GLO-40-350 v 2.0
2024-11-27

Tech note ASTM accuracy assessment

Author: Erwin Schoonderwaldt

1. Summary

This tech note gives an overview of the ASTM E3064 – 16 standard for assessing the accuracy of optical motion capture systems and its application to Qualisys systems. Alternative implementations are explained, including the full implementation using a 6DOF metrology wand and downscaled, derivative methods that can be used when such a wand is not available. This tech note describes how to perform the measurements and how to perform the analysis and create reports using the available resources (Matlab scripts and templates). The ASTM accuracy assessment is illustrated for a 24-camera Arqus A12 system in the mocap studio at Qualisys HQ. The results are compared with competitor systems of Vicon and Optitrack.

2. Introduction

2.1. Goals

- Provide insight into accuracy assessment of optical motion capture systems using the ASTM E3064 – 16 standard.
- Provide tips and tricks for optimizing the accuracy of a Qualisys motion capture system.
- Provide instructions and tools for performing accuracy assessment of Qualisys motion capture volumes based on this standard.
- Provide a comparison with competitor mocap systems (Vicon and Optitrack).

2.2. Background

The ASTM standard test method is described in the document [ASTM-standard]. The test method presents the metrics and procedures for measuring, calculating, and reporting the relative pose error of optical systems that compute the pose (position and orientation) of a rigid object in motion. The relative pose is obtained by using a metrology wand (also referred to as a *test artifact*) with two clusters of markers at a distance of about 300 mm, that is moved in a systematic way through the measurement volume. The positions of the markers should be known from an external source or measurement at a level of accuracy well beyond that of the motion capture system to be evaluated.

The method specifies three capture volumes for the assessment:
- Large volume: 12m x 8m x 2m (length x width x height)
- Medium volume: 9m x 6m x 2m
- Small volume: 3m x 2m x 2m

The metrics to be included in the report give insight into the error statistics of the position and orientation derived from the relative pose of the two clusters.

2.3. Structure of this tech note

This remainder of this tech note is structured as follows:
- Section 3: General information about error sources and optimization of the system for accuracy performance; explanation of the methodology for accuracy evaluation based on the ASTM standard.
- Section 4: Alternative implementations for a wider range of applications.
- Section 5: Accuracy evaluation of a 24-camera Arqus A12 system in the mocap studio.
- Section 6: Discussion of the results and comparison to competitor systems.

3. Methodology

3.1. Error sources and optimization of the system for accuracy performance

There are many possible sources of error for the estimation of position and orientation using an optical motion capture system. The most important ones fall into the following categories:

- Calibration quality (intrinsic and extrinsic)
- Camera properties (optical quality of the lens, sensor resolution, flash and exposure settings)
- Marker quality (shape, reflection properties, intensity distribution in case of active markers)
- Partial occlusions and merging of markers as observed by individual cameras, leading to errors in the marker center point estimation.

In order to optimize the system for accuracy performance, follow these guidelines:

- Warm up time
    - Make sure to power up the cameras at least 30 minutes before doing the calibrations and measurements.
- Calibration
    - Use an appropriate calibration method for the application (fixed or wand method).
    - Use a good quality calibration kit with a wand length appropriate for the size of the capture volume.
    - Make sure that the duration and the number of frames used for the calibration is sufficient. For example, for larger volumes the duration of the calibration should be sufficient to cover the whole capture volume. It is recommended to use at least 3000-6000 frames for the calibration (default setting in QTM is 1500).
    - Make sure to use a calibration that is close in time with the test measurements. The calibration quality may degrade over time due to mechanical factors (walls, camera mounts) and thermal fluctuations of the environment and within the cameras.
    - Make sure to minimize extra reflections. In some cases, 2D preprocessing using marker size filtering can improve the quality of the calibration.
- Linearization
    - Make sure to use a valid, if possible, recent linearization for the cameras.
    - For the best results, linearization should be performed after setting the focus and aperture on the cameras. When using motorized lenses, it is recommended to fix the lens settings. Make sure the cameras are warmed up when performing the linearization.
    - If possible, use the advanced calibration method (combined intrinsic and extrinsic calibration).
- Test artifact design (marker size, and attachment)
    - The marker configuration should be designed for optimal visibility and separation of the markers.
    - Make sure the markers of the respective clusters are well distributed (e.g. non-collinear) and well separated by using an appropriate marker size.
    - Use anti-reflective materials to avoid spurious reflections.
    - Use slim (but stable) mechanics to reduce occlusions of markers to a minimum.
- Camera configuration
    - Use an appropriate number of cameras and place them so that they provide good coverage of the capture volume and have sufficient overlap with each other. Any part of the volume should be covered by several cameras from different angles. Use sufficient margin between the cameras and the edges of the capture volume that needs to be covered.
- Camera settings
    - Make sure that the focus is set correctly.
    - If possible, use an aperture between f4 and f8 for a good depth of focus and minimum vignetting.
    - Make sure to optimize the camera settings for good marker detection.

- Avoid extra reflections. If needed, use exposure groups to avoid that the cameras see each other's reflections.

3.2. The Qualisys metrology wand

A metrology wand was made at Qualisys. The construction is based on a 300mm carbon wand. Two marker clusters consisting of 4 markers each were attached to the ends. The marker diameter was 14 mm.

Figure 1: Qualisys metrology wand.

The 3D marker positions were measured at Element Metech AB using a CMM. The uncertainty per component (X, Y, Z) was specified at 0.001 mm (1 micron) level.

The specified positions were used to create the rigid body definitions for the respective clusters (P and Q):

- For each cluster, the geometric center was calculated and used as the origin of the marker positions in the rigid body definitions of the respective clusters.
- The distance between the geometric centers of the clusters (used as reference length) was 346.712 mm.
- The average radius of the clusters was between 46 and 58 mm; the maximum distance to the center per point was 80 mm (within the ASTM specifications of maximum 100 mm).

It was found that using the cluster definitions based on the specified marker positions led to a persistent angular offset between the clusters in motion capture measurements. In order to obtain zero orientation offset, refined rigid body definitions of the clusters were used for the calculations. The refined positions of the markers were calculated from averaged pose measurements of the metrology wand obtained with a high-precision camera setup, using the nominal distance between the clusters to obtain the correct scaling.

The Element Metech service report (2022-02-10) and the rigid body definitions (original and refined) are available at QStore: \\qstore\Measurement_data\Accuracy_tests\2022-02-28 Metrology wand

3.3. Performing the assessment

3.3.1. Calibration

For the assessment, the volume should be calibrated as normal. When using the wand calibration method, a Qualisys calibration kit should be used that is in a good state and with a valid certificate.

The linearization of the cameras (intrinsic calibration) should be recent. For a fixed installation, make sure that the lens settings are fixed before performing the linearization. When using the wand calibration method, it is highly recommended to use the advanced calibration method, if possible.

It is recommended to perform multiple calibrations to provide insight into the repeatability. When the cameras are in stable positions and the test measurements are performed closely in time, they can be cross processed against multiple calibrations, giving insight into the quality and repeatability of the calibrations.

Make sure that the calibration covers the whole measurement volume to be assessed.

3.3.2. Grid measurements with test object

According to the ASTM test method, the wand should be moved in a regular grid X-Y with the wand in three different orientations, aligned in X, Y and Z directions, respectively. The distance between the grid lines should be less or equal to the length of the wand. This way of measuring gives full coverage of the volume with the wand in all basic orientations across all grid lines.

For pragmatic reasons, the grid measurement can be reduced. Since the wand is mounted on a stick, it is easiest to point the stick with the wand in front of you while walking along the grid. An equal sampling of the X, Y and Z orientations is then obtained as follows:

- Walk the grid in X direction while keeping the wand in horizontal orientation (parallel to the floor plane).
- Walk the grid in Y direction while keeping the wand in horizontal orientation.
- Walk the grid in either X or Y direction while keeping the wand in vertical orientation (perpendicular to the floor plane).

Depending on the size of the volume, the density of the grid lines can also be reduced to, for example, 1 grid line per meter, as long as it gives a representative coverage of the measurement volume.

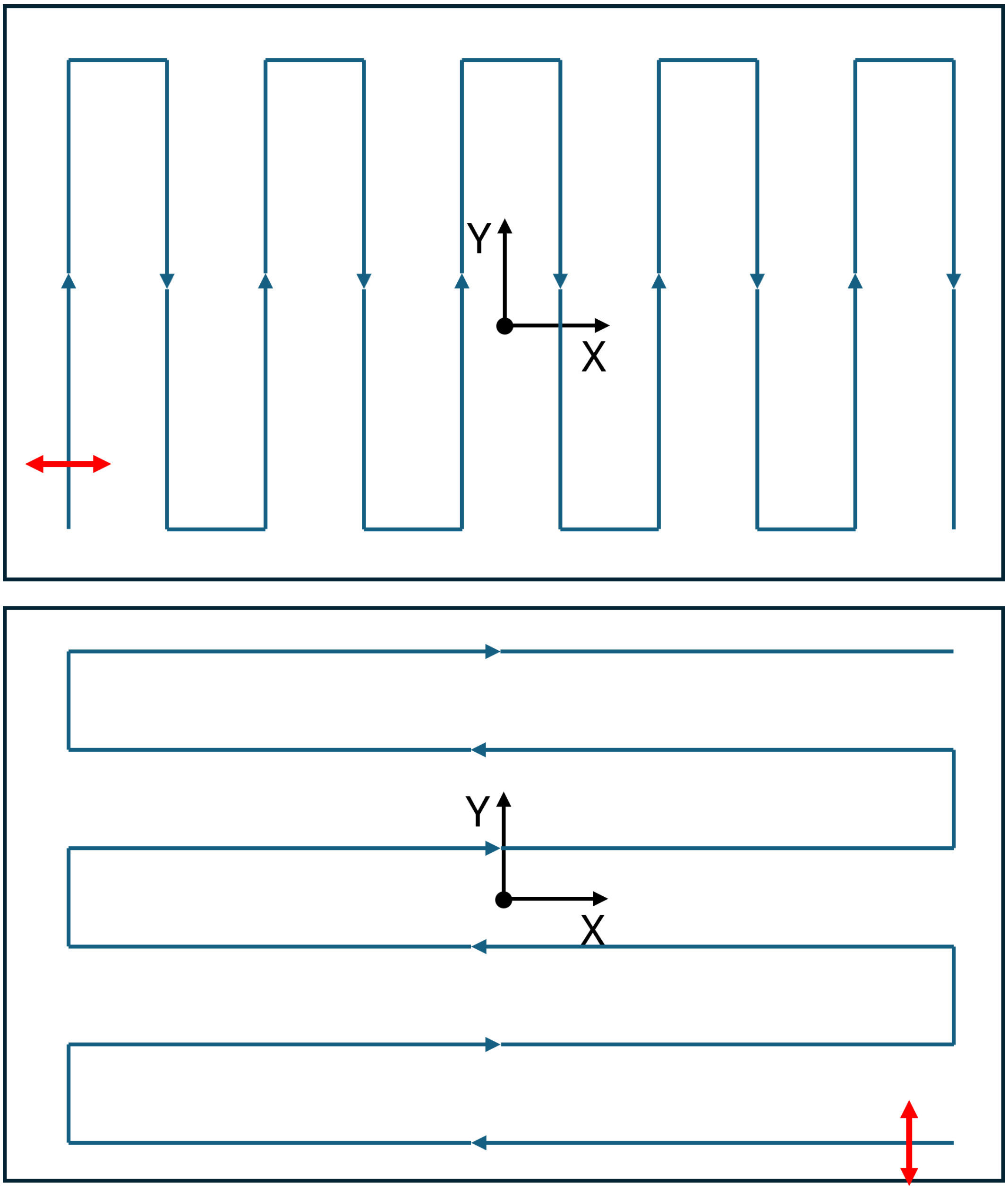

Figure 2: Illustration of grid to be traversed with the test artifact. Traverse the grid in X and Y direction with the wand (indicated by red double arrow) in horizontal orientation, then once more in either X or Y direction with the wand in vertical orientation.

3.3.3. Other variations

The ASTM standard test method is rather strict about the volume to be assessed (three specific volumes), and the performance of the grid measurement. Whereas this is beneficial for the comparability between evaluations, this may not always provide a good representation of measurement volumes for specific applications.

It is possible to adapt the method for a more representative evaluation for the specific application. Most important to keep in mind is to cover the measurement volume and vary the orientation of the wand for a representative sampling of the poses in the space.

3.4. Performing the calculations

3.4.1. Matlab resources

A set of tools for Matlab is available for extracting the data and performing the calculations [Q-MAT]. These tools include the following template scripts that can be used for the alternative implementations presented in Section 4:

- **Analyze_wand_file.mlx:**
  Live script that can be used for 6DOF metrology wand and 3D calibration wand. Applies to a single file and provides detailed information and graphs showing the error/repeatability distributions and a map showing how the error is distributed across the capture volume. The first section of the script allows the user to edit parameters that are specific for the capture volume and specify criteria for filtering/selecting the data. The results are presented in a report (Matlab live script output) and can be exported as PDF.
- **Aggregate_data_6d.m:**
  Script for extracting data from multiple files. Requires the use of a 6DOF metrology wand. The batch process is controlled by the admin.xlsx file. The script output is a new Excel file containing the metrics that are extracted from the measurements.

  The batch process is controlled through the information in the *calibrations* and *trials* sheets in the *Admin.xlsx* file, that need to be filled in by the user before running this script. These sheets contain obligatory columns that are needed for running the script, but the user can also add optional columns, allowing them to introduce additional conditions (for example, moving the test wand in different ways, use of different calibration wands, etc.). The analysis is iteratively performed (nested) across a set of sub-volume and wand-orientation selections that can be specified by the user in the parameters section of the script. This can give more detailed insight into the accuracy of different parts of the volume (e.g., central versus peripheral), or scaling errors in X, Y and Z directions. The results are collected in an Excel file, including meta information about the calibrations and trials (obligatory and the fields added by the user), selections (sub volumes and wand orientations), and the assessment metrics as specified in the ASTM standard.

  Statistics of the output can be obtained by using pivot tables, allowing to organize and filter the data and calculate the averages. This approach leads to a more statistically sound assessment of accuracy across calibrations and measurements and gives more detailed insight into the accuracy level in different areas of the capture volume (sub volumes).

  Note that the ASTM standard only requires a single calibration and measurement.
- **Cluster_accuracy_evaluation.mlx:**
  Live script that can be used for simplified 3D assessment using a fixed marker configuration, see Section 4.4. This assessment can be performed with any type of rigid object with markers attached to it and that can be tracked using 6DOF tracking, for example a ship model in a large water tank. A complete grid measurement can be obtained by combining multiple measurements that are combined in the analysis. The user needs to specify the parameters (trial(s) to be included, capture volume, maximum number of markers pairs used for the analysis, etc.) in the parameter section of the script. The analysis is performed on pairs of markers attached to the rigid body. The results are presented in a report (Matlab live script output) and can be exported as PDF.

The scripts include the option to use outlier removal. The advantage of using outlier removal is that the results are more representative for the statistical distributions and less dependent on infrequent extreme errors that can be introduced by measurement artifacts (e.g., merging of markers or other tracking artifacts at the edges of the volume). However, outlier removal is not included in the ASTM standard. When using it, it is good practice to mention that it has been used and report the number of outliers and maximum errors separately.

3.4.2. Processing of calibrations and measurements

When performing multiple calibrations and test measurements for analysis using the *Aggregate_data_6d* script, the measurements must be organized in folders per calibration used for processing them. It is recommended to process all measurements with all calibrations. If a calibration cannot be used for a trial, that trial should be omitted from the folder. The folder structure should look like this:

```
Cal_1\
        Trial_1.qtm
        Trial_2.qtm
        Trial_3.qtm
        …
Cal_2\
        Trial_1.qtm
        Trial_2.qtm
        Trial_3.qtm
        …
Etc.
```

Once the measurements are processed with the respective calibrations, review the tracking and labeling, and if needed fix problems. Make sure to use processing settings that give the best results across the whole range of files. Manual editing should be kept to a minimum. However, in case of the use of a calibration wand for the test measurements, the labeling of the markers should be done manually since AIM does not perform well for models with only two markers.

Once all QTM files are approved, export them to Matlab format, checking the 3D and 6DOF export options.

3.5. Reporting

The ASTM standard test method includes an appendix with an example report. The example report includes the following sections with specific information:

1. Test information
2. Instrument(s) and operator(s) information / optical tracking system
3. Ambient test conditions
4. Artifact characteristics
5. Test space
6. Test results

A template report, adapted from the example report presented in [ASTM-standard] that can be used for Qualisys evaluations is included with the resources [Q-MAT].

4. Alternative implementations

This section presents how the ASTM standard test method can be applied to a Qualisys system. Alternative implementations are discussed, starting with the full implementation using a special metrology wand, followed by scaled-down implementations when such a test artifact is not available.

The assessment metrics that can be calculated and reported are dependent on the type of test artifact used for the assessment, see Table 1.

Table 1: Assessment metrics that can be calculated and reported depending on the type of test artifact used. Errors statistics (deviations from a known reference value) can only be reported if the exact positions of the markers on the test artifact are known from an external measurement. If this requirement is not fulfilled, a full accuracy test is not possible, but instead the repeatability (deviations from the measured average) can be reported. The optional metrics, indicated with an asterix, are not listed as reportable metrics according to the ASTM standard, but can be reported to provide additional insight.

Test artifact type	**Metrology wand**	**6DOF wand**	**Calibration wand**	**Fixed distances**
Degrees of freedom	6DOF	6DOF	3D	3D
Described by ASTM standard	Yes	Yes	No	No
Error statistics results				
Distance				
*Average distance error**	optional		optional	
RMS distance error	Yes		Yes	
Max distance error	Yes		Yes	
Percentiles (99.7, 95, 50)	Yes		Yes	
Orientation				
*Average Euler angle error**	optional			
RMS orientation error	Yes			
Max orientation error	Yes			
Percentiles (99.7, 95, 50)	Yes			
Repeatability statistics results				
Distance				
*Estimated wand length**		optional		optional
RMS distance deviation		Yes		Yes
Max distance deviation		Yes		Yes
Percentiles (99.7, 95, 50)		Yes		Yes
Orientation				
*Average Euler angle offset**		optional		
RMS orientation deviation		Yes		
Max orientation deviation		Yes		
Percentiles (99.7, 95, 50)		Yes		

4.1. Full 6DOF accuracy assessment using a metrology wand

A full accuracy assessment according to the ASTM standard requires a 6DOF metrology wand with exact marker positions known from an external source.

It is good practice to perform multiple calibrations and measurements. If the camera positions are stable, the calibrations and measurements can be cross processed to increase the sample size.

The following Matlab scripts can be used for the analysis:

- The *Analyze_wand_file.mlx* live script can be used to perform a detailed analysis of a single file, specified in the script.
- The *Aggregate_data_6d.m* script can be used to extract data from multiple calibration-measurement combinations. Use a Pivot table in Excel to filter the data and calculate the average values (across calibrations and measurements) of the error metrics.

For creating the report, fill in the resulting values in the Excel report template.

4.2. 6DOF repeatability assessment using a 6DOF wand

Alternatively, a 6DOF wand can be used without known marker positions. This does not allow for assessment of the accuracy as compared to an external measurement but can be used to assess the repeatability of the system [ASTM, section 8.3].

For the measurements, create a representative rigid body definition for the respective clusters on the wand. The definition should be based on the average of a variety of recorded poses.

It is recommended to use a separate mocap system optimized for tracking the test artifact. For creating a representative rigid body definition, follow these steps:

- Record the test artifact in several static poses.
- Try not to block the view of the cameras, avoid merging of the markers in the camera views. If needed, use 2D preprocessing (e.g. software masks) to discard erratic 2D input (e.g. merged or partially occluded markers).
- The marker positions can then be calculated as follows:
  - Create a rigid body definition of the whole wand (all markers) based on the first pose (reference pose).
  - Calculate the marker coordinates of all poses in local coordinates (transform to reference pose).
  - Calculate the average marker positions of the wand across the poses and use these in the rigid body definition.
  - Split the rigid body definition into two separate clusters and translate them to their respective geometrical centers.

The distance between the clusters can be used as a reference distance for error statistics, but since the exact marker positions are not measured with an external system, the error statistics should not be included in the report. The repeatability (deviation from the average distance) can be reported.

The following Matlab scripts can be used for the analysis:

- Analyze_wand_file.mlx
  Refer only to the reported repeatability results (deviations).
- Aggregate_data_6d
  Refer only to the calculated repeatability results (deviations).

For creating the report, fill in the resulting values in the Excel report template.

4.3. Simplified 3D assessment using a calibration wand

The assessment can also be performed with a 3D calibration wand with only one marker at each end. For example, if you have two calibration wands, you can perform a calibration with one wand

and a test measurement with the other. This method is not described in the ASTM standard test method, but it can still give a useful indication of the accuracy performance of the system. For the accuracy calculation, the exact wand length from the calibration kit certificate can be used. However, this method does not give any insight into the measurement performance of orientation.

The following Matlab script can be used for the analysis:

- The *Analyze_wand_file.mlx* live script can be used to perform a detailed analysis of a single file, specified in the script. Make sure to specify the wand type as “3D” in the parameter section of the script.

The script output exported to PDF can be used as a report. The reported RMS deviation is comparable to the standard deviation of the wand in the calibration results.

4.4. Simplified 3D assessment using markers in a fixed configuration

Finally, if there is no 6DOF or 3D reference wand available, a scaled-down assessment can be performed using any rigid object with three or more fixed markers. This can be useful for large volumes, for example marine applications, which require the use of long-range active markers. The assessment gives insight into the coverage (6DOF fill level) and the standard deviation of fixed distance measurements (marker pairs) across the capture volume.

The grid measurement is performed by moving a rigid body with three or more markers through the capture volume.

The following Matlab script can be used for the analysis:

- The Cluster_evaluation_script.mlx live script can be used to perform a detailed analysis of one or more combined files.

The script output exported to PDF can be used as a report.

5. Accuracy test for a 24-camera Arqus A12 system

An accuracy assessment was performed in the mocap studio with a 24-camera Arqus A12 system. The project, including the analysis scripts and reports are available from the Qualisys File Library [QFL].

5.1. System setup

The system was set up as follows (see Figure 1):

- Capture volume: 12m x 8m x 2m (192 m³).
- Camera system: 24x Arqus A12
- Mounting height of the cameras: 2.8m – 3.2m

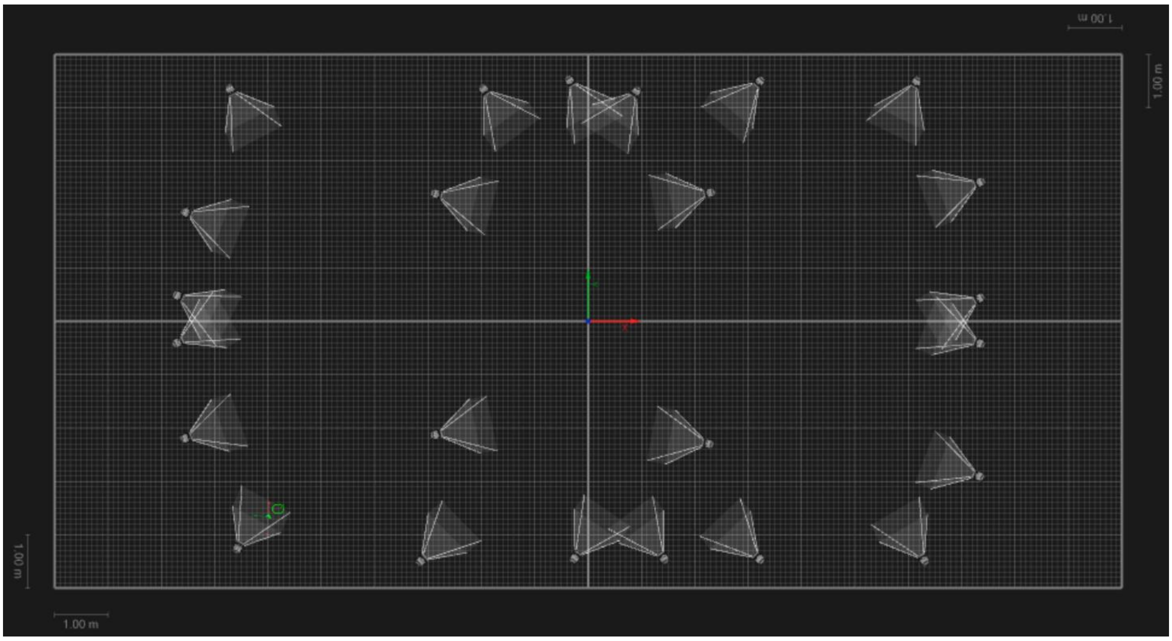

Figure 3: Mocap studio with 24x Arqus A12 camera system.

The capture volume was designed to cover the large volume specified by the ASTM standard. However, the cameras could not be placed with sufficient margin to the capture volume, so that the tracking at the edges of the volume was not optimal. For the analysis, a bounding box was applied limiting the analyzed volume to 11m x 7m x 2m.

5.2. Data collection

A total of three calibrations and four measurements were performed. Data was collected within a time span of 2:15 hours from the first calibration to the last measurement.

QTM version used for capturing: 2023.3 (12456).
Arqus firmware version: 0.05.005

5.2.1. Calibrations

Calibrations were performed with a 600mm carbon fiber wand kit. The calibration was done by systematically walking through the volume orienting the wand in X, Y and Z direction, respectively. The advanced calibration method was used, optimizing both the intrinsic and extrinsic parameters. The duration of the calibration was about 3 minutes. The number of frames used for the calibration was set to 6000 frames (about 1/3 of the total number of frames).

Even though care was taken to remove extra reflections before the calibration, there were still some present in some of the camera views. This seemed to affect one of the calibrations (calibration #3) in particular, which showed relatively high residuals for some of the cameras.

Applying a 2D marker size filter of 200 (min) – 3000 (max) subpixels in post processing led to a significant improvement of the calibration results. This preprocessing step was then applied to all three calibrations, so that they were processed in a consistent way.

The calibration results, average across the three calibrations were:

- Standard deviation of the wand: 0.43 mm
- Camera residuals: 0.43 mm

5.2.2. Measurements

The measurements were performed with the Qualisys metrology wand, described in Section 3.2. In total four measurements were performed, intertwined with the calibrations. The orientation of the metrology wand was varied: for two of the measurements, the wand oriented so that the local Z axis of the clusters was pointing upwards during the scanning of the grid in X-Y direction (wand parallel to the floor), while for the two other measurements the Z axis was pointing downwards.

5.3. Processing and analysis

For the analysis, all calibrations and measurements were cross processed as described in Section 3.4.2, yielding 12 combinations. Tracking and labeling of trajectories at the edges of the volume was negatively affected by the lack of camera coverage, especially at the turning points of the scan lines. For this reason, the prediction error in the 3D tracking settings was lowered from the default value of 25 to 5 mm, which led to an overall improvement of labeling and 6DOF fill levels. For the 6DOF calculations, the refined rigid body definitions of the cluster markers were used, see Section 3.2.

The files were exported to Matlab format. The calculations were performed using the *Analyze_wand_file.mlx* live script for analysis of single files and the *Aggregate_data_6d.m* script for collecting data for the statistical calculations. The results of the latter were used in a pivot table in Excel for the calculation of the average values of the assessment metrics across calibrations and measurements. For the calculations, a bounding box was applied to limit the analysis to an area of 11m x 7m, discarding the limited camera coverage at the edges of the volume.

QTM version used for processing: 2024.2 (13637).

5.4. Results

Figure 2 and Figure 3 show the error distributions for distance error and orientation error for measurement #1 processed with calibration #1. For this combination, the estimated wand length error was 0.05 mm, and the RMS distance error and deviation were 0.13 mm and 0.12 mm, respectively. The right pane of Figure 2 shows that the distance error is consistently low across the analyzed volume.

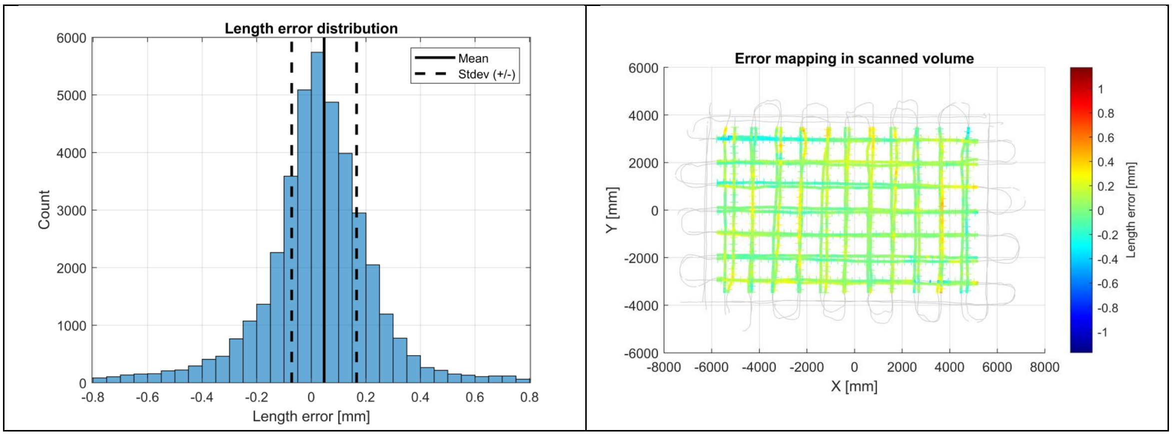

Figure 4: (left) Distance error distribution and (right) distance error mapping across the analyzed volume for the combination of calibration #1 and measurement #1.

The orientation error distribution represents the magnitude of the orientation difference between the respective clusters and is therefore positive only. The RMS of the orientation error distribution was 0.32 deg. The Euler angle distributions show that the orientation offset was close to zero, with average values of 0.02 to 0.05 deg and RMS values below 0.2 deg per component.

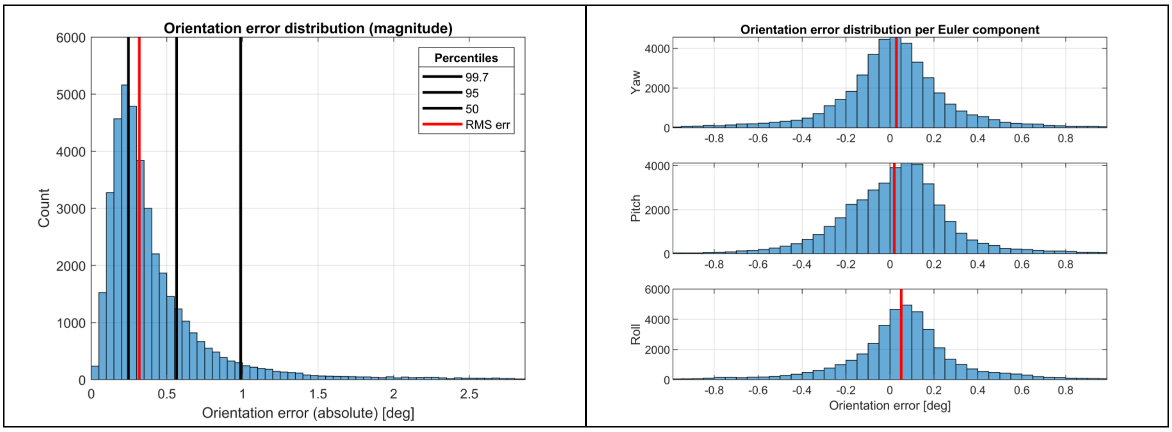

Figure 5: Orientation error distribution for (left) angular magnitude and (right) Euler angles for the combination of calibration #1 and measurement #1.

The accuracy assessment results for the aggregated data are summarized in Table 2 and Table 3 for distance and orientation, respectively. The results show good consistency across calibrations. The RMS distance error was 0.168 mm. The RMS deviation from the mean (standard deviation) was 0.146 mm. The RMS distance error includes the error of the estimated average wand length compared to the reference value, which amounted to 0.079 mm. The overall RMS orientation error was 0.356 deg.

Table 2: Accuracy assessment of distance for a 24x Arqus A12 system covering a 11m x 7m x 2m capture volume. Results are averaged across measurements per calibration. The lowest row shows the grand average across all calibration-measurement combinations.

Cal #	**No. of samples**	**RMS dist. Error [mm]**	**Max dist. Error [mm]**	**50 pctl [mm]**	**95 pctl [mm]**	**99.7 pctl [mm]**
#1	73044	0.150	1.339	0.082	0.294	0.663
#2	73013	0.175	1.400	0.093	0.353	0.678
#3	73022	0.180	1.389	0.098	0.361	0.676
Total (av.)	**219079**	**0.168**	**1.376**	**0.091**	**0.336**	**0.673**

Table 3: Accuracy assessment of orientation for a 24x Arqus A12 system covering a 11m x 7m x 2m capture volume. Results are averaged across measurements per calibration. The lowest row shows the grand average across all calibration-measurement combinations.

Cal #	**No. of samples**	**RMS orient. Error [deg]**	**Max orient. Error [deg]**	**50 pctl [deg]**	**95 pctl [deg]**	**99.7 pctl [deg]**
#1	73044	0.360	2.793	0.265	0.585	1.705
#2	73013	0.353	2.885	0.258	0.555	1.750
#3	73022	0.355	3.258	0.257	0.549	1.774
Total (av.)	**219079**	**0.356**	**2.979**	**0.260**	**0.563**	**1.743**

6. Discussion

6.1. ASTM standard test method: pros and cons

Assessment of the accuracy of optical motion capture systems is a complex topic. The performance of a system depends on many factors, and accuracy requirements may vary considerably depending on the application. For example, measuring the consistency of a robotic arm requires a high precision and repeatability of the measurement of fixed poses, whereas large volume marine systems may require measurement of poses at larger distance scales within a certain level of accuracy.

The ASTM standard test method discussed here presents one possible approach, in which the accuracy assessment is based on the relative pose of two clusters with a fixed relation that are moved through the entire capture volume. The method provides a set of metrics and test procedures for evaluating the performance of motion capture systems. For example, the method can be used to evaluate if newly installed motion capture systems meet the customer's requirements, or to evaluate the impact of new developments, such as improvements in tracking algorithms or calibration methods, by comparing new results with a baseline.

The validity of the ASTM standard is limited to distances or displacements comparable to the length of the used test artifact, which is sufficient for most biomechanical applications. The method does not address the accuracy of small displacements (e.g., repeatability of poses of a robotic arm), or the accuracy of large distances or displacements (e.g., traversed distance of a model in a large tank). A proper evaluation of such applications requires a different approach.

6.2. Discussion of results

The accuracy test presented here was performed for the standard large ASTM volume of 12m x 8m x 2m (192 m^3). However, since the studio did not allow for sufficient distance of the cameras to the edges of volume, the tracking at the edges was inferior. The current analysis was therefore limited to an area of 11m x 7m to give a better representation of the achieved accuracy.

The use of advanced calibration instead of standard calibration was important to achieve the presented results. For comparison, using the same calibration files processed with the standard calibration algorithm, the average distance RMS amounted to 0.67 mm (cf. 0.17 mm for advanced calibration).

The average length error of the test wand was 0.079 mm (0.036, 0.096 and 0.104 for calibrations #1, #2 and #3, respectively). This is within the specified accuracy interval (± 0.1 mm, 2SD) for the calibration wand.

The orientation RMS error was 0.36 deg. In terms of Euler angles, the RMS errors were below 0.2 deg per component (yaw, pitch, roll).

6.3. Comparison with competitors

Table 4 and Table 5 show a comparison of published ASTM assessment results from competitor systems. The results were obtained from the following sources:

- *Vicon #1 and #2*: results published on website in a blog and a FAQ section on accuracy assessment [Vicon].
- *Optitrack #1 and #2*: results published in a conference paper on accuracy assessment [ASTM-test1]. The paper does not specify the camera types used, but from the pictures and the specifications it can be guessed that it is an Optitrack Prime 41 system.
- *Optitrack #3 and Unknown*: results published in a book chapter on accuracy assessment [ASTM-test2] from the same first author as [ASTM-test1]. The Optitrack system is likely the same as the one in [ASTM-test2]. The other system for small volume measurements (2x2x1 m cell for robotic arm) could not be identified.

Table 4: Distance accuracy comparisons of published ASTM assessment results from competitor systems.

Brand	**Cameras**	**RMS dist. Err/dev [mm]**	**Max dist. Error [mm]**	**50 pctl [mm]**	**95 pctl [mm]**	**99.7 pctl [mm]**
Qualisys	24 Arq. A12	0.168 / 0.146	1,376	0.091	0.336	0.673
Vicon #1	24 Vant. V16	0.201 / 0.198	-	-	-	-
Vicon #2	24 Vant. V5	0.321	-	-	-	-
Optitrack #1	12 Prime 41	0.532	31.8	0.15	0.606	2.068
Optitrack #2	12 Prime 41	0.499	30.9	0.136	0.633	2.152
Optitrack #3	12 Prime 41	0.26	-	-	-	-
Unknown	8 cams, small volume	1.05	-	-	-	-

Table 5: Orientation accuracy comparisons of published ASTM assessment results from competitor systems.

Brand	**Cameras**	**RMS orient. Error [deg]**	**Max orient. Error [deg]**	**50 pctl [deg]**	**95 pctl [deg]**	**99.7 pctl [deg]**
Qualisys	24 Arq. A12	0.356	2.979	0.260	0.563	1.743
Vicon #1	24 Vant. V16	-	-	-	-	-
Vicon #2	24 Vant. V5	-	-	-	-	-
Optitrack #1	12 Prime 41	0.349	28.5	0.262	0.565	1.305
Optitrack #2	12 Prime 41	0.334	44.8	0.159	0.569	1.415
Optitrack #3	12 Prime 41	0.2	-	-	-	-
Unknown	8 cams, small volume	0.52	-	-	-	-

6.3.1. Vicon

Vicon presented an accuracy assessment using the ASTM standard and reported the results on their website [Vicon]. Here are the main comparisons:

- RMS distance error and deviation for the Arqus system was better than the Vantage V16 (0.201 mm for Vicon V16 system vs. 0.168 mm for a comparable Arqus A12 system). For the Vantage V5 system (lower resolution cameras) the errors were larger (0.321 mm).
- Vicon did not publish all reportable measures, for example, there are no metrics on orientation accuracy. This makes the comparison incomplete.
- In their marketing story about their assessment study Vicon poses strong claims (0.017 mm accuracy, “better than a human hair”) based on average distance error (average measured distance minus known reference value). However, this metric is not amongst the reportable assessment metrics according to the ASTM standard, and it is not a valid representation of accuracy (in Table 1 the average distance error is included as an optional metric).

Concerning the latter point, it was revealed in a Vicon webinar on this topic that they used volume scaling as an additional calibration step, which directly affects the average distance error. In fact, Vicon calibration devices are not individually certified, and to our knowledge the accuracy level of their calibration devices is not specified. Volume scaling can be applied to compensate for a possible scaling error. However, the volume scaling method requires knowledge of the exact distances between markers, and the way these reference values were obtained was not specified in the publicly available material.

Unlike Vicon, all Qualisys calibration wands are individually calibrated against a reference wand and the exact wand length is specified at an accuracy level of ± 0.1 mm (2 standard deviations). This means that the average distance error (scaling error) is expected to be within ± 0.1 mm for a normal wand calibration.

Vicon claims to have published a white paper on the ASTM accuracy assessments that can be obtained on request. Perhaps this paper includes more details about their testing, but we have been unsuccessful in our efforts to obtain a copy.

6.3.2. Optitrack

ASTM assessment results were published by Bostelman et al. in collaboration with NIST [ASTM-test1, ASTM-test2] verifying a 12 camera system at their robotics lab. The brand and camera type were not specified, but judging from the pictures and the specifications it is most likely an Optitrack Prime 41 system.

In terms of distance (Table 4), the results presented in [ASTM-test1] showed larger errors compared to the Qualisys and Vicon assessments. For the assessment presented in [ASTM-test2] not all results were included, but the RMS distance error was slightly better with a value of 0.26 mm, but still less accurate compared to Qualisys and Vicon assessments.

The orientation RMS error and the percentiles were rather similar in the Qualisys and Optitrack assessments. The large maximum error values presented in the Optitrack assessments (28.5 and 44.8 deg) should probably be considered as outliers.

Reference [ASTM-test2] included an assessment of a small-volume system used for tracking a robotic arm in a 2m x 2m x1m cell. Unfortunately, the brand and camera types were not specified, and could not be clearly identified. The RMS distance error was larger than 1 mm. The explanation for this large error is not clear, but it can be speculated that the standard test artifact used was not suitable for assessment of small capture volumes.

7. References and resources

[ASTM-standard] ASTM E3064 - 16: Standard Test Method for Evaluating the Performance of Optical Tracking Systems that Measure Six Degrees of Freedom (6DOF) Pose.
Url: https://www.astm.org/e3064-16.html

[Q-MAT] Matlab code and report template ASTM-based accuracy assessment alternatives for Qualisys. Repository: https://dev.azure.com/Qualisys/Dev/_git/fast-est-accuracy-matlab

[QFL] ASTM assessment project for 24-camera Arqus A12 system performed in the Mocap studio at Qualisys HQ, including QTM files, Matlab scripts, Excel files with calculated metrics and reports. Url: [not yet published; preliminary available at OneDrive: 2023-12-11 Mocap studio 24xA12-selection]

[Vicon] Vicon blog post on ASTM accuracy assessment, including link to webinar and FAQ section on accuracy. Url: https://www.vicon.com/resources/blog/vicon-study-of-dynamic-object-tracking-accuracy/

[ASTM-test1] Bostelman, R. , Hong, T. , Shah, M. and Legowik, S. (2016), Dynamic Metrology and ASTM E57.02 Dynamic Measurement Standard, Coordinate Metrology Society Conference (CMSC) 2016, Nashville, TN, US, [online], https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=921298 (Accessed January 17, 2024)

[ASTM-test2] Roger Bostelman, Joe Falco, Tsai Hong. Performance Measurements of Motion Capture Systems used for AGV and Robot Arm Evaluation. Roger Bostelman, Elena Messina, “Autonomous Industrial Vehicles: From the Laboratory to the Factory Floor”, STP1594, Chapter 7, May 2016., 2016. hal-01401480. Url: https://hal.science/hal-01401480/document