Accepted: 3 May 2024

DOI: 10.1111/ocr.12808

RESEARCH ARTICLE

Orthodontics & Craniofacial Research WILEY

Three-dimensional video recordings: Accuracy, reliability, clinical and research guidelines – Reliability assessment of a 4D camera

Gabriella Coppola[1] | Dominik Hänggi[1] | Gianluca Cassina[1] | Carlalberta Verna[1] | Nikolaos Gkantidis[2] | Georgios Kanavakis[1,3]

[1]Department of Pediatric Oral Health and Orthodontics, University Center for Dental Medicine Basel UZB, University of Basel, Basel, Switzerland

[2]Department of Orthodontics and Dentofacial Orthopedics, University of Bern, Bern, Switzerland

[3]Department of Orthodontics, School of Dentistry, National and Kapodistrian University of Athens, Athens, Greece

Correspondence
Georgios Kanavakis, Department of Orthodontics, School of Dentistry, National and Kapodistrian University of Athens, Thivon 2, 11527 Athens, Greece.
Email: gkanavak@dent.uoa.gr

Funding information
Schweizerische Zahnärzte Gesellschaft (SSO)

Abstract

Objectives: In addition to studying facial anatomy, stereophotogrammetry is an efficient diagnostic tool for assessing facial expressions through 3D video recordings. Current technology produces high-quality recordings but also generates extremely excessive data. Here, we compare various recording speeds for three standardized movements using the 3dMDface camera system, to assess its accuracy and reliability.

Materials and Methods: A linear and two circular movements were performed using a 3D-printed cube mounted on a robotic arm. All movements were recorded initially at 60 fps (frames/second) and then at 30 and 15 fps. Recording accuracy was tested with best-fit superimpositions of consecutive frames of the 3D cube and calculation of the Mean Absolute Distance (MAD). The reliability of the recordings were tested with evaluation of the inter- and intra-examiner error.

Results: The accuracy of movement recordings was excellent at all speeds (60, 30 and 15 fps), with variability in MAD values consistently being less than 1 mm. The reliability of the camera recordings was excellent at all recording speeds.

Conclusions: This study demonstrated that 3D recordings of facial expressions can be performed at 30 or even at 15 fps without significant loss of information. This considerably reduces the amount of produced data facilitating further processing and analyses.

KEYWORDS
3d imaging, 3dmd face-scan, stereophotogrammetry

1 | INTRODUCTION

The evaluation of the craniofacial structures is an essential part for reaching a diagnosis and performing outcome assessments in various medical and dental specialties including maxillo-facial surgery, orthodontics, plastic and reconstructive surgery. Such an evaluation is mostly done with the use of photographs and X-rays, which are static depictions of the face and hard cranial structures, respectively. Although these are still mostly acquired in two dimensions,[1] the use of three-dimensional imaging technology is becoming more available and has improved the diagnostic accuracy especially in challenging cases with high treatment demands.[2–5]

Nevertheless, there is often a need to perform quantitative measures of facial expressions. These are affected in cases with pathologies, such as in facial nerves, or after plastic facial surgery.[6] In addition, depressive disorders are associated with a decline in the intensity of facial expressions, which is often used to support the diagnosis.[7,8] In dentistry and orthodontics, there is a large interest in the smile and its association to tooth appearance and position. When young adults were asked to evaluate their own facial and smile attractiveness, they appeared more satisfied with wider smiles.[9,10] Also, an aesthetic improvement of the smile is the main motivation for seeking orthodontic treatment in children and adults.[11]

Several methods have been proposed over the years to study facial expression, qualitatively and quantitatively. The House-Brackmann grading system is a visual, qualitative scale that has been used widely in the past.[12] Due to the pure qualitative nature of this method, its use remained limited and more quantitative methods were explored. Using alginate impressions and plaster casts to reconstruct facial expressions was suggested; nevertheless, the method is highly demanding from a technical point of view for both, the operator and the patient, and it provides only static information.[13] The utilization of two-dimensional photographic and videographic technology presents a viable and reasonable solution, because the combination of these two tools provides static and dynamic information, is easy to perform and shows good reproducibility.[13,14] However, the lack of a third dimension could compromise the outcomes, leading to inaccurate and less informative assessments of facial expressions.[15]

Recent technological advancements in three-dimensional imaging led to the commercial development of cost-effective, easy to use 3D camera systems with high computational power, able to record facial movements in three dimensions. This adds the fourth component of 'motion' to a three-dimensional image, and thus these devices are often described as 4D cameras.[16,17] The generated videos can then be used to perform quantitative analyses of facial movements using special software.[17] For this purpose, multiple markers can be fixed on the subject's face and used as reference points to study changes on facial morphology. This method is referred to as a 'marker-based tracking system'. Despite its widespread use, the method introduces errors generated by the inaccurate placement of the markers at different time points.[18] To overcome this limitation, more advanced camera systems have been developed recently, which do not require the use of facial markers and allow fast recording of facial movements in high detail with speeds that reach 60 frames/second (fps). High fps values not only significantly increase the sensitivity of the recordings and provide highly accurate three-dimensional videos but also generate a very large amount of data to be saved and processed.[17] Decreasing the fps number produces less data, but may lead to important loss of information in cases where facial movements are very quick.[17]

Within this context, the most important factors limiting a more widespread adoption of 4D capturing systems in clinical settings are the costs of the camera system, the requirement of a computer with high processing speeds for three-dimensional image rendering, the substantial demand for data storage as well as the costs that are associated with such a demand. This is particularly important taking into account that the file size of an individual frame can range from 4 to 100 MB, depending on the device used and the program configuration.[16] Given that the computer specifications for high processing speeds and a large graphic card memory cannot be compromised if the acquired images are to be viewed properly and processed further, the most realistic option would be to reduce the amount of produced data. In order to do so, without jeopardizing the usability of marker-free 4D camera systems, the camera settings need to be adjusted to adequately capture facial movement, without generating excessive 'data waste'. The present investigation aims to assess the accuracy and reliability of a 4D camera system and provide guidelines for camera setting optimization based on different types of standardized movements.

2 | METHODS

The present study was conducted in the Department for Pediatric Oral Health and Orthodontics at the University Center for Dental Medicine Basel UZB, University of Basel in Switzerland. Ethical approval was waived, because the study does not include human subjects nor utilizes pre-existing patient information. To meet the primary objective of the study, a group of standardized movements, performed identically and repeatedly, was obtained through the use of a robotic arm (UR5e Universals Robots, Universal Robots GmbH, München, Germany). To reduce random error, all recordings were performed twice in one session (by two independent operators) and repeated at a second time point, 4 weeks later by the same two operators. The methods are described in detail below.

2.1 | Robotic arm

The specifications of the used robotic arm (Universal Robots GmbH, München, Germany) allow for medium-load applications (up to 5 kg) and ensure accurate, highly reproducible movements, which are pre-programmed in the built-in software. For this investigation, a 3D printed cube (45 × 45 × 45 mm) with a non-reflective surface was produced out of plastic material (Acrylonitrile Butadiene Styrene [ABS]) and was mounted on the robotic arm. The cube shape was selected because, under the current experimental design, it allows for accurate superimposition in three dimensions in relation to a sphere, for example, since the relative position of its three surfaces can be uniquely defined and approximated to a corresponding object in space. To ensure correct superimposition of consecutive frames, a cross mark was drawn on one of the surfaces for better identification of the cube orientation. The camera position was maintained fixed for all experiments and oriented in a way that its two-pods system provided a complete field of view of the robotic cube movement, in accordance with the eye-to-hand configuration method.[19]

Also, prior to the recordings, a 'hand-eye calibration'[19] of the robotic arm, the camera and their relative position was performed. This minimized possible errors due to differences in the experimental setting.

During the recordings, the cube was placed in front of the camera at a predetermined position, in a way that at least four surfaces and one edge were captured.

The camera performance was tested based on the recordings of the standardized individual cube movements in space. The following types of movements were programmed (Figure 1):

1. A linear movement between two points located on an orthogonal (perpendicular) plane to the recording direction of the camera.
2. A circular movement through the same points and plane mentioned above, hereafter called circular-orthogonal movement.
3. A circular movement between two points located on a plane parallel to the recording direction of the camera, hereafter called circular-parallel movement.

All movements were performed at the same speed of 100 mm/s. This was defined based on the duration and the extent of various facial movements, as reported in the literature (Table S1).[20–22]

2.2 | 4D camera system

The tested camera system was a four-dimensional stereoscopical camera that uses projected unstructured infra-red light to take three-dimensional pictures or videos (3dMDface, 3dMD, Atlanta, GA, USA). The system consists of two modular camera units (MCUs)

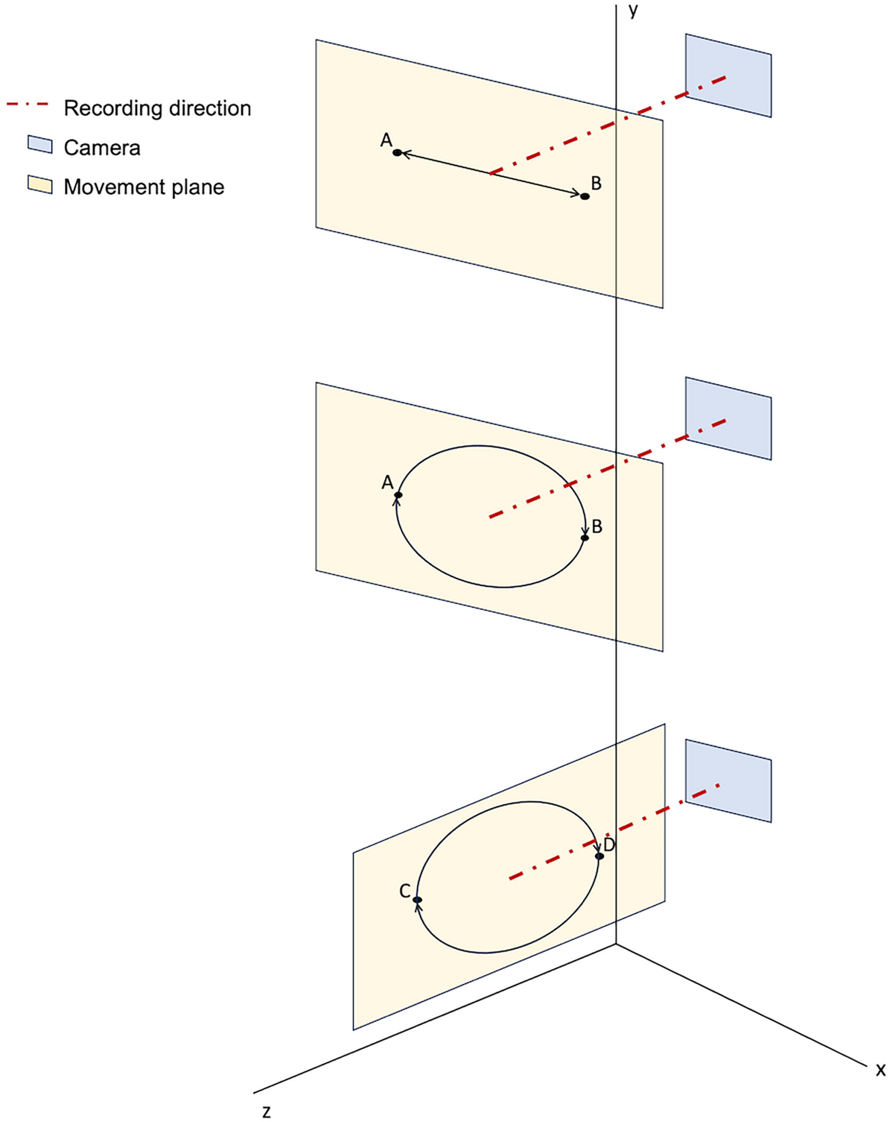

FIGURE 1 Schematic representation of performed movements. *Top*. Linear movement between two points (Point A – Point B), on an orthogonal plane to the recording direction of the camera. *Middle*. Circular movement between point A and point B on an orthogonal plane to the recording direction of the camera. *Bottom*. Circular movement between points C and D on a plane parallel to the recording direction of the camera.

synchronized with a robust LED lighting system. Each unit consists of three 1.3 mega pixel machine vision cameras (two greyscale and one colour one). The camera allows for individual photo acquisition and for video recordings at various speeds between 10 and 60 frames/second (fps).

2.3 | Accuracy and reliability of the 4D camera system

To evaluate the performance of the 4D camera system, a series of recordings were performed by two independent operators, who had previously been calibrated and trained in this system. At first, each operator performed a recording of all standardized movements (see above section 'Robotic arm') at 60 fps. From this, the frames corresponding to 2 seconds (120 frames in total) of the entire recording were extracted and saved as .obj files (60 fps group). To simulate recordings at 30 fps, every second frame of the initial 120 frames were selected and saved separately, forming a total of 60 frames (30 fps group). Similarly, to simulate recordings at 15 fps, every fourth frame was selected, forming a total of 30 frames (15 fps group).

The accuracy of the camera at different recording speeds was tested by measuring the Euclidean distances between the centroids[23] of three adjacent corresponding cube surfaces, depicted in different frames, using Viewbox 4 Software (version 4.1.0.1 BETA 64, dHAL software, Kifisia, Greece). For this purpose, each cube/frame was superimposed with the first one of the respective recording speed group, through the automated application of a best-fit approximation algorithm.[24] The movement required for this superimposition was recorded as the Euclidean distance metric used in the study. For each group (60, 30, 15 fps), the superimposition reference areas (100 triangles on each of the three depicted cube surfaces) were selected manually at the centre of each of the three cube surfaces of the first frame and were used for all superimpositions of the respective group for consistency reasons (Figure 2). After the completion of this process, the *x*, *y* and *z* translations of each object centroid, required for the best fit approximation of the cube surfaces of any frame to those of the first frame, were extracted to an Excel worksheet (Microsoft Excel, Microsoft ©, Richmond WA, USA) for further handling. For each frame, the square root of the sum of the square *x*, *y* and *z* distances was calculated and comprised the Euclidean distance metric used in the study. The outcomes were presented graphically and assessed qualitatively, according to the various pre-determined camera settings. All analyses and graphics were performed using IBM SPSS statistics for Windows (Version 27.0. Armonk, NY: IBM Corp.).

The reproducibility of the method was tested by comparing the outcomes between the two independent operators in all standardized movements and at all recording speeds. The inter-examiner error was presented using Bland–Altman plots,[25] as well as qualitative assessment.

The reliability of the method was assessed in two ways. In order to determine the intra-operator error, one operator repeated all superimpositions in the 60 fps group, 4 weeks after the initial processing. This assessment was also performed through Bland–Altman plots.

To assess the reliability of the camera performance, all standardized movements were recorded a second time at a speed of 30 fps and a third time at a speed of 15 fps. The outcomes of these additional recordings were compared to the outcomes of the extracted frames from the initial recording performed at 60 fps (30 fps group and 15 fps group, respectively). As a result, six additional comparisons were performed and presented graphically for qualitative assessment.

3 | RESULTS

The comparative assessment of the camera performance at various frames per second is displayed in Figures 3–5 for all recorded types of movements. The graph for the linear movements (Figure 3) shows that in all cases there was an initial stage when the cube accelerated to the pre-determined speed. During this stage, the change in cube position was sudden until it stabilized, at which point the Euclidean distances between frames become more subtle. For each fps group, the two identical, consecutive 'waves' represent two consecutive linear movements during which the cube moves between points A and B. Each 'wave' has an acceleration phase (ascending line), a linear phase (horizontal line) and a deceleration phase (descending line) followed by the next acceleration phase. The irregularity in the graph path during the linear phase represents the irregularity in cube movement produced by the robot. A comparative observation of the three lines in Figure 3 reveals that at faster recording speeds slightly smaller irregularities occur. However, the range in the irregularity of the movement stays within 1 mm in all three recording speeds. This means that there is minimal loss of movement data even when the camera is set at 15 fps.

Figures 4 and 5 depict the circular movements of the cube perpendicular (*x–y* plane) and parallel to the camera (*y–z* plane), respectively. Here, the irregularity in the movements increases as the recording speed decreases. Nevertheless, inaccuracies remain within the level of 1 mm, within which the recordings at 30 and 15 fps show insignificant loss of information.

In regard to the reproducibility of the camera recordings, the Bland–Altman plots of the measurements performed by two independent operators show very good agreement (Figures S1–S3), with differences consistently below 0.2 mm. The reliability of the camera recordings was tested through repeated recordings at 30 and 15 fps, as described above. The results of this assessment are presented graphically in Figures S4a–c and S5a–c. Each of these images presents a comparison between two lines. One (Recording 1) represents the movements displayed in frames extracted from the initial recording at 60 fps

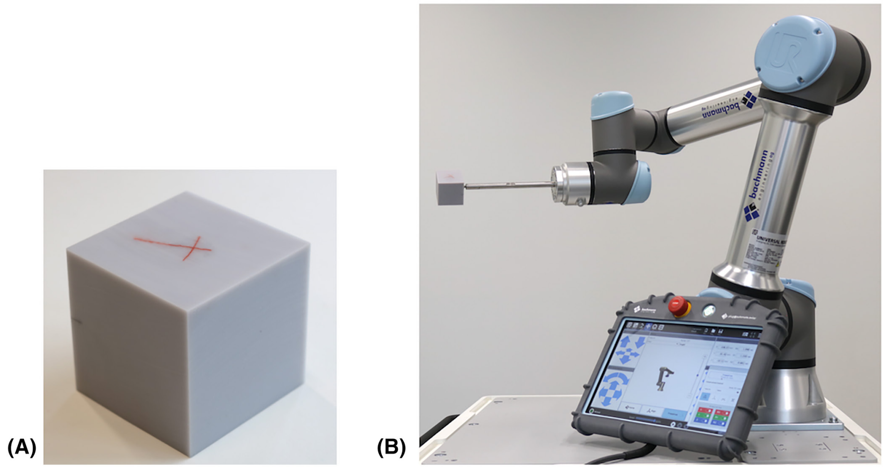

FIGURE 2 A, The 3D printed cube that was recorded during all standardized movements. The red cross helped with cube orientation. B, The cube mounted on the robotic arm using a custom-made extension. The extension provided enough distance between the cube and the robotic arm during its movements, so that the robotic arm was not included in the video recordings.

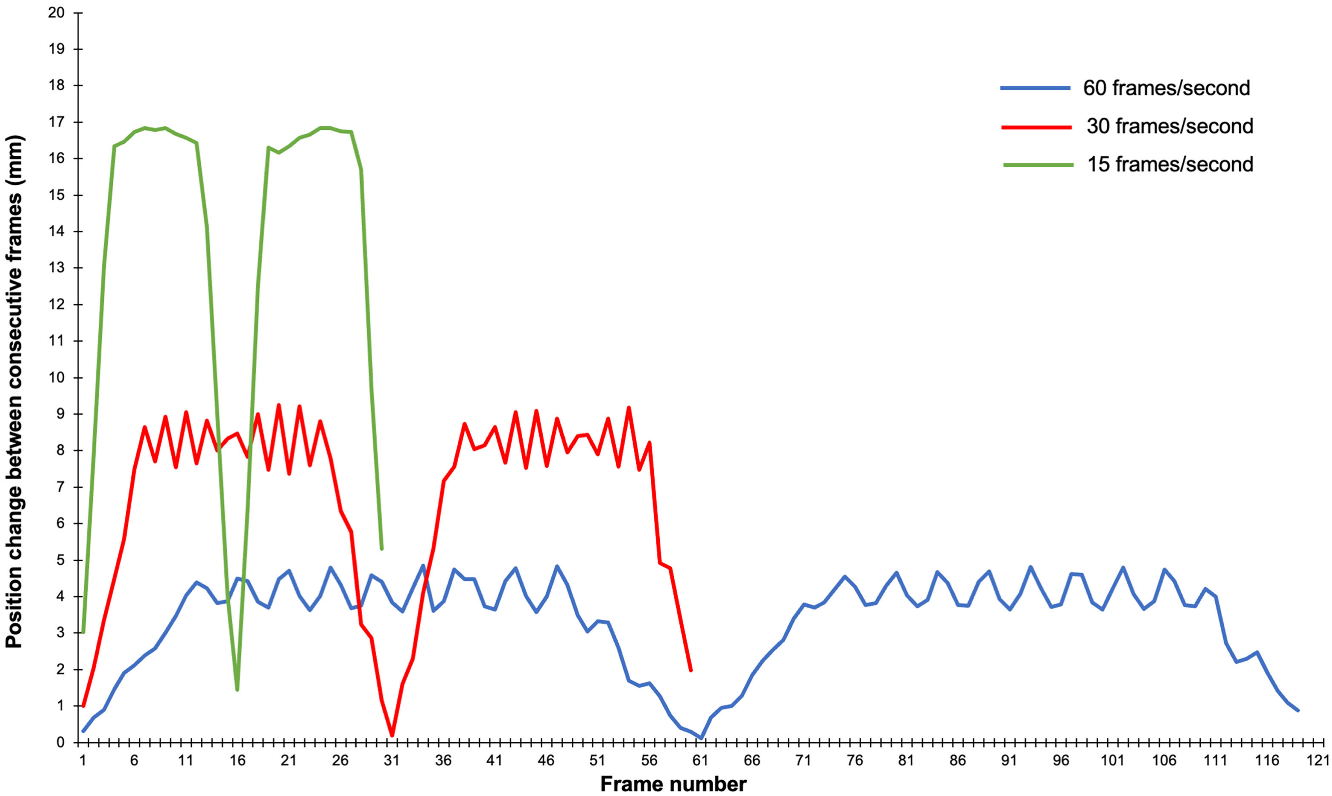

FIGURE 3 Positional change in mm between consecutive frames during linear movement of the object. Each line represents a recording speed.

that correspond to an equivalent recording at 30 fps or 15 fps, respectively. The other line (Recording 2) represents the movement captured during separate recordings at 30 fps or 15 fps revealing that for most movements the difference in movement increments between the two recordings is minimal. This indicates that the surface information provided by the camera is the same whether frames are extracted from a recording at maximum speed (60 fps) or whether the recording speed is reduced at 30 fps or 15 fps.

The reliability of the operator (intra-operator error) was assessed with Bland–Altman plots and was considered very good with differences not exceeding 0.4 mm even in extreme cases (Figures S6–S8).

4 | DISCUSSION

The purpose of this investigation was to perform a thorough assessment of the accuracy and reliability of a widely used 4D camera system (3dMD) in performing video recordings of standardized movements at various speeds. The results allow for valuable recommendations regarding the optimal camera speed settings (fps)

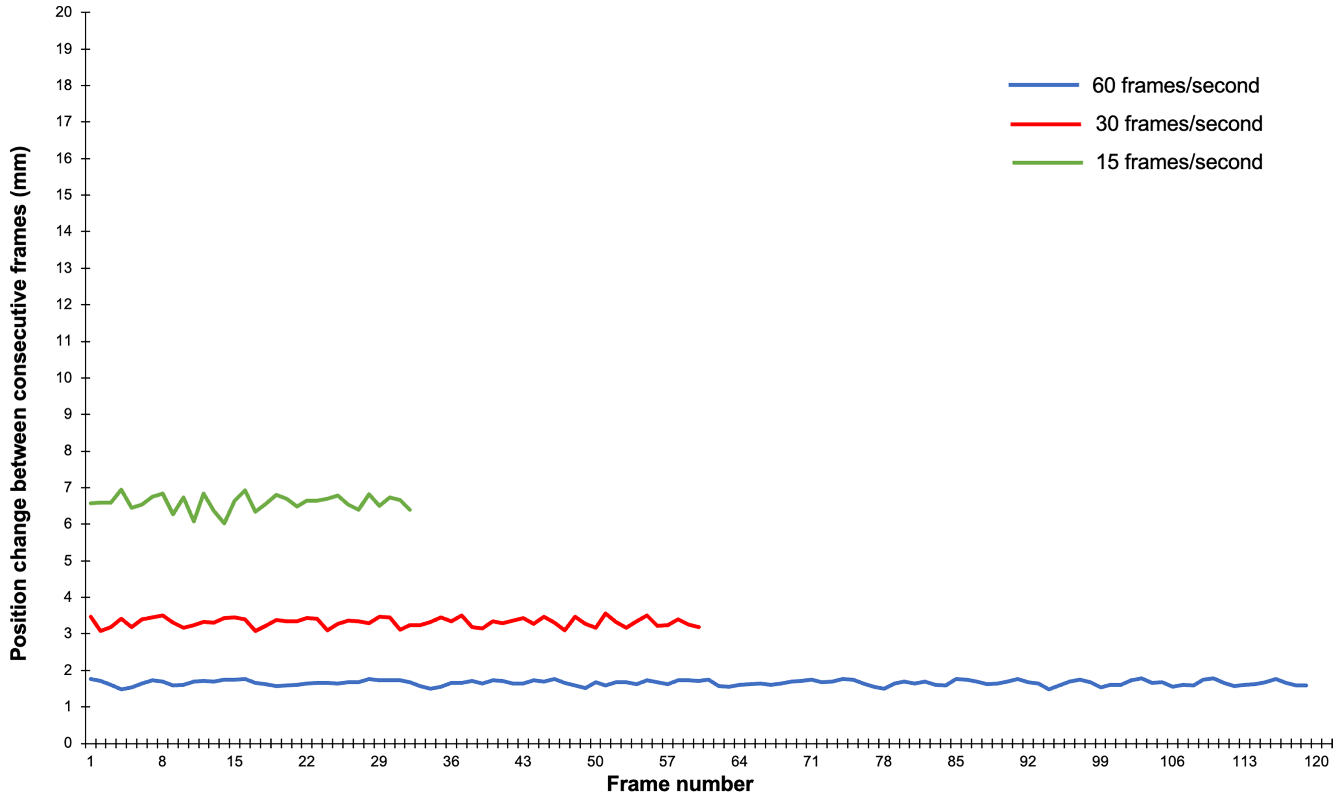

FIGURE 4 Positional change in mm between consecutive frames during circular movement of the object, on a plane perpendicular to the camera. Each line represents a recording speed.

when studying facial expressions in three dimensions. A reduction in frames per second without significant loss of information is desirable to facilitate the storing and processing of the generated data, which are usually quite large, and thus, time-, cost- and resource-consuming.

A 3D printed cube was used as the object performing the standardized movements in space by the robot. The configuration and the calibration method were set to minimize possible errors due to differences in the experimental setting. Nevertheless, it can be expected that small differences occur with every set up, positioning and calibration of the equipment. These differences are seen in the difference between inter- and intra-rater errors, with intra-rater error being larger than inter-rater error. The experiment was performed by two independent operators on the same day and did thus not require a new set up of the robotic arm. Intra-rater reliability, however, was evaluated with measurements performed 4 weeks after the first ones, which required a new set up, positioning and calibration of all equipment. This could have led to larger differences, which, however, are also inconsiderable.

The recordings were performed with the 3dMD camera system, which has shown excellent results in capturing facial volumetric data and facial movements.[26] Prior to stereophotogrammetry, the acquisition of three-dimensional facial surface information was only possible with direct anthropometry, a technique in which landmarks are placed directly on the face and measurements are performed in vivo in the form of linear and angular values.[27] Although anthropometry has been widely used, it can be reliable and is cost-effective, it also requires the operator to be trained, the patients to be collaborative, it is time-consuming, and does not allow for data storage and verification.[28] Thus, there are questions about its validity, especially when assessing volumetric changes such as facial swelling[26] or when studying facial expressions, which in themselves are highly variable.[29] In comparison to direct anthropometry, 3D stereophotogrammetry systems have showed satisfying accuracy and reliability in performing facial assessments and are equally precise to direct anthropometry in landmark identification.[30]

In the present study, we used the 3dMD camera to record a series of three-dimensional videos. When compared to other cameras, the chosen system demonstrated exceptional repeatability and precision, along with the highest capture and processing speed, and the lowest error in geometric representation. A single frame could take up to 8 seconds to be processed, a relatively short time compared to other cameras that require between 20 and 120 seconds. Nevertheless, it remains an important time-consuming process, especially considering a recording speed of 60 fps.[16–18] The use of video recordings for diagnostic purposes is still not widely used due to its limited diagnostic applications and the difficulties in data storage and processing. However, there are advantages in capturing the face dynamically rather than statically. In addition to expanding the diagnostic possibilities into capturing facial movements and expressions,[31] 3D videos are also less affected by erratic movements in restless patients, primarily in young children.[32]

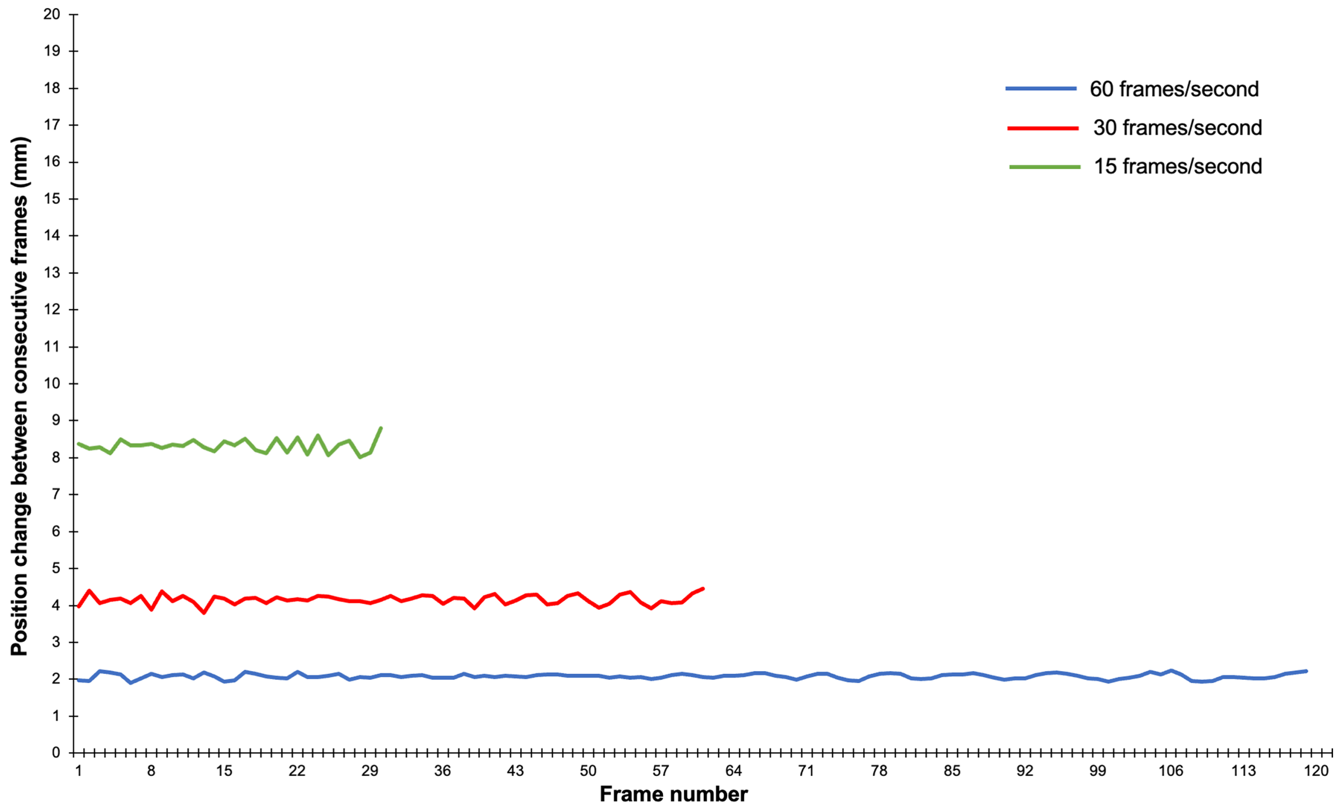

FIGURE 5 Positional change in mm between consecutive frames during circular movement of the object, on a plane parallel to the camera. Each line represents a recording speed.

Thus, instead of capturing a single 3D image of a face, a series of 3D frames are captured in the form of a video. This allows for both static and dynamic evaluations; a selection of a single video frame allows for static facial assessment and a selection of a series of frames provides dynamic data that can be used to study facial expressions or movements. Nevertheless, there are significant limitations relative to 3D video recordings that are related to the excessive amount of data generated. In reality, researchers and clinicians record long videos, capturing a series of facial expressions in order to build a more complete patient database that allows for various diagnostic analyses. For example, a 3D video recording of 1 minute at 60 fps generates 3600 3D frames. Each 3D frame, including texture and colour, is approximately 7 MB, leading to a 25 200 MB (25.2GB) video. This reduces to 12 600 MB (12.6 GB) if the speed is reduced to 30 fps and to 6300 MB (6.3 GB) if the recording speed is reduced to 15 fps. Therefore, adjusting the recording speed would significantly reduce the generated data. The present study showed that such a reduction would potentially not lead to significant loss of information when studying facial expressions. This finding has major implications regarding the computational resources and the associated processing requirements when using three-dimensional video recordings, especially in the case of large data sets or high clinical demands. On the basis of the above storage estimates, a video recording of a patient performing a multitude of complex facial movements could exceed a duration of 2 or 3 minutes, which would normally require more than 75 GB with a recording speed of 60 fps. Being able to save 50% of this huge storage demand without losing any potentially significant information could make a notable impact on the infrastructure of any clinical or research setting. This is particularly useful in centres that provide care for patients with craniofacial deformities or facial paralysis, and need to evaluate their ability to perform a series of facial movements at various points in time.[33] The same applies in large clinical centres providing medical or dental services to a large number of patients. A centre, for example, that serves 500 new patients per year and uses 1-minute 3D videos as a screening tool, could benefit from a reduction of 63 TB per year in storage demands. In addition, more and more medical research is based on exploring massive data sets as part of multi-centre clinical trials or 'big data' analyses. Being able to reduce recording speed of three-dimensional videos without compromising the quality and accuracy of the data, provides a vital option for research groups, with strong implications on funding demands. In countries with limited research funds, finding ways to minimize costs might be crucial in many cases. For the purposes of this study, a robotic arm was used to perform standardized movements in three dimensions, at a predetermined speed, which was higher than the reported speeds of the most common facial expressions.

Due to the scarce available information regarding the performance of 3D camera systems in capturing facial movements, an in vitro methodology was preferred in order to eliminate confounders related to the variability of human facial expressions. The variability of facial expressions and their reproducibility depend on their pattern, magnitude and speed. These characteristics also depend on whether an expression is verbal or non-verbal, on its voluntary/involuntary nature or if it is a micro- or macro-expression.[31] There is no agreement in the literature regarding the reliability of verbal and non-verbal expression. While there are reports that verbal expressions are more reproducible, because they are more refined and are performed more frequently,[31,32] others have found that non-verbal expressions are more reproducible. Five non-verbal expressions have been studied extensively, namely maximum smile, lip purse, cheek puff, maximum raising of the eyebrows and forceful eye-closure. It is considered that these expressions represent the muscle activity controlled by each of the main branches of the facial nerve.[29,34,35] The least reproducible has been found to be the 'cheek puff',[34] and the most reproducible the 'smile'.[36,37] Also, expressions requiring maximum muscle contraction tend to be more reproducible, such as a 'widest possible smile' in comparison to a spontaneous smile.[34] The same is true for expressions associated with increased muscle memory, such as a posed smile, which is a frequently performed movement, and thus, tends to be more reproducible than a spontaneous smile.[36,37]

The duration of an expression is the characteristic differentiating between micro- and macro-expressions or conventional ones. Micro-expressions usually last between 0.2[22] and 0.5 seconds[20,22,38] and are the main focus in research studying involuntary facial expressions. These are also known as 'leaked expressions' as they are extremely hard, if not impossible to suppress, they are very subtle and fast.[22] All facial muscle activities slower than micro-expressions are considered conventional or macro-expressions. The duration of smiling expression, for instance, which is the most common one, ranges between 0.5 and 0.75 seconds in various social contexts.[39] For the purpose of this study and based on the available literature, a movement speed that would represent the magnitude and speed of all possible facial expressions was necessary. With the selected movement speed of 100 mm/s, even the largest facial expressions, such as mouth opening, are included at large. Also, facial expressions are the result of facial muscle activity, whose minimal contraction speed is 0.1 second.[22] Given that with a recording speed of 15 fps a new frame is captured every 0.067 seconds, this speed would, in theory, be adequate to capture even the finest and fastest facial expressions. The results of this investigation confirm the above thought process, showing that even at recording speeds of 15 fps, the loss of information is always less than 1 mm. Considering also the significant reduction in the amount of data, it can be concluded that a recording speed of 15 fps is preferable to record and evaluate facial expression and movement. The authors realize the limitations of an in vitro study and the shortcomings related to not recording actual facial expressions, but rather a sequence of standardized, less complicated movements. Taking the high variability and unpredictability of human facial expressions into account, it might therefore be considered more clinically relevant, in critical cases, to perform video recordings at a speed of 30 fps, to eliminate even the slightest possibility for information loss. Ideally, future investigations should repeat the present study in humans and confirm the present guidelines for clinical use.

5 | CONCLUSION

This well-controlled, in vitro, methodological study shows that evaluations of three-dimensional facial movements can be performed with a recording speed of 15 fps or 30 fps, rather than 60 fps. This option provides accurate and reliable data for dynamic facial assessment without significant loss information and with a significant reduction in the amount of generated data.

AUTHOR CONTRIBUTIONS

G.C. contributed to data acquisition, analysis, and interpretation, drafted the first version of the manuscript and critically revised the final version of manuscript; D.H. contributed to data acquisition, analysis and interpretation and critically revised the manuscript; G.C. contributed to data acquisition and interpretation and critically revised the manuscript; C.V. contributed to data analysis and interpretation and critically revised the manuscript; N.G. contributed to study design, data analysis and interpretation, validation of the study and critically revised the manuscript. G.K. contributed to conceptualization, data analysis and interpretation, validation, wrote the final manuscript and supervised the project. All authors gave final approval and agreed to be accountable for all aspects of the work.

ACKNOWLEDGEMENTS

All the authors of the present papers declare that they have no financial or personal relationships with other individuals or organizations that could inappropriately influence or bias our work. This includes any potential conflicts of interest related to employment, consultancies, stock ownership, honoraria, expert testimony, patent applications, or grants or other funding. The manuscript represents original and unpublished work, and all sources used are appropriately cited. Any conflicts of interest, financial or otherwise, that might be perceived as influencing the results or interpretation of the work have been disclosed in this statement. If any changes in circumstances or affiliations could lead to potential conflicts of interest, this statement will be promptly updated, and the relevant parties informed.

FUNDING INFORMATION

This study was supported with research funding granted to Georgios Kanavakis by the Swiss Dental Association (SSO). Grant Nr.: SSO 320-20.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

All data sets are available at Zenodo: https://doi.org/10.5281/zenodo.8134866.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

ORCID

Gabriella Coppola https://orcid.org/0009-0004-4971-7245
Nikolaos Gkantidis https://orcid.org/0000-0002-3180-6272
Georgios Kanavakis https://orcid.org/0000-0001-7304-044X

REFERENCES

1. Graf CC, Dritsas K, Ghamri M, Gkantidis N. Reliability of cephalometric superimposition for the assessment of craniofacial changes: a systematic review. *Eur J Orthod*. 2022;44(5):477-490.
2. Kanavakis G, Häner ST, Matthey F, Gkantidis N. Voxel-based superimposition of serial craniofacial cone-beam computed tomographies for facial soft tissue assessment: reproducibility and segmentation effects. *Am J Orthod Dentofacial Orthop*. 2021;159(3):343-351.
3. Häner ST, Kanavakis G, Matthey F, Gkantidis N. Voxel-based superimposition of serial craniofacial CBCTs: reliability, reproducibility and segmentation effect on hard-tissue outcomes. *Orthod Craniofac Res*. 2019;23(1):92-101.
4. Wampfler JJ, Gkantidis N. Superimposition of serial 3-dimensional facial photographs to assess changes over time: a systematic review. *Am J Orthod Dentofacial Orthop*. 2022;161(2):182-197.e2.
5. Häner ST, Kanavakis G, Matthey F, Gkantidis N. Valid 3D surface superimposition references to assess facial changes during growth. *Sci Rep*. 2021;11(1):16456.
6. Grewe CM, Schreiber L, Zachow S. Fast and accurate digital morphometry of facial expressions. *Facial Plast Surg*. 2015;31(5):431-438.
7. Brozgold AZ, Borod JC, Martin CC, Pick LH, Alpert M, Welkowitz J. Social functioning and facial emotional expression in neurological and psychiatric disorders. *Appl Neuropsychol*. 1998;5(1):15-23.
8. Jaeger J, Borod JC, Peselow E. Facial expression of positive and negative emotions in patients with unipolar depression. *J Affect Disord*. 1986;11(1):43-50.
9. Horn S, Matuszewska N, Gkantidis N, Verna C, Kanavakis G. Smile dimensions affect self-perceived smile attractiveness. *Sci Rep*. 2021;11:2779.
10. Kanavakis G, Halazonetis D, Katsaros C, Gkantidis N. Facial shape affects self-perceived facial attractiveness. *PLoS One*. 2021;16(2):e0245557.
11. Samsonyanova L, Broukal Z. A systematic review of individual motivational factors in orthodontic treatment: facial attractiveness as the main motivational factor in orthodontic treatment. *Int J Dent*. 2014;2014:938274.
12. Jeong J, Lee JM, Cho Y, Kim J. Inter-rater discrepancy of the House–Brackmann facial nerve grading system. *Clin Otolaryngol*. 2022;47(6):680-683.
13. Saunders TR, Hansen NA. Synthetic casting tape as a facial impression tray material. *J Prosthet Dent*. 1995;74(2):169-170.
14. Wood DA, Hughes GB, Secic M, Good TL. Objective measurement of normal facial movement with video microscaling. *Am J Otol*. 1994;15(1):61-65.
15. Perez-Gomez V, Rios-Figueroa HV, Rechy-Ramirez EJ, Mezura-Montes E, Marin-Hernandez A. Feature selection on 2D and 3D geometric features to improve facial expression recognition. *Sensors*. 2020;20(17):4847.
16. Tzou CH, Artner NM, Pona I, et al. Comparison of three-dimensional surface-imaging systems. *J Plast Reconstr Aesthet Surg*. 2014;67(4):489-497.
17. Hallac RR, Feng J, Kane AA, Seaward JR. Dynamic facial asymmetry in patients with repaired cleft lip using 4D imaging (video stereophotogrammetry). *J Craniofac Surg*. 2017;45(1):8-12.
18. Popat H, Richmond S, Benedikt L, Marshall D, Rosin PL. Quantitative analysis of facial movement—a review of three-dimensional imaging techniques. *Comput Med Imaging Graph*. 2009;33(5):377-383.
19. Enebuse I, Ibrahim BKSMK, Foo M, Matharu RS, Ahmed H. Accuracy evaluation of hand-eye calibration techniques for vision-guided robots. *PLoS One*. 2022;17(10):e0273261.
20. Schmidt KL, Swearingen JMV, Levenstein RM. Speed, amplitude, and asymmetry of lip movement in voluntary puckering and blowing expressions: implications for facial assessment. *Mot Control*. 2005;9(3):270-280.
21. Cohn JF, Schmmidt KL. The timing of facial motion in posed and spontaneous smiles. *Int J Wavelets Multiresolution Inf Process*. 2004;2:121-132.
22. Yan WJ, Wu Q, Liang J, Chen YH, Fu X. How fast are the leaked facial expressions: the duration of micro-expressions. *J Nonverbal Behav*. 2013;37(4):217-230.
23. Zelditch M, Swiderski D, Sheets HD. *Geometric Morphometrics for Biologists: A Primer*. Academic Press; 2012.
24. Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell*. 1992;14:239-256.
25. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat*. 2007;17(4):571-582.
26. Van der Meer WJ, Dijkstra PU, Visser A, Vissink A, Ren Y. Reliability and validity of measurements of facial swelling with a stereophotogrammetry optical three-dimensional scanner. *Br J Oral Maxillofac Surg*. 2014;52(10):922-927.
27. Farkas LG, Deutsch CK. Anthropometric determination of craniofacial morphology. *Am J Med Genet*. 1996;65(1):1-4.
28. Wong JY, Oh AK, Ohta E, et al. Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images. *Cleft Palate Craniofac J*. 2008;45(3):232-239.
29. Lübbers HT, Medinger L, Kruse AL, Grätz KW, Obwegeser JA, Matthews F. The influence of involuntary facial movements on craniofacial anthropometry: a survey using a three-dimensional photographic system. *Br J Oral Maxillofac Surg*. 2012;50(2):171-175.
30. Ghoddousi H, Edler R, Haers P, Wertheim D, Greenhill D. Comparison of three methods of facial measurement. *Int J Oral Maxillofac Surg*. 2007;36(3):250-258.
31. Ju X, O'Leary E, Peng M, Al-Anezi T, Ayoub A, Khambay B. Evaluation of the reproducibility of nonverbal facial expressions using a 3D motion capture system. *Cleft Palate Craniofac J*. 2016;53(1):22-29.
32. Lane C, Harrell W. Completing the 3-dimensional picture. *Am J Orthod Dentofacial Orthop*. 2008;133(4):612-620.
33. Trotman CA, Faraway J, Hadlock TA. Facial mobility and recovery in patients with unilateral facial paralysis. *Orthod Craniofac Res*. 2020;23(1):82-91.
34. Sawyer AR, See M, Nduka C. Assessment of the reproducibility of facial expressions with 3-D stereophotogrammetry. *Otolaryngol Head Neck Surg*. 2009;140(1):76-81.
35. Alagha MA, Ju X, Morley S, Ayoub A. Reproducibility of the dynamics of facial expressions in unilateral facial palsy. *Int J Oral Maxillofac Surg*. 2018;47(2):268-275.
36. Dobreva D, Gkantidis N, Halazonetis D, Verna C, Kanavakis G. Smile reproducibility and its relationship to self-perceived smile attractiveness. *Biology*. 2022;11(5):719.
37. Dindaroğlu F, Duran GS, Görgülü S, Yetkiner E. Social smile reproducibility using 3-D stereophotogrammetry and reverse engineering technology. *Angle Orthod*. 2015;86(3):448-455.
38. Valstar MF, Pantic M, Ambadar Z, Cohn JF. Spontaneous vs. posed facial behavior: automatic analysis of brow actions. *Proc 8th Int Conf Multimodal Interfaces*. New York: ACM Press; 2006:162-170.

39. Schmidt KL, Ambadar Z, Cohn JF, Reed LI. Movement differences between deliberate and spontaneous facial expressions: zygomaticus major action in smiling. *J Nonverbal Behav*. 2006;30(1):37-52.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Coppola G, Hänggi D, Cassina G, Verna C, Gkantidis N, Kanavakis G. Three-dimensional video recordings: Accuracy, reliability, clinical and research guidelines – Reliability assessment of a 4D camera. *Orthod Craniofac Res*. 2024;00:1-10. doi:10.1111/ocr.12808