

Optimierung der vollautomatischen zeitkontinuierlichen Erkennung der Körperpose und Skiposition von Skispringern in Videoaufnahmen

– Eine weiterführende Überarbeitung des aktuellen Systems zur Schätzung der Absprung- und Flugparameter für einen verbesserten und autonomeren Praxiseinsatz

(AZ 072005/21-22)

Katja Ludwig¹, Rainer Lienhart¹ (Projektleitung), Sören Müller² & Sascha Kreibich²

¹Universität Augsburg, Lehrstuhl für maschinelles Lernen und Sehen

²Institut für Angewandte Trainingswissenschaft

1 Problem

Für die Leistungsdiagnostik von Skispringerinnen und Skispringern ist eine genaue Analyse der Körperposition während der Flugphase essentiell. Dabei sind die Positionen von Körperpunkten und daraus abgeleitete Körperwinkel relevant. Um Athletinnen und Athleten sowie Trainerinnen und Trainer zu entlasten, wurde ein für den Praxiseinsatz reifes System entwickelt, welches mittels hochmoderner tiefer neuronaler Netze wichtige Körperpunkte wie Gelenke und Lage der Ski automatisch Bild für Bild in Videos von Skispringern von der Seite bestimmt. Aus diesen erkannten Punkten werden im Anschluss die Flugparameter wie Skianstellwinkel und Flugkurve des Körperschwerpunktes der Skispringer automatisiert bestimmt.

Umfangreiche quantitative Analysen haben gezeigt, dass das aktuelle System zur Schätzung der abgeleiteten Leistungsparameter i. d. R. robust schätzt, aber unter schwierigen Bildbedingungen mit gelegentlichen Aussetzern bei den Ergebnissen zu rechnen ist. Dies trübt den regelmäßigen Einsatz, da nur manuell überprüfte Ergebnisse mit gutem Gewissen an die Trainerinnen und Trainer weitergeleitet werden können. Außerdem wäre es wünschenswert, zusätzliche Punkte auf den Skispringern/-springerinnen bzw. den Skiern definieren und erkennen zu können, ohne umfangreiche manuelle

Annotationen wie für die bisherigen Körperpunkte bereitstellen zu müssen.

Ziel des Service-Forschungsprojekts ist es deshalb, das aktuelle System weiter für den alltäglichen, autonomen Regelbetrieb zu verbessern, so dass der Aufwand für die manuelle Prüfung und eventuelle Korrektur auf ein Minimum sinkt. Zusätzlich sollen für die Leistungsdiagnostik Experimentiermöglichkeiten mit zusätzlichen, leicht zu definierenden Punkten auf Körper und Skiern ermöglicht werden.

2 Methode

Um das aktuelle System besser auf das Körpermodell anzupassen sowie eine Leistungsdiagnostik mit mehr Punkten zu ermöglichen, wurde zunächst eine kleine Menge an neuen Punkten erstellt, um die Möglichkeiten des Verfahrens auszuprobieren. Gleichmäßig verteilt wurden auf dem Nacken, Oberarm, Unterarm, Oberschenkel, Unterschenkel und Torso jeweils 3 zusätzliche Punkte erstellt und auf dem Ski jeweils 9 Punkte. Die Punkte wurden linear interpoliert zwischen die annotierten Körperpunkte gelegt. Entsprechend dieses Systems können beliebig viele Punkte mehr generiert werden. Ergebnisse haben jedoch gezeigt, dass diese Variante nicht zielführend ist, um beliebige Punkte zu erkennen. Deshalb wurde eine neue Darstellung entwickelt.

Dabei wird ein Vektor verwendet, der als Länge die Anzahl der annotierten Körperpunkte besitzt, also 11. Soll nun ein annotierter Körperpunkt x erkannt werden, enthält der Vektor nur Nullen und an der Stelle, die diesem Körperpunkt entspricht, eine 1 (soll der zweite Körperpunkt erkannt werden, wird an der zweiten Stelle eine 1 gesetzt). Für einen Körperpunkt zwischen zwei Punkten wird in den Vektor bei den Stellen der entsprechenden Punkte der Anteil an der Strecke geschrieben, wo der Körperpunkt liegen soll. Soll beispielsweise ein Körperpunkt erkannt werden, der zwischen Körperpunkt 3 und 4 liegt, und zwar zu einem Viertel auf der Strecke zwischen 3 und 4 (also näher an 3), dann wird bei Körperpunkt 3 0.75 eingetragen und bei Körperpunkt 4 0.25. Weitere Beispiele können der Abb. 1 entnommen werden. Mit dieser Darstellung können nun beliebige Zwischenpunkte, und damit unendlich viele, erkannt werden. Während des Trainings des Modells werden zufällige Zwischenpunkte generiert, um dem Modell die Logik dieser Vektoren beizubringen. Aus den vorhandenen annotierten Standardpunkten werden ebenfalls zufällig Körperpunkte ausgewählt, damit das Modell nicht nur mit diesen Punkten, sondern eben mit allen Varianten, funktioniert.

3 Ergebnisse

Zunächst wurde ein Training mit der kleinen Menge an neuen Punkten, wie oben beschrieben, durchgeführt. Die Ergebnisse für die neu definierten Punkte waren zufriedenstellend, jedoch wurde festgestellt, dass man keinen Zusammenhang zwischen benachbarten Körperpunkten in der internen Repräsentation des Netzes finden kann. Auf den 11 Standardpunkten konnte eine Percentage of Correct Keypoints (PCK) von 81,1 % erreicht werden, jedoch nur bei der Berechnung aller Punkte. Wenn nur die Standardpunkte berechnet werden sollen, wird nur eine PCK von 14,4 % erreicht. Auf allen Punkten (Standard- und zusätzliche Punkte) erreicht das System eine PCK von 86,0 %. Mit der bisher verwendeten Methode würde das bedeuten, dass mit der Anzahl der zusätzlichen Körperpunkten auch der Aufwand für das Netz steigt und man nicht dazu kommen kann, beliebige Zwischenpunkte zu erkennen. Ein weiteres Problem war die Abhängigkeit von den vorhandenen Körperpunkten. Man konnte entweder alle Punkte erkennen oder keine, es war nicht möglich, nur eine Untermenge der Punkte zu erkennen.

head	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
shoulder	0	1	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0
elbow	0	0	1	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0
hand	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
hip	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0.8
knee	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0.2
ankle	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
right ski tip	0	0	0	0	0	0	0	1	0	0	0	0	0.5	0.2	0	0	0
right ski tail	0	0	0	0	0	0	0	0	1	0	0	0	0.5	0.8	0	0	0
left ski tip	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0.7	0.5	0
left ski tail	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.3	0.5	0

Abb. 1: Beispiele für die Darstellung variabler Punkte in Vektorenform. Die ersten 11 Punkte sind die Standardpunkte. Deshalb sind alle Einträge des Vektors von Körperpunkt i Nullen bis auf den Eintrag an der Stelle i . Gemischte Punkte enthalten Anteile der Standardpunkte, die aufsummiert 1 ergeben müssen.

Auch dieses Problem konnte mit der neuen Darstellung behoben werden. Mit diesem Modell wird eine PCK von 81,8 % erreicht, egal ob nur die Standardpunkte oder auch zufällig generierte Punkte ausgewertet werden. Eine Auswertung aus Standard- und zusätzlichen Punkten (die jetzt aber zufällig sind und nicht fest, und damit für das Modell schwieriger zu erkennen) ergibt eine PCK von 84.4 %. Beispiele für Detektionen von Zwischenpunkten auf den Ski können Abb. 2 entnommen werden.

Abb. 3b (rechts) zeigt die Ergebnisse des neuen Systems in beiden Varianten. Zur Evaluation wird dabei die Metrik „PCA #“ (Percentage of Correct Angles) verwendet. Ein Winkel wird von dieser Metrik als korrekt bezüglich eines Schwellwertes # angesehen, wenn der Ergebnswinkel des vorgestellten Systems in Grad nicht mehr als # vom berechneten Winkel aus den annotierten Daten abweicht.

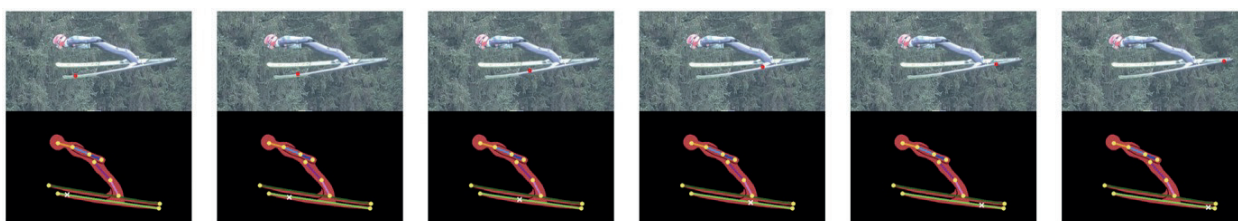


Abb. 2: Beispiele für Detektionen von Zwischenpunkten. Der gewünschte Punkt ist mit einem weißen Kreuz in der Silhouette markiert, der detektierte Punkt in Rot.

Unten ist der gewünschte Punkt in der Silhouette mit einem weißen Kreuz eingezeichnet, der vom Modell vorhergesagte Punkt ist im oberen Bild rot markiert. Zusätzlich wurde das System noch weiter angepasst für die Verwendung mit hochauflösenden Kameras. Am Ende konnten zwei unterschiedliche Systeme, eines basierend auf einem Faltungsnetz und eines basierend auf einem Transformer-Netz, erstellt werden. Beide Varianten erzielen bessere Ergebnisse als das bisherige System. Abb. 3a (links) zeigt die Ergebnisse des vorherigen Systems, nur ausgewertet auf hochauflösenden Bildern.

4 Diskussion

Im Vergleich zum bisherigen System erlangt das neue System eine höhere Genauigkeit. Das Faltungsnetz liefert durchwegs bessere Ergebnisse als das bisherige System, und erkennt sogar fast alle Flugparameter mit einer maximalen Abweichung von 3°.

Das Transformer-Netz liefert ebenfalls gute Ergebnisse, es ist etwas schlechter in der Präzision bei maximal 1° Winkelabweichung, jedoch erzielt es ab 2° Winkelabweichung bessere Ergebnisse.

Parameter	PCA 1	PCA 2	PCA 3	Parameter	PCA 1	PCA 2	PCA 3	Parameter	PCA 1	PCA 2	PCA 3
Lower Body	79.1	94.6	98.0	Unterkörper	78.5	97.7	99.4	Unterkörper	71.3	94.6	98.9
Upper Body	76.5	93.7	97.1	Oberkörper	82.2	97.7	100.0	Oberkörper	70.5	96.6	99.4
Body	91.4	96.8	97.7	Gesamtkörper	94.8	99.4	100.0	Gesamtkörper	92.3	99.4	100.0
Ski	91.4	96.3	98.0	Ski	96.8	98.6	99.4	Ski	94.6	99.4	99.4
Diff. LB/Ski	78.2	93.4	96.6	Diff. UK/Ski	79.9	95.1	98.9	Diff. UK/Skii	69.9	94.3	98.3
Gesamt	83.6	95.2	97.7	Gesamt	86.7	97.9	99.7	Gesamt	79.9	97.1	99.4

Abb. 3: Übersicht der PCA-Werte der einzelnen Flugparameter bei den Schwellwerten 1°, 2° und 3° für das bisherige System, das Faltungsnetz (links) und das Transformer-Netz (rechts)

Abschließend lässt sich also sagen, dass das in diesem Projekt entwickelte System eine sehr genaue automatisierte Schätzung der Flugparameter liefert, die so in der Praxis eingesetzt werden kann. Ausführliche Ergebnisse des Forschungsprojekts wurden in einem wissenschaftlichen Aufsatz (Ludwig, 2022) bei einer internationalen Konferenz publiziert.

5 Literatur

- Ludwig, K., Einfalt, M., & Lienhart, R. (2020). Robust estimation of flight parameters for ski jumpers. *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 1-6.
- Ludwig, K., Harzig, P., & Lienhart, R. (2022). Detecting arbitrary intermediate keypoints for human pose estimation with vision transformers. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*.
- Yanjie Li, et al. (2021). Tokenpose: Learning keypoint tokens for human pose estimation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*.
- Bowen Cheng, et al. (2020). HigherHRNet: Scale-aware representation learning for bottom-up human pose estimation. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.