European Journal of Orthodontics 28 (2006) 408–415
doi:10.1093/ejo/cjl024
Advance Access publication 11 August 2006

Reproducibility of soft tissue landmarks on three-dimensional facial scans

Jamie R. Gwilliam*, Susan J. Cunningham* and Tim Hutton**
*Orthodontic Department and **Unit of Dental Informatics, Eastman Dental Institute, University College London, UK

SUMMARY This study evaluated the reproducibility of 24 soft tissue landmarks on six three-dimensional (3D) facial scans. The scans were taken on a DSP400 facial scanner and were viewed using a customized software program. Intraoperator data were obtained by one researcher placing the 24 landmarks on all six scans a total of 30 times. Thirty different orthodontists of varying experience were then asked to place all 24 landmarks on each of the six facial scans in order to establish interoperator reproducibility. The standard deviations (SDs) from the mean were calculated from the data for each individual landmark in the *x*-, *y*-, and *z*-axes.

For the intraoperator data, 12 of the 24 landmarks were found to be reproducible to within a 1 mm SD for each plane of space. The interoperator data showed lower reproducibility with just two landmarks showing less than a 1 mm SD in all three planes of space.

Familiarity with 3D facial scans and associated software programs is important in improving reproducibility. In addition, the landmarks investigated in this study included those not often used. It is suggested that landmarks showing poor reproducibility for both inter- and intraoperator data should be avoided, if at all possible, or at least used with caution.

Introduction

The reproducibility of facial landmarks has been studied at length in two dimensions through the role of cephalometrics in orthodontics. However, as the face is a three-dimensional (3D) structure, the need to record its position in three dimensions has been highlighted (Ferrario *et al.*, 1996; Hajeer *et al.*, 2002). As computer software has developed, so too has the ability to visualize internal and external structures of the body. A number of methods of visualizing the face in three dimensions have been proposed, each with inherent advantages and disadvantages. For quantitative evaluation of facial morphology in three dimensions, it is preferable if little patient co-operation is required, especially if young children are involved. Furthermore, in studies of growth and development, methods should be non-invasive and non-ionizing so that repeat registrations are possible, and to allow control groups to be screened.

Two-dimensional imaging

Although 3D technology may be making groundbreaking progress in hospital and research settings, both primary and secondary care sectors continue to use predominantly two-dimensional imaging. Standard photographic and radiographic views remain the principle adjunct to orthodontic care. However, standardizing these images is difficult. Lateral cephalograms and photographs are frequently taken using different hardware which do not have directly comparable degrees of magnification.

The 3D imaging

A number of different image modalities have been developed over the years. Some have found greater acceptance in medicine and dentistry than others. Many have come under considerable scrutiny, including studies establishing the reproducibility of landmark placement (McCance *et al.*, 1993; Moss *et al.*, 1994; Kragskov *et al.*, 1997; Hajeer *et al.*, 2002). The use of laser scans and stereophotogrammetry have gained most favour in the orthodontic literature, principally because they are both non-invasive and non-ionizing.

Laser scans

Laser scanning is a non-invasive method of capturing 3D facial images and has been successfully used in studies of treatment outcome and relapse (McCance *et al.*, 1992a,b; Moss *et al.*, 1994; Morris *et al.*, 1998; Coward *et al.*, 2000; Ismail and Moss, 2002). Changes in dimensions between repeated scans or changes as a result of treatment are often shown by colour differentiation or colour 'maps'. The main disadvantage is that it takes between 8 and 10 seconds to capture an image successfully thus making distortion more likely, particularly with younger patients. The lack of soft tissue surface texture has also been highlighted as a possible drawback, although the use of a white light laser may prove more effective (Hajeer *et al.*, 2002). In addition, patients are scanned with their eyes closed, which may interfere with the natural facial expression and any landmarks placed around the eyes.

Stereophotogrammetry

Stereophotogrammetry is a method of obtaining an image by means of one or more stereo pairs of photographs taken simultaneously. The technique is not a new concept, having been discussed by Burke and Beard (1967). Initial methods of stereophotogrammetry required the use of a cephalostat-type device to hold the patient in position. However, the use of digital cameras and powerful computer software has revolutionized the technique. Ayoub *et al.* (2003) assessed this method of 3D imaging and reported favourable results. By contrasting its performance with a previously validated technique (3D contact ultrasonic measuring system), it was shown that the overall error between measurements was less than 0.6 mm. However, the images examined were of cleft study models with highly defined landmarks and a reduced surface area, as opposed to the face. The main advantages over other methods of 3D imaging were proposed by Ayoub *et al.* (1996) as rapid image capture, hence making it suitable for young patients (even babies), accurate identification of facial landmarks to within 0.5 mm, and immediate generation of a 3D display.

It is also useful to note that the 3D display produced with stereophotogrammetry has surface texture and subjects may recognize their own image more readily than with laser scanning. This feature of stereophotogrammetry is particularly important in the realm of patient consent, where, for example, recognition of oneself prior to surgery is vital if predictive images are to be discussed.

In their reproducibility study of landmarks on 3D facial scans, Hajeer *et al.* (2002), utilized pre- and post-operative scans of five orthognathic patients and identified 24 landmarks on each. In addition to using the anthropometric landmarks of Farkas (1994), they defined four new landmarks. Landmark identification was repeated three times with a 1 week interval between each session of digitization. The software program used gave the x-, y-, and z co-ordinates for each of the landmarks, and by finding the differences between the individual axes points, the mean differences from the three comparisons could be calculated. The accepted, but arbitrary, cut off point was 0.5 mm and landmarks that were not found to be reproducible were x-axis: left and right gonion, left and right zygion, and menton; y-axis: left and right gonion, left and right zygion, left and right tragion, and glabella; z-axis: left and right gonion, left and right tragion, menton, left and right otobasion inferius.

There was, however, no attempt at examining interoperator reproducibility in that study. Neither was any reason given as to why a period of 1 week was deemed appropriate for the interval of the landmark placements. It appears that investigators in this field cannot agree on how long an interval between landmark identification is required to reduce any effect of memory on landmark reproducibility.

The purpose of the present study was to investigate the reproducibility of various landmarks on 3D stereophotogrammetry facial images. In particular, the assessment of reproducibility of various landmarks between operators was undertaken as this information does not appear to have been published previously. A further aim was to determine which landmarks used on 3D facial images were most reproducible for both inter- and intraoperator data.

Materials and methods

Selection of facial images

Six faces were selected from the archive of stereophotogrammetry images at the Eastman Dental Institute. This was not a random process as the images were of differing qualities and many were of syndromic children. As the purpose of this research was to examine the repeatability of landmarks on non-syndromic patients, three adults and three children were selected whose images were deemed to be of reasonable quality by all three authors.

All scans were taken using a DSP400 facial scanner (3dMD LLC, Harefield, Middlesex, UK). A stereo picture of the subject is produced by the six cameras, all of which are at different angulations. The images produced were then articulated and displayed using a custom software program (ShapeFind) written by one of the authors (TH).

Selection of landmarks

The landmarks used by Hajeer *et al.* (2002) i.e. superior labial sulcus and zygion were combined with the classic points defined by Farkas (1994) to produce a total of 24 landmarks.

The landmarks selected represented those used most commonly in previous studies of 3D facial imaging. A combination of both midline and sagittal points were required, including those where there had been, or where it was anticipated that there would be, varying degrees of reproducibility. The number of landmarks was restricted to allow the operators to place all landmarks within an appropriate time scale of less than 1 hour.

Intraoperator assessment of repeatability

All 24 points were placed on each of the six chosen facial images a total of 30 times by one researcher (JRG). This figure (30 times) was chosen to allow statistical analysis of the data, while reducing the possibility of allowing reproducibility of landmark identification due to memory. The images were marked at intervals of 7 days to further reduce the potential for memory bias. All images were marked on the same computer under identical lighting conditions.

The data from the landmarked images were placed in a spreadsheet to give the *x*-, *y*-, and *z* co-ordinates.

Interoperator assessment of repeatability

Thirty orthodontists of varying levels of experience (equal gender distribution, age range 28–57 years, seven of which had a minimum 3 years postgraduate orthodontic training) were asked to place the same landmarks on all six faces. Supervision in the use of the software program was provided by one author, but no assistance was given in placing the landmarks. A description of each point was given according to Farkas (1994). There was no specified time limit for placement of landmarks.

Assessment of direction of error

To be able to determine the direction of error, it was necessary to describe a true horizontal and vertical reference line. Reference landmarks were used to define the Frankfort horizontal, with the placement of these undertaken by an experienced orthodontist (SJC). These points were reviewed 1 week after placement, affirmed and the lines connecting orbitale and assumed porion were joined bilaterally to create constructed Frankfort horizontal. A vertical axis was mathematically constructed from these four landmarks and, with such axes defined, it was possible to determine in which plane of space a particular landmark was most reproducible (see Appendix for details of axis construction).

Results

A representative graph of landmark errors on one image is shown in Figure 1, however, all facial images showed a similar pattern of errors.

In addition, the data for seven of the operators who were deemed more experienced were extracted and examined in comparison with those operators who were less experienced. There was no apparent increase in reproducibility of landmark placement with operator experience. Combined data are therefore presented.

Images showing the deviation from the mean can also be viewed in three dimensions, showing the scatter away from the mean 'ellipse of error'. Although by necessity shown in two dimensions, Figure 2 demonstrates the differing reproducibility of each landmark with the larger shapes denoting poorer reproducibility. It is particularly worth noting the poorer reproducibility of gonion on the profile view.

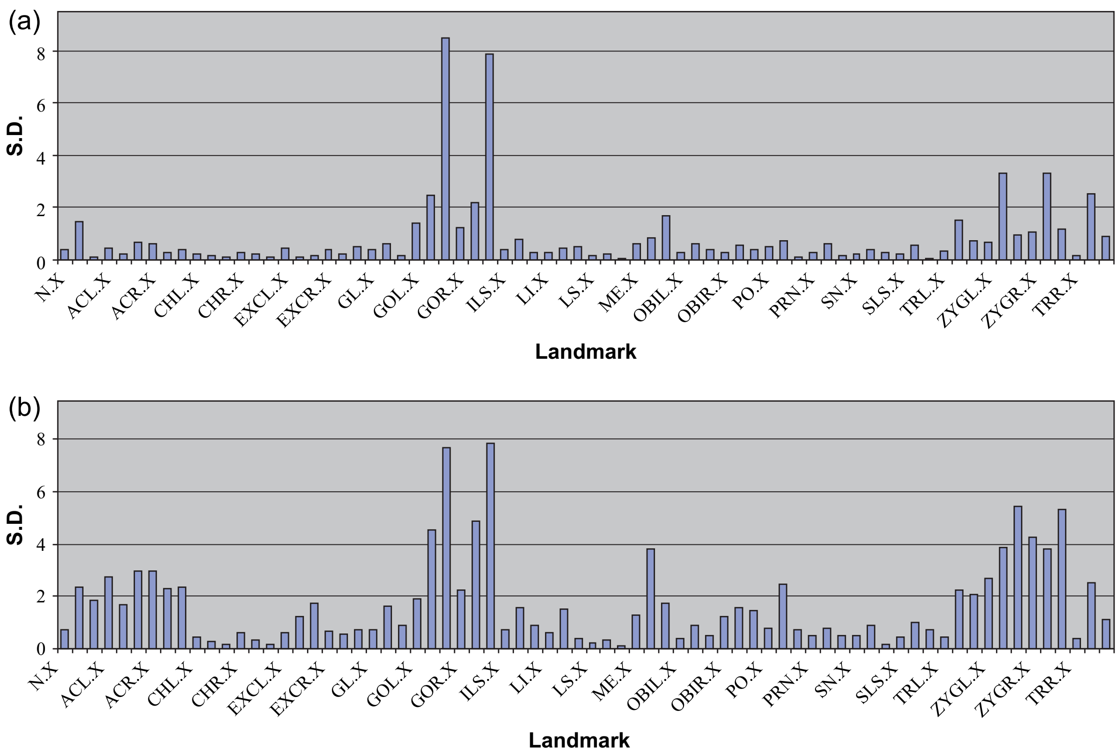

Figure 1 Intra- (a) and inter- (b) operator reproducibility data for one image (Image 1). L = Left; R = Right; X represents reproducibility in the *x*-axis for each landmark; and the consecutive bar charts represent the *y*- and *z*-axes. Please refer to Table 1 for landmark definition.

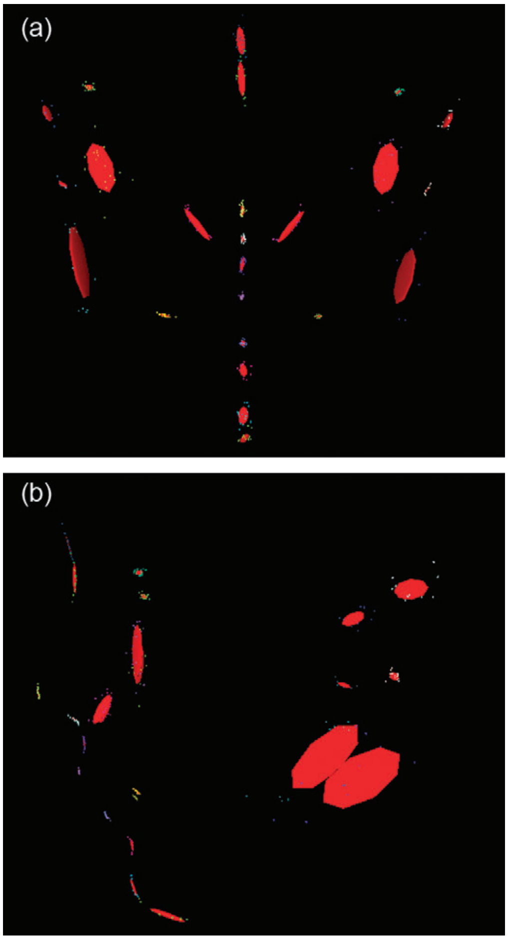

Figure 2 Examples from three-dimensional image of 'ellipse of error', (a) frontal and (b) profile view.

Table 2 highlights the number of landmarks, out of a total of 24, which were found to be reproducible to within the given SD range for all three planes of space. The high (less than 0.5 mm SD) or good (less than 1.0 mm SD) reproducibility for each facial image for both intra- and interoperator data are shown. In general, approximately twice as many landmarks showed good reproducibility for the intraoperator data set as opposed to the interoperator data.

Discussion

This study investigated the reproducibility of various soft tissue landmarks on 3D facial images. The reproducibility of some landmarks was found to be less than in a previous study (Hajeer *et al.*, 2002) where 20 out of the 30 landmarks used were found to be highly reproducible. However, it should be noted that those authors examined intraoperator data only.

Materials: facial scanner and software program

For some individuals, even if the image was considered of good quality, the facial surface was not completely captured, with voids evident in certain areas, particularly around the hairline and ears. In some cases, the image 'wire framework' showed large polygons, and landmarks placed on these will not be as specific as those placed in areas of high-density wire polygons. To minimize these errors, subjects were chosen in which the facial scanner had defined the majority of the face, with minimal voids in

Table 1 Landmark definitions.

Landmark	Definition
Alar curvature (or alar crest)	The most lateral point in the curved base line of each ala, indicating the facial insertion of the nasal wing base
Cheillion	The point located at each labial commissure
Exocanthion(Exc)	The point at the outer commissure of the eye fissure, located slightly medial to bony Exc
Glabella (Gl)	The most prominent midline point between the eyebrows, analogous to bony Gl on the frontal bone
Gonion (Go)	The most lateral point on the mandibular angle close to bony Go
Inferior labial sulcus	The deepest midline point on the labiomental fold, which determines the lower border of the lower lip or the upper border of the chin
Labrale inferius	The midpoint of the lower vermilion line
Labrale superius (Ls)	The midpoint of the upper vermilion line
Menton (Me)	The lowest median landmark on the lower border of the mandible, analogous to bony Me
Nasion (N)	The point in the midline of both the nasal root and the nasofrontal suture, always above the line that connects the two inner canthi, analogous to bony N
Otobaion inferius	The most inferior point on the earlobe, located at the attachment (junction) of the lobe to the face
Pogonion	The most anterior midpoint of the chin, located on the skin surface in front of the analogous bony landmark of the mandible
Pronasale	The most protruded point of the apex nasi identified in lateral view of the rest position of the head
Subnasale (Sn)	The midpoint of the angle at the columella base where the lower border of the nasal septum and surface of the upper lip meet: not identical to the bony point anterior nasal spine or nasospinale
Superior labial sulcus	The deepest midline point on the upper lip, which is located usually halfway between Sn and Ls
Tragion	The notch on the upper margin of the tragus
Zygion (Zyg)	The most prominent point on the cheek area beneath the outer canthus and slightly medial to the vertical line passing through it; differs from bony Zyg

Table 2 Number of landmarks (intra- and interoperator data) showing good or high reproducibility.

Scan no.	Landmarks within a SD < 0.5 mm		Landmarks with a SD between 0.5 and 1.0 mm		Total landmarks within a SD < 1.0 mm	
	Intraoperator	Interoperator	Intraoperator	Interoperator	Intraoperator	Interoperator
1	6	2	16	8	22	10
2	5	2	13	9	18	11
3	6	0	13	7	19	7
4	5	1	15	8	20	9
5	5	2	14	7	19	9
6	8	0	14	6	22	6

SD, standard deviation.

areas in which landmarks were to be placed. Nevertheless, the images did not show perfection as can be seen in Figure 3. This represents Image 5 and demonstrates the larger polygons in the ear region.

Familiarity with the ShapeFind program is also likely to play a part in an assessor's ability to place landmarks. In this study, despite training in the use of the program, the assessors who were less confident with the software tended to view the facial image primarily from the frontal or profile view rather than making full use of the program's zoom and rotate functions. Accurate landmark identification requires full 3D control by the operator in manipulating the images. Training in such skills, as in training for tracing of lateral cephalograms, should improve reproducibility of landmark placement.

Data co-ordinates

To be of clinical use, the reproducibility of each landmark must be elucidated in all three planes of space. To establish the *x*-, *y*-, and *z*-axes required an experienced clinician to place landmarks to form a true horizontal or Frankfort plane. Here lies a paradox of the evaluation: to assess reproducibility in different planes of space relative to the subject requires the placement of landmarks which themselves are prone to error. However, rather than have SDs produced around arbitrary axes, it was considered necessary to create this plane for all six subjects so that the mean distance from the relevant axes would have some meaning relative to the subject's face. In doing this, error is introduced as the plane does not exactly represent a true vertical or absolute horizontal. This is because the vertical axis was created by extrapolation of the two horizontal axes (external auditory meatus to orbitale, both left and right) and not from landmarks in a vertical plane. However, even if not absolutely accurate, these axes allow us to identify in which dimension each landmark has the best and worst reproducibility.

Reproducibility of landmarks: intraoperator data

As would be expected, one person's perception in terms of landmark reproducibility was better than between different assessors. In addition, it must be acknowledged that after landmarking the images 30 times, there may be some memory bias related to landmark placement. However, some landmarks are clearly less reproducible than others. Although there was some consistency, there was also some variation between the images. In addition, some landmarks showed a greater degree of reproducibility in different planes of space (*x*-, *y*-, or *z* axes).

Hajeer *et al.* (2002) suggested that a landmark was deemed highly reproducible if the SD from the centroid was 0.5 mm or less in all three planes of space. If these guidelines are followed, then for the intraoperator data, only the points cheilion (left and right), labrale superius, and exocanthion (left) were found to be highly reproducible for all images. This compares unfavourably with the data of Hajeer *et al.* (2002) data in which 20 out of the 30 landmarks were found to be highly reproducible.

However, if the SD is increased to 1.0 mm, then the landmarks, alar crest (right), exocanthion (right), glabella, inferior labial sulcus, labrale inferius, otobaion inferius (right), superior labial sulcus, and subnasale were all considered reproducible in all three planes of space for the intraoperator study (making a total of 12 out of 24 landmarks).

Reproducibility of landmarks: interoperator data

Among the 30 assessors, given a SD of 0.5 mm as being acceptable, none of the landmarks were found to be highly reproducible for all three axes on the six images. Even with a higher SD of 1.0 mm, only cheillion (right) and labrale superioris were found to be reproducible on all images for all three axes. No previous studies were found in the literature with which to compare these data.

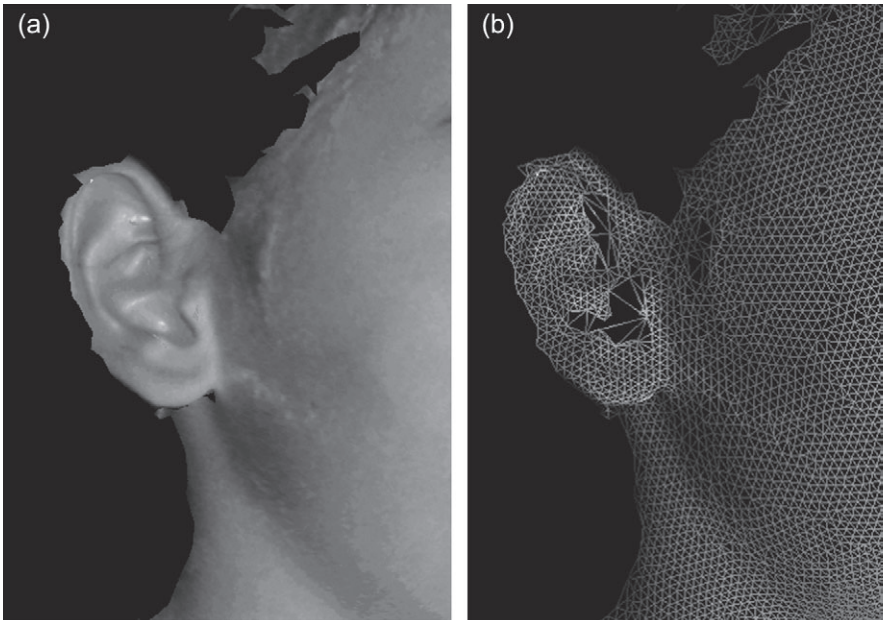

Figure 3 Image number 5 as seen with skin surface (a) and with the wire framework polygons, (b) which make up the subsurface structure of the image.

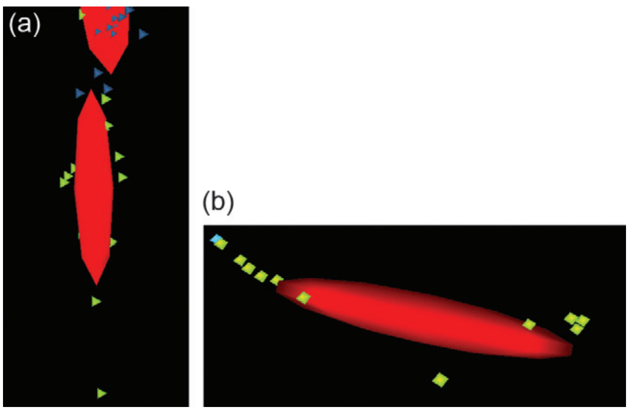

Figure 4 Plot showing an example of the scatter of landmark placement for (a) soft tissue nasion and (b) soft tissue menton. The red central areas indicate the placement of landmarks within one standard deviation, with the outliers in green.

Clinician experience

Although Houston (1983) suggested that the experience of the clinician plays an important role in the reproducibility of landmarks on cephalometric tracings, it seems this is not the case for soft tissue landmarks on 3D facial images. This may be due to the fact that even experienced clinicians have limited experience of placing landmarks on 3D computer images, compared with their considerable experience in tracing lateral cephalograms. It should, however, be noted that the sample size of experienced clinicians was small (n = 7) and further research comparing experienced versus non-experienced operators would be of benefit.

Specific landmark reproducibility

Each landmark was represented by a diagram to demonstrate its individual ellipse of errors. Data from the interoperator data set have been used for illustration, but the intraoperator data showed a similar pattern albeit with smaller SDs. The images are created from the placement of landmarks and form a 'cloud'-like appearance of scattered points in three dimensions. Each landmark has a particular pattern of error giving rise to the shapes seen in Figure 4a,b. The block shapes show the area where the majority of placed landmarks lie (within 1 SD) with the landmarks placed outside this illustrated as single points. In this way, it is easy to assess in which direction the majority of errors lie (x-, y-, or z-axis). In agreement with Hajeer *et al.* (2002), the landmarks gonion, zygion, and tragion showed unacceptable reproducibility.

Soft tissue nasion

Reproducibility of soft tissue nasion was relatively poor for both inter- and intraoperator data in the y-axis. One reason may be that soft tissue nasion can only be placed accurately with the patient in a natural head position (NHP) in lateral profile. This requires good manipulation skills in order to move the image to the correct position and also a good clinical knowledge of NHP. Failing to achieve this may mean that nasion is placed too high or too low vertically. However, in the x-axis it was highly reproducible (SD < 0.5 mm for all values in the intraoperator data set).

Alar crest

This landmark was generally positioned with fair reproducibility for the intraoperator data set. However, it was the vertical positioning that tended to be more consistent (all subjects had a SD below 0.5 mm) but in the other planes of space, the landmarks were found to be less reproducible. This might be a reflection of the contour of the nose in this area. For example, a landmark placed on the convexity of the nose will have different *z* dimensions to one placed in the alar fold. This will clearly depend on the operator's perception of the landmark definition.

Soft tissue gonion

Of all the landmarks, left and right soft tissue gonion were the least reproducible, in agreement with the findings of Hajeer *et al.* (2002). The lack of colour contrast and the potential for concealment of this area by the soft tissues are possible reasons for its poor reproducibility. Also, the flash from the camera tends to leave a shadow in this area, giving the angle of the mandible poorer definition.

The poor reproducibility of soft tissue gonion was particularly noticeable in the *y*- and *z*-axes. Horizontal control of soft tissue gonion (*x*-axis) is governed by the skin image and is less likely to be variable. Vertical height and depth of soft tissue gonion point are more prone to variability.

Soft tissue menton

It is predominantly in the *z*-axis that soft tissue menton had the poorest reproducibility. This result may not be altogether surprising as horizontal and vertical placement will be aided by the midline of the face and any previously plotted midline landmarks. The depth (or *z*-axis), however, will be more prone to error as the lowest point on the soft tissue of the mandible will be governed by the assessor's ability to place the subject in NHP.

With increasing use of 3D images in orthodontics, there is a need to know which landmarks show good levels of reproducibility and are therefore accurate for the purposes for which they are used. This study suggests that landmarks such as gonion and zygion would be unreliable and that information derived from these points should be treated with caution.

Conclusions

1. Reproducibility of 3D soft tissue landmarks varies depending upon the landmark being placed and the assessor placing the landmark. This study found a higher degree of reproducibility for a single assessor placing landmarks than between assessors. For good reproducibility, landmarks must be well-defined and clearly understood, with both written and illustrated information given to aid assessors.
2. Different facial landmarks have wide variation in their degree of reproducibility. Landmarks with well-defined borders or edges showed higher degrees of reproducibility than those placed on gently curving slopes.
3. It is important to become familiar with the software program used to view the images in order to improve landmark reproducibility.
4. Some landmarks should be used with caution, or not used at all, as their reproducibility may be questionable.

Address for correspondence

Dr J. R. Gwilliam
Eastman Dental Institute
University College London
256 Gray's Inn Road
London WC1X 8LD
UK
E-mail: jgwilliam@eastman.ucl.ac.uk

Acknowledgement

We thank Professor Peter Hammond for access to the Eastman Dental Institute image archive and for advice during the study.

References

Ayoub A F *et al.* 1996 Three-dimensional modeling for modern diagnosis and planning in maxillofacial surgery. International Journal of Adult Orthodontics and Orthognathic Surgery 11: 225–233

Ayoub A F *et al.* 2003 Validation of a vision-based, three-dimensional facial imaging system. Cleft Palate-Craniofacial Journal 40: 523–529

Burke P H, Beard F H 1967 Stereophotogrammetry of the face. A preliminary investigation into the accuracy of a simplified system evolved for contour mapping by photography. American Journal of Orthodontics 53: 769–782

Coward T J, Scott B J, Watson R M, Richards R 2000 Laser scanning of the ear identifying the shape and position in subjects with normal facial symmetry. International Journal of Oral and Maxillofacial Surgery 29: 18–23

Farkas L G 1994 Anthropometry of the head and face, 2nd edn. Raven Press, New York

Ferrario V F, Sforza C, Poggio C E, Serrao G 1996 Facial three-dimensional morphometry. American Journal of Orthodontics and Dentofacial Orthopedics 109: 86–93

Hajeer M Y, Ayoub A F, Millett D T, Bock M, Siebert J P 2002 Three-dimensional imaging in orthognathic surgery: the clinical application of a new method. International Journal of Adult Orthodontics and Orthognathic Surgery 17: 318–330

Houston W J B 1983 The analysis of errors in orthodontic measurements. American Journal of Orthodontics 83: 382–390

Ismail S F, Moss J P 2002 The three-dimensional effects of orthodontic treatment on the facial soft tissues—a preliminary study. British Dental Journal 192: 104–108

Kragskov J, Bosch C, Gyldensted C, Sindet-Pedersen S 1997 Comparison of the reliability of craniofacial anatomic landmarks based on cephalometric radiographs and three-dimensional CT scans. Cleft Palate-Craniofacial Journal 34: 111–116

McCance A M, Moss J P, Fright W R, Linney A D, James D R 1992a A three-dimensional soft tissue analysis of 16 skeletal Class III patients following bimaxillary surgery. British Journal of Oral and Maxillofacial Surgery 30: 221–232

McCance A M, Moss J P, Fright W R, James D R, Linney A D 1992b A three dimensional analysis of soft and hard tissue changes following bimaxillary orthognathic surgery in Skeletal III patients. British Journal of Oral and Maxillofacial Surgery 30: 305–312

McCance A M, Moss J P, Fright W R, James D R, Linney A D 1993 A three dimensional analysis of bone and soft tissue to bone ratio of movements in 17 Skeletal II patients following orthognathic surgery. European Journal of Orthodontics 15: 97–106

Morris D O, Illing H M, Lee R T 1998 A prospective evaluation of Bass, Bionator and Twin Block appliances. Part II-The soft tissues. European Journal of Orthodontics 20:663–684

Moss J P, McCance A M, Fright W R, Linney A D, James D R 1994 A three-dimensional soft tissue analysis of fifteen patients with Class II, division 1 malocclusions after bimaxillary surgery. American Journal of Orthodontics and Dentofacial Orthopedics 105: 430–437

Appendix

The origin and axes of the plane were calculated using the following formulae: Origin = (a + b) / 2,

x-axis = normalize (a − b) (where normalize makes the vector of unit length),

Corrected c = c − [*x*-axis × dot product (c-origin, *x*-axis)],

z-axis = normalize (corrected c − origin),

y-axis = normalize [cross product (*x*-axis, *z*-axis)].

As the *y*-axis should always be approximating up (the scanner captures the image with the subject sitting in a ‘normal’ position) the *y*-axis can be defined as

y-axis: y < 0 then *y*-axis = *y*-axis × (−1).