# Assessing the effects of time on cadaveric facial anatomy using conventional photogrammetry, stereophotogrammetry and computed tomography

Morgan Titmus[1], Hannah Radley[1], Paul Ellery[1], Gary Whittaker[1], Zhonghua Sun[1], Petra Helmholz[2]

[1]Curtin Medical School, Curtin University, Bentley, Western Australia GPO Box U1987, Perth WA 6845, Australia.
Corresponding author email: Morgan.Titmus@curtin.edu.au
[2]School of Earth and Planetary Sciences, Curtin University, Bentley, Western Australia

**Keywords:** anatomy, morphometric accuracy, model registration

**Abstract**
Body donation remains critically important for anatomical science, allowing examination of biological structures with three-dimensional (3D) context. However, body donors (cadavers) are a time-limited resource and the scarcity of body donors has prompted an interest in digital body preservation. Multiple imaging techniques (e.g., conventional photogrammetry [CPG], stereophotogrammetry [SPG] and computed tomography [3DCT]) can capture the 3D characteristics of a specimen indefinitely. Digital anatomical records provide an opportunity to measure anatomical structures in the absence of the physical specimen. In 2022, the face of a preserved body donor was digitally reconstructed using CPG and 3DCT. 28 months later, a repeat survey was performed using SPG and a series of facial landmarks were directly measured. The accuracy and stability of facial soft-tissues over time were measured using Euclidean landmark distances and cloud-to-mesh techniques. The results show that anatomical models produced by 3DCT and CPG produce similar facial measurements to those acquired by SPG and direct measurement at later timepoints. These data indicate that chemical fixation adequately stabilises facial anatomy over time, each sensor can be used interchangeably for facial measurement and models can be co-registered with minimal discrepancy.

## 1. Introduction

The human body has complex three-dimensional (3D) spatial relationships, explored through the scientific discipline of anatomy. Many institutions rely on body donors (cadavers) for teaching and research purposes. Human specimens are prepared for teaching using prosection, which is a specialised dissection conducted for anatomical demonstration. The use of body donors for learning is strongly supported by staff and students (Davis et al., 2014) but is not without challenges. Chemical fixatives, used to preserve biological specimens, can be harmful to human health (Viegas et al., 2010). Ventilation is needed to extract these chemicals, meaning student learning can only take place in specialised laboratories. Additionally, preserved tissues can dehydrate over time, impairing their appearance and usefulness. In Western Australia (WA), body donors are returned for cremation after a maximum of five years, requiring an ongoing need for provision and prosection of body donors.

To overcome the limitations of physical specimens, anatomists use various imaging sensors (e.g., photogrammetry [PG], computed tomography [CT], magnetic resonance imaging and structured light scanning) to ‘digitally preserve’ body donors (Vandenbossche et al., 2022). 3D visualisations can then be used to complement or replace physical specimens, providing flexible learning opportunities (Erolin, 2019; Titmus et al., 2025). The choice of imaging sensor depends on availability, cost, expertise and speed of acquisition; each has advantages and disadvantages. For example, CT data is quickly acquired, cross-sectional and can be 3D rendered (3DCT) to reconstruct internal or surface structures. However, the use of ionising radiation requires specialised facilities with on-site CT radiographers. Photogrammetry has a slower acquisition time but demonstrates photorealistic surface textures and can be performed in a standard anatomical laboratory.

The ability to use multiple sensors interchangeably for visualisation and measurement has several advantages. Features that are identifiable only on surface texture, but not through geometry (e.g., skin contusions) can be captured with PG, but not 3DCT (Bolandzadeh et al., 2013; Dixit et al., 2019). In contrast, 3DCT can capture and measure geometric features that are obscured from photogrammetric image capture, for example, internal structures and regions that are in contact with a surface; such as the skin of the back on a supine specimen (Mazonakis & Damilakis, 2016).

As 3DCT and PG data are complementary, it is ideal to acquire both and co-register the data. Data capture from the two sensors may happen at the same time or on separate occasions. Asynchronous capture may be suitable for rigid bone specimens, but soft tissues are problematic. Alterations in soft tissue can occur through tissue dehydration (loss of water leading to changes in tissue position), mechanical deformation or inconsistent specimen positioning. Kottner et al. (2017) produced a custom multi-camera rig (VirtoScan) that allowed simultaneous acquisition of CT and PG data during post-mortem examination. Their approach overcomes the problems of asynchronous capture but may be prohibitively expensive and complex for many anatomy facilities. Some institutions, including ours, do not have an on-site medical imaging facility, necessitating PG and CT data to be captured at different times and locations (Waltenberger, Rebay-Salisbury, & Mitteroecker, 2021).

Once generated, 3D visualisations are permanent anatomical records that persist after the body donor has been returned. If quantitative measures are to be derived from these data, there is a need to ascertain the accuracy of each sensor. Further, as multi-sensor imaging may be performed asynchronously, there is a need to establish whether measurements are stable over time or significantly affected by tissue dehydration or deformation.

Our laboratory developed three different visualisations of the same body, on separate occasions, and with different sensors (PG, stereophotogrammetry and 3DCT). This provided an opportunity to assess the stability of anthropometric soft tissue measurements over a 28-month period using both digital and direct methods. Three hypotheses were tested:

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-713-2026  

1. Anatomical measurements taken on PG models will not be significantly different to those obtained on 3DCT or by direct examination at any timepoint.
2. Measurements will be consistent both within and between examiners for models produced by the same sensor.
3. PG and 3DCT models obtained at different timepoints can be co-registered with minimal spatial discrepancy.

The paper is structured as follows: related work regarding the accuracy of different sensors in anatomy is provided in Section 2. The dataset used is introduced in Section 3, followed by the method in Section 4. Finally, the results are presented in Section 5 and the conclusion in Section 6.

## 2. Related Work

Numerous authors have used bone specimens to quantify the accuracy of different sensors. 3DCT bone measurements have been shown to be highly accurate, regardless of whether the bones are in situ or isolated (Stull et al., 2014). The authors found that the majority of 3DCT osteometric measurements were within a +/- 2mm range of the original specimen, which was considered acceptable for forensic applications. Waltenberger et al. (2021) compared the accuracy of 3DCT and photogrammetry using human pelves. They found the mean surface deviation between the two meshes was 0.17mm. Sub-millimetre accuracy has also been found when comparing bone measurements made physically and by photogrammetry (Feddema & Chiu, 2024). Measurement of human crania showed that photogrammetric measurements were slightly larger (mean difference = 1.28mm) relative to CT measurements, indicating these sensors could be used interchangeably (Donato et al., 2020). Overall, these findings suggest that bone measurements made with photogrammetry and CT should be highly correlated.

However, the majority of human anatomy is comprised of non-skeletal soft tissues (e.g., skin, muscle, adipose tissue, viscera). As soft tissues are less rigid than bone, measurements made on soft tissues may be less precise. Stereophotogrammetry (SPG) is an established technique to measure anatomical landmarks on facial soft tissues (Akan et al., 2021). Palmer, Helmholz, and Baynam (2020) compared the accuracy of facial SPG measurements performed by using *Cliniface* software to those obtained via direct anthropometry by an expert. Automated measurements in *Cliniface* were not significantly different to the expert’s and yielded an overall root mean square error of 4.87mm. This suggests that *Cliniface* can be used to establish a reference standard of measurement which is comparable to direct manual measurement. In this study, the authors used living participants and did not compare their findings to other (non-SPG) sensors. It is unknown if these results are generalisable to deceased body donors or measurements obtained by CPG or 3DCT.

If anatomical geometry can be shown to be accurate and stable across time, location and different sensors it would have important implications. The first is that quantitative anatomical measurements could be derived from digital records, instead of directly from the physical specimen. The second is that 3D visualisations from different timepoints and sensors could be co-registered to produce accurate, high-value educational materials.

## 3. Dataset

A body donor specimen was embalmed using chemical fixatives (formaldehyde and phenol) and stored in a 1% formalin solution. The body was donated through the University of Western Australia’s Body Bequest program and stored at Curtin University’s anatomy facility. The donor gave their consent for their remains to be used for teaching and research purposes. To respect the donor’s anonymity, no identifying facial imagery is presented in the paper. This study was approved by the Curtin University Human Research Ethics Committee (HRE2023-0260).

Three 3D models of the body were produced using different hardware / software configurations; a stereophotogrammetric model (SPG), a conventional photogrammetry model (CPG) and a 3D-rendered computed tomography model (3DCT). A reconstruction summary of all models, including materials used for each, is presented in Table 1. SPG and direct measurements (i.e., manual measurements on the cadaver) were performed in January, 2025 whereas CPG and CT scanning were performed in August, 2022.

The SPG model was produced using the Vectra H2 3D imaging system (Canfield Scientific, NJ, USA) with a stated 3D resolution of 1.2mm. The convergence of dual ranging lights was used during capture to confirm optimal object-lens distances. Three stereo images (total of six image perspectives) were processed with Canfield Vectra 3D software (Version 6.8.1; Canfield Scientific, NJ, USA) to produce a single merged model using a Dell computer (XPS 15 7590, CPU: Intel Core i7-9750H @ 2.6GHZ, RAM: 32 GB running Windows 11 Pro) and exported as object files (.obj) alongside texture and material files (.jpg / .mtl). No post-processing was applied.

The conventional photogrammetry model was produced using a standard structure-from-motion approach for CPG, adapted from the workflows described by Gitto et al., (2020) and Titmus et al., (2023). The body donor was positioned on a custom survey platform containing 10 ground control points (GCPs) that were located along the periphery of the body (head, upper limbs and lower limbs). The location of the GCPs were pre-calibrated using a Nikon D750 SLR with a Nikkor 24-70mm f/2.8 ED G AF-S lens and a scale bar. 3D reconstruction of the full-body specimen used Metashape Professional (Version 2.1.2; Agisoft LLC.) on a Dell Precision 7920 workstation (CPU – 2x Xeon Gold 6248R @ 3.00GHz (48 cores), GPU – 3x NVIDIA RTX 8000, RAM – 512 GB). Scaling was achieved by utilising the GCP’s.

CT data was acquired at the Western Australian State Mortuary using a Siemens Somatom Definition AS scanner with a slice thickness of 0.6mm. A full-body 3DCT model was produced using 3D Slicer (version 5.6.2) on a MacBook Pro (CPU: Apple M3 Pro, RAM:18GB) running macOS Sonoma (version 14.7.3). The skin was segmented from the CT image stack using Hounsfield unit thresholding, achievable as the skin was hyperdense (higher Hounsfield units) relative to the surrounding hypodense air (lower Hounsfield units). The model was decimated to 500K faces using Blender (version 4.0; Blender Institute, Amsterdam, The Netherlands) and had a coloured material and texture applied. As the CT model was cross-sectional, it had both an internal and external skin surface. The external surface was extracted using CloudCompare and used for subsequent stages

.

| Model | Imaging hardware | Lighting | Imaging parameters | Image acquisition time | Faces (following decimation and segmentation) | Surface texture |
|---|---|---|---|---|---|---|
| SPG | Vectra H2 3D imaging system | On-board flash unit | Parameters: Vectra software internal preset parameters<br>Total image number: 3X stereo images (total 6 perspectives) | ~3 minutes | 97 724 | Yes - photorealistic |
| CPG | Sony a7r mark 4 camera (61Mpix sensor, 35mm full-frame equivalent) and 50mm prime lens (Sony Sonnar T* E-Mount FE 55mm f1.8, Zeiss) | 8X studio lights (Godox SL-60W) | ISO:100-200<br>Aperture: f16-18<br>Shutter speed: automated<br>File format: RAW<br>Total image number (full-body):738 | ~6 hours | 12,403 | Yes - photorealistic |
| 3DCT | Siemens Somatom Definition AS CT scanner | Nil | Slice thickness: 0.6mm<br>Field of view: 780mm<br>kVp:120<br>mAs: 265<br>Total image number (full-body):2980 | ~80 seconds | 9,662 | Uniform colour applied during post-processing |

Table 1: Reconstruction summary of the digital 3D models used for measurement comparison. SPG – Stereophotogrammetry model, CPG – conventional photogrammetry, 3DCT – volume rendered three-dimensional computed tomography model. kVp – Kilovoltage peak, mAs – milliampere-seconds.

## 4. Method

### 4.1 Measurements of facial landmarks

To assess whether anatomical measurements were stable across models, facial landmarks were measured as they were discretely identifiable on the physical and digital specimens. Two examiners (total 32 years of anatomical experience) independently measured 11 distances on the 3D models and nine on the physical specimen (Table 2).

| Facial landmarks | Measure name | Bilateral? |
|---|---|---|
| Endocanthion to contralateral endocanthion | Intercanthal distance | No |
| Exocanthion to contralateral exocanthion | Outercanthal distance | No |
| Exocanthion to ipsilateral endocanthion | Palpebral fissure horizontal length | Yes |
| Chelion to contralateral chelion | Labial fissure length | No |
| Subnasale to labrale superius | Philtral length | No |
| Subalare to contralateral subalare | Subnasal width | No |
| Alar curvature point to pronasale | Nasal ala length# | Yes |
| Inferior ear lobule to superior helix | Vertical ear height | Yes |

Table 2: Facial landmarks and associated measurements used. Adapted from Palmer et al., (2020). #Nasal ala lengths were not measured directly via physical measurement.

Nasal ala lengths were not measured on the physical specimen as callipers were unable to contact the skin at both landmarks simultaneously. Prior to generating data, examiners had one practice session where a consensus for landmark placement was achieved. Measurements were repeated five times, with a typical interval of 24 hours between each measurement as per McMenamin et al. (McMenamin et al., 2014). In total, 420 measurements were obtained.

The SPG model was imported into Cliniface (Version 7.0.2205525, SIS research Ltd.) (Palmer et al., 2020) which automatically identified all relevant facial landmarks, except for ear landmarks which were manually placed. Examiners were permitted to manually alter landmarks if necessary. Distances between relevant landmarks were automatically extracted by the software. The software permits measurements to be performed via Euclidean distances or along the facial surface; only Euclidean measurements were extracted. The SPG model was designed to be a reference standard as it had a known spatial accuracy (1.2mm). Furthermore, the use of Cliniface minimised the number of landmarks that required manual placement.

CPG and 3DCT models were measured separately using the in-built measurement tool in CloudCompare (Version 2.13.2). Both examiners used the same pre-determined model orientations for landmark measurement but altered zoom/magnification to their preferred level.

Direct anthropometric measurements were performed using digital callipers (Acemeter 0-150mm digital callipers, China) with a stated resolution of 0.01mm and accuracy of 0.02mm. Between each set of measurements the body donor was submerged in 1% formalin. A summary of the digital models used for measurement is provided in Table 3.

| Method | Reconstruction software | Scale | Measurement software |
|---|---|---|---|
| SPG | Vectra H2 software (Canfield Scientific, NJ, USA) | Stereo-photogrammetry | Cliniface |
| CPG | Metashape Professional (Version 2.1.2; Agisoft LLC.) | Pre-calibrated control points | Cloud-Compare |
| 3DCT | 3DSlicer (version 5.6.2) | Inherently scaled | Cloud-Compare |

Table 3: Comparison of the techniques used to generate, scale and measure each 3D model.

### 4.2 Whole 3D model comparisons

The SPG, CPG and 3DCT models were segmented using CloudCompare so that equal segments of the face were visible in each model. The CPG and 3DCT models were aligned to the SPG model using the facial landmarks in Table 2, via a six-parameter Helmert transformation; no scaling factor was applied. This was followed by fine registration using the Iterative Closest Point (ICP) algorithm. Cloud-mesh distances (Helmholz et al., 2020), were calculated for the CPG and 3DCT models, using SPG as the reference standard. Scalar fields (heat maps) were used to visually identify sites where the SPG surface was discrepant with both the CPG and 3DCT. Cross sections were produced of the orbits and temporal regions with distances measured between each model’s surface.

### 4.3 Analysis

**4.3.1 Statistical analysis**: Statistical analysis was performed using R (Version 4.4.2) in the RStudio environment (version 2024.12.0) using the rstatix (version 0.7.2) and irr (version 0.84.1) packages. For all outcomes, data is presented as median (Mdn) +/- interquartile range (IQR), unless otherwise specified. Data normality was assessed using visual inspection of histograms and the Shapiro-Wilk test.

**4.3.2 Analysis of measurement reliability**: Intra-examiner reliability was calculated using an intraclass correlation coefficient (ICC) statistic; separate ICC’s were produced for each examiner by specimen type using all daily measurements. As intra-rater reliability was acceptable, each examiner’s daily results were averaged and used for inter-examiner reliability analysis, also producing ICC statistics. The primary outcome of interest is the aggregate accuracy of each specimen, as such, ICCs were not produced for each facial measure. When ICC was calculated for single measurements the ICC(A,1) statistic was used. For averaged measurements, the ICC(A,2) is reported using the McGraw and Wong convention (McGraw & Wong, 1996).

**4.3.3 Agreement and correlation between specimen type**: To assess whether measurements were consistent between the 3D models, Bland-Altman plots were produced. Pairwise comparisons (SPG Vs CPG, SPG Vs 3DCT, CPG Vs 3DCT) defined the limits of agreement between measurements (LOA; +/- 2 standard deviations from the mean of both methods) and calculated systematic bias (mean of the measurement differences between each method). As this was not a clinical study, no acceptable LOA threshold was defined *apriori*.

| Measure | Specimen | Median (mm) | IQR | Range (mm) |
|---|---|---|---|---|
| Intercanthal distance | 3DCT | 27.58 | 3.07 | 5.75 |
| | CPG | 30.88 | 0.50 | 1.23 |
| | Direct | 29.10 | 0.63 | 1.69 |
| | SPG | 29.74 | 1.18 | 2.41 |
| Left nasal ala length | 3DCT | 28.12 | 0.56 | 1.62 |
| | CPG | 30.63 | 0.95 | 1.84 |
| | SPG | 27.62 | 1.66 | 2.36 |
| Left palpebral fissure | 3DCT | 34.50 | 2.56 | 7.88 |
| | CPG | 30.26 | 0.77 | 1.56 |
| | Direct | 28.48 | 1.43 | 2.71 |
| | SPG | 29.56 | 0.38 | 2.01 |
| Left vertical ear height | 3DCT | 57.88 | 1.60 | 2.28 |
| | CPG | 59.43 | 0.17 | 0.55 |
| | Direct | 59.97 | 0.21 | 2.75 |
| | SPG | 59.46 | 1.27 | 2.27 |
| Labial fissure | 3DCT | 54.67 | 2.42 | 4.44 |
| | CPG | 52.59 | 1.23 | 2.44 |
| | Direct | 52.43 | 1.06 | 2.61 |
| | SPG | 52.39 | 0.77 | 1.73 |
| Outer canthal distance | 3DCT | 95.15 | 3.17 | 7.93 |
| | CPG | 92.57 | 0.65 | 1.88 |
| | Direct | 87.59 | 1.54 | 3.40 |
| | SPG | 89.21 | 0.28 | 0.91 |
| Philtral length | 3DCT | 14.08 | 0.52 | 2.21 |
| | CPG | 14.39 | 0.36 | 1.05 |
| | Direct | 14.04 | 0.24 | 1.11 |
| | SPG | 15.20 | 0.56 | 0.77 |
| Right nasal ala length | 3DCT | 30.17 | 0.60 | 2.26 |
| | CPG | 32.44 | 0.67 | 1.30 |
| | SPG | 28.27 | 0.09 | 0.45 |
| Right palpebral fissure | 3DCT | 35.52 | 2.78 | 5.86 |
| | CPG | 32.94 | 0.25 | 1.32 |
| | Direct | 31.01 | 0.64 | 2.48 |
| | SPG | 30.53 | 0.61 | 2.30 |
| Right vertical ear height | 3DCT | 56.29 | 0.51 | 1.79 |
| | CPG | 57.08 | 0.42 | 1.07 |
| | Direct | 57.89 | 0.32 | 0.82 |
| | SPG | 56.86 | 0.64 | 2.02 |
| Subnasal width | 3DCT | 22.48 | 0.45 | 1.39 |
| | CPG | 22.21 | 0.09 | 0.52 |
| | Direct | 19.64 | 0.35 | 1.13 |
| | SPG | 21.35 | 0.17 | 0.70 |

Table 4: Descriptive statistics for each inter-landmark measure, separated by specimen type. Pooled data for both examiners are presented.

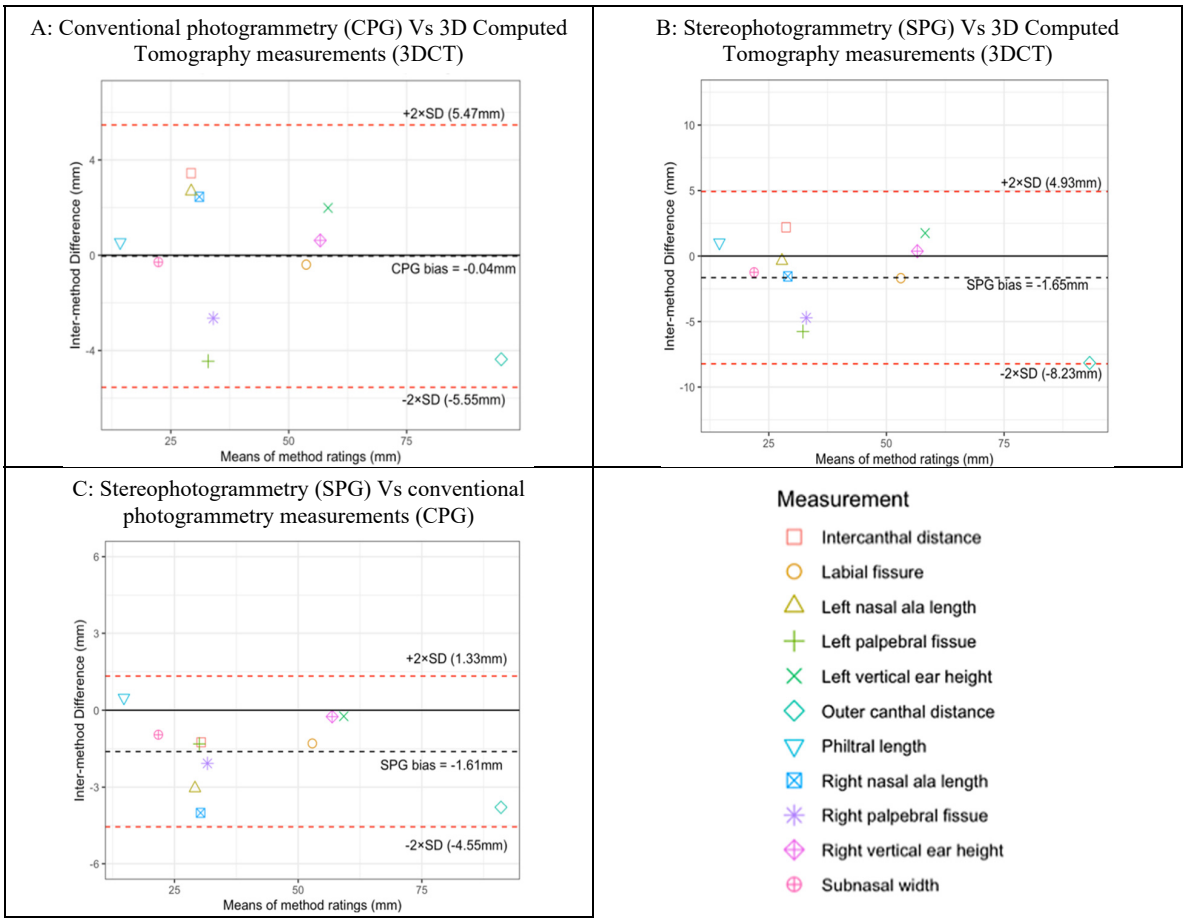


Figure 1: Bland-Altman plots demonstrating inter-method agreement. Y-axes show inter-method measurement differences of landmark distances (Panel A: CPG mean – 3DCT mean; Panel B: SPG mean – 3DCT mean; Panel C: SPG mean - CPG mean).

## 5. Results

### 5.1 Measurement differences between specimen types

Descriptive statistics for each measure name, separated by specimen type, are presented in Table 4. Data dispersion was lowest for the SPG right nasal ala measurements (IQR= 0.09, range = 0.45mm) and highest for the 3DCT outer canthal distance (IQR = 3.17, range = 7.93mm).

A Kruskal-Wallis test compared the residuals of the direct, 3DCT and CPG measurements (3DCT Mdn = 1.75mm±3.47, CPG Mdn = 1.25mm±0.83, Direct Mdn = 0.82mm ±0.9; Figure 2). No significant difference was found between any pair of groups ($\chi2(2) = 4.36$, $p = 0.113$). To assess for differences between the reference measurements (SPG) and remaining measurements (3DCT, CPG, Direct) a Kruskal-Wallis test was used. No significant difference was found between any specimen type (3DCT Mdn = 28.8mm±35.3, CPG Mdn = 26.3mm±32.7, Direct Mdn = 31.4mm ±29.6, SPG Mdn = 27.4mm±30.6; $\chi^2(3) = 0.179$, $p = 0.981$). A trend toward systematically larger measurements was observed for 3DCT measurements, this was further evaluated through whole-model comparisons in section 5.4.

### 5.2 Reliability of measurement within, and between, examiners

Intra-rater and inter-rater reliability was high across all specimen types (Table 5). 3DCT measurements exhibited the lowest inter-rater reliability out of all specimen types (ICCA,2 = 0.996); all other methods recorded excellent agreement.

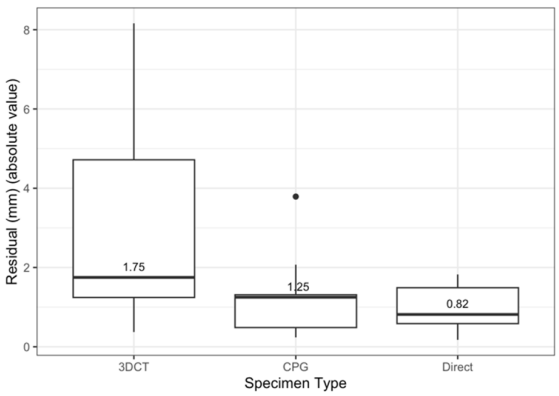


Figure 2: Residual comparison by specimen type. Absolute residuals generated by subtracting SPG measurements from those obtained by 3DCT, CPG and Direct (physical) measurements.

### 5.3 Measurement agreement and correlation between specimen types

Bland-Altman plots were produced to assess inter-method agreement between CPG and 3DCT measurements (Figure 1A), SPG and 3DCT measurements (Figure 1B) and SPG and CPG measurements (Figure 1C). All measures fell within two standard deviations of the mean of both methods, however, the comparison between SPG and CPG had the narrowest limits of agreement (SD = 1.47). Negative biases were observed for the SPG (reference) measurements when compared to the CPG (SPG bias = -1.61mm) and 3DCT (SPG bias = -1.65mm), indicating that CPG and 3DCT measurements produced systematically larger landmark distances. Minimal bias was noted between 3DCT and CPG measures (CPG bias = -0.04mm).

| **Intra**-rater reliability | **Vectra H2 (SPG)** | **Sony a7r iv (CPG)** | **CT imaging (3DCT)** | **Physical specimen (Direct)** |
|---|---|---|---|---|
| Examiner 1 [99% confidence interval] | 0.999 [0.998, 1] | 1 [0.999, 1] | 0.999 [0.996 ,1] | 0.999 [0.997 ,1] |
| Examiner 2 [99% confidence interval] | 1 [0.999 ,1] | 1 [0.999, 1] | 0.998 [0.993, 1] | 0.999 [0.998, 1] |
| **Overall** [99% confidence interval] | 1 [0.998, 1] | 1 [0.998, 1] | 0.996 [0.974, 0.999] | 1 [0.94, 1] |

Table 5: intra-rater and overall inter-rater reliability for each measurement method; ICC statistics presented.

Examiners’ average measurements (i.e., the mean of both examiners’ averaged daily landmark measures) were used to perform Spearman’s rank correlations between each specimen type (Table 6). Statistically significant correlations were found between measurements on all specimen types with strong positive associations.

| Correlation with: | **SPG (Vectra H2)** | **CPG (Sony a7r iv)** | **3DCT (CT imaging)** | **Direct (physical specimen)** |
|---|---|---|---|---|
| **SPG** | N/A | $r_s(9)$ =.97, p <0.05 | $r_s(9)$ =.95, p <0.05 | $r_s(7)$ =1, p <0.05 |
| **CPG** | $r_s(9)$ =.97, p <0.05 | N/A | $r_s(9)$ =.95, p <0.05 | $r_s(7)$ =1, p <0.05 |
| **3DCT** | $r_s(9)$ =.95, p <0.05 | $r_s(9)$ =.95, p <0.05 | N/A | $r_s(7)$ =.98, p <0.05 |
| **Direct** | $r_s(7)$ =1, p <0.05 | $r_s(7)$ =1, p <0.05 | $r_s(7)$ =.98, p <0.05 | N/A |

Table 6: Correlation coefficient matrix between measurement methods/models.

### 5.4 Whole model comparisons

To compare each 3D model in its entirety, cloud to mesh comparisons (CPG to SPG and 3DCT to SPG) were undertaken. A histogram of CPG-SPG distances (Figure 3) shows minimal difference between the two models (mean = 0.096mm, SD = 1.986; 95% of all points are below 2.30mm). Distances between 3DCT and SPG models (Figure 4) exhibit a larger mean discrepancy (mean = 0.412mm, SD = 1.221; 95% of all points are below 2.44mm). Examination of heat maps (scalar fields) demonstrated that both CPG and 3DCT models had larger dimensions than the reference SPG model at the right orbit and bilateral mid-temporal regions. A cross-sectional slice (0.89mm) at this site (Figure 5) showed divergent paths of all three model surfaces. At the right orbit, the largest CPG to SPG distance observed was 2.39mm; surface texture changes were seen in the same location (Figure 6) that suggest a loss of orbital volume. In the temporal regions, distances ranged between 2.47mm (left mid-temporal region) to 3.2mm (right anterior temporal region).

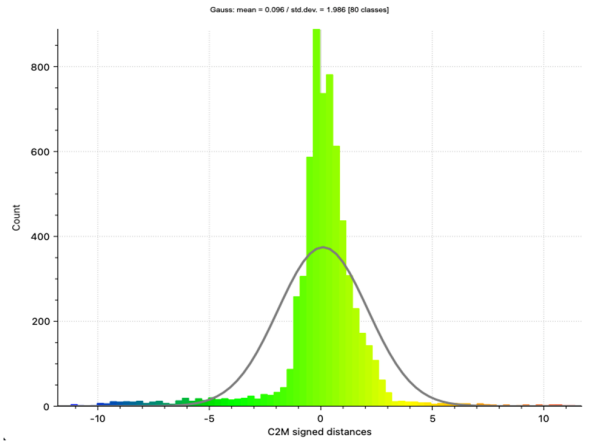


Figure 3: Histogram of Cloud to Mesh distances (x-axis in millimetres) between stereophotogrammetry and conventional photogrammetry models.

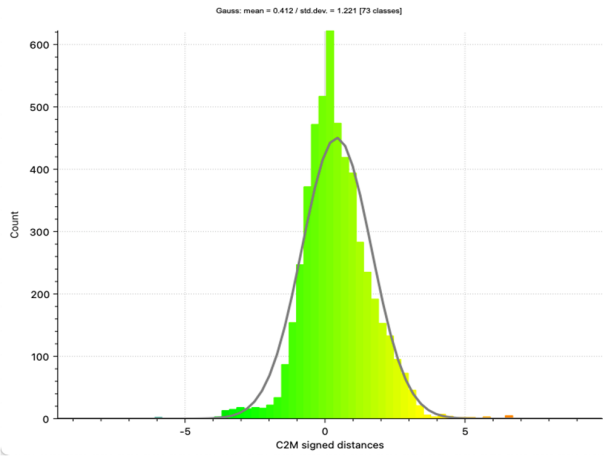


Figure 4: Histogram of Cloud to Mesh distances (x-axis in millimetres) between stereophotogrammetry and 3DCT models.

## 6. Conclusion

This study demonstrates that anatomical models produced by computed tomography and conventional photogrammetry produce similar facial measurements to those acquired by stereophotogrammetry and direct measurement at later timepoints. This provides preliminary evidence that a preserved body is sufficiently stable over time for accurate records to be generated by different imaging sensors. To our knowledge, this is the first time photogrammetry has been used to quantify soft tissue changes over time in a preserved human body donor.

Multiple sources of error can potentially occur within photogrammetry and CT imaging pipelines. The similarity observed between the SPG, CPG and 3DCT models on both landmark measurements (Table 4) and whole model comparisons (Figures 3 and 4) suggests that the CPG and 3DCT modelling pipelines used in this study were robust. For the CPG model, best-practice photogrammetry principles were followed, including the use of multiple imaging orbits which provided observational redundancy and a calibrated GCP network which produced accurate scaling. Furthermore, the high correlation between the direct and SPG measurements (Table 6) suggests that the use of stereophotogrammetry with the *Cliniface* software produced a suitable reference standard.

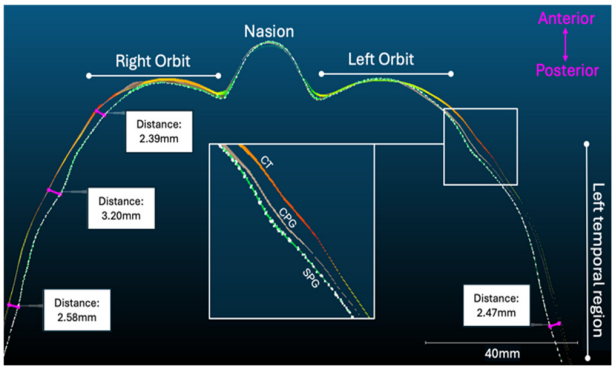


Figure 5: Cross-section through the orbits and temporal regions at the nasion, where maximum discrepancy was observed between the different models. The SPG surface (green surface, white cloud points) is convergent with the CPG (brown surface) and 3DCT surface (multi-coloured scalar field) at the nasion and mid-orbit. Divergent surfaces are present at the right lateral orbit (top-left image), right anterior temporal region (mid-left image) and bilateral mid-temporal regions (bottom left and right image).

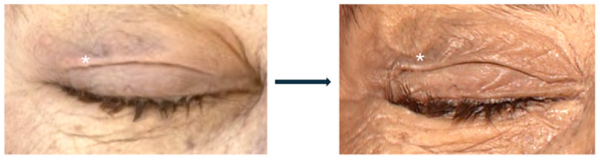

Figure 6: Comparison of right eye surface textures, as obtained by conventional photogrammetry (left) and stereo-photo-grammetry (right), 28 months later. The change in upper eyelid appearance suggests loss of orbital volume (white asterisk).

A small degree of tissue deformation occurred in a 28-month period, which appeared to be most prominent at the orbits and temporal regions (Figure 5). These regions contain large soft tissues, such as muscles, eyes and fat, which may be susceptible to dehydration over time. The same changes were not observed at the nasion where soft tissues are sparse, and the facial skeleton lies superficially. This pattern is consistent with the findings of Wilkinson and Tillotson (2012) who used laser scanning to assess facial decomposition in unpreserved cadavers. The authors found that orbital shrinkage and eyelid collapse occurs early, with little volume change at the upper and middle face. In this study, the apparent decrease in tissue volume did not significantly affect inter-landmark distances.

The 3DCT inter-landmark measurements had the largest measurement discrepancies (Table 4) and residuals (Figure 2), yet the whole-model comparisons showed the 3DCT geometry was highly consistent with the SPG model. The discrepancies may be due to variations in the examiners’ landmark placement for each repeated measure. Stull et al. (2014) found that 3DCT bone landmarks with high measurement errors were also error-prone via direct manual measurement, suggesting that the greatest source of measurement error is human error. Unlike measurements obtained through other methods (SPG, CPG and direct measurement), examiners could not derive surface texture information from the 3DCT surface. It is possible that the absence of skin texture made it difficult for examiners to identify surface landmarks. Future studies could investigate whether the presence or absence of surface texture on a photogrammetric model affects measurement reliability.

The findings of this study have implications for multi-sensor registration of facial models. We have shown that asynchronous capture by different sensors can produce models that can be closely co-registered, which is of benefit to institutions where simultaneous CT and PG capture is not possible, necessitating a time interval between image acquisition. However, there are potential drawbacks to this approach as repeated body positioning is required. Furthermore, facial tissues do not appear to deform uniformly, which results in discrepancies that become larger with time (figures 5 and 6). If perfect registration is required, we would recommend that simultaneous capture is attempted, or intervals between capture be kept as short as possible.

If 3D scans are used as quantitative anatomical records, they should be able to be measured accurately by multiple examiners and over multiple time periods. The intra-examiner and inter-examiner reliability was very high for all measurement methods used in this study (Table 5). Heike et al., (2009) reported interclass correlation coefficients for 30 craniofacial distances measured with stereophotogrammetry. For 26 measurements, the intra-rater reliability was over 0.95, whereas 21 measurements had an inter-rater reliability over 0.95. In contrast to our protocol which used five repeated measures, the authors took measurements twice on 40 living individuals. It is possible that a greater number of measurement repetitions have a training effect, leading to higher reliability. Omari et al., (2021) reported ICCs for four observers for physical, CT and PG landmarks. Measurements made in triplicate yielded ICCs ranging between 0.987 to 0.999. The high intra-rater and inter-rater reliability reported in their study is consistent with our findings. However, the reliability findings from our study should be interpreted with caution, as they were derived from two examiners on a single specimen.

This study has multiple limitations. Only a single specimen was used due to the limited availability of body donors. Formalin and phenol are common preservatives, but other alternatives exist such, as phenoxyethanol (Crosado et al., 2020), which may affect tissue volume differently over time. Inter-landmark measurements were confined to the face as a suitable reference imaging (SPG) was able to be generated for comparison. These findings cannot be generalised to other anatomical regions, particularly those with less bony support (e.g., anterior abdominal wall) and further research is required to validate the landmark accuracy and tissue stability in other regions.

3D visualisations of cadaveric facial anatomy can be successfully produced by computed tomography, conventional photogrammetry and stereophotogrammetry. There was no appreciable impact of imaging modality choice or time on landmark measurements, which means these techniques can be used in isolation, or together, to produce accurate anatomical records. The described imaging techniques can be executed asynchronously to develop models that are suitable for multi-

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-713-2026 | © Author(s) 2026. CC BY 4.0 License.

modal registration which are likely to be of benefit for anatomical education and quantitative research.

**Acknowledgements**

The authors sincerely thank those who donated their bodies to science, so that anatomical research could be performed. We also thank the technical staff in the Curtin University Anatomy Facility for their ongoing support, Curtin University’s 3D Hub for allowing us to use the VectraH2 imaging system and the Curtin Hub for Immersive Visualisation and eResearch (HIVE).

**References**

Akan, B., Akan, E., Sahan, A. O., Kalak, M. (2021). Evaluation of 3D Face-Scan images obtained by stereophotogrammetry and smartphone camera. Int Orthod, 19(4), 669-678. doi:10.1016/j.ortho.2021.08.007

Bolandzadeh, N., Bischof, W., Flores-Mir, C., Boulanger, P. (2013). Multimodal registration of three-dimensional maxillodental cone beam CT and photogrammetry data over time. Dentomaxillofacial Radiology, 42(2), 22027087.

Crosado, B., Loffler, S., Ondruschka, B., Zhang, M., Zwirner, J., & Hammer, N. (2020). Phenoxyethanol-Based Embalming for Anatomy Teaching: An 18 Years' Experience with Crosado Embalming at the University of Otago in New Zealand. Anat Sci Educ, 13(6), 778-793. doi:10.1002/ase.1933

Davis, C. R., Bates, A. S., Ellis, H., Roberts, A. M. (2014). Human anatomy: let the students tell us how to teach. Anat Sci Educ, 7(4), 262-272. doi:10.1002/ase.1424

Dixit, I., Kennedy, S., Piemontesi, J., Kennedy, B., Krebs, C. (2019). Which Tool Is Best: 3D Scanning or Photogrammetry – It Depends on the Task. In Biomedical Visualisation: Volume 1 (pp. 107-119). Cham: Springer International Publishing.

Donato, L., Cecchi, R., Goldoni, M., Ubelaker, D. H. (2020). Photogrammetry vs CT Scan: Evaluation of Accuracy of a Low-Cost Three-Dimensional Acquisition Method for Forensic Facial Approximation. J Forensic Sci, 65(4), 1260-1265. doi:10.1111/1556-4029.14319

Erolin, C. (2019). Interactive 3D digital models for anatomy and medical education. Biomedical Visualisation, 1-16.

Feddema, J. C., Chiu, L. Z. F. (2024). Accuracy and repeatability of 3D Photogrammetry to digitally reconstruct bones. Morphologie, 108(363), 100793. doi:10.1016/j.morpho.2024.100793

Gitto, L., Donato, L., Di Luca, A., Bryant, S. M., Serinelli, S. (2020). The Application of Photogrammetry in the Autopsy Room: A Basic, Practical Workflow. J Forensic Sci, 65(6), 2146-2154. doi:10.1111/1556-4029.14493

Heike, C. L., Cunningham, M. L., Hing, A. V., Stuhaug, E., Starr, J. R. (2009). Picture perfect? Reliability of craniofacial anthropometry using three-dimensional digital stereophotogrammetry. Plast Reconstr Surg, 124(4), 1261-1272. doi:10.1097/PRS.0b013e3181b454bd

Helmholz, P., Belton, D., Oliver, N., Hollick, J., & Woods, A. (2020). The Influence of the Point Cloud Comparison Methods on the Verification of Point Clouds Using the Batavia Reconstruction as a Case Study. IKUWA6 Shared Heritage: Proceedings of the Sixth International Congress for Underwater Archaeology, 370-381.

Kottner, S., Ebert, L. C., Ampanozi, G., Braun, M., Thali, M. J., Gascho, D. (2017). VirtoScan - a mobile, low-cost photogrammetry setup for fast post-mortem 3D full-body documentations in x-ray computed tomography and autopsy suites. Forensic Sci Med Pathol, 13(1), 34-43. doi:10.1007/s12024-016-9837-2

Mazonakis, M., and Damilakis, J. (2016). Computed tomography: What and how does it measure? Eur J Radiol, 85(8), 1499-1504. doi:10.1016/j.ejrad.2016.03.002

McGraw, K. O., Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. Psychological methods, 1(1), 30.

McMenamin, P. G., Quayle, M. R., McHenry, C. R., Adams, J. W. (2014). The production of anatomical teaching resources using three-dimensional (3D) printing technology. Anatomical sciences education, 7(6), 479-486.

Omari, R., Hunt, C., Coumbaros, J., Chapman, B. (2021). Virtual anthropology? Reliability of three-dimensional photogrammetry as a forensic anthropology measurement and documentation technique. Int J Legal Med, 135(3), 939-950. doi:10.1007/s00414-020-02473-z

Palmer, R. L., Helmholz, P., Baynam, G. (2020). Cliniface: Phenotypic Visualisation and Analysis Using Non-Rigid Registration of 3d Facial Images. ISPRS Archives, XLIII-B2-2020, 301-308. doi:10.5194/isprs-archives-XLIII-B2-2020-301-2020

Stull, K. E., Tise, M. L., Ali, Z., Fowler, D. R. (2014). Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images. Forensic Sci Int, 238, 133-140. doi:10.1016/j.forsciint.2014.03.005

Titmus, M., de Oliveira, B. I., Ellery, P., Whittaker, G., Radley, H., Radunski, M., Ng, L., Helmholz, P., Sun, Z. (2025). Using design thinking to create and implement a 3D digital library of anatomical specimens. Clin Anat, 38(4), 419-431. doi:10.1002/ca.24198

Titmus, M., Whittaker, G., Radunski, M., Ellery, P., IR de Oliveira, B., Radley, H., Helmholz, P., Sun, Z. (2023). A workflow for the creation of photorealistic 3D cadaveric specimens using photogrammetry. Journal of Anatomy, 243(2), 319-333. doi:10.1111/joa.13872

Vandenbossche, V., Van de Velde, J., Avet, S., Willaert, W., Soltvedt, S., Smit, N., Audenaert, E. (2022). Digital body preservation: Technique and applications. Anat Sci Educ, 15(4), 731-744. doi:10.1002/ase.2199

Viegas, S., Ladeira, C., Nunes, C., Malta-Vacas, J., Gomes, M., Brito, M., Mendonça, P., Prista, J. (2010). Genotoxic effects in occupational exposure to formaldehyde: A study in anatomy and pathology laboratories and formaldehyde-resins production. Journal of occupational medicine and toxicology, 5(1), 1-8.

Waltenberger, L., Rebay-Salisbury, K., Mitteroecker, P. (2021). Three-dimensional surface scanning methods in osteology: A topographical and geometric morphometric comparison. Am J Phys Anthropol, 174(4), 846-858. doi:10.1002/ajpa.24204

Wilkinson, C., & Tillotson, A. (2012). Post-mortem prediction of facial appearance. Craniofacial identification, 166-183.