

		指導教員	下園 真一 准教授
学生番号	192C1060	氏名	古賀 隆一
論文題目	GeoHash と kd-木を融合したデータ構造の設計		

## 1 はじめに

地図上の特定の領域に関するデータを取得したい場合, Open Street Mapなどのデジタルマップが用いられている。これを扱う情報処理システムでは、必要な領域、また、その周辺領域のデータのみをいかに効率よく取り出すことができるかが鍵になる。そこで、地図上の任意の領域を特定する GeoHash と相性が良いと思われる kd-木を組み合わせたデータ構造を提案する。

## 2 定義

GeoHash は、緯度と経度で交互に領域の分割を行う。kd-木は、x 軸方向、y 軸方向にデータ点をソートし、データ点の座標の中央値を軸として x 軸方向と y 軸方向で交互に分割を行う。GeoTree とは、kd-木に用いられている x 軸、y 軸を GeoHash で用いられている緯度、経度とみなし、さらに座標区間を 2 等分するという分割方法を組み合わせた GeoHash と kd-木を融合したデータ構造である。GeoTree は、集合ノードを葉とし、分割ノードを内部ノードとする根付き二分木で、以下を満たすものをいう。

- i) 根ノード  $r$  は長さ 0 の GeoHash  $h(r) = \varepsilon$  を持つ。
- ii) 根以外のすべてのノード  $v \neq r$  について、その GeoHash の長さは親  $\text{parent}(v)$  の GeoHash  $h(\text{parent}(v))$  の長さより 1 だけ長い。
- iii) 左の子の GeoHash が表す領域は、東または北の領域であり、右の子の GeoHash が表す領域は、西または南の領域である。

## 3 アルゴリズム

探索キーとなる GeoHash を  $b \in \{0, 1\}^*$  とする。 $b$  により与えられた矩形領域の範囲内にあるデータ点を探索する。

1. 木の根からスタートする。
2.  $b$  のビットによって左または右の子に進む。
3.  $b$  のビットに従った分岐が終了した時点で、到達したノードが葉であるか、内部ノードであるかによって、返すデータ点を求める。

## 4 実験

GeoTree を用いた、木の構築時間を計測する。条件として、今回用意したデータ点は、飯塚市のデータ点(17712 個)と福岡市のデータ点(696684 個)の 2 組である。最大深さとは、GeoTree を用いて構築される木の最大の深さである。また、それぞれの計測時間は、100 回の実行時間の平均とする。飯塚市のデータ点については、下の表のようになった。

表 1: データ点の数=17712

最大深さ	木の構築時間 [ms]	葉ノード数
1	9.6	1
10	19.4	1
20	30.6	1
30	39.4	89
40	68.9	9582
50	181.9	17699
56	312.2	17711

## 5 まとめ

本研究では、GeoHash と kd-木を融合した GeoTree というデータ構造を設計し、GeoHash を用いて領域を探索するアルゴリズムを作製した。座標値(地理点 2 点による矩形の指定)による領域探索は、作製できていない。これを作製することで、利便性が高いデータ構造が完成するのではないかと考えられる。

## 参考文献

- [1] 渡部有隆、プログラミングコンテスト攻略のためのアルゴリズムとデータ構造、マイナビ出版、2015.
- [2] GeoHash 変換と範囲の表示：<https://geohashsite.app/>
- [3] 直行領域探索：<https://www.slideshare.net/okuraofvegetable/ss-65377588>
- [4] ジオハッシュ：<https://ja.wikipedia.org/wiki/>

		指導教員	下園 真一 准教授
学生番号	192C1118	氏名	羽方 綾音
論文題目	曖昧照合と非決定性オートマトンによるメロディ検索		

## 1 はじめに

誰かの鼻歌を聞いたとき、曲のメロディはわかるが、曲名やアーティスト名はわからない、ということがよく起きる。そこで、メロディ概形、すなわち音の上がり下がりで検索するシステムが開発されている[1, 2]。

本研究は、この検索システムを相対音感に自信がないユーザーでも利用しやすいように、おおまかな音の上がり下がりで検索（曖昧照合）を可能にする。自信がないときに利用できる曖昧記号を導入し、C 言語等のライブラリで利用できる正規表現を用いるもの、ナイープマッチングに曖昧記号を処理する修正を加えたもの、そして NFA を利用した楽曲検索システムを開発する。実際には膨大な数の曲から検索する必要があるため、高速に検索できるアルゴリズムを目指す。

## 2 定義

概形文字列  $t$  は、MIDI ファイルの数値化された音程列から差分を取り、-127 以上 -2 以下を'-'に、-1 を'b'、0 を'=', 1 を'#', 2 以上 127 以下を'+'に置き換えた文字列のことである。音程マッチングは、これらの記号に加え、任意の文字列にマッチする記号'\*', {+, #} いづれかにマッチする記号'^', {-, b} のいづれかにマッチする記号'\_'の 8 種類の文字で検索メロディ  $s$  を入力し、 $s$  が  $t$  内で最初にマッチする位置を見つける問題である。

## 3 NFA の作成と動作

今回新たに開発した NFA の作成と動作アルゴリズムについて述べる。ユーザーが入力した検索メロディ  $s$  に対し、 $s$  が出現する概形文字列を受理する NFA  $M = (\Sigma, Q, \delta, q_0, F)$  を作成する。

- 状態を 0, 1, ..., | $s$ | =  $n$  とし、先頭 0, 末尾  $n$ , および  $s[i] = ^*$  となる状態  $i$  に自己ループを定義する。初期状態は 0 のみ、受理状態  $n$  のみとする。
- $s[i]$  が -, b, =, #, + のいづれかであれば、状態  $i$  から  $i+1$  への遷移を定義する。
- $s[i]$  が^(または\_)ならば、状態  $i$  から  $i+1$  への遷移を + と #(または - と b) について定義する。

このように作成された NFA を、概形文字列  $t$  に対し、以下のように動作させる。

- $S \leftarrow q_0$ . /\*  $S$  は現在の状態の集合. \*/
- $t$  から 1 文字  $a$  を読み込み、 $S \leftarrow \bigcup_{q \in S} \delta(q, a)$  と遷移する。/\*  $\delta : Q \times \sigma \rightarrow 2^Q$  \*/
- $S \cap F \neq \emptyset$  ならば、 $t$  を受理して停止する。そうでなければ、2 に戻る。
- $t$  をすべて読み終えたら、却下して停止する。

この NFA の状態集合による決定的動作を、ビット表現により実装した。プログラムでは GCC ビルトイン関数のビット操作命令を用いて、高速化を行っている。

## 4 実験と考察

5344 曲分の MIDI ファイルを用意し、その中の適当な 5 曲からメロディを取り出し、音程マッチングの実行時間を調べた。検索メロディの長さは 3~10 字の間で変化させ、曖昧記号を 0, 50, 100% の割合で入れた。ある曖昧記号 100% のメロディで検索を行ったときのそれぞれの実行時間(ミリ秒)を以下の表にまとめる。

メロディ長	3	5	7	9	10
正規表現	537.2	581.2	664	707.6	726.4
naive	384.6	375.6	379.6	375.2	376.6
NFA	382.4	381.4	379.8	380.2	383.6

結果、NFA を利用した方法では正規表現版より圧倒的に早く、ナイープマッチングよりはわずかに遅かった。また、正規表現は検索メロディを長くすることで実行時間が長くなり、曖昧さの割合が増えるほどグラフの傾斜が強くなった。反対に、ナイープマッチングや NFA は検索メロディの長さや曖昧さの割合にそれほど影響を受けないことが確認できた。

今後の改良点として、NFA では、さらに遷移で必要になるループ数を減らし、高速化することが考えられる。また、今回は精度に関して考慮しなかったが、音程変化を表すメロディ概形だけでなく音符の長さや区切りとしての休符なども用いることで精度向上が見込まれると考えられる。

## 参考文献

- [1] Desnys Parsons, *The Directory of Tunes and Musical Themes*, Bohmier Verlag, 1975.
- [2] 概形文字列とテンポの照合によるメロディ検索, 川村優太, 2021 年度卒業論文。

		指導教員	下園 真一 准教授
学生番号	192C1159	氏名	安永 純也
論文題目	B-木によるマイコンをターゲットとした データベースシステム		

## 1 はじめに

二色木など主メモリーに全データを常駐させることができ前提のデータ構造は、扱うデータ数が大きくなってくると、主記憶が小さいマイコンなどのエッジデバイスではそのままの使用が難しい。また、主メモリーが揮発性であることを考えると、停電などに対応するためバックアップを頻繁に行う必要がある。

本研究では、B-木を用いることで、2次記憶装置上にデータを置くマイコンのようなエッジデバイスに対応した耐障害性・冗長性・整合性のあるシステムを構築することを考える。

## 2 定義

B-木 [1] は、ブロックデバイス上にキーを含むレコードをストアするデータベースシステムで用いられており、2次記憶装置上で効率よく動くように設計された平衡探索木である。使用する記憶装置のブロックサイズに合わせてノードが保持できるレコードの数を最適化し、アクセスを最小限にすることができる。検索・追加・削除の操作における計算量は  $O(\log N)$  である。

一つのノードが格納できるキー数には上界と下界が存在し、B-木の最小次数と呼ばれる整定数  $t \geq 2$  を用いて表現できる。すなわち根を除くすべてのノードはキーを少なくとも  $t - 1$  個、子を  $t$  個持つ。また、どのノードも最大  $2t - 1$  個のキーを持つことができ、最大  $2t$  個の子を持つことができる。

## 3 バイナリファイル上への B-木の実装

本研究では、C++を用いてバイナリファイル上にデータをストアする B-木の実装をした。フラッシュメモリや SD カードを起動ディスクとするシステムであれば、データはすべてフラッシュメモリー上にストアされることになる。主メモリ上におけるノード数に限りがあることが前提となるため、操作において必要となるノードをその都度バイナリファイル上から得る。

バイナリファイルには全てのノードとノードの数を記録している。それぞれのノードには id が振られており、ノードが持つ子ノードへのポインタのかわりに子ノードの id を使っている。

## 4 実験

実装したデータベースシステムの動作を検証する実験を行った。電気錠システムなどで学生番号がデータとして入ることを想定し、キーの大きさは 8byte とした。 $t = 3$  とした空の木へのキーの追加数の違いによる木の構築時間の比較と、キーの数による検索時間の比較を行う。測定時間にはファイル入出力の時間も含まれている。キーの追加数は 10,000 個と 100,000 個、構築後の木に対して検索を 10,000 件行った。それぞれの実行時間は表 1 のようになった。検索時間は大きさ 10,000 の木で 1 件あたり平均  $0.021 \times 10^{-3}$  秒、100,000 の木で  $0.032 \times 10^{-3}$  秒となり、キーの登録数が 10,000 個や 100,000 個でも高速に動作することが分かった。

表 1: 構築および検索にかかった時間 [秒]

10,000 個		100,000 個	
構築	検索	構築	検索
2.24	$0.021 \times 10^{-3}$	20.9	$0.032 \times 10^{-3}$

## 5 考察

B-木を用いたバイナリファイル上でのデータベースシステムを構築することができた。レコードとしてキーのほかに IC カードチップの固有 ID などを持たせても、十分高速なまま使用できると考えられる。また、B-木の最小次数  $t$  を大きくすれば、レコードの数を増やしても木の高さを低く維持できる。

このシステムは、耐障害性・冗長性・整合性のあるシステムであるが、今回の研究では、実際に障害があった場合に正常に動くか調べることができなかった。また、頻繁に削除と挿入が行われた場合に必要となるガベージコレクションは実装していない。これらが今後の課題である。

## 参考文献

- [1] T. コルメン, C. ライザーソン, R. リベスト, C. シュタイン, 浅野哲夫 ら訳, アルゴリズムイントロダクション (改訂 2 版), 第 2 卷 アルゴリズムの設計と解析手法 近代科学社, 2007.

		指導教員	下薙 真一 准教授
学生番号	192C1179	氏名	李 昂
論文題目	デジタルマップを利用した ウォーキング・ジョギングコースの提示		

## 1 はじめに

広く使われるようになったスマートフォン等の地図アプリやカーナビゲーションシステムでは、目的地に向かう経路を所要時間や移動手段の種類等を考慮した上で第二、第三の選択肢を提案することが可能である。一方、ウォーキングやジョギングをする際には、目標距離を達成するコースを探す必要があるが、前述した経路検索システムは距離を指定した経路の検索ができない、この用途には向きである。そこで本研究では、ウォーキングやジョギングのコースを検索できるよう、ダイクストラのアルゴリズムをもとに開始地点と目標距離による経路検索を行うアルゴリズムを設計した。さらに、ウォーキングやジョギングをする際に目的地がある場合には、目的地に立ち寄ることのできる経路の検索を行うようにした。

## 2 ウォーキング・ジョギングコース

本研究でのウォーキング・ジョギングコースは、開始地点からスタートして開始地点に戻り、移動距離を目標距離に可能な限り近づけることを重視する。基本は、開始地点からある地点まで進んだ後、通ってきたルートを折り返して開始地点まで戻るコースで、目標距離に最も近いものとする。目的地を指定した場合は、開始地点から目的地まで進んだ後、移動距離が目標距離に近づくようなルートで開始地点まで戻るコースとした。

## 3 アルゴリズム

本研究では、重み付き無向グラフの 2 頂点間の最短経路を求める方法であるダイクストラのアルゴリズム [1] を利用してウォーキング・ジョギングコースの検索を行う。

開始地点と目標距離のみを指定した経路の探索では、目標距離の半分の距離と開始地点からの距離の差が最小になるような点を探す。そして、開始地点からのその点まで移動した後、折り返して開始地点まで戻るようにする。

目的地をふまえた経路の探索では、開始地点と目的地に加え、第 3 の点を通る 3 点移動を行う。これら 3 点の各 2 点間に對し、ダイクストラのアルゴリズムを計 3 回適用することになる。

## 4 実験結果

地図データとして範囲が北緯 33.7176 度から 33.6021 度、東経 130.722 度から 130.634 度の頂点数が 17712 個のデータを用意して実験を行った。このデータは福岡県飯塚市の地図情報である。九州工業大学の飯塚キャンパス正門前を開始地点とし、目的地を指定しない場合と目的地として新飯塚駅を指定した場合のそれぞれで目標距離を 7500m, 10000m, 15000 とした場合の 3 通りで検証をした。実際に求まった経路の距離、CPU 時間は表 1、表 2 にのようになり、目的地を指定しない場合と目的地を指定した場合の両方で求まった距離と目標距離がほぼ等しくなっていることが確かめられた。

表 1: 目的地を指定しない場合

目標距離 [m]	7,500	10,000	15,000
求まった距離 [m]	7,500	9,999	15,006
CPU 時間 [ms]	13,450	15,700	16,307

表 2: 目的地を指定した場合

目標距離 [m]	7,500	10,000	15,000
求まった距離 [m]	7,461	9,988	14,882
CPU 時間 [ms]	22,168	26,765	29,610

## 5 まとめ

本研究で利用したデジタルマップでは、駐車場やあぜ道などのように地図には記載されているものの、道として情報が登録されてない道がある。GPS の移動履歴を使い、これらの通り道を見つけ、マップデータの補完情報として活用できれば、より多くのルートを探索対象にし、より希望する距離に近いウォーキング・ジョギングコースを提示することができるようになる。また、ウォーキング・ジョギングをする際は、具体的な目的地が決まっていないものの、コンビニエンスストアやスーパーマーケットなどの特定の施設を経由したい場合がある。目的とする施設がある場合には、その施設を経由するルートを提案できるようになるとウォーキング・ジョギングコースの経路検索としてより優れたものになると考えられる。

## 参考文献

- [1] 渡部有隆, プログラミングコンテスト攻略のためのアルゴリズムとデータ構造, マイナビ出版, 2015.

		指導教員	中村 貞吾 准教授
学生番号	192C1109	氏名	長田 周也
論文題目	マルチエージェント強化学習を用いた 分業的行動の獲得		

## 1 はじめに

近年、マルチエージェント強化学習が注目され、アルゴリズムや報酬設計が多く研究されているが、そのエージェントがどのように協調行動をふるまうかはあまり議論されていない。本論文では、エージェントが他のエージェントと協力し、臨機応変にタスクを分業する行動の獲得を目的とする。

## 2 Cooperative Push Block 環境

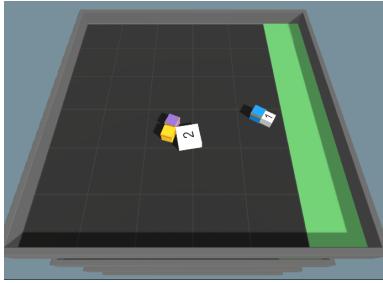


図 1: Cooperative Push Block 環境

本研究では Unity ML-agents[1] が提供する強化学習環境に変更を加えた Cooperative Push Block 環境を扱う(図 1)。3人のエージェントは 1-ブロックと 2-ブロックを協力してゴールへ運び、運んだブロックに書かれた数字分の報酬を受け取る。ブロックはその数字以上のエージェントが同時に押さなければ動かないため協力が必須の環境であり、また step ごとにマイナスの報酬が与えられるため、タスクを早く完了させるために臨機応変に 1-ブロックと 2-ブロックのどちらを押すかの分業行動も求められる。

## 3 分業の獲得

望月ら [2] は、臨機応変に分業体制をとることを創発的分業と呼び、教育現場においてそれを学習するために互いにタスクの状況を知ることが必要と述べている。その知見をマルチエージェント強化学習にも適用し、エージェントの視界を広げることで分業の獲得を目指す。また、創発的分業を行った際のメリットとしてタスクが早く終了することがあげられるため、タスクの早期終了を目的として、1 step のペナルティを増やし協調行動の変化に着目する。

## 4 実験

環境の 1step ごとのマイナス報酬を  $-0.5, -5, -30, -60, -120$  をそれぞれ最大ステップ数で割ったものに変更し、視野の大きさに対しタスクが終了するまでの平均エピソード長を調査し表 1 に示した。 $25 \times 25$  のフィールドにおいてエージェントの前方  $20 \times$  横方向  $20$  のみの視界を視野小、前方  $20 \times$  横方向  $25$  に後方  $5 \times$  横方向  $25$  を追加した視界を視野大とする。アルゴリズムに MA-POCA(Multi-Agent POsthumous Credit Assignment) を用い、2000 万回学習させた。なお、事前にカリキュラム学習でブロックを押すことができるエージェントを作成し、そのモデルに追加学習を行っている。

表 1: 平均エピソード長の比較

ペナルティ	0.5	5	30	60	120
視野大	37.39	36.22	35.26	34.76	34.08
視野小	41.88	39.58	38.39	39.27	42.64

視野大のエージェントはペナルティを増加させるほど平均エピソードの長さが短くなり、エージェントが効率よくタスクを成功させていることがわかる。また、実際のエージェントの動きに着目すると、30以下の少ないペナルティだと 1-ブロックを 2 人で押してしまう非効率な場面が多く登場したが、高いペナルティになると分業的行動が見られ、1-ブロックは 1 人で、2-ブロックは 2 人でそれぞれ近いエージェントが押し始めるような行動をみせ、分業行動が獲得できていた。その違いが平均エピソード長の差に表れていると考えられる。

## 5 まとめ

マルチエージェント強化学習において、時間ペナルティを与え行動が洗練されることで、分業行動が生まれることが判明した。今後の課題としては、難しいタスクにおいても分業行動が生まれるか、また最適の時間ペナルティをどのように設定するかなどがあげられる。

## 参考文献

- [1] Juliani et al, Unity: A general platform for intelligent agents, arXiv preprint arXiv:1809.02627, 2020.

		指 導 教 員	中村 貞吾 准教授
学生 番 号	192C1170	氏 名	山本 知佳
論 文 題 目	強化学習を用いた 花札の「こいこい」における AI プレイヤの作成		

## 1 はじめに

花札を用いて遊ぶ「こいこい」は不完全情報ゲームの一種である。本研究では、強化学習を用いて「こいこい」におけるどの手札を出すかの判断とあがるか否かの判断を行う AI プレイヤを作成した。

## 2 花札の「こいこい」

「こいこい」は 2 人のプレイヤが交互に札を取り合い、獲得した札で出来役を作り、得る点数を競うゲームである。交互に手札から 1 枚札を場に出し、同じ柄の札が場にあった場合は持ち札として獲得する。互いに持ち札を増やしていく、持ち札の中で役を作成できたプレイヤは、あがるかゲームを継続して更に高得点を目指すかの選択ができる。後者はこいこいをするといわれ、こいこいをした場合、追加で役を作成できれば更に多くの点を獲得できるが、相手が自身より早く役を作成してしまった場合は自身の得点がなくなるリスクがある。相手の手札の種類や、山にあるカードの種類はわからないため、相手の持ち札や手札の枚数を考慮して駆け引きを行う必要がある。

## 3 強化学習を用いたプレイヤの作成

「こいこい」をプレイする AI プレイヤについて、出し札を選択する部分を作成し、その後作成した出し札選択部を用いて、あがるか否か判断する部分を作成する。学習には DQN を用い、対戦相手にはランダム行動プレイヤとルールベース型プレイヤの 2 種を用いる。ニューラルネットワークの入力に用いる状態には、相手の情報がすべて見える状態にしてから学習を行い、学習が進むにつれ不完全情報に移行していく Oracle guiding を用いる。

### 3.1 出し札選択部の作成

毎ターンどの手札を出すのか選択する部分を作成する。学習において、役完成時にこいこいはせず必ずあがるものとする。報酬は即時報酬とゲーム終了時に与えられる遅延報酬の 2 種類を用いる。即時報酬はターン毎に与える。持ち札を増やすことのできる状況で持ち札を増やすことのできる手札以外を選択した場合は -1、持ち札を増やすことのできる手札を出した場合や、どの手札を選択してもそもそも持ち札を増やせないような場合は +1 を与える。また遅延報酬は、ゲーム終了時

に獲得した役得点 × 8 の報酬を与える。相手プレイヤが役を獲得した場合はその分マイナスの報酬を与える。

### 3.2 あがり判断部の作成

自身の役が揃ったときにこいこいをするか否か選択する部分を作成する。与える報酬は得た役の得点分だけ与えられ、相手が点を榨取した場合はその分マイナスの報酬が与えられる。また、出し札選択部分は 3.1 を通して作成した、Oracle guiding を用いず学習を行った場合のプレイヤを用いる。

### 3.3 実験結果

3.1 と 3.2 にて作成したプレイヤを用いて 1000 回対戦を行った場合の平均獲得点を以下の表 1 と表 2 に示す。表 1 より、2 種の対戦相手に対して勝ち越していることがわかる。あがり判断部分を含めた実験では、平均獲得点が 1.54 点となる出し札選択のプレイヤを用いているため、表 2 よりこちらでも獲得点を向上させていることがわかる。

表 1: 出し札選択部の平均獲得点

入力に用いた状態	対戦プレイヤ	平均獲得点
Oracle guiding	ランダム行動	0.26
	ルールベース型	0.22

表 2: あがり判断部分を含めた場合の平均獲得点

#### あがり判断部分の

入力に用いた状態	対戦プレイヤ	平均獲得点
Oracle guiding	ランダム行動	1.72
	ルールベース型	1.67

## 4 おわりに

本研究では DQN を用いて花札の「こいこい」を行うプレイヤを作成した。今後の課題として、モンテカルロ木法を組み合わせることによる、より平均獲得点を向上させられるようなプレイヤの作成や、数ゲームを通してプレイし、最終的な合計得点を競うようなルールでの「こいこい」を行うプレイヤの作成などが挙げられる。

## 参考文献

- [1] 佐藤直之、池田心、"花札のこいこいにおける方策勾配法と Neural Fitted Q Iteration の適用", GPW 2017.

		指導教員	中村 貞吾 准教授
学生番号	21221001	氏名	磯崎 遼斗
論文題目	レジスタンス: アヴァロンのプレイ環境構築と 信頼度に基づく AI プレイヤの作成		

## 1 はじめに

人狼より活発に他者の説得や協力が必要とされるゲームとして、レジスタンス: アヴァロンがあるが、人狼AI[1]に比べてAI開発が進んでいない。そこで本研究では、レジスタンス: アヴァロンにおいてAIがプレイできる環境構築とAIのプレイヤの作成を行う。

## 2 レジスタンス: アヴァロン

レジスタンス: アヴァロンは人狼と同様に会話を用いて進行していく正体秘匿系ゲームである。人狼との違いとして、途中でプレイヤが脱落しないこと、ゲーム最後に邪悪側が特殊役職を当てた場合逆転勝利できることなどが挙げられるため、戦略が制限される。また、人狼とは異なり、単純な投票行動だけでは意図が不明確であるため自分の意見への同意を得るには他者を納得させるような説明が必要となる。これらの点からレジスタンス: アヴァロンは人狼よりも他者への説得や協力が多く求められるゲームであるといえる。

## 3 説得行動

### 3.1 説得行動の表現

説得過程は、まず質問等で説得の鍵となる反応や情報を取得し、次に説得につながる説明や主張を行い、次に具体的な説得または説得の一部となる要望をする、という流れの繰り返しとみることができる。そこで本研究では、質問→主張→要望という流れを説得における基本単位とし、この繰り返しで説得行動を表現する。

### 3.2 割込み

ゲームをターン制で進行する場合、自分のターンにのみメッセージを送ることができる。しかし、自分のターン外で不利となる話し合いが行われた場合、反論するには自分のターンまで待たねばならず、即応的な反論に比べて効果が薄い。これを解消するために即応的な発言として割込みを実装した。割込みは反論したい場合以外にも、自分のターン終了後に割込むことで、発言の補強などにも使用することができる。

### 3.3 強度

発言の細かいニュアンスを表現するために、発言強度を実装した。発言強度は0から5の6段階に分かれしており、例えば0は「であることはない」、5は「であると確信している」のようなイメージとなっている。

## 4 信頼度

正体秘匿系ゲームでは、自分以外の陣営が分からぬいため今まで他プレイヤが行った行動や発言の結果を見て、そのプレイヤを信用するかどうか判断する必要がある。これを数値で表したものを作成する。

### 4.1 信頼度プレイヤ

信頼度に基づいて行動するAIプレイヤの作成を行った。このプレイヤは、最初に参加者全員の信頼度を0.47から0.52の幅でランダムで割り振り、信頼度の高い順にチーム編成を行う。信頼度の更新は各クエスト終了後に行い、クエストが成功した場合には参加したプレイヤ全員の信頼度を上げ、失敗した場合には参加したプレイヤ全員の信頼度を下げるようとした。

## 5 実験

### 5.1 プレイ環境

今回構築したレジスタンス: アヴァロンのシステムはRubyで作成し、サーバとプレイヤ人数分のクライアント間でTCP/IP通信を行う形で動作する。

### 5.2 対戦実験

クエスト結果から他プレイヤの信頼度を変化させる信頼度プレイヤ、チーム編成等をランダムで行うランダムプレイヤ、クエスト参加時に7割の確率で失敗に投票する邪悪プレイヤを使い対戦実験を行った。また、信頼度プレイヤの信頼度変化量は0.15とした。正義側3人を信頼度プレイヤ、邪悪側2人を邪悪プレイヤにした場合と、正義側3人をランダムプレイヤ、邪悪側2人を邪悪プレイヤにした場合で、それぞれ100試合した結果、信頼度プレイヤの勝率が38%、ランダムプレイヤの勝率が19%となった。この結果から信頼度はある程度勝率に関与していることが確認できる。

## 6 おわりに

本研究では、AIと人間がレジスタンス: アヴァロンをプレイできる環境と、信頼度に基づいたAIプレイヤを作成した。AIプレイヤでは信頼度を用いることで、正義側の勝率を19%上げることができた。今後の課題として、チーム編成の根拠の提示や信頼度を更新する場面の追加などがあげられる。

## 参考文献

- [1] 人狼知能プロジェクト, <http://aiwolf.org/>.

		指導教員	中村 貞吾 准教授
学生番号	21221002	氏名	出籠 聖弥
論文題目	FPS ゲーム「VALORANT」におけるプレイ映像からの勝敗予測		

## 1 はじめに

FPS ゲーム「VALORANT」は、様々な要素が勝敗に影響するため、どちらのチームが優勢なのか分かりづらい場合がある。そこで本研究では、大会観戦者や実況者の支援をするために、試合映像から有利不利状況を判定するシステムの構築を目的とする。それによって試合の解説にもつなげることが可能となる。

## 2 VALORANT

VALORANT は、爆弾を設置、起爆する攻撃側 5 名とそれを阻止する防衛側 5 名の合計 10 名のプレイヤーが、最大 25 ラウンド（1 ラウンド最大 145 秒、12 ラウンドで攻守交替、13 ラウンド先取で勝利）を攻防する。各プレイヤーはキャラクターを操作し、それぞれの独自の能力（アビリティ）や武器を使用して対戦する。

## 3 勝敗予測

本研究では、プレイ映像 [1] から攻守状況、残り時間、爆弾の設置状況、マップの種類、キャラクターの種類、武器の種類、アビリティの数、アルティメットアビリティの有無、残り体力、結果の情報を取得し、その情報を用いて機械学習により勝敗予測を行う。図 1 にプレイ映像の一例を示す。



図 1: プレイ映像の一例

### 3.1 データセット

プレイ映像から上で示した計 10 種の情報 421 次元を 1 秒ごとに取得し、合計で 5295 ラウンド分、約 40 万のデータセットを作成した。データセットのうち、解説者によって実際に解説されたラウンドをテストデータ（約 4 万）とし、それ以外を訓練データとした。

### 3.2 ネットワーク構成

取得した情報 421 次元のうち、正解ラベルの情報 1 次元を除いた 420 次元を入力に、勝率を出力するニューラルネットワークを作成した。中間層はノード数 256

の全結合層を 3 層とし、活性化関数には ReLU を使用している。

## 4 実験

訓練データの 1 割を検証用データとし、EarlyStopping により検証用データの正解率が改善されなくなるまで学習を行い、テストデータを用いて評価を行う。

### 4.1 モデルの評価

本研究では、人数状況 5 対 5 かつ残り時間 46 秒以上の状態を序盤、残り時間 45 秒以下の状態を終盤、それ以外を中盤としてテストデータを分け、勝敗予測結果の正解率を求めた。結果は、序盤は 0.629、中盤は 0.728、終盤は 0.820、全体で 0.733 となった。ラウンドの終盤になるほど勝敗予測の精度が高くなることが確認できる。

### 4.2 ラウンド別の評価

実際の解説と比較し、有利不利状況を捉えているかを評価した。縦軸を攻撃側の勝率、横軸をラウンドの経過時間とした、あるラウンドの攻撃側の勝敗予測を図 2 に示す。実際の解説では、「序盤から防衛側は武器的に不利だったが、アビリティの組み合わせ（約 25 秒後）により逆転した。」とある。図 2 から要因は解説できないが、有利不利状況は捉えていることが確認できる。

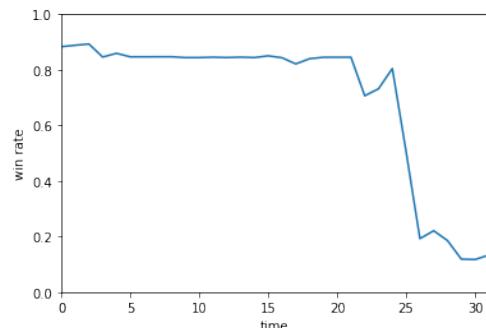


図 2: あるラウンドの攻撃側の勝敗予測

## 5 おわりに

本研究では約 7 割の精度で勝敗を予測するモデルを作成できた。今後の課題として、位置情報の適用、プレイヤーの技量の考慮などが挙げられる。

## 参考文献

- [1] VALORANT//JAPAN,  
<https://www.youtube.com/@VALORANTjp>

		指導教員	中村 貞吾 准教授
学生番号	21221004	氏名	岡本 龍馬
論文題目	麻雀における牌理に基づかない打牌に対する解説生成		

## 1 はじめに

近年、麻雀において、Suphx[1] や NAGA[2] といったトッププレイヤの実力に匹敵する麻雀 AI が開発されており、強いコンピュータプレイヤを作るという目標は実現されつつある。しかし、麻雀のプレーについて、AI によって言語化するための研究はまだ進んでいない。そこで、本研究では、麻雀の打牌に対する解説生成を目的とし、その打牌の利点や意図といった内容の推測や言語化を AI によって実現することを試みる。

## 2 牌理に基づかない打牌

牌理とは、最短で和了する手順のうち、和了する確率が最も高くなる手順を選ぶための理論である。麻雀において勝利するには、他のプレイヤよりも早く和了する必要があるため、牌理に基づく打牌に関しては「他のプレイヤよりも早く和了するために、その確率を最も高める打牌を行った」ということが解説として通じるが、そうではない打牌を解説するためには、牌理に基づく打牌を行なわなかった理由を推定する必要がある。

## 3 打牌に対する解説生成

### 3.1 解説生成の流れ

本研究では、牌理に基づかない打牌のうち、「リーチしているプレイヤに放銃しないことを意識した打牌」について注目し、以下の手順に基づいて解説を生成する。

1. リーチしているプレイヤがいる状況で、リーチしていないプレイヤが牌理に基づかない打牌を行った場合、その打牌と牌理に基づく打牌のどちらが安全かを推定する。
2. プレイヤの打牌の方が安全であると推定した場合、その打牌を「リーチしているプレイヤに放銃しないことを意識した打牌」と推定し、3 に進む。
3. プレイヤの打牌の方が安全であると推定した根拠を求め、それに基づいて解説文を生成する。

### 3.2 打牌の危険度推定

リーチしているプレイヤが 1 人以上存在する局面の情報を用いて、待ちパターンごとの危険度を推定するモデルの学習を行った。また、その結果をもとに、待ちパターンごとの危険度を用いて打牌の危険度を推定するモデルの学習を行った。モデルの学習には、それぞれニューラルネットワークを用いた。

### 3.3 特徴量の重要度測定

打牌の危険度推定の結果に対する根拠として用いる、特徴量の重要度について測定した。打牌の危険度を推定する際の特徴量の重要度については、Permutation Importance[3] を用いて測定した。また、待ちパターンごとの危険度を推定する際の特徴量の重要度については、テスト用のデータセットに対して待ちパターンごとの危険度推定を行い、ある特徴量について、その特徴量が 1 のデータに対する危険度の平均とその特徴量が 0 のデータに対する危険度の平均の差の絶対値をその特徴量の重要度とすることにより、測定した。

## 4 解説生成実験

牌理に基づく打牌の特徴量とそうでない打牌の特徴量を比較し、値が異なる特徴量のうち、最も重要度が高い特徴量を根拠とする解説生成実験を行った。本研究の手法を用いて生成した解説の例を以下に示す。

牌理に基づいて打つなら、打 8p だが、プレイヤ 1 に放銃しないことを意識して、プレイヤ 3 は、8p と比較すると、スジがノーチャンスでありリヤンメンに当たらない牌である、より危険度の低い打 9s とした。

リーチしているプレイヤ 1 に対して、牌理に基づく打牌よりもプレイヤ 3 の打牌の方が安全と推定し、その根拠としてリヤンメン待ちに当たらないこと、また、その根拠として、スジがノーチャンスであることを求め、これらの根拠に基づいた解説文が生成されている。

## 5 おわりに

本研究では、牌理に基づいていない打牌のうち、「リーチしているプレイヤに放銃しないことを意識した打牌」に対する解説の生成を行った。牌理に基づいていない打牌の理由として考えられるものは他にも多く存在するため、今後の課題として、それらの打牌についても解説を生成できるよう実験を行う必要がある。

## 参考文献

- [1] J.Li, et al. Suphx: Mastering Mahjong with Deep Reinforcement Learning, <https://arxiv.org/abs/2003.13590>, 2020.
- [2] DWANGO MEDIA VILLAGE, 深層学習麻雀 A-I 「NAGA」, [https://dmv.nico/ja/articles/mahjong-ai\\_naga/](https://dmv.nico/ja/articles/mahjong-ai_naga/), 2019
- [3] A.Altmann, et al. Random Forests, Machine Learning volume 45, p.5–32, 2001.

		指導教員	乃万 司 教授
学生番号	182C1016	氏名	出口 修己
論文題目	資料画像の視点・姿勢の調整が可能な人物イラスト制作支援システム		

## 1 はじめに

人物イラストを描く際、インターネット上や書籍から探したり自撮りしたりして集めた写真を、見本として参照することが多い。しかしながら、自身の描きたい人物のポーズの写真を探し出すことは難しく、類似したポーズの写真を見つけられたとしても、ポーズやアングルが完全に一致していることはまれである。

## 2 従来手法・類似手法

人物イラスト制作の資料を用意する方法には、デッサン人形や Clip Studio Paint などの仮想デッサン人形の利用もある。しかし、デッサン人形で意図するポーズを一から作るのは困難である。さらに実際の人体では困難なポーズを作ってしまうといった問題もある。

また、写真内の人体のポーズを取得する方法として、Asilla 3D Pose のような姿勢推定アルゴリズムを用いて写真の人物のポーズを抽出する手法がある [1]。しかし、従来の姿勢推定アルゴリズムを利用したシステムは行動分析を目的としたものが多く、ポーズやアングルを変更することは出来ない。

目的のポーズの写真、或いはそれに近い写真からポーズのデータを取得し、それを 3D 人体モデルに適用した後、ポーズやアングルの調整を行うことが出来れば、意図するポーズやアングルの人物イラストの見本を用意することが可能になる。

## 3 目的

本研究では、人物写真からその人物のポーズを取得し、そのポーズやアングルを調整した人物イラスト用の見本画像を生成できるシステムを開発する。

## 4 提案手法

上記の目的から、姿勢推定と 3 次元空間への描画という二つの処理を順に用いる。

姿勢推定では、人物の写真から人体のキーポイントの 3 次元座標を取得する。姿勢推定には、OpenPose[2] と MediaPipe[3] という二つの姿勢推定アルゴリズムを利用する。

3 次元空間への描画では、Blender と予め用意した人体モデル、ボーンを利用する。姿勢推定で取得した 3 次元座標を元に、予め用意した人体モデルの各ボーンのローカル回転角を設定し、ユーザに表示する。



図 1: 入力画像



図 2: 実験結果

## 5 実験

入力画像を図 1<sup>1</sup>、出力結果を図 2 に示す。正面からの見た目に関しては入力画像のポーズに近いものが 출력されたことが確認できた。しかし、z 座標(奥行)の影響を大きく受ける脚の開き方については、実際よりも大きく開いている印象があることが確認できた。

## 6 むすび

本研究では、人物写真からそのポーズやアングルを調整した人物イラスト用の見本画像を生成できるシステムの開発を目的とした。実現したシステムでは、自身の描きたいポーズに近い入力画像を用いることで、そのポーズの奥行に関してなど誤差はあるものの、類似したポーズをとる人体モデルを描写し、その人体モデルのポーズやアングルを調整できることを確認できた。本システムは使用する姿勢推定アルゴリズムの精度に大きく影響を受けるため、ポーズの精度向上を目的とするならば、他の姿勢推定アルゴリズムの利用、または新たな姿勢推定アルゴリズムの開発が必要となる。

## 参考文献

- [1] 株式会社 アジラ, 3 次元姿勢推定アルゴリズム「Asilla 3D Pose」のご紹介, <https://pr.asilla.jp/technology/a3dp-20211001>
- [2] Zhe Cao, et al., “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” arXiv:1812.08008, 2018.
- [3] Camillo Lugaressi, et al., “MediaPipe: A Framework for Building Perception Pipelines,” arXiv:1906.08172, 2019.

<sup>1</sup><https://github.com/CMU-Perceptual-Computing-Lab/openpose>

		指導教員	乃万 司 教授
学生番号	192C1034	氏名	大谷 信雄
論文題目	指示記号の記入による図形の対話的補正		

## 1 はじめに

近年、教育現場において電子黒板を用いて授業を行う機会が増えてきた [1]。しかし、算数や数学の授業で図形を示す際は、現在でも、黒板や紙などを用いることが多い。これは、電子的な方法では、その場で意図した図形を正確に描くことが難しいためである。市販のドローソフトや CAD ソフトにも、補助機能が備わっているが、授業中に電子黒板上で制約条件を目で確かめながら図形を描くのは、時間がかかりすぎる。本研究では、教育現場に適した速さと正確性で、電子黒板上の図形の作成を補助するシステムを開発する。

## 2 関連研究と関連製品

Pavlidis ら [2] や Wang ら [3] は、手書きのラフな図形から、制約条件を推測し正確な図形に補正する手法を提案した。しかし、正確な図形を作成する場合、制約条件は陽に与えたい場合が多く、これらの推測では不十分である。また、Bier ら [4] は、画面上でガイド線などを表示する手法を提案した。しかし、市販のソフトと同様に授業中に描くには時間がかかりすぎると不十分である。

## 3 ユーザーインターフェース

本システムは、画面上で手書きの記号で指示された線分間の垂直や平行、線分同士が端点で接続しているといった制約条件を保持しながら、対話的に図形を変化させていく。図 1 は線分 a, b, c において、予め垂直の制約条件  $b \perp c$  がある図形に対して、さらに別の垂直の制約  $a \perp b$  を付与して、 $b \perp c$  を維持したまま変化している様子である。

## 4 システムの実現

線分の位置を確定するためには、制約は高々 2 つあればよい。例えば、図 1 の最も右の図形において、線分 b は、線分 a の右側の端点を共有し、 $a \perp b$  となるようにすることで、位置を決定でき、線分 c は、線分 b の下側の端点を共有し、 $b \perp c$  となるようにすることで、位置を決定できる。本システムでは、このように、基準から図形間の関係を維持し、それらを基準から再帰的に走査することで、図形を描画していく。本システムでは、先に選択した線分を基準とすることで、

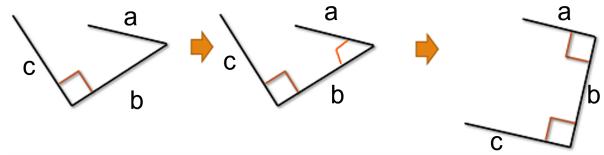


図 1: 垂直制約の追加による図形の変化

ユーザーが補正を制御できるようにする。

### 4.1 指示記号の認識

本システムには、図形描画、自由描画、指示記号編集、クリック編集の 4 つのモードがあり、そのうち、指示記号編集モードにおける書き込みを指示記号として認識する。また、指示記号書き込みの際に、近くにある線分を選択した線分として処理する。指示記号は、垂直と平行に対応しており、垂直の場合、2 つの線分を 1 度に選択をするものであり、平行は 2 つの線分を 2 つの書き込みで選択するものとして判別する。

## 5 むすび

本研究では、指示記号の記入によって図形を補正できるシステムを実現した。現在のシステムでは、端点が近い線分同士の接続、2 つの線分を垂直や平行にするといった補正のみしか行えず、指示記号の記入と線分の記入ではそれぞれ別のモードを選択する必要がある。今後は、指示記号として角度の指定や線分の長さの指定ができるようにしたい。また、図形と指示記号を区別できるようにし、モードの選択を不要にすることによって改善が必要である。

## 参考文献

- [1] 電子黒板の活用により得られる学習効果等に関する調査研究, 文部科学省, 2010, [https://www.mext.go.jp/a\\_menu/shotou/zyouhou/1297993.html](https://www.mext.go.jp/a_menu/shotou/zyouhou/1297993.html).
- [2] Pavlidis, et al., "An Automatic Beautifier for Drawings and Illustrations," Proc. of SIGGRAPH'85, 1985.
- [3] Wang, et al., "Exploring Sketch Beautification Techniques," Proc. of CHINZ '05, 2005.
- [4] Bier, et al., "Snap-Dragging," Proc. of SIGGRAPH'86, 1986.

		指導教員	乃万 司 教授
学生番号	192C1103	氏名	中島 悠貴
論文題目	自作した部品の選択による人体頭部モデリング		

## 1 はじめに

現在、メディアや建築、医療分野など、様々な場面で 3D モデルが利用されている。特に人体の 3D モデルはゲーム、アニメ、動画配信などで需要が大きい。モデルの制作にはポリゴンメッシュの頂点を挿入したり移動したりする方法や、形状を直接粘土のように扱う方法(スカルプトモデリング)があるが [1]、これらの場合で人体をモデリングするのは、初心者には困難である。本研究では、部品の制作と選択、また、合成により人体頭部をモデリングする手法を提案する。

## 2 関連研究

部品の組み合わせによりモデリングする手法は以前から知られている [2]。また、人体頭部も目、鼻、耳等の部品を考えることができるため、目や眉毛などの各部位ごとに部品を選択し配置することによって頭部のモデルを制作できる [3]。しかし、ユーザーが必要とする部品が用意されているとは限らない。

## 3 目的

本研究は部品選択による手軽さを残しつつ、必要に応じて部品を自作できる人体頭部用のモデルを開発する。

## 4 提案手法

ユーザーは部品(目と耳、顔)を選択しある程度意図するモデルを制作した後、部品を加工しより意図に近いモデルを制作することができる。この時、部位ごとに分けてモデリングすることができるためモデル制作が容易になる。また、一から部品を作成した部品でも既存の部品と組み合わせることができる。

### 4.1 自由度の削減

一般に 2 つの部品を合成するときは一方(例えば頭部全体)に対し、もう一方(例えば耳)の位置と方向には 6 自由度がある。しかし、本システムでは各部品に接続面を定めることによって 3 自由度に下げ、初心者のモデリングを補助している。

### 4.2 部品の合成

選択した部品を一つの頭部モデルとして合成する。合成するときに各部品の平面が顔の表面形状に沿うように変形させる必要がある。この変形には Blender の シュリンクラップ [1] を用いる。

## 5 システム

### 5.1 選択モード

ユーザーは両目と両耳、顔の計 5 パーツをそれぞれ自由に選択でき、部品の移動や回転もここで行う。

### 5.2 加工モード

ユーザーが直感的に 3D モデルを操作できるように 3D モデルを粘土のように扱うことができるスカルプトモデリングを採用し、Blender のスカルプトモード [1] を利用する。

選択モードで選択したモデルを Blender の 3D 空間に配置し、システム側で頭部全体との接続部(例、右耳なら左側の平面)には加工できない制限をかけ、ユーザーは制限のない部分に対してスカルプトモードで自由にモデリングすることができる。

### 5.3 作成モード

ユーザーは  $xz$  平面上で輪郭を作成し、目や耳などのカテゴリを指定する。その時選択したカテゴリが右耳の場合、左端の平面が残っていない場合はエラーを返し、接続部が存在するようにする。その後システムが入力された輪郭を  $y$  軸方向にスイープ操作で伸ばし、ユーザーは加工モードで加工する。

## 6 実現

実現の結果、部品の選択後にモデルの加工ができ、簡単にモデルを制作しながらもより意図するモデルを制作できるモデルであることを確認した。また、部位ごとに作成し、その場合に輪郭から作成できるため一から制作する場合でのモデリングの簡易化を確認した。

## 7 むすび

本研究では、部品を選択後に加工機能を加えることによってより意図するモデルを制作することができるモデルを実現した。しかし、接続面を無視した部品の移動が可能でありその部分の実装を今後の課題とする。また、口や髪などより多くの部位を実装する必要がある。

## 参考文献

- [1] Blender, <https://www.blender.org/>
- [2] Sebe, et al., “Rapid Part-Based 3D Modeling,” *Proc. of VRST ’05*, 2005.
- [3] Pixiv, <https://vroid.com/>

		指導教員	乃万 司 教授
学生番号	192C1112	氏名	西田 広大
論文題目	歪んだ面に映り込んだ顔の認識と プライバシー保護への応用		

## 1 はじめに

Youtubeなどの動画配信サイトにおいて、自身の顔を隠して活動している配信者は少なくない。しかし、意図せずこうした配信者の顔が動画内に映り込んでしまうトラブルが稀に生じている。この問題への対策として、動画内の顔を検出し自動でぼかしをかけるシステム(以下、ぼかしシステムと表記する)の利用を考えられるが、ぼかしシステムでも顔を検出できない場合がある。例えば、スプーンなどの歪んだ鏡面に映り込んだ顔は、映り込んだ顔も歪んでしまうため検出できないことがある。本研究では、歪んだ鏡面に映り込んだ顔を検出するモデルを作成し、さらにこの顔検出モデルを用いたぼかしシステムの実装を目指す。

## 2 関連研究

Zhao ら [1] は、画像の歪みを補正することで魚眼レンズで撮影した顔の検出精度を改善する手法を提案した。魚眼レンズで撮影した顔は画像の中心から放射状に歪むことが分かっているため補正しやすいが、歪んだ鏡面に映り込んだ顔は鏡面の形状や配置によって歪み方が異なるため、歪みの補正は難しいと考えられる。

## 3 提案手法

歪んだ鏡面に映り込んだ顔(以下、歪んだ顔と表記する)は補正が難しい。そこで本研究では、機械学習を用いて作成するモデルで歪んだ顔を補正せずに認識する手法を提案する。

### 3.1 画像の作成

歪んだ顔を検出するモデルを作成するためには、歪んだ顔の画像が必要になるが、このような画像を大量に集めることは難しい。そこで、本研究では Blender を用いて 3DCG で歪んだ顔の画像データを作成する。なお、本研究で作成する顔検出モデルは動画へのぼかし処理に使用するため、ここで作成する画像のサイズは一般的な動画の解像度と同じ  $1920 \times 1080$  ピクセルとする。画像の作成方法を図 1 に示す。背景を切り抜いた人物の写真を鏡面の 3D モデルの前に配置し、背景にパノラマ写真を配置する。これをレンダリングすることで歪んだ顔の画像を作成する。自然な顔や背景を 3D モデルで再現しようとすると大きなコストがかかるが、顔画像と背景画像を用いることで容易に歪ん

だ顔の画像を作成することができる。作成した画像を図 2 に示す。

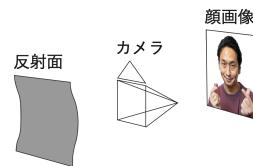


図 1: 画像の作成方法



図 2: 作成した画像

作成した画像内の歪んだ顔に対してアノテーションを行い、訓練データを作成する。

## 4 実装と実験

本研究では、YOLOv7 [2] の学習済みモデルにファインチューニングを行い顔検出モデルを作成する。有効性の検証のため、2種類の学習モデルを作成する。1つ目は、歪んだ顔の画像からなるデータセット(データセット 1)を用いて学習するモデル(モデル 1)である。2つ目は歪んでいない通常の顔画像からなるデータセット(データセット 2)を用いて学習するモデル(モデル 2)である。データセット 2 の作成でも Blender を使用し、顔画像が直接カメラから見えるように配置して作成する。データセット 1, 2 を用いて作成したモデル 1, 2 の検出精度を比較することで、提案手法の有効性を検証する。モデル 1, 2 の AP はそれぞれ 0.6837, 0.2022 となり、提案手法の有効性が確認できた。

## 5 むすび

本研究では、歪んだ面に映り込んだ顔を検出可能なぼかしシステムを実現した。このシステムにより、動画内に配信者の顔が映り込むことで生じるトラブルの減少が期待される。しかし、より実用性の高いぼかしシステムを実現するためには、本研究で扱わなかった種類の鏡面への映り込みにも対応する必要がある。

## 参考文献

- [1] He Zhao, et al. “RDCFce: Radial Distortion Correction for Face Recognition,” *CVPR 2020*, pp. 7721–7730, 2020.
- [2] Chien-Yao Wang, et al. “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” arXiv:2207.02696, 2022.

		指導教員	乃万 司 教授
学生番号	192C1165	氏名	山下 圭市
論文題目	採点機能付きカラオケで 満点を取るための歌声補正の研究		

## 1 はじめに

カラオケに精密採点などの採点ゲームが導入されるとから、カラオケは単なる娯楽に加えて点数を競い合う競技的な一面を持つようになった。今日では、カラオケの採点機能を用いて歌の上手さを競う THE カラオケ★バトルという TV 番組が成立していることからも、採点ゲームで満点を取りたい、あるいは目指したいという需要は存在する。

そこで、本研究では、人の歌声を満点を取れるように補正するシステムの構築を目指す。

## 2 手法

歌声の音声は主に、基本周波数、スペクトル包絡、非周期性指標の 3 つのパラメータで表される [1]。本研究では、音声解析合成ライブラリ PyWORLD を用いて楽譜に忠実に音声変換を行う。

### 2.1 パラメータの推定

PyWORLD では、入力した音声から基本周波数、スペクトル包絡、非周期性指標の 3 つのパラメータを推定し、独立に操作することで音声が変換される。今回は、基本周波数のパラメータのみを操作する。

### 2.2 基本周波数の変換

基本周波数は、周期的に繰り返される声帯振動のうち、最も短い間隔(基本周期)の逆数として定義され、人間が知覚する声の高さに概ね対応する。基本周波数は時々刻々と変化するため、音声を短時間ごとにフレームとして切り出し、各フレームについて基本周波数の推定を行う [2]。推定された基本周波数は 1 次元配列のデータとなる。この配列に対して、曲の BPM から各音符の実時間を計算により割り出し、それぞれの実時間に対応する配列の長さを特定する。例えば、4 分音符に対応する配列の長さが 4 とすると、2 分音符は 8、8 分音符は 2 となる。そこから、楽譜通りに各音符の音階に対応する周波数を代入していき、変換した基本周波数と、スペクトル包絡、非周期性指標から音声を合成することで楽譜通りのピッチ補正を行うプログラムが得られる。変換の例を図 1 に示す。

## 3 実験

実験では、歌いやすさや楽譜入手の容易さを考慮して SMAP の「世界に一つだけの花」を対象とする。

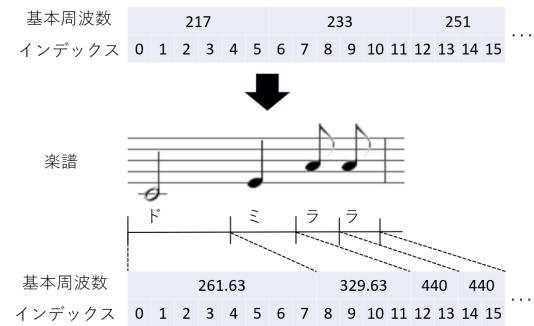


図 1: 基本周波数配列変換のイメージ

### 3.1 音声の前処理

録音した音声を変換する前に、歌い始めと歌い終わりの無声部分をカットする。今回は音声編集ソフトの Audacity を用いて無声部分を除去した。

### 3.2 実験方法および結果

ノート PC とスピーカーをカラオケボックスに持ち込み、DAM の精密採点 Ai という採点機能で実験を行った。本システムを用いて変換を行った音声をカラオケの歌い始めに合わせてノート PC で再生し、カラオケ店に常備されているマイクに向けて、スピーカーから音声を流した。マイクに認識されやすいよう、大音量で、マイクとスピーカーの距離は限りなく小さくなるようにした。

結果として、変換前の音声のスコア 74.207 点に対し、変換後は 92.588 点のスコアが得られた。

## 4 むすび

本研究では、楽譜に忠実にピッチ補正を行うというアプローチで満点を目指した。今回は音程のみの補正に終始したが、カラオケでは表現力やビブラートなどの採点項目もある。今後は歌声にしゃくりや、こぶし、ビブラートなどを加えられるよう改善していくたい。

## 参考文献

- [1] 森勢, “音声パラメータのデザイン,” 日本音響学会誌, 74 卷, 11 号, 2018
- [2] Morise, et al., “WORLD: A Vocoder-Based High-Quality Speech Synthesis System for Real-Time Applications,” IEICE Trans. Inf. & Syst., Vol.E99-C, No.7, 2016

		指導教員	乃万司 教授
学生番号	192C2059	氏名	橘慎二
論文題目	手書きマップを用いた絵双六対戦システム		

## 1 はじめに

将棋や双六といった古くから親しまれてきたアナログゲームには、デジタル技術が進んだ現代においても特有の利点が存在する。その内の1つがオリジナルゲーム制作の容易さである。計算機の知識が必要なデジタルゲームに比べて、アナログゲームの場合は最低限紙とペンさえあれば独自のデザインでゲーム制作に取り掛かることができる。一方、デジタルゲームであれば、ゲーム展開の自動処理や遠隔対戦などが可能になる。

本研究では、両者の特徴を生かし、紙上あるいはタブレット上に手書きされた絵双六を認識し遊ぶことができるシステムの構築を目指す。

## 2 関連研究

栗田ら [1] は、画像情報を用いた将棋における精度の高い局面の自動認識を実現した。また、コンピュータにおける紙などへの手書き入力の扱いは古くから研究されており [2]、子供に手書きでゲームをさせる研究なども報告されているが [3]、手書きのマップでゲームをするシステムは知られていない。

## 3 研究目的

本研究では、将棋のようにルールに基づいて定められた盤面を認識するのではなく、利用者によって柾目の形状や配置が異なる手書きの双六マップを認識し、そのまま遊ぶことができるシステム構築を目指す。

## 4 システムの構成

本システムでは、入力された双六マップ画像から柾目の座標や柾目同士の接続、柾目に書かれた指示といった情報を読み取ることで、双六マップをデジタル化して遊ぶことができる。

### 4.1 柾目の配置と接続の認識

入力されたマップ画像に対して輪郭検出を行い、輪郭の内側と外側に分ける。その後、内側の輪郭を柾目と判断し、外側の輪郭と内側の輪郭で重なっていない部分を柾目と柾目との接続部分と判断する。

### 4.2 柾目に書かれている指示の認識

光学式文字認識エンジンである Tesseract を使い、柾目に書かれている「スタート」、「ゴール」、「1マス進む」といった指示を認識する。認識されるべき文字列と実際に認識された文字列が異なる場合にも、レーベ

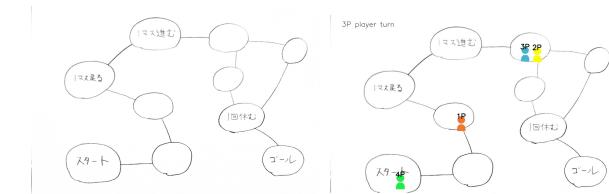


図 1: 入力画像



図 2: 出力結果

ンシュタイン距離を用いて 2 つの文字列の類似度を求めて認識された文字列を補正する。

### 4.3 双六対戦機能

双六マップから取得された情報を元に、手書きマップを用いた双六対戦を行う。サイコロを振った時の移動候補となる柾目の表示や、「1回休み」等の柾目に書かれた指示をコンピュータが自動で処理する。

## 5 実験

図 1 の入力画像を使って本システム上で処理を行った結果、柾目の座標と駒の初期位置、柾目に書かれた指示が正しく認識され、図 2 のように双六の駒を動かすことができた。

## 6 むすび

本研究では、手書きマップを認識し絵双六の対戦ができるシステムを実現した。現在のシステムでは、入力されるマップにおいて、隣り合う柾目は 1 本の線で繋がれていることを前提としたシステムになっているが、例えば柾目同士が直接隣合っているマップに対しても認識できるようにするなど、様々な構成のマップに対応できるように改善する余地がある。また、遠隔地間の対戦も実現したい。

## 参考文献

- [1] 栗田ら, “将棋を対象とした画像情報を用いた自動局面認識手法,” ゲームプログラミングワークショップ 2007 論文集, 2007.
- [2] Wellner, “Interacting with Paper on the DigitalDesk,” CACM, Vol. 36, No. 7, 1993.
- [3] Paulson, et al., “Sketch-Based Educational Games,” Proc. of IDC 2008, 2008.

		指導教員	乃美 正哉 講師
学生番号	192C1104	氏名	中野 智仁
論文題目	回帰分析によるヘッドホンの価格に関する統計調査		

## 1 はじめに

今日、市場に多く流通しているヘッドホンだが、商品のステータスは意外に多くの項目があり、その価格も性能によって高いものから安いものまで様々である。

本論文ではそんな多種多様なヘッドホンを商品表示から商品価格との相関を統計的な分析によって見出し、どのようなヘッドホンが一般的に市場で良いヘッドホンとされているのかを価格の面から考察していく。

## 2 回帰分析

本論文では、基準となる目的変数と予測を利用する説明変数との変数間の関係を数式として求める回帰分析という手法を用いる。

単回帰分析は1つの目的変数を1つの説明変数から予測する。目的変数を $y$ 、説明変数を $x$ として、以下のモデルを想定することになる。

$$y = b_0 + b_1x + e \quad (1)$$

一方、重回帰分析では1つの目的変数を複数の説明変数から予測する。目的変数を $Y$ 、 $p$ 個の説明変数を $x_1, x_2, \dots, x_p$ として、以下のモデルを想定することになる。

$$Y = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p + E \quad (2)$$

ここで $b_0$ は切片、 $b_1, b_2, \dots, b_p$ はそれぞれの変数の回帰係数、 $e, E$ は誤差を表す変数で平均0、分散 $\sigma^2$ の正規分布に従うことが仮定されている。

## 3 実験

目的変数をヘッドホンの本体価格とし、説明変数としてヘッドホンの性能を表す数値である、インピーダンス、音圧感度、最低周波数、最高周波数、ドライバ口径、コード長、最大入力電力、質量の計8個の変数データを取得して重回帰分析を行った。また、その中から回帰に有効な説明変数を選択することで、より精度の高い回帰式を導いた。

数値実験の結果、選択された説明変数は最大周波数、ドライバ口径、質量の3つとなった。最大周波数を $x_1$ 、ドライバ口径を $x_2$ 、質量を $x_3$ として以下にその回帰式を示す。また、予測の精度の指標となる自由度調整済み決定係数は0.4726となった。

### • 回帰式

$$Y = -73.29 + 0.001054x_1 - 6.619x_2 \\ -0.2313x_3 + E \quad (3)$$

次に独自で算出した、再生周波数の広さ、ドライバの質、音量の3個の変数を加えた計11個の説明変数を用いて再び重回帰分析、適切な説明変数の選択を行った。

数値実験の結果、選択された説明変数は最大周波数、ドライバ口径、質量、ドライバの質の4つとなった。最大周波数を $x_1$ 、ドライバ口径を $x_2$ 、質量を $x_3$ 、ドライバの質を $x_4$ として以下にその回帰式を示す。また、自由度調整済み決定係数は0.5224となった。

### • 回帰式

$$Y = 93.14 + 0.0005526x_1 - 3.291x_2 \\ -0.5420x_3 + 0.01749x_4 + E \quad (4)$$

## 4 おわりに

説明変数を8個使った際の分析の結果より、ドライバ口径と質量が目的変数と強い相関を持っていたという点からヘッドホンの価格帯はそのヘッドホンのドライバの規模によって大まかに決定されていると考えることができる。実際に、説明変数を追加して分析を行った際は変数選択により先ほどの3つの説明変数に加えてドライバの質が選択されていることから、ドライバの造りが価格に大きく影響しているのだと考えられる。また、最高周波数も価格との関係が強いため、市場ではドライバの規模が大きく、高周波数を再生できるのヘッドホンがより高級とされており、一般的にそれらの性能を向上させることに開発コストを掛けているのだと考えられる。

## 参考文献

- [1] 豊田秀樹、回帰分析入門 - R で学ぶ最新データ解析 -、東京図書株式会社、2012.
- [2] 価格.com、<https://kakaku.com/kaden/overhead/itemlist.aspx?pdf>
- [3] オーディオテクニカ、ヘッドホン・イヤホンを識る、<https://www.audio-technica.co.jp/headphone/navi/whatis/index.php>

		指導教員	乃美 正哉 講師
学生番号	192C1177	氏名	吉村 唯吹
論文題目	合計特殊出生率の地域差に関する重回帰分析		

## 1 はじめに

本研究では、各都道府県の合計特殊出生率について重回帰分析を用いて検討する。合計特殊出生率とは 15 歳から 49 歳までの女性の年齢別出生率を合計したものである。表 1 は各都道府県の合計特殊出生率をまとめたものであり、これを確認すると各都道府県で差があることが分かる。合計特殊出生率の地域差について、どのような社会的背景が影響を与えているかを分析する。

表 1: 各都道府県の合計特殊出生率

都道府県	合計特殊出生率	都道府県	合計特殊出生率
北海道	1.19	滋賀	1.44
青森	1.26	京都	1.20
岩手	1.37	大阪	1.28
宮城	1.25	兵庫	1.33
秋田	1.29	奈良	1.23
山形	1.39	和歌山	1.36
福島	1.49	鳥取	1.46
茨城	1.37	島根	1.55
栃木	1.43	岡山	1.39
群馬	1.38	広島	1.47
埼玉	1.28	山口	1.43
千葉	1.31	徳島	1.35
東京	1.12	香川	1.48
神奈川	1.28	愛媛	1.41
新潟	1.37	高知	1.29
富山	1.37	福岡	1.37
石川	1.40	佐賀	1.49
福井	1.55	長崎	1.50
山梨	1.31	熊本	1.58
長野	1.43	大分	1.50
岐阜	1.37	宮崎	1.61
静岡	1.43	鹿児島	1.56
愛知	1.43	沖縄	1.79
三重	1.40		

## 2 原理

本研究では、1 つの目的変数を複数の説明変数で予測する重回帰モデルを用いる。重回帰モデルは以下の式で表すことができる。

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + E_i$$

$\beta_0$  は母切片、 $\beta_1, \beta_2, \dots, \beta_p$  は母回帰係数を表している。また、誤差  $E_i$  は平均 0、分散  $\sigma^2$  の正規分布に互いに独立に従う ( $E_i \sim N(0, \sigma^2)$ ) ことが仮定される。

重回帰分析の後に変数選択を行う。その際の指標として AIC を用いる。回帰分析の場合には以下の式で求められる。

$$AIC = n \log\left(\frac{SS_e}{n}\right) + 2(q+1)$$

## 3 実験

各都道府県の「合計特殊出生率」を目的変数とし、「夫の初婚年齢」、「妻の初婚年齢」、「離婚率」、「婚姻率」、「他都道府県からの転入率」、「他都道府県への転出率」、「女性の就業率」、「平均賃金」、「平均家賃」、「10 大費目別消費物価地域差指数」の 10 個の説明変数で重回帰分析を行った。

しかし、これらの変数での重回帰分析では多重共線性が発生していると判断できたので、問題となった説明変数「他都道府県からの転入率」に変数変換を施した。「他都道府県からの転入率」を「他都道府県への転出率」で割った「転入率/転出率」を新たな説明変数とし、改めて 9 個の説明変数で重回帰分析を行った。また、AIC を基準に変数選択を行い、有意性があると考えられる変数を求めた。

実験の結果、「妻の初婚年齢」、「婚姻率」、「女性の就業率」、「平均賃金」の 4 つの説明変数が最も適切なモデルとなった。重回帰分析の結果を表 2、変数選択での最も適切なモデルの結果を表 3 に示す。

表 2: 重回帰分析の結果

決定係数	0.7084
自由度調整済決定係数	0.6375

表 3: 変数選択の結果

決定係数	0.6881
自由度調整済決定係数	0.6584

## 4 おわりに

本研究では、「合計特殊出生率」の地域差に関する重回帰分析を行った。重回帰分析と変数選択の結果、「妻の初婚年齢」、「婚姻率」、「女性の就業率」、「平均賃金」の 4 つの説明変数が「合計特殊出生率」に大きく影響を与えると考えられる。

課題として、複数年のデータを用いて重回帰分析を行い、長期的な要因を確認することが挙げられる。

## 参考文献

- [1] 豊田秀樹, 回帰分析入門—R で学ぶ最新データ解析, 2012.

		指導教員	乃美 正哉 教授
学生番号	21221009	氏名	吉田 大輝
論文題目	回帰分析を用いた野球における四球獲得能力の有用性と向上方法の検討		

## 1 はじめに

近年、MLB(メジャーリーグベースボール)では、選手のプレーの数値化と、データを用いた統計的分析が盛んに行われている。この流れの中で打率より出塁率のほうが得点相関が高いことがわかり、打率と出塁率の価値の逆転現象が発生している[1]。出塁率は、打者のヒットを打つ能力と四球を選ぶ能力を合わせた指標であり、出塁率の価値が高まることにより打者の四球を選ぶ能力にも注目する必要があると考える。本研究では、この打者の四球を選ぶ能力を四球獲得能力と呼称し、この四球獲得能力が持つ性質や他の打撃指標との関係を回帰分析を用いて分析する。

## 2 重回帰分析

まず、1つの目的変数を1つの説明変数で予測する単回帰モデルを以下に示す。

$$y_i = \beta_0 + \beta_1 x_i + e_i \quad (1)$$

これに対し、重回帰分析では、1つの目的変数を複数の説明変数で予測する重回帰モデルを用いる。重回帰モデルは以下のような式で表すことができる。

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + E_i \quad (2)$$

ここでの母切片は  $\beta_0$ 、母回帰係数は  $\beta_1, \beta_2, \dots, \beta_p$  で表している。また、誤差  $E_i$  は平均 0 分散  $\sigma^2$  の正規分布に互いに独立に従う ( $E_i \sim N(0, \sigma^2)$ ) ことが仮定される。

## 3 実験

本実験では、IsoD(=出塁率-打率)を四球獲得能力の指標として用いる。まず IsoD の有用性の検証のため、打撃総合指標である wRC+を目的変数、IsoD を説明変数として単回帰分析を行う。結果を表 1 に示す。

表 1: 単回帰分析の結果

切片	回帰係数	相関係数	決定係数
79.040	415.915	0.4317	0.1864

表 (1) より、IsoD と wRC+の間には弱めの正相関がみられることがわかる。

次に、IsoD を目的変数とした重回帰分析を行い、IsoD と他打撃指標の関係を検証する。ここで、似た打撃能

力を表す打撃指標を変数群としてまとめ、階層的重回帰分析を行う。変数群は「打者積極性」、「ミート力」、「長打力」、「打球方向」、「選球眼」の5つである。この中で、「打者積極性」は IsoD との関係が深いことが推測できるため、この「打者積極性」に他変数群を足して重回帰分析を行った際の自由度調整済み決定係数の変化に着目する。これらの重回帰分析による自由度調整済み決定係数をまとめたものを図 1 に示す。

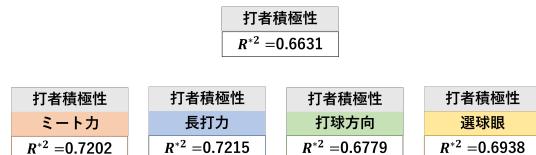


図 1: 階層的重回帰分析の結果

図 1 より、「ミート力」、「長打力」を加えることで自由度調整済み決定係数を大きく向上でき、よってミート力と長打力が IsoD と深い関係を持つことがわかる。

## 4 おわりに

本研究では、公開されている MLB の詳細な打者成績に対して、四球獲得能力に焦点を当てて回帰分析を適用することで、四球獲得能力の持つ性質や他打撃成績との関係を検証した。結果として、四球獲得能力は打者能力として有用であり、打者の積極性、ミート力、長打力が四球獲得能力と関係を持つ事を確認できた。この中で、ミート力と長打力による四球獲得能力の向上は、投手が勝負を避けたくなり、被投球のうちのボール球が増加するからだと説明できる。また、これらの能力の改善によって四球獲得能力を向上できると推測される。

今後の課題として、個人成績の移り変わりを確認し、実際に四球獲得能力が向上した選手データを集めて検証することや、今回行ったような回帰分析を効率化するアプリケーションを開発することがあげられる。

## 参考文献

- [1] Michael Monroe Lewis, Moneyball: The Art of Winning An Unfair Game W. W. Norton & Company, 2003
- [2] 豊田 秀樹, 回帰分析入門 – R で学ぶ最新データ解析 –, 東京図書株式会社, 2012.

		指導教員	二反田 篤史 准教授
学生番号	192C1013	氏名	石山 陽己
論文題目	ニューラルネットワークの表現学習体制の有効性検証		

## 1 はじめに

近年ニューラルネットワークに代表される深層学習は自然言語処理や画像処理など幅広い分野において優れた成果を残してきた。深層学習が従来技術と比べて特に優れていた点はデータの特徴表現そのものを獲得する表現学習を行うことができるという点である。ニューラルネットワークの性能は初期化スケールや学習率などの学習設定に大きく依存しており、パラメータ設定により表現学習の可否も決定される。本研究では表現学習に差がある NTK 領域と平均場領域のニューラルネットワークを比較し、また表現学習領域のニューラルネットワークと既存手法を比較する。

## 2 関連研究

ニューラルネットワークの最適化が保証される設定として NTK と平均場理論という二つの手法がある。これら二つの手法はパラメータのスケーリングが異なっており、その結果 NTK 領域では表現学習を行うことができず、平均場領域では表現学習を行うことができる。表現学習によって平均場領域ではより広い真のモデルを表現することが可能であり、その結果 NTK 領域よりも優れた性能を示すことが知られている。

## 3 表現学習

表現学習では学習段階においてデータに適忯的にカーネル関数を学習することができるため高い表現力を有している。本研究ではニューラルネットワークの設定について、表現学習可能な二つのパラメタリゼーションを使用する。一つは平均場理論であり、隠れ層を無限幅に近づけることでパラメータの分布を最適化する手法である。二つ目は最大更新パラメタリゼーションである。これは Pytorch のデフォルト設定では出力を抑えるために表現学習を行うことができないことを踏まえ、Pytorch のデフォルト設定において出力をスケルダウンさせ表現学習を可能にした手法である。

## 4 実験

### 4.1 平均場領域と NTK 領域の比較

それぞれの領域における二層ニューラルネットワークの人工データセット staircase における損失値を図 1 に示す。二つの領域には表現学習の可否に差があり、そ

の結果平均場領域の方が小さい損失値に収束している。

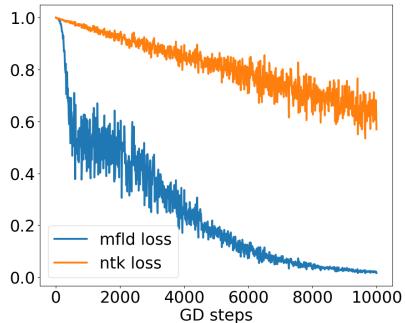


図 1: 損失値の比較

### 4.2 表現学習領域と既存手法の比較

表現学習領域における二、三層ニューラルネットワークと固定カーネル学習領域の二、三層ニューラルネットワーク、更に既存手法 (SVM, 勾配ブースティング法, ランダムフォレスト) の実データセットにおける分類精度の比較を表 1 に示す。この結果から、実データにおいても平均場領域は NTK 領域よりも高い分類精度を示している。また、全データセットで平均すると、表現学習領域は既存手法と比べて優れた分類精度を残していることがわかる。

表 1: 多クラス分類問題のテスト精度

手法	LETTER	MNIST	COTVTYPE
平均場二層 NN	0.921	0.985	0.914
NTK 二層 NN	0.676	0.973	0.898
$\mu P$ 三層 NN	0.975	0.986	0.957
NTK 三層 NN	0.751	0.978	0.903
SVM	0.851	0.974	0.822
ランダムフォレスト	0.965	0.972	0.950
勾配ブースティング	0.987	0.987	0.847

## 5 議論

本研究の実験において、表現学習によって平均場領域と NTK 領域には学習の様子や学習結果に差があることを示した。また、表現学習領域が既存手法と比べて遜色ない精度を残すことを確認した。

		指導教員	二反田 篤史 準教授
学生番号	192C1033	氏名	太田 潤
論文題目	一階微分情報に基づくデータ拡張最適化		

## 1 はじめに

教師あり学習においてデータ拡張は仮説関数の予測精度を向上させる。データ拡張最適化はその方策を最適化し、更なる予測精度の向上を実現する。データ拡張最適化手法の中でも、データセットサイズと、方策の良し悪しに起因する仮説関数の予測精度の高低に関して、旧来の手法よりも優れたスケーラビリティを持つ手法 [1] がある。本論文では、既存手法 [1] から得られる方策と同等に優れる方策を、既存手法 [1] よりも小さい時間計算量で出力できる手法を提案する。方策の優劣は、その方策に基づくデータ拡張を使った訓練で得られた仮説関数の予測精度の高低で評価する。

## 2 問題設定

### 2.1 教師あり学習

教師あり学習の目標は、全ての入力データに対して正しい予測データを出力する仮説関数の構築である。無限個のデータを使えばそれを最も厳密に行えるが、不可能である。実際は有限個のデータを使ってモデルを訓練し、最適な仮説関数を求める。

### 2.2 データ拡張

モデルの訓練に有限個のデータしか使えない実状を受け、データ拡張はデータを水増しする。データ拡張では、訓練で使うデータを確率的に変換し、実質的なデータ量を増やす。データ拡張の方策、つまり、「各変換の選択確率、実施確率、強度」は人間が経験的に定める。

### 2.3 データ拡張最適化

データ拡張最適化は、データ拡張の方策を数理的に最適化する。方策が異なると、拡張後データも異なり、訓練で得られる仮説関数の予測精度も上下する。数理的な探索を通して、人間の経験に基づく場合よりも効果的な方策を網羅的に探索し、発見できる。

## 3 先行研究

データ拡張最適化をハイパーパラメータ最適化として定式化し、勾配ベースの最適化手法で方策を最適化する手法 [1] は、分布の密度マッチングを行う手法や探索空間を縮小する手法に比べて、より優れたデータセットサイズ対予測精度のスケーラビリティを持つことが実験で示されている。

## 4 提案手法

提案手法は、データセットサイズ対予測精度のスケーラビリティに優れる既存手法 [1] と同様に、データ拡張最適化をハイパーパラメータ最適化として定式化し、勾配ベースの最適化手法で方策を最適化する。提案手法では、既存手法 [1] より小さい時間計算量で勾配の近似計算を行う。一方で、本研究の実験において、提案手法は既存手法 [1] による方策と同等に優れる方策を出力できることが示されている。

## 5 実験結果

表 1 では、既存手法 [1] による方策と提案手法による方策のそれぞれを使ったデータ拡張で訓練された仮説関数の予測精度を比較する。ベースラインは、方策を最適化しないデータ拡張のみを実施する場合を示す。WRN は WideResNet の略記である。表 1 より、時間計算量がより小さい提案手法は、既存手法 [1] と同等に優れる方策を出力していることが確認できる。

表 1: CIFAR10 の予測精度 [%]

モデル	ベースライン	既存手法 [1]	提案手法
WRN28-10	95.81	96.16	<b>96.18</b>
WRN40-2	93.87	94.36	<b>94.41</b>
WRN28-2	93.62	<b>94.13</b>	94.06

## 6 議論

提案手法は、より小さい時間計算量で、既存手法 [1] と同等に優れる方策を出力できることが示された。今後は、提案手法が既存手法 [1] と同様に、分布の密度マッチングを行う手法や探索空間を縮小する手法に比べて、より優れたデータセットサイズ対予測精度のスケーラビリティを持つか否かを検証することが課題になる。

## 参考文献

- [1] Ryuichiro Hataya, Jan Zdenek, Kazuki Yoshizoe, and Hideki Nakayama. Meta approach to data augmentation optimization. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 2574–2583, 2022.

		指導教員	二反田 篤史 準教授
学生番号	192C1069	氏名	齊藤 悠悟
論文題目	平均場ランジュバンダイナミクスを用いた画像生成		

## 1 はじめに

近年、画像生成の技術は機械学習の学習用データを集めめる際や新たなデザインを生み出すための手助けとして応用されている。これらには高品質な画像が必要である。より品質の高い画像を生成するために、GAN をはじめ様々なフレームワークが提案してきた。その中でも特に、画像を作るための生成器を使わない手法に着目する。生成器を使わない画像生成に Trust the Critic (TTC) という手法がある。これは生成器を使わず、高品質な画像の生成に成功しているが、メモリ量が大きくなることが懸念されている。そこで、平均場ランジュバンダイナミクスという手法を適用し、メモリ量を効率的に使った画像生成を提案する。

## 2 問題設定

### 2.1 Trust the Critic (TTC)

TTC は生成器を取り除いた画像生成ができる手法である。生成器を取り除く代わりに、適応的な学習率と学習済みのクリティック（識別器）に対して勾配降下法を用いてデータを繰り返し更新し、目標の画像に近づけていく。また、クリティックの勾配は目標の画像に近づくための方向を正確に示していることがわかっている。学習後は複数のクリティックの勾配を層のように重ね、1つの大きな生成器と見なすことで画像を生成している。

### 2.2 平均場ランジュバンダイナミクス (MFLD)

MFLD はある確率分布  $q$  を最適化するための手法である。この手法により、分布のデータを直接動かすことが可能である。さらに、更新する元の分布に対して裏にギブス分布と呼ばれる分布が生成されることがわかっている。また、元の分布とギブス分布の2つの分布は、学習を進めていくと最終的に近づけたい目標分布に収束することが知られている。

## 3 提案手法

TTC をベースとし、平均場ランジュバンダイナミクスを適用させた手法を提案する。本手法を用いることで、分布を直接移動させることができ、TTC で問題となっていたメモリ量をより効率的に使い画像を生成することができる。さらに、裏で目標分布に近似していたギブス分布からサンプリングを行うことでいつでも

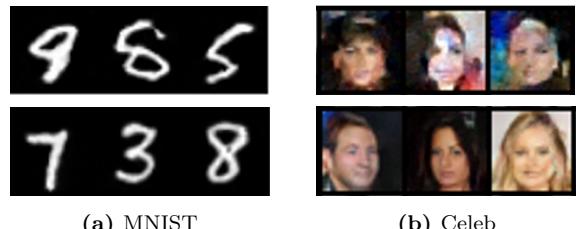
画像生成を行うことができる。

MFLD を使う手法では、クリティックの1回の学習に対して元の分布の更新も1回であるが、ギブス分布からのサンプリングがうまくいかない課題が挙がった。そこで、Entropic Fictitious Play (EFP) という手法を利用し、この課題を解決する手法を提案する。本手法では、クリティックの1回の学習に対して、分布を複数に分けてかつ複数回更新することを行なっている。

## 4 実験

### 4.1 目標分布に近づいた分布の画像

目標分布に近似していることを確認するために、最初に我々で用意する分布（ガウス分布）の最終的な形を画像として出力した。

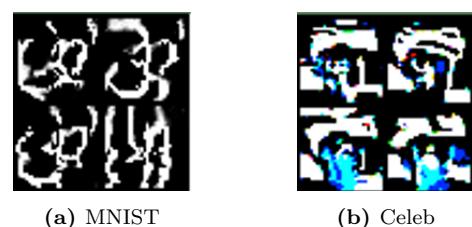


(a) MNIST (b) Celeb

図 1: 上：MFLD 下：MFLD + EFP

### 4.2 ギブス分布からのサンプリング

ギブス分布からのサンプリングを行った画像を示す。この実験は、MFLD + EFP の設定である。



(a) MNIST (b) Celeb

図 2: ギブス分布からのサンプリング画像

## 5 考察

図 1 よりノイズのガウス分布は目標分布に近似していることがわかり、MFLD と EFP の有用性を確認できた。特に、EFP を適用した実験は高品質な画像が出力できた。しかし図 2,3 よりギブス分布からのサンプリングは課題が残る形になった。

		指導教員	二反田 篤史 准教授
学生番号	192C1171	氏名	横尾 早香
論文題目	平均場ランジュバンダイナミクスを利用したアルゴリズミックアート		

## 1 はじめに

近年、深層学習は様々な分野で活用され成功を収め、特に AI を用いた画像描画には多くの注目が集まっている。既存研究により、コンピュータ・アートはアートスタイルが最適化アルゴリズムに強く依存することが知られている。ニューラルネットワークの最適化アルゴリズムとして、Mean-Field Langevin Dynamics (MFLD) がある。本研究では、MFLD を用いて三角形で画像描画する実験を行った。

## 2 関連研究

### 2.1 進化戦略を用いた画像描画

この既存手法では、進化戦略 (ES) を用いて、透明度のある三角形を配置することで画像生成を行なっている。遺伝的アルゴリズムで画像が反復的に進化していくため、アートの創造プロセスが得られることが特徴。

### 2.2 Mean-Field Langevin Dynamics

ニューラルネットワークの最適化をパラメータの分布最適化としてみなし最適化アルゴリズムとして、Mean-Field Langevin Dynamics (MFLD) がある。ギブス分布  $q_* \propto \exp(-f(\theta))$  からのサンプリング時に、一つのパラメータ  $\theta$  での位置が決まる従来の手法 Langevin Dynamics と比較して、移動位置が全体の分布に依存することが特徴としてある。MFLD では、式 (1) の最小化を考える。

$$\min_{q \in \mathcal{P}} \{L(q) = F(q) + \lambda \mathbb{E}_q[\log q(\theta)]\}. \quad (1)$$

ステップサイズ  $\eta$ 、ガウスノイズを  $\xi^{(k)}$  としたとき、パラメータ  $\theta$  の更新式を式 (2) に示す。

$$\theta^{(k+1)} = \theta^{(k)} - \eta \nabla \frac{\delta F}{\delta q} \left( q^{(k)} \right) \left( \theta^{(k)} \right) + \sqrt{2\lambda\eta} \xi^{(k)}. \quad (2)$$

## 3 MFLD を用いた画像描画

一つの三角形は 3 つの頂点座標  $(x_1, y_1, x_2, y_2, x_3, y_3)$  と、色  $(r, g, b)$ 、透明度  $(a)$  で表せるので、このパラメータを  $\theta$  とし、三角形の個数を  $M$ 、元のターゲット画像を  $\text{target}$ 、ピクセル数を  $W \cdot H$ 、 $\lambda_1$  を正則化係数、三角形の画像を  $h(\theta) \in \mathbb{R}^{W \times H}$  したとき、MFLD を用いた三角形での画像描画における目的関数を以下のように設定する。

$$F(q) = \frac{1}{W} \|\text{target} - \mathbb{E}_q[h(\theta)]\|^2 + \lambda_1 \|\theta\|^2. \quad (3)$$

式 3 に関して、MFLD を用いて最小化を考えることで画像を近似させる。

## 4 実験結果

### 4.1 MFLD を用いた画像近似

三角形の個数を 1000 個、学習を 10000 反復させ得られた生成画像とその生成過程を図 1 に示す。細かいテクスチャーと背景を三角形で表現してターゲット画像に近似できていることが視認できる。三角形の個数、反復回数を増やしたことでも既存手法と同等の結果が得られた。

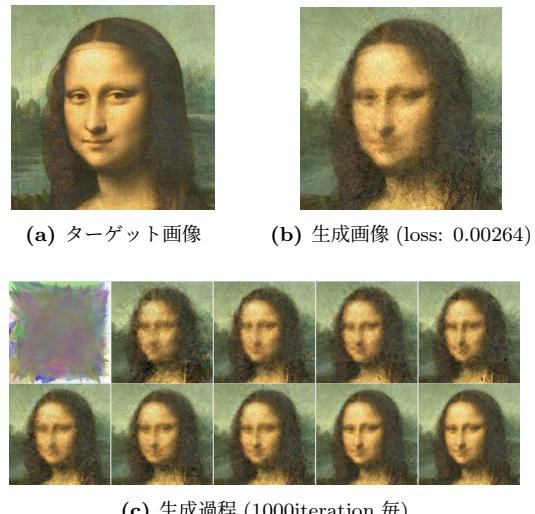


図 1: 三角形での画像近似

## 5 考察

MFLD 手法での三角形を用いた画像近似において、既存手法には劣るもの非常に高い精度で画像が近似できており、その精度は三角形の個数を増やす毎に増していった。これは、Mean-Field においては  $M \rightarrow \infty$  を想定しているため、MFLD の大域的収束性によって  $M$  の値を大きくすればするほど、反復回数を大きくすればするほど精度が高まったと考えられる。

		指導教員	二反田 篤史 準教授
学生番号	192C1175	氏名	吉原 啓人
論文題目	連合学習における平坦性を指向する最適化手法		

## 1 はじめに

近年、機械学習は様々な分野で研究・活用されており、その中でも特に深層学習は盛んに研究が進められている。この深層学習の発展に大きく貢献しているのがニューラルネットワーク（NN）と呼ばれる数理モデルである。NN の学習において明示的な正則化がなくても、暗に過学習が抑制される正則化が存在し、このような正則化を暗黙的なバイアス（Implicit bias）と呼ぶ。この暗黙的なバイアスは、最適化アルゴリズムの性質によるものだと考えられており、ここ数年の間で、特に確率的勾配降下法（Stochastic Gradient Descent）による暗黙的なバイアスの研究が盛んに行われている。この暗黙的なバイアスが確率的勾配由来のノイズによる平滑化効果に基づいていることから損失関数の形状について解が平坦な最適解である時、経験的に汎化性能が良くなると考えられている。いくつかの研究で、暗黙的なバイアスによって平坦性を指向するアルゴリズムが開発された。

## 2 既存研究

学習率を大きくした平均化確率的勾配降下法（Averaged Stochastic Gradient Descent）が、一般的な確率的勾配降下法よりも強く暗黙的なバイアスの効果を引き出し、より安定した平坦な最小値に収束することが示された。連合学習（Federated Learning）は、ローカルデバイスに学習データを分散させ、それぞれで最適化されたパラメータを集約して共有モデルを学習する手法である。連合学習の最適化アルゴリズムの一つに、Federated Averaging（FedAvg）がある。FedAvg は、それぞれのローカルデバイスで確率的勾配降下法を用いてパラメータ更新を行い、更新されたパラメータの平均を中心サーバーのモデルとするアルゴリズムである。

## 3 研究内容

この FedAvg と前述した平均化確率的勾配降下法の学習スキームが似通っていることに着目すると、学習率を大きくとることで、FedAvg が平坦性を指向するアルゴリズムになり得るのではないかという問い合わせが生まれる。本研究では、IID データに対して学習を行い、学習率に注目して FedAvg の汎化性能を確認した。

## 4 数値実験

### 4.1 実験 1

FedAvg が学習率を調整することで平坦な解に収束することを、簡単なモデルで確認した。学習率が小さいと鋭い解に収束するが、学習率が大きく取ることで平坦な解に収束している。

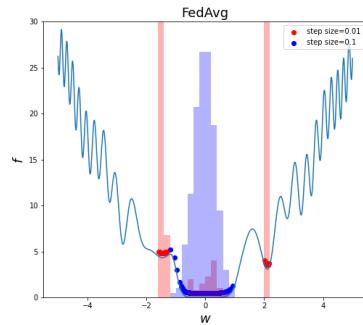


図 1: 異なる学習率を用いた FedAvg の比較。赤と青の点は、それぞれ  $\eta = 0.01$  と  $\eta = 0.1$  の収束した点を表しており、背景にそれぞれの色に対応したヒストグラムを描いている。

### 4.2 実験 2

ResNet18 を用いて CIFAR10 を学習した。初期の学習率を適切に大きく取りつつ学習率減衰を用いることで、汎化性能が高いパラメータに収束することを確認した。

学習率	0.1 → 0.02	0.04 → 0.008	0.01 → 0.002
TestAcc	88.02	86.99	82.55

表 1: CIFAR10 におけるテスト精度 [%]

## 5 議論

本研究では連合学習の代表的アルゴリズムである FedAvg が、IID データを扱う設定において平坦性を指向する最適化アルゴリズムであることを示した。実験を通して、学習率  $\eta$  の調整は汎化性能に大きく関わることが観察された。学習率を比較的大くとることで、確率的勾配降下法のバイアスを強めつつパラメータ平均によって収束を担保することにつながることが確認された。

		指導教員	二反田 篤史 準教授
学生番号	202C1059	氏名	坂口 瑞季
論文題目	平滑化効果を担保した確率的最適化法		

## 1 はじめに

現在、第3次AIブームにより機械学習の研究が隆盛を極めている。第3次AIブームは2006年にGeoffrey E. Hintonらによって機械学習手法のひとつである深層学習(Deep Learning)が発明されたことにより始まった。深層学習はニューラルネットワーク(Neural Network: NN)と呼ばれるモデルを用いて実装されており、ニューラルネットワークは膨大な量のパラメータを用いて表現される。この膨大なパラメータを損失関数の値を小さくするように最適化手法を用いて調整することが学習である。すべての最適化手法は損失関数の値が小さくなることを目指してパラメータを更新するが、最適化手法毎に汎化性能に差がある。これは、最適化手法毎に帰納的バイアス(Implicit Bias)が存在しており、学習の収束先が特徴付けられているからであると考えられている。

## 2 関連研究

最近になって、最小値の近傍の損失関数の形状が平坦である方が汎化性能が高いと注目されている。この平坦な領域を指向する最適化手法としてSharpness-Aware Minimization(SAM)などが提案されたが、ベーシックな最適化手法である確率的勾配降下法も平坦な領域を指向することが発見された。確率的勾配降下法には確率的勾配ノイズがあり、この確率的勾配ノイズが損失関数を平滑化することにより平坦な領域を指向することがKleinbergらの研究によりわかった。平滑化とは、関数の形をなめらかにすることであり、損失関数の形状が平坦な領域は残したまま、とがった領域を解消することができる。平滑化を行う確率的勾配ノイズは学習率との積となっており、確率的勾配降下法は学習を収束させるために終盤の学習率を小さく取る必要がある。つまり、学習の終盤では確率的勾配ノイズによる平滑化の効果が小さくなってしまう。

## 3 提案手法

本研究では、陽にノイズを加えることで損失関数の平滑化を保ちつつ学習を収束させることで汎化性能が向上するのではないかという疑問を解決するために実験を行った。本研究の最適化手法の更新規則はKleinbergらにより再解釈された確率的勾配降下法の式から導い

た。提案手法は勾配に依存したノイズを陽に加える。

$$\theta_{t+1} = \theta_t + \mu \frac{\nabla L_B(\theta_t)}{\|\nabla L_B(\theta_t)\|_2} - \eta_t \nabla_\theta L_B \left( \theta_t + \mu \frac{\nabla L_B(\theta_t)}{\|\nabla L_B(\theta_t)\|_2} \right). \quad (1)$$

FINITEは勾配に依存しない球面上の一様ノイズを陽に加える。 $\hat{\epsilon} \sim U(S^{d-1})$ とする。

$$\theta_{t+1} = \theta_t + \mu \hat{\epsilon} - \eta_t \nabla_\theta L_B (\theta_t + \mu \hat{\epsilon}). \quad (2)$$

## 4 数値実験

2つの提案手法に対していくつかの条件を変更しながら、確率的勾配降下法との精度の比較を行った。用いたデータセットはどちらもCifar100である。

### 4.1 パラメータ平均なし

提案手法は確率的勾配降下法と比べて高い精度を示した。FINITEは確率的勾配降下法と比べて低い精度を示した。

Model	SGD	提案手法	FINITE
ResNet18	$79.54 \pm 0.23$	$80.23 \pm 0.04$	$79.50 \pm 0.16$
WideResNet28-10	$81.79 \pm 0.25$	$83.19 \pm 0.07$	$81.39 \pm 0.12$

### 4.2 パラメータ平均あり

提案手法はWideResNet28-10のモデルを用いた時、確率的勾配降下法と比べて高い精度を示した。FINITEはどちらのモデルに対しても確率的勾配降下法とほとんど同程度の精度を示した。

Model	SGD	提案手法	FINITE
ResNet18	$81.05 \pm 0.07$	$81.10 \pm 0.07$	$81.12 \pm 0.16$
WideResNet28-10	$84.06 \pm 0.09$	$84.57 \pm 0.08$	$84.03 \pm 0.14$

## 5 考察

提案手法は確率的勾配降下法よりも高い精度を得ることができているが、これは勾配に依存したノイズを陽に加えることで平滑化が担保されており損失関数の形が平坦な領域を指向しているからであると考えられる。FINITEは確率的勾配降下法よりも低い精度となっているが、これは学習の状態と無関係なランダムなノイズを陽に加えているため平滑化の効果が小さいのだと考えられる。このことから加えるノイズによって精度に差が出ることがわかる。

		指導教員	井智弘 准教授
学生番号	192C1014	氏名	井芹 健人
論文題目	パラメタ化 BWT のオンライン構築の高速化		

## 1 はじめに

パラメタ化文字列とは、静的文字の集合  $\Sigma$  とパラメタ文字の集合  $\Pi$  の和集合上の文字列で、2つのパラメタ化文字列  $x$  と  $y$  は、片方の文字列を他方に変換する  $\Pi$  上の全単射が存在する場合、パラメタ化一致するという。パラメタ化一致は文字列照合を一般化する方法の一つであり、ソフトウェア保守や剽窃検出、遺伝子の構造解析などへの応用がある。

Burrows-Wheeler 変換 (BWT) は文字列の可逆変換の一つで、文字列圧縮や FM-index と呼ばれる索引に利用される。BWT をパラメタ化照合用に拡張したものをパラメタ化 BWT という。Hashimoto ら [1] は、与えられた文字列に対するパラメタ化 BWT をオンラインで  $O(n|\Pi|(\frac{\log n}{\log \log n}))$  時間で計算するアルゴリズムを提案した。ここで  $n$  は文字列の長さである。本論文では、Hashimoto らの手法の時間計算量を  $O(n \log |\Pi|(1 + \frac{\log |\Pi|}{\log \log n})(\frac{\log n}{\log \log n}))$  時間に改善する手法を提案する。

## 2 パラメタ化 BWT

$\infty$  を任意の整数より大きい記号とし、 $N_\infty = N_+ \cup \{\infty\}$  とする。パラメタ化文字列  $x$  に対して、 $x$  の p-string encoding [2] は  $x$  上の各パラメタ文字の出現を同じ文字が直前に現れるまでの距離（そこが初めての出現の場合は  $\infty$ ）で置き換えて得られる  $(N_\infty \cup \Sigma)^{|x|}$  の要素である。P-string encoding はパラメタ化一致する文字列を同じ記号列に変換するためパラメタ化照合で利用できる。

また、 $\phi(x)$  を次のように定義する： $x[1] \in \Sigma$  ならば  $\phi(x) = x[1]$ 、そうでないなら  $\phi(x)$  は  $x[1 : i]$  中のパラメタ文字の種類数。ただし、 $i$  は  $x$  中で文字  $x[i]$  が左から 2 番目に現れる位置（無ければ  $i = |x|$ ）。

$T$  を末尾に辞書順最小の終端文字  $\$$  をもつパラメタ化文字列とする。 $T$  の全接尾辞を p-string encoding の辞書順でソートしたときに順位が  $i$  番目の接尾辞の開始位置を  $RA_T[i]$  とすると、パラメタ化 BWT は  $L_T[i] = \phi(T[RA_T[i] - 1 :])$  であるような配列  $L_T$  と定義できる（ただし、 $RA_T[i] = 1$  のとき、 $L_T[i] = \$$  とする）。

## 3 提案手法の概要

パラメタ化 BWT のオンライン構築問題では、文字列  $S$  のパラメタ化 BWT  $L_S$  から  $cS$  ( $c \in \Sigma \cup \Pi$ ) のパ

ラメタ化 BWT  $L_{cS}$  を計算することを考える。 $S$  と  $cS$  の接尾辞集合は  $cS$  を除いて等しいため、この計算は、 $L_S$  の中の  $\$$  を  $\phi(cS)$  で置き換えた後に、全接尾辞中で  $cS$  の p-string encoding の順位  $k$  を計算し、 $L_S$  の  $k$  番目の位置に  $\$$  を挿入することで達成される。

Hashimoto ら [1] の手法からの改善点は主に二点ある。まず、Hashimoto らはパラメタ化 BWT を巡回文字列で定義していたため  $L_S$  を  $L_{cS}$  に更新する際に  $O(|\Pi|)$ 箇所の値の変化に対応する必要があった。提案手法では接尾辞に基づいた定義に変更することでこれを回避した。次に、順位  $k$  の計算において、Hashimoto らは  $cS$  の p-string encoding より辞書順が小さい接尾辞の数を数え上げる方法をとっており、ここに  $O(|\Pi|)$  回の操作を要求していた。提案手法では、 $cS$  の p-string encoding と辞書順が最も近い接尾辞を求めるという別のアプローチをとることで計算がより効率的に行えることを示した。

提案手法では、動的に挿入操作が起こる動的文書列中で rank/select クエリを始めとした様々なクエリ・操作に効率的に応えるために Navarro と Sadakane の動的ウェーブレット木 [3] を用いる。提案手法は、 $cS$  の順位計算において、アルファベットサイズ  $O(|\Pi|)$  の動的文書列上で  $O(\lg |\Pi|)$  回のクエリを要するため、オンライン構築の一文字あたりの時間計算量は  $O(\log |\Pi|(1 + \frac{\log |\Pi|}{\log \log n})(\frac{\log n}{\log \log n}))$  となる。

## 4 終わりに

長さ  $n$  のパラメタ化文字列に対してパラメタ化 BWT を  $O(n \log |\Pi|(1 + \frac{\log |\Pi|}{\log \log n})(\frac{\log n}{\log \log n}))$  時間でオンライン計算する方法を提案した。今後は、使用する動的データ構造の見直しやクエリ回数の削減によって、アルゴリズムをより高速化することを目指す。

## 参考文献

- [1] Daiki Hashimoto, Diptarama Hendrian, Dominik Köppl, Ryo Yoshinaka, and Ayumi Shinohara. Computing the Parametrized Burrows-Wheeler Transform Online. SPIRE2022, pp. 70–85, 2022.
- [2] Sung-Hwan Kim and Hwan-Gue Cho. Simpler FM-index for parametrized string matching. Information Processing Letters, 165:106026, 2021.
- [3] Gonzalo Navarro, Kunihiko Sadakane. Fully Functional Static and Dynamic Succinct Trees. ACM Transactions on Algorithmsms 10(3): 16, 2014.

		指導教員	井智弘 淋教授
学生番号	192C1075	氏名	柴 啓仁
論文題目	SOLCA の符号化において 文脈情報を用いる手法の提案		

## 1 はじめに

日々多くのデータをやり取りする現代の情報社会において、容量削減や転送の高速化を可能とし、扱いやすくする圧縮技術は必要とされている。圧縮したデータを完全に復元することができる圧縮である可逆圧縮には、多くのアルゴリズムが存在し、活用されている。その中でも文法圧縮は、繰り返し構造の多いテキストを効率よく圧縮する手法の一つであり、近年そのような構造のデータが多くあることから、データ圧縮分野において盛んに研究が行われている。

文法圧縮では、第一には、与えられた文字列を生成するなるべく小さい文法を求めることが主題となる。しかし、実際の圧縮率を高めるためには、その文法を効率的に符号化する手法を考えることも重要である。本研究では、文法圧縮の手法の一つである SOLCA が output する文法を符号化する際に、PPM のアイデアを用いる手法を提案し、計算機実験により符号長が大幅に改善できる可能性があることを示す。

## 2 SOLCA と PPM

SOLCA [1] は、局所一致分解を用いて文字列を段階的に圧縮していく文法圧縮手法をオンラインかつ最適な領域計算量で行えるように改良したものである。SOLCA の求める文法は後置順 SLP (POSLP) と呼ばれる形式をしており、POSLP は枝刈りした導出木 Post-Order Partial Parse Tree (POPPT) の形と葉ラベル列の情報を記録することで保存できる。アルファベットサイズを  $\sigma$ 、文法の変数の数を  $n$  とすると、POPPT の形は  $2n + 1$  ビットで表現でき、葉ラベルの列は  $n \lg(n + \sigma)$  ビットで表現できる。SOLCA では、各葉ラベル列に使用するビットをさらに切り詰めることで、葉ラベルの列全体をおよそ  $n \lg(n + \sigma) - n$  ビットで符号化している。

Prediction by Partial Match (PPM) [2] は文字列中の各文字が生起する確率を、その文字が出現する直前の文脈から推定し符号化に用いる手法である。

## 3 提案手法

SOLCA は、概念的には、文字列を段階的に圧縮していく。各段階では、局所一致分解という手法を用いて文字列中のどこを変数で置き換えるか決定するが、局所一致分解の性質により、局所的な周辺環境がある

程度等しければ同じ場所が変数で置き換えられることが保証される。本研究では、PPM のアイデアを用いて、変数が生成された直前の文脈からその変数の生起確率を推定し、葉ラベル列の符号化に用いる手法を提案する。ある変数の出現はその周辺環境に影響を受けるため、文脈からその変数の生起確率がより良く推定でき、符号長が削減できることが期待できる。

## 4 実験

提案手法を実装し、繰り返し構造の多いテキストに対して、葉ラベル列の符号長がどれだけ削減できるかを実験した。結果を表 1 に示す。提案手法において、PPM の次数（文脈の最大長）は 5 とし、葉ラベル列の符号長は、生起確率をもとにした自己情報量を推定値として用いた。

表 1: 葉ラベル列の符号長の比較。

使用コーパス	変数の数 ( $n$ )	SOLCA[bit]	提案手法[bit]	削減率(%)
einstein.de.txt	164,427	2,700,350	2,422,590	10.3
einstein.en.txt	423,001	7,515,800	7,044,240	6.27
world_leaders	826,402	15,482,700	10,900,600	29.6
kernel	2,516,227	51,166,500	37,636,600	26.4
influenza	3,343,870	69,374,600	51,501,600	25.8
coreutils	4,695,508	99,612,100	70,642,400	29.1
Escherichia_coli	5,813,764	125,332,000	89,529,100	28.6
cere	6,363,446	137,974,000	93,355,100	32.3
para	8,031,150	176,331,000	121,167,000	31.3

## 5 終わりに

今回の実験から、提案手法によって SOLCA の葉ラベル列の符号長が最大 32.3% 改善することが確認できた。結果が改善した要因として、対象データが繰り返しが多いテキストであり、PPM によって葉ラベルの生起確率が良く推定できたことが考えられる。しかしながら、今回推測の際に PPM の弱点でもある計算時間の増加が見られた。SOLCA の高速で圧縮できる強みを失ってしまったので、実装の高速化は今後の課題である。

## 参考文献

- [1] Yoshimasa Takabatake, Tomohiro I and Hiroshi Sakamoto A Space-Optimal Grammar Compression. 25th Annual European Symposium on Algorithms (ESA 2017), 67:1–67:15.
- [2] John G. Cleary and Ian H. Witten. Data compression using adaptive coding and partial string matching. IEEE Trans. Commun. 32(4): 396–402, 1984.

		指導教員	井 智弘 准教授
学生番号	192C1084	氏名	高坂 彰人
論文題目	文法圧縮の平衡化の実装と評価		

## 1 はじめに

単一の文字列を生成する文脈自由文法をその文字列の straight-line program (SLP) と呼ぶ。SLP  $G$  は文字列の圧縮表現とみなせる。近年、SLP 上で様々な操作を行う手法が提案されているが、そのいくつか（例えばランダムアクセスなど）は文法の導出木の深さ  $d(G)$  に依存する計算量をもつ。 $G$  のサイズ  $|G|$  のオーダーを変えることなく深さを抑えた文法に変換できれば、これらの操作を高速化することができる。Ganardi ら [1] は長さが  $n$  の文字列を生成する SLP  $G$  が与えられたとき、同じ文字列を生成し、 $|H| = O(|G|)$  かつ  $d(H) = O(\log n)$  である SLP  $H$  を  $O(|G|)$  時間で計算する手法を提案した。本研究は、この手法を実装し、実際に生成される文法サイズと導出木の深さを実験的に検証することを目的とする。

## 2 Symmetric centroid decomposition

SLP は閉路のない有効非巡回グラフ (DAG) で表現できる。SLP の開始記号は DAG 上で唯一の根ノードとなり、終端記号は葉ノードとなる。根ノードを  $r$ 、葉ノードの集合を  $W$  とする。任意のノード  $u, v$  に対して、 $\pi(u, v)$  で  $u$  から  $v$  への経路の数を表す。また、ノード集合  $V$  に対して、 $\pi(u, V) = \sum_{v \in V} \pi(u, v)$  とする。DAG  $D$  上の任意のノード  $v$  に対して

$$\lambda_D(v) = (\lfloor \log_2 \pi(r, v) \rfloor, \lfloor \log_2 \pi(v, W) \rfloor)$$

とする。定義より、 $u$  の子ノードの中で  $\lambda_D(u) = \lambda_D(v)$  となるノード  $v$  は高々一つしかなく、そのような親子を結ぶ辺の集合を  $E_{scd}(D)$  と書くと、 $E_{scd}(D)$  によって DAG をパスの集合に分解することができる。

## 3 平衡化アルゴリズム

変数列  $V$ （終端記号を含んでよい）に対して、 $V$  中の変数の展開長の和を  $\|V\|$  と書く。scd 中の一つのパス  $X_1, \dots, X_p$  に着目する。 $X_0$  から  $X_{p-1}$  まで順番に見て、パスの左側にいる子ノードを並べて得られる列を  $L$ 、 $X_{p-1}$  から  $X_0$  まで順番に見て、パスの右側にいる子ノードを並べて得られる列を  $R$  とすると、変数  $X_i$  の展開文字列は  $L$  の接尾辞の変数列と  $X_p$  および  $R$  の接頭辞の変数列を展開し繋げることで得られる。

平衡化アルゴリズムは、部分問題として、 $L$  を入力として以下を満たすような生成規則を作る。

- 新たに作られる変数は最大で  $3|L|$  個。
- 各生成規則の右辺の長さは高々 4。
- 任意の  $j$  ( $1 \leq j \leq |L|$ ) に対して、 $L$  の  $i$  番目から始まる接尾辞を導出する変数  $S_j$  を作る。
- $S_j$  から  $L$  上の記号  $x$  までの経路長は最大  $3 + 2(\log_2 \|S_j\| - \log_2 \|x\|)$ 。

また、 $R$  を入力として  $R$  の接頭辞を導出する変数  $P_k$  も同様に作り、 $X_i$  の生成規則の右辺を  $S_j X_p P_k$  の形で置き換える。全ての scd パスに対して上記の処理を行うと平衡化した SLP  $H$  が得られる。

## 4 実験

平衡化アルゴリズムを実装し、RePair で圧縮して得られる文法  $G$  を入力として平衡化後の文法  $H$  と比較を行った。RePair で圧縮する文字列として、天然痘の塩基配列から長さが 70 の部分配列をいくつか抜き出して用いた。

$d(G)$	$d(H)$	$\frac{\log_2 n}{d(H)}$	$ G $	$ H $	$\frac{ G }{ H }$
33	12	0.511	84	185	2.02
33	13	0.471	82	185	2.01
33	11	0.557	84	186	2.21

表 1: 導出木の深さとサイズの比較結果

## 5 おわりに

平衡化によって  $d(H)$  が  $O(\log n)$  となり、サイズも  $O(|G|)$  となっていることが確認ができた。今回は実装の実行速度の都合上、小規模データでの実験にとどましたが、この研究を進めることで、導出木の深さに依存する計算量を持つ問題に対して、文法の平衡化が有効な手段となるか判断する材料を与えることができる。今後の課題として、実装の改善や、深さやサイズの定数項を抑えるような手法の開発などが挙げられる。

## 参考文献

- [1] Moses Ganardi, Artur Jez and Markus Lohrey. Balancing Straight-line Programs. J. ACM 68(4): 27:1–27:40, 2021.

		指導教員	井智弘 淋教授
学生番号	192C1144	氏名	松本 健志
論文題目	制限付きブロック圧縮の実装と評価		

## 1 はじめに

圧縮索引とは、文字列を圧縮した状態のまま任意のパターンの検索を可能とするデータ構造である。[\[1, 2\]](#)では制限付きブロック圧縮と呼ばれる文法圧縮に基づいた索引が提案されている。制限付きブロック圧縮では、圧縮の各段階で局所一致性を持つ locally consistent parsing を用いて文法の変数で置き換える部分（ブロック）を決定していくが、その段階で置き換えるには展開長が長すぎる変数は置き換えに参加させないようにしている。この長さ制限によって、パターン検索時に利用できる良い性質を得ているが、長さ制限を厳しくすると圧縮性能としては劣化することが予想される。本研究では、制限付きブロック圧縮の実装を行い、長さ制限の度合いを変化させたときの圧縮性能の変化を評価した。

## 2 制限付きブロック圧縮

$S_0$  を圧縮対象の文字列としたとき、制限付きブロック圧縮では、連長圧縮と局所一致分解を交互に行いながら段階的に変数に置き換え圧縮を進める。 $k$  番目の段階では記号列  $S_{k-1}$  を受け取り、 $k$  が奇数のときは制限付き連長圧縮で、 $k$  が偶数のときは制限付き局所一致分解で  $S_k$  を作る。ただし、各段階で展開長がおよそ  $\alpha^{\lceil k/2 \rceil - 1}$  以上である記号集合  $R_k$  を置き換え対象から除外する（ $\alpha$  は 1 より大きい実数）。

$T = S_{k-1}$  が与えられたときの手続きは以下の通り。

### 2.1 制限付き連長圧縮

1. 全ての  $i \in [1..|T|)$  において  $T[i] \in R_k, T[i+1] \in R_k$  または  $T[i] \neq T[i+1]$  の時  $T[i]$  と  $T[i+1]$  の間にブロック境界を設置する。
2. ブロック境界で挟まれる  $T[i..i+m]$  の長さが 2 以上である場合  $T[i..i+m] = a^m$  を変数で置き換える。

### 2.2 制限付き局所一致分解

1.  $\pi$  を  $T$  上の記号を整数に置き換える単射の中からランダムに選ぶ。
2. 全ての  $i \in [1..|T|)$  において  $T[i] \in R_k, T[i+1] \in R_k$  または  $\pi(T[i])$  が極小値である場合、 $T[i]$  と  $T[i+1]$  の間にブロック境界を設置する。

3. ブロック境界で挟まれる  $T[i..i+m]$  の長さが 2 以上である場合、 $T[i..i+m]$  を変数で置き換える。

## 3 実験

制限付きブロック圧縮を実装し、長さ制限の具合を変更したときの文法サイズの変化を計測する。各段階の長さ制限を  $l_k := (\alpha)^{\lceil k/2 \rceil - 1}$  とし、 $\alpha$  の値を変化させたときの文法サイズの変化を表 1 に示す。

表 1: 結果

$\alpha$	fib41	Escherichia.coli	dblp.xml.00001.2	einstein.de.txt
1.1	251	9911314	340538	256569
1.2	245	9803377	345996	251088
1.3	225	9623319	338054	244728
1.4	229	9532864	332277	240038
1.5	229	9438291	325971	239362
1.6	226	9534864	319702	235931
1.7	223	9376055	331447	232426
1.8	222	9546707	323602	231274
1.9	227	9501777	323416	231532
2.0	221	9462420	322181	231394

## 4 終わりに

[2] で提案されている制限付きブロック圧縮を実装し、長さ制限の具合を変更したときの文法サイズの変化を計測した結果、制限を緩くすると文法サイズが小さくなる傾向があることを確認した。既存研究において、長さ制限はパターン  $P$  の 1 次出現を探索する際にチェックするパターンの分割位置を  $O(\log |P|)$ 箇所に抑えるために用いられる。制限を緩くすることによって、文法サイズが減少する代わりに、この分割位置の数が増加することが予想される。今後の課題として、長さ制限と分割位置の数の関係を計測し、文法サイズとのトレードオフを確認することが挙げられる。

## 参考文献

- [1] Tomasz Kociumaka, Gonzalo Navarro, Nicola Prezza. Towards a Definitive Measure of Repetitiveness. LATIN 2020: 207–219.
- [2] Tomasz Kociumaka, Gonzalo Navarro, Francisco Olivares. Near-Optimal Search Time in  $\delta$ -Optimal Space. LATIN 2022: 88–103.

		指導教員	佐藤 好久 教授
学生番号	192C1105	氏名	中野 尚仁
論文題目	位相的データ解析による銀河・星団の圧縮率の提案		

## 1 はじめに

近年、情報技術の発展に伴い、収集できるデータの数やその種類は大きく増加している。これに伴い、膨大なデータから有益な情報を抽出するデータ解析のさらなる発達が求められている。本論文の目的は、データ解析技術に対する新しい手法として注目されている TDA (Topological Data Analysis) 技術の有用性を確かめ、新たに応用できないかを考えることである。本研究の目的は、TDA を用いて、銀河・星団のデータ解析的指標として「圧縮率」を提案し、いくつかの銀河・星団に対して圧縮率の計算を行った。

## 2 TDA

有限個のデータからなる集合である点群をチェック複体  $\mathcal{C}(V; r)$ などを用いて「図形」として幾何的に表現する。これは、正の数  $r$  の増加に対して減ることなく増え続ける。このように  $r$  の増大列  $r_0 < r_1 < \dots < r_T$  に対応するチェック複体の増大列  $\mathcal{C}(V; r_0) \subset \mathcal{C}(V; r_1) \subset \dots \subset \mathcal{C}(V; r_T)$  をフィルトレーションという。

フィルトレーション  $\mathbb{K}$  の各時刻  $t$  における単体複体  $K_t$  の  $\mathbb{Z}_2$  係数ホモロジ一群  $H_k(K_t; \mathbb{Z}_2)$  を考え、これらホモロジ一群の時系列の表現として、フィルトレーション  $\mathbb{K}$  のパーシステントホモロジ一群  $PH_k(\mathbb{K})$ を得ることができる。そして、 $PH_k(\mathbb{K})$  の生成元のもつ birth-death の時間対として、パーシステント図を得ることができる。3次元散布図とそれを点群として得たパーシステント図の例を以下に示す。

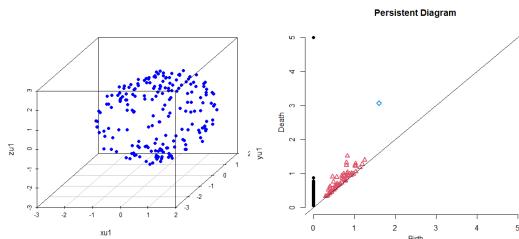


図 1: 3 次元の点群とパーシステント図

右図は左図の散布図を点群とするパーシステント図を可視化したものである。右図において、「●」は、0 次元、赤色の「△」は、1 次元、青色の「○」は、2 次元の生成元を表す。

## 3 実験方法

Gaia EDR3 から銀河・星団の恒星の位置データを抽出する。その際、パラメータは中心座標、半径 Radius、視差 parallax を指定する。中心座標は抽出したい天体名、半径は Radius= 3、または、Radius= 10、視差 parallax は  $6 \leq x < 9$  とする。

- (1) 銀河・星団の位置データを点群として TDA を行い、点群のパーシステント図を作成する。TDA を行うシステムとして、HomCloud を用いる。
- (2) 求めた 1 次元・2 次元のパーシステント図から位相的ノイズ  $N(\delta)$  を取り除く。取り除いた後、1 次元のサイクルの数に対する 2 次元サイクルの数の比を圧縮率として計算する。今回は、 $\delta = 0.025, 0.05, 0.075, 0.1$  で検証する。

## 4 実験結果

半径 Radius= 3 における各銀河・星団の圧縮率をまとめた図 1 を以下に示す。横軸は、今回実験を行った銀河・星団の天体名、縦軸にはそれぞれの、 $\delta$ ごとの圧縮率の値を取っている。半径 3 度のときは  $\delta$  の値の変化に応じて、圧縮率の値にばらつきが生じた。

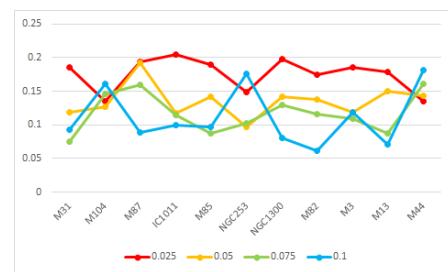


図 2: 半径 3 度の圧縮率

## 5 今後の課題

何らかの規則に従って圧縮率の値が変化していることを読み取ることはできなかった。今後の課題として、今回求めた値と相関関係をもつ統計量やパーシステント図の選別の際の最適なパラメータ  $l, u, \delta$ を見つけ出す必要がある。また、天体ごとの特徴を捉えられるようなデータの抽出が行えるよう修正を加えていきたい。

		指導教員	佐藤 好久 教授
学生番号	192C1133	氏名	古川健太郎
論文題目	素数判定法のアルゴリズムと計算処理時間の評価		

## 1 研究背景と目的

暗号システムは情報の安全性を確保するために利用している。暗号システムには素数を用いているものがあり、素数は目の見えない部分で我々の役に立っている。素数を作り出す関数はいまだに発見されていないため、作り出した整数が素数かどうか判定する必要がある。素数判定法はさまざまな種類存在するが、それは指数時間のアルゴリズムであり、入力する整数が大きい場合、莫大な時間を要した。しかし、2004年に世界初の多項式時間の決定性アルゴリズムである AKS 素数判定法が発表された。本実験では、AKS 素数判定法が多項式時間で行えるか検証し、従来の素数判定法と比較し、AKS 素数判定法を評価する。

## 2 実験方法

今回の実験では、フェルマーテスト、ミラーラビンテスト、試し割り法、AKS 素数判定法の 4 種類の素数判定法を用いる。方法として、用意した整数を  $N$  とし、2 から  $N$  までの整数の全ての素数判定を行い、その合計時間を計測する。そして、各桁ごとの計算処理時間の平均を算出する。

また、フェルマーテストとミラーラビンテスト、AKS 素数判定法は素数判定の精度の評価を 5 衡までの整数に対し行う。

AKS 素数判定法の多項式時間で行えるか詳しく調査するため、ランダムに生成された整数  $N$  を AKS 素数判定法で判定にかかる時間を測定し、平均を算出する。

## 3 結果

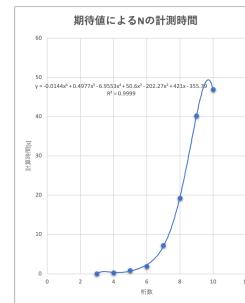
計測結果をグラフにまとめ、回帰分析を行った。従来の判定法は指數関数の回帰分析で精度の高い結果が出た。AKS 素数判定法は指數関数と多項式の回帰分析の両方の精度が高いと出力された。また、従来の素数判定法と AKS 素数判定法で、AKS 素数判定法が他の判定法と比べ多くの計算処理時間を要している。

フェルマーテストとミラーラビンテストの精度比較については、5 衡まで正確に判定できていたので 250000 まで判定する。フェルマーテストは 200000 を超えてから擬素数を判定する割合が増えてきたが、ミラーラビンテストは 250000 までの整数は正確に判定した。AKS 素数判定法は 250000 まで素数と合成数の正確な判定が出来た。

AKS 素数判定の整数  $N$  のみの判定をさせた結果について、多項式時間のアルゴリズムであり、指數時間ではないことが確認できたが、10 衡目の判定で測定値が急激に小さくなかった。

## 4 考察

実験結果により、AKS 素数判定法が多項式時間の決定性アルゴリズムであることは検証できた。しかし、この結果から入力した  $N$  のうち、合成数の割合が高いことで計算処理時間が減った可能性を考えた。そのため、測定結果のうち、計算処理時間の期待値を算出した。すると、桁数が上がるにつれて計算処理時間も上昇する結果となった。(下のグラフを参照) よって、合成数の割合が高かったと考えた。



さらに、AKS 素数判定法は従来の判定法と比べ、かなり多くの時間を要することが実験から分かった。これにより、現状の AKS 素数判定法では実際に使うことができず改良し、計算処理時間を短くする必要があると考えている。

## 参考文献

- [1] 佐藤好久, 初等整数論.
- [2] 佐藤好久, 代数学 I.
- [3] Manindra Agrawal, Neeraj Kayal, Nitin Saxena, [https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/primality\\_journal.pdf](https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/primality_journal.pdf), 2004.
- [4] 和田秀男, 素数全書, 朝倉書店, 2010.
- [5] 中村憲, 数論アルゴリズム, 朝倉書店, 2009.
- [6] 中村憲, 素数判定 PRIMES が決定性多項式時間 P である事の証明について, [https://www.jstage.jst.go.jp/article/sugaku1947/56/1/56\\_1\\_73/\\_pdf/-char/ja,2003](https://www.jstage.jst.go.jp/article/sugaku1947/56/1/56_1_73/_pdf/-char/ja,2003).
- [7] ユークリッド, ユークリッドの原論, 紀元前 3 年.

		指導教員	佐藤 好久 教授
学生番号	192C1139	氏名	又吉 美音
論文題目	ベイジアンネットワークを用いた アメリカンフットボールの得点戦略推定		

## 1 はじめに

本研究の目的はベイジアンネットワークを用いてアメリカンフットボールの解析をすることである。得点の獲得時に着目し、セイバーメトリクス的に試合情報を入手し解析を行う。アメリカンフットボールはアメリカを中心に世界で人気のあるスポーツである。攻守が明確に分かれている点から、データ解析が行いやすく、現在では様々な方法のデータ解析が行なわれている。アメリカンフットボールでのデータ解析における問題点は、戦況によって複雑に変化する状況である。ベイジアンネットワークには目的変数と説明変数を設定しないという特徴があり、新たな因果関係の発見が可能である。

## 2 結果

初めに、山登り法を用いてベイジアンネットワークを構築した。その結果から独立したノード、環境のノード、相関の強いノードを削除し、必要なノードのみを残した。最終的に決定したノードを用いて山登り法でベイジアンネットワークを構築した図が図 1 である。

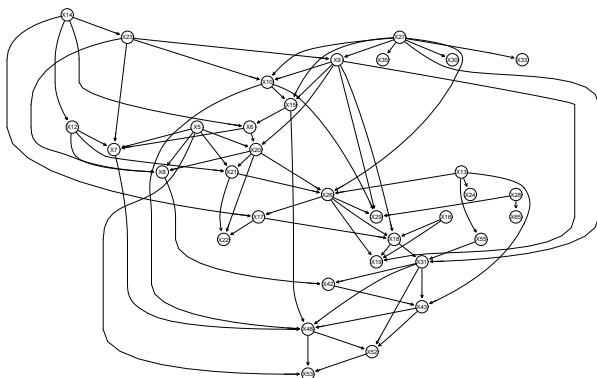


図 1: 構築したベイジアンネットワーク

次に、GS アルゴリズムを用いてベイジアンネットワークを構築した。GS アルゴリズムでは向きの無いアーケが発生した。また、特定の条件のみのデータからベイジアンネットワークを構築し、比較を行った。得られた DAG から表 1 のように条件付き確率を求めた。

最後に、条件付き確率を用いて得点獲得までのシミュレーションを行った。結果を表 2 に示す。この結果

表 1: 条件付き確率

	Play	G.Int	E.Int	G.Tyds	E.Tyds	頻度
1	pass	1	0	370 ~ 416	276 ~ 322	0.007
2	pass	1	0	370 ~ 416	323 ~ 369	0.00654
3	FG	0	0	370 ~ 416	323 ~ 369	0.00642
4	pass	1	0	370 ~ 416	229 ~ 275	0.00584
5	FG	1	0	370 ~ 416	276 ~ 322	0.00584

は、13 ヤード以下のパスプレイがタッチダウンする確率が一番高いことを示している。

表 2: 条件付き確率 2

	Play	yds	確率
1	pass	~ 13	0.37205
2	run	~ 13	0.24168
3	pass	14 ~ 26	0.09191
4	run	14 ~ 26	0.06123
5	pass	27 ~ 39	0.03202

## 3 考察

構築されたベイジアンネットワークには、想定と異なる点がいくつか存在した。一つはアーケの始点にならないはずのノードが始点となったことである。これは構築した山登り法がスコアベースのアルゴリズムであることが原因である。次に、独立したノードが発生した点である。これは、取得したデータに含まれる事象が少なかったことが原因である。

構築した結果から、勝敗とインターセプト数に因果関係があることが分かった。インターセプトはアメリカンフットボールにおいて勝敗の鍵を握る重要な要素であることが解析の結果からも判明した。

## 参考文献

- [1] J.Pearl, “Bayesian networks: A model of selfactivated memory for evidential reasoning,” In *Proceedings of the 7th Conference of the Cognitive Science Society*, pp.329–334, 1985
- [2] Marco Scutari 著, Jean-Baptiste Denis 著, 金 明哲 監修, 財津 亘 翻訳, R と事例で学ぶベイジアンネットワーク原著第 2 版, 共立出版, 2022

		指導教員	佐藤 好久 教授
学生番号	192C1160	氏名	柳澤 孝紀
論文題目	種数 1 の Lefschetz fibration における positive relator の生成システム		

## 1 研究背景と目的

有限個の Lefschetz 型特異ファイバーを許した曲面上の曲面束である 4 次元多様体を Lefschetz ファイバー空間という。特に、一般ファイバーの種数が 1 (2 次元トーラス  $T^2$  と同相) であり、底空間が球面  $S^2$  であるものは構造上研究対象として調べられることが多く、複素構造を付加したものは橢円曲面と呼ばれ、代数幾何的に重要な 4 次元多様体である。これを  $E(n)$  と書くこととする。 $E(n)$  への手術と Seiberg-Witten 不变量の手術公式はその上多くの微分構造をあたえる。 $E(n)$  上の微分構造に関しては [1, 2, 3] を参照してもらいたい。また、Gompf と Donaldson による仕事の帰結 ([4, 5]) により、Lefschetz ファイバー空間はファイバーが symplectic 部分多様体となる 4 次元 symplectic 多様体である。逆に 4 次元 symplectic 多様体は何回かの blow-up を行った後に、Lefschetz ファイバー空間となる。symplectic 多様体はケーラー多様体の一般化というだけでなく、幾何学的に豊富な構造を含んでおり、symplectic 幾何学は素粒子物理学における弦理論などにも深く関わる分野である。また、ゲージ理論に対しても相性が良く、Seiberg-Witten 不变量にもその特徴を残している。

Lefschetz ファイバー空間において有限個の Dehn ツイストの積が 1 となるものを positive ralator という。この positive ralator の計算や調査は、専門的な知識と共に長年の経験が必要とされており、非専門家にはなかなか手が出しづらいものとなっている。そこで、positive ralator を形式文法で表現し、それをもとに positive ralator を生成するプログラムを作成することで、長年の経験がなくとも気軽に positive ralator を扱うことが可能となる支援システムの作成を目的とする。

## 2 システムの生成

2 次元トーラス  $T^2$  の自己同相写像が作る写像類群  $M_1$  は、Dehn ツイスト  $\tau_m, \tau_l$  により定義され、関係式 (1)  $\tau_1\tau_2\tau_1 = \tau_2\tau_1\tau_2$ 、(2)  $(\tau_1\tau_2\tau_1^2\tau_2\tau_1)^2 = 1$  により表現される群である。これらの関係式 (1)、(2) を用いて種数 1 の positive ralator の形式文法を作成する。さらに、その形式文法と Mathematica を用いて、positive ralator を生成するプログラムを作成した。

### 定理

positive word 言語  $L((ab)^6)$  は次の形式文法で表される。

$$\begin{aligned} G &= \langle N, \Sigma, P, S \rangle \\ N &= \{S, A, A^{-1}, B, B^{-1}\} \\ \Sigma &= \{a, a^{-1}, b, b^{-1}\} \\ P &= \{S \rightarrow SS, S \rightarrow ABABABABAB, \\ &\quad ABA \rightarrow BAB, BAB \rightarrow ABA, \\ &\quad AB \rightarrow ABA^{-1}A, BA \rightarrow BAB^{-1}B, \\ &\quad A \rightarrow aAa^{-1}, A \rightarrow a^{-1}Aa, A \rightarrow bAb^{-1}, \\ &\quad A \rightarrow b^{-1}Ab, B \rightarrow aBa^{-1}, B \rightarrow a^{-1}Ba, \\ &\quad B \rightarrow bBb^{-1}, B \rightarrow b^{-1}Bb, A \rightarrow a, \\ &\quad A^{-1} \rightarrow A, B \rightarrow b, B^{-1} \rightarrow b^{-1}\} \\ S &= S \end{aligned}$$

このときの  $G$  は、文脈依存文法なので、 $L((ab)^6)$  は文脈依存言語である。

## 3 今後の課題

今回のシステムを用いて多くの positive ralator の例を作成し、データとしてまとめ、作成した positive ralator に対応する Lefschetz ファイバー空間について研究していきたい。また、種数 2 以上の場合でも使えるシステムを作成したいと感じた。

### 参考文献

- [1] R. Friedman and J. Morgan, "Smooth 4-manifolds and complex surfaces", Ergeb. Math. Grenzgeb. vol. 27, Springer-Verlag, 1994.
- [2] R. Friedman and J. Morgan, "Algebraic surfaces and Seiberg-Witten invariants", J. Algebraic Geom. vol.6(1997), pp. 445–479.
- [3] Y. Kametani and Y. Sato, "0-dimensional moduli spaces of stable rank 2 bundles and differentiable structures on regular elliptic surfaces" Tokyo J. Math. vol. 17 (1994), pp. 253–267.
- [4] R. E. Gompf and A. I. Stipsicz, "4-Manifolds and Kirby Calculus", Grad. Stud. Math. 20, Amer.Math. Soc., Providence, RI, 1999.
- [5] S. K. Donaldson, " Lefschetz pencils on symplectic manifolds", J. Diff. Geom., 53 (1999), pp. 205–236.

		指導教員	古川 徹生 教授
学生番号	192C1098	氏名	徳永 優希
論文題目	非線形振動現象の動的システムメタモデリング		

## 1 研究背景と目的

時間発展する現象を生成する微分方程式系を動的システムという。動的システム学習とは、観測された時系列から動的システムを同定する逆問題を解くことである。動的システム学習をすることで、新規の時系列を学習したシステムから生成したり、予測したりすることが可能になる。一方メタモデリングとは、複数の異なる学習タスクが与えられたとき、学習により得られたモデル集合をさらに高次モデルで学習することである。これにより、新規システムの生成やゼロショット学習が可能になる。

本研究の目的は、動的システムのメタモデリング法を開発することである。とりわけ本研究では非線形振動現象に注目し、非線形振動現象の動的システムメタモデリングを試みる。このようなメタモデリングは、多様な歩行や走行パターンの適応的な生成や、異なる身体間での模倣、ミラーシステムの実現などへの発展が期待できる。

本研究で開発するのは、Unsupervised Kernel Regression for Dynamics Modeling: UKRDM をメタモデリングに拡張した UKRDM<sup>2</sup> である。開発した UKRDM2 を Van der Pol 非線形振動に応用し、メタモデルにおいて Hopf 分岐が起きることを示す。

## 2 問題設定

$I$  個の動的システムから生成された時系列データ集合  $\underline{\mathbf{X}} = \{\mathbf{X}_i\}_{i=1}^I$  が与えられている状況で、それぞれのデータに対応する潜在状態変数  $\underline{\mathbf{Z}} = \{\mathbf{Z}_i\}_{i=1}^I$  とシステム特性パラメータ集合  $\Theta$ 、観測特性パラメータ集合  $\Phi$ 、そしてメタシステム  $F$  およびメタ観測変換  $G$  を推定する。

## 3 UKRDM<sup>2</sup>

UKRDM<sup>2</sup> は、UKR<sup>2</sup> [1] の学習構造を取り込むことによって動的システムメタモデリングを可能としたアルゴリズムであり、目的関数は

$$E_1[\underline{\mathbf{Z}}, \Theta] = \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \sum_{t=1}^{T-1} \|\mathbf{z}_{t+1}^{ij} - F(\mathbf{z}_t^{ij}, \theta_i)\|^2 \quad (1)$$

$$E_2[\underline{\mathbf{Z}}, \Phi] = \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \sum_{t=1}^T \|\mathbf{x}_t^{ij} - G(\mathbf{z}_t^{ij}, \phi_i)\|^2 \quad (2)$$

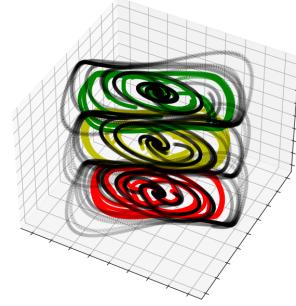


図 1: 推定した潜在状態変数を観測空間に写像した  $G(\mathbf{z}, \phi)$

というメタシステム  $F$  の二乗誤差とメタ観測変換  $G$  の二乗誤差を足し合わせた以下の式で表される。

$$E[\underline{\mathbf{Z}}, \Theta, \Phi] = \lambda E_1[\underline{\mathbf{Z}}, \Theta] + E_2[\underline{\mathbf{Z}}, \Phi] \quad (3)$$

このとき、 $\lambda$  は第一項と第二項の比率を決めるハイパーパラメータである。

## 4 非線形振動子を用いた実験

UKRDM<sup>2</sup> を用いて、非線形振動子の Hopf 分岐を再現できるかの検証を行なった。結果を図 4 に示す。図 4 から、それぞれの観測特性パラメータに対して、観測空間中の多様体が分離していることから、身体特性パラメータを推定できていることがわかる。また、潜在状態変数を観測空間中に写像した点を見していくと、それぞれが異なるアトラクタを生成していることがわかる。これにより、Hopf 分岐も再現できていると言える。

## 5 まとめ

本研究では、非線形振動子の Hopf 分岐を再現することを通して、UKRDM<sup>2</sup> の検証を行うことを目指した。今後の展望としては、歩行などの実データを用いての検証を行なっていきたいと考えている。

## 参考文献

- [1] 津野龍, 石橋英朗, 古川徹生, マルチタスク多様体モデリングの解くべき問題は何か～最尤推定と最適輸送距離のアプローチ比較～, 信学技報, Vol.120, No.403, pp.53-58, 2021.

# 九州工業大学 情報工学部 知能情報工学科 2022 年度 卒業論文 概要

		指導教員	和田親宗 教授
学生番号	192C1080	氏名	鈴木 春菜
論文題目	視覚障害者の道路横断を支援する情報提示方法の検討		

## 1 背景と目的

所属研究室では「視覚障害者の歩行支援装置」の研究開発が行われているものの、得られた情報の伝達方法について研究はまだなされていない。本研究の目的は、視覚障害者が単独で安全に道路横断が出来るような適切な情報伝達方法を調べることである。

## 2 提案手法

今回は、横断歩道までの接近と進行方向に関する情報を提示することを想定した。その情報を、200Hz の振動刺激を手首に、音声情報を骨伝導型イヤホンで、以下のパターンで提示した。

### (1) 横断歩道接近の提示

1. 歩道上を横断歩道に接近するまで断続的に振動を与える。
2. 歩道上を横断歩道に接近するまで連続的に振動を与え、横断歩道に近付くと振動が強くなる。

### (2) 進行方向の提示

進行方向が直進から逸れている間、振動刺激を提示すると共に、左右のずれを音声で提示する。振動は断続の場合と連続の場合の 2 パターンである。

## 3 結果

屋外の横断歩道で手法の評価実験をおこなった。この際、SUS(System Usability Scale)アンケートにも回答してもらった。

### (1) 横断歩道接近の提示

刺激提示から歩行停止までの時間を測定した。断続的な振動刺激提示の場合は平均 2.0[s]、連続刺激の場合は平均 1.4[s]であった。アンケート結果は、断続刺激の場合、平均 62.0 点、連続刺激の場合、平均 69.1 点となった。いずれの結果も連続提示の方が良い結果となった。

### (2) 進行方向の提示

直進から逸れた後、直進方向に修正するまでの時間を測定したところ、被験者によって結果が異なった。アンケート結果では、断続刺激は平均 66.6 点、連続刺激は平均 72.5 点となり、連続刺激の方が良い結果となった。

## 4 考察とまとめ

横断歩道接近の提示では、刺激の有無のわかりやすさのため、連続的な振動パターンが良いと考える。一方、進行方向の通知方法についても、連続的な振動パターンがより明快であるが改良の余地があると考える。

		指導教員	國近 秀信 准教授
学生番号	182C1155	氏名	松木 悠馬
論文題目	英語パラグラフの構成要素同定機能の改良 ～比較・対照パラグラフの種類同定～		

## 1 はじめに

英語初学者にとって、英語の論理展開法に従ったパラグラフライティングを行うことは難しい。これは論理展開法への理解が不十分で、日本語の文章との違いを把握できていないためである。Web や参考書にはパラグラフが大量にあり、これらを参考にすることが可能である。しかし、参考にするパラグラフが論理展開法に従っているか否かを判断するのは、初学者にとって困難である。そこで先行研究 [1] では、Web ページや参考書のパラグラフが論理展開法に則っているかを判断するシステムの実現のため、英語パラグラフの構成要素同定機能の実現を行った。さらに、パラグラフの構成要素同定機能の向上のためのパラグラフの種類同定機能の実現を行った [2]。本研究では、パラグラフの種類同定機能の改良、特に比較・対照パラグラフの種類同定機能の実現を目的とする。

## 2 パラグラフの構成要素同定

論理展開法に従ったパラグラフか否かを判断するため、パラグラフの構成要素同定を。英語行う。パラグラフの構成要素には Topic Sentence や Concluding Sentence 等の構成要素が存在する。ただしパラグラフの目的や種類によって構造が異なるため、パラグラフの種類ごとに典型的な構造を表すパラグラフ展開スキーマの構成要素同定を行う。構成要素同定の際には、文の位置や品詞、各構成要素特有の頻出語等を利用する。

## 3 パラグラフの種類同定

先行研究 [2] では、定義、列挙、比較、叙述の四つの種類のパラグラフを同定する機能を実現したが、パラグラフの頻出語を利用するのみであり、性能が十分ではなかった。そこで本研究では、パラグラフライティングに関する参考書や、条件考案用に用意したパラグラフをもとに、比較・対照パラグラフの特徴を把握し、改良することとした。

比較・対照パラグラフは二つの事物の類似点や相違点を述べるために用いられる。そこでは、まず Topic Sentence 内で二つのトピック A および B が出現し、その後の文章で、A および B が代名詞等により繰り返し出現するという特徴がある。

そこで本研究では、この特徴と、実際のパラグラフを参照しながら、以下の同定条件を定義し、実現した。なお、構文解析および語句の参照関係同定のため、Stanford CoreNLP[3] を利用した。

条件 1：3 回以上同じ単語・またはそれを指す単語が登場し、使用された代名詞が they や these などの複数形である。

条件 2：Topic Sentence 内で名詞句と名詞句を結ぶ and が存在する。

条件 1 と条件 2 を同時に満たす場合、比較・対照パラグラフと同定する。

## 4 評価

さまざまな種類のパラグラフを計 141 個用意し、比較・対照パラグラフの種類同定プログラムを適用して適合率と再現率を求めた。先行研究 [2] では適合率 40.8 %、再現率 55.6 % であったのに対し、本研究では適合率 71.4 %、再現率 27.7 % であった。本研究ではユーザーへの間違ったパラグラフの提示を避けるため、適合率を重視する。本研究により適合率の改善が確認できたため、有用性を高めることができたと考える。

## 5 おわりに

本研究では、パラグラフの構成要素同定機能の改良として、比較・対照パラグラフの種類同定機能の改良を行った。今後の課題は、比較・対照パラグラフの種類同定について適合率をより向上させるとともに、他の種類のパラグラフに関しても同定できるように機能を拡張することである。

## 参考文献

- [1] 藤崎 凌太, 英語パラグラフの構成要素同定機能の改良～比較・対照パラグラフを中心～, 2019 年度九州工業大学卒業論文 (2020).
- [2] 濱村 鳩吾, 英語パラグラフの構成要素同定機能のためのパラグラフ種類同定, 2020 年度九州工業大学卒業論文 (2021).
- [3] Stanford NLP Group: "Stanford CoreNLP", <https://stanfordnlp.github.io/CoreNLP/> (参照日:2022.02.09)

		指導教員	國近 秀信 准教授
学生番号	192C1002	氏名	阿部 聖史
論文題目	英語エッセイアウトラインの論理展開確認機能の改良		

## 1 はじめに

英語エッセイを記述する際には、手順や構成要素など複雑なルールに則り、説得力のある論理展開で書くことが重要である。しかし、エッセイライティングの知識が不十分である英語初学者にとって、説得力のある論理的なエッセイを記述することは困難である。先行研究[1, 2]では、英語エッセイライティングの知識が十分ではない初学者を対象とし、英語の論理展開法に沿ったアウトラインの作成を支援するためのアウトラインエディタおよび論理展開確認機能を実現した。しかし、論理展開確認機能において、ユーザが作成したアウトラインに対して正しく診断が行えないことがあるなどの問題があった。そこで本研究では、このような問題を解決するため、論理展開確認機能の改良を目的とする。

## 2 エッセイアウトライン作成支援

エッセイライティングでは、まずユーザが記述したい内容に相当するアイデアを書き出して整理する。その後、整理したアイデアの中から必要だと考えたアイデアを選択して並べることで、アウトラインを作成する。本研究では、英語エッセイに関する知識が不十分なユーザでも論理展開法に則ったアウトラインの作成が可能となるよう支援するため、ユーザが指定したエッセイの種類に対応したアウトラインのひな形を提示する。ユーザは、ひな形に対して、アイデアの追加、移動、削除、構造の変更等を行うことができる。そしてユーザがひな形の構造を変更した際に、作成したアウトラインが論理展開法に基づいたものとなっているか否かについて診断を行い、必要に応じて修正のための支援を行う。

## 3 エッセイアウトライン作成支援システム

本システムは、ラベル付きクラスタ図作成ツール、アウトラインエディタ、診断機能、および、英語エッセイ／パラグラフの論理展開法に関する情報を保持するエッセイ／パラグラフ展開スキーマデータベースから成る。ユーザは始めに、書きたい内容に相当するアイデアをラベル付きクラスタ図として列挙・整理する。そして、アウトラインエディタを用いて、ユーザが自ら指定した種類のエッセイのひな形へ、ラベル付きクラスタ図中のアイデアを当てはめ、アウトラインを作成

する。アウトライン作成後は、論理展開確認機能を用いてアウトラインの確認を行う。論理展開確認機能の診断項目には、アイデア不足、構成要素不足、構成要素過多、構成要素順序、異なる構成要素の使用、パラグラフの種類、エッセイの構成要素の種類があり、ユーザは、説得力のあるエッセイにするため、診断結果を基に、納得がいく範囲でアウトラインを書き直す。

## 4 論理展開確認機能の改良

アウトラインエディタに表示しているひな形には、パラグラフの主題の支持文に相当するアイデア Item と、その説明文に相当するアイデア Explanation of the item との関係を明示するための仕組み（以下、ダミーノード）がある。先行研究では、このダミーノード内の診断、および、パラグラフ内に複数のダミーノードが存在する場合の診断、エッセイの構成要素であるパラグラフの種類を変更した場合の診断が正しく行えないという問題があった。また、論理展開確認機能等が最新のアウトラインエディタに未対応であった。そこで本研究では、主に、ダミーノード内の診断、および、複数のダミーノードが存在する場合の診断が正しく行えるように改良した。具体的には、ユーザが作成したひな形のダミーノード内の要素の数だけスキーマと比較するように変更することで、ダミーノード内の要素全てを診断可能とした。また、ダミーノードが複数存在する場合には、ダミーノード内の構成要素の診断を再帰的に行うよう改良した。

## 5 おわりに

本研究では、英語エッセイアウトライン作成支援のための論理展開確認機能の改良を行った。今後は、本機能による指摘箇所について、アウトラインエディタおよびフィードバック画面で強調表示を行う機能を実装すること、および、論理展開確認機能の有用性について評価を行うことが必要である。

## 参考文献

- [1] 三好秀佳, エッセイ展開スキーマを用いた英語エッセイの論理展開確認機能の実現, 2015 年度九州工業大学卒業論文, 2016.
- [2] 山元莉帆, 英語エッセイライティングのためのアウトラインエディタの改良, 2019 年度九州工業大学卒業論文, 2020.

		指導教員	國近 秀信 准教授
学生番号	192C1019	氏名	因幡 繁伸
論文題目	英語パラグラフアウトライン作成支援システムの授業実践による有用性調査		

## 1 はじめに

英語パラグラフを説得力のあるものにするには、適切な論理展開で文章を書くことが必要である。しかし、それは知識の乏しい英語初学者にとって困難である。そこで先行研究 [1] にて、初学者を対象とした英語パラグラフアウトライン作成支援システムが実現された。本研究では、システムが有するクラスタ図診断機能と整合性確認機能の評価を実際の授業環境で行って機能の有用性を確認するとともに、システムを授業で使用するまでの課題を明らかにすることを目的とする。

## 2 英語パラグラフアウトライン作成支援

本研究のパラグラフライティングでは、ラベル付きクラスタ図を用いてアイデアの収集・整理を行った後、必要なアイデアを選んでアウトラインを作成する。しかし、ラベル付きクラスタ図の作成時、知識が不十分な英語初学者は必要な量のアイデア収集や適切なラベル付けが困難である。そこで本研究では、クラスタ図診断機能によって、ユーザが作成したラベル付きクラスタ図を診断し、修正支援メッセージを表示することで誤り解消を支援する。また、アウトライン作成作業はラベル付きクラスタ図の作成と行き来して繰り返し行われるため、両者の間に整合性の問題が生じるおそれがある。そこで本研究では、整合性確認機能が誤りを診断し、矛盾の解消を支援する。

## 3 システムの概要

本システムは、ラベル付きクラスタ図作成ツール、アウトラインエディタ、エッセイ／パラグラフ展開スキーマデータベースから成る。ユーザはまず、ラベル付きクラスタ図作成ツールを用いてアイデア整理を行う。このときクラスタ図診断機能が図の修正を支援する。次に、作成した図から必要なアイデアを選択し、アウトラインエディタに並べてアウトラインを作成する。その後、整合性確認機能が矛盾の解消を支援する。

## 4 有用性調査

本システムの有用性の評価として、英語授業の受講者である本学学部1年生76名を対象に調査を行った。本調査では、被験者が作成したラベル付きクラスタ図の誤り、および、ラベル付きクラスタ図とアウトライン間の整合性の問題について、自力での修正とクラス

タ図診断機能および整合性確認機能を使用した修正の結果を比較した。まずクラスタ図診断機能に関して、集計可能な36名のデータを分析した。書き出し段階では660個の誤りが指摘され、自力での修正にて誤りが4.7%減少し、629個となった。その後、機能を用いた修正によって誤りが63.4%減少し、230個となった。したがって、自力ではほぼ修正できなかった誤りが、機能の使用によって大幅に解消されており、本機能が誤り解消の一助になると判断できる。なお、修正されなかった誤りの主要な原因として、「誤りを修正しても新たに誤りが発生する」ことが挙げられる。そのため、誤りが発生しづらい修正手順を提示する仕組みが必要であると考える。

次に整合性確認機能に関しては、集計可能な22名のデータを分析した。アウトライン作成段階で指摘された229個の誤りが自力での修正にて8.7%減少した。その後、未解消の209個の誤りが機能を用いた修正により、67.0%減少し69個となった。以上より、本機能が整合性の誤り解消を支援できることが確認できた。誤りが未解消の場合のデータを確認した所、ラベル付きクラスタ図を適切に作成できない被験者は、誤り修正が難しく最後まで修正できない場合が多いことが分かった。よって、前段階のラベル付きクラスタ図作成を適切に行うことが重要であるといえる。

その他、アンケートにより、ラベル選択が難しいという課題が明らかとなった。そのため、ヘルプ内容の充実、ラベル選択時の説明表示など、ラベル理解・選択の支援が必要である。また、システムの操作性に関する指摘も確認された。操作性の問題はユーザの意欲低下に繋がるおそれがあるため、操作性の向上および操作方法を容易に理解できるような改良が必要である。

## 5 おわりに

本研究では、クラスタ図診断機能と整合性確認機能の評価および授業実践における課題の分析を行った。今後は、ユーザへのラベル理解支援の充実に加え、システムの操作性向上のための改良を行う必要がある。

## 参考文献

- [1] 山崎祥平、英語エッセイアウトライン作成支援のための整合性確認機能の有用性調査、2020年度九州工業大学修士論文、2021.

		指導教員	國近 秀信 准教授
学生番号	192C1123	氏名	福沢 唯心
論文題目	英語パラグラフの構成要素同定機能の改良 ～叙述パラグラフの種類同定～		

## 1 はじめに

英語初学者にとって、英語の論理展開法に従ったパラグラフライティングを行うことは難しい。これは結論が先に来るような形で展開される英語の論理展開法への理解が不十分であることなどが挙げられる。そこで、英語初学者が記述したいパラグラフの種類や関連したトピックの実例を示すことが支援方法の一つとして考えられるが、参考にするには適切な論理展開法に従ったパラグラフの数が少ないという問題がある。そのため先行研究 [1] では、WWW 上のパラグラフを利用し、入手したパラグラフの構造の適切性を判断するため、英語パラグラフの構成要素同定機能を実現した。本研究では、この構成要素同定機能の改良のため、叙述パラグラフの種類同定機能の改良を行う。

## 2 パラグラフの構成要素同定

パラグラフの検索、および、論理展開法に従っているか否かの判別のため、パラグラフの構成要素同定を行う。英語パラグラフにおいては、Topic Sentence や Concluding Sentence などの構成要素が存在するが、これはパラグラフの種類や目的に応じて異なる。そのため本研究では、パラグラフの種類ごとの典型的な構造を表したパラグラフ展開スキーマの構成要素を同定対象とする。この時、文の位置、各構成要素特有の頻出語とその品詞、パラグラフの種類等を利用する。

## 3 パラグラフの種類同定

先行研究 [2] では、定義、列挙、比較・対照、叙述の4種類のパラグラフを同定する機能を実現した。しかし、本機能の主な同定条件は、パラグラフの種類ごとの頻出語であり、性能が十分ではなかった。そのため本研究では、パラグラフライティングの参考書の説明やパラグラフをもとに、叙述パラグラフの特徴を調査し、改良することとした。

叙述パラグラフは、筆者の思い入れのある場所や人物を魅力的に記述するパラグラフである。同パラグラフの Topic Sentence では、トピックとそれがどんなものであるかという説明が出現するため、be 動詞が用いられる。その後の文章においては、人物の説明の場合はトピックおよびその代名詞を用いた文が連続し、場所の説明の場合は空間前置詞を用いた文が連続するという特徴がある。そこで本研究では、この特徴と実

際のパラグラフを参照した上で、以下の条件を定義・実装した。なお、構文解析および語句の参照関係同定のため CoreNLP[3] を利用した。

- (1) Topic Sentence 内において be 動詞が存在する。
- (2) 説明するトピックおよびその代名詞の Supporting Sentence 内での出現回数が Supporting Sentence の文章数に対して 120 % 以上である。
- (3) 空間前置詞および頻出語を含む文の Supporting Sentence 内での出現回数が Supporting Sentence の文章数に対して 80 % 以上である。

本システムは、(1) と (2)，もしくは (1) と (3) を同時に満たす場合、叙述パラグラフと同定する。

## 4 評価

さまざまな種類のパラグラフ計 110 個を用意し、叙述パラグラフの種類同定プログラムを適用をして適合率と再現率を求めた。本研究では学習者に対しパラグラフを提示するため、適合率を重視する。先行研究 [2] では適合率 11.8 %、再現率 33.5 % であったのに対し、本研究では適合率 41.1 %、再現率 70.0 % であった。よって、本研究により適合率の向上が確認できた。しかし、条件 (3) については、頻出表現を利用しておらず、叙述のパラグラフの構造的特徴を利用したものではないことから、条件の改善が必要だと考えられる。

## 5 おわりに

本研究では、パラグラフの構成要素同定機能の改良として、叙述パラグラフの種類同定機能の改良を行った。今後の課題は、叙述パラグラフの種類同定について、適合率をより向上させるとともに、他の種類のパラグラフについても同定が可能なように機能を拡張することである。

## 参考文献

- [1] 藤崎 凌太, 英語パラグラフの構成要素同定機能の改良～比較・対照パラグラフを中心～, 2019 年度九州工業大学卒業論文 (2020).
- [2] 濱村 鳩吾, 英語パラグラフの構成要素同定機能のためのパラグラフ種類同定, 2020 年度九州工業大学卒業論文 (2021).
- [3] Stanford NLP Group, "Stanford CoreNLP", <https://stanfordnlp.github.io/CoreNLP/> (参照日:2023.02.01).

		指導教員	國近 秀信 准教授
学生番号	192C1161	氏名	八尋 太嗣
論文題目	英語パラグラフの構成要素同定機能の改良 ～定義パラグラフの種類同定～		

## 1 はじめに

論理的かつ説得力のある英語のパラグラフを書くためには、英語の論理展開法に沿って書く必要があるが、英語初学者にとっては難しい作業である。この問題を解決するための方法の一つとして、適切な論理展開法に基づいたパラグラフを参照する方法がある。このとき参照するパラグラフは、著者が書きたい内容と構造に近いものが好ましい。Web サイトで検索すると望ましい内容のパラグラフは参照できるが、論理展開が適切か否かまでは判断できない。そのため先行研究では、構造を用いた英語パラグラフ検索システムの実現を目指し、英語パラグラフの構成要素同定機能を実現した[1]。本研究では、本機能の改良として、定義パラグラフの同定機能の改良を目的とする。

## 2 パラグラフの構成要素同定

英語パラグラフは”Introductory sentence”や”Topic sentence”などの要素から構成される。パラグラフは目的や種類によって構造が異なっており、その構成要素も異なる。本機能では、これら構成要素の場所と役割を同定し、タグを付与する。その際には、パラグラフの種類やキーワード、文脈、語句同士の関係性などを利用する。

## 3 定義パラグラフの同定機能の改良

先行研究[1]では、同定条件が緩いことにより、システムが誤って定義パラグラフ以外のものを定義パラグラフだと同定してしまったため、適合率が極端に低くなっていた。具体的には、同定条件として「パラグラフに”defined as”という述語、もしくは”名詞+be動詞+名詞句+関係代名詞”が出現した場合」を用いていた。

本研究では、より的確に定義パラグラフのみを同定するために、書籍や Web の情報を参照して定義パラグラフの特徴を調査し、以下の同定条件を考案・実装した。

条件 1：“defined as”，”define”，”definition”的いずれかが、パラグラフの 1, 2 文目に 1 回以上出現する。

条件 2：パラグラフ 1, 2 文目に Be 動詞が存在し、Be

動詞前後の語句のヘッド間に意味上の包含関係がある。

条件 3：パラグラフ 1, 2 文目に Be 動詞が存在し、Be 動詞前後の語句のヘッドの意味カテゴリが同一である。

本システムは、条件 1, 2, 3 のいずれかを満たす場合、定義パラグラフと同定する。ここで、条件 2 および 3 の意味カテゴリは、Wordnet[2] により求める。Wordnet に収録されていない単語については、その単語を英語版 Wikipedia[3] で検索し、その定義文から抽象概念を取得した上で、Wordnet を利用するよう実現した。

## 4 評価

システムが扱う全 11 種類のパラグラフを網羅した計 110 個のパラグラフを用意し、定義パラグラフの種類同定プログラムを適用して、適合率と再現率を求めた。先行研究[1]では適合率 8.8 %、再現率 66.7 %であったのに対し、本研究では適合率 40.0 %、再現率 60.0 % であった。本研究により、再現率を維持したまま適合率を上げることに成功した。

適切に判定できなかった場合の主な原因として、Wikipedia を利用して単語の定義文を取得する際の複数の意味を持つ単語の処理、および、意味上の包含関係を調べる際の一つ上位の意味カテゴリが複数存在する単語の処理が不十分であった点が挙げられる。

## 5 おわりに

本研究では、構成要素同定機能の改良として、定義パラグラフの同定機能の改良を行った。今後は、意味カテゴリの同定法の改良等により適合率をより向上させることが必要である。また、本システムの機能向上が、構成要素同定機能の精度向上にどれだけ影響したかを確認する予定である。

## 参考文献

- [1] 濱村 風吾, 英語パラグラフの構成要素同定機能のためのパラグラフ種類同定, 2020 年度九州工業大学卒業論文 (2021).
- [2] The Trustees of Princeton University, WordNet, <https://wordnet.princeton.edu/> (2023.02.10 参照) .
- [3] Wikimedia Foundation, Wikipedia, <https://www.wikipedia.org/> (2023.02.10 参照) .

		指導教員	坂本 比呂志 教授
学生番号	192C1058	氏名	古賀 秋音
論文題目	ニューラル機械翻訳のための LCP-Dropout の改良		

## 1 はじめに

サブワード分割は、機械翻訳における学習の前処理として一般的に用いられている手法である。形態素解析と比較して、日本語のように形態素解析が必要な言語ばかりではなく、英語などワード単位で分かち書きされている言語も存在することから、単語（トークン）単位で学習する従来手法よりも、さらに適切な分解をすることで翻訳精度がよくなることが知られている [1]。しかし、低資源言語に対して学習を行うとき、従来手法ではデータ拡張の観点から適していなかった。

その問題に対するアプローチとして、先行研究では文法圧縮手法の 1 つである Locally Consistent Parsing (LCP) を応用し、確率的に異なる分割パターンを生成することができる LCP-dropout を導入した。それにより、翻訳性能が向上した [2]。

本研究では、LCP-dropout のアルゴリズムを改良することにより、その有効性を検証するとともに、実装方法を改良し、計算時間の高速化を達成した。

## 2 提案手法

本手法では、LCP-dropout の符号割り振りを、以下のアルゴリズムに従った符号割り振りに改良した。

1. 文書中のアルファベット  $\Sigma$  について、「0」の符号が与えられたアルファベットを  $\Sigma_0$ 、「1」の符号が与えられたアルファベットを  $\Sigma_1$  とし、 $\Sigma_0 = \Sigma_1 = \emptyset$  として初期化する。
2.  $\Sigma$  からランダムに開始文字  $s$  を選択し、 $s \in \Sigma_0$  とする。
3. 右隣の文字  $x$  について、
  - (a)  $(x\Sigma_0 \text{ と } \Sigma_0 x \text{ の頻度}) \geq (x\Sigma_1 \text{ と } \Sigma_1 x \text{ の頻度})$  のとき、 $x \in \Sigma_1$  とする。
  - (b) それ以外のとき、 $x \in \Sigma_1$  とする。
4. ラベル列が「01」となる 2-gram を全て探索する。
5. 出現頻度上位  $k \in (0, 1)$  の 2-gram に対して、任意の確率で dropout を行い、2-gram を結合する。
6. 語彙数  $l$  を満たすまで、3. から 5. の手順を繰り返す。

また、文書の長さ  $w$  とすると、上記のアルゴリズムは時間計算量が  $O(|\Sigma||w|^2)$  となる。それに対してデータ構造を用いることにより、 $O(|w|^2)$  となるよう改良し、計算時間の高速化を行った。

## 3 実験

本実験では Transformer を用いて学習を行い、BLUE スコアを用いて定量的な評価を行った。加えて、計算時間の計測と評価を行った。

表 1: BLUE スコアによる性能評価 (%)

手法	BLUE
LCP-dropout (dropout 率 0.1)	35.8
LCP-dropout (dropout なし)	<b>36.6</b>
LCP-dropout (先行研究)	36.4

実験結果を表 1 に示す。先行研究と比較して、dropout なしの場合、スコアがわずかに優れた結果となったが、dropout 率 0.1 の場合は翻訳性能が低下した。

表 2: 計算時間の比較

	Time(s)
Before	67179.4
After	<b>13999.7</b>

次に、計算時間について表 2 に示す。計算時間に関して、約 4.8 倍の高速化に成功した。

## 4 まとめ

本研究では、LCP-dropout のアルゴリズムを改良し、翻訳性能について評価を行うとともに、時間計算量を縮小した。低資源言語に対する性能評価や、アルファベットサイズの大きい言語への適用が今後の課題である。

## 参考文献

- [1] Rico Sennrich *et al.*, Neural machine translation of rare words with subword units, *ACL*, pp. 1715-1725, 2016.
- [2] Keita Nonaka *et al.*, A Compression-Based Multiple Subword Segmentation for Neural Machine Translation, *Electronics* 2022, 11(7), 1014.

		指導教員	坂本 比呂志 教授
学生番号	192C1086	氏名	高野 耕大
論文題目	周波数領域を考慮した TransGAN による画像生成		

## 1 はじめに

画像生成の分野では、高品質な画像を生成するモデルが研究されている。関連研究としてこれまでに TransGAN[1] と周波数領域を考慮した画像生成 [2] が提案されている。

[1] は Transformer と GAN を組み合わせたモデルである。従来の CNN を用いたモデルと比べて高品質な画像を生成している。また、[2] は訓練画像に着目した画像生成手法である。学習初期では高周波を削減した画像を与える、訓練が進むにつれ高周波を含む画像で訓練することで高品質な画像の生成に成功した。

本研究では、この 2 つの手法がモデル内部と訓練画像の互いに影響し合わない点に着目し、TransGAN の訓練画像に周波数領域を考慮した訓練画像を与えることで、より高品質な画像を生成する手法を提案する。

## 2 提案手法

提案手法ではモデルに TransGAN を使用し、訓練画像に DCT 係数を用いて周波数領域を考慮した画像を与える。具体的に以下の 2 つの手法を提案する。

1. 高周波を削減した低周波を含む訓練画像を与える、学習が進むにつれ段階的に高周波を含む訓練画像に変化させる。
2. 低周波を削減した高周波を含む訓練画像を与える、学習が進むにつれ段階的に低周波を含む訓練画像に変化させる。

[2] より内部モデルを変更せず訓練画像に段階的な周波数削減を適用することで高品質な画像を生成できることが示されている。そのため、TransGAN に周波数領域を考慮した画像を与えることにより、高品質な画像が生成されることが見込める。

## 3 実験結果

モデルを CIFAR-10 で訓練した場合について、表 1 に実験結果、図 1 に訓練画像、図 2, 3, 4 に生成画像を示す。生成画像の識別のしやすさとクラスの多様さを示す IS は高い程性能が良く、生成画像と訓練画像の分布間の距離を示す FID は低い程性能が良い。

表 1 より提案手法 1 に比べ提案手法 2 の性能が上回ったが、その一方で既存手法に対しては、どちらの提案手法も IS, FID ともに劣っている。

表 1: 実験結果 (CIFAR-10)

	既存手法	提案手法 1	提案手法 2
IS ↑	5.97 – 6.90	3.86 – 5.10	5.26 – 5.50
FID ↓	41.4 – 59.7	71.2 – 92.0	67.2 – 73.8

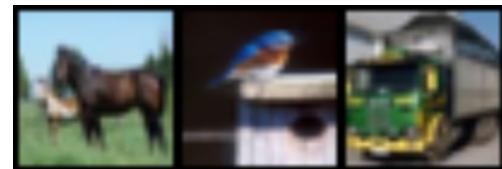


図 1: 訓練画像

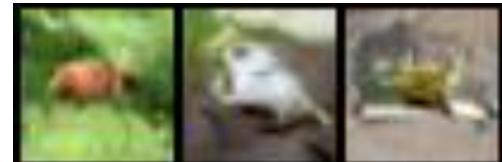


図 2: 既存手法による画像生成



図 3: 提案手法 1 による画像生成



図 4: 提案手法 2 による画像生成

## 4 終わりに

低周波を含む画像より高周波を含む画像から訓練することで、高品質な画像を生成できることが確認できた。一方で、どちらの提案手法も既存手法には劣っているため訓練画像とモデルの改善が今後の課題である。

## 参考文献

- [1] Yifan Jiang, et al. "TransGAN: Two Pure Transformers Can Make One Strong GAN, and That Can Scale Up", NeurIPS 2021(Poster).
- [2] 山之内 一隆 (2020), "周波数領域を考慮した GAN による画像生成" 九州工業大学情報工学部 知能情報工学科 卒業論文.

		指導教員	坂本 比呂志 教授
学生番号	192C1099	氏名	友利 拓誠
論文題目	構造的類似性指標に基づく伊勢型紙画像の分類		

## 1 はじめに

本研究室では、古典的手法を用いた伊勢型紙画像の分類自動化に取り組んでいる。伊勢型紙は背景である「地」と、地の上に描かれる植物や動物などの「図」によって構成されている伝統工芸品の一種である。型紙は、複雑な模様や抽象的な模様のものが多く、かつデータの総数が比較的少ない等の理由により、既存の手法や機械学習による分類は難しいことが知られている。

分類問題の一例として、型紙が地のみで構成されているか、地と図で構成されているかを分類する問題がある。本研究では、この問題を簡略化し、入力画像が規則的な模様か、あるいは不規則な模様かどうかの分類を行う。パッチ間の類似度を測る指標として構造的類似性指標を用いることで、型紙画像内の構造の揺れを標準偏差の値で表し、それを閾値に従い分類する。上記の手法により、型紙が規則的な模様か否かの分類を行うことを目的とする。

## 2 提案手法

### 2.1 構造的類似性指標

構造的類似性指標 (Structural Similarity Index, SSIM) は、画像間の類似度を測定するために使用される指標である [1]。この指標は、画像の全体的な品質を評価するために、画像の輝度、コントラスト、構造に基づく特徴を考慮して算出される。最終的に得られる値は  $[0, 1]$  上の実数値で表され、1 に近いほど画像間の類似度が高いとされる。

### 2.2 アルゴリズム

提案アルゴリズムを以下に示す。

1. ハイパーパラメータとして閾値とパッチサイズ  $n$  を定める。
2. 型紙画像に対して、リサイズ、平滑化、画像拡張処理を前処理として施す。
3. 型紙画像を  $n \times n$  のサイズで分割する。
4. 各パッチに対して、そのパッチを 1 ピクセルずつ横にスライドしながら切り出した  $n \times n$  の画像との類似度を算出する。
5. 各パッチごとに 4 で算出した類似度の最大値を求める。

6. 5 で得られた複数の最大値から標準偏差を算出し、閾値に従って分類する。

## 3 実験結果

実験に使用した型紙画像を並べた例を図 1 示す。しかし、左側は規則的な模様の型紙画像であり、右側は不規則な模様の型紙画像である。2.2 で示したアルゴリズムの 6 の工程で算出した標準偏差の値は、左側の画像が 0.02914 であり、右側の画像が 0.1945 であった。

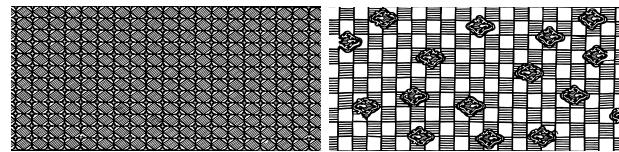


図 1: 実験に使用した入力画像の例 ((株)キヨーテック)

パッチサイズを 64 に設定し、閾値を 0.1 に設定した際に得られた混合行列を表 1 に示す。分類結果における F 値として、約 0.86 ほどの値が得られた。

表 1: 提案手法によって得られた混合行列

		予測	
		規則的	不規則
正解	規則的	<b>105</b>	22
	不規則	11	<b>73</b>

## 4 おわりに

本研究では、構造的類似性指標に基づき、型紙画像が規則的な模様か、あるいは不規則な模様かの分類を行なった。提案手法により、偏りなく一定の分類精度を得られたが、図が小さい型紙や、図が一定の間隔幅で出現するような型紙画像が誤分類となる傾向が確認できた。これらの模様を持つ型紙も正しく分類し、分類精度の向上を図ることが今後の課題である。

## 参考文献

- [1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol.13, no.4, pp. 600-612, 2004.

		指導教員	坂本 比呂志 教授
学生番号	192C1108	氏名	永尾 弥人
論文題目	Vision Transformer を用いた伊勢型紙画像の分類		

## 1 はじめに

本研究室では伊勢型紙の機械学習を用いた分類に取り組んでいる。伊勢型紙は背景である「地」と動物や植物など地の上に描かれる「図」から成っている。地として描かれる図形は抽象的なものが多い点や機械学習を行うには画像枚数が少ない点などの理由により分類は難しい問題であることが分かっている。訓練画像に対して処理を加え、画像の枚数を増やすことができればより良い学習が可能になると考えた。

分類に際しては Vision Transformer を用いる。本研究では画像枚数が少ない問題を解消するべく、訓練画像に対し分割処理や反転処理を行うことで画像枚数の水増しを行い、分類精度の向上を目的とする。

## 2 Vision Transformer

Vision Transformer[1] は画像のパッチを単語のように扱うことで自然言語処理分野で活躍していた Transformer を画像処理分野に応用したものである。従来の画像処理分野においては CNN をベースとした畠み込みを用いたアーキテクチャが主流となっていたが、Vision Transformer では畠み込みを行わず、計算量を大幅に削減できるという特徴がある。

## 3 提案手法

学習の際に訓練画像として元画像のみ、画像を 4 分割したもの、9 分割したもの、4 分割と 9 分割、反転させたものそれぞれを追加した計 5 つのパターンを Vision Transformer に読みませた場合の精度を比較する。4 分割を行った画像の例を図 1 に示す。

モデルとして事前学習された ViT モデルを用いファインチューニングによって画像を分類する。また、学習用画像に対しては必要に応じて一部切り出し、左右反転などのデータ拡張を行う。

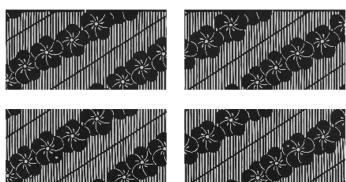


図 1: 4 分割処理を行った画像群 (KTS00803 (株) キョーテック)

## 4 実験結果

提案手法において 50 エポック学習した場合の損失値と分類精度を表 1 に示す。

表 1: Vision Transformer による分類精度

	損失値	分類精度
元画像	0.788	71.37%
4 分割	0.795	72.09%
9 分割	0.753	73.42%
4 分割+9 分割	0.810	72.26%
反転	0.707	74.38%

表 1 より画像を画像の分割、反転処理のすべてにおいて元画像以上の精度を達成できていることが分かる。特に反転画像においては損失値で約 0.08、分類精度で約 3% の向上を達成しており、明確な分類および分類精度の向上に寄与していると考えられる。

## 5 おわりに

本研究では Vision Transformer を用いた型紙画像の分類、および分類精度向上のための画像枚数の水増し処理を行った。画像に対し処理を行った場合分類精度が向上することが確認できた。適切なデータ拡張の選択や誤分類となった画像についても適切に分類し、精度向上を図ることが今後の課題である。また、本研究ではファインチューニングによる精度の高い型紙画像の分類を行うことは難しいことも確認できた。提案手法に加え地のみや図のみの画像を生成する手法 [2][3] 等を用い、型紙画像のみで大規模なデータセットを作成することができれば型紙画像分類における高精度のモデルが作成可能になると考えられる。

## 参考文献

- [1] Alexey Dosovitskiy, et al., An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, arXiv,2010.11929, 2020
- [2] 濱田 康弘 (2020), PCA を用いた伊勢型紙画像の図と地の分離、九州工業大学情報工学部 知能情報工学科卒業論文
- [3] 鬼木 梨帆 (2021), PCA を用いた伊勢型紙画像からの図の抽出、九州工業大学情報工学部 知能情報工学科卒業論文

		指導教員	坂本 比呂志 教授
学生番号	192C2113	氏名	脇山 巧基
論文題目	完全準同型暗号を用いた Grouping-CWC の秘匿計算		

## 1 はじめに

近年、蓄積したデータを機械学習手法を用いて解析し活用するデータ解析が積極的に行われている。一方で、積極的なデータの利活用に伴い、データに含まれる機密性の高い情報が流出するリスクも高まっている。たとえデータの通信・保管時に暗号化したとしても、解析時には一度復号する必要があるからである。そのような問題の解決策として、データを暗号化したまま計算を行う秘匿計算を用いたデータ分析の研究が行われている[1]。

本研究では、完全準同型暗号の一種である TFHE 方式[2]で暗号化されたデータに対して特徴選択アルゴリズム Grouping-CWC[3]（以下 gCWC）を適用する手法を提案する。従来手法[4]では参加者間で通信が必要だったが、本手法は通信することなく実行可能である。提案手法を実装し、アルゴリズムの評価と改善すべき点を考察した。

## 2 Grouping-CWC と完全準同型暗号

gCWC は、データを特徴ベクトルの接頭辞でグループ分けし、そのグループ分けをもとにクラス分類に有効な特徴を選択する特徴選択手法である。特徴選択を行うことで機械学習モデルの学習時間を短縮するなどの効果がある。

完全準同型暗号は暗号状態のままで加算・乗算を無制限に行える暗号方式である。TFHE 方式ではデータを bit ごとに暗号化し、任意の回数の bit 論理和・bit 論理積を行えるため、任意の回路を秘匿計算できる。

## 3 提案手法

提案手法ではデータにラベル付けを行うことでグループ分けを行い、特徴選択を可能にする。

まず、先頭の特徴から順に次の操作を行いデータにラベル付けをする。

- 直前のラベルを多項式を用いて読み取り、そのラベルに対応する構造体を取得する。
- 構造体とデータの値からデータに付けるラベルを決定する。
- すべてのデータのラベル付けが完了すると、直前の特徴のラベルと比較する。この時、すべて

のデータについてラベルが同じであれば、ラベル付けした特徴を選択から除外する。

その後、末尾の特徴から順に各特徴  $F_j$  ごとに次の操作を行い、クラス分けに有意な特徴を選択する。

- データごとに  $F_{j-1}$  と  $F_{j+1}$  のラベルを結合する。
- 結合させたラベルが同じでクラスラベルが異なるデータがほかに存在するかを調べる。そのようなデータがなければ  $s$ ,  $F_j$  を選択から除外する。
- 再びデータにラベル付けを行う。

終了時に残された特徴が選択結果である。計算量は  $O(mn^2 \log n)$  bit である。

## 4 実験結果

データ数  $n$ , 特徴数  $m$  における秘匿計算 gCWC の実行時間を計測した（表 1）。

表 1: 実行時間の変動 [sec]

$n \setminus m$	4	8	16	32
4	108.93	256.74	476.64	929.52
8	614.64	1086.43	2011.80	4062.16
16	2611.45	4354.30	9136.82	20220.30

## 5 終わりに

暗号化されたデータに対して特徴選択アルゴリズム gCWC を実装した。今後の課題として、ラベルの処理方法の改善・並列化での実装・アルゴリズムの多値多クラスデータへの拡張が挙げられる。

## 参考文献

- [1] 佐藤 宏樹, 馬屋原 昂, 石巻 優, 今林 広樹, 山名 早人, 完全準同型暗号のデータマイニングへの利用に関する研究動向, FIT2016 (第 15 回情報科学技術フォーラム), 2016.
- [2] I. Chillotti, N. Gama, M. Georgieva, and M. Izabach'ene., TFHE: Fast Fully Homomorphic Encryption over the Torus, CCGI19, pages 34–91, 2020.
- [3] 中島 紳輔, 特徴選択アルゴリズム CWC の高速化, 九州工業大学, 2021.
- [4] S. Ono, J. Takata, M. Kataoka, T. I, K. Shin, H. Sakamoto, Privacy-Preserving Feature Selection with Fully Homomorphic Encryption., Algorithms 15(7): 229 (2022),

		指導教員	大北 剛 淄教授
学生番号	192C1045	氏名	加藤 舜斗
論文題目	ローカルな画像パッチを用いた 血腫画像に対するマルチインスタンス学習		

## 1 はじめに

本研究は、脳血腫マーカーの一つであるハイポデンシティの画像認識タスクを解く。このタスクは、単純な CNN を用いた画像認識では精度が 0.6 に達せず認識が困難なタスクである。一つのやり方は、CNN とセマンティックセグメンテーションをジョイントで学習させ、精度を上げるという方法であった。本研究では、別の手法、マルチインスタンス学習 (MIL) という弱いラベル情報を仮定して行う学習法を探求する。

## 2 本研究における提案手法

本研究の問題に MIL を適用する際、脳血腫 CT 画像（全体画像）、脳血腫 CT 画像をグリッド状に分割した画像パッチ（ローカルな画像）という 2 種類の画像を考慮する。全体画像にはラベルが存在するが、ローカルな画像にはラベルが存在しないという設定である。つまり、このように全体画像とローカル画像を構成すると、MIL でのバグ集合とインスタンス（データ）がうまく表現できる。今回利用したモデルに対するデータローダーではデータセットのバグ内のインスタンス数が 20 を大きく超えるとデータセットが不安定となった。なぜ、不安定となるか検証した結果、MILにおいては、バグ内に一つでも陽性のインスタンスが存在するとバグとしてのラベルが陽性となり、バグ内のインスタンス数が大きくなればなるほど、陽性のインスタンスが含まれる可能性が高くなってしまうためであることがわかった。しかし、本研究の設定では 256 個のインスタンスでも、CT 画像ごとにバグを作成しているため陽性のインスタンスがバグに含まれる確率が上昇することはない。MIL の多くの文献では、このようなグリッド状に分割した画像パッチを利用する形ではないため、これは本研究の貢献と考える。さらに、このように設定したインスタンスを用いて、ローカルな分類器を解こうとすると、ラベルが未定のローカルな分類器が 256 個存在することになる。通常のマルチインスタンス学習器での最適化アルゴリズムでは複雑度が大きすぎるという問題がある。このため、半教師あり学習の設定として、ラベルが未定のローカルな分類器ではなく、ラベルを貼ったローカルな分類器として半教師あり的に解かせた。この点も本研究の貢献である。以上から、本研究の提案手法は Algorithm1 となる。

---

### Algorithm 1 ラベルの個数情報を用いた MIL

---

- 1: CNN を用いてローカルにラベル付けしたパッチに対して血腫か否かについて判別する分類器を作成
  - 2: データセットに対して MIL を行い、attention を含む特徴量の表現を求める
  - 3: データセットの各パッチを 1 で作成した分類器にかけて、各バグにおける陽性ラベルのパッチの個数情報、陰性ラベルのパッチの個数情報の表現を求める
  - 4: 2, 3 でそれぞれ求めた表現を結合して分類を行う
- 

## 3 実験

実験では、全体画像の枚数として書くと、訓練データを 383 枚、検証データを 48 枚、テストデータを 48 枚用了いた。ローカル画像としては、全体画像を  $16 \times 16$  の 256 枚に分割した画像となるため、それぞれの枚数の 256 倍となる。マルチインスタンス学習のため、エポック数を 20, 40, 60, 100, 200 と設定して、それぞれ学習を行った。検証データに対するスコアが良いモデルを取り出し、テストデータに対する分類を行った。結果を表 1 に示す。accuracy, f1score はそれぞれ 0.9166, 0.9130 となり、CNN を用いた場合に比べて 0.3 ほど高い。また、AUC の値は 0.92 であった。

表 1: epochs=40 のモデルによる分類結果

test accuracy	0.9166
test f1score	0.9130
AUC	0.92

## 4 おわりに

本研究では、脳血腫マーカーのハイポデンシティの画像認識タスクを、半教師ありマルチインスタンス学習の設定で解いた。今後の課題としては以下の 3 点が挙げられる。一つ目に、データセットの拡張を行ったときの評価実験を行いたい。二つ目に、他の急成長型の血腫に対する評価実験を行いたい。三つ目に、マルチラベルに対応させた上で評価実験を行いたい。

## 参考文献

- [1] 加藤 舜斗, 有村 公一, 飯原 弘二, 大北 剛. 半教師ありマルチインスタンス学習による脳血腫マーカーの分類. 第 106 回 MBL・第 77 回 UBI 合同研究発表会

		指導教員	大北 剛 淮教授
学生番号	192C1051	氏名	河津 水紀
論文題目	テスト時にリアルタイムビデオフレーム補間を実現する手法のマルチフレームの入力への拡張		

## 1 はじめに

フレーム補間はフレームレートの向上、医療、アニメーションの中割生成、スロー動画の作成、データ圧縮など様々な用途で研究が進められている [1]。身近な所ではテレビにおいてもフレーム補間は行われている。近年はテレビ、モニターにフレーム補間が機能として搭載され、デフォルトで設定されていることもある。アニメーションは制作にかかるコストが非常に高く、フレームレートが低くなりフレーム間の構造の変化が大きくなりやすい。また色調が単調でありテクスチャに乏しいという特徴もあり、アニメーションのフレーム補間は実写動画と比較して困難である。

### 1.1 フレーム補間

フレーム補間とは連続するフレームによって構成される動画に対して、新しいフレームを生成して内挿することである。この内挿フレームは通常その前後の 2 枚のフレームから生成されるが、マルチフレームという 3 枚以上のフレームを入力として用いるアプローチもあり、フレーム間の複雑な動きを捉えることができる。フレーム補間には、オプティカルフローという画素の動きを表すベクトルを用いて、前後の画像を変形して合成するという手法があり、精度のよいオプティカルフロー推定手法が必要となっている。フレーム補間の精度は向上してきているが、生成フレームにはブレや歪みが発生していることも多く、特にアニメーションでは顕著である。フレーム補間のために様々なアルゴリズムが提案されてきたが、その中の RIFE[2] はリアルタイムビデオフレーム補間に向けて提案されたオプティカルフロー推定アルゴリズムであり、他の手法と比較して高速で品質の高いフレーム補間を可能としている。本手法では RIFE に着目し、マルチフレームの入力に拡張する手法を提案する。

## 2 マルチフレームの入力への拡張の提案

### 2.1 RIFE の構成

RIFE は IFNet を用いてフローを推定し、内挿フレームを合成する。IFNet は、3 つの IFBlock からなっており、複数の IFBlock を通して段階的にオプティカルフローの解像度を上げて推定するアルゴリズムとなっている。また IFBlock は、一つ前の IFBlock の出力を入力にすることでフローやマスクの解像度を上げてい

る。IFBlock は、変形された画像とフローとマスクを入力にとり、畳み込みとアップサンプリングによって解像度を上げたフローとマスクを出力しており、このフローをもとに入力の変形画像を再度変形する。この最後の IFBlock で変形された画像とマスクをもとに内挿フレームを合成している。

### 2.2 マルチフレームの入力に拡張する方法

RIFE をマルチフレームの入力に対応させる方法として、複数の IFNet を用いて 4 枚のフレームを 2 枚ずつ処理することを繰り返している。入力フレームを  $I_0, I_1, I_2, I_3$  として、 $I_0$  と  $I_1, I_1$  と  $I_2, I_2$  と  $I_3$  の 2 枚ずつのフレームの組でそれぞれ IFNet に入力し、その出力を 2 枚ずつ IFNet に入力することを繰り返して全体で 4 枚のフレームを処理している。また  $I_0$  と  $I_1, I_1$  と  $I_2, I_2$  と  $I_3$  の 3 組から  $I_{0.5}, I_{1.5}, I_{2.5}$  の 3 枚のフレームを合成し、 $I_{0.5}, I_{1.5}, I_{1.5}, I_{2.5}$  の 2 組から  $I'_1, I'_2$  の 2 枚のフレームを合成する。最終的にこの  $I'_1, I'_2$  を用いて  $I'_{1.5}$  を合成する。通常 IFNet は、IFBlock で計算したフローで画像を変形し、その変形した画像を次の IFBlock に入力しているが、提案する手法では  $I'_1, I'_2$  と同時に  $I_1, I_2$  を変形しており、 $I'_1, I'_2$  の変形画像を IFBlock の入力としてフローの推定に用い、 $I_1, I_2$  の変形画像は内挿フレームの合成に利用している。

## 3 おわりに

本論文では、テスト時にリアルタイムビデオフレーム補間を実現する手法である RIFE のマルチフレームの入力への拡張を行う手法の提案を行った。提案した手法では RIFE をマルチフレームの入力に対応させることができた。一方で、IFNet を複数用いる本手法は計算時間の増大を招いた。また本論文での実験では、オリジナルの RIFE の学習済モデルを用いたが、マルチフレームの入力に対応したデータセットで新たに学習することが必要だと思われる。

## 参考文献

- [1] 河津, 大北.:2D アニメーションのフレーム補間に関する  
サーベイ, 第 47 回 IBISML 研究会, pp.19-24, 2022.
- [2] Huang, Zhewei and Zhang, Tianyuan and Heng, Wen and Shi, Boxin and Zhou, Shuchang.:Real-Time Intermediate Flow Estimation for Video Frame Interpolation, In Proceedings of the European Conference on Computer Vision (ECCV), 2022.

		指導教員	大北 剛 淄教授
学生番号	192C1102	氏名	中島 崇晴
論文題目	複雑な生成機構を持った敵対的生成ネットワーク		

## 1 はじめに

日本の漫画やアニメは国内のみならず世界において広く普及する文化の一つとなった。マンガ家の制作過程に直接関係するような、キャラクタの考案過程は非常に興味深い。このキャラクタの考案過程について、深層学習で生成することはできないかということを考えたい。本論文では、キャラクタの考案過程を実現することを目的とする。人物の実写画像をその人の特徴を残したままある特定の漫画あるいはアニメの画風を取り入れ変換させるような技術の実現をさせたい。人間と、漫画やアニメのようなキャラクターを別々のものとして区別する理由は、たとえば、漫画やアニメのようなキャラクターの目は人間の目とは、かなりの差があることを原因とするからである。

## 2 GAN の技術

以上の目的を達成する基本技術としては GAN を使用する。GAN の中にはそういった何らかの画風を取り入れ実写顔画像を編集するものが多数存在する [4]。近年、GAN の成長は著しく生成画像の精度を上げている。一般に、写真を生成するマシンは多数開発されており、このような中で対象をアニメーションなどのキャラクタ/線画をターゲットにする。styleGAN-NADA[1], jojoGAN[2], Blend GAN[3] などが提案されている。これらにはいくつかの問題点が指摘されている。1つ目に、漫画やアニメのようなキャラクターは人間のものと乖離していることが多いという点である。特に目である。2つ目にアニメの非現実的な顔の形や目の形を人に取り入れるのが難しいという点が挙がっている。3つ目に、顔が正面を向いてないといけないなどの顔の角度の問題も存在する。本研究では、これらの問題点の改善に取り組みたいと考えている。

## 3 スタイルトランスファーの人間の評価手法

本研究で取り組む内容はスタイルトランスファーと呼ばれている。画像を違う画像の様相に変換する。人物画像、ポートレート画像を中心としたものに焦点を置く。

スタイルトランスファーの問題点として挙げられるのが評価指標を示すのが難しいことである。これは一般的の GAN の評価が難しいことと類するが、一方、もともと、写真の顔を入力して、アニメ顔を GAN で出力し

た場合に、このアニメ顔の評価が難しいという根本的な問題に関わることもある。このような状況のために人間の評価指標を設定した。

本研究ではまずこの評価指標を設定したうえで関連研究の手法によって生成された画像を使い、その評価指標を使用し設定した評価指標の有用性を試す。関連研究では JoJoGAN と Blend GAN を動かし、生成された画像に対して評価を行う。その際変換するものは人の顔の構造に近い絵画と人の顔の構造に少し遠いアニメの絵を用いて行う。つまり、実際に出力した結果を見て、理想的な出力にするにはどのような評価指標を定めたら良いかについて考察した。

## 4 データセット

われわれのターゲットは、アニメ顔を生成したいという目的があるため、顔に関するデータを独自に集めた。Pinterest に掲載されている有名人の絵と、これに類似した有名人の写真を収集しペア画像データセットとして用意した。

## 5 終わりに

われわれの設定した評価指標はいずれも人の絵画に対するスタイルトランスファーは評価が高いという結果を得たが、アニメの絵になると点数は低くなるという結果を得た。これはいずれも実写の画像を学習に用いていることが原因だと考える。アニメの絵を用いた学習を行うことによって評価の点数は高くなると考える。その際に必要なのがデータセットの収集が問題となる。著作権の問題等もあり、そういったところも解決しなくてはならない。

## 参考文献

- [1] Gal, R., Patashnik, O., Maron, H., Chechik, G., Cohen-Or, D.: Stylegan-nada: Clip-guided domain adaptation of image generators 2021
- [2] Chong, Min Jin, and David Forsyth. JoJoGAN: One Shot Face Styling. arXiv preprint arXiv2022
- [3] Liu, M., Li, Q., Qin, Z., Zhang, G., Wan, P., Zheng, W.: Blendgan: Implicitly gan blending for arbitrary stylized face generation. In: Advances in Neural Information Processing Systems 2
- [4] 中島崇晴, 大北 剛, 人物実写画を特定の漫画の画風に沿うように変換し画像を出力する技術のサーベイ, 第 47 回 IBISML 研究会, pp.13-18, 2022.

		指導教員	大北 剛 淄教授
学生番号	21221003	氏名	大野 友暉
論文題目	Structure from Motion: 理想の動きを可視化する 3D 人体モデル作成		

## 1 自由視点からの 3D の人の動きを作成する技術と NeRF の技術

「自由視点からの 3D の人の動きを作成する技術」はここ 10 年という間に急速に発展した技術である。写真をベースとして姿勢(ポーズ)推定を行う深層学習の技術をそもそもの発端とする。しばらくして、3D の姿勢(ポーズ)推定を行う技術へと発展する。さらに、シェイプの推定へと進んだ。その後、のっぺら棒のシェイプに対して、テクスチャを貼り付けることにより、本物らしく見せる技術が誕生した。それぞれのステップをまとめることにより、「自由視点からの 3D の人の動きを作成する技術」が完成した。つまり、この「自由視点からの 3D の人の動きを作成する技術」は、以下の 4 ステップからなる技術を用いて実現可能ということになる。1 ステップ目は、入力したポーズからシェイプを作成する。2 ステップ目は、シェイプから法線マップを作り、法線マップからテクスチャマップを予測する。3 ステップ目は、予想したテクスチャマップをシェイプに貼り付ける。4 ステップ目は、シェイプを標準空間に変形する、というプロセスである。

この技術は、さらに、ポーズとカメラの位置、方向、シェイプを入力とする NeRF[1] と呼ばれるレンダリング技術と結び付いた。この技術によりライティングを導入して、より現実的な描写を行うレンダリングを可能とする技術へと発展した。

## 2 3D 人体モデルを使った人の任意の動きの生成

動きとは、人が歩く動作、人が転ぶ動作などを指す。ただし、単純な動きを行う 3D 人体モデルを作成するために正確な姿勢推定やシェイプ推定を行う必要があり実際に 3D 人体モデルを作成するには様々な問題がある。そこで、本研究では、3D 人体モデルに動作を行わせるためにどのような技術が必要となるかを考え、このような技術を開発することを目標とする。

まず、われわれが考えたのは、2 つの異なる動きを仮定した場合に、これら 2 つの動きを滑らかに接続することである。もしこのような接続を滑らかに行うことができれば、一般性を失なわずに、任意の動きをさせることができたと結論づけることができるからである。たとえば、人が歩く動作と人が走る動きを接続

して、「歩いていた人が走り出す」という動きを生成できたとすると、他の動きも同様の方法で滑らかに結合できるであろうからである。

## 3 シェイプの比較評価

3D 人体モデルを作成するに当たって 2 種類のシェイプ作成手法の比較評価を行った。図 1 左のシェイプは右のシェイプに対してなめらかなシェイプを作成する事が出来た。ただし、右のシェイプは服の形状や体形などの元の人物の形を忠実に再現している。

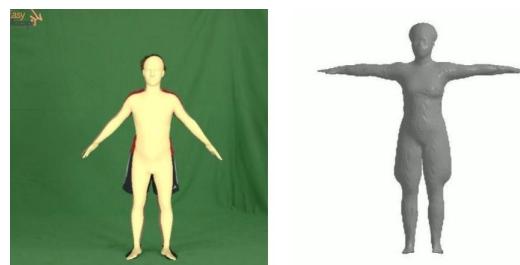


図 1: シェイプの比較評価

## 4 実装

実装には姿勢推定と 3D モデル作成のツールを用いた。ただし姿勢推定の出力データと 3D モデル作成に必要な入力データは異なるデータの変換を行った。姿勢推定ツールの出力データとして Rh, Th, poses, shapes を得る。3D モデル作成では translation, joints\_RT, rotation が必要である。translation は姿勢推定ツールの出力データ Th をそのまま使うことが出来る。Rh から回転行列を求め、その結果を転置する事で rotation を得る。joint\_RT は poses の回転行列を計算して joint と結合することで求める。joint はシェイプの頂点座標と ration との行列の積に translation を足すことで得られる。

## 参考文献

- [1] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision*, 2020.
- [2] 大野友暉, 大北剛. 動きからの構造: 2d 画像からの 3d テクスチャの再構築.

		指導教員	宮野 英次 教授
学生番号	192C1047	氏名	金海 匠馬
論文題目	積算温度によるイチジクの収穫時期予測		

## 1 研究背景

食料・農業・農村基本計画（令和 2 年 3 月 31 日閣議決定）において、データ駆動型の農業経営により消費者ニーズに的確に対応した価値を創造・提供する農業（FaaS（Farming as a Service））への変革を進めるため、デジタル技術を活用した様々なプロジェクトを「農業 DX 構想」として取りまとめることが計画された。その後、令和 3 年 3 月 25 日に、農林水産省は、農業・食関連産業の関係者が農業 DX を進める際の羅針盤として、また、取組全体を俯瞰する見取り図として活用できるように「農業 DX 構想」を取りまとめた[1]。農業 DX 構想の中では、少子高齢化や人口不足が進む中、農業従事者や消費者の間で様々な課題を抱えていることが指摘されており、今後さらに深刻化すると考えられている。現在の農業では、ロボット、AI、IoT など先端技術の導入により、作業の省力化・自動化・高度化を進めることができることが課題となっているが、それらを実際に活用した農業を行っている農業経営体の割合は全体の 2 割に満たない水準となっているとの指摘がある[1]。

## 2 本研究の内容と結果

本研究では、(1) 積算温度を算出するための気温データ、(2) 農業試験場で実際に植え付けから収穫までを行ったイチジクの生育状況が記録されたイチジクデータ、(3) 非定常的な特性を持つ時系列データをモデル化する自己回帰和分移動平均モデル（Auto regressive integrated moving average model, 以下では ARIMA モデル）[4] を用いて実験を行う。(1) 気温データは、国土交通省気象庁が保持している過去の気象データ[2] を用いる。使用する気温データは、2010 年から 2021 年のものとする。(2) イチジクデータは、本研究で収穫時期予測を行うにあたり、論文[3] で示された積算温度と収穫時期の整合性を確かめるための正解データとして用いる。使用する項目は、着果日と収穫日とする。(3) ARIMA モデルは、古典的な時系列モデルだが最も実用的かつ実績のあるモデルとされている。

本研究で行った研究の内容・結果を以下に示す。

### (1) 着果日から収穫日までの積算温度の分布

イチジクデータのすべての着果日と収穫日を取得し、その間の積算温度を算出、積算温度全体の標準偏差と

ヒストグラムによる分布の出力をを行う。データにはばらつきがあるものの、積算温度が 2170 °C付近で収穫されているものが多いことを示す。

### (2) 重み付き気温データによる収穫時期予測

日平均気温を取得した後に重みと気温の積を求め、積算温度・正解率の算出、実際の 2021 年の気温に重みをつけた積算温度全体の標準偏差とヒストグラムによる分布の出力をを行う。重み付与に際して、積算温度、重みの下限・上限の調整を行う。重み付与の方法については、気温を横軸、重みを縦軸として、イチジクの生育適温とされている 15 °C以上 30 °C未満の間の 22.5 °Cと重みの上限の値を頂点に上に凸となる二次関数を作成する。また、15 °Cと 30 °Cの時に重みの下限の値に達し、15 °C以下と 30 °C以上では、重みの下限の値で一定となるように設定する。上記の補正と収穫時期との差異を示す。

### (3) ARIMA モデルによる収穫時期予測

過去の気温データをトレーニングデータとして与え、収穫する年の気温を ARIMA モデルによって予測し、積算温度・正解率の算出を行う。また、中央移動平均の移動幅を 1 から 9 に設定したものと組み合わせた、ARIMA モデルによる気温遷移予測の実験も行う。結果はすべて正解率が 0%と精度が芳しいものではなかったが、中央移動平均の移動幅を大きくする毎に精度が上がることを示す。

## 参考文献

- [1] 農林水産省, 2021, 「農業 DX 構想～「農業×デジタル」で食と農の未来を切り拓く～」, 農林水産省ホームページ, (2023 年 2 月 12 日取得, <https://www.maff.go.jp/j/press/kanbo/joho/210325.html>)
- [2] 国土交通省気象庁, 2022, 国土交通省気象庁ホームページ, (2023 年 1 月 28 日取得, <https://www.data.jma.go.jp/gmd/risk/obssl/index.php>)
- [3] 矢羽田第次郎, 野方仁.「結果節位の異なるイチジク果実の発育および形質変動と気温との関係」, 園芸雑誌 (J. Japan. Soc. Hort. Sci) 69(2), pp.202–207, 2000
- [4] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, G.M. Ljung. Times series analysis: forecasting and control. 5th Ed., Wiley (2015)

		指導教員	宮野 英次 教授
学生番号	192C1065	氏名	是永 華奈
論文題目	変更制約付き活動選択問題に対するアルゴリズム		

## 1 研究背景

日々の生活において様々なスケジューリング問題が見られる [1]. 大学における講義時間割作成問題は、担当教員が同じ時間に講義を行わないようにするなどの制約条件を満たして、学生が履修する際に 1 日に受講する講義数を出来るだけ均一化することや、出来るだけ 4 限以内に多くの講義を開講することを目的とするなどのスケジューリング問題として捉えることが出来る。例えば、時間割作成において、昨年度の割当講義を出来るだけ利用しながら、より良い時間割を求めるような場合がある。本論文では、そのような初期解からの変更によるスケジューリング問題を扱う。

## 2 変更制約付き活動選択問題

本論文では、会議室の利用のように 1 つのリソース（会議室）に対して、ある同一時間帯には複数のタスク（利用者）が割り当てられることが無いという制約条件の下で、できるだけ多くのタスクを割り当てることを目標とした活動選択問題の派生問題について考える。活動選択問題の目的は、互いに両立可能な活動から構成される集合で、サイズが最大のものを求めることがある。従来の活動選択問題では、初期状態として何も活動が割り当てられないような「ゼロの状態」から出来るだけ多くの活動（区間）を割り当てる問題であり、貪欲法に基づく多項式時間アルゴリズムで最適な割り当てを求めることが出来る [1]。

本論文では、従来の活動選択問題に対し初期解が与えられた下での変更制約付き活動選択問題を考える。また、入力として (i) 1 日単位として活動区間が与えられる場合と (ii) 与えられる活動区間が円状に影響ある場合の、この 2 つの問題に対するアルゴリズムを設計する。

(i) 24 時間以内の活動選択を考えた場合、活動を時間軸上の直線として表現することが出来る。ここで、24 時間以内という時間制約がある活動選択を変更制約付き区間活動選択問題と呼ぶ。

(ii) 制限を設けずに、24 時以降のある活動と次の日の活動に重なりがある場合、活動区間が円状に影響を及ぼす。それゆえ、時間軸を円状に取り、それぞれの活動は円弧区間として表現することができ、時間制限がない活動選択を変更制約付き円弧活動選択問題と呼ぶ。

変更制約付き活動選択問題において、 $S^0$  及び  $S$  をそれぞれ初期独立集合とアルゴリズム解とする。また、変更した頂点数を示す関数  $f(S, S^0) = |S^0 \setminus S|$  と定義する。これは初期解  $S^0$  には含まれるが、アルゴリズム解  $S$  には含まれない頂点数を示す関数であり、初期解  $S^0$  から削除された頂点数を示す。目的解となる  $S$  は  $|S^0 \setminus S| \leq k$  を満たすように得られ、これは初期解  $S^0$  から取り除いた頂点数が入力された上限  $k$  以下を満たす [2]。以下に変更制約付き活動選択問題の定義を示す。

### 変更制約付き活動選択問題

入力：  $n$  個の区間集合  $I = \{I_1 \dots I_n\}$ , 初期独立区

間集合  $S^0$ , 整数  $k$

目的：  $|S^0 \setminus S| \leq k$  を満たす最大サイズの独立区間集合  $S^* \in S$

## 3 本論文の結果

本論文では変更制約付き活動選択問題に対して、それぞれ以下のアルゴリズムを示す。

1. 変更制約付き区間活動選択問題に対して、動的計画法に基づく  $O(kn)$  時間で最適な独立区間集合を求めるアルゴリズムを示す。
2. 変更制約付き円弧活動選択問題に対して、区間の場合の上記の  $O(kn)$  時間アルゴリズムをサブルーチンとして用いることにより、 $O(kn \times |S^0|)$  時間で最適な独立円弧区間集合を求めるアルゴリズムを示す。

## 参考文献

- [1] Fanica Gavril. Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph. *SIAM J. Comput.*, 1(2):180–187, 1972.
- [2] Onur Şeref, Ravindra K Ahuja, and James B Orlin. Incremental network optimization: Theory and algorithms. *Operations Research*, 57(3):586–594, 2009.

		指導教員	宮野 英次 教授
学生番号	192C1089	氏名	田中 駿一
論文題目	最長ラン部分文字列問題に対する 近似アルゴリズム		

ゲノム解析における主要な目的は、生命体のゲノムが持つ遺伝情報を比較、研究することである。そのため、ゲノムの完全な配列情報を決定するゲノム解読（genome sequencing）が主要な手続きとなっている。完全なゲノム情報が得られてから、ゲノム解読技術は革新的に進んできた。最新のDNA列の解読技術の多くは以下のような手続きで行われている。まず、DNA列のランダムな位置から数十億にもなる断片情報をショットガン解読（shotgun sequencing）により読み込む。次に、ランダムに取り出された断片情報（リード, read と呼ばれる）をコンティグ（contig）と呼ばれる連続列の集合を生成するために、組み立てアルゴリズム（assembly algorithm）を用いて結合する。その後、スキャフォールド（scaffolding）と呼ばれる段階で、コンティグを正しい順番に並べる。スキャフォールドにおける代表的な手法は、関連したサンプルの2つ以上の不完全に組み上げられた列を比較することでコンティグを整列するというものである [2]。

最近、スキャフォールド段階を最適化問題として定式化した最長ラン部分文字列問題（LRS）を、Schrinnerらが提案している [3, 4]。LRS は以下のように定義される：ランは、連続する同じ文字の最大部分文字列のことである。文字集合  $\Sigma$  上の文字列  $S = s_1 \dots s_n$  に対し、 $S$  のラン部分文字列  $S'$  は  $\sigma \in \Sigma$  であるすべての文字に対し、最大1つのランを含む部分文字列である。LRS 問題とは、文字列  $S$  が与えられたとき、 $S$  のすべてのラン部分文字列の中で最長となるラン部分文字列  $S^*$  を求めることである。例えば、文字集合  $\Sigma = \{a, b, c\}$  上の文字列  $S = abacacbbab$  を考える。 $a, b, c$  の出現数はそれぞれ4回、4回、2回である。また次のことが分かる。(i) 文字  $a$  についてランは、1文字目の  $a$ , 3文字目の  $a$ , 5文字目の  $a$ , 9文字目の  $a$ , この4つのランが存在する。(ii) 文字  $b$  についてランは、2文字目の  $b$ , 7, 8文字目の  $bb$ , 10文字目の  $b$ , この3つのランが存在する。(iii) 文字  $c$  についてランは、4文字目の  $c$ , 6文字目の  $c$ , この2つのランが存在する。

入力  $S$  の LRS 問題に対する最適解は  $S^* = aaccbbb$  である。例えば、 $S^*$  中の長さ2の左端のラン  $aa$  は  $S$  の左端の部分文字列  $aba$  から2番目の文字  $b$  を削除することによって得られる。 $S^*$  はラン部分文字列であるので、 $S^*$  は各文字最大1つのランを含んでいる。 $S^*$  の

長さは7であり、 $S' = aaacbbb$  も長さが7であるため、こちらも最適解となる。

LRS 問題は NP-困難 [4] であり、入力文字列の各文字の出現数が最大2回の時でさえ APX-困難であることが知られている [1]。LRS 問題には、入力文字列の各文字の出現数が  $k$  回以下の場合、 $\min\{|\Sigma|, k\}$ -近似アルゴリズムが存在する [1]。本論文では、LRS 問題に対する以下の改善した近似アルゴリズムを設計する。

- 各文字の出現数が最大  $k$  回に制限した場合の LRS 問題に対して、多項式時間で動作する  $\frac{k+1}{2}$ -近似アルゴリズムを示す。本アルゴリズムの近似精度保証  $\frac{k+1}{2}$  は最適解の長さの上界を従来よりも詳しく調べることにより与えている。また、この近似精度保証の解析が厳密であることを示す。すなわち、設計した  $\frac{k+1}{2}$ -近似アルゴリズムに対して、実際に  $\frac{k+1}{2}$  倍の長さを持つような入力例が存在することを示す。
- 上記の設計した  $\frac{k+1}{2}$ -近似アルゴリズムは、 $k = 2$  のとき、 $\frac{3}{2}$ -近似となる都合の悪い入力文字列が存在する。この都合の悪い例に対してより良い解を出力するようにアルゴリズムを完全することにより、多項式時間で動作する  $\frac{4}{3}$ -近似アルゴリズムを設計する。さらに、 $\frac{4}{3}$ -近似アルゴリズムに対して、実際に  $\frac{4}{3}$ -近似となってしまうような、解析にとってタイトな文字列が存在することを示す。

## 参考文献

- R. Dondi and F. Sikora. The longest run subsequence problem: further complexity results. In *CPM 2021*, pp.14:1–14:15 (2021)
- J. Luo, Y. Wei, M. Lyu, Z. Wu, X. Liu, H. Luo, and C. Yan. A comprehensive review of scaffolding methods in genome assembly. *Briefings in Bioinformatics*, 22(5) (2021)
- S. Schrinner, M. Goel, M. Wulfert, P. Spohr, K. Schneeberger, and G.W. Klau. The longest run subsequence problem. In *WABI 2020*, pp.6:1–6:13 (2020)
- S. Schrinner, M. Goel, M. Wulfert, P. Spohr, K. Schneeberger, and G.W. Klau. Using the longest run subsequence problem within homology-based scaffolding. *Algorithms for Molecular Biology*, 16(1):11 (2021)

		指導教員	宮野 英次 教授
学生番号	192C1140	氏名	松下 瑠花
論文題目	ハッピーセットゲームの勝利条件と引き分け条件		

## 1 研究背景

将棋や囲碁、チェスなどの二人で行う対戦型のゲームについて、その戦略に関する研究や、コンピュータゲームとの実装などが行われている [3]。特に、オセロ（またはリバーシ）、ニム（石取りゲーム）、三目並べ（○×ゲーム、Tic tac toe）は、二人零和有限確定完全情報ゲームと呼ばれている。先に述べたように、二人零和有限確定完全情報ゲームでは、勝ち・負け、または引き分けが確定的に、有限回で決定するために勝利条件と引き分け条件の観点からの研究が行われている。

本論文では、グラフ上で二人のプレイヤーで行う新しい二人零和有限確定完全情報ゲームの一つとしてハッピーセットゲームを提案し、勝利条件や引き分け条件について考える。

## 2 ハッピーセット問題とハッピーセットゲーム

文献 [1, 2]において、朝廣らにより最大ハッピーセット問題が提案されている。入力はグラフ  $G = (V, E)$  と整数  $k$  である。頂点  $v$  とそのすべての隣接頂点が頂点部分集合  $S \subseteq V$  に入るとき、 $v$  をハッピーと呼ぶ。最大ハッピーセット問題の目的は、ハッピー頂点の数が最大となるような  $k$  頂点の集合  $S$  を探索することである。

本論文では、ハッピーセット問題をゲームとして表現したハッピーセットゲームを提案する。ハッピーセットゲームは、与えられたグラフを盤面として二人のプレイヤーが互いに手を打つ二人零和有限確定完全情報ゲームである。勝利条件は最終的にハッピーである頂点をより多く得ることである。

図 1 は、盤面を 5 頂点からなるパスグラフとして、ハッピーセットゲームを行った場合の最終決着の一例を示している。先手の手を白、後手の手を黒とする。このとき、白は 3 個並んでいるため、真ん中がハッピーとなっている。一方で後手は隣接頂点が黒である場所がないためハッピーが生じておらず 1 対 0 で先手の勝利である。

## 3 本論文の結果

本論文の主要な結果は以下である。

- 一般のグラフにおいて、その頂点数が偶数のときの引き分け条件について考える。(i) 与えられ



図 1: ハッピーセットゲーム

たグラフに完全マッチングが存在する場合に、後手が必ず引き分けに出来ることを示す。(ii) グラフの対称性を考えて、対称なグラフについては後手が引き分けに出来ることを示す。

- サイクルグラフについては、頂点数が偶数のときには上記 1 の結果より、後手が引き分けに出来ることがわかるが、同様に頂点数が奇数の時には引き分けに出来ることを示す。
- パスグラフについては、頂点数が 5、または 7 の場合には先手必勝となる手順が存在することを示す。
- 1 ラウンド型のハッピーセットゲームで偶数パスの最大得点差を示す。ここで、1 ラウンド型のハッピーセットゲームとは以下のような派生版である。 $n$  頂点からなるグラフを考える。まず、先手は  $\frac{n}{2}$  頂点を選択する。次に、後手が残った  $\frac{n}{2}$  頂点を選択することでゲームが終了するというそれぞれの手順が 1 回のみのハッピーセットゲームである。

## 参考文献

- [1] Yuichi Asahiro, Hiroshi Eto, Tesshu Hanaka, Guohui Lin, Eiji Miyano, and Ippei Terabaru. Complexity and approximability of the happy set problem. *Theoretical Computer Science*, 866:123–144, 2021.
- [2] Yuichi Asahiro, Hiroshi Eto, Tesshu Hanaka, Guohui Lin, Eiji Miyano, and Ippei Terabaru. Parameterized algorithms for the happy set problem. *Discrete Applied Mathematics*, 304:32–44, 2021.
- [3] Robert Aumann and S. Hart, editors. *Handbook of Game Theory with Economic Applications*, volume 1. Elsevier, 1 edition, 1992.

		指導教員	尾下 真樹 教授
学生番号	182C1073	氏名	黒木 冬悟
論文題目	統計モデルを用いたキャラクターの 格闘動作制御の検証		

## 1 背景と研究目的

格闘ゲームは、状況に応じてコマンド入力を実行することの難しさやゲームスピードの加速化などが原因で初心者が上達しにくい。この問題の対策の1つとして、スマホなどのタッチデバイスで手軽に遊べる格闘ゲームを作り、初心者でも気軽に遊べる格闘ゲームを作るという方法を考えてみる。尾下研究室が過去に行ったタッチデバイスによるキャラクタの動作制御に関する研究では、統計的な学習モデルを用いた動作生成を行い、動作ごとに異なる学習モデルを用意してユーザーのタッチ入力に対応した学習モデルを選択できるインターフェースを実現したりしている[1]。しかし、従来の研究は動作制御の開発がメインであるため、作成した格闘動作を実際に格闘ゲームにそのまま適用できるかが分からない。加えて、マルチタッチによるキャラクタの動作制御の研究自体があまり多くないので他の動作制御の方式を用いたインターフェースの比較が困難である。そこで本研究では、統計モデルがどのような動作を生成することが出来るのかについて検証することを目的とする。

## 2 統計モデルによる動作生成

先行研究[1]で用いられる統計モデルにもとづく動作生成手法を実装し、モーションキャプチャーで収集した動作データを使った学習を行い、適切な学習モデルを得られたか検証する。今回の検証用のシステムでは、入力された動作データと制約条件にもとづいて姿勢生成を行う。各タッチ入力によるキャラクタを制御するための制約条件を入力としてキャラクタの姿勢を出力する機能を開発する。姿勢生成までの処理の流れとして、まず、多数のサンプル姿勢を低次元の潜在空間にマッピングして統計モデルを構築する。統計モデルを作成したら、関節の制約条件を入力としてサンプル姿勢を探査し、最も制約条件に近い潜在空間の位置と姿勢を出力する。統計モデルの探索で姿勢を出した後、一般的な逆運動学計算を適用して制約条件をより満たす姿勢に変化させる。また、連続的な姿勢生成を行う場合は、潜在空間の位置や関節の制約条件が指定されたときに潜在空間の目標位置を設定して、潜在空間内の現在位置を更新させる。

## 3 学習モデルの検証

動作の範囲が異なる学習モデルを複数作成して、それぞれ学習処理や姿勢生成が適切にできているか、一つの統計モデルでどの程度の範囲の動作を自然な動きで生成できるかを調べた。異なる身体部位で同じ格闘動作を実行する学習モデルや、同じ身体部位で複数の異なる格闘動作を実行する学習モデルは、2つの動作データを使用した場合、両方とも潜在空間に反映されて元の動作データと同じ姿勢生成を行いうことが出来た。異なる身体部位で、複数の異なる格闘動作を実行する学習モデルについては、右半身の身体部位で行う動作のみ、または左半身の身体部位で行う動作のみの学習モデルの場合だと、全ての動作データの姿勢を生成することが出来た。しかし、右半身の身体部位と左半身の身体部位で行う動作を混合させた学習モデルの場合だと、どちらか片方の動作データの姿勢しか生成されなかった。移動と格闘動作を実行する学習モデル」については、格闘動作の姿勢生成は実行できるが移動動作については動作の途中の姿勢までしか生成されなかつた。また、潜在空間の点を選択したときにキャラクタの向きが反対になった姿勢が生成される学習モデルも存在した。さらに、足を用いた格闘動作を含む学習モデルの場合、移動動作と格闘動作が合成された姿勢が生成された。連続的な姿勢生成については、潜在空間における速度場が正しく生成できていなかつたため、実装が間に合わなかつた。

## 4 今後の課題

今後の課題として、3つ以上の動作データを1つの統計モデルで扱える姿勢生成や、より高精度に潜在空間を探索することができる動作制御の作成を目指す。また、マルチタッチ入力による複数の関節の制約条件を入力することで複数の身体部位を操作できるようにする。これらの課題を達成すれば、マルチタッチ操作による格闘ゲームの実現に近づけるだろう。

## 参考文献

- [1] Masaki Oshita, "Multi-Touch Interface for Character Motion Control Using Example-Based Posture Synthesis", 20th International Conference on Computer Graphics, Visualization and Computer Vision 2012 (WSCG 2012), pp.213-222, 2012 .

		指導教員	尾下 真樹 教授
学生番号	192C1030	氏名	惠良 健平
論文題目	関節可動域を考慮した人体モデルの物理シミュレーション		

## 1 はじめに

動作の主な取得手段にモーションキャプチャーがあるが、落下や衝突などの危険な動作を取得することは演技者に負担がかかる。このような負担はない危険動作の取得手段として物理シミュレーションがある。

物理シミュレーションでは関節の可動域を制限しなければ、可動域を超過した不自然な動作を生成してしまう。特定の軸に対して上限と下限を与える制限手法はあるが、このような回転制限では細かい可動域の制限ができない。また、関節が大きく回転して可動域を超えてからでは制限が間に合わない。

これらを解決するため、本研究では幾何学的可動域を用いて改良した物理シミュレーション手法を開発した。

三自由度の回転を制限するため、回転を点で表現して可動域を多角体領域として扱う Yamane らの手法 [1] を利用した。図 1 のように回転をリンク方向 ( $a_y, a_z$ ) ねじれ角  $\alpha$  に分け、点のように表現して、多角体の内外判定で可動域超過を判定する。

## 2 提案手法

受動的制限は、シミュレーションの各剛体を対象とする動力学計算が行われ、関節の拘束条件が適応される後の段階で超過が発生した場合に、超過点を最も近い可動域内の点に近づける計算を行って修正回転を計算し、現在の回転を修正する。

能動的制限では、シミュレーションで衝撃が発生した段階でコントロールとして呼び出し、各関節で 3 軸まわりの角速度を取得し、閾値より大きい角速度が生じた場合に回転を相殺する修正トルクを加える。結果として次の動力学計算からトルクが反映される。

## 3 開発システム

本提案手法を ODE をベースとして開発した。関節回範囲は、三次元の可動域を三角柱集合を三角形の頂点とその高さの上限・下限の値の組み合わせを用いて表す。これを入力として与える。シミュレーションで用いる多関節体は BVH 形式から骨格を取得し構成する。基本的な制御としては直立制御を行う。衝撃パターンとして落下と球体衝突を用意した。図 2 はシミュレーションの実行結果例である。

ODE での実装にあたって、受動的制限では、ODE の任意軸回りの可動角を制限する機能を拡張し、現在

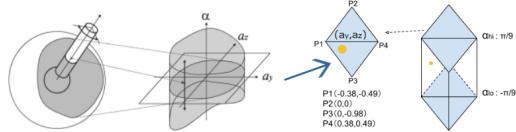


図 1: 回転を表す点とその可動域

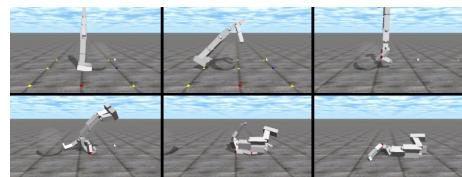


図 2: 制限を加えた小球衝突シミュレーション

の回転に基づいて姿勢に合わせ、3 軸で可動域を動的に更新する。

## 4 実験結果

幾何学的な可動域を与えた受動的制限により、制限を全くかけない場合と比較して膝が逆に折れることなどなく、可動域を制限した動作が生成できた。一部の動作例で能動的制限により、能動的制限がない場合と比較して過剰な回転が抑えられ、結果として回転に抗うような動作になることが確認できた。

## 5 まとめ

物理シミュレーションにおいて幾何学的な可動域を用いた可動域制限の検証を行った。受動的制限の効果により、一見しては自然な可動域を守る動作が生成できた。一部の結果で関節超過が発生するような動作で能動的制限が機能した。現状、能動的制限は動作が不安定である。単独の関節ではなく、複数関節に跨る制御を行うような能動的制限のパターンを増やすことでこれらの解決が期待される。

## 参考文献

- [1] Katsu Yamane, Katsu Yamane "Natural Motion Animation through Constraining and Deconstraining at Will" IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 9, NO. 3, JULY-SEPTEMBER 2003

		指導教員	尾下 真樹 教授
学生番号	192C1046	氏名	門崎雄紀
論文題目	メタボールを用いた衝突回避のための動作変形手法		

## 1 はじめに

コンピュータゲームではキャラクターの腕や頭などが壁をすり抜けることがあるため、壁をすり抜けないように動作を変形する技術が求められる。尾下の研究[1]では動作全体を格子で覆い、障害物に応じて格子を変形させ、変形した格子にしたがってキャラクターの関節を移動させることで障害物を回避した動作を生成していた。しかし、この方法では格子の構築や変形の処理で計算時間が大きくなってしまうため、より高速な手法が求められる。人体の動作と障害物の情報を入力として障害物を回避するように入力動作を変形した動作を出力するシステムを開発する。動作は連続する姿勢で表され、姿勢は全関節の回転と腰の位置・向きにより表される。

## 2 提案手法

本手法では先行研究の手法[1]と同様に、動作を構成するすべての姿勢の関節点を順運動学計算により求め、関節点に対して障害物を避けるように位置の変換処理を行う。移動した関節点の位置を基に、逆運動学により姿勢をもとめ、変形後の動作を生成する。

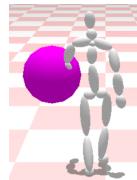
メタボールによるモデリングの考え方を応用し、障害物から関節を遠ざけるような力を及ぼす領域とを定義する。障害物の中に入った関節を移動させる距離と方向を決める関数を障害物関数と呼ぶ。一般的に1つの障害物に対し、複数のメタボールを設定することで密度空間を定義する。すべてのメタボールが障害物関数を持ち、障害物関数の和を計算することで、メタボール領域に入った関節点の移動量が計算できる。

## 3 実験結果

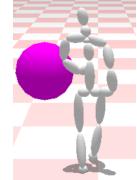
歩行動作と複数の異なる障害物をシステムに入力し、意図したとおりの出力動作が得られるかを検証する。図1のように手が障害物に入る例では、動作は自然に変形された。図2のように多関節体の両端に障害物がある例では、腕が移動しすぎて胴体と腕が重なってしまう問題が発生した。

## 4 まとめ

メタボールの考え方に基づく動作変形手法を開発し、有効性の検証を行った。実験結果から多くの場合、意図した動作が得られることがわかった。それと同時に

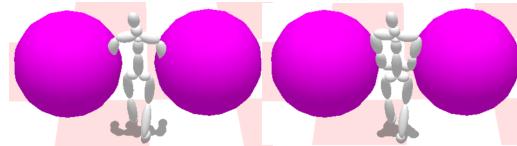


(a) 入力動作

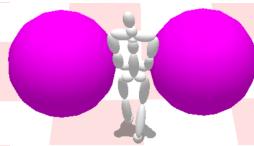


(b) 出力動作

図 1: 実験 1 手が障害物に入る場合



(a) 入力動作



(b) 出力動作

図 2: 実験 2 多関節体の両端に障害物がある場合

問題があることもわかった。腕が移動しすぎて胴体と腕が重なってしまう問題は、胴体にもメタボールを設定することで解決できる可能性が考えられる。

## 参考文献

- [1] Masaki Oshita, “Lattice-Guided Human Motion Deformation for Collision Avoidance”, Motion in Games 2017, Short Paper, 6 pages, Barcelona, Spain, November 2017.

		指導教員	尾下 真樹 教授
学生番号	192C2029	氏名	河野 俊汰朗
論文題目	仮想カメラと立体フィギュアを使った カメラワーク編集		

## 1はじめに

CG 映像を作成するには、カメラワーク（カメラの動き）を作成する必要がある。しかし、一般的なカメラワークの作成方法では、ある時刻におけるカメラや人物の位置と向きを設定する、ということを繰り返す必要がある。そのため、直感的に操作することができず難しい。この問題を解決するために、カメラワークを簡単に作成できる技術が求められている。

Disney の仮想カメラを使った手法 [1] では、人物等の配置が現実空間で行えない、広い空間が必要であるという問題がある。

小野坂らの AR 技術を用いたカメラワーク作成手法 [2] では、シーンの確認が現実空間で行えない、人物の位置と向きを変更できないという問題がある。

本論文では、カメラワークの作成と人物の配置の決定を、直感的かつ手軽に実現するシステムを開発する。

## 2 提案手法

本論文で開発するシステムは、狭い空間でもカメラワークが作成できるように、シーン上の人物を立体フィギュアで表現し、実際のシーンを縮小して再現する(図 1)。立体フィギュアは人物の動作を表し、動作データの中から分かりやすい特徴的な姿勢を利用者が選び、この姿勢を 3D プリンタで印刷することによって制作する。

仮想カメラは、シーン上のカメラの位置と向きを設定するために、利用者が手を持って動かす。モニターディスプレイに表示される、仮想カメラに映る映像の中で、人物の動作を再生することで、人物の動作を確認しながらカメラワークを作成することができる。

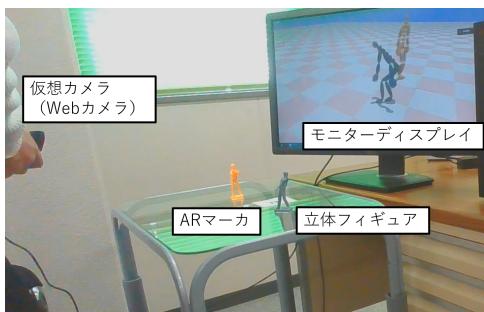


図 1: 実写画像

## 3 開発システム

本論文では、利用者の仮想カメラや立体フィギュアの操作を入力とし、カメラの動作データを出力するシステムを開発する。

立体フィギュアや仮想カメラの位置と向きを取得するため、Web カメラと AR マーカを使用する。立体フィギュアの底面に AR マーカを貼り、これを人物配置用カメラで撮影することで、配置を取得する。テーブルの上に AR マーカを貼ることで、仮想カメラの位置と向きを取得する。

利用者はまず、立体フィギュアを配置して、シーンの人物と動作の配置を決定する。次に、仮想カメラを操作して、人物の動作を見ながら、カメラの動作を記録し、カメラワークを作成する。最後に人物とカメラの動作を再生して、カメラワークや画面上の人物の配置、人物の動作を確認する。必要に応じて、人物の配置を変更し、仮想カメラを操作してカメラワークを作成し直す。

## 4 実験結果

歩行動作と歩行後にしゃがむ動作を用意して、2 体の立体フィギュアを作成し、1 人ずつ順に動作を行うシーンを想定し動作テストを行った。本システムを使って 2 人の動作を順に追うようなカメラワークを作成し、人物の動作に合わせたカメラワークを実現することができた。しかし、仮想カメラで AR マーカを認識しにくくなる場合、カメラの位置と向きのブレが発生した。

## 5 まとめ

提案手法により、人物の動作に合わせたカメラワークの作成が行えた。AR マーカが仮想カメラの画角に映らない場合があるカメラワークを作成することができない点や、AR マーカの認識精度が照明や立体フィギュアの色に影響される点は、未解決の課題である。

## 参考文献

- [1] ディズニー・スタジオ公式. “「ライオン・キング」MovieNEX まるで VR ゲーム？！驚きの撮影方法”. YouTube. 2019-11-29. [https://www.youtube.com/watch?v=CGbTiQ\\_1-6M](https://www.youtube.com/watch?v=CGbTiQ_1-6M).
- [2] 小野坂 明生, 新藤義昭, “Augmented Reality (AR) 技術を用いた CG アニメーションのカメラワーク制作システムの開発”, 第 10 回情報科学技術フォーラム, 第 3 分冊, 408 ページ, 2011

		指導教員	岡部 孝弘 教授
学生番号	192C1057	氏名	倉知 美帆
論文題目	多波長・多方向光源下の单一偏光画像からの表面素材識別		

## 1 はじめに

物体表面の素材や状態の識別は、外観検査の自動化や質感の認識において重要である。画像に基づく手法は、非破壊かつ非接触で画素ごとの識別を実現できる。

画像の各画素の輝度値は、被写体の素材に固有な性質である表面粗さや分光反射率に依存する。そのため、特定の方向・波長の光源下では、素材ごとに見えが異なる場合がある。Wang ら [1] は、撮影時の照明環境を最適化することで素材識別を行う手法を提案した。

また、物体表面からの反射光の偏光状態は、鏡面/拡散反射率や屈折率など、素材に固有な性質に依存する。Wolff ら [2] は、偏光解析により分離される反射成分を手掛かりとし、金属と非金属の識別を行った。

本稿では、照明環境の工夫と偏光情報の活用による素材識別手法を提案する。様々な色・方向の光源下における偏光画像を用いることで、性質の類似した素材の識別精度を向上させる。さらに、光源強度に非負値制約を課した提案手法は、ワンショットの識別が可能であるため、動物体にも適用可能である。

## 2 提案手法

本手法では、多波長・多方向光源下におけるモノクロ偏光画像を利用する。これは、1 画素につき、水平面に対して、 $0^\circ, 45^\circ, 90^\circ, 135^\circ$  の 4 偏光成分の輝度値を捉えた画像である。物体表面上のある点において、 $i$  番目の角度の偏光成分に着目したとき、 $L$  個 ( $L = C$  色  $\times D$  方向) の光源を、明るさを 1 として、順に点灯し、観察される輝度値を並べた  $L \times 1$  のベクトルを  $\mathbf{x}_i$  とする。4 偏光成分を結合した  $4L$  次元のベクトル ( $\mathbf{x}_1^T, \mathbf{x}_2^T, \mathbf{x}_3^T, \mathbf{x}_4^T$ ) を特徴とし、 $4L$  次元の特徴空間を考える。この特徴空間における線形識別面により素材を識別する。

一方、多波長・多方向光源下におけるワンショット画像について考える。光源強度が  $\mathbf{w} = (w_1, w_2, \dots, w_L)^T$  のとき、ある偏光成分の輝度値は、 $\mathbf{w}^T \mathbf{x}_i$  と記述できる。重み  $w'_i$  により、4 偏光成分の輝度値を線形結合すると、 $L$  種類の光源下における偏光情報を

$$I' = \sum_{i=1}^4 w'_i \mathbf{w}^T \mathbf{x}_i \quad (1)$$

のように表すことができる。この式にバイアス項  $b$  を付加し、全体を 0 とすると、式 (2) が得られる。この式



図 1: 識別対象となる 4 種類のプラスチック

表 1: 識別率 [%]

		提案手法	光源のみ	偏光のみ
PMMA	PET	100.0	98.7	69.8
PMMA	ABS	100.0	100.0	100.0
PMMA	PP	100.0	100.0	100.0
PET	ABS	100.0	100.0	100.0
PET	PP	100.0	100.0	100.0
ABS	PP	100.0	100.0	66.8

は、 $4L$  次元特徴空間における超平面を表す。

$$I' + b = \sum_{i=1}^4 w'_i \mathbf{w}^T \mathbf{x}_i + b = 0 \quad (2)$$

この超平面を識別面とし、結合係数  $\mathbf{w}$ ,  $w'_i$  を最適化することで、2 クラス識別を行う。これにより、識別に適した光源強度と偏光成分の重みが得られる。本手法では、識別面の結合係数が互いに独立でなく、通常のソフトマージン SVM では解くことができないため、二次計画法により、交互最適化を行う。なお、ワンショット識別のため、光源強度に非負値制約を課す。

## 3 実験・結果

4 色  $\times$  5 方向の計 20 種類の光源下において、図 1 に示す 4 種類のプラスチックを対象に実験を行った。表 1 に、提案手法、光源強度のみを最適化したとき、および、偏光成分成分の重みのみを最適化したときの識別率を示す。どちらか一方のみの最適化では識別に限界がある素材に対しても、良好な結果が得られた。

## 4 まとめ

本稿では、照明環境の工夫と偏光情報の活用により、素材識別を行う手法を提案し、その有効性を示した。今後は、対象素材を増やし、多クラス識別に取り組みたい。

## 参考文献

- [1] C. Wang and T. Okabe, "Joint optimization of coded illumination and grayscale conversion for one-shot raw material classification," In Proc, BMVC2017, 2017.
- [2] L. B. Wolff, "Polarization-based material classification from specular reflection," IEEE TPAMI, Vol.12, No.11, pp.1059–1071, 1990.

		指導教員	岡部 孝弘 教授
学生番号	192C1100	氏名	中川 俊輔
論文題目	偏光ライトフィールドの解析による鏡面の法線推定		

## 1 はじめに

コンピュータビジョンにおいて、鏡は視点の拡張に用いることができる。カメラの視線方向は鏡によって変換され、変換後の視線方向は鏡の位置や形状（法線）によって決定される。したがって、鏡によって変換される視線集合（ライトフィールド）[1] を利用するためには、鏡の法線を獲得しておくことが必要となる。

平面鏡を用いる場合、鏡の法線推定はチェスボードを鏡に映し出し、単一画像中のチェスボードの複数のコーナー点とその鏡映点の位置関係から、共通の鏡面法線を推定可能である。一方曲面鏡を対象とする場合、ピクセルごとに法線が異なるため、実像と鏡像の対応関係のみでは鏡面の法線を推定するのは困難である。

そこで本稿では、既存のライトフィールドを偏光情報に拡張した偏光ライトフィールドを導入する。そして鏡映点の対応関係に加え、偏光角から求められる電場の方向の鏡映関係を用いて解析を行い、任意曲面の鏡の法線を 1 ピクセルごとに推定する手法を提案する。

## 2 提案手法

本手法では図 1 のように、出射光が完全偏光である液晶ディスプレイの画面を映し出すように曲面鏡を配置して、偏光カメラで撮影する。ここで照射・観測される偏光光線の集合を偏光ライトフィールドと定義し、これらを解析することで鏡面法線を推定する。

ある光線を考えるとき、一般にその幾何学的な記述は、始点  $p$  と方向ベクトル  $v$  によって定義できる。本研究ではこの光線を偏光光線に拡張する。座標系の  $z$  軸に垂直な平面において、平面を通過する光線の電場の方向が  $k = (\cos \alpha, \sin \alpha, 0)^\top$  で表されるとき、光線における電場の振動方向は以下の偏光面の法線

$$h = (v \times k) / \|v \times k\| \quad (1)$$

によって表現できる。この拡張により、鏡面法線に関して偏光面の法線と視線方向のそれについて、偏光光線の鏡映関係としての拘束を定式化できる。

図 1 において偏光面の法線の鏡映関係を考える。まず、 $v'_c$  とカメラ原点を始点とし、画像平面上で観測された電場の方向  $k'_c = (\cos \alpha', \sin \alpha', 0)^\top$  によって張られる、偏光面の法線方向  $h'_c$  を求められる。一方で、ディスプレイ上の点  $p_d$  に対応する既知の電場の方向

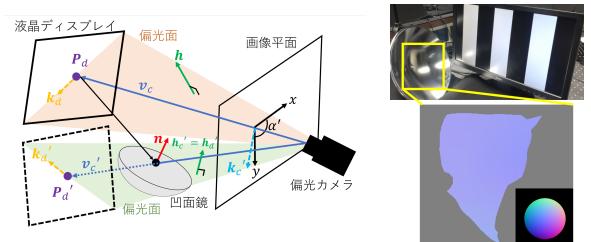


図 1: 撮像系における偏光光線 図 2: 法線推定結果

$k_d = (\cos \beta, \sin \beta, 0)^\top$  は、鏡面の法線  $n$  によって与えられるハウスホルダー行列で  $p'_d$  に鏡映変換することができる。したがって、 $v'_c$  と  $k_d$  を鏡映変換したベクトルによって張られる、偏光面の法線方向  $h'_d$  を求められ、これは上述の  $h'_c$  と等しい関係にある。

同様に、 $v'_c$  に対応する偏光光線の方向の鏡映関係を考えたとき、既知のディスプレイ上の点  $p_d$  に対する鏡映点  $p'_d$  は、ハウスホルダー行列  $H$ 、カメラ座標系の原点から鏡映点における曲面鏡の接平面に下ろした垂線の長さ  $\ell$ 、鏡面法線  $n$  を用いて求めることができる。

以上から、 $h'_c$  と  $h'_d$  が一致すること、および  $v'_c$  と  $p'_d$  の単位ベクトルが一致することを利用して、 $n$  と  $\ell$  の最適化を行い、鏡面法線を推定する。

## 3 実験

液晶ディスプレイの画面を凹面鏡に映し出すように配置し、ディスプレイの画面にグレイコードのパターン画像 [2] を映しながら、凹面鏡上の鏡映点を同時に偏光カメラで撮影した。撮影画像からピクセルの対応ごとに偏光面の法線を計算し、提案手法で示した最適化で鏡面の法線推定を行った。推定した鏡面の法線方向をカラーマップとして図 2 に示す。図 2 より凹面鏡上の法線は良好に推定されていることが確認できる。

## 4 まとめ

本稿では、鏡映点の対応関係と電場の方向の鏡映関係から任意曲面の鏡面法線推定の手法を提案した。今後は滑らかさ制約を導入し、より頑健に法線を推定する。

## 参考文献

- [1] M. Grossberg, and S. Nayar, “The raxel imaging model and ray-based calibration”, IJCV, 61(2):119–137, 2005.
- [2] S. Yamazaki, M. Mochimaru, and T. Kanade, “Simultaneous self-calibration of a projector and a camera using structured light”, In Proc. PCS2011, pp. 67–74, 2011.

		指導教員	岡部 孝弘 教授
学生番号	192C1153	氏名	三好 優輝
論文題目	偏光照明環境による実物体の光沢の操作		

## 1 はじめに

実物体の材質があたかも変化したように見えを操作する技術は質感編集と呼ばれる。任意照明環境下における実物体の光沢（鏡面反射成分）の操作は AR, MR への応用が期待される。

一般に、鏡面反射光の明るさは物体への入射光の偏光方向に大きく依存する一方、拡散反射光の明るさはほとんど依存しない。この性質を用いて、鏡面反射光のみを強調／抑制できると考えられる。

Ma ら [1] は、球状照明下で偏光を用いて鏡面反射光のみを抑制し、反射成分を分離する手法を提案した。しかし、この手法では、カメラの前に偏光板が必要である上に、光源の偏光方向は固定されており、鏡面反射光の強調は困難である。

そこで本稿では、任意照明環境下において裸眼で観測される実物体の光沢を操作する手法を提案する。提案手法では、色や明るさ、偏光方向を制御可能な偏光照明システムを構築する。さらに構築した照明システムを用いて所望する光沢の観測を実現する。

## 2 提案手法

一般に、入射光と物体表面の法線が張る入射面に対して、垂直／平行な偏光（s 偏光／p 偏光）では反射率が異なる。鏡面反射光は視線方向と光源方向の二等分方向が法線と一致するときに最も強く観測される性質があるため、入射角と反射角が等しくなる光源方向を持つ入射光の偏光方向を、s 偏光や p 偏光に近づけると、鏡面反射光の強調や抑制ができる。

図 1 に偏光照明システムのモデルを示す。本システムでは特に、偏光方向の操作を透過型液晶を用いて行う。透過型液晶に入力した画像の画素値に応じて、偏光方向を最大 90° の範囲で変化させることができる。その後、反射の前後で偏光方向を保つ偏光反射スクリーンに投影した反射光により実物体を照明する。なお、事前の較正により、プロジェクタ画像座標-透過型液晶座標-法線間の対応関係を求める [2]。

提案手法では、画素値が 0 から 255 までの一様なグレースケール画像を透過型液晶に入力し、入射光の偏光方向を変化させながら参照物体（球）を撮影する。撮影画像の画素毎に 256 個の輝度値と透過型液晶入力画素値の関係について、ルックアップテーブルを作成／

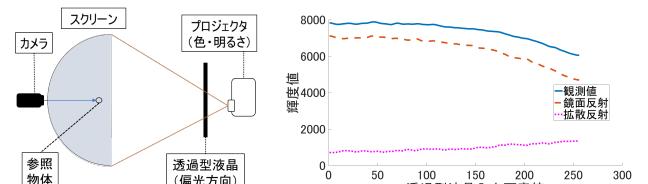


図 1: 偏光照明システム

図 2: 観測輝度値変化例

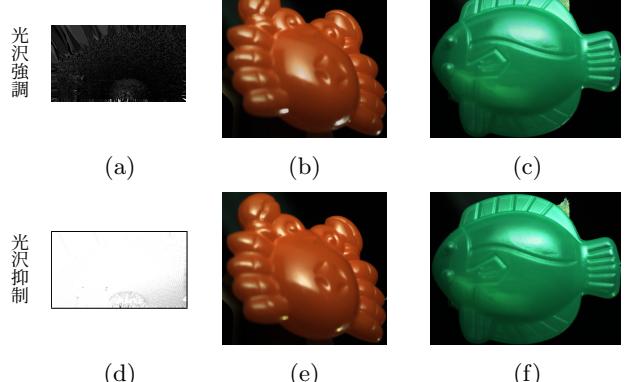


図 3: 作成パターンと撮影画像

参照して所望の光沢を与える入力パターンを作成する。

## 3 実験

実際に、観測される光沢が透過型液晶への入力により操作できることを確認するため、平面スクリーンを用いた実画像実験を行った。図 2 に、球の撮影画像のある画素の輝度値変化を示す。図 3(a) / (d) の光沢を強調／抑制するパターンを用いた形状未知の実物体の観測結果をそれぞれ図 3(b) と (c), (e) と (f) に示す。これらの撮影画像から、光沢部分の輝度値変化が確認でき、作成パターンが定性的に有効であることが分かる。

## 4 むすび

本稿では、偏光を用いて鏡面反射成分を強調／抑制する手法を提案し、実画像実験を通してその有効性を確認した。半球状のスクリーンへの拡張や、偏光方向の変化量を最大 90° から 180° へ拡張し、任意の偏光方向の光を入射させることなどは今後の課題である。

## 参考文献

- [1] W. C. Ma, et al. "Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient", In Proc. EGSR2007, pp.183–194, 2007.
- [2] S. Yamazaki, et al. "Simultaneous self-calibration of a projector and a camera using structured light", In Proc. PCS2011, pp.67–74, 2011.

		指導教員	岡部 孝弘 教授
学生番号	192C1174	氏名	吉田 百花
論文題目	偏光と陰に基づくワンショット法線推定		

## 1 はじめに

コンピュータビジョン分野において、形状復元は重要な課題である。特に画素毎の法線の推定は、工業製品の外観検査や文化財の保存・調査等に応用できる。

偏光に基づく形状復元手法として、1枚の偏光マルチスペクトル画像から形状と屈折率を同時推定する Huynh らの手法 [1] がある。しかし、偏光解析による法線は方位角に関して  $180^\circ$  の不定性があるため、物体形状が滑らかかつ凸物体であることを仮定して不定性を解決している。また、陰に基づく形状復元として、Kriegman らの手法 [2] がある。しかし、陰境界の交点においてのみ法線推定が可能であり、密に推定できない。

本稿では、偏光解析によって法線に関して方位角の2つの解と天頂角を推定し、陰に基づく解の選択により、物体形状に関する滑らかさなどの制約を課すことなく画素毎の法線を推定する。また、不定性の解決のために光源方向の最適化も行う。

## 2 提案手法

本手法では、RGB3色かつ異なる方向の3光源を同時に照射した被写体をカラー偏光カメラで撮影する。異なる方向の光源で照らされた被写体をカメラの各バンドで独立に捉えることにより、ワンショットの法線推定を実現する。なお、ランバートモデルに従い、屈折率が既知の被写体を仮定する。

被写体表面の法線は天頂角と方位角を用いて表される。4つの偏光成分から入射光の反射と透過の割合であるフレネル透過率  $r^2$  が求められる。十分な明るさで照らされていてかつ鏡面反射が観察されない被写体表面上の点において、天頂角  $\theta$  は、被写体の屈折率  $\eta$  とフレネル透過率を用いて、

$$\theta = \arcsin \frac{\eta \sqrt{1 - r^2}}{\sqrt{\eta^2 - 2r\eta + 1}} \quad (1)$$

と表される。偏光解析による方位角は、最大輝度が観測される偏光角である位相角と同じもしくは  $180^\circ$  異なる角度をとることから、方位角には  $180^\circ$  の不定性が存在する。

観測輝度は光源方向と法線の内積に依存し、照らされている部分での光源方向と法線の内積は正である。また、陰部分での内積は負であるため、観測輝度は0となる。

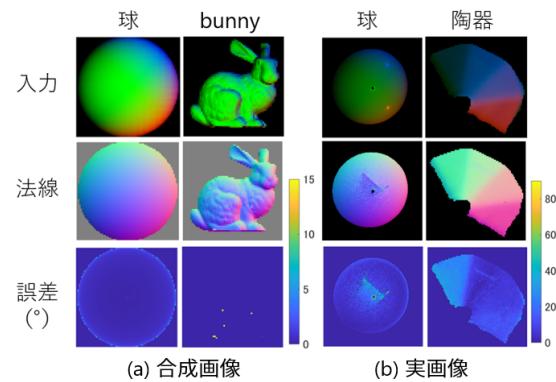


図 1: 実験結果

る。光源方向と法線候補の内積を取り、観測輝度と整合性のある法線候補を選択することで、法線の方位角に関する  $180^\circ$  の不定性を解消する。

偏光解析のために、G 光源は視線方向から照らし、R, B 光源を用いて陰に基づく不定性解消を行う。R, B の光源方向について、2つの法線候補と光源方向の積の正負が一致するとき、不定性を解くことができない。そこで、法線と光源方向のなす角度を用いて、R, B の光源方向を最適化する。

## 3 実験

光源方向の最適化により、3つの光源方向が直交するような光源方向が最適であることが分かった。したがって、合成画像実験、実画像実験ともに光源方向がそれぞれ直交する3光源を用いて推定行った。図 1 に推定結果を示す。合成画像、実画像とともに定性的に概ね良好な結果が得られているが、実画像実験に関して改善の余地がある。

## 4 むすび

本稿では、偏光解析によって得られる法線の方位角の2つの解と天頂角を観察される陰の整合性に基づくように選択することで、画素毎の法線を一意に決定した。また、提案手法における光源の最適化も行った。鏡面反射や陰を考慮した拡張、偏光解析手法の改善による実画像実験での精度向上は今後の課題である。

## 参考文献

- [1] C. Huynh, A. Robles-Kelly, and E. Hancock, “Shape and Refractive Index from Single-View Spectro-Polarimetric Images”, IJCV, 101(1), pp.64–94, 2013.
- [2] D. Kriegman and P. Belhumeur, “What shadows reveal about object structure”, In Proc. ECCV1997 (LNCS1407), pp.399–414, 1997.

		指導教員	岡部 孝弘 教授
学生番号	192C2107	氏名	山田 悠稀
論文題目	ライトトランスポート獲得のための 照明パターンと復号処理の同時最適化		

## 1 はじめに

光源から出た光がシーンで反射・散乱し、カメラなどの観測装置に達するまでの光の伝搬はライトトランスポート (LT: Light Transport) と呼ばれる。LT は、照明条件を変化させながら撮影した大量の画像そのものであり、その獲得は照明シミュレーションなどの応用に重要である。本研究では、ディスプレイ-カメラシステムの LT を効率良く獲得する手法を考える。

これまでに、複数光源を同時に点灯した様々な照明パターン下の撮影画像から復号処理により単一光源下画像を求める符号化照明が提案されている。特に、光源を線形独立な組み合わせで同時に点灯し、線形演算で復号を行う多重化照明 [1] では、理想的なノイズ伝播モデルに基づいて、S 行列に基づく照明パターンが最適であることが示されている。しかし、このモデルでは拡散反射の低周波性 [2] や画素値の量子化などの実画像の性質が考慮されていないため、撮影画像の拡散反射成分がほぼ一定になってしまい、単一光源下画像を高精度に復号することは困難である。

そこで本稿では、実画像を用いたデータ駆動型のアプローチで、実画像の性質を考慮した照明パターンと復号処理を同時に最適化する。提案手法では、畳み込みニューラルネットワーク (CNN) における  $1 \times 1$  の畳み込みカーネルを用いて照明パターンを表現できることに着目し、CNN の枠組みで同時最適化を行う。

## 2 提案手法

照明パターンと復号処理の最適化のために、照明層と復号ネットワークからなる図 1 のようなネットワークを考える。学習時は、単一光源下画像を入力として、照明層において、それらを組み合わせた照明パターン下画像を作成する。復号ネットワークでは、照明パターン下画像を入力として、単一光源下画像を復号する。

照明層では、 $1 \times 1$  の畳み込みカーネルにより照明パターンを表現する。ここで、重ね合わせの原理より、任意の照明パターン下画像は各单一光源下画像の線形結合で表現される。結合係数は各光源の明るさに等しく、CNN の  $1 \times 1$  の畳み込みカーネルで表現できるため、照明パターンの学習は照明層における  $1 \times 1$  の畳み込みカーネルの学習で行える。さらに、実際の撮影画像にノイズが含まれることから、照明層の出力である線形

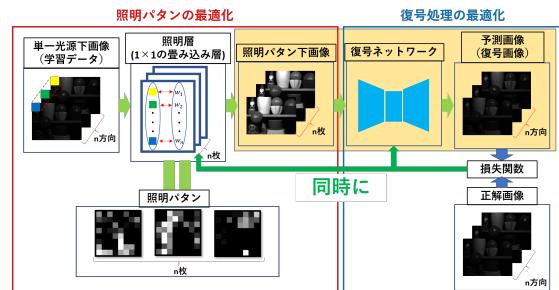


図 1: 照明パターンと復号処理の同時最適化ネットワーク

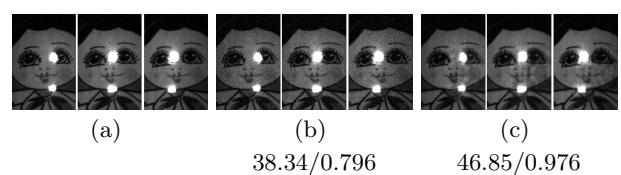


図 2: 正解画像と復号画像

結合画像にノイズを付与して、照明パターン下画像とする。具体的には、正規分布ノイズだけではなく一様分布ノイズを加えることで、量子化の影響を表す。

テスト時には、図 1 の黄色の領域のように、最適化された照明パターン下で実際に撮影した画像を復号ネットワークに入力し、単一光源下画像を復号する。

## 3 実験

提案手法の有効性を示すために実画像を用いた実験を行った。図 2 に (a) 正解画像、(b) S 行列に基づく多重化照明の結果、(c) 提案手法の結果 (数値は PSNR/SSIM) を示す。図 2(b) がノイズの影響を大きく受け、ざらざらした画像になっているのに対して、図 2(c) ではノイズが低減されていることが分かる。また、定量的な比較においても、提案手法が S 行列に基づく多重化照明よりも有効であることが分かる。しかし、鏡面反射成分の周辺の復号には改善の余地がある。

## 4 むすび

本稿では、LT 獲得のための照明パターンと復号処理の同時最適化手法を提案した。今後の課題として、偏光カメラを用いて反射成分を分離し、成分ごとに復号を行うネットワークを作成することで精度向上を目指す。

## 参考文献

- [1] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "A theory of multiplexed illumination", In Proc. ICCV2003, Vol.3, pp.808–815, 2003.
- [2] R. Ramamoorthi, and P. Hanrahan, "A signal-processing framework for inverse rendering", In Proc. SIGGRAPH2001, Vol.1, pp.117–128, 2001.

		指導教員	嶋田 和孝 教授
学生番号	192C1155	氏名	本島 永佳
論文題目	深層学習を用いた評価文からの評価対象抽出		

## 1 はじめに

世の中には、様々な事物に対する評価を含む文章が存在し、これらを評価文と呼ぶ。評価をうける企業や自治体といった組織は、サービスの改善のために評価文を分析する必要がある。評価文の分析では、肯定的か否定的かという分析だけでなく、評価の対象や評価表現などの抽出も行う必要がある。しかし、これらの分析を人手で行った場合、コストがかかる。そこで、評価文からの適切に評価対象を自動抽出することを目指す。

本研究では、評価対象抽出に深層学習モデルを使用し、本タスクにおける深層学習モデルの有用性を検証する。

## 2 データセット

本研究では、栗原ら [1] によって作成された評価対象-評価表現データセットを使用する。このデータセットは、Twitter から収集した日本語ツイートから構成されている。何らかの事物に評価を行っているツイートに対して、評価対象と評価表現を表す部分文字列のアノテーションを行っている。本研究では、このデータセットのなかでも評価対象を含む 4773 ツイートのみを扱う。

また、本タスクでは IOB2 方式でタグ付けを行う。図 1 にタグ付けの例を示す。

## 3 評価対象抽出モデル

CRF, BERT, BERT+CRF の 3 つのモデルを使用する。CRF とは、構造学習によって系列とラベルの関係を学習する識別モデルである。今回、CRF として sklearn-crfsuite<sup>1</sup>を使用する。素性として、単語の表層形、係り受け、品詞、分散表現の 4 つを用いる。BERT は、事前学習モデルの 1 つである。大規模なデータセットを事前学習し、単語の汎用的な特徴を習得しているため、小規模なデータセットに対しても高い精度が得られる。本研究では、東北大学が公開している日本語事前学習済み BERT モデル<sup>2</sup>を使用する。以降、このモデルを BERT<sub>wiki</sub> とする。BERT<sub>wiki</sub>+CRF は、BERT<sub>wiki</sub> にラベルの遷移を学習する CRF 層を追加したものである。以降、このモデルを BERT<sub>wiki</sub>+CRF とする。

図 1: 評価対象およびタグ付けの例。評価対象を橙字で示す。

表 1: 評価対象の抽出精度

モデル	完全一致			部分一致		
	Pre	Rec	F1	Pre	Rec	F1
CRF	<b>0.554</b>	0.427	0.483	0.602	0.429	0.501
BERT <sub>wiki</sub>	0.438	0.321	0.370	0.691	0.545	0.605
BERT <sub>wiki</sub> +CRF	<b>0.554</b>	<b>0.540</b>	<b>0.547</b>	<b>0.713</b>	<b>0.654</b>	<b>0.682</b>

## 4 実験

データセットは 8:1:1 に分割し、学習データ・検証データ・テストデータにあて、10 分割交差検証を行う。なお、CRF においては検証データを用いないものとする。評価は、完全一致と部分一致の 2 点から行う。完全一致は評価対象にあたる系列がすべて一致しているか、部分一致は O タグ以外の評価対象のラベルがどれだけ一致しているかを意味する。

結果を表 1 に示す。CRF と BERT<sub>wiki</sub> の精度を比較すると、完全一致では CRF が BERT<sub>wiki</sub> を上回り、部分一致では BERT<sub>wiki</sub> が CRF を上回る結果となった。CRF は、素性によって適切な範囲でラベルが付与されたと考えられる。BERT<sub>wiki</sub> は、評価文中で評価対象が出現する位置をおおよそ予測できているという特徴が見られた。また BERT<sub>wiki</sub>+CRF は、すべての指標において BERT<sub>wiki</sub> を上回る結果となった。特に完全一致において、精度の大きな向上が見られ、CRF 層を追加することの有用性が確認された。

## 5 おわりに

本論文では、評価文からの評価対象抽出タスクに対して、BERT、および BERT に CRF と同等の機能を組み合わせた BERT+CRF を使用した。今後は評価表現の抽出も含めたマルチタスクに対して、深層学習を用いることが考えられる。

## 参考文献

- [1] 栗原理聰、水本智也、乾健太郎：“Twitter による評判分析を目的とした評価対象-評価表現データセット作成”，言語処理学会第 24 回年次大会発表論文集, pp. 344–347 (2018).

<sup>1</sup><https://github.com/mansweet/sklearn-crfsuite>

<sup>2</sup><https://github.com/cl-tohoku/bert-japanese>

		指導教員	嶋田 和孝 教授
学生番号	192C1172	氏名	横山 正秋
論文題目	トピック情報と BERT を用いた 音声認識誤りの検出		

## 1 はじめに

自動音声認識は、人間の話す音声を認識し自動的にテキスト化する技術である。音声認識結果に含まれる誤りは後段のタスクの精度に悪影響をもたらすため、認識誤りを検出することは重要なタスクである。

認識誤りの検出にあたり、文脈情報に加えてトピック情報を用いることで、検出が容易になると考える。以下の例文の太字は音声認識結果の誤りである。

- (1) 幼児の認識の発達レーサー<sup>1</sup>について実験した。
- (2) 投手<sup>2</sup>へのインタビューです。

人間は例文(1)の文脈中の「レーサー」に対し違和感を持ち、誤りの可能性を感じ取ることができる。さらに、例文(1)のトピックとして年齢に関する議論がなされていると仮定すれば、誤りの確信を持ちやすい。例文(2)では、周囲の発話のトピックが政治に関する内容であると判明しなければ、「投手」が誤りであるかを判断できない。このように、前後を含む発話全体がどのような語群を含むかというトピック情報を追加することで、より多くの誤りを検出可能であると考えられる。

本研究では認識誤りに対して、文脈情報とトピック情報を組み合わせ、系列ラベリング問題として検出を行う。

## 2 モデル

BERT は文脈依存の単語埋め込み表現を内部に持ち、文脈と異なる語の検出に有効であると考えられる。BERT-CRF は、BERT の単語埋め込み表現を CRF に入力し、ラベルを推測するモデルである。CRF が系列間での依存関係を学習するため、系列ラベリング問題に対して有効である。BERT, BERT-CRF をベースラインとする。

提案手法の BERT-LDA-CRF は、BERT の単語埋め込み表現ベクトルと、LDA によるトピックベクトルを結合したベクトルを CRF の入力とするモデルである。LDA は各文書の持つトピックを推定し、各トピックに対する確率を出力する。出現頻度が閾値未満の出現頻度の単語を、各文書の BoW 作成時の素性とする。トピック数は LDA モデルの生成時に指定する。

表 1: 各モデルにおける認識誤り検出結果。

モデル	topic	Pre	Rec	F1
BERT	-	<b>0.676</b>	0.450	0.541
BERT-CRF	-	0.627	0.634	0.630
BERT-LDA-CRF	50	0.642	<b>0.644</b>	0.643
出現頻度閾値=0.3	100	0.655	0.638	<b>0.646</b>

## 3 実験

『日本語話し言葉コーパス』(CSJ)[1]を用いて、手法評価のデータセットを作成する。CSJ には音声データと人手によって書き起こされた転記テキストが収録されている。音声データを文ごとに分割し、Google Speech Recognition<sup>3</sup>を用いて音声認識データとする。転記テキストと音声認識データを比較し、ラベルを付与する。本研究では、学会講演のうちコア講演の 70 講演を用いる。この 70 講演を用いて、LDA モデルを作成する。出現頻度閾値を 0.3, トピック数を 50, 100 と変化させて実験を行った。

BERT, BERT-CRF, BERT-LDA-CRF の各モデルに対して、データセットを用いて 5 分割交差検証を行い、精度を比較する。

表 1 に実験結果を示す。太字は最高値である。BERT-CRF は BERT に対して高い Recall, F 値を達成した。BERT-LDA-CRF はすべてのトピック数において、ベースラインに比べて高い精度を達成した。このことから、音声認識誤りの検出において、トピック情報の追加は有効であるといえる。

## 4 おわりに

本研究では、音声認識誤りに対し、文脈情報とトピック情報を同時に用いて検出を行った。BERT に対して、系列情報を用いることのメリットが確認された。BERT の単語埋め込み表現と LDA のトピックを連結し、CRF による系列ラベリングを行うことで、認識精度の向上が確認された。今後は、異なる条件でのトピック情報の追加による認識誤りの検出精度向上を目指したい。

## 参考文献

- [1] 前川喜久雄. 『日本語話し言葉コーパス』の概要. 日本語科学, Vol. 15, pp. 111–133, 2004.

<sup>1</sup> 実際の発話では齢差。

<sup>2</sup> 実際の発話では党首。

<sup>3</sup> <https://pypi.org/project/SpeechRecognition/>

		指導教員	嶋田 和孝 教授
学生番号	21221005	氏名	竹尾 匠貴
論文題目	マルチタスク学習を用いた複数項目を持つレビューの評価値推定		

## 1 はじめに

近年のデジタル化の進展に伴い、さまざまな製品やサービスに対するレビュー文が Web 上で蓄積されている。製品やサービスを提供する側にとって、レビューデータを分析して得られた知見を生かすことは有用である。したがって、蓄積された大量のレビュー文を自動的に分析できることが望ましい。

Nakamuta ら [1] は、複数の評価項目（以降、アスペクト）を擁するレビューデータにおいて、アスペクトごとに評価値を推定した。Nakamuta らはレビュー文に対してアスペクト分類を実施し、レビュー文を BoW 化した上で SVR によって推定モデルを構築した。しかし BoW では文脈を考慮できないという問題がある。

本研究では、文脈を考慮可能な深層学習モデルである BERT を用いて各アスペクトごとに評価値推定モデルを構築する。また、モデルの学習においてそのアスペクト固有の情報を学習する一方、他アスペクトの評価値推定においても有用な情報を学習しているはずである。そこで、各アスペクトの評価値推定という類似したタスクを解くモデル間でマルチタスク学習を行うことで、さらなる評価値推定の精度向上を期待する。

## 2 提案手法

BERT は大規模テキストデータで事前学習を行った汎用言語モデルである。大規模テキストから得られる知識を用いることで、評価値推定精度の向上を期待する。

提案手法を図 1 に示す。BERT は transformer 層を 12 層重ねたモデルで、図 1 にも示す通り大きく 3 段階に分けられる。本研究では、そのうち中間層のパラメータのみを各推定モデルで共有してマルチタスク学習を行うことで、各アスペクトの評価値推定モデルの精度向上を期待する。12 層の transformer 層の分割割合は Pahari ら [2] に倣い 1:7:4 とする。

評価値推定には、Nakamuta らが作成した分類モデルによって自動付与された、アスペクト分類済みレビューデータを用いる。各アスペクトの評価値推定モデルには、レビュー文のうち推定したいアスペクトのタグが付与された文のみを連結して入力する。ただし、レビュー文に推定したいアスペクトのタグが付与された文がない場合、全レビュー文を連結して入力する。

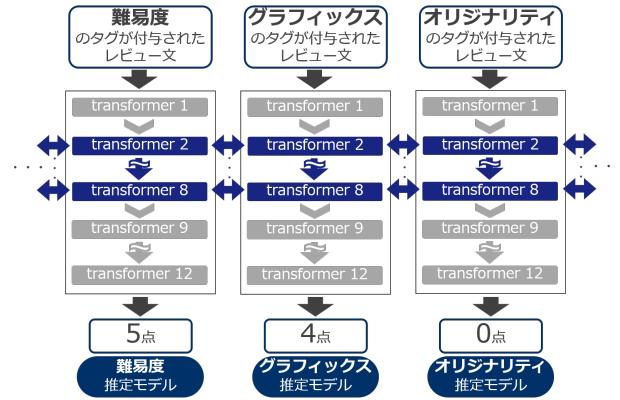


図 1: 提案手法の概要.

## 3 実験及び考察

レビューデータとして Web サイト<sup>1</sup>から収集した 3464 件を用いる。このレビューデータは 7 つのアスペクトを持ち、それぞれに 0~5 点の 6 段階で評価値が付与されている。レビューデータを 3:1:1 に分割した上で 5 分割交差検証を行い、推定評価値と正解評価値の RMSE の平均を算出した。結果は、Nakamuta らの手法で 0.925、提案手法で 0.787 となり、マルチタスク学習を行う提案手法において Nakamuta らの手法より大きく改善した。各アスペクトの評価値推定という、類似タスクを解くモデル間におけるマルチタスク学習の有効性が明らかになったといえる。

## 4 おわりに

本研究では、BERT を用いてマルチタスク学習を実施し、複数の評価項目を擁するレビューデータの評価値推定を行った。先行研究に比べて大きな改善が見られ、マルチタスク学習の有効性を確認した。

## 参考文献

- [1] T.Nakamuta and K.Shimada. Multi-aspects rating prediction using aspect words and sentences. In *Proceedings of the 29th PACLIC*, pp. 513–521, 2015.
- [2] N.Pahari and K.Shimada. Multi-task learning using bert with soft parameter sharing between layers. In *12th SCIS & 23rd ISIS*, pp. 1–6, 2022.

<sup>1</sup><http://ndsmk2.net>

		指導教員	嶋田 和孝 教授
学生番号	21221006	氏名	中本 さや香
論文題目	機械学習に基づく日本語小論文の文書力および論理性の自動採点		

## 1 はじめに

近年、入試などにおいて思考力や表現力を評価するために小論文が採用される。小論文のような記述問題は一般的に手作業による採点が行われる。しかし、手作業による採点には時間や人的なコストが大きいこと、採点ミスの可能性や平等な評価の難しさなど様々な問題が存在する。そこで自動採点システムが注目されている。高精度なモデルの開発により、採点に要する人員の削減と高品質な採点の実現が期待される。

本研究では竹内ら [1] によって開発された日本語小論文データセットを対象とした高精度な採点システムの構築を目指とする。論理性と文書力の 2 つの観点に着目した複数の自動採点システムを作成し、精度の比較検証を行う。

## 2 提案手法

文書力スコア予測システムと論理性スコア予測システムについて提案手法を説明する。それぞれ特微量ベース、深層学習ベースの予測システムを作成する。

### 2.1 文書力スコア予測

文書力は、指定字数に対する文字数や誤字数、文章の質といった観点に基づいて採点されている。

特微量ベースでは文字数割合、誤字数、漢字割合、漢字画数ベクトルを特微量として採用する。これらの特微量を結合して正規化を行い、ランダムフォレスト回帰モデルによって予測を行う。

深層学習ベースでは事前学習言語モデルである BERT を採用する。小論文を BERT に入力し、回帰的に点数を予測する。モデルの損失関数には平均二乗誤差を採用する。

### 2.2 論理性スコア予測

論理性は、文章の論理構成や問に対する適切な解答といった観点に基づいて採点されている。

特微量ベースでは 2.1 節と同様、ランダムフォレスト回帰モデルを使用して予測を行う。論理性に対しては、文字数や形態素数などの 20 個の特微量 (20 features) と Bag of Words (BoW) を採用する。

深層学習ベースでは 2.1 節と同様、BERT を使用して予測を行う。また、BERT に人手特微量を追加する、ハイブリッド手法のシステムも作成する。

表 1: 文書力スコア予測結果

手法	QWK
ランダムフォレスト回帰	<b>0.912</b>
BERT	0.620

表 2: 論理性スコア予測結果

手法	QWK	
ランダム	20 features	0.496
フォレスト	BoW	0.603
回帰	20 features + BoW	0.608
		0.596
BERT	+ 20 features	0.587
	+ BoW	<b>0.632</b>
	+ 20 features + BoW	0.622

## 3 実験

モデルの精度評価には QWK を使用する。

文書力スコア予測の実験結果を表 1 に示す。精度が最高となったのは特微量ベースのモデルであった。また、特微量ベースは人手採点者の QWK と比較しても非常に良い結果となった。

論理性スコア予測の実験結果を表 2 に示す。精度が最高となったのは BERT + BoW であった。論理性スコアの予測では BoW で文の表層的な内容を取得し BERT で文脈を考慮することで精度が向上した。

## 4 おわりに

本研究では、人手採点済み小論文データセットを使用し、特微量ベースおよび深層学習ベースの手法で自動採点システムを作成し、比較と考察を行った。今後は論理性スコア予測について、談話構造や文章の接続などに注目し、より高精度なモデルの構築を行いたい。

## 参考文献

- [1] 竹内孔一, 田口雅弘, 稲田佳彦, 飯塚誠也, 阿保達彦, 上田均. 小論文自動採点システム評価のための日本語小論文データの構築. 電気情報通信学会信学技報, Vol. 121, No. 178, NLC2021-15, pp. 40–44, 2021.

		指導教員	平田 耕一 教授
学生番号	192C1003	氏名	阿部 大祐
論文題目	根付きラベル付き木の編集距離の定数下限距離		

## 1 はじめに

Web マイニングに対する HTML や XML データなどの半構造データ、またバイオインフォマティクスにおける RNA や糖鎖データといった木構造データの比較はデータマイニングにおける重要なタスクの一つとされている。これらのデータにおける 2 つの根付きラベル付き木（以後、単に木という）の間の非類似度を表す距離として最も有名なものの一つに編集距離がある。編集距離の計算は無順序木では MAX SNP 困難 [4] となる。一方、葉距離ヒストグラム、次数ヒストグラム、ラベルヒストグラム、兄弟ヒストグラムの L1 距離 [1, 2] は編集距離に対して定数倍下界が存在するため、編集距離を近似することができる。

## 2 ヒストグラム

本論文では、編集距離をより効率良く近似するため、木  $T$  の葉距離ヒストグラム  $h_f(T)$ 、次数ヒストグラム  $h_d(T)$ 、ラベルヒストグラム  $h_l(T)$  [2]、および、兄弟ヒストグラム  $h_s(T)$  [1] を利用する。そして、二つの木  $T_1$  と  $T_2$  に対するそれぞれのヒストグラム  $h_f$ 、 $h_d$ 、 $h_s$  の L1 距離をヒストグラム距離といい、それぞれ、 $\delta_f$ 、 $\delta_d$ 、 $\delta_l$ 、 $\delta_s$  と表す。これらの L1 距離は、以下のように木の編集距離  $\tau_{\text{TAI}}(T_1, T_2)$  の定数下限を与える。

$$\begin{aligned} \delta_f(T_1, T_2) &\leq \tau_{\text{TAI}}(T_1, T_2), \\ \delta_d(T_1, T_2) &\leq \tau_{\text{TAI}}(T_1, T_2), \\ \frac{\delta_l(T_1, T_2)}{3} &\leq \tau_{\text{TAI}}(T_1, T_2), \\ \frac{\delta_s(T_1, T_2)}{4} &\leq \tau_{\text{TAI}}(T_1, T_2). \end{aligned}$$

また、 $\delta_f$ 、 $\delta_d$ 、 $\delta_l$ 、 $\delta_s$  の最大値を  $\delta$  として以下のように与える。

$$\delta(T_1, T_2) = \max \left\{ \frac{\delta_f(T_1, T_2)}{2}, \frac{\delta_d(T_1, T_2)}{3}, \frac{\delta_l(T_1, T_2)}{4}, \frac{\delta_s(T_1, T_2)}{4} \right\}.$$

## 3 実験

木のデータセットとして、N-glycans、all-glycans、CSLOGS、dblp、SwissProt、TPC-H、Auction、Nasa、Protein、University を用いる。このとき、図 1 は計算困難な計算距離に対する、多項式時間で計算可能かつ最も一般的な編集距離の変種である孤立部分木距離 [3]

と  $\delta$  との差を横軸、その差を持つ木の組の頻度を縦軸としたヒストグラムである。

図 1 より、ばらつきがあるデータセットはあるものの、どれも  $x$  軸において 0 の方に多くデータが固まっている。より正確に編集距離の近似ができる。

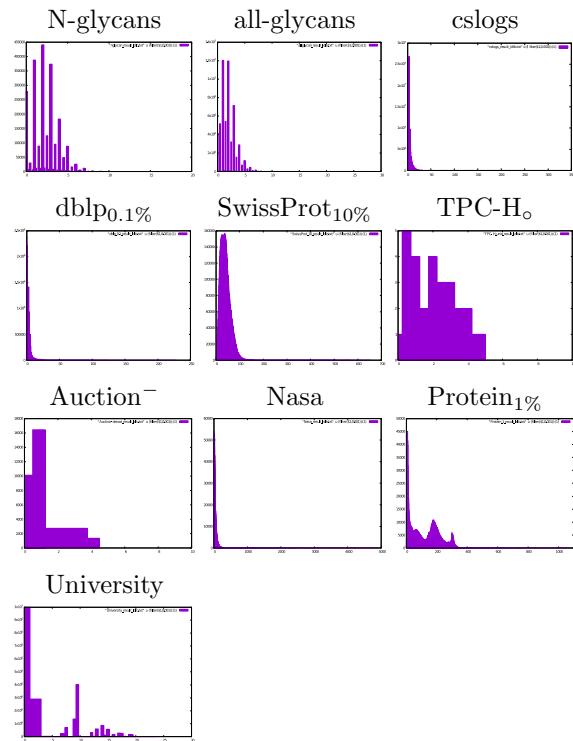


図 1: 孤立部分木距離と距離  $\delta$  の差のヒストグラム

## 参考文献

- [1] T. Aratsu, K. Hirata, T. Kuboyama: *Sibling distance for rooted ordered trees* Proc. ALSIP'08, Working Notes of PAKDD Workshops, 101–112, 2008.
- [2] F. Li, H. Wang, J. Li, H. Gao: *A Survey on Tree Edit Distance Lower Bound Estimation Techniques for Similarity Join on XML Data*, Proc. SIGMOD'13, 29–39, 2013.
- [3] J. T. L. Wang, K. Zhang: *Finding similar consensus between trees: An algorithm and a distance hierarchy*, Pattern Recog. 34, 127–137, 2001.
- [4] K. Zhang, T. Jiang: *Some MAX SNP-hard results concerning unordered labeled trees* Inform. Process. Lett. 49, 249–254, 1994.

		指導教員	平田耕一 教授
学生番号	192C1027	氏名	内野由喜
論文題目	グラム染色画像からの白血球貪食推定の改良		

## 1はじめに

グラム染色 [1] は、感染症の初期診療に用いられる手法である。患者から採取した検体材料に存在する細菌を色素によって染め上げ、色と形状から菌種を推定する。安価で容易に施行できるが、ゴミも一緒に染色され、目視による判別を行うため、熟練した検査技術が必要である。現在、熟練した検査者が不足しているため、機械学習によるシステムによって検査を支援することで、その問題を解決することができる。

吉原 [3] は、グラム染色画像からの白血球貪食の分類と検出を行った。そして、分類は高い精度で実現できたが、YOLOv5 による検出の精度はそれほど高くなかった。特に、画像数が少ないことが課題として挙げられた。

そこで本研究は、検出精度の向上を目的とし、画像数の増加と新たな検出器を導入する。

## 2 物体検出

本研究では、吉原 [3] と同様に、白血球が貪食している貪食像クラス、貪食しているようで実際にはしていない擬貪食像クラス、貪食していない非貪食像クラスの 3 クラスに分ける。そして、喀痰のグラム染色画像 211 枚から、学習用画像 6 割、検証用画像 2 割、テスト用画像 2 割によるホールドアウト検証により結果を評価する。矩形数の内訳は、貪食クラスが 254 個、擬貪食クラスが 256 個、非貪食クラスが 831 個である。画像数の増加では、画像の絶対数を増やすと共に、学習を行う際に画像数を 6 倍に水増しする。検出器は、[3] で最も高性能だった YOLOv5 に出力層を追加して、改良したモデルに加えて、最新モデルである YOLOv7[2] を用いて 3 クラスの物体を検出し、その性能を評価する。

物体検出の評価指標として、領域の重なり具合を表す IoU(Intersection over Union) を用いる。IoU の閾値が 0.5 以上の平均適合率である AP50 と IoU が 0.75 以上の平均適合率である AP75、さらに、IoU を 0.5 から 0.95 まで 0.05 刻みで変化させたときの平均適合率の平均である mAP の 3 つを使用する。

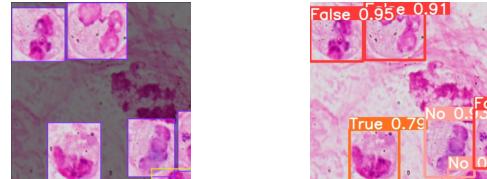
## 3 検出結果

表 1 は、各検出モデルによる AP50, AP75, mAP の値である。また、図 1 と図 2 は検出例であり、赤が貪食

像 (True), 青が擬貪食像 (False), 黄が非貪食像 (No) を表している。

表 1: 白血球貪食の検出結果

モデル	AP50	AP75	mAP
YOLOv5s	0.649	0.545	0.443
YOLOv5x	0.658	0.575	0.484
YOLOv7	0.654	0.514	0.448
YOLOv7x	0.656	0.499	0.453



正解矩形 検出結果

図 1: 誤検出が存在する検出例



正解矩形 検出結果

図 2: 未検出が存在する検出例

表 1 から、YOLOv5x のモデルが一番高く、AP50 が 0.658、AP75 が 0.575、mAP が 0.484 であった。しかし、これらの値は、[3] の結果よりは高いものの、際立ってよい値とまでは至らなかった。図 1 の左下の白血球のように誤検出することや、図 2 のように未検出が多く存在したことが原因の一つとして考えられる。

## 参考文献

- [1] 田里大輔、藤田次郎、できる！見える！活かす！グラム染色からの感染症診断～検体採取・染色・観察の基本とケースで身につく診断力（レジデントノート別冊）、羊土社（2013）。
- [2] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, arXiv:2207.02696(2022).
- [3] 吉原教平、グラム染色画像からのキャンピロバクターと白血球貪食の分類と検出、九州工業大学大学院情報工学府修士論文（2021）。

		指導教員	平田 耕一 教授
学生番号	192C1043	氏名	菓子野 右京
論文題目	グラム染色画像からのグラム陰性桿菌検出の改良		

## 1 はじめに

グラム染色とは、1884年にデンマークの学者である Hans Christian Joachim Gram によって考案された微生物塗抹検査法で染色性（藍色・紫色、赤色・桃色）や形状（球状、棒状、単体、対など）から菌種を推定する。本研究ではグラム陰性桿菌に対して深層学習を用いた検出を行う。グラム陰性桿菌とはグラム染色によって赤、ピンク色に染まる細長い棒状の菌のことであり、本研究では緑膿菌 (PA), キャンピロバクター (CA), 大腸菌 (EC), ピロリ菌 (HP), インフルエンザ菌 (HI), 肺炎桿菌 (KP) を取り扱う。加えてグラム陽性桿菌であるウェルシュ菌 (CP) とコリネバクテリウム (CD) も扱う。図にグラム陰性桿菌とグラム陽性桿菌の例を示す。

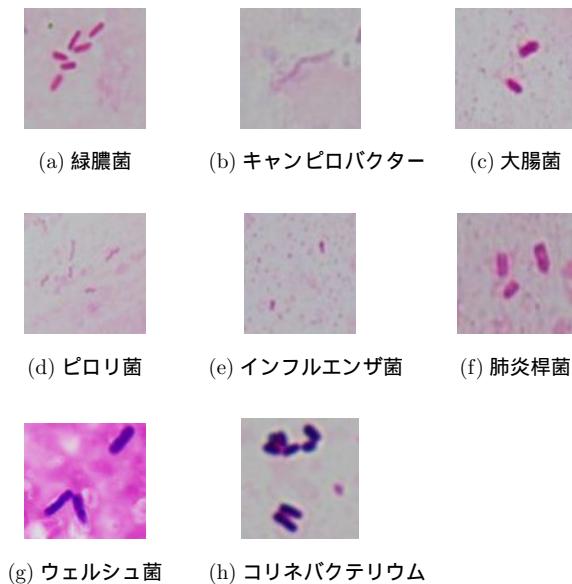


図 1: グラム陰性桿菌とグラム陽性桿菌

## 2 物体検出

物体検出手法として YOLOv5 [1], YOLOv7 [3] を用いる。YOLO は物体検出アルゴリズムの一つであり、YOLOv5, YOLOv7 は YOLO のバージョン 5, 7 である。YOLO では画像を小さな正方形の領域に分割し、その領域ごとに物体の検出を行う。

実験には専門家の監修のもとアノテーションした画像 483 枚を使用する。表 1 は菌種ごとの画像数と矩形数である。この画像を菌種別に学習用画像 6 割、検証

用画像 2 割、テスト用画像 2 割に振り分け、学習用画像には  $x$  軸反転、 $y$  軸反転により画像数を 3 倍に拡張する。また、検出器の改良として、YOLOv5s, YOLOv5x, YOLOv7, YOLOv7x に対して、全画像の 4 分割 (4D), 学習用画像の回転 (R), それらの組合せ (4D+R), 出力層の増加 (O-layer) を適用する。

表 1: 菌の画像枚数と矩形数

菌名	画像枚数	矩形数
緑膿菌	97	6527
キャンピロバクター	93	625
大腸菌	98	5371
ピロリ菌	50	1804
インフルエンザ菌	31	6189
肺炎桿菌	42	3287
ウェルシュ菌	13	156
コリネバクテリウム	59	3204

## 3 検出結果

表 2: 最良の検出器

菌	改良	適合率	再現率	AP50	AP75	mAP
PA	v7x-4D+R	0.736	0.676	0.705	0.397	0.385
CA	v5s-O-layer	0.726	0.499	0.564	0.301	0.290
EC	v7-4D	0.876	0.874	0.925	0.681	0.581
HP	v7-4D	0.755	0.692	0.770	0.524	0.445
HI	v7-4D	0.864	0.715	0.878	0.311	0.427
KP	v5x-4D	0.863	0.888	0.889	0.775	0.615
CP	v5s	0.811	0.992	0.979	0.873	0.734
CD	v7x-4D+R	0.828	0.907	0.946	0.737	0.609

菌ごとの最良の検出器を表 2 に示す。表 2 より [2] の結果に比べて AP50 が最悪でも 0.229, AP75 が最悪でも 0.159 上回ることができた。またキャンピロバクターとウェルシュ菌を除くすべての検出器は、全画像の 4 分割を適用している。このことから、全画像の 4 分割により、小さなグラム陰性桿菌を見逃すことなく検出できていると考えることができる。

## 参考文献

- [1] G. Jocher: YOLOv5, <https://github.com/ultralytics/yolov5>, 2020.
- [2] 王春妍: グラム染色画像からのグラム陰性桿菌の検出, 九州工業大学情報工学部知能情報工学科卒業論文, 2022.
- [3] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao: YOLOv7, arXiv:2207.02696, 2022.

		指導教員	平田 耕一 教授
学生番号	192C1132	氏名	船津 遼太郎
論文題目	グラム染色画像からのゲックラー分類推定の改良		

## 1はじめに

グラム染色は、検体中の細菌を染め上げることで菌の種類を特定する手法であり、感染症の初期治療に用いられる。検体が喀痰である場合、1,000 倍視野の顕微鏡検査に値する品質かどうかを判定する方法として、100 倍視野顕微鏡のグラム染色画像から、扁平上皮細胞と白血球の個数に応じて、表 1 のように 6 つのクラスに割り当てるゲックラー分類が知られている。

表 1: ゲックラー分類

クラス	扁平上皮細胞 [個]	白血球 [個]	品質
1	>25	<10	低品質
2	>25	10-25	低品質
3	>25	>25	低品質
4	10-25	>25	高品質
5	<10	>25	最高品質
6	<25	<25	不明

## 2 ゲックラー分類の推定方法

ゲックラー分類の推定として遠藤 [1] は、Faster R-CNN と YOLOv5x [3] の学習モデルを利用して、100 倍視野のグラム染色画像における扁平上皮細胞と白血球を検出することで、ゲックラー分類を推定した。そこで本研究では、YOLOv5 に加えて YOLOv5 を改良した YOLOv5-O のそれぞれの 4 つの学習モデル、および、YOLOv7 [4] の 2 つの学習モデルを利用してゲックラー分類を推定し、それらの推定結果を比較する。

## 3 ゲックラー分類の推定の結果

表 2 は、YOLOv5l-O, YOLOv7x によるゲックラー分類の推定に対する、扁平上皮細胞と白血球の検出における AP50, AP75, mAP の値である。白血球は扁平上皮細胞と比べて非常に小さく、白血球と同じ大きさのゴミも存在しているため、ゴミを白血球と誤って検出してしまい、値が低くなったと考えられる。

表 3 は、YOLOv7x によってゲックラー分類を行い、画像の品質ごとに分類した結果である。

表 2: 学習モデルごとの 3 つの指標の平均

	学習モデル	AP50	AP75	mAP
EHC	YOLOv5l-O	0.715	0.616	0.520
	YOLOv7x	<b>0.908</b>	<b>0.751</b>	<b>0.635</b>
WBC	YOLOv5l-O	<b>0.424</b>	<b>0.119</b>	<b>0.183</b>
	YOLOv7x	0.367	0.062	0.140

表 3: YOLOv7x のゲックラー分類の推定結果 [枚]

正解 推定	1-3	4	5	6
	1-3	90	1	0
4	0	12	0	0
5	0	6	29	1
6	1	0	2	22

表 3 より、テスト画像の品質が本来とは異なる品質のクラスに推定された画像が数枚あった。このうち、正解クラスが 1-3 と 4 の誤推定は扁平上皮細胞の誤検出や未検出が、正解クラスが 5 と 6 の誤推定は白血球の誤検出や未検出が原因だと考えられる。

## 参考文献

- [1] 遠藤徳: グラム染色画像からのゲックラー分類の推定, 九州工業大学大学情報工学部知能情報工学科専攻卒業論文, 2022.
- [2] R. Geckler, D. Gremillon, C. McAllister, C. Ellenbogen: *Microscopic and bacteriological comparison of paired sputa and transtracheal aspirates*, J. Clin. Microbiol. **6**, 396–399, 1977.
- [3] G. Jocher, YOLOv5, <https://github.com/ultralytics/yolov5>.
- [4] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao: *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*, Proc. CVPR'22, 2022.

		指導教員	平田 耕一 教授
学生番号	192C1169	氏名	山本有藍
論文題目	新型コロナウイルス アミノ酸配列における近隣位置解析		

## 1はじめに

本研究では、新型コロナウイルス (SARS-CoV-2) の4つの構造タンパク質であるスパイクタンパク質 (S), 膜タンパク質 (M), エンベロープタンパク質 (E), ヌクレオカブシドタンパク質 (N) のアミノ酸配列に対して、位置間の距離を計算し、変異位置の近隣位置を解析する。

## 2アミノ酸配列の近隣位置解析

アミノ酸配列  $x$  の  $i$  番目と  $j$  番目のアミノ酸をそれぞれ  $x[i]$  と  $x[j]$  とする。このとき、 $hyd(x[i])$  で  $x[i]$  のハイドロバシー指標 [2],  $iso(x[i])$  で  $x[i]$  の等電点 [2] とする。また、 $pam(x[i], x[j])$  で  $x[i]$  と  $x[j]$  の PAM250 [1, 2] の値とし、 $blosum(x[i], x[j])$  で  $x[i]$  と  $x[j]$  の BLOSUM50 [1, 2] の値とする。さらに、これらを用いたアミノ酸  $x$  における位置  $i$  と  $j$  間の距離  $D_k$  ( $1 \leq k \leq 5$ ) を以下のように定義する。

$$\begin{aligned} D_1(x[i], x[j]) &= \begin{cases} 0 & x[i] = x[j] のとき, \\ 1 & それ以外. \end{cases} \\ D_2(x[i], x[j]) &= |hyd(x[i]) - hyd(x[j])|. \\ D_3(x[i], x[j]) &= |iso(x[i]) - iso(x[j])|. \\ D_4(x[i], x[j]) &= pam(x[i], x[j]). \\ D_5(x[i], x[j]) &= blosum(x[i], x[j]). \end{aligned}$$

そして、アミノ酸配列集合  $S$  における位置  $i$  と  $j$  の距離  $d_k(i, j)$ 、および、その正規化  $S_k(i, j)$  ( $1 \leq k \leq 5$ ) を以下のように定義する。ここで、 $\max d_k$  と  $\min d_k$  は  $d_k$  の最大値と最小値である。

$$\begin{aligned} d_k(i, j) &= \sum_{x \in S} D_k(x[i], x[j]) \\ S_k(i, j) &= \frac{d_k(i, j) - \min d_k}{\max d_k - \min d_k} \times 100. \end{aligned}$$

## 3 実験

実験は、NCBI で公開されている新型コロナウイルスのアミノ酸配列を使用する。2021年4月1日から2022年5月31日までに収集された配列をデルタ株のアミノ酸配列、2021年11月1日から2022年7月29日までに収集された配列をオミクロン株のアミノ酸配列として分類する。デルタ株とオミクロン株のアミノ酸配列の本数はそれぞれ 834,870 本と 393,655 本である。スパイクタンパク質の主要な変異位置の中から、デ

ルタ株は 19, 452, 681、オミクロン株は 69 の位置に対して、距離を計算し、正規化した距離が近いものを近隣位置とした結果を表 1 と表 2 に示す。

表 1: デルタ株のスパイクタンパク質の変異位置との近隣位置

S 変異位置	距離指標	1	2	3
19	$d_1$	E(38)	N(277)	M(105)
	$d_2$	S(681)	S(190)	S(905)
	$d_3$	S(452)	S(681)	S(190)
	$d_4$	S(156)	S(190)	S(681)
	$d_5$	S(452)	S(190)	E(38)
452	$d_1$	M(105)	N(277)	S(995), N(107)
	$d_2$	S(19)	S(681)	S(190)
	$d_3$	S(19)	S(681)	S(190)
	$d_4$	S(19)	S(157)	N(681)
	$d_5$	S(19)	S(157)	S(190)
681	$d_1$	N(107), N(277)	M(150)	S(995), S(983), S(1000), N(149), M(105)
	$d_2$	S(190)	S(905)	S(355), M(72), M(150)
	$d_3$	S(19)	S(452)	S(190)
	$d_4$	S(190)	M(198)	S(355)
	$d_5$	S(190)	M(198)	N(277), M(105)

表 2: オミクロン株のスパイクタンパク質の変異位置との近隣位置

S 変異位置	距離指標	1	2	3
69	$d_1$	N(356)	S(1064)	S(1048)
	$d_2$	N(356)	S(1048)	S(1064)
	$d_3$	S(1064)	N(356)	S(1271)
	$d_4$	N(356)	S(1064)	S(1048)
	$d_5$	N(356)	S(1048)	S(1064)

表 1 の変異位置 19 に対し、S(452), S(681) といったデルタ株で主要なスパイクタンパク質の変異位置が確認された。続いて、変異位置 452 に対し、S(19) が確認でき、変異位置 19 の近隣位置であった S(681) も確認された。さらに、S(681) では、等電点  $d_3$  で S(19), S(452) の 2 つが確認できた。

表 2 の変異位置 69 の近隣位置では、全ての非類似度で N(356), S(1064) が確認された。この 2 つの近隣位置は、アミノ酸の文字列と性質が似ている位置だと考えられる。

## 参考文献

- [1] 阿久津達也, バイオインフォマティクスの数理とアルゴリズム, 共立出版 (2007).
- [2] 有坂文雄, バイオサイエンスのための蛋白質科学入門, 裳華房 (2004).

		指導教員	徳永 旭将 准教授
学生番号	192C1012	氏名	石橋 涼輔
論文題目	リザバーコンピューティングモデルを用いた 線虫の嗅覚神経細胞における膜電位応答の再現		

## 1 はじめに

一般的なニューラルネットワークにおいては、個々のニューロンの応答は単純な線形変換と非線形変換の組み合わせでモデル化されている。このようなモデルを形式ニューロンと呼ぶ。しかし、現実の生物の神経細胞応答はより複雑であることも多く、形式ニューロンでは神経細胞が生物の学習に果たす機能の本質を正しく模倣できていないと考えられる。実際に、線虫 *C. elegans* (以後、線虫と称する) の嗅覚神経細胞 AWA のライブセルイメージングによって、AWA は形式ニューロンとは異なり、時間的に変化する環境情報を鋭敏かつ定量的に表現できることが示唆された [1]。AWA のように複雑な応答を返す神経細胞の挙動を、時系列データの処理が可能な小規模で計算コストの低いモデルで再現することができれば、それらを複数結合してより生体の神経系に近いネットワークを構成できる。そこで本研究では、リザバーコンピューティングモデルの一つである Echo State Network (ESN) [2] を用い、線虫の嗅覚神経細胞の一つである AWA の膜電位応答の再現を行う。

## 2 研究方法

本研究の目的は、ESN を用いて線虫の嗅覚神経細胞の外部刺激に対する神経応答を再現することである。実験に使用した ESN は時系列データの処理に適した機械学習モデルであり、出力層のパラメータでのみ学習を行う。そのため、同規模の Recurrent Neural Network と比べ学習時の計算コストを抑えられることが特徴の一つである。

実験で再現する神経応答として、線虫の鼻先に匂い物質の一種であるジアセチル水溶液を流出させたときの膜電位応答を観測したものを用いる。鼻先に流出させたジアセチル水溶液は原液を希釈したものである。これらのデータは蛍光顕微鏡を用いて光学的に計測したライブセルイメージングデータであるが、本研究ではこのデータから時系列データを作成し、実験に使用した。

本実験では、匂い刺激を模した人工波形に対する ESN の出力が目標データである線虫の膜電位応答を再現するように、ESN の出力層のパラメータを最適化する。次に、様々な人工波形に対して学習済み ESN が適當

な応答を示すかを確認し、ESN による膜電位応答が再現可能であるかを検証する。

## 3 実験

本実験では、学習における目標データとして線虫の膜電位応答の観測データを用いた。ただし、観測データには光退色によって膜電位応答とは関係ないトレンドが発生するため、前処理としてその除去を行った。また、入力データにはジアセチル水溶液の濃度変化を模した人工波形を使用した。今回は入力パターンの異なる 2 種の学習データを用いて、学習したモデルの未知データに対する応答の評価(実験 1)と学習したモデルの濃度依存性の再現度の評価(実験 2)を行った。

結果として、実験 1 ではモデル学習時に与えられていない未知の入力データに対しても、濃度変化時の特徴的な応答を再現することができた。また、実験 2 では入力データの大きさに応じた応答の違いを再現することができた。ただし、どちらの実験においても、入力の大きさが変化しない区間で応答が上手く再現できなかった。

## 4 まとめ

本研究では、ESN を用いてジアセチル水溶液に対する線虫の嗅覚神経細胞 AWA の膜電位応答の再現を行った。実験結果より、ジアセチル水溶液の濃度変化時の特徴的な応答の再現はできたが、入力が定常的な状況における膜電位の変動は再現できなかった。今後の展望として、この問題を改善し ESN の汎化性能を向上させるため、より多くの学習データの用意やデータに対する適切な前処理の摸索を行う。そのほか、ESN の出力層を複数使用して神経細胞の複数のイオンチャネルを再現することで、生物学的に妥当な構造を持つモデルを提案することを目指す。

## 参考文献

- [1] Jamine Hooi-Min Teo, Itsuki Kurokawa, Yuuki Onishi, Noriko Sato, Tomohiro Kitazono, Terumasa Tokunaga, Manabi Fujiwara, and Takeshi Ishihara. Behavioral forgetting of olfactory learning is mediated by interneuron-regulated network plasticity in *caenorhabditis elegans*. *Eneuro*, 2022.
- [2] 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*, 148(34):13, 2001.

		指導教員	徳永 旭将 准教授
学生番号	192C1077	氏名	城間 亮
論文題目	外部視覚注視機構による深層学習の異常検出能力と 注視領域生成方法の関係について		

## 1 研究の背景と目的

異常検知とは、データの中から期待される振る舞いをしないデータを見つけ出す技術であり、近年深層学習を用いた手法がよく用いられる。深層学習を用いた手法には教師あり異常検知、教師なし異常検知が存在する。これらの手法には、データセットの持つ正常異常のバランスの悪さや異常度の設計などの問題点が存在する。

Layer-Wise External Attention Network (LEA-Net) [1] は、画像分類タスクにおける Convolutional Neural Network (CNN) の性能を向上させる Visual Attention Mechanism のアイデアを活用した異常検知手法である。LEA-Net は教師なし異常検知での結果を異常の事前知識として教師あり異常検知に活用した Layer-wise External Attention Mechanism を使用することで、教師ありと教師なしの異常検知の特徴をうまく統合することができる。LEA-Net の課題点として、データセットに応じた適切な Attention の適用方法の模索が必要となっており、本研究では異なる教師なし異常検知手法で生まれる結果の違いによる異常検知性能の変化の傾向を探る。

## 2 研究方法

LEA-Net での異常検知に用いる教師なし異常検知は、画像を再構成し、元画像との差分を取るというものである。比較する手法は、そのまま画像を再構成する方法、白黒画像をカラー画像にする色の再構成を行う方法、そして画像の一部分を隠しそれを復元させる欠損補完を活用した再構成を行う手法の 3 つである。また、LEA-Net の性能向上を最大化した場合を見るために、異常領域の正解画像を理想的な教師なし異常検知の結果とした場合も比較する。

## 3 実験

本実験では、MV AD Dataset [2] と Plant Village [3] での異常検知性能を比較する実験を行った。これらのデータセットは、正常画像に対する異常画像の割合は 30%ほどであり、各データの総数はおよそ 400 枚である。教師なし異常検知手法は画像再構成、色の再構成、欠損補完による再構成で、それに加えて正解画像の 4 種で行った。LEA-Net に用いる Layer-wise External Attention Mechanism のベースとなる CNN には

ResNet18 を使用した。実験には層化 5 分割検証を行い、評価指標として各分割の F 値を平均したものを使用した。加えて、MVTec に関しては Layer-wise External Attention Mechanism に加えて Self-Attention Module を導入した場合についても検証を行った。

再構成方法の比較では、どの手法でもベースラインより性能が向上していた。教師なし異常検知手法の比較をすると、各データの F 値の上昇幅を平均すると色の再構成は 9.0%、欠損補完による再構成は 2.5%、画像の再構成は 2.3%であり、色の再構成が一番性能を上げる手法となっていた。また、正解画像を利用した場合はどのデータでも性能が大きく上昇しており、各データの平均では 13.2% の上昇幅であった。Self-Attention Module を導入した場合では、先ほどの性能をさらに引き上げる結果となり、3 手法においては倍近い性能の上昇幅を見せ、正解画像でも性能の向上が見られた。

## 4 今後の課題と展望

本研究では、LEA-Net で使用される Attention Mechanism の解明のために、異なる異常領域の生成方法による異常検知性能の比較を行った。結果は色の再構成を行った場合が異常検知性能を一番引き出せる場合が多くかった。また、追加で Self-Attention の導入を行った場合にも性能を向上させることができた。今後の展望として、より多様なデータセットを用いた実験や、Layer-wise External Attention Mechanism と Self-Attention との関係性の模索が必要である。また、LEA-Net に用いる CNN は他のモデルに変えることができるので、最新の CNN を用いることでさらなる性能向上も達成可能であると考えられる。

## 参考文献

- [1] Katafuchi, Ryoya, and Terumasa Tokunaga. “LEA-Net: Layer-Wise External Attention Network for Efficient Color Anomaly Detection.” arXiv, September 12, 2021. <https://doi.org/10.48550/arXiv.2109.05493>.
- [2] Bergmann, Paul, Michael Fauser, David Sattlegger, and Carsten Steger. “MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection.” In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 9584–92. Long Beach, CA, USA: IEEE, 2019. <https://doi.org/10.1109/CVPR.2019.00982>.
- [3] Hughes, David P., and Marcel Salathe. “An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics.” arXiv, April 11, 2016. <https://doi.org/10.48550/arXiv.1511.08060>.

		指導教員	徳永 旭将 准教授
学生番号	192C1152	氏名	三山 晃季
論文題目	分布を仮定しない半教師あり画像セグメンテーションに対する mixup の有用性の検証		

## 1 はじめに

深層学習の登場により画像認識分野は目覚ましい発展を遂げており、様々な教師あり画像セグメンテーション手法が提案されている。これらの手法において、セグメンテーション性能の向上には、大量のアノテーションされたデータを必要とする。しかし、大量のデータを用意するためには作業コストが高いという問題がある。そこで、少量の部分的アノテーション画像を使用した半教師あり画像セグメンテーション技術:Cost-effective Segmentation with Partial Patch Annotations(CoSPA)[1] が存在する。CoSPA では部分的なアノテーションは全てユーザが与えることを想定している。しかし、データ拡張を導入することでさらにユーザの作業負担を低減できる可能性がある。

## 2 mixup-CoSPA

本研究では、ユーザが作成した部分的アノテーションに対し、データ拡張手法である mixup[2] を導入することで CoSPA のセグメンテーション性能を向上させることができるか、検証を行った。特に、半教師あり 2 値分類のコスト関数として提案されている拡張 Focal Loss とクラス事前罰則項を組み合わせたコスト関数[3]について、mixup に適した形に拡張した。具体的には、mixup 導入時のクラス事前確率を、モデルが予測する混合ラベルの期待値を用いて定式化した。ラベルなしデータにおける正例の割合を  $\hat{\pi}$  とすると、ラベルなしパッチ画像から 1 組を選択するとき、確率変数は次のようになる。

$$P_{unlabel} = \begin{cases} \hat{\pi}^2 & \text{if } (y_i, y_j) = (1, 1) \\ (1 - \hat{\pi})^2 & \text{elseif } (y_i, y_j) = (0, 0) \\ 2\hat{\pi}(1 - \hat{\pi}) & \text{otherwise} \end{cases} \quad (1)$$

ここで  $\{y_i, y_j\} \in \{0, 1\}$  はラベルなしデータに対する正例、負例を仮定した場合のフラグであり、正例の場合は 1 を、負例の場合は 0 をとる。また、これらの確率変数のとる値  $y$  は

$$y = \begin{cases} 1 & \text{if } (y_i, y_j) = (1, 1) \\ 0 & \text{elseif } (y_i, y_j) = (0, 0) \\ \lambda y_i + (1 - \lambda) y_j & \text{otherwise} \end{cases} \quad (2)$$

と表せる。以上の式 1 と 2 の各組の積の和より、ラベルなしデータにおける出力の期待値は  $\hat{\pi}$  と表すことができる。よって、クラス事前確率  $\pi$  との一致度から導出する正則化項  $r$  は

$$r := \log(1 - |\pi - \hat{\pi}|) \quad (3)$$

のように表すことができる。ここで  $|\cdot|$  は絶対値記号である。 $r$  は、 $\pi$  と  $\hat{\pi}$  が一致する場合、損失は 0 になる。よって、Binary Crossentropy( $l_{BCE}$ ) と式 3 を用いたコスト関数  $l$  は

$$l := l_{BCE} + \delta r \quad (4)$$

と表せる。 $\delta$  は  $r$  の大きさを調整するハイパーパラメータである。

## 3 実験

CoSPA に mixup を導入することの有用性を検証するため、実データを用いたセグメンテーション比較実験を行った。比較手法として、パッチ画像の PN 分類、パッチ画像の PN 分類 + mixup、パッチ画像の半教師あり 2 値分類 + mixup + クラス事前罰則項、パッチ画像の半教師あり 2 値分類 + mixup + 拡張されたクラス事前罰則項の 4 通りについて性能を評価した。実験の結果、パッチ画像の PN 分類 + mixup が最もセグメンテーションの性能が良い結果となった。

## 4 まとめ

本研究では、分布を仮定しない半教師あり画像セグメンテーションに対する mixup の有用性の検証を行った。本実験の結果、提案した手法では、教師あり学習に mixup を使用した場合と比較するとセグメンテーションの性能が低下することがわかった。これは、mixup という手法自体が正則化の効果を持っており、更にクラス事前罰則項を加えたことで過剰な正則化が発生したためと考えられる。

## 参考文献

- [1] Tokunaga Terumasa, Nakanishi Keiichi, Katafuchi Ryoya, Miyama Kohki, and Watanabe Shou. Cost-effective deep image segmentation with partial patch annotations based on semi-supervised learning. 2022.
- [2] Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412*, 2017.
- [3] 徳永旭将 中西慶一. 半教師あり二値分類のためのクラス事前確率を用いたコスト関数の提案. *MIRU2022*, 2022.

		指導教員	徳永 旭将 准教授
学生番号	192C1180	氏名	渡邊 渉
論文題目	部分的アノテーションを用いた半教師あり学習による 3D のセグメンテーションの研究		

## 1 はじめに

深層学習を用いた教師あり学習による手法により、画像セグメンテーションタスクは発展を遂げた。それらの手法は、モデルの汎化性能向上のために、訓練時に大量のデータとピクセルレベルのアノテーションを必要とする。しかし、ピクセルレベルのアノテーション作成は作業コストが高いという問題がある。3D 空間でのボクセルレベルのアノテーション作成が必要となるため、2D セグメンテーションよりもアノテーション作成の作業コストが桁違いに高い。そのため近年では、訓練時に大量のデータとピクセル/ボクセルレベルのアノテーションを必要としない、半教師あり学習による画像セグメンテーション手法が研究されている。

## 2 先行研究

CoSPA (Cost-effective Segmentation with Partial Annotations) [1] は、ピクセルレベルの GT (Ground Truth) を必要としない、パッチ画像による部分的なアノテーションと半教師あり学習に基づくセグメンテーション手法である。なお、CoSPA は 2D 画像を対象としているため、ここでは本手法を 2D-CoSPA と呼ぶ。2D-CoSPA では、パッチ画像による部分的なアノテーション用いた訓練とラベル無しデータを用いた PNU 分類 [2] を行うことにより、アノテーションの作業コストを下げる。パッチ画像は Positive (前景), Negative (背景), Unlabeled (ラベル無し) の 3 クラスから構成される。訓練段階では PNU 分類の損失関数を最小化することで、分類モデルを作成する。テスト段階ではラスタスキャンを行い、パッチ単位での推論を行い、最終的にピクセルレベルのセグメンテーションを行う。

## 3 提案する 3D セグメンテーション手法

本研究では 2D-CoSPA を 3D データに対応するよう拡張した 3D セグメンテーション手法、3D-CoSPA を提案する。3D-CoSPA は、2D-CoSPA を 3D データに対応するよう拡張した手法である。2D-CoSPA ではパッチ画像単位でセグメンテーションを行うのに対し、3D-CoSPA では奥行や時間軸を追加した、立体パッチ画像単位でセグメンテーションを行う。立体パッチ画像を用いることで、3D データの空間構造や時間構造を特徴量とすることが期待される。

## 4 実験

提案する 3D セグメンテーション手法の有用性を検証するために、立体画像内に正八角錐を描画し、セグメンテーション対象と同じ輝度の 1vx のノイズを段階的に付与したデータと、オーロラを 3 秒ごとに撮影した動画像を用いて実験を行なった。時間構造利用の有用性の検証のため、時間構造を訓練に用いない 2D-CoSPA を比較手法として用いた。また、アノテーションコストの比較のため、教師あり学習によるセグメンテーション手法である U-Net[3] を用いた。

3D-CoSPA は 50% 程度のノイズでは影響を受けずに立体画像のセグメンテーションが可能であった。3D-CoSPA が時間構造を利用しない 2D-CoSPA に比べて IoU が高いことから、時間構造を利用することにより、汎化性能が向上した。また、3D-CoSPA は U-Net に訓練画像 64 枚用いて訓練を行なった場合と同程度の IoU であった。3D-CoSPA ではアノテーションの作成に 15 分要するのに対し、U-Net で 3D-CoSPA と同程度の性能を出すために必要な 64 枚のピクセルレベルのアノテーションを作成すると 128 分要することから、アノテーションコストは 3D-CoSPA が優位であった。

## 5 まとめ

本研究ではピクセル単位の GT を必要としない、立体パッチ画像による部分的なアノテーションを用いた半教師あり学習による 3D セグメンテーション手法の提案を行なった。3D-CoSPA によって立体画像のセグメンテーションが可能であることや、時間構造利用の有用性、そして教師あり学習に比べてアノテーションコストの優位性があることを確認できた。

## 参考文献

- [1] Terumasa Tokunaga, Keiichi Nakanishi, Ryoya Katafuchi, Kohki Miyama, and Shou Watanabe. Cost-effective deep image segmentation with partial patch annotations based on semi-supervised learning. 2022.
- [2] Tomoya Sakai, Marthinus Christoffel Plessis, Gang Niu, and Masashi Sugiyama. Semi-supervised classification based on classification from positive and unlabeled data. In *International conference on machine learning*, pages 2998–3006. PMLR, 2017.
- [3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18, pages 234–241. Springer, 2015.

		指 導 教 員	徳永 旭将 准教授
学生 番 号	192C2078	氏 名	早川 季寿
論 文 題 目	異常検知精度の向上を目的とした 外部注視領域の特微量選択処理の比較検証		

## 1 はじめに

機械学習が進歩したことにより、異常検知技術に深層学習手法が用いられるようになった。機械学習による異常検知は、検査にかかる人的コストの削減や、人為的ミスによる見逃しを防ぐ上で重要な技術である。ところが、深層学習のために十分な量の異常データを収集することは、一般に作業コストが高い。一方、画像認識分野では、学習の効率性を向上させる技術として視覚的注視機構が注目されている、特に、自身の中間出力である特微量に基づき注視誘導を行う自己注視機構は、画像認識性能を向上させることに成功している。

近年、自己注視機構とは異なる視覚注視機構：外部視覚注視機構と、それに基づくモデル Layer-wise External Attention Network (LEA-Net) [2] が提案された。外部注視機構は、自己注視機構と異なり、外部のネットワークから得た特微量 (Attention Map) を用いて注視誘導を行う。そのため、注視点の違いから、単独のモデルでは実現できない多様な性能向上が期待できる。一方、Attention Map が異常領域を過剰に強調してしまうこともあり、常に異常検出性能の向上に繋がるとは限らない。そこで本研究では、外部視覚注視機構により、より安定的に異常検出性能を向上させるために必要な Attention Map 生成過程の特微量選択について検討を行う。

## 2 先行研究

Attention Branch Network (ABN) は、注目点を誘導するための特微量として、自身のモデルから作成する Attention Map を使う。自己注視機構は、自身のモデルを分岐させて得た特微量を用いて注視誘導を行う。しかし、性能向上は自身のモデルに依存しているため、大幅な性能向上は期待できない。

LEA-Net は、入力に対する Anomaly Map を生成する外部ネットワークである教師なしモデルと 2 つのネットワークから構成される外部注視機構を持つ教師ありモデルを統合した異常検知手法である。自己注視機構と違い、他モデルから注視誘導を行い、注視点の違いを利用し性能向上を期待する。

## 3 提案する特微量選択処理

提案する特微量選択処理は、正規化、大津の二値化、Attention Map に対する学習パラメータの付与の 3 種類

の特微量選択処理を検証する。

正規化では、外部から得た特微量の値の幅を  $\vec{x} \in [0.5, 1]$  から  $\vec{y} \in [0, 1]$  にする処理であり、特微量のスケールを変更し、異常領域をより強調し、正常領域の強調を制限する。大津の二値化 [1] により決定した閾値による処理は、閾値を自動で決定し、閾値により異常だと思われる箇所以外の特微量を 0 にすることで、異常領域部分のみの特微量を強調する。

外部から得た特微量に対し、新しいパラメータを付与する処理では、学習時にパラメータを学習することで、外部から得た特微量の重要度を調整する。

## 4 実験

提案する特微量選択処理を用いた LEA-Net の汎化性能の評価のため、比較実験を行った。比較のため、案する 3 通りの手法と先行研究で用いられた手法を用いた。正常画像に対して異常画像が少量な不均衡データセットを用いて学習を行ったところ、提案した手法で性能が向上したことや変化したことを確認出来た。特に、正規化を特微量選択処理に用いた結果が、他の結果よりも優位である場合が多くかった。

## 5 まとめ

本研究では、外部から得た Attention Map に対して、3 通りの特微量選択処理による Attention Map の生成方法を提案し、比較検証を行った。LEA-Net では、正規化に特微量選択処理を行うことで異常検知精度が向上することを確認できた。

## 参考文献

- [1] 大津展之. 判別および最小 2 乗規準に基づく自動しきい値選定法. 電子情報通信学会論文誌 D, Vol. 63, No. 4, pp. 349–356, 1980.
- [2] 片渕凌也. 深層学習による異常検知精度の向上を目的とした外部ネットワークからの注視領域の誘導機構の導入. 2022.

		指 導 教 員	徳永 旭将 准教授
学生 番 号	202C1081	氏 名	高島 功佑
論 文 題 目	Echo State Network による 線虫の嗅覚神経細胞のカルシウムイオン応答		

## 1 はじめに

リカレントニューラルネットワークは、時系列データを処理するために、ニューラルネットワーク内に再帰的な構造を組み込んだ時系列処理モデルである。ただし、従来の多くのリカレントニューラルネットワークは、ネットワーク内の全てのパラメータを再帰的に学習するため、計算時間がかかり学習コストが高い。また生物学の分野では、生物の神経システムの動作原理を解明するための生物モデルとして、線虫 *C. elegans* (線虫) が用いられることがある。線虫は神経細胞がわずか 302 個しかないにも関わらず、嗅覚受容体が 1200 種以上存在し、嗅覚情報処理能力がヒトやイスよりも高いことが知られている。

このような背景により、本研究では生物の個々の神経細胞が実現する情報表現を、数理モデルから明らかにしたいと考えた。またハードウェア実装までできると、身体性まで含めた生物の情報処理の仕組みの解明につながると期待する。よって計算コストが小さな神経細胞モデルを構築する必要があり、そのような技術の一つにリザバーコンピューティングが存在する。リザバーコンピューティングとは、いくつかの特殊なリカレントニューラルネットワークモデルから派生して生じた一般的な概念であり、固定パラメータを用いることで、学習に要する計算速度を高速にした機械学習の技術である。これを用いることで、学習コストの低い時系列処理モデルによる線虫の神経細胞のモデリングが可能になると考えられる。そこで本研究は、リザバーコンピューティングの代表的なモデルの一つである Echo State Network (ESN) と呼ばれるモデル [1] を用いて、生物モデルである線虫の神経活動の再現を行う。

## 2 研究手法

本研究では、線虫の神経活動を ESN により再現する。ここで線虫の神経活動として、匂い刺激であるジアセチル水溶液を与えた時の、線虫の嗅覚神経細胞である AWA のカルシウムイオン応答を用いる。これらのデータは蛍光顕微鏡を用いて、線虫に特定の周波数の光を照射することで光学的に計測したライブイメージングデータである。本研究ではライブイメージングデータから時系列データを作成し、実験における正解データとする。また、ESN はモデルの状態変化の速度を調節

するリーケ率をはじめとした様々なハイパーパラメータを調整することで、時系列パターン認識を行う。本実験では、匂い刺激を模した人工波形に対する ESN の出力が正解データである線虫のカルシウムイオン応答の時系列データを再現するように、リッジ回帰により学習パラメータを最適化する。次に、様々な人工波形に対する学習済み ESN の推論結果が、適当な応答を示すかを確認し、ESN によるカルシウムイオン応答が再現可能であるかを検証する。

## 3 実験

本研究では、匂い刺激であるジアセチル水溶液を模した人工波形を入力データ、カルシウム濃度変化の時系列データを正解データとしてデータセットを作成した。ここで、正解データは異なるジアセチル水溶液の濃度に対する、カルシウム濃度変化の時系列データである。実験結果より、ハイパーパラメータであるリーケ率を調節することで、ジアセチル水溶液の濃度の変化に対するカルシウムイオン応答の変化の特徴をうまく表現できることがわかった。しかし、カルシウムイオン応答の強さの再現がうまくいかない場合や、ジアセチル濃度の強さを変化させる際に、出力データが正解データより少し強くなるような傾向があった。

## 4 まとめ

本研究では、ESN を用いた線虫の嗅覚神経細胞のカルシウムイオン応答の再現を行った。実験結果より、ハイパーパラメータであるリーケ率を変更することで、正解データに対する予測がうまくいくことがわかった。ただし、リーケ率の設定によって、時系列の前半のみ、もしくは後半のみ再現がうまくいくというトレードオフになる場合があった。そのため今後の展望として、リーケ率を時間方向に可変にすることで再現を行うことを目標とする。またリーケ率を時間方向に可変にすることの妥当性の検証も兼ねて、今後の課題として、様々な入力に対して実験を行う必要があるとも考えた。

## 参考文献

- [1] Herbert Jaeger. The “echo state” approach to analysing and training recurrent neural networks-with an erratum note. German National Research Center for Information Technology GMD Technical Report, 148(34):13, 2001.

		指導教員	斎藤 寿樹 准教授
学生番号	192C1005	氏名	有吉 優聖
論文題目	タンパク質連接ネットワークの中心性とランダムコイル指標の関係性調査		

タンパク質は、人間や動物の生命を担う基本的な物質である。人間の体内には約 3 万から 10 万種のタンパク質が存在するとされており、それぞれが独自の機能をもっている。これらのタンパク質は立体構造のかたちも様々である。そして、このかたちは機能と関連があることが知られている。よって、タンパク質の構造を知ることは生命の仕組みを理解する上で大切なことである。

近年タンパク質の立体構造は様々な実験で得られており、タンパク質構造データバンク (Protein Data Bank:PDB) と呼ばれるデータベースに保管されている。このデータベースに保管されている立体構造の多くのデータは、X 線結晶構造解析法と核磁気共鳴 (Nuclear Magnetic Resonance:NMR) 法という 2 つの解析方法によって得られてたものである。X 線結晶構造解析法は解析するにあたってタンパク質を結晶化する必要があり、常温では解析できない。一方、NMR 法は溶液中でタンパク質を解析することができるので、人間の体内の環境に近い状態で解析を行うことができる。しかし、NMR 法によるタンパク質の構造の予測が本当に正しいものであるかを検証する尺度が確立されていない。

ANSURR(Accuracy of NMR Structures Using Random coil index and Rigidity) 法とは、NMR によって予測された立体構造が正しいかどうかを検証するためを開発されているソフトウェアである。ANSURR 法ではランダムコイル指標 (Random Coil Index:RCI) と柔軟な領域と合成部分構造トポグラフィ (Floppy Inclusions and Rigid substructure Topography:FIRST) という手法によって求められる指標を用いている。柔軟性と立体構造の間にはつながりがあるとされており、この 2 つの手法は、ともにタンパク質の各アミノ酸において定義される局所的剛性 (柔軟性) を評価するものである。FIRST は、NMR 法で推定されたタンパク質の立体構造を元に、組合せ剛性理論を用いて導き出す。一方で、RCI は NMR から得られる情報の一つである化学シフトという数値を用いて得られる。しかし、タンパク質によっては一部の原子間の化学シフトが算出できないものがある。その場合は算出されている化学シフトのみを使用して RCI の値が計算される。ゆえに、RCI の信頼性が低くなり、ANSURR の検証スコ

アの信頼性も低くなってしまっている。

そこで、本研究では化学シフトの算出されている割合に関わらず、安定した検証ができる指標を得ることを目的としている。つまり ANSURR に使用されている指標である RCI を別の指標に置き換えることを目的としている。タンパク質連接ネットワークの中心性は柔軟性を表すことが可能であると示唆されている。そこで本実験では、タンパク質連接ネットワークの中心性と RCI との関係性を調べた。そして、RCI をタンパク質連接ネットワークの中心性に置き換えることができないかを調べた。

RCI とタンパク質の中心性の値はタンパク質の各アミノ酸ごとに求められる。本実験では、それぞれのタンパク質に対して、これらの数値のピアソンの積率相関係数を求めた。求めた相関係数の中には 0 という値になるものがあった。相関係数が 0 になっているものは、RCI の全てのアミノ酸に対する値は 0 になっていることがわかった。これは化学シフトが得られていないことにより RCI のデータの欠損が起きていることが影響すると考える。また、今回は個々のタンパク質の中心性と RCI の関係性だけではなく、全体での関係性を調べるために、各タンパク質に対して求めた相関係数の平均値を取った。この値からタンパク質の中心性と RCI の間にはやや相関があることが分かったが、これらの関係性は見出すことができなかった。

今後の課題として、本実験で求めた RCI とタンパク質の中心性の相関係数を化学シフトや ANSURR の値と比較して関連性を調べていく必要がある。また現在 ANSURR に使用されている RCI をタンパク質連接ネットワークの中心性に置き換えた場合にどのような違いがあるのかを知りたい。そこで「RCI と FIRST」と「FIRST とタンパク質連接ネットワークの中心性」との関係性についても調べていく必要があると考える。

		指 導 教 員	斎藤寿樹 準教授
学生 番 号	192C1031	氏 名	大石瑛輝
論 文 題 目	中学校・高等学校用時間割作成アプリケーションの開発 —ZDD を用いた時間割作成アルゴリズムの実用化—		

## 1 はじめに

中学校・高等学校において、時間割作成は教務部に配属された教員が行う。作成期間は学校の規模によるが、1週間を超えることもある。実際に時間割を作成する教員は、手作業かアプリケーションを使用して時間割作成を行っている。手作業では複数人の教員を長時間拘束することになる。また、既存アプリケーションを使用している学校では、入力の複雑さ等により利用者が満足できていないのが現状である。

松尾による ZDD を用いた時間割作成アルゴリズム [1] は、教員数、教科数、1日の时限数、学科数、クラス数、教科名と各教科の必要时限数、教員の担当教科と担当クラスを入力として与えると、手作業より短い時間で、1週間分の時間割を出力する。このアルゴリズムは、同入力・同条件の場合、一度の計算で複数の時間割案を提示することが可能である。

本研究では、時間割作成アルゴリズムを用いて、中学校・高等学校用の時間割を作成するアプリケーションを開発する。このアプリケーションは、時間割作成アルゴリズムに必要な情報を入力すると、1週間分の時間割案を複数出力する。既存アプリケーションで出力される時間割は1つである。それが理想とかけ離れている場合、再計算を行う必要がある。また、出力された時間割が完璧であることは少なく、編集を行う必要がある。その時間は、出力時間割が理想の時間割と異なるほど長くなる。1週間分の時間割案を複数出力する機能は、ユーザーの理想に近い時間割が出力される可能性を高める。したがって、このアプリケーションは再計算時間と出力時間割の編集時間を短縮することができる。

## 2 調査と方針

アプリケーションの開発方針を定めるために既存アプリケーションの調査と現地調査を行った。その結果、既存アプリケーションはデータの入力が複雑であるという問題点が判明した。また苦労点として、アプリケーションで時間割のベースを作成した後の業務にも時間がかかるという点が挙げられた。具体的にその業務とは、ベースを編集して完成に近づける作業や、完成後に県へ提出する作業等である。したがって、アプリケーションの開発方針は、入力の複雑さを解消すること、時

間割作成後の業務時間を短縮することの2点とした。

## 3 実装

調査結果をふまえて、時間割作成業務のコストを減らすためのアプリケーションを開発した。機能は、教員情報設定・単位種類情報設定・時間割出力の3つである。教員情報設定では、教員情報の登録と登録情報の編集を行う。単位種類情報設定では、単位種類情報の登録と登録情報の編集を行う。時間割出力では、データベースに登録された情報をもとにアルゴリズムを動作させ、時間割作成を行う。作成が終了すると、複数の時間割案が表示される。時間割案はスプリットビューで表示され、各案の比較が可能である。また各案に対して、csv ファイルでの出力が可能である。

## 4 おわりに

中学校・高等学校の時間割業務を効率化するためのアプリケーション開発について示した。まず、アプリケーションを開発するために調査を実施し、その調査結果に基づいて開発方針を決定した。結果として、入力情報をもとにアルゴリズムを動作させ、複数の時間割案を表示させるアプリケーションを開発した。これは開発方針である、時間割作成後の業務時間短縮の一助となるアプリケーションであると言える。今後の課題としては、もう一つの開発方針である入力のしやすさを達成することが挙げられる。入力の改善案としては、教員登録の際に、あらかじめクラスと教科を全列出しておき、それを選択することで登録できる方式が考えられる。また、初期起動時に入力する情報を誘導することで、クラスや教科に関する入力を少なくできるのではないかと考える。加えて、アプリケーションを実用的なレベルにするには、標準機能の搭載が必要である。既存アプリケーションに標準搭載されている機能がなければ、強みに注目してもらうことはできない。したがって、標準機能のなかで実装できていない時間割編集機能と、詳細な条件設定の中で時間割アルゴリズムが対応している、連続授業・時限指定授業の設定機能を実装する必要がある。

## 参考文献

- [1] 松尾朋卓. Zdd を用いた時間割作成アルゴリズムの実装. 2021.

		指導教員	斎藤寿樹 准教授
学生番号	192C1067	氏名	後藤廣樹
論文題目	時系列データの離散化による類似度計算手法の提案		

## 1 序論

時系列データは時間とともに変動しているデータで、株価、気温などを始めとして様々な機器による観測で数多くの時系列データが蓄積されている [1]。こうした時系列データからデータの特徴を捉えたり、情報を取り出したりすることを時系列解析といい、こうした膨大な時系列データを効率的に処理する技術が求められている。これまでの時系列解析では、時系列を周期ごとに大きさを見るスペクトル分析や相関関数等を用いた時系列データの類似度計算が行われている。しかし、こうした解析技法は、非定常時系列データに見られるような、時系列データに重畠している複数の周波数特性信号間における相関関係やパターン認識には向きである。

時系列データの多くは様々な周波数の波の重なりと見ることができる。そのため、時系列データから特定の周波数を抽出し、その発生時間を離散的に表現することで、特定の周波数間における相関を容易に見ることが期待できる。本研究では時系列データの周波数を離散化するのに区間グラフと文字列の列を用いる。

区間グラフは数直線上の区間の集合を表現として持つ離散構造である [2]。区間グラフの頂点は区間と対応し、2つの頂点間に辺がある必要十分条件は対応する区間に重なりを持つことである。時系列データにおける各事象の発生時間を区間として捉えることにより、時系列データは区間グラフとして表現することができる。また、文字列の列は、複数の文字列を列で表現した離散構造である。

本研究では、時系列データの類似度を区間グラフと文字列の列の2つの離散構造へとモデル化し、2つのデータに対する類似度を定義する。本手法では2つの離散化したデータの共通となる部分を求めて類似性を表す。

本手法を実装し、人工的な時系列データに対して実験を行った結果、2つの手法ともに、周波数の重なりをもとに時系列データの類似性をある程度評価する事ができた。

## 2 手法

時系列データは複数の周波数特性を持つ波の重なりと見ることができる。観察したい事象により、その周

波数等は異なり、多くの場合事象ごとにその周波数は概ね定まっている。そこで、周波数解析で一定の振幅以上で発生している観察する事象に応じた各周波数の波を抽出する。

区間グラフの手法は、周波数解析により取り出された各周波数の波が発生している時間を区間表現で表し、区間グラフを作成する。グラフ同士の類似性は2つのグラフの最大共通誘導部分グラフから計算する。

文字列の列の手法は、1つの時系列データに対して、一定時間毎に特定の周波数が発生しているかを表現した文字列の列を作成する。時系列データを一定時間毎に分割し、分割した時間の中で発生している周波数を文字として文字列に追加する。文字列の列の距離から類似度を計算する。

## 3 実験

本研究では提案手法の有用性について検証するため、人工的な時系列データを100個作成して実験を行った。2つの提案手法を実装し、全てのデータの組み合わせに対して類似度を計算して比較を行った。

区間グラフの手法に関しては時間に依らず、2つのデータ間に似たような周波数の重なりがあった場合に類似度が高くなり、似たような周波数の重なりが少ないときに類似度が低くなったり、また、区間グラフの辺が多くなった場合、計算に時間がかかり処理がタイムアウトとなったり。

文字列の列の手法については、同じ時間で2つのデータ間に似たような周波数の重なりがあった場合に類似度が高くなり、似たような重なりが少ないとときに類似度が低くなったり。

## 4 結論

本研究では時系列データを区間グラフと文字列の列で表現し、類似する時系列データを検索する手法を定義した。今回の実験では、周波数の重なりをもとに時系列データの類似性を評価する事ができた。

今後は既存手法と比較するとともに、実際の時系列データを用いて本手法の有用性について検証する。

## 参考文献

- [1] 北川源四郎. 時系列解析入門. 岩波書店. 2005.
- [2] M.C. Golumbic. Algorithmic Graph Theory and Perfect Graphs. Elsevier. 2004.

		指導教員	斎藤 寿樹 准教授
学生番号	192C1130	氏名	藤村 海都
論文題目	福祉施設向け送迎スケジュール生成機能付き 利用予約システムの開発		

現代社会において、送迎サービスを提供する福祉施設や保育施設が増加している。これらの施設は仕事の忙しい両親や保護者の負担を軽減することに努め、子どもたちのスムーズな移動をサポートするとともに、安全な環境で過ごすことができるよう努めている。これらの施設の多様なサービスは、現代社会において非常に重要な役割を果たしている。

しかし、これらの送迎サービスを提供する施設の多くは送迎スケジュールを紙で収集し、配布するといった作業に膨大な時間と労力がかかっている。このような手間を省くことで施設スタッフの労力を他の重要なタスクに集中させることができ、効率的な運用を実現することができる。このような状況にも関わらず、多くの施設は資金が十分にないことなどから DX 化を進めることができていない。本研究ではそれらの問題を解決するために、安価に運用できシンプルな機能を持つ利用予約システムを開発した。

システムでは施設の利用日を簡単に登録できる機能や、各利用者をそれぞれの家に送迎する際の送迎スケジュールを瞬時に作成できる機能がある。これらの機能により長期的に見て大幅な業務時間の短縮、入力ミスの減少などの改善を見込める。

対象の施設へシステムの詳細について説明会を行い、このシステムの導入が決定した。また説明会を行なった事業所の職員に向けて、システムの導入についてのアンケートを行なった。アンケートにはシステムに期待する点、懸念する点、システム導入後の業務時間の変化の予想の 3 つの項目を用意した。期待する点については業務の効率化、入力ミスの減少などの意見が多い結果となった。懸念する点については導入による事業所職員、施設利用者保護者への負担が増えることなどとなった。また業務時間についてはシステムの導入により 30 分ほど短縮しそうという意見が半分を占めていたのに対し、1 時間以上増加しそうという意見が残りを占めていた。これらは導入初期に施設利用者保護者へ説明、対応する負担が理由だと考える。システム導入により長期的に見ると業務時間が短縮することは考えられるが、初期導入の負担を重要視するという傾向がみられた。

今後は 2023 年夏頃を目標に事業所職員、施設利用者を含めた全員のシステム利用を開始する予定である。

またシステムの保守・運用、システム内で使用している配送計画アルゴリズムの改善をしていく必要がある。さらにシンプルでわかりやすいデザインを採用し、システムのユーザー体験を向上させていく必要もある。

		指導教員	新見 道治 准教授
学生番号	192C1032	氏名	大島 奨太
論文題目	本棚に並ぶ漫画本背表紙からの文字領域抽出		

## 1 はじめに

近年、多くの漫画本が発行され、個人でたくさんの漫画本を所有することも少なくない。このような多くの漫画本を整理・管理することは手間がかかる。この問題を解決するために書籍の背表紙から情報を取得し、整理・管理を行う手法について研究が行われている。背表紙から情報を取得する手法には様々なものがあるが本研究では背表紙に書かれている文字情報を用いる方法を採用することとし、その前段階として背表紙画像から文字領域を正確に抽出することを目的とする。具体的な方法としてはまず画像の RGB 情報を用いた 2 値化を行い、複数の部分からなる文字を 1 つの塊として扱えるようにクロージング処理を行う。その後、輪郭検出を用いて文字領域の抽出を行う。

## 2 手法の流れ

本研究では以下のような流れで処理を行う。

1. 画像の 2 値化
2. クロージング処理
3. 輪郭検出を用いた文字領域抽出

画像の 2 値化では RGB プレーンから生成した濃淡画像に閾値処理を行い 3 枚の 2 値画像を生成した。さらに生成した 2 値画像の前景と背景を入れ替えた反転 2 値画像を生成し、合計 6 枚の 2 値画像を生成した。この画像に  $5 \times 5$  のカーネルを用いたクロージング [1] 処理を行い、複数部分からなる文字の部分間に前景画素を生成することで複数部分からなる文字を 1 つの塊として扱えるようにした。最後に輪郭検出を行い、抽出した輪郭に外接する矩形を生成することで文字領域の抽出を行った。抽出結果は図 1 のようになり、青で描かれた矩形が抽出された文字領域である。右の 6 冊は文字部分を矩形で囲むことができており文字領域の抽出ができているが文字以外の部分も文字領域として抽出してしまっている部分がある。加えて左の 3 冊は文字部分を矩形で囲むことができていない。

## 3 クロージング処理におけるカーネルについての検証

クロージング処理におけるカーネルのサイズや形状を変化させることで文字領域の抽出結果が変化するか



図 1: 文字領域の抽出結果

を検証した。今回は  $7 \times 7$ ,  $9 \times 9$ ,  $11 \times 11$  の正方形のカーネル、縦書きの背表紙文字に対して有効と考え、 $7 \times 5$ ,  $9 \times 5$ ,  $11 \times 5$  の縦長のカーネルを用いて検証を行った。カーネルを大きくすることで複数の文字を 1 つの文字領域とすることができますが同時に文字以外の前景領域との間に前景画素を生成してしまうため文字の状況に合わせて最適なカーネルを設定する必要があることがわかった。

## 4 対話処理を用いた 2 値化

文字部分を確実に前景とする処理が必要だと考え、HSV 色情報を用いて画像の指定した部分の色と近い色を前景とする処理を対話的に行えるようなプログラムを作成した。近い色と判定する HSV それぞれの条件を最適なものに変更する必要があるが指定した色を持つ部分を前景とする画像の生成に成功した。

## 5 まとめ

一連の処理により不完全だが文字領域の抽出を行うことができた。今回用いた背表紙画像以外でも文字領域抽出ができるのか検証するとともにクロージング処理におけるカーネルを最適なものとする方法や対話処理を用いた 2 値化により生成した画像を用いて文字部分を確実に前景とする方法の検証などを行う必要がある。

## 参考文献

- [1] [http://labs.eecs.tottori-u.ac.jp/sd/Member/oyamada/OpenCV/html/py\\_tutorials/py\\_imgproc/py\\_morphological\\_ops/py\\_morphological\\_ops.html?highlight=dilate](http://labs.eecs.tottori-u.ac.jp/sd/Member/oyamada/OpenCV/html/py_tutorials/py_imgproc/py_morphological_ops/py_morphological_ops.html?highlight=dilate)

		指導教員	新見 道治 准教授
学生番号	192C1092	氏名	田中 大志
論文題目	画像のベクトル化による潜在的興味の発見を用いた衣服の推薦システム		

## 1 はじめに

EC サイトなどで服を購入したい場合は膨大な服から嗜好に合う服を見つけなければならず、ユーザーにお勧めの商品を提示する推薦システムがあれば便利である。本稿では、単語のベクトル同士の計算により発見した潜在的興味を用いた推薦システムの研究 [1] を、画像ベースの衣服の推薦システムに適用することを考える。

## 2 提案する推薦システムの概要

今回提案する推薦システムでは、まず服画像をベクトル化する。このベクトルが存在する特徴空間を服特徴空間とする。また、ユーザーが所有する複数の服ベクトルの平均ベクトルをそのユーザーの潜在的興味と定義する。この潜在的興味と服特徴空間内で類似度が高い服を推薦する。

## 3 服画像のベクトル化による服特徴空間

多クラス分類型ニューラルネットワーク (MCNN) にカラー画像を入力として与えると、そのカラー画像の各クラス分類の確率が出力結果として得られる。この出力結果をベクトルと考える。今回は服の特徴を表すために、“形”，“模様”，“サイズ感”の 3 つを学習させた MCNN の出力を用いて、服画像のベクトル化を行い、34 次元服特徴空間を作成した。現実空間に存在する服は、この特徴空間内の 1 点にマッピングできる。MCNN には EfficientNetV2-B0 を用いた。

## 4 潜在的興味の発見と推薦方法

潜在的興味とは、ユーザー自身が気づいていない興味のことである。本研究では、34 次元服特徴空間において、ユーザーが所有する複数の服ベクトルの平均ベクトルを潜在的興味と定義する。この潜在的興味と類似度の高い現実空間の服を選び出し、ユーザに推薦する。

## 5 提案する推薦システムの有用性

図 1 は、異なる雰囲気のユーザーが所有する 34 次元の服画像ベクトルを主成分分析により 2 次元で可視化した。このグラフから、34 次元服特徴空間において服の雰囲気を捉えていることがわかる。また推薦においては、特徴空間内でユーザーの平均ベクトル（潜在的

興味）と近い位置に存在する服を推薦するので、そのユーザーの雰囲気を捉えた推薦が行えると言える。ここで図 2 に対する推薦結果を図 3 に示す。

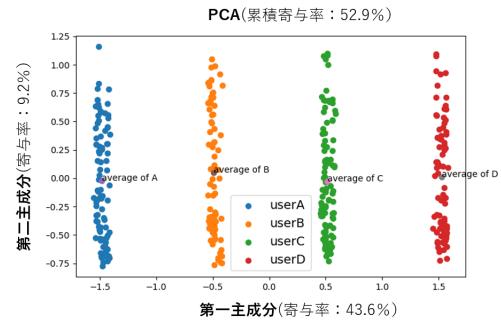


図 1: 服特徴空間の可視化



図 2: 所有する服画像



図 3: 推薦結果

また推薦動作を数値化し、推薦スコアとした。ランダムに服画像を提示するよりも推薦スコアが高くなかった。

## 6 まとめ

提案する推薦システムはユーザーの雰囲気を捉えた推薦ができることが示せた。今後は、実際にユーザーに利用してもらっての評価を基にした推薦システムの改良を行っていく。

## 参考文献

- [1] 開地 亮太, 檜垣 泰彦, ”潜在的興味に基づく観光地推薦システムの試作” 電子情報通信学会技術研究報告, vol.115, no.138, pp.29–34, 2015.

		指導教員	新見 道治 准教授
学生番号	192C1156	氏名	森 啓輔
論文題目	HSV 色空間に基づく空領域の分類		

## 1 はじめに

大気汚染は排出ガスや大気汚染物質の排出によって引き起こされ、近年では世界的に重要視されている環境問題となっている。しかし、危険性を理解している人々は少ない。本研究では、大気汚染の危険性を社会に浸透させるために、スマートフォンで空画像を撮影し、ニューラルネットワークを用いて大気汚染の度合いを検出することを目的としていた。その過程で、空画像から青空領域を抽出し大気汚染を検出することで、検出精度を上げることができると考えた。そこで、本研究は、空画像から青空領域を抽出することを目的とする。

## 2 RGB 色空間と HSV 色空間を用いた雲領域の検出

空画像から直接青空を検出する手法は、従来の研究では見つかっていない。しかし、RGB 色空間を用い、(B-R)/(B+R) で表される値を参照する雲検出の手法は存在する [1]。この RGB 色空間を用いた雲検出と HSV 色空間による雲検出を組み合わせた手法を用いて、雲を検出して雲領域のマスク画像を作成し、それを原画像にマスク処理することで空画像からの青空領域の抽出を行った。本実験では、福岡大学と九州工業大学に設置されている高所カメラで撮影された空画像を画像データとして用い、実験を行った。その結果、一部の雲が検出できない場合や、青空領域を雲領域と誤検出してしまう問題が発生した。

## 3 HSV 色空間を用いた青空領域の検出

雲領域の検出において、一部の雲領域が検出できない場合や、青空領域を雲領域と誤検出してしまう場合があった。そこで、青空領域の HSV 色空間の画素値の平均値と分散値を調査することで、画素値の分布を調査し、直接青空領域を検出する手法を考えた。



図 1: 青空領域を抽出した画像

[2] による、正規分布するデータは、平均土  $2 \times$  標準偏差の範囲中に約 95 パーセントのデータが含まれるという法則をもとに、画素値が正規分布すると仮定し、HSV 色空間の閾値の設定することで青空領域の直接検出を行った。その結果、青空領域を検出することに成功した。出力結果は図 1 のようになっている。

## 4 追加実験

追加実験では、空画像の空領域を 4 レベルに分割する手法を考えた。以下では青空のレベルを CSL(Clear sky level) と呼ぶ。CSL は、空領域に対する雲の混在量によって、レベル分けを行う。本実験では、飯塚市内でスマートフォンを用いて撮影した空画像を用いて実験を行った。空領域の分割には HSV 色空間の閾値を用いる。その過程で、撮影する機材や、画角によって、空領域の明度が変化することがわかり、空領域の検出において H と S の値が強く関係していることが判明した。そして、V の値をすべての範囲に設定することで、空領域の分割をうまく行うことに成功した。

## 5 まとめ

本研究では、空画像から青空領域を抽出するために、HSV 色空間の青空領域の画素値の調査し、画素値の分布をつかむことで、青空領域の抽出を行った。また、追加実験では、検出した空領域を雲の混在量のレベル別に分割する手法を考え、それを実行した。これらの実験で、空検出において、H と S の画素値が特に重要なことがわかり、V の値は撮影時の画角や機材によって変化する場合があることが判明した。本研究の目標は、スマートフォンで撮影した画像を対象としていくため、今後の課題として、スマートフォンの画像データに適応した閾値となるよう、再度閾値の見直しを行い、実験および改良を行っていく必要がある。

## 参考文献

- [1] Yamashita, M., Yoshimura, M. and Nakashizuka, T., "Cloud cover estimation using multitemporal hemisphere imageries", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 35, 826 – 829, 2004
- [2] いちばんやさしい、医療統計, "正規分布とは?わかりやすく標準偏差とのかんんけいもガウス分布に関して解説", [https://best-biostatistics.com/summary/normal\\_dist.html](https://best-biostatistics.com/summary/normal_dist.html), 2022 年 11 月 29 日 時点.

		指導教員	新見 道治 准教授
学生番号	192C1168	氏名	山寺 慧司
論文題目	印刷耐性を有するアルティメイトリンクに関する研究		

## 1 はじめに

我々の研究室では、アルティメイトリンクという現実世界におけるある意味的なまとまりに対して、サイバー空間における付随/補足情報をつなぎ合わせる技術を提唱している。この技術を利用して紙面に印刷された文字列に対して URL を埋め込み、抽出を行うことで Web ページを表示するシステムの提案がなされている [1]。しかし、文献 [1] では漢字やひらがなを含む文字列からの情報の抽出には失敗している。そこで、本研究では漢字やひらがなを含む文字列にも対応した印刷物のシステムの改良を行う。

## 2 印刷物に対するシステム概要

印刷物のシステムの処理の大まかな流れは、埋め込みと抽出である。埋め込みでは、文字画像から文字認識を行い、文字列を 1 文字ずつに分割して各文字ごとに細線化し、前景画素に順番づけを行う疑似筆順を計算する。疑似筆順を用いて埋め込み情報に対応した色のついた円を描画する。抽出では、埋め込みを行った文字を印刷した紙面をカメラで撮影したもので行う。撮影した画像から再度文字認識と文字列を分割し、各文字上の円から埋め込み情報に割り当てた色ごとに円を検出し、疑似筆順を利用して色ごとに対応するビット列を順番に並べることで埋め込んだ情報を抽出する。

## 3 先行研究における問題点

先行研究での問題点は、抽出処理の文字列の分割で文字画像によっては図 1 のように正しく分割が行われないことがある。

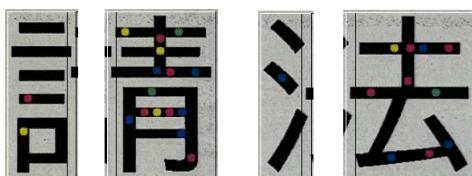


図 1: 分割失敗

また、図 2 のように円と円の距離が近すぎて円の検出に失敗するということがある。

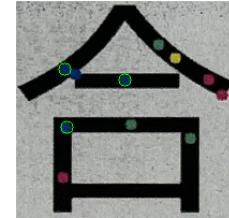


図 2: 円検出失敗

## 4 提案する改良手法

文字列の分割手法を、文字認識の結果を利用したものに変更する。文字認識の結果、 $n$  文字認識し、文字列全体の横幅の長さが  $X$  とすると  $(X/n)$  を 1 文字分の横幅の長さとして文字列の分割を行う。

また、埋め込み処理の細線化前に文字の収縮処理を行うことで、円の描画時に円と円の距離が近くなりすぎるのを防ぐ。

## 5 提案手法による実験結果

提案した文字列の分割手法により、正しく文字列の分割が行われ情報の抽出に成功した。また、細線化前の文字の収縮処理により、円と円の距離が近くなりすぎるのを防ぎ、適切に円の検出が行われ情報の抽出に成功した。

## 6 まとめ

印刷物にアルティメイトリンクを利用したシステムを改良して、漢字やひらがなを含む文字画像にも対応する手法を提示した。今後の課題として、画数の多い複雑な形の漢字を含む文字画像の場合に頑健なシステムの構築や、背景色が白色以外の場合や、蛍光灯の下以外の照明条件の場合でも正しく処理が行われるかの評価を行う必要がある。

## 参考文献

- [1] 牟田賢弘, 新見道治, 山内幸治, "意味的なまとまりを持った複数文字を対象とした印刷耐性を有するアルティメイトリンク", マルチメディア情報ハイディング・エンリッチメント研究会, March.2022

		指導教員	新見 道治 准教授
学生番号	21221007	氏名	浜川 怜
論文題目	ユーザ指向な英語発声学習補助ツールに関する研究		

## 1 はじめに

近年、日本の英語教育目標は英語によるコミュニケーション能力の養成へ変化し、それに伴い教育現場での授業のあり方も変化が求められている。特にスピーキング能力を向上させるための授業スタイルの変化が求められる。しかし、英語教育者自身の英語力不足や ALT などの人材不足から、現場の教育は目標に対して追いついていない現状である。特に、発音指導については生徒一人ひとりに適切な指導を行うのは困難である。そこで本研究では、英語初学者の発音学習を支援する補助ツールの提案・技術検証を行う。

## 2 モデル

本研究には、テキストから音声を生成するモデルである Tacotron2 を応用した Waris Quamer[1] らのアクセント変換のモデルを用いた。図 1 にモデルの概要図を示す。なお、本研究ではアクセント変換を行う Seq2Seq Model(Translator) と Translator の入力のひとつであるアクセント埋め込みベクトルを生成する Accent Encoder 以外は学習済みモデルを流用する。

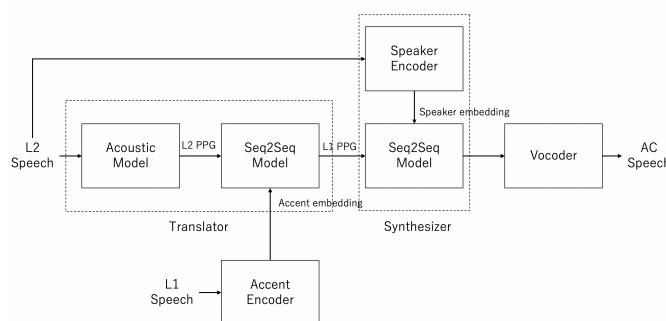


図 1: モデル概略図

## 3 実験

実験には L2-ARCTIC に含まれる音声とこのコーパスと同じ内容を日本人が読み上げた音声データを学習データとする。ARCTIC を変換後のターゲットとなる音声として用いる。なお、1 人あたり 400 文の音声データを学習データとする。そして Waris Quamer[1] の手法に基づきバッチサイズを 32、最適化手法に Adam を採用し、学習率  $10^{-3}$  として学習を行う。

表 1: Seq2Seq(Translator) のパラメータ

Component	Params
PPG	144 D
Accent Emb	144 D
3×Conv	512 5 × 1 kernels
2×pBiLSTM 256	cells/direction
Location Sensitive	512-dim context; 32
Attention	31×1 conv kernels
PreNet	2×FC; 512units
2×LSTM	2048 cells
Linear(PPG)	1×FC; 144units
Linear(stop-token)	1×FC; 1units
PostNet	5×1 Conv; 1024 5×1 kernels
PPG	144 D
stop-token	2 D

## 4 結果

学習データに含まれる文章をカタカナ発音で入力した場合は、正しくネイティブ発音の音声に変換されることが確認できた。しかし、学習データに含まれない文章を読み上げた場合、はじめの 1~2 単語以外の発話内容を維持出来ていないネイティブ発音の音声が生成されてしまう。

## 5 まとめ

日本人特有のカタカナ発音の音声をアクセント変換によって、ネイティブ発音の音声を生成できることが確認できた。学習データに含まれない文章は正しく生成出来ていない問題が残存するが、この問題は学習データ量をより多くしたり、学習データをより細かい単語・フレーズ単位で与えることで改善が見込める。

今後の課題としては、上記の問題を解決し、さらに英語初学者が自身で発音の矯正をする手掛かりとなるような、発音の謝りを指摘する手法の検討・開発があげられる。

## 参考文献

- [1] Waris Quamer, Anurag Das, John Levis, Evgeny Chukharev-Hudilainen, Ricardo Gutierrez-Osuna, "Zero-Shot Foreign Accent Conversion without a Native Reference" Interspeech, 2022, pp.4920-4924

		指導教員	本田 あおい准教授
学生番号	192C1016	氏名	市原亜美
論文題目	TF-IDF による顧客コミュニティサイト サービス向上のためのキーワード抽出		

## 1 背景・目的

近年、自然言語処理分野は急速に発展しており、多くの企業は、顧客コミュニティサイト内で自然言語処理を利用したよくある質問コーナーを活用している。しかし、専門性の高い事業を行っている企業においては、独自のコーパスを作成する必要があり、実用化が難しいという問題がある。

そこで、本研究では、実用化への第一歩として、タグ付けの自動化を試みることで、効率化を図り、コミュニティサイトにおけるサービス向上を目指した。

## 2 TF-IDF 法

求める TF-IDF の値はある文書  $j$  に登場するある単語  $i$  の重要度とする。まず、TF-IDF の値は以下のようになる。

$$tfidf_{i,j} = tf_{i,j} \cdot idf_i$$

ただし、

$$\left\{ \begin{array}{l} tf_{i,j} = \frac{\text{単語 } i \text{ の文書 } j \text{ における出現回数}}{\text{文書 } j \text{ の単語数}} \\ idf_i = \log \frac{\text{総文書数} + 1}{\text{単語 } i \text{ の出現する文書数} + 1} \end{array} \right.$$

## 3 検証

行政向けのサポート業務を行う N 社のコミュニティサイトにおけるよくある質問コーナーの質問、回答リストをデータとして用いた。質問文ごとに形態素解析を行い、TF-IDF の値を求めた。その結果、TF-IDF の値の上位 5 位以内の単語をキー

ワードとして抽出した。

## 4 考察

9 個、182 個のデータから抽出されたキーワードを以下の表 1 に示す。

表 1 抽出されたデータ

データ数	1位	2位	3位	4位	5位
9 個	多子	軽減	方	サービス	—
182 個	多子	軽減	方	アセスメント	アップ

表 1 より、データ数を増やすごとに「サービス」という一般的な用語の順位が下がったことがわかった。これにより、データ数が多いほど精度は上がると考えられた。また、既存の関連ワードと比較しても精度が上がったことがわかった。

## 5 おわりに

コミュニティサイトサービス向上のためには TF-IDF を用いたキーワード抽出を目的として検証を行った。抽出されたキーワードは既存の関連ワードと概ね一致しており、有用性があった。一方で、TF-IDF の値を用いたキーワード抽出の実用化までは、改善の必要がある。

## 6 参考文献

- [1] <https://tech.at-iroha.jp/?p=1576>
- [2] <https://scarlet09libra.hatenablog.com/entry/2021/02/24/210856>
- [3] <https://gotutiyen.hatenablog.com/entry/2020/09/10/181919>

		指導教員	本田 あおい 准教授
学生番号	192C1116	氏名	野田部 恵央
論文題目	レコメンドシステムにおける 様々な距離による類似度の測定		

## 1はじめに

近年、急速な情報社会の進展に伴い、得られる情報の量が膨大になっている。数多の情報の中から、ユーザーの必要な情報を探し出すことは非常に困難である。映画についてこの問題を考える。動画配信サービスなどには、数多くの映画が存在し、その中から見たい映画を自力で見つけ出すことは困難である。この問題を解決するために行われているのが、推薦システムの開発である。推薦システムの方式には、協調フィルタリングと呼ばれる手法が主流となっている。推薦を行う際にはあらかじめ、ユーザー同士の類似度を測定しておく必要があるが、何を重視するかによってユーザーの類似度は変わってくる。そこで本稿では、協調フィルタリングを用いつつ、複数の距離関数を比較して類似度にどんな差（特徴）が出るか調べた。

## 2用いるデータセット

本研究を行うにあたって用いる実データセットとして Group Lens が提供している Movie Lens Data Set を使用する。概要を表 1 に示す

表1 Movie Lens Data Setの概要

評価対象 (映画数)	9719 (本)
評価人数	610 (人)
評価数	100,836 (件)
Time Stamp	100,836 (件)

本実験では、包除積分とショケ積分の計算をする際に、 $\lambda$ -ファジイ測度を用いるため、ある程度実験データを絞る必要がある。そこで、表1で示したデータに対して、見た映画の数が900本未満の人を除く、かつ900本以上映画を見た全員がレビューを残した映画のみを表示する処理を行い実験を行った。

## 3 実験

実験の流れは次の通りである

対象となる17ユーザから2人のユーザを選ぶ。  
選んだユーザが対象となる映画10本に対して出した評価を見る。

互いが出した評価に対し、ユークリッド距離、マンハッタン距離、包除積分、ショケ積分の4つを用いて距離にどのような違いが見られるかを測定する。測定結果は以下の表2のようになった。上に行くほどユーザーの類似度が高い。

表2 各距離による類似しているユーザーの組

ユークリッド距離	マンハッタン距離	包除積分( $\lambda=1$ )	包除積分( $\lambda=-0.5$ )	ショケ積分( $\lambda=1$ )	ショケ積分( $\lambda=-0.5$ )
(599, 606)	(599, 606)	(599, 606)	(599, 606)	(599, 606)	(599, 606)
(249, 414)	(249, 610)	(249, 610)	(249, 610)	(249, 610)	(249, 610)
(182, 606)	(182, 606)	(414, 610)	(380, 414)	(249, 414)	(182, 606)
(298, 474)	(298, 414)	(249, 414)	(182, 606)	(414, 610)	(298, 474)
(380, 414)	(298, 474)	(182, 606)	(298, 474)	(380, 414)	(380, 414)
(249, 387)	(380, 414)	(298, 474)	(249, 414)	(182, 606)	(414, 610)
(249, 610)	(414, 610)	(380, 414)	(414, 610)	(298, 474)	(249, 414)
(387, 603)	(387, 606)	(387, 606)	(387, 606)	(387, 606)	(387, 606)
(298, 387)	(182, 606)	(387, 606)	(182, 606)	(387, 606)	(182, 606)
(298, 606)	(387, 603)	(182, 599)	(182, 599)	(414, 599)	(414, 599)
(182, 387)	(414, 599)	(414, 599)	(387, 503)	(182, 599)	(387, 603)
(414, 610)	(249, 380)	(249, 380)	(288, 380)	(288, 380)	(599, 610)
(182, 599)	(288, 380)	(288, 606)	(599, 610)	(249, 380)	(288, 380)
(474, 610)	(298, 605)	(380, 610)	(380, 610)	(380, 610)	(380, 610)
(249, 506)	(380, 610)	(599, 610)	(288, 606)	(288, 380)	(249, 380)
(298, 387)	(599, 610)	(288, 380)	(249, 380)	(599, 610)	(298, 606)
(298, 599)	(182, 387)	(182, 387)	(288, 448)	(182, 387)	(274, 298)
(298, 610)	(274, 298)	(599, 603)	(414, 606)	(599, 603)	(414, 606)
(387, 599)	(288, 448)	(274, 298)	(274, 298)	(288, 448)	(288, 448)
(599, 610)	(414, 606)	(414, 606)	(599, 603)	(414, 606)	(599, 603)

## 4 おわりに

ユーザーの嗜好に全体的に少しづつ差がある場合は、ユークリッド距離を使うことでレコメンドが適切に行われやすく、ユーザーの嗜好に全体的な差はさほどないものの、一部の対象で大きな差が見られる場合にはマンハッタン距離や  $\lambda > 0$  の包除積分、ショケ積分を使うことで適切なレコメンドが行われるのではないかと考察できる。また、大きな差が見られるのが複数あり、そのほかでは評価に差があまりないときは  $\lambda < 0$  の包除積分、ショケ積分が有効であると考えられる。今後の展望としては、 $\lambda$  ファジイ測度ではなく、一般的のファジイ測度を使うことで、映画による異なる重みづけを可能にしていきたいと考えている。

## 5 参考文献

- [1] 本田あおい、包除積分数理モデルを用いた多変量データ解析、2018
- [2] 本田あおい、岡崎悦明、新しいファジイ積分の提案と評価問題への応用、2009
- [3] David Goldberg, David Nichols, Briaan M. Oki, Douglas Terry: Using collaborative filtering to weave an information tapestry(1992)

		指導教員	本田 あおい 準教授
学生番号	192C1120	氏名	林田 一吹
論文題目	単語の重要度を考慮した FAQ 検索システム		

## 1 はじめに

コールセンターのオペレーターがユーザーからの問い合わせに対応する際は、過去の問い合わせ履歴や Frequently Asked Question (FAQ) を参照して回答することが多い。しかし、迅速な対応が求められる業務で、膨大なデータの中から正確な回答を見つけることは困難であり、オペレーターの大きな負担となっている。本研究では、FAQ データをコーパスとして、質問文の文書中の単語の重要度と検索クエリ中に含まれる単語の重要度を比較し、文書の類似度を算出する検索システムを試作した。

## 2 提案手法

### 2.1 提案手法 1:TF-IDF 法

1 つ目の手法として、TF-IDF を用いて文書に含まれる単語の重要度を算出した。TF-IDF とは、文書内に出現する単語について、単語の出現頻度 (TF 値) と単語の逆文書頻度 (IDF 値) の 2 つの情報から、その単語の重要度を算出する手法である。ある文書  $d_j$  に出現する単語  $t_i$  について、出現回数を表す関数を  $f$  とすると TF 値は以下のようになる。

$$tf(t_i, d_j) = \frac{f(t_i, d_j)}{\sum_{t_k \in d_j} f(t_k, d_j)} \quad (1)$$

続いて、ある文書集合における単語  $t_i$  について、単語  $t_i$  が出現する文書数を  $df$ 、総文書数を  $N$  とすると IDF 値は以下のようにになる。

$$idf(t_i) = \log \frac{N}{df(t_i) + 1} \quad (2)$$

TF-IDF 値は TF 値と IDF 値をかけ合わせて計算される。

$$tfidf(t_i, d_j) = tf(t_i, d_j) \cdot idf(t_i) \quad (3)$$

### 2.2 提案手法 2:BM25

2 つ目の手法として、TF-IDF の発展的手法である Okapi BM25(以下 BM25) を用いて、文書の順位付けを行った。 $D$  を文書全体の集合、 $D$  の要素である文書を  $d$ 、 $q$  を検索クエリ、 $idf(q_i)$  を単語  $q_i$  の IDF 値、 $f(q_i, D)$  を文書  $d$  中の単語  $q_i$  の出現頻度、 $|d|$  を文書  $d$  の単語数、 $avg(dl)$  を全文書の平均単語数とする。

$$score(q, d) = \sum_i idf(q_i) \cdot \frac{(k_1 + 1)f(q_i, d)}{f(q_i, d) + k_1(1 - b + b \frac{|d|}{avg(dl)})} \quad (4)$$

$score(q, d)$  の値が大きいほど、文書  $d$  の重要度が高いことを示す。ここで、 $k_1$  は 0 以上の実数、 $b$  は 0 以上 1 以下の実数をとるパラメータである。

## 3 評価実験

実験データとして、株式会社ニックの FAQ サイトから、170 件の FAQ データを提供していただき、FAQ コーパスを作成した。評価指標には、MRR (Mean Reciprocal Rank) を用いた。MRR は、上位 5 個の検索結果のうち最も上位にある正解の順位の逆数の平均値である。結果を表 1 に示す。MRRにおいて、BM25 の手法を用いた検索システムが、TF-IDF を上回る結果となった。しかし、両手法とも MRR 値が 0.5 以上であることから、平均的に 2 件目までに適合文章があることが期待できる。

表 1: MRR 値の結果

	MRR 値
TF-IDF	0.8315
BM25	0.8953

## 4 おわりに

単語の分散表現を用いた検索システムの構築で、検索意図にマッチする FAQ が上位表示されることを確認した。今後の展望は、FAQ データよりもデータ数の多いコールセンターのお問い合わせデータをコーパスとして、検索システムの構築、性能の評価を行う予定である。

## 参考文献

- [1] sklearn.feature\_extraction.text.TfidfVectorizer  
[https://scikit-learn.org/stable/modules/generated/sklearn.feature\\_extraction.text.TfidfVectorizer.html](https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html)
- [2] summarization.bm25 – BM25 ranking function  
[https://radimrehurek.com/gensim\\_3.8.3/\\_summarization/bm25.html](https://radimrehurek.com/gensim_3.8.3/_summarization/bm25.html)
- [3] 現代版 TF-IDF である Okapi BM25 の原理について (前半)  
<https://zenn.dev/wm3/articles/233dccc6e54dab>
- [4] 現代版 TF-IDF である Okapi BM25 の原理について (後半)  
<https://zenn.dev/wm3/articles/31dd38542861eb>

		指導教員	本田 あおい 准教授
学生番号	192C1124	氏名	福島 慎央
論文題目	包除積分ニューラルネットワークの設計と実装		

## 1 背景・目的

近年、機械学習の手法の一つであるディープラーニングがデータ分析の分野で高い成果を上げている。しかし、複雑であるために解釈性が損なわれるという問題がある。本研究では、非加法的測度であるファジイ測度と多項演算によって定義される包除積分を用いてネットワークの設計、実装を行い、学習後の重みを解釈可能な値として得ることを目的とする。

## 2 設計

包除積分はメビウス変換により重回帰の形に式変形することができるため、ニューラルネットワークの重みの計算でメビウス変換の値を求め、包除積分を得ることでモデルの解釈を行う。

包除積分ニューラルネットワークモデルはデータの入力層、データの前処理を行う層、包除積分の式を表す層、そして出力層の大きく分けて4層からなるネットワークモデルとなる。図2.1は入力を3変数としたときの包除積分ニューラルネットワークモデルである。

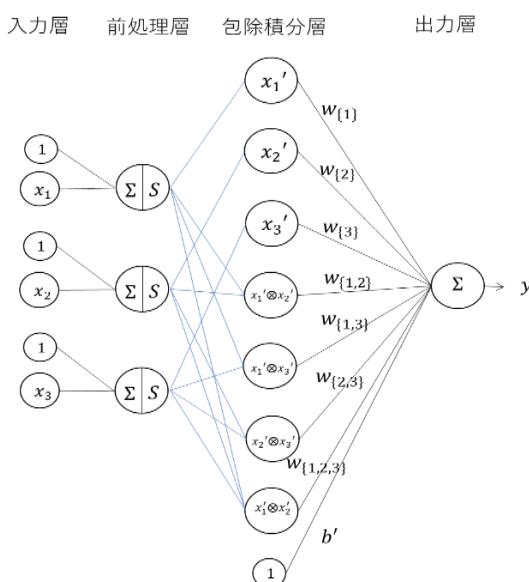


図2.1 包除積分ニューラルネットワーク

入力層の各説明変数は前処理層において重み、バイアスによる正規化の後、Sigmoid

関数を用いることで[0,1]区間に変換し、t-ノルムを[0,1]状に定義することで代数積をただの積の多項演算とできる。包除積分を表す層では[0,1]で得られた各説明変数からt-ノルム演算を行いその出力と重みの積を出力層へ渡す。

## 3 実装

包除積分ニューラルネットワークはPythonのオープン機械学習ライブラリであるPyTorchを使い、torch.nn.Moduleを継承したクラスにより実装を行う。

作成するクラスは大きく分けて4つあり、前処理層を表すクラス、出力層を表すクラス、包除積分を表す層でt-ノルム演算を行うクラス、上記3つのクラスを用いた包除積分ニューラルネットワーク全体を表すクラスである。図3.1にクラス図を示す。

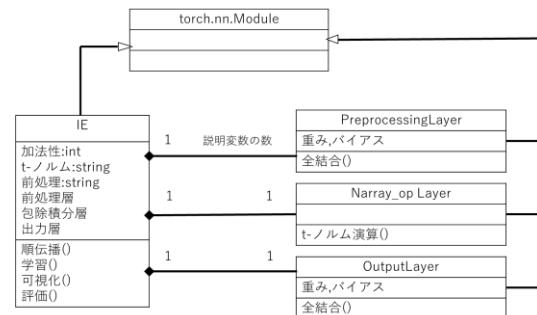


図3.1 クラス図

## 4 おわりに

本研究では、非加法的測度であるファジイ測度と多項演算により定義される包除積分を用いてネットワークの設計、実装を行った。代数積や論理積などのt-ノルムはパラメータを持っていないためクラスとして実装する必要はないが、拡張性を考慮し、torch.nn.Moduleを継承したクラスとして実装した。また、t-ノルムにはパラメータを持ったパラメトリックt-ノルムも存在し、今後それらを実装する際に、クラスのコンストラクタに更新可能なパラメータとして設定することで、学習の際にパラメトリックt-ノルムのパラメータの更新も行うことができると考えられる。

		指導教員	本田 あおい 準教授
学生番号	192C1166	氏名	山下 兼巧
論文題目	包除積分ニューラルネットワークの性能評価と重みの解釈		

## 1はじめに

本研究では、非加法的測度であるファジィ測度と、多項演算で表される包除積分を用いて、説明可能なネットワークモデルを構築する。さらに、包除積分の加法性条件やt-ノルム、前処理を変えたときのモデルの性能の比較とニューラルネットワークモデルであるMLP Regressorとの性能比較を行うことで高い精度が得られているか検証する。また、ゲーム理論で使われるシャープレイ値とグループを考慮したOwen Valueの2つからモデルの解釈を行う。

## 2包除積分ニューラルネットワークの構築

包除積分ニューラルネットワークは入力層、前処理層、包除積分層、出力層の4層で構成される。図1は入力を3変数としたときの全加法の包除積分ニューラルネットワークモデルである。

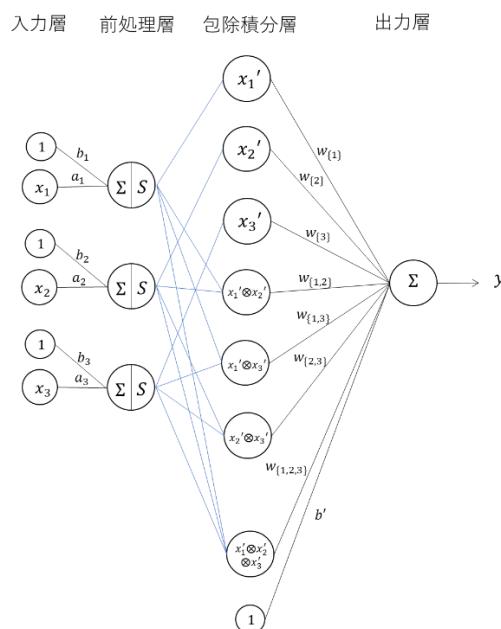


図1 全加法のモデル図

## 3実験

代数積を用いた包除積分ニューラルネットワーク

に関して入力数6と10のときに加法性条件を変えた場合の決定係数R<sup>2</sup>、計算速度、パラメータ数について比較した。

表1 入力数6と10で加法性条件を変えたときの比較

	入力数					
	6			10		
	R <sup>2</sup>	計算速度[sec]	パラメータ数	R <sup>2</sup>	計算速度[sec]	パラメータ数
1加法	0.8428	276.1	19	0.8543	294.3	31
2加法	0.8675	339.7	34	0.8818	466.7	76
全加法	0.8909	552.4	76	0.8987	4766.5	1044

表2 入力数6のシャープレイ値

		説明変数の番号					
		1	2	3	4	5	6
シャープレイ値	1加法	-0.0562	0.0548	0.0501	0.0445	0.0516	0.0774
	2加法	-0.0636	0.0565	0.0794	0.0560	0.0745	0.0924
	全加法	-0.7040	0.2012	0.6811	0.5653	0.3014	0.4401
Owen Value	1加法	-0.3375	0.3285	0.3006	0.2669	0.3096	0.4644
	2加法	-0.5467	0.2789	0.4504	0.2686	0.3823	0.4348
	全加法	-0.6956	0.2072	0.6657	0.0772	0.3087	0.4129

表3 入力数6と10でのMLPモデルの性能

	入力数					
	6			10		
	R <sup>2</sup>	計算速度[sec]	パラメータ数	R <sup>2</sup>	計算速度[sec]	パラメータ数
MLPモデル1	0.8136	0.3948	801	0.8329	0.4527	1201
MLPモデル2	0.9003	2.304	8001	0.9012	2.652	12001

上記の結果から包除積分ニューラルネットワークは加法数を増やすほど精度が上がっていくが、加法数、パラメータ数を増やすことによる精度の上がり方は徐々に緩やかになっていくと考えられる。また、MLP Regressorと近い精度を出せていることが分かる。また、計算時間に関してはパラメータ数が大きく影響することが分かったので、精度、計算速度の面から包除積分ニューラルネットワークを使用する際は2加法か、3加法程度までモデルを構築するのが良いと考えられる。

## 4おわりに

本研究では、加法性条件、t-ノルム、前処理を変えたときのモデルの性能の比較、ニューラルネットワークモデルであるMLP Regressorとの性能の比較を行った。他のt-ノルムや加法性条件等についてもより詳細に比較・実験を行っていき、ブラックボックスモデルの原理解明などについても考えていただきたい。

		指導教員	武村 紀子 准教授
学生番号	192C1028	氏名	江角 翼
論文題目	座圧情報を用いたグループ活動における話者推定		

## 1 はじめに

近年、アクティブラーニングや PBL などの知識活用型の学習が注目されており、教育の場でグループ活動を行う機会が増えている。しかし、グループ活動において教員が全学生の状態を把握するのは困難であり、適切な声掛けや公正な成績評価が難しいといった問題がある。この問題を解決するために、誰がいつ、どの程度発話したかという情報は有益であると考える。これまでに音声情報を用いた話者推定の研究が数多くなされているが、マイクの位置やノイズ、プライバシーの問題がその実用を妨げている。そこで、本研究では、プライバシー情報が少なく、安定した計測が可能な座圧情報を用いて話者推定を行う。

## 2 データベースの作成

図 1 のように 4 名のグループ（計 8 グループ、32 名）で 25 分間のディスカッションタスク（KJ 法）を 2 回行い、タスク中の座圧データ、音声データを用いて、話者推定データベースを作成する。ここでは、座布団型の座圧センサを用いて  $16 \times 16$  箇所の圧力値を計測し（図 2 参照）、音声データは各被験者が装着したピンマイクにより収集する。音声データにはノイズが多く含まれるために、Word for the web のトランスクリプト機能を用いて発話区間を切り出し、発話区間のマイク音量から話者を特定する。 $16 \times 16$  画素の座圧画像データ（fps=5）に話者ラベルを付与したものを話者推定モデルの学習・評価に使用する。



図 1: グループ学習の様子

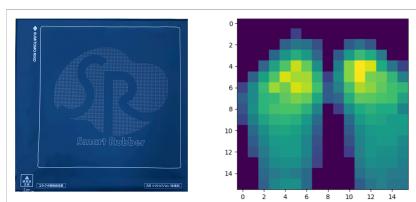


図 2: 座圧センサと座圧画像

## 3 話者推定モデル

1 枚の座圧画像を入力とするモデルとして VGG-16[1]、1 秒間の座圧画像列（5 枚）を入力とするモデルとして 3D-CNN[2] を採用する。どちらのモデルも画像認識問題において高い性能を示しているモデルである。また、話者推定には姿勢だけではなく動き情報も重要であると考え、1 フレーム前との差分画像を入力とし、動き情報を明示的に与えた場合についても検証を行う。

## 4 評価実験

グループ学習を行った 8 グループについて、Leave one-group out 交差検証を行った。ここでは、3 値分類（自分が発話/他人が発話/発話なし）および 2 値分類（自分が発話/他人が発話または発話なし）について評価実験を行う。表 1 に実験結果を示す。

表 1: 実験結果 (macro-F1 score)

		3 値分類	2 値分類
VGG-16	座圧	0.306	0.486
	差分	0.395	0.623
3D-CNN	座圧	0.452	0.717
	差分	<b>0.456</b>	<b>0.728</b>

VGG-16 より、時系列を考慮した 3D-CNN の方が、また、座圧の元画像より差分画像を入力した方が良い結果となった。以上より、話者推定には動作情報が重要なことが分かる。

## 5 おわりに

本研究では、座圧情報を用いた話者推定手法を提案し、評価実験により、その有用性を示した。今後はグループ活動を行う他者の座圧データやマルチモーダルデータの使用を検討する。

## 参考文献

- [1] K. Simonyan, et al., "Very deep convolutional networks for large-scale image recognition," *ICLR*, 2015.
- [2] S. Ji, et al., "3d convolutional neural networks for human action recognition," *IEEE TPAMI*, 2012.

		指導教員	武村 紀子 准教授
学生番号	192C2001	氏名	朝枝 彩夏
論文題目	顔画像における個人差を考慮したユーザ状態推定		

## 1 はじめに

表情は人の内部状態をよく表すと言われており、顔画像を用いた表情認識や感情推定に関する研究が数多く行われている。しかし、顔の作りや表情表出方法には個人差があり、このような個人差は推定精度に悪影響を及ぼす。特に、集中や眠気、疲労といった表情に表れにくい曖昧な内部状態は個人差の影響を強く受ける。そこで本研究では、顔画像における個人差を考慮することで、曖昧な内部状態の推定精度の向上を目指す。

## 2 提案手法

本研究では、深層学習モデルにおいて、顔の特徴  $V_{face}$  から個人の特徴  $V_{id}$  を減算することで内部状態の特徴  $V_{state}$  を抽出する、偏差モジュール [1] を用いる（図1参照）。偏差モジュールの個人識別モデルについては、個人が識別できるように事前に学習し、提案ネットワークを学習させる際は、重みを固定する。顔認識モデルは個人識別モデルと全く同じ構造で、個人識別モデルの重みを初期値として学習させる。偏差モジュールにより得られた  $V_{state}$  (512次元) に対し、全結合層からなる状態推定モジュールを用いて、内部状態の推定を行う。

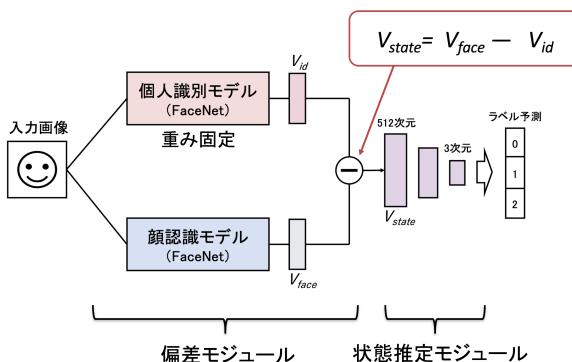


図 1: ネットワークモデル (提案手法)

## 3 実験と考察

本研究では、推定する内部状態を e-Learning 中の学習者の覚醒度とし、評価実験を行った。学習、評価用のデータベースとして、ノートパソコンを用いて講義動画を視聴する被験者の顔画像データ（27名、約20時間分、fps=30）を用いた。各顔画像データには、被

験者の状態に応じて3種類のラベル（眠っている/眼そう/起きている）が付けられている。ここでは、顔画像データから MTCNN[2] を用いて顔領域を切り取った画像を提案モデルへの入力とする（図2参照）。また、学習データにおけるラベルの偏りを防ぐため、アンダーサンプリングを行った。



図 2: MTCNN 適用例

27名の被験者を9グループに分割し、Leave One-group out 交差検証を行った。表1に提案手法および偏差モジュールの代わりに顔認識モデルのみを用いた比較手法における平均 macro-F1 の結果を示す。なお、本実験では、状態推定モジュールの全結合層の数を変化させた場合についても比較を行った。表1より、状態推定モジュールの全結合層数は2層が最も良く、どの場合も提案手法が比較手法の精度を上回った。

表 1: 提案手法及び比較手法の平均 macro-F1

状態推定モジュール	提案手法	比較手法
512 → 3	<b>0.5260</b>	0.4866
512 → 128 → 3	<b>0.5553</b>	0.5053
512 → 128 → 32 → 3	<b>0.5086</b>	0.4739

## 4 おわりに

顔画像から曖昧な内部状態の推定を行う際に、偏差モジュールを用いることで個人差の影響を低減させる手法を提案した。e-Learning 中の学習者の覚醒度推定問題について評価実験を行い、本提案手法の有用性を示した。

## 参考文献

- [1] W. Zhang, et al., "Learning a facial expression embedding disentangled from identity," *CVPR*, 2021.
- [2] K. Zhang, et al., "Joint face detection and alignment using multitask cascaded convolutional networks," *SPL*, 2016.

		指導教員	武村 紀子 准教授
学生番号	192C2087	氏名	古川 栄
論文題目	歩容映像における個人差を考慮した疾患推定		

## 1 はじめに

高齢化社会に伴い、歩行障害患者の増大が問題となっている。歩行障害をひきおこす代表的な疾患として、腰部脊柱管狭窄症（以下、LCS）や頸椎症性脊髄症（以下、CSM）が挙げられる。本研究では、患者の歩行映像から作成した平均シルエット画像（Gait Energy Image, 以下、GEI）[1] を用いて、LCS と CSM の疾患推定を行う。しかしながら、患者の GEI には図 1 のように疾患の特徴だけでなく個人の特徴も含まれており、そのような個人の特徴は疾患推定の精度を低下させる。そこで、変分オートエンコーダ（Variational Auto Encoder, 以下 VAE）[2] を用いて疾患特徴から個人特徴を分離し取り除くことで、精度の高い疾患推定を目指す。

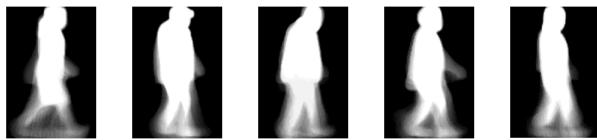


図 1: 個人差のある GEI

## 2 提案手法

本手法では潜在変数に意味付けが可能な Guided-VAE[3] を用いる。提案手法のネットワークを図 2 に示す。VAE の潜在変数を 2 つに分割し、一方を疾患が識別できるが個人は識別できないように、他方を個人は識別できるが疾患は識別できないように学習させ、前者の潜在変数のみを用いて疾患推定を行う。

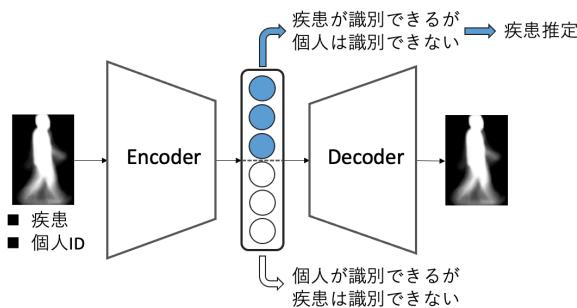


図 2: 提案手法のネットワーク

## 3 実験と考察

LCS 患者 142 人（GEI:49505 枚）、CSM 患者 133 人（GEI:28053 枚）、健常者 332 人（GEI:28817 枚）の歩容映像データを用いて、LCS/LCS 以外、CSM/CSM 以外、疾患（LCS, CSM）/健常の 3 つの識別問題について評価実験を行った。比較手法として、潜在変数の一方を疾患が識別できるように、他方を疾患が識別できないように学習させた場合（比較手法 1），潜在変数を分割せずに（変数の数は半分）疾患が識別できるように学習させた場合（比較手法 2）の 2 手法と比較した。実験結果を表 1 に示す。

表 1: 実験結果（5 分割交差検証、平均 F1score）

	LCS/その他	CSM/その他	疾患/健常
提案手法	<b>0.863</b>	<b>0.822</b>	0.952
比較手法 1	0.855	0.808	0.960
比較手法 2	0.851	0.789	<b>0.963</b>

LCS 推定器および CSM 推定器では提案手法が最も精度が高くなり、潜在変数において個人の特徴を分離させる本手法の有用性を示せた。一方、疾患推定器についてどの手法も推定精度が非常に高く飽和状態になっているため、ネットワーク構造が単純で学習が容易である比較手法の精度が高くなかった。

## 4 おわりに

本研究では、Guided-VAE を用いて歩容特徴（GEI）の個人差を考慮した疾患推定手法を提案した。評価実験により、潜在変数において個人特徴を分離させる提案手法の有用性を示した。今後は、本手法を多クラス分類に拡張させる。

## 参考文献

- [1] J. Han, et al., "Individual recognition using gait energy image," *IEEE TPAMI*, 2005.
- [2] D. P. Kingma, et al., "Auto-encoding variational bayes," *ICLR*, 2014.
- [3] Z. Ding, et al., "Guided variational autoencoder for disentanglement learning," *CVPR*, 2020.

		指導教員	武村 紀子 准教授
学生番号	21221008	氏名	藤村 昇矢
論文題目	座面角度が制御可能なスマートチェアの開発		

## 1 はじめに

近年、長時間のデスクワークによる集中力の低下や疲労感の増加が問題視される中、スマートチェアが注目を集めている。一般的なスマートチェア [1] は座面に搭載されたセンサによりユーザの姿勢や疲労状態を分析し、アラームを通知する。そのため、ユーザは通知に応じて、手動で座面角度を調整したり、着座姿勢を正す必要がある。そこで本研究では、自動的に姿勢を矯正させるため、座面角度を自動で制御可能なスマートチェアを開発する。

## 2 スマートチェアの開発

本スマートチェアは座面（天板）と土台を繋ぐ 2 台のアクチュエータを伸縮させることで座面角度を制御する（図 1 参照）。動作部（チェア本体）の機構は Kutzbach-Gruebler 方程式に基づき、停止時は自由度が 0、稼働時は自由度が 2（座面の前後左右方向）となるように設計した。また、制御部では、制御用 PC で作成したスマートチェア制御プログラム上の制御信号・命令をネットワークコンバータ、アクチュエータ用ドライバを介して電動アクチュエータへと送信する。

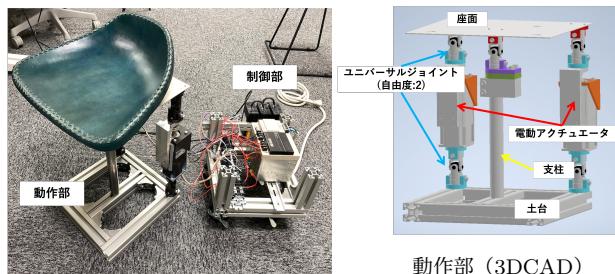


図 1: 開発したスマートチェア

## 3 評価実験

### 3.1 実験内容

開発したスマートチェアの使用感や効果を検証するため、被験者（20 代 3 名）に対して評価実験を行った。着席状態で講義動画を視聴する被験者に対し、5 分おきに座面角度を  $3^\circ$  変化させる。その際の姿勢の変化をスマートチェアに取り付けた座圧センサ [2] および 2 台のカメラにより計測する。また、実験後にスマートチェアの使用感に関するアンケートを実施する。

### 3.2 実験結果

評価実験中の姿勢変化を図 2 に示す。座面を前傾させた場合は、被験者の姿勢もそれに伴い前傾していた。一方、座面を左右に傾けた場合、被験者は傾けた方向と逆方向に重心を移動させ、姿勢を垂直に保っていた。また、アンケートの結果、スマートチェアの動作音や振動については気にならない人が多数派であった。

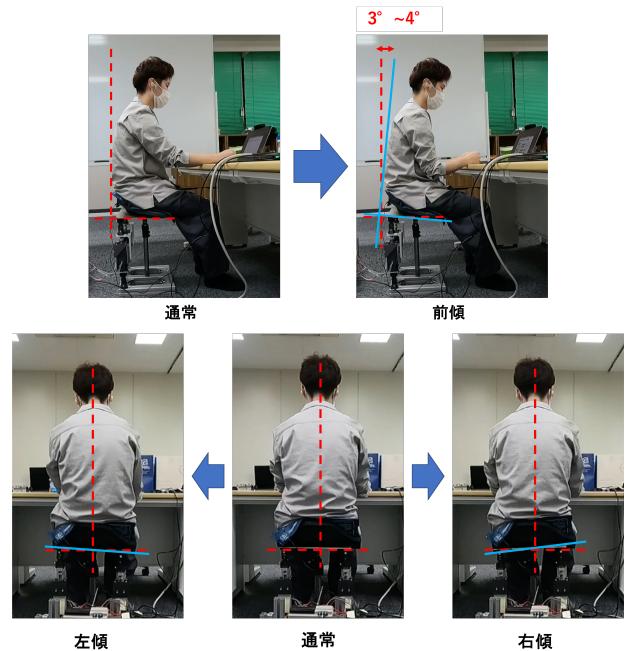


図 2: 評価実験中の姿勢変化

## 4 おわりに

本研究では、自動で座面角度を制御可能なスマートチェアの開発を行い、評価実験により使用感や座面角度変化による効果について検証した。今後の課題は、本実験結果に基づいたユーザへの適切な姿勢評価およびフィードバック方法の実装である。

## 参考文献

- [1] Teruhiro MIZUMOTO et al. Design and implementation of sensor-embedded chair for continuous sitting posture recognition. *IEICE Transactions on Information and Systems*, E103.D(5):1067–1077, 2020.
- [2] 体圧分布センサー（体圧測定器）SR ソフトビジョン - 住友理工. <https://www.sumitomoriko.co.jp/product/health/srsrv/>, [n.d.]. (Accessed on 31/01/2023).

		指 導 教 員	田向 権 教授
学生 番 号	192C1149	氏 名	溝 忠剛
論 文 題 目	直前の探索行動を元に未経験の行動経路を生成するための海馬に触発されたモデル		

## 1 はじめに

人間はいつ、どこで、何をという一連の情報をエピソード記憶として断続的に結びつけ脳内に記憶する。貯蔵されたエピソードは、意図して想起される。それによって人間は状況に合わせた柔軟行動ができる。海馬という脳部位がエピソード記憶には重要である。

海馬の機能を摸倣したニューロモルフィックハードウェアモデルを構築することで、低消費電力でかつ日常生活における柔軟な行動計画を行えるサービスロボットの実現が期待できる。ニューロモルフィックハードウェアへ適用可能な海馬モデルとして、中川ら [1] は、迷路探索課題において場所とキューバー信号を結び付けてイベントとし、イベント発生順を記憶するモデルを提案している。ただし、中川らの海馬モデルは報酬を 1 個しか扱えず、また経路順に沿ったイベントを想起する機能だけであった。そこで本研究では複数の報酬による学習ができるように改良を施し、新規経路生成と行動計画を行えるようにした。

## 2 方法

本提案モデルは中川の海馬モデル [1] の拡張モデルである。本研究では、迷路課題において、複数の報酬を想定し、ドーパミン濃度に依存したシナプス可塑性を導入した。これにより報酬の有無で大きく発火値に差が出るようになることを期待した。したがってドーパミン増減の検証には既存のモデルと増減するモデル 2つを用いて学習と記憶、そして想起を行い報酬のある地点と報酬のない地点の発火値の差を比べた。本研究で想起は学習して記憶した結果から行列計算を用いて自分が入力した座標と頭方位に関連する経路を出力することである。

新規経路生成と行動計画のアルゴリズムは学習して記憶したものから時系列の順方向想起と逆方向想起した発火値の結果を用いて行動が確率に従って行われるものを作成した。順方向想起は学習、記憶したときに経験した時系列通りの情報が得られ、逆方向想起は学習、記憶時の時系列とは逆順の結果が得られるため、これら 2つの結果を用いることで新規経路生成が期待できる。それらを用いて実際に行動を確率に従わせたこと理由は、今後情報の種類や数が増えた場合でもそれら全てが確率に収束するという一つの目的につなが

るため、そういったケースでも柔軟に行動計画を行えると考えたからである。アルゴリズムは説明した確率に加えて報酬が得られなかった経路の情報を減衰させることをベースに作成し検証タスクでは同じスタート地点から 20 回アルゴリズムをもとに試行した。

## 3 結果

ドーパミン増減を導入したことでの報酬の有無ではっきりとした違いが出るようになった。

アルゴリズムの検証タスクにおいても複数の新規経路を生成し行動を行うことができることを確認できた。

## 4 考察

ドーパミンの増減を導入して報酬によって柔軟かつはっきりと発火値が変動することがわかったため、今後学習して記憶する過程では複数回報酬を獲得するようなエピソードを想定することができると考えられる。

アルゴリズムの検証タスクの結果から発火値の確率による新規経路生成と行動計画ができていたため、今後新たに多くの情報を加えたとしても学習して記憶するモデルか行動計画アルゴリズムに適用することで柔軟に新規経路生成と行動計画を行うことができるようになったと考えられる。また本研究を通してモデルを適用するハードウェア、学習して記憶するモデル、新規経路生成と行動計画のアルゴリズムという 3つが新しいサービスロボットへの応用に近づくことができたと考えられる。

## 参考文献

- [1] Hiroki Nakagawa,Katsumi Tateno,Kensuke Takada,Takashi Morie.A Neural Network Model of the EntorhinalCortex and Hippocampus for Event-Order Memory Processing, IEEE access, Volume 10, 2022.

九州工業大学 情報工学部 知能情報工学科 2022年度 卒業論文 概要

		指導教員	础崎 賢一
学生番号	192C1023	氏名	井上 凌
論文題目	クラスを単位とした遠隔プログラミング演習支援システムにおけるアカウント管理に関する研究		

1 はじめに

近年、遠隔授業が普及してきたが、対面授業と異なり、学生の活動状況を把握しづらいため、演習を行うに当たって苦慮する教員も多い。そこで我々は学生の学習状況把握を目的とした遠隔プログラミング演習支援システムを研究している。本研究では、遠隔プログラミング演習支援システムで多数の学生の演習を支え、学生の学習状況を取得・可視化するためのアカウント管理方法を提案する。

2 遠隔授業

## 2.1 遠隔授業の種類

遠隔授業には同期型と非同期型の2種類の授業形態が存在するが、本研究では同期型の遠隔授業を対象としている。

## 2.2 遠隔授業の問題点

教師から学生の姿が見えないため、演習時間に学生が何を言っているかわからず、適切な指導を行えないという問題がある。

### 3 遠隔プログラミング演習システム

遠隔プログラミング演習支援システムでは、学生にWEBで動作するIDE(統合開発環境)を提供する。ファイルのコンパイル・実行環境はサーバ上に構築されており、すべての学生が統一されたコンパイル・実行環境を利用できる。また、編集やコンパイルなどの学生の演習状況はサーバに記録され、その情報を可視化することで、教員が学生の学習状況を一目で把握することができるようしている。

4 実現手法

先行研究では、学生の学習状況の記録や可視化といった基本機能の評価用のシステムを構築したが、多数の学生が利用できる状態ではなかった。そこで本研究では実用に向けて、多数の学生がそれぞれ独立した環境で演習を行えるようアカウントシステムの試作や性能評価を行った。

## 4.1 データベースの実装

学生の学習状況を可視化するために、学生の活動状況を取得し、サーバに保存する必要がある。そのためにデータベースを作成した。結果を図1に示す。

データベースに保存する情報は、一定時間毎の編集情報、コンパイル時の情報、実行時の情報が存在する。そして可視化する情報には、クラス全体の学習状況、学生個人の学習状況、活動時の学生の画面の復元が存在する。

保存する情報としては、その時の学生のプログラムを学生 id やその日時と共に、srcprog テーブルに保存し、また、現在のプログラムと直前のプログラムの文字数の変化量を学生 id や比較した 2 時刻のプログラム id と共に、srcsizes テーブルに保存する。今回は、表示のためにプログラムを再度計算する手

間を考慮し、予めプログラムの文字数を計算し、結果をテーブルに保存している。

これらを利用して多数の学生の情報を取得して可視化することに成功した。

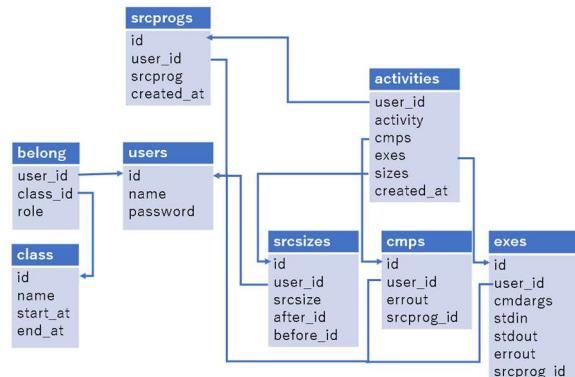


図1 データベースの構造

## 4.2 負荷テスト・評価

多数の学生の利用を想定した負荷テストを行った。サーバとして使用したPCの性能を次に示す。

OS : CentOS, CPU : Intel(R) Core(TM) i5-2500, 4-Core, 3.30GHz, 記憶装置 : SSD, メモリ : 24GB

また、かける負荷として、「学生の編集情報をDBへ保存するリクエスト」、「コンパイルし、その記録をDBへ保存するリクエスト」をクライアントからサーバへ高速に送信する。このときコンパイルするソースファイルは、104行のものを使用した。

今回の負荷テストでは、10分間継続してリクエストをサーバに送信し、このときの Load Average や CPU 使用率、エラーが起きないかを調査する。負荷テストの結果を表1に示す。

表1 負荷テストの結果

リクエスト内容	処理可能速度	最高Load Average	CPU使用率
編集情報	100回/s	2.35	9~11
コンパイルと保存	40回/s	3.51	38~40

この結果から、秒間 100 回学生の編集情報を保存でき、秒間 40 個のリクエストを処理できる。100 人クラスの授業において、毎分学生の情報を取得し、各学生が毎分 1,2 回コンパイルのリクエストを行う。よって、サーバは負荷に耐えられることがわかった。また、Load Average や、CPU 使用率を参照してもサーバへの負荷は問題ないことが確認できた。

5 まとめ

本研究では、多数の学生を対象とした遠隔プログラミング演習支援システムのアカウント管理システム、情報管理システム、クラス単位での運用をするための性能評価を示した。

		指導教員	磯崎 賢一
学生番号	192C1085	氏名	高瀬 峻成
論文題目	大規模地震時の建物の損傷調査に向けた 三次元点群データの分散処理に関する研究		

## 1 はじめに

近い将来日本では都市部を中心とした大規模地震が起こると想定されている。地震後、被災地では建物状況把握のため損傷調査が行われる。しかし、現行の調査方法には時間と安全性の面で問題がある。本研究では損傷調査に三次元点群データとWEBブラウザを用いることで、大規模な損傷調査を迅速かつ安全に分散処理できるシステムについて提案する。

## 2 現行の震災時建物損傷調査の問題点

現行の調査方法では、日本全国から調査員を集め建物1つ1つを目視で確認する。そのため、大規模地震の被災建物すべての調査には多くの人出と時間が必要で、建物倒壊による二次災害の恐れがある。

## 3 現行の震災時建物損傷調査の問題解決法の提案

### 3.1 三次元点群データの利用

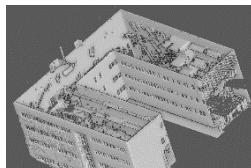
三次元点群データはレーザースキャナーにより計測される。レーザースキャナーでは短時間で一度に複数建物の三次元点群データを容易に計測できる。また電子データであるため、ネットワーク上でのデータの共有が容易である。そこで、計測された三次元点群データを計測者が個別に処理し、ネットワークで集約する損傷調査システムを提案する。

### 3.2 WEB ブラウザの利用

大地震における点群データはその量が膨大であるため、それらをサーバーに集め一括で処理することは処理能力的にも記録容量的にも不可能である。このため、損傷分析のためのデータ処理を分散して行う仕組みや、計測精度や密度を維持しながら計測データ容量を大幅に縮減する仕組みが必要である。本研究ではWEBブラウザで稼働する損傷分析システムを構築し上記のような仕組みの実現法を提案する。

## 4 点群データを損傷調査に適した形にする処理

今回使用した点群データを図1に示す。これは講義棟を計測した点群データであり、点数は90,309,957点、データ量は1.54GBである。



### 4.1 建物点群の抽出

計測した三次元点群データには、複数の建物や建物以外のデータが存在している。損傷調査を行うためには1つの建物のみの点群データが必要であるため、基盤地図情報が提供する建物の外周線情報をもとに建物構成点群のみを抽出するシステムを作成した。

### 4.2 RANSACによる壁面推定

後で示す計測データの高精度化・軽量化の処理を行うために、壁面ごとの点群の分割が必要である。そこでRANSACを用いて建物の壁面を推定し抽出する機能を作成した。これにより、建物の点群データ

から壁面単位でデータを取得できるようにした。

### 4.3 座標軸に沿うような面の回転

以後の処理では、分析処理を単純化するために、壁面の平面がXY平面上に乗るように正規化することが求められる。そこで推定した平面の方程式とともに壁面の平面を正規化できるようにした。

### 4.4 統計処理

計測した点群データには、3-5mm程度の計測誤差があり、そのままでは1mm程度の変位量の検出を必要としている損傷分析には利用できない。この問題を解決するために、計測平面をメッシュ分割し、メッシュ内の点群ごとに真値推定を行い、計測データを高精度化するようにした。これにより、精度向上、データ量の削減、および各部位で一意な計測結果を得られるようになれた。

## 5 結果

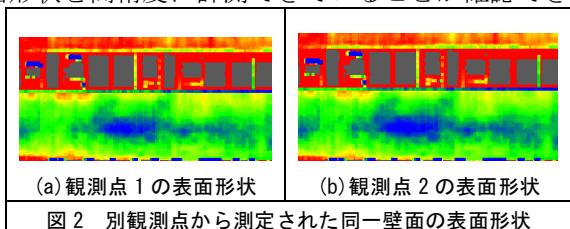
実行環境を表1に示す。

表1 プログラム実行環境

CPU	メモリ	使用環境
Intel(R) Core(TM) i5-2500	16GB	GoogleChrome

### 5.1 建物の表面形状

建物の壁面の高精度化データについて、別地点から観測された同一壁面の精度比較を行った。図2を見ると、ほぼ同様な表面形状(凹凸)が求められていることがわかる。これより点群データで壁面の表面形状を高精度に計測できていることが確認できた。



### 5.2 点群の処理速度

作成したシステムでのファイル読み込みや建物抽出、平面推定及び13面の点群の抽出と座標軸への統一処理、統計的処理における真値推定の合計処理時間は51分ほどであった。損傷調査には迅速な評価が求められるが、評価を行うためのデータ処理が十分な速度で行われていると考えられる。またこれらの処理で生成されたデータ量は58MBと元データの約4%であり、大幅な削減ができていることがわかる。

## 6 まとめ

本研究では、三次元点群データとWebブラウザを用いた損傷調査に向けた点群データの分散処理システムを提案し試作した。また、その有効性を確認した。

		指導教員	磯崎 賢一
学生番号	192C1145	氏名	松本 玲音
論文題目	震災前後を想定した二時期の建物三次元点群データの差分による損傷検出手法に関する研究		

## 1 はじめに

大地震が発生し建築物に被害が発生した際、建築物の損傷調査が行われる。しかし、現在の損傷調査には調査員が被災地に赴き目視で行うことにより、多くの人手や時間といったコストがかかる、目視での高層の建築物の調査が困難である、二次災害のリスクがあるといった問題が発生する。そこで我々は、レーザースキャナにより非接触かつ高速な計測により得られる三次元点群データを用いた損傷検出手法を提案する。

## 2 我々が提案する損傷検出手法

我々は、点群データを用いて建築物の壁面に生じた損傷を、機械的に一括に検出することを目的としている。しかし、損傷を単純に建物表面の凹凸であるとみなした場合、建築物の意匠や施工上の凹凸と、損傷による凹凸との判別が困難であるという問題がある。その問題を解決する方法として、我々は被災前後の二時期の点群データを用いた損傷検出手法を提案した。この方式では、被災前後で変化の起こらない意匠による凹凸は取り除かれ、変化が生じた損傷による凹凸のみが取り出されるはずである。

現状、この提案方式の有用性は縦2[m]、横3[m]の実験用コンクリート壁面ですでに確認している。本研究では、提案手法の実用性を確認するために、実建築物の壁面を対象に提案手法による損傷検出手法を行った。

## 3 提案手法の具体的処理内容

### 3.1 点群データの高精度化処理

点群データを建築物の損傷調査に用いる場合、損傷によるわずかな変位量を検出しなければならないため、サブミリ単位の精度が求められる。レーザー計測により得られた点群データには、3-5ミリ程度の計測誤差が生じ、要求精度を満たせないためそのまま用いることはできない。そこでまず、点群データに対して高精度化のため真値推定を行う。

### 3.2 二時期の点群データの位置合わせ

レーザースキャナによる計測では、各計測で座標値に微小なずれが生じてしまう。よって二時期での比較を行う前に計測対象の位置合わせを行わなければならない。位置合わせでは、二時期の壁面を鉛直、水平方向の座標軸に平行になるよう回転移動を行った後、対応する部位が重なるよう平行移動させる処理を行う。

### 3.3 二時期の点群データの差分による損傷検出

最後に、高精度化処理・位置合わせを行った二時期の点群データを用いて奥行方向の差分を求ることで変化が生じた損傷部位のみを取りだす。

## 4 実験と結果

### 4.1 実験方法

我々が提案している損傷検出手法の有用性を確認するために、実建物を対象として、震災による損傷を模した損傷模擬材を設置した壁面と設置する前の壁面の二時期の点群計測を行い提案方式の評価を行った。まずは、損傷模擬材の断面図を図1に示す。断面図のように、損傷模擬材は部位により厚みを変更している。

**図1 損傷模擬材断面図**  
壁面  
その中で最も薄い部位は1[cm]  
程度である。

次に、実験に用いた点群データを図2に示す。なお、損傷模擬材の可視化のため反射強度に基づいて色付けを行った点群データを表示する。

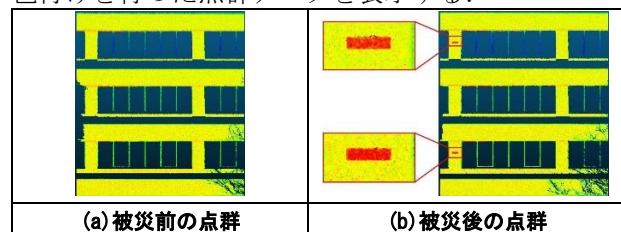


図2 実験に用いた点群データ

この二時期の点群データを対象として高精度化処理・位置合わせ処理を行い、差分を計算した結果5[mm]以上の点を損傷とみなし、今回のような損傷による小さな凹凸に関しても抽出できるか確認した。

### 4.2 実験結果

被災後の点群データのみで奥行方向の変位量に基づいて色付けを行った結果、損傷模擬材はほとんど確認できなかった。しかし、二時期での点群データに対し高精度化処理を行い、奥行方向の差分が5[mm]以上の点を求めた結果、一時期の結果と比較すると損傷部分をはっきりと検出することができた。また、厚みの違いについても確認できるほど高精度な検出ができた。結果を図3に示す。

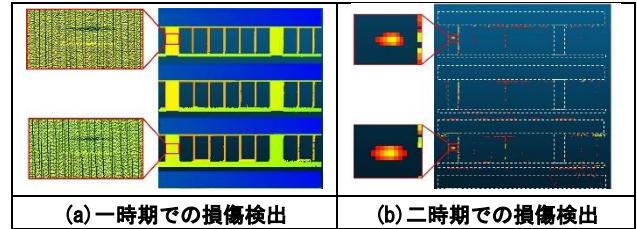


図3 実験結果

## 5 まとめ

実際の建築物の壁面を対象として、我々が提案している二次期の点群データの差分による損傷検出手法の評価を行った。本稿で行った評価により、損傷箇所の高精度な検出が行えることが確認できた。

		指導教員	磯崎 賢一
学生番号	192C1164	氏名	山㟢 韶希
論文題目	WEB ブラウザを使用した CityGML の 3 次元表示システムに関する研究		

## 1 はじめに

CityGML は 3 次元都市モデルを記述するデータフォーマットの国際標準規格である。CityGML は建物をはじめとする都市の多様な地物の幾何情報に意味情報を付与できる XML ベースの形式となっている。しかし国際規格にもかかわらず、CityGML に直接対応したアプリケーションがないため、3 次元表示するためには他のグラフィックデータ形式に変換する必要があった。そこで本研究では、CityGML を直接読み込み 3 次元表示できる WEB ブラウザシステムを提案する。

## 2 CityGML の概要

### 2.1 CityGML について

CityGML は各地物の幾何情報だけでなく、用途や種類、災害リスク等の意味情報を記述できる。これらの情報を使用することで、各種シミュレーションや都市計画への活用や分析を 3 次元都市モデル上で行うことが可能となる。

### 2.2 LoD について

CityGML ではモデルの表現に LoD (Level Of Detail) を使用している。CityGML における LoD とは意味情報からモデルの重要度を考慮して設定されたモデルの詳細度を指す。LoD0 から LoD4 までの詳細度があり、CityGML の記述者は目的や用途に合わせて適切なモデルを選択できる。

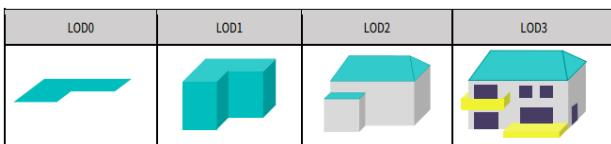


図 1 建築物における LoD の例

### 3 データ解析とメッシュ生成の手法

本研究では意味情報を含めて建物等の地物を可視化することを想定しているため、内部処理では、地物ごとに意味/幾何情報を格納するオブジェクトを作成し、操作や表示に使用する。

#### 3.1 幾何情報の取得

地物の幾何情報は各 LoD を示すタグの最下層にあるタグ<gm1:posList>の中に、一つの面の頂点座標として地理座標系で表現されている。一般的な表示システムは通常平面直角座標系で表されるため、WEB ブラウザ上で 3D モデルを作成する際に扱いやすい平面直角座標系に変換した。建物の屋根や壁を構成する面はポリゴンに分割し、LoD ごとにオブジェクトの幾何情報として記録した。

#### 3.2 意味情報の取得

地物の種類や用途など意味情報を関係するタグから取得し、オブジェクトに格納する。意味情

報によっては別ファイルの情報と紐づける処理を作成した。

### 3.3 3 次元都市モデルの表示

JavaScript のライブラリ「Three.js」を使用して地物のオブジェクトに格納された幾何情報を使用して、都市の 3D モデルを表示するシステムを構築した。表示システムには、地物ごとや LoD 毎に表示するモデルを選択できる仕組みを実装した。

## 4 表示の結果と評価

### 4.1 実験

3 章で述べた手法を基に CityGML を 3 次元表示した。地図の拡大縮小や回転、表示するオブジェクトの切り替え、建築物の意味情報を利用した色分けが可能となっている。

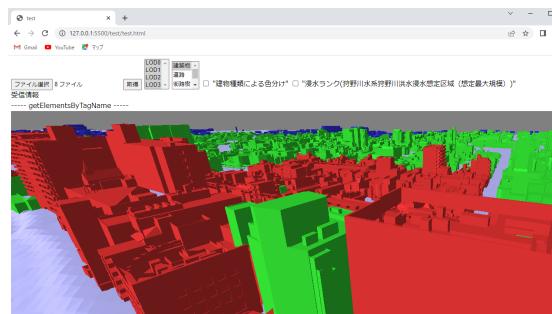


図 2 WEB ブラウザ上で都市モデルの表示

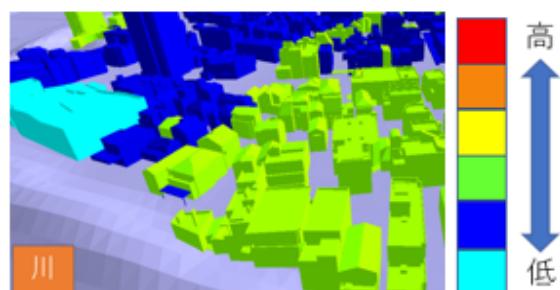


図 3 洪水に対する浸水ランクで色分け

### 4.2 評価

建物、街路樹、道路それぞれの各 LoD の特徴となる構造やテクスチャが国土交通省が提供している PLATEAU VIEW と同様の幾何形状で表示されていることを確認した。また、WEB ブラウザ上から LoD 毎の表示や意味情報による色分けを行えることから、意味情報を活用できることを確認した。

## 5 まとめ

本研究では CityGML 形式を直接読み込み 3 次元表示を行える WEB ブラウザシステムを提案した。

		指導教員	藤本 晶子 准教授
学生番号	192C1021	氏名	井上 一成
論文題目	磁気嵐に伴う宇宙デブリ軌道落下量モデルの検討		

## 1 はじめに

近年、人類の宇宙利用が加速しており、人工衛星の数は、現在までに世界で約 1.4 万機にものぼる [1]. その打ち上げ数も年々増加している [1] ため、地球軌道上の運用中の人工衛星や宇宙デブリの総数が、近年飛躍的に増加している [1]. 一方で、これらの地球軌道上の宇宙オブジェクトの多くは、高度 90km から 800km の熱圏と呼ばれる領域の大気密度変動の影響を大きく受ける。太陽活動の影響を受けて極域大気がジュール加熱され、その結果として地球規模の大気密度の増加が大気抵抗の増大へつながる。大気抵抗の増大は人工衛星や宇宙デブリの軌道長半径減少を引き起こす。例として、2022 年 2 月 3 日にスペース X 社が打ち上げたスターリンク衛星が磁気嵐発生に伴い、大気圏へ再突入した [2]. 軌道制御できない宇宙デブリについて、その軌道長半径の減少とその原因となる大気抵抗変動を支配する太陽活動との関係をモデル化することで、太陽活動から宇宙デブリ軌道落下量をモニタリングできることを目指す。本研究では、モデル化にあたり、説明変数としてどの物理量が重要かについて解析を行う。

## 2 宇宙デブリ軌道落下量モデル検討のための実験

宇宙デブリ軌道落下量の説明変数として利用できる物理量を選定するために、(1) K-means アプローチと、(2) 探索的アプローチを実施する。

(1) K-means アプローチ：太陽風と地磁気活動指数のデータと宇宙デブリのデータを用いて、K-means によるクラスタリングを行う。

(2) 探索的アプローチ：太陽風と地磁気活動指数のデータと宇宙デブリの軌道落下量データを用いて、次の 4 つの探索的データ解析を行う。

1. 宇宙デブリのサイズと軌道落下量
2. 宇宙デブリの初期高度と軌道落下量
3. 磁気嵐規模と軌道落下量
4. ジュール加熱と軌道落下量

## 3 結果

(1) K-means アプローチ：宇宙デブリ軌道落下量の規模でクラスタリング分類を試みたが、期待するような結果は得られなかった。

## (2) 探索的アプローチ：

- 宇宙デブリのサイズが小さい方が、大気抵抗の影響が大きいため軌道落下量は増大する。
- 宇宙デブリ初期高度及び磁気嵐規模は軌道落下量への影響が小さい。
- 図 1 で示すように、CME タイプの磁気嵐の場合、増大したジュール加熱と軌道落下量の間に正の相関関係があった。

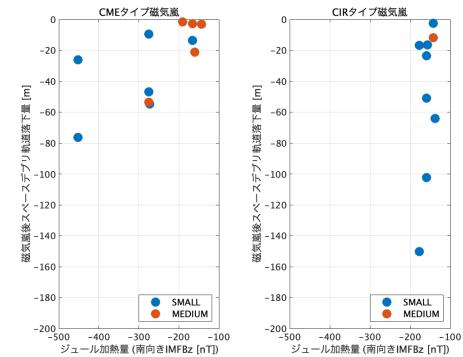


図 1: ジュール加熱と宇宙デブリ軌道落下量

## 4 おわりに

本研究では、太陽風と地磁気活動指数データ及び宇宙デブリ軌道落下量データを用いて K-mean によるクラスタリングと探索的データ解析を行った。その結果、宇宙デブリ軌道落下量モデルの説明変数として、宇宙デブリサイズ、ジュール加熱（南向き IMFBz）、磁気嵐タイプが有用であることがわかった。今後の展望として、本実験では、高度 600km 以上を対象としたが、ジュール加熱の影響が大きい、より低高度の宇宙デブリ軌道落下量についての実験も行い、落下量モデルを作成し、宇宙デブリ軌道落下モニタリングを実施する。

## 参考文献

- [1] Our World in Data, (参照日:2023/02/07), <https://ourworldindata.org/>
- [2] Dang, T., Li, X., Luo, B., Li, R., Zhang, B., Pham, K., et al. (2022). Unveiling the space weather during the Starlink satellites destruction event on 4 February 2022. Space Weather, 20

		指導教員	藤本 晶子 准教授
学生番号	192C1107	氏名	中村 駿仁
論文題目	ソフトウェアラジオデバイスを用いた電離圏シンチレーション計測器の開発		

## 1 はじめに

電離圏は高度 60 ~ 600km に分布する大気の領域である。電離圏に関する電波伝搬の特性は GNSS 測位や短波放送等と幅広い分野で利用されている。しかし、宇宙天気現象による電離圏の乱れは電離圏を通過する信号の振幅・強度変化を引き起こし、我々の社会システムに悪影響を与える。この影響を予測・抑制するためには継続的な電離圏環境のモニタリングが必要となる。本研究では、電離圏モニタリングシステムとして GPS 信号を用いたシンチレーション計測器の開発を目的とする。また、図 1 に示す観測手法のうち、既存の FMCW と統合したシステムとすることを視野に入れると、ソフトウェアラジオデバイスを使用した独自での計測器およびシステム開発を行った。

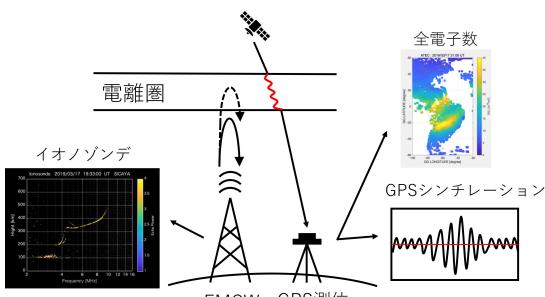


図 1: 主要な電離圏環境の観測手法

## 2 ソフトウェアラジオデバイスによる電離圏シンチレーション計測装置の提案

本論文で提案するシステムの概略を図 2 に示す。アンテナより受信した GPS 信号を、Raspberry Pi をベースに作成した GPS シンチレーション計測器によって信号処理を行うことで信号強度等のデータを取得する。それらのデータを RINEX format に従って出力し、そのファイルを入力として  $S_4$  指数を算出する。

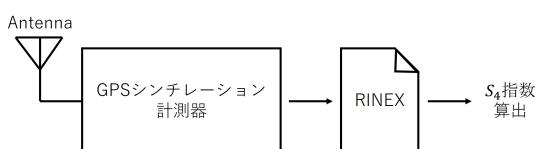


図 2: 電離圏シンチレーション観測システムの概略

## 3 電離圏シンチレーション観測装置による観測実験

提案した電離圏シンチレーション観測装置による実際の観測実験を実施した。また、同じシステムの計測器により計測された 2022 年 10 月 22 日の磁気嵐について、20 日から 10 日間の観測データを元に  $S_4$  指数の算出を行った(図 3)。 $S_4$  指数は、 $I$  を電波強度、 $\langle \cdot \rangle$  を平均値のオペレータとして、次の式で表される [1]。

$$S_4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}}$$

22 日の磁気嵐発生前後の各指標値や測位誤差量の推移を比較すると、磁気嵐発生後には  $S_4$  指数平均と測位誤差量の推移にて最大値の増加が見られた。

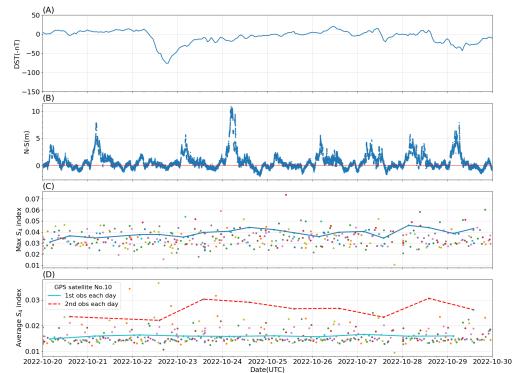


図 3: Dst 指数、GNSS 誤差量 (南北)、 $S_4$  指数の最大値と平均値 (B,C 中の線は GPS 衛星 10 番の観測値)

## 4 おわりに

本実験では、ソフトウェアラジオによる GPS シンチレーション計測器の構築ならびに取得したデータより  $S_4$  指数を算出する一連のシステムが動作することを示した。今後の展望として、GPS 信号の受信により適した環境での継続観測や、FMCW と統合した観測システムの開発、電離圏モニタリングアプリケーションの作成などが挙げられる。

## 参考文献

- [1] E. J. Fremouw, R. L. Leadabrand, et al. Early results from the dna wideband satellite experiment —complex-signal scintillation. *Radio Science*, Vol. 13, No. 1, pp. 167–187, 1978.

		指導教員	藤本 晶子 准教授
学生番号	192C1137	氏名	前田 愛乃
論文題目	自然言語処理を用いた宇宙天気ニュース記事検索システム		

## 1 はじめに

太陽は、生命にとって危険な X 線・紫外線などの電磁波、高エネルギー粒子、あるいは太陽風を噴出することがある。特に、太陽フレアやそれに伴う高エネルギー粒子、コロナ質量噴出、高速の太陽風、およびフィラメント放出に伴う太陽風などが地球や生活に与える影響は大きい。<sup>[1]</sup>

過去に発生した宇宙天気のアーカイブデータを利活用し、宇宙天気災害をもたらす宇宙天気現象の特徴を参考し、現在の宇宙天気状況について評価することは、宇宙天気災害対策につながる。その際、低次元の多くの数値データではなく、数値を言語化した高次元の文章アーカイブデータを利活用することが、宇宙天気の概況を知る上で最も有効な手段であると考える。そこで、文章データに自然言語処理を用いてアプローチし、宇宙天気ニュース記事検索システムを開発することを提案する。

## 2 宇宙天気ニュース記事検索システム

本研究のデータは、ウェブサイト「宇宙天気ニュース」<sup>[2]</sup>から引用する。こちらのウェブサイトは見出し文と詳細解説文で構成された今日の宇宙天気の記事である。2003/11/18 から今日に至るまでの過去記事が存在し、本研究では、2003/11/18 から 2023/1/19 までの記事の見出し文をデータとして扱う。見出し文を形態素解析し、入力した任意の日の宇宙天気見出し文と似た過去の宇宙天気ニュース記事の見出し文を検索するシステムを開発する（図 1）。見出し文は、TF-IDF によってベクトル化し、文書間をコサイン類似度計算によって 1 に近い順番、つまり似ていると判断された順番に出力する仕組みを採用する。

## 3 宇宙天気ニュース記事検索実験

3 つの実験を行う。実験 1 は、動詞を除外する場合としない場合のキーワードによる比較実験と、形状詞を除外する場合としない場合のキーワードによる比較実験である。本研究では、動詞を除外する場合と、形状詞を除外しない場合が最も検索見出し文と類似した見出し文が検出された。実験 2 は、宇宙天気辞書を適用しない場合と適用する場合の比較実験である。本研究では、宇宙天気辞書を適用する場合が最も検索見出し文と類似した見出し文が検出された。実験 3 は、2022

年 12 月 1 日～31 日のそれぞれの検索見出し文について、出力された検出見出し文の上位 3 件を主観評価し、宇宙天気ニュース記事検索システムの有効性を示すシステムの評価実験である。主観評価の結果好ましい値が出たため、システムの有効性が明らかとなった。

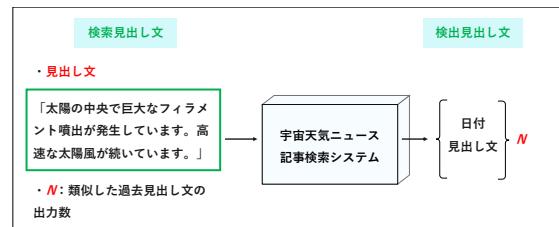


図 1: 宇宙天気ニュース記事検索システム

## 4 おわりに

本研究では、自然言語処理を用いて宇宙天気ニュース記事検索システムを開発し、キーワードによる比較実験と宇宙天気辞書による比較実験を行った。不必要な品詞だった”補助記号”, ”助詞”, ”助動詞”に, ”動詞”や”形状詞”も加え、それらを除外してキーワードの組み合わせを変える比較実験を行った。その結果、動詞も除外するキーワードの組み合わせの場合に、最も検索見出し文と類似した見出し文が検出された。また、宇宙天気専門用語に特化した辞書である宇宙天気辞書を適用することで、宇宙天気辞書を適用しない場合と比べて、検索見出し文と類似した見出し文が検出された。これにより、検索見出し文に形態素解析を行っただけのシステムから宇宙天気ニュース記事検索システムの精度が向上したことが言える。

## 参考文献

- [1] 名古屋大学学術機関リポジトリ, 太陽地球圏環境予測 オープン・テキストブック, (参照日:2023/01/24) , [https://nagoya.repo.nii.ac.jp/search?search\\_type=2&q=1621213872938](https://nagoya.repo.nii.ac.jp/search?search_type=2&q=1621213872938).
- [2] 篠原学, 宇宙天気ニュース, (参照日:2022/09/15) , <http://www.kagoshima-ct.ac.jp/>.

		指導教員	藤本 晶子 准教授
学生番号	192C1146	氏名	儘田 龍一
論文題目	文字列圧縮・照合技術を用いた磁気圏静電孤立波動検出		

## 1 はじめに

磁気圏では頻繁にパルス状に孤立した双極子型波形の電場変動「磁気圏静電孤立波動(ESW: Electrostatic Solitary Waves)」が発生している。この ESW は、磁気圏プラズマ粒子どうしのエネルギー交換としての波動-粒子相互作用において重要な役割を担う。ESW 発生分布や波動特性を分析することは、磁気圏で起きている物理素過程の理解につながる。ESW 検出方法の先行研究として、(1) 画像フィルター処理方法 [1] と (2) 時系列フィルター処理方法 [2] がある。(1) は波形の数値データを画像として処理するため、数値データを直接処理するより処理コストが上がる。またフィルターサイズに制限があり、検出できる ESW 波形特性の制約が生じる。(2) は ESW に重畳する高周波成分と背景場の低周波成分を除去できるが、厳密な ESW 波形区間の抽出はできず、手作業での波形検出となる。そこで本研究では文字列連長圧縮・照合技術を応用した、低計算コストかつ検出できる波形の制約のない ESW 検出モデルを提案する。

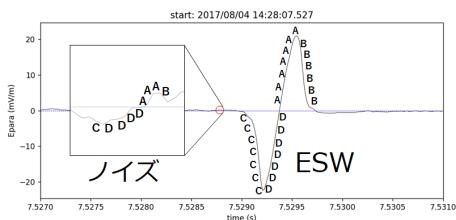


図 1: ESW とノイズの違い

## 2 磁気圏静電孤立波動検出モデル

磁気圏静電孤立波動検出モデル(ESW 検出モデル)は、(1) 数値-連長圧縮文字列変換器 (2) ESW 波形文字列検索器 (3) 特徴量算出器 (4) ESW 判別器の 4 つの処理部から構成される。本モデルでは、ESW の概形を文字列で表現することで、ESW 電場数値データから ESW 波形区間を検出するタスクを文字列照合の問題として取り扱う。ESW の典型パターンの波形長(=データ点数)がノイズ波形長に比べて長い(=データ点数が多い)という特徴がある(図 1)。この ESW 波形長の特性を、文字列の連長圧縮率  $cps$ 、文字列長  $len$  として定量化し、上記(4)の ESW 判別器における特徴量として扱う。ESW の典型パターンの波形の連長圧縮後

の文字列は「CDAB」となる。この「CDAB」の 1 文字目と 3 文字目の連長数をそれぞれ  $hl, tl$  とし、ESW 判別器の特徴量として選択する。ESW 判別器は、説明変数に上記の特徴量  $cps, len, hl, tl$ 、目的変数に ESW の正解ラベル・不正解ラベルを与えた時の分類木学習によって判別条件を獲得する。

## 3 実験

本研究では、(A) ESW 判別器の判別条件の選定実験と (B) ESW 検出モデルの性能実験の 2 つの実験を行う。ESW データとして、あらせ衛星プラズマ波動観測器による磁場に並行な電場データ  $E_{\parallel}$  を使用する。期間は 2017/8/4 14:28:04 から 2017/8/4 14:28:09 の計 5 秒間を用いる。5 秒のデータ長は 0.1 秒間隔を一区切りとして扱い、学習データは任意の 0.1 秒区間を 10 件(計 1 秒)、検証データは残りの 4 秒分のデータとする。(A) ESW 判別器の判別条件の選定実験では、ESW 判別器の 4 つの特徴量  $cps, len, hl, tl$  の組み合わせを変えた場合の 15 通りの決定木(分類木)の学習を行う。決定木の学習結果に対し再現率  $r$ 、適合率  $p$ 、F 値  $f$  を算出し、これらの値が比較的高い、特徴量条件 7 組を選出する。(B) ESW 検出モデルの性能実験では、(A) で選出した 7 組の特徴量条件の組み合わせそれぞれを用いて ESW 検出を行う。その後、それぞれの組み合わせでの検出性能を  $r, p, f$  により評価する。評価が良かった ESW 検出モデルは、結果 1:  $cps \leq 17$  のとき  $r = 0.56$ ,  $p = 0.87$ ,  $f = 0.68$ 、結果 2:  $len \geq 36, hl \geq 9$  のとき  $r = 0.71$ ,  $p = 0.37$ ,  $f = 0.49$  であった。

## 4 おわりに

本研究では、パルス波 ESW の検出モデルを提案した。本検出モデルにより、先行研究の抱える問題点を克服することができた。一方で、検出性能に関しては改善の余地がある結果となった。今後は本研究の提案したモデルを基にアンサンブル学習を行うなどして検出性能の向上を目指す。

## 参考文献

- [1] H.Kojima, Y.Omura, H.Matsumoto, K.Miyaguti and T.Mukai, Automatic Waveform Selection method for Electrostatic Solitary Waves, Earth Planets Space Vol52, pp.495-502, 2000.
- [2] 八木大介, NICT サイエンスクラウドを用いた科学衛星観測データの並列分散解析環境の構築, 金沢大学, 2015, 修士論文.

		指導教員	齊藤 剛史 教授
学生番号	182C1074	氏名	桑木 俊輔
論文題目	イベントベースカメラを用いた単語読唇に関する研究		

## 1 はじめに

イベントベースカメラは生物が光を感じる仕組みを模したセンサである。イベントベースカメラは高時間分解能やハイダイナミックレンジなどの特長を持ち、輝度の変化を画素毎に独立して非同期に検知し、データを出力する。従来型カメラを用いた、顔や機械読唇に関する研究は進んでいるが、イベントベースカメラを用いた研究は未だ少ない。本研究室では昨年、イベントベースカメラを用いて、機械読唇の分野の一つである単音読唇の研究が行われた [1]。単音読唇の場合は、イベントベースカメラを用いることで精度が向上した。本研究ではその応用として、イベントベースカメラを用いて単語読唇に取り組む。また、各実験に用いるデータは独自で収集した。

## 2 イベントを用いた画像生成

従来型カメラは一定の時間間隔で全画素の輝度を画像という形で記録する。一方、イベントベースカメラは非同期的に画素毎のしきい値以上の輝度変化を「イベント」と呼ばれる形で記録する。イベントには正と負の極性が存在し、イベントが発生した画素の時間、座標、極性が outputされる。本研究では、画素毎に記録されたイベントの極性に応じて、画素値を決定した画像を実験に用いる。また、従来型カメラから得られる画像をフレーム画像、イベントを用いて生成した画像をイベント画像と表記する。図 1 にそれぞれの画像例を示す。

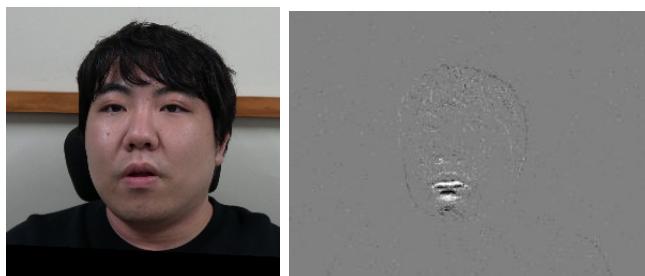


図 1: 2 種類の画像

## 3 単語読唇

単語読唇は母音や店名を含む 101 単語を認識対象とし、MS-TCN[2] を用いて認識実験を行う。21 名の

被験者を集め、101 単語の発話シーンを 5 パターン撮影し、計 10605 シーンを収集した。グレースケールの 29.97fps のフレーム画像 (frame)、29.97fps で生成したイベント画像 (event) を実験に用い、認識率を比較した。表 1 にそれぞれの画像を用いた場合の平均認識率を示す。

表 1: 画像毎の平均認識率

画像種類	平均認識率
frame	0.900
event	0.894

表 1 より、フレーム画像とイベント画像の単語読唇における認識率はほぼ同じであることが分かる。イベント画像の場合は口の動きによってエッジの検出のしやすさが変わってくるため、イベントベースカメラを用いても必ずしも精度が向上するとは言えない。

## 4 まとめ

本研究ではイベントベースカメラを用いた単語読唇に取り組んだ。イベントより生成した画像で深層学習を行った認識実験の結果、従来型カメラより得られる画像を用いた認識実験と同程度の認識率が得られることを確認できた。今後は、early fusion や late fusion で二つの画像を組み合わせて学習を行い、精度を検証する。二つの画像を組み合わせることで精度の向上を目指すことができると考えられる。また、今後は他の認識手法との比較や、イベントベースカメラを用いた文章読唇にも取り組む。

## 参考文献

- [1] Tatsuya Kanamaru, Taiki Arakane, and Takeshi Saitoh. Isolated single sound lip-reading using a frame-based camera and event-based camera. *Frontiers in Artificial Intelligence*, Vol. 5, , 2022.
- [2] Brais Martinez, Pingchuan Ma, Stavros Petridis, and Maja Pantic. Lipreading using temporal convolutional networks. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6319–6323. IEEE, 2020.

		指導教員	齊藤 剛史 教授
学生番号	192C1041	氏名	甲斐 千裕
論文題目	マスク顔画像に対する単語読唇		

## 1 序論

2020 年以降、新型コロナウイルス感染症拡大の影響で人々は日常的にマスクを着用するようになった。マスク着用時も読唇が可能なのかという疑問の声が上がったため、本研究ではマスク着用時における読唇の可能性について検証する。

## 2 認識手法

読唇処理は、顔領域検出処理、顔特徴点検出処理、口唇領域抽出処理および認識処理から構成される。マスク顔画像用の顔領域および顔特徴点検出も検証する。

### 2.1 顔領域検出

本研究では、(1)dlib に実装されている HOG 特徴量による顔検出、(2)OpenCV に実装されている Haar-Like 特徴量による顔検出、(3)MobileNet による顔検出、(4)SSD → MobileNetV2 による顔検出の四つの顔検出法を検証した。(3)のみマスク顔画像用に学習されたモデルである。本研究で収集したマスク顔画像から 40 シーンを選出し、4 手法における顔検出の成否を目視で判定した。結果として(1)～(4)の顔検出成功 rate はそれぞれ 29.6%、30.6%、99.4%、70.8% であったため、(3)を採用した。

### 2.2 顔特徴点検出

dlib の顔特徴点検出器をもとに、(a) 目のみの 12 点、(b) 目および眉の 22 点、(c) 目、眉、頬の 27 点、(d) 顔特徴点全体の 68 点を検出する四つのモデルを検証した。(a), (b), (d) の学習データは SSSD[1] の非マスク顔画像から合計 306 枚を選定した。(c) は本研究で収集したマスク顔画像より合計 200 枚を選び、人手で特徴点を与えてモデルを学習した。目の特徴点位置の推移を目視で確認する定性的評価、目の  $y$  座標の標準偏差を計算する定量的評価を行った。結果、検出精度が最も高かった(c)を採用した。

27 点モデルで検出した目の位置を基準に正規化処理を適用して  $96 \times 96$  画素の口唇 ROI 画像とした。

### 2.3 認識処理

先行研究 [2] で非マスク顔画像に対して高認識精度を得た、3D-Conv+ResNet-18 により特徴を抽出して MS-TCN で認識するモデルを利用し、単語単位の読唇に取り組む。モデルの入力は  $64 \times 64$  画素の口唇 ROI である。

## 3 評価実験

### 3.1 実験条件

発話シーンの収集のため、図 1 のマスク 9 種類を 20 名に配布し、合計 5,400 シーンを用意した。比較用の非マスク顔画像は、SSSD の 20 名より合計 4,500 シーンを利用した。発話内容は、SSSD から挨拶 15 単語を選定した。深層学習モデルは読唇用公開データセット LRW で事前学習した後、Fine Tuning を適用した。

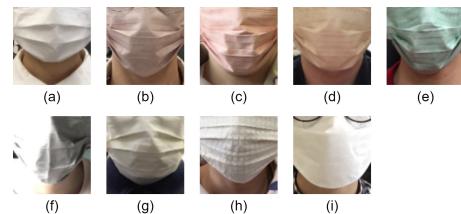


図 1: マスク画像

### 3.2 実験結果

マスク顔画像、非マスク顔画像、両画像を含めた 3 条件に対し 10 分割交差検証で評価した。学習データとテストデータが共にマスク顔画像の場合、平均認識率は 88.3% であった。これはマスク着用時でも単語読唇が可能であることを意味する。非マスク顔画像では 97.8%，混在する場合は 94.1% であった。

## 4 結論

本研究では、マスク顔画像に対する単語読唇に取り組んだ。認識実験の結果、マスク顔画像のみを学習データおよびテストデータとした場合 88.3% の精度が得られた。このことからマスク顔画像でも単語読唇が可能であることが分かった。今後の課題として、現状マスク着用時の発話シーンが少ないため、更なる発話シーンの収集が挙げられる。

## 参考文献

- [1] 齊藤剛史, 窪川美智子. SSSD : スマートデバイスを用いた読唇技術向け日本語データベース. 信学技報, 第 117 卷, pp. 163–168, 2018.
- [2] 荒金大清, 齊藤剛史. Temporal convolutional network を用いた単語読唇. 第 11 回バイオメトリクスと認識・認証シンポジウム (SBRA2021) , pp.29–33, 2021.

		指導教員	齊藤 剛史 教授
学生番号	192C1068	氏名	権藤 優季
論文題目	深層学習を用いた 神経難病患者の口形認識に関する研究		

## 1 はじめに

コミュニケーションによる意思疎通は私たちの日常生活において不可欠であるが、筋萎縮性側索硬化症(ALS)などの難病を患うとその手段が次第に制限される。そこで、発声や道具を必要としないコミュニケーション手段の一つとして口文字コミュニケーションがある。口文字コミュニケーションでは、口の形と瞬きなどの合図により文字を表現する。それらの文字を組み合わせて単語や文章を生成して意思を伝える。しかし、ALSの患者は病状が進行すると口形の形成が困難になり、介助者が正確に口形を読み取ることが難しくなると円滑に会話のためには練習が必要となる場合がある。そこで読唇技術と口文字コミュニケーションを組み合わせた技術に着目する。本研究では画像認識技術を用いて“あ”, “い”, “う”, “え”, “お”の5母音と“閉唇”的計6口形の認識を目的とし、深層学習モデルを用いて認識精度を検証する。

## 2 認識手法

発話シーンを収集し、それらのデータに対して正規化処理を施す。その後、Dlib を用いて顔特徴点を取り出す。これらの特徴点に基づき口元周辺領域を切り出した口唇 ROI (Region of Interest; ROI) を抽出する。最後に、抽出した口唇 ROI を入力として与え、深層学習モデルの ViT[1] を用いて口形を認識する。学習方法として、モデル全体を再学習する手法である Fine tuning と学習データの水増しを行う手法である Data Augmentation を適用する。

## 3 認識実験

先行研究 [2] よりデータを追加し、神経難病患者 31 名の発話シーンに対して認識実験を行う。ViT におけるハイパラメータ調整実験では、パッチサイズが  $4 \times 4$ , Transformer Encoder 数が 8 の条件において最も高い認識精度を得た。このハイパラメータの条件下において行った Fine tuning を適用する実験では、適用した場合に最も高い認識精度 76.96%を得た。全被験者の平均認識率から作成した混同行列を図 1 に示す。6 口形ごとの認識率では“閉唇”的認識率が最も高く、“い”的認識率が最も低い結果となった。また、類似した口形で誤認識が目立った。さらに、すべての被験者から二名の組み合わせを選び、混同行列の要素の相関

を全パターンで求めた。認識率が高い被験者の組み合わせにおいて相関が強く、低い被験者の組み合わせでは相関が弱いことが確認できた。

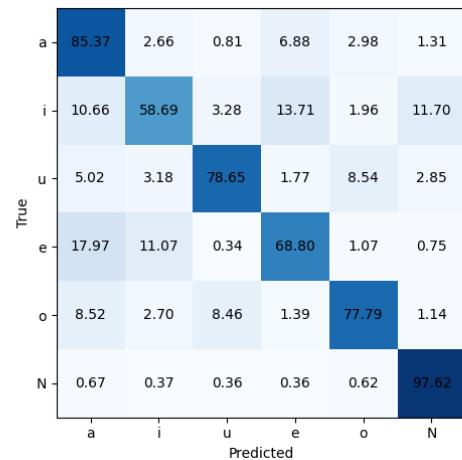


図 1: 平均認識率

## 4 おわりに

本論文では、神経難病患者を対象とした口形認識に関する研究を行った。そこで、先行研究 [2] よりも発話シーンを追加し、神経難病患者 31 名の発話シーンを用いて ViT における認識実験に取り組んだ。ハイパラメータの調整実験や Fine tuning の適用実験より、パッチサイズが  $4 \times 4$ , Transformer Encoder 数が 8 の条件において、Fine tuning を適用させた場合に最も高い認識率 76.96%を得た。今後も継続的に発話シーンを収集していくことで認識率の向上や不特定話者認識の可能性が期待でき、システムの実利用を十分に検討することができる。今後の展望として、時系列での情報を考慮した他の深層学習モデルを用いた口形認識や、口文字コミュニケーションを実現するシステムの開発を検討している。

## 参考文献

- [1] Alexey Dosovisliy et al. An image is worth 16x16 words: Transformers for image recognition at scale, *arXiv preprint arXiv:2010.11929*, 2020.
- [2] 中村祐哉, 神経難病患者のコミュニケーション支援のための深層学習を用いた口形認識に関する研究, 九州工業大学 情報工学部 システム創成情報工学科 修士論文. 2022.

		指導教員	齊藤 剛史 教授
学生番号	192C1096	氏名	玉井 龍斗
論文題目	眼球運動を利用した 眼鏡型スイッチの改良に関する研究		

## 1 序論

治療が極めて困難な指定難病の一つに筋萎縮性側索硬化症（ALS）がある。ALSは症状が進行するにつれて筋肉が動かせなくなっていくため運動障害、呼吸障害などの症状が現れる他、会話や手話、筆談を利用した他者とのコミュニケーションが困難になる。介護の現場では意思伝達の手段としてナースコールや意思伝達装置が用いられ、これらに使用スイッチとしてALSの末期症状でも比較的障害が軽微な眼球運動を利用したスイッチが開発された[1]。本研究では開発されたスイッチの実証実験でのALS患者や介護者からの改善要望をもとにシステムの改良を行う。

## 2 眼瞼運動によるスイッチ機能の評価実験

追加したスイッチ機能を評価するために、健常者10名を対象に被験者実験を行った。目を閉じるよう指示するアナウンスを用意し、ランダムなタイミングで12回流した。定量的に評価するために、真陽性( $TP$ )、偽陰性( $FN$ )、偽陽性( $FP$ )を計測した。また、再現率 $R = TP / (TP + FN)$ 、適合率 $P = TP / (TP + FP)$ 、F値 $F = 2PR / (P + R)$ を算出した。

実験の結果、全被験者の $TP$ ,  $FN$ ,  $FP$ はそれぞれ115, 5, 7であり、 $R$ ,  $P$ ,  $F$ を算出した結果はそれぞれ0.958, 0.943, 0.950であった。 $R$ は0.958であり、目を閉じてはいるが、閉じている時間がしきい値時間に達していないことがあった。 $P$ は0.943であり、意図していない普段の瞬きがしきい値時間に達してしまったことや、目を閉じている途中で一度開いていると判断されることがあった。

## 3 目画像解析の評価実験

目の一部がカメラの画角外にある時にカメラを移動するよう指示する機能の追加を試みる。目の一部が画像外にある画像を学習した検出器による目の輪郭点検出の評価実験をした。

### 3.1 データセット

先行研究[1]で利用されたデータセットD1, D2と、D1, D2の各目画像から $100 \times 60$ [pixels]を切り出し、 $120 \times 80$ [pixels]にリサイズして、目の一部が画像外にある画像を5枚ずつ作成して構築したデータセットD1a, D2aを用いる。

### 3.2 実験条件

CNNモデルとしてVGG16とXceptionを使用し、さらにVGG16の全結合層の後にDropoutを追加したモデル(VGG16-D)と畳み込み層の後にBatch Normalizationを追加したモデル(VGG16-BN), Xceptionの全結合層の後にDropoutを追加したモデル(Xception-D)の五つのモデルを使用した。輪郭点の検出には8点検出するモデルと同時に瞳孔中心点も検出するモデルを検討した。学習パターンとしてD1aのみ学習する場合(D1a), D1aとD2aをまとめて学習する場合(D1a+D2a), D2aで学習した後にD1aで再学習する場合(FT)の3種類を試みた。

### 3.3 実験結果

表1に輪郭点検出精度の上位3条件の結果を示す。輪郭点8点と瞳孔中心点を同時に検出し、XceptionにD1a+D2aを学習させた場合に最も精度が高く、平均誤差が7.62pixelであった。先行研究では最も精度が良いモデルは平均誤差が2.84pixelであったので比較すると本研究では高い精度が得られなかった。しかし、目画像外の輪郭点や左右方向が切れている目画像など特定の条件では高い精度が得られた。

表1: 輪郭点検出の上位5条件

CNNモデル	特徴点数	学習パターン	平均誤差 [pixel]
Xception	9	D1a+D2a	7.62
VGG16-D	8	D1a+D2a	7.78
Xception	9	FT	8.00

## 4 結論

本論文では改善要望をもとにシステムの改良をし、追加した機能の評価実験と機能追加のために目の一部が切れた目画像の輪郭点検出の精度評価をした。

今後は改良した眼鏡型スイッチをALS患者に利用してもらい、実際の介護現場での声をもとに、より使いやすいものになるよう改良を続ける。

## 参考文献

- [1] 坂本一樹, 齊藤剛史, 伊藤和幸. 瞳孔中心検出に基づく眼球運動を利用したALS感情の眼鏡型スイッチの開発. 信学技報, Vol. 121, No. 203, pp. 18 - 23, 2021.

		指導教員	齊藤 剛史 教授
学生番号	192C1135	氏名	古谷 優樹
論文題目	手術室スタッフの頭部姿勢推定に関する研究		

## 1 はじめに

安全で質の高いサービスを要求される医療現場では、チーム内スタッフ間で迅速かつ適切なコミュニケーションが必要である。この時、言語によるコミュニケーションのみならず、視線情報などの非言語行為によるコミュニケーションも利用される。本研究では、映像情報を用いて手術時における非言語行為を解析すること目的とし、手術室スタッフの頭部姿勢の推定に取り組む。頭部姿勢推定は数多くの手法が提案されているが、メディカルキャップやマスクを装着した人物を対象とした研究はなされていない。そのため、既存の手法では手術室スタッフの頭部姿勢が正しく推定できない。本研究では、この問題を解決するために、独自に頭部姿勢推定用のデータベースを効率よく構築する。頭部姿勢推定は監視カメラなどの広角カメラで撮影された動画像に対して行うことを考慮し、撮影には通常のビデオカメラに加え、全方位撮影が可能な 360 度カメラも用いる。

また、画像への畳み込み演算より特徴を抽出する Convolutional Neural Network (CNN) を用いた機械学習による頭部姿勢推定の精度評価と、実際の手術室で撮影されたシーンの頭部姿勢推定に取り組む。

## 2 データセットの構築

複数台のカメラで撮影を行う場合、カメラごとに撮影開始時刻は異なるため、すべての映像に対してライトを用いた同期処理を行う。また、データセットには被写体が設定した角度を向いた状態の画像を用いるため、ターンテーブルが停止しているフレームをフレーム間差分を用いて検出する。頭部領域の抽出にはOpenPose で得られる目、耳、鼻の座標を用いる。特定の向きに対して Yaw 角の値が一意に定まらないため、Pitch 角と Yaw 角を 3 次元極座標  $(r, \theta, \varphi)$  から 3 次元直交座標  $(x, y, z)$  へ変換し、座標を正解データとしてアノテーションする。

## 3 頭部姿勢推定

頭部姿勢推定には CNN モデルを用いる。CNN モデルは畳み込み演算を行う畳み込み層、プーリング層、全結合層で構成されたニューラルネットワークである。CNN は姿勢推定分野で高い性能を示し、現在でも多く

利用されている。本研究では CNN モデルの VGG16, MobileNet V3[1] を用いる。

## 4 実験

本研究では 10 名の頭部画像 19,200 枚からなる頭部姿勢推定用のデータセットを構築した。構築したデータセットの例を図 1 に示す。



図 1: 構築したデータセットの例

評価方法として 1 人抜き法を用いて実験を行い、誤差の平均を示す指標である Mean Absolute Error (MAE) を用いて精度の評価を行った。VGG16 では層の深さが十分でなく、正しく学習が行えなかった。一方、MobileNet V3 では 1 人抜き法を用いて実験を行った結果 MAE は  $\pm 5.67^\circ$  となりメディカルキャップを装着した頭部に対しても姿勢推定が可能であることが分かった。また、実際の手術室スタッフの頭部姿勢推定を行い、後ろ向きの頭部画像に対して構築したデータセットを学習した MobileNet V3 での推定は既存のデータセットと手法での推定精度を上回ることが定性的に確認できた。

## 5 結論

フレーム間差分での静止フレームの抽出とライトの点滅による映像の同期を行い、効率のよいデータセットの構築に取り組んだ。また、マスクとメディカルキャップを装着した人物の頭部画像を学習することで、頭の向きにかかわらず手術室スタッフの頭部姿勢推定が行えた。今後は最新手法やより層の深いモデルで学習し、頭部姿勢推定の精度評価を行う。また、データ数を増やし手術室スタッフの頭部姿勢推定の精度向上を目指す。

## 参考文献

- [1] Andrew Howard et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF international conference on computer vision, pp. 1314 – 1324, 2019.

		指導教員	齊藤 剛史 教授
学生番号	192C1162	氏名	山口 剛
論文題目	国会審議映像に対する発言者の頭部姿勢推定		

## 1 序論

日本の政治制度は議会制民主主義の体制をとっており、国会は必要不可欠なものとなっている。一方、政治家や閣僚は発言の際に原稿を読んでいるのか、自身の言葉を伝えているのかの判断がなされておらず、原稿を棒読みしていると指摘されることがある。本研究では、国会における発言者の頭部姿勢を推定することで、発言者が原稿を読んでいるか否かを判定することに取り組む。

## 2 頭部姿勢推定

### 2.1 国会審議映像

国会審議映像には、図 1 のように発言者が顔を下に向けて発言しているシーンと顔を上げて発言しているシーンが含まれている。本研究では手動で発話シーンの切り出しを行い、実験に用いる。



(a) 下を向いているシーン (b) 顔を上げているシーン

図 1: 国会審議映像の例

### 2.2 提案手法

顔検出および顔特徴点検出には、dlib[2] と Face Mesh[3] を用いる。また、複数人の人物が映っている映像で発言者のみを頭部姿勢推定対象とするため、目視で発言者の位置を確認し、手動で発言者を切り出す。

本研究では 3 種類の提案手法を用いる。顔検出の成否を利用する手法では、dlib より提供されている顔検出モデルを用いて発言者の顔検出を行い、顔検出が失敗したフレームでは発言者が下を向いていると推定する。

頭部姿勢角度を利用する手法では、Face Mesh を利用して取得した発言者の顔特徴点の座標と、カメラ内部パラメータ行列から頭部姿勢を示すオイラー角を求め、しきい値を用いて頭部姿勢推定を行う。

CNN を用いる手法では、ImageNet にて事前学習を行った ResNet-50[4] を利用し、発話シーンから発言者の顔領域のみを切り出したフレーム画像を入力し頭部姿勢推定を行う。

## 3 実験

本研究では、国会審議映像検索システム [1] より 2010 年 1 月 18 日から 2021 年 3 月 5 日までの 5,900 本の国会審議映像を取得した。取得した映像から代表的な発話シーンを 20 種類切り出し、頭部姿勢についてアンケート調査を行い実験に用いた。3 種類の実験手法による推定精度を表 1 に示す。表 1 より、ResNet-50 を用いる場合が最も精度が高いため、テストデータに対する頭部姿勢推定の定性的評価には ResNet-50 を用いる。

また、頭部姿勢推定精度比較に用いたものとは異なる発話シーンをテストデータとし、ResNet-50 を用いて頭部姿勢推定を行い、定性的に評価する。

表 1: 実験結果

Method of face detection	Accuracy
dlib	0.716
Face Mesh and Euler angles	0.843
ResNet-50	0.910

## 4 結論

本研究では、国会審議映像における発話シーンに対し、顔検出の成否を利用する手法、頭部姿勢角度を利用する手法、CNN を利用する手法での発言者の頭部姿勢推定を行ない、CNN を利用する手法を用いてテストデータとした発話シーンの頭部姿勢推定を行った。今後は頭部姿勢だけでなく、発言者の身体全体の動きや音声を用いた状況解析を行う。

## 参考文献

- [1] 国会審議映像検索システム. <https://gclip1.grips.ac.jp/video/meeting>.
- [2] Davis E King. Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, Vol. 10, pp. 1755–1758, 2009.
- [3] Yury Kartynnik, Artsiom Ablavatski, Ivan Grishchenko, and Matthias Grundmann. Real-time facial surface geometry from monocular video on mobile gpus. *arXiv preprint arXiv:1907.06724*, 2019.
- [4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning. *Image Recognition*, Vol. 7, 2015.