

骨格情報と LLM を用いた 個別対応型姿勢評価手法

桂陸登¹ 岡部誠¹

受付日 2025 年 8 月 31 日, 採録日 2025 年 8 月 31 日

概要: 本研究では, ユーザが自身の理想とする姿勢を基準として登録しておき, 30 分ごとにその時点での姿勢と理想の姿勢を比較することで, 姿勢の変化や崩れを検出, 指摘するシステムを提案する. システム起動時, Web カメラを起動する. ユーザはまず座った状態で自身の理想の姿勢をとり, Enter キーを押す. すると, MediaPipe により Web カメラの映像のみから Json ファイル形式で骨格情報が保存される. それから 30 分後, その時点の姿勢の骨格情報の保存が行われる. そして, 保存された骨格情報は ChatGPT に送信され, 姿勢の評価と改善案の出力が行われる. 以降, 30 分経過する度に, その時点の姿勢と理想の姿勢との比較評価が繰り返される.

キーワード: 姿勢推定, 大規模言語モデル, 健康促進

Personalized Posture Evaluation Method Using Skeletal Information and a LLM

RIKUTO KATSURA¹ MAKOTO OKABE¹

Received: August 31, 2025, Accepted: August 31, 2025

Abstract: In this study, we propose a system that allows the user to register their ideal posture as a reference and, by comparing it with their current posture every 30 minutes, detects and points out changes or deterioration in posture. When the system is launched, the webcam is activated. The user first sits in their ideal posture and presses the Enter key. Then, using MediaPipe, skeletal information is extracted solely from the webcam image and saved in JSON file format. After 30 minutes, the skeletal information of the current posture is recorded again. The saved skeletal data is then sent to ChatGPT, which outputs a posture evaluation and suggestions for improvement. This comparison between the current posture and the ideal posture is repeated at 30-minute intervals thereafter.

Keywords: Pose estimation, LLM, Health promotion

1. はじめに

近年, パソコンやスマートフォンの普及により, 長時間にわたって同じ姿勢で作業を行う機会が増加している. 特に在宅勤務やオンライン授業の普及に伴い, デスクワーク中心のライフスタイルが一般化しつつある. このような状況の中, 猫背や前傾姿勢などの悪い姿勢が慢性的な肩こりや腰痛, 椎間板ヘルニアといった健康被害の原因となることが報告されている[1]. そのため, 日常的に自分の姿勢を

モニタリングし, 必要に応じて姿勢を修正することの重要性が高まっている.

姿勢の評価を行う手法としては, 専門家による視診や, センサースーツ・加速度センサなどを用いた計測[11]が一般的である. しかし, これらの方法は高価であり, 装着が複雑であるといった課題を抱えている. そこで近年注目されているのが, カメラ映像から人の骨格座標を取得し, 姿勢を自動的に評価する方法である. 特に, OpenPose[2]や MediaPipe[3]などの骨格推定技術の発展により, 一般的なカメラ映像からリアルタイムで骨格座標を取得することが可能になってきている. なお, 本研究の初期段階では,

¹ 静岡大学
Shizuoka University, 3-5-1 Johoku, Chuo-ku, Hamamatsu, Shizuoka,
432-8561, Japan

ChatGPT[4]に2枚のWebカメラ画像(図1)を与え、自然言語によって姿勢評価を行う方法も試みた。しかし、評価結果(図2)は抽象的かつ曖昧な表現にとどまり、再現性や基準の明確さに欠ける傾向があった。これにより、定量的かつ安定した姿勢評価の実現には、骨格座標に基づく客観的な手法の導入が不可欠であると判断した。そこで本研究では、MediaPipeを用いた骨格推定によって姿勢を定量的に把握し、より明確な評価基準を構築する方針を採用した。

また、これまでの研究の多くは、身体の側面や後方からの映像を用いて姿勢を評価する手法が主流であった。しかし、一般的なノートパソコンやデスクトップパソコンに搭載されているWebカメラはユーザの正面に設置されており、側面や後方からの撮影は複数台のカメラが必要となり現実的でない場合が多い。したがって、正面からの映像のみを利用して姿勢評価を行う手法の確立は、実用的な姿勢モニタリングシステムの開発において重要な課題である。判定の手法についても、多くは一律の基準によって姿勢の良し悪しを評価しており、個人差を考慮していない。これにより、無理に一般的な良い姿勢に合わせようとし、かえって身体を痛めてしまうことも考えられる。

本研究では、Webカメラのみでユーザ自身の理想の姿勢の骨格情報を取得し評価の基準とする、個人差に対応した姿勢改善システムの構築(図3)を目的とする。本研究の成果により、専門的な機器や知識がなくとも、誰もが手軽に自身の姿勢を評価できる環境の実現が期待される。実際にユーザスタディを行い、高い評価を得た。これについては第4章で詳しく説明している。

本論文では、まず第2章において関連研究について概観し、次に第3章で本研究におけるシステムの構成と姿勢判定のアルゴリズムについて述べる。第4章では本システムの評価の結果と考察を示し、第5章で本研究のまとめと今後の課題について述べる。最後に、提案システムを用いたユーザスタディのユーザごとの結果を付録として記載する。

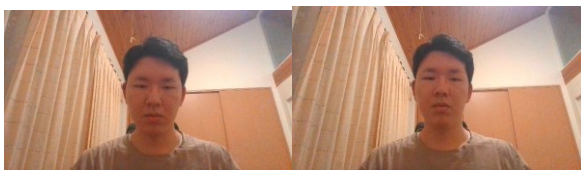


図1 ChatGPTに入力した画像：評価の基準の画像(左)と評価の対象の画像(右)

Figure 1. Images provided to ChatGPT: reference image for evaluation (left) and target image for evaluation (right)

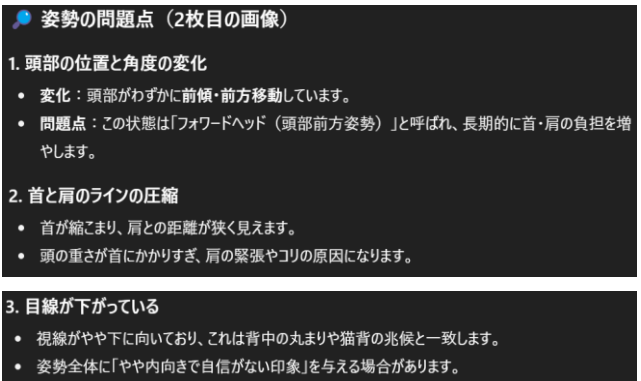


図2 画像のみで行ったChatGPTによる姿勢評価
Figure 2. Posture evaluation by ChatGPT using images only

2. 関連研究

2.1 姿勢推定・形状推定

姿勢推定や形状推定の研究は多く行われている。コンピュータビジョン、人工知能、機械学習などの最新の研究成果が発表される場であるCVPRにはこれらに関する論文が多く掲載されている。以下で、CVPRに掲載されたものを中心に姿勢推定の研究を紹介する。

小野寺ら[5]は、MediaPipeによる姿勢推定を活用し、座位姿勢の改善支援を目的とした健康支援システムを開発した。本システムでは、取得した骨格情報から姿勢を分類・評価し、ユーザにフィードバックを提供することで、良姿勢の維持を促す。さらに、評価結果は記録・可視化され、継続的な姿勢改善への動機づけとなる。一方で、姿勢評価の基準が限定的であり、多様な個人差や動的な動作に対応する柔軟性に課題が残る。

Jiong Wang, Fengyu Yangら[6]は現実世界の多様な条件下での3D人物姿勢推定をベンチマークするための新しい大規模データセット「FreeMan」を作成した。この研究は、現在のデータセットが持つシーンの多様性の不足、行動や身体スケールの限定性、スケラビリティの課題といった制約に対処しています。FreeManデータセットは、8台のスマートフォンを使用して、多様なシナリオ、アクション、照明条件、および被写体をカバーし、1100万フレームで構成されています。また、効率的なデータ収集と正確なアノテーションを可能にする半自動パイプラインを開発しました。本研究では、このデータセットを用いた単眼3D人物姿勢推定、2Dから3Dへの変換、多視点3D人物姿勢推定、およびニューラルレンダリングという4つのベンチマーク結果も提供しており、FreeManが現実世界でのアルゴリズムの性能評価に適していることを示しています。

Yasiru Ranasinghe, Deepti Hegde, Vishal M. Patelら[7]は、単眼カメラ画像からの3次元物体検出と姿勢推定を、拡散モデルを用いて実現する新しいフレームワークMonoDiffを提

案した。従来の手法では、ステレオカメラやLiDAR[8]などの追加モダリティを必要とすることが多く、コストやデータ準備の負担が課題となっていた。MonoDiffでは、ガウス混合モデル（GMM）に基づく初期化と2D再投影規制を導入することで、追加データを使用せず単眼画像のみで高精度な3D境界ボックスと姿勢を推定可能とする。KITTI[9]およびWaymoデータセット[10]での実験では、既存の最先端手法を上回る精度を達成した。今後の課題としては、リアルタイム処理性能のさらなる向上、拡散過程の高速化、より軽量のモデル設計などが挙げられ、特に自動運転など時間制約のある応用への適用に向けた最適化が期待される。

Hang Jinら[11]は、Azure Kinectという深度センサとアンサンブル学習を用いて、座っている人の姿勢分類と長時間座位の検出をリアルタイムで実現するシステムSitPoseを提案している。対象となる姿勢は、良い姿勢・猫背・前傾・左右傾き・仰向け・立位の計6種類であり、センサから得られる9か所の上半身関節座標と、それらから計算される9つの角度特徴量を用いて分類を行う。分類器にはSVM、決定木、MLPの3種を採用し、それらをソフト投票により統合することで精度を高めている。実験では33,409件のデータを用いて評価され、精度・再現率・F1スコアのいずれにおいても高い性能を示し、実時間処理に対応可能であることが確認された。成果として、実際の家庭やオフィス環境でも利用可能な姿勢検出システムを実装し、姿勢変化のアラート機能などを含むGUIベースの支援アプリケーションとして応用している。今後の課題としては、身体遮蔽や照明条件など現実環境におけるセンサ精度のばらつきへの対応、姿勢改善へのフィードバック内容の高度化、さらには学習モデルの軽量化によるエッジデバイスへの展開が挙げられる。

Yunseo Nohら[12]は、人物の性格（Human Character）と体幹（Human Core）の理解を通じて不自然な姿勢を検出する手法「H3Net」を提案した。H3Netは、従来の姿勢推定手法が3D関節情報や単純な姿勢ラベルに基づいていたのに対し、人間の体幹軸の安定性やキャラクター性（動き方の傾向）を加味することで、異常姿勢の文脈的理解を実現している。モデルはTransformerベースで構成されており、空間・時間・身体構造の各次元から特徴を抽出し、不自然な動作や異常姿勢を高精度に検出する。実験では、異常姿勢検出のベンチマークデータセットで既存手法を上回る性能を達成した。今後の課題としては、多様な性格や動作傾向を持つ個人への適応性や、日常的な軽微な異常検出への対応が挙げられる。

Shanxin Yuanら[13]によると、本研究では拡散モデル（Diffusion Model）を用いたマルチ人物姿勢推定手法「DiffusionRegPose」を提案している。従来の姿勢推定手法は検出と回帰を個別に行っていたが、本手法は拡散過程を通じて人物位置と3D姿勢を同時に推定するエンドツーエ

ンドなアプローチを特徴とする。Human3.6M[14]やCMU Panoptic[15]など複数のベンチマークで高精度な推定性能を示し、従来手法を上回る結果を達成した。一方で、拡散過程に要する計算コストや推論速度の課題が残っており、実時間処理への応用にはさらなる最適化が求められる。

2.2 様々な応用例

近年、カメラを用いない姿勢推定も行われている。カメラ映像の代わりに信号やRGB-D映像などを用いて姿勢推定を行っている研究や姿勢推定を応用した技術の研究について紹介する。

Kangwei Yanら[16]は、Wi-Fi信号（CSI: Channel State Information）を活用して、複数人物の3次元姿勢を推定するエンドツーエンドの手法を提案している。従来の姿勢推定はRGBカメラやLiDAR[8]に依存する傾向があったが、本手法は光の影響や視界遮蔽に強く、プライバシー保護の観点でも有利なWi-Fi信号を用いる点が新規性である。提案モデルは、CSI信号を受信した2次元マップを入力とし、空間特徴を抽出するバックボーンと、人物の3D位置とキーポイントを推定するPose Estimatorで構成される。損失関数には、Heatmap Loss, Depth Regression Loss, およびスケルトン構造に基づくPose Lossを統合し、人物の構造的な一貫性を考慮している。さらに、姿勢推定の精度を高めるために、複数Wi-Fiアンテナの信号を統合するマルチチャネル処理も導入されている。実験では、リアルな屋内環境で撮影されたRGB映像とCSI信号からなるWi-Poseデータセットを構築し、最大3人までのマルチパーソン3D姿勢推定を高精度に実現している。今後の課題としては、より複雑な動作への対応、デバイス間の汎化性能向上、およびリアルタイム性の強化が挙げられる。

Zhouら[17]は、RGB-D画像から未知物体の6D姿勢を推定する手法として「MatchU」が提案した。MatchUは、事前に3Dモデルを用意することなく、RGB-D画像とクエリ画像間の特徴点をマッチングし、それに基づいて物体の6D姿勢（位置・姿勢）を推定できる点に特徴がある。姿勢の推定には、対応点の幾何関係を活用した変換行列の推定が用いられる。NOCS-REAL295やYCB-Videoデータセット[18]を用いた評価では、既存手法と比較して高い精度を示した。一方で、遮蔽や背景の複雑さ、物体形状の類似性が高い場合には、マッチング精度が低下しやすく、こうした状況下での頑健性向上が今後の課題として挙げられている。

Longteng Kongら[19]によると、本研究では、カメラの内部パラメータやキャリブレーション情報が存在しない状況でも、複数視点の画像から人物の位置と姿勢を正確に登録するための枠組みが提案されている。提案手法は、Bird's Eye View（BEV）空間上における人間の空間的分布と、各カメラの相対的な構成を同時に最適化する点に特徴がある。従来の3Dポーズ推定手法がカメラキャリブレーションに依

存していたのに対し、本研究はキャリブレーション不要の柔軟な環境構築を可能にし、屋外や設置制限のある現場での応用性を高めている。Human3.6M[14]や Panoptic[15]などの大規模データセットで高精度なポーズ推定を実現した点も成果の一つである。一方で、推定精度は初期化精度や人数密度に影響されるため、多人数・密集環境下での頑健性や拡張性が今後の課題として挙げられる。

Wang ら[20]は、3D 人体メッシュ復元の精度向上を目的として、姿勢変換に着目した新手法「PostureHMR」を提案した。本手法は、入力画像の姿勢を標準姿勢（canonical pose）に変換することで、HMR（Human Mesh Recovery）モデル[21]の不確実性を低減する。具体的には、回転情報を表現する SMPL パラメータ空間において、Transformer ベースの姿勢変換モジュールを導入し、安定かつ高精度なメッシュ再構成を実現している。実験では、既存の手法と比較して、複雑なポーズや視点変化に対しても高い再現性を示した。一方で、標準姿勢への変換を前提とするため、極端なポーズや学習外分布への一般化には課題が残る。

Xu ら[22]は、レンズレスイメージング環境下での人体姿勢および形状推定を可能にするエンドツーエンド手法「LPSNet」を提案した。本手法は、レンズを用いない画像から人体の 3D ポーズとメッシュ形状を復元するもので、光線再構成と HPS（Human Pose and Shape）推定を統合した構造を特徴とする。従来困難とされていたレンズレス画像からの高精度推定を実現し、新たな応用可能性を示した。一方で、再構成精度は照明条件や装置配置に依存しやすく、現実環境への一般化には課題が残る。

3. 提案手法

3.1 システムの概要

本研究では、ユーザが自身の理想とする姿勢を基準として登録しておき、30 分ごとにその時点での姿勢と理想の姿勢を比較することで、姿勢の変化や崩れを検出、指摘するシステムを提案する。システム起動時、Web カメラを起動する。ユーザはまず座った状態で自身の理想の姿勢をとり、Enter キーを押す。すると、MediaPipe により Web カメラの映像のみから Json ファイル形式で骨格情報が保存される。それから 30 分後、その時点の姿勢の骨格情報の保存が行われる。そして、保存された骨格情報は ChatGPT に送信され、姿勢の評価と改善案の出力が行われる。以降、30 分経過する度に、その時点の姿勢と理想の姿勢との比較評価が繰り返される。以下、図 3 はシステムの概要図である。



図 3 システムの概要図

Figure 3. Overview of the system

3.2 MediaPipe

本節では、姿勢判定のための骨格座標の取得に用いる MediaPipe について紹介する。MediaPipe は Google によって開発されたオープンソースのマルチモーダル機械学習パイプラインであり、特にリアルタイムな骨格検出や手・顔のランドマーク検出などにおいて高い精度と処理速度を誇る[23]。

本研究では、MediaPipe の Pose モジュールを用いて、Web カメラ映像から 33 点の骨格ランドマーク（図 4 左）[24]をリアルタイムで取得し、姿勢評価の基礎データとして活用している。図 4 右は Web 映像に対して適用した例である。Pose モジュールは人物の頭部・肩・胴体・四肢に対応する各関節の 3 次元座標を、映像からフレームごとに取得する機能を備えている。Web カメラのみで骨格情報を得られ、処理速度が速いため、本研究における骨格特徴量の抽出に適していると考え、採用に至った。

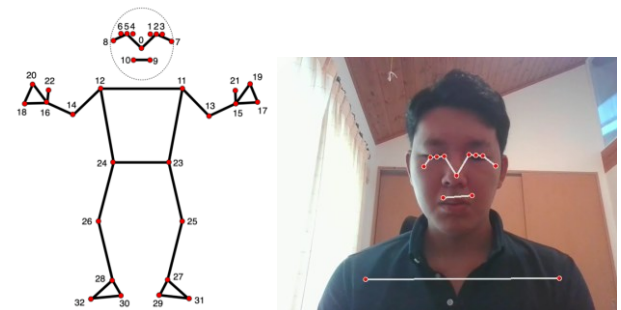


図 4 Pose モジュールの骨格ランドマーク（左）と実際の適用例（右）

Figure 4. Skeletal landmarks of the Pose module (left) and an example of its application (right)

3.3 骨格情報の保存

骨格情報の保存は、ユーザが任意のタイミングで理想の姿勢を登録することから始まる。具体的な流れを説明する。まず、ユーザは理想の姿勢を取り Enter キーを押すことで任意のタイミングで維持したい姿勢の保存を行う。座標を保存する時間は 30 秒間で、1 秒間につき 30 フレーム、つまり 900 フレーム分の座標の数値データを保存する。そして、理想の姿勢の保存から 30 分後、自動で評価する姿勢の保存が行われる。評価する姿勢も MediaPipe により 1 フレーム毎に骨格情報を保存する。

なお、姿勢の評価には 900 フレーム全てを比較するのではなく、各部位の x 座標、y 座標、z 座標それぞれの中央値を計算し、維持したい姿勢の中央値と評価する姿勢の中央値を比較することで姿勢の評価を行う。ここで、中央値を用いるのは、平均値に比べ外れ値に強いからである。これにより、30 秒の姿勢保存中に大きな動きをしても、評価に大きな影響が出ないようになっている。

3.4 ChatGPT による姿勢評価のためのプロンプト

本システムでは、3.3 にて保存された維持したい姿勢および評価する姿勢の骨格情報とプロンプト (図 5) を ChatGPT に送信し姿勢の評価を受け取る。

図 5 のプロンプトの構成について説明する。1 段落目では ChatGPT に姿勢の専門家として MediaPipe による骨格座標データを評価するよう指示している。2 段落目では、ChatGPT に入力されるデータを明示している。3 段落目では、評価内容に姿勢のずれの程度、ずれている部分の特定、改善案の提示を含めるよう指示している。4 段落目では、評価に用いる骨格情報の範囲を指定している。システムを試運転していたところ、評価で Web カメラ映像に移っていなかった膝について言及されたことから、図 2 左の 0~22 の関節の骨格情報のみを評価に用いることとした。最後の段落では、日本語で出力するよう指示している。

```
system_prompt = """You are a professional posture correction expert.
Based on 3D body landmark coordinates from MediaPipe, analyze the posture of a person.

You are given two sets of time-series 3D landmark data:
one is a 'ideal (comfortable)' posture and one is the 'current target' posture.

Compare them and:
1. Evaluate how much the target posture deviates from the ideal one.
2. Identify specific joints or body parts that are misaligned.
3. Suggest clear and simple posture improvements.

Only use landmark points with IDs from 0 to 22.
Do not use any landmarks above ID 22 — this includes hips (IDs 23 and 24),
knees, ankles, and feet. Ignore those points entirely in your analysis.

Only give practical advice. Output in Japanese."""
```

図 5 座標データの評価に用いたプロンプト

Figure 5. Prompt used for evaluating coordinate data

3.5 実装

本章では本システムの実装について説明する。まず、システムの実行は Windows 11 を搭載した Intel Core i7 (第 11 世代) ノートパソコン上で行った。メモリは 16GB、GPU は非搭載 (または内蔵 GPU) であり、すべての処理は CPU 上で実行された。また、使用したプログラミング言語は Python3.10 である。

骨格情報の取得に関する実装を説明する。はじめに、OpenCV ライブラリを用いて Web カメラを起動し、映像を 1 フレームずつ取得する。次に、MediaPipe の solutions.pose モジュールを用いて、姿勢推定用の Pose モデルを初期化し、各フレームに対して骨格推定を実行する。これにより全身で 33 点の関節の 3 次元骨格情報を取得し、JSON 形式で保存する。実際には保存は 30 秒間行われ、1 秒につき 30 フレームあることから、1 回の保存で 900 フレーム分の骨格情報が保存されることとなる。評価の際には、各関節の 3 次元の座標(x, y, z)それぞれの 900 フレーム分の中央値を抽出する。平均値より、大きな動きや保存中の離席など外れ値に対して頑健であるため、中央値を評価に取り入れる。中央値を抽出する処理は関数を用いて行い、維持したい姿勢、評価する姿勢それぞれの骨格情報の中央値を ChatGPT

に送信して姿勢評価を行う。

姿勢の評価には、OpenAI が提供する ChatGPT API を活用している。先ほど述べた姿勢の中央値を読み取り、プロンプトと共に API に送信することで、姿勢の変化や改善点に関するフィードバックを自然言語でターミナル上で取得する。最後に、30 分ごとに評価する姿勢を更新し、維持したい姿勢との比較評価を行うようループ処理を行っている。なお、API の利用にはインターネット接続と OpenAI の API キーが必要であり、プログラム実行前に適切に設定しておく必要がある。

4. 結果と考察

4.1 本システムによる姿勢評価の出力と考察

本章では、本システムにて出力された判定結果について紹介する。本システムでは動画の保存は行わないため、実際には以下のような画像が出力されることはないが、参考のため掲示する (図 6)。図 6 は維持したい姿勢と評価する姿勢の骨格情報の保存と同時に撮影した Web カメラ映像から切り抜いた画像である。左に比べ右の画像の方がやや前傾姿勢になっていることが分かる。

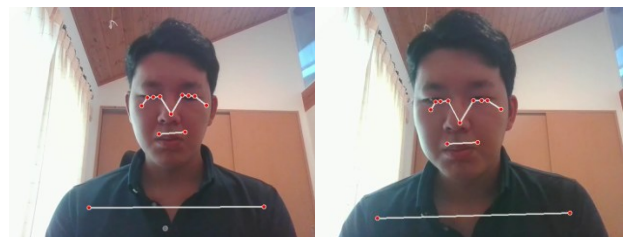


図 6 理想の姿勢 (左) と 30 分後時点の姿勢 (右)

Figure 6. Ideal posture (left) and posture after 30 minutes (right)

次に、図 6 の姿勢に対して、本システムを使用して得られた評価について説明する。

まずは、維持したい姿勢 (理想の姿勢) と評価する姿勢 (現在の姿勢) のずれについて出力される (図 7)。内容を見ると、背中と肩の高さについて、「現在の姿勢では、肩が理想の位置よりもかなり前方に出ています。特に、上半身全体が前傾しており、肩が丸まった状態になっています。これにより、背中が丸くなり、猫背のような姿勢になっています。」と出力されている。図 6 の画像を見ると、上半身が前傾しており評価が正しいことが分かる。このように画像のみでは分かりづらい変化を、3 次元の座標で傾き等を計算することで本システムでは評価することができた。続く頭の位置、腰の位置と傾き、腕の位置についても同様に、図 6 のような画像では分かりづらい変化を捉えられていることが分かる。

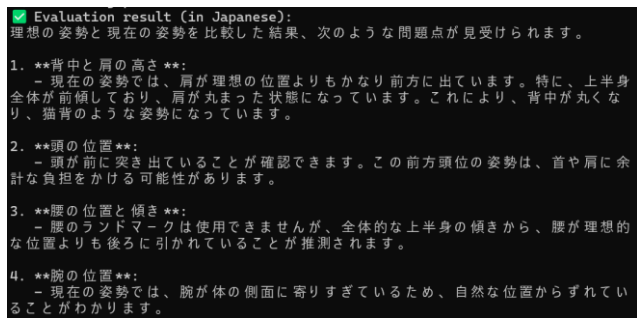


図7 部位ごとの姿勢のずれ

Figure 7. Postural deviation by body part

改善のためのアドバイスは図8のような出力となった。内容を見ると、肩甲骨を引き寄せるというトピックには、「背中の肩甲骨を軽く引き寄せるように意識し、肩を後ろに引くことで、自然な肩の位置を取り戻すことができます。」とあり、改善に向けた具体的な身体の動かし方が出力されていることが分かる。頭の位置を正す、胸を開く意識のトピックについても的確な指示が出力されており、姿勢改善の参考になったといえる。

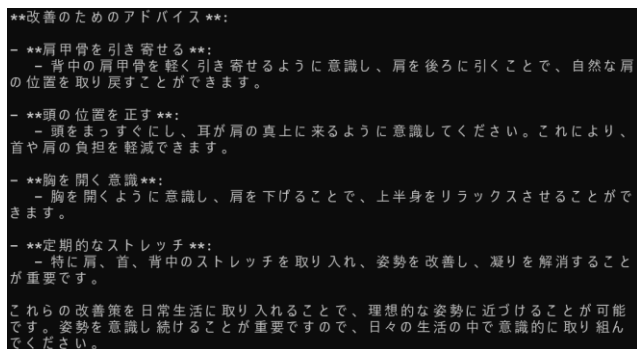


図8 改善のためのアドバイス

Figure 8. Advice for improvement

最後に図7、図8の結果に対する考察を述べる。図7の結果より、図6のような画像だけではしづらい評価を、3次元座標を用いて正しく評価できていることが分かった。図8では部位ごとの改善の動きが事細かに出力され、実際に行うことで参考になることを確認した。

また、本研究がパソコン作業中の姿勢の変化を対象に評価していることで、4.2で紹介するユーザスタディでもほとんどのユーザが上半身に前に傾いているという評価が出たが、「肩甲骨を寄せましょう」のような決まった形で出力されることはなく、個人にあった改善の動きが出力されていることが分かった。

4.2 ユーザスタディ

提案システムの有用性および使用感を評価するために行ったユーザスタディを行った。本章ではその実施方法、得られた結果、および考察について述べる。

4.2.1 実施手順

被験者は計8名であり、全員が日常的にパソコン作業を

行う大学生または大学院生である。ユーザスタディは以下の手順で実施した。なお、ユーザそれぞれの評価結果は付録に掲載した。

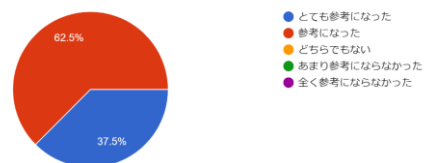
1. 動画の収集：各被験者には、通常のパソコン作業を30分間行い、その間Webカメラで自身を撮影した動画を提出してもらった。
 2. 評価の実施：提出された動画の冒頭30秒および終盤30秒を抽出し、本システムと同様の評価基準に基づき、姿勢の変化や改善点を自動で解析した。評価結果はテキストファイル形式で被験者に返却された。
 3. アンケート調査：被験者は、返却された評価結果を確認した後、アンケートに回答した。
- アンケート項目は以下の2種類である。

- ・5段階評価項目（1=全くそう思わない～5=とてもそう思う）
 - 判定結果の指摘は参考になったか
 - 改善案を試して役に立ったか
 - 実用性はあると思ったか
- ・自由記述項目
 - 本システムを使用して感じたこと、改善点など

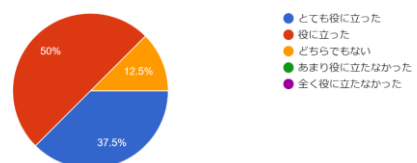
4.2.2 アンケートによる評価

アンケートの集計結果について説明する。以下は5段階評価の結果である。

判定結果の指摘は参考になりましたか？
8件の回答



改善案を試して役に立ちましたか？
8件の回答



実用性はあると思いましたか？
8件の回答

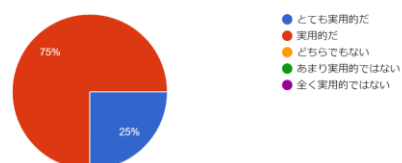


図9 アンケートの結果

Figure 9. Questionnaire results

これらを見ると、被験者に対する本システムの姿勢の評価、改善案が役に立ち、実用性があることが分かる。

また、自由記述による意見からは以下のようなポジティブな傾向が見られた：

- どこが、どのようにずれているのかが的確に出力されている。また、改善の提案に関して、「肩甲骨を寄せるイメージを持つと良い」のように、具体的な改善案で、わかりやすいと感じた。ノートパソコンのカメラで、実際は自動でやってくれるそうなので、実用的なシステムだと感じた。
- これがリアルタイム（30 分ごと）に知らせてくれるのはデスクワークが多くなる社会人とか大学生にとっては役に立つと感じた。
- 改善のアドバイスなどは、とても分かりやすく実践しやすいと思いました。
- 指定されている点や改善案は納得できるものが多かった

一方で、一部のユーザからは、「テキストだけ表示されて、リラックスした姿勢と現在の姿勢の差異を言われてもピンとこなかった。それぞれの代表的な画像も一緒に載っていると、確かにそうになっているかどうか判断できると思う。」や「1. 姿勢の逸脱評価と 2. 具体的なずれのある関節や部位において現在の姿勢が真逆の評価をされていてどちらが正しいのか不明な部分がありました。AI の回答に一貫性がないのがまだ改善点としてあげられるのではないかと思います。」のような意見もあり、改善案の視覚的な出力や AI の一貫性について考える必要性を感じた。

5. まとめと今後の展望

本研究では、姿勢の乱れがもたらす健康リスクに着目し、MediaPipeを用いた骨格情報の検出技術と ChatGPT API を組み合わせることで、リアルタイムかつ簡易に姿勢を評価できるシステムを提案した。本システムでは、30 秒間の姿勢データを取得し、はじめに保存する理想の姿勢とその 30 分後時点の姿勢を比較し、ずれと改善案の指摘を行った。システムの実行結果については、姿勢のずれについては姿勢の保存と同時に撮影した Web カメラ映像を確認することにより正しいことを評価した。また、改善のためのアドバイスについても、実行することで、参考になることを定性的に評価した。また、ユーザスタディにてアンケートを集計し、本システムに実用性があることが示せた。

本システムの課題点としては出力がテキスト形式のみであることが挙げられる。実際に、ユーザスタディのアンケートでも「テキストだけではピンと来ない」というような意見も見受けられた。また、定性的な評価しか行っていないため、AI の信頼性や評価の一貫性といった課題が残る。

今後は、改善のためのアドバイスを 3D 人体モデルによる

アニメーションや画像を用いて出力することで、視覚的に分かるシステムにし、日常生活で使えるアプリケーションへの応用が考えられる。また、姿勢の長期的な変化を追跡できるようなフィードバック機能や、個人の姿勢履歴を記録し、評価に活用するようなパーソナライズ機能の追加を目標としている。

参考文献

- [1] スポーツ庁：「日本人の座位時間は世界最長「7」時間！座りすぎが健康リスクを高める あなたは大丈夫？その対策とは・・・」（オンライン），DEPORTARE，2019-10-11，入手先〈<https://sports.go.jp/special/value-sports/7.html>〉（参照 2025-07-20）。
- [2] OpenPose（オンライン），入手先〈<https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/index.html>〉（参照 2025-07-14）。
- [3] Google AI for Developers，“MediaPipe ソリューションガイド”（オンライン），2025-02-26，入手先〈<https://ai.google.dev/edge/mediapipe/solutions/guide?hl=ja>〉（参照 2025-07-14）。
- [4] OpenAI，ChatGPT（オンライン），入手先〈<https://openai.com/ja-JP/chatgpt/overview/>〉（参照 2025-07-14）。
- [5] 小野寺 拓実，福田 直人，斉藤 大樹，五十嵐 透，“姿勢推定 AI を活用した健康支援システムの開発”，IEICE Technical Report, 2023.
- [6] Jiong Wang, Fengyu Yang, Bingliang Li, Wenbo Gou, Danqi Yan, Ailing Zeng, Yijun Gao, Junle Wang, Yanqing Jing, Ruimao Zhang, “FreeMan: Towards Benchmarking 3D Human Pose Estimation under Real-World Conditions”, In CVPR 2024.
- [7] Yasiru Ranasinghe, Deepti Hegde, Vishal M. Patel, “MonoDiff: Monocular 3D Object Detection with Diffusion Models”, In CVPR 2024.
- [8] AIST：産業技術総合研究所，“#話題の〇〇を解説 “LiDAR” とは？ 科学の目でみる，社会が注目する本当の理由”（オンライン），産総研マガジン，入手先〈https://www.aist.go.jp/aist_j/magazine/20220928.html〉（参照 2025-07-28）。
- [9] KIT，“The KITTI Vision Benchmark Suite”（オンライン），入手先〈<https://www.cvlibs.net/datasets/kitti/>〉（参照 2025-07-28）。
- [10] WAYMO，“WAYMO Open Dataset”（オンライン），入手先〈<https://waymo.com/open/>〉（参照 2025-07-28）。
- [11] Hang Jin, Xin He, Lingyun Wang, Yujun Zhu, Weiwei Jiang, Xiaobo Zhou, “SitPose: Real-Time Detection of Sitting Posture and Sedentary Behavior Using Ensemble Learning With Depth Sensor”, In CVPR 2024.
- [12] Seungha Noh, Kangmin Bae, Yuseok Bae, Byong-Dai Lee, “H3Net: Irregular Posture Detection by Understanding Human

- Character and Core Structures”, In CVPR 2024.
- [13] Dayi Tan , Hansheng Chen , Wei Tian , Lu Xiong ,
“DiffusionRegPose: Enhancing Multi-Person Pose Estimation
using a Diffusion-Based End-to-End Regression Approach”, In
CVPR 2024.
- [14] Institute of Mathematics of the Romanian Academy ,
“Human3.6M Dataset” (オンライン) , 入 手 先
〈 <http://vision.imar.ro/human3.6m/description.php> 〉 (参 照
2025-07-28).
- [15] CMU Panoptic Dataset, “CMU Panoptic Dataset” (オンライン) ,
入 手 先 〈 <http://domedb.perception.cs.cmu.edu/> 〉 (参 照
2025-07-28).
- [16] Kangwei Yan, Fei Wang, Bo Qian, Han Ding, Jinsong Han,
Xing Wei, “Person-in-WiFi 3D: End-to-End Multi-Person 3D
Pose Estimation with Wi-Fi”, In CVPR 2024.
- [17] Junwen Huang, Hao Yu, Kuan-Ting Yu, Nassir Navab, Slobodan
Ilic, Benjamin Busam, “MatchU: Matching Unseen Objects for
6D Pose Estimation from RGB-D Images”, In CVPR 2024.
- [18] GitHub , “YCB_Video_toolbox” (オンライン) ,
〈 https://github.com/yuxng/YCB_Video_toolbox 〉 (参 照
2025-07-28).
- [19] Zekun Qian, Ruize Han, Wei Feng, Song Wang, “From a Bird’s
Eye View to See: Joint Camera and Subject Registration without
the Camera Calibration”, In CVPR 2024.
- [20] Yu-Pei Song, Xiao Wu, Zhaoquan Yuan, Jian-Jun Qiao, Qiang
Peng, “PostureHMR: Posture Transformation for 3D Human
Mesh Recovery”, In CVPR 2024.
- [21] Axross Recipe, “HMR2.0(4D Humans)を用いて 3D Human
Mesh で人物を追跡するレシピ” (オンライン) , 入 手 先
〈 <https://axross-recipe.com/recipes/1176> 〉 (参 照 2025-07-28).
- [22] Haoyang Ge, Qiao Feng, Hailong Jia, Xiongzheng Li, Xiangjun
Yin, You Zhou, Jingyu Yang, Kun Li, “LPSNet: End-to-End
Human Pose and Shape Estimation with Lensless Imaging”, In
CVPR 2024.
- [23] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan
“MediaPipe: A Framework for Building Perception Pipelines.”
arXiv preprint arXiv:1906.08172, 2019.
- [24] Google AI for Developers, “姿勢ランドマーク検出ガイド”,
2025-01-13 , 入 手 先
〈 https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker?hl=ja 〉 (参 照 2025-07-14) .