

令和7年度 電子情報工学科

卒業研究中間発表会

講演予稿集

令和7年11月7日（金）8:50~

専攻科棟1F 視聴覚室(S-102)

令和7年度 電子情報工学科 卒業研究中間発表会

講演予稿集 目次

発表時間：1名発表（発表7分+質疑応答3分）、2名発表（発表11分+質疑応答4分）

出欠確認・注意事項	8:40	築地教員	
第1セッション	8:50~10:10	座長：築地教員	
1. 視線推定によるマウス代替インターフェースの構築			8:50
	山本 伊吹	(渡邊研究室)	
2. 音響信号による死角領域の物体検出			9:01
	根岸 慧伍	(渡邊研究室)	
3. 単眼深度推定を用いた2D顔認証における精度向上の評価と考察			9:12
	真庭 正宗	(渡邊研究室)	
4. LLMを用いたVR操作手法の有効性に関する研究			9:23
	豊巻 煌	(渡邊研究室)	
5. ラスベリーパイを用いたタンブラー自動洗浄機の試作			9:34
	浅川 武揚	(大豆生田研究室)	
6. 高効率な入力を目的としたキー配列の検討			9:45
	田部井 珂維人	(大豆生田研究室)	
7. 平衡二分探索木を用いたインデックス無制限の遅延評価セグメント木の試作			9:56
	藤田 恭輔	(大豆生田研究室)	
第2セッション	10:30~11:50	座長：市村教員	
8. マーカーレスモーションキャプチャを用いた、フォーム変化予測モデルの開発			10:30
	土谷 琢磨	(荒川研究室)	
9. 3D姿勢推定を応用したアニメーションからのボーン抽出AIの開発			10:41
	武田 幸寛	(荒川研究室)	
10. ミキシング支援AIの開発			10:52
	新井 春登	(荒川研究室)	
11. 言語処理を用いた人狼ゲームにおける形勢評価			11:03
	金井 佑心	(荒川研究室)	
12. 衛星画像情報と出没実績を融合したハイブリッド特徴量によるクマ出没予測モデルの開発			11:14
	田村 伊織	(西山研究室)	
13. 家庭支出のショック伝搬を示すABM			11:25
	飯尾 龍矢	(西山研究室)	
14. 学校内の落とし物アプリ			11:36
	メイマーカイン, 飯島 慶士	(西山研究室)	

第3セッション	12:50~14:25	座長：川本教員	
15. マイクアレイとスマートウォッチ連携によるリアルタイム音源定位システムの開発			12:50
	西場 真翔 (大墳研究室)		
16. パーキンソン病患者の小声症を支援するリアルタイム音声フィードバックシステムの提案			13:01
	遠山 蓮 (大墳研究室)		
17. 足の甲を使用する体表点字システムの検討			13:12
	鈴木 煌大 (大墳研究室)		
18. 大腸検査ロボットのシミュレーション環境の構築			13:23
	笹川 竜誠 (市村研究室)		
19. エッジAI デバイスを用いるオフライン対話型ロボットの試作			13:34
	中澤 啓人 (市村研究室)		
20. 群馬県全域におけるからっ風の風力へ山岳地形がもたらす影響に関する検討			13:45
	後藤 天慈, 齋藤 優月 (雑賀研究室)		
21. 群馬県全域におけるからっ風の風向へ山岳地形がもたらす影響に関する検討			14:01
	阿久澤 聖蒼 (雑賀研究室)		
22. ホップフィールドモデルにおける連想記憶の性能向上に関する検討			14:12
	渡邊 祥 (雑賀研究室)		

第4セッション	14:40~16:15	座長：荒川教員	
23. ダイレクトサンプリング方式 SDR 受信機におけるアンチエイリアス BPF に関する検討			14:40
	秋山 珀空 (築地研究室)		
24. 電流モード制御 DC-DC コンバータにおける電流帰還回路モデルに関する検討			14:51
	樋口 晃太 (築地研究室)		
25. 0.6umCMOS オペアンプ回路の設計			15:02
	福島 彩斗 (築地研究室)		
26. 音声対話型デジタルサイネージシステムの開発			15:13
	黛 琵琶 (川本研究室)		
27. 非流暢発話検出に基づく発表練習支援システムの開発			15:24
	山本 康生 (川本研究室)		
28. アドリブに対応するライブパフォーマンス支援システムの開発			15:35
	余 夏輝 (川本研究室)		
29. 集合の種類と濃度の階層			15:46
	江畑 太一 (北田研究室)		
30. OpenFOAM を用いた空調効率の研究			15:57
	佐相 冨太郎, 中瀬 亮 (宇治野・高橋研究室)		

講評 **大墳学科長** 16:15

予鈴 1名発表 発表終了1分前(6分 1回) 発表終了(7分 2回) 質疑終了(10分 3回)
 2名発表 発表終了1分前(10分 1回) 発表終了(11分 2回) 質疑終了(15分 3回)

1：視線推定によるマウス代替インターフェースの構築

渡邊研究室 山本伊吹

1 はじめに

現代社会において、PC やスマートフォンをはじめとするデジタルデバイスは、業務、教育、コミュニケーション、エンターテインメントなど、あらゆる場面で不可欠な基盤となっている。これらのデバイスを効率的に利用するためには、人間がコンピュータに意図を伝えるための入力インタフェースが極めて重要な役割を担う。中でも「視線」は、人間の注意や意図を最も直接的かつ高速に反映する情報である。人が何かを見るとき、そこには「情報を得たい」「操作したい」といった意図が内在している。この視線情報を入力として利用できれば、マウスカーソルを「動かそう」と意識するよりも先に、見ている場所にカーソルを移動させることが可能となり、究極のハンズフリー・インタフェースを実現できる可能性を秘めている。

2 研究目的

本研究の目的は、カメラを用いて顔の画像から視線推定を行い、画面上のどこを見ているかを検出し、マウスの代わりとなるインターフェースを構築することである。また、視線入力課題である意図しない誤選択などを減らすシステムの開発を目指す。

視線推定を行う方法として、MediaPipe Face Landmark と OpenCV を使用し、マウス操作は PyAutoGUI を使いシステムの構築を行う。

3 関連研究

関連研究として、角膜に照射された光源の反射像を用いて視線を推定する角膜反射法 [1] やGoogle にビデオカメラを固定して装着し頭部を固定した状態で、虹彩中心を検出し5点のキャリブレーションを行って視線推定を行う『画像処理を用いた虹彩中心検出による視線入力装置の開発に関する研究』[2] がある。近年では、機械学習を用いた手法が提案されている。これには、モデルベース (Model-based) とアピランスペース (Appearance-based) の大きく2つに分けることができ、3D 眼球モデルを適用し推定する手法 [3] や直接画像から視線を推定する手法 [4] が提案されている。

4 関連技術

4.1 MediaPipe Face Landmark[5]

MediaPipe Face Landmark とは、Google 社が提供するオープンソースの機械学習ライブラリ MediaPipe の1つである。画像や動画、リアルタイム映像から人間の顔の特徴点を468点と左右の瞳の特徴点を5点ずつ検出ができる。図1に顔全体のランドマークと図2に左目のランドマークを示す。左目のランドマークには、目の輪郭と虹彩の輪郭、虹彩の中心の特徴点が含まれている。

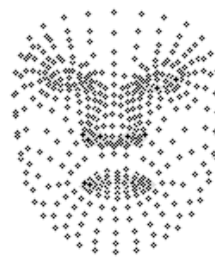


図1 顔全体のランドマーク (特徴点)

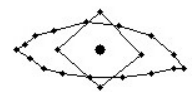


図2 左目のランドマーク

4.2 PyAutoGUI[6]

PyAutoGUI は PC の操作を自動化するためのライブラリであり、マウスやキーボードの操作も行うことができる。

5 提案手法

本研究では、現在左右の虹彩中心のランドマークと目周辺ランドマークを使用している。必要に応じて、使用するランドマークを追加する予定である。

5.1 視線推定

視線推定は以下の手順で行っている。

- 目のランドマーク検出
目尻、目頭、上下まぶたの頂点、虹彩中心を使用している。
- 相対的な視線特徴量の計算
検出された目のランドマークから、目内部での虹彩中心の相対的な位置を計算。

- キャリブレーション

ユーザに既知の点を注視してもらい、相対的な視線特徴量と画面上のターゲット座標の対応関係を学習させる。

- マッピング

学習した対応関係から、新しい相対的な視線特徴量を与えられたときに、画面上のどこを見ているかをRBF(Radial Basis Function)を使って推定する。

5.2 マウスとしての操作

カーソル操作は視線推定を用い、マウスとしての操作はPyAutoGUIを使用する。

5.3 GUI操作

PyAutoGUIと顔を使ったジェスチャを使用し、マウスと同等の操作を対応させる。可能であれば、キーボード入力も実装する予定である。

6 進捗

現在、視線推定を行うシステムと瞬き検出まで進んでいる。図3に示すように、カメラをディスプレイの下に置き、ディスプレイと被験者との距離を60cmとし、できるだけ動かないようにしてもらい、図4に示す測定点について視線推定の実験を行った。解像度1920×1080ピクセルで30fpsのカラーカメラを使用し、実験を行った。

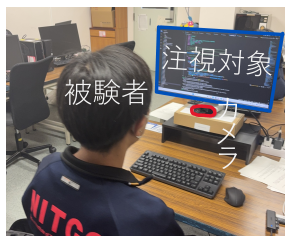


図3 実験環境

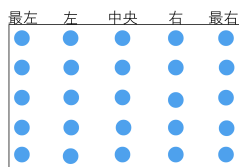


図4 視線推定の測定点

頭を固定した状態で画面上に16点の視線推定を行ったときの精度を表1に示す。

表1 視線推定の精度実験結果

	最左	左	中央	右	最右
x 誤差 [px]	65.125	68.000	74.833	67.000	37.125
x 誤差 [mm]	17.870	18.659	20.533	18.384	10.187
y 誤差 [px]	113.08	138.29	148.29	160.83	124.41
y 誤差 [mm]	31.030	37.947	40.690	44.132	34.139

表1より、画面の縦方向の精度が横方向の精度よりも悪いことがわかる。また、画面中央の列の精度が縦横両方ともに他の列よりも悪いことがわかる。

図5に精度調査に参加した人ごとの精度の差を示す。

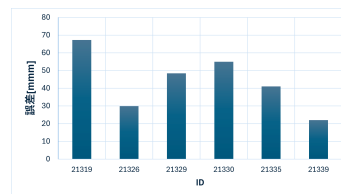


図5 視線推定の個人差

図5より、視線推定の精度は個人によって精度が大きく異なることがわかる。これは、頭の動きを考慮できていないため、目の位置が測定中に移動してしまったことが原因と考えられる。

7 今後の予定

現在、頭の向きや目の高さを一定にしないと視線推定を正確に行えないこととマウスとして使えるほどの視線推定の精度が得られていないことが課題である。今後、3D眼球モデルの考慮などマウスとして使える程度の精度の向上を目指す。また、クリックなどの操作を誤認識の少ないジェスチャの選定と実装を行う。最終的には、作成したシステムの精度調査と使用感についての評価を行う。

参考文献

- [1] Guy Thomas Buswell: "How people look at pictures: a study of the psychology and perception in art", 1935, 2025/10/30
- [2] 米沢徹也: "画像処理を用いた虹彩中心検出による視線入力装置の開発に関する研究", 熊本大学大学院学位論文 (2010), https://www.kumamoto-u.ac.jp/daigakujouhou/jouhoukoukai/gakuironbun/kougaku_kou/kougaku_kou_file/kougaku_kou459ronbun.pdf, 2025/10/30
- [3] 中松 由香梨 他: "3次元 Active Appearance Model を利用した視線方向推定", 画像の認識・理解シンポジウム (MIRU2011), <https://ipsj.ixsq.nii.ac.jp/record/77722/files/IPSJ-MIRU2011082.pdf>, 2025/10/30
- [4] Kyle Krafka 他: "Eye Tracking for Everyone", CVPR 2016, https://people.csail.mit.edu/khosla/papers/cvpr2016_Khosla.pdf, 2025/10/30
- [5] MediaPipe ソリューションガイド: Google, <https://ai.google.dev/edge/mediapipe/solutions/guide?hl=ja>, 2025/10/30
- [6] Welcome to PyAutoGUI's documentation!: Al Sweigart, <https://pyautogui.readthedocs.io/en/latest/>, 2025/10/30

2：音響信号による死角領域の物体検出

渡邊研究室 根岸 慧伍

1 はじめに

内閣府が作成している交通安全白書 [1] によると、過去 10 年間、毎年出会い頭事故が全体の 24~25% であり、追突事故に次いで 2 番目に高い割合を占めている。出会い頭事故が発生する原因として、交通ルールの無視、安全確認の不十分、死角からの飛び出しなどが挙げられる。その中で、安全確認の不十分や死角からの飛び出しによる事故はシステムによってその発生確率を低減することが可能であると考えた。

そこで、見通しの悪い交差点などで発生する死角領域の物体を検出するシステムを開発することで安全確認の支援を行い、交通事故の発生件数の減少が期待される。

2 研究目的

本研究では、音響信号によって死角領域の物体を検出することを目的としている。壁や物体に反射した音響信号を解析することで、センサが直接視認できない位置にある物体を検出し、死角による認識不足を補う。関連研究では、光を使った研究もあるが装置が非常に高価であり設置環境にも制約が多い。そのため、本研究では低コストかつ汎用的に利用できる音響信号を活用し、見通しの悪い交差点などにおける安全確認の支援を目指す。

3 関連研究

関連研究として、Faccio,D,Velten,A, Wetzstein,G らによる『Non-line-of-sight imaging』[2]がある。この研究では、SPAD (Single-Photon Avalanche Diode) などをを用いることで、死角内の物体形状を 3 次元的に復元することに成功している。しかし、使用している装置が非常に高価であるため一般的な環境での運用が困難である。そこで、本研究では低コストな構成として、スピーカとマイクを用いた音響信号によって死角領域の物体検出を行う。音響信号を用いた研究には、D.B.Lindell,G.Wetzstein, V.Koltun らによる『Acoustic Non-Line-of-Sight Imaging』[3]がある。この研究では、音響信号によって死角領域の物体を検出・再構成することに成功している。しかし、この研究ではマイクとスピーカを 16 個使い、位置を移動させながら計測を行っているため、情報取得には数分を要している。そこで、本研究ではマイクとスピーカの数进行限定し、あらかじめ正形状に配置することで計測の高速化を図る。

4 提案手法

Raspberry Pi に接続したスピーカとマイクを使用することで、死角領域の物体検出を行う。図 1 に構成図を示す。

スピーカとマイクをそれぞれ 3×3 の正形状に配置することで情報取得時間の短縮を図

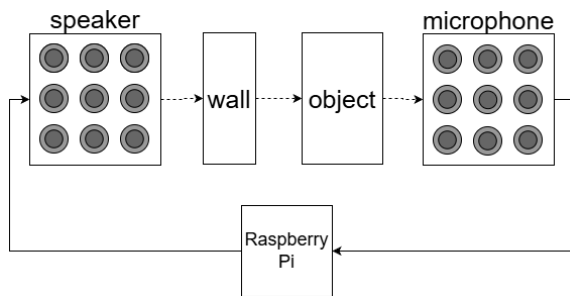


図 1: 構成図

る。また、超音波を用いることで、環境音などの雑音の影響を受けにくく、安定した物体検出を可能にする。図 2 にスピーカとマイクを用いた死角物体検出の模式図を示す。

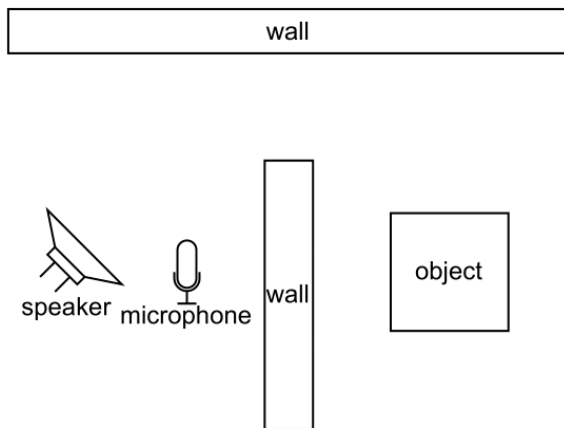


図 2: 死角物体検出の模式図

まず、スピーカから音響信号を発生させ、対象空間に送信する。壁や障害物に反射して戻ってきた信号を、マイクで受信する。受信した信号に対して短時間フーリエ変換 (STFT) や離散フーリエ変換 (DFT) を適用し、反射波の到来時間や周波数成分を解析することで、センサの視認できない死角領域に存在する物体の位置を推定する。

5 進捗

現在、スピーカを実装し動作確認まで終了している。

今後は、マイクの実装、マイクとスピーカの複数同時制御を行い、死角領域の物体検出の実験を進める予定である。

参考文献

- [1] Cabinet Office, Government of Japan 交通安全白書
 <<https://www8.cao.go.jp/koutu/taisaku/index-t.html>>(2025/10/29 参照)
- [2] Faccio, D., Velten, A. & Wetzstein, G., “Non-line-of-sight imaging”, Nature Reviews Physics, vol.2,no.6, pp. 318–327,2020.doi:10. 1038/s42254-020-0174-8.
- [3] D. B. Lindell, G. Wetzstein and V. Koltun, “Acoustic Non-Line-Of-Sight Imaging,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach,CA,USA,2019,pp.6773-6782,doi: 10.1109/CVPR.2019.00694.

3: 単眼深度推定を用いた 2D 顔認証の精度向上の評価と考察

渡辺研究室 真庭正宗

1 はじめに

近年、顔認証技術は施設の入退室管理やスマートフォンのロック解除などさまざまな場面で利用されている。顔認証技術はその方式により 2D または 3D 顔認証に大別されるが、2D 顔認証は専用機器を必要とせずに単眼カメラ一つだけで導入できるという点から広く普及している。

しかし、2D 顔認証技術は顔の平面的な情報のみを参照しているため、顔の向きが変化すると認証精度が低下し、誤認識が発生しやすくなるという課題が指摘されている。一般的に Yaw 角が ± 45 度を超えると認識精度が下がり始め、 ± 60 度からは急激に認識精度が低下することが、論文で明らかになっている。[1] (正面画像から 45 度変わると約 75% に低下)

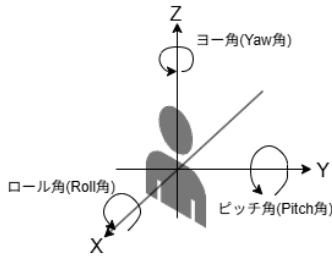


図 1: 3次元における角度の名称

本研究では、単眼深度推定モデルを用いて顔の立体的な構造を捉えることで、より角度変化に強い 2D 顔認証システムの実現を目指す。従来の 2D 顔認証システムに後付けする形で角度変化への耐性をつけるため、専用の計測機器などを導入する費用を抑えつつ、より自然な角度での認証を実現する。

2 研究目的

本研究の目的は、従来の 2D 顔認証システムより角度変化に堅牢なシステムを作成することである。

単眼カメラ単体からの情報により顔の立体的な情報を取得することで、2D 顔認証の弱点を補うことができると考えた。

単眼深度推定モデルは一枚の画像から深度マップを作成することができるため、鼻や目などの立体的な位置関係から顔の向きを推定し、それを認証に取り入れること

で、認証精度の向上が期待できると考えた。

そこで DepthAnything V2[2][3] などの単眼深度推定モデルを活用し、既存の単眼カメラ環境でも適用可能な「立体情報を用いた顔認証手法」を実現することで、姿勢変化への耐性と実用性の両立を図る。

3 関連研究

中島寛らによる研究論文「2次元画像と3次元立体形状を用いたハイブリッド顔認証技術」[4]では、2D データのみの場合、顔の向きが変化したときに認証性能が低下する問題に対して、受動型ステレオビジョンを用いた 3D 顔計測技術と 2D 顔認証を組み合わせることで、計算時間と認証性能の両立を実現し、実用性の高い手法として評価されている。

本研究では、受動型ステレオビジョンを用いず、単眼カメラ単体からの情報により顔の立体的な情報を取得することで、2D 顔認証の弱点を補うことができると考えた。

4 実験環境

表 1 に実験環境を示す。

表 1: 実験環境

器具名	規格	製造会社
CPU	Intel Core i7-8700 3.20GHz	Intel
GPU	NVIDIA GeForce GTX 1660 Ti	NVIDIA
RAM	16.0GB	
Web カメラ	C270 HDWeb カメラ, 720p 30fps	logicool
Python	3.10.11	
単眼深度推定モデル	DepthAnything V2	

顔検出や姿勢推定には MediaPipe、既存の顔認証システムの実装には、face_recognition を用いる。

5 提案手法

Web カメラで利用者の顔検出を行い、顔領域のみのフレームを切り取る。その後切り取ったフレームに対して深度推定および顔の特徴量の抽出を行い、すでに保存されている顔データと比較し、一定の閾値を超える類似度が確認された場合に本人として認証を行う。その際、顔の角度により 2D 顔認証単体の場合と深度マップを組み合わせた場合で認識率や誤認識率にどのような差がでるか検証と考察を行う。

6 進捗

DepthAnything V2 によるリアルタイムの深度推定の実装を行った。図 2 にもとの顔領域のフレームと、深度推定を行ったフレームを示す。

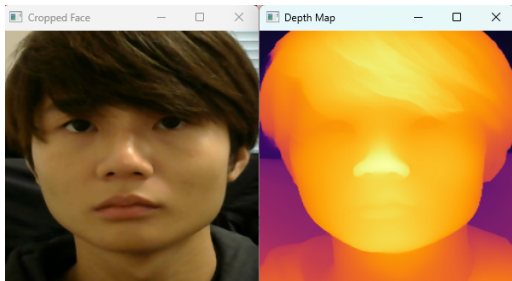


図 2: DepthAnything V2 による深度推定

鼻の形や背景、輪郭や口などの奥行きを捉えられることがわかった。実行において、特に深度推定における遅延はなく、平均で 46.23FPS を観測できたため十分にリアルタイム性があることが確認できた。

2D 顔認証単体の場合の各角度に対する認証精度に対しての実験を行った。実験の手順としては、被験者の顔画像を登録し、Yaw 角-90 度から-60 度の顔の角度に対して 30 回顔検出を行い、その際の顔認証成功率を記録した。これを各角度 (-60~-40 度, -40~-10 度, -10~10 度, 10~40 度, 40~60 度, 60~90 度) について行った。現段階では、発表者を含め 5 人に実験を行った。表 2, 3 に実験結果を示す。

表 2: 2D 顔認証の Yaw 角の変化に対する認証率

Yaw 角 [deg]	成功 [回]	失敗 [回]	認証率 [%]
-90~-60	50	100	33.3
-60~-40	141	9	94
-40~-10	149	1	99.3
-10~+10	150	0	100
+10~+40	149	1	99.3
+40~+60	122	28	81.3
+60~+90	71	79	47.3

表 3: Yaw 角の変化に対する平均ユークリッド距離

Yaw 角 [deg]	平均ユークリッド距離
-90~-60	0.47928
-60~-40	0.40262
-40~-10	0.30296
-10~+10	0.18952
+10~+40	0.29516
+40~+60	0.42422
+60~+90	0.49172

表より、face_recognition を用いた顔認証において、論

文で確認した通り Yaw 角の絶対値が大きくなるにつれ認証率が低下することが判明した。Web カメラに対してほとんど正面顔である-40~+40 度の角度において認証率は 99.3% を超えており、逆に角度がつく (-90~-60, +60~+90) とその認証率は 50% を下回り顔認証として信頼できない程度の認証率になってしまうことが理解できた。

face_recognition で設定されている顔の特徴量が同一人物のものであるかを判定するユークリッド距離の閾値 [5] は 0.6 であり、表 3 より各角度に対する平均ユークリッド距離を参照すると正面顔に近いほどその値は小さく、データベースに保存されている顔画像の特徴量と Web カメラで捉えている顔の特徴量が非常に類似していることがわかる。

また、-90~-40 度, 40~90 度において平均ユークリッド距離が 0.6 を下回っている理由として、被験者を違う人物として認識していることが原因であると考えられる。

7 今後の展望

DepthAnything V2 によるリアルタイムの深度推定から、顔の深度情報を読み取りその情報を認証に使うことで、現在課題となっている Yaw 角の絶対値が大きい場合の誤認証率を減少させることを目指す。

参考文献

- [1] ResearchGate: 「Face Recognition Performance in Facing Pose Variation」, https://www.researchgate.net/publication/323480551_Face_Recognition_Performance_in_Facing_Pose_Variation, (閲覧年月日 2025-10-25)
- [2] Lihe Yang et al.: 「Depth Anything V2」, <https://github.com/DepthAnything/Depth-Anything-V2?tab=readme-ov-file>, (閲覧年月日 2025-11-02)
- [3] kdl-di: 「Depth Anything V2 で zero-shot 単眼深度推定」, <https://kdl-di.hatenablog.com/entry/2024/09/20/103047>, (閲覧年月日 2025-11-02)
- [4] 中島 寛ら: 「2 次元画像と 3 次元立体形状を用いたハイブリッド顔認証技術」, https://www.azbil.com/jp/corporate/pr/library/review/pdf/2012_01_3.pdf, (閲覧年月日 2025-10-25)
- [5] Adam Geitgey: 「face_recognition package」, https://face-recognition.readthedocs.io/en/latest/face_recognition.html, (閲覧年月日 2025-10-30)

4：LLM を用いた VR 操作手法の有効性に関する研究

渡邊研究室 豊巻 煌

1 はじめに

近年,VR や AR をはじめとする空間コンピューティング技術は,訓練,設計,医療,エンターテインメントなど多岐にわたる分野での活用が急速に進んでいる.この空間内での没入感や作業効率を高める上で,オブジェクトを直感的かつ効率的に操作するインターフェースの確立は重要である.

現在主流の操作手法は物理的なコントローラである.コントローラは直接的な操作感を提供する一方で,操作が複雑化しやすい.また AR/MR グラスにおいては,ハンドジェスチャーによる操作が注目されているが,長時間の使用による身体的疲労や操作の精度といった点が課題として残る.

大規模言語モデル(LLM)を利用した操作手法は,キーワードのマッチングに留まらず,発話の背後にあるユーザーの包括的な意図を理解し,抽象的,複合的,文脈依存の指示であっても,適切な一連のコマンドへと展開することが可能となる.本研究では,LLM を VR 空間における操作インターフェースとして実装し,その有効性を従来手法と比較検証することを目的とする.

2 研究目的

本研究の目的は,VR 空間における操作インターフェースとして,LLM に基づく自然言語解釈手法の有効性を,従来のコントローラおよび音声での入力を行う音声コマンドと比較し,定量・定性的に評価することである.

特に,単純な操作から複数の手順を組み合わせた複雑な操作まで,タスクの難易度に応じて各手法の優位性がどのように変化するかを明らかにする.評価指標として,タスクの遂行効率や正確性といった客観的指標に加え,システムユーザビリティスケール(SUS),raw TLX による主観的な使いやすさや操作主体感(SoA)を測定する.これにより,VR インターフェースとしての LLM の有効性と課題を考察する.

3 関連研究

関連研究として,自然言語の指示を具体的な空間操作に変換する Wenlong Huang らによる VoxPoser[1]がある.この研究は自然言語からロボットの制御を行うものであり,LLM の持つ高い言語理解能力を示している.

また,医療 VR シミュレーションにおいて,自然言語処理(NLP)ベースの音声アシスタントを用いた Mohit Nayak らによる研究[2]がある.この研究では,NLP ベースの音声アシスタントを用いた場合,従来のボタン操作と比較してユーザビリティと認知的負荷の面で優位性を持つことが示唆されている.

他に,LLM を VR の物体操作インターフェースに組み込んだ Xiangzhi Eric Wan らによる VR Mover[3]がある.この研究では,操作をオブジェクトの生成と移動に絞り,オブジェクトの情報,プレイヤーの視覚,ジェスチャーを組み合わせることで,自然言語によって VR 空間内のオブジェクトを操作できることが示されている.

本研究では,VR 空間における操作の難易度に応じて,各手法の優位性がどのように変化するかを明らかにする

4 提案手法

本研究では,3 種類の操作インターフェースを実装し,同一の VR 環境下で参加者による評価実験を行う.

本システムの構成図を図 1 に示す.

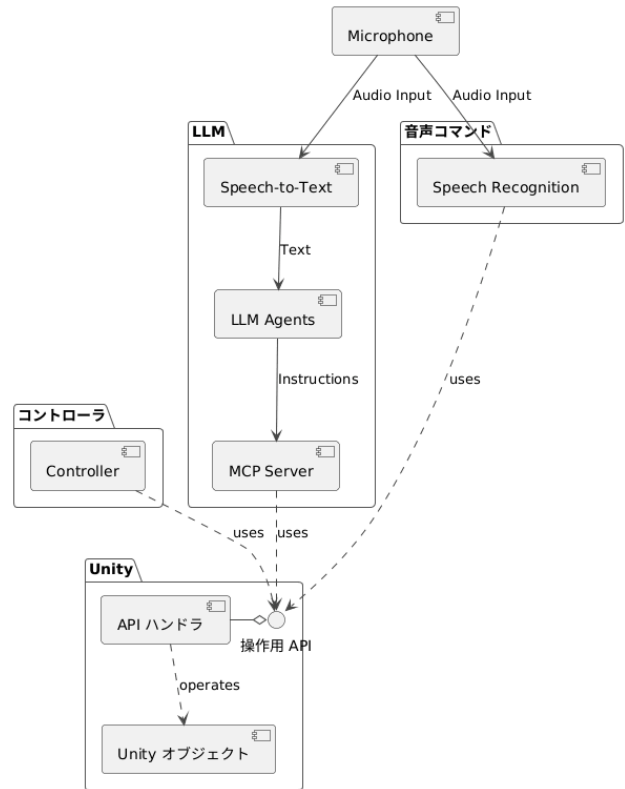


図 1 システム構成図

4.1 実装するインターフェース

コントローラ VR コントローラのポインティングによるオブジェクト選択と,ボタン操作によるプロパティ変更や移動を行う.

音声コマンド 事前に定義したコマンドリストに基づき,音声認識結果と完全一致した場合に操作を実行する.

LLM ユーザーの発話を文字起こしし,LLM Agent に入力する.また,Model Context Protocol を用いて LLM Agent へ VR 空間での操作を提供することで,LLM による操作の実行を行う.

4.2 評価タスク

タスクの複雑性に応じて各インターフェースの特性を評価するため,以下の課題群を設定する.

単純タスク群

1. 単純選択: 目の前にある特定のオブジェクトを選択する.

2. 単純なプロパティ変更: 選択中のオブジェクトの色やサイズを変更する.
3. 2次元 UI の操作: 空間に表示されたボタンの押下・テキストの入力を行う.

複雑タスク群

1. 連続操作: 複数の操作を連続して行う.
2. 属性・関係性の理解: 他のオブジェクトとの比較や関係性に基づいた操作を行う.
3. 一括操作: 条件に合致する全オブジェクトを同時に操作する.

4.3 評価指標

定量的指標としてタスク完了時間・操作実行回数 / 操作試行回数・エラー率・学習効率(試行回数による完了時間の変化)を用いる.

主観的指標として SUS・raw TLX・SoA を用いる.

5 進捗

現在, 本実験に先立ち, 実験計画の妥当性を検証するため予備実験を実施した.

5.1 予備実験の目的と方法

本実験を正確かつ円滑に実施するため, 予備実験の目的を

- 実験システムの動作検証
- 参加者への指示の明確さの確認
- タスク難易度と所要時間の見積もり
- データ収集プロセスの確認

と設定した.

対象: 3名

手順: 本実験で実施予定のタスクを, 各操作方法を一回ずつ実施してもらった.

測定: 測定項目をシステムログとして記録し, 実験後にはアンケートへの回答によるフィードバック収集を行った.

5.2 問題点

予備実験の結果, 開発したシステムが動作し, ログデータが収集できることを確認できた. 一方で以下の主要な問題が明らかになった.

5.2.1 コントローラ

コントローラでは, プロパティ変更やオブジェクトの移動方法がわかりにくいという問題があった.

5.2.2 音声コマンド

音声コマンドでは, 参加者が定義済みコマンドを覚えられないという問題, 冗長な文法や現在までに認識されているコマンドが確認できないという指摘があった.

5.2.3 LLM

LLM では, 発話してから LLM Agent が操作を実行するまでに平均 20 秒かかり, 参加者が認識失敗と処理中を判断できなかった.

5.2.4 複雑タスク群

複雑タスク群では, タスクの指示が曖昧さや直前に行った操作によって, 指示が一意に定まらなかった.

5.2.5 評価指標

「操作実行回数 / 操作試行回数」の定義が甘かった. 特に LLM で「言い直し」や「処理の中断」を求めた場合, それを失敗とカウントするのか, 操作の一部とカウントするのか曖昧だった.

5.3 改善策

上記の問題に基づき, 以下の改善策を考案した.

1. 入力・選択状態を統一的に表示するインターフェースの追加
2. 音声コマンドリストの表示
3. 音声コマンド用文法の変更
4. ローカル LLM をクラウド LLM へ変更
5. タスク指示文を一義的な表現へ修正
6. ユーザーが発話を開始してから, LLM Agent が最終的な操作を完了するまでを 1 試行と定義

6 今後の展望

本研究の最終目的である LLM を用いた VR 操作手法の有効性評価に向け, 今後は以下のステップで研究を行う.

まず, 予備実験で特定された問題点に対応するため, 提示した改善策をシステムに実装する.

次に, 改良した実験システムを用いて本実験を実施する. 実験では, コントローラ, 音声コマンド, LLM の 3 手法を, 単純タスク群・複雑タスク群それぞれを複数回実行・比較し, タスク完了時間, エラー率などの定量的指標, および SUS, raw TLX, SoA といった主観的指標を取得する.

本実験で得られたデータの分析では, 特にタスクの複雑性が各操作手法の評価指標に与える影響に焦点を当てる.

参考文献

- [1] Wenlong Huang, Chen Wang, Ruohan Zhang, Yunzhu Li, Jiajun Wu, Li Fei-Fei: "VoxPoser: Composable 3D Value Maps for Robotic Manipulation with Language Models", CoRL 2023, Oral 4: LLM, 6
- [2] Mohit Nayak, Jari Kangas, Roope Raisamo (2025): "A Study of NLP-Based Speech Interfaces in Medical Virtual Reality", Multimodal Technol. Interact., Volume 9, Issue 6 (June 2025), 50
- [3] Xiangzhi Eric Wang, Zackary P. T. Sin, Ye Jia, Daniel Archer, Wynonna H. Y. Fong, Qing Li, Chen Li (2025): "Can You Move These Over There? An LLM-based VR Mover for Supporting Object Manipulation", UIST '25, 1. Interacting in XR, 3

5 : raspberry pi を用いたタンブラー自動洗浄機の試作

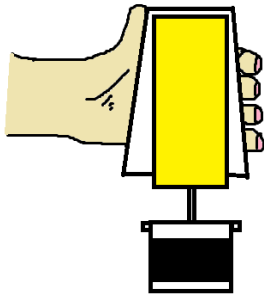
大豆生田研究室 浅川 武揚

1. はじめに

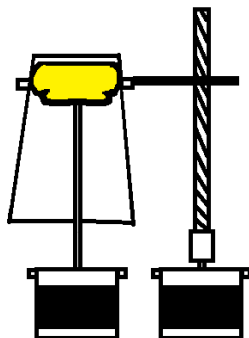
タンブラーは飲料を注ぐ容器として使われていて、使用後のタンブラーは食洗器に搬入し洗浄ができるが、搬入前に油汚れを落としてから使うことが望ましいため、乳製品等を使用した後のタンブラーの油汚れはスポンジを使って落とすが必要になる。

この作業を簡略化する電動グラスウォッシャーは図1(a)のようなブラシによる物理的な洗浄手法がある。物理的な洗浄手法をとるグラスウォッシャーは一定速度で回転する巨大なヘッドに逆さに持ったグラスを押し当てて洗浄する。

そこで図1(b)のようにブラシを小型化し、グラスの押し当て及びヘッドの回転速度等をラズベリーパイ(※)でコンピュータ制御することでグラスを押し当てる時間や強さを見極める手間を省き、さらには確実な洗浄プロセスを確立できると考えられる。本研究ではこのプロセスを確立し、さらに実用性について検証することを目的とする。



(a)電動グラスウォッシャー



(b) タンブラー自動洗浄機

図1 : 各種洗浄機

※

(使うマイコンは都合上ラズベリーパイからアルデューノ(以下 arduino uno)に変更となった。)

2. システム概要

図2にシステム構成図を示す。

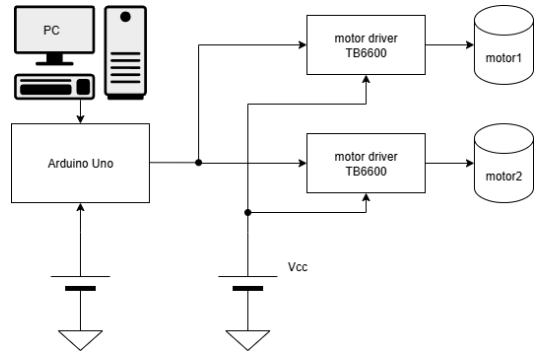


図2 : システム構成図

図1(b)のとおり、装置を制御するにはヘッドを回転させるモータとコップの上下位置を調整するためのリニアに動力を与えるモータの制御が必要になる。そのため二つのモータドライバを用意しそれらを1台の arduino uno のボードから制御信号を、直流安定化電源からモータ回転用の電力を送る。

また、arduino uno のボードはパソコンと接続しプログラムで命令が送られるようにする

3. 進行状況および実験方法

現在、図2のシステム構成の内、PC、arduino uno、モータドライバ2台、電源、片方のステッピングモータを入手している。

今後は、リニアと一体となったステッピングモータの入手、ブラシの製作及びモータへの取り付け、装置の構成、制御プログラムの作成の順で進めていく予定である。また、制御プログラムが完成でき次第以下の実験群を行う予定である。

① 汚れに対する効果の検証

ヘッドを湿らせ、油汚れが付いたタンブラーに対して洗浄を行い、洗浄による効果を画像データとして記録する。

② 装置のメンテナンスの必要性の検証

洗浄を行い続けて、装置に問題が生じないか調

べ、問題が生じた際の解決策について検討する。

③ 配置換えによる性能差の検証

デフォルトでは装置はタンブラーの注ぎ口が下向きの状態で洗浄するが、タンブラーが横向きであったり上下逆の状態での洗浄動作を行った場合パフォーマンスにどのような影響があるか検証する。

5 参考文献

¹食洗機の使い方ガイド！食器の入れ方から使用後の掃除までを解説/参照 2025-11-02/

<<https://panasonic.jp/life/housework/100114.html>>

²Glass Polishers/参照 2025-11-02/

<<https://bestinthebar.com/product/glass-polisher/>>

6：高効率な入力を目的としたキー配列の検討

大豆生田研究室 田部井珂維人

1. はじめに

従来のキー配列は、キーと文字の対応関係である論理配列が日本語入力に対して必ずしも最適なものではない。そこで、手指の運動特性及び文字頻度分布の観点から、日本語入力に対して最適なキー配列を検討する。この検討結果に基づいて、日本語入力の高効率化が可能となる。さらに、特定分野に対する入力ごとに最適なキー配列の検討も行う。

2. 従来のキー配列

キー配列は、キーボード上のキーの物理的配置及び各キーへの文字の論理的割り当てを指す。本研究では特に、キーと文字の対応関係である論理配列に着目する。

QWERTY 配列は、図 1 に示す JIS X 6002 の一部を構成し、ローマ字を用いた日本語入力に使用される標準的なキー配列である^[1]。しかしながら、当該配列は米国のタイプライタ配列を基礎として策定されたものであり、日本語入力の実情に即した最適化が十分に行われていたとは言い難い。

一方、JIS X 6002 を構成するもう一つの配列であるカナ配列は、かな文字を直接入力する方式であり、ローマ字入力方式と比較して打鍵数が少ないという利点を有する。しかし、五十音順を基礎として設計されているため、打鍵効率や手指の運動の観点から必ずしも最適とはいえない。この問題に対して、左右の手の交互の打鍵や同じ指の連続を文字頻度分布に基づき最適化した、図 2 に示す JIS X 6004 が制定された^{[2][3][4]}が、普及には至らず廃止された。



図 1: JIS X 6002 におけるキー配列¹²



図 2: JIS X 6004 におけるキー配列¹²

¹ <https://keyboard-layout-editor.com> により作成

² 文字キー以外の一部キーの位置を市販キーボードに近い配置へ変更

このように、日本語入力におけるキー配列の最適化は古くから課題とされており、これまでに親指シフト^[5]や大西配列^[6]など、複数の配列が提案されている。

本研究においては、日本語入力の高効率化を目的とし、手指の運動特性及び文字頻度分布の観点から、日本語入力に用いるキー配列の最適化を検討する。さらに、特定分野に対する入力ごとに最適なキー配列の検討も行う。

3. 研究内容・手法

JIS X 6004 など、多くの入力方式は、コーパスからの n-gram データの集計と、手指の運動特性の調査を基に、最適なキー配列の探索を行っている。本研究においても、同様の手法を用いる。具体的な手順を以下に示す。

1. Wikipedia, 青空文庫, SNS や小説投稿サイトなどのインターネット上の文献から日本語データを取得する。
2. n-gram 頻度を算出し、文字連鎖の傾向を調査する。
3. 手指の運動特性の調査を行い、キーボードの各キーの押しやすさのモデルを設計する。
4. ローマ字入力やかな入力など、入力方式を検討する。
5. 自動割り当てアルゴリズムの検討を行い、キー配列を探索する。
6. 既存方式との比較・評価を行う。

3.1. 日本語データの取得

インターネット上から日本語データを取得する。取得したデータは漢字かな交じり文であるため、平仮名・ローマ字への変換を行う。

3.2. n-gram

n-gram は、文字又は単語単位の n 個の並びをマルコフモデルで表す統計的手法である^[7]。例として、「吾輩は猫である。名前はまだ無い。」における文字単位の bigram(2-gram)は以下のように表すことができる。

「吾輩」「輩は」「は猫」「猫で」「であ」「ある」「る。」「名」「名前」「前は」「はま」「まだ」「だ無」「無い」「い。」

日本語入力に関するキー配列の最適化には、かな文字単位・ローマ字の文字単位での n-gram 頻度解析が用いられる。n-gram 頻度は、文字や文字列が出現する確率分布を近似的に表すものである。

3.3. キーボードの各キーの押しやすさのモデル設計

キーボードには、個々のキーの押しやすさに差があり、また打鍵の連続によって押しづらくなるキーの組み合わせも存在する。前者の典型的な例はホームポジションに配置されたキーであり、後者の例としては‘ZA’のような同一の指による連続打鍵があげられる。

本研究では、押しやすさのモデルを自己ループを含む重み付き完全有向グラフとして定式化する。

- キーの集合: $V = \{v_1, v_2, \dots, v_n\}$
- キー遷移の集合: $E \subseteq V \times V$
- キー遷移に対するコストの写像: $w : E \rightarrow \mathbb{R}_{\geq 0}$

以上により、グラフ G を式 1 で定義する。

$$G = (V, E, w) \quad (1)$$

3.4. 入力方式の検討

カナ入力方式、ローマ字入力方式又はその他の入力方式を検討する。

3.5. 自動割り当てアルゴリズム

キー配列の評価関数を以下のように定める。

- 文字の集合: $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$
- 配列写像: $\varphi : \Sigma \rightarrow V$ 全単射
- bigram 頻度の写像: $p : \Sigma \times \Sigma \rightarrow \mathbb{R}_{\geq 0}$
- 評価関数:

$$J(\varphi) = \sum_{\sigma_i, \sigma_j \in \Sigma} p(\sigma_i, \sigma_j) w(\varphi(\sigma_i), \varphi(\sigma_j)) \quad (2)$$

この評価関数は、文字列中に σ_i, σ_j が順に連続する確率とそのときのキー遷移の押しづらさの積の総和となる。キー配列の最適化は、押しづらいキー遷移が発生する確率は小さく、押しやすいキー遷移が発生する確率は大きくすることを目指す。したがって、評価関数を最小化する問題となる。

$$\text{minimize } J(\varphi) \quad (3)$$

これは、QAP(二次割当問題)であり、NP困難として知られている^[8]。キー配列の割当問題においては、集合 V が大きく、現実的な時間で解くことは難しい。そのため、配列写像 φ の探索には、アニーリングや遺伝的アルゴリズムなど、ヒューリスティック手法を用いることとする。

3.6. 比較

前述の JIS X 6002, JIS X 6004, 親指シフトや大西配列のほか、DvorakJP や月配列など既存の方式との比較を、製作したモデル上で行う。

4. 進捗状況と今後の予定

現在は、インターネットから取得した日本語データの平仮名及びローマ字への変換のプログラムを作成している。今後、先述の内容を並行して進める予定である。

参考文献

- [1] 日本産業規格: JIS X 6002:2011, 『情報処理用けん盤配列』, 日本産業標準調査会, 2011, <https://www.jisc.go.jp/app/jis/general/GnrJISSearch.html>
- [2] 渡辺定久: 『カナタイプストにおける指の運動特性について』, 情報処理学会研究報告ヒューマンコンピュータインタラクション(HCI), 1983, 17(1983-HI-011), p. 1-10, 1983-07-13, <https://ipsj.ixsq.nii.ac.jp/records/37879>
- [3] 渡辺定久: 『カナタイプストにおける指の運動特性について(続報)』, 情報処理学会研究報告ヒューマンコンピュータインタラクション(HCI), 1986, 74(1986-HI-009), p. 1-8, 1986-11-12, <https://ipsj.ixsq.nii.ac.jp/records/37809>
- [4] 日本電子工業振興協会, 『日本語情報処理の標準化に関する調査研究報告書[昭和 59 年度]』, 1985, <https://dl.ndl.go.jp/pid/12606676/>
- [5] 情報処理学会, “親指シフトキーボード”, <https://museum.ipsj.or.jp/computer/word/0006.html>
- [6] 大西拓磨: “大西配列 | ローマ字をもっと打ちやすく”, <https://o24.works/layout/>
- [7] 中川裕志: “N-gram モデル”, https://ocw.u-tokyo.ac.jp/lecture_files/is_01/6/notes/ja/langinfo6.pdf
- [8] 久保幹雄: “2 次割当問題”, <https://scmopt.github.io/opt100/34qap.html>

7: 平衡二分探索木を用いた インデックス無制限の遅延評価セグメント木の試作

大豆生田研究室 藤田 恭輔

1. はじめに

本研究では、「平衡二分探索木を用いたインデックス無制限の遅延評価セグメント木の試作」と題して、従来は列長が固定長であった遅延評価セグメント木(後述 2.1 節を参照)や遅延評価の困難だった動的セグメント木(後述 2.3 節を参照)について、平衡二分探索木上に縮約した区間を乗せて表現し、列長を可変としインデックス(以下、キーと表記する)の定義域に関して制限をなくすことを提案する。

データ構造の試作を C++ で行い、性能を評価する。

2. 既存のデータ構造

2.1. 遅延評価セグメント木

数列において、区間に対して演算を作用させたり、区間に対する演算結果を取得したりすることが、数列長を N としてクエリ毎 $O(\log N)$ できるデータ構造である。¹なお、セグメント木上で扱える演算体系は、モノイドと呼ばれるものに限られる。

モノイドとは、単位元が存在し、かつ結合法則を持つ演算が定義された集合のことである。集合を S 、 S の任意の元を a, b, c 、単位元を e 、二項演算を f とし、それぞれ式(1)(2)に示すような性質が要求される。また、群よりも制約が弱く、逆元が存在を要求されない。

$$a = f(a, e) = f(e, a) \quad (1)$$

$$f(f(a, b), c) = f(a, f(b, c)) \quad (2)$$

2.2. 平衡二分探索木を用いた連想コンテナ

キーとそれに対応するヴァリュー値を格納する連想配列である。実装上は、赤黒木や AVL 木、Splay 木等で構築した平衡二分探索木においてキー順に各要素を格納することで実現され、要素数を N としてクエリ毎 $O(\log N)$ で各処理が可能である。

ある。

2.3. 動的セグメント木

本実験で作成を目指すものから、遅延評価をなくしたものである。平衡二分木上に、入力を受けたインデックスのノードを生成して、データ管理を行う。クエリ数を Q とし、全体の時間計算量 $O(Q \log Q)$ で処理が可能である。

2.4. 遅延伝搬反転可能乱択平衡二分木

区間演算や区間反転、および数列に対する割り込み

挿入などを対数オーダー(平均計算量)で可能としたデータ構造である。

また、先述の平衡二分木とは異なり、ランダムな頂点を根とすることで期待値的に高さの平均を要素数を N とし $O(\log N)$ とするデータ構造である。

アクセスには、先頭から数えて何番目かというインデックス(したがって、0 以上要素数未満の整数のみのキー)またはイテレータを用いる。したがって、任意に設定できる主キーによるアクセスができない。

本研究では、任意のキーに対するヴァリューを格納するようなデータ構造を提案する。ただし、キーとキーとの差が 1 の所に割り込み挿入して後ろのキーをずらすような処理は出来ないデータ構造である。

3. 提案手法

任意の整数 L, R ($L < R$) に対して、

- ・区間 $[L, R)$ に対する区間加算
- ・区間 $[L, R)$ に対する区間乗算
- ・区間 $[L, R)$ に対する区間代入
- ・区間 $[L, R)$ に対する区間と取得

を行うことができるよう、次のようなデータ構造を作成する。

3.1. データ構造の仕様

平衡二分探索木として、赤黒木²を用いる。数直線全体(はじめ $(-\infty, \infty)$ で初期化するものとする)を木に持ち、入力されていない領域は全要素が 0 であるものとして扱う。縮約された区間をキーに持つ二分探索木の各ノードは次の情報をもつものとする。一般的なセグメント木の実装とフォーマットを合わせるため、区間は半閉半開区間とした。

- ・赤または黒の色
- ・縮約区間の左端と右端
- ・区間内の各要素のヴァリュー
- ・部分木が管理する数列の区間の左端と右端
- ・部分木が管理する数列の範囲における総和
- ・遅延評価値
- ・左の子ノードへのポインタ
- ・右の子ノードへのポインタ
- ・親ノードへのポインタ

3.2. 区間演算と遅延評価の集計について

縮約されている区間を分割しなくてはならないようなクエリを受けた場合、演算を適用する前に分割すべき縮約された区間のノードを用意する。挿入は赤ノードで行い、赤黒木の制約が満たされていない場合、木を回転して挿入操作を完了させる。操作の一例を図1に示す。縮約された区間の分割処理の完了後、従来の遅延セグメント木と同様の要領で上位セグメントの未処理の遅延評価を行ってから演算を適用し、区間演算の適用後、深い方から浅い方へ向かってセグメントの再集計を行う。

4. 評価手法

4.1. 計算量について

区間を新たに1つ作るごとに、元々あった1つの区間が最大3つに分かれるか、元々あった2つの区間が最大4つに分かれるケースが最も管理すべき縮約された区間を増やすパターンである。一例を図2に示す。また、このことから、クエリごとに最大2つしか平衡二分探索木のノードは増加せず、ノード数の上界は、クエリ数を Q として $O(Q)$ で抑えられる。

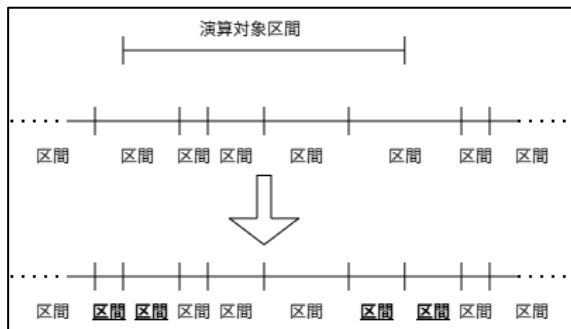


図2 区間が2個増える例

そして、管理すべき区間は、与えられたクエリに登場する最小のキーと最大のキーとの間だけである。そのため、その範囲内でどのような複雑な処理が行われたとしても、ノード数は最大のキーと最小のキーとの差分を K として K 以下であるといえる。したがって、ノード数の上界は $O(K)$ でも抑えられる。

ノード数が $O(\min(Q, K))$ であることから、木の高さは赤黒木の性質より最大で $O(\log(\min(Q, K)))$ である。したがって、計算量はクエリ毎 $O(\log(\min(Q, K)))$ と言える。

4.2. 正当性の評価とオーバーフローについて

正しく計算が行えているかどうかを確認するために、配列上で愚直に1要素ずつ区間加算や区間和取得を行うプログラムを作成する。それと今回試作したデータ構造を利用したプログラムが、同じ入力ケースに

対して同じ出力を返すかを、何ケースも調査することで、今回試作したデータ構造が動作しているかを確認する。なお、ランダムテストでは、完全にバグが無いことを証明することは不可能である。

また、今回はC++で64ビット整数を使って、キーおよびヴァリューを管理しているため、大きな値を管理しようとするときオーバーフローを起こすが、理論上は多倍長整数を用いてキーもヴァリューも上限なく管理することが可能である。

5. 進捗と今後の課題

5.1. 現在の成果

遅延評価セグメント木より原理が平易な、セグメント木(遅延評価セグメント木と比較すると、区間演算が出来ず一点更新のみが可能なデータ構造)での実装はすでに完成した。

また、これについて、4.2節で述べた手法を用いてランダムテストを行い、ランダムテストで洗い出せたバグは全て取り除いた。

5.2. 今後の課題

当初予定していた、平衡二分探索木を用いたインデックス無制限の遅延評価セグメント木の完成を目標とし、実装を行う。

また、完成したら、各種メゾットやイテレータなどの周辺機能の実装を行うほか、任意のモノイドに対応するよう拡張も行いたい。

参考文献

- [1] 秋葉拓哉、岩田陽一、北川宜稔『プログラミングコンテストチャレンジブック第2版』株式会社マイナビ(2012) p.153
- [2] 【公式】ピカラ：『Red-Black Tree by Java & Python -- これで分かった赤黒木』
<http://www.pikara.ne.jp/okojisan/rb-tree/index.html>
2025/11/01 閲覧

8：マーカーレスモーションキャプチャを用いた、 フォーム変化予測モデルの開発

荒川研究室 土谷 琢磨

1. はじめに

野球における投手の投球フォームは、選手の競技成績や能力に極めて大きな影響を与える要素である。投球動作は、全身の関節を連動させる運動連鎖の集大成であり、肩や肘に大きな負荷を伴う。投球障害（野球肘、野球肩など）はスポーツ医学上の深刻な問題であり、その発生要因の多くが、疲労の蓄積に伴う投球フォームの再現性の低下や崩れに起因することが、近年の研究で指摘されている[1]。現在、指導現場における選手のコンディション把握は、指導者の経験や選手の主観的な訴えに頼る部分が大きく、客観的な根拠が不足している。この主観的評価の限界は、疲労が限界に達する前の微細なフォーム変化を見逃す要因となり、選手の潜在的な怪我のリスクを増加させている。

本研究は、この課題を解決するため、モーションキャプチャによる客観的な動作データと、主観的な疲労度を統合する。そして、疲労の蓄積という時間的な要素を考慮した、個人特化型の予測モデルを構築し、投球フォームの具体的な変化量を事前に予測する。これにより、客観的な数値に基づき選手の潜在的な傷害リスクを把握することを可能にし、疲労が限界に達する前の適切なタイミングでの介入（指導や休養）を支援する技術的基盤の確立を目指す。

2. 先行研究

先行研究では、フォーム解析や生理学的計測を用いたパフォーマンス予測が進めら

れているが、疲労によってフォームがどれだけ変化するかを事前に予測するモデルは未だ確立されていない [2]。また、従来の光学式モーションキャプチャは、設置が難しく、直接接触を伴う場面や、野外で行うスポーツの計測は困難であると結論付けられているものがある[3]。

本研究は、これらの課題に対し、以下の3点で新規性を持つ。

- ① 先行研究が疲労の有無を分類・検出するのに対し、本研究は「未来のフォーム変化の偏差」を予測する。これにより、疲労が顕在化する前の介入のタイミングを最適化する。
- ② 従来の計測の課題であった直接接触を伴う計測の困難さに対し、非接触型のマーカーレスモーションキャプチャを用いることで、野外での簡便なデータ収集を実現し、データ取得の制約を克服する。
- ③ 選手一人ひとりに最適化された独立したモデルを構築し、最高水準の予測精度を追求する。

3. 提案手法

3.1. 概要

本提案手法では、投球練習時の動作データをもとに、将来的な疲労によるフォームの変化を予測するモデルを提案する。図1に提案手法の流れを示す。本手法は「動作学習部」と「変化予測部」から構成され、動作学習部では取得したデータを基に予測モデルを構築し、変化予測部では構築した予測

モデルを用いて、変化予測を行う。

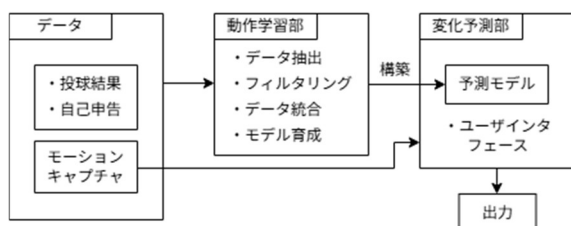


図1:提案手法の処理の流れ

3.2. データ取得

2~5名の投手を対象に、4週間以上、週2,3回の頻度でデータ収集を行う。マーカーレスモーションキャプチャによる客観的データ(関節角度、座標など)、球速、回転数と主観的疲労度を取得対象とする。実際のデータ取得の様子を図2に示す。



図2:データ取得の様子(仮)

3.3. データ抽出・フィルタリング

骨盤の移動や股関節の角度から投球開始、肩の回旋やひじの屈折から投球終了を定義し、投球動作を切り出す。時間正規化を行うため、線形補間を用いて、各データのフレーム数の統一を行う。練習開始直後のデータをベースラインと定義し、その後の偏差を予測の対象として定義する。また、動作データと主観データの統合を行う。

3.4. 予測モデルの構築

RNNでは、長期間のデータ処理に不向きであるため、ゲート構造を持ち、情報の取捨

選択が可能なLSTMを採用する。フォームには個人差があるため、個人特化モデルを作成し、予測精度を高める。また、比較用に混合モデルも作成する。過学習を防ぐための早期終了、損失関数などの実装も検討している。

3.5. モデル評価

RMSEや決定係数などの評価指標を用いて、モデルの精度検証を行う。また、各個人モデルや混合モデルとの比較を行う。

4. 進捗

現在は、データ取得を行うため、人間(ヒト)研究倫理審査に向けての準備を進めている。また、自身のフォームを対象に、データ抽出・フィルタリングについてのシステムの試作を行っている。

5. 今後の展望

人間(ヒト)研究倫理審査が通り次第、データの収集を開始する。収集と並行して予測モデルの構築を進める。

参考文献

- [1] 東善一, 松井知之, 山口弘佑, 平本真知子, 宮崎哲哉, 山崎勢那, 瀬尾和弥, 来田宣幸, 森原徹, 丸太町リハビリテーションクリニック, 京都府立医科大学附属病院リハビリテーション部, 京都工芸繊維大学基盤科学系: 球速および制球のパフォーマンスアップに必要な投球フォームとエクササイズについて: 元プロ野球投手かつ理学療法士に対する聴き取り調査から, <https://www.jstage.jst.go.jp/article/rjsp/15/0/15_2218/_pdf/-char/ja>
- [2] Yi-Wei Ma, Jiann-Liang Chen, Chia-Chi Hsu, Ying-Hsun Lai: Design and Analysis of a Pitch Fatigue Detection System for Adaptive Baseball Learning, <<https://pmc.ncbi.nlm.nih.gov/articles/PMC8711585/>>
- [3] 西山勇毅: 投球障害予防支援システムのための投球フォーム変化検出アルゴリズムの提案, <<https://www.sfc.wide.ad.jp/thesis/2011/files/tetujin-publish-thesis.pdf>>

9：3D 姿勢推定を応用したアニメーションからのボーン抽出

AI の開発

荒川研究室 武田 幸寛

1. はじめに

本研究では、3DCG 映像を拡散モデルによってアニメ調に変換した映像を学習データとして使用することで、アニメーションからボーンを抽出する AI を実現する手法を提案する。

2. 3D 姿勢推定とは

3D 姿勢推定とは、身体の動きや指先の細かな角度、第三者から観測した場合に重なっている腕などを 3D 認識する技術である[1]。スマートフォンなどで撮影した人物の映像から、その人物がどのような動きをしているかを捉え、これをリアルタイムで 3DCG モデルのアバターに反映させて動かすといったことが可能である。この 3D 姿勢推定には、深層学習によって構築された 3D 姿勢推定モデルが用いられる[2]。モデルの学習には、2 次元画像の人物(学習データ)と、その人物の 3 次元における各関節の座標(教師データ)があり、2 次元画像からモデルが推定した座標との誤差を教師データと評価することでネットワークを学習させていく。また、姿勢の推定には、その時間でのフレームのみを用いるのではなく、時間的な一貫性を持たせるために、前後のフレームの姿勢との関係についても考慮することで、抽出される動きの正確性や滑らかさが向上する。

学習データの取得には、実際にモーションキャプチャで人間の動きを撮影した得られた 2 次元映像と、3 次元空間における座標をデータセットとする実データを用いる方法や、3DCG のアバターをコンピュータ上で動かし、この映像とソフトから得られるアバターの関節の情報をデータセットとする合成データを用いる方法がある[1]。実データを用いる方法では、学習データである映像が実際の人間の映像であるため、最

終的に動きを抽出したい映像とのギャップが少ないという利点がある。しかしモデルの学習には大量のデータセットが必要となるため実際にデータセットを作成するのは困難である。対して合成データを用いる方法では、学習データで用いる 3DCG と実際の人間にはギャップが大きくなるという欠点があり、データセットの作成はコンピュータ上で行えるため容易であるという利点がある。

3. アニメーションからの抽出について

アニメ調の映像から動きを抽出するためのアプローチの研究は過去に行われている[2]。この研究では、まず、アニメ調の映像に対して実際の人間の映像をもとに学習したモデルでは関節の座標を正確に抽出することはギャップにより困難であることが分かった[3]。しかし、ギャップを緩和するためにアニメ調の映像を学習データとし、これに関節の座標をラベル付けてデータセットを作成するのは困難であるため、合成データを用いることとなった。合成データは、3DCG キャラクターソフトウェアである MikuMikuDance のモデルの映像と、このソフトより得られる MMD モーションキャプチャデータ(3D 関節座標)を用いる。学習には、MMD モデルによって得られた大量の合成データを用いて、畳み込みニューラルネットワークをベースとした簡単なポーズ推定モデルを学習させ、次に手書きのアニメ調イラストに対して 2D 関節座標をラベル付けした少量のデータセットを用いて学習させるという 2 段階で行われた。この研究により得られたモデルを用いることで、アニメ調の映像をもとに正確な 2D 関節座標の抽出を行うことが出来た。しかし、

この研究では抽出される座標が2D空間上のものであるため奥行きといった情報が得られない他、使用した合成データが3DCG由来の映像であるため、アニメ調の映像とのギャップは存在していたため動きの抽出には限界があった。図1に、前述した先行研究のシステムの実行プロセスを示す。ボーンを抽出したいアニメ調映像を学習済みの畳み込みニューラルネットワークモデル(CNN)に入力すると、各関節点が存在する確率を示すヒートマップを作成する。このヒートマップをもとに、最も信頼性の高い関節点の座標を決定し、これをデータとして保存する。

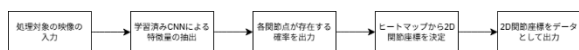


図 1:先行研究のシステムの実行プロセス

4. 本研究の課題と目標

先行研究では、合成データを学習に用いることはアニメ調の映像から動きを抽出するアプローチとして有効であることが示された。そこで、本研究では、異なる合成データの作成とそれを用いた学習方法を提案する。本研究の一番重要な課題としては、学習用の映像と実際にボーンを抽出したいアニメ調映像との間のギャップを小さくし、抽出する座標情報を3次元上でのものにするということとする。目標としては、アニメ調の映像を読み込ませると、この映像に即したボーンを3Dソフトウェアで扱うためのFBX形式で出力されるという動作である。具体的には、学習済みモデルに処理対象のアニメーションを入力することで、一度2D姿勢を推定し、この連続した2D姿勢の情報を用いて、さらに2D姿勢から3D姿勢にリフトアップするという2段階の処理によってボーンを抽出する。この2D姿勢から3D姿勢にリフトアップするという手法は、より高精度のボーンを抽出するための基本的な手法である[4]。図2に、本研究で提案するシステムの実行プロセスを示す。まずはボーンを抽出したいアニメーションを学習済みモデルに入力し、これに

より2D関節座標を得る。次に、連続した2D関節座標から3D関節座標にリフトアップし、これより各関節の回転や動きといったパラメータを計算する。これらのパラメータと骨格情報を結合し、ボーンの動きを作成してFBX方式に変換する。

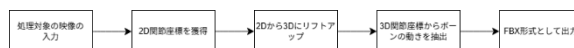


図 2:本研究で提案するシステムの実行プロセス

合成データの作成には、blenderでモデルを動かし、この映像とモデルの各関節の3D座標を取得する。この映像に対して、構造保持型画像変換という技術を用いて、モデルの身体的な情報(ポーズ)を変えずに描画スタイルを3DCG映像からアニメ調の映像に変換する。構造保持型画像変換とは、画像から画像に変換する拡散モデルで、特に元の画像の3D構造を高い精度で維持するものである。この変換モデルを用いて作成したアニメ調の映像を学習データ、そしてblenderより得られた3D関節座標を教師データとするデータセットを用いることで、さらにギャップを減少させた奥行きの情報も抽出できる3D姿勢推定モデルを学習できる。最終的な目標は、あらゆる描画スタイルで描かれたあらゆる身体的特徴を持つキャラクターからボーンを抽出することだが、本研究は合成データの作成手法の検証に焦点を置きたい。よって、作成する合成データの人物を「スーツ姿の男性」といったように固定して行う。これにより、ある特定のキャラクターのボーン抽出に必要な学習量が特定できる他、学習量の不足によるボーン抽出の失敗を予防できる。

参考文献

- [1]知財図鑑：3D姿勢推定の概要，<<https://chizaizukan.com/property/403/>>，(2025/10/30)
- [2] L. Kumarapu and P. Mukherjee: AnimePose: Multi-person 3D pose estimation and animation, rXiv preprint, arXiv:2002.02792, (2020).
- [3]P. Khungurn, E. Chuangsuwanich, and J. Fakcharoenphol : "Pose Estimation of Anime/Manga Characters: A Case for Synthetic Data", Proc. MANPU '16, pp. 1-7 (2016).
- [4]G. Moon, J. Y. Chang, and K. M. Lee:"A Simple Yet Effective Baseline for 3D Human Pose Estimation",Proc. ICCV, pp. 5240-5248, (2017).

10：ミキシング支援 AI の開発

荒川研究室 新井 春登

1. はじめに

近年、AI 技術の発展により音楽制作分野においても自動マスタリングや自動ミキシングなどのツールが登場している。これらのツールは、経験の浅い制作者でも一定のクオリティを持つ音を得られるという利点を持つ一方で、制作者の意図やジャンル特性を十分に反映できないという課題がある。

また、既存の自動ミキシングシステムは結果としての音量・EQ バランスなどを出力するのみで、その判断根拠が不明瞭な“ブラックボックス的処理”となっている場合が多い。したがって、制作意図を反映しつつ、AI による支援の透明性を確保する手法の構築が求められている。

2. 先行研究との相違点

本研究の特徴は、単なる自動化ではなく、「ユーザ入力に基づく意図反映型 AI ミキシング支援システム」である点にある。

従来の AI ミキシングが“結果だけを提示するブラックボックス”であるのに対し、本研究では音響特徴量・メタ情報・推定理由を組み合わせ、透明性と柔軟性を両立した支援を実現する。

先行研究や商用ツール（例：iZotope Neutron[1]、Sonible smart EQ 4[2]など）は、AI が音響信号を解析し、自動でパラメータを調整することに重点を置いている。一方、本研究では、ユーザが入力する音楽情報（ジャンル・BPM・雰囲気など）を考慮に入れたうえで、AI が推奨ミキシング方針を提示するという人間と AI の協調型支援を志向している。

さらに、音響特徴量の抽出においても、スペクトルを7帯域（Sub-bass～Brilliance）に細分化し、より詳細な帯域エネルギー分布を取得することで、楽曲の質感を多角的に解析する。このように、主観的な制作意図と客観的な音響解析を統合する点において、従来研究とは明確な差異を有している。

3. 研究目的

本研究の目的は、音響特徴量とユーザ入力情報を統合し、ミキシング方針を提案する AI 支援システムを設計することである。これにより、初心者がジャンルや雰囲気に適したミキシングバランスを得やすくするとともに、経験者にとっても客観的な指標を提供す

ることを目指す。

4. システム構成

本システムは主に以下の3段階で構成される。各モジュールは、音響特徴量の客観的分析とユーザ意図の主観的情報を統合するために設計されている。

図1にシステム構成図を示す。

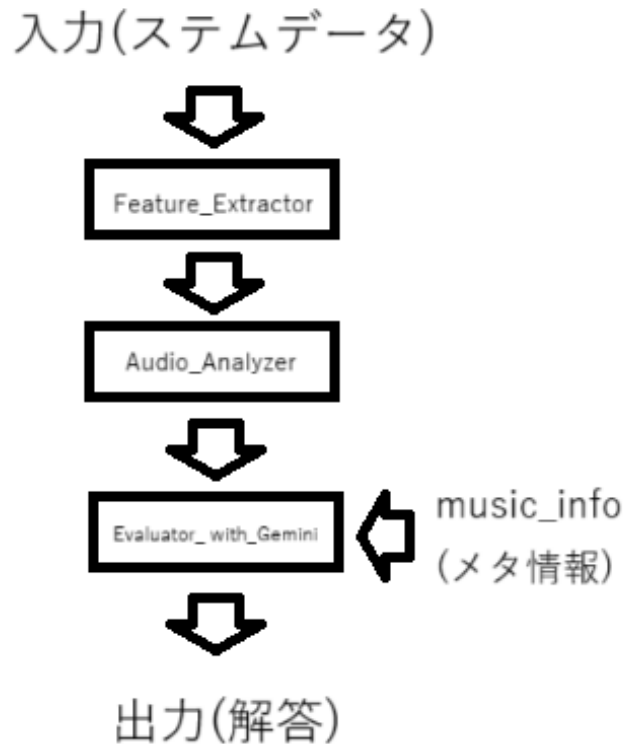


図1:システム構成図

4-1. Feature Extractor (特徴量抽出)

各トラック（例：Drums, Bass, Piano など）から音響特徴量を抽出する。

解析項目は、平均ラウドネス、7帯域エネルギー、スペクトル重心、ロールオフ、ステレオ幅、トランジェント密度、ダイナミックレンジ などである。これらの数値は、音響的な傾向を定量的に把握する基礎情報として利用される。

4-2. Audio Analyzer (音響解析モジュール)

本モジュールは、各トラック（ステムデータ）の

WAV ファイルを入力として受け取り、それぞれの音響特徴量を抽出・統合し、1 曲分の特徴情報を JSON 形式で出力する役割を担う。

具体的には、feature_extractor.py によって算出された平均ラウドネス、7 帯域の周波数エネルギー、スペクトル重心、ロールオフ、ステレオ幅、トランジェント密度、ダイナミックレンジといった特徴量を集約し、曲全体のバランス傾向をデータ構造として整理する。

解析対象は「曲フォルダ」として構成され、各フォルダ内には複数のトラック(例:drums.wav、bass.wav、piano.wav など)が含まれる。ファイル名から自動的に楽器名を抽出し、トラック単位の特徴を階層的に整理することで、後段の評価・提案モジュール(Gemini)に入力可能な形式の JSON データを生成する。

このように、本モジュールは複数トラックの音響情報を一元化する“中核的解析部”として機能し、AI によるミキシング支援のための定量的基盤を提供している。

4-3. Evaluator with Gemini (評価・提案モジュール)

Audio Analyzer が生成した特徴量 JSON および楽曲メタ情報(=作曲者が入力する「この楽曲はどのジャンル/どんな雰囲気を目指しているか」「どの楽器を主役にしたいか」などの意図記述テキスト)を入力とし、Google Gemini API を用いて統合的な音響評価とミキシング提案を行う。

Gemini は特徴量の傾向と楽曲意図を関連付けて解析し、「どの帯域を強調すべきか」「どのトラックの音量や EQ を調整すべきか」といった推奨設定を自然言語で出力する。

また各提案には、AI が導いた推定理由を付与することで、ブラックボックス的な自動処理ではなく透明性の高い支援を実現している。

図 2 に、手動で作成した本提案システムの出力の例を示す：

```
■曲意図 (ユーザー入力/メタ情報)
「この曲はLo-Fi HipHop。ドラムは乾いた質感で前に出したい。ベースは控えめでOK。」
■AI提案 (Gemini出力例)
Drum (Kick/Snare) : 2kHz~5kHzのアタック成分が埋もれている。
→ +3dB程度のブーストでタイトさを確保できる。
Bass : 60Hz近辺が過多で、ドラムとの分離が悪い。
→ -2dBの減衰 or HPFを70Hz開始で濁りを軽減。
Piano/Pad : ステレオ幅が広すぎて中域を占領している。
→ ステレオ幅を0.7→0.5へ、ボーカル帯域を開ける。
(理由)
本メタ情報に記述された「ドラムが前」「ベースは控えめ」を満たすためには、低域はBassではなくKickで存在感を担保する設計が適切である。したがって、Bassを少し引き算し、Drumのアタックを強める処理を推奨した。
```

図 2 提案システム出力結果 (仮)

5. 今後の展望

本発表までに、提案システムが「ユーザ意図を反映しながら、適切な音響的補正を行えるか」を検証する。ここでのユーザ意図は、上記のメタ情報入力テキスト(例:「この曲は低域をタイトにしたい」「ボーカルの中高域を前に出したい」等)として明示的に与える。Geminiはこの意図を条件として参照することで、「ただの自動補正」ではなく「意図を尊重した方向の補正」を返す。すなわち、ブラックボックス的な最終出力精度ではなく、AIが“意図を理解して支援する”過程の有効性を評価することを目的とする。

また本システムの有効性を検証するため評価実験として、「元ミックス → 意図的な崩し → AI提案 → 手動で反映 → 結果ミックス」

の流れを複数曲で実施し、①スペクトル距離およびステレオイメージ距離などの客観指標と、②“意図を反映できているか”に関する主観評価(ミキシング経験者3~5名程度の5段階評価)を比較する。

これにより、ブラックボックス的な最終出力精度ではなく、AIが“意図理解を伴って支援する”プロセスの有効性を評価する。

6. まとめ

本研究では、ユーザの制作意図を反映しながら、AIがミキシング支援を行う新しい枠組みを提案した。

従来の自動ミキシングの限界であった“ブラックボックス化”を回避し、解析結果の根拠を可視化することで、音楽制作におけるAI支援の新しい形を示すものである。

7. 参考文献

[1]:iZotope 公式サイト Neutron 5 (2025年11月1日閲覧)

<https://www.izotope.jp/jp/products/neutron-5/>

[2]:sonible 公式サイト smart:EQ 4 (2025年11月1日閲覧)

<https://www.sonible.com/smarteq4/>

11：言語処理を用いた人狼ゲームにおける形勢評価

荒川研究室 金井 佑心

1. はじめに

人狼ゲームは、複数のプレイヤーが「村人」または「人狼」の役職を持ち、互いの発言や行動を通じて相手の正体を推理する心理ゲームである。プレイヤーは、議論や投票を通して他者の信頼度や疑い度を評価し、推理の結果に基づいて投票を行う。本ゲームは、情報の不完全性、嘘やブラフの存在、議論の多様性といった特徴を持つため、人間の思考過程や意思決定の分析対象としても注目されている。[1]

近年、人工知能による「人狼知能」の研究が進み、発言解析に基づく形勢評価や行動予測のモデル化が試みられている。しかし、従来の手法では定型文や固定ルールに依存した解析が中心であり、実際のプレイヤーの多様な発言表現や表記ゆれに対応できていないことが課題であった。

本研究では、日本語の自然言語処理技術（正規表現に基づくキーワード解析）を用いて、ゲーム中の発言から各プレイヤーの疑い度を定量化することを目指す。さらに、疑い度を基に投票行動の確率推定を行い、予測と実際の投票結果との比較を通じて形勢評価の精度を検証することを目的とする。

2. 関連研究

従来の人狼知能研究では、議論ログを解析してプレイヤーの役職や疑い度を評価する手法が報告されている[2]。これらの手法は、主に定型文ログに基づくルール型解析や単純なスコア集計を行うものであり、表現の揺れや言い換え、複雑な文章構造には十分対応できていない。

同じ意味の表現でも「怪しい」「あやしい」「疑わしい」など複数の言い回しが存在する場合、従来手法では一部の表現しか検知できず、疑い度の評価が不正確になることがある。また、肯定的表現（信頼・白など）や否定的表現（疑い・黒など）の扱いも限定的であり、人間の直感に近い形勢判断を反映できないという課題がある。

本研究では、正規表現ベースの部分一致解析と辞書拡張を用いることで、表記ゆれや言い換えを考慮した柔軟な疑い度算出を試みる。また、疑い度に基づく投票確率の推定を導入することで、単純なスコア比較では得られないプレイヤーの投票傾向や形勢の動的変化を分析可能とすることを目標とする。

3. 方法

本提案手法では、人狼ゲームにおいて現局面までのプレイログ(テキストデータ)を入力することにより、次回投票における各プレイヤーが投票で処刑される確率を計算して出力する。提案手法の処理の流れを図1に示す。

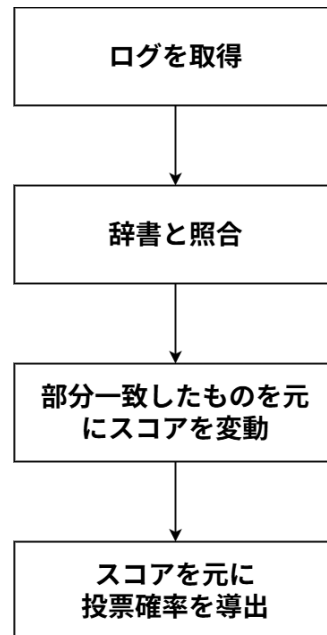


図1 提案手法の処理の流れ

図1における『スコア』とは、各プレイヤーが他のプレイヤーに対し、どれぐらい強く人狼ではないかと疑っている強さを表す数値であり、プレイヤーの人数をサイズとする2次元配列によって表される。(4節表1、表2参照)

4. 試作システムによる予備実験

4.1 データ準備

予備実験では、まず半自動生成された短文ログを用いて手法を検証した。

形式：発言者, 発言内容

発言例：「Agent[01] は怪しいと思う」

キーワード辞書には、肯定・否定表現、怪しい・信頼などの語彙を50語以上登録し、部分一致で1回のみ加算するルールとした。これにより、同じ単語が複数回出現してもスコアの計算は1回に制限される。

4.2 キーワード辞書による疑い度計算

発言文中のキーワード出現を検知し、その言葉に応じて対象の Agent の疑いスコアを加算もしくは減算する。部分一致を用いることで表記揺れや言い換えに対応させている。

また、対象 Agent が明記されていない場合はスコアの変動なしとし、疑いスコアは2次元配列（行：発言者、列：対象者）で保持させる。

本研究では重みを付けず、一定の疑いスコア(1.0)の計算を行う。

4.3 個人疑い度を元にした投票確率モデル

累積疑いスコアに基づき、各プレイヤーが特定の対象に投票する確率を計算

本研究では一番スコアが高い Agent に確率 100 を割り振る方式をとる。ただし、自分への投票確率は必ず 0 とする。

これにより、実際の投票行動と比較可能な確率分布が得られる。

4.4 実行結果

テストデータとして以下のようなログを使用した。プレイヤーは4人、人狼は1人の想定で行った。

Agent[01],Agent[02]が怪しいと思う
Agent[02],Agent[03]は信用できる
Agent[03],Agent[01]は人狼っぽい
Agent[04],Agent[03]は信じる
Agent[02],Agent[01]は黒かも

その結果、以下のような2次元配列、投票確率となった。

表1 疑いスコア

	01	02	03	04
01	0	1	0	0
02	1	0	-1	0
03	1	0	0	0
04	0	0	-1	0

表2 投票確率

	01	02	03	04
01	0.0	100	0.0	0.0
02	100	0.0	0.0	0.0
03	100	0.0	0.0	0.0
04	50	50	0.0	0.0

発言内容から疑っているであろう人物には正の比率、信じている人には負のスコアがかかり、投票確率も最も高いスコアの Agent に 100 を割り振っている。

4.5 考察

2次元配列によって、疑われている人と疑っている人の関係を詳細にすることが可能になっている。

投票確率推定により、疑いスコアだけでは分からない動的な投票傾向の分析が可能になっている。これらのデータに基づき、現局面において村人陣営、人狼陣営のどちらがどの程度優勢であるかを示すことも可能と考えられる。

5. まとめ

本研究では、正規表現によるキーワード解析と投票確率モデルを組み合わせることで、人狼ゲームにおける形勢評価手法を構築した。

半自動生成ログを用いた初期実験により、疑い度の算出と投票確率の推定が可能であることを確認した。

今後は、実ゲームログへの応用やキーワード辞書の拡張、さらにはプレイヤーの心理的要素を取り入れた精緻な形勢判断モデルの構築が課題である。

参考文献

- [1] 竹本春太, 三浦莉, 樋渡海斗, 他, 「人狼ゲームのログデータを用いた戦術の分析」, 第75回電気・情報関係学会九州支部連合大会, 04-2A-09, 2022.
- [2] 荒川達也, 宮田洋輝, 市村智康, 他, 「会話ログ・進行ログから陣営有利度を評価する人狼ゲームの数理モデル」, 群馬高専レビュー Vol.39, pp.55-62, 2020.

12: 衛星画像情報と出没実績を融合したハイブリッド特徴量による地域別クマ出没予測モデルの構築

西山研究室 田村伊織

1 はじめに

近年、人里におけるクマの出没件数は継続的に増加しており、地域住民の生活圏における安全確保と生態系保全の両立が重要な社会課題となっている。従来のクマ出没予測モデルは、主に地形因子や過去の出没実績といった静的または履歴的な特徴量に依存してきたが、クマの出没は餌資源の変動や植生の遷移といった動的な環境要因に強く影響を受けるため、予測精度に限界があった。特に、広範囲かつリアルタイムな植生の変化を捕捉することが、モデルの性能向上における鍵となる。

本研究の目的は、この課題を解決するため、衛星画像から得られる多時期の環境指標（正規化植生指標：NDVI）と、過去の時空間的な出没実績データを統合したハイブリッド特徴量を設計し、これに基づく高度な機械学習モデル（XGBoost）【1】を用いて、地域別のクマ出没予測モデルの構築手順を確立することである。本アプローチは、より精度の高いリスクマップの生成を可能にし、効果的な早期警戒システムへの応用を目指す。

2 ハイブリッド特徴量に基づく予測モデルの構築手順

予測モデルは、環境、実績、地形の3つの特徴量カテゴリを統合したハイブリッド特徴量ベクトルを Gradient Boosting Decision Tree (GBDT) 【2】モデルに入力することで構築する。予測の空間単位は 500m メッシュとし、モデル構築の全てのステップでデータの次元と計算負荷の最適化を図る。

2.1 環境特徴量 (NDVI) の抽出と削減

環境特徴量の抽出においては、データ量の削減と予測への寄与度を考慮し、植生指標として、植生の活性度を最もよく示す正規化植生指標 (NDVI) 【3】に特化する。高解像度な Sentinel-2 Level-2A プロダクトをデータソースとし、Google Earth Engine (GEE) 環境下で自動的な前処理パイプラインを構築する。

1. **NDVI 算出:** NDVI は、植物の葉緑素による近赤外域 (NIR) の高反射と赤色光域 (Red) の低吸収という特性を利用し、以下の計算式で算出される。

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

ここで、NIR は近赤外域の反射率、Red は赤色光域の反

射率を示す。この NDVI 値を-1 から 1 の範囲で正規化し、緑の豊かさを示す指標として利用する。

2. **空間集約:** 対象期間 (4 月~11 月) の Sentinel-2 画像に対し、クラウドマスキング処理を適用後 NDVI を算出し、これを予測グリッドと重ね合わせる。各月のグリッド内 NDVI の平均値および標準偏差を算出し、時系列テーブルデータとして格納する。

3. **時系列特徴量の削減:** データ次元を抑えるため、時系列データからクマの生態に最も関連性の高い以下の3つの統計指標に次元を削減する。

- **植生最大豊度:** 活動期全体の NDVI の最大値。
- **植生変化の安定性:** 活動期全体の NDVI の標準偏差。
- **直近の環境状態:** 予測時点の直前 1 ヶ月間の NDVI の平均値。

2.2 実績特徴量の設計と集計

実績特徴量は、緯度・経度および出没日付の情報を含む過去の出没データを基に、予測グリッド単位での**出没密度**として集計する。基礎データセットは、**群馬県が公開しているクマ出没マップ (Esri ArcGIS ベースのダッシュボード) **から解析・構築する。マップからの直接的なデータダウンロードが制限されているため、ブラウザの開発者ツールを用いた API 解析や、特定のパイロットエリアに絞った手動収集・集計などの手法を組み合わせて、出没位置と日付の基礎データセットを構築する。

データ量の削減と予測への有効性を考慮し、以下の2つの指標に絞って集計する。

1. **長期的な出没傾向:** 各 500m メッシュにおける**過去 3 年間の総出没回数**。常習的な出没地域やクマの恒常的な通り道をモデルに認識させるために利用する。
2. **短期的なホットスポット:** 予測時点の**直近 1 ヶ月間**の**出没回数**。直近の異常な行動や一時的な餌資源の変化による出没を捉え、早期警戒の感度を高めるために利用する。

表1 ハイブリッド特徴量のカテゴリと構成要素

カテゴリ	特徴量	役割
環境（動的）	NDVI 最大豊度、NDVI の標準偏差、直近 NDVI 平均	餌資源の季節変化と現況
実績（履歴）	過去3年出沒回数、直近1ヶ月出沒回数	行動と緊急性
地形（静的）	平均標高、標高標準偏差、平均傾斜角、最大傾斜角	移動の制約と生息地の複雑性

2.3 地形特徴量の設計

地形特徴量として、数値標高モデル (DEM) から標高と傾斜角を抽出し、各予測グリッドで空間統計量に集約する。

- **標高:** グリッド内の平均標高 ($Elevation_{mean}$) と、標高の標準偏差 ($Elevation_{std}$) を算出し、標高レベルと起伏の複雑さを同時に表現する。
- **傾斜角:** グリッド内の平均傾斜角 ($Slope_{mean}$) および最大傾斜角 ($Slope_{max}$) を算出し、移動の容易性と峻急な避難場所の有無をモデルに提供する。

2.4 予測モデルの構築と評価手順

統合されたハイブリッド特徴量ベクトルをモデルに入力し、予測モデルを構築する。目的変数 (Y) は、予測対象期間に出没が「あったか (1)」または「なかったか (0)」を示す二値分類ラベルとする。

1. **モデルの選択:** 非線形なデータ構造を効果的に処理できる**XGBoost (Extreme Gradient Boosting)**を採用する。
2. **データ分割:** 収集した過去のデータを、学習データセットとテストデータセットに**時系列で分割**し、未来の予測能力を検証可能なデータ構造とする。
3. **ハイパーパラメータ最適化:** 学習データセットに対して k -fold 交差検証 (5-fold) を実施し、グリッドサーチなどの手法を用いて最適なモデルパラメータを探索し、モデルの汎化性能を最大化する。
4. **評価指標:** モデルの性能評価には、適合率 (Precision)、再現率 (Recall)、そして両者のバランスを示す**F1 スコア**を用いる。また、モデルの識別能力を総合的に示す**曲線下面積 (AUC)**も主要な指標とする。
5. **モデル解釈:** モデルが予測を行う際にどの特徴量が最も重要であったかを定量的に把握するため、**SHAP (SHapley Additive exPlanations)** 値を用いて、ハイブリッド特徴量それぞれの予測への**寄与度**を分析する。

評価に用いる F1 スコアの計算式は、Precision と Recall を用いて次のように定義される。

$$F1 \text{ Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

3 まとめと今後の展望

本研究は、衛星画像 NDVI、出沒実績、地形情報を統合したハイブリッド特徴量に基づくクマ出沒予測モデルの構築手順を確立した。この手順により、従来のモデルでは捉えられなかった環境の動的な変化を予測因子として取り込むことが可能となり、予測精度の飛躍的な向上を目指す。

今後の展望として、以下の技術的課題に取り組む。モデルの汎化性能を高めるため、NDVI 以外の分光指数 (NDWI、EVI) や、気象データ (気温、降水量) を統合したマルチモーダルな特徴量設計を検討する。

参考文献

- [1] 著書名: とことん DevOps 「XGBoost を使って機械学習を始めてみよう」、<https://devops-blog.virtualtech.jp/entry/20250312/1741767926>
- [2] 著書名:Tomokazu Rikioka 「ゼロから始める勾配ブースティング決定木の理論」、<https://zenn.dev/dalab/articles/9c843f0ec8aabf>
- [3] 著書名: 国土地理院 「植生指標データについて」
<https://www.gsi.go.jp/kankyochiri/ndvi.html>

1. 研究背景

近年の金融市場は、急激な価格変動、投機的バブル、集団的過熱感など非線形的で複雑な現象を示す。これらは従来の一般均衡モデルや合理的期待形成モデルでは十分に説明できないと指摘されている。一方、行動経済学や計算社会科学の進展に伴い、個々の主体（エージェント）の異質性や相互作用を考慮するエージェントベースモデル（Agent-Based Model: ABM）が注目されている。

ABM は、マクロの市場挙動が、ミクロの意思決定ルールや心理の変化から創発するという視点に基づく。特に、消費者心理や群集行動が価格形成に与える影響を分析する上で有用である。

2. 研究目的

本研究の目的は、消費者心理が市場価格の安定性に与える影響をシミュレーションを通じて分析することである。特に、楽観・悲観といった心理状態の変動が、どのように市場価格のボラティリティや収束性、バブル形成に寄与するかを検討する。本報告では、研究準備段階としてモデル設計、参照文献整理、用語定義を中心に報告する。

3. エージェントベースモデル（ABM）の概要

ABM は、多数の自律的主体が局所情報に基づき意思決定を行うモデルであり、全体の動態は各主体の相互作用によって創発する。ABM の特徴として以下が挙げられる：

- ・ミクロ行動のルールを明示的に記述可能
- ・非線形相互作用を自然に扱える
- ・不均一なエージェント特性を設定可能
- ・政策介入や制度変更の効果を数値的に検討可能

本研究では、金融市場の消費者をエージェントとして仮定し、期待形成、情報伝播、意思決定、取引行動をモデル化する。

4. 主要専門用語の説明

● 期待形成（Expectation Formation）
経済主体が将来の価格や収益を予測する過程。合理的期待仮説ではなく、適応的期待や社会的影響を含む非合理要素を考慮する。

● ボラティリティ（Volatility）
価格変動の大きさを表す尺度。心理の変化がこの値にどの程度反映されるかが重要となる。

● 群集行動（Herding Behavior）
他者の行動を模倣する現象。市場の加熱や急落を引き起こす要因となり得る。

参考文献

- [1] 人工物工学入門
藤田豊久・太田順 著
(2015)
- [2] 複雑ネットワーク入門
今野紀雄・井手勇介 著
(2008)
- [3] Python コンピュータシミュレーション入門 人文・自然・社会科学の数理モデル
橋本洋志・牧野浩二 著
(2021)

● 価格形成 (Price Formation)

需給バランスに基づいて市場価格が決定されるメカニズム。本研究では、エージェントの売買意思決定から価格変化が生じる設定とする。

5. 研究設計

現状、シミュレーションは未実行であり、以下の要素を設計中である：

- ・エージェント特性（心理パラメータ・期待形成ルール）
- ・市場メカニズム（需要・供給・価格更新式）
- ・心理伝播構造（近傍モデル、ネットワークモデル検討中）
- ・評価指標（価格変動幅、収束速度、極端現象の発生頻度）

モデルは Python および Mesa フレームワークを用いて実装予定であり、今後は逐次的なパラメータ検討と検証手続きを進める。

6. 今後の計画

- ・ Mesa を用いた基本モデル実装
- ・ 心理パラメータ変更による振る舞い比較
- ・ 文献との対応関係検証 (Lux-Marchesi モデル等)
- ・ 結果の視覚化と分析

特に、単純な期待更新ルールから始め、徐々に複雑な社会的影響モデルへ発展させる予定である。

14：AIを活用した落とし物管理アプリの開発

西山研究室 メイマーカイン・飯島 慶士

1. はじめに

近年、学校や公共施設における落とし物・忘れ物の管理は多くの人々にとって身近な課題であり、管理者の負担や誤返却のリスクが問題となっている。特に、従来のシステムでは本人確認が不十分な場合が多く、落とし主のなりすましによる誤受け取りが発生する可能性がある。

本研究では、これらの課題を解決し、より安全かつ効率的に落とし物を管理できるスマートフォンアプリをFlutterで開発することを目的とする。

2. 研究目的

本研究の目的は、AI技術を用いた落とし物と拾得物の自動マッチング機能を備えたアプリケーションの開発である。利用者が紛失物を登録し、拾得者が拾得物を登録することで、両者の情報を照合・共有できる仕組みを構築する。

また、AIによる画像・テキスト情報の類似度分析を行い、落とし主である確率が高い場合に通知を行うシステムを目指す。

3. システム構成

本システムは、Flutterを用いて開発するスマートフォンアプリとして設計している。アプリ内では、一般ユーザと管理者(admin)が同一端末上で利用できるようにした。図1に全体のシステム構成を示す。

まず、登録されたデータはすべてローカルデータベース(SQLite)に保存され、インターネット接続がなくても動作する構成とした。

管理者は専用パスワードでログインし、拾得物の登録を行う。一方、一般ユーザはアカウントを作成し、自身の紛失物を登録できる。管理者は、拾得された物品を受け取り、同様にアプリ上で情報を登録する。管理者専用のログイン機能を実装しており、拾得物情報の入力・編集・削除が可能である。これにより、データの信頼性と一貫性が保たれる。

登録された落とし物・拾得物データはAIマッチングエンジンにより照合される。AIは、画像認識モデルとテキスト類似度解析を組み合わせ、類似度を算出し、類似度が80%以上の場合には、紛失者の登録メールアドレス宛に通知を送信する仕組みとした。

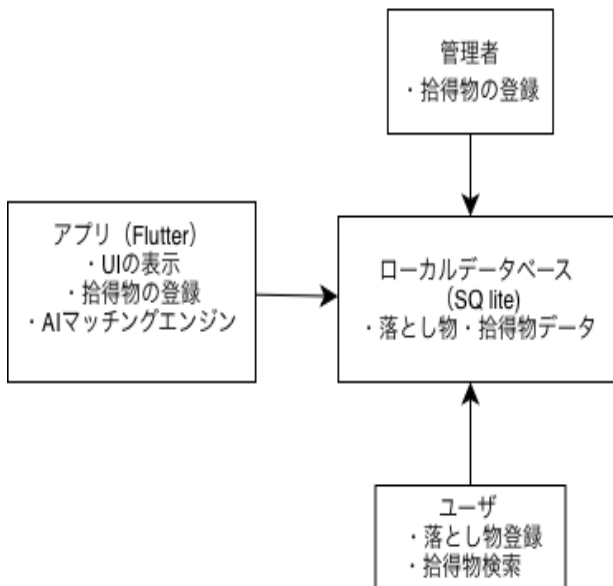


図1：AIを用いた落とし物管理アプリのシステム構成

4. 実装内容と進歩

現在までに以下の機能を実装している。

ホーム画面 UI 設計：登録された落とし物・拾得物の一覧を表示。

ログイン画面の実装：一般ユーザーのアカウント作成および admin ログイン機能を実装。

ローカル DB の準備：データ登録・参照機構を設計中。

今後は、登録情報（タイトル、説明、写真、場所、日付など）の保存機能を整備し、ユーザーが容易に落とし物情報を登録・閲覧できる環境を構築する予定である。

5. 今後の予定

今後の課題として、AIによる落とし物と拾得物のマッチング機能を実装する。画像認識とテキスト解析を併用し、登録データ間の類似度を算出する。類似度が高い場合は落とし主へ通知を行い、迅速な照合と返却を支援する仕組みを構築する。また、UI/UXの改善やデータ整合性の確認を進め、最終的には群馬工業高等専門学校内での実用化を目指す。

参考文献

[1] Flutter公式ドキュメント,
<https://flutter.dev/>

[2] SQLite 公式ドキュメント,
<https://www.sqlite.org/docs.html>

15: マイクアレイとスマートウォッチ連携によるリアルタイム音源定位システムの開発

大墳研究室 西場 真翔

1 はじめに

聴覚障害者は聴力レベルが低く、日常における様々な音を認識することが難しい。特に、踏切や緊急車両のサイレンをはじめとする危険を知らせるための報知音が認識できないことは、重大な事故に繋がる危険性がある。この問題に対しては、2020年にワシントン大学 Dhruv Jain らがスマートウォッチを使用して利用者に報知音等の発生を知らせる Sound Watch[1]を開発・発表した。Jain らの Sound Watch は報知音の発生を検知することに焦点を当てており、音源定位の表示については研究されていなかった。本研究では、マイクアレイとスマートウォッチを連携させることにより、リアルタイムに報知音の発生方向を利用者に示すシステムを開発し、技術的な障壁を考察・提案する。

2 開発手法

2.1 機器連携

本研究ではマイクアレイで受け取った音声情報をもとにPCで処理をし、その結果をスマートウォッチに表示することで利用者に警告することを想定している。今回使用する Pixel Watch (スマートウォッチ) は、スマートフォンとの同時使用が前提として設計されている。また、マイクアレイは4chの入力を受け付けるデバイスで使用する必要があり、一般的なスマートフォンはこれに対応していない。したがって本研究においては、PC、スマートフォン、スマートウォッチの3つの機器の使用が必須であり、これらデバイス同士の通信方法を定める必要がある。図1に本研究におけるシステムの全体構成とデバイス間の通信方式を示す。

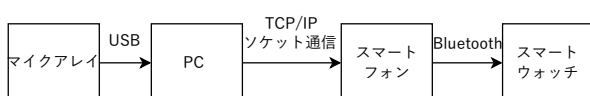


図1: システムの全体構成と通信方式

本システムでは、マイクアレイとPC間をUSB接続、PCとスマートフォン間をTCP/IPソケット通信、スマートフォンとスマートウォッチ間をBluetoothで通信することにより、データの送受信を行う。この通信方式を採用することで、システム全体としてインターネット環境を必要とせず、使用場所を限定しない実用的な動作ができる。

2.2 報知音検出と音源定位

報知音の検出と音源定位の取得を行う音声処理部については、報知音の種類を識別するAIと音源定位を計算するアルゴリズムの組み合わせによって実現する。しかし、単にこれらを組み合わせるだけでは目的の報知音よりも大きな別の環境音が存在する場合、音源定位アルゴリズムが環境音の方向を出力してしまう問題が考えられる。本研究では、AIによる報知音検出と周波数フィルタリングを連携させた特定の音源に対する定位の取得システムを提案する。図2にPCでの音声情報の処理のフローチャートを示す。

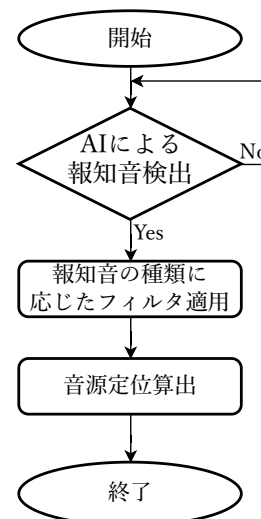


図2: 音声処理のフローチャート

音声処理部の処理の流れとしては、AIによって報知音の種類を特定後、その報知音の主要な周波

数帯のみを取り出すフィルタをかけることでノイズを抑制するというものである。フィルタ済みの音源に対して音源定位の算出アルゴリズムを実行することで、環境音による誤認識を回避することを目標とする。

2.3 ユーザへの警告

報知音検出後のユーザへの警告方法として、スマートウォッチの画面表示による音源定位の可視化とバイブレーションパターンによる音源の概略方向の通知の2つを提案する。スマートウォッチの画面表示方式としては、スマートウォッチ着用者の頭の向きを基準とした報知音の発生方向を、矢印により8方向で知らせるUIを採用する。また、バイブレーションパターンによる通知は、主に画面表示を確認できない状況において、報知音のおおよその発生方向をユーザに知らせることを目的としており、バイブレーションに使用するのは単一の振動モータであるというハードウェア的制約から、4方向の通知を採用する。これら通知方法の認識満足度については、複数の被験者に対してアンケート調査を実施する予定である。

3 予想される技術的障壁

本研究を進めるにあたって、複数の技術的障壁が予想される。まず、警告のリアルタイム性について、本プロセスではマイクアレイから得られた音声から報知音を検出し、その定位を計算するプロセスにおいて数百ミリ秒程度のラグが生じるものと予想される。さらに、本システムは日常での利用を想定しているため、外環境でのAI認識精度やバッテリーの持続性の問題が考えられる。また、ユーザへの警告方法についても、バイブレーションの強弱や、音源定位の可視化方法など、多くのユーザが瞬時に認識できるものを検討する必要がある。今後開発を進めていくにあたっては、これら予想される問題を視野に入れ、それぞれの障壁に対して評価および対策ができるようなアルゴリズムの検討が必要である。

4 接続方式の妥当性の評価

4.1 実験方法

図1で示したシステム構成において、PC→スマートフォン→スマートウォッチ間の通信がリアルタイム警告に適していることを検証するため、PC側の送信時刻とスマートウォッチ側の受信時刻

を記録し、その差を取ることで遅延時間の測定を行う。送信するデータ形式は文字列とする。送信時刻は、PCのpythonプログラム上でtimeモジュールを用いることで、TCP/IPソケット通信が確立されたタイミングでの時刻を記録する。受信時刻は、currentTimeMillisメソッドを使用することでスマートウォッチが文字列を取得したタイミングでの時刻を記録する。時刻の取得形式はUNIXエポックタイムとし、各機器の時刻は事前にNTPにより同期する。

4.2 実験結果

表1に本実験によって得られた通信遅延の測定結果を示す。5回の実行による平均遅延時間は381.8msであった。本システムにおいてこの結果は、リアルタイム性を確保するのに有効な遅延時間であると考えられ、今回提案した通信経路が妥当なものであることが確かめられた。今後は、音声処理部と警告表示部での遅延時間も合算したシステム全体での遅延時間を測定し、評価および遅延の改善を行なっていく必要がある。

表 1: 通信遅延測定結果

実行回数	送信時刻 [ms]	受信時刻 [ms]	通信遅延 [ms]
1	1,761,732,486,853	1,761,732,487,205	352
2	1,761,732,511,793	1,761,732,512,056	263
3	1,761,732,517,854	1,761,732,518,169	315
4	1,761,732,522,443	1,761,732,522,879	436
5	1,761,732,531,007	1,761,732,531,550	543

5 まとめと今後の予定

今回得られた実験結果より、PC、スマートフォン、スマートウォッチの機器連携に要する通信時間が、リアルタイム警告に適していることが確かめられた。今後は、マイクアレイから得られた音声情報をPCで処理する機構の開発を進めていく。それにあたっては、報知音検出に使用するAIモデルの選定や、音源定位アルゴリズムの騒音環境下における精度評価が必要である。

参考文献

- [1] Jain, D., Ngo, H., Patel, P., Goodman, S., Findlater, L., Froehlich, J. E., Sound-Watch: Exploring Smartwatch-based Deep Learning Approaches to Support Sound Awareness for Deaf and Hard of Hearing Users, Proceedings of ASSETS 2020. <https://doi.org/10.1145/3373625.3416991> .

16: パーキンソン病患者の小声症を支援するリアルタイム音声フィードバックシステムの提案

大墳研究室
遠山蓮

1. はじめに

パーキンソン病 (PD)[1] は、脳内のドパミン不足により体の動きが障害される神経変性疾患である。その症状の一つである構音障害の中でも、声が小さくなる「小声症」[1] は、多くの患者に見られる。小声症は、呼吸筋や声帯の筋肉の動きが悪くなることに加え、患者自身が自分の声が小さくなっていることに気づきにくい、という感覚フィードバックの問題を伴うことが特徴である。この感覚の齟齬は、日常生活におけるコミュニケーションの障害となり、患者の QOL を著しく低下させる要因となる。

近年、高齢化社会の進展に伴い PD 患者数は増加傾向 [2] にあり、小声症への対策は重要性を増している。既存のリハビリテーションとしては、言語聴覚士の指導のもとで行う専門的な訓練が存在するが、訓練室以外の様々な場面で応用するのが難しい。

本研究では、患者自身が声量の低下に気づきにくいという問題に着目し、日常生活においてに声量を観測し、適切な声量での発話を習慣化することを支援するツールとして、リアルタイムで音声フィードバックシステムを開発する。

2. 関連研究

パーキンソン病の小声症に対するリハビリテーションとして、最もエビデンスが確立されているものに Lee Silverman Voice Treatment (LSVT LOUD) [1] がある。これは言語聴覚士の指導のもと、「Think LOUD!」をスローガンに、発声努力のレベルを再調整し、患者自身の声量に対する感覚（感覚フィードバック）を改善することを目指す集中的な訓練法であり、日本国内でもその有効性が報告されている。しかし、LSVT LOUD は専門家による高頻度の対面指導を基本とするため、訓練効果を日常生活の様々な場面へ適用させることや、訓練終了後も効果を維持することが課題として指摘されている。

この先行研究を踏まえ、本研究では特に日常生活にお

ける声量の維持を支援することに着目する。リアルタイムの音声フィードバックを提供することで、専門的な訓練場面以外でも日常生活でも患者自身が声量を意識し、適切な発声を習慣化することを支援できるという点で、先行研究とは異なるアプローチをとる。

3. システム

3.1. システム構成

本システムは、話者にヘッドセットマイクを装着してもらい、そこから取得した音声进行分析するアプリケーションである。具体的には、プログラミング言語 Python を用いてマイクから音量を取得・分析し、利用者の声量が事前に設定した基準値を指定した秒数下回った場合に、振動による触覚的なフィードバックを本人に返す。これにより、利用者は会話中に自身の声量の低下を即座に認識し、修正することが可能になる。

本システムは、「音量分析機能」と「フィードバック機能」の二つの主要機能を持つ。システム構成図を図 1 に示す。

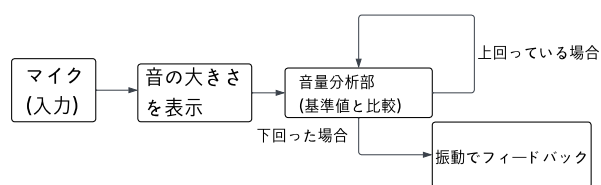


図 1 システム構成図

3.2. 主要機能の実装

3.2.1. 音量分析機能

【実装方法】プログラミング言語 Python の音声処理ライブラリ (Pyaudio や librosa など) を使い、マイク API からリアルタイムで音声データを取得する。取得したデータはデシベル (dB) 値に変換してリアルタイムに処理・表示する仕組みを実装する。

【課題】マイクが利用者の声だけでなく周囲の環境騒音も拾ってしまうため、騒がしい場所では正確な声量測定が困難になる可能性がある。

3.2.2. フィードバック機能

【実装方法】音量分析機能で得られた dB 値が、設定された基準値を下回った場合に、PC のビープ音や（将来的には）振動デバイスによって利用者に通知する仕組みを実装する。

【課題】音声の取得から分析、フィードバックまでに処理遅延が発生し、会話のテンポが遅れる可能性がある。

3.2.3. 基準値の決定

システム構成における小声区域の判断に用いる基準値の決定にあたり、実験参加者 5 名に集団での自然な会話状況を再現するため 30 分間自由に会話してもらおう。その際、主要な分析対象者となる 1 名にヘッドセットマイクを装着してもらい、本人の発話を記録する。記録した 30 分間のデータから、平均値を算出する。算出した平均値を、以降のデータ比較の基礎となる「基準値」として設定する。

3.2.4. システムの検証

まず、システムの検証用として、意図的に声量を落とした「小声」の区間を複数含むテスト音声を作成する。これは、システムが検出するべき低音量状態を、既知のタイミングで発生させるためのものである。

次に、テスト音声を開発した装置に入力する。そして、音声データ上の「小声」のタイミングと、装置から振動によるフィードバックが正しく出力されるかを確認する。これにより、システムが想定通りの動作をするか検証する。

4. 評価

4.1. 評価方法

開発後に測定を行い、システムの有用性や使いやすさを確かめるために以下のような評価を行う。参加者に自身が発話した音声記録（記録方法は後述）を聞き返してもらい、同時に装置がフィードバックを行ったタイミングを提示する。そして、参加者自身が「小声になっていた」と認識しているタイミングと装置から振動によるフィードバックが返ってきたタイミングが一致しているかを評価してもらおう。

4.2. 評価用データの取得

リアルタイムで自然な発声の小声に対してフィードバックを行う状況を作るのは難しいため以下のような測定を行う。

自然な小声の発生を促すため、以下の 2 点の工夫を用いる。

- 研究目的の偽装：

参加者には「パーキンソン病患者の小声検出」という本来の目的は伝えず、「マイクの性能評価のための音声収録」など、異なる目的を伝え、同意を得た上で測定に協力してもらおう。

- 認知的負荷課題の付与：

参加者の意識を発話から逸らすため、思考を要する複雑な作業（例：連続的な計算問題、間違い探しなど）を課す。これにより、意識が課題遂行に集中し、発話への注意が低下することで、自然な小声が出やすい環境を構築する。

以下のような手順で測定を行う

1. 参加者にヘッドセットマイクを装着させる。
2. 上記の認知的負荷課題を 15 分程度行ってもらおう。
3. その際、思考過程や解答を口頭で発し続けてもらい、その発話データをすべて記録する。

5. まとめと今後の予定

本研究は、パーキンソン病患者が自覚しにくい「小声症」を支援するため、リアルタイムで音声フィードバックを行うシステムを開発するものである。今後は基準値を決定するための音声データを収集する。次に、プログラミング言語 Python を使い、PC 上でマイク入力を利用したシステムを実装し、音量分析とフィードバックの基本動作を確認する。PC での動作検証後、スマートフォン端末への移植と最適化を行う。

参考文献

- [1] 西尾正輝 (2014) パーキンソン病に伴うディサースリアの訓練・治療。
- [2] 筑波大学 神経内科（脳神経内科）パーキンソン病の患者数について。

17：足の甲を使用する体表点字システムの検討

大墳研究室 鈴木 煌大

1. はじめに

盲ろう者の情報伝達手段の一つとして、体表への振動で点字を伝える体表点字¹⁾がある。本研究室では耳たぶ、肩、手の甲などを使用した体表点字システムを検討してきたが、一昨年度から新たな部位として足裏を対象としたシステムについて検討してきた。これまでは、足の裏に振動モータをあてて座位で実験を行っていたが、立位では振動を感じづらいことが分かってきた。今回は、立位においても安定して振動刺激を知覚できる部位として足の甲に着目した。その中で先端、中央、側面、足首のうち最も有効な知覚位置を検討する。

2. 体表点字システム

本研究室では体表点字システムに注目し、昨年度から両足の裏の2か所の振動で伝達を行う2点式体表点字システムの研究を行ってきた。通常の点字は6点からなる点字マスで構成されるが、本システムではこの6点を横方向に2点ずつ3段に分割し、1つの点字を3回の振動で表現する。図1に「え」を体表点字で表した例を示す。

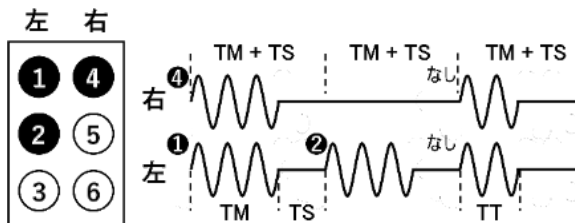


図1: 「え」の振動例

「え」の点字の点の有無の組み合わせは、上段から両方・左・なしとなっている。このような組み合わせで点字マスを2つの振動モータの振動で表現する。振動の継続時間をTM、次の振動までの時間をTS、そして短振動の継続時間をTTとする。

昨年度は、この2点式体表点字システムのより有効な部位を検証するため以下の装置を用いて実験が行われた。図2に実験時の実験装置と測定器具を示す。図2の青矢印が指示しているのは、サンダルに触覚刺激素子として振動モータを取り付けたものである。

振動モータの振動を制御し、振動の読み取りを記録するアプリケーションプログラムは、Raspberry Pi上のPythonを用いて開発された。ログ機能として、測定開始・終了時刻や振動時間を表すパラメーター、実験参加者の回答、正しい答え、そして正誤を記録する。

本研究ではこの装置で実験を行う。

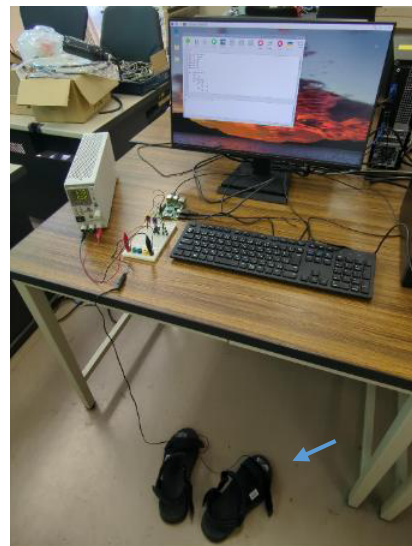


図2: 実験時の実験装置と測定器具

3. 足の甲での振動認識特性に関する実験

一昨年度の研究では足の裏が振動の認識に良いとされていたが、これは座位によって検証されており、立位ではあまり有効でないことが分かった。そこで本研究では、立位において有効な部位として足の甲に着目し、そのうちの部位が最も振動を認識しやすいか調べるため、足の甲の先端、中央、側面、足首の4

か所に振動を提示する実験を行う。振動提示位置を図3に示す。

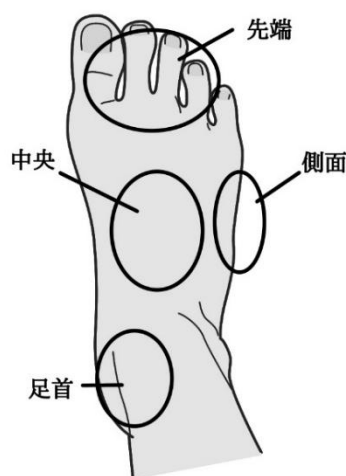


図3 振動位置

参加者には装置を装着してもらい、回答のためのテンキーを渡す。表1に各部位に対応する数字を示す。

表1 部位ごとの対応する数字

部位	対応する数字
つま先	7
中央	4
側面	5
足首	1

回答の際は「7」がつま先、「4」が中央、「5」が側面、そして「1」が足首に対応している。初めに、参加者がテンキーを用いて任意の部位を振動させ、その感覚を確認する機会を設ける。また、本実験では知覚しやすい120Hzの振動²⁾を100ミリ秒振動させる。

振動の確認後、提示された振動がどの位置であるかキーボードを用いて回答するテストを立位にて2回を行う。振動は4か所×各4回の16回提示され、振動する位置の順番はランダムである。振動の提示から被験者の回答までの時間を記録する。また、被験者の回答から次の振動の提示までには、2秒から5秒の間でランダムの間隔があげられる。図4にこの実験のタイミングチャートを示す。

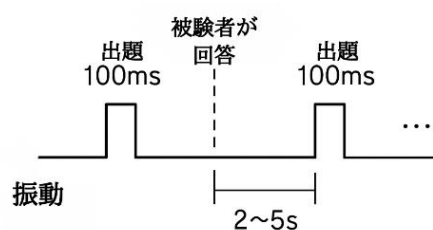


図4 振動提示および回答間のタイミング

テスト終了後、被験者に振動の感じ方について聞き取りを行い、振動位置による認識しやすさの違いなどについて意見を収集する。

4. 現在の状況

本研究は、3年前に行われた足裏部位の実験を発展させたものであり、当時使用していたNintendo Switchに代わって新しい装置を用いる。この変更に伴い、測定プログラムを新しい環境に合わせて再構築している。

現在は過去の研究生が作成した実験装置およびPythonについて学習中であり、実験装置の仕組みおよび昨年度の卒業研究に用いられたソースコード³⁾を理解した。

5. 今後の予定

本研究では人間の振動の感じ方を検証する必要があるため、被験者を募りデータを集めることでより正確な実験結果を目指して進めていく。

参考文献

- 1) 瘡師丈・ほか2名, 新しい体表点字の呈示方法:1点式体表点字の検討 第四報, 第21回日本福祉工学会学術講演会, (2018), pp.109-110.
- 2) Ohtsuka S., A Study of Making Handheld Game Console Vibration Notification Device, 2021 IEEE 10th Global Conference on Consumer Electronics, (2021), pp.30-31.
- 3) 森川遥菜, 足の裏を使用する体表点字システムの検討, 2024年度卒業研究報告書, pp.31-34.

18:大腸検査ロボットのシミュレーション環境の構築

市村研究室 笹川竜誠

1. はじめに

近年のがん死亡の原因部位は、大腸が女性と男性において、それぞれ1位、2位であり、男女合計の場合、2位と非常に多い[1]。大腸がんは早期の発見と治療を行うことで大幅に生存率を上げることができるため、定期的な検査は大変重要である。現在、多くの大腸の検査において、内視鏡が広く用いられる。しかし、胃部と異なり大腸のような複雑な経路で行われる内視鏡検査の場合、医師の高度な操作技術が要求され、医療事故等の危険性も高い。

この問題を解決するため、様々な手法が開発されている。中村らは、空気圧制御の人工筋肉を用いてミミズ等の蠕動運動を模擬する大腸内視鏡搭載型検査ロボットを開発している[2]。また、太田らは、カプセル内視鏡に磁石を内蔵する尾ひれを取り付け、ひれを交流磁場で揺動させることで推進力を得る自走式カプセル内視鏡を開発している[3]。しかし、能動的な機構部を有する大腸検査ロボットは小型化が難しく、その研究開発例は非常に少ない。そこで、岡野らは、小型の電動アクチュエータを採用し、ミミズ等の蠕動運動を機械的に模擬する大腸検査ロボットを試作した[4]。しかし、作業環境が生体内という特殊性から、実際の大腸を用いる実験を頻繁に行うのは難しい。

本研究は、大腸検査ロボットをソフトウェア上で3次元モデルとして再現し、特殊な生体内における動作を確認できるシミュレーション環境の開発を目的としている。

2. ロボットの構成と動作

本ロボットは、肛門から挿入され、蠕動運動で大腸内を自走する。3D CAD Solid works で設計した試作ロボットの外観を図1に示す。本ロボットは、前後に対称な構成となっており、胴体の前部と後部のそれぞれに4本の腕を有する。これらの腕

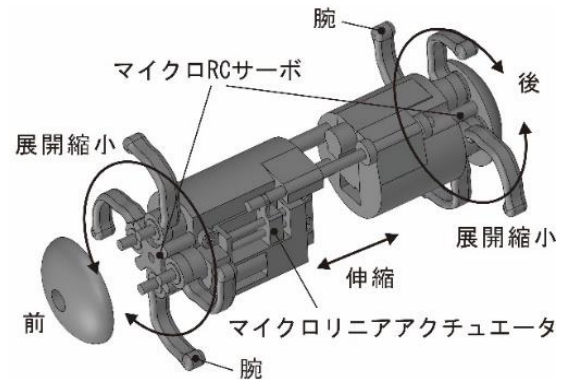


図1. ロボットの外観

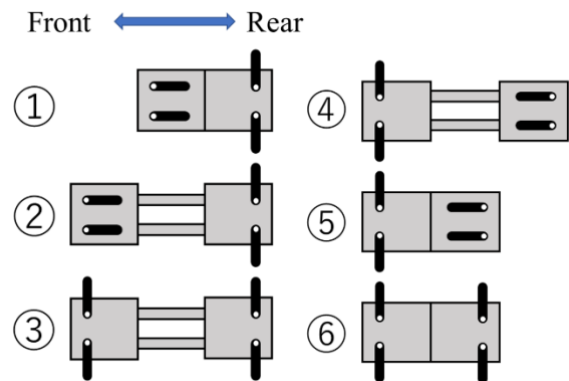


図2. 基本動作①~⑥(1サイクル)

は、マイクロ RC サーボで駆動され展開折畳する。さらに、胴体の前部と後部は、2本の丸棒によるスライド機構(以下、前後連結部)を用いて前後方向に沿って動くように拘束されており、マイクロリニアアクチュエータでロボットの全長を伸縮させることが可能である。また、腕の展開と折畳における動作円の直径は、それぞれ20mm、40mmである。また、前後連結部の収縮と伸長におけるロボットの全長は、それぞれ69mm、76mmである。

本ロボットは、以下の①~⑥の基本動作を順次実行し、これを1サイクルとして前後進する。試作

ロボットにおける各基本動作を図 2 に示す。

- ①前部の腕を折畳し腸壁から開放
- ②前後連結部を伸長し前部を前方へ移動
- ③前部の腕を展開し腸壁に固定
- ④後部の腕を折畳し腸壁から開放
- ⑤前後連結部を収縮し後部を前方へ移動
- ⑥後部の腕を展開し腸壁に固定
- ⑦「①」へ戻る

3. シミュレーション環境

3.1. 物理演算エンジン

本実験は、大腸壁面の軟体を表現するため、ビデオゲームエンジンの Unreal Engine5 を用いる。Unreal Engine 5 は、高精度な物理シミュレーションが可能な Chaos 物理演算エンジンを搭載しており、剛体、衝突、関節運動、布や軟体、破壊表現など多様なシミュレーションをリアルタイムで再現できる点が特徴である。

3.2. 大腸検査ロボットのモデルの作成

大腸検査ロボットの物理シミュレーションを行うため、まずロボットの 3D モデルの作成を行った。3D モデルの作成のため、無償で利用可能な 3DCG 作成ソフトウェアの Blender を用いた。Blender は 3D モデリング、アニメーション、物理演算、レンダリングなどを統合的に行えるオープンソースツールである。作成したロボットのモデルを図 3 に示す。

3.3. シミュレーションの構築

本シミュレーションを、以下の①～④の手順で構築する予定である。

- ① 作成したロボットのモデルを Unreal Engine5 へインポート
- ② 各パーツの接続設定
- ③ 腕の展開と胴体の伸縮をキーボード操作と関連付け
- ④ 物理演算を適応し、生体の腸内壁を模擬したフィールドでシミュレーションを実施

現在の進捗状況は、②を作成中である。

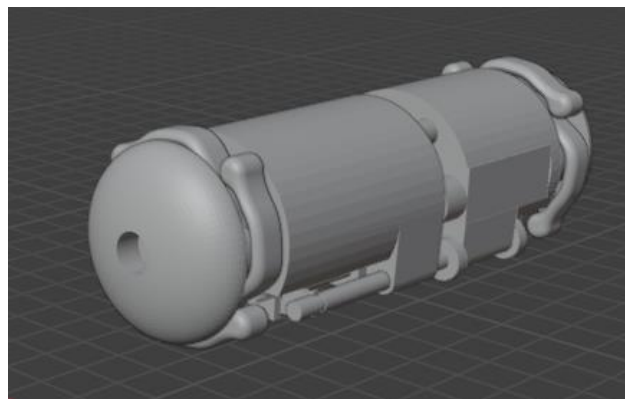


図 3 大腸検査ロボットの 3D モデル

①について、Blender で作成したデータを fbx 形式でエクスポートし、これを Unreal Engine5 へインポートした。②について、Unreal Engine5 の Physics Constraint を用いた。腕と胴体、前部と後部をコンストレイントアクタで接続し、腕と胴体間はロボットのロール方向のみ、前後部間はロボットの前後方向のみに動かせるよう制限をつけた。コンストレイントとは、ジョイントの位置や角度に制限をつけることである。コンストレイントアクタは、直動モータと回転モータの 2 種類のコンポーネントがあり、モータの速度を制御できる。

4. 今後の予定

ロボットの各パーツの接続設定、および各動作とキーボード操作の関連付けを行う。さらに、生体の腸内壁を模擬したフィールド作成し、シミュレーション環境を構築する。本シミュレータを用いて、大腸検査ロボットの動作について物理シミュレーションを行う。また、Unreal Engine5 で大腸壁面とロボット間の相互干渉をどの程度の精度で模擬できるか検証を行う。その結果、精度の低い場合は、他のシミュレーション手法への見直しの検討を行う。

参考文献

- [1] 国立がん研究センターがん情報サービス「がん統計」(厚生労働省人口動態統計), https://ganjoho.jp/reg_stat/statistics/stat/summary.html (参照日:2024/10/29)
- [2] 中村太郎:“空気圧システムを用いた大腸内視鏡挿入デバイスの開発”, 計測と制御, vol. 56, no. 4, pp. 281-285, 2017.
- [3] 太田和寛, 能田貞治, 竹内利寿, 進藤 康則, 大塚 尚武, 樋口和秀:“自走式カプセル内視鏡の開発-現状と今後の課題”, 臨床消化器内科, vol. 30, no. 13, pp. 1675-1679, 2015.
- [4] 岡野泰一郎, 小堀洋史, 荒川達也, 市村 智康:“電動アクチュエータを用いる蠕動運動型大腸検査ロボットの提案”, 日本機械学会ロボティクス・メカトロニクス講演会 2023 講演論文集, 2P2-C07(1-3), 2023.

19: エッジ AI デバイスを用いるオフライン対話型ロボットの試作

市村研究室 中澤啓人

1. はじめに

近年、AI を用いる製品が数多く流通し、誰でも容易に AI を活用できる環境が整いつつある。特に、人間とコミュニケーションを取れる AI は注目を集めており、大規模言語モデル(以下、LLM)を用いる OpenAI 社の ChatGPT や Alphabet 社の Gemini 等はスマートフォンでも利用が可能のため、日常で使用される機会が非常に増えている。AI の発展に伴い、音声対話ができるロボットも増えており、代表的なものに図 1 に示す「ポケとも」[1]や「Romi」[2]がある。

しかし、これらロボットの返答内容の生成はクラウド上の LLM に依存しているため、オンライン環境以外で対話を行うことができず、利用場所は限定される。また、費用面で API の利用にかかるコストを考慮しなければならず、セキュリティ面でプライバシーに配慮した会話をする必要がある。本研究は、エッジ AI デバイスを用いてオフライン環境で対話が可能なロボットを試作する。また、カメラによる顔追尾機能も実装し、人との対話をより自然にできないか検証する。

2. Jetson Orin Nano Super

Jetson Orin Nano Super[3](以下、Jetson)は、NVIDIA 社が開発した小型エッジ AI デバイス[4]向けの計算機である。外観を図 2 に示す。特に AI を用いるための GPU 性能を重視しているため LLM や VLM を用いることに適している。大きさは 103mm×90.5mm×34.77mm であり、小型ロボットに搭載するのが可能なサイズとなっている。本研究では、本計算機を用いてシステムを構成する。



図 1: ポケとも(左)と Romi(右)



図 2. Jetson Orin Nano Super

3. システムの構成

本研究で試作するロボットは、以下の動作の実現を目的としている。動作の流れを図 3 に示す。

- ・マイクを用いた文字起こしシステム
- ・ローカル LLM による対話生成システム
- ・音声による返答出力システム
- ・カメラを用いる顔追尾システム
- ・ROS を用いた首振りシステム

以上の機能を全てオフライン環境で実現できるように実装する。対話にはローカル LLM を用いる。なお、各システムを Docker 上のコンテナで作成し、それぞれを独立に管理する。

3.1. ローカル LLM

本研究では MLC LLM コンテナを使用し、対話内容を生成するモデルには TinyLlama, または Alphabet 社の Gemma-2B(以下、Gemma)の 2 つで検証し、より対話に適するものを採用する。使用するパラメータは TinyLlama が 11 億、

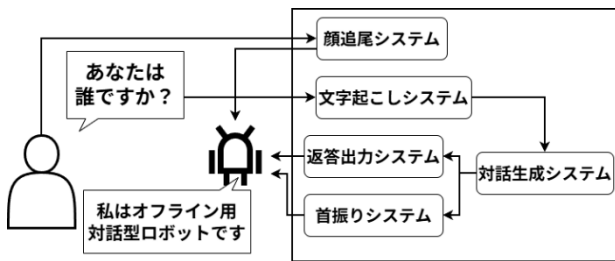


図 3. オフライン対話型ロボットの動作

Gemma は 20 億であるため非常に小さいモデルである。そのため、対話以外に実装を想定している顔の追尾や首振り機能に対してメモリ等のリソースを確保しやすい。

4. 各機能の実装状況

4.1. 対話生成システム

現在使用している Jetson のメモリは 8GB でありリソース不足が懸念されるため、OS を CUI 環境に設定し Gemma の実装を試みた。しかし、実装途中で GPU 用のメモリ（以下、VRAM）が上限に達してしまった。また、SSD 領域の 1 部を仮想メモリとして 8GB 確保して試みたが、VRAM としては利用されず、Jetson が再起動した。そのため、Gemma を用いる対話機能の実装を断念した。次に、TinyLlama を用いてコマンドプロンプト上で対話をするシステムを作成した。TinyLlama を用いる場合、対話を行えるが同じ質問で複数回の対話を行うと答えが一通りになる問題が生じた。例えば、「what is the date?」の回答は、日付については答えず、「I'm happy to provide the most recent information on the date」と、質問と噛み合わない回答を繰り返してしまう。また、単純な四則演算でも回答が不正確であり、人と対話を行う上で十分な性能を有しないのが分かった。

4.2. 顔追尾システム

カメラを用いるリアルタイムシステムを実装する前に、保存済みの動画において物体を検出するプログラムを作成した。本プログラムは、新たなデータにアノテーションを付与できる。実際に犬の画像にアノテーションを付与したモ



図 4. アノテーションを付与したモデルによる推定結果

デルを用いて動画内の犬を検出した結果を図 4 に示す。これは動画を秒間 1/60 間隔で区切り、切り取った各フレームにアノテーションの付与したものであり、動画内に入っている犬を認識し、「dog」というラベルの付与ができています。

5. まとめと今後の課題

現在のシステムではローカル LLM の実装のみでメモリ・リソースが不足する。また、TinyLlama では対話において快適さを得ることができない。今回、Jetson をシステムの計算機として採用したが、このようなエッジデバイスでは VRAM の増設を行えず、AI モデルを扱うのは容易でない。今後のシステム開発においても、限定されたメモリで如何に各システムを実装できるかが課題となる。具体的に、Jetson で更なるメモリの調整を行う、または 8GB 以上の VRAM を搭載する新たな Jetson を採用し、Gemma による対話生成システムを実装できないか検証する。対話生成システムを実装後、文字起こしシステムおよび返答出力システムの作成を行う。

参考文献

- [1] ポケとも, <https://poketomo.com/>, (2025/10/29 参照)
- [2] Romi, <https://romi.ai/>, (2025/10/27 参照)
- [3] Jetson Orin Nano Super 開発者キット, <https://www.nvidia.com/ja-jp/autonomous-machines/embedded-systems/jetson-orin/nano-super-developer-kit/>, (2025, 10/31 参照)
- [4] 白石洋一: ”エッジ AI デバイスの最新動向” エレクトロニクス実装学会誌 vol23 No20, 特集/インテリジェント実装技術の最新動向, 2020.

20：群馬県全域におけるからっ風の風力へ山岳地形がもたらす影響に関する検討

雑賀研究室 後藤天慈 齋藤優月

1. はじめに

日本列島は、山脈と平野が複雑に組み合わさった地形を有しており、この特殊な地形構造が原因となり数多くの気象現象が発生している。日本各地では、その地方特有の局地風が知られている。具体的には、やまじ風、やませ、広戸風、からっ風などが挙げられる。このうち、からっ風は、主に関東地方や東北地方で冬期に吹く乾燥した北西から吹く風のことを指す。からっ風に関して、気象学的観点からの研究が重ねられてきたが、「からっ風」の定義に関しては、研究者や論文によって多少の相違がみられる。

2. 関連研究と本研究の目的

西と日下[1]による研究では、からっ風がおろし風の一種であることを示し、群馬県北部の三国山脈の影響により一般のおろし風よりも山から平野に吹き下ろす性質を持つという説を提唱した。昨年度の新井[2]の卒業研究では、群馬県各地の風力分布・風向分布のデータと観測地点の地形からからっ風への山岳の影響を検討した。なお、風力の定義は気象庁の定義に従っている。

本卒業研究では、風力分布と風向分布のうち、風力分布について、新井の卒業研究において利用した観測地点を含む群馬県の全ての観測地点において風力分布を調査して、からっ風への山岳の影響を検討することを目的とする。

3. からっ風の定義

からっ風と呼ばれるための要件には、研究者や論文により多少の差異があるが、本卒業研究では、西と日下の研究による以下のからっ風の要件を用いる。

- ① 冬型の気圧配置で風向が北から西
- ② 平均風速が 8.0[m/s]以上
- ③ 相対湿度が 40%未満
- ④ 吹走期間が 11 月 1 日から 3 月 31 日までの期間

4. 過去のからっ風の吹走状況の調査

草津(北緯 36 度 37.0 分,東経 138 度 35.5 分,標高 1223m)、西野牧(北緯 36 度 14.7 分,東経 138 度 42.4 分,標高 375m)、藤原(北緯 36 度 51.8 分,東経 139 度 03.5 分,標高 700m)、田代(北緯 36 度 27.8 分,東経 138 度 27.8 分)、の 4 地点について調査を行った。八木と野口の卒業研究[3]によって抽出された 2009~2023 年までの期間の前橋市の、3 節の要件を満たすからっ風の吹走日を参考に、気象庁の過去の気象データ[4]を用いた。

一例として、以下に前橋市のからっ風の吹走日における前橋市、草津市、および田代の風力分布を示す。

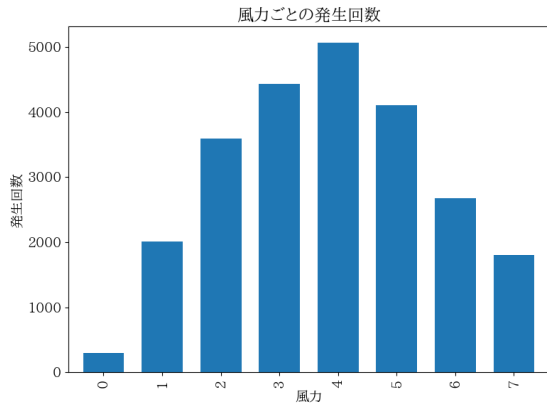


図 1: 前橋市の風力分布

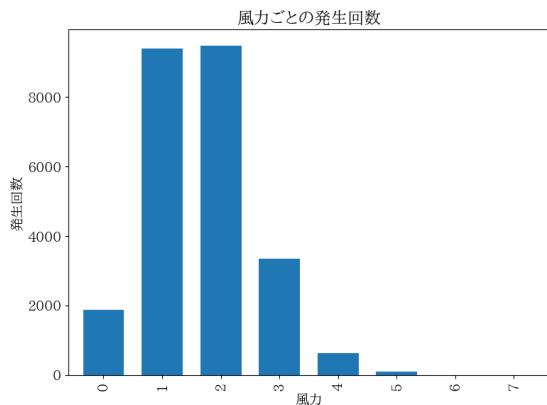


図 2: 草津市の風力分布

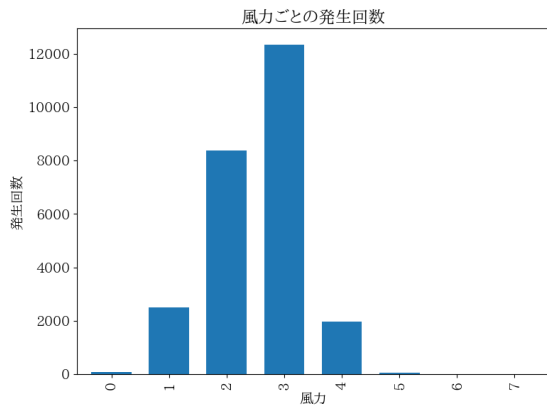


図 3: 田代の風力分布

グラフから、草津・田代は風力の小さいものが多いが、前橋は風力の大きいものが多いことが読み取れる。これは、草津・田代が山岳に近い場所に位置しているため風力が小さく、前橋は平野に位置しているため、風力が大きくなっていると考え

られる。西と日下による研究では、凸性を有する山脈および平坦部(半盆地)を持つ簡単な地形に対する Weather Research and Forecasting(WRF)モデルを用いたシミュレーションより、半盆地領域において特に強い風が発生するという結果と、上記の3つの観測地点における風力分布とおおむね一致していることを確認した。

5. まとめと今後の課題

これまで、藤原、西野牧、草津、田代の4地点について風力分布の調査を行い、その特徴について検討した。

今後は、他の観測地点での調査を行い、そこからからっ風へ山岳地形が与える影響について検証することを予定している。

参考文献

- [1]A.Nishi and H.Kusaka : "Effect of Mountain Convexity on the Locally Strong "Karakaze" Wind", Journal of the Meteorological Society of Japan. Ser. II, Vol.15, pp.787-803, 2019
- [2] 新井 堯太, 令和 6 年度卒業論文, 群馬工業高等専門学校, 2025
- [3]八木聡太, 野口晃宏, 令和 5 年度卒業論文, 群馬工業高等専門学校, 2024
- [4]気象庁「過去のデータ検索」
(<https://www.data.jma.go.jp/stats/etrn/index.php>)

21：群馬県全域におけるからっ風の風向へ山岳地形がもたらす影響に関する検討

雑賀研究室 阿久澤聖蒼

1. 背景と目的

群馬県では冬季になると乾燥した北西の強風、いわゆる「からっ風」が発生し、地域の気候的特徴の一つとなっている。また、凸性を有する山脈および半盆地状の平坦な地形からなるシンプルなモデルに対する WRF モデルを用いた大規模数値シミュレーション[1]から、群馬県南西部に広がる平坦な地形において、強風が吹くとの予測が得られている。また、新井[2]によって、群馬県のいくつかの観測地点における風力分布および風向分布の気象データ[3]から、群馬県南西部にある平坦な地形における観測地点において北西からの強風が吹くこと、みなかみなどの山岳部にある観測地点では谷あいを吹くギャップ風が存在することが明らかにされた。

本卒業研究では、群馬県に存在する気象庁の全ての観測地点において、風向分布を調査することで、群馬県における山岳地形がからっ風におよぼす影響を明らかにすることを目的とする。

2. からっ風の定義

気象学的な観点からからっ風と呼ばれるための要件は論文により多少の差異がみられるが、本研究では西と日下の研究による以下の要件を用いる。

- (1) 冬型の気圧配置で風向が北～西からの強風が吹く。

- (2) 平均風速が 8.0[m/s]以上である。
- (3) 相対湿度が 40%未満である。
- (4) 吹走期間が 11 月 1 日から 3 月 31 日までの期間である。

3. 局地風について

局地風には、気流が山を越える時にその麓や斜面で吹く風をおろし風（図 1）、谷の出口で吹く風をギャップ風（図 2）という。からっ風は一般的におろし風であるといわれている。



図 1 おろし風のイメージ図[2]



図 2 ギャップ風のイメージ図[2]

22：ホップフィールドモデルにおける連想記憶の性能向上に関する検討

雑賀研究室 渡邊 祥

1 はじめに

近年、生成 AI をはじめとした人工知能が急速に発展し、様々な分野・場所で活躍している。それらの基礎となるのが、ニューラルネットワークである。

ニューラルネットワークとは、生物の脳に存在する神経系をコンピュータで再現したものである。生物の神経系では、多数のニューロンが結合して働いている [1]。ニューロンはそれぞれ電気信号を発生することができ、別のニューロンから受け取った電気信号をもとに、また別のニューロンに電気信号を送ることで脳が成り立っている。このニューロンや、その結合をデータで再現すれば、コンピュータでも生物の脳を再現できるという考えでニューラルネットワークは生まれた。

ホップフィールドモデルとは、ニューラルネットワークの一種である。このモデルは、ノイズがかかったパターンが入力されたとき、はじめに記憶したパターンのうち近いものを想起するネットワークである。

人間が無制限に物事を記憶できないのと同様に、コンピュータによって再現したニューラルネットワークにおいても無制限に物事を記憶することはできない。また、ホップフィールドモデルに記憶させたパターンによっては、少ないパターン数でも想起に失敗する可能性がある。今回の研究では、これらの問題を軽減し、ホップフィールドモデルの性能向上を図ることが目的である。

2 ホップフィールドモデルについて

ホップフィールドモデル [2] では、記憶したいパターンのビット数と同じ数のニューラルユニットを用意する。今回は、 K ビットのパターンを覚えさせるものとする。各ユニットは、自分自身以外のユニットすべてと結合を持つものとする。ニューラルユニットは、 ± 1 の値をとる。ネットワークに記憶させたい M 個のパターンのうち

m ($m = 1, 2, \dots, M$) 番目を、

$$\mathbf{u}^{(m)} = \begin{bmatrix} u_1^{(m)} \\ u_2^{(m)} \\ \vdots \\ u_K^{(m)} \end{bmatrix}$$

とする。また、 i 番目のユニットから j 番目のユニットへの重み係数 w_{ij} を ij 要素として持つ重み係数行列 \mathbf{W} を以下のように定める。

$$\mathbf{W} = \sum_{m=1}^M (\mathbf{u}^{(m)} \mathbf{u}^{(m)T} - \mathbf{I})$$

ホップフィールドモデルでは、この重み係数行列をもとに状態更新を行うことによって、少しずつ目的のパターンに近づいていく。状態更新は、以下の式に示す方法で更新する。

$$\begin{aligned} \mathbf{v}[n+1] &= \mathbf{W}\mathbf{u}[n] \\ \mathbf{u}[n+1] &= \frac{\mathbf{v}[n+1]}{|\mathbf{v}[n+1]|} \end{aligned}$$

ここで、 $\mathbf{u}[n]$ は、時点 n におけるパターンを表す。

ホップフィールドモデルは、以下の式に示すエネルギー関数が定義されている。ホップフィールドモデルは、エネルギーが極小値をとるように動作する。

$$E_n = -\mathbf{u}^{(n)T} \mathbf{W} \mathbf{u}^{(n)}$$

ここで、 $\mathbf{u}^{(n)}$ は n 番目のパターンを示す。

3 何も手を加えない場合のシミュレーション

今回、以下のような条件において、ホップフィールドモデルによる連想記憶再現するプログラムを作成し、想起がどのように収束するかシミュレーションを行った。

- 記憶させるパターンは、 16×16 のグリッド、つまり 256 ビットのパターンとする。
- パターンのそれぞれのビットは、半分の確率で $+1$ 、残りの半分の確率で -1 になるように生成する。

- 記憶させるパターンは、10 個、30 個、50 個で試す。
- 初期パターンと目的のパターンの内積を変化させながら、想起が成功するかどうかを調べる。

4 シミュレーションの結果

図 1 に、10 個のパターンを記憶させた時の、初期パターンと目的のパターンの内積の変化を示す。

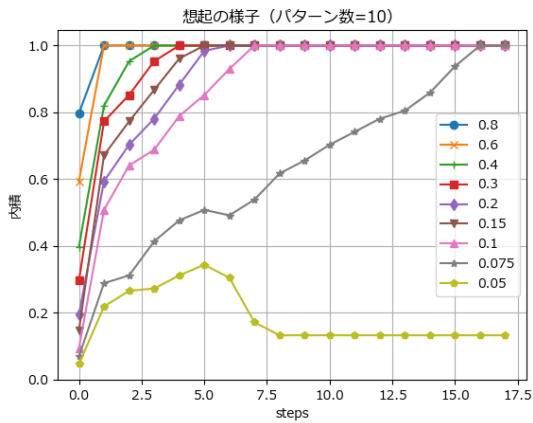


図 1 記憶させたパターンが 10 個のときにおける内積の変化

図 2 に、30 個のパターンを記憶させた時の、初期パターンと目的のパターンの内積の変化を示す。

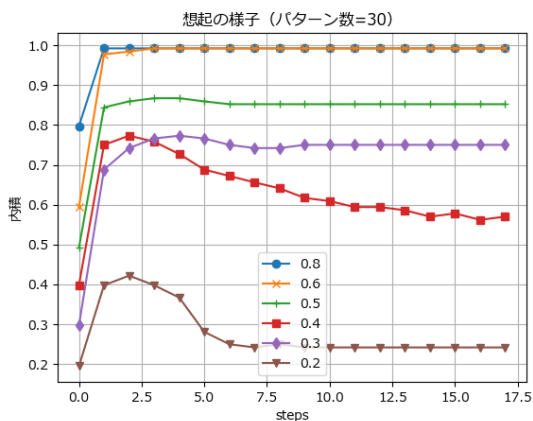


図 2 記憶させたパターンが 30 個のときにおける内積の変化

図 3 に、50 個のパターンを記憶させた時の、初期パターンと目的のパターンの内積の変化を示す。

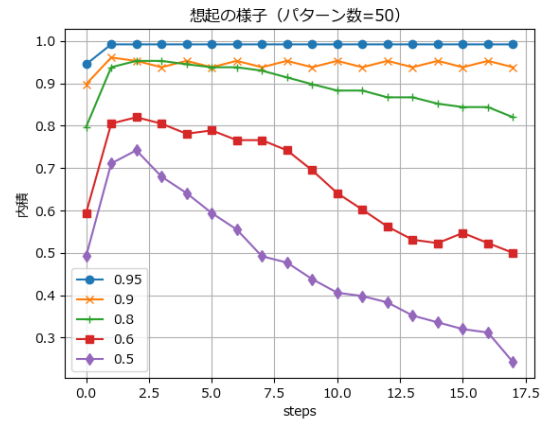


図 3 記憶させたパターンが 50 個のときにおける内積の変化

図から分かる通り、初期パターンと記憶させたパターンの内積が大きいと想起が成功しやすくなる。また、同じ内積でも、記憶させたパターン数が少ない場合は想起が成功しやすい。

5 今後の課題

今後は、過去の研究や文献をもとに、ホップフィールドモデルについての理解を深め、ホップフィールドモデルの性能向上のために、状態更新における入出力関数を調整するなどの工夫を施し、それについて検討する。また、このホップフィールドモデルの研究が最新の人工知能に活かせることについて考える。

参考文献

- [1] 中野 馨 編著, 阪口 豊, 森田 昌彦, 新妻 素直, 脇本 良則, 田中 宏一良, 藤田 啓介, 衛 作人 共著. ニューロコンピュータの基礎. コロナ社, 1990 年, pp.37-49
- [2] 八名 和夫, 鈴木 義武 共著. ニューロ情報処理技術—基礎と応用—. 海文堂, 1992 年, pp.53-57

23：ダイレクトサンプリング方式 SDR 受信機における アンチエイリアス BPF に関する検討

築地研究室 秋山 珀空

1. はじめに

ソフトウェア無線機 (SDR: Software Defined Radio) は、ハードウェア構成を再構成可能な FPGA などにて信号処理を実現することにより、柔軟な無線通信システムを構築できる方式である。FPGA は内部の論理回路をソフトウェアで書き換えて再構成できるハードウェアデバイスであり、これにより通信方式や処理内容を動的に変更できる点が広義の SDR に含まれる。本研究では、ダイレクトサンプリング方式を採用し、アナログフロントエンド部でのエイリアシング防止を目的としてアンチエイリアスバンドパスフィルタ (BPF: Band Pass Filter) の設計および特性評価を行う。特に、FM 帯域 (76.1~94.9 MHz) を対象に、バターワース型 BPF の回路シミュレーションを実施した結果を報告する。

2. ダイレクトサンプリング方式の課題

従来の受信機では、局部発振器とアナログミキサを用いて受信信号を中間周波数 (IF: Intermediate Frequency) に変換した後、信号の復調処理を行うスーパーヘテロダイン方式が一般的である。これに対し、図 1 に示すダイレクトサンプリング方式では、アンテナから受信した高周波信号を直接 AD コンバータ (ADC: Analog to Digital Converter) に入力し、デジタル化後に復調・フィルタ処理を行う。この方式ではアナログミキサや IF 段が不要になり、構成が単純化されるという利点がある。一方で、ADC には高いサンプリング周波数と分解能が要求され、折り返し雑音 (エイリアシング) の防止が課題となる。

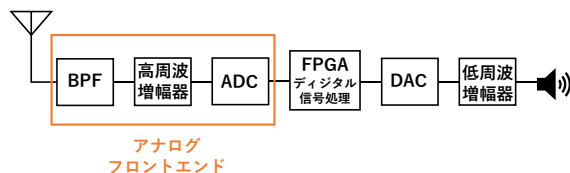


図 1 ダイレクトサンプリング方式

AD 変換では、入力信号の周波数成分がサンプリング周波数 f_s の $1/2$ (ナイキスト周波数) を超えると、折り返して $f_s - f$ の位置に現れ、図 2 のようにスペクトルの重複を引き起こす。これにより復調後の信号に歪

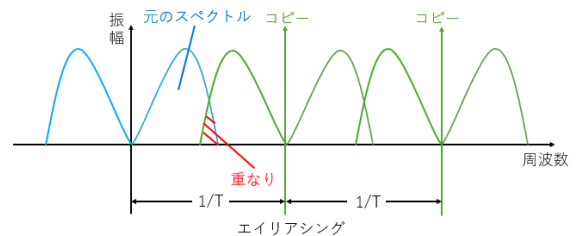


図 2 エイリアシング

みやノイズが生じる。この問題を防ぐため、AD 変換前にアンチエイリアスフィルタを配置し、不要な高周波成分を除去する必要がある。ダイレクトサンプリング方式 SDR では、アンテナと ADC の間に BPF を挿入し、目的帯域 (本研究では FM 帯) のみを通させる。本研究では、通過帯域内で平坦な利得特性をもつバターワース型 BPF を設計対象とし、LTspice を用いてその振幅特性および群遅延特性を解析した。

3. 代表的なアナログフィルタの特性

アンチエイリアス BPF には、通過帯域特性や群遅延特性の異なる複数の設計法が存在する。代表的な 3 種類を以下に示す [2]。

(1) バターワース型

- 特徴：通過帯域の振幅特性が最も平坦
- 利点：リップルがなく、減衰量が制御しやすい
- 欠点：減衰の傾きが緩やか
- 用途：一般的な通信や音響など幅広い

(2) ベッセル型

- 特徴：群遅延が最も一定
- 利点：過度応答性 (立ち上がり/立ち下がり) が良い
- 欠点：減衰の傾きが最も緩やか
- 用途：音声・映像信号処理など波形を重視するもの

(3) チェビシェフ型

- 特徴：通過帯域または阻止帯域にリップルを持つ
- 利点：減衰の傾きが急峻
- 欠点：通過帯域と阻止帯域の両方にリップルがある
- 用途：狭帯域フィルタ、選択度が重要な通信機器等

これらの特徴を比較すると、バターワース型は一般的で設計が容易な折衷型、ベッセル型は波形忠実度重

視、チェビシェフ型は遮断特性重視と位置付けられる。本研究では、まずバターワース型を対象にLTspiceで特性解析を行う。

4. フィルタ設計

4.1. 設計仕様

本研究では、FM放送帯域を通過帯域とし、通過帯域端を76.1MHzおよび94.9MHzとした。アンテナおよび回路の特性インピーダンスは50Ωとした。

4.2. フィルタ形式の選定

アナログBPFの設計方法には、上で説明した三種類が代表的だが本発表では、まず通過帯域内の振幅特性が平坦であり、回路設計が容易な3次バターワース型BPFの設計・評価を行う。

4.3. 素子値の決定

T入力型のN次バターワース型BPFのLC直列回路の素子値は以下の式(1),(2)を用い、LC並列回路の素子値は式(3),(4)を用いて求められる[2]。

$$L_n = \frac{Z}{2\pi\Delta f} \cdot g(n) \quad [H] \quad (1)$$

$$C_n = \frac{1}{Z \cdot (2\pi f_0)^2} \cdot \frac{1}{g(n)} \quad [F] \quad (2)$$

$$L_n = \frac{Z \cdot 2\pi\Delta f}{(2\pi f_0)^2} \cdot \frac{1}{g(n)} \quad [H] \quad (3)$$

$$C_n = \frac{1}{Z \cdot 2\pi\Delta f} \cdot g(n) \quad [F] \quad (4)$$

ここでg(n)は正規化したバターワース・フィルタの素子値を示している。これを式(5)に示す[2]。

$$g(n) = 2 \sin \left[\frac{(2n-1)\pi}{2N} \right] \quad n=1,2,\dots,N(Nは次数) \quad (5)$$

4.4. LTspiceによる回路構成

4.3節で示した式を用いて設計した3次バターワース型BPFを図3に示す。

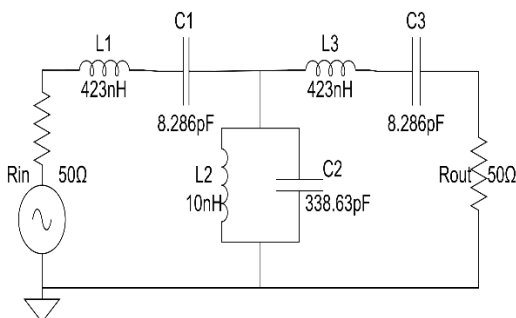


図3 シミュレーション回路

5. シミュレーション結果

5.1. 振幅特性

図4に示す振幅特性より、-6dBでの平坦性は見られるが通過帯域が79MHzから93MHzと想定より狭くなっていることが確認できる。

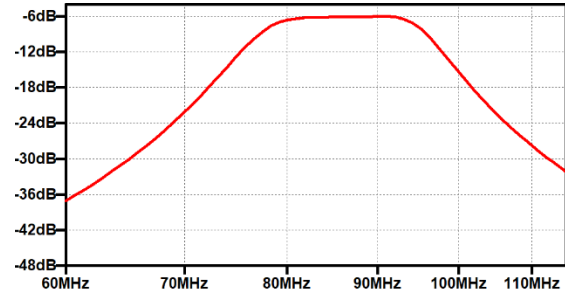


図4 アンチエイリアス BPF の振幅特性

5.2. 群遅延特性

図5に示す群遅延特性より、通過帯域中央付近では群遅延特性の平坦性を確保しているものの、帯域端付近で平坦性が失われていることが確認できる。

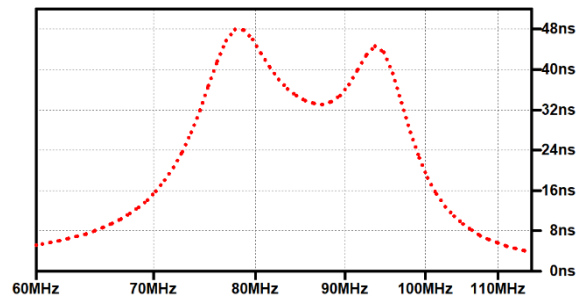


図5 アンチエイリアス BPF の群遅延特性

6. まとめ

振幅特性における平坦性は確認できたが、想定のとおり通過帯域より狭いことが確認できた。群遅延特性では実際の帯域端付近で遅延が増加するというバターワース型の特徴が見られた。

今後は、バターワース型フィルタの再設計また、バツセル型フィルタを新たに設計し、振幅特性・群遅延特性を比較する。また、FPGA上での実装を見据え、デジタル領域でのフィルタ特性との整合性についても検討する予定である。

参考文献

- [1] 太郎丸 真 / 阪口 啓:「SDR:ソフトウェアで作る無線機的设计法」,科学情報出版株式会社,pp.156-159(2016年).
- [2] 森 栄二:「LCフィルタの設計&制作」,CQ出版株式会社,pp.145-176(2007年).

24：電流モード制御 DC-DC コンバータにおける電流帰還回路モデルに関する検討

築地研究室 樋口 晃太

1. はじめに

電流モード制御方式を用いた DC-DC コンバータは、高速な負荷応答特性および電流制御の安定性により広く利用されている。しかしながら、電流帰還経路の多くは集積回路内部に実装され、外部から内部構成や帰還特性を直接把握することは困難である。したがって、この制約下で実機の挙動をシミュレーション回路として再現するためには、まず、外部応答に基づいて帰還係数や補償特性といった制御パラメータを推定する手法が必要となる。さらに、推定したパラメータをそのまま受け入れ、実機の動作を等価的に再現できる回路モデルが不可欠である。本研究では、電流モード制御 DC-DC コンバータにおける電流帰還回路を制御電源（従属電源）により抽象化し、外部応答から推定される制御パラメータを独立に与えるだけで実機特性を再現可能な電流帰還等価回路モデルの構築を検討する。

2. 電流モード制御方式の原理

図 1 に電流モード制御を用いた降圧形 DC-DC コンバータの全体回路を示す。電流モード制御方式では、回路のインダクタ電流 i_L を電圧に変換した後、スロープ補償回路により傾きが補正される。そして、出力電圧と基準電圧の誤差を増幅したエラーアンプ出力との比較により、スイッチング波の時比率 D が決定する。図 2 に電流モード制御動作のタイミングチャートを示す。出力電圧はスイッチング波を LC フィルタで平均化することで得られる。したがって、出力電圧 V_o は次式によって決定する。

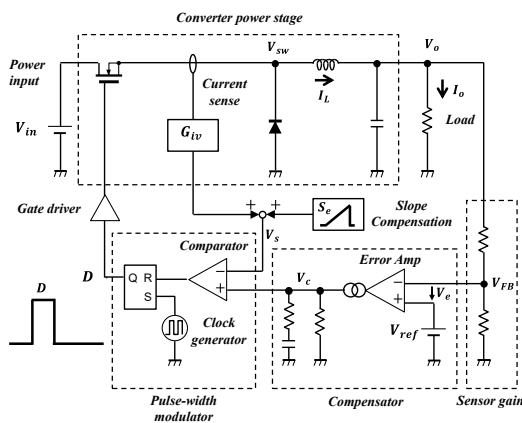


図 1：電流モード制御方式の原理図

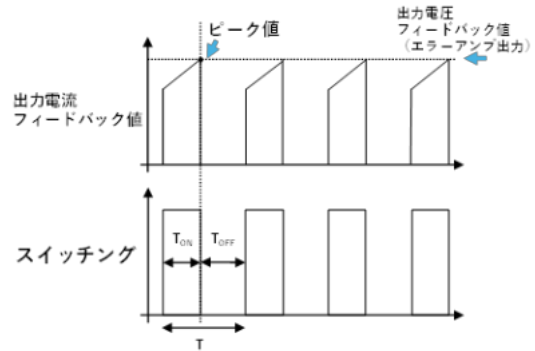


図 2：電流モード制御方式のタイミングチャート

$$V_o = \frac{T_{on}}{T_{on} + T_{off}} V_{in} \equiv DV_{in} \quad (1)$$

このようにして、電流モード制御ではインダクタ電流を帰還させることで出力電圧を一定に保つように時比率 D を制御する [1,2].

3. 制御電源を用いた電流帰還モデルの構築

本研究で提案する制御電源を用いた電流帰還モデルを図 3 に示す。 V_{in} と V_{sw} の間に流れるインダクタ電流 i_L は電流制御電流源 CCCS (Current Controlled Current Source) を用いて検出し、検出した電流は電圧源に接続した抵抗 R_s を介して電圧に変換される。スロープ補償回路は電圧制御電流源 VCCS (Voltage Controlled Current Source) および定電流 I_T とコンデンサ C_T を用いて構成している。コンデンサ C_T に定電流 I_T を充電する電圧を VCCS で電流に変換し、抵抗 R_s を介して電圧に変換している。スロープの傾きは定電流 I_T およびコンデンサ C_T の値で調整が可能である。抵抗 R_s によって変換された電圧 V_s がコンパレータへの入力電圧となり、出力電圧と基準電圧の誤差を増幅したエラーアンプ出力と比較することにより、スイッチング波の時比率 D が決定される。

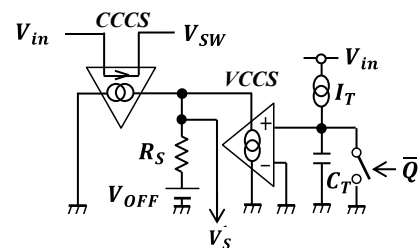


図 3：制御電源を用いた電流帰還モデル

4. 実機検証およびシミュレーション検証

本研究で提案した制御電源による電流帰還モデルを用いたシミュレーション動作波形と先行研究[3]にて開発実装した PCB 評価基板を用いた実機動作波形の比較検証を行う。なお、先行研究[3]では電流モード制御方式の降圧形 DC/DC コンバータ IC には ROHM 社の BD9329A を用いている。

$V_{in}=12V$, $V_o=3.3V$, $L=10\mu H$, $C_o=22\mu F$, $I_o=1.0A$ の条件にて、先行研究で開発した実機動作におけるスイッチング波形を図 4、出力リップル電圧の波形を図 5 に示す。

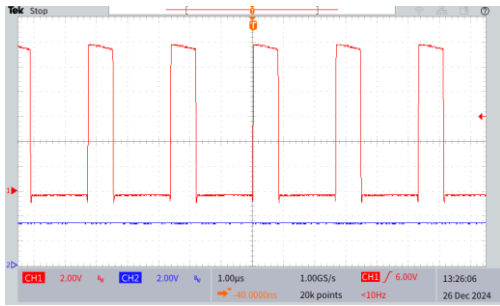


図 4：実機動作のスイッチング波形[3]

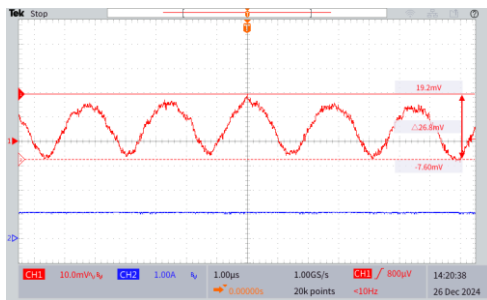


図 5：実機動作の出力リップル電圧[3]

次に、本研究で提案したモデルを使用したシミュレーション回路を図 6 に示す。図 6 のシミュレーション回路に使用したパラメータを表 1 に示す。

表 1：シミュレーション回路（図 6）のパラメータ

Parameter	Value	Parameter	Value
V_{in}	12 V	R_1	27 k Ω
V_o	3.3 V	R_2	10 k Ω
R_{on}	150 m Ω	G_{ma}	300 $\mu A/V$
L	10 μH	V_{ref}	0.9 V
R_L	3.3 Ω	R_o	10 M Ω
$C_{o1,2}$	20 μF	R_C	7.5 k Ω
$r_{c1,2}$	5 m Ω	C_C	3300 pF
L_C	10 nH	R_S	125 m Ω
r_L	20 m Ω	F_{CLK}	380 kHz

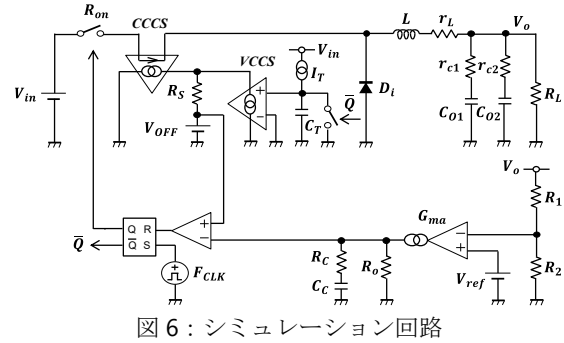


図 6：シミュレーション回路

図 7 にシミュレーションから得たスイッチング波形、図 8 に出力リップル電圧を示す。図 4 および図 5 の結果と比較して、概ね実機の動作を再現する波形を得ることができた。

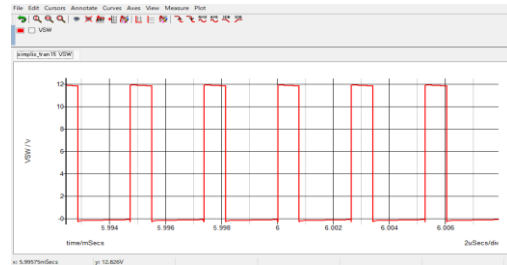


図 7：シミュレーションによるスイッチング波形

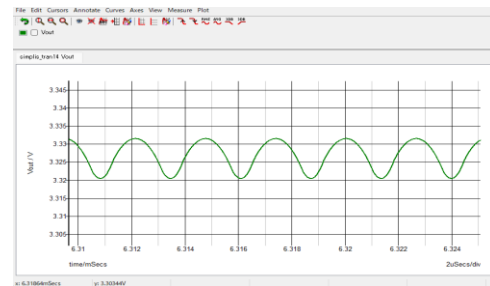


図 8：シミュレーションによる出力リップル電圧

5. まとめ

これまでに、動作するシミュレーション回路を作成し、実機で得られたスイッチング波形および出力リップル電圧と近い波形を得ることができた。今後はシミュレーション回路と実機を用いて更に比較検証を行い、電流帰還等価モデルの構築を行なっていく。

参考文献

- [1] RHOM 株式会社, Tech Web Hand Book, スイッチングレギュレータの基礎
- [2] RHOM 株式会社, Tech Web Hand Book, スイッチングレギュレータの特性と評価方法
- [3] 木下橋平, 電流モード制御 DC-DC コンバータにおける電流帰還モデルパラメータの抽出, R6 群馬工業高等専門学校本科卒業研究, 2024

25 : 0.6 μm CMOS オペアンプ回路の設計

築地研究室 福島 彩斗

1. はじめに

オペアンプ (演算増幅器、Operational Amplifier) は、アナログ電子回路において最も基本的かつ重要な機能ブロックの一つであり、フィルタ、増幅器、発振回路、AD・DA 変換回路など幅広い応用に用いられている。また、オペアンプは理想増幅器モデルを通じて負帰還や安定性、周波数特性を学ぶ格好の題材であり、アナログ回路設計の基礎を身につけるうえで中心的な役割を果たしてきた。一方で、実際の集積回路として CMOS プロセス上にオペアンプを設計およびレイアウトし、その動作をトランジスタレベルで検証する機会は多くない。そこで、本研究では、東京科学大学が主催する演算増幅器設計コンテストの試作部門への参加を念頭に置き、0.6 μm CMOS プロセスを対象とした 2 段構成オペアンプの設計に取り組むこととする。具体的には、差動入力段と増幅出力段からなる 2 段 CMOS オペアンプを対象とし、直流利得や利得帯域幅積、位相余裕などの基本特性をコンテスト所定の仕様値以上とすることを目標とする。なお、試作部門ではフェニックス社 0.6 μm CMOS モデルパラメータを用いるが、機密情報に該当するため、本発表ではシミュレーション部門で公開されている TSMC 社 0.18 μm CMOS モデルパラメータに置き換えて報告する。

2. 2 段構成オペアンプの概要

本研究で対象とする CMOS オペアンプの構成を図 1 に示す。図 1 のオペアンプは差動増幅回路とソース接地増幅回路を組み合わせた 2 段構成のオペアンプである。

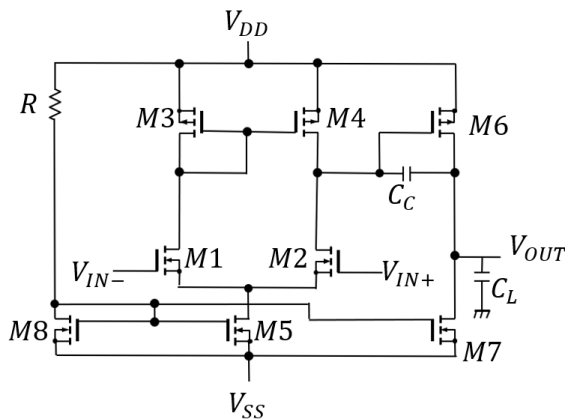


図 1 : 2 段構成オペアンプの概要図

3. MOS キャラクタライズ

飽和領域における MOS のドレイン電流は以下の式で表される。

$$I_D = \frac{\mu C_{ox}}{2} \left(\frac{W}{L} \right) (V_{GS} - V_T)^2 \left(1 + \frac{V_{DS}}{V_A} \right) \quad (1)$$

ここで、 μ はキャリア移動度、 C_{ox} はキャパシタの単位面積当たりの容量、 W はチャンネルの幅、 L はチャンネルの長さ、 V_T は MOS のしきい値電圧、 V_A はチャンネル長変調効果の度合いを表すアーリー電圧[1]である。このうち I_D を計算するためには μC_{ox} 、 V_T 、 V_A は求める必要がある。シミュレーションモデル用のプロセスパラメータの値を直接使用しても、(1)式とは MOS モデルの計算式が違うので計算結果が合わない。そのため、これらはシミュレーション結果から抽出するキャラクタライズを行う。

求め方は以下の通りである。 L は本研究で想定している 0.6 μm 、 W は今後計算しやすいように L の 10 倍の 6 μm 、 V_{DS} は 5V に設定する。はじめに、(1)式を V_{GS} で微分し、 I_D を V_{GS} で微分したトランスコンダクタンス g_m を求める。求めた g_m の接線から外挿で V_T を求める。次に、 g_m を V_{GS} で微分して求めた g_m の傾きと V_A から μC_{ox} を求める、という流れである。この流れを、スクリプトで抽出する。

まず、 V_{GS} 、 $g_m - I_D$ 特性から NMOS の g_m の傾きと、 V_T を求める。 V_{GS} を 0~5V に変化させて測定する。次に、 $V_{DS} - I_D$ 特性から I_D を V_{DS} で微分したドレインコンダクタンス g_d を求めて、 V_A を導出する。 $V_{GS} - V_T$ を 0.2V 程度に設定し、 V_{DS} を 0~5V に変化させて測定する。求めた g_m の傾きと V_A を (2)式に代入して μC_{ox} を求める。

$$\mu C_{ox} = \frac{\frac{\partial g_m}{\partial V_{GS}}}{\frac{W}{L} \left(1 + \frac{V_{DS}}{V_A} \right)} \quad (2)$$

その後、(1)式を用いて手計算した I_D とシミュレーションした I_D がどの程度一致しているかを確認する。PMOS でも以上の測定を行う。しかし、 V_{GS} 、 V_{DS} 、 V_T は負の値になるため、絶対値を取る必要がある。

図 2~5 に NMOS と PMOS の V_{GS} 、 $g_m - I_D$ 特性、 $V_{DS} - I_D$ 特性測定回路とスクリプトを示す。また、これらの測定回路とスクリプトから得られた g_m の傾き、 V_T 、 V_A 、 μC_{ox} を表 1 に、これらを用いて手計算した I_D とシミュレーションした I_D の誤差を表 2 に示す。

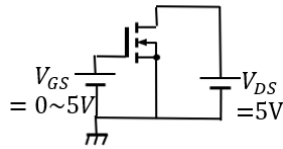


図2 : $V_{GS}, g_m - I_D$ 特性測定回路(NMOS)

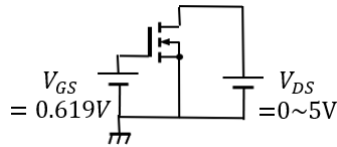


図3 : $V_{DS} - I_D$ 特性測定回路(NMOS)

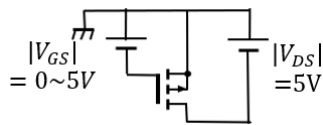


図4 : $V_{GS}, g_m - I_D$ 特性測定回路(PMOS)

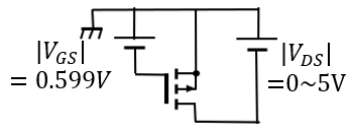


図5 : $V_{DS} - I_D$ 特性測定回路(PMOS)

表1 : NMOS と PMOS の g_m の傾き、 V_T 、 V_A 、 μC_{ox}

	g'_m [mS/V]	V_T [V]	V_A [V]	μC_{ox} [$\mu A/V^2$]
NMOS	2.472	0.419	12.19	175.3
PMOS	0.639	0.399	12.13	45.25

表2 : I_D の手計算とシミュレーションの誤差

	V_{DS} [V]	I_D [μA]		誤差 [%]
		手計算	Sim	
NMOS	5.0	49.41	51.58	+4.4
PMOS	5.0	12.78	13.37	+4.6

表2より、誤差が5%以内で一致することが確認できた。シミュレーションから求めた μC_{ox} 、 V_T 、 V_A は有効な値であることがわかる。

4. カレントミラー回路の設計

図1のM5、M7、M8部分はカレントミラー回路で構成されている。カレントミラー回路[1]とは、もとにするMOSに電流が流れると、もう片方のMOSにW/Lの大きさ分だけ電流が流れる回路である。図1でのも

とになるMOSはM8である。式に表すと(3)式になる。 I_{Dn} はn番目のMOSのドレイン電流、 $(W/L)_n$ はn番目のMOSのW/Lである。

$$I_{D8} : I_{D5} : I_{D7} = \left(\frac{W}{L}\right)_8 : \left(\frac{W}{L}\right)_5 : \left(\frac{W}{L}\right)_7 \quad (3)$$

図1のカレントミラー回路の設計方法を以下に述べる。例とする回路を図6に示す。

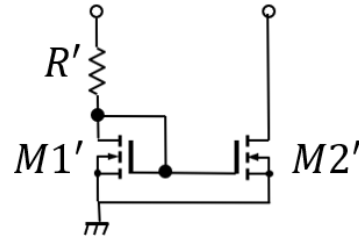


図6 : カレントミラー回路

まず、MOSに流す電流の大きさ、 V_{DD} 、 $V_{GS} - V_T (= V_{eff})$ を決める。これと、先ほど求めた V_T 、 V_A 、 μC_{ox} と以下の2つの式を用いてRとM1'、M2'のWを求め、Lは固定する。

$$W = \frac{2I_D}{\mu C_{ox} \cdot V_{eff}^2 \cdot \left(1 + \frac{V_{DS}}{V_A}\right)} \quad (4)$$

$$R = \frac{V_{DD} - (V_{eff} + V_T)}{I_D} \quad (5)$$

これらを実際に計算し、シミュレーションで確認した結果を以下に述べる。 $I_D = 50\mu A$ 、 $V_{DD} = 5.0V$ 、 $V_{eff} = 0.20V$ とする。また、Lは $0.6\mu m$ で固定とする。(4)、(5)式より、 $W = 6.84\mu m$ 、 $R = 87.62k\Omega$ となった。シミュレーションすると、M1'に流れた電流は $49.87\mu A$ 、 $V_{GS} (= V_{eff} + V_T)$ は $0.630V$ となった。M2'に流れた電流も $V_{DS} = 0.630V$ でM1'に流れた電流と同じ値になったことから、カレントミラーとして機能していることがわかる。

5. まとめ

今回はCMOSオペアンプ回路設計に必要なMOSのキャラクタライズを行い、カレントミラー回路のシミュレーションによってキャラクタライズの結果を検証した。今後は、図1にある2段構成オペアンプの全MOSのW/L比の計算、設計したオペアンプのシミュレーションを行う予定である。

参考文献

- [1] 吉澤浩和 : CMOS OP アンプ回路 実務設計の基礎、pp.10、18、19、29、30、75、76、87、109 (2007)

26：音声対話型デジタルサイネージシステムの開発

川本研究室 黛 琵琶

1 はじめに

近年、公共施設や商業施設にデジタルサイネージ（以下、サイネージ）が多く見られる。サイネージとはディスプレイを用いて情報発信を行うシステムで、電子看板や電光掲示板とも呼ばれる。従来の看板やポスターに比べ、印刷や張替えの作業の削減や映像を用いて情報発信ができるなど多くの利点を持つ。さらに、タッチパネルを用いたインタラクティブ性が注目されていて、実際に商業施設ではタッチパネルを用いて施設のエリアマップを表示及び検索することができるものも存在している。しかし、このタッチパネルでは設置や維持に関する問題やアクセシビリティの低さといった問題が存在し、この問題を解決する方法として音声対話が挙げられる。

現状、音声を活用したサイネージを使用している店舗や施設はほとんど見ることがない。過去には NTT と PDC による接客型デジタルサイネージの提案 [1] や凸版印刷による駅で接客サイネージ「BotFriends®Vision」の実証実験 [2] が行われているがこれらは製品として市場に出ていない。また、サイネージとしてタッチパネルや音声対話のユーザビリティ向上に関する研究は多く行われているものの、それぞれの入力方式がサイネージの評価にどのような影響を与えるのかを検討している研究は行われていない。本研究では、タッチパネルと音声対話がサイネージに対してどのような影響を与えるのかを調査し、音声対話の有効性を検証することが目的である。

2 タッチパネルと音声対話の比較

タッチパネル式サイネージは大規模商業施設のフロアマップやフードコートの店舗案内などに用いられている。しかし、タッチパネル式は目的の情報を得るために多数のアクションが必要であり、また、幼児や高齢者などのタッチ操作に慣れない層が使いこなせないというデメリットがある。対して、音声対話式は複雑な操作が不要で係員と話しているように情報を受け取ることができるため、簡単に使えてアクセシビリティが高い。また、音声対話式サイネージは通常のディスプレイとカメラを設置すれば運用可能である点も優位である。

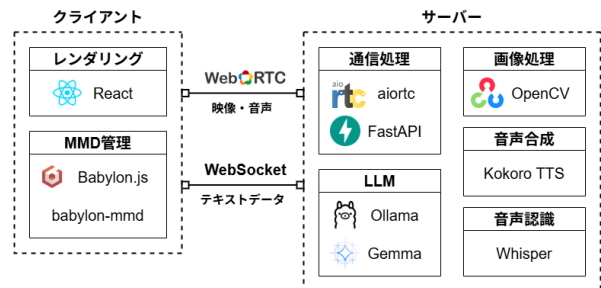


図1 技術スタック

3 開発システム

既存の音声対話式サイネージとして、名古屋工業大学の李研究室が開発した MMDAgent-EX がある [3]。これは C や C++ などと比較的古い技術が用いられているため実装可能な画面制御が限られている。そのため本研究では MMDAgent-EX は用いず、表示部分には画面制御が実装しやすい Web 技術を用いたシステムを開発し、タッチパネルと音声対話の比較実験を行う。本研究で開発する音声対話式サイネージは、低負荷の処理をクライアント側に、高負荷の処理をサーバ側に分割したクライアント・サーバ方式のシステムとする。クライアント側はサイネージに搭載されたカメラ映像の配信と対話エージェントの表示、サイネージコンテンツの表示を行う。サーバ側は映像の受信とユーザの発話に対する返答処理を行う。本システムの技術スタックを図1に示す。

4 提案システムの製作状況

現時点でクライアントとサーバとともに基本的な会話が行える音声対話式サイネージを製作した。

4.1 クライアント

図2にクライアントの画面を示す。対話エージェントには MMD モデル [4] を使用し、MMD 制御には babylon-mmd ライブラリ、映像と音声の送受信には WebRTC 通信、サーバからの MMD 制御処理のための WebSocket 通信を行う機能を実装した。

4.2 サーバ

サーバではクライアントとの通信機能と会話区画検出 (VAD) や音声認識、返答生成、音声合成を実装した。



図2 クライアントのスクリーンショット

level	エラー率 [%]								
	10ms			20ms			30ms		
	W	P	B	W	P	B	W	P	B
0	3	3	3	4	0	0	4	0	0
1	3	3	3	4	0	1	4	0	0
2	3	3	3	4	1	2	3	0	0
3	1	2	2	4	0	2	3	1	0

表1 ノイズ加算後に非音声と判定された割合

VADにはWebRTC VADを用いて0.02秒のフレームに対して音声検出を行っているが、現段階では物音や環境音などのノイズを検出してしまふ問題がある。予備実験として、声優統計コーパス [5] より取得した6.76秒音声に対して、そのままの音声と音声対雑音比が10dBのホワイトノイズ・ピンクノイズ・ブラウンノイズをそれぞれ加算した計4つの音声を用意し、VADの非音声フィルタリングレベルとフレーム幅を変化させながら音声検出を行った。なお、フィルタリングレベルは0が最も弱く、3が最も強く働く。検出結果に対して、ノイズ加算前の音声で非音声と判定されたフレームがノイズ加算後に音声と判定された割合と、この逆の非音声が音声と判定された割合を取得した。この実験結果を表1, 2に示す。なお、表中のW, P, Bはそれぞれホワイトノイズ、ピンクノイズ、ブラウンノイズを表す。結果より、ノイズが含まれた音声に対する非音声フィルタリングは、levelが1かつフレーム幅が10msのときに最

level	エラー率 [%]								
	10ms			20ms			30ms		
	W	P	B	W	P	B	W	P	B
0	0	27	43	0	87	100	0	100	100
1	0	0	0	0	30	91	0	92	100
2	9	9	9	11	46	83	12	79	100
3	0	9	9	10	34	71	11	82	100

表2 ノイズ加算後に音声と判定された割合

も有効であることがわかった。以降のシステムではこれらの値を採用する。しかし、これは環境音のような常に発生しているノイズにのみ有効であり、物音や咳払いのような短いノイズに対しては有効ではない。そのため、例えば“500ms以上継続している音を音声とみなす”のような実装で対策を施すの必要があり、今後はこの実験に取り組む予定である。

5 まとめと今後の予定

現段階では、基本的な音声対話が行える音声対話式サイネージの開発とVADのパラメータに関する予備実験を行った。しかし、VADに関しては短いノイズの対策が行えていない。今後はVADの問題解決や表示コンテンツに合わせた対話を可能にするためのLLMの検索拡張機能(RAG)の実装などに取り組む。そして、音声対話式サイネージの有用性を検証するためにタッチパネル式と音声対話式の比較を行い、アンケート調査によってそれぞれの満足度を比較する実験を行う予定である。

参考文献

- [1] (株) NTT, “自然な音声対話が可能「接客型デジタルサイネージ」を開発”, <https://group.ntt.jp/newsrelease/2015/01/21/150121a.html> (2025/10/27 閲覧).
- [2] (株) 凸版印刷, “凸版印刷、「BotFriends®Vision+」を提供開始”, https://www.holdings.toppan.com/ja/news/2020/12/newsrelease201202_3.html (2025/10/27 閲覧).
- [3] Akinobu Lee, “MMDAgent-EX”, <https://mmdagent-ex.dev/ja/> (2025/10/17 閲覧).
- [4] Sour 暄, “Sour 式初音ミク Ver.1.02”, <https://bowlroll.net/user/193408> (2025/10/20 閲覧).
- [5] y.benjo and MagnesiumRibbon, “Voice-Actress Corpus”, <http://voice-statistics.github.io/> (2025/11/04 閲覧).

27: 非流暢発話検出に基づく発表練習支援システムの開発

川本研究室 山本 康生

1 はじめに

発話の際、無意識のうちに「あー」や「そのー」といったような意味を持たない単語を発声してしまう場合がある。このような語は一般的に「フィラー (filler)」と呼ばれ、発話の流暢性を損なう原因の1つである。このフィラーによって相手が話の内容を理解しやすくなる [1] とときもある反面、多すぎてしまうと話の理解度の低下 [2] や、不快感を与える原因となりうる。特に流暢性が求められる口頭発表では、発表者の発話においてフィラーを減らすことは、聴衆の理解度向上や心地よい発表につながる。

フィラー検出技術については、フィラーにラベルを付けて検出するもの [3] や、大語彙連続音声認識エンジン Julius を用いてフィラー検出する手法 [4] など様々なものが取り組まれている。さらに、「UMU」と呼ばれる学習プラットフォームの「uShow」 [5] のように発表練習のためのサービスは存在するが、発話の流暢性が中心ではなく、また、オンライン上での利用であれば音声情報が外部に流出する可能性がある。そこで本研究では、フィラー検出を中心としたオフラインで使用可能な発表練習支援システムの実現を目指す。

2 フィラー検出システムの試作

発表練習の音声からフィラーを検出するプロトタイプシステムを試作するにあたり、「音声認識エンジンで音声を文字データに変換すること」、「文章からフィラーとなる部分を検出する」ことの2段階での処理により、実現することを検討した。

音声認識エンジンについては、次の3つを要件として考え、技術選定を行った。

- 1) フィラーを含む非流暢発話も省略することなく文字データに変換できること。
- 2) 発表練習の音声情報が外部に流出することを避けるため、オフラインで処理が完結すること。
- 3) 発話タイミングの時間を取得可能であること。

候補となる音声認識エンジンとして、多方面で利用される whisper [7], Speech Recognition [8], vosk [9] の3つについて比較した。このうち、whisper は単語ごとの発話の開始時間と終了時間を取得できなく、Speech Recognition はオフライン利用ができない。そのため、プロトタイプシステムにおいては vosk を採用した。なお、vosk を使用する際、認識モデルとしては提供されるモデルの中で性能がよい「vosk-model-ja-0.22」を使用した。

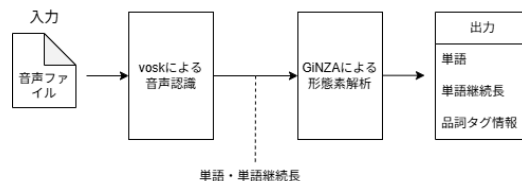


図1 システム構成図

表1 vosk+GiNZA でのフィラー検出精度

含まれるフィラー数	音声認識した	20
	音声認識していない	10
計		30
検出したフィラー数	フィラー	10
	誤検出	2
計		12
検出できなかったフィラー数		20

フィラーか否かを判別するための文章の形態素解析のライブラリとして GiNZA を用いた。GiNZA の他にも形態素解析を行えるライブラリは存在するが、GiNZA の単語間の文法的な構造の解析がフィラー検出に役立つと考えて GiNZA を採用した。

これらの機能を組み合わせ、音声入力に対し、フィラーを検出するプロトタイプシステムを構築した。システム構成を図1に示す。

3 試作システムのフィラー検出精度検証

プロトタイプシステムにおけるフィラーの検出精度を確認するための実験を行った。検証に使用する音声は発話であり、検証音声全体で30個のフィラーを含むよう設計されたデータである。プロトタイプシステムによる検証データのフィラー検出精度をまとめた結果を表1に示す。表1より、vosk による音声認識で認識できなかったフィラーを含めると33.3%のフィラーを検出でき、認識できたフィラーのみであれば、50.0%のフィラーを認識できた。「あの」などのような指示語と同じ発音のフィラーが検出できなかったため検出率が低くなったと考えられる。また、今回は vosk の標準の出力をそのまま使用したところ音声認識の漏れによって、33.3%のフィラーを取り出すことができなかったため、カスタム辞書にフィラーの追加や他の音声認識ライブラリによる補完で認識精度の向上を図る。

4 フィラーと指示語の判別

vosk と GiNZA を使っただけでは精度が低かったため、指示語かフィラーかを判断できるように改善する。

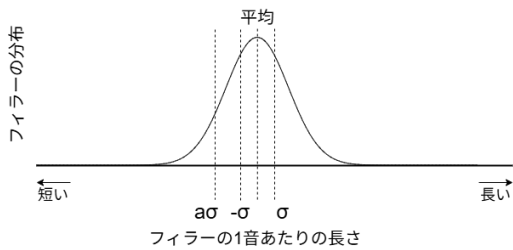


図2 考え方の正規分布

4.1 判別方法

フィルターの「あの」や「その」は指示語の「あの」や「その」と比較して、発声したときの長さが長くなる傾向が見られた。そこで、各単語での1文字あたりの発声時間の平均をとり、その長さでフィルターと指示語を判別できるといふ仮説を立てた。

まず、検出された各フィルターの1文字あたりの発声時間を、voskによって得られた各単語の発声開始時間と発声終了時間の出力から計算し、その値を標準化してから図2のような正規分布を仮定して、調べたい単語がフィルターか否かを判別した。このとき、調べる対象は指示語として使われる連体詞とし、正規分布において範囲を考える際に標準偏差にかかる係数(a)を、誤ってフィルターと判別された語と誤ってフィルターではないと判別された語の数の交点を求めて定めた。音声ファイルとして、複数の文に合わせてフィルターが7個と指示語が20個含まれた音声と、フィルターが11個と指示語が15個含まれた音声の2つを用意して行った。

4.2 閾値の計算

正規分布における範囲を定めるときに標準偏差にかかる係数を算出するために、係数を閾値として変化させたときのフィルターがフィルターではないと判断された割合(T-F)と、フィルターではない語がフィルターだと判断された割合(F-T)を求めて、その交点を調べた。標準偏差にかかる係数を0.0から0.5の区間で0.1ごとに変化させて、T-FとF-Tの値を求めてそれをそれぞれシグモイド関数にフィッティングさせて近似曲線を求めるところ、式(1)の通りになった。

$$\begin{aligned} T-F : y &= \frac{1}{1 + e^{3.211x+1.979}} \\ F-T : y &= \frac{1}{1 + e^{-8.983x+6.8019}} \end{aligned} \quad (1)$$

この2つ関数の交点を求めたところ、0.3955であった。そのため、閾値を0.3955として、標準化によって平均値が0で標準偏差が1になったことから、平均値 $-0.3955 \times$ 標準偏差に基づき、標準化された各単語の1文字あたりの平均時間が -0.3955 以上の指示語をフィルターとする。

4.3 判別精度の検証

4.2で求めた値をもとに、フィルターか指示語かを判断したときの精度を表2に示す。ただし、音声はvoskとGiNZAのみで検証したときと同じものを使用し、GiNZAによる検出した語も含める。表2より、voskに

表2 指示語とフィルター判別精度

含まれるフィルター数	音声認識した	20
	音声認識してない	10
	計	30
検出したフィルター数	フィルター	20
	誤検出	2
	計	22
検出できなかったフィルター数		10

よる音声認識で認識できなかったフィルターを含めると、66.7%フィルターを検出でき、認識できたフィルターのみであれば、100%という値となった。誤検出したフィルターが2つあり2つとも「まで」という単語であった。この「まで」に対してのみ2回解析が行われており、1回目でフィルター判定され、2回目で接続詞と判定されたが結果が上書きされなかったため最終的にフィルターと判定されたと考えられる。

5 まとめと今後の課題

GiNZAの解析処理によるフィルター検出と、それだけでは検出できない「あの」や「その」といった、指示語と同じ音を持つフィルターの検出を行った。

今後は、音声認識によるフィルターの認識漏れの改善、フィルターの検出精度をより正確に測るための複数人の音声データを使用した検証と、フィードバック手法の評価を行う。

参考文献

- [1] Hiromitsu Nishizaki *et al.*, “The Effect of Filled Pauses in a Lecture Speech on Impressive Evaluation of Listeners”, Proc. Interspeech2007, pp.2673-2676, 2007
- [2] Junko Kondo, “Japanese Learners’ Oral Narratives: Linguistic Features Affecting Comprehensibility”, 『世界の日本語教育』, pp.53-74, 国際交流基金リポジトリ, 2004
- [3] Koharu Horii *et al.*, “End-to-End Spontaneous Speech Recognition Using Disfluency Labeling”, Proc. Interspeech2022, pp.4108-4112, 2022
- [4] 宗宮 充弘 他, “話し言葉音声中のフィルター検出精度と音声認識率の関連性”, 第7回情報科学技術フォーラム, pp.191-192, 2008
- [5] Umu, “uShow”, 2025年10月30日閲覧, <https://m.umu.co/product/u-show>
- [6] Megagon Labs, “GiNZA”, 2025年10月16日閲覧, <https://megagonlabs.github.io/ginza/>
- [7] OpenAI, “Whisper”, 2025年10月30日閲覧, <https://openai.com/ja-JP/index/whisper/>
- [8] PyPI, “Speech Recognition”, 2025年10月30日閲覧, <https://pypi.org/project/SpeechRecognition/>
- [9] Alpha Cephei, “vosk”, 2025年10月16日閲覧, <https://alphacephei.com/vosk/>

[28] アドリブに対応するビジュアルジョッキシステムの構築

川本研究室 余 夏輝

1 はじめに

ライブパフォーマンスにおいて、音楽と映像や照明といった演出の同期は極めて重要である。例えば、業界標準の VJ ソフトウェアである Resolume[1] は、DJ が提供する安定した BPM (Beats Per Minute) に映像クリップやエフェクトの再生速度を同期させたり、タイムライン機能を用いて事前に構成した映像ショーを再生したりすることで、この同期を実現している。

しかし、これらの手法は、音楽の構造やテンポが事前に決まっていることを前提としている。そのため、人間のソロ演奏で生じるテンポの揺れや、アドリブによる予期せぬ展開 (小節のスキップや繰り返しなど) が発生した場合、システムが持つタイムラインと実際の演奏との間にずれが生じ、同期が破綻するという根本的な課題があった。

そこで本研究は、音楽の構造をリアルタイムに解析し、演奏の動的な変化に追従することで、音楽と演出の同期を維持できるライブパフォーマンス支援システムを開発することを目的とする。

2 リアルタイム追従システムの構築

本研究では、演奏とお手本音源を照合し、次のパート切り替えタイミングを予測するアルゴリズムを提案する。実装には、Python の音声分析ライブラリである librosa[2] を用いた。

2.1 音響分析条件

本研究で扱う入力音源は、WAV 形式のモノラル音声ファイルである。音響分析の主要なパラメータは以下の通りである。次の分析を始めるまでにずらす時間をフレーム周期、1 フレームのサンプルの数をフレーム長とする。

- 標本化周波数: 22050 Hz
- フレーム周期: 約 23.2 ms (512 サンプル)
- フレーム長: 約 92.9 ms (2048 サンプル)

2.2 音楽特微量と類似度計算

音響特微量として、楽曲のハーモニー情報を捉えるクロマベクトル [3] を採用した。クロマベクトルは、音のエネルギーを 12 音階の強度分布として表現した 12 次元のベクトルである。これにより音色に左右されずにリアルタイム追従を行うことができる。

2 つの音源の時間的な対応付けには、DTW (動的時間伸縮法) を用いる。本稿では、DTW の計算コストを削減するため高速な近似アルゴリズムである FastDTW[4] を採用した。ベクトルの類似度は、音量の違いに対応できるコサイン距離を用いる。コサイン距離 $d_{cos}(\mathbf{a}, \mathbf{b})$ は、2 つのベクトル \mathbf{a}, \mathbf{b} の内積をそれぞれの大きさを正規化したものであり、式 (1) で定義される。

$$d_{cos}(\mathbf{a}, \mathbf{b}) = 1 - \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \quad (1)$$

この尺度を用いることで、音量に左右されない安定した類似度計算が可能となる。

2.3 探索アルゴリズム

リアルタイム予測において計算時間と精度のバランスが課題となる。そこで本稿では、以下 4 つの探索手法を設計し、その性能を比較検証した。

- A: **逐次探索法** (A とおく) 探索範囲を 1 フレームずつ走査し、最適な位置を特定する基本手法。
- B: **データ軽量化法** (B とおく) DTW で比較する特微量を、時間方向に半分の間引くことで、1 回あたりの DTW 計算量を削減する手法。
- C: **スキップ探索法** (C とおく) 探索範囲を 4 フレーム (≒ 0.1 秒) 間隔で走査することで、DTW の実行回数を削減する手法。
- D: **複合探索法** (D とおく) データ軽量化法 (B) とスキップ探索法 (C) を同時に適用し、さらなる高速化を図る手法。

3 予測性能の検証実験

3.1 実験設定

提案手法の有効性を検証するため、リアルタイム追従を模擬したシミュレーションを行った。データには、筆

者自身がピアノで演奏した『きらきら星』のWAVファイルを用いた。4つの手法それぞれに対して現在の演奏位置を特定し、テンポ比を用いて切り替えの予測を行った。

音源を0.5秒ごとに進めながら、各アルゴリズム(A-D)が現在位置を予測する性能を評価した。予備実験で複数の窓幅を試行した結果、予測精度と計算時間のバランスが最も良かった0.8秒を本報告の窓幅として採用する。

なお、本実験における計算時間の測定には、以下の仕様を持つPCを使用した。

- CPU: Intel Core i7-6800K (3.40 GHz)
- RAM: 64.0 GB

3.2 結果と考察

図1に予測精度、図2に計算時間の比較をそれぞれ示す。

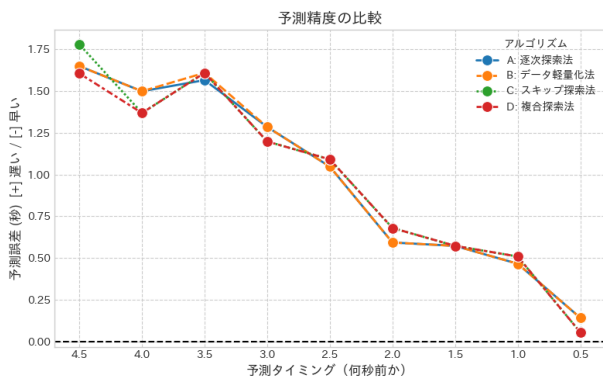


図1 : 予測精度の比較

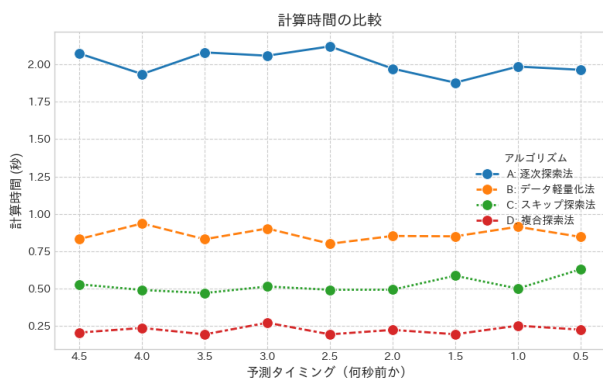


図2 : 計算時間の比較

図1より、提案した4手法の予測精度に大きな差はないことが分かる。一方、図2を見ると、計算時間には大きな差が確認された。特に、パート切り替えの0.5秒前

の時点で、逐次探索法(A)が約2.0秒を要したのに対し、複合探索法(D)は約0.25秒で計算を完了した。

音と映像の同期に関する先行研究では、映像が音に対して266ms先行しても、視聴者は統計的に有意な違和感を覚えないと報告されている[5]。本実験の複合探索法(D)は、0.5秒前の予測において予測誤差が56msであり、266msを下回る精度を達成している。さらに、その計算時間は0.25秒であるため、予測実行からイベント発生まで0.25秒の猶予があり、リアルタイム処理が可能である。

以上の結果から、複合探索法(D)は、人が違和感を覚えない精度を維持しつつ、計算時間をリアルタイム処理が可能な水準まで短縮できることを実証した。

4 まとめと今後の課題

本研究では、リアルタイム演奏追従のための効率的な探索アルゴリズムとして、データ軽量化とスキップ探索を組み合わせた複合探索法を提案した。実験の結果、本手法が精度を保ったまま、計算時間を大幅に短縮できることを示した。

しかし、本検証は一楽曲のみによるものであり、適当なアルゴリズムであったかは判断ができない。また、表示切替等の残りの処理の遅延許容量が0.25秒未満である。そのため、余裕を持たせるように高速化しテンポやジャンルが異なる複数の楽曲を用いて本手法の有効性を追加検証する必要がある。また、その上で、次の優先課題である「小節のスキップ・繰り返し」といった、より高度なアドリブに対応するアルゴリズムの開発へと進む。

参考文献

- [1] Resolume B.V.: Resolume VJ Software, <https://www.resolume.com/> (2025年10月29日閲覧).
- [2] B. McFee, et al.: “librosa: Audio and music signal analysis in python,” *Proc. 14th Python in Science Conf. (SciPy)*, pp. 1-7, 2015.
- [3] M. Müller, et al.: “Chroma-based statistical audio features for audio matching,” *Proc. IEEE WASPAA*, 2005.
- [4] S. Salvador and P. Chan: “FastDTW: Toward accurate dynamic time warping in linear time and space,” *Proc. KDD '07*, 2007.
- [5] 林 真由佳: “映像と音の時間ずれに対する違和感は曲の種類によって異なるか?,” 日本女子大学大学院人間社会研究科紀要, Vol. 28, pp. 71-78, 2022.

29：集合の種類と濃度の階層

北田研究室 江畑 太一

1. はじめに

集合という概念は、19世紀に数学者ゲオルク・カントールによって体系化され、無限集合や濃度の考え方を通して数学に新しい視点をもたらした。

本研究の目的は、無限集合の概念を理解し、有限集合との違いや、無限の種類(可算・非可算)を具体例を交えながら考察することである。

本発表では、べき集合をとる操作が無限の中に階層を作り出すことや、自然数と実数の集合の濃度の比較、カントル集合が実数の集合と同じ濃度をもつことを紹介する。

2. 用語説明

2.1 集合

- ・集合：ある性質を満たすものをひとまとめにした対象
- ・元(要素)：集合を構成する1つ1つの対象
- ・ \mathbb{N} ：自然数全体の集合 (例：1,2,3 \in \mathbb{N})
- ・ \mathbb{Z} ：整数全体の集合 (例：-1,0,1 \in \mathbb{Z})
- ・ \mathbb{Q} ：有理数全体の集合 (例： $\frac{2}{3}, 0.5, 0.333\dots \in \mathbb{Q}$)
- ・ \mathbb{R} ：実数全体の集合 (例： $\sqrt{2}, e, \pi \in \mathbb{R}$)
- ・有限集合：要素の個数が有限である集合
(例： $\emptyset, \{\emptyset\}, \{0,1\}, \{x,y,z\}, \{A,B,C,D\}$)
- ・無限集合：有限集合でない集合 (例： $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$)
- ・1対1対応：2つの集合間の要素の漏れも重複もない対応のこと (全単射)

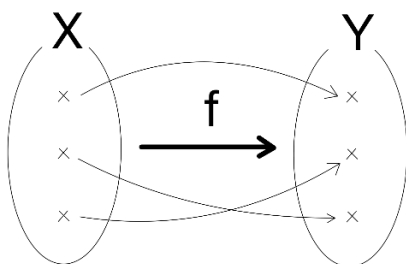


図1:1対1対応

2.2 濃度

X, Y を集合とする。

- ・濃度が等しい： X, Y の間に1対1対応があること
($\#X = \#Y$, (濃度は#をつけて表す))
- ・ $\#X \leq \#Y$ ： X から Y への1対1写像(単射)が存在すること
- ・可算集合： \mathbb{N} と同じ濃度を持つ無限集合
($\#\mathbb{N} = \aleph_0$, 例： $\mathbb{N}, \mathbb{Z}, \mathbb{Q},$ 直積集合 $\mathbb{N} \times \mathbb{N}$)
- ・高々可算集合：有限か可算な集合 ($\#X \leq \aleph_0$)
- ・連続体の濃度： \mathbb{R} の濃度 ($\#\mathbb{R} = \aleph, \aleph_0 < \aleph$)
- ・非可算集合：濃度が \aleph_0 より大きい集合
(例： $\mathbb{R}, \mathbb{R} \setminus \mathbb{Q}$ (無理数全体), \mathcal{C} (カントル集合))

3. べき集合

べき集合とは、ある集合において、すべての部分集合を集めた集合

例えば、

$$A = \{1,2\} \text{ のとき、}$$

$$\mathfrak{P}(A) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$$

どんな集合(無限集合も含む)のべき集合ももとの集合より濃度が大きくなることが知られている。 $\#A < \#\mathfrak{P}(A)$
→無限には階層がある。

$$\#A < \#\mathfrak{P}(A) < \#\mathfrak{P}(\mathfrak{P}(A)) < \dots$$

4. 実数の集合 \mathbb{R} は可算集合か?

[定理] 実数の集合 \mathbb{R} は可算集合ではない

この証明の前に、开区間(0,1)に属する点全体の集合を $\mathbb{R}(0,1)$ と表すと、 \mathbb{R} と $\mathbb{R}(0,1)$ は1対1対応するため、同じ濃度をもつことに注意する。

「集合 $\mathbb{R}(0,1)$ は可算集合ではない」ことを証明するために、対角線論法を用いる。

$\mathbb{R}(0,1)$ は可算集合と仮定すると、 $0 < x < 1$ を満たす全ての实数は、 \mathbb{N} と 1 対 1 対応するため、

$$\mathbb{R}(0,1) = \{x_1, x_2, x_3, \dots, x_n, \dots\}$$

と表すことができる。各 x_n を無限小数による10進展開で表し、

$$x_1 = 0.\alpha_1\alpha_2\alpha_3 \dots \alpha_n \dots$$

$$x_2 = 0.\beta_1\beta_2\beta_3 \dots \beta_n \dots$$

$$x_3 = 0.\gamma_1\gamma_2\gamma_3 \dots \gamma_n \dots$$

...

$$x_n = 0.\mu_1\mu_2\mu_3 \dots \mu_n \dots$$

とする。このとき、次のような無限小数 \tilde{x} をとってみる。

$$\tilde{x} = 0.\omega_1\omega_2\omega_3 \dots \omega_n \dots$$

ここで、 $\omega_1, \omega_2, \omega_3, \dots, \omega_n, \dots$ は、0でも9でもなくて、かつ、 $\omega_1 \neq \alpha_1, \omega_2 \neq \beta_2, \omega_3 \neq \gamma_3, \dots, \omega_n \neq \mu_n, \dots$ を満たしているとする。 \tilde{x} は、確かに0と1の間の実数を表しており、したがって $\tilde{x} \in \mathbb{R}(0,1)$ であるが、 \tilde{x} は、

$x_1, x_2, x_3, \dots, x_n, \dots$ のどれとも一致しない。

これは明らかに矛盾である。したがって、 $\mathbb{R}(0,1)$ は可算集合ではない。

つまり、 $\aleph_0 < \aleph$ (\aleph は \mathbb{R} の濃度)である。

5. カントル集合

5.1 手順

1. 区間 $[0,1] = \{x \mid 0 \leq x \leq 1\}$ を考える。
2. 区間 $[0,1]$ から真ん中の $\frac{1}{3}$ の区間(つまり $(\frac{1}{3}, \frac{2}{3})$)を取り除く。
3. 残った2つの区間 $[0, \frac{1}{3}]$ と $[\frac{2}{3}, 1]$ のそれぞれから、再び真ん中の $\frac{1}{3}$ を取り除く。
4. これを無限に繰り返す。

このようにして开区間をどんどん取り除いていくとき、最後まで除かれないう残った点全体のつくる集合を、カントル集合(C で表す)という。

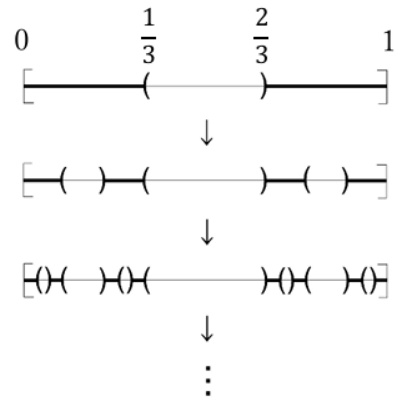


図2: カントル集合の構成過程

5.2 性質の説明

- C の中には区間(連続部分)が1つもない。
- 全体の長さは0になる。
- 無限に多くの点が残っていて、濃度 \aleph をもつ。

5.3 $\#C = \aleph$ である理由

カントル集合は、3進法で1を含まない数の集合としても表せる。

例:

$$0.0222\dots \rightarrow \bigcirc \quad (1 \text{ を含まない})$$

$$0.2122\dots \rightarrow \times \quad (1 \text{ を含むため除外})$$

この表現において、2を1に置き換えることで、カントル集合の点と、2進法で表した区間 $[0,1]$ のすべての実数が1対1対応する。

よって、 $\#C = \aleph$ である。

6. 今後の課題

今後の課題として、非可算集合の中でも濃度の違いをより詳しく調べ、無限集合の階層構造について理解を深めたい。

また、濃度の演算や、それに有用なベルンシュタインの定理についても学習していきたい。

参考文献

- [1] 志賀浩二: 「集合への30講」, 朝倉書店

30 : OpenFOAM を用いた空調効率の研究

宇治野・高橋研究室 佐相 冨太郎・中瀬 亮

1. まえがき

近年、オフィスや住宅などの室内環境において、快適性と省エネルギー性の両立が強く求められている。特に空調システムは建物全体のエネルギー消費の大きな割合を占めており、効率的な空調設計は重要な課題である。本研究では、空調条件の違いが室内温度分布や気流にどのような影響を及ぼすのかを、数値流体解析 (CFD) を用いて検討する。

現段階では、直方体の単純な部屋モデルを作成し、今後の解析に向けて OpenFOAM 環境を整備した段階である。

2. 概要

2.1 OpenFOAM

OpenFOAM は、オープンソースの数値流体解析 (CFD) ソフトウェアであり、流体の流れや熱伝達、化学反応などをシミュレーションできる。利用者が物理モデルや境界条件を柔軟に設定・拡張できる点が特徴である。

他の市販 CFD ソフトウェアと比べると、GUI がやや扱いにくい一方で、計算の自由度が高く、ライセンス費用が不要であるため、研究・教育目的に広く利用されている。また、ソースコードが公開されているため、独自の解析モデルを組み込みたい場合にも適している。

本研究では、空調の給気口や吸気口の位置、風量などを変化させた条件でシミュレーションを行い、室内の温度分布・気流分布を比較することを最終的な目的としている。そのため、熱伝達や乱流モデルを扱える OpenFOAM は非常に有効であると考えら

れる。

2.2 CFD (数値流体解析) と他手法との比較

CFD は、流体の運動を数値的に解析する手法であり、基礎となるナビエ-ストークス方程式を離散化して近似的に解くものである。CFD では主に「有限差分法 (FDM)」「有限要素法 (FEM)」「有限体積法 (FVM)」の3つが代表的である。

- ・有限差分法 (FDM) : 格子点上で微分方程式を直接近似する方法であり、アルゴリズムが単純で計算が高速な一方、複雑な形状の境界条件を扱いにくいという欠点がある。

- ・有限要素法 (FEM) : 構造解析などで広く用いられ、メッシュを柔軟に変形できるため複雑な形状にも対応できる。しかし、要素間での計算が煩雑になり、計算コストが高くなる傾向がある。

- ・有限体積法 (FVM) : 流体解析で最も一般的な手法であり、各セルに対して保存則を適用して計算を行う。計算の安定性と汎用性のバランスが良く、複雑な境界にも対応しやすい。

OpenFOAM はこの有限体積法を採用している。FVM では、各セルの表面を通過する流束を計算し、体積内の物理量変化を評価することで、質量・運動量・エネルギーの保存を確実に保つことができる。この特徴により、空調のように流れと熱の両方を扱う解析に適している。

また、FEM と比較すると、メッシュ生成や計算式の設定が比較的簡便であり、乱流モデルを容易に導入できる点も実用的である。特に建築分野では、室内空気分布の解析や煙流動シミュレーションなどに広く利用

されている。

2.2.1 有限体積法における保存則の仕組み

FVM の最も重要な優位性は、物理量の保存則を離散化の過程で厳密に満たす点にある。CFD の基礎となるナビエ-ストークス方程式などは、質量、運動量、エネルギーといった物理量の保存則を表現したものである。FVM では、解析領域を多数のセルに分割し、個々のセルに対して保存則を適用する。これは、特定の物理量 ϕ について、セルへの流入量と流出量が、セル内での発生量と釣り合っていることを意味する。定常状態における、一つのセル V における物理量 ϕ の保存則は、全表面を通る流束の総和を用いて以下の離散化された形式で表現される。

$$\sum_f F_{in} - \sum_f F_{out} + S_{cell} = 0$$

- ・ $\sum_f F_{in}$: セルのフェイス (表面) から流入する物理量 ϕ の流束の総和。
- ・ $\sum_f F_{out}$: セルのフェイスから流出する物理量 ϕ の流束の総和。
- ・ S_{cell} : セル内部での物理量 ϕ の発生源 (ソース)。

この式の物理的な意味は、
(流入流束) - (流出流束) + (内部発生) = 0
すなわち、セル内で物理量が失われたり増えたりすることなく、厳密に保存されていることである。FVM の離散化手法では、あるセルから流出した流束 F_{out} は、隣接するセルの流入流束 F_{in} として完全に等しく計算される。このため、全計算領域のセルを合計しても、内部の界面を通過する流束は全て相殺し合い、領域全体の境界を通過する流束と内部での発生源のみが全体として釣り合う。この性質が、FVM が流体解析において高い信頼性を持ち、特に空調のようにエネルギー保存が重要となる熱流体解析に広く採用される理由である。

3. 現段階での解析環境構築

現段階では、OpenFOAM を用いて直方体状の室内モデルを作成し、解析環境の整備を行っている。モデルは一辺が数メートルの直方体空間とし、今後の解析で給気口・排気口を設置して気流シミュレーションを行う予定である。以下に現時点での解析環境を図1に示す。

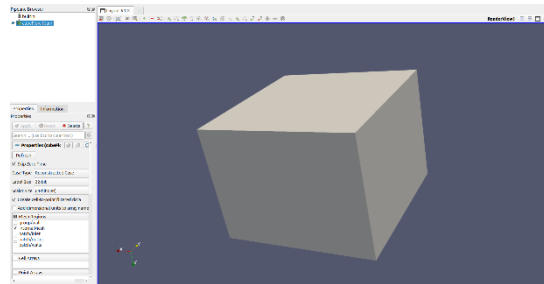


図1 : 現時点での解析環境

メッシュ生成には「blockMesh」を用い、単純な構造でセル数を制御できるようにした。次のステップとして、壁面条件 (断熱・非断熱) や初期温度分布を設定し、熱移動を含めた解析を行う予定である。これにより、空調の位置や風量が室内温度分布に及ぼす影響を定量的に評価できる見込みである。

4. 今後の展望

今後は、作成したモデルを基に給気口や排気口の位置を変えた複数条件の解析を実施し、室内の温度・流速分布を比較する。また、エネルギー効率や快適性の観点から評価指標を設定し、どの配置が最も効率的な空調設計となるかを検討する。

さらに、将来的にはより現実的な室内構造をモデル化し、実験データとの比較を通じて解析精度の検証を行う予定である。

参考文献

日本建築学会編『建築環境工学』丸善版, 2016.