American Journal of Orthodontics & Dentofacial Orthopedics

THREE-DIMENSIONAL FACIAL CAPTURE USING A CUSTOM-BUILT PHOTOGRAMMETRY SETUP: DESIGN, PERFORMANCE, AND COST

--Manuscript Draft--

Manuscript Number:	
Article Type:	Techno Bytes
Corresponding Author:	Hans L. L. Wellens, DDS. In private practice Brugge, Belgium
First Author:	Hans L. L. Wellens, DDS, PhD
Order of Authors:	Hans L. L. Wellens, DDS, PhD
	Hanne Hoskens, Ir
	Peter Claes, Ir, PhD
	Anne Marie Kuijpers-Jagtman, DDS, PhD, FDSRCSEng
	Alejandra Ortega-Castrillón, Ir, PhD
Abstract:	Abstract: Introduction: Although stereo-photogrammetry is increasingly popular for scanning faces three-dimensionally, commercial solutions remain quite expensive, limiting their accessibility. We propose a more affordable, custom-built photogrammetry setup (SF3D), and evaluate its variability within- and between-systems. Methods: 29 subjects and a mannequin head were imaged three times using SF3D and a commercially available system (3dMDFace). Next, an anthropometric mask was mapped visco-elastically onto the reconstructed meshes using Meshmonk. Within-systems shape variability was determined by calculating the RMSE of the Procrustes distance between each of the three subject's scans and the subject's ground truth (calculated by averaging the mappings after a non-scaled generalized Procrustes superimposition). Inter-system variability was determined by similarly comparing the ground truth mappings of both systems. Two-factor Procrustes ANOVA was used to partition the inter-system shape variability to better understand the source of the discrepancies between the facial shapes acquired by both systems. Results: The RMSE of the within-system shape variability for 3dMDFace and SF3F were 0.52 +/- 0.07 mm, and 0.45 +/- 0.17 mm, respectively. The corresponding values for the mannequin head were 0.42 +/- 0.02 mm, and 0.29 +/- 0.03 mm, respectively. The between-systems RMSE was 1.6 +/- 0.34 mm for the study group, and 1.38 mm for the mannequin head. Two-factor ANOVA indicated that variability attributable to the system was expressed mainly at the upper eyelids, nasal tip and alae, and chin. Conclusions: The variability values of the custom-built setup presented here were competitive to a state-of-the-art commercial system at a more affordable level of investment.

THREE-DIMENSIONAL FACIAL CAPTURE USING A CUSTOM-BUILT PHOTOGRAMMETRY SETUP:

DESIGN, PERFORMANCE, AND COST

Corresponding author: Hans L. L. Wellens[1] DDS, PhD

Hanne Hoskens[2,3] Ir

Peter Claes[2,3,4,5,6] Ir, PhD

Anne Marie Kuijpers-Jagtman[7,8] DDS, PhD, FDSRCSEng

Alejandra Ortega-Castrillón[2,3] Ir, PhD

e-mail address corresponding author: wellens.hans@telenet.be

1 In private practice, Groene-Poortdreef 16, 8200 Sint-Michiels, Belgium.
2 Department of Electrical Engineering, ESAT/PSI, KU Leuven, Leuven, Belgium.
3 Medical Imaging Research Center, UZ Leuven, Leuven, Belgium.
4 Department of Human Genetics, KU Leuven, Leuven, Belgium.
5 Murdoch Children's Research Institute, Melbourne, Australia.
6 Department of Biomedical Engineering, University of Oxford, Oxford OX3 7DQ, United Kingdom.
7 Department of Orthodontics, University Medical Center Groningen, The Netherlands.
8 Faculty of Dentistry, Universitas Indonesia, Jakarta, Indonesia.

Highlights

- A custom-built photogrammetry system for 3D facial capture was presented: SF3D.
- SF3D's vertex count (resolution) was 10 fold higher compared to 3dMDFace.
- SF3D was slightly more precise, while also slightly more variable compared to 3dMDFace.
- Overall, SF3D's performance matched and at times surpassed that of 3dMDFace.

AJODO Techno bytes section

Abstract:

Introduction: Although stereo-photogrammetry is increasingly popular for scanning faces three-dimensionally, commercial solutions remain quite expensive, limiting their accessibility. We propose a more affordable, custom-built photogrammetry setup (SF3D), and evaluate its variability within- and between-systems. **Methods:** 29 subjects and a mannequin head were imaged three times using SF3D and a commercially available system (3dMDFace). Next, an anthropometric mask was mapped visco-elastically onto the reconstructed meshes using Meshmonk. Within-systems shape variability was determined by calculating the RMSE of the Procrustes distance between each of the three subject's scans and the subject's ground truth (calculated by averaging the mappings after a non-scaled generalized Procrustes superimposition). Inter-system variability was determined by similarly comparing the ground truth mappings of both systems. Two-factor Procrustes ANOVA was used to partition the inter-system shape variability to better understand the source of the discrepancies between the facial shapes acquired by both systems. **Results:** The RMSE of the within-system shape variability for 3dMDFace and SF3F were 0.52 +/- 0.07 mm, and 0.45 +/- 0.17 mm, respectively. The corresponding values for the mannequin head were 0.42 +/- 0.02 mm, and 0.29 +/- 0.03 mm, respectively. The between-systems RMSE was 1.6 +/- 0.34 mm for the study group, and 1.38 mm for the mannequin head. Two-factor ANOVA indicated that variability attributable to the system was expressed mainly at the upper eyelids, nasal tip and alae, and chin. **Conclusions:** The variability values of the custom-built setup presented here were competitive to a state-of-the-art commercial system at a more affordable level of investment.

INTRODUCTION

In orthodontics, facial esthetics has traditionally been scrutinized using the lateral cephalogram's profile outline combined with classic two-dimensional facial photographs. With the introduction of cutting-edge three-dimensional (3D) imaging techniques into the orthodontic and/or craniofacial diagnostic toolset, (structured light) photogrammetry setups and 3D facial images derived from full-size CBCT exposures[1] have increasingly been adopted for this purpose. Three-dimensional CBCT does carry an increased radiation burden compared to traditional 2D radiology, especially if image quality is of primary concern. Combined with the ALARA principle[2,3], this has so far precluded its use as de facto imaging solution for orthodontic diagnosis, at least in Europe[4]. Additionally, the restraints required to immobilize the patient's head during image capture using CBCT may potentially obscure facial regions of interest such as the forehead and/or chin area. Combined with ethical objections associated with repeatedly exposing patients to ionizing radiation for growth monitoring purposes, treatment follow-up or outcome assessment, this entails there might be a bright future for non-ionizing methods for diagnosing facial esthetics, growth or treatment change, such as photogrammetry.

A number of studies report on the accuracy and reliability of various commercially available photogrammetry solutions applied in an orthodontic/craniofacial setting. These include both *active* stereo-photogrammetry solutions (which illuminate the patient's face with invisible structured light patterns to provide the features required for interpreting and reconstructing the face's three-dimensional geometry) from manufacturers such as AxisThree[5], and *passive* ones (which reconstruct the scene directly from visual cues present in the acquired image), from Canfield (e.g. VECTRA)[6–11] and 3DiD[5,12–15]. *Hybrid* stereophotogrammetry solutions, which combine both techniques to achieve an optimal result[5] have been presented as well by 3dMD[16–23]. Even fairly low-cost solutions, such as David's SLS-2[24], Fuel3D's Scanify[25,26], and Microsoft's Kinect[27] have been investigated.

Interestingly, some of the aforementioned studies use direct anthropometry (i.e. caliper and measuring tape) as the required golden standard, notwithstanding the notable variability of the latter[8,12,17,22,28]. Aside from the human face's general sparsity in terms of clearly definable landmarks, direct anthropometry is additionally hampered by skin compressibility, and slight changes in facial expression. Some studies attempt to minimize the effects of landmark identification error by pre-labelling the facial surfaces[8,9,11–13,15,17,18,23,29–32]. Both tissue compressibility and facial pose variation can be circumvented by performing measurements on a mannequin head[8,11,12,15,17,18,20,23,24], or plaster casts[13,25]. Other studies replace direct anthropometric measurements with (repeated) digital ones [18,21–23,29], use electromagnetic digitizers, coordinate measuring machines[13,15,30], or other stereophotogrammetry devices [10,11,15,17,24,25,27,33] as the gold standard for comparison.

One problem very few studies offer a solution to, is the very feature-sparse nature of the human face, which entails broad regions having few landmarks at which accuracy and reliability can be gauged. One relatively straightforward solution might be to use elastic deformation of a standard anthropometric mask to densely sample and model the entire facial surface with a very large number of landmarks which, by virtue of the elastic deformation, are effectively homologous[34].

In general, cutting edge technologies such as 3D facial capture typically demand a significant premium over traditional, two-dimensional methods. Budgetary constraints often deprive orthodontic departments and private practices alike from access to these technologies. Aside from cost, one of the major impediments to the democratization of this technology has traditionally been the imposing complexity of the photogrammetric algorithms involved in reconstructing the 3D scene. This has changed somewhat with the introduction of relatively affordable, high quality multi-base photogrammetry software. In the latter, image reconstruction proceeds from a relatively large set of images taken from multiple slightly different viewpoints. This approach was made possible due to fairly recent, fast-paced innovations in the field, combined with the ever-increasing availability of relatively low-cost computational power. Combined, this begs the question of whether it would be possible to design, build and test a custom-built photogrammetry-based setup for 3D facial capture, and to assess and report on its accuracy. The latter is the aim of this study.

MATERIALS AND METHODS

Stereo-face3D (SF3D) custom-built system:

The system, custom-built by the first author (HW), consists of fourteen Canon EOS 1200D DSLR cameras (Canon, Tokyo, Japan), mounted on a square aluminum frame (measuring one by one meter), assembled from industry standard system profiles (45 x 45 mm cross-section, with a 10 mm slot) (Motedis, Ensdorf, Germany) (Fig 1). The frame is attached to a similarly constructed, wheel-mounted support assembly (measuring 1.12 m wide, 1.7 m high and 0.7 m deep), which provides a working surface for the control switches and laptop/computer, and houses the power supplies, USB hubs and electronics (Fig 1 and 2). Two height-adjustable Bosch Rexroth lifts (Bosch Rexroth, Lohr am Main, Germany) allow the frame's height to be adjusted over a distance of approximately 0.4 m.

The cameras are mounted in a semi-hemispherical arrangement in order to better accommodate the human facial form (Fig 1, A). This is accomplished both by applying a mild 'inset' to the four central cameras (i.e. positioning them slightly 'out of plane', further away from the patient) (Fig 2, blue arrows), as well as by the more forward position of the outermost cameras, resulting from their inward rotation around the 28 mm round aluminum struts they are connected to (Fig 1, A).

Adjustable camera mounts allow for precise control over the camera positions and angles (Multi-mount 6, Vanguard, Guangdong, China).

Aside from the cameras, the frame also supports three remote controlled, high CRI, 5600K LED panels (Godox, Shenzhen, China) which provide uniform, shadow-free illumination. These are located on the frame's upper and left and right corners (2 x Godox LED500LW), and on its lower-middle section (1 x Godox LED308W) (Fig 1, A and 2). These panels are low-weight, dimmable, flicker-free and do not generate heat. They are powered by a separate Meanwell HRPG-150-15 enclosed power supply unit (15V, 10A) (MeanWell, GuangZhou, China). Patient positioning is facilitated by 2 diode line-lasers (Picotronic, Koblenz, Germany) on either side of a 20 by 20 cm mirror, positioned approximately in the center of the frame (Fig 1, A and 2). Furthermore, three eye-safe 660 nm random pattern (dot) lasers provide additional texture to the relatively feature-sparse human face (SL-660-S-C, Osela, Lachine, Canada), mounted slightly obliquely from the upper left and right corner, and centrally from below (Fig 1, A and 2).

Removing cameras from the setup to replace depleted batteries is highly undesirable in a carefully calibrated setup. This is avoided by using Canon DC-10 DC couplers (Canon, Tokyo, Japan), which in turn are fed directly from a Meanwell SP-320-7.5 V power supply unit (40 A) (MeanWell, GuangZhou, China). Apart from the power cable, each camera requires one USB cable for image transfer and one 2.5 mm jack cable for camera focusing/triggering. The USB cables are connected to two 7-port industrial USB hubs, which are powered by the 7.5V MeanWell power supply unit. The same power supply also feeds the electronics (after appropriate downregulation of the voltage), consisting of an Arduino Uno microcontroller (https://www.arduino.cc/) in combination with pushbuttons/relays for controlling the positioning and random pattern lasers, and 8-bit shift registers/optocouplers for focusing and triggering the cameras (Fig 1, B). The electronic components were soldered on three Adafruit Perma-Proto full-size PCB boards (Adafruit, New York, USA).

Camera settings and image retrieval are controlled using Smartshooter 3 GRID software (https://kuvacode.com/smart-shooter) installed on a laptop, while the imported images are reconstructed using 3DFlow's 3DF Zephyr PRO (multibase) stereophotogrammetry software on a desktop computer (https://www.3dflow.net/3df-zephyr-pro-3d-models-from-photos/).

Workflow

Preparing for image-capture typically involves removing the cameras' lens covers, powering up the setup and laptop, flipping the 'camera reset switch' on the working surface (thus providing power to the cameras too). The LED panels light up when the setup is powered on, after which their

brightness can be adjusted using a remote control. The whole startup takes between one to two minutes, and does not have to be repeated when capturing multiple subjects sequentially.

After seating the subject centered in front of the setup, a headband displaying four machine-vision markers is fitted. This headband, which serves to scale the reconstructed facial mesh to life size, is positioned such that as much of the forehead as possible remains exposed, while also ensuring that the markers are visible to a sufficient number of cameras (at least three, preferably more). Any loose hair is tucked away behind it in the process. The positioning lasers are then activated, and the subject is instructed to look into the centrally placed mirror with the nose slightly tilted upwards. The frame's height and anteroposterior position are subsequently adjusted to align the projections of both laser crosses on the subject's facial midline subnasally, thus providing an easy visual cue for both patient and practitioner to confirm proper positioning (Fig 3). Upon instructing the subject to maintain a 'neutral (i.e. relaxed) facial expression', the image acquisition button is pressed, after which all cameras automatically and simultaneously focus and trigger. An example of the images acquired by each camera is presented in Fig 4. The process is then continued on a computer fitted with a sufficiently powerful NVIDIA graphics card (NVIDIA, Santa Clara, USA). After loading the images into 3DF Zephyr Pro, the reconstruction into a 3D mesh takes about 5 to 15 minutes to complete, depending on the desired mesh resolution and system specifications (at which point the presence of the imaging subject is no longer required)

After image acquisition, the setup can be left on by simply shutting off the cameras and lights, to speed up image capture for the next subject.

System validation

Assessing the accuracy of the SF3D setup ideally required both intra- and inter-system evaluations, for which we had access to both the frequently used 3dMDFace and Vectra H1 systems. Since the latter requires three acquisitions from different angles to perform one facial reconstruction while both SF3D and 3dMDFace are used as stationary, single shot systems, we opted to use the 3dMDFace system for comparison. Briefly, the 3dMD system is a hybrid structured-light stereo-photogrammetry system consisting of three (stereo-)pairs of two cameras each, with one pair positioned centrally in front of the patient, while the other two are placed on either side.

A study group of thirty volunteers of diverse ethnicity was recruited [redacted], using as exclusion criterion having undergone any facial surgical interventions and having dense facial hair such as mustache and/or beard. The age and sex distribution of the sample is given in Table I. To account for the highly variable nature of human facial expression, image acquisitions were repeated three times for each

individual consecutively, using the methodology presented above. Furthermore, the technical baseline performance of both systems, defined as the performance in the absence of biological variability, i.e. facial pose variability, was assessed and compared by scanning a mannequin head three times consecutively with each system.

All acquisitions were mapped using MeshMonk[35], an open-source software toolbox which allows for spatially dense (i.e. high-resolution) registration of 3D surfaces. MeshMonk non-rigidly (i.e. visco-elastically) maps an anthropometric mask (i.e. a 'landmark template') onto the previously generated 3D facial surfaces, thus establishing high-resolution configurations of quasi-landmarks, which are homologous across subjects[34,36,37]. Homologous, in this context refers to the position of each quasi-landmark relative to all other quasi-landmarks being identical in all individuals[37]. In contrast to similar studies focusing on smaller sets of manually placed landmarks, information is thus provided on the entire facial surface, including regions which are rarely evaluated, like the eyelids and nostrils.

Both the variation within and between systems under investigation was calculated by superimposing the Meshmonk-generated, mapped facial meshes using a generalized Procrustes analysis without scaling[38] (GPA): the meshes were therefore translated to the origin and rotated to minimize the squared distance between the corresponding (quasi-)landmarks, but not scaled to centroid size, effectively superimposing them in size-and-shape space. Furthermore, each subject's true facial shape (i.e. the ground truth) was established by calculating the mean of the three GPA-superimposed acquisitions. The Procrustes distance to the ground truth (i.e. the Euclidean distance between two corresponding landmark configurations of Procrustes coordinates) was then used to quantify shape variation, reported as the root-mean-square error (RMSE). This approach was applied both to the test subjects' mapped reconstructions, as well as those of the mannequin head. A schematic representation of the acquisitions is shown in Fig 5.

System validation then proceeded by calculating and comparing 3dMD's and SF3D's precision, defined here as the mean difference in RMSE between repeated measures of the same subject, comparable to the approach reported in Aldridge et al[39]. For *intra-system* validation, the three acquisitions of each subject as well as the mannequin head were compared to the corresponding averaged shape by calculating the RMSE values. *Inter-system validation* involved comparing the average shape of each set of three acquisitions per subject across both systems by calculating their RMSE values, as well as the normal distances between these average shapes. The two average shapes of the mannequin head generated using both systems were similarly compared. Furthermore, a two-factor Procrustes ANOVA[40] was used to partition the inter-system shape variation in order to better understand the factors of the discrepancies between the facial shapes acquired with both

photogrammetry systems. This is comparable to Aldridge et al.'s assessment of system error (defined here as the proportion of total variance attributable to a particular factor)[39]. For a detailed explanation of the use and interpretation of two-factor ANOVA, the reader is referred to the pertaining studies by Claes et al.[41,42].

RESULTS

Fig 6 provides an example of surface reconstructions of the mannequin head, as acquired by the 3DMD system (Fig 6, left) and the SF3D setup (Fig 6, right). Facial meshes from the latter system on average contained about 350.000 vertices; those originating from the 3DMDFace system had about 34.000. Since a headband with machine vision markers was used in the SF3D system, the amount of forehead included in the analysis had to be reduced somewhat. In order to effectively compare the facial surfaces generated by both systems, the facial area depicted in red in Fig 7 was therefore used for all mappings of the reconstructed meshes. One subject's set of SF3D generated facial meshes exhibited very strong artifacts in the region of the forehead upon mapping. On closer inspection of the accompanying photographs, this proved attributable to the headband being positioned too low in combination with hear sticking out from under it. This subject was therefore removed from further analysis.

The *intra-system variation*, quantified as the RMSE between each repeated acquisition/reconstruction and the corresponding average shape per subject, is shown in Fig 8. The unfilled boxplots and compact boxplots represent the 3dMDFace and SF3D system, respectively. The study group's mean RMSE value was 0.52 (SD 0.07) mm when imaged by 3dMDFace and 0.45 (SD 0.17) mm for SF3D. When imaging the mannequin head, 3DMDFace's RMSE was found to be 0.42 (SD 0.02) mm, while the corresponding value for SF3F was 0.29 (SD 0.03) mm.

Inter-system variation, computed as the RMSE between the average shape of each subject across both systems, is presented in Fig 9. The average RMSE of the study group was 1.6 (SD 0.34) mm, while the mannequin's corresponding value was 1.38 mm. To better understand the location and direction of these differences, a heatmap of the normal distances between the two systems' overall mean shape (i.e. that of the entire study group) was plotted in Fig 10, A. The corresponding results for the mannequin head are depicted in Fig 10, B. A further partitioning of the shape differences between the study group's mean shapes as generated by 3dMDFace and SF3D was performed with a two-factor ANOVA, as shown in Fig 12. All values are normalized per column.

All the RMSE means and standard deviations of the inter- and intra-systems variation for subjects and mannequin head are listed in Table II.

When partitioning the facial shape differences between both systems into individual variability versus system using the two-factor ANOVA, the inter-subject variations were mainly located at the nose, the lower-posterior part of the cheeks and to a lesser extent the chin (Fig 12, upper row). The choice of system was mainly expressed at the level of upper eyelids, the nasal tip and alae, and the chin area (Fig 12, second row). The very low random error found (Fig 12, fourth row) confirmed that the main source of variation indeed were individual differences in facial morphology as well as the system used.

DISCUSSION

In this study we presented a custom-built stereo-photogrammetry solution for 3D facial capture, and assessed its performance 'in vitro' (on a mannequin head) and in vivo (in the study group) both within-, and between systems (comparing to the commercially available 3dMDFace system).

When visually comparing the custom built SF3D system to the 3dMDFace system (Fig 6, top left and right), it is quite obvious that the higher vertex count provided by SF3D - 350.000 on average compared to 34.000 by 3dMDFace - resulted in a more detailed facial mesh. Given that the 3dMDFace system covers not only the face but also the shoulder region, its actual facial vertex count is even lower. The resulting difference in detail is most obvious around the mouth, nose and eyes. On the other hand, the forehead was much better represented in the 3DMDFace meshes. This was a consequence of difficulties in recovering the position of the machine vision markers in the SF3D generated images using the 3DZephyr software package. Although this problem has since been resolved, for the purpose of this study the headband required for scaling the SF3D generated meshes had to be positioned fairly low, partly obscuring the forehead (Fig 6, upper right image).

On closer scrutiny of the images mapped by Meshmonk, the differences in detail between both systems become less obvious (Fig 6, lower panes). This is attributable to the anthropometric mask's resolution which currently contains about 7.160 vertices. While this vertex count makes perfect sense based upon the tradeoff between resolution and computational load, the resulting mappings do present a severe down-sampling of the high-resolution meshes generated by the SF3D system. As a result, although overall facial shape is well preserved by the mapped representations, the details of the surfaces acquired with both facial scanners, and the SF3D system in particular, look significantly smoothed out after mapping.

With respect to the *intra-system variation*, the SF3D setup exhibited lower RMSE yet higher variability values for both the study group (0.6 mm lower mean RMSE and 0.1 mm higher SD, Table II) and the mannequin head (0.12 mm lower RMSE and a 0.01 mm higher SD, Table II) compared to 3dMDFace. Both systems' lower RMSE values when acquiring the mannequin head were not unexpected, due to the absence of biological variability associated with changes in facial pose. Interestingly, when comparing the *differences in RMSE means and SDs within systems*, the influence of facial pose variability was found to be larger for the SF3D (mean RMSE difference of study subjects versus mannequin head: 0.160 versus 0.098 mm for SF3D and 3dMDFace respectively, SD: 0.136 versus 0.051 mm respectively, Table II). This might tentatively be attributed to the increased resolution of the SF3D system, allowing smaller variations in facial morphology to be picked up, thus explaining why the intra-system RMSE SD values between both systems differed by only 0.013 mm for the mannequin head, whereas an almost 8-fold larger value was found for the study group (0.1 mm, Table II).

Similarly, for the *inter-system variability*, a lower mean RMSE value was found for the mannequin head compared to the study group, corroborating the influence of the biological error when comparing facial surface meshes. Interestingly, the study group's heatmap of normal distances indicated that the 3dMDFace system consistently positioned the upper eyelids slightly more posteriorly and the lower eyelids slightly anteriorly compared to the SF3D system (Fig 10, A). Similarly, the nasal tip and alae, lips, lower-posterior part of the cheeks and chin were also positioned slightly more anteriorly by the 3dMDFace system compared to the SF3D (Fig 10, A).

The results of this study indicate that it is possible to design and build a multi-base stereo-photogrammetry setup which rivals and, in some aspects, surpasses commercially available solutions for 3D facial capture, albeit at considerable effort. Overall, the errors reported for both systems under investigation (introduced by facial pose variability and/or system choice) were found to be quite small and well within the accepted limits of the geometric morphometric community. The basic question therefore amounts to whether the aim (high-quality 3D facial capture at 'reduced' cost) justifies the means (the time and effort invested in acquiring all system components and putting them together). For the final build of the SF3D system, we spent about 8.900 USD on materials and 4.900 USD on software, totaling 13.800 USD without VAT. This would appear to compare favorably to the 3dMDFace system which, as configured in the present study, amounted to more than 50.000 dollars (in 2013) including delivery, installation, training and software (but not including annual maintenance fees). Although, the amount mentioned above for SF3D does not include the cost of labor pertaining to assembly, anyone aiming to replicate this project would benefit from not having to 'reinvent the wheel', at considerable savings in both cost and time. Furthermore, there is a significant leeway in terms of 'how good is good enough' with regard to the resolution obtained by the SF3D system: it may

very well be sufficient to reduce the number of cameras in the currently proposed system from 14 to 10, equally lowering the number of required DC couplers, mounts, cables and electronics. Although this would reconstruct the face to a somewhat lower resolution compared to SF3F, it would still be considerably higher than what can be obtained using commercially available alternatives. Further cost-reductions might be obtained by omitting the random pattern lasers, although we found this to be quite detrimental to the obtained mesh quality, mostly so in pre-pubertal children with very feature-sparse faces (e.g. in the absence of freckles and pimples, wrinkles etc.).

A possibly more fundamental question is what resolution we really need for 3D facial capture in an orthodontic/craniofacial-surgical context. The results of the current study suggest the SF3D system is slightly more precise while also slightly more variable compared to the 3dMDFace system (Table II), both of which are probably attributable to the increased resolution of the former compared to the latter system. We would argue the increased resolution makes sense if the focus of research or diagnosis is aimed towards (the effects of treatment on) the chin, lips and paranasal areas, since these regions exhibited the largest between-systems differences (apart from the eyelids). This might particularly be the case for orthodontic treatment involving extractions or functional appliances, craniofacial surgical patients and cleft lip and palate patients, if the number of patients is limited such that computational time is less important. Studies involving higher-resolution meshes such as those generated by the SF3D system should probably use an anthropometric mesh with a higher vertex count, in order to be able to maintain most of the original meshes' detail in the mapped versions thereof. Furthermore, the increased variability associated with the use of higher resolution mesh representations of the human face may be countered by averaging multiple 3D captures of the same patients, thus preserving detail while improving reliability.

Aside from overall cost and algorithmic complexity, another major impediment to the democratization of stereo-photogrammetry has traditionally been the absence of (freely) available software tools for comprehensive facial mesh analysis. Although the reconstruction software provided with the various 3D scanners provides some analytical tools, their utility is usually limited to performing iterative closest point algorithms and deriving distance heatmaps. Since the introduction of sliding semi-landmarks[43,44] and visco-elastic mapping[36,37,34] to the geometric morphometric toolbox, multiple software packages such as Geomorph[45], Morpho[46], mesheR (https://github.com/zarquon42b/mesheR/blob/master/R/mesheR-package.R), and Meshmonk[35] have been made available free of charge. Their application to the generated datasets does require some (mild) programming effort. Alternatively, Halazonetis' Viewbox software (http://www.dhal.com/) also has some built-in functions for sliding semi-landmark analysis of facial meshes.

With regard to the limitations of this study, it needs to be pointed out only adults were included into the study group. The variability in facial expression we report is therefore expected to be smaller compared to studies including babies and young children[47].

CONCLUSION

1. The StereoFace3D system was found to be slightly more precise, while also slightly more variable, compared to the 3dMDFace system, both of which are probably attributable to the higher resolution of the StereoFace3D system.
2. The errors reported for both systems under investigation (introduced by facial pose variability and/or system choice) were found to be small and well within the accepted limits of the geometric morphometric community.
3. The performance of the custom-build stereo-photogrammetry setup equaled and in some aspects even surpassed, commercially available solutions for three-dimensional facial capture.

REFERENCES

1. Fourie Z, Damstra J, Gerrits PO, Ren Y. Accuracy and repeatability of anthropometric facial measurements using cone beam computed tomography. *Cleft Palate-Craniofacial J.* 2011;48(5):623–30.

2. Anon. Radiation Protection series publications - Energy - European Commission. *Energy*. Available at: /energy/en/radiation-protection-publications. Accessed February 24, 2018.

3. Nederlandse Vereniging van Orthodontisten. Richtlijn Orthodontie Röntgen. 2018;[in press].

4. Halazonetis DJ. Cone-beam computed tomography is not the imaging technique of choice for comprehensive orthodontic assessment. *Am. J. Orthod. Dentofac. Orthop.* 2012;141(4):403-407.

5. Tzou C-HJ, Artner NM, Pona I, et al. Comparison of three-dimensional surface-imaging systems. *J. Plast. Reconstr. Aesthetic Surg.* 2014;67(4):489–97.

6. de Menezes M, Rosati R, Ferrario VF, Sforza C. Accuracy and reproducibility of a 3-dimensional stereophotogrammetric imaging system. *J. Oral Maxillofac. Surg. Off.* 2010;68(9):2129–35.

7. Othman SA, Ahmad R, Mericant AF, Jamaludin M. Reproducibility of facial soft tissue landmarks on facial images captured on a 3D camera. *Aust. Orthod. J.* 2013;29(1):58–65.

8. Metzler P, Sun Y, Zemann W, et al. Validity of the 3D VECTRA photogrammetric surface imaging system for cranio-maxillofacial anthropometric measurements. *Oral Maxillofac. Surg.* 2014;18(3):297–304.

9. Andrade LM, Rodrigues da Silva AMB, Magri LV, Rodrigues da Silva MAM. Repeatability Study of Angular and Linear Measurements on Facial Morphology Analysis by Means of Stereophotogrammetry. *J. Craniofac. Surg.* 2017;28(4):1107–11.

10. Camison L, Bykowski M, Lee WW, et al. Validation of the Vectra H1 portable three-dimensional photogrammetry system for facial imaging. *Int. J. Oral Maxillofac. Surg.* 2018;47(3):403–10.

11. Gibelli D, Pucciarelli V, Cappella A, Dolci C, Sforza C. Are Portable Stereophotogrammetric Devices Reliable in Facial Imaging? A Validation Study of VECTRA H1 Device. *J. Oral Maxillofac. Surg. Off. J. Am. Assoc. Oral Maxillofac. Surg.* 2018;76(8):1772–84.

12. Winder RJ, Darvann TA, McKnight W, Magee JDM, Ramsay-Baggs P. Technical validation of the Di3D stereophotogrammetry surface imaging system. *Br. J. Oral Maxillofac. Surg.* 2008;46(1):33–7.

13. Khambay B, Nairn N, Bell A, Miller J, Bowman A, Ayoub AF. Validation and reproducibility of a high-resolution three-dimensional facial imaging system. *Br. J. Oral Maxillofac. Surg.* 2008;46(1):27–32.

14. Naudi KB, Benramadan R, Brocklebank L, Ju X, Khambay B, Ayoub A. The virtual human face: superimposing the simultaneously captured 3D photorealistic skin surface of the face on the untextured skin image of the CBCT scan. *Int. J. Oral Maxillofac. Surg.* 2013;42(3):393–400.

15. Kook M-S, Jung S, Park H-J, et al. A comparison study of different facial soft tissue analysis methods. *J. Cranio-Maxillo-fac. Surg.* 2014;42(5):648–56.

16. Weinberg SM, Scott NM, Neiswanger K, Brandon CA, Marazita ML. Digital three-dimensional photogrammetry: evaluation of anthropometric precision and accuracy using a Genex 3D camera system. *Cleft Palate-Craniofacial J.* 2004;41(5):507–18.

17. Weinberg SM, Naidoo S, Govier DP, Martin RA, Kane AA, Marazita ML. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the Genex and 3dMD imaging systems with one another and with direct anthropometry. *J. Craniofac. Surg.* 2006;17(3):477–83.

18. Lübbers H-T, Medinger L, Kruse A, Grätz KW, Matthews F. Precision and accuracy of the 3dMD photogrammetric system in craniomaxillofacial application. *J. Craniofac. Surg.* 2010;21(3):763–7.

19. Ort R, Metzler P, Kruse AL, et al. The Reliability of a Three-Dimensional Photo System-(3dMDface-) Based Evaluation of the Face in Cleft Lip Infants. *Plast. Surg. Int.* 2012;2012:138090.

20. 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):922–7.

21. Nord F, Ferjencik R, Seifert B, et al. The 3dMD photogrammetric photo system in cranio-maxillofacial surgery: Validation of interexaminer variations and perceptions. *J. Cranio-Maxillo-fac. Surg.* 2015;43(9):1798–803.

22. Dindaroğlu F, Kutlu P, Duran GS, Görgülü S, Aslan E. Accuracy and reliability of 3D stereophotogrammetry: A comparison to direct anthropometry and 2D photogrammetry. *Angle Orthod.* 2016;86(3):487–94.

23. Hong C, Choi K, Kachroo Y, et al. Evaluation of the 3dMDface system as a tool for soft tissue analysis. *Orthod. Craniofac. Res.* 2017;20 Suppl 1:119–24.

24. Secher JJ, Darvann TA, Pinholt EM. Accuracy and reproducibility of the DAVID SLS-2 scanner in three-dimensional facial imaging. *J. Cranio-Maxillo-fac. Surg.* 2017;45(10):1662–70.

25. Ritschl LM, Roth M, Fichter AM, et al. The possibilities of a portable low-budget three-dimensional stereophotogrammetry system in neonates: a prospective growth analysis and analysis of accuracy. *Head Face Med.* 2018;14(1):11.

26. Liu C, Artopoulos A. Validation of a low-cost portable 3-dimensional face scanner. *Imaging Sci. Dent.* 2019;49(1):35–43.

27. Maués CPR, Casagrande MVS, Almeida RCC, Almeida M a. O, Carvalho F a. R. Three-dimensional surface models of the facial soft tissues acquired with a low-cost scanner. *Int. J. Oral Maxillofac. Surg.* 2018;47(9):1219–25.

28. Düppe K, Becker M, Schönmeyr B. Evaluation of Facial Anthropometry Using Three-Dimensional Photogrammetry and Direct Measuring Techniques. *J. Craniofac. Surg.* 2018;29(5):1245–51.

29. Lübbers H-T, 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):171–5.

30. de Menezes M, Rosati R, Allievi C, Sforza C. A photographic system for the three-dimensional study of facial morphology. *Angle Orthod.* 2009;79(6):1070–7.

31. Naini FB, Akram S, Kepinska J, Garagiola U, McDonald F, Wertheim D. Validation of a new three-dimensional imaging system using comparative craniofacial anthropometry. *Maxillofac. Plast. Reconstr. Surg.* 2017;39(1):23.

32. Weinberg SM. 3D stereophotogrammetry versus traditional craniofacial anthropometry: Comparing measurements from the 3D facial norms database to Farkas's North American norms. *Am. J. Orthod. Dentofacial Orthop.* 2019;155(5):693–701.

33. Zhao Y-J, Xiong Y-X, Wang Y. Three-Dimensional Accuracy of Facial Scan for Facial Deformities in Clinics: A New Evaluation Method for Facial Scanner Accuracy. *PloS One* 2017;12(1):e0169402.

34. Claes P, Walters M, Clement J. Improved facial outcome assessment using a 3D anthropometric mask. *Int. J. Oral Maxillofac. Surg.* 2012;41(3):324–330.

35. White JD, Ortega-Castrillón A, Matthews H, et al. MeshMonk: Open-source large-scale intensive 3D phenotyping. *Sci. Rep.* 2019;9(1):6085.

36. Claes P. A robust statistical surface registration framework using implicit function representations-application in craniofacial reconstruction. (Thesis) 2007.

37. Claes P, Walters M, Vandermeulen D, Clement JG. Spatially-dense 3D facial asymmetry assessment in both typical and disordered growth. *J. Anat.* 2011;219(4):444–455.

38. Dryden IL, Mardia KV. *Statistical Shape Analysis*. Wiley; 1998.

39. Aldridge K, Boyadjiev SA, Capone GT, DeLeon VB, Richtsmeier JT. Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am. J. Med. Genet. A.* 2005;138A(3):247–53.

40. Klingenberg CP, Barluenga M, Meyer A. SHAPE ANALYSIS OF SYMMETRIC STRUCTURES: QUANTIFYING VARIATION AMONG INDIVIDUALS AND ASYMMETRY. *Evolution* 2002;56(10):1909–20.

41. Claes P, Walters M, Shriver MD, et al. Sexual dimorphism in multiple aspects of 3D facial symmetry and asymmetry defined by spatially dense geometric morphometrics. *J. Anat.* 2012;221(2):97–114.

42. Claes P, Reijniers J, Shriver MD, et al. An investigation of matching symmetry in the human pinnae with possible implications for 3D ear recognition and sound localization. *J. Anat.* 2015;226(1):60–72.

43. Gunz P, Mitteroecker P. Semilandmarks: a method for quantifying curves and surfaces. *Hystrix Ital. J. Mammal.* 2013;24(1):103–9.

44. Pomidor BJ, Makedonska J, Slice DE. A Landmark-Free Method for Three-Dimensional Shape Analysis. *PLoS ONE* 2016;11(3).

45. Adams DC, Otárola-Castillo E. geomorph: an r package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* 2013;4(4):393–9.

46. Schlager S. Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In: *Statistical shape and deformation analysis* Elsevier; 2017:217–256.

47. Maal TJJ, Verhamme LM, Loon B van, et al. Variation of the face in rest using 3D stereophotogrammetry. *Int. J. Oral Maxillofac. Surg.* 2011;40(11):1252–7.

Figure captions:

Figure 1: (A) Frontal-oblique (patient side) and (B) rear view (operator side) of the custom-built StereoFace 3D system: (1) camera mounting frame, (2) support frame, (3) Bosch-Rexroth lift modules, (4) LED panels, (5) random pattern lasers, (6) positioning lasers.

Figure 2: CAD renderings of the StereoFace 3D system (oriented similarly to the views in Fig 1) showing the camera mounting frame as well as the wheel mounted, height-adjustable support structure: (1) camera mounting frame, (2) support frame, (3) Bosch-Rexroth lift modules, (4) LED panels, (5) random pattern lasers, (6) positioning lasers. The blue arrows indicate the slight 'inset' applied to the four central cameras.

Figure 3: Mannequin head showing the two eye-safe positioning lasers projecting one cross each on the patient's face. The optimal working distance is obtained when both crosses coincide on the patients' midline, just below the nose.

Figure 4: Example of the 14 camera views, as acquired from the mannequin head. The red dots are generated by the random pattern lasers, which provide additional texture to the relatively feature-sparse human facial surface.

Figure 5: Schematic representation of the acquisitions performed by each system. Each subject was scanned three times, after which the images were registered using MeshMonk. The mapped results were then superimposed using a generalized Procrustes superimposition without scaling (GPA). The resulting average shape was considered to be the ground truth facial shape per subject.

Figure 6: The top pane depicts the original, reconstructed meshes of the mannequin head, as acquired by the 3dMDFace system (left) and the SF3D system (right). The bottom pane displays the corresponding registered results using MeshMonk.

Figure 7: Superimposition area (in red) overlaid on the mapped mannequin head. Only the region colored red was considered in the current investigation.

Figure 8: Intra-system variation expressed in RMSE (mm). The unfilled boxplots represent the 3dMDFace system, while the compact boxplots represent the SF3D setup. Left: RMSE of the study group with a mean RMSE of 0.41 (SD 0.15) mm for SF3D and a mean of 0.51 (SD 0.07) mm for 3dMDFace. Right: RMSE values of the mannequin head, with a mean of 0.29 (SD 0.03) mm for SF3D and 0.42 (SD 0.01) mm for 3dMDFace. The difference between means of the study group and the mannequin can be attributed to biological error (i.e. facial pose instability) and/or registration error.

Figure 9: Inter-system variation expressed in RMSE, with a mean of 1.63 +/-0.34 mm for the study group (left) and 1.38 +/- 0.25 mm for the mannequin head (right).

Figure 10, A: Heatmap depicting the normal distances between the average facial shapes of the study group (A) as generated by 3dMDFace and SF3D showing the inter-system variation. Yellow colors indicate 3dMDFace positioned the mesh outwards compared to SF3F, while blue suggests 3dMDFace's mesh is positioned inward relative to that of SF3D.

Figure 10, B: Corresponding heatmap depicting the normal distances between the average facial shapes of the mannequin head (B) as generated by 3dMDFace and SF3D (inter-system variation).

Figure 11: Two-factor ANOVA inter-system shape variation following an isotropic model. P-values using 100 permutations: white $P < 0.001$; gray $P < 0.05$; black $P \geq 0.05$ ns. MS (mean square) is the sum of squares divided by the appropriate degrees of freedom, reflecting the magnitude value. F (F-ratio) is the MS divided by an appropriate error MS, reflecting the relative magnitude or strength of the effect. The interaction term is used as error term for

the main effects of individuals and sides, while the actual error term is used for the interaction term.

Figure 1

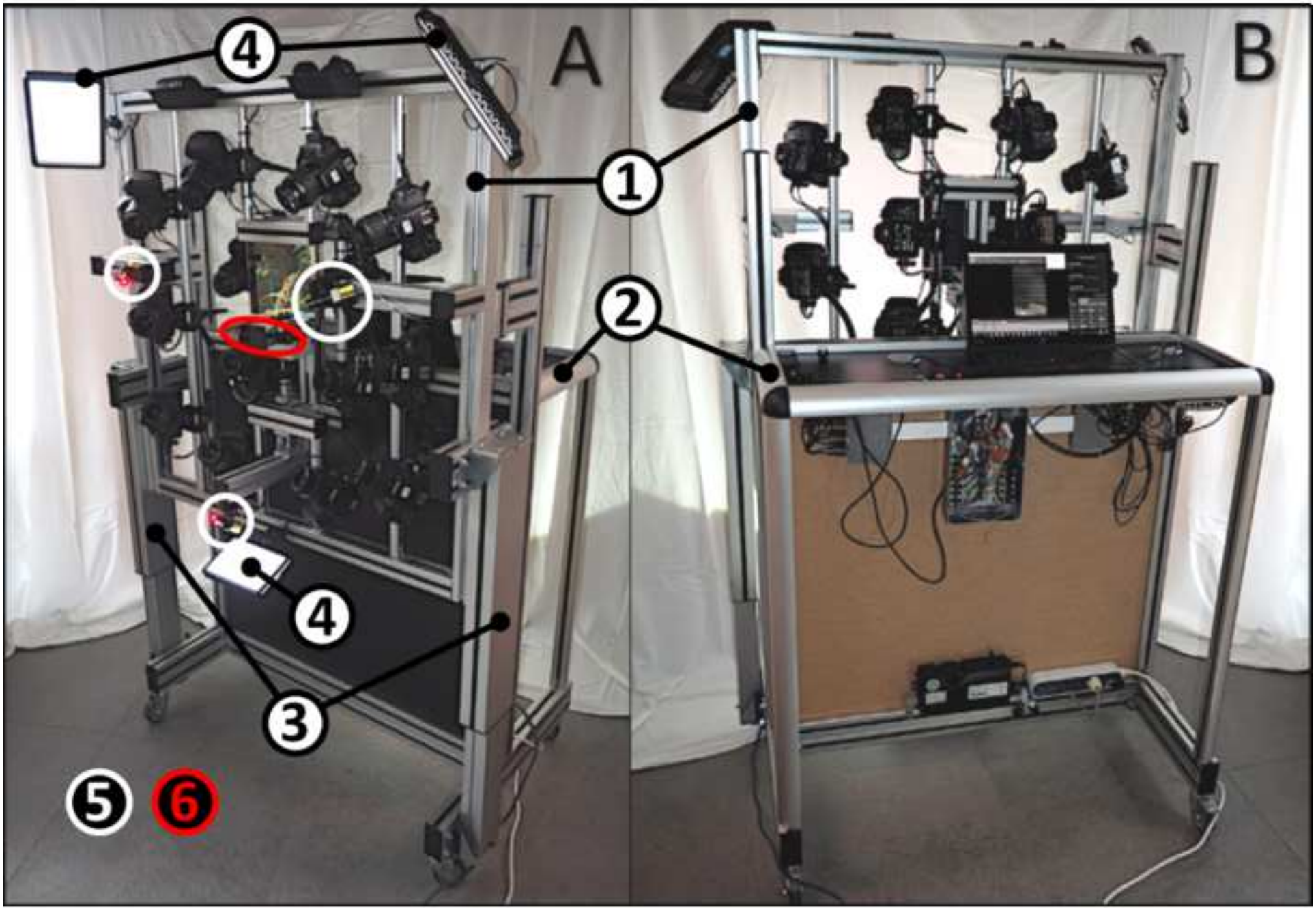


Figure 2

Click here to access/download;Figure;Fig2_Setup_oblique_front_rear_revised.tif

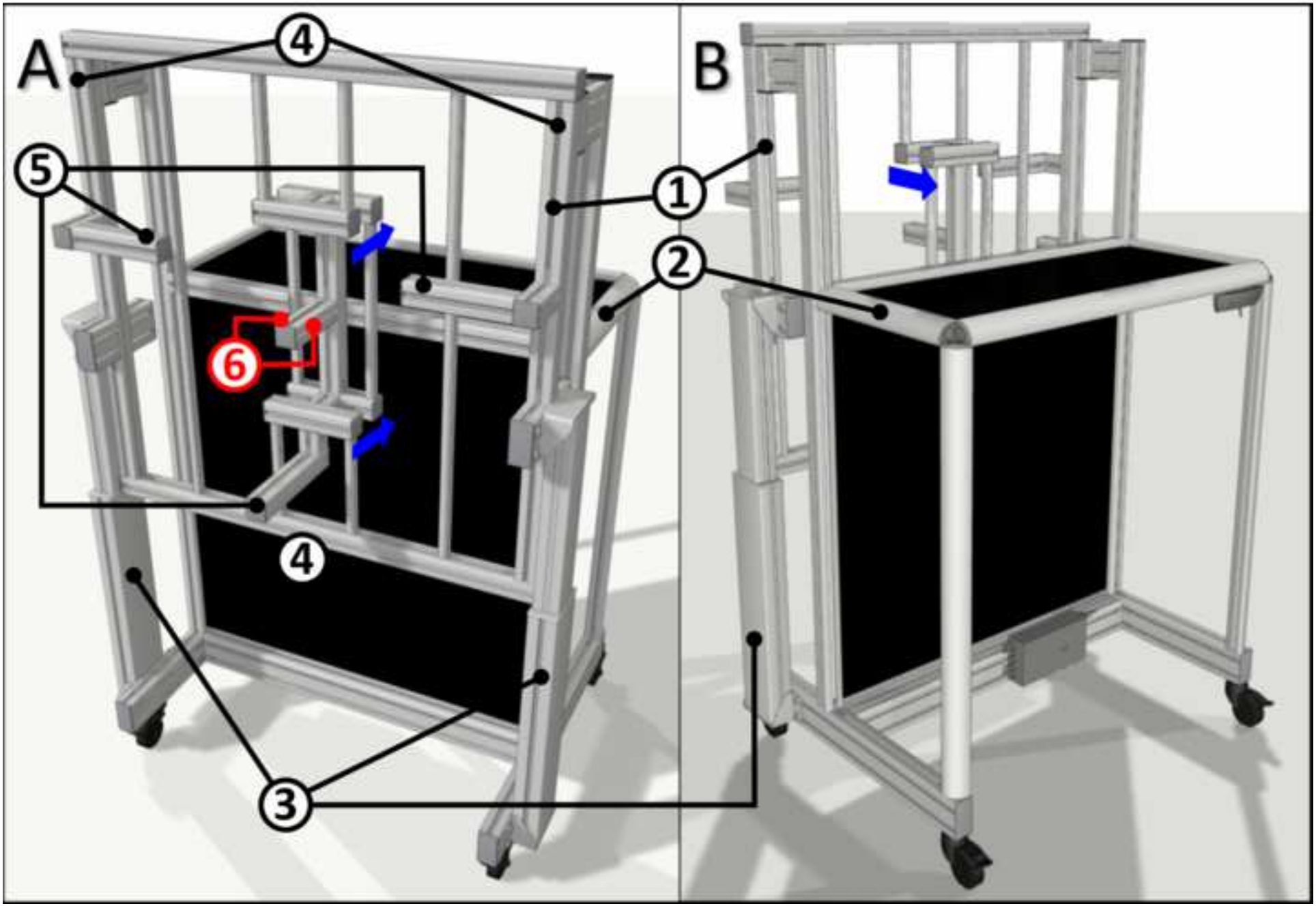


Figure 3

Click here to access/download;Figure;Fig3_Positioning_lasers.png

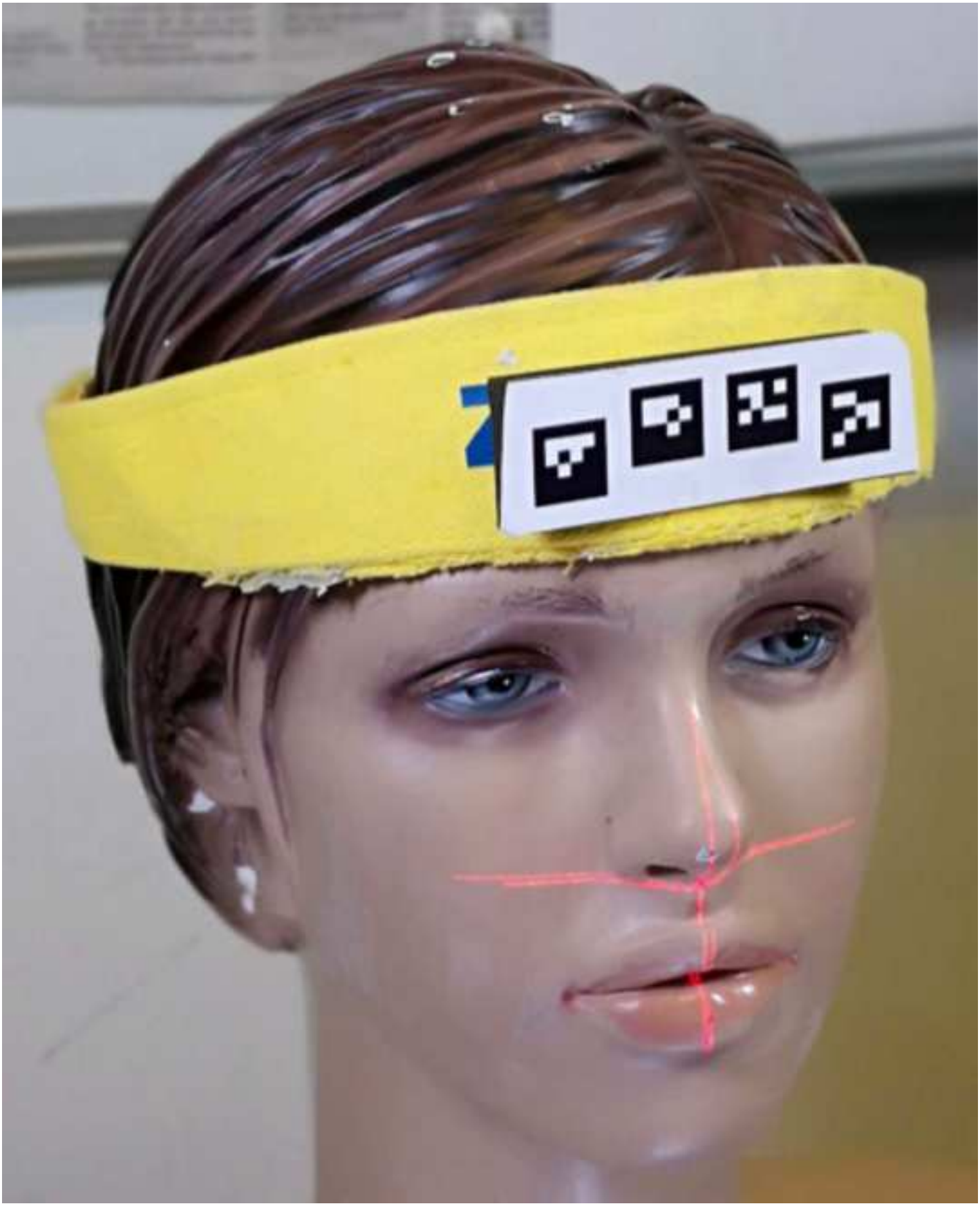


Figure 4

Click here to access/download;Figure;Fig4_Angie_14.tif

Figure 5

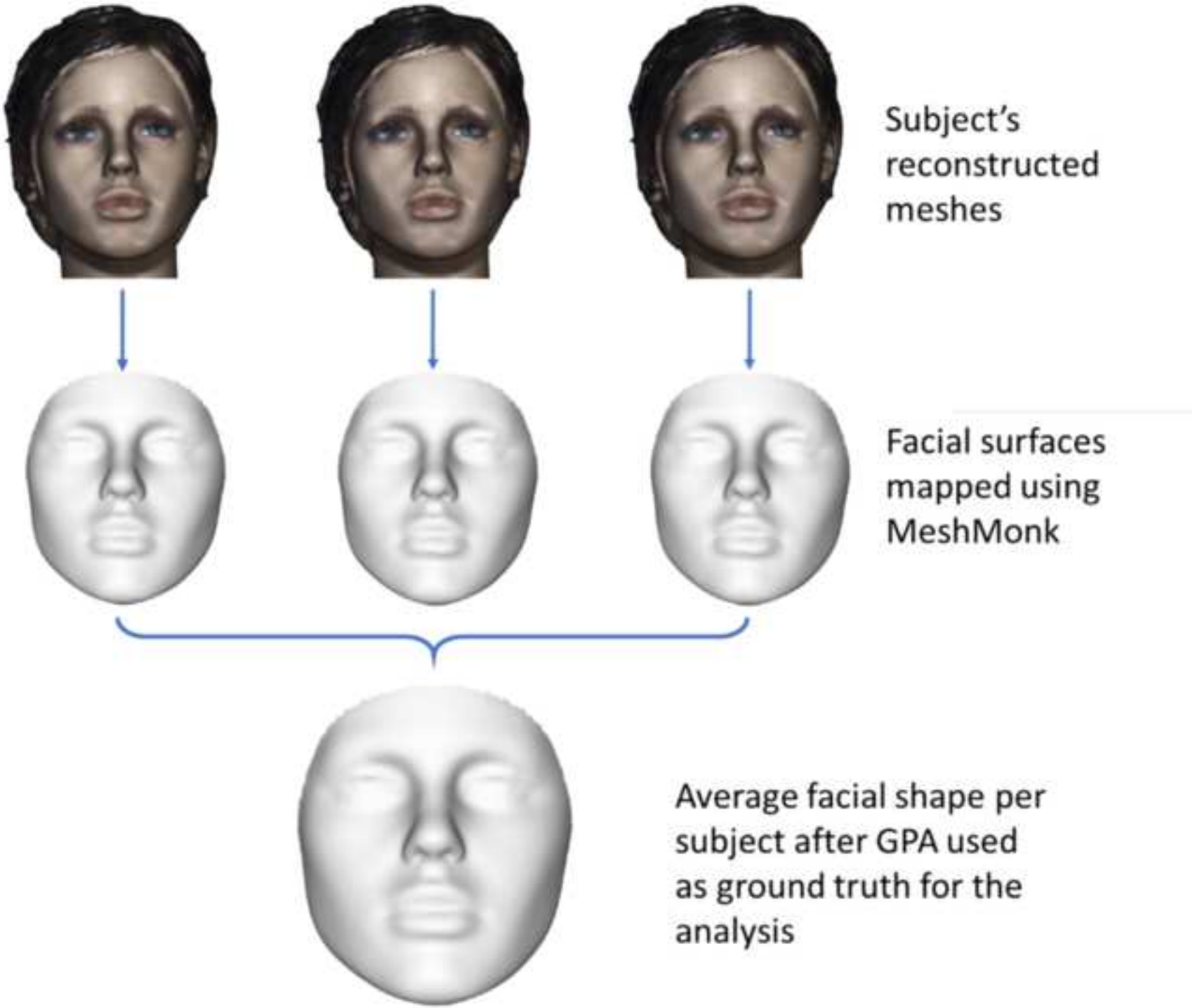


Figure 6

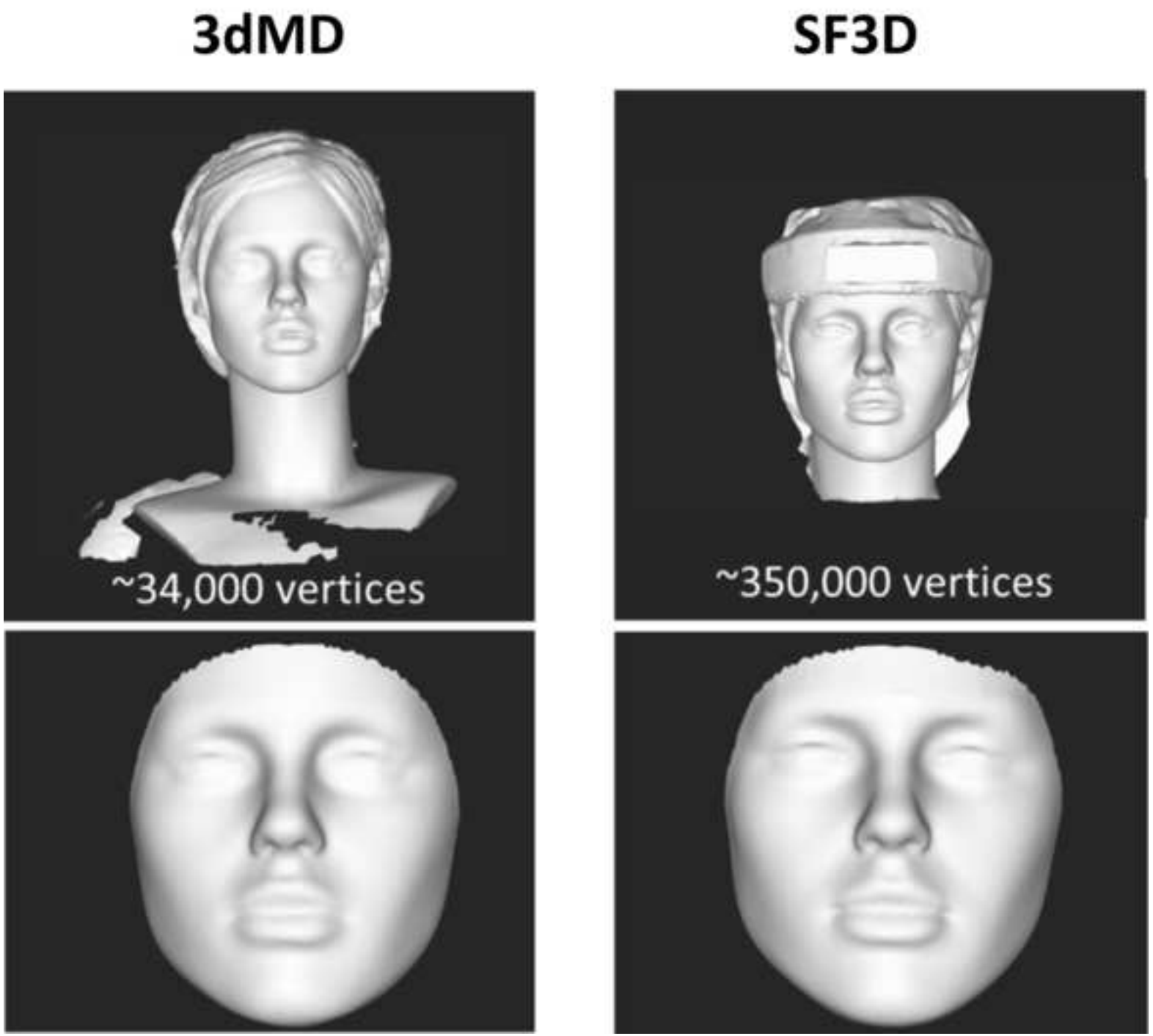


Figure 7

Click here to access/download;Figure;Fig7_superimposition_area.tif

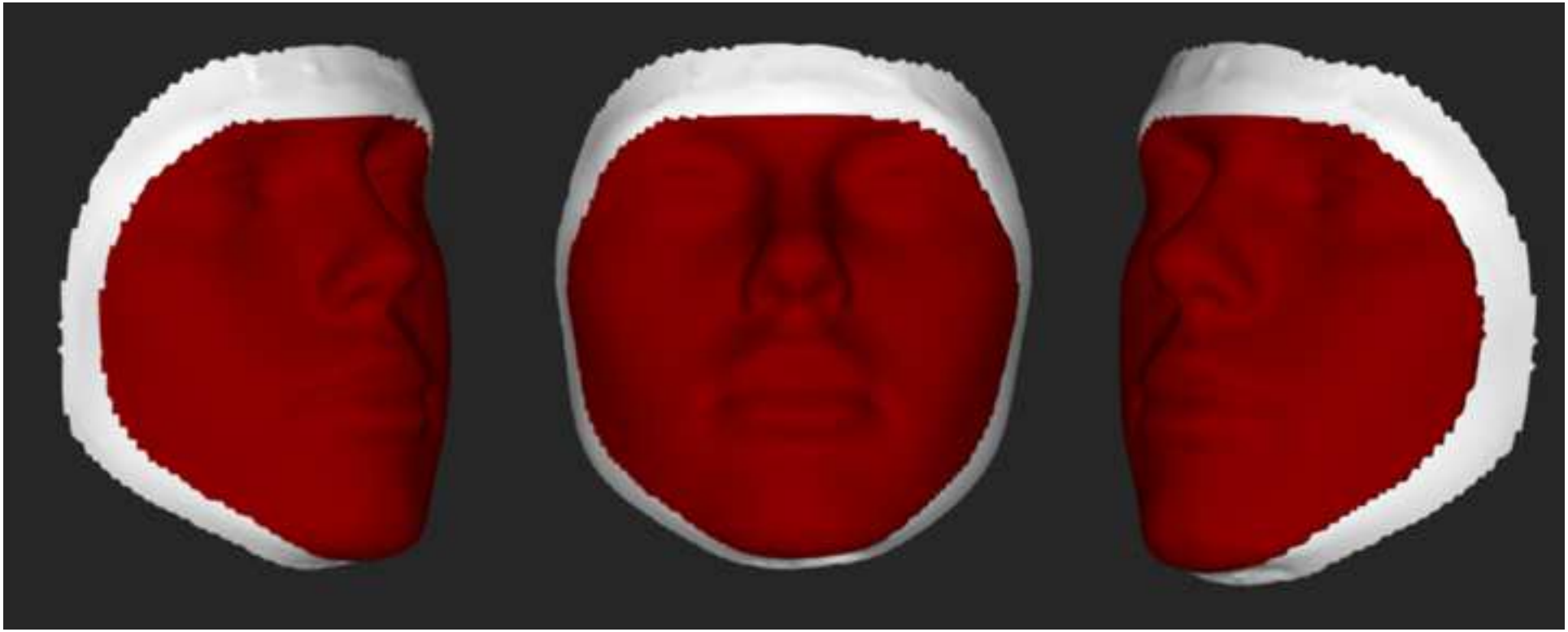
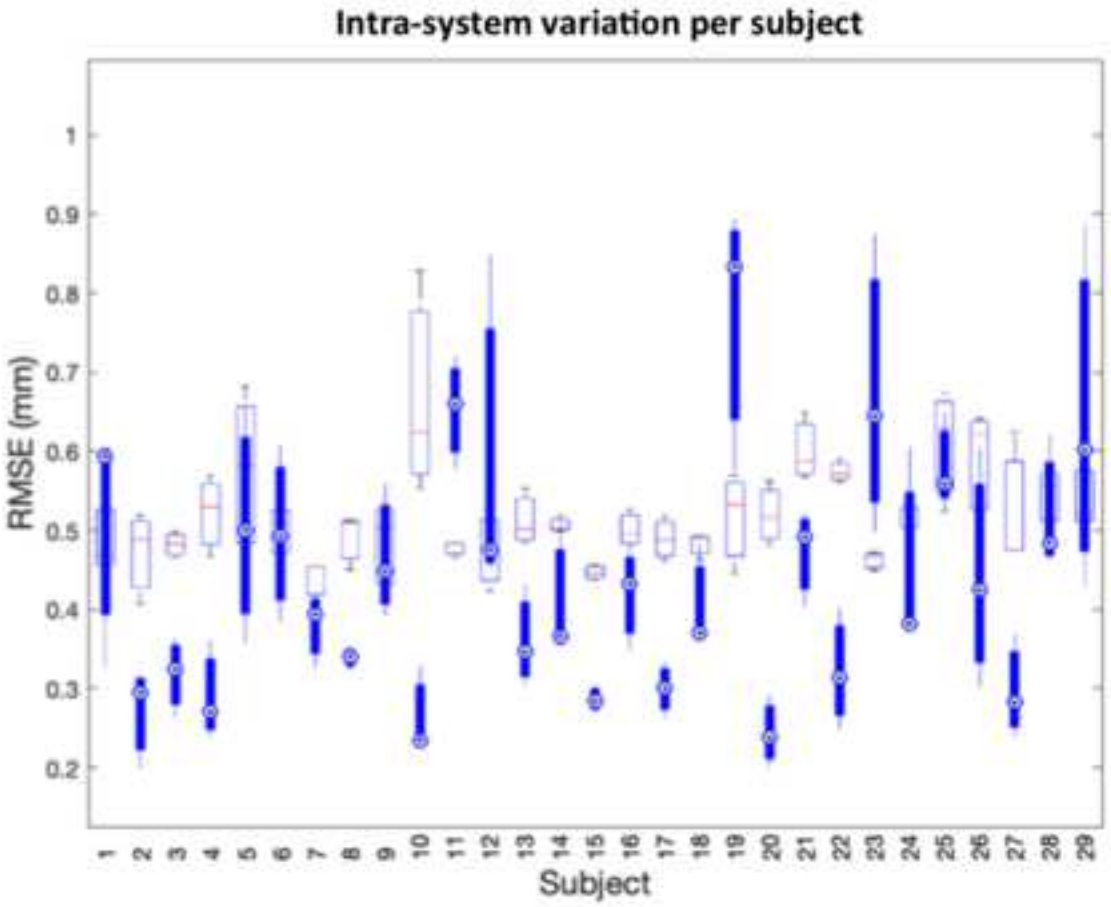


Figure 8



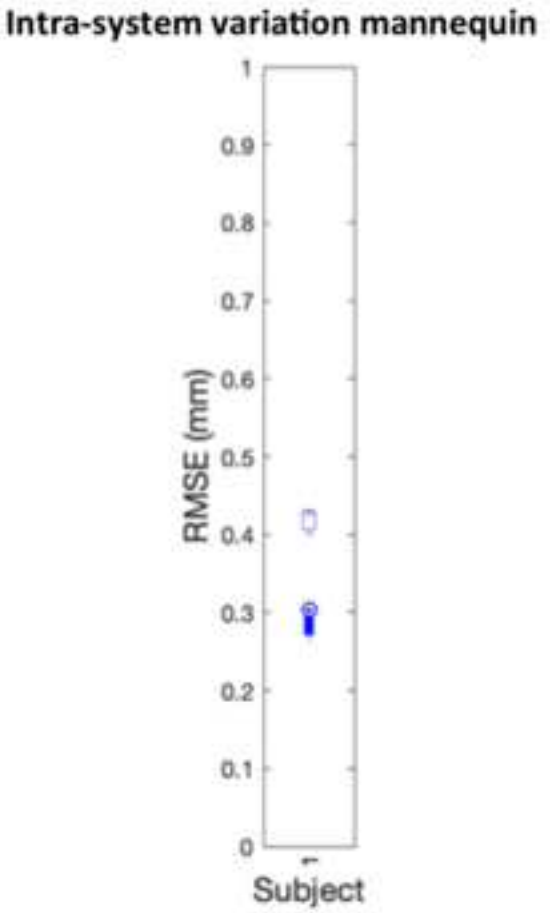
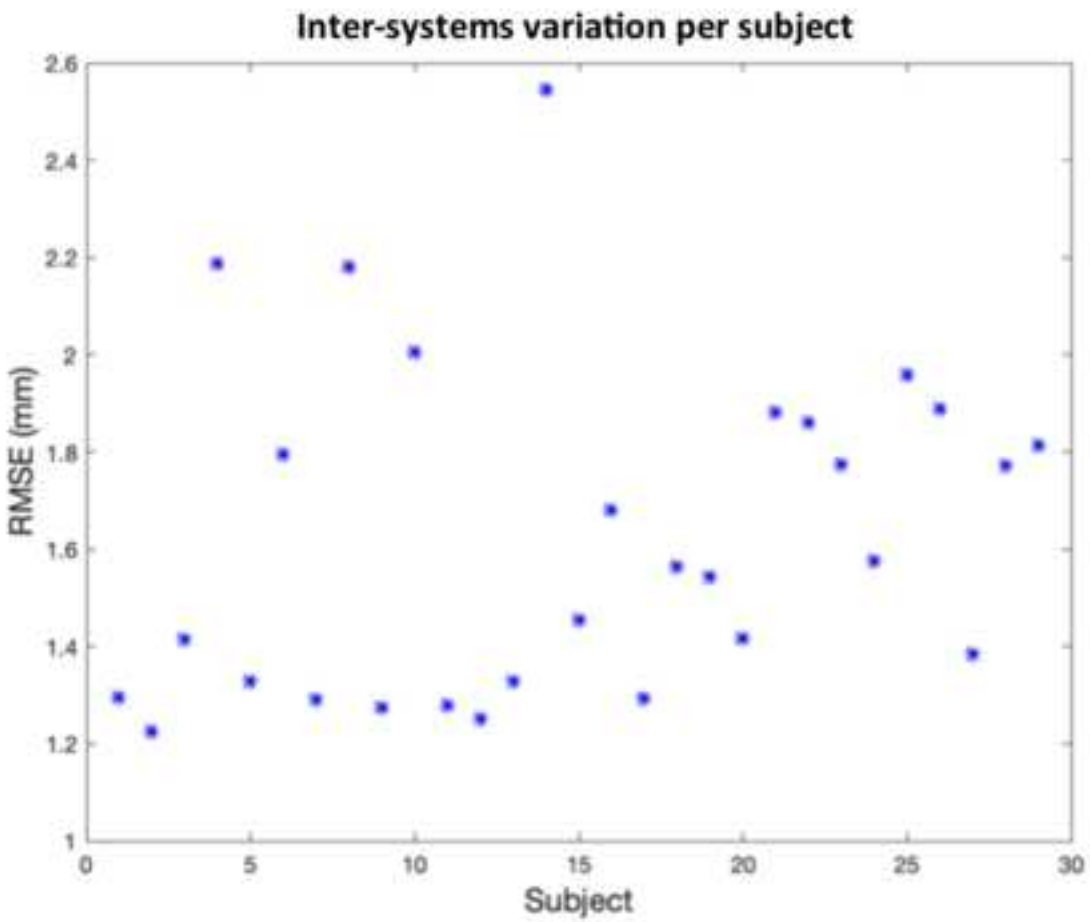


Figure 9

Click here to access/download;Figure;Fig9_Inter_camera_variation.tif



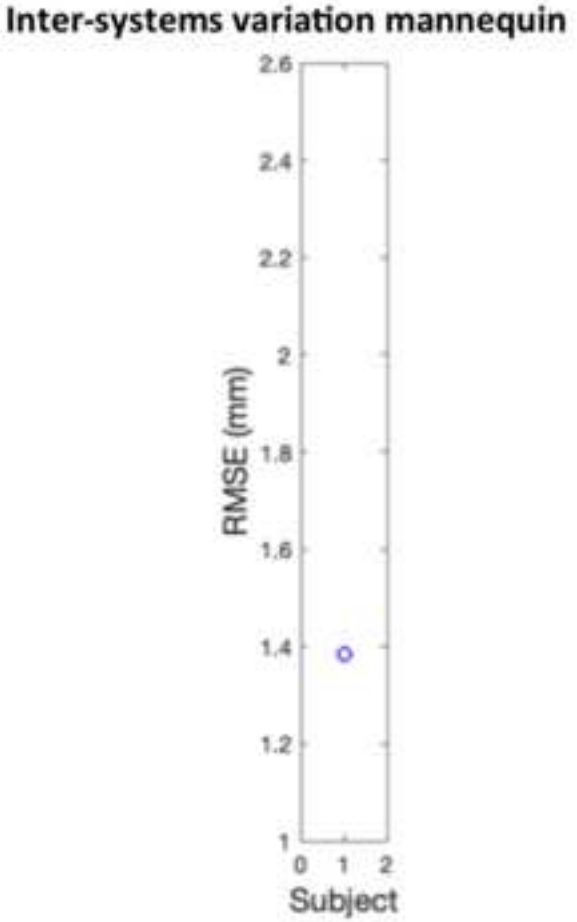


Figure 10

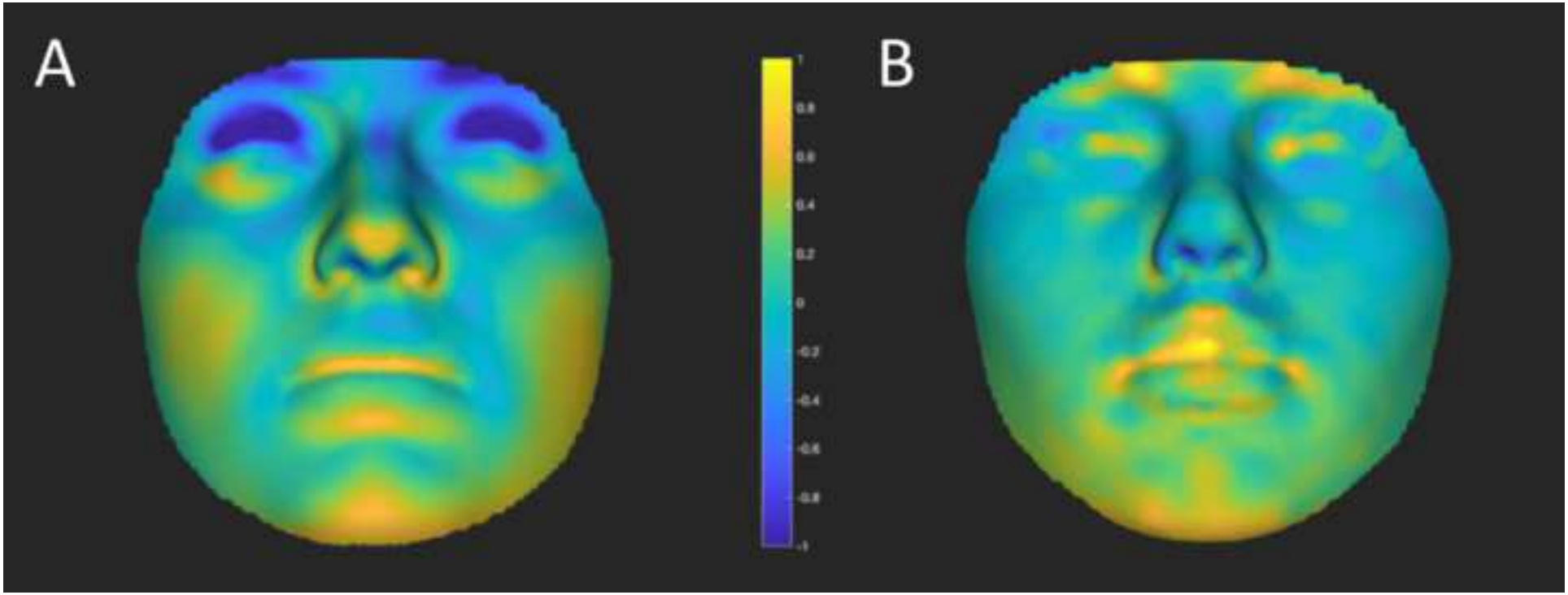
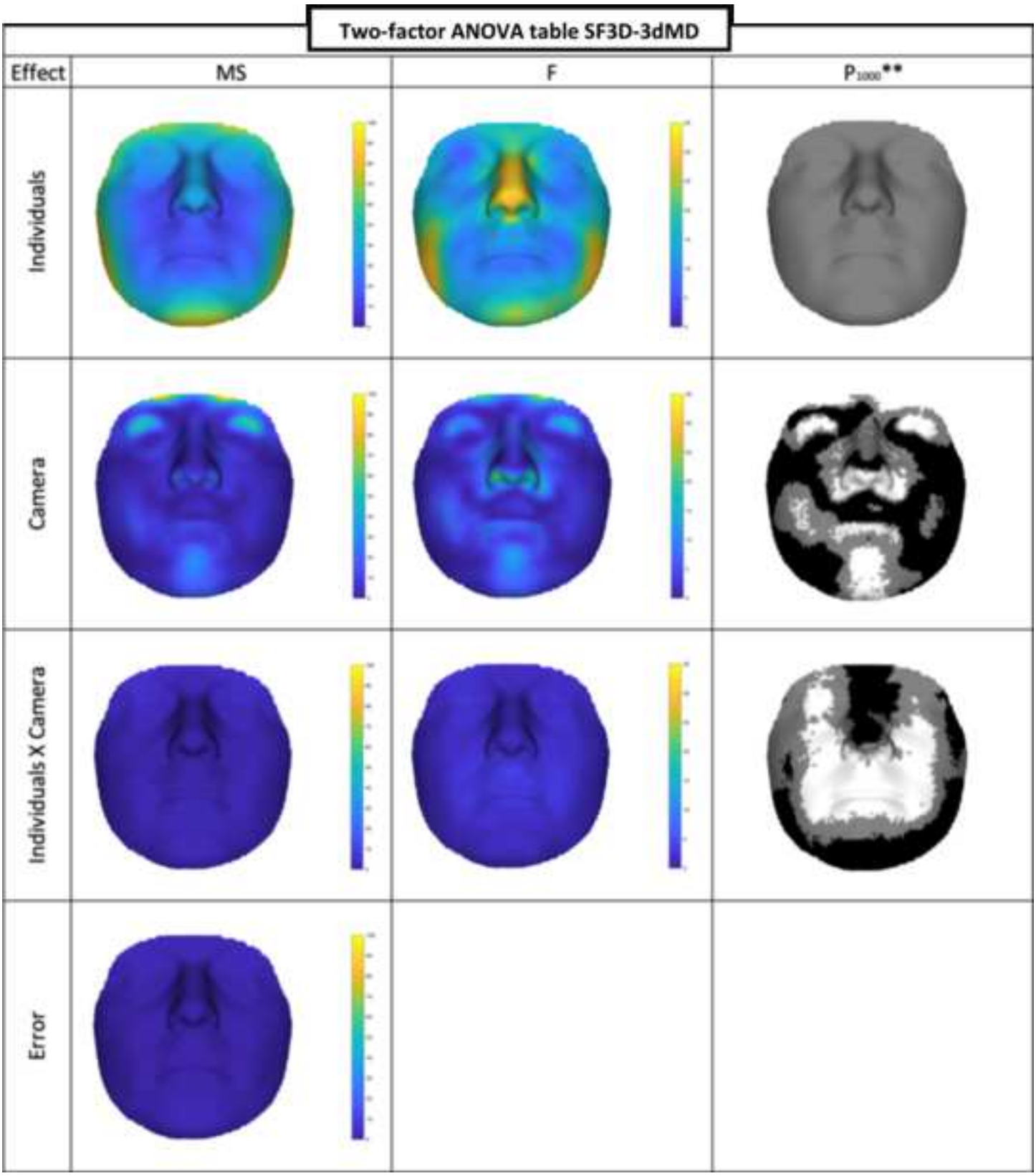


Figure 11

Click here to access/download;Figure;Fig11_Two_factor_anova.tif



Table

Table I: Age and sex distribution of the study group

		Age(years)			
	N	mean	SD	min	max
Total	29	28.3	3.8	21	38
Males	13	30	3.9	24	38
Females	16	27	3.3	21	34

Table II: Intra- and Inter-systems shape variability, expressed as RMSE (mm)

Intra-systems variability:

	Study group				**Mannequin head**		
RMSE (mm)	mean	SD	min	max	mean	SD	min
3dMD	0.517	0.068	0.407	0.828	0.419	0.016	0.400
SF3D	0.455	0.166	0.198	0.895	0.295	0.029	0.262
Diff.	0.062	-0.098			0.124	-0.013	

Inter-systems shape variability:

RMSE (mm)	mean	SD	min	max
Study group	1.630	0.342	1.225	2.546
Mannequin head	1.384			
Diff.	0.246			

max
0.432
0.317

Click here to access/download

Multimedia Files

StereoFace 3D.mp4

American Journal of Orthodontics and Dentofacial Orthopedics

COPYRIGHT STATEMENT

Must be signed by ALL authors

"The undersigned author(s) transfers all copyright ownership of the manuscript [title of article] to the American Association of Orthodontists in the event the work is published. The undersigned author(s) warrants that the article is original, does not infringe upon any copyright or other proprietary right of any third party, is not under consideration by another journal, has not been published previously, and includes any product that may derive from the published journal, whether print or electronic media. The author(s) confirm that they have reviewed and approved the final version of the manuscript. I (we) sign for and accept responsibility for releasing this material.

The corresponding author must be named HANS WELCKER
(Type name) Signature

Each author's name must be typed underneath the signature.

Alejandra Ortega Castrillon. Alejandra Ortega C.

PETER CLAES

Hanne Hoskens. Hanne Hoskens

Date: