Journal of Anatomy

J. Anat. (2011) **219**, pp444–455 doi: 10.1111/j.1469-7580.2011.01411.x

Spatially-dense 3D facial asymmetry assessment in both typical and disordered growth

Peter Claes,[1,2] Mark Walters,[3] Dirk Vandermeulen[2] and John Gerald Clement[1]

[1]*Melbourne Dental School, The University of Melbourne, Melbourne, Victoria, Australia*
[2]*K.U. Leuven, Medical Imaging Research Center (MIRC), Faculty of Engineering, Department of Electrical Engineering – ESAT, Center for Processing Speech and Images – PSI, Herestraat, Leuven, Belgium*
[3]*Cranio-Maxillo-Facial Unit, Princess Margaret Hospital for Children, Subiaco, Perth, WA, Australia*

Abstract

Mild facial asymmetries are common in typical growth patterns. Therefore, detection of disordered facial growth patterns in individuals characterized by asymmetries is preferably accomplished by reference to the typical variation found in the general population rather than to some ideal of perfect symmetry, which rarely exists. This presents a challenge in developing an asymmetry assessment tool that is applicable, without modification, to detect both mild and severe facial asymmetries. In this paper we use concepts from geometric morphometrics to obtain robust and spatially-dense asymmetry assessments using a superimposition protocol for comparison of a face with its mirror image. Spatially-dense localization of asymmetries was achieved using an anthropometric mask consisting of uniformly sampled quasi-landmarks that were automatically indicated on 3D facial images. Robustness, in the sense of an unbiased analysis under increasing asymmetry, was ensured by an adaptive, robust, least-squares superimposition. The degree of overall asymmetry in an individual was scored using a root-mean-squared-error, and the proportion was scored using a novel relative significant asymmetry percentage. This protocol was applied to a database of 3D facial images from 359 young healthy individuals and three individuals with disordered facial growth. Typical asymmetry statistics were derived and were mainly located on, but not limited to, the lower two-thirds of the face in males and females. The asymmetry in males was more extensive and of a greater magnitude than in females. This protocol and proposed scoring of asymmetry with accompanying reference statistics will be useful for the detection and quantification of facial asymmetry in future studies.

Key words: 3D imaging; anthropometric mask; facial asymmetry; robust least-squares.

Introduction

> Our notion of symmetry is derived from the human face
> Blaise Pascal (1623–1662)

Bilateral symmetry often occurs in organisms, including humans, and is defined with respect to reflection or mirroring across the midsagittal plane. This plane divides a perfectly bilaterally symmetrical organism into equal right and left halves, and thus equates to the plane of symmetry. Any deviation from perfect symmetry is defined as asymmetry, and during development in vertebrates imbalances in growth will inevitably result in a degree of asymmetry (Hamada et al. 2002). Though departure from symmetry is a property of the individual, patterns of asymmetry are studied at the level of the population and can be grouped into three categories (Palmer & Strobeck, 1986; Palmer, 1994): first, directional asymmetry (DA), representing the consistent greater development of characteristics in a population on one side of the body than the other; second, antisymmetry (AS), where the greater development is not consistently biased to one side only, but occurs on both sides with approximately equal frequency; and third, fluctuating asymmetry (FA), resulting in the inability of a characteristic to develop in a predetermined way (Van Valen, 1962). In a variety of contexts, however, including this study, a score of overall (Klingenberg & McIntyre, 1998) or total (Richtsmeier et al. 2005) asymmetry per individual is required. This score encodes for the combined magnitude of DA and FA (or AS) in an individual.

The face is the most identity-coding part of the human body (Smeets et al. 2010). Humans are adept at detecting

Correspondence
Peter Claes, Melbourne Dental School, The University of Melbourne, Melbourne, 3010 Victoria, Australia. E: peter.claes1979@gmail.com; peter.claes@esat.kuleuven.be

Accepted for publication *13 June 2011*
Article published online *11 July 2011*

subtle differences in facial appearance from which an individual's gender, ethnicity, age, intelligence, emotional state and health is often unconsciously inferred. Facial asymmetry is similarly cognitively universally detectable; for example, relatively rare conditions such as hemifacial microsomia are depicted in anthropological collections throughout the world (Poswillo, 1989).

Mild facial asymmetries are common in typical growth and development (Ferrario et al. 1995; Ercan et al. 2008) and are, for example, an important factor in mate selection (Baudouin & Tiberghien, 2004), with some asymmetry features accepted as a trait of beauty (Zaidel & Cohen, 2005). Severe and pathological asymmetries, on the other hand, are a feature of disordered growth as a consequence of congenital and/or environmental causes (Yu et al. 2009). They can interfere with normal function and aesthetic appearance (Rossi et al. 2003). An accurate, objective and quantified means to assess an individual's asymmetry as a component of typical variation and/or as a consequence of disordered growth is therefore beneficial. It is, furthermore, important that any such assessment protocol is applicable for both typical and disordered growth patterns without any modification and assumptions.

Different approaches for calculating facial asymmetry and the associated midsagittal plane have been presented in the past. One main approach is to divide the face into left and right hemifaces first in order to compare them subsequently. This comparison is traditionally done using conventional morphometric measurements, like distances, angles, areas or ratios. Differences of pair-wise corresponding measurements are taken on both hemifaces separately (Ferrario et al. 2001; Baudouin & Tiberghien, 2004) or distances perpendicular to the midsagittal plane (Ferrario et al. 1994), which is then assumed to be known, then quantify asymmetry. However, these conventional morphometric measurements often fail to represent the complete spatial arrangement of interest (Slice, 2007). More recently, a more spatially complete form analysis to compare 'half-forms' using EDMA (Euclidean Distance Matrix Analysis), as suggested in Cole (2001), can be used to study asymmetry. This technique is used to compare L–R hemifaces in the studies of Ferrario et al. (1995) and Ercan et al. (2008).

Alternatively to the separation into hemifaces, asymmetry can also be assessed by comparing the complete face with a mirror image of itself, which is simply a reflection of the face with respect to an arbitrary plane. This approach is considered more informative, for example, in dentofacial deformities (McIntyre & Mossey, 2002). One such protocol, grounded in geometric morphometrics, is given by Klingenberg et al. (2002) where form or shape is represented as a (landmark) configuration consisting of the coordinates of manually indicated and ordered landmarks. These landmarks are homologous points or points of correspondence on an object that match between and within populations (Dryden & Mardia, 1998), generating compatible configurations. Furthermore, they often have developmental, functional, structural or evolutionary significance (Richtsmeier et al. 2002). The original configuration is then transformed into its reflected configuration, by mirroring with respect to an arbitrary plane [Klingenberg et al. (2002) use the $x = 0$ plane]. The landmarks in the reflected configuration are relabelled to pair them with their homologous landmark in the opposite hemifacial region. This establishes a mirror configuration that is again compatible with the original configuration for further superimposition. A rigid Procrustes alignment of the original configuration and this reflected and compatible configuration is then carried out. Rigid Procrustes alignment minimizes the distance between homologous landmarks by a rigid (only translations and rotations) least-squares superimposition. An individual overall score of asymmetry can be obtained from the remaining distances (errors) between superimposed landmarks of the original and mirror configurations [e.g. root-mean-squared-error (RMSE)]. The consensus configuration defined as the average positions of the superimposed original and reflected landmarks is, by construction, symmetrical with respect to a plane, and is used to analyse variation of symmetry in populations. The thus generated symmetry plane is an estimate of the midsagittal plane providing a frame of reference with a clear anatomical meaning (Klingenberg et al. 2002).

Because faces are three-dimensional (3D) structures, their asymmetry assessment should be performed in 3D. This is facilitated through the availability of 3D scanning technology where surface scanners are the preferred choice when imaging healthy participants because images are captured rapidly, safely, cost-efficiently and non-invasively (Aeria et al. 2010). The Klingenberg protocol is applicable to both 2D and 3D data, but in order for it to be used on faces in both typical and disordered developmental conditions some modifications are required. Firstly, owing to the lack of anatomically discrete features in regions of the face, manually indicated landmarks provide only a sparse representation, and salient features of the facial form can be overlooked (Thomas, 2005). The demand to detect, quantify and visualize both subtle and severe asymmetries in discrete regions of the face requires more complete facial representations. This, however, presents the challenge for automated, standardized and spatially-dense point indications as well as their relabelling or their re-indication after reflection. Secondly, an overall score of asymmetry based on the RMSE, for example, only measures the severity or degree of asymmetry. Furthermore, we would also like to quantify the overall extent or proportion of the facial area considered to be asymmetrical. For example, the degree of asymmetry introduced by a facial pimple can be very high, but its extent is only limited. Finally, the least-squares superimposition is erroneous for subjects with severe asymmetry as least-squares solutions are sensitive to large differences in a

few landmarks. This is also known as the Pinocchio effect within shape analysis (Zelditch et al. 2004). It generates an unknown bias in the assessment for which no compensatory adjustment is possible.

In this work a spatially-dense and robust 3D facial asymmetry assessment protocol is proposed that is applicable for assessing faces associated with both typical and disordered growth. To achieve this, a previously established anthropometric mask (AM) mapping and robust superimposition strategy are employed to obtain complete facial representations in a standardized manner useful for the derivation of population statistics. A robust superimposition strategy assures assessments are undertaken without any bias for both typical and disordered growth patterns. Hence, it provides detail on the obvious as well as the more subtle asymmetries. Furthermore, it allows for the extraction of a novel numerical score, here defined as: 'the relative significant facial asymmetry' expressed as a percentage used to quantify the overall proportion of the facial area considered to be asymmetrical. This type of numerical data is provided in conjunction with the already known degree of asymmetry encoded in the RMSE. The protocol was applied to a database of 3D facial images from young healthy individuals to obtain reference statistics for typical males and females. The same technique, without any modification or specific knowledge, is then applied to three individuals with known disordered growth patterns all characterized by severe facial asymmetry.

Materials and methods

Participants

Three-dimensional (3D) facial images of healthy young adults between the ages of 18 and 25 years of admixed self-reported ancestry were made available from a library of facial scans, comprising The Western Australian 3-dimensional Facial Reference Range for Children and Adolescence. Subjects completed a questionnaire on relevant health history and population affinity. Exclusion was made on the basis of self-reported prior surgery or the diagnosis of a syndromic condition impacting on the face. The study cohort of 3D images consisted of 128 males and 231 females. The male and female cohorts were treated as separate populations. Subjects were instructed to display a neutral facial expression in their natural head position whilst sitting upright when scanned. The 3D facial images consisted of a spatially-dense set of 3D points connected to form a wireframe made of polygons completely representing the 3D facial form. The precision and repeatability of the 3dMDface™ (two pod) System used for scanning was previously tested and validated to be sub-millimetre (Aldridge et al. 2005).

Three individuals expressing different and severe facial asymmetries were imaged by the same camera system and image capture protocol: first, a pair of female monozygotic twins discordant for hemifacial microsomia; second, a person with a right hemi-mandibular hypertrophy presenting with a discrepancy in the right lower mandibular border, deviated chin and occlusalcant; and third a young adult male with a history of growth restriction subsequent to radiation therapy to the jaw in infancy.

AM, mapping and mirroring

An AM, which is essentially a spatially-dense extension of the facial anthropometric landmarks defined by Farkas (1981), was fitted to each 3D facial form as illustrated in Fig. 1. The mask (Fig. 1b), covering the facial area of interest, consisted of spatially-dense, uniformly sampled (equally distanced at ~2 mm), unpaired and non-symmetrical points on an averaged facial form calculated over the healthy young adults following a bootstrapping strategy given in Claes (2007; section 5.3). The AM was mapped onto the 3D facial forms equivalent to indicating conventional landmarks. The dense nature of the mask comprising ~10 000 points realistically precluded any manual indication of the points. Therefore, the need for an automated mapping strategy was mandated. The mapping strategy, which is based on the work of Chui & Rangarajan (2000, 2003), was developed and validated previously on faces (Claes, 2007; chapter 5). The mapping strategy is akin to fitting an elastic mask onto a solid facial statue through a geometry-driven mapping of geometrically or anatomically corresponding features onto each other. As an initial starting point the mask was roughly aligned (rotated and translated) based on an ordered and approximate

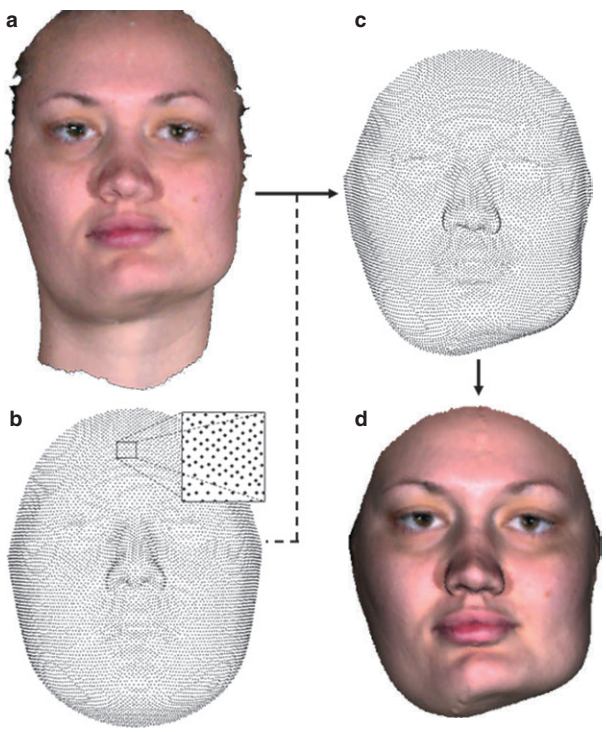

Fig. 1 Anthropometric mask (AM) mapping. (a) Example 3D facial scanner output of a young woman presenting with hemi-mandibular hypertrophy, rendered as a solid surface with texture. (b) The AM configuration consisting of a spatially-dense set of quasi-landmarks (zoom window). (c) The AM after mapping represents the same facial form as the scanner output (a), but using a configuration. (d) The configuration in (c) rendered as a solid surface with texture.

indication of the centres of right and left eye, nose tip, and right and left mouth corner. Then, by allowing iteratively more flexibility in the elasticity of the mask, initially larger, but gradually more local and more subtle, differences were accommodated. This process was continued until the mask (Fig. 1b) fitted the 3D facial form (Fig. 1a) and, as a result, defined the underlying facial structure using the standardized and predetermined template points (Fig. 1c,d). The resulting dense set of points, mapped in a quasi-anatomical manner, provided a dense set of corresponding quasi-landmark indications over all of the 363 3D facial forms. The key to understanding is that the same 3D facial form, previously represented using spatially-dense 3D points captured as a 3D image by the 3dMDface™ scanner (Fig. 1a), was now represented as a configuration of spatially-dense quasi-landmarks in Fig. 1c,d. This allowed asymmetry assessments from different individuals to be standardized in a spatially-dense way.

The construction of the 'mirror' configuration is illustrated in Fig. 2 following two different quasi-landmarks (one black and one white). First, the sign of the *x*-coordinate of the quasi-landmarks in the configuration was changed (Fig. 2a). Although this choice is completely arbitrary, it can be seen here that the *x*-axis represented an approximate left–right anatomical axis resulting from participants sitting upright in their natural head position during scanning. The resulting reflected configuration (Fig. 2a, right) correctly represented the mirror facial form, but lost compatibility with the original configuration. To re-establish homology, the original configuration was subsequently mapped onto the mirror facial form (represented using the reflected configuration) in the exact same way the AM had been mapped onto the original facial form previously. The starting point (based on the eye centres, nose tip and mouth corners) used here ensured a correct mapping of the original configuration independent from the reflection plane used. Doing so created a mirror configuration in Fig. 2b, not only representing the same mirror facial form as in Fig. 2a (right), but also being compatible with the original configuration in Fig. 2a (left) or Fig. 1c,d. It was now ready to be superimposed onto the original configuration as explained in the next section.

Note that the starting point needed before the mapping of the original configuration onto the mirror 3D facial form can be created by relabelling reflected eye, mouth and nose landmark indications on the original configuration following Klingenberg's protocol. The relabelling step for the spatially-dense quasi-landmarks is replaced here with a re-indication (mapping). Alternatively, one could opt to start from an AM consisting of a (L–R) paired vs. unpaired spatially-dense landmark configuration, changing the indication back into the possibility of a relabelling. Therefore in both scenarios, the proposed protocol can be seen as a spatially-dense extension of Klingenberg's protocol.

Robust superimposition

For a spatial assessment of asymmetry (Fig. 2c) to be made, a robust superimposition to eliminate orientation and position differences of the original and mirror configuration is required. When confronted with severe asymmetries, it is important to perform the superimposition based only on quasi-landmarks that are located in symmetrical areas of the face (e.g. the black

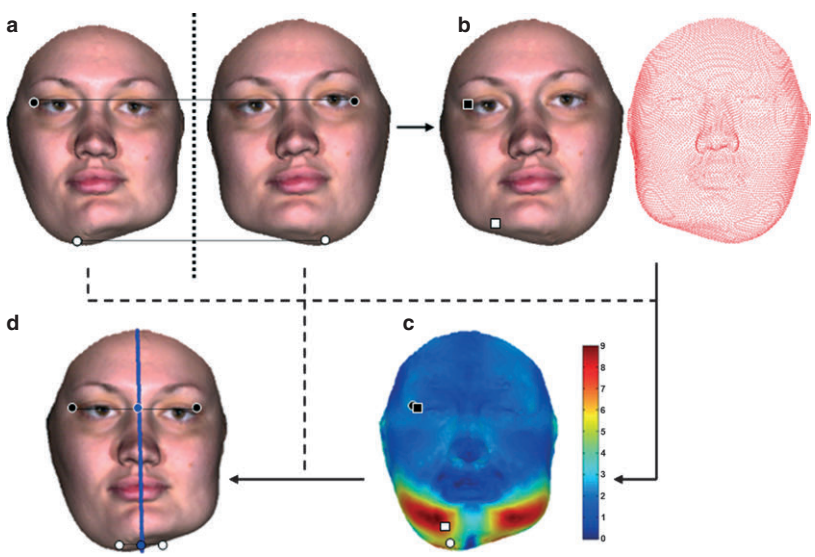

Fig. 2 Protocol steps. (a) Step 1: reflection (right) of original configuration (left) with two unpaired quasi-landmark examples located in different regions of the face. The white quasi-landmark is located in a symmetrical region of the face, whilst the black quasi-landmark is located in a highly asymmetrical region of the face. The reflected quasi-landmarks lost their homology with the original ones. (b) Step 2: re-indication of quasi-landmarks onto the mirror facial form. The indicated square white/black quasi-landmark is now homologous to the spherical white/black quasi-landmark of the original configuration in (a) (left). (c) Step 3: robust superimposition and spatial assessment of asymmetry. Homologous quasi-landmarks are robustly superimposed. As a result, a better fit is obtained for the white quasi-landmark compared with the black quasi-landmark. The remaining distances of all quasi-landmarks are visualized using a colour map ranging from 0 mm (blue) to 9 mm (red). (d) Step 4: midpoint construction and midsagittal plane estimation. Applying the estimated superimposition parameters from the previous step to the reflected configuration in (a) (right) allows the construction of the midpoints for each quasi-landmark as the average of the landmark itself and its aligned reflected landmark. A robust estimate of the midsagittal plane can be obtained by fitting a plane through the midpoints.

quasi-landmark throughout Fig. 2). A robust strategy found in Van Leemput et al.(2001) was adapted for superimposition purposes in Claes (2007; chapter 3 and app. B). The result is an iterative two-step optimization procedure, illustrated in Fig. 3, whereby both the current superimposition and the asymmetrical area estimations are alternated until no more change is observed in either one.

The detection/estimation of the asymmetrical areas was based on using the concept of outliers. These are data samples that are considered atypical to the majority of the data comprising the inliers. In this scenario, outliers/inliers are considered quasi-landmarks located in the asymmetrical/symmetrical areas of the face illustrated as white/black regions in Fig. 3. A statistically relevant and meaningful estimation of outliers vs. inliers was defined, and is explained in more detail in Appendix I. The result was a confidence value for each quasi-landmark, reflecting the confidence of such a point being an inlier (value closer or equal to 1, symmetrical) or an outlier (value closer or equal to 0, asymmetrical). The confidence values were subsequently used as weights to update the superimposition in a weighted least-squares manner until no more change was observed. For the sake of simplicity in the further analysis we will use the 1-complement (one minus the confidence value), or atypicality value, as a measure for the presence of local asymmetry.

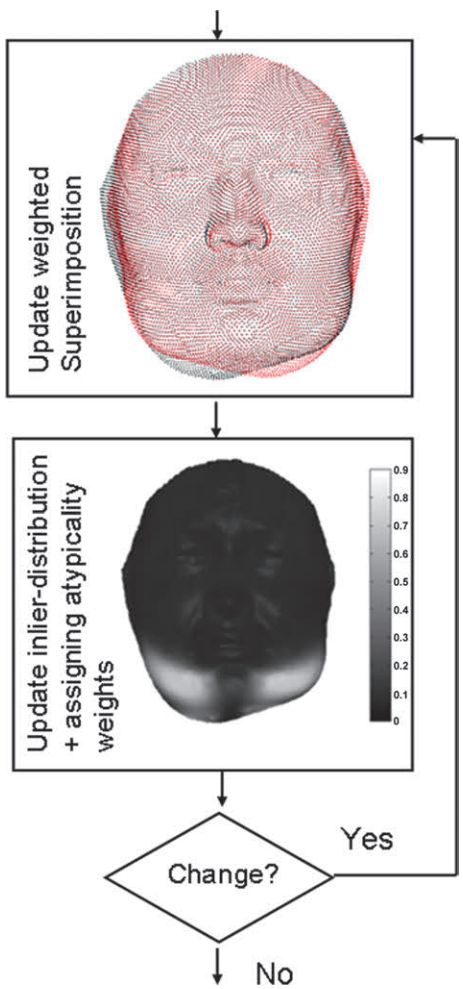

Fig. 3 A two-step robust superimposition: both superimposition and atypicality weights are iteratively updated until no more change in either is observed.

Midpoint construction and midsagittal plane estimation

An estimate of the midsagittal plane was also of interest, and was generated as a byproduct of the protocol as follows. Firstly, after the robust superimposition the same alignment parameters were applied to the reflected (but not homologous!) configuration (Fig. 2a, right). Subsequently, the midpoints for each quasi-landmark were constructed as the average location of a quasi-landmark and its reflected and aligned copy (Fig. 2d). Finally, an estimate of the midsagittal plane was then obtained by fitting a plane through all the midpoints.

Scoring, analysis and visualization of found asymmetries

The magnitude of spatial discrepancies between corresponding quasi-landmarks in the original and mirror configurations after superimposition was calculated. This was visualized as a colour map projected onto the original configuration (Fig. 2c). This 'distance map' was summarized by a RMSE, which incorporates both the variance and bias (average) of the asymmetry as an error in mm. These outputs reflect the localized and overall degree of asymmetry, respectively, found in an individual.

The atypicality values of the quasi-landmarks after superimposition could be visualized as a grey map projected onto the original configuration. An example of an atypicality map is shown in Fig. 3 and highlights regions similar to those in the distance map in Fig. 2c, hence making it redundant in terms of visual feedback and so is omitted in further illustrations. However, the colour (grey) mapping and scale are different and informative. Here the grey-value indicates whether a quasi-landmark is considered asymmetrical or not [more like a 'yes' (white) or 'no' (black) situation]. The atypicality map was summarized as the average of all atypicality values, which is similar to the amount of 'yes' situations divided by the total amount. This reflects localized and overall (in percentage) relative significant asymmetry (RSA), respectively, in an individual, and can be used as a means to quantify the proportion of the asymmetry on the facial form.

Asymmetry described here was computed for the healthy cohort of young adults in the study to establish typical population reference indices. First, differences of overall RMSE and RSA between males and females were tested using a Wilcoxon rank sum test to avoid underlying distribution assumptions. It was expected that the distribution would be non-Gaussian, because both RMSE and RSA are strictly positive. Second, homologous quasi-landmark asymmetry atypicality values over multiple individuals were averaged and visualized as a colour map projected onto the average male and female consensus configurations (from a generalized Procrustes superimposition). This reflected the occurrence of localized RSA in a normal population. Third, homologous quasi-landmark asymmetry magnitudes in the distance maps over multiple individuals were summarized using the RMSE and visualized as a colour map projected onto the average male and female consensus configurations. This reflected combined localized variance and bias (average) in facial asymmetry found in a typical population.

Results

Descriptive statistics of overall asymmetry scores for the typical growth of male and female populations are given in Table 1. The Wilcoxon rank sum test showed that RMSE scores for females (median = 0.94) differed significantly from males (median = 1.05; $W = 37\,287$, $z = -4.558$, $P < 0.001$, $r = -0.24$). Similarly, a Wilcoxon rank sum test showed that RSA scores for females (median = 9.39) differed significantly from males (median = 10.05; $W = 36\,951$, $z = -4.915$, $P < 0.001$, $r = -0.26$).

Table 1 Descriptive statistics of overall asymmetry scores in a young healthy population with typical growth patterns; the RMSE and RSA scores for males and females separately.

	Males		Females	
	RMSE (mm)	RSA (%)	RMSE (mm)	RSA (%)
Mean	1.09	10.20	0.94	9.53
Median	1.05	10.05	0.94	9.39
Std. deviation	0.28	1.17	0.225	1.21
Minimum	0.60	8.35	0.48	6.67
Maximum	1.96	13.99	1.71	13.15
Range	1.35	5.63	1.23	6.49
Interquartile range	0.32	1.64	0.31	1.71

RMSE, root-mean-squared-error; RSA, relative significant asymmetry.

The localized asymmetry assessment summaries for typical growth of male and female populations are given in Fig. 4. In general it was observed that both higher occurrence and degree of asymmetry for both males and females were localized around the mid and lower facial regions. The percentage of localized RSA illustrated that for females the highest frequency (25%) of RSA occurred around the angles of the mouth. For males, frequency of RSA around the angles of the mouth is also high, but lower. The localized RMSE values for males were consistently larger than those for the females. This outcome is reflective of male faces inherently being larger with more prominent features than in females. The asymmetry of the male compared with the female nose also involved the nasal bridge. Other differences exist and can be seen in Fig. 4. Male and female facial structures differ, which is more commonly known as sexual dimorphism in facial form (Ferrario et al. 1993; Thornhill & Gangestad, 2006). Hence, the localized as well as the previously detected overall differences in asymmetry were not unexpected.

The asymmetry assessments for the three subjects with diagnosed asymmetrical growth patterns, including one unaffected twin sibling, are depicted in Fig. 2 for one case and in Fig. 5 for all the other cases. In all these subjects the midpoints were plotted to visually confirm the robustness of the protocol. Overall two-valued asymmetry scores are reported as (RMSE, RSA). The first example shown in Fig. 5 (top row) is the unaffected twin (1.19 mm, 10.7%) of the young woman with hemifacial microsomia (Fig. 5, second

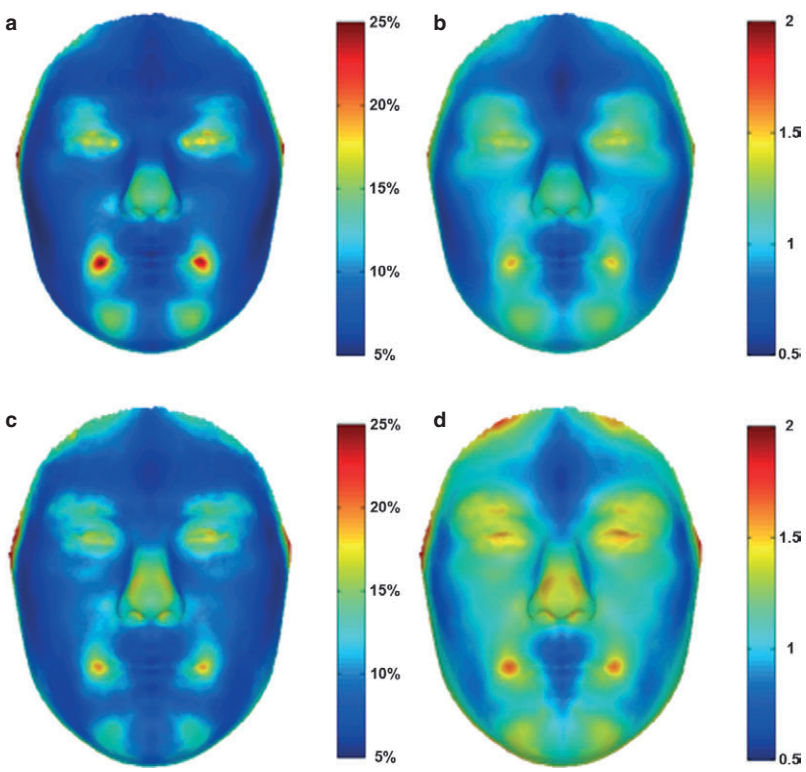

Fig. 4 Localized asymmetry in a young healthy population. (a,c) Occurrence of RSA for females and males, respectively, visualized using a colour range from 5% (dark blue) to 25% (dark red). (b,d) RMSE values per quasi-landmark in the AM for females and males, respectively, visualized using a colour range from 0.5 mm (dark blue) to 2 mm (dark red).

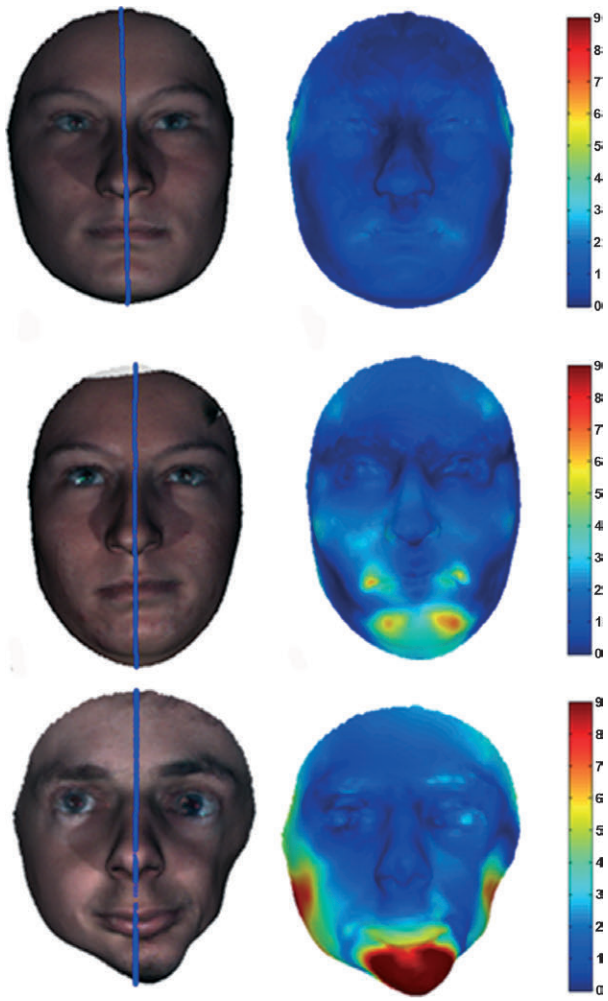

Fig. 5 Midpoints extraction (blue line) and distance maps of subjects with growth disorders resulting in facial asymmetries. Top row: twin sibling of a young woman with hemifacial microsomia (second row). Bottom row: young man with severe growth restriction subsequent to radiation therapy in childhood. Distance maps, depicting local spatial discrepancy magnitudes between the original and mirror configuration after superimposition, visualized using a colour range from 0 mm (blue) to 9 mm (red).

row). The area of the face affected by asymmetry is low with minimal magnitude and well within typical range. For the affected twin (1.96 mm, 13.9%; Fig. 5, second row) the deviation of the mandible and asymmetry of the malar region associated with growth restriction is readily discernable with the additional information on the magnitude of the differences being outside the typical range. The young woman used for illustration purposes throughout the Materials and methods section presented with hemi-mandible hypertrophy, resulting in overgrowth of the right hemi-mandible and subsequent displacement of the mandible thereby generating asymmetries to the mandible, lip, nose and alar base. All these are detected (Fig. 2) and quantified by the protocol (3.24 mm, 18.8%) with numeric feedback that is correctly higher than the previous example. The asymmetries in a young male subject with severe growth restriction as a consequence of radiation therapy are depicted in Fig. 5 (bottom row). The two-valued asymmetry score was (4.21 mm, 23.1%), which clearly indicated both the biggest degree of abnormality and extent of asymmetry compared with the other examples, and which was also well outside the typical range. It should be noted that the distance map appeared asymmetrical in this case because of the underlying asymmetry between the left and right part of the configuration used to display them. It should also be noted that the atypicality maps are not depicted here as their contribution to the visual feedback and localization of the asymmetry is similar to the distance maps provided.

Discussion

Bilateral facial symmetry was defined with respect to the midsagittal plane. Because asymmetry involves absence or violation of symmetry, the unambiguous definition of this midsagittal plane becomes problematic. For example, Ferrario et al. (1994) observed that overall the symmetry plane is not located in facial midline landmarks. When confronted with cases of severe asymmetry, a plane that divides the face into perfectly left and right parts does not exist or is not a plane anymore. The last case in Fig. 5 is a good example where the chin is displaced entirely to one side of the face. Asymmetry assessment protocols requiring a predefined midsagittal plane are not useful here.

In this study the aim was to define a protocol useful in the investigation of facial asymmetry according to an unknown midsagittal plane in both normal and disordered growth patterns. Statistics of typical asymmetry are important as a reference. Individual asymmetry can then be compared with this reference under the condition that the same assessment protocol is applied. The protocol defined by Klingenberg et al. (2002) for object symmetry was used as a methodological foundation upon which modifications were made in the use of an AM and a robust superimposition utilizing a robust weighted least-squares approach. The AM facilitated automated use of spatially-dense quasi-landmarks over a range of subjects. Subtle and severe asymmetries in discrete regions of the face could be detected, quantified and statistically described. Unlike the original least-squares solution, the weighted least-squares superimposition described here was robust against increasing asymmetry. This allowed the use of the protocol without modification or prior knowledge for the detection and measurement of a wide range of both mild and severe asymmetries. Furthermore, it provided a novel measurement of RSA in addition to the commonly known degree of symmetry.

The biggest challenge when working with spatially-dense facial representations is to obtain compatible configurations beyond 'true' landmarks, which have a name and are uniquely defined. Following the original and broad definition of semi-landmarks, that is, points that do not have

names, but that correspond across all cases of similar but variable shapes (Bookstein, 1997; Andresen et al. 2000), quasi- and semi-landmarks are essentially the same. The challenge for both types of landmarks is to find a mapping function that establishes one-to-one correspondences and therefore generates compatible configurations. Despite the strong similarity, we would like to make a slight distinction between quasi- and semi-landmarks. In the life-sciences, semi-landmarks became more narrowly defined relative to other features based on selection criteria (Zelditch et al. 2004; e.g. relative to true landmarks). In contrast, quasi-landmarks are simply defined as spatially-dense points sampled on a continuous surface independently from other features. The difference in definition results in a different mapping strategy. Indeed, semi-landmarks can be pre-assigned on a set of shapes, because of their relative definition (Bookstein, 1997). Subsequently, semi-landmarks are allowed to slide along tangent directions to the surface (Gunz et al. 2005). The sliding is most frequently done by minimizing either bending energy or Procrustes distance (Perez et al. 2006). The result is that the pre-assigned semi-landmarks are 'relaxed' or corrected to remove any tangential variation following the criterion of choice. Stated differently, through the relative definition of semi-landmarks, one-to-one correspondences are known and are subject to correction using, for example, sliding methods. In contrast, quasi-landmarks are not pre-assigned on all shapes. Therefore, one-to-one correspondences from quasi-landmarks are not known beforehand and have to be found in a slightly different way.

Finding a mapping function between two or more 3D shapes is commonly known as '3D registration' in computer science. The goal of a registration algorithm is to find the geometrical relationship (one-to-one correspondences) between shapes following a predefined transformation model. As such, the previously mentioned sliding methods for semi-landmarks are registration algorithms. A popular registration algorithm, without pre-assigned correspondences, is the Iterative Closest Point (ICP) procedure (Besl & McKay, 1992). ICP is an iterative two-step algorithm in which candidate correspondences and transformation model parameters are updated until no more change in both is observed. For every point on the reference shape the candidate correspondence is defined as being the closest point on the target shape (which is different to the sliding techniques). In its original form (Besl & McKay, 1992), ICP assumes a rigid transformation model (only translations and rotations are allowed). However, closest points under a rigid model do not guarantee structural correspondence. For example, the tip of a tip-tilted nose will most likely not correspond to the tip of a hooked nose, as another point might be closer in proximity. The rigid model does not account for local shape differences when searching for correspondences. This can still work for faces that are highly similar, but it rapidly becomes incorrect when dealing with realistic facial variation. As a result, correspondences under a rigid model are at best approximate because they do not directly quantify spatial differences between homologous structures of interest. One interesting approach to deal with this problem is to use limited landmark information to deform [using Thin-Plate-Splines (TPS)] one facial form closer to the other and then taking the closest points as corresponding points (Hutton et al. 2003; Hammond et al. 2004). In doing so, the anatomical knowledge of 'true' landmarks is roughly interpolated for points in-between them and the mapping gains anatomical relevance. The TPS deformation is a non-rigid instead of a rigid transformation model, and allows compensating for local shape differences when searching for correspondences. The same TPS non-rigid model is also used in the sliding method based on minimal bending energy as well as in the mapping strategy used in this study. In contrast, however, the mapping strategy used here (Claes, 2007) is a non-rigid extension of the original ICP algorithm. Therefore, no pre-assigned correspondences or 'true' landmark information is required. Instead, iteratively more flexibility in the elasticity (bending energy) of the transformation model is allowed such that initially larger, but gradually more local and more subtle, differences are accommodated for when searching for correspondences. It is important to note that different registration algorithms can lead to a different result, which is demonstrated for sliding methods in (Perez et al. 2006).

In a superimposition framework for asymmetry assessment, an additional challenge is to obtain a compatible mirror configuration for the original configuration. When working with spatially-dense facial representations, different strategies exist. An alternative option to the mapping technique used here is to start from and re-label paired-landmarks as defined in the Klingenberg protocol. This requires a specially constructed and symmetrical AM. As mentioned in the Materials and methods, this is certainly a good substitute to be used in the protocol. Alternatively, one could manually indicate a limited set of corresponding landmarks onto both facial forms (Yu et al. 2009) and use only those to perform the superimposition. This is often provided as a tool in commercial 3D software, but introduces subjectivity and manual indication errors. Another alternative is to use the original ICP technique (Benz et al. 2002; Zhang et al. 2006; Tang et al. 2008) that combines the search for correspondences and superimposition (with the rigid model) into one algorithm. However, as previously mentioned, this does not guarantee left–right structural correspondence and results in a degrading of the plausibility of the assessment outcome. Finally, Hammond et al. (2008) reflect and re-label manually indicated landmarks on the original facial form to obtain their mirror versions. These are then used to deform (TPS) the mirror facial form similar to the original facial form. Subsequently, closest points are taken as corresponding points, as described previously (Hutton et al. 2003; Hammond et al. 2004).

The need for caution when using the original least-squares superimposition in the localization of severe asymmetry has already been mentioned by Klingenberg & McIntyre (1998). However, at the time, the known repeated median method that was used for resistant fitting (Rohlf & Slice, 1990), which is similar to robust superimposition, lacked proper mathematical underpinning and convergence behaviour compared with their original least-squares solution, and was therefore not advocated as an alternative. The requirement for robustness has been further acknowledged, and in a superimposition framework this has been addressed by eliminating the influence of severe asymmetries in different ways. One approach is to perform the superimposition on a predefined assumed symmetrical area of the face like the nose ridge in Tang et al.(2008), or by carefully indicated landmarks in Yu et al. (2009). Another strategy is to perform a least-squares superimposition of the mirror configuration first, then remove the asymmetrical parts according to a threshold and redo the superimposition using the remaining parts (Benz et al. 2002). Both these approaches try to achieve the same result as the robust superimposition proposed in this study. However, they introduce subjectivity by the obligation to manually select a region, landmarks or threshold of interest. In contrast, the proposed technique is an adaptive superimposition based on a meaningful statistical significance and outlier detection. Furthermore, the underlying math is very similar to the least-squares solution, but it is weighted. It is interesting to note that a superimposition-free, but also robust, approach to assess asymmetry is given in Prima et al. (2002) and Combès & Prima (2008), where the focus is directly on the midsagittal plane extraction. An algorithm is used to extract the symmetry plane that is not influenced by asymmetries as a robust estimate for the midsagittal plane around which the face can then be mirrored and compared.

A comparison of our findings with related work has to be done whilst taking into account study differences. Many studies on facial asymmetry have been conducted with in-house assessment protocols, databases and imaging modalities, which makes direct comparison between reports difficult. More importantly, most studies found in the literature concentrated on DA only, which is defined on the level of a population. For example, a lot of work has focussed on determining the dominant side of the face with no real consensus emerging, as discussed by Ercan et al. (2008). In contrast, this study aimed to relate an individual's asymmetry as a component of typical variation and/or as a consequence of disordered growth. As stated in Ferrario et al. (1994), this requires a completely different approach. Therefore, a total score of asymmetry per individual as suggested by Klingenberg & McIntyre (1998) is used. In this work no separation of the total scores into DA and FA over the typical population is performed. This makes the comparison of males vs. females in this study, for example, difficult to relate to other studies.

More than a decade ago, Ferrario et al. (1995) stated that by adding more landmarks a better understanding of the facial form could be obtained. Here a spatially-dense assessment allowed highly localized comparisons of asymmetry between different regions in the face. In contrast, previous studies (Ferrario et al. 1994; Shaner et al. 2000; Haraguchi et al. 2002) were limited to the analysis and comparison of regions in the face described as upper, middle and lower third, with the latter two being more asymmetrical. Their studies, typically using prominent landmarks on facial features only, could not include the facial regions between the discrete features and were therefore limited in the comparisons they made. Depiction of asymmetry as in Fig. 5 is more complete and allows for more detailed comparisons. Overall, in general terms, the asymmetries detected in the presented study for both males and females are not dissimilar to those of previous studies, where the middle and lower thirds of the face express more asymmetry.

Hajeer et al. (2005) used a comparable protocol to ours in 3D. This is very similar to the Klingenberg protocol but based only on 19 landmarks, the majority (11) of which are indicated on the midline. The protocol was applied to three groups before and after different types of orthognathic surgery. A mean-squared-error (MSE) score was used as a measure of asymmetry degree comparable to our RMSE score by simply taking the root. Similar to this study, their asymmetry scores were not normally distributed, and hence the median, range and inter-quartiles were reported, which after surgery were still higher ($\sim$ 0.5) than the values of our typical population provided in Table 1. However, for comparison purposes, their scores were in dimensionless units and not in mm, as all the faces were scaled to a common size (which was not given) before assessment of asymmetry. Furthermore, all the 19 landmarks used are located in the more asymmetrical areas of the face seen in Fig. 5, influencing their median score over all the landmarks towards a higher value. Hence, it is hard to conclude whether or not the reported asymmetry scores after surgery in their study deviate from our typical values. Interestingly, Hajeer et al. (2005) advised the need for caution to be taken into account when comparing preoperative with postoperative asymmetry scores. This was because the achievement of 'best-fit', using a least-squares superimposition was influenced by severe asymmetries in the preoperative situations. This resulted in spurious changes and reduction of the asymmetry postoperatively in regions not affected by the surgical intervention.

Another comparable, but spatially-dense, protocol was developed by Benz et al. (2002) and used by Hartmann et al. (2007) on a single male. The protocol uses a spatially-dense facial representation in combination with the ICP technique for the superimposition of the original configuration with the mirror configuration. The mean absolute distance was used as a score for the degree of asymmetry, and a value of 0.738 mm was reported. This score is similar to the presented RMSE score and is on the lower end of our normal

population range. However, this can be explained, as the definition of closest points used in the ICP technique typically leads to the underestimation of the real error or difference, as noted by Mollemans et al. (2007).

A final comparable protocol is given in Hammond et al. (2008). This protocol also uses spatially-dense facial representations but now in combination with a TPS non-rigid model based on indicated landmarks as previously explained. An asymmetry index for an individual is computed as the Euclidean distance between dense surface model representations of the original facial form and its mirror form.

The RSA score proposed here, being a byproduct of the adapted weighted least-squares superimposition, is intended to complete the overall assessment of asymmetry. The higher this score the more significant asymmetry is observed, and this can be used to indicate the proportion of abnormal asymmetry as shown in the cases used in this study. The RSA is a relative score such that a small increase in value can already represent a big increase in the extent of the asymmetry. This makes the score independent of differences in scale, but not always straightforward to interpret. A population-based overall and local significant asymmetry score also expressed as a percentage should be defined according to typical asymmetry for which the statistics provided here can be useful in future work.

By focusing on the asymmetry measured in a particular individual, it is possible to relate that person with regard to typical ranges of asymmetry measured across many individuals. This is highly relevant in clinical practice, for example. Furthermore, the robust estimate of the midsagittal plane can also prove to be useful in order to fine tune any diagnosis and treatment planning. Conversely, population biologists are more interested in the separation of the asymmetry into DA and FA. Although not reported here, a measure of DA, for example obtained by analysing differences on average facial forms (as in Hammond et al. (2008)) is perfectly possible. In future work, populations (instead of just an individual) with similar disordered growth are to be collected. Then these can be assessed in comparison with normal populations both in terms of DA and FA separately. The protocol can be expanded for this along the lines of the original Klingenberg protocol (Klingenberg et al. 2002). As suggested by one of the reviewers, another interesting future study is the effect of facial expression on asymmetry assessments.

Conclusion

Spatially-dense facial asymmetry assessment calibrated against typical growth variation is important in the diagnosis of abnormal growth patterns. A robust protocol applicable without any modification or prior knowledge in both common and disordered situations is therefore important, and has been proposed. Our scoring of asymmetry with accompanying reference statistics also provides the basis for future studies.

Acknowledgements

This work is supported by the Australian Research Council (ARC) grant DP0772650 and by funding provided by the Princess Margaret Hospital Foundation in Perth, Western Australia. Technical stages of the work were supported by the Flemish Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT Vlaanderen), the Research Program of the Fund for Scientific Research - Flanders (Belgium) (FWO) and the Research Fund K.U. Leuven. The authors would also like to thank Dr Paul Sillifant, Dr David Gillett and Dr James Savundra for providing the clinical data used in this study. We also like to thank all the participants and their permission to use their images.

Ethics approval

(1) *The Characterization of 3D Facial Profile in Young Adult Western Australians* was granted from the Princess Margaret Hospital for Children (PMH) ethics committee (PMHEC 1443/EP) in Perth, WA, Australia. (2) *Establishment of Identity from Quantitative Analysis of Facial Characteristics* (*digital 3D facial modelling*) was granted from the University of Melbourne, human research ethics committee (HESC 050550.1) in Melbourne, VIC 3010, Australia.

A graphical user interface for the robust superimposition of two compatible configurations together with the anthropometric mask (AM) used in this study is made available at http://mirc.uzleuven.be/downloads/index.php. Currently, the software used to map the AM onto given facial forms cannot be distributed as it makes use of other commercially protected software.

Author contributions

P.C. and D.V. developed the technical details behind the anthropometric mask mapping and robust weighted least-square superimposition. P.C. re-worked and applied the fundamental techniques to obtain asymmetry assessments. M.W. collected the clinical and required facial data to build up population normative statistics. P.C., D.V. and J.C. situated everything within current morphometry and techniques. M.W. and J.C. supported the anatomical relevance of the work. P.C. wrote the manuscript, with input and revisions from M.W., J.C. and D.V.

References

Aeria G, Claes P, Vandermeulen D, et al. (2010) Targeting specific facial variation for different identification tasks. *Forensic Sci Int* **201**, 118–124.

Aldridge K, Boyadjiev SA, Capone GT, et al. (2005) Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am J Med Genet A* **138 A**, 247–253.

Andresen PR, Bookstein FL, Conradsen K, et al. (2000) Surface-bounded growth modeling applied to human mandibles. *IEEE Trans Med Imaging* **19**, 1053–1063.

Baudouin JY, Tiberghien G (2004) Symmetry, averageness, and feature size in the facial attractiveness of women. *Acta Psychol (Amst)* **117**, 313–332.

Benz M, Laboureux X, Maier T, et al. (2002) The symmetry of faces. In: *Vision, Modeling, and Visualization*. (eds Greiner G, Niemann H, Ertl T, et al.), pp. 332–339. Amsterdam: IOS Press.

Besl PJ, McKay ND (1992) A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* **14**, 239–256.

Bookstein FL (1997) Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med Image Anal* **1**, 225–243.

Chui H, Rangarajan A (2000) A new algorithm for non-rigid point matching. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (ed. the IEEE Computer Society), pp. 44–51. Hilton Head Island, SC, USA: The Printing House.

Chui H, Rangarajan A (2003) A new point matching algorithm for non-rigid registration. *Comput Vis Image Underst* **89**, 114–141.

Claes P (2007) A robust statistical surface registration framework using implicit function representations: application in craniofacial reconstruction. In: *Faculteit ingenieurswetenschappen, department Elektrotechniek, afdeling PSI*. K.U. Leuven, Belgium: Leuven.

Cole T (2001) Further applications of EDMA. In: *An Invariant Approach to the Statistical Analysis of Shapes*. (eds Lele SR, Richtsmeier JT), pp. 263–284. London: Chapman & Hall.

Combès B, Prima S (2008) New algorithms to map asymmetries of 3D surfaces. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*. (eds Metaxas D, Axel L, Fichtinger G, et al.), pp. 17–25. Berlin: Springer.

Dryden IL, Mardia KV (1998) *Statistical Shape Analysis*. Chichester, England: John Wiley.

Ercan I, Ozdemir ST, Etoz A, et al. (2008) Facial asymmetry in young healthy subjects evaluated by statistical shape analysis. *J Anat* **213**, 663–669.

Farkas LG (1981) *Anthropometry of the Head and Face in Medicine*. New York: Elsevier.

Ferrario VF, Sforza C, Pizzini G, et al. (1993) Sexual dimorphism in the human face assessed by euclidean distance matrix analysis. *J Anat* **183**, 593–600.

Ferrario VF, Sforza C, Poggio CE, et al. (1994) Distance from symmetry: a three-dimensional evaluation of facial asymmetry. *J Oral Maxillofac Surg* **52**, 1126–1132.

Ferrario VF, Sforza C, Miani A, et al. (1995) A three-dimensional evaluation of human facial asymmetry. *J Anat* **186**, 103–110.

Ferrario VF, Sforza C, Ciusa V, et al. (2001) The effect of sex and age on facial asymmetry in healthy subjects: a cross-sectional study from adolescence to mid-adulthood. *J Oral Maxillofac Surg* **59**, 382–388.

Goodall C (1991) Procrustes methods in the statistical analysis of shape. *J R Stat Soc Series B Methodol* **53**, 285–339.

Gunz P, Mitteroecker P, Bookstein FL (2005) Semilandmarks in three dimensions. In: *Modern Morphometrics in Physical Anthropology* (ed. Slice D), pp. 73–98. New York: Kluwer Academic.

Hajeer MY, Mao Z, Millett DT, et al. (2005) A new three-dimensional method of assessing facial volumetric changes after orthognathic treatment. *Cleft Palate Craniofac J* **42**, 113–120.

Hamada H, Meno C, Watanabe D, et al. (2002) Establishment of vertebrate left-right asymmetry. *Nat Rev Gen* **3**, 103–113.

Hammond P, Hutton TJ, Allanson JE, et al. (2004) 3D analysis of facial morphology. *Am J Hum Genet* **126A**, 339–348.

Hammond P, Forster-Gibson C, Chudley AE, et al. (2008) Face-brain asymmetry in autism spectrum disorders. *Mol Psychiatry* **13**, 614–623.

Haraguchi S, Takada K, Yasuda Y (2002) Facial asymmetry in subjects with skeletal class III deformity. *Angle Orthod* **72**, 28–35.

Hartmann J, Meyer-Marcotty P, Benz M, et al. (2007) Reliability of a method for computing facial symmetry plane and degree of asymmetry based on 3D-data. *J Orofac Orthop* **68**, 477–490.

Hutton TJ, Buxton BF, Hammond P, et al. (2003) Estimating average growth trajectories in shape-space using kernel smoothing. *IEEE Trans Med Imaging* **22**, 747–753.

Klingenberg CP, McIntyre GS (1998) Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. *Evolution* **52**, 1363–1375.

Klingenberg CP, Barluenga M, Meyer A (2002) Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution* **56**, 1909–1920.

McIntyre GS, Mossey PA (2002) Asymmetry of the parental craniofacial skeleton on orofacial clefting. *J Orthod* **29**, 299–305.

Mollemans W, Schutyser F, Nadjmi N, et al. (2007) Predicting soft tissue deformations for a maxillofacial surgery planning system: from computational strategies to a complete clinical validation. *Med Image Anal* **11**, 282–301.

Palmer AR (1994) Fluctuating asymmetry: a primer. In: *Developmental Instability: its Origins and Implications* (ed. Markov T), pp. 335–364. Dordrecht, The Netherlands: Kluwer.

Palmer AR, Strobeck C (1986) Fluctuating asymmetry: measurement, analysis, patterns. *Annu Rev Ecol Syst* **17**, 391–421.

Perez SI, Bernal V, Gonzalez PN (2006) Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *J Anat* **208**, 769–784.

Poswillo DE (1989) Myths, masks and mechanisms of facial deformity. *Eur J Orthod* **11**, 1–9.

Prima S, Ourselin S, Ayache N (2002) Computation of the mid-sagittal plane in 3-d brain images. *IEEE Trans Med Imaging* **21**, 122–138.

Richtsmeier JT, Deleon VB, Lele SR (2002) The promise of geometric morphometrics. *Am J Phys Anthropol* **45**, 63–91.

Richtsmeier JT, Cole T, Lele SR (2005) An invariant approach to the study of fluctuating asymmetry: developmental instability in a mouse model for down syndrome. In: *Modern Morphometrics in Physical Anthropology*. (ed. Slice D), pp. 187–212. New York: Kluwer.

Rohlf F, Slice D (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst Zool* **39**, 40–59.

Rossi M, Ribeiro E, Smith R (2003) Craniofacial asymmetry in development: an anatomical study. *Angle Orthod* **73**, 381–385.

Shaner DJ, Peterson AE, Beattie OB, et al. (2000) Assessment of facial asymmetry in medically normal and syndrome-affected individuals by analysis of landmarks and measurements. *Am J Med Genet* **93**, 143–154.

Slice D (2007) Geometric morphometrics. *Annu Rev Anthropol* **36**, 261–281.

Smeets D, Claes P, Vandermeulen D, et al. (2010) Objective 3D face recognition: evolution, approaches and challenges. *Forensic Sci Int* **201**, 125–132.

Tang XM, Chen JS, Moon YS (2008) Accurate 3D face registration based on the symmetry plane analysis on nose regions. In: *16th European Signal Processing Conference (EUSIPCO)*. Lausanne, Switzerland.

Theobald DL, Wuttke DS (2006) Empirical Bayes hierarchical models for regularizing maximum likelihood estimation in the matrix Gaussian Procrustes problem. *Proc Natl Acad Sci USA* **103**, 18 521–18 527.

Thomas CDL (2005) Three-dimensional quantification of facial shape. In: *Computer-Graphic Facial Reconstruction*. (eds Clement JG, Marks MK), pp. 55–78. London: Elsevier Academic Press.

Thornhill R, Gangestad SW (2006) Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and woman. *Evol Hum Behav* **27**, 131–144.

Van Leemput K, Maes F, Vandermeulen D, et al. (2001) Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Trans Med Imaging* **20**, 677–688.

Van Valen L (1962) Study of fluctuating asymmetry. *Evolution* **16**, 125–142.

Yu Z, Mu X, Feng S, et al. (2009) Flip-registration procedure of three-dimensional laser surface scanning images on quantitative evaluation of facial asymmetries. *J Craniofac Surg* **20**, 157–160.

Zaidel DW, Cohen JA (2005) The face, beauty, and symmetry: perceiving asymmetry in beautiful faces. *Int J Neurosci* **115**, 1165–1173.

Zelditch ML, Swiderski DL, Sheets HD, et al. (2004) *Geometric Morphometrics for Biologists: a Primer*. New York: Elsevier Academic Press.

Zhang L, Razdan A, Farin G, et al. (2006) 3d face authentication and recognition based in bilateral symmetry analysis. *J Vis Comput* **22**, 43–55.

Appendix I

A statistically relevant and meaningful estimation of outliers vs. inliers was defined as follows:

- In a Procrustes Maximum Likelihood (ML) estimation of the superimposition a Gaussian perturbation model is assumed (Goodall, 1991; Theobald & Wuttke, 2006). Inliers are quasi-landmarks following this assumption, therefore allowing updating the superimposition. Furthermore, they also allow estimating any unknown perturbation model parameters if needed. Under the simplifying assumption that the vectorial differences E between homologous (original and mirrored) landmarks are isotropic, independently and identically distributed the standard deviation of a Normal distribution with zero mean was estimated. This constitutes the inlier-distribution $f(E) = N(E, 0, \sigma)$.
- Outliers on the other hand are atypical compared with inliers and hence cannot be explained by the inlier-distribution. In other words, they violate the Gaussian perturbation model assumption. This results in an unbounded bias of outliers in the original Procrustes ML estimation, known as the Pinocchio effect. Instead, following the work of Van Leemput et al. (2001), outliers are considered a fraction $(1 - P)$ with $0 < P < 1$ of the data that are drawn from a rejection class that is assumed to be uniformly distributed. This constitutes the outlier-distribution $f'(E) = 1/L$ bounding the contribution of outliers to a fixed value.
- The confidence of a quasi-landmark to be an inlier is then defined as the probability of being an inlier divided by the total probability: $W = f(E)P/(f(E)P + f'(E)(1 - P))$.
- Again following the work of Van Leemput et al. (2001), the level L of the outlier-distribution and the fraction of the data $(1 - P)$ with $0 < P < 1$ expected to be outliers are combined into a single parameter λ. Subsequently, λ itself is re-parameterized using K and is made dependent on the standard deviation of the inlier-distribution. The final parameter K can be regarded as the choice of a statistical significance level.

The result was a confidence value W for each quasi-landmark, reflecting the confidence of such a point being an inlier (value closer or equal to 1, symmetrical) or an outlier (value closer or equal to 0, asymmetrical). A quasi-landmark in the original configuration was gradually considered an outlier if the distance to its homologous quasi-landmark on the mirror configuration approached or became more than twice the standard deviation of the inlier-distribution. This corresponds to $K = 2$ or a significance value of $P = 0.05$. The confidence values were subsequently used as weights to update the superimposition in a weighted least-squares manner until no more change was observed. To conclude, the robust superimposition procedure can be written in pseudo-code as:

E_j = difference between homologous landmarks.

W_j = landmark weight or confidence value.

L = number of landmarks and j = landmark index.

$K = 2$, significance level.

Initialize

All landmarks are inliers.

$W_j = 1$ for $j = 1, \ldots, L$.

Start

1) *Update weighted-LS superimposition.*

$$\min \sum_{j=1}^{L} W_j E_j^2$$

2) *Update asymmetry estimation.*

A) Update inlier-distribution parameters.

$$\sigma = \sqrt{\frac{\sum_j^L W_j E_j^2}{\sum_j^L W_j}}$$

B) Update outlier-distribution parameters.

$$\lambda = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2}K^2\right)$$

C) Update confidence values.

$$W_j = \frac{N(E_j, 0, \sigma)}{N(E_j, 0, \sigma) + \lambda}$$

Change?

Yes, re-iterate.

No, STOP.