

手首検出を用いた高齢者見守り・依頼送信システムの提案

広瀬研究室
c1231470 辻啓太

令和 8 年 1 月 13 日

概要

日本では全国的に少子高齢化と過疎化が進行しており、とくに農山村地域ではその影響が深刻化している。地域の高齢化率上昇に伴い、日常的な生活支援の担い手不足、交流機会の減少、緩やかな見守り体制の維持が困難になるといった課題が顕在化している。既存の高齢者見守りシステムは常時監視型が中心であり、プライバシー負担や導入コストの高さ、高齢者自身の意思が反映されにくいといった問題を抱えている。また農山村地域では通信環境や機器管理の負担から、これらのシステムが普及しにくい点も課題として指摘されている。そこで本研究では、高齢者が必要時に自発的に援助を依頼できる仕組みとして、カメラ映像と MediaPipe Hands による指先検出を用いた非接触 UI を提案する。(328 字)

目次

第 1 章	序論	5
1.1	研究背景	5
1.2	本研究の位置づけ	6
1.3	モーションキャプチャに関連する基本用語	6
第 2 章	関連研究	9
2.1	高齢者支援システムの研究動向	9
2.2	身体動作・ジェスチャー認識を用いた支援技術	10
第 3 章	モーションキャプチャの歴史	11
3.1	モーションキャプチャの歴史的発展	11
3.2	深層学習ベースの姿勢推定の登場	11
3.3	MediaPipe Hands による手指モーションキャプチャ	13
3.4	姿勢推定技術の分類と歴史的発展	14
3.5	見守り分野への応用と課題	14
第 4 章	依頼・見守り機能の設計	15
4.1	設計方針	15
4.2	画面構成	17
4.3	操作の流れ	18
第 5 章	指先検出と判定モデル・データの処理	19
5.1	簡易モーションキャプチャの発展	19
5.2	MediaPipe Hands	19
5.3	滞留判定アルゴリズム	20
5.4	システム全体構成	21
5.5	Sinatra API	22
5.6	データベース構造	23
5.7	構成	24
第 6 章	評価と考察	27
6.1	評価観点の再整理	27
6.2	現時点での観察と問題点	27

6.3	日向地区で得られたフィードバックと考察	28
6.4	今後の改善と評価計画	28
第 7 章	結論	31
参考文献		33

第1章

序論

1.1 研究背景

日本の農山村地域では、高齢化率の上昇と若年層の流出が全国平均を大きく上回るペースで進行しており、農林水産省の白書では「地域社会の担い手不足」と「生活支援機能の弱体化」が深刻な課題として指摘されている[1]。特に農山村地域では、単身高齢者世帯が増加し、家族・近隣住民による日常的な見守りや軽作業の支援が従来のように期待できない状況が報告されている。また同白書では、地域外からの関係人口が地域活性化の鍵であると示されており、外部人材が高齢者支援に関わる仕組みづくりの重要性が強調されている。特に、遠隔地に家族が暮らし、身近な支援者を確保できない高齢者にとって、日常的な安否確認や軽作業支援を受ける機械は限られている。このような地域においては、地域外の若年層や外部ボランティアが地域と関わる「交流人口」を拡大、継続的な支援の担い手を確保することが求められている。ここでいう交流人口とは、観光客のように短期的に訪れる滞在者や、地域活動に参加する関係人口を含む、定住者以外で地域と継続的に関わりを持つ人々の総称である。農林水産省の白書でも、農村の活性化には交流人口の増加が重要とされ、地域外の人材が農作業支援や日常的な交流に参加することが、支援体制の維持や高齢者の生活支援の補完として位置づけられている。

一方で、技術を用いた高齢者支援システムは都市部を中心に多数提案されているものの、農山村地域に適した形で設計された事例は少ない。特に、生活状況を共有する際の「プライバシー保護」は重要であり、高齢者が抵抗なく利用できる設計と、地域外の若年層が参加しやすい支援インターフェースの両立が課題となっている。

1.2 本研究の位置づけ

本研究は、農山村地域における高齢者支援において、DX（デジタル技術）を活用した新しい見守りのあり方を検討するものである。既存の見守りシステムは生活センサや常時監視型の方式が主流であるが、プライバシー負担や導入コスト、さらには高齢者自身の意思を反映しにくい点など多くの課題が指摘されてきた。また農山村地域では通信環境・機器管理の制約から都市部と同様の技術を適用しにくい状況にあり、地域特性に適合した見守り手法の再設計が求められている。

このような背景から、「必要なときに自ら意思を示せる」「負担の少ない操作で援助を依頼できる」仕組みの重要性が高まっている。そこで注目されているのが、人の身体動作を利用して意思や状況を伝えるモーションキャプチャを利用したインタフェースである。身体動作による入力、触れたり話したりすることが難しい状況でも利用でき、高齢者の多様な身体能力に対応しやすいことから、見守り支援の新しい手段として期待されている。

次節では、本研究で扱う身体動作入力技術を理解するため、モーションキャプチャに関する基本的な概念について概説する。

1.3 モーションキャプチャに関連する基本用語

本節では、本研究で扱う MediaPipe Hands や姿勢推定技術を理解する上で必要となる基礎用語を整理する。

1.3.1 ランドマーク (Landmark)

ランドマークとは、人の手指や身体の特定位点を表す「基準点」を意味し、モーションキャプチャ分野では身体動作を記述する最小単位として用いられる。MediaPipe Hands では指先・関節など 21 点が定義され、各ランドマークは正規化座標 (x, y, z) として取得される。これらの点の配置と推移を追跡することで、ジェスチャーや姿勢を解析できる。MediaPipe Hands が採用する座標系は「右手系 (Right-handed coordinate system)」であり、次の特徴を持つ。

- x 座標：画像左端を 0、右端を 1 とした水平方向の正規化値
- y 座標：画像上端を 0、下端を 1 とした垂直方向の正規化値
- z 座標：カメラからの距離を表し、手前方向が負値になる

この座標系では、 x と y は 2 次元画像上の位置関係を表し、 z は手の前後動作を推定するために利用される。また、右手系の座標定義により、人差し指先端（ランドマーク ID 8）を含む 3 次元位置関係が整合的に扱えるため、指先の動きや押下動作判定などの計算が容易になる。

さらに、MediaPipe の正規化座標は入力画像の解像度に依存しないため、異なるデバイス（スマートフォン、タブレット、Web カメラ）間でも統一された判定が可能である。本研究では、ランドマーク ID 8（人差し指先端）を指先座標として取得し、UI ボタン領域との位置関係の判定に利用した。

1.3.2 ポーズ推定 (Pose Estimation)

ポーズ推定とは、画像や動画から人間の関節位置を推定する技術の総称である。従来は反射マーカークラウドや深度センサが必要であったが、近年は深層学習の発展により RGB カメラのみで精度の高い推定が可能となった。

代表的な技術として、Posenet, MoveNet, MediaPipe Pose などがある。これらはいずれもリアルタイムかつ単眼カメラで動作可能である。

本論文の第 3 章では、こうしたポーズ推定技術の系譜と特徴、および簡易モーションキャプチャとしての発展過程について詳細に整理する。

1.3.3 スケルトンモデル (Skeleton Model)

人体を関節点 (ランドマーク) とそれを結ぶ骨格線で表現したモデルを指す (図 1.1)。スケルトンは動作解析・ジェスチャー認識・アニメーション生成などに広く用いられ、手指に特化したものが MediaPipe Hands である。

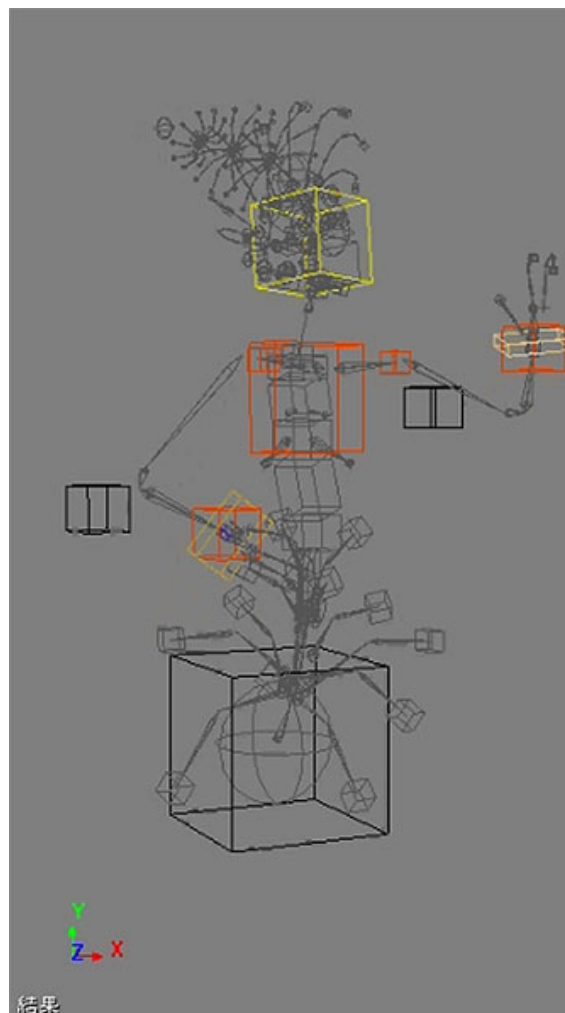


図 1.1 スケルトンアニメーション (神戸芸術工科大学より引用)[2]

第2章

関連研究

2.1 高齢者支援システムの研究動向

高齢者見守り技術は、IoT・センサネットワーク・画像処理・ロボティクスなど幅広い領域で進展してきた。なかでも、高齢者が住まう居宅内にセンサを配置し、生活動作データを収集して異変を検知するタイプの研究が多くみられる。温度センサ・人感センサ・加速度センサなどを用いる方式は、カメラを必要としないためプライバシー負担が少なく、導入コストも比較的低い。この方式は孤独死防止や在宅ケアの補助として実運用された例も報告されている。

しかし一方で、センサ設置型システムは「生活状況の抽象的推定」に留まる傾向があり、高齢者の意思・困りごと・希望を直接的に取得することは難しい。また、異変検知はできても「困っている内容を具体的に伝達する」「支援者へ能動的に依頼する」といった双方向性のある支援には繋げにくいという課題が存在する。加えて、センサの常時設置は設置工数・機器管理の負担・費用面で導入障壁となりやすい。特に農山村地域では通信環境や保守コストの問題も重なり、都市部と比べて普及が進みにくい状況にあると指摘されている。

東京都健康長寿医療センター研究所の小池らは「高齢者見守りセンサーに関する研究の現状と課題」(2012)[3]として、国内における見守り技術の発展経緯を俯瞰し、生活センサ方式の強みと弱みを体系化し、高齢者の生活サポートに見守りセンターの活用が必要であることを示した。しかしながら見守りシステムの多くが「異常検知・行動推定に偏る」点や、個々の高齢者の意思や感情が十分に反映されないこと、さらにプライバシー配慮が十分に考えられていないことを示している。この整理は本研究の位置づけを考える上で重要であり、「高齢者自身が操作し意思を表明できる UI」を設計する意義を裏付ける基礎資料となる。

本研究で扱う“依頼・見守り機能”は、常時監視ではなく“本人の意思による操作を前提とする”アプローチを軸としており、この先行知見が指摘する課題—「自律的意思の反映」「プライバシー保護」「身体的・認知的負荷の低減」—を解決する位置付けにあるといえる。

2.2 身体動作・ジェスチャー認識を用いた支援技術

見守り研究の次の流れとして、「センサが見守る」から「本人が動作で意思を伝える」方向への発展が挙げられる。ジェスチャーや手指検出による UI は、視覚言語的な入力を可能にし、ボタン押下や音声入力が困難な高齢者でも操作可能である点が評価されている。特に近年は、Kinect や深度カメラ・MediaPipe のような機械学習ベースの姿勢推定エンジンの発展に伴い、身体動作をインタラクション手段として活用する研究が増加している。

その代表例として、「Elderly Care Based on Hand Gestures Using Kinect Sensor」(2021) [4] が挙げられる。本研究は高齢者が手の動きだけで支援要求を送信できるシステムを構築しており、〈手を挙げる→食事依頼〉〈手を振る→緊急要求〉といった操作が可能である。これは「身体動作そのものを入力装置とみなす UI 設計」の成功例であり、本研究の方向性と最も近い研究領域といえる。特に、キーボードやスマートフォン操作が難しいユーザーに対して、自然動作で意思伝達できる点は本研究の MediaPipe Hands による操作設計と通底している。

ただし Kinect 方式は RGB カメラと赤外線 (IR) 深度センサを備え、ユーザの骨格位置を高精度に推定できるデバイスである (図 2.1)。そのため、深度センサを必要とし環境制約が大きい。また機器価格・設置空間・距離制約などから、農山村の家庭環境での導入には適合しづらい面も指摘される。一方、本研究の方式はノート PC・タブレット・スマートフォンのカメラなど一般的な機材のみで動作可能であり、機器依存性が低いという優位性を持つ。この点で本研究は「高齢者が自宅で使える現実的な入力設計」という観点から既存研究との差別化が図られている。

以上の関連研究から、本研究は既存の見守り研究が抱える課題（常時監視・意思反映の難しさ）を踏まえつつ、身体動作を操作インタフェースとする近年の知見を活かし、より低負荷・少機材・プライバシー配慮の方向で発展させた位置にあると整理できる。

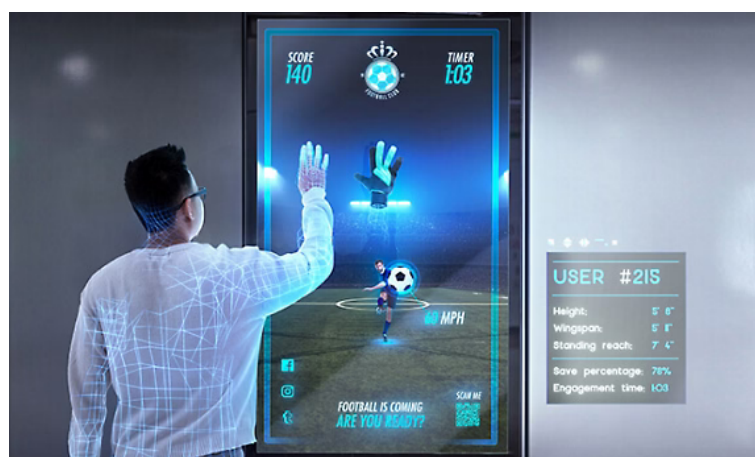


図 2.1 Microsoft Kinect DK(公式サイトより引用)[5]

第3章

モーションキャプチャの歴史

3.1 モーションキャプチャの歴史的発展

モーションキャプチャは、人の身体動作を計測し、デジタルデータとして取得する技術として発展してきた。当初のモーションキャプチャは映画制作やバイオメカニクス研究など専門領域に限られ、赤外線カメラと反射マーカを用いる方式が主流であった。この方式は精度が高い一方で、高価な専用スタジオを必要とし、一般利用には適さなかった [6]。

2000 年代に入るとゲーム産業の発展とともにモーションキャプチャの応用が拡大し、Microsoft Kinect に代表される「深度センサ方式」が登場した。Kinect はマーカを不要とし、深度カメラと機械学習モデルにより骨格推定を実現した点で革命的であった。これにより身体動作推定は一般家庭でも実行可能な技術となり、非接触インタフェース研究が急速に進展した [7]。しかし Kinect は専用ハードウェアに依存し、明るさ、距離、設置空間など環境制約が大きいという問題が残された。

その後、AI を用いた深層学習に基づく姿勢推定が急速に発展し、RGB カメラのみで骨格推定を行う「カメラ単体による モーションキャプチャ」へと潮流が移行した。

こうした技術的変遷によって、モーションキャプチャは「スタジオで行う特殊計測」から「日常利用可能な身体入力技術」へと変化し、医療、スポーツ、福祉、教育など幅広い領域に応用されるようになった。本研究が扱う MediaPipe Hands もこの流れの中に位置づけられる。

3.2 深層学習ベースの姿勢推定の登場

3.2.1 OpenCV を基盤とした画像処理技術

モーションキャプチャの一般化に大きく寄与した技術のひとつに OpenCV がある。OpenCV は 1999 年に Intel によって公開されたオープンソースのコンピュータビジョンライブラリで、画像処理・特徴点抽出・機械学習などの豊富な機能を備える [8]。深層学習以前の姿勢推定では、HOG 特徴量、Harris コーナー検出、オプティカルフローなど、OpenCV に基づく手法が多く用いられてきた。現在でも画像前処理や描画、リアルタイム処理の基盤として広く活用されており、深層学習ベースの推定技術と併用されることが多い。

- HOG 特徴量入力画像の各ピクセルで勾配 (大きさと方向) を計算し、小さなセルごとに勾配方向のヒストグラムを集計し特徴量として用いる手法であり、歩行者検出を中心とした物体検出全般に利用された [9]。

- Harris コーナー検出全方向に対し、画素位置の移動量に対する画素値の違いを示すコーナー点を抽出する方法である [10]。
- オプティカルフロー物体やカメラの移動によって生じる隣接フレーム間の物体の動きの見え方のパターンである。フレーム間の明度変化に基づき画素の動きを推定する [11]。

3.2.2 PoseNet の登場と Web 化

2018 年に Google が公開した PoseNet は、単眼 RGB 画像から 17 関節の骨格を推定する軽量モデルであり、Web ブラウザ上で動作可能な点が特徴である。TensorFlow.js によって GPU がない環境でもリアルタイム推定が可能であり、「ブラウザで動くモーションキャプチャ」を一般化した最初のモデルといえる [12]。

PoseNet が示した重要な意義は以下の通りである。

- 専用デバイスを不要とする姿勢推定の実現
- Web アプリケーションへの容易な組み込み
- 教育・福祉領域での活用可能性の拡大

このモデルはモーションキャプチャの民主化に大きく貢献し、後続研究の基盤となった。

3.2.3 MoveNet による高精度・軽量化

PoseNet の後継として Google が公開した MoveNet は、高速・高精度かつ軽量のモデルとして設計されている。特に素早い動作の MoveNet Lightning と高精度な MoveNet Thunder の 2 種があり、使用環境に応じて速度と精度を選択できる点が特徴である [13]。

MoveNet の特徴は以下に整理される。

- スマートフォンでも 30fps 以上の高速推定が可能
- 複数人推定 (MultiPose) にも対応
- 正確な関節角度の推定に優れる

MoveNet により、廉価な端末でも安定したモーションキャプチャが行えるようになり、スポーツ解析や医療リハビリなどで利用されている。

3.3 MediaPipe Hands による手指モーションキャプチャ

モーションキャプチャ技術の中でも、手指動作の推定は特に難易度が高い。指は細く関節が多いため、従来の骨格推定よりも高精度な推定モデルが必要とされる。

Google が公開する MediaPipe Hands は、この課題に対する代表的な解法の一つであり、単眼 RGB カメラのみを用いて 21 点の手指ランドマークをリアルタイムに推定する軽量モデルである [14]。MediaPipe Hands は手検出モデルとランドマーク推定モデルを組み合わせた二段階構成を採用しており、遮蔽や姿勢変化に強い推定が可能である。また、GPU だけでなく CPU でも高速に動作し、スマートフォン、Web アプリケーション、ノート PC など多様な環境に適用できる点が特徴とされている。

本研究では指先 (ID 8) を利用し、UI ボタンへの「滞留判定」によって非接触操作を実現している。

表 3.1 MediaPipe Hands における 21 ランドマークの対応表

ID	ランドマーク名 (部位)
0	手首 (Wrist)
1	親指 CMC 関節 (Thumb CMC)
2	親指 MCP 関節 (Thumb MCP)
3	親指 IP 関節 (Thumb IP)
4	親指指先 (Thumb Tip)
5	人差し指 MCP 関節 (Index MCP)
6	人差し指 PIP 関節 (Index PIP)
7	人差し指 DIP 関節 (Index DIP)
8	人差し指指先 (Index Tip)
9	中指 MCP 関節 (Middle MCP)
10	中指 PIP 関節 (Middle PIP)
11	中指 DIP 関節 (Middle DIP)
12	中指指先 (Middle Tip)
13	薬指 MCP 関節 (Ring MCP)
14	薬指 PIP 関節 (Ring PIP)
15	薬指 DIP 関節 (Ring DIP)
16	薬指指先 (Ring Tip)
17	小指 MCP 関節 (Pinky MCP)
18	小指 PIP 関節 (Pinky PIP)
19	小指 DIP 関節 (Pinky DIP)
20	小指指先 (Pinky Tip)

3.4 姿勢推定技術の分類と歴史的発展

モーションキャプチャの歴史を整理する際には、各時代で用いられた姿勢推定技術を「どの情報を基盤として推定を行うか」という観点で分類できる。本節では、代表的な技術である (1) マーカーベース方式、(2) 深度センサベース方式、(3) RGB カメラベース方式（深層学習）の三つに分類し、その歴史的発展を概説する。

■(1) マーカーベース方式 赤外線カメラと反射マーカーを利用した従来方式であり、映画制作やバイオメカニクス研究で 1990 年代から広く用いられてきた。複数台の IR カメラを用いてマーカーを三角測量し、cm 単位の高精度な 3 次元位置を推定できるが、高価な専用スタジオを必要とし、一般環境では使用が難しい。

■(2) 深度センサベース方式 2010 年代の Kinect を代表とする方式で、RGB 画像に加えて深度画像（距離情報）を用いて骨格推定を行う。深度情報を持つことで人物領域の抽出が容易になり、マーカーなしで関節位置を推定できる技術として普及した。ただし、深度推定精度は照度や距離に依存し、屋外や広い室内では性能低下が起きやすい。

■(3) RGB カメラベース方式（深層学習） 深層学習の発展により、単眼 RGB カメラのみで姿勢推定を行う「RGB ベース姿勢推定」が主流となった。PoseNet, MoveNet, MediaPipe などのモデルは、CNN によって関節位置（キーポイント）を直接推定し、機器依存性の低いモーションキャプチャを実現している。本研究が利用する MediaPipe Hands も同系統に属し、指先を含む細かな手指ランドマークの推定に特化している。

以上のように、モーションキャプチャは (マーカー) → (深度) → (RGB + 深層学習) という技術的遷移を経て発展しており、近年では「安価な機材」「環境制約の少なさ」「リアルタイム推定」が重視されている。この流れの中で、本研究が対象とする手指推定技術は、見守り分野において現実的に導入可能な技術として位置づけられる。

3.5 見守り分野への応用と課題

カメラベース姿勢推定の普及により、モーションキャプチャは見守り・介護支援の領域にも応用され始めた。従来の常時監視型システムでは、生活状況を「推定」するだけであり、高齢者の意思が直接反映されない問題があった。また、照度条件・カメラ距離・背景の複雑さなど、環境要因による精度低下といった課題により、特に農山村地域では通信環境や設置環境の制約が大きい。

第 4 章

依頼・見守り機能の設計

4.1 設計方針

高齢者が利用する UI では、操作内容が単純で分かりやすい、判断負荷・身体負荷のいずれも最小限であることが求められる。特に農山村地域での単独生活者に適用することを前提とした場合、「迷ったら操作できない」「押せないと伝えられない」といった状況を避ける必要がある。本研究では、以下の設計原則に基づき UI を構築した。

- **3 ボタン以内のシンプルな画面構成**

選択肢が増えるほど認知負荷は高まるため、機能は送信・確認・挨拶に集約した。

- **誤操作を防ぐ滞留判定（2 秒）**

一瞬触れただけでは操作と認識せず、指先が一定時間重なった場合のみ押下とする。これにより「テレビに指が触れただけで送信」「誤って動いた」などの誤動作を低減する。

- **身体動作として自然な操作モデル**

クリック動作・ボタン押下の代替として、現実の「指を重ねて押す」動作を採用し、スマートフォン操作が難しい高齢者にも直感的な操作体験を提供する。

- **カメラ映像を背景とする UI 配置**

自身の手の位置が視覚的に把握できるため、ボタンに重ねるという目的行動が理解しやすい。

- **配置位置・色彩の明確化**

UI は背景と混ざらない高コントラスト配色を採用し、文字は視認性の高い大きさ（18 ポイント以上）に統一した。これは WCAG(Web Content Accessibility Guidelines) のアクセシビリティ指針においても視認性の確保のため少なくとも 18 ポイントまたは、14 ポイントの太字が必須であると明記されている [15]。高齢者の視力差を考慮し、透明度とハイライト色により「触れている状態」が即時に判別できるようにした。

- **余計な遷移を排除した一画面完結型 UI**

画面遷移・ポップアップ・入力操作を要求しないことで認知負荷を最小化し、「見れば使い方がわかる」状態を目標とした。

- **操作結果の即時フィードバック**

押下判定時には色変化とメッセージ表示を行い、操作が成功したか迷わせない仕組みとした。

- **信頼関係に応じた情報開示の調整**

高齢者が安心して支援を依頼できるよう、支援者との関係性や距離に応じて開示する情報の範囲を段階

的に制御する仕組みを導入した。特に、依頼者自身が「どこまで情報を見せるか」を選択できるように、開示意欲度をスライダーで設定可能とした。

本研究における UI 設計原則の抽出は、農山村地域という利用環境の制約を踏まえて行った。高齢者は認知機能の観点や視力、細かい操作など身体的衰退と行った面から操作項目を最小限に整理することが重要である [16]。これを反映し、機能構成は 3 要素に限定した。また、高齢者は意図しない動作や手指の揺れが生じやすく [17]、誤操作防止機構の必要性が指摘されてきたため、滞留時間を用いた押下判定を導入した。高齢者向けインタフェースでは、身体動作に基づく自然な操作が理解しやすく、ボタン押下より手を伸ばす、かざす等現実世界に近い動きのほうが理解しやすいことが報告されている [18]。

加えて、農山村地域の単独生活者を念頭に置くと、複雑な画面遷移や入力手順は操作失敗の要因となることから、一画面で完結する構成が適している。視覚機能の多様性への配慮としては、高コントラスト配色や大きめの文字サイズを採用し、操作結果が即時に理解できるフィードバックも欠かせない要素となる。これらの要件を総合的に整理した結果、認知負荷の軽減、誤操作の抑制、視認性の確保、自然動作との整合性、環境制約への適応という観点から、上記の 7 原則を設計方針として適用した。

4.1.1 開示意欲と信頼関係に基づく設計

本研究では、依頼者が安心して支援を要請できる環境を整えるために、支援者との関係性や距離に応じて情報の開示範囲を段階的に調整する仕組みを導入した。特に、依頼者自身が「どこまで情報を見せるか」を選択できるように、開示意欲度をスライダーで設定可能とした。

この設計は、アーヴィン・アルトマンとダルマス・テイラーによる社会的浸透理論に基づいている。社会的浸透理論では、自己開示は信頼関係の深まりとともに段階的に進むとされており、物理的な距離よりも心理的な距離（信頼度や親密さ）が安心感に大きく影響する。

実際、置賜農業高校でのアンケートでは、「いくら近くても知らない人に情報を渡すのはためらう」といった声が聞かれた。こうした意見は、物理的距離だけでは支援の受け入れやすさを測れないことを示している。

そこで本研究では、物理的距離、信頼度、開示意欲の 3 要素を統合した「情報開示距離」という概念を導入し、依頼者のプライバシーと安心感の両立を図った。これにより、依頼者は自身の意志に基づいて情報の開示範囲を調整でき、より柔軟で信頼に基づいたマッチングが可能となる。

4.2 画面構成

本システムの画面には、カメラ操作を通じて操作できる以下の 3 つの主要ボタンを配置している。

- 朝の挨拶
- 服薬確認
- 依頼を送信

まず「朝の挨拶」は、単なるコミュニケーション機能ではなく、高齢者の生存確認を兼ねる重要な行為である。日々の軽い交流を通じて、家族が高齢者の生活状況を把握できる。

次に「服薬確認」は、農山村地域の高齢者において服薬アドヒアランスの低下が課題となりやすい点を踏まえたものである。特に高齢者は、複数の病気を同時に持つ併存疾患服により薬の数が増える傾向にある。そのため薬忘れの可能性が高まりそのリスクが危険視されている [19]。本システムではボタンによる簡易確認を導入することで、過度な監視を伴わずに日々の服薬状況を共有できる。

最後に「依頼を送信」は、高齢者が必要時に外部支援者へ作業支援を求めるための主要な手段として位置づけられる。農山村地域では除雪、草刈り、買い物支援、家具移動といった日常的な作業であっても、近隣に支援者が確保しにくい状況が多い。そのため、依頼内容を本人が選択して送信できる機能は、生活維持に不可欠な支援を確実につなぐ役割を果たす。また、身体的作業に限らず、日常的な会話やちょっとした相談など、心理的支えを必要とする場面にも活用でき、多様な生活ニーズに対応できる点に特徴がある。

以上の理由から、本研究ではこれら 3 機能を「最小限かつ必要性が高い操作項目」と位置づけ、UI を単純化することで認知負荷と誤操作のリスクを低減し、高齢者にとって実用的な構成とした。

図 4.1 に実際の画面を示す。

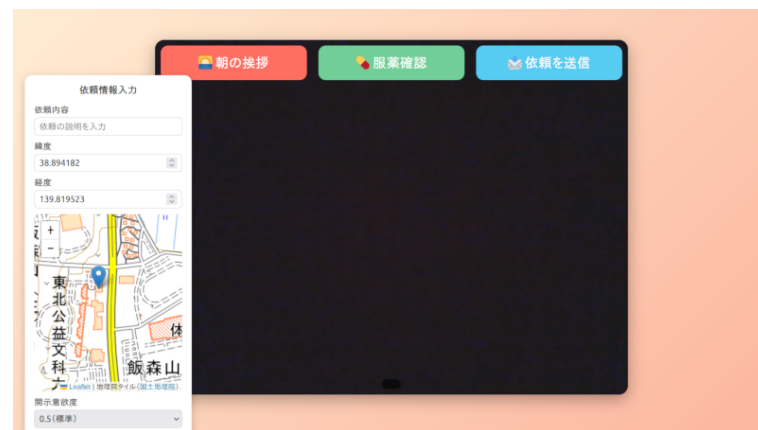


図 4.1 システムの操作画面

4.3 操作の流れ

図 4.2 は、本システムにおける利用者操作とシステム処理の概略フローを示す。利用者はカメラ起動後に表示された仮想ボタン上に手を重ねることで操作を開始する。利用者が指先を所定のボタン領域に 2 秒以上滞留させると、システム側で滞留判定が成立し、Sinatra API へ依頼情報が送信される。API 受信後、サーバは受信データを SQLite3 に記録し、送信完了のフィードバックを利用者画面へ返す。図は「利用者の具体的操作」および「システムの処理（判定・送信・記録・フィードバック）」が連続する様子を簡潔に示している。

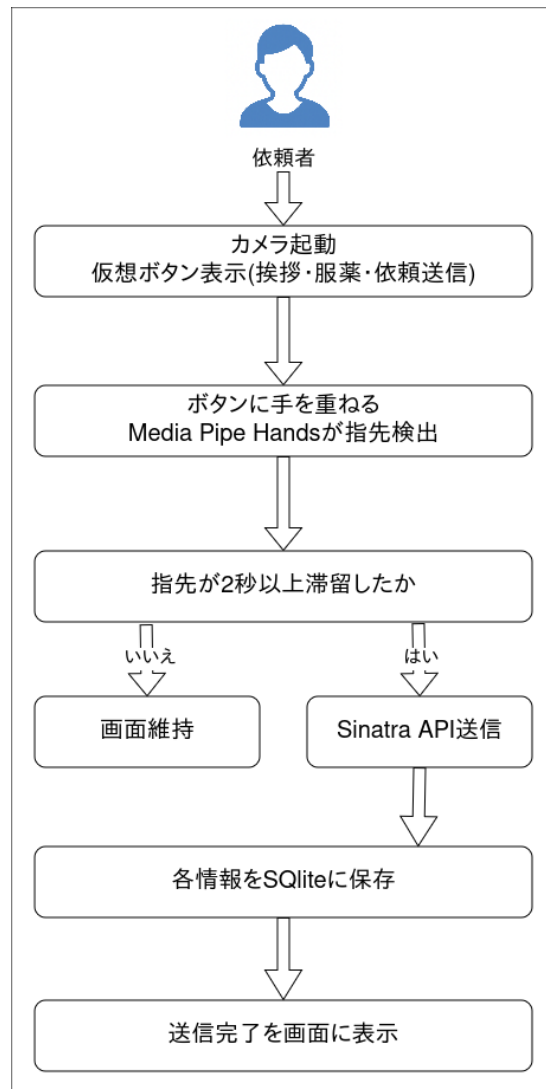


図 4.2 依頼操作のフロー

第 5 章

指先検出と判定モデル・データの処理

5.1 簡易モーションキャプチャの発展

近年のモーションキャプチャ技術は、専用スタジオを必要とする高精度計測から、一般的なカメラのみで動作を推定する軽量な方式へと大きく変化してきた。特に深層学習を用いた姿勢推定の発展により、RGB カメラだけで骨格点を検出する「簡易モーションキャプチャ」が広く普及した。PoseNet や MoveNet に代表される全身姿勢推定モデルは、人間の主要関節を高速に推定でき、スマートフォンやノート PC でも実行可能である。これに対し、手指の微細な動きを扱う研究では、より高精度なランドマーク推定が求められ、指関節 21 点を推定できる MediaPipe Hands が注目されている。また、OpenCV によって「画像を読み込む・輪郭を検出する・領域を追跡する」といった処理が容易になった。

5.2 MediaPipe Hands

MediaPipe Hands では 21 点の指関節を座標として取得できる。本研究では指先のランドマーク ID (8 番) を採用し、以下の処理を行う：

1. カメラ画像を OpenCV で取得
2. MediaPipe に渡して指位置を取得
3. ボタン領域との矩形判定
4. 2 秒間の滞留判定

5.3 滞留判定アルゴリズム

以下にボタンの押下に関するアルゴリズムを示す。

5.3.1 座標空間および変数の定義

MediaPipe Hands は、カメラ画像を幅 W 、高さ H とする 2 次元画像平面上で、指関節を正規化座標 (x, y) として返す。このとき、 $x, y \in [0, 1]$ の範囲であり、 $(0, 0)$ は画像左上、 $(1, 1)$ は画像右下に対応する。

本研究で扱う主な変数を以下に示す。

- $x(t)$, $y(t)$: 時刻 t における指先ランドマーク (ID=8) の正規化座標
- R : ボタン領域を表す軸平行長方形

$$R = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$$

- $P(t)$: 時刻 t に指先が R 内に存在するかを表す判定関数
- T : 押下と判定するために必要な滞留時間 (本研究では $T = 2$ 秒)
- f : 処理フレームレート (本研究では $f = 60$ fps の想定)
- t_0 : 指先がボタン領域に侵入し、その後連続して滞留し続けたとみなされる最初の時刻

座標空間を図 5.1 に示す。

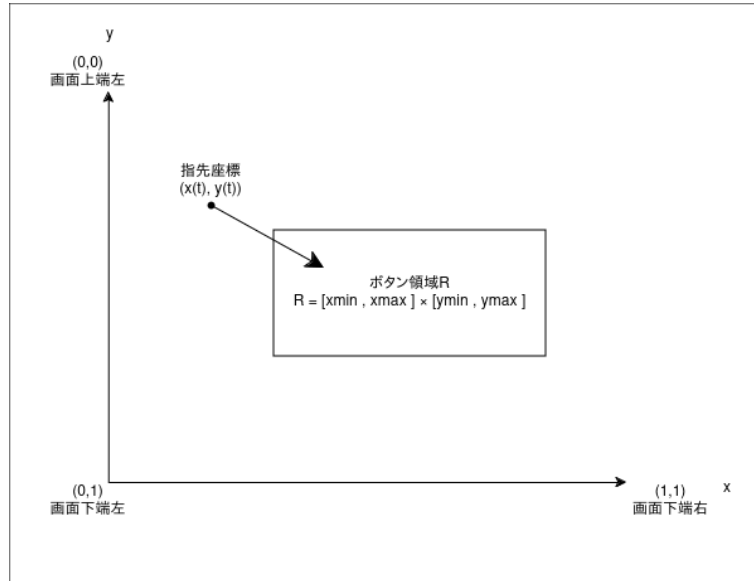


図 5.1 MediaPipe の正規化座標系とボタン領域 R の関係

5.3.2 滞留判定の定式化

判定関数 $P(t)$ は以下のように定義する。

$$P(t) = \{ 1 \mid (x(t), y(t)) \in R \} \text{ それ以外}$$

押下判定条件は次式で与えられる。

$$\sum_{t=t_0}^{t_0+T} P(t) = fT$$

本研究では $T = 2$ 秒, $f = 60$ fps としたため, 押下判定基準は以下のようになる。

$$\sum_{t=t_0}^{t_0+2s} P(t) = 120$$

これは, 指先ランドマークがボタン領域 R 上に連続して 120 フレーム存在した場合に押下が成立することを意味する。

5.4 システム全体構成

本研究で構築した非接触操作システムは, 利用者側で動作するクライアントアプリケーションと, 依頼データを受信・保存する Sinatra ベースのサーバアプリケーションから構成される。

クライアント側では, Web ブラウザ上で実行される JavaScript アプリケーションが MediaPipe Hands を用いて指先ランドマークを取得し, UI ボタンとの重なりを検出する。一定時間の滞留が確認されると, 操作内容 (挨拶, 服薬確認, 支援依頼) が JSON 形式として生成され, HTTP POST によりサーバへ送信される。

サーバ側は Sinatra を用いた軽量 Web API と SQLite3 データベースで構成される。Sinatra は POST リクエストを受信すると, 依頼種別, 送信時刻, 処理状態をデータベースに記録し, クライアントに JSON 応答を返す。

5.5 Sinatra API

本研究では支援依頼データをサーバに送信し記録するために、Sinatra を用いた簡易 Web API を構築した。Sinatra は Ruby 製の軽量 Web フレームワークであり、複雑な設定を行わずに HTTP 通信を受け付けるサーバ機能を実装できる点が特徴である。ここで利用する HTTP 通信の方式のうち、「POST」とはクライアントからサーバへデータを送信するためのメソッドであり、依頼内容の登録処理に用いられる。

POST 通信で送られるデータは JSON (JavaScript Object Notation) 形式を採用した。JSON はキーと値の組み合わせを持つ軽量データ形式であり、プログラミング言語に依存せず扱いやすいため Web API で広く用いられている。

依頼データはサーバ側で SQLite3 に保存される。SQLite3 はファイル単体で動作する軽量データベース管理システムであり、専用サーバを必要としないことから、小規模なアプリケーションに適している。本研究では「依頼種別」「時刻」「処理状況」などを保持するテーブルを用意し、Sinatra API から受信したデータを逐次記録する。

以下に依頼送信の具体的な POST リクエスト例を示す。

POST リクエスト例

```
POST /request
Content-Type: application/json
{
  quad "type": "help",
  quad "time": "2025-01-01 10:00",
  quad "location": \{ "lat": 38.9, "lng": 139.8 \},
  quad "trust": 7,
  quad "willingness": 5
}
```

Sinatra 側ではこの JSON データを受信し SQLite3 データベース に記録した上で、支援者向け UI に反映することで、依頼状況をリアルタイムに共有可能としている。

5.6 データベース構造

本研究の依頼管理機能では、Sinatra と SQLite3 を用いている。農山村地域のようにサーバ設備が限られる環境でも運用できる点から採用した。

本研究が扱うシステムでは、依頼情報を管理するためのテーブル `requests` を 1 つ用いる。本節では、データベースの模式図（スキーマ）を表形式で示す。

カラム名	型	説明
<u>id</u>	INTEGER	主キー（自動採番）
type	TEXT	依頼の種類（greet / med / help）
timestamp	TEXT	依頼が送信された時刻（YYYY-MM-DD HH:MM）
latitude	REAL	緯度
longitude	REAL	経度
trust	INTEGER	信頼度（1～10）
willingness	INTEGER	開示意欲（1～10）
status	TEXT	処理状況（未処理 / 対応中 / 完了）

表 5.1 依頼データベーススキーマ（requests テーブル）

依頼には、位置情報（緯度・経度）や、依頼者が設定した信頼度・開示意欲といった情報も含まれる。これらはマッチング時の情報開示距離の計算に用いられるため、データベースにも記録されるように設計している。

5.7 構成

5.7.1 非接触 UI の構築と滞留判定処理

本研究の非接触 UI は、MediaPipe Hands を用いて指先座標を取得し、その座標がボタン領域に連続して $T = 2000$ ms 滞在した場合に押下操作と判定する。この滞留判定は、操作に不慣れな高齢者でも誤操作を避けつつ、意図的な操作を確実に検出するための工夫である。

MediaPipe Hands の導入は、JavaScript アプリケーション上で以下のように構成される：

- **手指検出の初期化と設定**：Hands インスタンスを生成し、検出精度や手の数などのパラメータを設定する。
- **検出結果の処理と滞留判定**：人差し指先端（ID=8）の正規化座標を画面座標に変換し、仮想ボタン領域との重なりを検出する。一定時間（2000ms）滞留した場合に押下と判定し、依頼データを JSON 形式でサーバに送信する。

以下に、滞留判定と送信処理の主要部分を示す：

滞留判定と送信処理

```
const tip = lm[8];
const x = tip.x * canvasEl.width;
const y = (1 - tip.y) * canvasEl.height;

if (hitButton(x, y)) {
  stayTime += deltaTime;
} else {
  stayTime = 0;
}

if (stayTime > 2000) {
  sendRequest(type);
}
```


5.7.2 スクリーンショット機能（実装中）

現在、押下時の画面を自動でキャプチャし、依頼データとともに送信するスクリーンショット機能を実装中である。この機能により、依頼時の状況を視覚的に共有でき、支援者がより適切に対応できるようになることが期待される。

スクリーンショットは、Canvas API を用いて現在の映像フレームを画像として取得し、Base64 形式でエンコードして JSON に含める。以下にその処理例を示す：

スクリーンショットの取得と送信（実装中）

```
const canvas = document.createElement("canvas");
canvas.width = videoEl.videoWidth;
canvas.height = videoEl.videoHeight;
const ctx = canvas.getContext("2d");
ctx.drawImage(videoEl, 0, 0);
const screenshot = canvas.toDataURL("image/png");

fetch("/request", {
  method: "POST",
  headers: { "Content-Type": "application/json" },
  body: JSON.stringify({
    type: "help",
    time: getCurrentTime(),
    image: screenshot
  })
});
```


第 6 章

評価と考察

本研究は依頼送信・見守り補助を目的としたジェスチャ UI を実装したが、現段階では高齢者ユーザを対象とした実験には至っていない。そのため、本章では「現在すぐに測定可能な項目」に限定し、評価指標の妥当性と今後の改善課題を整理し、あわせて日向地区で得られた知見を述べる。

6.1 評価観点の再整理

これまでの議論を踏まえ、本研究の UI において「高齢者の主観に依存せず、短時間で客観的に測定可能な項目」を評価対象として再定義する。本研究では以下の 2 点を主要評価指標とする。

1. ボタン押下成功率

指先座標が正しく取得され、滞留判定を満たして押下が成立した割合を測定する。これはログ解析により定量的に取得可能であり、環境差（照明・背景）に対する頑健性の指標ともなる。

2. 誤反応率（誤判定率）

ボタン領域 R に入っていないにもかかわらず押下が成立してしまう事例を記録し、全押下判定数に対する割合として算出する。

いずれも MediaPipe Hands が生成するフレームごとのランドマーク座標と、ボタン領域判定ログを保存することで定量評価が可能であり、特別なユーザ実験を必要としない。

6.2 現時点での観察と問題点

現時点での動作確認により、以下の傾向が確認されている。

- 押下成功率は比較的良好

指先がボタン中央に重なるケースでは、2 秒滞留判定が正しく機能し、押下判定は安定して成立した。

- 指先追跡の安定性に課題

MediaPipe Hands による指先座標が照明条件や手の角度によって揺らぐ場面があった。特に逆光環境や背景とのコントラストが低い場合に座標ノイズが顕著となった。

- 誤反応（誤判定）が発生する

ボタン領域に完全に入っていないにもかかわらず、指先が境界に近づくとき押下判定が成立する例が見られた。

これらの結果より、本研究の UI は操作成立そのものは比較的安定しているが、誤反応率と追跡安定性は改善すべき主要課題として位置づけられる。

6.3 日向地区で得られたフィードバックと考察

日向地区での試行的な設置とヒアリングを通じて、システムの運用に関する実用的な示唆が得られた。以下に主な意見と考察を示す。

- **設置場所と導入初期の動線設計**

高齢者が最初にどこで操作するかを明確にする必要があるとの指摘があった。具体的には、壁掛けや玄関先の机の上など、日常動線上に自然に組み込める場所への設置が望ましい。

- **家族との連携と依頼の主体**

一人暮らしの高齢者にとっては、システムを通じて依頼を出すよりも、家族に直接連絡する方が心理的ハードルが低いという意見があった。そのため、本人だけでなく家族が代理で依頼を出せるような設計が重要である。

- **信頼と距離の関係性**

「いくら近くても知らない人ならいやだ」「どんなに遠くても家族なら助けてほしい」といった声があり、物理的距離よりも心理的な信頼関係が重視される傾向が確認された。本研究で導入した「情報開示距離」の設計方針は、この点と整合しており、今後は信頼度の定義や可視化方法の工夫が求められる。

- **高齢者の新規行動への抵抗感**

農山間地域の高齢者は、新しい操作や仕組みに対して強い抵抗感を持つことが多く、「今のままだと使ってもらえない」という率直な意見もあった。そのため、操作の簡略化や導入初期のサポート、継続利用を促す仕掛け（例：スタンプ機能など）を通じて、徐々に慣れてもらう工夫が必要である。

- **顔が見えることの安心感**

「顔が見えるのは良いと思う」「顔色や調子がわかる」といった意見もあり、カメラ映像を通じた非言語的な情報共有が信頼形成に寄与する可能性が示唆された。

これらの意見は、単なる UI の操作性だけでなく、地域の文化的背景や心理的要因を踏まえた設計の重要性を示している。今後は、こうした現場の声をもとに、より実用的で受け入れられやすい支援環境の構築を目指す必要がある。

6.4 今後の改善と評価計画

誤反応率の低減と追跡安定性の向上に加え、実用性と継続利用を高めるために、以下の改良と機能追加を計画している。

1. **ボタン領域の安全マージン導入**

ボタン領域 R をそのまま用いるのではなく、外側に非反応領域（dead zone）を設定し、境界付近での誤判定を抑制する。

2. **環境条件による補正**

明度変動に対して自動的に露出補正を行うなど、OpenCV を用いた前処理により MediaPipe Hands の安定性を向上させる。

3. スクリーンショット機能の実装

押下時の画面を自動でキャプチャし、依頼データとともに送信する機能を実装中である。これにより、依頼時の状況を視覚的に共有でき、支援者がより適切に対応できるようになることが期待される。

4. ログイン・スタンプ機能の導入

高齢者の継続的な利用を促すため、ログイン時にスタンプが付与される仕組みを導入予定である。日常的な利用に対する小さな達成感や楽しさを提供することで、操作への心理的抵抗を軽減することを目指す。

5. 家族連携アカウント（親子アカウント）の実装

一人暮らしの高齢者が自ら操作することが難しい場合を想定し、遠方の家族が代理で依頼を登録・管理できる機能を追加する。家族が依頼を出すことで、支援のきっかけを作りやすくし、システムの利用促進と安心感の向上を図る。

これらの改善後は、以下の項目について再評価を行い、システムの実用性と信頼性を検証する予定である。

- 押下成功率
- 指先追跡の安定性
- スクリーンショットによる状況伝達の有効性
- 継続利用率（ログイン頻度・スタンプ取得数）
- 家族アカウントによる依頼登録の実用性

第7章

結論

本研究では農山村地域の交流人口支援システム WhoMatch において，高齢者が自発的に支援を依頼できる「依頼・見守り機能」を実装した。MediaPipe Hands を活用した非接触 UI は，身体操作だけで意思表示できるため，従来の常時監視・センサ依存型の見守り方式とは異なり，プライバシーに配慮しながら自律的なコミュニケーションを促す可能性を示した。

今後は本 UI の操作実験を実施し，認識精度・誤操作率・安定性の改善を行う。特に指先検出が困難な環境への適応，ボタン領域の動的拡張，照明条件の変化への補正は重要な課題である。

本研究は設計段階としての成果であり，今後の実装改善と評価により，高齢者の意思表示を支え見守り負担を軽減する実用的な UI への完成を目指す。

参考文献

- [1] 農林水産省. “第 1 節 農村人口の動向と地方への移住の促進”, 『令和 5 年度食料・農業・農村白書』, 2025.https://www.maff.go.jp/j/wpaper/w_maff/r5/r5_h/trend/part1/chap5/c5_1_00.html, (参照 2025-11-26).
- [2] 神戸芸術工科大学. “神戸芸術工科大学紀要 芸術工学 2008”.https://www.mlib.kobe-du.ac.jp/bulletin/kiyou_old/08/work/01-04.html, (参照 2026-01-13).
- [3] 小池 高史, 野中 久美子, 他. “高齢者見守りセンサーに関する研究の現状と課題”, 老年社会科学, 2012, 34 巻, 3 号, p.412-419, https://www.jstage.jst.go.jp/article/rousha/34/3/34_412/_article/-char/ja/, (参照 2025-11-26).
- [4] R. Gomes, M. H. Kolekar, and S. L. Nalbalwar. “Elderly Care Based on Hand Gestures Using Kinect Sensor”, Computers, vol.10, no.1, p.5, 2021.<https://www.mdpi.com/2073-431X/10/1/5>, (参照 2025-11-27) .
- [5] Microsoft. “ Azure Kinect DK ”.<https://azure.microsoft.com/ja-jp/products/kinect-dk#x945b4a9ee3834b98abdc524c716e2afd>, (参照 2025-12-06) .
- [6] Kishor Das, Thiago de paula Oliveira, John Newell. “ Comparison of markerless and marker-based motion capture systems using 95% functional limits of agreement in a linear mixed-effects modelling framework”, scientific reports, vol. 13, 2023. <https://www.nature.com/articles/s41598-023-49360-2>, (参照 2025-12-09) .
- [7] Microsoft. “ About Azure Kinect DK”.<https://learn.microsoft.com/en-us/previous-versions/azure/kinect-dk/about-azure-kinect-dk>, (参照 2025-12-09) .
- [8] OpenCV. “ Open Computer Vision Library”.<https://opencv.org/>, (参照 2025-12-07) .
- [9] 山崎 俊彦. “画像の特徴抽出:Histogram of Oriented Gradients(HoG)”, 映像情報メディア学会誌, 2010, 64 巻, 3 号, p.322 -329, https://www.jstage.jst.go.jp/article/itej/64/3/64_3_322/_article/-char/ja/, (参照 2025-12-09) .
- [10] 鳥取大学. “Harris コーナー検出”.https://labs.eecs.tottori-u.ac.jp/sd/Member/oyamada/OpenCV/html/py_tutorials/py_feature2d/py_features_harris/py_features_harris.html, (参照 2025-12-09) .
- [11] 鳥取大学. “オプティカルフロー (Optical Flow)”.https://labs.eecs.tottori-u.ac.jp/sd/Member/oyamada/OpenCV/html/py_tutorials/py_video/py_lucas_kanade/py_lucas_kanade.html, (参照 2025-12-09) .
- [12] Google AI for Developers. “姿勢ランドマーク検出ガイド”.https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker?hl=ja, (参照 2025-12-09) .
- [13] Google Colab. “movenet.ipynb”.<https://colab.research.google.com/github/tensorflow/>

- docs-l10n/blob/master/site/ja/hub/tutorials/movenet.ipynb?hl=ja, (参照 2025-12-09) .
- [14] Google AI for Developers. “手のランドマーク検出ガイド”.https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker?hl=ja, (参照 2025-12-09) .
- [15] ウェブアクセシビリティ基盤委員会 (WAIC). “Web Content Accessibility Guidelines (WCAG) 2.2 (日本語訳)”.<https://waic.jp/translations/WCAG22/>, (参照 2025-12-09) .
- [16] 総務省情報通信政策研究所. “高齢者のユーザビリティに配慮した ICT 利活用環境に関する調査研究”.https://www.soumu.go.jp/main_sosiki/joho_tsusin/b_free/pdf/0703_hk1.pdf, (参照 2025-12-09) .
- [17] 国立研究開発法人 国立長寿医療研究センター. “手の震え”.<https://www.ncgg.go.jp/hospital/navi/03.html>, (参照 2025-12-09) .
- [18] 長谷部 尚志, 芦澤 寛樹, 山本 昇志. “ジェスチャー認識を備えた直感的な画像インターフェイス”, 社会法人映像情報メディア学会技術報告, 2012, 34 巻, 10 号, p.29-32, https://www.jstage.jst.go.jp/article/itetr/34.10/0/34.10_29/_article/-char/ja/, (参照 2025-12-09) .
- [19] 一般社団法人 くすりの適正使用協議会. “あなたのくすりいくつ飲んでいませんか”.<https://www.rad-ar.or.jp/knowledge/post?slug=polypharmacy>, (参照 2025-12-09) .