



HAL
open science

Comparison of marker-based and markerless motion analysis: robustness to camera reduction

Antoine Muller, Baptiste Boulaire, Anaïs Chaumeil, Nicola Hagemeister, Raphaël Dumas, Thomas Robert

► **To cite this version:**

Antoine Muller, Baptiste Boulaire, Anaïs Chaumeil, Nicola Hagemeister, Raphaël Dumas, et al.. Comparison of marker-based and markerless motion analysis: robustness to camera reduction. 2024. hal-04737955

HAL Id: hal-04737955

<https://hal.science/hal-04737955v1>

Preprint submitted on 15 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Comparison of marker-based and markerless motion analysis: robustness to camera reduction

Antoine Muller *

Univ Eiffel, Univ Lyon 1, LBMC UMR_T 9406, F-69622 Lyon, France

Baptiste Boulaire

Laboratoire d'Innovation Ouverte en Technologie de la Santé (LIO-ETS), École de Technologie Supérieure and Centre de Recherche du CHUM, Montréal, Canada

Anaïs Chaumeil

Univ Eiffel, Univ Lyon 1, LBMC UMR_T 9406, F-69622 Lyon, France

Nicola Hagemeister

Laboratoire d'Innovation Ouverte en Technologie de la Santé (LIO-ETS), École de Technologie Supérieure and Centre de Recherche du CHUM, Montréal, Canada

Raphaël Dumas

Univ Eiffel, Univ Lyon 1, LBMC UMR_T 9406, F-69622 Lyon, France

Thomas Robert

Univ Eiffel, Univ Lyon 1, LBMC UMR_T 9406, F-69622 Lyon, France

***Corresponding author.** Email: antoine.muller@univ-lyon1.fr

Keywords: video-based analysis; 3D reconstruction; OpenPose; direct linear transformation

1. Introduction

Because it requires no participant instrumentation, markerless motion analysis opens up a wide range of applications. The approaches commonly used (e.g., in OpenCap (Uhlrich et al. 2023)) are similar to those used more conventionally with markers: 1) detection of the 2D position of keypoints for each camera; 2) reconstruction of their 3D positions. The difference lies in the first step where the markerless uses a machine learning approach which almost systematically identifies all the points, although with a potentially large error (Chaumeil et al. 2023).

The identification of almost all the points means that analyses can be carried out with very few cameras. To move towards field analyses, many studies propose using only two cameras. However, the accuracy of 3D

points reconstruction from few 2D information (few cameras) embedding potentially large errors should be questioned. The impact of reducing the number of cameras on the accuracy of the 3D keypoint locations has been partly assessed in the literature, with contrasted results: the impact seems minimal when the reduced set of cameras is optimal (Uhlrich et al. 2023; Pagnon et al. 2021), but appeared strongly dependent on the quality of the 2D estimation (Chaumeil et al. 2023). Therefore, a more systematic analysis is still needed.

The aim of this study was to compare the marker-based and markerless motion analysis in terms of robustness to a reduction in the number of cameras, both in terms of the position of the reconstructed 3D point and the possibility of reconstructing this point when a 2D keypoint is not identified or occultation occurs.

2. Methods

2.1 Experimental protocol

One female (22 years old, 167 cm, 55 kg) participated to the motion analysis experiment after signing a consent form approved by the institutional review board. One random sampled instant of time of a static trial in a T-pose was used for this study. The pose was simultaneously captured by a marker-based and by a markerless motion capture system. For each system, 10 cameras were equally distributed around the participant (Qualisys Miquis M3 for marker-based and Qualisys Miquis Video for markerless) at approximatively 2 m. Cameras were arranged in groups of 2 (one from each system). Both systems were synchronised and calibrated using Qualisys Track Manager (QTM - Qualisys AB, Sweden, v2021.1.2). For the marker-based system, 46 markers were placed on the whole body of the participant following the ISB recommendations.

2.2 Data analysis

For the markerless analysis, the 2D position of 20 keypoints and its associated confidence were obtained with OpenPose (v1.7.0) using the BODY25B model. The confidence is directly given by the machine learning algorithm and is a value between 0 and 1. As proposed by OpenPose, when a confidence was lower than 0.05, the associated keypoint was considered as not identified.

For the marker-based analysis, the 2D position of 46 keypoints (corresponding to markers) were obtained with QTM. To be consistent with both analyses, a confidence of 1 was associated to each 2D position.

For a given set of cameras, 3D point reconstruction was performed using a weighted direct linear transformation (Pagnon et al. 2021). For each keypoint, the weight was the associated confidence.

The robustness of each system was assessed by evaluating the influence of removing one camera of a set of n cameras ($n \in \llbracket 3,10 \rrbracket$) using two variables: 1)

the distance in the 3D point reconstruction; 2) the possibility to reconstruct the 3D point.

The first variable corresponded to the distance between the 3D point reconstructed with n cameras and the one with $(n-1)$ cameras.

The second variable corresponded to the percentage of cases where the 3D point reconstruction was possible with $(n-1)$ cameras if it was possible with n cameras.

For all combinations of a set of n cameras, the two variables were assessed for all combinations of a subset of $(n-1)$ cameras. It corresponded to about 2.6M combinations of $n \rightarrow n-1$ cameras.

3. Results and discussion

Figure 1 displays the robustness of both systems for each combination of $n \rightarrow n-1$ cameras.

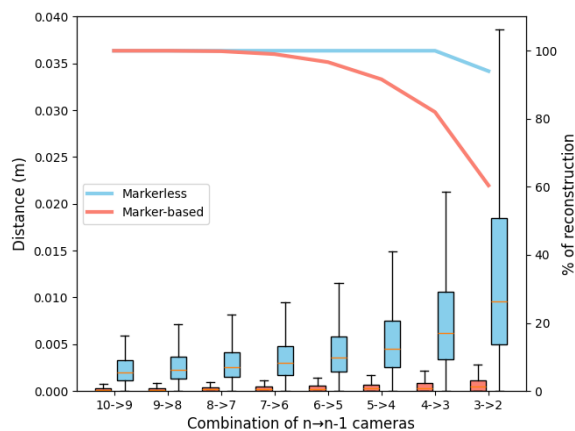


Figure 1. Evaluation of the distance in the 3D point reconstruction and the percentage of possible reconstruction for all combinations of $n \rightarrow n-1$ cameras for the marker-based and the markerless motion capture systems.

For the marker-based analysis, the mean distance was lower than 1 mm for all combinations and the percentage of 3D point reconstruction had a minimal value of 60% for the 3→2 combination.

For the markerless analysis, the mean distance was between 16 mm (3→2 combination) and 2 mm (10→9 combination) and the percentage of 3D point reconstruction was below 100% only for the 3→2 combination.

The observed results show that marker-based and markerless motion capture behaved differently when the number of cameras was reduced.

Using a marker-based approach, from the moment a point is seen by two cameras, the position of the 3D point is only slightly affected by the number of cameras. The obtained distances are in the same order of magnitude as the observations of Conconi et al., (2021) on static tests. However, given the method used to detect the points, the number of occultations can be significant and a large number of cameras is thus needed. For the static pose analyzed, only the

configurations with at least 8 cameras ensure 100% of the 3D point reconstruction.

Using a markerless approach, only 3 cameras can reconstruct the position of all points whatever the set of cameras used. However, with only a couple of cameras, the 3D position of the reconstruction points is sensitive (more than 1 centimeter) to the camera's combination used. This behavior can be explained by the low specificity but high sensitivity of the 2D keypoint identification algorithm used in markerless approaches: keypoints are almost always detected but with possible large errors, even if the associated confidence is high (Chaumeil et al. 2023).

4. Conclusions

Marker-based and markerless analyses behave differently when the number of cameras is reduced. For a markerless analysis, even if only a couple of cameras is sufficient to reconstruct the 3D position of a point, its position can vary by several centimeters depending on the camera combination used. Even though this behavior was only observed in a single static pose, these results appear to be consistent with observations made on the accuracy of 2D keypoint identification (Chaumeil et al. 2023) across a variety of movements. So, although very attractive, particular attention needs to be paid to the results obtained with markerless analysis using a small number of cameras.

4. Conflicts of interests

The authors declare no conflict of interest during this study.

References

- Chaumeil, A., Muller, A., Dumas, R., & Robert, T. (2023). Effect of a confidence-based weighted 3D point reconstruction for markerless motion capture with a reduced number of cameras. *Comput. Methods Biomech. Biomed. Eng.: Imaging Vis.*, 1-8. doi:10.1080/21681163.2023.2292067
- Conconi, M., Pompili, A., Sancisi, N., & Parenti-Castelli, V. (2021). Quantification of the errors associated with marker occlusion in stereophotogrammetric systems and implications on gait analysis. *J. Biomech.*, 114, 110162. doi:10.1016/j.jbiomech.2020.110162
- Pagnon, D., Domalain, M., & Reveret, L. (2021). Pose2Sim: An end-to-end workflow for 3D markerless sports kinematics—Part 1: Robustness. *Sensors*, 21(19), 6530. doi:10.3390/s21196530
- Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., ... & Delp, S. L. (2023). OpenCap: Human movement dynamics from smartphone videos. *PLoS Comput. Biol.*, 19(10), e1011462. doi:10.1371/journal.pcbi.1011462

Received date:07/04/2024

Accepted date: 28/06/2024

Published date: XX/XX/2024

Volume: 1

Publication year: 2024