

自動生成された評価ルールを用いたリアルタイム筋力トレーニング指導システム

鈴木貴之^{†1} 岡部誠^{†1}

概要: 健康意識の高まりに伴い自宅でのトレーニングが普及しているが、専門家の指導がない環境では、不適切なフォームによる怪我のリスクや運動効果の低下が課題となっている。本研究では、ユーザーが提供した手本動画を学習する筋力トレーニングのためのリアルタイムフィードバックシステムを提案する。動画から抽出した「最適な1回」のデータを基準とし、LLMによる評価ルールの生成と、動的時間伸縮法(DTW)による同期を組み合わせることでリアルタイムにフィードバックするシステムを実現した。これにより、従来のルールベース方式では困難であった「種目の多様性」と「ユーザー個別の身体特性」に対応した、具体的なパーソナルトレーニング体験を提供する。

1. はじめに

近年、予防医学や健康寿命向上への関心からフィットネス市場は急速な拡大を見せている。米国スポーツ医学会(ACSM)の2024年トレンド調査においても「ウェアラブル技術」や「モバイルエクササイズアプリ」が上位を占めており[1]、テクノロジーを用いた客観的な運動管理への需要は社会的標準となりつつある。特に、場所や時間を選ばない自宅トレーニングは有効な手段として広く定着している。

しかし、専門家の監視がない環境でのトレーニングには、誤ったフォームによる筋骨格系の怪我のリスクが伴う。このリスクに対し、既存のフォーム評価システムの多くは、開発者が事前に定義した固定的な判定を行うアプローチが一般的である。しかし、評価すべき「重要な関節」や「判定ルール」は種目ごとに大きく異なるため、開発者が事前にルールを定義していない種目には対応できないという課題がある。この課題を解決する手段として、近年登場したビデオ言語モデル(VLM)による指導は汎用性が高いものの[2]、計算コストと遅延が大きい。また、動画から解析した各関節の角度や座標のデータをLLMに送ることで言語フィードバックを得る手法もある[3]が、LLMの応答に時間がかかるため、これらの手法はリアルタイムのフィードバックには不向きである。

そこで本研究では、ユーザーが用意した手本動画を学習し、その動作に特化した指導を行うシステムを提案する。本システムは、LLMを用いて動作の意味を解釈して評価ルールを自動生成し、実行時には軽量なアルゴリズムで動作判定を行うハイブリッド構成をとる。これにより、「多様な種目への適応」と「リアルタイム指導」の両立を目指す。

2. ユーザーインターフェース(UI)

本システムのユーザーインターフェース(UI)は、

「手本登録」「登録データの確認」「トレーニング実行」の3つの画面から構成される。

ユーザーは、これらの画面を順に利用することで、手本動作の登録から、実際のトレーニングとフィードバックまでを直感的に行うことができる。

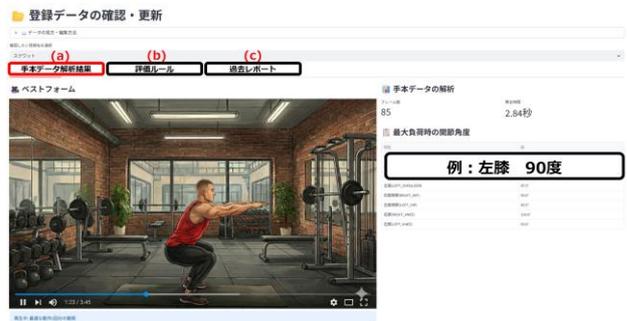


図1 登録データの確認画面のUI

(1) 手本登録画面と登録データの確認画面

まずユーザーは、「手本登録」画面において、行いたいトレーニング種目名と手本となる動画ファイルを入力する。

本論文では例としてスクワットを取り上げる。ユーザーは、YouTubeなどで公開されているプロトレーナーによるスクワットの解説動画などを、手本動画としてシステムに入力する。手本動画を登録すると、その動画から抽出された情報が「登録データの確認・更新」画面(図1)に表示される。本画面では、機能ごとにタブが分かれており、「手本データ解析結果タブ」や「評価ルールタブ」などを確認することができる。

「手本データ解析結果タブ」(図1(a))では、画面左側に動作1回分の手本動画が表示され、右側には解析結果が数値として一覧表示される。ここでは、1回の動作にかかる時間や、動作中に重要となる関節角度などを、ユーザーが直感的に把握できる。例えばスクワットの場合、1回の動作

^{†1} 静岡大学

時間は2.84秒となっており、また、最も負荷がかかる姿勢における膝の角度は90度や腰の角度は90度などが表示されている。つまり、ユーザーは「どのようなフォームが基準になっているのか」を視覚的に確認できる。

「評価ルールタブ」(図1(b))では、トレーニング中に提示されるフィードバック内容が一覧として表示される。例えば、「腰が丸まっている」「膝の曲げが浅い」といった指摘がどのような条件で提示されるのかを、事前に確認することが可能である。

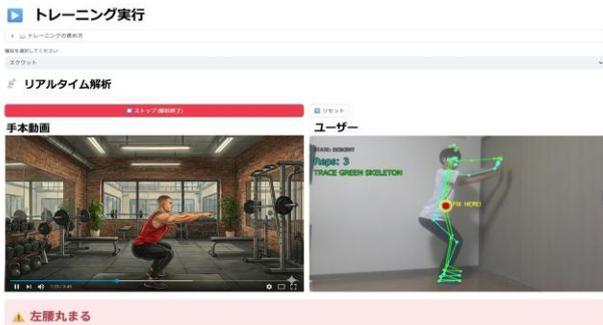


図2 トレーニング実施画面のUI

(2) トレーニング実施画面

「トレーニング実行」画面(図2)では、ユーザーがスタートボタンを押すことでトレーニングを開始する。

画面左側には登録した手本動画が表示され、右側にはWebカメラを通して取得したユーザー自身の映像がリアルタイムで表示される。これにより、ユーザーは手本動作と自分の動作を同時に比較しながらトレーニングを行うことができる。ユーザーの映像上には、手本動作の姿勢を示す半透明の骨格モデル(ゴースト)が重ねて表示される(図3(a))。ユーザーがゴーストに合わせた姿勢を取ることで、トレーニングが開始される。2回目以降の動作では、ユーザーの動作に合わせてゴーストがリアルタイムに同期して動く(図3(b))。これにより、ユーザーは鏡を見る感覚で、自身のフォームと手本動作との差を視覚的に確認しながら、ペースの維持やフォーム修正を行うことができる。フォームの乱れが検知された場合には、該当する関節部位(例:膝や腰)が画面上で強調表示される。同時に、「左腰が丸まっている」といったフィードバックが画面下部に表示され、音声による読み上げも行われる。これにより、ユーザーは視覚と聴覚の両面から、直感的に指導を受けることができる。

トレーニング終了後には、全体の動作を振り返るための結果画面が表示される。ここでは、フォームの崩れやすかった傾向や、次回に向けた改善点が提示され、ユーザーは自身のトレーニング内容を客観的に把握できる。これらの結果は、登録データの確認画面にある「過去のレポート」からも、いつでも確認可能である。

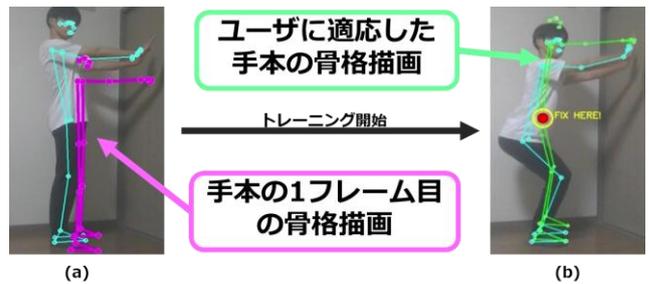


図3 半透明の手本の骨格モデル

スクワット16回、お疲れ様でした！最後までやり遂げたことは素晴らしいです。データから見ると、中盤までは比較的安定した深さでトレーニングできていましたね。特に、前半の数回は動作の時間も安定しており、集中力が見て取れます。しかし、中盤から「右肘の伸展」が頻繁に見られ、後半には「左体幹が起きすぎる」という左右のバランスの崩れも現れました。特に最後の16回目は深さが大幅に浅くなり、疲労によるフォームの崩れが顕著です。次回は、バーを握る力を意識的に抜いて両肘をリラックスさせ、腹圧を保ちながら左右均等に体幹を使い、深さを維持できる回数でセットを終えることを目標にしましょう。

図4 トレーニング実行後のフィードバック内容

3. システム実装

本システムは、入力デバイスとして単眼Webカメラのみを使用し、Streamlit上で動作する。処理フローは以下の通りである。また、本システムで用いるLLMはGemini API(gemini-2.5-flash)である。

(1) 評価ルールの定義

本システムにおける「評価ルール」とは、入力された手本動画の動作特性をLLMが解釈し、計算可能なパラメータセットとして構造化したものである。具体的には、まず動作の周期性を捉えて回数をカウントするための「基準点(例:腰の高さ)」が定義される。その上で、フォームの良し悪しを決定づける「各関節の重要度」、動的時間伸縮法(DTW)[4]で特徴量として用いる「関節名」、および手本軌道から許容される「誤差範囲」が設定される。さらに、単なる数値判定にとどまらず、ズレの方向に応じた具体的な「フィードバック内容(例:鋭角すぎる場合は『腰が丸まっている』と指摘)」や、関節が重要となる「動作フェーズ(例:ボトム時)」も記述される。これらを一つの構造化データとして集約することで、種目ごとに異なる「正解」の定義を、汎用的なアルゴリズムで処理可能な形式へと変換している。

(2) 手本データの抽出と評価ルールの生成

まずLLMを介して、入力された種目名から動画解析のための主要関節(例:スクワットなら腰)を決定する。次に、入力された動画から主要関節のY座標(高さ)を抽出し、ピーク検出アルゴリズムを用いて動作を1回ごとに分割する。分割された動作の中から、平均的な軌道に最も近いデータを「手本データ」として採用する。

最後に、手本データのボトムフレーム（最大負荷時）の画像と関節角度情報を LLM に入力し、種目特有のフィードバック内容や関節ごとの重要度を設定ファイルとして出力する。

(3) リアルタイム同期とフィードバック

トレーニング実行中は、MediaPipe を用いてユーザーの骨格座標を取得する。動作の同期には、動的時間伸縮法（DTW）を用いる。これにより、動作速度の個人差を吸収し、フレーム単位での正確な対応付けを行う。フィードバック内容は、同期された全フレームにおける手本との角度差分に対し、LLM が設定した「関節ごとの重要度」と「フェーズ（下降・ボトム・上昇）の重み」を掛け合わせたスコアを動作 1 回分の全フレームにわたり累積加算することで決定する。具体的には、評価項目（例：「左膝の角度」）ごとに累積スコアを算出し、動作終了時にそのスコアが最も高い項目を「最重要な改善点」としてフィードバックを行う。これにより、一瞬のノイズに過剰反応することなく、動作全体を通して最も修正すべき癖を的確に指摘することが可能となる。また、負荷の高いフェーズの重みを大きく設定することで、重要な局面でのミスを優先的に指摘する仕組みとなっている。フィードバック時に用いる音声合成は、Windows OS に標準搭載されている音声合成エンジン Microsoft Speech API (SAPI) を使用している。

(4) LLM による事後フィードバック

トレーニング実行後、システムは「動作 1 回にかかる時間」や「最も頻発したミス」、「主要関節の高さの推移」の集計データを再び LLM に入力する。これにより、「後半に疲れから腰が浮きやすくなっています」といった、セッション全体を通じた傾向分析や、次回に向けた改善アドバイスを生成し、ユーザーに提示する。

4. まとめ

本研究では、任意の動画を手本として利用可能な適応型フォーム矯正システムを開発した。LLM による「ルールの構造化」と DTW による「時間的同期」を組み合わせることで、専門知識や事前学習を必要とせず、ユーザーの理想とするフォームに基づいた直感的な指導を実現した。現在の課題は、ランジ等の左右非対称な反復動作への対応である。今後は、動作解析ロジックの拡張による多種目への適応や、ユーザーテストを行いフォーム改善効果の定量的検証を行うことで、本システムの実用性を高めていく方針である。

参考文献

- [1] A'Naja, M. N., Reed, R., Sansone, J., Batrakoulis, A., McAvoy, C., & Parrott, M. W.. 2024 ACSM worldwide fitness trends: future directions of the health and fitness industry. ACSM's Health & Fitness Journal. 2024, 28(1), 14-26.
- [2] Ashutosh, K., Nagarajan, T., Pavlakos, G., Kitani, K., and Grauman, K.. ExpertAF: Expert actionable feedback from video. In Proceedings of the Computer Vision and Pattern Recognition

Conference. 2025, pp.13582-13594.

- [3] 高橋汰汰, 岡部誠. 後方伸身宙返りの段階的学習をサポートする練習支援システム. WISS 2025, 2025, p.142-149.
- [4] S. Sempena, N. U. Maulidevi, and P. R. Aryan. Human action recognition using dynamic time warping. In Proc. ICEEI. 2011, pp.1-5, IEEE.