ言語化する．以下にその例を示す(図2 参照)．

テンプレート例	条件	条件に合う編集距離の例
どちらも下落の動きを示し、[データ1]の方が下げ幅が大きい	編集距離の値がどちらも全て負の値である下落の動きで「類似の動きをしている」と判断され、また[データ1]の方が編集距離の絶対値が大きかった場合	データ1[-4,-2]、データ2[-2,-1]
どちらも同様に上昇ののち下落する動きを示しているが、[データ1]の方が下落の下げ幅が大きい	編集距離の値がどちらも正の値である山形の動きで「類似の動きをしている」と判断され、また編集距離の負の値のみが異なり[データ1]の方がその絶対値が大きかった場合	データ1[2,-6]、データ2[2,-3]
[データ1]は上昇ののち下落する動きを示しているが、[データ2]は逆に下落ののち上昇する動きを見ている	「対称の動きをしている」と判断され、[データ1]の編集距離が正の値のち負の値をとる山形の動きで、[データ2]の編集距離が負の値のち正の値をとる谷形の動きをする場合	データ1[3,-2]、データ2[-3,2]

図2 言語化に使用するテンプレートの例

・実験

2011年12月5日の日経平均17業種別株価(全18個)について、それぞれ、9:00~15:00(休憩時間11:30~12:30)を5分足でとってきた時系列データ(データ数:62)を使用し、それらの時系列データ間の関係を言語化する．

言語化の手順を以下に示す．

- step 1. データの入力
- step 2. 相関係数によるタイプ分け
- step 3. 比較
- step 4. 言語化

結果を図3に示す．

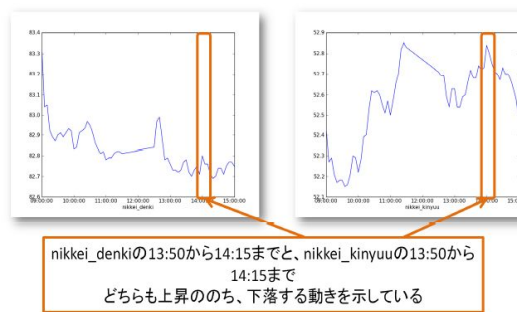


図3 低相関における類似した動作の言語化

・被験者実験

言語化の正当性を検証するために、20代の女子学生13人に言語化結果に対するアンケートを行う被験者実験を行った．その結果を図4に示す．

タイプ	(i)類似の動きをするもの	(ii)対称の動きをするもの	(iii)関連性がないもの
抽出された特徴点	(A)大まかに見て、特に似た動きをする箇所	(B)細かく見て、逆に似た動きをする箇所	(A)大まかに見て、特に似た動きをする箇所
抽出された特徴点	0% 0% 0%	0% 0% 0%	8% 0% 0%
抽出された特徴点	38% 62%	23% 77%	10% 89%
抽出された特徴点	100%	100%	92%
抽出された特徴点	100%	100%	92%
抽出された特徴点	100%	100%	92%

図4 アンケートによる被験者実験結果

(2) 人の動作を対象にした確率的言語生成への取り組み

人の動作を捉えるために、本研究ではMicrosoft社が製造しているKinectカメラを用いて、人の骨格情報の時系列データを観測する．人の骨格情報の取得と同時に人の動作に関する物体を追跡した時系列データをパーティクルフィルタを用いて取得する．

研究概要を図5に示す．

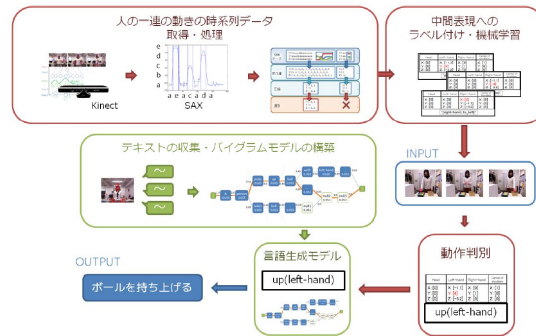


図5 動画を入力とする確率的テキスト生成の枠組み

以下、各部について説明する．

・時系列データの処理

人の骨格と物体の色を追跡することで得られた時系列データは、SAXを使い、文字列に変換する．本研究では動作判別の精度を高めるため、一般にデータを等間隔に分割するところを、各データに動的計画法を用い、尤もらしい区切りを取得することで、よりデータに沿った文字列を取得する．取得した文字列からデータの変化量を抽出し、その後の処理における機械学習に適用し易いように圧縮を行い、ノイズと思われる不要データを排除することにより最終的に時系列データのパターンを抽出する(図6参照)．

・中間表現

テキスト生成では、時系列データと自然言語文をつなぐ中間表現を用いることでテキスト生成に使う言語資源を選択する．中間表現は表1のように定義する．

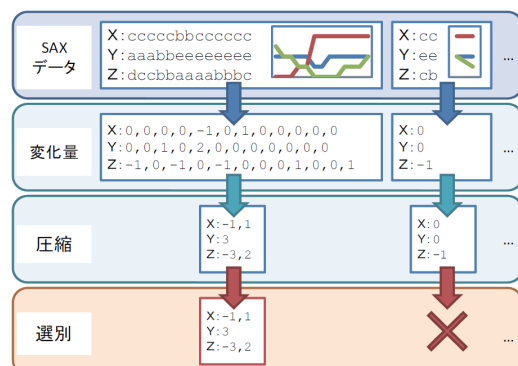


図6 時系列データのパターン抽出

多い．そこで，動作には組み合わせ構造があると仮定し，ドメイン適用による言語資源の拡充を試みる．ここでは，ドメイン適用に最小二乗推定を用いる．

要素トピックの確率分布パラメータを ϕ_i ($i = 1, \dots, K$)，観測データの確率分布パラメータを ψ_j ($j = 1, \dots, M$) とする．また，どの観測データも，down-front-hand-right といった 4 つの要素トピックの混合であると仮定する．要素トピックの混合行列を， $M \times K$ 行列である A とし，式(2) の最小化問題を解くことで推定する．ここで， $\Phi = (\phi_1, \dots, \phi_K)^T$ ， $\Psi = (\psi_1, \dots, \psi_M)^T$ である．

$$\hat{\Phi} = \min_{\Phi} \|\Psi - A\Phi\|^2 = A^+ \Psi \quad (2)$$

いま，対象とする動作は 20 あり，それらの動作には図 9 のような組み合わせ構造があるとする．



図 9 実験における動作組み合わせ構造

要素トピック数 $K = 9$ ，観測カテゴリ数 $M = 20$ とする．また，動的計画法を用いて尤もらしい文を生成するには，バイグラムの遷移確率と各単語の出現確率を求める必要があるため，要素トピックの確率分布パラメータと観測データの確率分布パラメータを遷移確率と出現確率の各々用意し，推定を行う．20 ある動作の内一つの動作の持つ言語資源を削除し，上述したドメイン適用を用いてその言語資源を推定するという実験を全ての動作を対象に行った．20 の動作内 6 動作に対して，言語資源を用いて生成された文の上位 3 件を表 3 に，他の動作の資源のみを用いて生成された文上位 3 件を表 4 に示す．

表 3 言語資源を用いた生成文の上位 3 件

動作	生成文	尤度
1	● 右手, を, あげる, . . . , null4, null5, null6, null7, null8, null9,	3.43e-30
	● 右手, を, あげる, . . . , null5, null6, null7, null8, null9, null10,	1.92e-30
	● 右手, を, 上, に, あげる, . . . , null4, null5, null6, null7,	6.54e-31
2	● 右手, を, 下げる, . . . , null4, null5, null6, null7, null8, null9,	2.10e-30
	● 右手, を, 下, に, 下ろす, . . . , null4, null5, null6, null7,	1.76e-32
	● 右手, を, 上, から, 下, に, 下ろす, . . . , null4, null5,	1.85e-33
3	● 左手, を, あげる, . . . , null5, null6, null7, null8, null9, null10,	7.72e-31
	● 左手, を, あげる, . . . , null4, null5, null6, null7, null8, null9,	4.41e-31
	● 左手, を, 上, に, 挙げる, . . . , null5, null6, null7, null8,	7.36e-32
⋮	⋮	⋮
18	● 右足, を, 下ろす, . . . , null4, null5, null6, null7, null8, null9,	4.48e-31
	● 右足, を, 横, から, 下げる, . . . , null4, null5, null6, null7,	1.48e-32
	● 右足, を, 横, に,) , 下げる, null7, null8, null9, null10,	1.87e-33
19	● 左足, を, 横, に, あげる, . . . , null6, null7, null8, null9,	4.57e-30
	● 左足, を, 横, に, あげる, . . . , null7, null8, null9, null10,	1.23e-30
	● 左足, を, 横, に, あげる, . . . , null5, null6, null7, null8,	1.17e-31
20	● 左足, を, おろす, . . . , null4, null5, null6, null7, null8, null9,	2.18e-31
	● 左足, を, 横, に,) , おろす, . . . , null8, null9, null10,	6.22e-33
	● 左足, を, 横, に, 下ろす, . . . , null4, null5, null6, null7,	3.24e-33

表 4 他の動作の言語資源のみを用いた生成文の上位 3 件

動作	生成文	尤度
1	● 右手, を, あげる, . . . , null5, null6, null7, null8, null9, null10,	2.38e-31
	● 右手, を, あげる, . . . , null4, null5, null6, null7, null8, null9,	1.66e-31
	● 右手, を, 上, に, あげる, . . . , null5, null6, null7, null8,	1.69e-32
2	● 右手, を, 下げる, . . . , null4, null5, null6, null7, null8, null9,	2.33e-31
	● 右手, を, 下げる, . . . , null5, null6, null7, null8, null9, null10,	1.00e-31
	● 右手, を, 上, から, 下げる, . . . , null4, null5, null6, null7,	6.85e-33
3	● 左手, を, 上, に, あげる, . . . , null6, null7, null8, null9,	1.47e-31
	● 左手, を, あげる, . . . , null5, null6, null7, null8, null9, null10,	9.02e-32
	● 左手, を, 上, に, あげる, . . . , null5, null6, null7, null8,	1.48e-32
⋮	⋮	⋮
18	● 右足, を, 下げる, . . . , null4, null5, null6, null7, null8, null9,	1.86e-32
	● 右足, を, 下げる, . . . , null5, null6, null7, null8, null9, null10,	3.79e-33
	● 右足, を, 横, に, 下ろす, . . . , null4, null5, null6, null7,	1.63e-33
19	● 左足, を, 横, に, あげる, . . . , null6, null7, null8, null9,	5.92e-32
	● 左足, を, あげる, . . . , null5, null6, null7, null8, null9, null10,	1.93e-32
	● 左足, を, 横, に, あげる, . . . , null4, null5, null6, null7,	2.55e-33
20	● 左足, を, 下ろす, . . . , null4, null5, null6, null7, null8, null9,	2.89e-32
	● 左足, を, 下ろす, . . . , null5, null6, null7, null8, null9, null10,	6.53e-33
	● 左足, を, 横, に, 下ろす, . . . , null4, null5, null6, null7,	6.43e-34

・考察

表 3 と表 4 を比較してみると，若干表現は違うものの，他の動作の言語資源のみで人の動作を表現する文が生成出来ていることが確認できる．

4 . 研究成果

(1) 複数時系列データ間の特徴点の抽出

時系列データを捉える設定した 3 つの項目とも，「とても一致している」「一致している」と答えた方が 9 割を超えている．これにより，実際のデータの振る舞いを言語によって表わすことができていると考えられる．これにより提案手法の有効性が確認できた．

(2) 人の動作を対象にした確率的言語生成への取り組み

，動画像中の人の動作を表現する確率的言語生成の枠組みを提案した．Kinect ビデオで抽出された人の動作およびパーティクルフィルタで取得された物体の軌跡は，時系列データとして取得され，いくつかの次元圧縮手法を適用することで機械学習に適した形に変換された後，対数線形モデルで人の動作を表す中間表現の対応関係が学習される．また観測された人の動きを表現するために，被験者実験によって集められた自然言語文に基づきバイグラムモデルを構築し，動的計画法を適用することで，文生成に単語数の制限をつけずに自然言語文生成を行うことができた．さらに，言語資源が十分でないときを想定し，最小二乗推定による言語資源の拡充も試みた．提案手法に基づく実験を通じ，その結果より提案手法の有効性を検証した．

5. 主な発表論文等

(研究代表者, 研究分担者及び連携研究者には下線)

[雑誌論文](計 2 件)

北島理沙, 小林一郎, 文書上の単語対を素性とした潜在的トピック推定, 日本知能情報ファジィ学会論文誌, Vol. 25, No. 1, pp. 501-510, 2013. 査読有,
https://www.jstage.jst.go.jp/article/jsoft/25/1/25_501/_article/-char/ja/
Risa Kitajima and Ichiro Kobayashi, Latent Topic Estimation Based on Events in a Document, pp. 603-610, Journal of Advanced Intelligence and Intelligent Informatics, Vol.16 No.5 July 2012. 査読有,
<http://www.fujipress.jp/finder/xslt.php?mode=present&inputfile=JACII001600050006.xml>

[学会発表](計 11 件)

小林瑞季, 小林一郎, 麻生英樹, 人と物とのインタラクションを記述する確率的言語生成への取り組み, 3P-7, (学生奨励賞), 第76回情報処理学会全国大会, 北千住, 東京, 2014年3月12日.
一瀬詩織, 小林一郎, 岩爪道昭, 田中康司, RDF リソースのリンク構造を考慮した SPARQL クエリ検索結果のランキング手法, WI2-2013-33, ARG Web インテリジェンスとインタラクション研究会 (ARG SIG-WI2), 横浜, 2013年12月14日.
Mizuki Kobayashi, Ichiro Kobayashi, Hideki Asoh, and Sergio Guadrrama, A Probabilistic Approach to Text Generation of Human Motions extracted from Kinect Videos, the International Conference on Computer Science and Applications (ICCSA'13), San Francisco, 2013年10月24日.
小林瑞季, 小林一郎, 麻生英樹, Kinectにより観測された人の動作を説明する確率的言語生成への取り組み, 第3回 ARG Web インテリジェンスとインタラクション研究会, WI2-2013-38, ARG WI2 No.3, 横浜, 2013年12月14日.
小林瑞季, 小林一郎, 麻生英樹, 動画像中の人の動作を表現する確率的言語生成に関する取組み, 2D5-0S-03b-3, 第27回人工知能学会全国大会, 富山, 2013年6月5日.
鈴木聡子, 小林一郎, 疑似ラベルを教師信号とした潜在ディリクレ配分法への取り組み, 1F3-4in, 第27回人工知能学会全国大会, 富山, 2013年6月5日.
Mizuki Kobayashi and Ichiro Kobayashi, An Approach to Linguistic Summarization based on Comparison

among Multiple Time-series Data, The 6th International Conference on Soft Computing and Intelligent Systems & The 13th International Symposium on Advanced Intelligent Systems, Paper ID:F1-31-5, Kobe, 2012年11月22日.
立川華代, 小林一郎, 共起関係に基づく事前知識を用いた潜在的トピック抽出の取り組み, 4I1-R-9-3, 第26回人工知能学会全国大会, 山口, 2012年6月12日.
小林瑞季, 小林一郎, 複数の時系列データの比較に基づく言語化の試み, 4K1-0S-2-1, 第26回人工知能学会全国大会, 山口, 2012年6月12日.
小林瑞季, 小林一郎, 複数の時系列データの関連性に基づく言語化のための基礎的考察, 第22回 Web インテリジェンスとインタラクション研究会, 東京, 2012年3月17日.
小林瑞季, 小林一郎, 複数の時系列データの関連性発見に基づく言語化の一考察, 第74回情報処理学会全国大会, 名古屋, 2012年3月7日.

6. 研究組織

(1) 研究代表者

小林 一郎 (KOBAYASHI, Ichiro)
お茶の水女子大学・大学院人間文化創成科学研究科・教授
研究者番号: 60281440

(2) 研究分担者

岩爪 道昭 (IWAZUME, Michiaki)
独立行政法人情報通信研究機構・知識創成コミュニケーション研究センター・研究マネージャー
研究者番号: 80319756

(3) 連携研究者

なし