

Tracking von Gruppeninteraktionen im Raum

Diese Arbeit befasst sich mit dem Tracking von Gruppeninteraktionen im Raum bzw. der Evaluation bestehender Trackingverfahren in Bezug auf nonverbal kommunikative Elemente. Dafür wurden zuerst bestehende Datensätze analysiert, Dabei wurden drei passende Datensätze gefunden. Anschliessend wurden die Trackingansätze bezüglich deren Eignung für ein Gruppentracking vorevaluiert. Auch hierbei wurden drei Trackingansätze bzw. Kombinationen von Trackingansätzen ausgewählt, zwei davon wurden anschliessend trainiert und mittels eines eigenen Evaluationsverfahrens evaluiert.

Analyse der geeigneten Datensätze

Die Analyse der geeigneten Datensätze erfolgte analog mithilfe der Dokumentationen der Datensätze, ausserdem wurden alle gefundenen Datensätze angeschaut um diese besser analysieren zu können.

Die Parameter für geeignete Datensätze wurden folgendermassen festgelegt:

- Die Datensätze sollen dreidimensionale Daten enthalten.
- Es müssen mehrere Personen pro Bild im Datensatz vorhanden sein.
- Je mehr Gelenkposen vorhanden sind, desto besser ist der Datensatz geeignet.

Dabei wurden drei Datensätze als für dieses Projekt geeignet evaluiert: Diese sind:

- 3DPW (Marcard u.a. o. J.): Dieser Datensatz enthält Bilder zweier Personen. Mittlere Abweichung: 16mm. Es werden jeweils 24 Gelenke pro Person getrackt (Marcard u.a. 2018) .
- MuCo-3DHP (Mehta; Sotnychenko; u.a. o. J.): Dieser Datensatz enthält Bilder von maximal 4 Personen. Mittlere Abweichung: 80.7mm. Es werden jeweils 28 Gelenke pro Person getrackt (Mehta; Rhodin; u.a. 2017; Mehta u.a. 2018).
- MuPoTS-3D (Mehta; Sotnychenko; u.a. o. J.): Dieser Datensatz enthält Bilder von maximal 3 Personen. Mittlere Abweichung: 126.3mm. Es werden jeweils 17 Gelenkposen pro Person getrackt (Mehta; Rhodin; u.a. 2017; Mehta; Sridhar; u.a. 2017) .

Analyse der geeigneten Ansätze

Die Analyse der geeigneten Ansätze erfolgte analog mithilfe der Dokumentationen der Ansätze, ausserdem wurden die Ansätze nach Verfügbarkeit und Vollständigkeit überprüft.

Die Parameter für die geeigneten Ansätze waren:

- Ein 3D-Tracking ist sinnvoll.
- Es müssen mehrere Personen getrackt werden können.

Dabei wurden folgende Ansätze als geeignet evaluiert:

- 3D Multi-person Pose Estimation from Single RGB Image (Moon; Chang; Lee 2019): Es werden bei diesem Ansatz mithilfe von drei Netzwerken die 3D-Gelenkpositionen mehrerer Personen getrackt (Moon; Chang; Lee 2019)
- Eine Kombination eines 2D-Ansatzes mit dem Ansatz 3D Pose Baselines (Martinez u.a. 2017): Der 3D Pose Baselines Ansatz schätzt 3D-Gelenkposen mithilfe von 2D-Gelenkposen (Martinez u.a. 2017) .

- Eine Kombination eines 2D-Ansatzes mit VideoPose3D (Pavlo u.a. 2019): Der VideoPose3D-Ansatz benutzt 2D-Gelenkposen mehrerer Zeitstände um daraus 3D-Gelenkposen eines Zeitstandes zu schätzen (Pavlo u.a. 2019).

Training und Evaluation der Ansätze

Es wurden 2 der 3 Ansätze angepasst, trainiert und evaluiert. Dies waren die Ansätze:

- 3D Multi-person Pose Estimation from Single RGB Image (Moon; Chang; Lee 2019).
- 3D Pose Baselines (Martinez u.a. 2017)

Fürs Training des 3D Multi-person Pose Estimation from Single RGB Image Ansatzes (Moon; Chang; Lee 2019) wurde den einzelnen Gelenkpunkten Gewichte hinzugefügt, wodurch erwartet wurde, dass Gelenke mit höheren Gewichten genauer getrackt werden können. Die Evaluationsresultate dieses Ansatzes zeigen, dass das Gewicht der einzelnen Gelenkposen keine Rolle spielen. Abb. 1 zeigt die Evaluation, wenn sämtliche Gelenkpunkte mit demselben Gewicht trainiert werden. Abb. 2 zeigt die Evaluation bei unterschiedlichen Gewichten der Gelenkpunkte.

Die Evaluation des 3D-Pose Baseline (Martinez u.a. 2017) Ansatzes ist hingegen nicht brauchbar. Dies liegt daran, dass die Bilder des Testdatensatzes eine andere Dimension aufweisen als jene des Trainingsdatensatzes.

```
Joint Head_top has an MPJPE of 203.290443377769 mm and was trained with a weight of 1.
Joint Thorax has an MPJPE of 142.91158834170548 mm and was trained with a weight of 1.
Joint R_Shoulder has an MPJPE of 157.68390182082746 mm and was trained with a weight of 1.
Joint R_Elbow has an MPJPE of 206.06498734079523 mm and was trained with a weight of 1.
Joint R_Wrist has an MPJPE of 250.70575935271847 mm and was trained with a weight of 1.
Joint L_Shoulder has an MPJPE of 157.66107614036196 mm and was trained with a weight of 1.
Joint L_Elbow has an MPJPE of 203.95467049745884 mm and was trained with a weight of 1.
Joint L_Wrist has an MPJPE of 242.54601548252984 mm and was trained with a weight of 1.
Joint R_Hip has an MPJPE of 128.43710762844194 mm and was trained with a weight of 1.
Joint R_Knee has an MPJPE of 166.47496189955186 mm and was trained with a weight of 1.
Joint R_Ankle has an MPJPE of 218.94134002906173 mm and was trained with a weight of 1.
Joint L_Hip has an MPJPE of 127.00161786875354 mm and was trained with a weight of 1.
Joint L_Knee has an MPJPE of 191.4156524697458 mm and was trained with a weight of 1.
Joint L_Ankle has an MPJPE of 242.42689705586142 mm and was trained with a weight of 1.
Joint Pelvis has an MPJPE of 89.66148323235021 mm and was trained with a weight of 1.
Joint Spine has an MPJPE of 106.16102797470853 mm and was trained with a weight of 1.
Joint Head has an MPJPE of 154.61663478463245 mm and was trained with a weight of 1.
Total MPJPE is 175.879715605722.
```

Abb. 1: Evaluation mit gleichen Gewichten

```
Joint Head_top has an MPJPE of 192.48725607281403 mm and was trained with a weight of 0.5.
Joint Thorax has an MPJPE of 134.24121656690187 mm and was trained with a weight of 0.5.
Joint R_Shoulder has an MPJPE of 151.64251156336977 mm and was trained with a weight of 0.5.
Joint R_Elbow has an MPJPE of 210.02572910931227 mm and was trained with a weight of 0.5.
Joint R_Wrist has an MPJPE of 224.81538403394396 mm and was trained with a weight of 1.
Joint L_Shoulder has an MPJPE of 138.24576401020227 mm and was trained with a weight of 1.
Joint L_Elbow has an MPJPE of 198.8183900474 mm and was trained with a weight of 1.
Joint L_Wrist has an MPJPE of 211.7415664410442 mm and was trained with a weight of 1.
Joint R_Hip has an MPJPE of 119.19466456204144 mm and was trained with a weight of 2.
Joint R_Knee has an MPJPE of 141.61117663691246 mm and was trained with a weight of 2.
Joint R_Ankle has an MPJPE of 191.8994876251945 mm and was trained with a weight of 2.
Joint L_Hip has an MPJPE of 112.56996718384859 mm and was trained with a weight of 2.
Joint L_Knee has an MPJPE of 164.95903504133884 mm and was trained with a weight of 4.
Joint L_Ankle has an MPJPE of 220.74312920036297 mm and was trained with a weight of 4.
Joint Pelvis has an MPJPE of 83.04214350020531 mm and was trained with a weight of 4.
Joint Spine has an MPJPE of 96.23695921839042 mm and was trained with a weight of 4.
Joint Head has an MPJPE of 135.63632791059524 mm and was trained with a weight of 1.
Total MPJPE is 160.46533580746697.
```

Abb. 2: Evaluation mit unterschiedlichen Gewichten

Referenzen:

- Marcard, Timo u.a. (o. J.): 3DPW | Real Virtual Humans. Online im Internet: URL: <https://virtualhumans.mpi-inf.mpg.de/3DPW/> (Zugriff am: 03.03.2020).
- Marcard, Timo u.a. (2018): „Recovering Accurate 3D Human Pose in the Wild Using IMUs and a Moving Camera: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part X.“ In: , S. 614–631. Online im Internet: DOI: 10.1007/978-3-030-01249-6_37
- Martinez, Julieta u.a. (2017): „A Simple Yet Effective Baseline for 3d Human Pose Estimation.“ In: 2017 IEEE International Conference on Computer Vision (ICCV), (2017). Online im Internet: DOI: 10.1109/ICCV.2017.288
- Mehta, Dushyant u.a. (o. J.): Monocular 3D Human Pose Estimation In The Wild Using Improved CNN Supervision. Online im Internet: URL: <http://gvv.mpi-inf.mpg.de/3dhp-dataset/> (Zugriff am: 16.03.2020).
- Mehta, Dushyant; Rhodin, Helge; u.a. (2017): Monocular 3D Human Pose Estimation In The Wild Using Improved CNN Supervision. International Conference on 3D Vision (3DV 2017). Online im Internet: URL: <http://gvv.mpi-inf.mpg.de/3dhp-dataset/> (Zugriff am: 03.03.2020).
- Mehta, Dushyant u.a. (2018): Single-Shot Multi-Person 3D Pose Estimation From Monocular RGB. International Conference on 3D Vision (3DV 2018). Online im Internet: URL: <http://gvv.mpi-inf.mpg.de/projects/SingleShotMultiPerson/> (Zugriff am: 16.03.2020).
- Mehta, Dushyant; Sridhar, Srinath; u.a. (2017): VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera. ACM Transactions on Graphics (SIGGRAPH 2017). Online im Internet: URL: <http://gvv.mpi-inf.mpg.de/projects/VNect/> (Zugriff am: 20.03.2020).
- Moon, Gyeongsik; Chang, Juyong; Lee, Kyoung Mu (2019): „Camera Distance-aware Top-down Approach for 3D Multi-person Pose Estimation from a Single RGB Image.“ In: The IEEE Conference on International Conference on Computer Vision (ICCV).

Autor/in: Marc Vögele

Examinator/in: Prof. Dr. Susanne Bleisch

Experte/n: Stefan Cavegn