

18th SICE Symposium on Computational Intelligence

June 17-18, 2021, Online

第18回コンピューターショナル・インテリジェンス研究会

講演論文集

期 日：2021年6月17日(木)，18日(金)

会 場：オンライン開催

SICE[®]

主 催：計測自動制御学会 システム・情報部門

企 画：コンピューターショナル・インテリジェンス部会，知能工学部会

協 賛：システム制御情報学会，日本知能情報ファジィ学会，進化計算学会，電気学会，
情報処理学会，日本神経回路学会，日本機械学会，人工知能学会，ヒューマン
インタフェース学会，電子情報通信学会，IEEE Computational Intelligence
Society Japan Chapter (CISJ)，IEEE Systems, Man, and Cybernetics Society Japan
Chapter

カタログ番号 21PG0005

著作権 © 2021

公益社団法人計測自動制御学会(SICE)
〒101-0052 東京都千代田区神田小川町 1-11-9 金子ビル 4 階

カタログ番号 21 PG 0005

著作権は、計測自動制御学会がもっている
ので、個人の使用のための複製以外の
目的で掲載の記事の一部または全文を
複製する場合には、著作権者に許可を求
め規定の複製料を支払うこと。

発行日：2021年6月17日

発行者：公益社団法人計測自動制御学会 システム・情報部門
コンピューターショナル・インテリジェンス部会，知能工学部会

第 18 回コンピューテーショナル・インテリジェンス研究会プログラム

6月17日(木)

13:05~14:15 (3件) 司会 礪川 悌次郎 (兵庫県立大学)

- 1 加齢における脳波のマルチフラクタル性と機能的神経ネットワーク構造の関係性の分析 1
○安藤桃, 信川創 (千葉工業大学), 菊知充 (金沢大学),
高橋哲也(福井大学, 金沢大学, 魚津神経サナトリウム)
- 2 カオスの同期性に基づく Echo State Network の性能評価 8
○江波戸雄大, 信川創 (千葉工業大学), 西村治彦 (兵庫県立大学)
- 3 カオス-カオス間欠性振動子を用いたメモリ機構の検討 14
○信川創 (千葉工業大学), 我妻伸彦 (東邦大学), 西村治彦 (兵庫県立大学),
稲垣圭一郎 (中部大学), 山西輝也 (福井工業大学)

14:30~16:05 (4件) 司会 森 禎弘 (京都工芸繊維大学)

- 4 デュアルパーティクルフィルタによるドライバモデルパラメータのオンライン推定 18
○鈴木宏典 (日本工業大学), 藤原翔 (日本工業大学大学院)
- 5 骨格の時系列変化情報を用いた歩行人物の行動推定 23
○酒井俊権, 礪川悌次郎, 上浦尚武 (兵庫県立大学)
- 6 VR 環境におけるフリック文字入力に適したフィードバック手法の検討 27
○飯田春樹, 今井順一 (千葉工業大学)
- 7 視線情報に基づく視認推定とその視覚探索タスク支援への応用 32
吉田竣亮 (兵庫県立大学), 施真琴 (ATR, 立命館大学),
内海章 (ATR), ○山添大丈 (兵庫県立大学)

6月18日(金)

10:00~11:15 (3件) 司会 信川 創 (千葉工業大学)

- 8 ファジィ測度を用いたアクティブラーニングによる効率的な分類学習 38
○木村慶豪, 濱上知樹 (横浜国立大学)
- 9 土壌断面画像群に対する画素輝度値の補正法 44
○前川新司, 礪川悌次郎, 木村敏文 (兵庫県立大学), 池野英利 (福知山公立大学),
大橋瑞江, 上浦尚武 (兵庫県立大学)
- 10 深層学習による死後肺 CT 画像を用いた説明可能な溺死鑑別システムに関する研究 49
○坂本奨太 (東北大学), 張曉勇 (仙台高等専門学校),
本間経康, 川住祐介, 白井章仁, 小河原輝正, 舟山真人, 杉田典大 (東北大学)

加齢における脳波のマルチフラクタル性と 機能的神経ネットワーク構造の関係性の分析

○安藤桃 信川創 (千葉工業大学) 菊知充 (金沢大学) 高橋哲也 (福井大学, 金沢大学, 魚津神経サナトリウム)

Analysis of Relationships between Multi-Fractal Characteristics and Functional Connectivity for Electroencephalography in Aging

*M. Ando and S. Nobukawa (Chiba Institute of Technology)

M. Kikuchi (Kanazawa University)

T. Takahashi (University of Fukui, Kanazawa University, Uozu Shinkei Sanatorium)

Abstract— Complexity analysis is a useful approach for alterations of cortical network detecting amputations in senile psychiatric disorders that result with impaired neural activity interactions among a wide range of brain regions. Multi-fractal analysis and multi-scale entropy analysis are typical methods for capturing the complexity of neural activity in the brain. Multi-fractal analysis plays an important role in detecting changes in neural activity; the multi-fractal analysis is superior in to other methods to detect alteration of complexity. However, the relationship between the multi-fractal characteristic of brain activity and functional neural networks has not been fully studied. Therefore, the purpose of this study was to investigate the relationship between the complexity of neural activity and functional connections in aging EEG. As a result, it was found that multi-fractal property can classify aging most.

Key Words: EEG signal, multifractal, phase lag index

1 序論

脳内の神経活動の時間的変動の複雑性は、知覚・思考などの脳情報処理において重要な役割を担っている¹⁾。脳内の神経活動を捉える手法として脳波が挙げられる。脳波は、神経細胞が複雑に結合している神経ネットワークで構成された脳神経系の活動を捉えることができる。神経ネットワークの特徴を捉える手法としては、脳活動内の適度なゆらぎが脳機能を増進させるという確率共鳴の理論^{2, 3, 4)}により、複雑性指標を用いて各部位の脳波のゆらぎのダイナミクスを解析するもの⁵⁾や脳の部位間の同期の程度を定量化した機能的結合のネットワーク構造に着目したもの⁶⁾が挙げられる。脳の複雑性を捉える手法として、古くから決定論的カオスと相関次元やリアプノフ指数などのフラクタル次元に着目した研究が行われてきた^{7, 8)}。

また、システムの同期性と複雑性の間には複雑な関係があると示唆されている。Ibáñez-MolinaらはEEGの複雑性と同期性の間には負の相関があるが、神経ネットワークの同期性が弱い領域においては、複雑性と同期性の間には正の相関があることが示された⁹⁾。このことは、複雑性と同期性の関係性は、神経活動の相互作用が生成する複雑な非線形性を含有することを示唆している^{9, 10)}。

複雑性の分析は広範囲の脳領域間の神経活動の相互作用の障害をもたらす老齢性の精神疾患の皮質切断を検出するための有益なアプローチである¹¹⁾。決定論的カオスと相関次元やリアプノフ指数などのフラクタル次元に着目した研究により老齢性の精神疾患をもつ患者の神経活動の複雑性の低減が報告されている^{7, 12, 13, 14, 15, 16, 17, 18)}。また、EEGダイナミクスは、記憶機能や認知機能、知覚機能などがtheta波やbeta波、gamma波の各バンドの各時間スケールにおいて異

なる役割を果たしている¹⁹⁾。このことから、各バンドや時間スケールにおける複雑性もこれらの脳機能を反映する¹⁰⁾。

複雑性の時間スケール依存性としてマルチフラクタル (Multifractal : MF) 解析とマルチスケールエントロピー解析 (Multiscale entropy : MSE) が挙げられる^{20, 8)}。アルツハイマー病に対する複雑性解析としてはMF解析の方が老齢性の精神疾患を捉えることができるとされており、脳内の神経活動の変化を検出する上で重要な役割を担っている¹¹⁾。マルチフラクタル性は他の手法と比較して複雑性検知に優れているが、脳活動のマルチフラクタル性と機能的な神経ネットワークとの関係は十分に研究されていない。そこで本研究では、加齢の脳波を対象に脳波の複雑性と機能的結合の関係性の検討を行う。

2 解析方法

2.1 被験者

本研究の被験者は18名の若年者(21~35歳)と32名の高齢者(52~67歳)である。高齢者は、非喫煙者であり、投薬を行っていない。また、過去に癲癇または頭部外傷を含む医学的または神経学的状態を有する被験者やアルコールまたは薬物依存の経歴がある被験者は除外した。すべての被験者は、研究に対する説明の後、インフォームドコンセントを得ている。また、本研究は金沢大学の倫理委員会によって承認され、ヘルシンキ宣言に従って実施した。

2.2 脳波計測

脳波データを記録及び前処理するための手法は研究で報告され確立されている²¹⁾。脳波を計測する際は、被験者は防音の録音室に座り、部屋の照明が制御された。脳波計測は、国際10-20法という電極配置から16電極(Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2,

F7, F8, Fz, Pz, T5, T6) を用いた。両耳朶結合を参照結果として脳波活動を計測した。脳波測定には日本光電社の EEG-4518 を用い、眼球運動は、双極心電図 (EOG) を使用して観測した。サンプリング周波数 200Hz と 2.0-60Hz のバンドパスフィルターを使用して記録した。電極/皮膚コンダクタンスインピーダンスは、各電極で 5kΩ 未満に制御した。被験者の脳波信号は、目を閉じた安静状態で 10-15 分間測定した。ビデオ監視システムを使用して、参加者の覚醒を視覚的に検査し目を閉じて覚醒状態のエポックのみが測定されたことを確認した。また、脳波信号および EOG レコードの目視検査により、目を閉じて覚醒状態で記録された脳波信号時系列セグメントが特定された。MSE 解析と MF 解析は、連続する 50 秒 (10000 データポイント) のエポックに対して実行した。

2.3 マルチフラクタル解析

マルチフラクタル解析では、離散ウェーブレット変換の係数から導出されたウェーブレットリーダーが広く使用されている^{22, 23)}。マルチフラクタル解析は Hölder 指数によって、データ X の構造を特徴づけている部分構造のフラクタル次元をスペクトルによって表現する解析方法である。離散信号 $X(t)$ の離散 Wavelet 係数は (1) 式で示される。

$$d_X(j, k) = \int_R X(t) 2^j \psi_0(2^{-j}t - k) dt \quad (1)$$

$$(j = 1, 2, \dots, k = 1, 2, \dots)$$

ここで ψ_0 はコンパクトサポートな mother wavelet 関数とする。1次元の wavelet leaders は (2) 式で表される。

$$L_X(j, k) = \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(j, k)| \quad (2)$$

ここで、 $\lambda = \lambda_{j,k} = [k2^j, (k+1)2^j]$ はスケール 2^j による時間間隔を、 $3\lambda_{j,k-1} = \cup \lambda_{j,k} \cup \lambda_{j,k+1}$ は隣接時間を表す。特異値スペクトルは、wavelet leaders によって (3) 式で示される^{22, 23)}。

$$D(h) = \inf_{q \neq 0} (1 + qh - \zeta_L(q)) \quad (3)$$

ここで、 h と q はそれぞれ Hölder 指数と wavelet leaders のモーメントを示す。スケーリング指数 $\zeta_L(q)$ と構造関数 $S_L(q, j)$ は (4) 式と (5) 式で示される。

$$\zeta_L(q) = \liminf_{j \rightarrow 0} \left(\frac{\log_2 S_L(q, j)}{j} \right) \quad (4)$$

$$S_L(q, j) = \frac{1}{n_j} \sum_{k=1}^{n_j} |L_X(j, k)|^q \quad (5)$$

ここで、 n_j はスケール 2^j の場合の X のサンプル数を示す。本研究では、 $D(h)$ の評価のために、滑らかさの指標として一次キウムラント c_1 を使用し、マルチフラクタル性の指標として二次キウムラント c_2 を使用した。本研究では、マルチフラクタル解析を Wavelet Toolbox of MATLAB (<https://jp.mathworks.com/products/wavelet.html>) によって実施する。

2.4 マルチスケールエントロピー解析

脳波の時系列の複雑性の時間スケールへの依存性の評価を行うためにマルチスケールエントロピー (Multiscale entropy: MSE) 解析を行う。確率変数 $\{x_1, x_2, \dots, x_N\}$ に対して、サンプルエントロピーは (1) 式のように定義される。

$$h(r, m) = -\log \frac{C_{m+1}(r)}{C_m(r)} \quad (6)$$

$C_m(r)$ は $|x_i^m - x_j^m| < r (i \neq j, i, j = 1, 2, \dots)$ の確率を示す。 x_i^m は m 次元ベクトル $x_i^m = (x_i, x_{i+1}, \dots, x_{i+m-1})$ を示す。MSE 解析においては、 x_1, x_2, \dots, x_N を (2) 式で、粗視化した時系列 y_j に対して行う。

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \quad (1 \leq j \leq \frac{N}{\tau}) \quad (7)$$

ここで、 $\tau (\tau = 1, 2, \dots)$ は時間スケールを表す。

2.5 Phase Lag Index 解析

位相同期を測定するために Phase Lag Index (PLI) を求め、同期信号の特性を定量的に推定した。EEG 信号は delta 波 (2~4Hz), theta 波 (4~8Hz), alpha 波 (8~13Hz), beta 波 (13~30Hz), gamma 波 (30~60Hz) といった 5 つの周波数帯域に分割した。時間 t と点 ϕ_a での各帯域分割信号は、ヒルベルト変換を用いて位相 ϕ_a と振幅 $Aa(t)$ で示される。また、時間 t_i で 2 つの異なる点 a と b を持つ信号間で観測された位相の差 $\Delta\phi_{ab}(t_i)$ は (8) 式と (9) で示される²⁴⁾。

$$\Delta\phi_{ab}(t_i) = \phi_a(t_i) - \phi_b(t_i) \quad (8)$$

$$\Delta\phi_{mod}(t_i) = \Delta\phi_{ab}(t_i) \pmod{2\pi} \quad (9)$$

いくつかの信号 T に対する 2 つの観測点 a と b の間の信号の PLI は (10) 式で示される。

$$PLI_{ab} = \left| \frac{1}{T} \sum_{i=0}^T \text{sign}(\Delta\phi_{mod}(t_i)) \right| \quad (10)$$

(8) 式と (9) 式から、 $\Delta\phi_{ab}(t_i)$ が 0 であり、 $\Delta\phi_{mod}(t_i) = 0$ であるため、異なるソースで信号が観測された場合の PLI の値は 0 となる。また、電気双極子の反対側にある点での観測は、(8) 式で $\Delta\phi_{ab}(t_i) = \pi$ となり、信号源は双極子モデルに従うと仮定される。これにより $PLI_{ab} = 0$ になる。(10) 式から、ノード次数 (ND) と呼ばれる他の電極 $b = 1, 2, \dots, K (b \neq a)$ を介した任意の電極 a の平均 PLI は、(11) 式で示される。ここで K は電極の総数を表し、 $K = 16$ である。

$$ND_a = \frac{1}{K-1} \sum_{b=1, b \neq a}^K PLI_{ab} \quad (11)$$

2.6 統計解析

c_1 と c_2 では、若年者と高齢者に有意差があるかを判断するために repeated measure ANOVA (ANOVA) を行った。被験者間要因として若年者と高齢者のグループを使用し、被験者内要因として 16 電極を使用した。そして事後 t 検定を使用して、若年者と高齢者および電極ごとの相互作用の主効果を評価した。ここで、 c_1 と c_2 ($q < 0.05$) での多重比較の t 値に Benjamini-Hochberg の誤検出率補正を適用している。

サンプルエントロピーでも若年者と高齢者に有意差があるかを判断するために ANOVA 解析を行った。被験者間要因として若年者と高齢者のグループを使用し、被験者内要因として 16 電極と時間スケールを使用した。そして事後 t 検定を使用して、若年者と高齢者および電極と時間スケールごとの相互作用の主効果を評価した。

PLI の ND についても若年者と高齢者に有意差があるかを判断するために ANOVA 解析を行った。被験者間要因として若年者と高齢者のグループを使用し、被験者内要因として 16 電極を使用した。そして事後 t 検定を使用して、若年者と高齢者および電極ごとの相互作用の主効果を評価した。

同期性と複雑性の関係の評価するために、 c_1 、 c_2 と PLI の ND の間の相関係数 R を使用した。また、MSE 解析と MF 解析の関係の評価するために、 c_1 、 c_2 とサンプルエントロピーの間の相関係数 R を使用した。MSE 解析と ND の関係の評価するためにもサンプルエントロピーと PLI の ND の間の相関係数 R を使用した。

Receiver operating characteristic (ROC) カーブを使用して高齢者の分類を行った。主成分分析は次元削減の前処理として使用し、 c_1 、 c_2 、サンプルエントロピーと PLI の ND に基づくロジスティック回帰を使用した。ROC カーブは主成分を使用し若年者と高齢者を分類する能力を評価するために使用した。識別の精度は、ROC カーブの下面積 (AUC) によって評価した。また、過学習を防止するために 5 分割交差検証を使用した。この時、AUC=1.0 は完璧な識別に対応し、AUC=0.5 はランダムな識別に対応している。

3 解析結果

3.1 マルチフラクタル解析

本研究では、若年者と高齢者に対してマルチフラクタル解析を行った。若年者と高齢者の c_1 と c_2 の ANOVA 解析結果を Table 1 に示す。結果から、 c_1 と c_2 において高い主効果と Group \times node との相互作用が見られた。事後 t 検定では、F3, Fz, F4, C3, C4, P3, Pz, P4, O1, O2, F7, F8, T6 の 13 か所において c_1 の値が有意に低く、Fp1, Fp2, F3, Fz, F4, C3, C4, Pz, P4, O1, O2, F7, F8, T6 の 14 か所において c_2 の値が有意に高いという結果が見られた。

事後 t 検定の結果を Fig. 1 に示す。

3.2 マルチスケールエントロピー解析

本研究では、若年者と高齢者に対して MSE 解析を行った。若年者と高齢者のサンプルエントロピーの ANOVA 解析結果を Table 2 に示す。結果から、主効果は見られなかったが Group \times scale での相互作用が見られた。事後 t 検定では、時間スケール

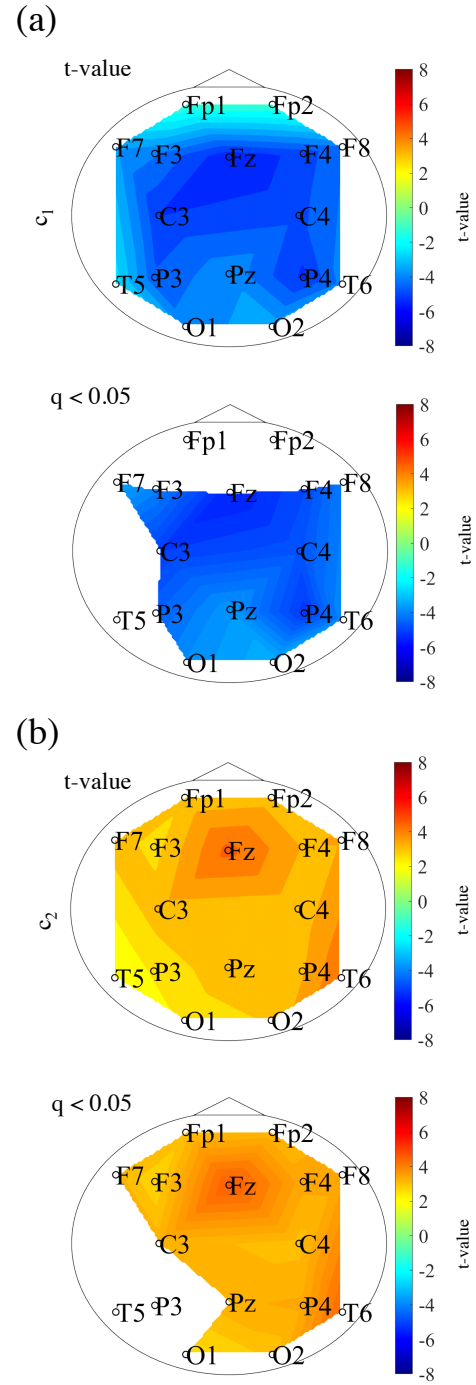


Fig. 1: (a) t values between the younger and older groups. The warm (cold) color represents higher (smaller) c_1 values of older than those for younger group. The left and right correspond to the t -value and t -value satisfying the false discovery rate (FDR) correction criteria $q < 0.050$. c_1 of the older group had significantly lower values at F3, Fz, F4, C3, C4, P3, Pz, P4, O1, O2, F7, F8 and T6. (b) t -value between the older and younger groups warm (cold) color represents higher (smaller) c_2 values of older than those for younger. The left and right correspond to the t -value and t -value satisfying the false discovery rate (FDR) correction criteria $q < 0.050$. c_1 of the older group had significantly higher values at Fp1, Fp2, F3, Fz, F4, C3, C4, Pz, P4, O1, O2, F7, F8 and T6.

Table 1: younger vs. older repeated measure ANOVA analysis results (F value (p value)) in multifractal (MF) analysis results, F and p value with $p < 0.05$ are represented by bold characters.

	Group	Group \times node
c1	$F = 25.25$ ($p = 7.41 \times 10^{-6}$)	$F = 1.73$ ($p = 0.13$)
c2	$F = 22.23$ ($p = 2.11 \times 10^{-5}$)	$F = 1.73$ ($p = 0.11$)

1~5(0.005~0.025[s]) で高齢者のサンプルエントロピーが有意に低いという結果が見られた。事後 t 検定の結果を Fig. 2 に示す。

3.3 PLI 解析

本研究では、若年者と高齢者に対して PLI 解析を行った。若年者と高齢者の各 band の PLI の ANOVA 解析結果を Table 3 に示す。結果から、alpha 帯域においては前頭部分において若年者の方が高く、gamma 帯域においては高齢者の方が高いことがわかった。事後 t 検定の結果、delta 帯域では P4 と C3 のペアにおいて高齢者の方が有意に高く、alpha 帯域では Fp1 と F7, Fp1 と Fz, Fp2 と Fz のペアにおいて高齢者の方が有意に低い、そして gamma 帯域では Fp1 と Fz, Fp2 と Fz, F4 と P4, F4 と Pz のペアにおいて高齢者の方が有意に高いということがわかった。

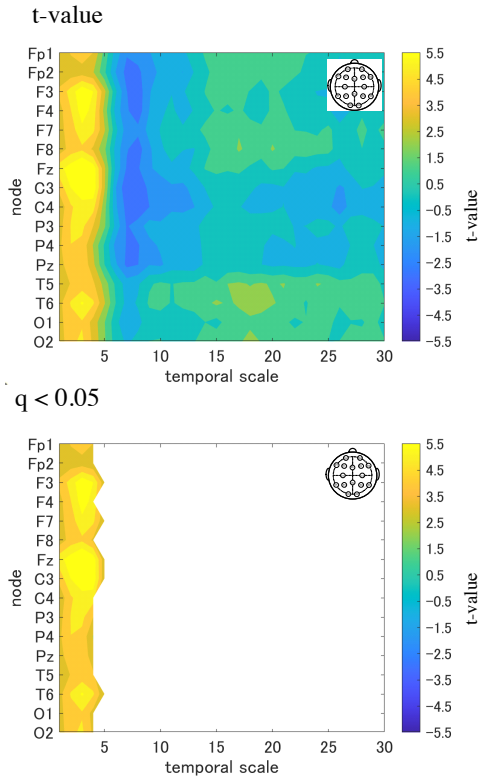


Fig. 2: Multi-scale entropy analysis in younger and older group. The horizontal axis represents the temporal-scale factor, τ . t -value between the older and younger groups (left part). The warm (cold) color represents a higher (smaller) sample entropy value for older than that for younger. The t -value satisfying the FDR correction criteria $q < 0.050$. Significantly smaller sample entropy of low temporal scale regions 1 to 5 (0.005 to 0.025 s).

3.4 相関係数

PLI の ND とサンプルエントロピーの関係を評価するため相関係数 R を求めた。PLI の ND とサンプルエントロピーの間に強い相関は見られなかった。PLI の ND とサンプルエントロピーの相関係数の結果を Fig. 3 に示す。

c_1 , c_2 とサンプルエントロピーの相関関係については、それぞれ速い時間スケールにおいて強い相関関係があり、サンプルエントロピーが上昇すると c_1 は減少し、 c_2 は上昇することがわかった。 c_1 , c_2 とサンプルエントロピーの相関係数の結果を Fig. 4 に示す。また、 c_1 , c_2 と PLI の ND の間の相関関係については、強い相関は見られなかった。

3.5 ROC カーブ

c_1 , c_2 , サンプルエントロピーと PLI の ND の分類能力を評価するために ROC カーブを求めた。サンプルエントロピーは、時間スケール 1~5 の値を平均化している。第一主成分から第三主成分を使用して求めた ROC の結果を Table 4 に示す。 c_1 と ND_{α} の AUC は組み合わせを行わない 8 種類のデータの中で一番大きく AUC=0.86 であった。全てのデータの中で最も AUC が高いのは c_1 と c_2 を組み合わせたデータであり、AUC=0.89 である。また、PLI の ND の中で最も分類精度が高い (AUC=0.86) ND_{α} を組み合わせて AUC を求めたところ分類精度が向上した。しかし、 c_1 と c_2 と ND_{α} を組み合わせた場合は分類精度が下がった。

4 考察

本研究では、加齢の脳波を対象に脳波の複雑性と機能的結合の関係性の検討を行った。マルチフラクタル解析では脳波信号の滑らかさの指標である c_1 が加齢に伴って減少し、マルチフラクタル性の指標である c_2 は加齢に伴って上昇した。マルチスケールエントロピー解析では、高齢者の方が小さい時間スケールでサンプルエントロピーが増加した。PLI 解析では、若年者と高齢者の間に大きな差は見られなかった。相関係数では、 c_1 と PLI の ND, c_2 と PLI の ND, サンプルエントロピーと PLI の ND の間に大きな相関は見られなかった。 c_1 , c_2 とサンプルエントロピーの間には、それぞれ速い時間スケールにおいて強い相関関係があり、サンプルエントロピーが上昇すると c_1 は減少し、 c_2 は上昇することがわかった。分類精度では、 c_1 が一番高い分類精度を示し、 ND_{α} と組み合わせることによって分類精度が向上した。

第一に c_1 , c_2 とサンプルエントロピーの間に強い相関が見られた理由について考える必要がある。 c_1 は時系列信号全体の滑らかさの指標として用いられており、 c_2 は時系列信号の局所的な変動を評価している²⁵⁾。それに対してサンプルエントロピーは時系列信号の滑らかさと局所的な変動を全て同一の尺度により評価している。そのため、サンプルエントロピーが上昇すると

Table 2: AD vs. HC repeated measure ANOVA results (F value (p value)) in multi scale entropy (MSE) analysis results, F and p value with $p < 0.05$ are represented by bold characters.

Group	Group \times node	Group \times scale	Group \times node \times scale
$F = 3.37$ ($p = 0.073$)	$F = 1.51$ ($p = 0.21$)	$F = 24.94$ ($p = 5.8 \times 10^{-9}$)	$F = 2.08$ ($p = 0.034$)

Table 3: Younger vs. older repeated measure ANOVA analysis results (F value (p value)) in Phase Lag Index (PLI) analysis results, F and p value with $p < 0.05$ are represented by bold characters.

	Group	Group \times node
delta	$F = 4.18$ ($p = 0.05$)	$F = 1.26$ ($p = 0.26$)
theta	$F = 0.04$ ($p = 0.83$)	$F = 0.98$ ($p = 0.45$)
alpha	$F = 1.09$ ($p = 0.30$)	$F = 3.95$ ($p = 8.9 \times 10^{-4}$)
beta	$F = 0.10$ ($p = 0.76$)	$F = 2.89$ ($p = 5.8 \times 10^{-3}$)
gamma	$F = 4.17$ ($p = 0.05$)	$F = 3.35$ ($p = 2.1 \times 10^{-3}$)

Table 4: The area under the ROC curve (AUC) for c_1 , c_2 , sample entropy, and PLI ND. In this case, c_1 , c_2 , and PLI, each 1st-3rd principal components was used separately. We evaluated ROC in the case using all these values; the results show the enhancement of classification ability (AUC = 1.00).

	AUC
c_1	0.86
c_2	0.79
sample entropy	0.84
ND _{delta}	0.60
ND _{theta}	0.55
ND _{alpha}	0.86
ND _{beta}	0.74
ND _{gamma}	0.62
c_1 & c_2	0.89
sample entropy & ND _{alpha}	0.84
c_1 & ND _{alpha}	0.87
c_2 & ND _{alpha}	0.86
c_1 & c_2 & ND _{alpha}	0.84

c_1 が減少し, c_2 が上昇するといった c_1 と c_2 に対して逆の相関関係が見られたと考えられる.

第二に PLI の ND の alpha と c_1 , c_2 のそれぞれを組み合わせることによって分類精度が向上する理由について考える必要がある. PLI の ND と c_1 , c_2 の間には相関関係があまり見られなかった. そのため脳波の複雑性と機能的結合は相補的な関係を持っていると考えられる. この相補的な関係が分類精度の向上に寄与している可能性があると考えられる.

第三に PLI の ND の alpha と c_1 , c_2 を全て組み合わせた場合に分類精度が下がる理由について考える必要がある. 本研究では, ロジスティック回帰による学習を行い, AUC を求めた. 学習するデータの特徴量が多く, データセットが小さい場合には過学習が生じる恐れがある. PLI の ND の alpha と c_1 , c_2 を全て組み合わせた場合には, 次元が高くなりすぎてしまい, 本研究のデータセット数では過学習を起こしてしまっている可能性が考えられる. そのため, 過学習の影響により分類精度が下がってしまっていると考えられる. し

たがって今後の研究では, 効率的な次元削減についての検討が必要となる.

5 結論

本研究では, 脳波の複雑性と機能的結合の関係性を明らかにすることを目的として, 若年者と高齢者の脳波を対象にマルチフラクタル解析, マルチスケールエントロピー解析, PLI 解析を行った. 結論として, マルチフラクタル解析とマルチスケールエントロピー解析では加齢を分類することができるが, 時系列信号全体の滑らかさと局所的な変動を別々の指標で捉えるマルチフラクタル解析の方が分類精度が高いことがわかった. また, 脳波の複雑性と機能的結合の間には相補的な関係があり, それらを組み合わせることによって分類精度が向上することがわかった. 本研究の結果は, 脳波の複雑性と機能的結合の相補的な関係が加齢による神経活動の変化を捉える上で補完的な役割を果たすことを示していると考えられる.

参考文献

- 1) Nobukawa, Sou, Haruhiko Nishimura, and Teruya Yamanishi, : Temporal-specific complexity of spiking patterns in spontaneous activity induced by a dual complex network structure, Scientific reports, 9, 1/12 (2019)
- 2) Sinha, Sitabhra, : Noise-free stochastic resonance in simple chaotic systems, Physica A: Statistical Mechanics and its Applications, 270, 1-2, 204/214 (1999)
- 3) Van der Groen, Onno and Tang, Matthew F and Wenderoth, Nicole and Mattingley, Jason B : Stochastic resonance enhances the rate of evidence accumulation during combined brain stimulation and perceptual decision-making, PLoS computational biology, 14, 7, e1006301 (2018)
- 4) Kurita, Yuichi and Sueda, Yamato and Ishikawa, Takaaki and Hattori, Minoru and Sawada, Hiroyuki and Egi, Hiroyuki and Ohdan, Hideki and Ueda, Jun and Tsuji, Toshio, : Surgical grasping forceps with enhanced sensorimotor capability via the stochastic resonance effect, IEEE/ASME Transactions on Mechatronics, 21, 6, 2624/2634 (2016)
- 5) Nobukawa Sou and Yamanishi Teruya and Nishimura Haruhiko and Wada Yuji and Kikuchi Mitsuru and Takahashi Tetsuya : Atypical temporal-scale-specific fractal changes in Alzheimer 's disease EEG and their relevance to cognitive decline, Cognitive neurodynamics, 13-1, 1/11 (2019)
- 6) Takahashi, Tetsuya and Yamanishi, Teruya and Nobukawa, Sou and Kasakawa, Shinya and

- Yoshimura, Yuko and Hiraishi, Hirotohi and Hasegawa, Chiaki and Ikeda, Takashi and Hirose, Tetsu and Munesue, Toshio and others, : Band-specific atypical functional connectivity pattern in childhood autism spectrum disorder, *Clinical Neurophysiology*, **128**, 8, 1457/1465 (2017)
- 7) Kantz Holger and Schreiber Thomas : Nonlinear time series analysis, **7**, (2004)
 - 8) Takahashi Tetsuya : Complexity of spontaneous brain activity in mental disorders, *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, **45**, 258/266 (2013)
 - 9) Ibáñez-Molina, Antonio José and Iglesias-Parro, Sergio and Escudero, Javier, : Differential effects of simulated cortical network lesions on synchrony and EEG complexity, *International journal of neural systems*, **29**, 4, 1850024 (2019)
 - 10) Nobukawa Sou and Yamanishi Teruya and Kasakawa Shinya and Nishimura Haruhiko and Kikuchi Mitsuru and Takahashi Tetsuya : Classification Methods Based on Complexity and Synchronization of Electroencephalography Signals in Alzheimer ' s Disease, *Frontiers in Psychiatry*, **11**, (2020)
 - 11) Ando Momo, and Nobukawa Sou, Kikuchi Mitsuru, Takahashi Tetsuya, : Identification of Electroencephalogram Signals in Alzheimer's Disease by Multifractal and Multiscale Entropy Analysis, (2021) (to be published)
 - 12) Jeong Jaeseung : EEG dynamics in patients with Alzheimer's disease, *Clinical neurophysiology*, **115-7**, 1490/1505 (2004)
 - 13) Woyshville Mark J and Calabrese Joseph R: Quantification of occipital EEG changes in Alzheimer's disease utilizing a new metric: the fractal dimension, *Biological psychiatry*, **35-6**, 381/387 (1994)
 - 14) Besthorn C and Sattel H and Geiger-Kabisch C and Zerfass R and Förstl H : Parameters of EEG dimensional complexity in Alzheimer's disease, *Electroencephalography and clinical neurophysiology*, **95-2**, 84/89 (1995)
 - 15) Jelles B and Van Birgelen JH and Slaets JPJ and Hekster REM and Jonkman EJ and Stam CJ: Decrease of non-linear structure in the EEG of Alzheimer patients compared to healthy controls, *Clinical Neurophysiology*, **110-7**, 1159/1167 (1999)
 - 16) Bluth Charles and Bluth James : Health care information system, (2004)
 - 17) Smits Fenne Margreeth and Porcaro Camillo and Cottone Carlo and Cancelli Andrea and Rossini Paolo Maria and Tecchio Franca : Electroencephalographic fractal dimension in healthy ageing and Alzheimer ' s disease, *PloS one*, **11-2**, e0149587 (2016)
 - 18) Al-nuaimi Ali H and Jammeh Emmanuel and Sun Lingfen and Ifeachor Emmanuel : Higuchi fractal dimension of the electroencephalogram as a biomarker for early detection of Alzheimer's disease, 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2320/2324 (2017)
 - 19) Klimesch Wolfgang and Sauseng Paul and Hanslmayr Simon and Gruber Walter and Freunberger Roman : Event-related phase reorganization may explain evoked neural dynamics, *Neuroscience & Biobehavioral Reviews*, **31-7**, 1003/1016 (2007)
 - 20) Yang Albert C and Tsai Shih-Jen : Is mental illness complex? From behavior to brain, *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, **45**, 253/257 (2013)
 - 21) Mizuno Tomoyuki and Takahashi Tetsuya and Cho Raymond Y and Kikuchi Mitsuru and Murata Tetsuhito and Takahashi Koichi and Wada Yuji : Assessment of EEG dynamical complexity in Alzheimer ' s disease using multiscale entropy, *Clinical Neurophysiology*, **121-9**, 1438/1446 (2010)
 - 22) Wendt Herwig and Abry Patrice: Multifractality tests using bootstrapped wavelet leaders, *IEEE Transactions on Signal Processing*, **55-10**, 4811/4820 (2007)
 - 23) Jaffard, Stéphane and Lashermes, Bruno and Abry, Patrice : Wavelet leaders in multifractal analysis, *Wavelet analysis and applications*, 201/246 (2006)
 - 24) Stam, Cornelis J and Nolte, Guido and Daffertshofer, Andreas : Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources, *Human brain mapping*, **28**, 11, 1178/1193 (2007)
 - 25) Ihlen, Espen Alexander Fürst EAFI : Introduction to multifractal detrended fluctuation analysis in Matlab, *Frontiers in physiology*, **3**, 141 (2012)

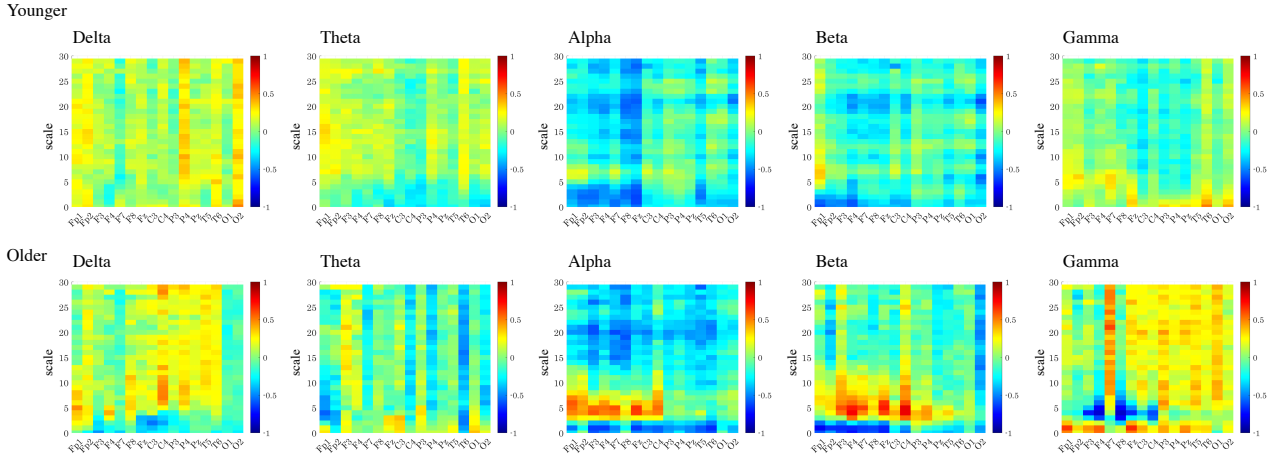


Fig. 3: Correlation coefficient for each band of PLI ND for younger (upper part). Warm (cold) colors indicate a high (small) correlation coefficient. Correlation coefficient for each band of PLI's ND for older. Warm (cold) colors indicate a high (small) correlation coefficient (lower part).

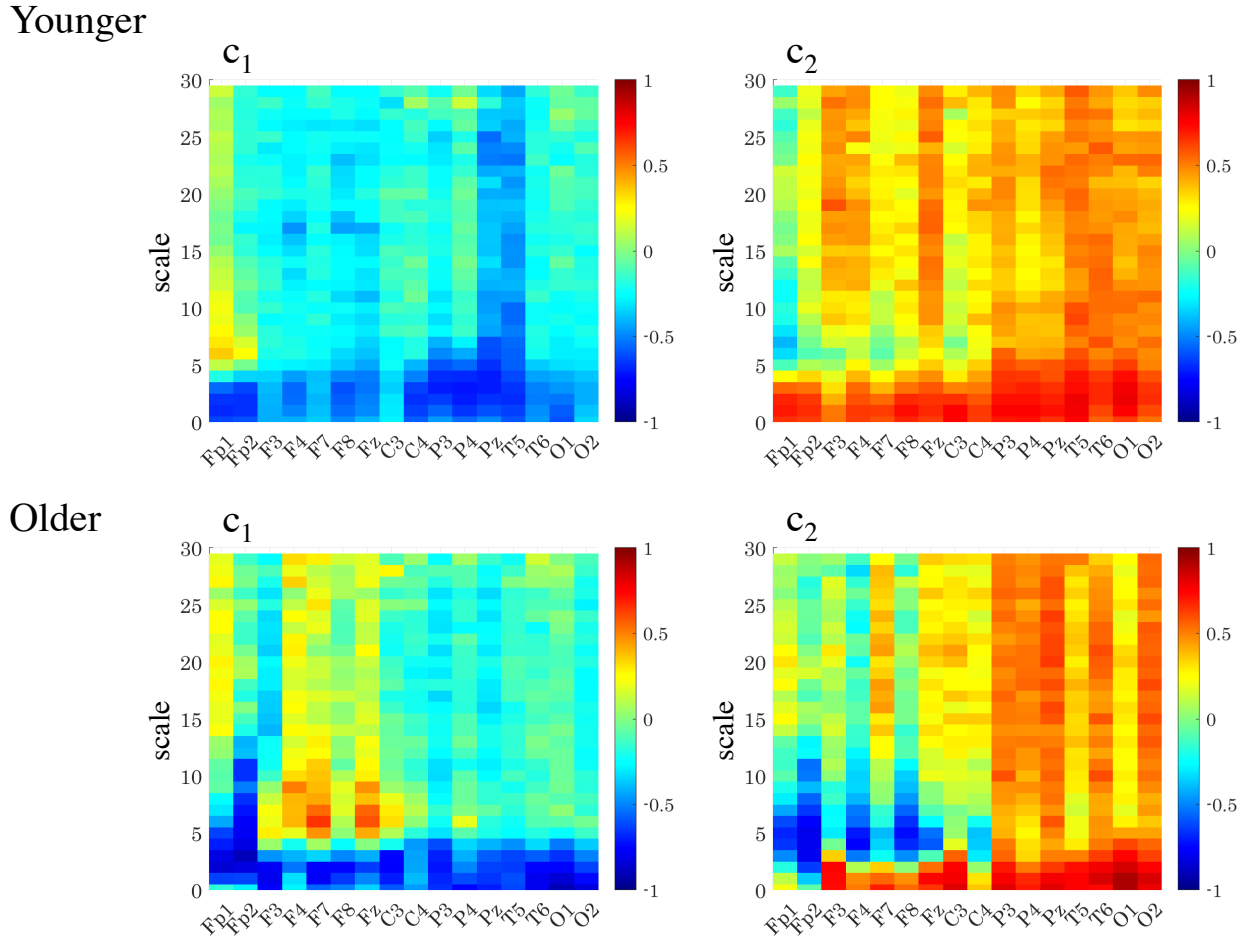


Fig. 4: Correlation coefficient between the sample entropy and c_1 for younger group (upper left parts). The figure on the right shows the correlation coefficient between the sample entropy and c_2 for younger group (upper right parts). Warm (cold) colors indicate a high (small) correlation coefficient. The figure on the left below shows the correlation coefficient between the sample entropy and c_1 for older group (upper left parts). The figure on the right shows the correlation coefficient between the sample entropy and c_2 for older group (upper right parts). Warm (cold) colors indicate a high (small) correlation coefficient.

カオスの同期性に基づく Echo State Network の性能評価

○江波戸雄大 信川創 (千葉工業大学) 西村治彦 (兵庫県立大学)

Evaluation for Echo State Networks by Indexes for Chaotic Synchronization

*Y. Ebato, S. Nobukawa (Chiba Institute of Technology) and H. Nishimura (University of Hyogo)

Abstract– An echo state network (ESN) is a reservoir computing framework consisting of an input layer, a reservoir, and a readout layer. A reservoir is a recursive network comprising neuron models. Time-series learning for ESNs requires a high memory capacity for storing the historical information of past inputs. However, generally used neurons are incapable of storing time history by themselves. Therefore, historical information gained by introducing internal neural dynamics can enhance the memory capacity of ESNs. In this context, we hypothesized that the evaluation of the functions of internal-neural decay factors and optimal balances between the decay factors of chaotic neurons can provide useful results. Therefore, to validate this hypothesis, we investigated the performance of an ESN using a reservoir comprising a chaotic neural network (ChESN). The proposed approach is expected to have wide applications in reservoir computing.

Key Words: echo state networks, transverse Lyapunov exponent, synchronization

1 はじめに

リザーバーコンピューティングは、リザーバーという再起的ネットワークの時空間パターンを時系列データの学習に用いる手法である^{1,2,3,4}). この手法は他の再起的ネットワークを用いる手法^{5,6})と比べて極めて学習効率が高いことが特徴である(レビューは^{7,8,9})を参照のこと). エコステートネットワーク (echo state network:ESN) はリザーバーコンピューティングフレームワークの一つである (ESN の概念図を Fig. 1 に示す). ESN ではリードアウト部のシナプス荷重のみを調整することで学習が可能である; 対照的に, backpropagation through time や real-time recurrent learning などの手法では, 再起的ニューラルネットワークの全てのシナプス荷重の調整が必要になる. この学習に必要なシナプス荷重の削減によって ESN の効率が向上する.

過去数十年にわたって行われてきたこの分野の様々な研究の中で, ESN の設計原理が検討されてきた^{1,10}). リザーバーネットワークの重みの隣接行列の最大固有値で定義されるスペクトル半径が 1.0 に近づくと, 入力時系列の記憶容量が最大になることが知られている^{1,10}). それによって, ESN の学習性能が向上する. このスペクトル半径の最適化に加えて, 河合は最近, リザーバーにスモールワールドトポロジー (脳の典型的なニューラルネットワークのトポロジーにヒントを得たもの) を適用することで, ESN の性能がさらに向上することを示した¹¹).

ネットワークの重みやトポロジー以外に, リザーバーの内部の神経ダイナミクスに関する時間履歴性に注目することも, 多様な神経ダイナミクスを実現し, また ESN の記憶容量を高めるために有効だと考えられる^{12,13,14}). 不応性, 内部フィードバック, 外部入力の変衰因子を持つカオスニューロンモデルは, 多様なニューロン内部ダイナミクスを持つ典型的なニューロンモデルである¹⁵). このニューロンモデルで構成されたカオスニューラルネットワークは, 動的学習・記憶想起や¹⁶), 外部からの微弱な刺激による記憶想起 (カオス共鳴) など^{17,18}), 学習・記憶機能の強化に広く利用されている. さらに最近, 堀尾らはカオスニューロンで構成されたリザーバーを提案した¹³). この ESN は, ニューロン内部

のダイナミクスを利用して, 過去の情報を長期間にわたって記憶することができる. これに対して, 一般的に用いられる, ロジスティック出力関数を利用したアナログニューロンで構成された ESN は, それ自体では入力履歴を記憶することができない¹⁹). この ESN は, 多様なニューロンの内部ダイナミクスを利用することで, スペクトル半径が小さい場合でも高い分類精度を実現する¹³). しかしながら, この ESN では, 不応性, 内部フィードバック, 外部入力の変衰係数が同じ値であるという制約を課している¹³).

そこで我々は, カオスニューラルネットワークを用いた ESN (本研究では ChESN と呼ぶ) において, 変衰係数の個々の機能や変衰係数間の最適なバランスを評価することで, さらに性能が向上するのではないかと仮説を立てた. そこで, この仮説を検証するために, 3つの変衰因子を持つ ESN の非線形時系列予測課題における性能を調べ, その動的特性を最大リアプノフ指数と, 横断的リアプノフ指数, 短期記憶容量の観点から評価した.

2 手法

2.1 エコステートネットワークとカオスニューロン

ESN とその学習手順を説明するために, Fig. 1 を見ていく. リザーバーは N_x 個のニューロンで構成される. $\mathbf{x}(t) \in \mathbb{R}^{N_x} (\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_{N_x}(t)]^T)$ はリザーバーニューロンの発火状態である. $\mathbf{W} \in \mathbb{R}^{N_x \times N_x}$ はリザーバーニューロンの隣接行列である. N_u 次元の時系列データである入力信号 $\mathbf{u}(t) \in \mathbb{R}^{N_u}$ とバイアス b は入力行列 $\mathbf{W}^{\text{in}} \in \mathbb{R}^{N_x \times (N_u+1)}$ に適用される. ここで, 入力 $\mathbf{u}(t)$ とバイアス b はリザーバーの全てのニューロンに適用される. ESN の出力 $\mathbf{y}(t) \in \mathbb{R}^{N_y}$ は出力ニューロンから出力の隣接行列 $\mathbf{W}^{\text{out}} \in \mathbb{R}^{N_y \times (N_x+N_u+1)}$ を用いて計算される. \mathbf{W}^{out} はリザーバーの全てのニューロンと現時点の入力 $\mathbf{u}(t)$, バイアス b の線形変換で定義される. \mathbf{W}^{out} は後に示すように, 教師信号を用いたリッジ回帰で学習する.

具体的な ESN のダイナミクスと学習手法を説明する. リザーバー状態 $\mathbf{x}(t)$ のダイナミクスは (1) 式で表される.

$$\mathbf{x}(t+1) = f(\xi(t+1) + \eta(t+1) + \zeta(t+1)), \quad (1)$$

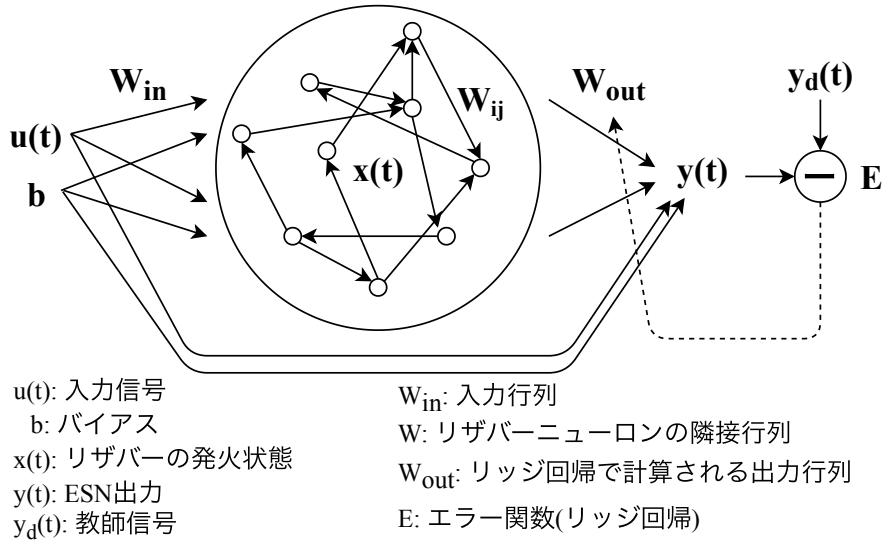


Fig. 1: Overview of the structure of an echo state network (ESN). In the learning process of an ESN, only the synaptic weights in the readout layer are adjusted. Consequently, reductions in the size of the learning synaptic weights lead to high efficiency. All the vectors in this figure are vertical vectors.

$$\xi(t+1) = k_e \xi(t) + \mathbf{W}^{in} \begin{bmatrix} \mathbf{u}(t+1) \\ b \end{bmatrix}, \quad (2)$$

$$\eta(t+1) = k_f \eta(t) + \mathbf{W} \mathbf{x}(t), \quad (3)$$

$$\zeta(t+1) = k_r \zeta(t) - \alpha \mathbf{x}(t) + (1 - k_r) \theta. \quad (4)$$

ここで、 $\xi(t) \in \mathbb{R}^{N_x}$ と $\eta(t) \in \mathbb{R}^{N_x}$ 、 $\zeta(t) \in \mathbb{R}^{N_x}$ は $\mathbf{x}(t)$ と同じように縦ベクトルである。 $k_e (0 \leq k_e < 1)$ と $k_f (0 \leq k_f < 1)$ 、 $k_r (0 \leq k_r < 1)$ はそれぞれ外部入力、フィードバック入力、不応性の減衰係数である。 α は不応性のスケールパラメータである。 θ はしきい値ベクトルを示している。本稿ではこのベクトルの要素は共通の値 θ とする。活性化関数 f はハイパボリックタンジェントである。ここで $k_e = k_f = k_r = k$ という制限を課すと、堀尾らが用いている単純化したカオスニューロンモデルに対応する¹³⁾。 $k_e = k_f = k_r = \alpha = \theta = 0$ という制限では、ロジスティック出力関数を利用したアナログニューロンモデルに対応する。

出力 $\mathbf{y}(t)$ は (5) 式で計算される

$$\mathbf{y}(t) = \mathbf{W}^{out} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \\ b \end{bmatrix}. \quad (5)$$

\mathbf{W}^{out} は学習期間 T の教師信号 $\mathbf{y}_d(t)$ を用いたリッジ回帰で得られる:

$$\mathbf{W}^{out} = ((\mathbf{X}^T \mathbf{X} + \beta \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y}^{teach})^T, \quad (6)$$

$$\mathbf{X} = \left[\begin{array}{c} \left(\begin{array}{c} \mathbf{x}(t+1) \\ \mathbf{u}(t+1) \\ b \end{array} \right) \dots \left(\begin{array}{c} \mathbf{x}(t+T) \\ \mathbf{u}(t+T) \\ b \end{array} \right) \end{array} \right]^T, \quad (7)$$

$$\mathbf{Y}^{teach} = [(\mathbf{y}_d(t+1)) \dots (\mathbf{y}_d(t+T))]^T, \quad (8)$$

ここで $\mathbf{X} \in \mathbb{R}^{T \times (N_x + N_u + 1)}$ は状態収集行列である。 $\mathbf{Y}^{teach} \in \mathbb{R}^{T \times N_y}$ は状態に対応する教師信号の収集行列である。 β は正則化係数である。 \mathbf{I} は単位行列を表している。

2.2 パラメータ設定

我々の実験では、ESNのパラメータを次のように設定した。リザーバーの次元は $N_x = 1000, 100$ 。入力信号の次元は $N_u = 1$ 。出力の次元は $N_y = 1$ 。スカラーであるバイアス b は 1.0。入力行列 \mathbf{W}^{in} は密行列であり、この行列の要素は $[-1, 1]$ の一様分布から抽出した。リザーバー行列 \mathbf{W} は、 $N_x = 1000$ のときエッジ数 50、 $N_x = 100$ のときエッジ数 5 のスパース行列とし (結合率: 0.05)、この行列の要素は 0 を中心とする一様分布から抽出した。ESNの主要なパラメータである $\rho(\mathbf{W})$ は \mathbf{W} をスケールリングすることで設定する。スペクトル半径の値を r とするためには (9) 式で計算する。

$$\mathbf{W} = r \frac{\mathbf{W}_0}{\rho(\mathbf{W}_0)}. \quad (9)$$

ここで、 \mathbf{W}_0 はスケールリングする前の \mathbf{W} であり、 $\rho(\mathbf{W}_0)$ は \mathbf{W}_0 のスペクトル半径である。学習では、最初の 500 タイムステップを burn-in 期間、続く 4000 ステップを学習期間、続く学習期間後の 500 ステップをテスト期間とした。リッジ回帰の正則化係数 β は 0.2 とした。

2.3 非線形時系列予測

時系列予測タスクに用いる時系列は (10) 式の Mackey-Glass 方程式から生成した²⁰⁾。

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{1+x(t-\tau)^{10}} - 0.1x(t). \quad (10)$$

この実験では τ を 32 とした。Mackey-Glass 方程式の解は初期値 0.2 として 4 次のルンゲクッタ法で求めた。求めた時系列データは平均 0、標準偏差 1 にスケールリングし、入力信号とした。このタスクでは、時間 t の入力信号は $\mathbf{u}(t)$ 、教師信号 $\mathbf{y}_d(t)$ は $\mathbf{u}(t+1)$ である。

2.4 評価指標

ESNの性能評価には、非線形時系列予測タスクの normalized root mean squared error (NRMSE)、と短期記

憶容量²¹⁾のタスク性能を用いた。短期記憶容量の性能はリザーバーの記憶性能の指標として用いる。また、リザーバーのダイナミクスの評価指標には最大リアプノフ指数²²⁾と最大横断的リアプノフ指数^{23,24)}を用いた。最大リアプノフ指数はリザーバーダイナミクスの軌道不安定の指標である。最大横断的リアプノフ指数はリザーバーを構成するニューロンの同期不安定性の指標である。

非線形時系列予測のタスクでは NRMSE を性能指標として用いる。

$$\text{NRMSE} = \sqrt{\frac{\sum_{t=1}^T (\mathbf{y}(t) - \mathbf{y}_d(t))^2}{T \sigma^2(\mathbf{y}_d)}}. \quad (11)$$

ここで、 $\mathbf{y}(t)$ は ESN の出力であり、 $\sigma^2(\mathbf{y}_d)$ は教師信号の分散である。

短期記憶容量 (MC) は τ ステップ前の入力と出力の決定係数で定義される。ここで、入力信号 $\mathbf{u}(t)$ は $[-1, 1]$ の一様分布から抽出したランダム値とし、教師信号 $\mathbf{y}_d(t)$ は $\mathbf{u}(t - \tau)$ で与えられる。MC は (13) 式で計算する。

$$\text{MC} = \sum_{\tau=1}^T \text{MC}_\tau, \quad (12)$$

$$\text{MC}_\tau = \frac{\text{cov}^2(\mathbf{u}_\tau, \mathbf{y})}{\sigma^2(\mathbf{u}_\tau) \sigma^2(\mathbf{y})}. \quad (13)$$

ここで、 $\text{cov}(\mathbf{u}_\tau, \mathbf{y})$ は教師信号と ESN の出力の共分散である。 $\sigma^2(\mathbf{u}_\tau)$ と $\sigma^2(\mathbf{y})$ はそれぞれ教師信号、出力の分散である。この実験では T は 50 に設定した。

最大リアプノフ指数 λ は (15) 式で計算する。

$$\lambda_k = \frac{1}{\tau} \ln \left(\frac{\|\delta_k^{(t_i=\tau)}\|}{\|\delta_k^{(t_i=0)}\|} \right), \quad (14)$$

$$\lambda = \frac{1}{K} \sum_{k=1}^K \lambda_k, \quad (15)$$

ここで、 $\delta_k^{(t_i=0)}$ ($k = 1, 2, \dots, K$) は k 番目の初期摂動を表している。ここで、最初の初期摂動 $\delta_k^{(t_i=0)} \in \mathbb{R}^{3N_x}$ は L2 ノルムが δ_0 となるようにスケールしたベクトルである。 $t_i \in [0: \tau]$ は初期摂動の時間発展を示している。 $\delta_k^{(t_i=\tau)}$ は (16) 式で計算する。

$$\delta_k^{(t_i=\tau)} = X'_k(t) - X(t)|_{t=t_0+k\tau}, \quad (16)$$

$$X'_k(t) = X(t) + \delta_k^{(t_i=0)}|_{t=t_0+(k-1)\tau}, \quad (17)$$

$$X(t) = \begin{bmatrix} \xi(t) \\ \eta(t) \\ \zeta(t) \end{bmatrix}. \quad (18)$$

ここで、基準軌道 $X(t) \in \mathbb{R}^{3N_x}$ はリザーバーのカオスニューロンの内部状態の縦のベクトルの連結である。 $X'_k(t) \in \mathbb{R}^{3N_x}$ は k 番目の摂動軌道である。 $k(> 1)$ 番目の初期摂動は (19) 式で計算される。

$$\delta_k^{(t_i=0)} = \frac{\|\delta_1^{(t_i=0)}\|}{\|\delta_{k-1}^{(t_i=\tau)}\|} \delta_{k-1}^{(t_i=\tau)}. \quad (19)$$

この実験では時間発展 τ は 1. K は 1000. δ_0 は 10^{-6} とした。

最大横断的リアプノフ指数 λ_\perp は摂動ベクトルとその時間微分の内積を用いて (20) 式で計算する。

$$\lambda_k^s = \frac{\delta_k^s(\tau_\perp) \cdot \frac{d\delta_k^s(\tau_\perp)}{dt}}{|\delta_k^s(\tau_\perp)|^2} \quad (20)$$

ここで、 \cdot は内積を表している。 $\delta_k^s(\tau_\perp)$ ($k = 1, 2, \dots, K$) は、 k 番目の初期摂動 $\delta_k^s(0)$ の期間 $t_\perp \in [0: \tau_\perp]$ での時間発展を表している。摂動は $t = t_0 + (k-1)\tau_\perp$ で $\delta_k^s(0)$ に初期化する。そして最大横断的リアプノフ指数の K 回の試行で平均化したものは (21) 式で計算する。

$$\lambda_\perp = \frac{1}{\tau_\perp K} \sum_{k=1}^K \lambda_k^s \quad (21)$$

時間発展 τ_\perp は 5. K は最大リアプノフ指数と同じく 1000. $\delta_k^s(0)$ は 10^{-6} とした。

3 結果

3.1 通常の ESN とカオスニューロンを用いた ESN の性能比較

非線形時系列予測のタスクを用いて、ChESN と通常の ESN のスペクトル半径 $\rho(W)$ に対する NRMSE を比較した。その結果が Fig. 2 (a) である。ここで、カオスニューロンパラメータは $k_e = 0, k_f = 0.88, k_r = 0.9, \alpha = 0.9, \theta = 0$ に設定した。この結果から、 $\rho(W) \lesssim 1.0$ で ChESN の性能が通常の ESN より優れていることがわかる。それぞれのモデルの性能の最大値は、ChESN で NRMSE = 0.0023 ($\rho(W) = 1.0$)、通常の ESN で NRMSE = 0.0053 ($\rho(W) = 1.3$) となった。Fig. 2 (b) は ChESN と通常の ESN のスペクトル半径に対する短期記憶容量の比較結果である。この図から、どちらの ESN も MC の最大値は NRMSE が最小となったスペクトル半径 1.0 付近でとっていることがわかる。この図では ChESN と通常の ESN で最大の MC にほとんど差は見られない。そこでより詳しい MC の特徴を調べるために、MC の時間スケール τ 依存 (式 12 で与えられる) を調べた結果が Fig. 3 である。この図から、 $20 \lesssim \tau \lesssim 50$ で、通常の ESN より ChESN の方が MC_τ が高いことがわかる。このことから、ChESN は通常の ESN よりも、より長期の記憶を持つと言える。

3.2 ESN の性能とカオスニューロンの減衰係数の関係

ここでは、ChESN の性能とカオスニューロンの減衰係数、 k_f と k_r の依存性について調べた。この実験では $k_e = 0, \alpha = 0.9, \theta = 0$ とした。これらのパラメータは Fig. 2, 3 でのパラメータと対応する。リザーバーのニューロン数は Fig. 2, 3 より小さい $N_x = 100$ とした。Fig. 4 は $\rho(W) = 1.0$ とし、 k_f ($0 \leq k_f \leq 0.98$) と k_r ($0 \leq k_r \leq 0.98$) を変化させたときの NRMSE である。その結果、NRMSE は $k_f \approx 0.84, k_r \approx 0.88$ の時に最小となっていることがわかる。そしてさらに、 $0.1 \leq \rho(W) \leq 1.4$ における最適な k_f と k_r の値も調べた。Fig. 5 (a), (b), (c), (d) はそれぞれ、各スペクトル半径において、NRMSE が最小となる k_f ($0 \leq k_f \leq 0.98$) と k_r ($0 \leq k_r \leq 0.98$) の組み合わせを用いたときの、NRMSE, k_f と k_r の値、最大リアプノフ指数、最大横断的リアプノフ指数を示している。

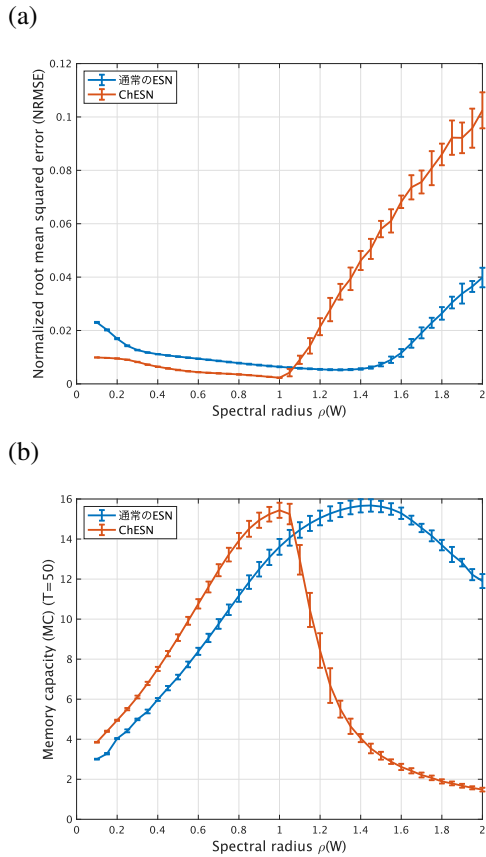


Fig. 2: カオスニューロンを用いた ESN (ChESN) と通常の ESN の、リザーバのスペクトル半径 $\rho(W)$ に対する性能の比較 (a) ESN の出力と教師信号の Normalized root mean squared error (NRMSE) (b) ESN の Memory capacity (MC). 実線と誤差棒はそれぞれ 10 回の試行の平均と標準偏差を示している. $\rho(W) \lesssim 1.0$ で, ChESN の性能は通常の ESN より高くなっている.

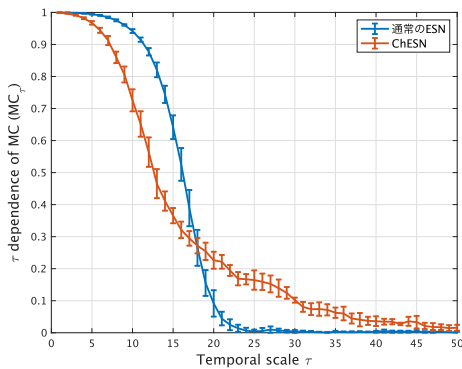


Fig. 3: ChESN と通常の ESN の時間スケール τ 依存. 実線と誤差棒はそれぞれ 10 回の試行の平均と標準偏差を示している. $20 \lesssim \tau \lesssim 50$ で ChESN は通常の ESN より高い性能を示している.

この図から, 適切な k_f と k_r の値は $\rho(W)$ の値に依存して変化し, NRMSE は $\rho(W) \approx 1.0$, $k_f \approx 0.84, k_r \approx 0.88$ で最小となっていることがわかる. また, 各 $\rho(W)$ で k_f と k_r が最適値であるときの最大リアプノフ指数は $\lambda \approx 0$ であり, 概ねカオスの縁となっている. 最大横断的リアプノフ指数 λ_\perp は NRMSE とよく対応しており,

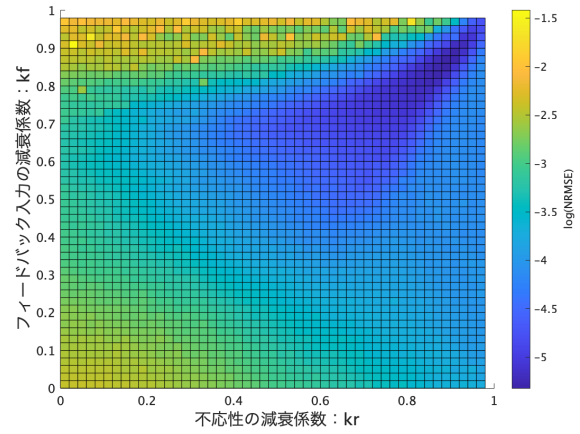


Fig. 4: $\rho(W) = 1.0$ における, k_f と k_r に対する NRMSE. この結果は 10 回の試行の平均である. NRMSE は $k_f \approx 0.84$, $k_r \approx 0.88$ のとき最小となっている.

λ_\perp が最も高いとき NRMSE は最小となっている.

4 考察

ここまでの実験で, ChESN と通常の ESN の性能比較と, ChESN の性能と減衰係数の関係について調べた. まず, 減衰係数をもつカオスニューロンを用いた ChESN と, 減衰係数を持たないアナログニューロンを用いた通常の ESN の非線形時系列予測の性能と短期記憶容量の性能を比較した. その結果から, ChESN は通常の ESN より高い時系列予測性能を持ち, また, 長い時間スケールでの記憶性能が高いことがわかった. このことから, カオスニューロンの減衰係数はリザーバの長期の記憶性能を向上させ, それによって時系列予測タスクの性能が向上したと考えられる. そのため ChESN は時系列予測のような長期の入力の保持が必要なタスクに有効だと考えられる. また, 時系列予測タスクでの最適な減衰係数の組み合わせとスペクトル半径について調べた. その結果, 単体でカオスなダイナミクスをもつカオスニューロンを用いた ChESN でも, 通常の ESN と同じくスペクトル半径 1.0 付近で性能が最大になることがわかった.

次に ChESN のリザーバのダイナミクスについて調べるために最大リアプノフ指数と最大横断的リアプノフ指数を評価した. Fig. 5(c), (d) は各スペクトル半径で減衰係数 k_f, k_r を最適値としたときの最大リアプノフ指数, 最大横断的リアプノフ指数である. Fig. 5(c) を見ると, k_f, k_r が最適値のとき最大リアプノフ指数は 0 に近い値 (カオスの縁) をとっていることがわかる. 通常の ESN ではスペクトル半径 1 付近でカオスの縁となり, 性能が最大になることが知られている²⁵⁾. 対して ChESN は減衰係数を持つことによってスペクトル半径 1 未満でもカオスの縁となり, 多様なダイナミクスを表現できる. このことは Fig. 2 のスペクトル半径 1 未満で ChESN の性能が通常の ESN より高くなったことの一つの説明になると考えられる. Fig. 5(d) は各スペクトル半径で k_f, k_r が最適値のときの最大横断的リアプノフ指数である. この最大横断的リアプノフは Fig. 5(a) の NRMSE とよく対応している. 最大リアプノフ指数ではこのような対応は見られないため, 最大横断

的リアプノフ指数は ESN の性能を評価するための新たな指標になると考えられる。このような対応が見られるのは、同期不安定がダイナミクスの表現力の高さを示しているからだと考えられる。リザーバーを構成するニューロンは、スペクトル半径が小さすぎる場合、入力信号によって同期してしまう。また、スペクトル半径が高すぎる場合、ニューロン同士の結合によって同期してしまい、これらの同期によってダイナミクスの表現力が下がると解釈できる。

今後の課題としては、スペクトル半径によって変化する k_f , k_r の最適値がどのように決まるのか調べること、最大横断的リアプノフ指数と ESN の性能の関係をもっと詳しく調べるなどがある。そのために各スペクトル半径で k_f , k_r を独立に変化させたときの各性能指標の変化などを調べる予定である。

5 結論

本稿では、ニューロン内部ダイナミクスに減衰係数をもつカオスニューロンを用いることによって ESN の性能が向上すること、ESN の性能評価に横断的リアプノフ指数が有効である可能性を示した。ESN にカオスニューロンを用いることはスペクトル半径に対する性能のロバストさや、長期の入力の保持が必要なタスクの性能の向上に役立つ一方で、調整するパラメータ数が増えることが欠点である。今後は、今回示した横断的リアプノフ指数などの性能評価を用いて、適切なカオスニューロンパラメータ設定の指針を明らかにする必要がある。

謝辞

本研究の一部は、JSPS 科研費 基盤研究 (C) (研究課題番号 20K11976) の助成を受けている (H.N.).

参考文献

- 1) Herbert Jaeger. The “echo state” approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, Vol. 148, No. 34, p. 13, 2001.
- 2) Wolfgang Maass, Thomas Natschläger, and Henry Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, Vol. 14, No. 11, pp. 2531–2560, 2002.
- 3) Grzegorz M Wojcik and Wieslaw A Kaminski. Liquid state machine and its separation ability as function of electrical parameters of cell. *Neurocomputing*, Vol. 70, No. 13-15, pp. 2593–2597, 2007.
- 4) Sou Nobukawa, Haruhiko Nishimura, and Teruya Yamanishi. Pattern classification by spiking neural networks combining self-organized and reward-related spike-timing-dependent plasticity. *Journal of Artificial Intelligence and Soft Computing Research*, Vol. 9, , 2019.
- 5) Paul J Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, Vol. 78, No. 10, pp. 1550–1560, 1990.
- 6) Ronald J Williams and David Zipser. A learning algorithm for continually running fully recurrent neural networks. *Neural computation*, Vol. 1, No. 2, pp. 270–280, 1989.
- 7) Mantas Lukoševičius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, Vol. 3, No. 3, pp. 127–149, 2009.
- 8) Mantas Lukoševičius, Herbert Jaeger, and Benjamin Schrauwen. Reservoir computing trends. *KI-Künstliche Intelligenz*, Vol. 26, No. 4, pp. 365–371, 2012.

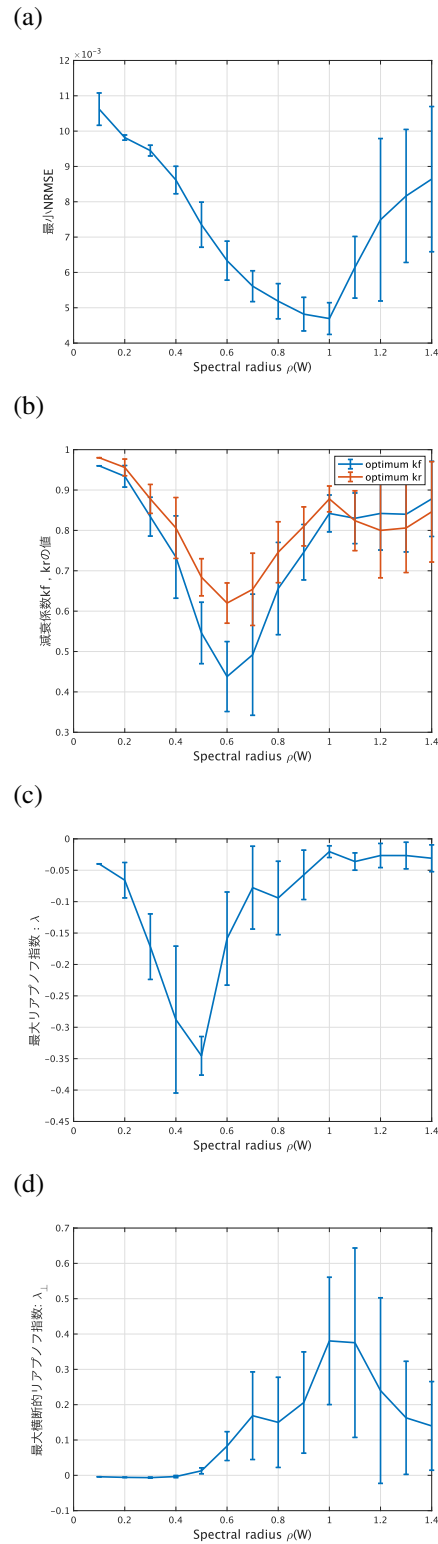


Fig. 5: ChESN の性能とカオスニューロンの減衰係数 k_f と k_r の関係. (a) k_f ($0 \leq k_f \leq 0.98$) と k_r ($0 \leq k_r \leq 0.98$) の組み合わせでの最小 NRMSE. (b) Fig. (a) の最小 NRMSE に対応する k_f と k_r . (c) Fig. (a) の最小 NRMSE に対応する各スペクトル半径 $\rho(W)$ での最大リアプノフ指数. (d) Fig. (a) の最小 NRMSE に対応する各スペクトル半径 $\rho(W)$ での最大横断的リアプノフ指数. 実線と誤差棒はそれぞれ 10 回の試行の平均と標準偏差を示している。

- 9) Gouhei Tanaka, Toshiyuki Yamane, Jean Benoit Héroux, Ryosho Nakane, Naoki Kanazawa, Seiji Takeda, Hidetoshi Numata, Daiju Nakano, and Akira Hirose. Recent advances in physical reservoir computing: A review. *Neural Networks*, Vol. 115, pp. 100–123, 2019.
- 10) Izzet B Yildiz, Herbert Jaeger, and Stefan J Kiebel. Revisiting the echo state property. *Neural networks*, Vol. 35, pp. 1–9, 2012.
- 11) Yuji Kawai, Jihoon Park, and Minoru Asada. A small-world topology enhances the echo state property and signal propagation in reservoir computing. *Neural Networks*, Vol. 112, pp. 15–23, 2019.
- 12) Herbert Jaeger, Mantas Lukoševičius, Dan Popovici, and Udo Siewert. Optimization and applications of echo state networks with leaky-integrator neurons. *Neural networks*, Vol. 20, No. 3, pp. 335–352, 2007.
- 13) Yoshihiko Horio. Chaotic neural network reservoir. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–5. IEEE, 2019.
- 14) Shu-xian Lun, Xian-shuang Yao, and Hai-feng Hu. A new echo state network with variable memory length. *Information Sciences*, Vol. 370, pp. 103–119, 2016.
- 15) Kazuyuki Aihara, T Takabe, and Masashi Toyoda. Chaotic neural networks. *Physics letters A*, Vol. 144, No. 6-7, pp. 333–340, 1990.
- 16) Haruhiko Nishimura, Naofumi Katada, and Yoshihito Fujita. Dynamic learning and retrieving scheme based on chaotic neuron model. In *Complexity and diversity*, pp. 64–66. Springer, 1997.
- 17) Haruhiko Nishimura, Naofumi Katada, and Kazuyuki Aihara. Coherent response in a chaotic neural network. *Neural Processing Letters*, Vol. 12, No. 1, pp. 49–58, 2000.
- 18) Sou Nobukawa, Haruhiko Nishimura, and Teruya Yamanishi. Evaluation of chaotic resonance by lyapunov exponent in attractor-merging type systems. In *International Conference on Neural Information Processing*, pp. 430–437. Springer, 2016.
- 19) David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, Vol. 323, No. 6088, pp. 533–536, 1986.
- 20) Michael C Mackey and Leon Glass. Oscillation and chaos in physiological control systems. *Science*, Vol. 197, No. 4300, pp. 287–289, 1977.
- 21) Herbert Jaeger. Short term memory in echo state networks. *GMD Technical Report*, Vol. 152, , 2002.
- 22) Thomas S Parker and Leon Chua. *Practical numerical algorithms for chaotic systems*. Springer Science & Business Media, 2012.
- 23) Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge Nonlinear Science Series. Cambridge University Press, 2001.
- 24) Sou Nobukawa, Haruhiko Nishimura, Teruya Yamanishi, and Hirotaka Doho. Induced synchronization of chaos-chaos intermittency maintaining asynchronous state of chaotic orbits by external feedback signals. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. 102, No. 3, pp. 524–531, 2019.
- 25) Joschka Boedecker, Oliver Obst, Joseph T Lizier, N Michael Mayer, and Minoru Asada. Information processing in echo state networks at the edge of chaos. *Theory in Biosciences*, Vol. 131, No. 3, pp. 205–213, 2012.

カオス-カオス間欠性振動子を用いたメモリ機構の検討

○信川創 (千葉工業大学) 我妻 伸彦 (東邦大学) 西村治彦 (兵庫県立大学)
稲垣圭一郎 (中部大学) 山西輝也 (福井工業大学)

Memory Mechanism utilizing Oscillations with Chaos-Chaos Intermittency

*S. Nobukawa, (Chiba Institute of Technology), N. Wagatsuma (Toho University),
H. Nishimura (University of Hyogo), K. Inagaki (Chubu University),
and T. Yamanishi (Fukui University of Technology)

Abstract— Recent progress for application of stochastic resonance showed that the presence of additive noise enhances the memory functions in the memory elements with bi-stable oscillations even under the extremely low power consumption. In addition to additive noise, the deterministic chaos induces a similar phenomenon like stochastic resonance, which is known as chaotic resonance. Chaotic resonance emerges in the non-linear dynamical systems with chaos-chaos intermittency where the chaotic orbit moves among separated attractor regions through attractor-merging bifurcation. In previous studies, higher sensitivity of chaotic resonance than stochastic resonance was reported. In this context, we hypothesized that memory devices based on chaotic resonance would realize a novel device for storing memory with lower power consumption than the case based on stochastic resonance. In this study, to prove this hypothesis, we proposed a memory system through attractor-merging bifurcation by approach adjusting internal system parameter under noise-free as chaotic resonance and one for applying stochastic noise like stochastic resonance. Through the comparison of ability between them, we revealed that the former exhibits the higher memory storing ability than latter even if weaker memory storing input signals. The approach by chaotic resonance might facilitate the development memory devices that was previously limited to the application of stochastic resonance.

Key Words: Chaotic resonance, feedback control, noise, synchronization.

1 はじめに

近年の人工知能やニューロモルフィック・コンピューティング、ビッグデータ解析の普及に伴い、世界で保存されるデータ量は飛躍的に上昇している¹⁾。このような大量のデータを扱うために、相変化メモリ^{2, 3, 4, 5)}などの高密度で高速なデータ転送能力を備えたメモリデバイスの開発が進んでいる⁶⁾。さらにこのような高性能メモリデバイスには、低炭素社会の達成に向けた低消費電力化も同時に達成されなければならない^{7, 8, 9)}。

このような低消費電力なメモリデバイスを実現する機構の一つとして、確率共鳴¹⁰⁾ (レビュー^{11, 12, 13)}を参照のこと)の機構を利用したものが提案されている^{14, 15)}。特に、IbáñezらとStotland & Di Ventraは、双安定な振動子により構成されたメモリデバイスが極めて低消費電力な状況下において、確率共鳴の効果で機能し得ることを示した^{14, 15)}。メモリーデバイスにおいては、各素子は‘0’/‘1’に対応した双安定状態を持っている。メモリの保存前には、これらの素子は‘0’に初期化されており、保存段階では、特定の素子を‘1’に遷移させる。確率共鳴機構を利用したメモリデバイスでは、この状態遷移に必要な外部信号強度が背景ノイズを利用することで低く抑えられる^{14, 15, 16)}。

付加的な確率的ノイズに加えて、システムの内的ゆらぎである決定論的カオスも確率共鳴と類似の現象を引き起こし、カオス共鳴と呼ばれている^{17, 18, 19, 20, 21, 22, 23, 24, 25)} (レビュー^{12, 13, 26)}を参照のこと)。カオス共鳴は、カオス-カオス間欠性と呼ばれる複数のカオスアトラクタ間を動的に遷移するカオス軌道を持つシステムで観測され、特にカオス共鳴の信号応答性はカオス-カオス間欠性を生成するアトラクタ併合分岐付近で最大化することが知られている (レビュー^{12, 26)}を参照のこと)。このカオス共鳴の弱入力に対す

る信号感度は、確率共鳴と比較して高いことが報告されているが^{27, 23)}、これまでに定常な周期入力に対する応答性の評価がほとんどである^{28, 29, 30, 31, 22, 23, 32)}。よって、メモリデバイスにおいて現れる‘0’から‘1’状態への遷移のような過渡的なカオス共鳴の応答特性については不明な点が多い。

このような中で、我々はカオス共鳴の機構を利用することで、確率共鳴の場合よりも低い消費電力の状況であっても高い記憶能力を有するメモリデバイスが実現できるという仮説を立てた。この仮説を検証するために、我々の先行研究³³⁾で、アトラクタ併合によってカオス共鳴を誘起する最も単純なシステムである cubic 離散写像において¹²⁾、メモリの保存機構が実現することを示した。しかし、このメモリ保存機構の性能である状態遷移に必要な最低信号強度の確率共鳴との比較は未だ行われていない。さらに、カオス共鳴の背景ノイズに対するロバスト性の検証も必要となる。そこで、本研究では我々の先行研究に基づき³³⁾、前述の点について評価を行う。

2 手法

2.1 モデル

cubic 離散写像は、カオス-カオス間欠性を持つ最も単純なモデルであり、カオス共鳴の評価に用いられる (レビュー¹²⁾を参照のこと)。本研究では、状態遷移を促す $c_i(t)$ ($i = 1, 2, \dots, N$) の信号を持つ N 個の cubic 離散写像によって構成されるシステム:

$$x_i(t+1) = F(x_i(t)) - c_i(t) + D_a \xi(t), \quad (1)$$

$$F(x) = (ax - x^3) \exp(-x^2/b), \quad (2)$$

を扱う。ここで、 $x_i(t)$ は i 番目の cubic 写像の時系列を表しており、 a はカオス-カオス間欠性を調整する内

部パラメータで、パラメータ b については $b = 10.0$ に設定する¹²⁾。 D_a は付加的なガウス型白色ノイズ $\xi(t)$ の強度を示す ($\xi(t)$ の平均値と標準偏差はそれぞれ、0 と 1.0)。アトラクタ併合分岐をパラメータ a で制御する場合は、 a 値を調整する。一方、アトラクタ併合分岐を付加的ノイズで調整する場合は、 a 値をカオス-カオス間欠性が生じない値 ($a = 2.8$) に設定する。

メモリの保存の前段階 ($0 < t < t_s$: t_s と t_e はメモリ保存の開始と終了時刻を表す。) では $c_i(t)$ は 0 に設定されている。 $x_i(0)$ の初期値は負に設定される。メモリ保存段階では $c_i(t) = A\eta_i$ の信号が入力される。ここで η_i は、‘0’ と ‘1’ のメモリパターンであり、‘1’ の生起確率 Λ のポアソン過程により生成される。そして、メモリ保存終了後には $c_i(t) = 0$ となる。ここで、保存するメモリパターンを $\{c_1, c_2, \dots, c_N\}$ とする。一定レベル以上の正の $c_i(t)$ の信号が入力されると、 $x_i(t)$ は $x_i(t) > 0$ の領域に遷移する。本研究では、メモリ保存時間に関連する t_s と t_e はそれぞれ、100 と 200 に設定し、ポアソン過程の Λ は 0.1 に設定する。

2.2 評価指標

アトラクタ併合分岐の判定には、付加的ノイズと入力信号がない状況下 ($D_a = 0, A = 0$) におけるアトラクタ併合の条件として、 $F(f_{\max, \min}) - x_d$ を用いる。ここで、 x_d は cubic 離散写像のアトラクタ分離点であり $x_d = 0$ に設定する。 $f_{\max, \min}$ は $x = x_d$ 近傍の極大・極小値を示す。 $F(f_{\max}) - x_d < 0, F(f_{\min}) - x_d > 0$ は、 $x_i > x_d$ と $x_i < x_d$ 領域間を軌道が遷移するカオス-カオス間欠性を示す³¹⁾。一方、 $F(f_{\max}) - x_d > 0, F(f_{\min}) - x_d < 0$ では、 x_i の初期値に従って、 $x_i > x_d$ か $x_i < x_d$ の領域にトラップされる。

メモリ保存性能の評価には、 η_i と保存されたパターンとして $x_i \geq 0$ ならば 1, $x_i < 0$ ならば 0 としたビット列 X_i との間のビットエラー比 (BER) で評価する。

3 結果

ノイズ・フリーな環境下で、内部パラメータ a を調整しアトラクタ併合を制御した場合の BER に対する信号強度 A と内部パラメータ a の依存性を Fig. 1 に示す。アトラクタ併合 ($F(f_{\max, \min}) - x_d = 0$) が生じる $a = 2.839$ 付近で $A \gtrsim 4.0 \times 10^{-3}$ の領域で $\text{BER} \approx 2.0 \times 10^{-3}$ が達成される様子が確認できる。さらにこの性能のノイズに対するロバスト性の評価として、内部パラメータ $a = 2.839$ の場合の BER の入力信号強度 A とノイズ強度 D_a への依存性を Fig. 2 に示す。ノイズ強度 D_a が増加すると、 $\text{BER} \approx 2.0 \times 10^{-3}$ 程度の低い BER を達成する最小の入力信号強度 A は増加し、最終的に、 $D_a \gtrsim 2.0 \times 10^{-4}$ の領域では $\text{BER} \approx 2.0 \times 10^{-3}$ を満足する $A - D_a$ の領域は消失する。

次に、内部パラメータ a をアトラクタ分離の状態に設定し ($a = 2.8$)、付加的ノイズによるアトラクタ併合制御の場合の BER を評価する。 Fig.3 (a) は、BER の入力信号強度 A とノイズ強度 D_a を示している。アトラクタ併合が生じる $D_a = 7.0 \times 10^{-3}$ 付近で、 $A \gtrsim 0.03$ の信号強度で $\text{BER} \lesssim 2.0 \times 10^{-2}$ が達成される様子が確認できる。さらにこれらの BER 値に基づき、 $10^{-4} \leq D_a \leq 10^{-2}$ の設定で、最小の BER 値と対応する D_a 値を Fig.3 (b) に示す。この結果から、ノイズ・フリーな条件下で $a = 2.839$ のアトラクタ併合付近における $A = 10^{-3}$ の

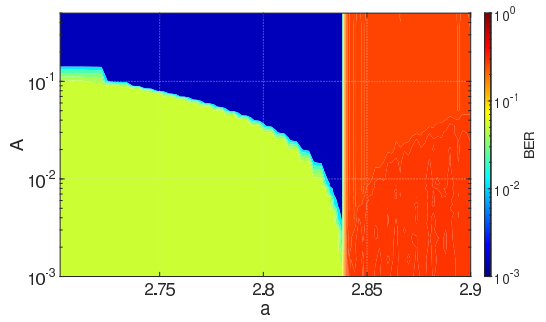


Fig. 1: Dependency of BER on the input signal strength A and the internal parameter a . Around $a = 2.839$ where attractor-merging bifurcation arises, in $A \gtrsim 4.0 \times 10^{-3}$, $\text{BER} \approx 2.0 \times 10^{-3}$ is achieved. ($D = 0, b = 10$)

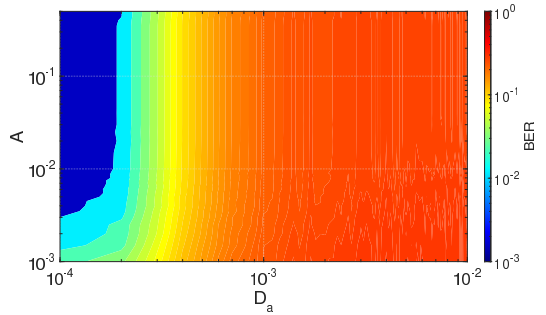


Fig. 2: Dependency of BER on the input signal strength A and the strength of additive stochastic noise D_a at $a = 2.839$ where BER minimizes under noise-free condition. Increasing the noise strength D_a , the lower limit of input signal strength A where $\text{BER} \approx 2.0 \times 10^{-3}$ achieves increases.

信号強度で達成された $\text{BER} \approx 4.0 \times 10^{-3}$ と同程度の BER を付加的ノイズで達成するには、 $A \gtrsim 0.4$ の信号強度が必要であることがわかる。

4 おわりに

本研究では、カオス-カオス間欠性を有する cubic 離散写像の集合において、アトラクタ併合分岐付近の高い信号感度を利用したメモリ保存機構を提案した。アトラクタ併合を制御する手法としては、cubic 離散写像の内部パラメータ a により制御する手法と従来の確率共鳴と同様に付加的な確率ノイズの印加による手法を用いて、それらの性能の比較を行った。その結果、前者の制御手法では、低消費電力に結びつくより微弱な入力信号によるメモリ保存が達成された。さらにその性能は、一定レベルのノイズの存在下においても維持されることが確認された。今後の課題としては、Chua 回路などのカオス-カオス間欠性を示す非線形回路を利用した本メモリ機構の実装とその性能評価が挙げられる。

謝辞

本研究の一部は、JSPS 科研費 基盤研究 (C)(18K11450)(TY)、基盤研究 (C)(20K11976)(NH) の助成を受けている。

参考文献

- 1) C. S. Hwang, “Prospective of semiconductor memory devices: from memory system to materials,” *Ad-*

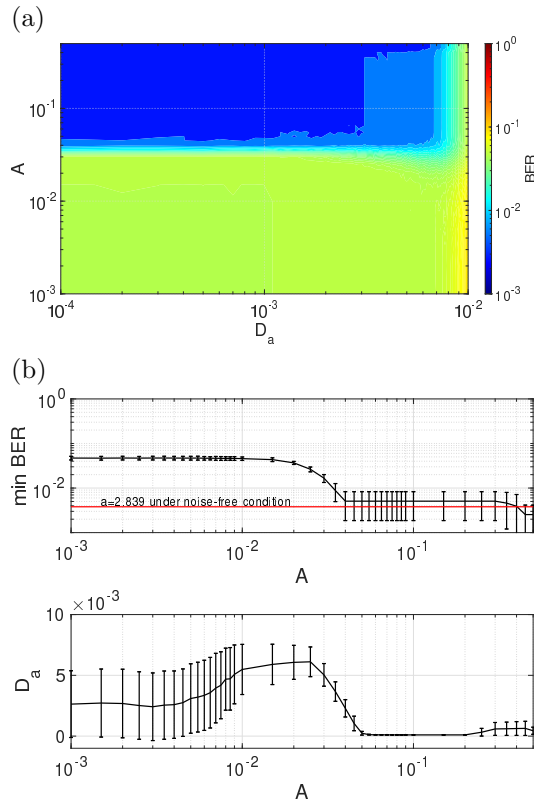


Fig. 3: (a) Dependency of BER on the input signal strength A and the strength of additive stochastic noise D_a . Around $D_a = 7.0 \times 10^{-3}$ where attractor-merging bifurcation arises, in $A \gtrsim 0.03$, $\text{BER} \approx 2.0 \times 10^{-2}$ is achieved. (b) Input signal strength A dependency of minimum BER in $10^{-4} \leq D_a \leq 10^{-2}$ (upper part). A dependency of appropriate strength of additive stochastic noise D_a (lower part). ($a = 2.8, b = 10$)

vanced Electronic Materials, vol. 1, no. 6, p. 140056, 2015.

- 2) X.-B. Li, N.-K. Chen, X.-P. Wang, and H.-B. Sun, "Phase-change superlattice materials toward low power consumption and high density data storage: Microscopic picture, working principles, and optimization," *Advanced Functional Materials*, vol. 28, no. 44, p. 1803380, 2018.
- 3) W. Zhang, R. Mazzarello, M. Wuttig, and E. Ma, "Designing crystallization in phase-change materials for universal memory and neuro-inspired computing," *Nature Reviews Materials*, vol. 4, no. 3, pp. 150–168, 2019.
- 4) A. Lotnyk, M. Behrens, and B. Rauschenbach, "Phase change thin films for non-volatile memory applications," *Nanoscale Advances*, vol. 1, no. 10, pp. 3836–3857, 2019.
- 5) J. Zhu, T. Zhang, Y. Yang, and R. Huang, "A comprehensive review on emerging artificial neuromorphic devices," *Applied Physics Reviews*, vol. 7, no. 1, p. 011312, 2020.
- 6) Y. Frégnac, "Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain?" *Science*, vol. 358, no. 6362, pp. 470–477, 2017.
- 7) S. K. Garg, C. S. Yeo, and R. Buyya, "Green cloud framework for improving carbon efficiency of clouds," in *European Conference on Parallel Processing*. Springer, 2011, pp. 491–502.

- 8) B. Saha, "Green computing: Current research trends," *International Journal of Computer Sciences and Engineering*, vol. 6, no. 3, pp. 467–469, 2018.
- 9) C. Jin, X. Bai, C. Yang, W. Mao, and X. Xu, "A review of power consumption models of servers in data centers," *Applied Energy*, vol. 265, p. 114806, 2020.
- 10) R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *Journal of Physics A: mathematical and general*, vol. 14, no. 11, p. L453, 1981.
- 11) A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: a universal concept in nonlinear sciences*. Cambridge university press, 2003, vol. 12.
- 12) V. S. Anishchenko, V. Astakhov, A. Neiman, T. Vadivasova, and L. Schimansky-Geier, *Nonlinear dynamics of chaotic and stochastic systems: tutorial and modern developments*. Springer Science & Business Media, 2007.
- 13) S. Rajasekar and M. A. F. Sanjuán, *Nonlinear resonances*. Springer, 2016.
- 14) S. Ibáñez, P. Fierens, R. Perazzo, G. Patterson, and D. Grosz, "On the dynamics of a single-bit stochastic-resonance memory device," *The European Physical Journal B*, vol. 76, no. 1, pp. 49–55, 2010.
- 15) A. Stotland and M. Di Ventra, "Stochastic memory: Memory enhancement due to noise," *Physical Review E*, vol. 85, no. 1, p. 011116, 2012.
- 16) V. Kohar and S. Sinha, "Noise-assisted morphing of memory and logic function," *Physics Letters A*, vol. 376, no. 8-9, pp. 957–962, 2012.
- 17) T. Carroll and L. Pecora, "Stochastic resonance and crises," *Physical review letters*, vol. 70, no. 5, p. 576, 1993.
- 18) —, "Stochastic resonance as a crisis in a period-doubled circuit," *Physical Review E*, vol. 47, no. 6, p. 3941, 1993.
- 19) S. Sinha and B. K. Chakrabarti, "Deterministic stochastic resonance in a piecewise linear chaotic map," *Physical Review E*, vol. 58, no. 6, p. 8009, 1998.
- 20) S. Sinha, "Noise-free stochastic resonance in simple chaotic systems," *Physica A: Statistical Mechanics and its Applications*, vol. 270, no. 1-2, pp. 204–214, 1999.
- 21) S. Zambrano, J. M. Casado, and M. A. Sanjuán, "Chaos-induced resonant effects and its control," *Physics Letters A*, vol. 366, no. 4-5, pp. 428–432, 2007.
- 22) S. Nobukawa and N. Shibata, "Controlling chaotic resonance using external feedback signals in neural systems," *Scientific reports*, vol. 9, no. 1, p. 4990, 2019.
- 23) S. Nobukawa, N. Shibata, H. Nishimura, H. Doho, N. Wagatsuma, and T. Yamanishi, "Resonance phenomena controlled by external feedback signals and additive noise in neural systems," *Scientific reports*, vol. 9, no. 1, pp. 1–15, 2019.
- 24) H. Doho, S. Nobukawa, H. Nishimura, N. Wagatsuma, and T. Takahashi, "Transition of neural activity from the chaotic bipolar-disorder state to the periodic healthy state using external feedback signals," *Frontiers in Computational Neuroscience*, vol. 14, p. 76, 2020.
- 25) Y. He, Y. Fu, Z. Qiao, and Y. Kang, "Chaotic resonance in a fractional-order oscillator system with application to mechanical fault diagnosis," *Chaos, Solitons & Fractals*, vol. 142, p. 110536, 2021.
- 26) S. Nobukawa and H. Nishimura, "Synchronization of chaos in neural systems," *Frontiers in Applied Mathematics and Statistics*, vol. 6, p. 19, 2020.
- 27) H. Nishimura, N. Katada, and K. Aihara, "Coherent response in a chaotic neural network," *Neural Processing Letters*, vol. 12, no. 1, pp. 49–58, 2000.

- 28) V. S. Anishchenko, M. Safonova, and L. O. Chua, "Stochastic resonance in the nonautonomous chua's circuit," *Journal of Circuits, Systems, and Computers*, vol. 3, no. 02, pp. 553–578, 1993.
- 29) S. Nobukawa, H. Nishimura, and N. Katada, "Chaotic resonance by chaotic attractors merging in discrete cubic map and chaotic neural network," *IEICE Trans A*, vol. 95, no. 4, pp. 357–366, 2012.
- 30) S. Nobukawa, H. Nishimura, and T. Yamanishi, "Evaluation of chaotic resonance by lyapunov exponent in attractor-merging type systems," in *International Conference on Neural Information Processing*. Springer, 2016, pp. 430–437.
- 31) S. Nobukawa, H. Nishimura, T. Yamanishi, and H. Doho, "Controlling chaotic resonance in systems with chaos-chaos intermittency using external feedback," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 101, no. 11, pp. 1900–1906, 2018.
- 32) N. Shibata and S. Nobukawa, "Synchronization of chaos-chaos intermittency controlled by external feedback and stochastic noise," in *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications*, vol. 2020. The ISCIE Symposium on Stochastic Systems Theory and Its Applications, 2020, pp. 17–22.
- 33) S. Nobukawa, N. Wagatsuma, H. Nishimura, K. Inagaki, and T. Yamanishi, "Novel approach for memory storage systems with chaos-chaos intermittency," in *The International Conference on Emerging Techniques in Computational Intelligence*. IEEE, 2021 (to be published).

デュアルパーティクルフィルタによる ドライバモデルパラメータのオンライン推定

○鈴木宏典（日本工業大学） 藤原翔（日本工業大学大学院）

Online Estimation of Driver Model Parameters by Dual Particle Filter

* H. Suzuki (Nippon Inst. of Tech.) and K. Fujiwara (Graduate School of Nippon Inst. of Tech.)

Abstract— Driver models which is generally recognized as car-following models are the key components of traffic simulation tools. Nevertheless, the model parameters are regarded as constant or distributed randomly with specific probability density functions although they are changing in real-time depending on traffic situations and other stochastic factors. This paper attempts to make an online estimation of these unobservable parameters by using extended dual particle filter. This approach enables us to identify the reaction delay time included in the driver model as an implicit variable. The online estimation requires no linearization of state-space models. Numerical analyses based on driving simulator experiment showed that the proposed approach succeeded in describing individual driver's behavior by estimating the model parameters in real-time. In addition, the identified parameters are able to express the real acceleration precisely including the rapid deceleration scenarios.

Key Words: Online parameter estimation, Driver model, Dual particle filter

1 はじめに

自動車交通流の推定や予測には、交通流シミュレーション技術が活用されている。短時間で動的に変化する交通状態をオンラインで推定し、将来の交通状態を最適化するように的確な交通制御を行うためには、交通流シミュレーションの高精度化がより一層求められる。交通シミュレーションは、マクロ型とミクロ型に大きく大別される。Cell Transmission Model¹⁾に代表されるマクロ型は、自動車交通流を流体として捉え、ある一定区間の道路セグメントの密度、速度、交通量を計算するものである。ミクロ型は、所望のタイムスキャンニングで、車両1台ごとの加速度、速度、位置を計算し、加速度を計算する部分がドライバの追従モデル（以下、ドライバモデル）として定義されている。

これまで、ドライバモデルのパラメータは定数として与えることが一般的であった。しかし、モデルパラメータはドライバ個人によって異なるだけではなく、同一のドライバであっても交通状態によって異なることは想像に難くない。近年、自動運転技術の急速な開発に伴う計測デバイスの進展により、車両走行時のデータ取得が容易になっていることから、実世界で走行する車両のドライバモデルパラメータをオンラインで推定することができれば、より精度の高い交通シミュレーションの実現が期待される。

ドライバモデルパラメータのオンライン推定は、著者の知る限り 2006 年の論文²⁾が最初である。Hoogendoorn ら²⁾は、モデルパラメータに反応時間を含む 2 つの線形モデル（1st GHR モデル³⁾と Helly モデル⁴⁾を対象とし、パーティクルフィルタの粒子を unscented 変換した unscented particle filter⁵⁾によりモデルパラメータの推定を試みた。しかし、対象は単純な線形モデルであり、推定するパラメータは車両1台ではなく複数台を対象としたものであった。Monteil ら⁶⁾は、Intelligent Driver Model (IDM)⁷⁾と呼ばれる非線形のドライバモデルのパラメータを拡張カルマンフィルタ (EKF) でオンライン推定したが、IDM はドライバ特性の重要なファクターである反応時間を含まず、EKF で非線形モデルを線形化する際の精度が課題で

ある。Woo ら⁸⁾は、反応時間を含む非線形型の GHR モデル³⁾を対象に、3 つのパラメータをレーベンバーク・マーカート法、反応時間をプリミティブな最適化手法で推定した。反応時間だけ異なる手法が用いられた理由は、反応時間だけは偏微分できない implicit なモデルパラメータであるためと推測されるが、一貫性の観点では疑問が残る。その他、パラメータのオンライン推定に関するほとんどの先行研究⁹⁾は、非線形モデルの線形化が必要な EKF を利用し、反応時間を含まない IDM を対象としている。また、すべての先行研究において、加速度を計算するモデルパラメータの推定にも関わらず、パラメータの妥当性は加速度ではなく、時間で積分した速度や位置で評価されている。このため、必ずしも適切な加速度を生むパラメータであるか、その妥当性が十分検証されているとは言えない。

そこで本研究では、implicit な反応時間を含む非線形のドライバモデルを対象とし、モデルの線形化を必要としないデュアル型のパーティクルフィルタを用いて、モデルパラメータのオンライン推定を試みる。また、速度や位置ではなく加速度の計測値で推定値の妥当性を評価する。

2 ドライバモデル

Fig. 1 に、追従走行する 3 台の車両を示す。車群先頭から i 台目車両の加速度、速度、位置、車間距離をそれぞれ、 a_i, v_i, d_i と定義すると、Gazis-Herman-Rothery (GHR) モデル³⁾は、車両の加速度を次式で定義する。

$$a_i(k) = \alpha \frac{v_i^m(k-T)}{d_i^l(k-T)} (v_{i-1}(k-T) - v_i(k-T)) \quad (1)$$

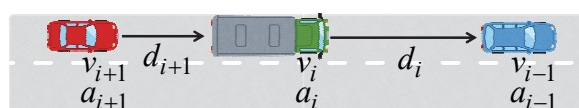


Fig. 1: A three-vehicle car-following.

ここに、 k はタイムステップ、 α, ℓ, m はモデルパラメータ、 T は反応時間である。

GHR モデルは、stimulus-response (刺激反応) モデルの一つであり、先行車との速度差及び車間距離を刺激、4つのパラメータをドライバ個人の感度として、その反応である加速度を出力するモデルである。

ℓ 及び m は、それぞれドライバの車間距離と自車速に対する感度、 α は相対速度に対する感度を表す。 T は反応時間であるが、問題は、式(1)に定義される通り、 T が α, ℓ, m とは異なり、時制を定義するインデックスの役割を果たすことである。このため、例えば拡張カルマンフィルタで T の推定を試みようとしても、 T 自体の偏微分が困難なため、現実的ではない。したがって、 T のオンライン推定には、モデルの線形化を伴わない手法が求められる。

3 デュアルパーティクルフィルタ

3.1 状態空間モデル

状態変数を $\theta_k = [\alpha(k), \ell(k), m(k), T(k)]^T$ 、観測変数を $y_k = a_i(k)$ 、入力変数を $x_k = [v_i(k), v_{i-1}(k), d_i(k)]^T$ とすると、状態空間モデルは以下のように定義される。

$$\theta_k = \theta_{k-1} + r_k \quad (2)$$

$$y_k = g(\theta_{k-1}, x_{k-T}) + n_k \quad (3)$$

r_k 及び n_k は、時刻 k でのシステム誤差と観測誤差である。非線形関数 g は式(1)のドライバモデルである。

3.2 デュアルパーティクルフィルタ

先行研究⁵⁾に従うと、一般に、デュアルフィルタは状態変数とパラメータを交互に推定するシステムとして定義される。すなわち、状態変数をオンライン推定する状態フィルタと、パラメータをオンライン推定するパラメータフィルタの2つのフィルタを交互に動作させるシステムである。

本研究では、このデュアルフィルタの考え方を拡張したオンライン推定システムを提案する。まず、式(2)を個別のパラメータに分割する。

$$\begin{cases} \alpha_k = \alpha_{k-1} + r_{\alpha,k} \\ \ell_k = \ell_{k-1} + r_{\ell,k} \\ m_k = m_{k-1} + r_{m,k} \\ T_k = T_{k-1} + r_{T,k} \\ y_k = g(\theta_{k-1}, x_{k-T}) + n_k \end{cases} \quad (4)$$

次に、推定するパラメータそれぞれに対して一つずつパーティクルフィルタを用意する (PF1~PF4)。例えば、PF1で α を推定する場合、一般的な以下の手順を実行する。

(a) Generation

粒子 $\tilde{\alpha}_k^{(j)}$ をシステムモデルで生成する。 j は粒子番号を示す ($j=1, \dots, M$)。

$$\tilde{\alpha}_k^{(j)} = \hat{\alpha}_{k-1}^{(j)} + r_{\alpha,k} \quad (5)$$

(b) Prediction

生成した粒子 $\tilde{\alpha}_k^{(j)}$ を観測モデル g に入力し、観測変数の一期先予測 $\tilde{y}_k^{(j)}$ を得る。ここで、 α 以外の ℓ, m, T は一つ前の時刻の推定値を入力する。

$$\tilde{y}_k^{(j)} = g\left(\tilde{\alpha}_k^{(j)}, \hat{\ell}_{k-1}, \hat{m}_{k-1}, \hat{T}_{k-1}, x_{k-\hat{T}_{k-1}}\right) + n_k \quad (6)$$

(c) Weight

各粒子の予測値 $\tilde{y}_k^{(j)}$ と実際の観測 y_k を比較し、適当な関数 (ここではガウス分布) を与えて、両者が近い値であれば大きい重み、そうでなければ小さい重み $\tilde{w}_k^{(j)}$ を与える。

$$\tilde{w}_k^{(j)} = N\left(y_k - \tilde{y}_k^{(j)}, \mathbf{n}_k\right) \quad (7)$$

(d) Resampling

必要に応じて各粒子の重み $\tilde{w}_k^{(j)}$ に従って粒子 $\hat{\alpha}_k^{(j)}$ を再選択する。すなわち、重みの小さい粒子は排除し、重みの大きい粒子からさらに「子粒子」を再選択して、尤度の高い粒子のみを残す。

$$\hat{\alpha}_k^{(j)} \sim \begin{cases} \tilde{\alpha}_k^{(j)} & \text{with prob. } \tilde{w}_k^{(1)} \\ \tilde{\alpha}_k^{(j)} & \text{with prob. } \tilde{w}_k^{(2)} \\ \vdots \\ \tilde{\alpha}_k^{(j)} & \text{with prob. } \tilde{w}_k^{(M)} \end{cases} \quad (8)$$

(e) Update

すべての粒子から期待値 $\hat{\alpha}_k$ を計算する。

$$\hat{\alpha}_k = \frac{1}{M} \sum_{j=1}^M \hat{\alpha}_k^{(j)} = \sum_{j=1}^M \tilde{w}_k^{(j)} \tilde{\alpha}_k^{(j)} \quad (9)$$

そして、推定された $\hat{\alpha}_k$ を次のパーティクルフィルタ

PF2に入力し、同様に、 $\hat{\ell}_k$ を計算する。この作業を m, T に対しても繰り返してPF3とPF4を実行し、最後に時刻を進める。図2にこの処理のプロセスを示す。

PF3では、現時刻 k の $\hat{\alpha}_k$ と $\hat{\ell}_k$ 、PF4では、 $\hat{\alpha}_k$ と $\hat{\ell}_k$ に加え \hat{m}_k が入力され、最終的に、 $k=k+1$ として、 \hat{T}_k を含む4つのパラメータがPF1にフィードバックされる。

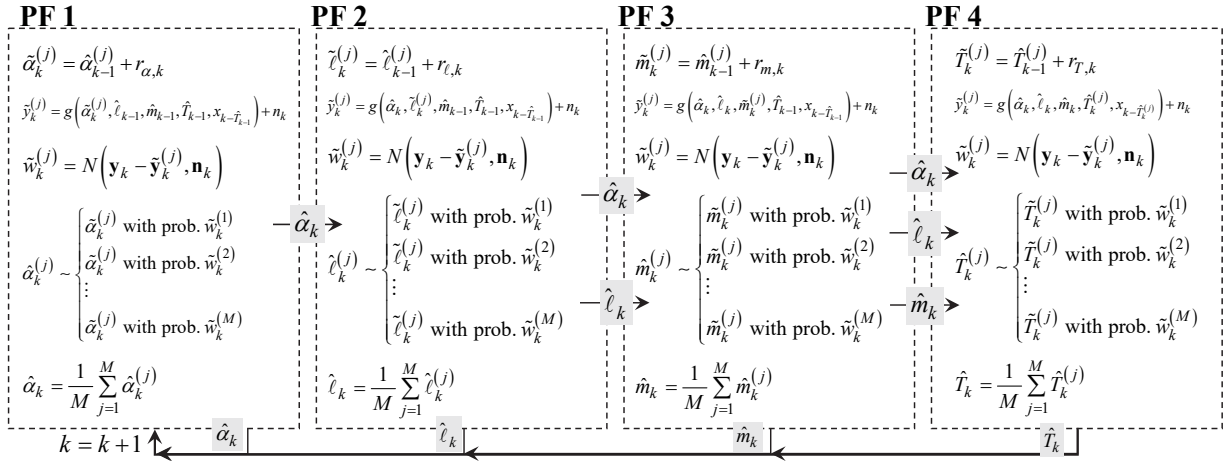


Fig. 2: Estimation procedure by extended dual particle filter.

4 数値計算

4.1 ドライビングシミュレータ実験

評価データを取得するため、アストジェイ株式会社製の定置型ドライビングシミュレータ（以下、DS）を用い、被験者実験を実施した。市街地を想定した街並みの全長約 5 km、幅 7 m の片側一車線の対面通行の直線道路を模擬した。自車は、約 5 分間の走行の内、0 ~ 20 m/s の速度で加減速を繰り返す 3 台の先行車群に追従する。3 台目車両の速度プロファイルを Fig. 3 に示す。

実験参加者は、普通自動車第一種運転免許を保有する者 14 名（平均年齢 21.6 歳、標準偏差 0.610 歳、平均運転歴 3 年）とし、事前に日本工業大学人を対象とする研究倫理委員会の承認及び、インフォームドコンセントを得て実施した。

4.2 評価データ

DS 実験において、実験参加者が実際に運転した 4 台目の車両データを評価データとする。すなわち、3 台目車両との車間距離 d_4 、相対速度 $v_3 - v_4$ から、4 台目の加速度 a_4 が得られた際の、4 台目ドライバのモデルパラメータをオンラインで推定することとした。入力変数である v_3 , v_4 , d_4 については、実環境においてはノイズが含まれているため、毎ステップ、カルマンフィルタによりノイズ除去した値を使用した。

4.3 計算条件

システム誤差の共分散は、 α, ℓ, m についてはいずれも 0.002、 T についても 0.002 とした。尤度を求める際のガウス分布の分散は 0.3 とした。粒子の数は 200 個とし、毎回リサンプリングを行った。

α, ℓ, m の初期値はいずれも 1.0、 T の初期値は 1.5 とした。

4.4 パラメータ推定結果 (α, ℓ, m)

一例として、実験参加者 A の 2 回目、実験参加者 B の 2 回目の走行試験の推定結果 (α, ℓ, m) を Fig. 4 に示す。3 つのパラメータすべてに対して数値が大きい

ほど感度が高くなるよう表現するため、 ℓ の値にはマイナスの符号を付けて表記した。

A, B ともに、パラメータの値はさほど大きく変動していないものの、B については後半にかけて α が低下し、代わりに m が増加した。これは、ドライバ A が相対速度に対する感度を下げ、自車速に対する感度を上げて走行する状態に変化したことを示唆している。一方、A では、時間の推移とともに α と m は増加し、 ℓ は低下したことから、B は相対速度と自車速の感度は高め、車間距離への感度を下げて運転する傾向にあることがわかる。

このように、ドライバモデルパラメータは、同一の走行環境であっても実験参加者毎に異なる値を示し、同一のドライバでも走行状況に応じて時間的にも変化していることが明らかとなった。

4.5 パラメータ推定結果 (T)

同様の実験参加者の反応時間 T の推定結果を Fig. 5 に示す。B は、前半は反応が短い後半は反応が長くなっている。自車速に対する感度を上げたことで、反応が短くなくとも追従するように行動を変化させた可能性が示唆される。A は、3 つのパラメータは顕著に変化させない代わりに反応時間を頻繁に変化させて追従していることが推察される。

4.6 パラメータの妥当性

推定されたパラメータが正しく加速度を表現できているのか、Fig. 6 を用いて確認する。同一の走行環境

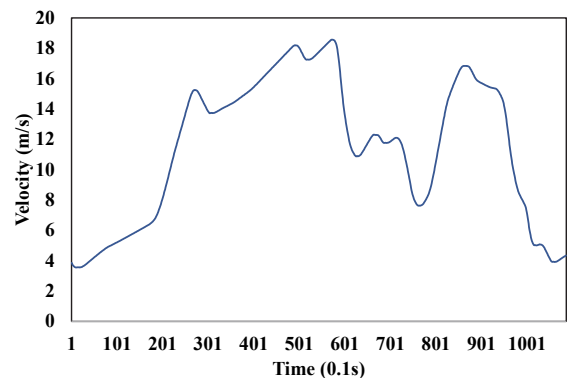


Fig. 3: Speed profiles of three preceding vehicles.

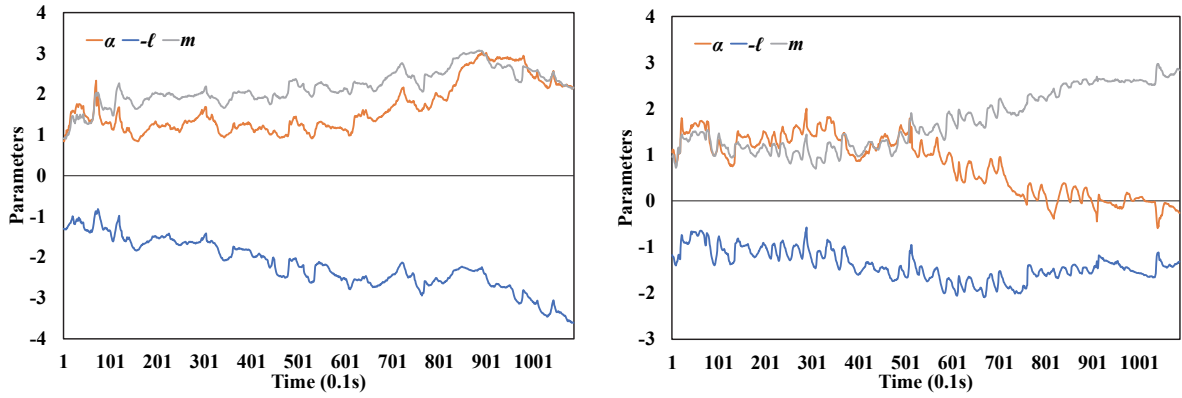


Fig. 4: Estimated parameters (α , ℓ and m) (Left: participant A, right: participant B).

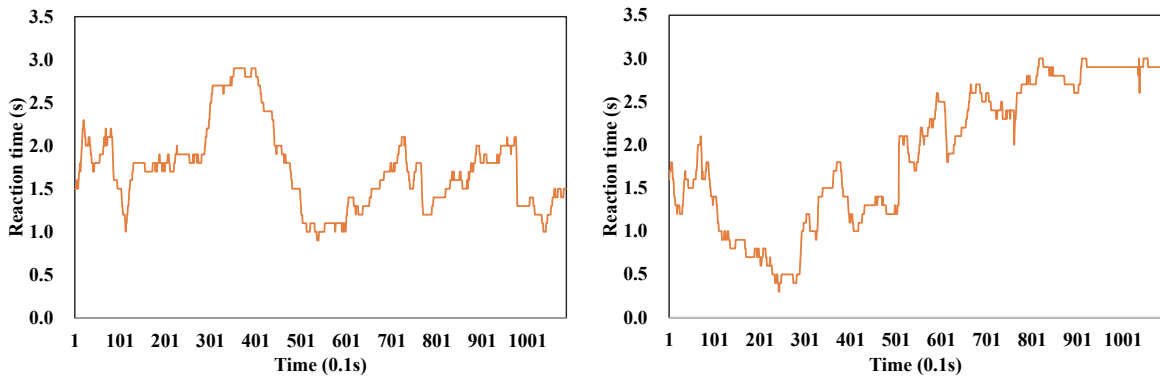


Fig. 5: Estimated reaction time (T) (Left: participant A, right: participant B).

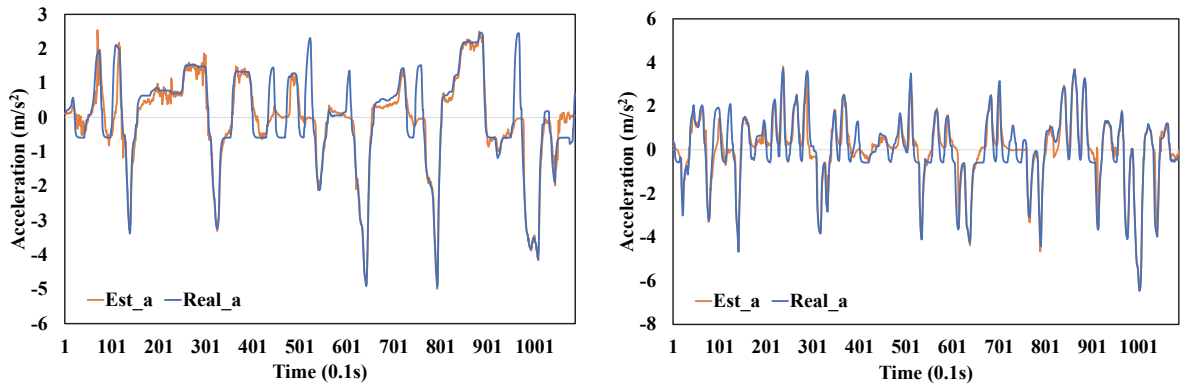


Fig. 6: Comparison between estimated and real acceleration (Left: participant A, right: participant B).

であっても、Bは非常に小刻みに加減速を繰り返す、Aは必要な時に大きな減速を行っている。いずれも、細かい部分においては加速度の再現が十分ではない箇所もあるものの、特に強い減速を行う場面では、正しくその最大減速度も推定できていることから、推定された4つのパラメータの妥当性は保証されていると考えられる。挙動の異なるドライバに対しても、今回の提案手法は、その加速度を高い精度で再現できるポテンシャルを持つことが明らかにされた。

5 おわりに

今回、デュアルフィルタの考え方を拡張したデュア

ルパーティクルフィルタを用いて、ドライバモデルパラメータをオンライン推定する試みを行った。数値計算の結果、4つのパラメータを統一されたプロセスの中で推定できたこと、ドライバ状態を表す重要な反応時間も線形化を伴わずに推定できたこと、推定されたパラメータは高い精度でドライバ固有の加速度を再現できることを確認した。

また、同一の走行環境であってもドライバの走行特性が異なることや、同一のドライバであっても走行環境に応じて走行特性を変化させている可能性も示唆された。

しかしながら、Hoogendoornら²⁾が指摘している通

り、パーティクルフィルタの共分散の違いによりパラメータの推定結果が異なることも経験上明らかとなっている。今後は、より多くのデータでの検証が必要であると同時に、共分散の値そのものも推定すべきパラメータに含めて精度の向上を図る必要がある。

謝辞

本研究は、鈴木シンワート株式会社・若林翔氏（前日本工業大学大学院）の多大なる協力の元、科研費（基盤研究(C) 19K04926）の補助を受けて実施した。ここに記して深謝の意を表す。

参考文献

- 1) Daganzo, C.F.: The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory, *Transportation Research Part B*, Vol. 28, No. 4, p. 269-287 (1994)
- 2) Hoogendoorn, S., S. Ossen, M. Schreuder, B. Gorte: Unscented Particle Filter for Delayed Car-Following Models Estimation, *Proc. of IEEE ITSC*, p. 1598-1603 (2006)
- 3) Chandler, R.E., R. Herman, and E.W. Montroll: Traffic Dynamics: Studies in Car-Following, *Operations Research* 6 (2), p. 165-184 (1958)
- 4) Helly, W.: Simulation of Bottlenecks in Single Lane Traffic Flow. in *International Symposium on the Theory of Traffic Flow*, (1959)
- 5) Haykin, S.: *Kalman filtering and Neural Networks*, Adaptive and Learning Systems for Signal Processing, Communications, and Control, ed. S. Haykin, John Wiley & Sons, Inc. (2001)
- 6) Monteil, J., N. O'Hara, V. Cahill, M. Bouroche: Real-Time Estimation of Drivers' Behaviour, *Proc. of IEEE ITSC*, p. 2046-2052 (2015)
- 7) Kesting, A., M. Treiber and D. Helbing: Enhanced Intelligent Driver Model to Assess the Impact of Driving Strategies on Traffic Capacity, *Philosophical Transactions of the Royal Society A*, 368, p. 4585-4605 (2010)
- 8) Woo, H., H. Madokoro, K. Sato, Y. Tamura, A. Yamashita, H. Asama: Estimation of Operation Characteristics of Following Drivers Based on Car Following Model, *Transactions of Society of Automotive Engineers of Japan*, Vol. 51 (2), p. 304-309 (2020) (in Japanese)
- 9) 例えば, Buyer1, J., D. Waldenmayer1, N. Sußmann1, R. Zöllner, J. M. Zöllner: Interaction-Aware Approach for Online Parameter Estimation of a Multi-lane Intelligent Driver Model, *Proc. of IEEE ITSC*, p. 3967-3973 (2019)

骨格の時系列変化情報を用いた歩行人物の行動推定

○酒井俊権 磯川悌次郎 上浦尚武 (兵庫県立大学)

An Estimation Scheme for Pedestrian Behaviors from Their Skeleton Signals

*T. Sakai, T. Isokawa, and N. Kamiura (University of Hyogo)

Abstract— This paper presents a scheme for estimating the behaviors of pedestrians from these images, in order to construct behavioral models for pedestrians in crossing roads. The proposed scheme first extracts the structure of the body for pedestrian by using the OpenPose system that extracts the skeleton structures of humans from the image. After collecting the structural changes in a short duration, such as angles of knees and toes, these time series are used for classification of behaviors. The proposed scheme is evaluated by using actual images of a pedestrian, and it is shown that a k -neighbor classifier achieves the best performances for estimating his behaviors.

Key Words: Pedestrian, behavior estimation, time series

1 はじめに

自動運転技術は、交通事故の減少、過疎地域での高齢者の移動補助、ドライバーの負荷軽減や交通渋滞の緩和など様々な問題の解決策として近年非常に注目されており、自動車メーカーのみならず様々な研究機関において研究開発が加速している。しかし、自動運転技術が実用化した際に、自動運転システム自体がどの範囲の予期せぬ事象に対応しうるのかという予測が立てられていないことから、自動運転の安全性や効果に疑問の声も多くあり、重大な事故を起こす可能性も指摘されている。

加えて自動運転の技術がどの程度実用化された際に、どの程度の効果があるのかが実証されていない点も問題として挙げられている。運転の自動化レベルは米運輸省道路交通安全局によると5つのレベルに分類される。レベル1(運転者支援)では運転モード(道路や環境条件など)が許せば、操舵あるいは加減速実行のいずれかはシステムが担うことができる。レベル2(部分的運転自動化)では、操舵および加減速実行の両方についてシステムがドライバーに代わり行うことができるが、ドライバーは常に走行環境とシステムを監視しておく必要がある。レベル3(条件付運転自動化)では、レベル2においてドライバーの監視義務がなくなるという点が異なる。ただし、この場合においても運転タスクのバックアップはドライバー自身で行う必要がある。レベル4(高度運転自動化)・レベル5(完全運転自動化)では、どのような状況においても運転操作はシステムが行い、異常事態が生じた場合でもシステムがバックアップを行う。レベル4では、対応できる運転モードに制約がある場合があるが、レベル5においてはどのような運転モードにおいてもシステムが自動運転を行うことができる。しかし、自動運転のレベルがどの程度に達すればどの程度安全性に影響があるのかという予測はされていないのが現状である。このため自動運転の及ぼす効果や運転の自動化のレベルの変化による安全性への影響の変化を事前にシミュレーションすること求められている。

また、交通死亡事故の大半を占める一般道路で自動運転を実用化することが国家目標となる中で、技術的には複雑な交通環境下で歩行者を認識し歩行行動を正

確に予測することが課題となっている。また、道路周辺の構造物や天候など複雑な環境下で歩行者認識システムの性能を評価するためには、仮想環境での高度なシミュレーションが有効である。歩行者認識システムは企業間の競争によって開発が進められているが、シミュレーションに組み込む歩行者の行動モデルは協調して開発することが重要かつ喫緊の課題となっている。

しかし、現在シミュレーションに使用するための車や信号機などの機械についてはそれらのモデルは存在しているが、歩行者の行動モデルとして適したものはほとんど存在しないのが現状である。これは行動モデルの元となる人間の実際の行動の解析がなされていないことが原因として挙げられる。手動にて解析した結果は、例えば^{1,2)}において検討されているが、多くの歩行者動画像について手動にて解析することは困難である。このため行動モデルを制作するためには、まず人間が実際どのように行動しているのかを機械が正しく識別するシステムが必要となる。

そこで本研究では、歩行者モデルを構成する第一段階として、まず実際の歩行行動を行う歩行者の画像からその行動を推定する手法を提案する。提案手法では、動画像からの人間の骨格推定ライブラリであるOpenPose³⁾を用い、取得した動画上の各フレームでの歩行者の関節の位置情報を取得し、その短時間時系列情報を入力とした分類器を用いることによって認識を行う。

2 歩行行動の推定手法

本論文で提案する歩行者行動の推定方法を図1に示す。提案手法では、まず撮像した歩行者の動画像を静止画像フレームとして取り出し、各画像フレームに対して歩行者の骨格構造およびその座標値を得る。短時間の画像フレーム群から座標情報を取り出すことにより、歩行者の各箇所(頭部、腕、足など)についての時系列信号が得られる。この時系列信号に対して欠損値を補完した後に、識別器によりこの時系列信号を入力とした際の行動クラスを推定する。この時系列信号の取り出しとそこから推定された行動を全画像フレームに対して繰り返して行う。

まず、歩行者の骨格情報の取り出しについては、OpenPose³⁾を用いて人間の姿勢情報を取得する。OpenPoseは動画像から人間の姿勢を推定するライブ

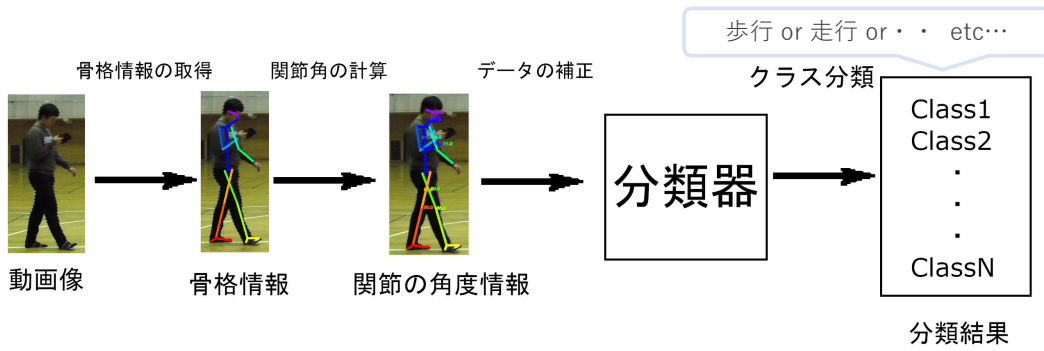


Fig. 1: Processing flow in the proposed scheme



Fig. 2: An example of output from OpenPose system³⁾

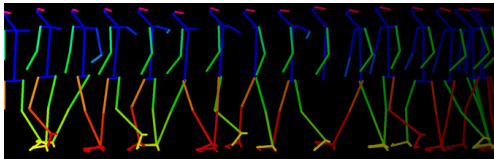


Fig. 3: A set of extracted skeleton structure for a pedestrian

ラリであり、画像中の身体部位の位置のみではなく、部位間の関係性もエンコードすることによって人間の姿勢情報の推定精度を向上させたものである。姿勢情報は入力画像中の座標で表された25個の関節点で表されている。例として、OpenPoseによって出力された関節位置をつなぎ合わせることで骨格として画像に重ね合わせて表示した画像を図2に示す³⁾。

OpenPoseにて抽出した歩行行動の関節位置を数フレーム間隔で重ねてプロットしたものを3に示す。また多様な測定条件においてOpenPoseによって取得できる25個の関節位置からの解析は困難であるため、提案手法では、文献⁴⁾にて用いられている手法にて関節角度を算出した。この方法では特定の骨を表す関節位置のペアをボーンとし、同一関節点を持つボーン同士の内積から角度を求めることで0から 2π の範囲で求めた。取得画像と関節角の例を図4に示す。本提案手法では、主要な17個の関節について関節角データを求め、学習および評価データとして用いた。

取得した動画画像についてOpenPoseを用い関節座標を取得した後に、関節座標から各関節角を求め分類に使用する関節角データとした。また欠損値や誤差などの影響を低減するため、得られた関節角の補正を行った。関節角の補正方法として次に示す三つの方法を検討した。

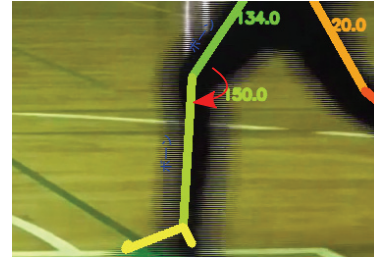


Fig. 4: Calculation for joint angles

補正方法1 (SC1) OpenPoseで関節が取得できていないフレームは1フレーム前の値を使用

補正方法2 (SC2) 補正方法1を行った後、過去5フレームの関節角の移動平均を値として使用

補正方法3 (SC3) 補正方法1を行った後、過去8フレームの関節角の移動平均を値として使用

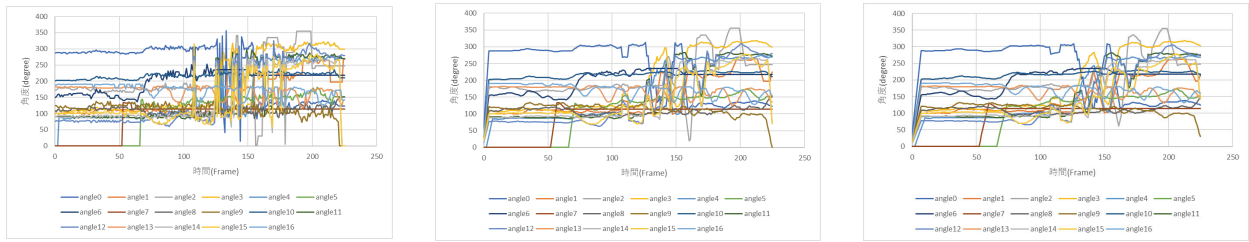
取得した関節角の時系列信号について、各補正手法を適用した結果を図5に示す。Fig. 5(a)に示す補正方法1を行ったデータでは欠損値が正しく補正されていることが分かる。またFig. 5(b), Fig. 5(c)にそれぞれ示す補正方法2, 3を行ったデータでは、移動平均を取ることで補正方法1よりもさらに波形が滑らかに変化していることが確認できる。

この時系列信号を20ステップずつ取り出し、識別器に入力することにより行動の識別(推定)を行う。本論文では、いくつかの識別器を用いて識別性能の比較を行う。識別器として、 k 近傍法(k -NN)、ランダムフォレスト⁵⁾(RF)、SVM⁶⁾、ロジスティック回帰⁷⁾(LR)、多層パーセプトロン⁸⁾(MLP)という5種類の識別器を性能評価に用いた。

3 実験

3.1 データセット

本研究では、行動分類について分類対象となるクラスを、手前、奥、左、右からの4方向からの、歩行、走行、スマートフォンを見ながらの歩行(スマホ歩き)、首を振りながらの歩行(脇見歩行)の4種類の動作パターンの計16種類に加えて、静止状態、転倒、転倒後、の3種類を加え計19種類と設定した。この設定したクラスについて、歩行者1名に対してそれぞれの動画撮影を行いフレームごとの画像を取り出した。動画の撮影にはPanasonic HC-V480MSを用い、解像度



(a) Signals with SC1

(b) Signals with SC2

(c) Signals with SC3

Fig. 5: Time series for joint angles with three types of compensation

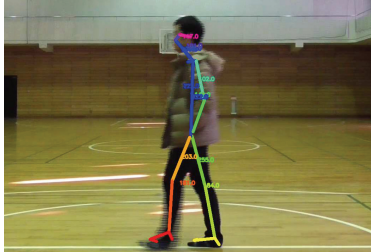


Fig. 6: A pedestrian in the experiment

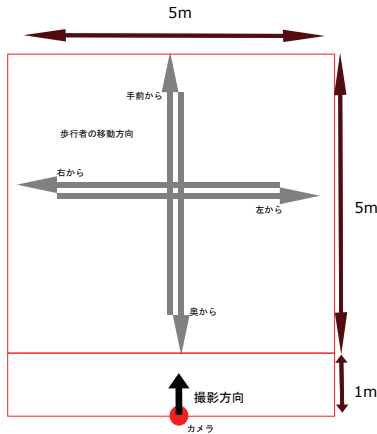


Fig. 7: Configuration for a camera and a target pedestrian

1920×1080, フレームレート 30fps とした. カメラおよび歩行者の位置関係 (撮影環境) を Fig. 7 に示し, 実際の撮影画像について OpenPose により骨格情報を取り出した図を Fig. 6 に示す. このような撮影を行い, 各クラスと画像のフレーム数について Table 1 に示す. また撮影データの 8 割を学習データ 2 割をテストデータとして使用した.

3.2 分類結果

テストデータを用い分類精度を求めた結果を Table 2 に示す. この結果より, どの補正方法を行ったとしても k -近傍法および, ランダムフォレストは 95% を超える高い正答率で行動分類を行うことができていることが分かる. また, 補正方法 3 によるデータ補正および k -近傍法による分類が 97.7% と最も精度が高くなることが示された.

次に, 深層学習 (ニューラルネットワーク) 手法を用いる分類手法により 19 クラス分類を行った結果について示す. ここでは, RNN⁹⁾, LSTM¹⁰⁾, CNN+LSTM¹¹⁾, ConvLSTM¹²⁾ の 4 種類の分類手法を用いた. 各手法の分類結果について Table 3 に示す. 実験結果より ConvLSTM を用いたものが 94.4% と最も高い精度が得られた. しかしながら, k -近傍法などの分類手法と比べ

Table 1: The numbers of image frames for each behavior class

Behavior	#. frames
walk from right	7175
run from right	2495
walk from left	4736
run from left	2157
walk from back	2560
run from back	1518
walk from front	1669
run from front	1356
walk on smartphone from right	2545
walk from right with looking aside	2397
walk on smartphone from left	3230
walk from left with looking aside	2781
walk on smartphone from back	3665
walk from back with looking aside	1713
walk on smartphone from front	2661
walk from front with looking aside	2055
stand still	3377
fall down	527
rise after falling down	3377
Total number of frames	49605

Table 2: Classification accuracy for each classifier, where k -NN, RF, SVM, LR, and MLP, are respectively k -Nearest Neighbor, Random Forrest, Support Vector Machine, Logistic Regression, and Multilayer Perceptron

Classifier	Compensation method		
	SC1	SC2	SC3
k -NN	0.961	0.973	0.977
RF	0.964	0.976	0.975
SVM	0.512	0.569	0.482
LR	0.816	0.850	0.862
MLP	0.914	0.940	0.947

ると低い精度であり, 深層ニューラルネットワークを用いても精度の向上は確認できなかった.

以上の結果より, k -近傍法による分類手法を用いることが動画から人間の行動を認識するにあたり最も適切であることがわかった.

最後に, 学習データに偏りがあるためすべてのクラスで正しく分類されているかを確認するため, 最も精度が高くなった結果について (k 近傍法+補正方法 3), 混同行列を算出し, 各クラスについての真陽性率, 真陰性率をヒートマップにより可視化したものを Fig. 8 に示す. この結果より, すべてのクラスで正しく分類

Table 3: Classification Accuracy using Deep neural networks(%)

Classifier	Compensation methods		
	SC1	SC2	SC3
RNN	0.781	0.826	0.836
LSTM	0.873	0.881	0.898
CNN+LSTM	0.913	0.937	0.933
ConvLSTM	0.933	0.934	0.944

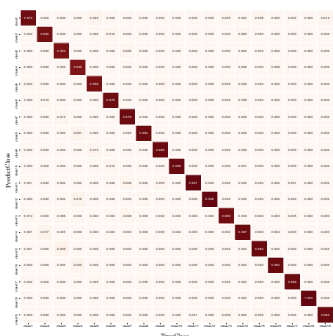


Fig. 8: Confusion matrix by k -NN with SC3

されていることがわかり、 k -近傍法は高い精度で分類が可能であることが示された。

4 まとめ

本研究では動画からの歩行者の行動認識を行う手法を提案し、その有効性を実歩行者動画をを用いて評価を行った。本手法では、OpenPoseを用い人間の関節座標を取得したのち、関節座標について補正を行い、その短時間時系列座標変化を識別器を用いてクラス識別を行う。実験結果より k -近傍法により分類することにより高精度で行動推定ができることが示された。

今後の課題として、より現実に即して撮影されたデータや様々な距離や角度で撮影されたデータに対する精度の検証、およびより多クラスの分類での検証が挙げられる。また、この提案手法を用いて多数の歩行者動画について行動推定を行い、歩行者の周囲状況など

を考慮に入れた歩行者行動モデルを構成していくことが必要である。

参考文献

- 1) 青木ほか: 予防安全支援システム効果評価シミュレータ (ASSESS) のための歩行者行動特性の解析, 自動車技術会論文集, 42(5), 1199/1204 (2011)
- 2) 稲垣ほか: 生活道路における子どもの横断判断特性に関する実験的考察, 土木学会論文集 D3 (土木計画学), 71(5), 665/671 (2015)
- 3) Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh: OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields, arXiv preprint 1812.08008 (2018)
- 4) K. Inada, T. Ishikawa and H. Watanabe: Evaluation of Action Synchronization and Shape for Multiple Dancers Using OpenPose, 映情学技報, 42(44), 21/23 (2018)
- 5) L. Breiman: Random forests, Machine learning, 45(1), 5/32 (2001)
- 6) V. N. Vapnik, and C. Cortes: Support-vector networks, Machine learning, 20(3), 273/297 (1995)
- 7) D. R. Cox: The regression analysis of binary sequences, Journal of the Royal Statistical Society: Series B (Methodological), 20(2), 215/232 (1958)
- 8) V. D. Malsburg: Frank Rosenblatt: principles of neurodynamics: perceptrons and the theory of brain mechanisms, Brain theory, 245/248 (1986)
- 9) I. Sutskever, O. Vinyals, and Q. V. Le: Sequence to sequence learning with neural networks, Advances in neural information processing systems, 3104/3112 (2014)
- 10) S. Hochreiter and J. Schmidhuber: Long Short-term Memory, Neural computation, 9, 1735/1780 (1997)
- 11) J. Donahue, L. A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell: Long-term Recurrent Convolutional Networks for Visual Recognition and Description, arXiv preprint CoRR abs/1411.4389 (2014)
- 12) S. Xingjian, Z. Chen, H. Wang, D. Yeung, W. Wong, and W. Woo: Convolutional LSTM network: A machine learning approach for precipitation nowcasting, Advances in neural information processing systems, 802/810 (2015)

VR 環境におけるフリック文字入力に適した フィードバック手法の検討

○飯田春樹 今井順一（千葉工業大学）

Investigation of Feedback Modalities Suitable for Flick Input in VR Environments

* H. Iida and J. Imai (Chiba Institute of Technology)

Abstract— Modern virtual reality (VR) technique allows us to use a movement of a user's fingers in the real world for input in the VR environments without wearing any special devices. The method for flick text input in the VR environments using this technique has been proposed. However, this method has the drawback of high input error rate. In this study, we focus on modalities of feedback for VR flick input and investigate their effectiveness on the input error rates and usability.

Key Words: Virtual reality, Flick input, Feedback modalities

1 はじめに

近年, Oculus Quest¹⁾といった安価なヘッドマウントディスプレイ (HMD) の登場により, VR (Virtual reality) が一般的に普及し始めている. さらには, Unity3D や Unreal Engine4 といった VR 開発環境が整いつつあることから, 今後多種多様な VR コンテンツが登場することが考えられる.

VR コンテンツの中には, テキストチャットやコンテンツの文字検索といった場面で文字入力が必要となる場面が存在する. しかし, VR コンテンツを体験するためには利用者が HMD を装着する必要がある, 現実世界の視界が HMD によって遮蔽されてしまうため, 物理的なキーボードを利用した文字入力を行うことが難しい. そこで現在 VR での文字入力には専用のコントローラを利用する手法が一般的に用いられている (Fig.1 参照). しかし, 小さなキーをコントローラで正確に狙うことは難しく, 入力速度も遅いという問題がある.

そこで, 正確かつ高速な文字入力を目指して, ユーザの指の動きを VR 空間上に反映させる機器を利用したフリーハンド文字入力手法が盛んに研究されている^{2),3)}. しかし, それらの研究は独自の入力方法を考案するもの²⁾や QWERTY 配列のキーボードを対象にするもの³⁾が多く, 操作の習熟に時間がかかる, 若年層を中心に QWERTY 配列キーボードに苦手意識がある人が増加している⁴⁾などの課題があった.

それらの課題を解決するため, フリック入力方式に着目し, VR でのフリック文字入力インタフェースの開発に取り組んでいる研究がある^{3),5)}. フリック入力の利点として, モバイルデバイスなどで広く利用されており操作の習熟に時間がかからない点や QWERTY 配列のキーボードに比べ高速に日本語文字入力が行える点などがある. しかし, これらのインタフェースは, 誤入力率が高く, 実用的とは言い難かった.

そこで本研究では, VR フリック入力時のフィードバックの種類に着目し, 多様なフィードバックを付与することによって誤入力率低下と操作性の向上を目指す.

2 VR環境におけるフリック入力

本研究では, ハンドトラッキング機能が内蔵された VRHMDであるOculus Questを用いて, 先行研究⁵⁾で開発さ



Fig. 1: Conventional text input in VR environments.

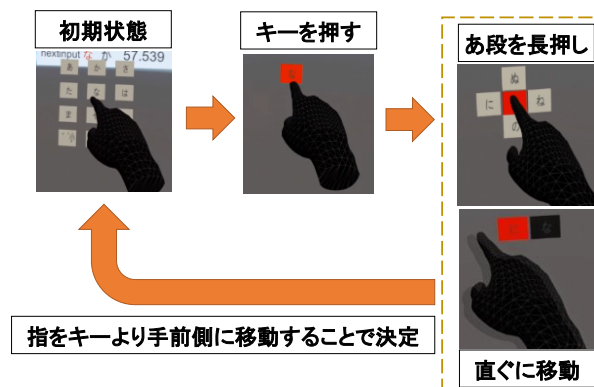


Fig. 2: Flick text input in VR environments.

れたVR環境においてフリーハンドでひらがなを入力できるフリック入力インタフェースを再現する. ひらがなは五十音46文字, 濁音15文字, 半濁音5文字, 捨て仮名10文字の合計71文字を入力対象とし, 変換機能などは考慮しないものとした.

作成したインタフェースにおいてひらがなを入力するまでの手順をFig. 2に示す.

2.1 キー

各キーには当たり判定があり, ハンドトラッキングにより検出したVR環境内のユーザの手とキーの接触を検出できる. このキーをFig. 3のように12個配置する. また, キーの間隔について先行研究⁵⁾では仮想キー

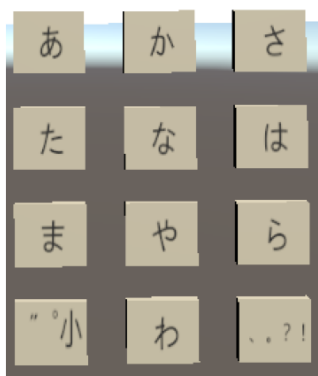


Fig. 3: Flick keyboard.

ボードの適切なキーの大きさに関する小沢らの研究成果⁶⁾から縦横のキー間隔を約26.3mmと約23.4mmと設定していた。本研究でもこのキー間隔を採用する。

2.2 入力

12個のキーのうち、「あ段」が表示されているキーとVR空間内のユーザの手が接触した時点から入力を開始し、指先がキーよりも手前側に移動した時点で入力を確定する。

入力文字は、確定時点での指先が、最初に触れた「あ段」のキーを中心とする範囲のどの領域にあったかで判定する (Fig. 4)。Fig. 4の領域0が最初に触れた「あ段」のキーの範囲を表す。それ以外の領域は式 (2.1), (2.2) を境界として分割され、領域1 (左) が「い段」,

2 (上) が「う段」, 3 (右) が「え段」, 4 (下) が「お段」に対応する。

$$x + y = 0 \quad (2.1)$$

$$x - y = 0 \quad (2.2)$$

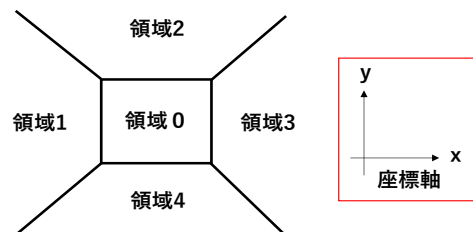
また、Fig. 3の最下段左側のキーは押されると一文字前に入力された文字を取得し、その文字が濁音/半濁音に変換できれば濁音/半濁音に、濁音/半濁音/捨て仮名が取得されている時に押されると清音に変換する。

「あ段」を表示する10個のキーと「濁音/半濁音/捨て仮名」を変換するキーを用いることで目標の71文字を入力することができる。「や行」の領域1, 3や「わ行」の領域3, 4及び最下段右側のキーについては、本研究の対象外ではあるが、Google日本語入力を参考に各種記号を割り当てた。

3 提案手法

先行研究⁵⁾では、実験結果からフリック入力に習熟しているユーザはインタフェースを見ずに入力を行っているのではないかとこの考察がなされており、習熟したユーザには視覚的なフィードバックの効果が薄い可能性が示唆されていた。

フィードバックの有無が操作性に与える影響について様々な研究が行われている^{7,8)}。和氣ら⁷⁾は、タッチパネルにおける操作反応音の有効性について調査を行っている。実験結果から、不正確操作の低減について操作反応音が有効であることが示されている。また、Wuら⁸⁾は、VR空間の文字入力について、物理キーボード、仮想キーボード、触覚フィードバックを付与した



- ・領域0 : あ段のキーの範囲
- ・領域1 : 領域0を除く範囲で $x - y < 0$ かつ $x + y < 0$
- ・領域2 : 領域0を除く範囲で $x - y < 0$ かつ $x + y > 0$
- ・領域3 : 領域0を除く範囲で $x - y > 0$ かつ $x + y > 0$
- ・領域4 : 領域0を除く範囲で $x - y > 0$ かつ $x + y < 0$

Fig. 4: Regions for key input discrimination.

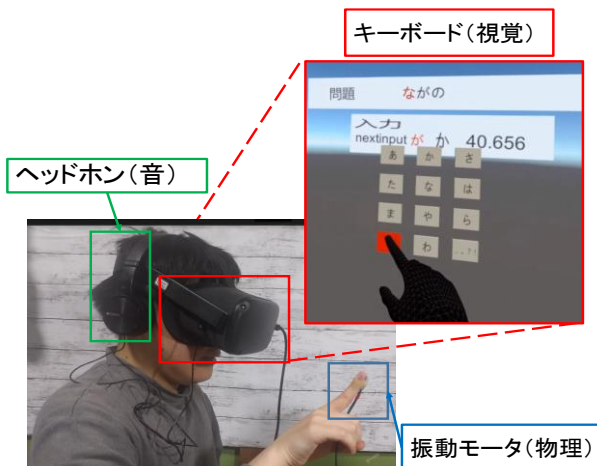


Fig. 5: System overview.

仮想キーボードの比較を行い、仮想キーボードの場合は触覚フィードバックを付与すると性能が高まることを示した。これらの研究から操作性の向上に視覚以外のフィードバックが有効に働くことが考えられる。

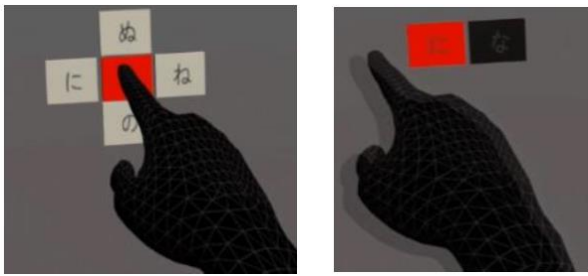
そこで本研究では、VR環境においてひらがなを入力できるフリック入力インタフェースにおいて、先行研究⁵⁾で実装されていた視覚的フィードバックに加えて、聴覚的フィードバック及び物理的フィードバックを付与し、入力速度の向上や誤入力率の低下といった操作性の向上を目指す。提案手法を利用して文字入力を行っている様子をFig. 5に示す。

3.1 視覚的フィードバック

本研究では視覚的フィードバックとしてキーの展開とキーの色変化を採用する。これらはモバイルデバイスなどのフリック入力形式キーボードでも広く採用されており、それを再現することでモバイルデバイスでの入力感覚に近い操作感をユーザに与えるためである。また、どのキーをユーザが操作しているのかわかりやすくするためと誤入力を減らすために、任意の段のキーを操作している時には、他段のキーを非表示にし、操作できないようにした。

①キーの展開

「あ段」のキーが一定時間押され続けると、キーの上下左右に「い～お段」が表示される (Fig. 6 (a))。また、キーに触れてから一定時間内に領域0から離れるとキーは一部分しか展開されない (Fig. 6 (b))。



(a) Touch

(b) Flick

Fig. 6: Visual feedback.

②キーの色変化

「あ段」のキーが押されてから入力が確定するまでの間、指先が存在する領域の色を変化させる。また、キーの展開が一部分しか行われていない場合は、選択された領域に加えて「あ段」の色もFig. 6 (b)のように変化させる。

3.2 聴覚的フィードバック

聴覚的フィードバックはヘッドホンから音を出力することで行う。音の出力は1文字の入力に対して2回行う。まず、1回目はキーに指先が触れた瞬間に行う。使用する音にはモバイルデバイスでの聴覚的フィードバックに用いられているような電子音を用いた。2回目はキーから指先が離れた瞬間に行う。こちらは、キーボードで文字入力を行っている感覚を想起させることを狙い、物理キーボードの打鍵音を模した音を用いた。状況に合わせてそれぞれ別の音を出力することでユーザに現在の操作状況をフィードバックする。

3.3 物理的フィードバック

物理的なフィードバックではArduinoで制御された振動モータをユーザの指先に取り付け、VR空間内でキーに指先が触れている間振動を続けることで、キーに触れている感覚をフィードバックする。

4 評価実験

本研究で追加したフィードバックの効果を確認するため、VR環境でのタイピングゲームによる実験を行った。

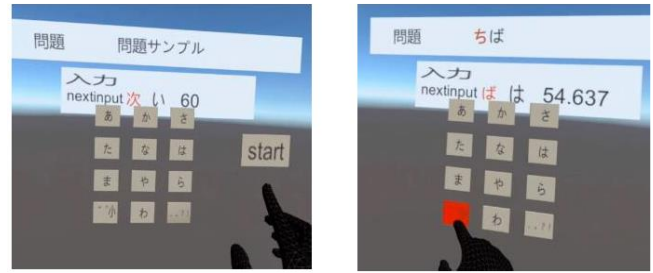
4.1 概要

タイピングゲームの画面をFig. 7に示す。はじめに画面上には「スタート」ボタンが表示されており、このボタンを押すとゲームが開始される (Fig. 7 (a))。ゲームでは、ランダムに出題される単語を入力し、正しい入力が行われると次の単語が表示されるという試行を60秒間繰り返す (Fig. 7 (b))。出現する単語は先行研究⁵⁾に倣って日本の都道府県名とした。

実験参加者は15歳から25歳までの男女16名 (男性12名、女性4名) である。

4.2 条件

実験条件として、次に示す4条件と設定する。



(a) Start

(b) Text input

Fig. 7: Typing Game.

1. 視覚的フィードバックのみ (先行研究と同等)
2. 視覚的フィードバック+聴覚的フィードバック
3. 視覚的フィードバック+物理的フィードバック
4. 視覚的フィードバック+聴覚的フィードバック+物理的フィードバック

4.3 評価方法

上記の4条件に対し、入力速度や誤入力率といった客観評価とアンケートによる主観評価の2つの面から評価を行う。入力速度を表す単位は一分間に何文字入力できたかを表現する CPM (Characters per minutes) とし、誤入力率は誤入力回数をキーの入力回数で割った値とした。

アンケートの項目を以下に示す。普段フリック入力をどの程度使用しているのか、各フィードバックに気付いたかまた効果的に感じたか、どの条件のフィードバックが最も使いやすかったかなどを評価する。

1. 普段フリック入力をどの程度利用しているか
2. インタフェースの操作方法はすぐ理解できたか
3. どの段が入力しにくかったか
4. 視覚的フィードバックに気付いたか
5. キーの色変化は入力の助けになったか
6. キーの展開は入力の助けになったか
7. 音によるフィードバックに気付いたか
8. キーに触れた時の音は入力の助けになったか
9. キーから離れた時の音は入力の助けになったか
10. 物理的フィードバックに気付いたか
11. 物理的フィードバックは入力の助けになったか
12. VRで日本語入力を行う際に本研究のインタフェースを使いたいと思ったか
13. どの条件のものを使いたいと感じたか

アンケート項目の1, 2, 4~12において5段階評価を行い、項目3は、あ~お段の選択式、項目13は条件1から条件4の選択式とした。

4.4 客観評価の結果と考察

4条件の CPM と総キー入力回数の平均値を Fig. 8、誤入力率を Fig. 9 にそれぞれ示す。

CPM と総キー入力回数について、条件1 (視覚的フィードバックのみ) が最も少なく、条件2 (視覚+聴覚)、条件3 (視覚+物理)、条件4 (視覚+聴覚+物理) の順に CPM 及び総キー入力回数の両方が増加していることが確認できる。このことから、本研究で付与した聴覚、物理的フィードバックが入力速度の向上

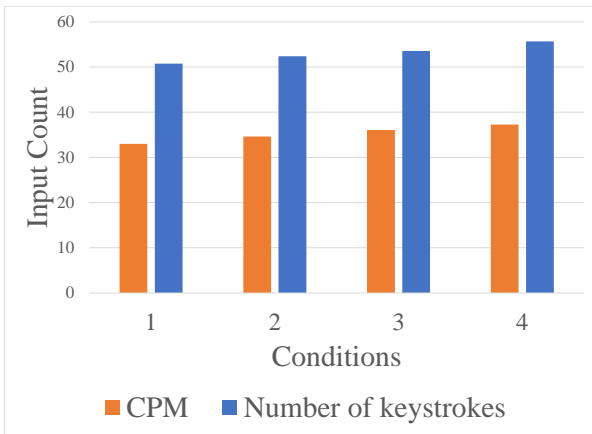


Fig. 8: CPM and number of keystrokes.

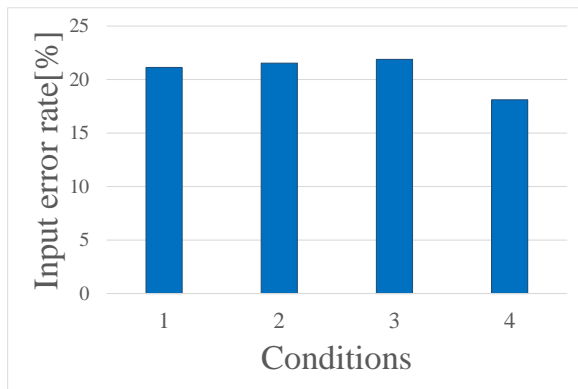


Fig. 9: Input error rate.

に寄与する可能性が示唆された。一方、各条件間に統計的に有意な差は確認できなかった。

誤入力率については、条件3が最も誤入力率が高く、条件2、条件1、条件4の順に誤入力率が低下している。こちらも視覚、聴覚、物理の3種類のフィードバックすべてを付与した条件4が最も誤入力率が低くなった。しかし、各条件間に統計的優位差はみられなかった。

まとめると、条件4（視覚+聴覚+物理）がCPM及び総キー入力回数と誤入力率について最も評価が高くなり本研究で提案したフィードバックの効果が示唆された。一方、いずれの条件間でも統計的に有意な差は見られなかった。

実験中、手の形やフリック入力の癖などの個人差からハンドトラッキングを正確に行えない実験参加者もいることが確認された。そのため、正しくハンドトラッキングが行えるようにフリック入力をするために手を動かすとHMDに搭載されたカメラから手全体を撮影できるようなキーボードのレイアウトを設計するなどの対策が必要と考えられる。

誤入力率の算出に関して、今回の実験では誤入力回数を「正しい入力ではない文字を入力した回数」と設定したため誤入力率が高くなった可能性がある。例えば「とうきょう」という単語を入力しようとした際「き」の入力を正しく行えていないことにユーザが気付かずに「よ」の入力を繰り返し行うなど、どこで間違えたのかユーザが気付いていない場面が散見された。今後の実験では、間違えたタイミングでユーザに気付かせ

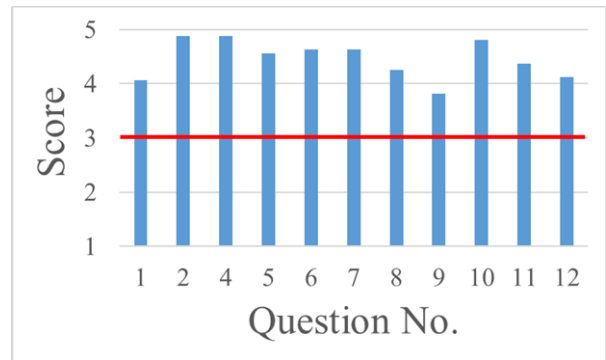


Fig. 10: Results of Questionnaire.

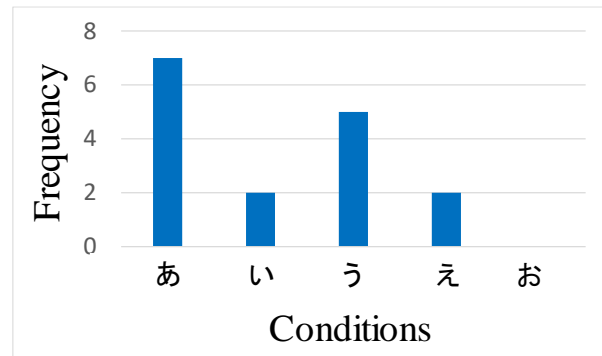


Fig. 11: Answers for Question 3.

る仕組みを導入する必要があると考えられる。

4.5 主観評価の結果と考察

5段階評価を行った項目について結果をまとめたものをFig. 10に示す。評価の中央値である3と有意水準5%でt検定を行った所すべての項目において有意な差が見られた。

設問2で有意な差が確認されたことから、インタフェースの使い方は直感的に理解できるものであったことが示された。また、設問4~11の結果から本研究で実装したフィードバックについて正しく認知され、主観的には効果的だと実験参加者が感じていることが確認できた。さらに設問12の結果から、提案したインタフェースが実験参加者から見て主観的には利用したいものであったことを示した。

設問3の結果を表したグラフをFig. 11示す。設問3の結果から実験参加者が最も入力が難しかったと感じたキーはあ段であることが分かった。これは、あ段を入力するために、領域0の部分に触れている指をキーに対して垂直に手前側に引く必要があり、その際に、他の領域に指が動いてしまうと誤入力になってしまうためだと考えられる。次に入力を行い難いと感じたキーはう段であった。これは、キーボード上部に配置されている「あ」「か」「さ」行において、上側にフリックしようとする手がカメラの測定範囲外に飛び出てしまい、ハンドトラッキングが正確に行えず誤入力が発生していたことが原因だと考えられる。

設問13の最も使いたいインタフェースについての結果をFig. 12に示す。Fig. 12から最も使いたいインタフェースとして先行研究の再現である条件1（視覚的フィードバックのみ）が最も低く、条件2（視覚+音）

と条件3（音+物理）が同等となり、条件4（視覚+音+物理）が最も高い評価を得たことが確認できた。

まとめると、アンケート結果から以下3つのことが確認された。

1. 本研究で付与した3種のフィードバックは実験参加者に正しく認識され、主観的には入力の手助けになっていると感じている。
2. キーの位置や入力時の無意識の手の動き、ハンドトラッキングの精度の問題により「あ段」や「う段」が入力し難いと感じる人がいた。
3. 最も利用したいフィードバックは視覚と音、物理のすべてを組み合わせたフィードバック方法だった。

5 おわりに

本研究では、VR環境での文字入力手法の一つであるフリック入力形式の誤入力率が高いという問題点を改善することを目的として、視覚的フィードバックに加え、音による聴覚的フィードバックと振動による物理的フィードバックを付与し、タイピングゲームによる実験を行った。

実験の結果、客観的評価からは視覚、聴覚、物理の3種のフィードバックを複合した提案手法が視覚的フィードバックのみを行っていた先行研究に比べ入力速度と誤入力率の面で優れている傾向が示唆された。また、アンケートによる主観評価からは付与したフィードバックが入力の助けになっていると実験参加者が感じ、最も利用したいフィードバックが本研究で提案した視覚、聴覚、物理の三種のフィードバックを組み合わせたものであることが示された。以上の結果から、VRフリック文字入力において、本研究で提案した多様なフィードバックを付与することは、誤入力率の低減と操作性を向上させることに寄与すると考えられる。

今後の課題として、キーボードのレイアウトの最適化や実験に用いるタイピングゲームの改良、振動モータの無線化などが挙げられる。

参考文献

- 1) Oculus Quest, <https://www.oculus.com/quest-2/> (参照 2021-6-10)
- 2) 二本松, 中村: ピンチ動作を用いた VR 環境向け文字入力方法, 研究報告コンピュータビジョンとイメージメディア, 2018-CVIM-210, 17, 1/4 (2018)
- 3) 喜多, 小倉, Bista, 高田: Leap Motion を用いた VR 上での文字入力手法の検討, 情報処理学会報告, 2019-HCI-181, 21, 1/7 (2019)
- 4) 長澤: 大学生のスマートフォンと PC での文字入力方法—若者が PC よりもスマートフォンを好んで使用する理由の一考察—, コンピュータ&エデュケーション, 43, 67/72 (2017)
- 5) 福仲, 謝, 宮田: VR 環境におけるフリック入力形式インタフェースの開発, 情報処理学会報告, 2019-HCI-182, 3, 1/8 (2019)
- 6) 小澤, 梅沢, 大沢: 空中におけるつまむ動作を用い

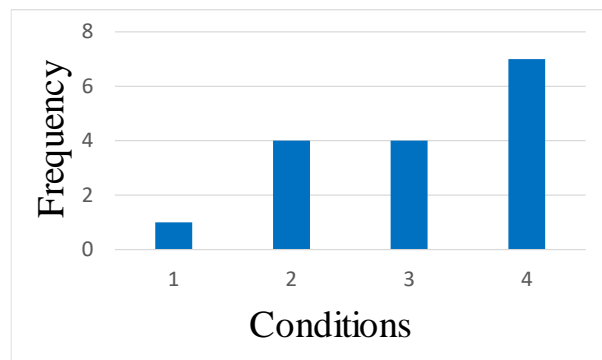


Fig. 12: Answers for Question 13.

た効率的な文字入力の検討, 情報科学技術フォーラム講演論文集, 14, 3, 389/390 (2015)

- 7) 和氣, 今井, 西崎, 光本, 長田: タッチパネル操作における操作反応音の有効性～視覚フィードバック有無の観点から～, ヒューマンインタフェース学会論文誌, 19, 1, 61/68 (2017)
- 8) Wu, Hsu, Lee, Smith: Virtual Reality Keyboard with Realistic Haptic Feedback in a Fully Immersive Virtual Environment, Virtual Reality, 21, 1, 19/29 (2017)

視認推定に基づく視覚探索タスク支援のための VR環境における有効性評価

吉田 峻亮 (兵庫県立大学) 施 真琴 (ATR/立命館大学)
内海 章 (ATR) ○山添 大丈 (兵庫県立大学)

Gaze-based estimation of target visual cognition and its application to visual search assistance

S. Yoshida (University of Hyogo), M. Sei (ATR/Ritsumeikan University),
A. Utsumi (ATR) and *H. Yamazoe (University of Hyogo)

Abstract— To assist tasks that require visual search, such as driving a car, guiding people’s attention is important. In particular, in situations that the surrounding environment change from time to time, it is important to provide attention guidance at the appropriate timings. For this purpose, we need to estimate whether the person has visually recognized the target objects (target visual cognition) or not, and to provide appropriate support for visual search tasks in response to the visual cognition. In this research, we propose a method for efficient assistance of visual search tasks by guiding attention based on target visual cognition estimation, assuming a scene in which multiple visual search tasks are required continuously. In this paper, we evaluate the effectiveness of the assistance of visual search tasks based on visual cognition using a VR game environment.

Key Words: Target visual cognition, Assistance of visual search task, Visual cognition difficulty, Virtual reality environment

1 はじめに

自動車運転などの様々なタスクにおいて、人は視覚・聴覚など様々な感覚を通じて情報を収集し、周辺状況を把握しながら、それに基づき適切な行動を決定している。一般に、視覚を通じて得られる情報は特に多く、人は全ての視覚情報を取得し処理することができないため、適切に注意を配分することで、情報の取捨選択を行っている。例えば、交通量の多い交差点で右折する状況を考えると、運転手は信号や対向車、歩行者、右折先の道路環境等に様々な対象に注意を払う必要がある。しかし、全ての対象に同時には注意を向けることはできないため、変化する状況に対応しながら、注意すべき対象に適切なタイミングで注意を向ける必要がある。本研究では、こういった自動車運転のようなタスクにおいて、適切に注意誘導を行うことでタスクの支援を行うことを検討している。

人の注意を誘導する手法については、これまでに多くの研究が行われている^{1, 2, 3)}。自動車運転時のように時々刻々と変化する周辺状況において効率的に注意誘導を行うためには、適切なタイミングで注意誘導を行うことが必須である。そのためには、人(運転手)が注視対象を認識したかどうか(視認)を推定し、それに対応して、適切に注意誘導の対象を切り替えることが重要となる。

そこで本研究では、上で述べた交差点での右折シーンのような、複数の対象を連続して探索し、注意を向ける必要があるシーン(複数の視覚探索タスクを連続して行うシーン)を想定し、複視認推定に基づいて注意誘導を行うことで、効率的な視覚探索タスク支援の枠組みの実現を目指す。支援を行うタスクとして、Fig.1に示すVRゲームを作成した。このゲームでは、実験参加者(プレイヤー)はアルファベットが表示された大



Fig. 1: Experimental environment

量のBoxが接近してくるシーンにおいて、指定されたターゲットのBoxを発見し、そのBoxのみを切ることがタスクとなる。前のターゲットを切った後に、次のターゲットへの注意誘導を行うのではなく、前のターゲットを視認した時点で、次のターゲットへの注意誘導を行うことで、より早い段階での注意誘導を実現し、それにより効率的なタスク支援を目指すのが、提案手法の基本的なアイデアである。

本稿では、上述のVRゲーム環境を対象として、視認難易度に基づく視認推定と、それをを用いた視覚探索タスク支援を提案し、実験によりその有効性を示す。

2 視認推定に基づく視覚探索タスク支援

本章では、本稿で提案する視認推定に基づく視覚探索タスク支援の枠組みについて述べ、その前提となる視認推定について簡単に紹介する。



Fig. 2: Preliminary experimental environment

2.1 視認推定

人が注視対象を視覚的に認識したかどうかを推定する視認推定について、これまでにいくつかの研究が行われている^{4, 5, 6, 7)}。視認を推定する基本的な考え方は、注視対象の領域内を一定時間以上注視したかどうかで視認の有無を推定するというものである^{4, 5)}。しかし、視認推定に一定時間の注視が必要なため、視認推定に時間がかかるという問題がある。また、注視対象の視覚特性に関わらず、一定の条件で視認の有無を推定しており、対象の視覚特性が大きく異なる場合には、正確に視認推定ができないといった問題もある。

これに対し、阪口らは注視対象の視覚特性を踏まえた、より正確な視認推定を実現するため、注視対象の視認の難しさ(視認難易度)と視認の関係を分析した^{6, 7)}、その結果、視認難易度が低い(見やすい)対象の場合には、対象と注視方向との距離が大きい段階で視認ができていないのに対し、視認難易度が高い(見にくい)場合には、対象と注視方向との距離が小さくならないと視認ができないといったように、視認難易度によって視線の挙動が異なることを明らかにした。本研究では、これらの知見に基づく視認推定を行い、人の視認状態に応じて注意誘導を行うことで、効率的な視覚探索タスク支援を目指す。

2.2 視認に基づくタスク支援

次に、本研究で提案する視認に基づくタスク支援について説明する。例えば、自動車運転においては、案内標識を発見・認識すると、それに従って進路を変更し、その後、進行方向に対応する標識を発見し、さらに進路を変更する、といったように、次々と現れる案内標識や信号など、適切に発見し、それに基づき行動を変化していく必要がある。

本研究では、このような複数の視覚探索タスクを連続して行うシーンを想定し、視認推定に基づく注意誘導を行うことで、効率的な視覚探索タスク支援の枠組みを目指している。視認推定に基づき注意誘導の対象を切り替えていくことで、視認後の行動の完了を待つことなく、より早い段階で次の注意を向けるべき対象への注意誘導が可能となり、効率的なタスク支援が実現できる。

3 視認難易度に基づく視認推定のための予備実験

視認推定に基づく視覚探索タスクの支援を実現するための視認推定手法としては、Sakaguchiらの視認難易度に基づく視認推定⁶⁾の知見を利用する。ただし、実験環境が異なり対象の視覚特性が異なるため、そのま

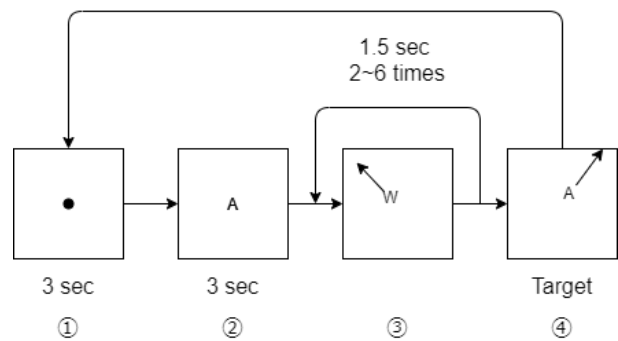


Fig. 3: Flow of preliminary experiment

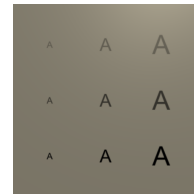


Fig. 4: Difference of visual cognition difficulty

ま本研究に適用することはできない。そこで、今回実装したVRゲーム(Fig.1)において視認推定を実現するために、視認難易度と視認時の視線行動との関係を分析する予備実験を行った。

3.1 予備実験の概要

予備実験では、実験参加者は椅子に座っており、VR環境内(Fig.2)で実験参加者前方5[m]の壁平面上に移動する文字が出現するシーンにおいて、指定された文字が表示された時にコントローラのボタンを押下するタスクを行ってもらい、異なる視認難易度の文字を表示し、その際の視線行動を記録することで、VR環境中の文字探索タスクにおける視認難易度と視線行動との関係性を明らかにし、その関係性をVRゲームにおける視認推定に利用する。

実験環境の構築には、HMD(HTC VIVE Pro Eye)、VR用コントローラ(HTC VIVEコントローラ2018)を利用し、Unity(2019.4.10f1)、VR用アセット(SteamVR、ViveSR)を用いて実装した。実験中、HMDの位置・姿勢が仮想環境中の仮想カメラの位置・姿勢として反映される。上述の通り、仮想空間中の実験参加者前方5[m]に壁が存在し、壁の中心点位置(文字の出現位置)は実験開始時にHMDと等しい高さに初期化した。

3.2 予備実験の手順

Fig.3に予備実験の流れを示す。まず、①壁の中心に黒い注視点が3秒間表示され、それを注視してもらう。その後、②ターゲットとなる文字が中心に3秒間表示される。消失後、③1.5秒周期で、ターゲットではない文字の出現と中心から離れる方向への移動が2~6回繰り返される。最後に、④ターゲットとなる文字を出現し、同様に中心から離れる方向に移動する。実験参加者には、ターゲットとなる文字を視認したタイミングで、コントローラのボタンを押下してもらうよう指示しており、以上で1回の試行が終了となる。ステップ③の文字の出現・移動では、壁上の中心点を中心とする半径0.5[m]の円上にランダムに文字が出現し、中心から離れる方向に秒速3[m]で移動する。中心から

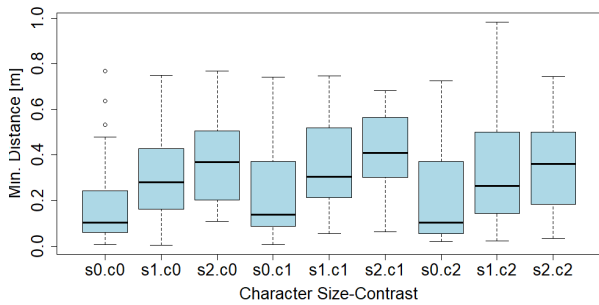


Fig. 5: Minimum distance between target and gaze with respect to visual cognition difficulty (VCD)

Table 1: ANOVA results of minimum distance between target and gaze

	Df	SS	MS	F	p
size	2	1.02	0.51	11.61	<.01 *
cont	2	0.03	0.02	0.34	0.71
size*cont	4	0.09	0.02	0.50	0.73

*: $p < .01$

2.5[m] 離れると文字が消失する。

視認難易度としては、フォントサイズ3段階 (s0: 16, s1: 32, s2: 48)、コントラスト (色の α 値) 3段階 (c0: 0.3, c1: 0.6, c2: 1.0) を組み合わせた計9段階とした。Fig.4に視認難易度の違いによる見え方の違いを示す。ターゲットとなる文字は、VRゲームでの利用を想定し、大文字と小文字の判別がしやすい9文字 (A, B, D, E, F, G, H, J, M) とした。Fig.3の試行9回の繰り返しを実験1セットとし、実験1セットには、ターゲット文字と視認難易度をそれぞれ1回ずつを含む。ターゲット文字と視認難易度の順序は、実験参加者間でカウンターバランスを取った。

実験には、21~23歳の男性5名が参加した。実験の概要を説明し、実験参加への同意を口頭で取った後、視線取得のためのキャリブレーションを行った。その後、練習試行を行ってもらった後、実験5セットを行ってもらった。

3.3 結果と考察

取得データから、文献^{6,7)}と同様に、文字が提示される壁平面上でのターゲット文字の位置と注視位置との最接近距離を計測した。また、ボタン押下の時点では視認ができていないことから、文字提示から視認までの時間として、ボタン押下時間を計測した。視認難易度条件ごとの最接近距離の結果を Fig.5 に、視認までの時間の結果を Fig.6 に示す。

これらの結果に対し、それぞれ2要因分散分析を行った。結果を表1と表2に示す。最接近距離については、ターゲットの文字サイズ (size) に関して、視認までの時間については、ターゲットの文字サイズ (size) とコントラスト (cont) の両方で有意差が確認できた。以上より、視認難易度が高くない場合 (ターゲット文字が大きい、コントラストが高い) 場合には、ターゲットと注視方向の角度が大きい段階や、より早いタイミングで視認可能であることがわかった。

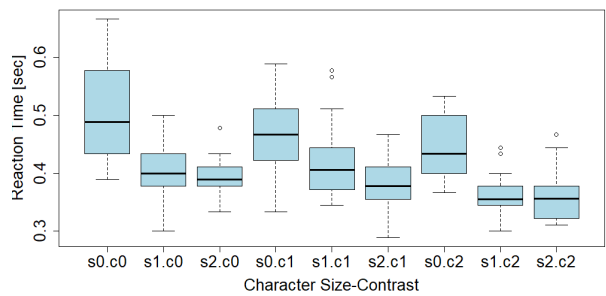


Fig. 6: Duration from appearance of target to subject's recognition of target with respect to VCD

Table 2: ANOVA results of time from appearance of target to subject's recognition of target

	Df	SS	MS	F	p
size	2	0.35	0.17	58.28	<.01 *
cont	2	0.07	0.04	12.16	<.01 *
size*cont	4	0.02	0.01	1.84	0.12

*: $p < .01$

4 視認に基づく視覚探索タスク支援の評価実験

次に、視認に基づく視覚探索タスク支援の有効性を示すため、VRゲームにおけるタスク支援実験について述べる。

4.1 VRゲームの概要

実験対象のVRゲームでは、Fig.7-8に示すように、実験参加者 (プレイヤー) に対してアルファベットが表示された大量のBoxが接近してくるというシーンにおいて、対象となる文字が表示されたBoxを発見し、これをVRコントローラで切ることが目的である。

本研究では、複数の視覚探索タスクを連続して行う状況を対象とするため、2つのターゲットを連続して切るタスクとしている。Fig.7に示すように、Boxに表示されるアルファベットの大半は小文字であるが、いくつかの大文字が表示されたBoxが出現する。これが1つ目のターゲットである。1つ目のターゲットの出現から一定時間 (0.5秒または1.0秒) 経過後には、1つ目のターゲットに表示されたアルファベットの小文字が表示されたBoxが出現する (Fig.8)。これが2つ目のターゲットであり、実験参加者は1つ目のターゲットを発見し切った後に、対応する2つ目のターゲットを発見し切るというタスクを繰り返し行ってもらう。1つ目のターゲットは同時に複数出てくるが、同時に複数のターゲットペアを対象にすることはできず、1つ目のターゲットを切った後は、それに対応する2つ目のターゲット切らなければならないこととした。つまり、AとBが同時に出現したときに、A-B-a-bという順番で切ることは許されず、A-aもしくはB-bと切るように指示した。

4.2 ゲームにおけるタスク支援

タスク支援では、2つ目のターゲットとなる小文字が表示されたBoxの色を変更することで、ターゲットを視認しやすくする。以下では、この色の変更を支援表示と呼ぶ。以下の4種類のタスク支援を実装した。

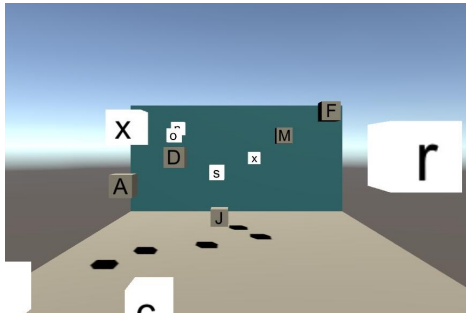


Fig. 7: Scenes including 1st targets
(Uppercase letters A, D, J, F, M appear)

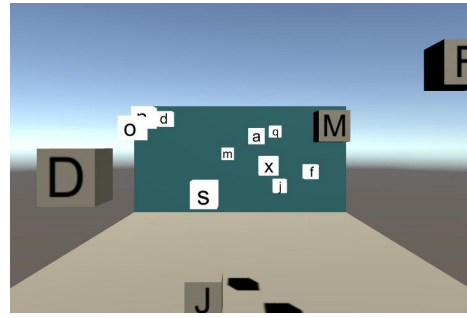
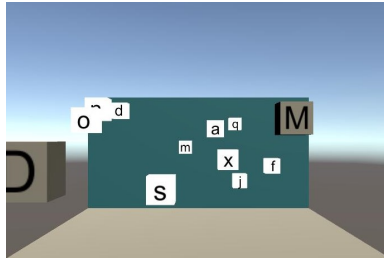
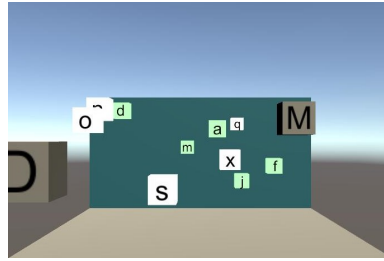


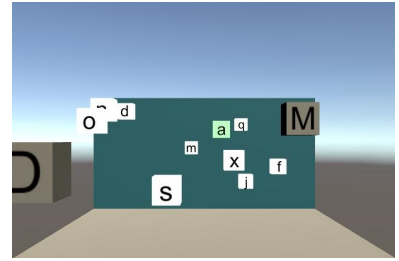
Fig. 8: Scenes including 2nd targets
(Lowercase letters a, d, j, f, m appear)



(a) without assistance



(b) assistance to all target candidates



(c) assistance to target corresponding to selected 1st target

Fig. 9: Assistance to 2nd target candidates (Assisted target candidates are shown as green boxes)

- (a) タスク支援を行わない場合 (None)
- (b) 2つ目のターゲット候補すべてに支援表示 (All)
- (c-1) 1つ目のターゲットを切った時点で、対応する2つ目のターゲット Box のみに支援表示 (Cut)
- (c-2) 1つ目のターゲットを視認後に、対応する2つ目のターゲットのみに支援表示 (Gaze)

記号はそれぞれ Fig.9 の記号に対応し、(c-1) と (c-2) の違いは、支援表示のタイミングの違いである。(c-2) が本稿で提案する視認に基づくタスク支援である。

4.3 実験環境と実験手順

予備実験と同じシステム構成で、実験環境を構築した。VR 環境中には、HMD とコントローラの位置・姿勢が反映される仮想のカメラ、コントローラが存在し、HMD の移動に伴って実験参加者は VR 空間中を移動できる (Fig.1)。コントローラにはその中心から Z 軸 (正面) 方向へ延びる長さ 1.05m, 幅, 高さ 2.5cm の直方体の当たり判定を持つ光の棒が固定されており、実験参加者はこれをターゲットに接触させることでターゲットの Box を切る。

ターゲットとなる Box は 1 辺が 25[cm] の立方体であり、基準点 (ゲーム開始時の実験参加者位置) から Z 軸 (正面) 方向 10[m] の位置にある壁から出現する。壁の中心 (高さ 1.25m) から縦 ± 0.75m, 横 ± 1.25m の範囲から 0.1 秒ごとにランダムに出現し、5[m/s] で実験参加者に接近する。その際、Box の X 座標 (横方向)、Y 座標 (縦方向) は変化しない。アルファベットは実験参加者から見える面に表示される。ターゲットとなる Box に表示されるペアのアルファベットとしては、大文字と小文字の判別がしやすい 9 文字 (A, B, D, E, F, G, H, J, M) を用いている。1 目目のターゲット候補は、同時に 5 つ出現し、一定時間 (0.5 秒または

Table 3: Threshold angles for visual cognition estimation

Font size	Angle [deg]
16	2.46
32	3.86
48	4.38

1.0 秒) 後に小文字が出現する。Box の色は、3 色に分かれており、灰色が 1 目目のターゲットとなる大文字が表示された Box, 緑が支援表示された Box であり、その他の Box は白色である。

この実験における視認難易度としては、予備実験で最接近距離 (角度) に関して有意差が確認できたフォントサイズのみを考慮することとした。予備実験と同様にフォントサイズ 3 段階 (s0: 16, s1: 32, s2: 48) を準備し、実験条件によって 1 目目のターゲットに表示される大文字のフォントサイズを変化させた。

視認については、前章の予備実験で得られた関係性をを用いる。具体的には、実験参加者の視線方向と実験参加者から見たターゲット Box の方向とのなす角が、視認難易度ごとに設定された基準角度よりも小さくなった場合に視認と判断する。Table 3 に視認難易度 (フォントサイズ) と基準角度の関係を示す。基準角度は予備実験で得られた最接近角度の結果に基づき計算したものである。

4.4 実験

実験には 21~23 歳の男女 15 名 (女性 1 名, 男性 14 名) が参加した。実験の前には、実験参加者に実験内容等について説明し、口頭で実験参加へ同意を得た。説明の際、支援については、ゲーム内で緑色の Box 表示として支援がされることのみを説明し、支援が行われる基準については説明しなかった。また、1 番目のターゲットとなる大文字の Box が灰色であることも説明した。その後、視線取得のためのキャリブレーションと

Table 4: Confusion matrix of visual cognition (VC) estimation

	VC	non-VC	total
cut targets	1227	14	1241
other targets	1663	3396	5059
total	2890	3410	6300

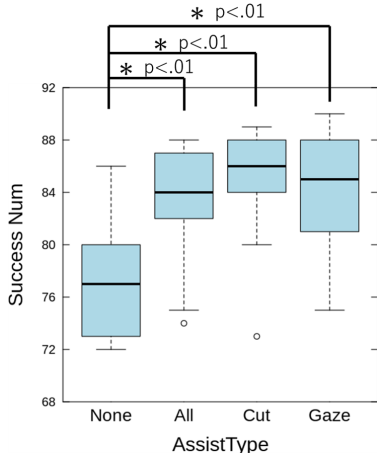


Fig. 10: Number of task completions

練習の後、実験を行った。

実験1セットでは、ターゲットとなる大文字・小文字のペアが視認難易度3段階ごとにそれぞれ10回ずつ、計30ペアが出現する。実験参加者は、4種類の支援についてそれぞれ3回ずつ、計12セットの実験を行い、支援方法や視認難易度の順番はカウンターバランスを取った。実験中の視線方向と頭部姿勢、切断行動に関するデータを取得した。

以下では、タスク支援によるタスク達成に対する影響として、タスクの成功数とタスク達成までの時間に着目して分析を進める。また、実験参加者の1名が、本実験における支援の意味を正しく理解していなかったため、この参加者に関するデータを分析から除外した。

4.5 実験結果

4.5.1 視認推定精度とタスク成功数の変化

ここでは、提案するタスク支援によるタスク成功数への影響を分析するが、その前提となる視認推定の精度を評価する。視認後に支援を行う Gaze 条件のデータから、視認推定の精度を評価した。同時に出現する5つのターゲット候補の内、切ったものを視認対象、それ以外を非視認対象とし、それらに対して視認と判断したかどうかを評価した。混同行列を Table 4 に示す。結果より、視認推定の正解率は73.4%となった。再現率は98.9%となっており、実験参加者が切った対象はほぼすべて視認対象として推定できていることがわかる。一方で、非視認対象に対する誤認識が多いことから正解率が低下している。この原因として、ターゲット候補5つのうち、実験参加者が切った対象のみを視認対象として精度を計算したために、実際には切っていない対象も視認している可能性が考えられる。以上より、視認難易度と最接近角度との関係に基づく視認推定によって、実験参加者の視認が推定できることが確認できた。

次に、支援方法によるタスク成功数(ペアを正しい順番で切れた数)への影響について分析する。結果を Fig.10

Table 5: ANOVA of task completions

	Df	SS	MS	F	p
assist	3	486.2	162.07	8.38	<.01 *

*: p<.01

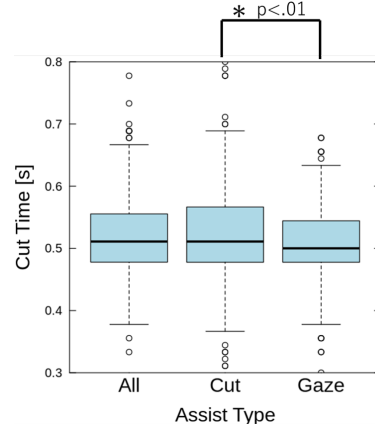


Fig. 11: Duration between two-target cuttings

に示す。この結果に対して、分散分析を行った結果を表5に示す。支援方法(assist)の条件間に有意差が見られたため、さらに多重比較を行った。結果は Fig.10 に示しており、支援なし(None)と支援あり(All,Cut,Gaze)の間には有意差が見られたが、支援ありの条件間では有意差が見られなかった。このことから、支援表示によるタスク成功数への支援効果は確認できたものの、視認に基づく支援の効果は確認できなかった。

4.5.2 タスク完了時間と支援タイミング

次に、支援によるタスク完了時間への影響を分析する。支援条件3種類(All, Cut, Gaze)について、1つ目のBoxを切ってから2つ目のBoxを切るまでにかかった時間を分析した。結果を Fig.11 に示す。また分散分析を行った結果を Table 6 に示す。有意差が見られたため多重比較を行った結果も Fig.11 に示す。結果より、Cut 条件と比べて視認後に支援の Gaze 条件の間に有意差が見られ、差はわずかではあるものの、1つ目のBoxを切ってから2つ目のBoxを切るまでにかかった時間が短くなることが確認できた。

次に、タスク完了時間に有意差が見られた2条件(Cut 条件, Gaze 条件)について、さらに分析を行った。まず、支援が決定するタイミングの比較として、2つ目ターゲットの出現から、そのターゲットへの支援表示が行われるまでの時間差を評価した。ここでは、支援方法と視認難易度(フォントサイズ Size)の2要因を考慮した。結果を Fig.12 に、2要因分散分析の結果を Table 7 に示す。また交互作用(size * assist)が確認されたため、単純主効果の分析を行った。結果を Fig.12 に示す。結果より、Cut 条件と比べて視認に基づき支援を行う Gaze 条件で支援タイミングが有意に早く、さらに、Gaze 条件で視認難易度が低い(文字が大きい)ほど支援タイミングが有意に早いことがわかる。以上よ

Table 6: ANOVA results of duration between two-target cuttings

	Df	SS	MS	F	p
assist	2	0.04	0.02	3.90	<.05 *

*: p<.01

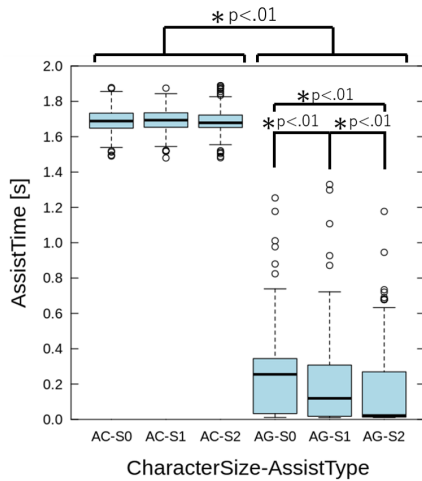


Fig. 12: Duration from target appearance to task assistance

Table 7: ANOVA results of duration from target appearance to task assistance

	Df	SS	MS	F	p
size	2	1.48	0.74	29.64	<.01 *
assist	1	1323	1323.2	14956.7	<.01 *
size*assist	2	1.01	0.50	26.06	<.01 *

* : p<.01

り、視認に基づく支援によって早い段階で支援が可能となることが示された。

次に、1つ目のターゲットを切ってから2つ目のターゲットを切るまでの時間差について詳細に分析する。ここまでの分析で、タスクが容易な場合にはタスク支援の効果が現れにくい可能性があるため、1つ目と2つ目のターゲットの出現間隔が0.5秒のターゲットペアに着目し。また、ターゲットの出現位置によってもタスクの難易度が異なると考えられるため、2つ目のターゲットの出現位置を中心からの距離によって3段階 (D0: 1m 以内, D1: 1-2m, D2: 2m 以上) に分けて分析を行った。結果を Fig.13 に、2要因分散分析を行った結果を Table 8 に示す。結果より、2つ目のターゲットが周辺領域に出現する場合には、2つ目のターゲットを切る時間が有意に遅くなること、Cut 条件と比べて Gaze 条件では、2つのターゲットを切る時間差が有意に早くなることが確認できた。以上より、視認に基づく支援によってタスクのパフォーマンスが向上することが確認できた。

5 おわりに

本稿では、VR 環境での視覚探索タスクを対象として、視認難易度に基づく視認推定と、それに基づく視覚探索タスク支援を提案し、その有効性を評価した。実験結果より、視認推定に基づく注意誘導によって、早

Table 8: ANOVA results of detailed analysis of duration between two-target cuttings

	Df	SS	MS	F	p
assist	1	0.04	0.04	15.14	<.01 *
dis	2	0.54	0.27	107.94	<.01 *
assist*dis	2	0.01	0.01	1.85	0.18

*: p<.01

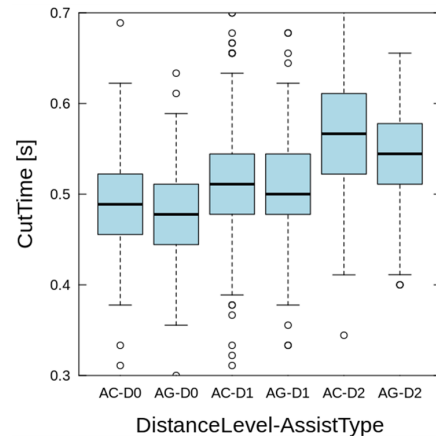


Fig. 13: Detailed comparison of duration between two-target cuttings

い段階での注意誘導が可能となること、また視認に基づく視覚探索タスク支援によって、タスクのパフォーマンスが向上することを確認した。このことから、提案する視認推定に基づく視覚探索タスク支援を行うことで、自動車運転のように時々刻々と変化する環境において、適切に注意誘導の対象を切り替えることが可能となることが期待できる。

本稿で述べた視認推定では、視認難易度と最接近角度の関係のみを利用しているが、坂口らの研究^{6,7)}では、視認と視認対象に対する追従時間や最接近時の移動方向の差の関係も示されており、これらの関係も含めた分析を進めていく。さらに、自動車運転などのより現実に近いシーンにおけるタスク支援についても検討を進めていく。

謝辞

本研究は、JSPS 科研費 18K11383, 20K21824, 21K11968 の助成を受けて実施した。

参考文献

- 1) Maglio, P. P. and Campbell, C. S.: Tradeoffs in displaying peripheral information, *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 241-248 (2000).
- 2) Ono, F., Okada, R., Takada, H. and Matsuura, Y.: Visual warning method using attentional induction by light flashing, *FISITA World Automotive Congress 2008, Congress Proceedings-Mobility Concepts, Man Machine Interface, Process Challenges, Virtual Reality*, Vol. 1, pp. 154-162 (2008).
- 3) Arun, S., Murugappan, M. and Sundaraj, K.: Hypovigilance warning system: A review on driver alerting techniques, *ICSGRC 2011*, pp. 65-69 (2011).
- 4) Fletcher, L. and Zelinsky, A.: Driver inattention detection based on eye gaze—Road event correlation, *The international journal of robotics research*, Vol. 28, No. 6, pp. 774-801 (2009).
- 5) Tawari, A., Møgelmoose, A., Martin, S., Moeslund, T. and Trivedi, M.: Attention estimation by simultaneous analysis of viewer and view, pp. 1381-1387 (2014).
- 6) Sakaguchi, H., Utsumi, A., Susami, K., Kondo, T., Kanbara, M. and Hagita, N.: Analysis of relationship between target visual cognition difficulties and gaze movements in visual search task, *SMC 2017*, pp. 1423-1428 (2017).
- 7) 阪口栄穂, 内海 章, 須佐見憲史, 近藤公久, 神原誠之, 萩田紀博: 視線挙動と対象の視認難易度に基づく視認推定手法の検討, *信学技報*, pp. HIP2017-96 (2018).

ファジィ測度を用いたアクティブラーニングによる効率的な識別学習

○木村慶豪 濱上知樹 (横浜国立大学)

Efficient classification learning using Fuzzy Measured Active Learning

*Keigo Kimura and Tomoki Hamagami (Yokohama National University)

Abstract– One of the learning methods for efficient learning of classification of specialized images is active learning, in which the data with the highest learning efficiency is interactively queried. The efficiency of active learning depends greatly on query strategy. In this study, there are two types of data that are necessary to improve the accuracy of the model. In order to obtain the data explicitly, we proposed a query strategy using fuzzy measure which quantifies the certainty of attribution. The results of active learning on multiple datasets showed that the proposed strategy can provide more efficient data than existing strategies.

Key Words: Active Learning, Fuzzy Measure, Assisted Reproductive Technology

1 はじめに

様々な分野で画像データが蓄積されるのに伴い、画像中の物体識別を機械学習を用いて学習し、利用者の判断を支援をする需要が増大している。

機械学習による画像中の物体識別では物体のクラスをラベル付けするアノテーションを必要とする。しかし、専門画像のアノテーションは高度な知識を必要とするため作成できるラベル付きデータの数が限られる。また、十分にデータが取得できたとしてもカメラの距離に起因する被写体の大きさ、ライティング、解像度によって画像特徴が変化する。このため学習データが撮影環境に依存する課題があり、撮影環境ごとに新たにアノテーションを行う必要がある。

よってアノテーション回数を減らすため、少数のラベル付きデータで利用者の目的に合致した機械学習モデルを作成できる効率的な学習手法に需要がある。

利用者の目的に合致したモデルを少数のラベル付きデータで学習させるためには、利用者と機械学習モデルが対話を行い、逐次的にアノテーションする学習手法であるアクティブラーニングが有効であると考えられる。本稿では、アノテーションすべきデータを明示的に取得するためにファジィ測度に着目した新しいアクティブラーニングを提案する。

2 関連技術

2.1 半教師あり学習

半教師あり学習は少量のラベル付きデータと大量のラベルなしデータを用いる学習手法である¹⁾。画像データのように、蓄積は容易であるがラベル付けに時間と労力が必要なデータの場合に半教師あり学習が有効となる。

半教師あり学習の手法に自己訓練や共訓練がある¹⁾。自己訓練では少量のラベル付きデータを学習し、その後ラベルなしデータを推論する。推論の確度が高いデータにモデル自らがラベル付けを行う。共訓練では2つのモデルを用いてそれぞれを独立した特徴量で学習させる。学習後、それぞれのモデルはラベルなしデータを推論し、推論の確度が高いデータにラベルを付与し、ラベル付きデータに追加する。

半教師あり学習を活用できる分野の1つに生殖医療支援における精子評価システムがあげられる。生殖補助医療の一種である顕微授精は採取された精子から正常な形状かつ運動性に優れた精子を選定し、直接卵子に注入する手法である。この作業は胚培養士が行うが、顕微鏡中から精子を発見し評価を行うため多大な時間と専門的な技術を必要とする。顕微授精において顕微鏡画像の撮影は容易である一方、撮影された精子のラベル付けには胚培養士の高度な知識と時間が必要であるため大量のアノテーションが難しい。そこで半教師あり学習を適用することで、少ないアノテーション回数で胚培養士の支援が可能な機械学習モデルを作成できると考える。

自己訓練や共訓練は少量のラベル付きデータと大量のラベルなしデータで学習できる一方で、データの追加はモデルの推論に依拠する。このため推論が誤っていた場合には反対に汎化性能の低下に繋がる。また、推論の確度が低いデータはラベル付きデータに加わらないためそのようなデータを学習する回数が少なくなる。顕微鏡画像のように特徴量が多い画像データではこの課題が顕著に現れると予想される。以上から機械学習モデルの精度向上には推論が不確かであったデータを選択し、対応するクラスを利用者に問い合わせることがより重要であると考え、アクティブラーニングに着目した。

2.2 アクティブラーニング

アクティブラーニングはモデルの性能向上に寄与すると考えられるデータを選択し、クラスを神託(Oracle)に問い合わせる半教師あり学習の一種である。モデルの性能向上に有効なデータはデータセット中の少数であり、そのようなデータを優先的にラベル付けすることで少ない教師データ数で十分な汎化性能を出せる仮定に基づいている。アクティブラーニングの手法の1つにPool Basedがある。Pool Basedな手法をFig.1に示す。

Pool Basedな手法では、学習データはすでにラベルが付与されているLabeled Pool、ラベルが付与されていないUnlabeled Poolに分けられる。モデルは最初にLabeled Poolを学習する。その後Unlabeled Poolを推

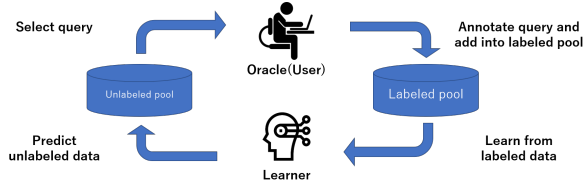


Fig. 1: Pool Based Active Learning

論し、その中で精度向上に有効とされるデータを選択し、そのデータが属するクラスを Oracle に問い合わせる。Oracle はデータにアノテーションを行い、Labeled Pool に追加する。追加された Labeled Pool をモデルが再度学習する。

モデルが Unlabeled Pool からデータを選択し、Oracle に問い合わせ、Oracle がアノテーションを行う一連の作業を繰り返し行うことで、アノテーション回数を抑えつつ精度向上に有効となるデータのみを教師データとする。

精度向上に有効とされるデータの選び方を Query Strategy と呼ぶ。タスクや学習済みのラベル付きデータセットの数に応じて最も有効となる Query Strategy は変化する。

2.3 ファジィ測度

ファジィ測度とは帰属の確かさを表す測度である。データ x のクラス c_i へのファジィ測度を $g(c_i|x)$ と表現し、 $[0, 1]$ の値を取る。ファジィ測度が大きいほどデータがクラスへ帰属している確かさが大きい。ファジィ測度は次の性質を満たす。

$$g(\emptyset) = 0 \quad (1)$$

$$g(\Omega) = 1 \quad (2)$$

$$E \subset F \Rightarrow g(E) \leq g(F) \quad (3)$$

ファジィ測度は加法性を仮定しない。そのため全てのクラスに対するファジィ測度を総和したときに1になるとは限らないため、確率測度よりも推論の様子を正確に表現できる。

3 提案手法

ラベルなしデータは十分に存在すること、撮影環境に応じて画像の特徴が変化する課題に着目して学習する画像データ毎にモデルを作成し、精度改善に有効なデータを逐次的に加えるアクティブラーニングを用いた学習モデルを採用し、また必要となるデータの特性からファジィ測度を用いた Query Strategy を提案する。

モデルの精度改善に必要なデータを考察する。少量のデータセットで学習したとき、識別が困難となるデータは決定境界近傍や未学習の領域のデータである。提案手法はそのような2種類の識別が困難なデータを追加して学習器の推定する密度分布を改善することを目的とする。前者はクラス間の密度分布の分離に寄与し、後者はクラスの密度分布の拡大に寄与する。

代表的な Query Strategy に Uncertainty Sampling がある。この手法はラベルなしデータを学習器で推論した事後確率 $P(y|x)$ を用いて各データに対する推論の不確か度を計算する²⁾。不確か度が高いデータを判別が難しいデータとみなして Query とする。しかし、確率測度を用いた Uncertainty Sampling は密度分布の分

離と拡大に寄与するデータを区別できない。例えば2つのデータが存在し、一方はすべてのクラスに属していると判断したデータ、もう一方はどのクラスにも属していないと判断したデータとする。この場合、確率測度では同じ値となる。

このように異なる不確かさを確率測度では混同するため Query が一方のデータに偏る場合がある。

そこで、2種類の精度改善に寄与するデータを明示的に Query とするためにファジィ測度を用いた Query Strategy を提案する。各クラスへのファジィ測度 $G(\mathbf{x})$ は式4のように計算される。

$$G(\mathbf{x}) = \{g(c_1|\mathbf{x}), g(c_2|\mathbf{x}), \dots, g(c_N|\mathbf{x})\} \quad (4)$$

ファジィ測度は確率測度と異なり加法性を仮定しない。このため、全てのクラスに属するようなデータに対しては高くなるように、どのクラスにも属さないようなデータに対しては低くなるようにファジィ測度が算出される。前者は分布の分離に寄与するデータであり、後者は分布の拡大に寄与するデータであるためファジィ測度によって2種類の精度改善に寄与するデータを明示的に取り出すことが可能となる。本手法では分布の分離に寄与するデータ \mathbf{x}_{sep}^* と拡大に寄与するデータ \mathbf{x}_{exp}^* を(5)式で定義する。

$$\mathbf{x}_{sep}^* = \arg \max_x \left\{ \sum_i g(c_i|\mathbf{x}) - \max_i g(c_i|\mathbf{x}) \right\} \quad (5)$$

$$\mathbf{x}_{exp}^* = \arg \min_x \{ \max_i g(c_i|\mathbf{x}) \}$$

分布の分離に寄与するデータの指標として各クラスへのファジィ測度の合計値から最大のファジィ測度を引いた値を使用する。値が大きいほど複数のクラスへの帰属の確かさが大きいと意味しており、判断が難しいデータであったといえる。また、分布の拡大に寄与するデータの指標として最大となるクラスへのファジィ測度を使用する。最大のファジィ測度が小さいほどいずれのクラスにも属していないと判断しているため未学習の領域に存在するデータであったといえる。本手法では式5で示した2種類の不確か度に従って Query を取り出す。

4 実験

4.1 実験1

入力次元数や難易度の異なる、様々な分野のデータセットを使用して比較実験を行った。

4.1.1 目的

アクティブラーニングの有効性は学習するデータセットの入力次元数や難易度に応じて変化する。入力次元数や難易度に依存せずに包括的な比較を行うために行った。また提案手法が幅広いデータセットにおいて、ラベル付きデータの少ない学習初期でより学習効率の高いデータを与えられるかを確認するために行った。

4.1.2 データセット

Table.1に示す10のデータセットを使用した。これらのデータセットはUCI Machine Learning Repositoryに掲載されている³⁾。

Table 1: Datasets for experiment 1

Name	#Instances	#classes	#attributes	#nominal
Abalone	4177	3	8	1
Car Eval	1728	4	6	6
Census Income	48842	2	14	8
Coverttype	581012	7	54	0
Dry Bean	13611	7	17	0
internet	65532	4	12	0
Letter Recog	20000	26	16	0
Mushroom	8124	2	22	22
Pen Based	10992	10	16	0
Poker Hand	1025010	10	10	10

実験では、データセットをアクティブラーニング用のデータとテストデータに分割して使用した。分割の割合は8:2とした。アクティブラーニング用のデータはLabeled Pool, Unlabeled Poolにさらに分割され学習器はLabeled Poolを学習する。学習後、Query Strategyに従いUnlabeled PoolからLabeled Poolに追加して再学習を行わせる。この作業は、ラベルが与えられていないデータからモデルの精度改善に最も寄与するとされるデータを選択し、Oracleに問い合わせる作業に対応している。

4.1.3 学習器の設定

学習器は各クラスに1つの勾配ブースティング決定木を作成し、ファジィ測度

$$\mathbb{G}(\mathbf{x}) = \{g(c_1|\mathbf{x}), g(c_2|\mathbf{x}), \dots, g(c_N|\mathbf{x})\} \quad (6)$$

推論する。各クラスへのファジィ測度 $g(c_i|\mathbf{x})$ は独立して計算される。学習では決定木に対応するクラスのデータが与えられたときには1となるように、対応しないクラスのデータが与えられたときには0となるように学習する。勾配ブースティング決定木の深さを2、ブースティングの回数を50とした。確率測度を用いるQuery Strategyでは $\mathbb{G}(\mathbf{x})$ にソフトマックス関数を適用し、確率測度とした。

4.1.4 比較した Query Strategy

実験では以下に示す4つのQuery Strategyを比較した。

- FuzzyMeasuredUncertaintySampling(Proposed)
- Entropy
- LeastCertainty
- RandomSampling(BaseLine)

提案手法ではファジィ測度をもとにモデルが予測する密度分布の分離と拡大の尺度を計算し、上位のデータを1:1の割合でQueryとした。

Entropyでは事後確率 $P(y|\mathbf{x})$ をもとにエントロピーが最大となるデータをQueryとした。

$$\mathbf{x}_{ENT}^* = \arg \max_{\mathbf{x}} \left\{ - \sum_i P(y_i|\mathbf{x}) \log P(y_i|\mathbf{x}) \right\} \quad (7)$$

Least Certaintyでは最大となるクラス y^* の事後確率 $P(y^*|\mathbf{x})$ が最小となるデータをQueryとした。

$$\mathbf{x}_{LC}^* = \arg \min_{\mathbf{x}} P(y^*|\mathbf{x}) \quad (8)$$

ベースラインである RandomSampling では Unlabeled Pool からランダムに Query とする。

4.1.5 評価指標

データセットの難易度に依存しない評価指標として Ranking Curve を使用する⁴⁾。

アクティブラーニングの性能はQueryの追加回数を横軸に、Accuracyを縦軸にとった学習曲線を用いて評価する。Accuracy向上の傾きが大きいほどより精度向上に寄与しているデータをQuery Strategyによって与えられることを意味している。しかし、複数のデータセットの学習曲線を平均して評価を行うと手法間の相対的な順位関係が失われる。

そこで、Queryの追加回数を横軸に、Accuracyの順位を縦軸にとったグラフを評価する。順位を比較することでAccuracyが小さい場合にもQuery Strategy間の相対関係を失わずに表現できる。順位が良いほど相対的に良いQuery Strategyと言えるが、優位性の程度はRanking Curveからは判別できない。

4.1.6 実験手順

データセットを1つ選択して、アクティブラーニングで逐次的にデータを追加、学習を行わせ、イテレーション毎にテストデータを使用してAccuracyの測定を行った。イテレーションは20回行った。このとき、始めに与えるデータの数は100、イテレーション毎に追加するデータの数はデータセットに応じて、クラス数の5倍とした。

Accuracyを測定した後に各イテレーションでQuery Strategy間でAccuracyの順位付けを行った。順位付けを全てのデータセットで行った後にイテレーション毎のAccuracyの順位の平均をとり、Accuracy Curveとした。

4.1.7 実験結果

Fig.2に実験結果を示す。Ranking Curveから学習初期のイテレーションでは、提案手法が最も順位が高く、精度向上に寄与するデータを与えられているとわかる。これは提案したQuery Strategyがデータ数の少ないときに特に有効であると考えられるが、追加すべきデータと指標が高いデータが対応しているかは追加で確認する必要がある。

学習が進行するに伴い、既存手法のLeast Certaintyと順位が変動するようになり、最終的にほとんど同じ順位になった。使用したデータセットそれぞれでの学習曲線を見たところ、使用したデータセットは難易度

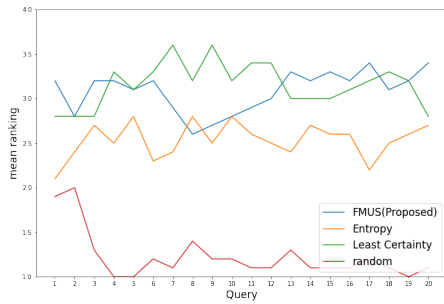


Fig. 2: Ranking Curve

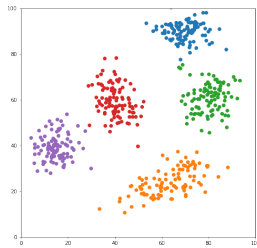


Fig. 3: Dummy Data

が低かったため Accuracy が飽和した場合や難易度が高く、学習が不安定であった場合があった。

なお、randomで精度改善の効率が相対的に悪かった理由は明瞭に識別できるデータを加えているからである。ただ1つのクラスのみ高いデータはモデルが明確な決定境界を持っていることを意味しているため、加えても密度分布の改善に大きくは寄与しないと予想される。結果として改善に有効なデータが少なくなったため、精度向上が遅くなったと考えられる。

4.2 実験2

ダミーデータをアクティブラーニングで学習したときに Query Strategy によって取り出された Query に対して学習器が与えた評価の傾向を調査した。

4.2.1 目的

実験1で提案手法が学習初期において学習効率の高い Query を与えることができたが、追加すべき2種類のデータをそれぞれ分離の指標、拡大の指標によって取り出せているか不明であった。このため、加えられたデータと使用した指標の間に整合性があるかを確認するために行った。

4.2.2 データセット

平均と共分散行列の異なる5種類の2変量正規分布から各300のデータをサンプリングした。サンプリングしたデータを2次元平面上にプロットした図を Fig.3 に示す。

学習器はある2次元ベクトルが与えられた時にどの分布からサンプリングされたデータかを識別する学習を行う。

4.2.3 比較した Query Strategy

以下に示す3つの手法を比較した。詳細は実験1と同様とした。

- FuzzyMeasuredUncertaintySampling(Proposed)

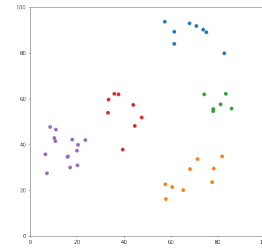


Fig. 4: Initial Labeled Data

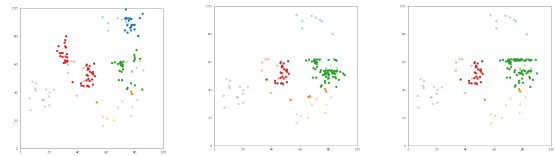


Fig. 5: Queries from each Query Strategy

- Entropy
- LeastCertainty

4.2.4 評価指標

共通の初期データを与えたときに、各 Query Strategy が取り出した Query の位置特性と学習器が与えたファジィ測度を評価する。共通の初期データは45、Queryの大きさは100とした。初期データの散布図を Fig.4 に示す。

4.2.5 実験結果

Fig.5に各 Query Strategy によって取り出された Query を示す。Entropy や Least Certainty によって取り出されたデータはどちらも同じような位置に存在するデータから取り出されているとわかる。これらのデータは提案手法によって取り出された Query にも存在する。これらのデータに対して学習器が推論したファジィ測度を Fig.6 に示す。ファジィ測度はどのクラスに対しても低いファジィ測度を与えているため、拡大の指標が高いデータを表している。Entropy や Least Certainty によって取り出された Query も同様のファジィ測度を予測しているため、Entropy や Least Certainty では拡大に寄与する未学習の領域のデータに Query が偏るとわかった。

一方で、提案手法では Entropy や Least Certainty が取り出さなかった領域のデータが存在した。これらのデータは分離の指標が高いデータであった。Fig.7にその領域に属するデータに対するファジィ測度の推論結果を示す。Fig.7より、分離の指標が高いデータは複数のクラスに対して高い値を推論している。これより、加えるべきデータと指標の対応づけができていないと確認できた。

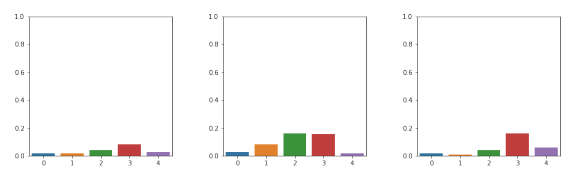


Fig. 6: Fuzzy Measure of query from exp metric

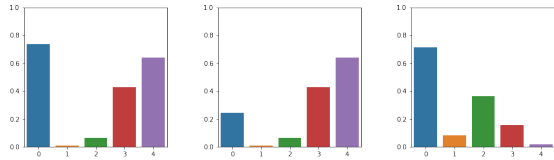


Fig. 7: Fuzzy Measure of query from sep metric

Table 2: Datasets for experiment 2

Name	#Instances	#classes	#attributes
Pen Based	10992	10	16
Mnist	70000	10	784
Cifar-10	60000	10	3072

2種類の領域に存在するデータと学習器が与えたファジィ測度との関係から、提案手法は指標をもとに2種類のデータを明示的に取り出せ、結果として Entropy や Least Certainty と比較してより精度向上に寄与しやすい Query を与えられるとわかった。また、Entropy や Least Certainty によって取り出される Query は未学習の領域に存在するデータに偏ると確認できた。

4.3 実験 3

クラス数は同一で特徴次元数の異なる3つのデータセットをアクティブラーニングで学習し、各データセットでの Accuracy の推移を調べる実験を行った。

4.3.1 目的

データセットの特徴次元数や難易度に応じてアクティブラーニングの学習推移は変化する。入力次元数の変化に応じた学習推移の変化、提案手法の優位性と入力次元数の関係を明らかにするために本実験を行った。

4.3.2 データセット

実験ではクラス数が同一で入力次元数の異なる3つのデータセットを使用した。データセットの詳細を Table.2 に示す

アクティブラーニングでは、始めに与えるデータを Pen Based, Mnist では 100, Cifar-10 では 1000 とした。またイテレーション毎に追加するデータは Pen Based, Mnist では 100, Cifar-10 では 1000 とした。

4.3.3 評価指標

横軸にアクティブラーニングのイテレーション数を、縦軸に Accuracy の推移をプロットした学習曲線を3つのデータセットに関して作成し、評価する。イテレーション間の傾きが大きいほど効率よく精度が改善できており、より精度向上に寄与する Query を与えられていることを表している。

また、学習初期における精度向上の速度を確認するために1回目の Query 追加時の Accuracy を評価する。

4.3.4 実験結果

各データセットにおける Query Strategy の Accuracy の推移を Fig.8-10 に示す。

また、1回目の Query を追加して学習した後の Accuracy を Table.3 に示す。太字が最も Accuracy の高かった手法である。

Table.3 から入力次元数が少ない Pen Based では Least Certainty が学習初期において最も精度向上に寄与するデータを与えられていると確認できた。一方

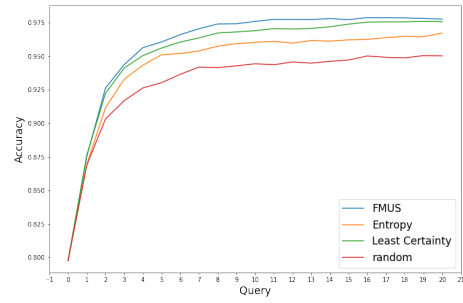


Fig. 8: Learning Curves (Pendigits)

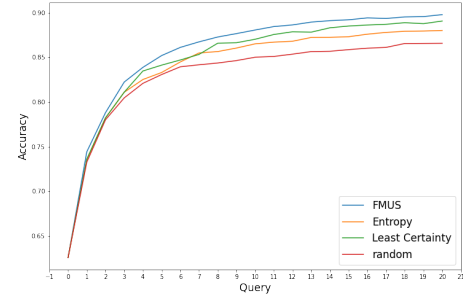


Fig. 9: Learning Curves (Mnist)

で Mnist や Cifar-10 のような入力次元数の多いデータセットでは提案手法が最も Accuracy が高い。さらに各データセットにおいて、提案手法と Least Certainty の Accuracy の比を取ると Pen Based では 0.9989, Mnist では 1.011, Cifar-10 では 1.015 となっており入力次元数が増えるほど提案手法の有効性が増す関係が得られた。

これは入力次元数が多いほどそれぞれのデータ間の距離が大きくなるが原因と予想される。データ間の距離が大きい場合、未学習の領域に存在するデータが取得しやすくなり、反対に複数のクラスで混同するようなデータは取得しづらくなる。ゆえに密度分布の分離に寄与するデータを明示的に取得できる提案手法は入力次元数の増大に伴い優位性が増大したと考えられる。

また、入力次元数が少ない場合には拡大に寄与するデータを入力次元数が多い場合には分離に寄与するデータを重視する Query Strategy がより有効であると予想できる。

5 おわりに

画像の識別をアクティブラーニングにより学習しようとした場合、物体の判別の難しさから追加するデータは予測する密度分布の分離に寄与するデータと密度分布の拡大に寄与するデータを加える必要があり、ファジィ測度を使用してモデルの精度向上に寄与する2種類のデータを明示的に取り出す Query Strategy を提案した。

Table 3: Accuracy after 1st Queries

Dataset	FMUS	Entropy	LC
Pen Based	0.8753	0.8692	0.8762
Mnist	0.7441	0.7351	0.7358
Cifar-10	0.3544	0.3474	0.3490

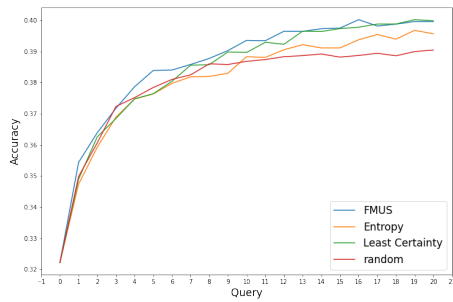


Fig. 10: Learning Curves (Cifar-10)

実験1および実験2から、提案手法は決定境界近傍のデータと未学習の領域に存在するデータを明示的に取り出せて、結果として既存の確率測度を用いた Query Strategyと比較してより精度の向上に寄与しやすいデータを与えられるとわかった。反対に、確率測度を用いた Query Strategyは密度分布の拡大に寄与するデータを優先的に Query とし、決定境界近傍に存在する密度分布の分離に寄与するデータを提示しないとわかった。

実験3では、入力次元数と提案手法の優位性に関する実験を行い、入力次元数が多いほど優位性が増す相関があると確認できた。入力次元数が多くなり、データ間の距離が大きくなると決定境界近傍のデータが相対的に少なくなり、明示的に取得できる提案手法以外では取得できなかったからであると予想される。

また、この実験3から入力次元数が少ない場合には密度分布の拡大に寄与するデータを、入力次元数が多い場合には密度分布の分離に寄与するデータを優先的に Query とする Query Strategy がより精度向上に寄与しやすいデータを与えられると考えている。

参考文献

- 1) V. Jothi Prakash, Dr. L.M. Nithya: A Survey on Semi-Supervised Learning Techniques, International Journal of Computer Trends and Technology (IJCTT) 8(1):25-29, February (2014)
- 2) Burr Settles: Active Learning Literature Survey, Computer Sciences Technical Report 1648 (2009)
- 3) Dua, D. and Graff, C.: UCI Machine Learning Repository <http://archive.ics.uci.edu/ml>. Irvine, CA: University of California, School of Information and Computer Science. (2019)
- 4) Davi Pereira-Santos, Ricardo Bastos Cavalcante Prudêncio, André C.P.L.F. de Carvalho: Empirical investigation of active learning strategies, Neurocomputing, 326-327号 15/27 (2019)

土壌断面画像群に対する画素輝度値の補正法

○前川新司¹ 磯川悌次郎¹ 木村敏文² 池野英利³ 大橋瑞江² 上浦尚武¹

(¹兵庫県立大学大学院工学研究科 ²兵庫県立大学環境人間学部 ³福知山公立大学情報学部)

A Calibration Scheme for Brightness of Pixels in Cross-Sectional Soil Images

*S. Maegawa¹, T. Isokawa¹, T. Kimura², H. Ikeno³, M. Ohashi², and N. Kamiura¹

(¹Graduate School of Engineering, University of Hyogo ²School of Human Science and Environment, University of Hyogo ³The University of Fukuchiyama)

Abstract— This paper presents a scheme for calibrating the pixel intensities for soil images taken by flatbed scanner buried in the ground. First, our proposed scheme first detects the region of gray balance cards in the image by using template matching method, and then obtains average value of gray scale level for these cards. Then, calibration of pixel intensities are carried out by using a intensity curve that is made by pixel intensities of these cards. Experimental results using a set of actual soil images in several observation sites showed that the proposed scheme was effective for compensation of brightness distribution.

Key Words: Soil image, template matching, brightness calibration

1 はじめに

植物の根系は様々な観点から重要な役割を持っている。植物単体という面から見れば、根系は土壌から水分や栄養素を吸収する重要な生理学的機能を担っており、樹木根の細根は高い養分吸収活性を備えている。また生態システムとしての観点からは、樹木根の細根は発生・成長から枯死・分解にいたるまでのターンオーバーが速く、生態系内における物質循環に大きな影響を与えていると考えられている¹⁾。そして、森林生態系内における炭素循環の年間の一次生産部分は40~70%は細根が占めており²⁾、森林生態系やその生態系の物質循環の変化を明らかにするためには根系の特に細根のフェノロジーについて調べることが極めて重要である³⁾。しかしながら、植物の葉や枝などの地上部のフェノロジーが明らかになっているのに対して、根系のフェノロジーは地中に埋まっていることもあり、未だほとんど理解されていない。

現在、地中の根系のフェノロジーを知るために、地中の根を掘り出さずに観測する方法の一つとしてフラットベッドスキャナを用いたスキャナ法⁴⁾が用いられている。このスキャナ法では従来法よりも広い範囲の土壌断面を継続的に撮影が可能である。しかし、この方法で得られた画像群はスキャナ容器の壁面や土壌の状態によってその輝度分布などが多様に変化する。そのため、この画像群から細根の検出を行いその変化を捉えるには、撮影された画像群の輝度分布を均一化することが有効であると考えられる。スキャナ法では画像個別の輝度の状態を知る手がかりとして、白色、灰色、黒色の3枚のグレーバランスカードを重畳して撮影が行われている。しかしながら、このような方法で撮影され始めたのがここ数年のこともあり、このカードの輝度情報を用いて輝度分布を補正する方法については、未だほとんど検討されていない。

本研究では、スキャナ法で撮影された画像から自動的にバランスカード領域を検出し、そこから得られる輝度情報を用いて画像の輝度分布を補正する手法を提案する。また、いくつかのサイトにおいて実際にスキャ

ナ法で撮影された画像に提案手法を適用し、その有効性を評価する。

バランスカード領域の検出方法としては、テンプレートマッチング法でカードの内一つを検出し、その周囲の輝度分布情報により残りのカード領域の検出を行う。カード領域検出方法の他の手法として、画像の二値化を行うことによりカード領域を抽出する方法が一般的である。しかしながら、本研究で用いる土壌画像ではカード以外の画素値の分布が土の粒などの陰影などで大きく変化しているため、前処理を行ったとしてもカード領域をそのまま検出することは簡単ではない。そこで本研究では、二値化処理を行わずにカード領域の検出を行う方法としてテンプレートマッチング法を採用した。

検出が成功した後に、得られたカード領域から輝度値を取得し輝度補正に用いた。輝度補正法として、画像中で得られたカードの輝度値と標準光下でのカードの輝度値から多項式回帰によって補正関数となる曲線を求め、画像の輝度値を補正することにより行う。輝度補正の別の手法として、輝度分布のヒストグラム平坦化など輝度値全体に対して一律な処理を施すことが考えられる。しかしこのような方法を用いた場合、ある画像内で明暗が場所によって異なっているときには、輝度値の平坦化した後では暗い箇所と明るい箇所は同時に明るくなることになり輝度の違いがそのまま反映され、過度な補正がかかることになる。また画像ごとに元の輝度分布は異なっており、その差異がそのまま反映されるので画像群の輝度分布均一化は期待できない。よって、輝度補正の処理としては過度な輝度補正を防ぎ輝度分布を均一化させるため、輝度値の目標値を設定しそれに合わせて補正を行う。

2 原理

2.1 スキャナ法

土壌画像を取得する手法の従来法としてミニライゾトロン法などが存在する。ミニライゾトロン法とは、透明な管容器を用意しその中に特殊な撮影機器を入れ、管



Fig. 1: An example of soil image taken by a flatbed scanner. Gray-scaled balance cards are embedded on the scanner and they are also imaged.

内表面 360 度を土壤画像として取得する手法である¹⁾。この手法は観察管が設置できる場所ならどこでも実験実施が可能で、繰り返し多数の土壤画像を撮影できることから多くの研究者に利用されてきた。しかし短所として、観察面が小さく、観察対象とする根系の本数やサイズが限られることや、多くの場合、観測装置の設置や記録は手作業で行われ、データ取得に手間がかかることなどが挙げられる。

本研究の対象となる土壤断面画像を取得するスキャナ法とは、地面に直方体の透明容器を埋設し、その中にフラットベッドスキャナを設置することにより透明容器の表面に分布する根の生育状態を観察する手法である⁴⁾。撮像された画像の例を Fig. 1 に示す。この画像の右上部分には、土壤の状態により変化する画像輝度を補正することを目的としてグレイバランスカードと呼ばれるカードを埋設している。このバランスカード部分を拡大した図を Fig. 2 に示す。埋設しているバランスカードは 3 種類あり、それぞれ黒色、灰色、白色のカードである。

スキャナ法の利点としては次の点が挙げられる。まず、スキャナ法で得られる画像のサイズは約 20cm × 30cm と従来法の約 30 倍も大きくなる。そのため撮影できる根の本数が多く根の広がりや分岐の程度なども把握できる。また、特殊な観測装置を必要とする従来法と比べて、市販のフラットベッドスキャナを用いることができるためコストを低く抑えることができる。さらに、外部電源などに接続することで連続的な撮影が可能であり、また撮影頻度を制御することで根の微細な成長、枯死状況を観測することができる。このスキャナ法により細根に対してどれほどの長期観測に適用できるかの実験も行われており、詳細かつ長期的なデータが高頻度に取得可能なことが示されている⁵⁾。

2.2 テンプレートマッチング

テンプレートマッチング法は画像中に存在するテンプレート画像 (部分画像) の位置を発見する方法である⁶⁾。この手法はテンプレート画像を入力画像全体にスライドさせ、テンプレート画像と画像の注目領域とを比較する。出力は各画素がテンプレート画像と注目画素の近傍領域の類似度を表すグレースケール画像になる。入力画像のサイズが $(W \times H)$ であり、テンプレート画像のサイズが $(w \times h)$ である場合、出力画像のサイズは

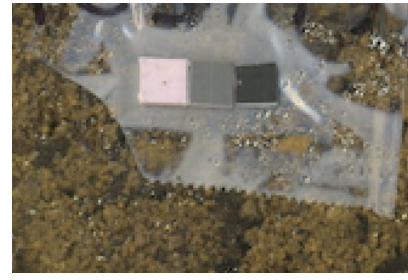


Fig. 2: An example of gray-scaled balance cards. There are three types of cards (black, gray, and white cards).

$((W - w + 1) \times (H - h + 1))$ となる。テンプレートマッチングを行うことにより、類似度が最大となる画素の位置を得ることができる。テンプレート画像に最も似ている領域を表す矩形の左上の画素は類似度が最大となる画素の位置であり、領域のサイズは (w, h) となる。テンプレートの比較法として、本研究では SSD (Sum of Squared Difference) を用いる。

SSD は「画素値の二乗誤差」により画像間の類似度を評価する⁶⁾。この指標を用いると、値が最小になる場所が類似度の最も高いということになる。入力画像の画素値を $I(x, y)$ 、テンプレート画像の画素値を $T(x, y)$ とし、テンプレート画像の幅を w 、高さを h とする。走査位置を (dx, dy) とすると SSD の値は式 (1) として定義される。

$$SSD(dx, dy) = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} (I(dx + x, dy + y) - T(x, y))^2 \quad (1)$$

3 提案手法

本論文にて提案する輝度補正方法では、土壤と一緒に撮影されたグレイバランスカードの輝度情報を用いて画像全体の輝度補正を行う。本提案手法では、最初に土壤画像中にあるグレイバランスカード (Fig. 2) の各カードの位置を特定する。その後、各カードの画像領域からカードの輝度値を算出し、それらの値と標準照明環境下で撮影したグレイバランスカードの輝度値の関係から輝度補正曲線を作成する。最後に、この輝度補正曲線に従い土壤画像全体の輝度補正を行う。

3.1 グレイバランスカード領域の検出

まず、土壤画像中におけるグレイバランスカードが画像中のどこに存在するかを探索する。探索方法として、テンプレートマッチング法とその周囲領域の探索を組み合わせた方法を提案する。

カード検出の具体的な処理は、最初に白カードのテンプレート画像を用いてテンプレートマッチングにより白カード領域の検出を行う。その後に出検した白カードの位置を手掛かりとして平均輝度分布の情報から残りの二つのカードの領域検出を行う。灰色カードと黒色カードの検出においては、予め設定したカード領域内での輝度平均値と分散値を算出し、それらが一定範囲内の値であれば、灰色カードもしくは黒色カードの候補点として記録する。その候補点と白色カードの左上座標を基準にユークリッド距離を算出し、距離が一定範囲内かつ最も輝度分散値が小さい領域の候補点

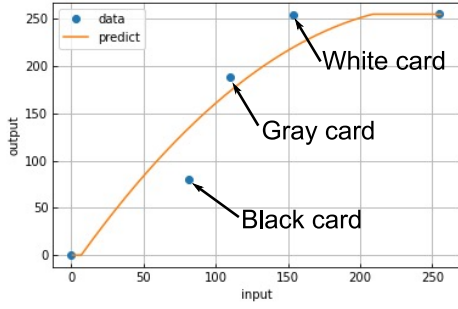


Fig. 3: An example of brightness calibration curve

を対象のカードの左上座標として決定する．以上の手順により，全てのグレーバランスカードの領域を検出する．

3.2 画像の輝度補正法

先述のグレーバランスカード領域が検出成功した画像に対して，白・灰・黒カード領域内の輝度値から輝度補正を行う．具体的には，対象画像の色空間を HSV(H:色相, S:色彩, V:輝度) 空間に変換し，そこから輝度成分 (V) の情報を取り出す．この輝度成分に対して階調変換を行う．この階調変換に必要な補正関数となる多項式曲線は抽出したカード群の輝度値と，標準光下で撮像されたカード群輝度値から画像ごとに算出する．標準光下でのカード輝度値は，カード試料各色3枚を標準環境下の室内で撮影し，輝度の平均値を用いる．本研究で用いる標準のカード輝度値は複数回値を取った結果を平均し白:254, 灰:188, 黒:80とした．

一例として，ある土壌画像について抽出したカード輝度値を横軸，標準光下でのカード輝度値を縦軸としてプロットした関係を Fig. 3 に示す．ここでは輝度値が 0 と 255 の点もプロットしている．この 5 点から多項式回帰により 2 次の回帰曲線を生成する．ただし，回帰曲線は 255 より大きい値の場合は 255 に，0 より小さい値の場合は 0 に輝度値を固定する．

また，輝度補正後の画像群に対して輝度分布の均一化がなされているかを評価する．評価方法として，各サイトの画像群から 1 枚の画像を選び，同サイトにある他の画像群の輝度分布の間で類似度を算出する．類似度はヒストグラムの相関関数として式 (2) および式 (3) により計算される．

$$d(H_i, H_j) = \frac{\sum (H_i(I) - \bar{H}_i)(H_j(I) - \bar{H}_j)}{\sqrt{\sum (H_i(I) - \bar{H}_i)^2 \sum (H_j(I) - \bar{H}_j)^2}} \quad (2)$$

$$\bar{H}_k = \frac{1}{N} \sum_l H_k(l) \quad (3)$$

ここで， H_k はヒストグラムの値， N はヒストグラムのビン数である．また $i \cdot j$ は画像が何枚目かのものを示す．算出される数値は $-1 \sim 1$ の範囲の値を取り，1 に近いほど画像ヒストグラム間の類似度が高いことを示す．

4 実験結果

本節では，提案手法を実際の土壌画像に対して適用し，その有効性を評価する．撮影された土壌画像は撮影

Table 1: Detection accuracy for finding gray balance cards

Observation site	#. images	#. successful
Ryukoku (Bright)	9	9
Ryukoku (Dark)	20	17
Shiiba	10	10
Terasawa	10	2

箇所と撮影状態から 4 サイトに分割し，龍谷 (明)，龍谷 (暗)，椎葉，手良沢とした．画像データ数はそれぞれ，9, 20, 10, 10 である．

4.1 グレーバランスカード領域の検出結果

まず，テンプレートマッチング法と輝度値の平均・分散値の情報を用いてグレーバランスカードの検出精度を評価する．全画像データについて，カード領域を取得した場合の検出成功率を Table 1 に示す．

この結果から龍谷 (明) と椎葉の画像群では，全ての画像に対してカード領域の検出が成功していることがわかる．しかし，龍谷 (暗) と手良沢についてはいくつかの画像においてカード検出が失敗している．この時，白色カードについては検出はできてはいたが，灰色もしくは黒色カードの片方か両方の検出ができていなかった．これらの画像については，バランスカード周辺部分がカードを封入している部分にある水蒸気により曇っているという共通点が見られた．また，誤検出した箇所は灰色カードでは水蒸気などで曇っている部分を，黒色カードでは離れた土壌部を検出していた．

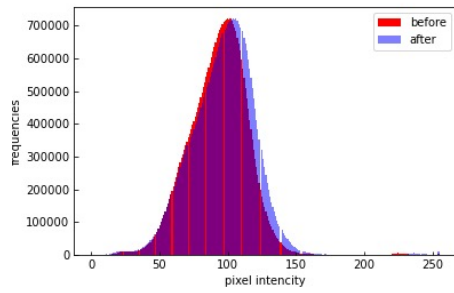
4.2 画像の輝度補正結果

次に，カード領域の検出に成功した画像群に対して，提案手法により輝度補正を行った結果を示す．傾向として場所ごとに補正関数が類似する傾向があるため，各サイトについて代表して一つの土壌画像に対しての補正結果として輝度ヒストグラムの変化と補正関数を示す．補正関数の結果では，縦軸を補正後輝度値で横軸を補正前輝度値として，中央の 3 点について左からそれぞれ黒・灰・白の輝度値を示している．

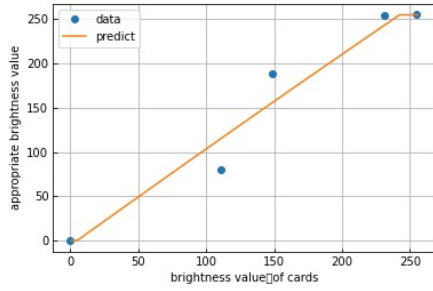
最初に，龍谷 (明) のサイトにおける補正前後の輝度値ヒストグラムと補正に用いた補正曲線をそれぞれ Fig. 4(a) と Fig. 4(b) に示す．Fig. 4(a) より，龍谷 (明) ではヒストグラムからほとんど変化がないことから輝度値補正はほとんど行われていないということがわかる．また，Fig. 4(b) に示す補正関数についてみると，この補正曲線は傾きがおよそ 1 である線形関数が得られている．このような補正関数が得られた理由として，補正前の白輝度値の数値が高いことが挙げられる．

同様に，龍谷 (暗) のサイトについて，輝度値ヒストグラムと補正曲線の図を，それぞれ Fig. 5(a) と Fig. 5(b) に示す．龍谷 (暗) では画像の輝度値は全体的に大きくなる方向にヒストグラムは推移していることがわかる．Fig. 5(b) の補正関数からは，龍谷 (明) の時と比べ白輝度値の値が低下していることがわかる．

椎葉のサイトにおける輝度値ヒストグラムと補正曲線を，それぞれ Fig. 6(a) と Fig. 6(b) に示す．これらの図より，椎葉では画像の輝度値は全体的に大きくなる方向にヒストグラムは推移しているが低輝度に値が集中している様子が見られる．Fig. 6(b) では龍谷 (暗) と同様の関数が得られていることがわかる．

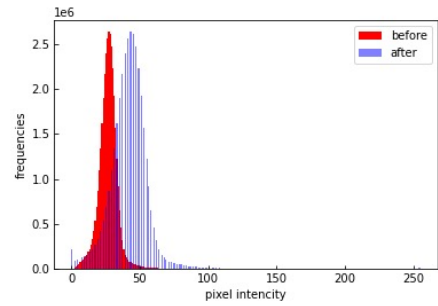


(a) Histogram before/after calibration

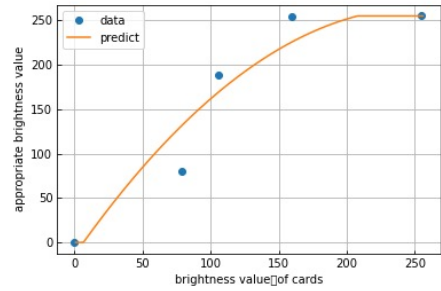


(b) Calibration curve

Fig. 4: Histograms for pixel intensities and calibration curve in Ryukoku (Bright) images

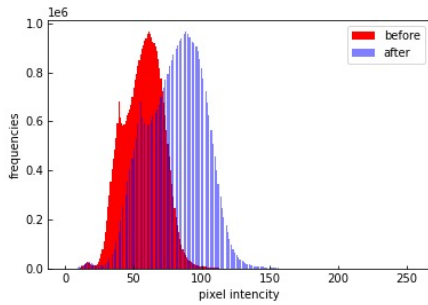


(a) Histogram before/after calibration

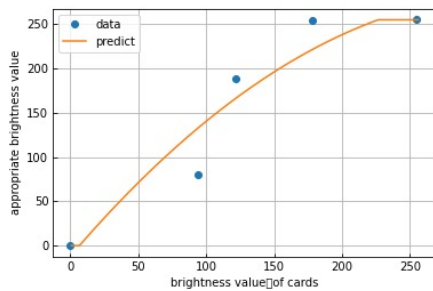


(b) Calibration curve

Fig. 6: Histograms for pixel intensities and calibration curve in Shiiba images

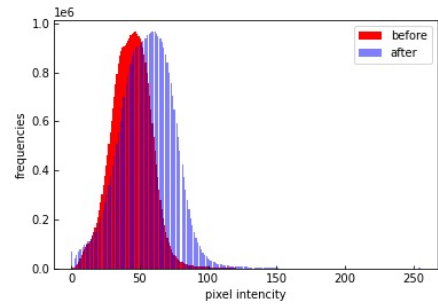


(a) Histogram before/after calibration

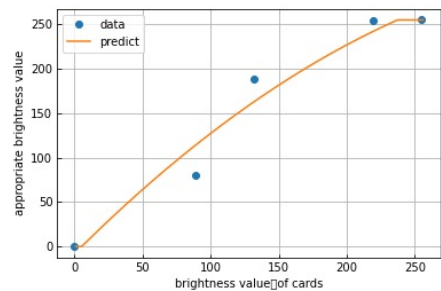


(b) Calibration curve

Fig. 5: Histograms for pixel intensities and calibration curve in Ryukoku (Dark) images



(a) Histogram before/after calibration



(b) Calibration curve

Fig. 7: Histograms for pixel intensities and calibration curve in Terasawa images

最後に、手良沢のサイトの画像に対する輝度ヒストグラムと補正曲線をそれぞれ Fig. 7(a) および Fig. 7(b) に示す。ヒストグラムについては、他のサイトと同様に輝度分布の変化が見られる。補正関数としては、龍谷(暗)や椎葉より傾きが小さい関数が得られていることがわかる。

最後に、輝度補正後での土壌画像の輝度値の均一化

がなされているかを検証するため、サイトごとに土壌画像の1枚を基準とし、同サイトにある他すべての土壌画像との輝度値のヒストグラムの類似度を算出した。先に、比較のために補正前と補正後の土壌画像で各サイトにおけるヒストグラム類似度の平均値と標準偏差を算出した結果を Table 2 に示す。表の結果より、龍谷(明)は類似度は補正前後で高い類似性を示している

Table 2: Similarity evaluation of the histogram before/after calibration

Observation Site	Averaged similarity	
	Before calibration	After calibration
Ryukoku (Bright)	0.996 ± 0.000	0.900 ± 0.014
Ryukoku (Dark)	0.916 ± 0.008	0.561 ± 0.066
Shiiba	0.058 ± 0.003	0.244 ± 0.298
Terasawa	-0.017 ± 0.000	0.580 ± 0.000

が数値は減少している。この原因は、龍谷(明)の画像群は同じ箇所を異なる時間で撮影した画像群であることや補正関数が一次関数に近いこと、類似度の平均値が高く補正前後での変化も小さいと考えられる。また、龍谷(暗)では補正前後で類似度が減少している。原因としては、この画像群については輝度分布が一様でなかったため補正効果を均一に受けることができなかったことが考えられる。

次に、椎葉での補正後の類似度は上昇しているため、輝度の補正効果が確認できる。しかしながら、補正後の標準偏差の数値からわかるように、類似度が高い画像と低い画像が混在した結果となった。類似度が低い画像が得られた原因として、椎葉では補正前の輝度分布の低い傾向があり、本補正手法では Fig. 7(b) の補正関数に示すように、低輝度値の場合に補正効果が作用しにくくなる。このため、特に輝度分布が低い値に集中する画像とそうでない画像との間の補正効果に違いが生じたためであると考えられる。手良沢については、対象となる画像は2枚であるが類似度は補正前後で高くなっている。このサイトについては、バランスカード領域検出が上手くいっておらず類似度のための比較画像枚数が非常に少ないため今後変化する可能性がある。

5 まとめ

本研究では、複数の環境で撮影された土壌画像に対して、一緒に埋設したグレーバランスカード領域の検出とこのカードの輝度情報を用いた輝度補正手法の提案し、実土壌画像を用いて評価を行った。カード領域の検出の結果から、提案手法により多くの画像に対応できることが示された。一方で、手良沢でのカード領域は2枚のみ検出できた。検出できなかった画像に関してはカード周囲が水蒸気により曇っているという共通点があることを確認している。これにより、輝度分布情報がカード領域のそれと近い状態の領域が別々に発生していたことが原因であると考えられる。

また、補正関数を用いた輝度補正に関しては輝度分布全体を輝度が大きくなる方向に移動させることがで

きたことを確認した。しかし、画像全体として輝度が極端に低い値に集中している椎葉などについては十分な補正を行うことができなかった。最後に補正後の画像群について、輝度値の分布について類似度を算出した。これについても椎葉を除くサイトの画像群について類似度が高くなることを確認した。また補正前の類似度からの変化については龍谷(明)と龍谷(暗)は減少はしたが、椎葉や手良沢などについては増加していることが確認できた。

今後の課題として、カード領域の検出精度や輝度補正での輝度の均一化性能向上を目指していくことが挙げられる。それと同時に、グレーバランスカードの埋設方法を改善することや、より補正を行いやすくするためのバランスカードの選定などを行っていく必要があると考えられる。

参考文献

- 1) 野口享太郎, 阪田匡司, 溝口岳男, 高橋正通: ミニリゾトロン法とコアサンプリング法による細根量の解析, 日本森林学会発表データベース, vol.114, p.295, 2003
- 2) T. Kume, M. Ohashi, N. Makita, L. K. Kho, A. Katayama, I. Endo, K. Matsumoto, and H. Ikeno: Image analysis procedure for the optical scanning of fine-root dynamics: errors depending on the observer and root-viewing window size, *Tree Physiology*, 1/12, 2018
- 3) L. Radville, M. L. McCormack, E. Post, and D. M. Eissenstat: Root phenology in a changing climate, *Journal of Experimental Botany*, 67(12), 3617/3628, 2016
- 4) M. Dannoura, Y. Kominami, H. Oguma, Y. Kanazawa: The development of an optical scanner method for observation of plant root dynamics, *Plant Root*, 2, 14/18, 2008
- 5) 仲畑了, 中村瞳, 檀浦正子, 大澤晃: フラットベッドスキャナー法による細根動態の経年的評価, 日本森林学会大会学術講演集, 126, 92, 2015
- 6) R. Brunelli: Template matching techniques in computer vision: theory and practice, John Wiley & Sons, 9/40, 2009

深層学習による死後肺CT画像を用いた説明可能な溺死鑑別システムに関する研究

○坂本奨太[†] 張暁勇^{††,†} 本間経康[†] 川住祐介[†] 白井章人[†] 小河原輝正[†]
舟山真人[†] 市地慶[†] 杉田典大[†] 吉澤誠[†]
([†]東北大学, ^{††}仙台高等専門学校)

Explainable Deep Learning-Based System for Discrimination of Drowning in Post-Mortem Lung CT Images

*Shota Sakamoto[†], Xiaoyong Zhang^{††,†}, Noriyasu Homma[†], Yusuke Kawasumi[†], Akihito Usui[†],
Terumasa Ogawara[†], Masato Funayama[†], Kei Ichiji[†],
Norihiro Sugita[†], and Makoto Yoshizawa[†]
([†]Tohoku University, ^{††}National Institute of Technology, Sendai College)

Abstract— Diagnosis of drowning is a difficult task in forensic medicine. Our previous study proposed deep learning (DL) to analysis the post-mortem lung CT for drowning diagnosis. The experimental results have demonstrated its effectiveness. However, due to the black-box nature of the DL, the output of the proposed method is not explainable and is restricted its clinical use. In this study, we applied Grad-CAM and Deep SHAP, two visualization methods of DL, to analyze the contribution of an input image feature region to the diagnosis of drowning. As a preliminary result, the Grad-CAM and Deep SHAP shown their capabilities of extracting several typical features of drowning, and providing an explainable diagnosis decision.

Key Words: Deep Learning, Explainable AI, Postmortem CT

1 はじめに

溺死の一般的な定義は日本救急医学会¹⁾によれば、「浸水あるいは浸漬により窒息をきたした状態を溺水といい、溺水により死亡したものを溺死という」と与えられている。

世界保健機構 (WHO) の 2014 年の報告書によると、世界中で年間約 37 万人が溺死であると診断される²⁾。日本においても、2019 年の厚生労働省の統計³⁾より、不慮の溺死及び溺水は 7690 人と報告され、不慮の事故死の中でも高い割合を占める。

死因究明の場で、溺死診断は除外診断と呼ばれる事が多い。その診断は死亡時の周囲の状況、一般に認められている非特異的な解剖所見などに基づいて行われる。その結果、溺死は法医学的に診断が困難とされている。

溺死診断に対し、死後画像診断 (Autopsy imaging : Ai) が有効とされ、溺死所見を画像特徴から求める研究が以前より行われている⁴⁾⁻¹⁰⁾。しかし、溺死所見の

科学的根拠や評価指標の不足¹¹⁾に加え、専門医の不足などから、医師の負担が大きいという問題があった。

上記に対し、先行研究¹²⁾では、深層学習の利用による医師の負担軽減を提案した。実験では、深層畳み込みニューラルネットワーク (Deep Convolutional Neural Network : DCNN) による、死後肺 CT 画像を用いた溺死鑑別の検証が行われた (Fig.1)。提案された手法は、鑑別性能を評価する Area Under the Curve (AUC) の指標において、 $AUC = 0.82$ を示し、剖検上十分な性能が示唆された。

しかし、DCNN モデルの出力は溺死の予測値のみであり、鑑別結果の信頼性が大きな課題であった。実際の臨床現場において、画像診断の専門出ない法医学者自身が読影を行っている場合も少なからずある¹³⁾。従って、そのような現場で信頼性を担保するためには画像診断の非専門医である法医学者にも理解可能な根拠の提示が必要であると考えた。

そこで本研究では、DCNN の説明手法として提案されている、Grad-CAM¹⁴⁾による注目領域の可視化に加え、Deep SHAP¹⁵⁾による注目特徴の可視化を応用し、非専門医でも解釈可能な診断根拠の解析を試みた。

2 DCNN 説明手法について

ここでは、本研究で用いた DCNN モデルの一つである VGG16¹⁶⁾を基にした DCNN の構造の説明に加え、DCNN 説明手法である、Grad-CAM¹⁴⁾による注目領域の可視化と、Deep SHAP¹⁵⁾による注目特徴の可視化について説明していく。

2.1 VGG16

VGG16 は Oxford 大学の Visual Geometry Group が開発した VGG モデルの一種である。その構造を Fig.2 に示す。

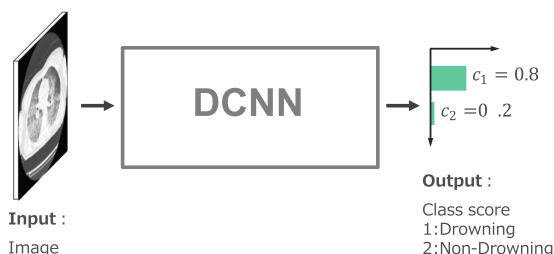


Fig. 1: DCNN-based method to classify drowning.¹²⁾ c_1 is DCNN's output for drowning class and c_2 is DCNN's output for non-drowning class. Each output score means probability that the input image belongs to the class.

VGG16は13層の畳み込み (convolution) 層と5層のプーリング (pooling) 層を持つことを特徴する。畳み込み層では畳み込み演算が行われ、入力マップから特定の特徴を抽出する働きを担う。また、学習により畳み込み層の重みが修正されていく事で、識別に有用な特徴を学習する。プーリング層では対象とする特徴の位置変動に対するロバスト化をする働きを担う。今回は最大プーリング (Max pooling) が用いており、入力マップの局所領域中の最大値を利用したダウンサンプリングの処理を行う。

畳み込み層とプーリング層の繰り返しにより抽出された特徴をもとに、全結合層 (fully-connected layer) でクラス判定が行われ、各クラスに対する出力 (Fig.2 y_1, y_2) が得られる。その後ソフトマックス (softmax) 関数の処理を得て、各クラスに属する確率が出力される。

2.2 Grad-CAM

Grad-CAM¹⁴⁾ では、DCNN 中の勾配が大きな箇所がモデルの予測値への寄与をするという考えに基づき、注目領域を示すヒートマップを算出する。ここでの勾配は、DCNN 中の最後の畳み込み層の出力 (特徴マップ, Feature map) に対する予測値の変化を意味している。畳み込み層は入力画像から特徴抽出を行う役割を持ち、DCNN 中に多数存在する。また、特徴マップの例を Fig.2 中に示す。

計算には、DCNN の出力値 y_1 と、特徴マップの2つを用いる。次式で出力 y_1 に対する、特徴マップの各次元 k の貢献度を計算する。

$$\alpha_k = \frac{1}{Z} \sum_{i=1}^I \sum_{j=1}^J \frac{\partial y_1}{\partial m_{ij}^k} \quad (1)$$

上式で、 α_k は DCNN の出力 y_1 に対する特徴マップ中の次元 k の重みを示す。 α_k が大きいほど予測に貢献している事を表す。 (I, J, K) は特徴マップの (高さ, 幅, 次元数) であり、 $Z = I \times J$ を示す。従って、 y_1 に対する特徴マップの各要素 m_{ij}^k の勾配を求め、次元ごとに平均し、重み α_k としている。

次に、各次元ごとに、重み α と特徴マップ m の積和演算を行う。

$$\mathbf{H} = \text{ReLU}\left(\sum_{k=1}^K \alpha_k m^k\right) \quad (2)$$

上式において、積和演算の結果を \mathbf{H} とする。また、ReLU 関数を通す事で、予測値 y_1 に貢献した部分のみを残している。

最後に、 \mathbf{H} を線形補間を用いて元の入力画像サイズに拡大する事で、DCNN の予測根拠を示すヒートマップが生成される。

2.3 Deep SHAP

Deep SHAP¹⁵⁾ は、各入力要素に対する SHapley Additive exPlanation (SHAP) 値¹⁵⁾ を DeepLIFT¹⁷⁾ のアルゴリズムに基づいて算出する手法である。この手法は、学習後のモデルに対し、可読表現を用いた説明が可能な手法として知られている。

以下では、計算より得られる SHAP 値の原理について説明する。

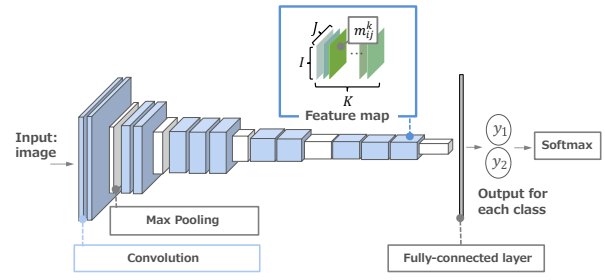


Fig. 2: Architecture of VGG16¹⁶⁾ Feature map is output of last convolutional layer in VGG16. I : height, J : width, K : channel, m_{ij}^k : element of feature map at $(i, j, k) \in (I, J, K)$.

SHAP 値は共同ゲーム理論における Shapley 値¹⁸⁾ と等価な指標であり、各入力特徴が予測に与える貢献度を統一の尺度で表した指標とされる。ここで、説明したいモデルを $f(\mathbf{x})$ 、入力要素を $\mathbf{x} = (x_1, x_2, \dots, x_n)$ とおき、それを説明するモデルを $g(\mathbf{x}')$ として次式で表す。

$$g(\mathbf{x}') = \omega_0 + \sum_{i=1}^n \omega_i x'_i \quad (3)$$

上式で、 \mathbf{x}' は \mathbf{x} 中の非ゼロ要素を 1 に置き換えたものを示す。 ω_i は、 $g(\mathbf{x}')$ の係数であり、次の3つの条件を満たす場合に、予測値 $f(\mathbf{x})$ に対する x_i の貢献度を示す指標としてただ一つ決定する。

1. $f(\mathbf{x}) = g(\mathbf{x}')$
2. $x'_i = 0 \Rightarrow \omega_i = 0$
3. $f'_x(x') - f'_x(x' \setminus i) \geq f_x(x') - f_x(x' \setminus i)$
 $\Rightarrow \omega_i(f', \mathbf{x}) \geq \omega_i(f, \mathbf{x})$

上記において、 \Rightarrow は十分条件を示し、 $B \Rightarrow A$ の場合、 B が成立すれば A も成立する事を表す。また、条件3において、 $f'_x(x') \equiv f'(h(x'))$ を示す。 $h(x')$ は、 $h(x') = \mathbf{x}$ を行う関数を示す。 $f_x(x' \setminus i)$ は $x'_i = 0$ とおいた場合を示す。従って、 $f'_x(x') - f'_x(x' \setminus i)$ は、 x'_i と対応する x_i の、予測値 $f(\mathbf{x})$ に対する限界貢献度を意味する。そのため条件3では、特徴 x_i の限界貢献度が変化した際に、 ω_i も同様の変化、もしくは無変化である事を条件としている。

以上の条件を満たす時の ω_i は次式で定義されている。

$$\omega_i(f, \mathbf{x}) = \sum_{z' \subseteq \mathbf{x}'} \frac{|z'|!(s - |z'| - 1)!}{s!} [f_x(z') - f_x(z' \setminus i)] \quad (4)$$

上式において、 s は \mathbf{x} の要素数、 z' は \mathbf{x}' の非ゼロ要素の部分集合であり、 $|z'|$ は非ゼロの数を示す。この $\omega_i(f, \mathbf{x})$ が SHAP 値となる。

また、上記の SHAP 値は定義通りだと計算コストが膨大になる問題がある。そこで説明対象モデル f の線形性を過程する事で、次式の近似を行っている。

$$f_x(z') \equiv f(h(z')) \doteq E[f(z|z_d)] \quad (5)$$

ここで、 d は z' 内の非ゼロの要素インデックスを示す。従って、 f に対する条件付き期待値関数を用いた近似を行い、計算コストを軽減させている。

本研究では、説明対象モデル f は DCNN モデルに該当し、SHAP 値を深層学習に特化させた手法で計算する Deep SHAP を用いている。Deep SHAP で用いられている DeepLIFT アルゴリズムの Rescale Rule¹⁷⁾ による線形近似を f の非線形部に適用している。その結果、 f 全体を線形モデルに近似して、SHAP 値を計算する。

3 実験手法

3.1 データセット

データセットとして、東北大学オートプシーイメージングセンターにおいて撮影された溺死 128 症例、非溺死 135 症例の死後肺 CT 画像を用いた。スライス中で肺の陰影が不鮮明な画像を除き、溺死 2047 枚、非溺死 2160 枚を訓練に、溺死 400 枚、非溺死 400 枚をテストに使用した。また、前処理として、大津法¹⁹⁾と単純なフィルタ処理による背景ノイズの削除を行った。その例を Fig.3-b に示す。

3.2 DCNN の学習と評価

本実験では DCNN モデルの一つである VGG16¹⁶⁾ (Fig.2) を用いた。全結合層の代わりに Gloval Average Pooling 層²⁰⁾を適用し、損失関数にはクロスエントロピー誤差、最適化関数には Adam²¹⁾を使用した。また、VGG16 の学習の際に、Fine-tuning²²⁾を用いた。Fine-tuning とは別のデータセットで行った訓練済みモデルに対し、最終層のみ対象の出力数に付け替え、対象データセットで再度訓練をする手法である。今回は Pytorch が提供している訓練済みモデルに対し、死後肺 CT 画像で訓練を行った。なお、訓練時の死後肺 CT 画像を入力する際に、複数の画像処理をランダムで適用した。その内容を以下に示す。

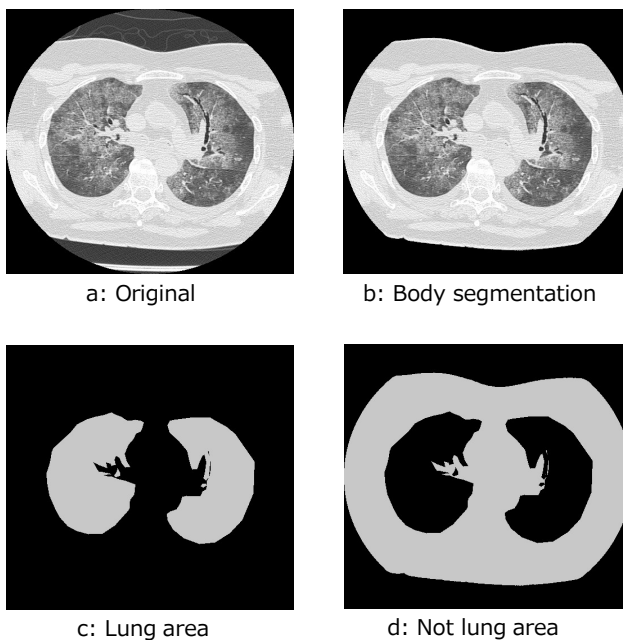


Fig. 3: Example of each image

- 水平方向の反転
- 垂直方向の反転
- 回転（最大 10°）
- 拡大と縮小（0.8 倍から 1.2 倍）

訓練後、テストデータを用いて鑑別性能の評価を行った。評価指標に、真陽性率と偽陽性率より受信者動作特性 (Receiver Operating Characteristics : ROC 曲線) と曲線下面積 (Area Under Cover : AUC) を用いた。AUC 値は 1.0 に近い程高い分類性能を示す。真陽性率 (True positive rate) は、溺死症例の中で溺死と判断できた割合、偽陽性率 (False positive rate) は、非溺死症例の中で、誤って溺死と判定した割合を示す。

3.3 診断根拠の解析

訓練後の DCNN モデルに対し、Grad-CAM と Deep SHAP を用いた診断根拠の可視化を行った。それぞれの可視化結果に対して分析を行い、診断根拠の解釈を試みた。

3.3.1 Grad-CAM を用いた解析

Grad-CAM の結果で得られるヒートマップに対し、以下の式より、肺領域 (Fig.3-c) への注目度 (Attention degree) を算出し、全体的な傾向を調査した。

$$\text{注目度} = \frac{\text{注目領域} \cap \text{肺領域}}{\text{注目領域}} \times \text{重み} \times 100 \quad (6)$$

上式において、注目領域は Grad-CAM により算出されたヒートマップ各値を昇順に整列し、小さい方から全体の 80% の位置にある値 (80 パーセンタイル値) を閾値として 2 値化した領域をとす。重みはヒートマップの強度を示し、肺と重なるヒートマップ領域の平均値 (0~1 に正規化) した値を用いる。以上より、注目度の計算を行っている。

同様の手法で肺以外 (Fig.3-d) に対する注目度も算出した。

3.3.2 Deep SHAP を用いた解析

Deep SHAP による注目特徴の可視化を行った。各入力要素に対する SHAP 値の算出結果に対し、最大値 255 で正規化を行った。その後、98 パーセンタイル値を閾値として 2 値化を行い、注目特徴とした。

上記の過程で得られた注目特徴に対し、溺水を疑う際に考慮される代表的な CT 画像所見⁶⁾⁻⁹⁾を参考にして比較を行った。以下の (ア) ~ (オ) に各所見の内容を示す。

- (ア) 副鼻腔内の液体貯留
- (イ) 気道内の液体貯留
- (ウ) 肺スリガラス影
- (エ) 胸膜腔液体貯留
- (オ) 胃などの消化管内の多量液体貯留と拡張

上記項目の内、今回用いた肺 CT 画像で確認できる (イ) ~ (エ) の所見に注目した。ここで、(イ) 肺スリガラス影 (ground glass shadow) とは肺野内に見られる白いモザイク上の所見を示す。

また、Deep SHAP の結果は、広い範囲に散布する場合が想定されるため、Grad-CAM との共通領域を求める事で、解釈可能性の向上を図った。

4 実験結果

VGG16による鑑別検証の結果として、テストデータに対し $AUC = 0.90$ (Fig.4) を示し、訓練後のモデルは高い鑑別性能を示した。

十分な鑑別性能を確認できた VGG16 に対し、Grad-CAM と Deep SHAP を用いて診断根拠の可視化とその分析を行った。なお、ここでは溺死確率を向上させる原因にのみ着目した。

4.1 Grad-CAM に基づく解析結果

全テストデータに対し、Grad-CAM による可視化と注目度 (Attention degree) の算出を行った。Table.1 にそれぞれの注目度の平均と相関を示した。また、Grad-CAM による可視化の例を Fig.5 に示す。

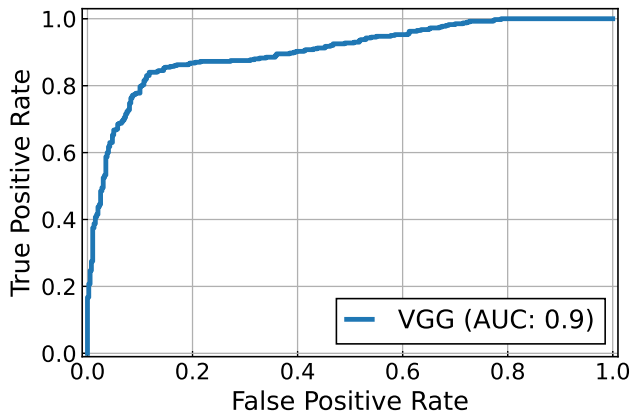


Fig. 4: ROC curve for drowning classification task

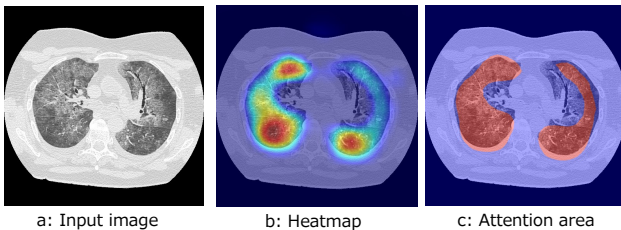


Fig. 5: a: Drowning case. b: Grad-CAM's result. The redder the area, the more evidence for prediction of drowning. c: Binarize the heatmap and the threshold was 80 percentile value when heatmap values were arranged in ascending order.

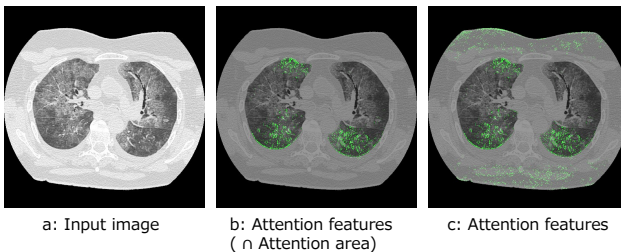


Fig. 6: Example of Attention features (made from Deep SHAP). Green points are Attention features in input image (a), and indicate contribution to the prediction of drowning. Attention features (b) is more interpretability than original (c).

Table 1: Attention degree for each area. dr: drowning case, ndr: not drowning case, correlation: VGG16's output for drowning class and Attention degree

		dr	ndr	total
lung	average	47.9	30.1	39.0
	correlation	0.74	0.55	0.77
not lung	average	3.22	4.6	3.91
	correlation	-0.26	-0.058	-0.26

注目度の平均値から、溺死クラスに対するヒートマップは主に肺野内にある事が分かった。また、Table.1 では、溺死の予測値と肺への注目度の間に正の相関が見られ、肺野内のある画像所見に対して強く反応する傾向を示唆した。

4.2 Deep SHAP に基づく解析結果

全テストデータ中から、撮影条件を考慮して各クラス 100 枚を選択し、Deep SHAP による解析を行った。

Fig.6 に注目特徴 (Attention features) の可視化結果を示した。Fig.6-b では、Grad-CAM による注目領域を用いてセグメンテーションを行う事で、肺野内の注目特徴 (緑色) のみを残し、解釈可能性の向上を確認できた。これを受けて以下では、Deep SHAP の結果に対し、注目領域でセグメンテーションしたものを注目特徴として解析を行った。

Fig.7 では、溺死症例を溺死と鑑別 (True Positive, TP) できたケースを 2 例、非溺死症例を非溺死と鑑別 (True Negative, TN) できたケースを 2 例示した。3.3.2 章の肺 CT における溺水所見 (イ) ~ (エ) に基づいて評価を行うと、(ウ) 肺スリガラス影 (Fig.7-a1, 赤矢印) を診断根拠の一つとする事が確認できた。また、Fig.7-d2 では、スリガラス影に比べ濃度が高い所見 (Fig.7-d1, 青矢頭) には注目を示さず、スリガラス影と区別ができていた事を確認できた。

しかし、(イ) 気道内の液体貯留、(ウ) 胸膜腔液体貯留、を示す結果は確認できなかった。

5 考察

ここでは、偽陽性 (False Positive, FP) と、溺死確率を減衰させている要因について考察を行った。

5.1 偽陽性について

可視化結果の解析より、訓練後のモデルは主にスリガラス影に注目して鑑別をする事が確認できた。しかし、スリガラス影は非溺死症例にも見られる所見であり、溺水特有の所見ではない^{4), 10)}。そこで、誤って溺死と鑑別した偽陽性例 14 例 (100 例中) を調査した所、いずれもスリガラス影に対し強い反応を示している事が確認できた。

しかし、訓練モデルは高い分類性能 ($AUC=0.9$) を示しており、スリガラス影がある場合でも正確に鑑別ができていた例がいくつもある。従って、訓練後のモデルはある特定パターンのスリガラス影に強く反応している事が考えられた。実際に、特定のスリガラス影のパターンを溺死の特徴と示す報告があり^{4), 7)}、今後

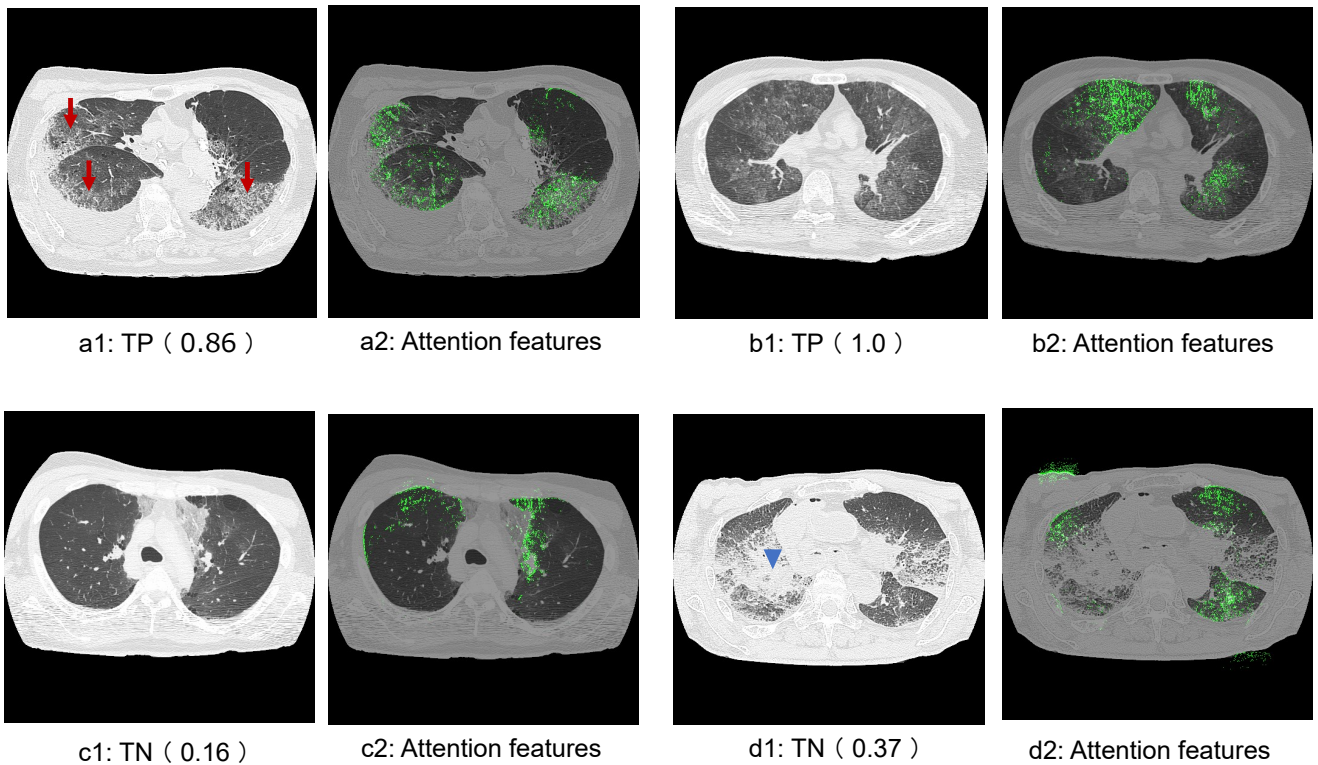


Fig. 7: Example of Attention features (Green points) for drowning. TP is True Positive, TN is True Negative, and each value is model's output for drowning class (a1, b1, c1, d1). Attention features (a2, b2, c2, d2) are tend to show ground glass shadow (GGO). Red arrows (a1) indicate GGO. Blue arrowhead (d1) indicates high CT values area in lung which is not GGO.



Fig. 8: Example of Attention features for non-drowning (Yellow points). These are tend to show respiratory tract. However, some of them are near the boundary between body and background.

分析を進めていく事で、医学的により高度な解釈につながる可能性が考えられた。

5.2 溺死確率を減衰させる特徴

溺死確率を減衰させる特徴について、同様の手法で分析を行った。その結果、傾向の一つとして液体貯留がない気道に対して反応を示す傾向が確認できた。Fig.8 にその例を示す。Fig.8 のように、気道内が空洞なケースは、各症例 100 例の中で、溺死症例 18 例、非溺死症例 49 例が確認でき、いずれに対しても Fig.8 同様の反

応を確認できた。これを受け、訓練後のモデルは、(イ) 気道内の液体貯留を考慮して溺死鑑別をしている事が考えられた。

また、Fig.8 でも示すように、胴体と背景の境界部付近に反応を示すケースを確認したが、医学的な解釈につなげる事は現状できなかった。今後分析を進め、用いた説明手法のエラーか否かといった事を検討していく必要がある。

6 おわりに

本研究では、深層学習の説明手法である Grad-CAM と Deep SHAP を応用し、非専門医でも解釈可能な診断根拠の解析を試みた。その結果、訓練後のモデルは、肺スリガラス影、気道内の液体貯留といった溺水の代表的な画像特徴に基づき溺死鑑別を行っていた。従って、画像診断診断の非専門医に対しても解釈可能な診断根拠を提示でき、信頼性が担保できる可能性を示唆した。

今後は、専門医の見解も含め、より高度な診断根拠の解析を行っていく事、また、解析の結果を基に鑑別性能の向上を目指していく。

7 謝辞

本研究の一部は、科研費 18K19892, 19H04479, 20K08012 ならびに東北大学スマートエイジング学際重点研究センターの助成を受けた。

参考文献

- 1) Japanese Association for Acute Medicine, (2019).
- 2) World Health Organization Global report on drowning: preventing a leading killer (2014).
- 3) Ministry of Health, Labor and Welfare Vital Statics (2019).
- 4) Usui, Akihito and Kawasumi, Yusuke and Funayama, Masato and Saito, Haruo: Postmortem lung features in drowning cases on computed tomography, *Japanese Journal of Radiology*, **32-7**, 414/420 (2014).
- 5) Kawasumi, Yusuke and Onozuka, Naoki and Kakizaki, Ayana and Usui, Akihito and Hosokai, Yoshiyuki and Sato, Miho and Saito, Haruo and Ishibashi, Tadashi and Hayashizaki, Yoshie and Funayama, Masato: Hypothermic death: Possibility of diagnosis by post-mortem computed tomography, *European Journal of Radiology*, **82-2**, 361/365 (2013).
- 6) Levy, Angela D and Harcke, H Theodore and Getz, John M and Mallak, Craig T and Caruso, James L and Pearse, Lisa and Frazier, Aletta A and Galvin, Jeffrey R: Virtual autopsy: two-and three-dimensional multidetector CT findings in drowning with autopsy comparison, *Radiology*, **243-3**, 862/868 (2007).
- 7) Christe, Andreas and Aghayev, Emin and Jackowski, Christian and Thali, Michael J and Vock, Peter: Drowning-post-mortem imaging findings by computed tomography, *European radiology*, **18-2**, 283/209 (2008).
- 8) Van Hoyweghen, Astrid JL and Jacobs, Werner and de Beeck, Bart Op and Parizel, Paul M: Can post-mortem CT reliably distinguish between drowning and non-drowning asphyxiation?, *International journal of legal medicine*, **129-1**, 156/164, (2015).
- 9) Vander Plaetsen, Sarah and De Letter, Els and Piette, Michel and Van Parys, Geert and Casselman, Jan W and Verstraete, Koenraad: Post-mortem evaluation of drowning with whole body CT, *Forensic science international*, **249**, 35/41 (2015)
- 10) 舟山真人, 斎藤春夫: Aiはどこまで事実に向かえるか, 31/43, 医歯薬出版株式会社 (2014)
- 11) 日本医学放射線学会, 北海道大学大学院医学研究院, 死因究明教育研究センター: 死後画像読影ガイドライン 2020 年版, 127/131, 金原出版株式会社 (2020).
- 12) Noriyasu Homma, Xiaoyong Zhang, Amber Qureshi, Takuya Konno, Yusuke Kawasumi, Akihito Usui, Masato Funayama, Ivo Bukovsky, Kei Ichiji, Norihiro Sugita, Makoto Yoshizawa: A Deep Learning Aided Drowning Diagnosis for Forensic Investigations using Post-Mortem Lung CT Images, 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 1262/1265 (2020).
- 13) 今井裕, 高野英行, 山本正二: Autopsy imaging ガイドライン第3版, 株式会社ベクトル・コア, 26/29 (2015).
- 14) Selvaraju, Ramprasaath R and Cogswell, Michael and Das, Abhishek and Vedantam, Ramakrishna and Parikh, Devi and Batra, Dhruv: Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE international conference on computer vision*, 618/626 (2017).
- 15) Lundberg, Scott and Lee, Su-In: A unified approach to interpreting model predictions, *arXiv preprint arXiv:1705.07874* (2017).
- 16) Simonyan, Karen and Zisserman, Andrew: Very deep convolutional networks for large scale visual recognition challenge, *arXiv preprint arXiv:1409.1556* (2014).
- 17) Shrikumar, Avanti and Greenside, Peyton and Kundaje, Anshul: Learning important features through propagating activation differences, In *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, 3145/3153 (2017).
- 18) Shapley, Lloyd S: A value for n-person games, *Classics in game theory*, **69** (1997).
- 19) Otsu, Nobuyuki: A threshold selection method from gray-level histograms, *IEEE transactions on systems, man, and cybernetics*, **9-1**, 62/66 (1979).
- 20) Li, Zhi and Wang, Shui-Hua and Fan, Rui-Rui and Cao, Gang and Zhang, Yu-Dong and Guo, Ting: Network In Network, *International Journal of Imaging Systems and Technology*, **29-4**, 577/583 (2019).
- 21) Kingma, Diederik P and Ba, Jimmy: Adam: a method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014).
- 22) S. J. Pan and Q. Yang: A Survey on Transfer Learning, *Knowledge and Data Engineering, IEEE Transactions on*, **22-10**, 1345/1359, (2010).