

動的特徴を用いた歩容認証：RNN 及び SVM の性能比較

熊埜御堂, 裕太 / KUMANOMIDO, Yuta

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編 / 法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

58

(開始ページ / Start Page)

1

(終了ページ / End Page)

4

(発行年 / Year)

2017-03-31

(URL)

<https://doi.org/10.15002/00014215>

動的特徴を用いた歩容認証 -RNN 及び SVM の性能比較-

A GAIT RECOGNITION SYSTEM USING DYNAMIC FEATURES.
-COMPARING THE PERFORMANCE OF RNN AND SVM-

熊埜御堂 裕太

Yuta KUMANOMIDO

指導教員 平原誠

法政大学大学院理工学研究科応用情報工学専攻修士課程

This paper presents a gait recognition system that uses recurrent neural networks (RNNs) and support vector machines (SVMs) for identifying individuals. Our system extracts the spatiotemporal features of distances between the waist and various joint positions obtained by a Kinect sensor. These spatiotemporal features are invariant for a walking subject. To verify our system performance, we conducted tests using the data of 12 individuals. The data were divided into two datasets for training and testing. The RNNs and SVMs were trained for classification using the training dataset. SVMs achieved an average accuracy of over 99% for the test dataset, whereas the average accuracy of RNNs was 94%.

Kew Word : RNN,SVM,Kinect,gait recognition

1. はじめに

本研究では歩行時における関節座標の時系列を Kinect より取得し、それらを用いて人物識別を行うシステムを提案する。

Kinect を用いた従来研究としては歩行の周期に合わせ情報を切り出すような人手を必要とするもの[1]や、1 周期分の歩行情報を必要とするものが多かった[2]。

また Kinect の配置によって識別率に変動が起きてしまうため[3]、撮影時の方向を 1 方向に限定した識別も多い[4]。

本研究ではこれらの問題を克服するため、被撮影者との距離や撮影方向に依存しない骨格や関節の動きを用いて人物の識別を行う。

識別器には時系列特徴を学習できるリカレントニューラルネットワーク (RNN) 及び高い汎化能力を持つサポートベクターマシン (SVM) を用い性能を比較する。

2. 提案システム

2.1 特徴抽出

地面に対して水平に設置した Kinect で 12 名の歩行を 30 フレーム撮影する。これは約 2 歩分(1 周期)に相当する。

また性能評価のために被撮影者のうち 3 名に対しては、高さ 2.5m の位置から見下ろすように設置した Kinect と

高さ 0.5m から見上げるように設置した Kinect を用いて同様のフレーム間、歩行の撮影を行う。

各フレームでは、関節の 3 次元座標値が得られるが、これらを腰の中心関節点から手先、手首、肘、肩、足、膝、腰の左右それぞれと首までの関節間距離 d_j ($j=1, \dots, 15$) に変換して特徴量とする。

各特徴について 3 連続フレームごとに中央値を取ることによって 30 フレームを 10 フレーム (以下 1 シーンと呼ぶ) に間引く。図 1 のように 1 シーンには開始フレームの異なる 5 連続フレームが 6 つある。この連続フレームを T=5 フレーム歩行パターン (約半周期) と呼ぶ。同様に 3 連続フレーム (8 つ) とした場合を T=3 フレーム歩行パターン (約 1/3 周期) と呼ぶ。

識別器は 2 つのフレーム歩行パターンに対して別々に作成する。



図 1. 歩行パターンの取得例

2. 2 RNN の構成

ある時刻 t での RNN の構成を図 2 に示す. これをフレーム数 $T(T=3, 5)$ の分だけ展開して層状のネットワークのように動作させる.

RNN の入力層には前述の 15 特徴を受ける 15 素子と閾値を扱う(常に 1 を出力する)素子の計 16 素子を入力素子 $x_i (i=0, \dots, 15)$ として設けた. 隠れ層は 50 素子を設けた. 出力層には識別対象者に対して発火する素子 y_1 とそれ以外の非対象者に対して発火する素子 y_2 の 2 つを設ける.

人物の識別は入力に対してこれら二つの素子のどちらが強く発火したかによって行う. RNN に対してテスト歩行パターンを入力し, 各素子の出力を見比べ $y_1 > y_2$ であれば入力は識別対象者であったと判別し $y_2 > y_1$ であれば非対象者であったと識別する.

RNN の学習には Back Propagation Through Time (BPTT) 法を用いる. BPTT 法は出力と教師信号との二乗誤差のフレーム数分の和を最小化するように学習していく. 目的関数 C は, 出力素子の数を S とすると

$$C = \frac{1}{2} \sum_{t=1}^T \sum_{k=1}^S [y_k(t) - d_k(t)]^2 \quad (1)$$

で与えられる.

本実験では歩行パターンの最終フレーム T にのみ教師信号($d_1(T), d_2(T)$)を与えるため,

$$C = \frac{1}{2} \sum_{k=1}^2 [y_k(T) - d_k(T)]^2 \quad (2)$$

となる. 識別対象者である場合に教師信号 (0.9, 0.1) を, そうでない場合には (0.1, 0.9) を与えた. 以上の RNN を識別対象者ごとに作成する. 以下被撮影者 $P (P=A, B, \dots, L)$ を識別対象者とする RNN を P -RNN と呼ぶ. 3 フレーム歩行パターン用と, 5 フレーム歩行パターン用との P -RNN をそれぞれ作成する.

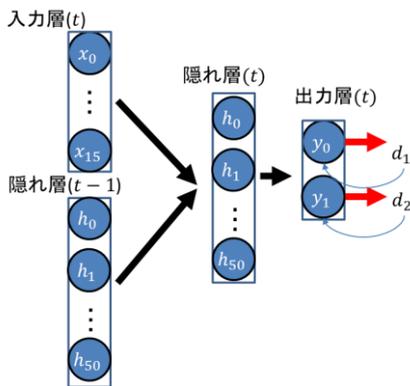


図 2. 本実験における RNN の構成図

2. 3 NAG (Nesterov's Accelerated Gradient)

BPTT 法の学習には勾配降下法を用いる. 勾配降下法は目的関数を各パラメータで偏微分し, その値を元にパラメータの修正を繰り返す. ネットワークのパラメータをまとめて θ とすると修正量 $\Delta\theta$ は

$$\Delta\theta = -\varepsilon \frac{\partial C}{\partial \theta} \quad (3)$$

となる. ここで $\varepsilon (0 < \varepsilon < 1)$ は修正量を調整する小さな値の定数である.

しかし勾配降下法は一般的に収束が遅く学習に時間がかかる. そのため, 本システムでは学習の効率化を目的として NAG[2] を RNN の学習に適用する.

学習の高速化の一つとしてパラメータの修正に過去の修正量を慣性項(Momentum)として加えることで, 学習を高速に進める方法がある. しかし慣性項は過去の修正量が大きい場合に目的の収束地点を飛び越えてしまう欠点を持っている.

NAG[4] はこの慣性項に対し, 現在のパラメータではなく, 次(未来)のパラメータの推定値を計算してから修正を行うことにより目的の収束地点を飛び越えないよう修正を行うものである. $k+1$ 回目の更新量 $\Delta\theta_{k+1}$ は目的関数 $C(\theta + \Delta\theta_k)$ に対する勾配を用いて, 次式で求めることができる.

$$\Delta\theta_{k+1} = \gamma \Delta\theta_k - \varepsilon \frac{\partial \theta(w_k + \Delta w_k)}{\partial w} \quad (4)$$

ここで γ は前回修正量の影響力を調整するものであり 1 以下の正の値である.

2. 4 SVM の構成

SVM はカーネル関数を用いて非線形分類に対応させる. カーネル関数にはガウシアンカーネルを使用する. ガウシアンカーネルを用いた SVM では, 誤分類に対するペナルティの大きさを制御するコストパラメータ C に加えガウシアンカーネルの分布の幅を制御する γ (gamma) という二つのパラメータを持つ.

本実験では C, γ 共に $10^{-3} \sim 10^3$ の範囲においてグリッドサーチを行い全テストデータに対する平均正答率が最も高かった値 $C=10, \gamma=10$ を使用する.

なお歩行パターンには上記 3 及び 5 フレーム歩行パターンに加え 1 フレームの歩行パターンを追加する.

RNN 同様に 12 名の識別対象者の SVM を作成する. 以下被撮影者 $P (P=A, B, \dots, L)$ を識別対象者とする SVM を P -SVM と呼ぶ

3. 実験

以下の全ての実験において、識別器の学習には正面から撮影した歩行パターンのみを使用し、各人物の識別器(識別対象者 P-SVM, 識別対象者 P-RNN)を作成する。

実験 1:

学習には正面からの歩行パターンのみを用い、テストにも正面からの歩行パターンを用いる。

12名の人物(被撮影者 A~L)の各々から 10 シーンを撮影し、9 シーンを学習歩行パターンとし、残りの 1 つのシーンをテスト歩行パターンとして学習、識別を行った。10 通りの識別を繰り返して平均の正答率(Accuracy), 適合率(Precision), 再現率(recall)を求めた。

実験 1: 結果

被撮影者 12 名の 5 フレームテスト歩行パターンを入力とした際の各 RNN の正答率, 適合率, 再現率とその平均を表 1 に示す。図 3 は、3 および 5 フレーム歩行パターンの場合の結果をまとめたものである。

同様に被撮影者 12 名の 5 フレームテスト歩行パターンを入力とした際の各 SVM の正答率, 適合率, 再現率とその平均を表 2 に示す。SVM では 3 フレームおよび 5 フレーム歩行パターンに加え、1 フレーム歩行パターンを入力とした学習、識別を行った。図 4 は、1, 3 および 5 フレーム歩行パターンの場合の結果をまとめたものである。

表 1 5 フレーム歩行パターンを入力とした RNN 識別結果

	正答率	再現率	適合率
A-RNN	0.954	0.900	0.667
B-RNN	1.000	1.000	1.000
C-RNN	0.999	1.000	0.984
D-RNN	0.953	0.800	0.686
E-RNN	0.913	0.900	0.486
F-RNN	0.939	0.867	0.591
G-RNN	0.899	0.917	0.447
H-RNN	0.950	0.950	0.633
I-RNN	0.971	0.900	0.783
J-RNN	0.961	0.933	0.700
K-RNN	0.933	0.900	0.563
L-RNN	0.939	0.883	0.589
平均	0.951	0.913	0.677

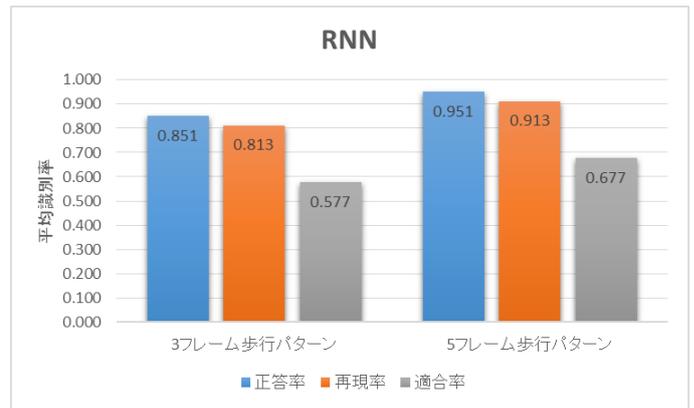


図 3. 各フレームテスト歩行パターン RNN 識別結果

表 2 5 フレームの歩行パターンを入力とした SVM 識別結果

	正答率	再現率	適合率
A-SVM	1.000	1.000	1.000
B-SVM	1.000	1.000	1.000
C-SVM	1.000	1.000	1.000
D-SVM	1.000	1.000	1.000
E-SVM	0.990	0.883	1.000
F-SVM	0.986	0.867	0.871
G-SVM	0.994	0.950	0.986
H-SVM	0.994	0.950	0.986
I-SVM	0.981	0.767	0.900
J-SVM	0.996	0.967	0.986
K-SVM	1.000	1.000	1.000
L-SVM	0.994	0.933	1.000
平均	0.995	0.943	0.977

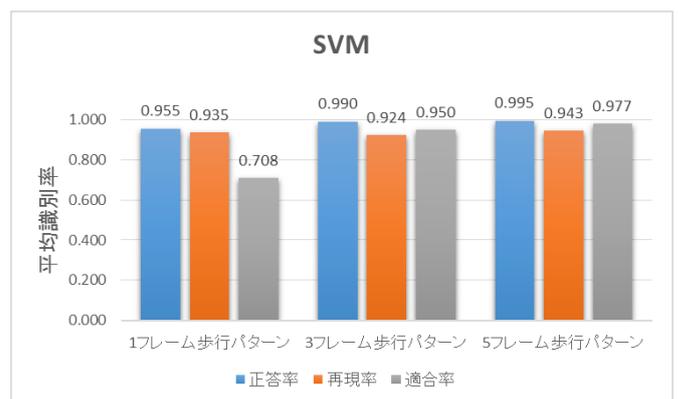


図 4. 各フレームテスト歩行パターン SVM 識別結果

実験 2:

学習には正面からの歩行パターンのみを用い、テストには撮影方向の異なる歩行パターンを用いる。

3名の被撮影者 A, B, C の正面からの歩行パターンで識別器を作成後、各識別器に対して撮影方向の異なる次の歩行パターンを入力した。1つ目は斜め方向から撮影した

テスト歩行パターン, 2つ目は上から見下ろすように撮影した歩行パターン, 3つ目は下から見上げるように撮影した歩行パターンとした. 以上の3種類の歩行パターンをそれぞれの識別器へ入力し, 正答率を求める.

実験2: 結果

正面5フレーム歩行パターンで学習を行った識別器A, B, C-RNNとA, B, C-SVMに対して斜め方向, 上方向, 下方向より撮影したテスト歩行パターンを入力した場合の正答率を表3に示す.

表3. A, B, C-SVM, RNNに対して被撮影者A, B, C斜め, 上, 下からの撮影テスト歩行パターンを入力した際の正答率

	A-SVM正答率	A-RNN正答率
A斜め方向	0.870	0.750
A上方向	0.717	0.650
A下方向	0.783	0.667
	B-SVM正答率	B-RNN正答率
B斜め方向	0.870	0.967
B上方向	0.717	0.917
B下方向	0.870	0.900
	C-SVM正答率	C-RNN正答率
C斜め方向	0.762	0.900
C上方向	0.848	0.717
C下方向	0.930	0.900

4. 考察

(1) SVMによる識別

実験1より平均正答率が99%を超える結果となった. 学習, 識別に扱う歩行パターンのフレーム数を1, 3, 5と増やしていくに従って正答率, 再現率, 適合率, 全ての指標が高くなる結果となった.

また実験1の結果に比べ, 実験2の結果が悪くなったこともわかる. このことからKinectを用いた歩行による人物の識別は撮影方向に依存することが確認できた.

(2) RNNによる識別

SVMによる識別に比べて, 正面からの撮影方向ではどの指標においても劣る結果となった. 特に適合率が70%程度になっていることがわかる. これは非対象者のテスト歩行パターンを入力とした際に識別対象者であると誤識別してしまっていることを示す. しかし再現率についてはSVMと同程度のものや高い値を出しているものがある. これは全ての識別対象者のテスト歩行パターンに対してRNNが高水準で正しく判断できていることを示す. このことから, 未学習のテスト歩行パターンを入力した際, RNNは識別対象者と識別してしまいやすい学習が進んでいると考える. 識別対象者の歩行パターンに対して高い比重を与えて学習させることで, 厳しく識別境界面をフィッティングすれば識別性能が向上するだろう.

5. おわりに

今回の実験では二つの識別器において各人物ごとに学習を行ったものの, パラメータについては全て統一して実験を行った. 各人物ごとに個別にパラメータの設定を行うことで識別率の向上を目指せるはずである.

また複数の静的な特徴量を組み合わせることで識別率の向上を行った研究もある[6]. 本研究では動的な特徴量だけに注目したが, 顔や身長といった静的な身体的特徴を組み合わせることで識別率が向上するだろう.

謝辞: 本研究を進めるにあたり, ご指導を頂いた修士論文指導教員の平原誠准教授に心より感謝いたします.

参考文献

- [1]Aniruddha Sinha, Kingshuk Chakravarty, Brojeshwar Bhowmick: Person Identification using Skeleton Information from Kinect, International Conference on Advances in Computer-Human Interactions, pp.101-108, 2013.
- [2]遠山卓也, 山本正信: Kinectを用いた個人歩行認証, 電子情報通信学会総合大会講演論文集, pp.94, 2014.
- [3]松本圭祐, 真部雄介, 菅原研次: 骨格情報を用いた歩容認証における有効なカメラ配置の検討, 人工知能学会全国大会論文集 Vol.27, pp.1-4, 2013.
- [4]高田憲一, 北須賀 輝明, 有次正義: マーカーレスモーションキャプチャ装置を用いた歩容による個人識別法の検討, 情報処理学会研究報告ユビキタスコンピューティングシステム, Vol.35, No.9, pp.1-7, 2012.
- [5]Mohammed Ahmed, Naseer Al-Jawad Azhin Sabir: Gait recognition based on Kinect sensor, Proc. SPIE 9139, Real-Time Image and Video Processing 2014
- [6]I. Sutskever, J. Martens, G. E Dahl and G. E. Hinton: On the importance of initialization and momentum in deep learning, In Proceeding of the 30th International Conference on Machine Learning (ICML-13), 28, pp.1139-1147, 2013
- [7]熊埜御堂裕太, 骨格情報を用いたリカレントニューラルネットによる歩容認証, 知的システム研究会, 2015.
- [8]熊埜御堂裕太, リカレントニューラルネットを用いた歩行パターンによる個人識別, ファジィシステムシンポジウム講演論文集 Vol131, pp737-740, 2015
- [8]熊埜御堂裕太, 動的特徴を用いた歩容認証-RNN及びSVMの性能比較-, 第43回ファジィ・ワークショップ, 2017(予定)