

Computer Vision-based Action Quality Assessment
Incorporating Human Expert's Knowledge:
Focusing on the Figure Skating Jump

July 2024

Seiji Hirose

| 報告番号 | ㊦乙 第 号 | 氏 名 | 廣澤 聖士 |
|--|--------|-----|-------|
| <p>主 論 文 題 名 : Computer Vision-based Action Quality Assessment Incorporating Human Expert's Knowledge: Focusing on the Figure Skating Jump (人間の専門知識を適用したコンピュータビジョン手法に基づく身体動作の質の評価—フィギュアスケートのジャンプを対象に)</p> | | | |
| <p>(内容の要旨)</p> <p>コンピュータビジョン分野における Action Quality Assessment (AQA) タスクは、映像中の人間の動作の品質を評価するものであり、特にスポーツパフォーマンスの評価への応用が期待されている。これまでの AQA 研究では、フィギュアスケートのプログラム全体の映像から最終得点を予測する取り組みが見られたが、個々のジャンプスコアを予測した先行研究はない。ジャンプの得点は、質を示す GOE (Grade of Execution) と難易度に基づく基礎値の合計によって決まるため、高い GOE を獲得する質の高いジャンプは競技で勝利するために不可欠である。しかし、GOE の基準は解釈の余地があり、審判による主観的な評価である。ジャンプの質に関わる特徴を明らかにし、精度の高い予測モデルを開発することは、競技関係者にとって有益だろう。本研究では、トラッキングシステムから得られる運動学的特徴量に加えて、映像特徴量の入力時に競技の専門家である審判と選手の視線情報を活用して冗長な情報を削減し、精度を向上させるモデルを提案することを目的とした。具体的には以下の3つの検討事項に取り組んだ。最初に2次元 Direct Linear Transformation 法を用いたハンドクラフトベースの運動学的特徴量に基づく機械学習モデルの開発を行った。次に出来栄評価を行う際の専門家(選手・審判)の視覚行動の違いを明らかにした。最後に運動学的特徴量に加えて、専門家の視線を学習することで情報削減を行ったディープラーニングベースのモデルを提案した。</p> <p>第1章では、本研究の背景について概説した。</p> <p>第2章では、3つの検討事項に基づく関連研究について概説した。</p> <p>第3章では、ハンドクラフトベースの運動学的特徴量に基づき、機械学習モデルを用いて GOE に影響を与える運動学的特徴を明らかにした。具体的には、2019年世界選手権女子ショートプログラムにおけるダブルアクセルジャンプの水平飛距離、最大跳躍高、着氷後の滑走速度を取得し、GOE に影響する3つの変数の対比を明らかにするために主成分回帰分析を行った。その結果、水平距離と着氷後速度が跳躍高に比べて大きいこと、および着氷速度に比べて水平飛距離が大きいことの2つの主成分が有意であり、高い GOE に寄与することが明らかとなった。一方で、第3章のアプローチでは解釈しやすい特徴量とモデルを使用しているものの、得られた特徴量が大きく関与するのは6つの評価基準のうち2つであり、残りの評価基準については人間が特徴量を作成することが困難であると考えられる。そこで、放送映像の特徴量を入力特徴量に加えることで、より包括的な評価基準を考慮できるモデルを提案する。放映映像には評価に不要な情報が含まれているにもかかわらず、人間の専門家がジャンプを評価する際には、視覚的に情報を絞り込んで削減することが可能であり、模倣した機構を取り入れることで精度向上が期待できる。</p> <p>第4章では、審判と選手がジャンプを評価する際の視線配置の特徴を明らかにし、視線位置を利用して映像から得られる特徴量を削減する機構を取り入れた予測モデルを提案した。前提となる専門家の視線配置について、選手群と審判群は GOE 評価の判定精度に有意差はないが、視線配置には有意差があり、両者とも上半身に最も視線を配置しつつも、選手は顔に、審判は下肢により注目することが明らかになった。これらの視線位置をモデルに適用したところ、両群の専門家の視線位置を利用した場合に人間の専門家の予測精度とベースラインモデルの精度を上回っており、人間の専門家の知識と数理モデルの能力を組み合わせることで、より高い精度が得られることを実証した。</p> <p>第5章には、結論として各検討事項で得られた内容をまとめ、研究の成果を要約した。</p> | | | |

Thesis Abstract

No. _____

| | | | |
|--|---|------|-----------------|
| Registration Number | <input checked="" type="checkbox"/> “KOU” <input type="checkbox"/> “OTSU” No. _____ *Office use only | Name | HIROSAWA, Seiji |
| Thesis Title | | | |
| Computer Vision-based Action Quality Assessment Incorporating Human Expert’s Knowledge: Focusing on the Figure Skating Jump | | | |
| Thesis Summary | | | |
| <p>In computer vision, action quality assessment (AQA) tasks evaluate video action quality and can be applied to sports performance evaluation. A common application is the prediction of final scores from entire figure skating programs. However, no previous studies have predicted individual jump scores, which are important to competitors due to high competition weight. Jump scores are calculated by the sum of base values based on difficulty and grade of execution (GOE), indicating performance quality. Therefore, performing high-quality jumps to achieve a high GOE is crucial for winning competitions. However, the GOE criteria are open to interpretation and clarifying the factors influencing jump quality, and developing a prediction model benefits skaters and coaches. First, we proposed a handcraft-feature-based model using the 2D Direct Linear Transformation method. Second, we clarified the differences in visual behavior between skaters and judges and proposed a deep learning-based model with information reduction utilizing the expert's gaze.</p> <p>Chapter 1 summarizes the background.</p> <p>Chapter 2 outlines related work based on the three points for consideration.</p> <p>Chapter 3 proposes a machine learning model to identify how kinematic features influence GOE. Specifically, we obtained the double axel jump's horizontal distance, vertical height, and landing speed in the ladies' short program at the 2019 World Championships. The results demonstrated that greater horizontal distance and landing speed compared with vertical height and greater horizontal distance compared with landing speed contributed to higher GOE. Although the approach in Chapter 3 uses features and models that are easy to interpret, the handcraft-based features meet only two GOE evaluation criteria, and creating features for the remaining evaluation criteria would be challenging for humans.</p> <p>Chapter 4 proposes a model that can consider comprehensive evaluation criteria by adding broadcast video to the input feature. Although the video contains unnecessary information, human specialists can visually reduce information when evaluating jumps. First, we clarified the eye movements of judges and skaters and proposed a model that uses the specialists’ gaze location. Results showed skaters focused more on the face, whereas judges focused on the lower extremities. Both specialists’ gaze locations were applied to the model, demonstrating the highest accuracy. The model outperformed specialists’ predictions, and the baseline model indicated that combining human specialist knowledge with mathematic model capabilities can yield higher accuracy.</p> <p>Chapter 5 summarizes the results of this study.</p> | | | |