



RESEARCH

Open Access

# Dysmorphometrics: the modelling of morphological abnormalities

Peter Claes<sup>1,3\*</sup>, Katleen Daniels<sup>1</sup>, Mark Walters<sup>2</sup>, John Clement<sup>3</sup>, Dirk Vandermeulen<sup>1</sup> and Paul Suetens<sup>1,4</sup>

\* Correspondence: Peter.

Claes@esat.kuleuven.be

<sup>1</sup>Department of Electrical Engineering - processing of speech and images - medical image computing, KU Leuven, Herestraat 49 Bus 7003, Leuven, 3000, Belgium

Full list of author information is available at the end of the article

## Abstract

**Background:** The study of typical morphological variations using quantitative, morphometric descriptors has always interested biologists in general. However, unusual examples of form, such as abnormalities are often encountered in biomedical sciences. Despite the long history of morphometrics, the means to identify and quantify such unusual form differences remains limited.

**Methods:** A theoretical concept, called dysmorphometrics, is introduced augmenting current geometric morphometrics with a focus on identifying and modelling form abnormalities. Dysmorphometrics applies the paradigm of detecting form differences as outliers compared to an appropriate norm. To achieve this, the likelihood formulation of landmark superimpositions is extended with outlier processes explicitly introducing a latent variable coding for abnormalities. A tractable solution to this augmented superimposition problem is obtained using Expectation-Maximization. The topography of detected abnormalities is encoded in a dysmorphogram.

**Results:** We demonstrate the use of dysmorphometrics to measure abrupt changes in time, asymmetry and discordancy in a set of human faces presenting with facial abnormalities.

**Conclusion:** The results clearly illustrate the unique power to reveal unusual form differences given only normative data with clear applications in both biomedical practice & research.

**Keywords:** Geometric Morphometrics, Dysmorphometrics, Procrustes ML-estimator, Pinocchio effect, robust statistics, abnormalities, outlier-processes

## Background

Morphometrics involves the measurement of morphology based on quantitative descriptions [1]. Different definitions of form exist but the most commonly adopted is that form is defined as size and shape independent of position and orientation [2]. Morphometric methods are designed to measure form and variations in form and have been extensively used in evolutionary, developmental and systematic biology. In those contexts, they provide information on phylogenetic relationships and the evolutionary development of organisms [3]. They also allow for taxonomic discrimination of sampled populations to test whether these were drawn from different (sub)species or not [4]. However, in biomedical contexts unusual from instances such as abnormalities [5] are often encountered and of interest. For example abnormalities may include

deformations in form due to congenital malformation and/or environmental constraints as well as abrupt changes in form associated with traumatic injuries and surgical interventions. The occurrence of an abnormality is more than just variation in form. Consequently, testing for hypothesized “unusual form” represents another type of challenge and is dependent on the definition used. Here, we follow a pattern analysis perspective where an abnormality is defined as a pattern in the data that does not conform to some expected behavior [6].

Morphometric analysis underwent a conceptual revolution during the 1980s and 1990s [7]. As a result, a variety of methods and approaches have evolved with some clear conceptual and technical differences. Continuous boundary representations, such as Fourier descriptors e.g., have been employed to quantify form based on outline data [3,8]. In contrast, geometric morphometrics [9] uses homologous landmarks, these being defined as “a point of correspondence on an object that matches between and within populations” [10]. Within these landmark-based approaches three different ways to deal with the confounders of position and orientation have been proposed [11]: (1) Superimposition [12,13], involving placing the landmark data into a common frame of reference, (2) Deformation [14-17], where form differences are described in terms of deformation fields of one object into another and (3) Linear Distances [18], where all possible distances between landmarks and not their absolute position and orientation are measured. Advances over the past few decades have made morphometric analysis easier, more accessible and therefore more attractive to use in practical studies. For example, with the advent of three-dimensional (3D) tomographic imaging and rapid surface scanning, the quantification of form can now be performed indirectly and virtually both in 2D and 3D [19].

However, none of the existing approaches in morphometrics are particularly well suited to deal with abnormalities as such and hence are of limited use in their measurement. One reason for this is that unusual form differences typically introduce a Pinocchio effect characterized by a substantial, but localized, form change or difference [4]. For example, congenital malformations often have a discontinuous impact on form. A second and more important reason for the limited use of morphometrics, so far, is that there is no real mechanism in place to separate unusual from usual form differences. To address this dilemma, a new theoretical concept and associated modeling methodology, termed dysmorphometrics, is proposed. The topography of form abnormalities is encoded in a dysmorphogram, which is used to visualize and facilitate any subsequent quantification of the detected form differences of interest.

The manuscript is organized as follows: We start with the morphometric background. First, landmark superimposition following a likelihood formulation is shortly introduced. Secondly, the influence of and current solutions to the Pinocchio effect are given. The next section introduces and defines dysmorphometrics and extends the superimposition accordingly to deal with form abnormalities. Furthermore, we illustrate the relationship with the known solutions for the Pinocchio effect. Throughout the results and without loss of generality, we use human faces as example biological forms of interest (Appendix A). We define three different but typical types of questions related to facial form in clinical practice starting from 3D surface scans, e.g., and dysmorphometrics is applied to answer these questions: measuring abrupt form changes, asymmetry, and discordancy of features in human faces presenting abnormalities.

A discussion centered on comparisons with current morphometric techniques follows and future possibilities, limitations, and work of dysmorphometrics conclude the article.

## Methods

### Morphometric Background

#### Landmarking

Homologous *landmarks* are fundamental to geometric morphometrics as a biostatistical primitive. They are often defined as precise locations on biological forms that hold some developmental, functional, structural, or evolutionary significance [11]. However, owing to the lack of anatomically discrete features in regions of the face (cheeks and forehead for example), landmarks can only provide a sparse representation of the complete facial form and salient features can be overlooked [20]. Therefore, in addition to using true landmarks, points defined by relative locations (*pseudo-landmarks*), for example “the point of highest curvature”, or defined relative to other landmarks (*semi-landmarks* [21]), for example “halfway between the corners of the eyes”, can be used as well [19]. Alternatively, a spatially-dense indicated set of landmarks can be obtained using an anthropometric mask and mapping technique [22-24]. The latter provides a spatially-dense set of *quasi-landmark* indications, which are used here and are essentially obtained using a non-rigid surface registration (mapping) of a predefined facial template (anthropometric mask) [25].

Independent of the kind of landmarks used, in order to compare different form instances the same homologous or corresponding landmarks are indicated. Mathematically, a form instance is then represented as a (landmark) configuration consisting of the coordinates of a list of ordered landmarks:  $C = \{l_j | j = 1, \dots, K\}$  with  $K$  the number of landmarks and the ordering index  $j$  defining the unique label of the associated landmark  $l_j = (x_j, y_j, z_j)$  in 3D. Hence, a single form is defined as a  $K \times 3$  matrix.

#### Superimposition

When comparing a particular configuration or a group of configurations to another, any and all orientation and/or position differences are considered unimportant. However, both orientation and position influence the actual recorded coordinates. As mentioned previously, several approaches exist to deal with this problem; one of which is superimposition. Superimposition has the advantage of being quite intuitive and easy to visualize. It accounts for the confounding variables related to position and orientation through the optimal transformation of further analysis. In the context of a so called ordinary analysis, given two configurations  $C$  and  $\hat{C}$ , differences in form are defined based on the residual spatial differences  $D$  after superimposition. This can be formally written as:

$$C = T(\hat{C}, \theta) + D \quad (1)$$

with  $T(\cdot, \theta)$  a transformation model encoded by a set of parameters  $\theta$ , that are reflecting rotation, translation and scale operations (Appendix B). Finding the optimal superimposition requires the estimation of the optimal transformation parameters  $\theta$  such that  $D$  reflects true differences in form. Following a likelihood formulation of this problem requires a statistical model with an associated distribution that is assumed to have generated the observed data [26]. The likelihood of transformation parameters is

expressed as:  $L(\theta) = \Pr(C, \hat{C}|\theta) = \Pr(D|\theta)$  and the optimal superimposition is obtained following a Maximum Likelihood (ML) – Estimator :  $\hat{\theta} = \arg \min_{\theta} l(\theta)$  where  $l(\theta)$  is the negative log-likelihood (NLL):  $l(\theta) = -\log L(\theta)$ .

In morphometrics the Procrustes ML-estimators from Goodall [27], assuming a Gaussian perturbation model, are the most commonly known and used. Here, the form difference  $D$  is modeled as a zero-mean matrix of Gaussian displacements  $D \approx N_{K \times 3}(0, \Omega)$ . In its most general form, local perturbations for each landmark may be unequal and may even be correlated with each other as specified by the elements in the covariance matrix  $\Omega$ . According to Goodall [27] this covariance or so-called *model metric* accommodates two sources of variation in the observed data: measurement error and variation in shape.

As a *superimposition metric* the covariance can be user-defined or may be set to an estimate of the model metric. In the latter case any unknown parameters in the model metric are to be estimated on top of the unknown transformation parameters. This can prove to be challenging and a series of simplifications is therefore typically introduced: (a) under the assumption that the directions of variation are the same between different landmarks the full covariance matrix can be factorized using the Kronecker product  $\Omega = \Sigma \otimes \Xi$ .  $\Sigma$  is a  $K \times K$  covariance matrix for the rows of  $D$  reflecting variances and correlations among landmarks.  $\Xi$  is a  $3 \times 3$  covariance matrix for the columns of  $D$  or the dimensions and  $D \approx N_{K,3}(0, \Sigma, \Xi)$ . (b) Under the assumption that the displacements around a landmark are isotropic the dimension covariance matrix becomes the identity matrix  $D \approx N_{K,3}(0, \Sigma, I_3)$ . (c) Under the assumption that landmark differences are independent, the landmark covariance matrix is a diagonal matrix and (d) under the assumption that the landmark differences are identically distributed (i.e. they are homoscedastic) we have  $D \approx N_{K,3}(0, \sigma^2 I_K, I_3)$ . Because of the independency assumption, the probability  $\Pr(D|\theta)$  can be factorized into a product of individual landmark displacement ( $d_j$ ) probabilities  $\prod_{j=1}^K \Pr(d_j|\theta)$  (e) Finally, under the assumption that the distribution is the standard normal distribution we have  $D \approx N_{K,3}(0, I_K, I_3)$ . This results in the following series of NLL simplifications:

$$l(\theta)_{(a)} = \frac{1}{2} \left\| T(\hat{C}, \theta) - C \right\|_{\Sigma^{-1}, \Xi^{-1}}^2 + \frac{K}{2} \log |\Xi| + \frac{3}{2} \log |\Sigma| \quad (2)$$

$$l(\theta)_{(a,b)} = \frac{1}{2} \left\| T(\hat{C}, \theta) - C \right\|_{\Sigma^{-1}, I_3}^2 + \frac{3}{2} \log |\Sigma| \quad (3)$$

$$l(\theta)_{(a,b,c)} = \sum_{j=1}^K \left( \frac{1}{2\sigma_j^2} \left\| T(\hat{l}_j, \theta) - l_j \right\|^2 + \log \sqrt{2\pi} \sigma_j \right) \quad (4)$$

$$l(\theta)_{(a,b,c,d)} = \sum_{j=1}^K \left( \frac{1}{2\sigma^2} \left\| T(\hat{l}_j, \theta) - l_j \right\|^2 + \log \sqrt{2\pi} \sigma \right) \quad (5)$$

$$l(\theta)_{(a,b,c,d,e)} = \sum_{j=1}^K \left\| T(\hat{l}_j, \theta) - l_j \right\|^2 \quad (6)$$

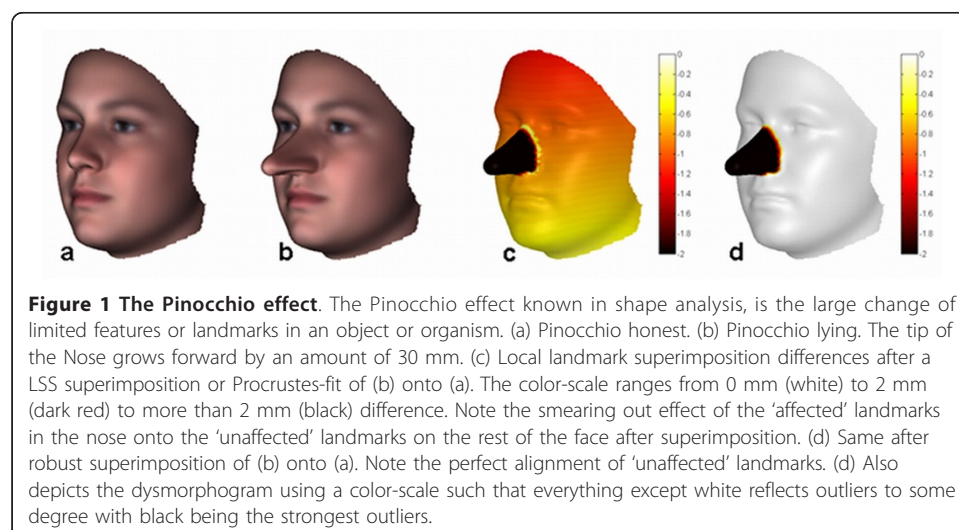
Note that any constant term (independent from all parameters) in these simplifications has been omitted as they do not influence the ML-estimation. Also note that the most simplified version with the NLL given in (6) is the frequently used Least Sum of Squares (LSS) superimposition [12].

The Procrustes solutions to the superimposition problem are desirable in shape analysis because of the straightforward link between form differences as squared residual errors and variances. This linkage with conventional multivariate statistics is known as the *Morphometric Synthesis* [28]. The Procrustes ML-estimators are also used when comparing more than two configurations simultaneously in a generalized setup (introducing an additional sum over the different configurations in the NLL's). In this generalized analysis different configurations are superimposed to a consensus configuration, typically an estimated geometric mean.

#### The Pinocchio effect

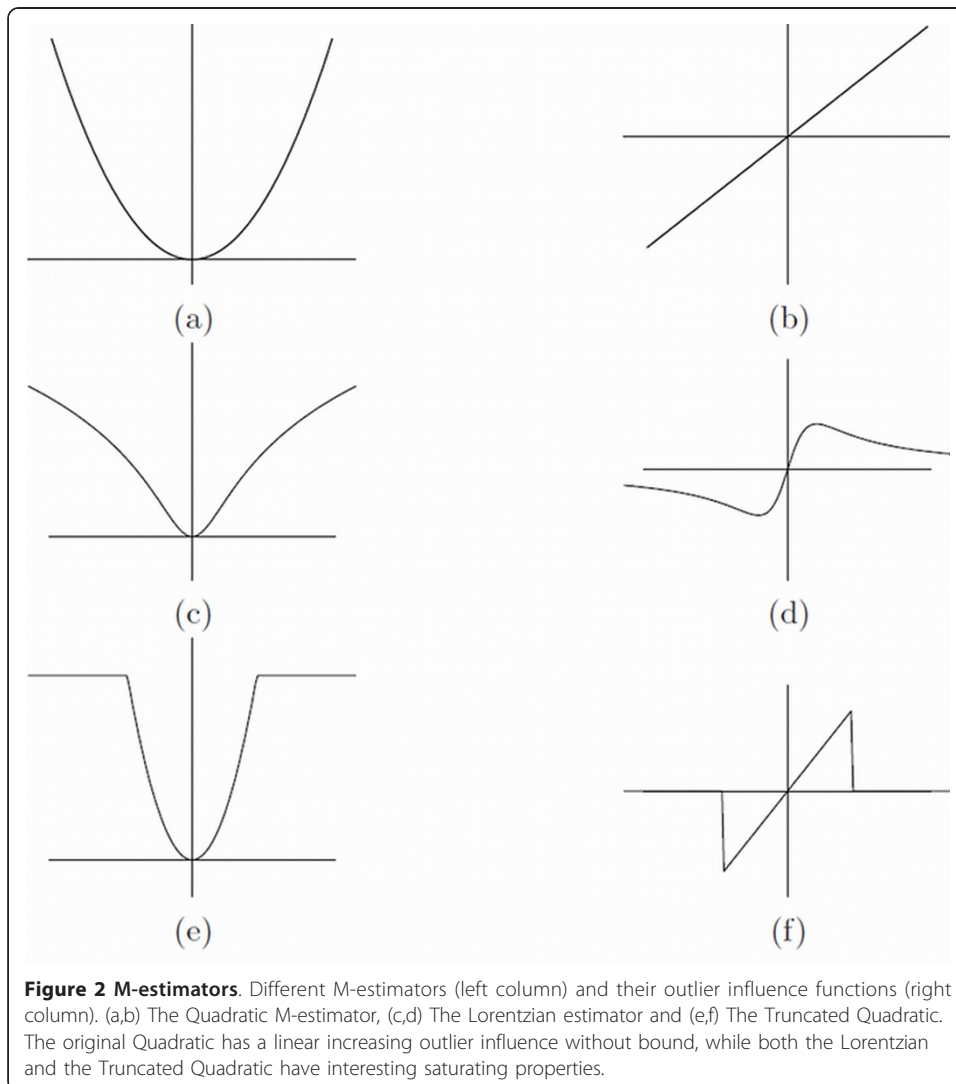
In order to illustrate the challenge in studying form abnormalities we use the well-known Pinocchio effect in shape analysis [4]. This effect is characterized by a substantial, but localized, form change or difference, as depicted in Figure 1. The Pinocchio effect influences a Procrustes superimposition using (6), e.g., between the honest (Figure 1(a)) and lying state (Figure 1(b)) of Pinocchio. The large differences of the 'affected' nose landmarks are contaminating the optimal placement of the 'unaffected' remaining landmarks as well. As a result all the landmarks are improperly aligned (Figure 1(c)) and any residual differences do not reflect true differences in form. An analysis of the Pinocchio effect can be obtained by casting the superimposition into an M-estimator formulation:

$$\hat{\theta} = \operatorname{argmin}_{\theta} \sum_{j=1}^K \rho(d_j) \quad (7)$$



The  $\rho$ -function is a loss function having a unique minimum when the residual error  $d_j = \|T(\hat{l}_j, \theta) - l_j\|$  is zero. The NLL in (6) for example is equivalent to using a Quadratic M-estimator with a  $\rho$ -function defined as  $\rho(x) = x^2$  and illustrated in Figure 2(a). An analysis of the Pinocchio effect is then performed by examining the equivalent M-estimator through its influence function [29], which characterizes the bias that a particular measurement has on the solution and is proportional to the derivative of the  $\rho$ -function. For the Quadratic M-estimator the influence function is  $\psi(x) = x$  which is depicted in Figure 2(b). It can be seen that the influence of substantial form differences increases linearly and without bound, as expected.

A robust alignment, not influenced by the Pinocchio effect, is illustrated in Figure 1 (d), which can only be obtained using an appropriately robust superimposition and associated estimator. Chronologically, the first such robust technique in morphometrics was suggested in [13] and is the so-called resistant-fit using a repeated median (RM) estimator. However, this technique fails to be successful in practice as it exhibits a high time complexity whilst convergence behavior is unclear and dependent on the



initial alignment [10]. Another robust median-based estimator is the Least Median of Squares (LMedS) estimator [30], where the sum in (7) is simply replaced by the median.

Alternatively, as suggested by [10], any type of robust estimator can be used such as, for example, robust M-estimators for which a wide range of possible choices has been suggested within the field of robust statistics. Good choices are the redescending M-estimators that have an influence function that is increasing near the origin but decreasing towards zero again further away from it. Two such examples are the Lorentzian and the Truncated Quadratic (TQ) estimator whose  $\rho$ - and influence functions are illustrated in Figure 2(c,d) and 2(e,f) respectively. They both completely ignore gross form differences during superimposition but do not completely ignore moderately large differences as the median approaches do. Note that different M-estimators imply different underlying distributions that are not necessarily the same as the Gaussian distribution used in Procrustes ML-estimators.

Only recently, tractable and constrained solutions to the model metric estimation for heteroscedastic Gaussian distributed landmarks were formulated allowing for the modelling of individual variances per landmark. Essentially, the model metric can be adapted to allow for larger mobility (higher variance) in the affected nose landmarks and lower mobility (lower variance) in the invariant substructure or unaffected landmarks. The result is a variance weighted LSS solution. In [31] the NLL in (4) without the log term is used and is theoretically linked with a scaled mixture model of Gaussians enabling “large-scale” tolerant superimposition. Indeed, a scaled mixture of Gaussians is a Heavy-tailed distribution known to provide robustness. The shape of the Heavy-tailed distribution is dependent on the prior distribution constraining the variance estimations. In the case of an inverse gamma prior distribution, e.g., the Student T-model is obtained [31]. Similarly, but within a generalized analysis, a complete covariance model metric estimation also allowing for landmark dependency or correlation (using the NLL in (2)) is given in [26].

### **Dysmorphometric Extensions**

#### ***Motivation: The Pinocchio dilemma***

Morphometrics is primarily focused on form related questions such as: What is the average form and what are the patterns of variation around it within a population, or how do groups differ in shape and what is the functional importance of those differences [9]? The morphometric dilemma raised by the Pinocchio effect is whether to include or exclude this case completely or partially from the study, as the morphometric approach studies only typical variation and co-variation over all the observed landmarks in populations and not any atypical localized differences. If Pinocchio is unique within a population sample, the strength of his uniqueness will influence the analysis of typical variability in the population. For the model metric estimation of a population with heteroscedastically distributed substructures, e.g., a higher variance around the nose is estimated because of Pinocchio. In other words, the model metric is stretched to include Pinocchio and this inclusion might not be desirable. One possible course of action is to use a robust procedure to superimpose first, analyze the residuals and to ignore suspicious landmarks in a subsequent Procrustes analysis as missing data [10].

Dysmorphometrics, in contrast to morphometrics, is focused on questions such as: What makes an individual different and what are the patterns of variation of that particular difference over a group of individuals? For example, it strives to provide an answer to the question ‘what specific feature makes Pinocchio different?’ In such a context, the only differences of interest are the ‘affected’ landmarks and a robust superimposition as in Figure 1(d) is preferred, if not crucial. Stated differently, in morphometrics, typical (co-)variances are the variables of interest and the atypical variations are the ‘nuisance’ variables needing to be discarded, whereas, by contrast in dysmorphometrics, the variables of interest are the atypical variations with the typical variations being the confounding variables. However, the challenge remains to define and identify atypical variation whilst compensating for the confounding typical variation.

#### **Modelling form differences as outliers**

A morphological abnormality can be seen as a difference in form that is inconsistent, discordant, and/or atypical. This definition inherently implies two important aspects. Firstly, they are only relatively defined and can only exist given prior knowledge of what is normal, consistent, harmonious, and/or typical. Hence, a representative norm is required. Secondly, the difference is required to be significantly distant from the norm. Employing this definition, the key point is thus to model form differences as outliers.

Most often the studied abnormalities are spatially localized and do not extend over the whole region of interest. In morphometrics such abnormalities will be represented by an ensemble of landmarks positioned ‘suspiciously’ compared to the norm, while the other landmarks are positioned as predicted by the norm. In order to distinguish the ‘suspicious’ or outlier landmarks from the ‘normal’ or inlier landmarks, during a superimposition, only the inlier landmarks need to be optimally aligned in the LSS sense and used to estimate the model metric if necessary. Hence, an appropriate robust superimposition and model metric estimation is required. Although many of the estimator choices mentioned in the previous section can be used to obtain such a robust placement of landmarks they, unfortunately, model outliers only implicitly. In other words, these superimpositions are robust against outliers, as in not influenced by them, but they do not allow for a meaningful outlier flagging/detection mechanism. Furthermore, they do not allow for a (robust) Procrustes model metric estimation as their underlying distributions typically deviate from Gaussian distributions. Instead, dysmorphometrics models outliers explicitly using outlier processes. A problem formulated in terms of explicit outlier-processes can be converted or viewed in terms of robust estimators [32]. An outlier process formulation, however, is more general than the original robust estimator. For example, due to the explicit nature of the outlier process, constraints on the spatial organization of the outliers can be formulated [33]. Based on the relationship with robust estimators defined in [32], equivalent M-estimators can be determined to analyze the influence of outliers on the superimposition.

#### **Extending ML-estimators with outlier-processes**

In practical applications, the assumed perturbation model in ML-estimators is only an approximation to reality, and estimation of the parameters should not be severely affected by the presence of unusual form differences. Indeed, the probability of an outlier according to the perturbation model is very low because the *outlier cannot be explained* by the model and therefore  $\lim_{Pr(d_j|\theta) \rightarrow 0} \log Pr(d_j|\theta) = -\infty$ . In order to



deal with outliers a principled approach is adapted from [33]. The underlying idea is to model inliers and outliers as separate random variables. Inliers/Outliers are landmarks whose displacements have been generated by an inlier/outlier-process with associated inlier/outlier-distribution. The complete process, having generated all displacements, with associated complete distribution combines both inlier- and outlier-processes through the introduction of a latent variable  $z_j$  that signals whether the residual error  $d_j$  is an outlier ( $z_j = 0$ ) or inlier ( $z_j = 1$ ). This is achieved using mixture modelling and the result is an augmented superimposition in which landmarks are superimposed and outliers are flagged at the same time.

Given the observed landmarks, a joint estimation is now presented where besides the transformation parameters  $\theta$  also the outlier map of latent variables or unobserved data  $Z = \{z_j | j = 1, \dots, K\}$  is to be estimated. A popular tool for statistical estimation problems involving unobserved or incomplete data is the Expectation-Maximization (EM) algorithm ([34,35]). Herein, the likelihood of transformation parameters can be re-written in terms of the latent variables:  $Pr(D|\theta) = \sum_Z Pr(D, Z|\theta)$ . EM then produces a sequence of parameter updates  $\{\hat{\theta}^{(t)} | t = 0, 1, \dots\}$  by alternating two steps: the E-step and the M-Step. In the E-step, the latent variables are estimated using the conditional expectation resulting in the so-called Q-function. In the M-step, the Q-function is maximized generating a new update for the transformation parameters. Note that the Q-function is the equivalent robust M-estimator for the superimposition with an explicit outlier-process formulation.

The challenge is to formulate all likelihoods and distributions involved taking into account certain assumptions and prior knowledge of the augmented superimposition problem. In Appendix C we give such a detailed formulation and derive a practical extended Procrustes (Ext-P) ML-estimator. Here, we only summarize the practical Ext-P ML-estimator.

#### **A practical extended Procrustes ML-estimator**

As an example ML-estimator extension, we start from the Procrustes ML-estimator assuming a zero-mean Gaussian perturbation model with i.i.d landmark displacements (using the NLL in (5)). Hence, the model metric reflects additive white Gaussian noise (AWGN) with a single parameter  $\sigma$ . This constitutes the inlier-distribution  $Pr_i(x)$ . For the outlier-distribution several choices can be considered. If some knowledge about the associated distribution of the outlier generating process is given then this can be used. In most cases, however, this knowledge is not known and could be estimated as well [33,36]. However, we simply assume the outlier-distribution to be uniform such that  $Pr_o(x) = \delta$  with  $0 < \delta < 1$ . In this simplified case, the Q-function, written down as an M-estimator (7), has a  $\rho$ -function equal to:

$$\rho(d_j) = \frac{b_j}{2\sigma^2} d_j^2 + b_j \log \sqrt{2\pi} \sigma - (1 - b_j) \log \delta \quad (8)$$

$$b_j = E_{Pr(Z|D, \hat{\theta}^{(t)})}[z_j] = \frac{Pr_i(d_j | \hat{\theta}^{(t)})}{Pr_i(d_j | \hat{\theta}^{(t)}) + \lambda} \quad (9)$$

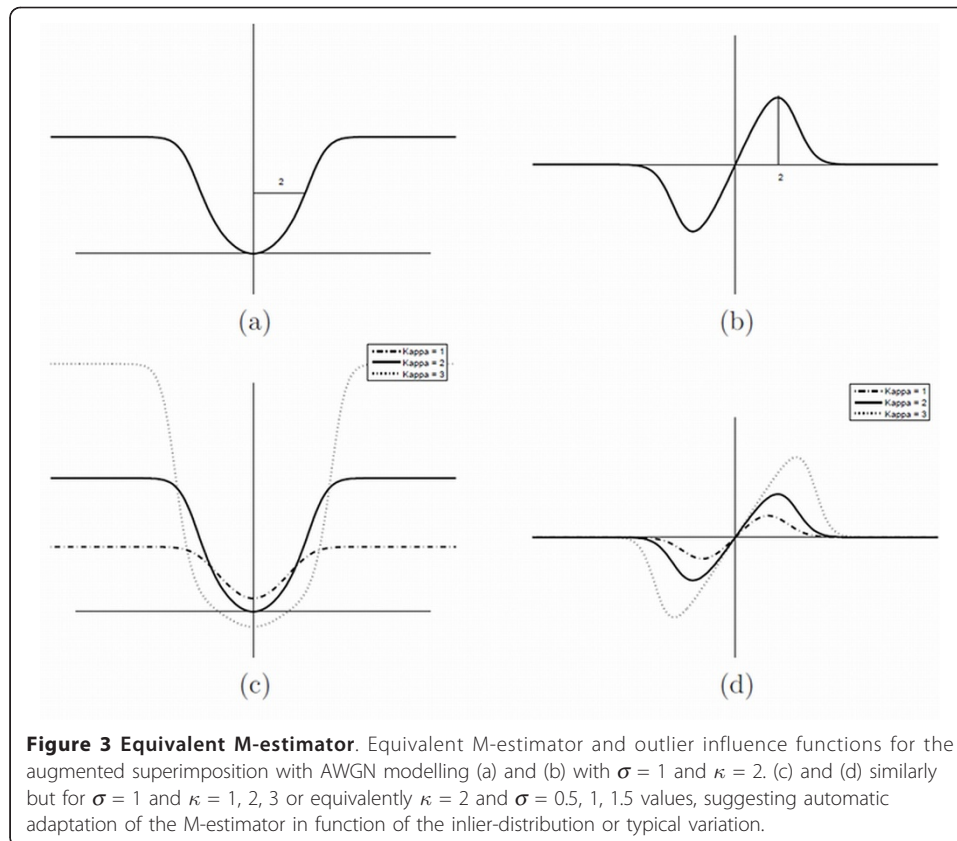
$$\lambda = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\kappa^2\right) \quad (10)$$

In (8) we see that for inlier landmarks ( $b_j = 1$ ) the Ext-P ML-estimator is equal to the original Procrustes ML-estimator (5). Furthermore, an outlier landmark ( $b_j = 0$ ) does not contribute to the estimation of the transformation parameters  $\theta$  or model metric parameter  $\sigma$  because  $\log \delta$  is independent of both. From (9), we see that the estimation of the latent variables during superimposition is dependent on the estimated  $\sigma$  of the inlier-distribution and on a single parameter  $\lambda$  that codes for prior knowledge on the outlier distribution. Changing  $\lambda$  changes the amount of outliers versus inliers. A higher value suggests more outliers, while a lower value suggests more inliers. A proper choice of this parameter is therefore required, but can prove to be challenging. This choice is made easier and, more importantly, statistically relevant in (10) following [37].

In statistics, an observation can be called atypical with respect to a given normal distribution if its (Mahalanobis) distance exceeds a predefined threshold. Therefore, in (10)  $\lambda$  is made inlier-distribution dependent and is re-parameterized using the more interpretable prior outlier parameter  $\kappa$ . The actual choice of  $\kappa$  is equivalent to the choice of a statistical significance level above which local form differences are considered atypical compared to the typical Gaussian (inlier-) distribution. For example  $\kappa = 2$  and  $\kappa = 3$  suggest a significance level of  $p = 0.05$  and  $p = 0.001$ , respectively. Figure 3 depicts the  $\rho$ -function (8) (Figure 3(a)) and its influence function (Figure 3(b)). It is a continuous approximation of the truncated quadratic function (thus having better second derivate properties, which is important for reasons of optimization and convergence) and resembling the Tukey-biweight function. Note that the width of the  $\rho$ -function is equal to  $\kappa \times \sigma$  and thus completely defined in terms of the inlier-distribution and prior outlier significance parameter. In Figure 3(c) and Figure 3(d) the  $\rho$ - and influence functions are depicted for different values of  $\kappa$  with fixed  $\sigma$  or, equivalently, different values of  $\sigma$  with fixed  $\kappa$ . It is observed that, apart from the intuitive choice of  $\kappa$ , a meaningful and adaptive outlier flagging mechanism is provided. Indeed, inlier landmarks, encoding for typical form differences, are fitted in the LLS sense as in the original Procrustes-fit and thus provide an estimate of typical form variation following the assumed perturbation model. Outlier landmarks are then identified in terms of being significantly atypical compared to the estimated typical variation. The resulting Ext-P ML-estimator with significant outlier detection is used throughout the results section with a fixed parameter setting  $\kappa = 2$ , reflecting a common choice for statistical significance ( $p = 0.05$ ) in biology.

## Results

In this section, we focus on the biological relevance of the results only and additional computational comparisons such as time complexity can be found in Appendix D. Throughout the results, human faces are used as example biological forms of interest (Appendix A). They are the biological billboard of our identity, underlying genes, and environmental exposures. Our visual-cognition system is very adept and capable of discriminating between different faces and identifying or classifying facial abnormalities.

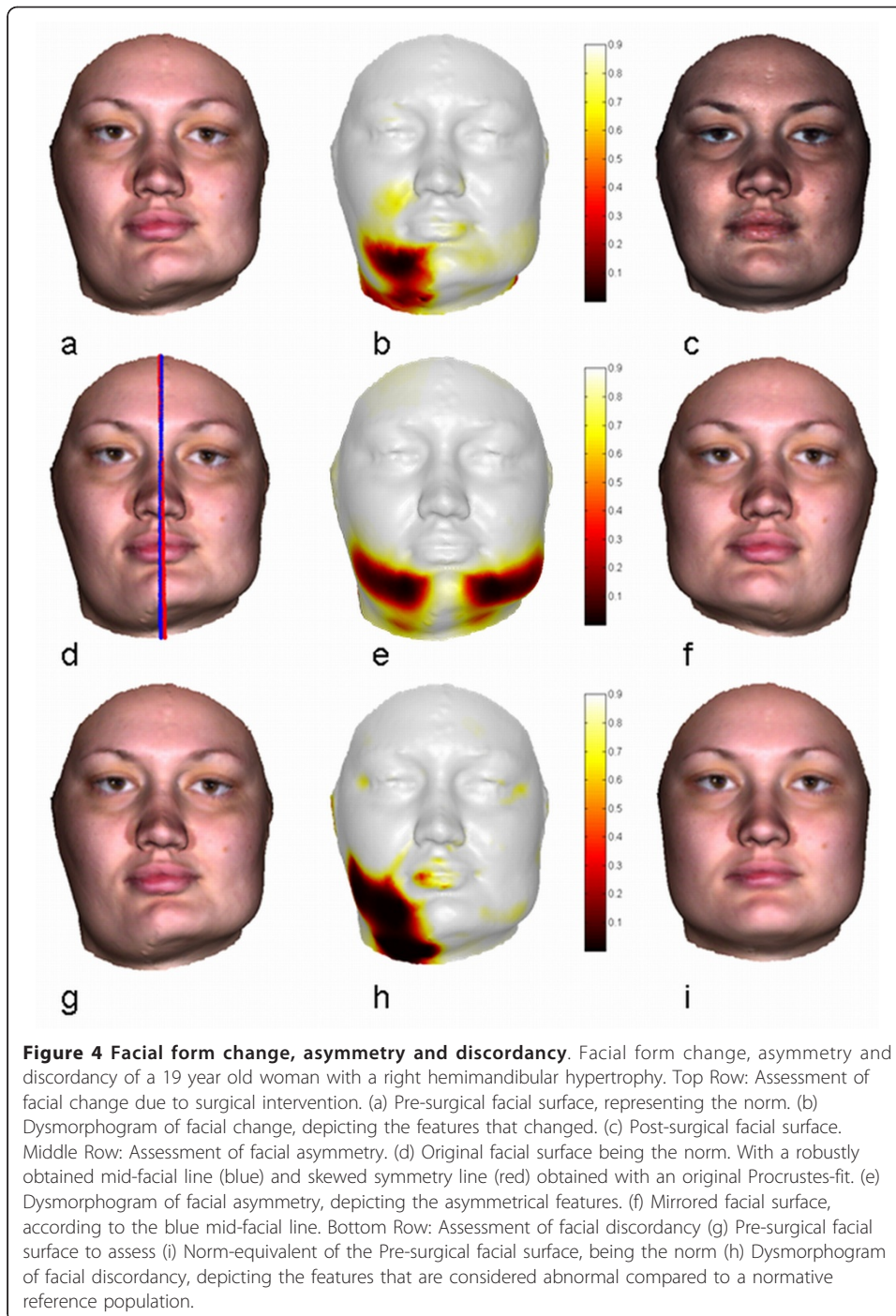


Furthermore, facial abnormalities have been studied in depth in the literature enabling knowledge-based validation of our results.

#### Dysmorphometrics for abrupt facial changes

Abrupt changes in form may occur for many reasons and are highly variable making their measurement challenging. The Pinocchio effect is such a (synthetic) toy-example. Here, the true change was perfectly identified in the dysmorphogram (Figure 1(d)) using the proposed Ext-P ML-estimator. Biomedical examples of abrupt changes in facial form include deformations associated with surgical interventions or traumatic events. The capacity to measure the effects of surgical treatment over time has been long sought after to audit outcomes and evaluate relapse [23]. Such measurements provide feedback and insight allowing for reflection on surgical improvements. An example of a surgical intervention is illustrated in the top row of Figure 4. It shows a 19 year old woman treated for facial asymmetry coincident with right hemimandibular hypertrophy. The discrepancy in the lower mandibular border was corrected with an osteotomy and a wedged Le Fort1 osteotomy to resolve occlusal cant. 3D images were taken pre- (Figure 4(a)) and post-treatment (Figure 4(c)).

Visual comparison of the pre and post-surgery images clearly shows regions that have been altered by the intervention. In order for the superimposition to reflect true changes in form, it has to be performed based on the unaltered regions only. A dysmorphometric solution to this is to define the 'norm'; in this case, e.g., the pre-surgical situation. Then, an Ext-P ML-estimator is used for the superimposition of the post-



treatment situation onto the norm. The idea is that outlier-landmarks are situated in regions that changed, while inlier-landmarks are located in the unchanged regions. The variation encoded in the inlier-landmarks should ideally be zero as in the Pinocchio example, but in reality this is never the case. This is because errors in the 3D scanning and quasi-landmark mapping procedures introduce noise or small perturbations on the landmark locations. Here, the level of this noise is considered typical variation (of no interest) around the norm configuration and is modeled by the AWGN inlier-

distribution of the Ext-P ML-estimator. Outliers or atypical differences are then defined as significantly different with regard to the estimated noise-level. The resulting dysmorphogram is shown in Figure 4(b). It depicts a continuous-valued  $[0,1]$  spatial map of the local regions that are significantly different ('yes', non-white regions) or not ('no', white regions). According to the intervention that took place these accurately reflect the anticipated anatomical changes. Similar examples can be found in related work [23].

#### **Dysmorphometrics for facial asymmetry**

Bilateral symmetry often occurs in organisms and is defined with respect to reflection across the midsagittal plane dividing a perfectly bilaterally symmetrical organism into equal right and left halves. During development in vertebrates imbalances in growth will inevitably result in some degree of asymmetry. Mild facial asymmetries are thus common in typical growth and development [38]. Severe and pathological asymmetries, on the other hand, are a feature of disordered growth as a consequence of genetic and/or environmental causes [39].

A protocol, grounded in geometric morphometrics, for the measurement of objects displaying bilateral symmetry consists in undertaking a LSS Procrustes-fit of landmark configurations and their mirror configurations [40,41]. However, when the asymmetry in the face increases locally and develops abnormally, as in the pre-surgical presentation of the case shown in Figure 4(d), the original Procrustes-fit becomes influenced by it. To obtain a robust assessment of asymmetry the superimposition involved has to be done based on the more symmetrical regions only. A dysmorphometric solution is to define the original configuration as the 'norm'. Then, an Ext-P ML-estimator is used for the superimposition of the mirror configuration onto the norm. The idea now is that outlier-landmarks are situated in regions that are asymmetrical, while inlier-landmarks are located in symmetrical regions. Again, errors in the 3D scanning and quasi-landmark mapping procedures introduce noise on the landmark locations modeled by the AWGN inlier-distribution.

The resulting dysmorphogram of the asymmetry case is shown in Figure 4(e). As a by-product, after the superimposition, a robust estimate of the midsagittal plane can also be obtained (blue line in Figure 4(d)). This 'extended' asymmetry assessment protocol has been used in related work [22,42]. There, the goal was to detect disordered facial growth patterns in individuals characterized by asymmetries with reference to the individual asymmetry variation found in the general population rather than to some ideal of perfect symmetry, which rarely exists.

#### **Dysmorphometrics for facial discordancy**

Given a population of interest, an individual is concordant with that population if it is within the boundaries of variation of the population. Stated differently, a concordant form is in harmony with the population. Form discordancy, on the other hand, is the lack of harmony. Biologically, this is manifested as the type of facial abnormalities that are best known in craniofacial disorders and dysmorphologies. Assessment of such disorders affecting facial morphology is typically performed compared to 'normality'. This, however, presents two major challenges. The first is to define normality or harmony in

facial form. The second is to define the dysmorphic face with respect to normality by identifying and localizing the discordancy in the form of the face.

A morphometric approach to this problem is to establish both normal and abnormal population databases [43] and to assign a given individual to the most plausible population, conceptually similar to a taxonomic discrimination in systematic biology. Normal variation can be learned starting from a proper reference dataset consisting of healthy individuals without pathology. The collection of abnormal population data however often proves to be more difficult and impractical particularly when dealing with rare and highly variable situations. A dysmorphometric formulated solution approaches the problem without the need to compile databases representing abnormal populations as follows: Firstly, the normal population is chosen as the 'norm'. Secondly, an Ext-P ML-estimator is used for the superimposition of a configuration under investigation onto the norm. Here, the 'population norm' is represented as a Point Distribution Model (PDM) [10] using a principal component analysis (PCA) around the consensus configuration of the population (Appendix B). One of the advantages of using a PDM is that the within-population variation becomes part of the transformation model used during superimposition. As such, typical variation in the population is treated and compensated for in the exact same way as other confounding variables of no interest like orientation and position. The result of this superimposition is the creation of a norm-equivalent configuration, which is the harmonious counterpart of a given form configuration and can be considered as an individual-specific typical or normalized reference. Stated differently, the consensus configuration of the population is allowed to change within the boundaries of typical variation to reflect the given configuration as much as possible.

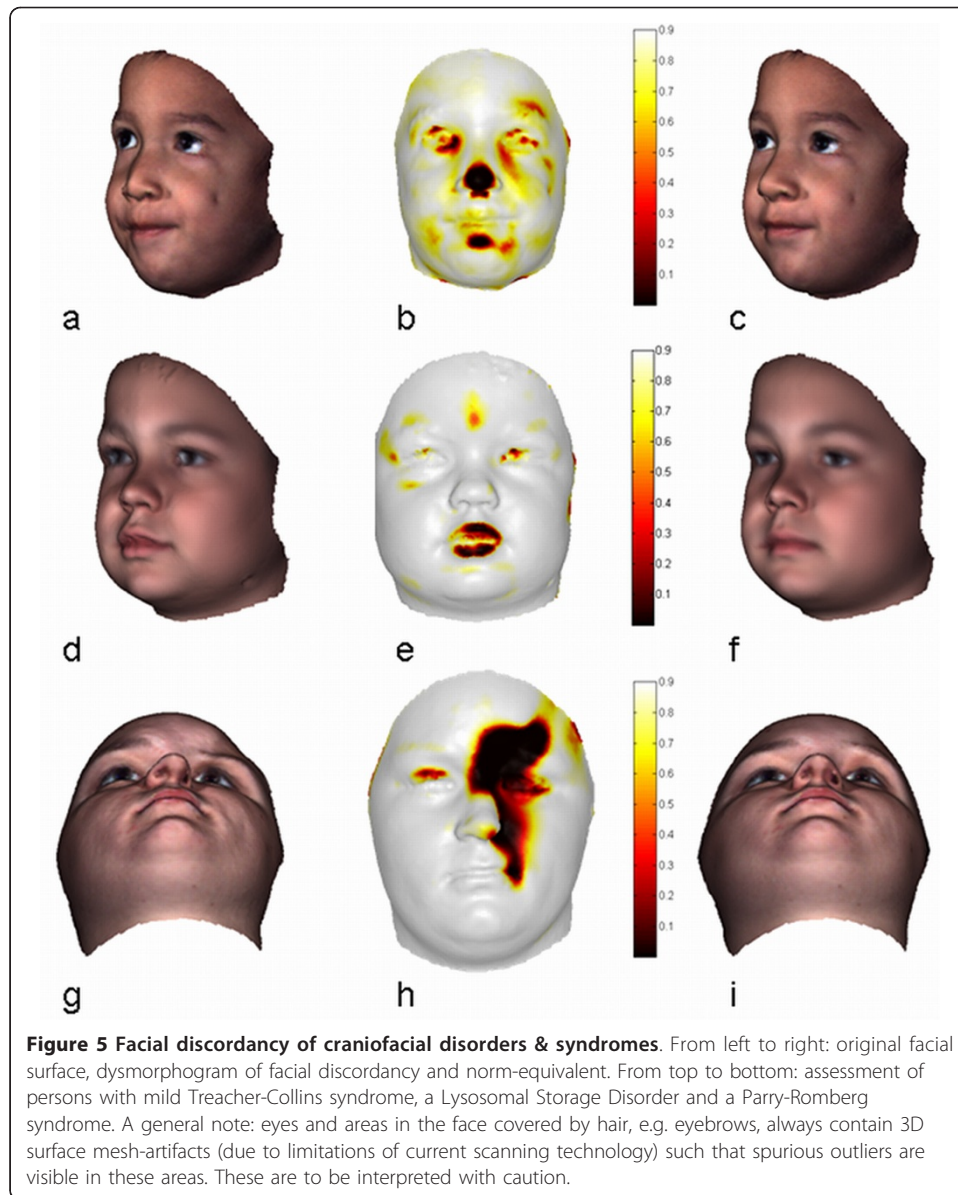
The superimposition or the creation of the norm-equivalent of a given dysmorphic face has to be performed on regions in the face that are in harmony with the normal population. From a dysmorphometric point of view, the idea is that outlier-landmarks are situated in regions that are discordant, while inlier-landmarks are located in concordant regions. Again, errors in the 3D scanning and quasi-landmark mapping procedures introduce noise or small perturbations on the landmarks locations. The level of this noise is considered typical variation (different to the population variation) and as before is modeled by the AWGN inlier-distribution. Outliers or atypical differences are then defined as significantly different w.r.t. the norm-equivalent configuration based on the estimated noise-level, therefore reflecting significant discordancy after compensating for within-population variation. Note that by using the PDM within the transformation model of the superimposition, the simple but practical Ext-P ML-estimator only modelling AWGN can be used to separate landmarks into concordant and discordant.

For the lying Pinocchio toy-example (Figure 1(b)) the norm-equivalent according to a clinically normal population turns out to be the honest Pinocchio (Figure 1(a)) (which was the consensus of that population) with a dysmorphogram (Figure 1(d)) correctly identifying the nose as significantly discordant. Assessment of facial discordancy is further demonstrated in four different cases shown in Figure 4 (bottom row) and Figure 5. From left to right, the facial configuration under study, the dysmorphogram and the norm-equivalent configuration are displayed. The norm-equivalent of each case is clearly a patient-specific norm, which has not been previously available and is proving

to be of value in clinical practice for treatment planning and auditing purposes. The first case (Figure 4(g,h,i)) concerns a 19 year old woman with right hemimandibular hypertrophy, treated for facial asymmetry. The dysmorphogram, as expected, localizes the highly asymmetric right lower mandibular as discordant but indicates a displacement as well of the mandible to the contralateral side and alveolar compensation in the maxilla on the affected side. The second case (Figure 5(a,b,c)) is a child with a (mild) Treacher-Collins Syndrome (TCS), a rare genetic disorder characterized by craniofacial deformities with occurrence prevalence of 1:10,000 [44]. The dysmorphogram revealed regions known to be affected in TCS including malar, zygomatic, and periorbital regions. In addition, the nasal tip was highlighted, which has previously been described but not considered a major symptomatic feature (GeneTests, <http://www.genetests.org>). The third case (Figure 5(d,e,f)) is a person suffering from Lysosomal Storage Disease (LSD, MPSII), which is a rare inherited metabolic disorder that results in accumulations in glycosaminoglycans (GAGs) with a prevalence of <1:100,000 [45]. People with this condition have been described as having 'coarse' facial features; the dysmorphogram demonstrated the fullness to the lips for example caused by accumulated GAGs. The last case (Figure 5(g,h,i)) is a person with a Parry-Romberg Syndrome. This is a rare hemifacial atrophy disorder characterized by progressive degeneration of the subcutaneous tissues and fat that can also involve bone, cartilage and muscle. This progressive condition often affects the left maxillae adjacent to the nose progressing to the corner of the mouth, around the eyes and brow but may vary from case to case as it appears to occur randomly with unknown etiology [46]. The dysmorphogram illustrates the extent of the disease in this young woman, in whom the affected area appears to be confined to distinct connective tissue septa in the face. All these cases illustrate features that have been previously known as distinctive features of the condition. However, new and highly relevant spatial information can be identified and quantified as well. This transforms previously descriptive dysmorphology to much more informative quantitative dysmorphometrics.

## Discussion

The study of form using quantitative, morphometric descriptors is fundamental to many biological studies. Emerging technologies of 3-dimensional (3D) scanning and geometric morphometrics are providing the means to establish objective criteria which can be used for phenotypic investigations in combination with epigenetic investigations [47,48]. Here, we have introduced dysmorphometrics as a means to identify and quantify unusual form differences like abnormalities that are typically encountered in biomedical sciences. Dysmorphometrics builds upon existing geometric morphometric techniques, but models form differences explicitly as outliers. Applying a test for outliers is referred to as a test of discordancy [49] and the practical Ext-P ML-estimator clearly resembles known statistical tests for outliers. For example, the same result could have been obtained by applying a Grubbs' test using the inlier-distribution without the need to explicitly define an outlier-process and associated distribution. However, the explicit case is more general and open to further assumptions or prior knowledge about the outliers. The potential of this is strong and is still to be explored. The key is to model the patterns and their analysis as biologically relevant based on the right assumptions.



The requirement for robustness during the analysis of unusual form variation has been acknowledged by many others. For example, [50] advised the need for caution to be taken into account when comparing preoperative with postoperative asymmetry scores in faces. This was because the achievement of ‘best-fit’, using the LLS Procrustes superimposition, was influenced by severe asymmetries in the preoperative situations. This resulted in spurious changes and unrealistic reduction of the asymmetry post-operatively in regions not affected by the surgical intervention. A popular way to address this problem in practice is to work with “stable” landmarks only, based upon which the superimposition is then performed. For example, the superimposition is done on predefined areas in the face like the nose ridge in [51] or by carefully indicated landmarks as in [39] and in [52]. Another strategy is to perform a LSS superimposition first, then remove parts according to a threshold and redo the superimposition using the remaining parts [53]. All these approaches try to achieve the



same result as a robust superimposition. However, they introduce subjectivity by the obligation to manually select a region, landmarks or threshold of interest. They are heuristic and different for different superimposition tasks. Unfortunately, at the time, the known repeated median method that was used for resistant fitting [12] lacked proper mathematical underpinning and convergence behavior compared to their original LSS solution and was therefore never advocated as an alternative. In contrast, the proposed modelling methodology of dysmorphometrics results in an adaptive and robust superimposition with meaningful outlier detection. The underlying mathematical model is an extension of the original theory of Procrustes ML-Estimators [27]. Furthermore, it is shown that behavior against outliers equates that of re-descending M-estimators, which are a popular and efficient class of estimators in robust statistics. However, the difference here is that the underlying Gaussian perturbation model assumption is kept intact. Alternatively, this is also achieved using the previously mentioned techniques of heteroscedastic variance estimation during superimposition [26,31]. However, these techniques by themselves do not allow for a robust model metric estimation, as the metric is stretched to include any higher variances caused by outliers. This inclusion does not enable separating outliers from inliers and so might be undesirable for further analysis. An interesting extension, however, would be to combine both the techniques of heteroscedastic variance estimation and dysmorphic outlier flagging.

In the assessment of discordancy, which inspired the development of the concept of dysmorphometrics, contrasts in strategy with morphometric analysis are apparent. For example, [43] and [54] build distributions of form for both typical and atypical populations. To assess an individual a closed-classification is then performed. However, the individual is always attributed to either one of the given populations, even in the case when it does not belong to any of them. Furthermore, this morphometric approach loses the power to individualize and can only visualize population differences expressed as a difference between averages or a net difference. In contrast, a dysmorphometric approach enables an open-classification, which is less restrictive and has a greater range of applications. It also allows for an individual-specific assessment and visualization of a problem or hypothesis. To conclude, the essence of dysmorphometrics is the ability to identify and measure the unknown abnormality, if any; with a norm reflecting what is known. This enables an alternative research strategy. Whereas morphometrics is typically used in a deductive research approach, where a general theory is formed and typical data observations are collected to test hypotheses, dysmorphometrics, on the other hand, provides for an exploratory research approach. An initially unknown, quantified observation is then made that can be subject to further testing by directed data collection from which a general theory is derived. A good example is the extent of the Parry-Romberg disease appearing to be confined by known connective tissue septa in the face. Hypotheses to explain this single observation are to be validated by collecting similar case data.

The use of dysmorphometrics is limited by two constraints. Firstly, but most importantly, an appropriate norm must be established. Hence, a straightforward comparison between two different individuals without taking population variation into account, e. g., cannot be done because neither one of the two individuals can be chosen as the norm. Additionally, in case of a population-based norm, it may prove challenging to

decide if an individual belongs to the norm or not and will be entirely dependent on the application. Secondly, the number and extent of form abnormalities cannot exceed the breakdown point of the outlier detection scheme employed. The breakdown point of a detection scheme is the percentage of data allowed to be outliers before they become undetectable.

## Conclusions

Morphological abnormalities are often encountered and of great interest in biomedical sciences and include, e.g., deformations in form due to congenital malformation and/or environmental constraints as well as large changes in form associated with traumatic injuries and surgical interventions. In current morphometrics, the means to identify and quantify such unusual form differences remains limited. To address this shortcoming, dysmorphometrics was introduced, which is a novel and unexplored concept that augments current geometric morphometrics to deal with, quantify and spatially map form abnormalities to facilitate their analysis. Essentially, an abnormality implies the existence of a data pattern that cannot be explained by typical patterns. As such, dysmorphometrics models form differences explicitly using outlier processes, resulting in adaptive and robust superimpositions. Furthermore, dysmorphometrics builds upon existing techniques such that the underlying mathematical model is a straightforward extension of the original theory of Procrustes ML-Estimators.

Throughout the results, outliers were defined above an estimated noise-level, as modeled in a practical Ext-P ML-estimator, therefore reflecting significant differences in form with a biological meaning. The results are unique and illustrate the power of this technique to reveal unusual form differences given only normative data either representing a single individual or a population. In the case of a population norm, variation within the population is considered of no interest and is treated as a confounding variable just like orientation and position differences. Dysmorphometrics can generate an individualized quantification of form abnormalities and leads to alternative and more informative population comparisons and research strategies. The analysis of multiple dysmorphograms for example can lead to novel and unexplored statistics in morphometric analysis.

## Appendix A: Facial Data Acquisition

Human faces are used as the biological data of interest for which appropriate Ethics approval was received: (1) *The Characterization of 3-Dimensional 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, Australia.

3D facial images of 800 healthy young people as well as patients between the ages of 5-25 were collected using a 3dMD facial scanning system. An anthropometric mask (AM) [22,23] was mapped onto the 3D facial images in a (quasi) anatomical manner using a non-rigid surface registration algorithm based on implicit functions described and validated in [25]. As a result all the points defined in the AM were automatically

indicated on each facial surface separately in a consistent way. Consequently, spatially-dense (~10.000) quasi-landmarks are known on all 3D facial images.

### Appendix B: Transformation Model

The transformation model stipulates how landmark configurations can be placed onto each other and reflects prior knowledge on the superimposition problem. Sometimes, not all transformations are feasible or realistic so that certain fitting constraints can be imposed to ensure that the superimposition behaves according to the prior knowledge. These constraints, inducing transformation regularizations, can be modeled probabilistically using a Gibbs prior distribution [10] on the parameters  $\theta$ , restricting the space of possible solutions:

$$Pr(\theta) = \frac{1}{Z} e^{-||L(\theta)||^2} \quad (11)$$

This distribution expresses the probability of a certain transformation parameter setting  $\theta$ , within the range of possible parameters, favoring more plausible settings.  $Z$  is a normalization constant and  $L$  is an operator defined on the space of parameter values representing the regularization as a squared norm. Inclusion of the prior model into the likelihood formulated superimposition follows a Bayesian inference strategy and results in multiplying  $Pr(D|\theta)$  by (11). After taking the negative log likelihood, an extra term  $||L(\theta)||^2$  is therefore added to equations (2-6).

#### Single norm transformation model

The simplest transformation model is the class of rigid transformations, only compensating for overall differences in pose by global translation, rotation and, if shape instead of form comparison is wanted, also scaling:

$$T(\hat{C}, \theta) = sR\hat{C} + t \quad (12)$$

$R$  is a rotation matrix parameterized using three Euler angles,  $t$  is a 3D translation vector and  $s$  is a scaling factor. This transformation model is well-known and extensively used in geometric morphometrics. It is also used in dysmorphometrics when dealing with a norm that is defined based on a single configuration as in the scenarios of time related form changes and asymmetry assessments of individuals. Generally, no particular rotation, translation and/or scaling is favored, hence due to the rigidity of the transformation model the regularization simply reduces to  $||L(\theta)||^2 = 0$ , such that (11) is a constant term (and therefore omitted) not influencing the ML-estimation of the parameters  $\theta$ .

#### Population norm transformation model

In the case of a population defining the norm, the same rigid transformation model (12) can still be used. In this scenario, a Generalized Procrustes-fit of all the configurations in the population is performed and the norm consists of the resulting consensus configuration in combination with the full covariance matrix  $\Omega$  around it. The covariance matrix is the model-metric coding for typical shape variation amongst individuals in the normal population and can be plugged straight into a Procrustes ML-estimator without the need to make any of the simplifying assumptions to create equations (2-6).

Doing so, measures the Procrustes distance between the consensus configuration and a given configuration relative to the within population variation known as the Mahalanobis distance in multivariate statistics as stated in [55] whilst referring to [56]. The advantage is that the model metric is given based on a population sample and does not need to be updated or estimated during superimposition. Furthermore, the analysis is a clear multivariate analysis taking into account the complete configuration and within population variation to determine harmony or lack thereof. The disadvantage, however, is that it becomes difficult to localize the discordant regions in the configuration and one is limited to keeping or rejecting the complete configuration as a whole. Furthermore, the Mahalanobis distance requires an inversion of the full covariance matrix that becomes computationally expensive and even practically impossible when dealing with a vast amount of landmarks.

An alternative for a population norm is the use of a Point Distribution Model (PDM) [10]. A PDM is a model for representing the mean geometry of a configuration and some statistical modes of geometric variation inferred from a training set of configurations. First a Generalized Procrustes-fit of all the configurations in the population is performed as before. The consensus configuration represents the mean geometry. Subsequently, a principal component analysis is done on the aligned data generating a ranked set of principal components representing the modes of geometric variation [57]. Doing so, allows for new configurations to be constructed based on the geometric mean and a linear combination of modes of geometric variation:

$$\hat{C} = \bar{C} + \sum_{k=1}^D U_k c_k \quad (13)$$

$\bar{C}$  is the consensus configuration,  $U_k$  is the  $k^{\text{th}}$  eigenvector or principal component of the covariance matrix  $\Omega$ ,  $c_k$  is the  $k^{\text{th}}$  PDM parameter and reflects the loading or contribution of the  $k^{\text{th}}$  principal component and  $D$  is the number of principal components used in the PDM. It is quite common to use only the top part of the principal components explaining, e.g., 98% of the total variance in the population (under the assumption that the last 2% corresponds to biologically insignificant variance due to random errors or artifacts). Finally, plugging (13) into (12) defines the PDM based transformation model  $T(\bar{C}, \theta)$  for a population norm. It starts from the consensus configuration  $\bar{C}$  and allows compensating for rotation, translation, scaling and typical variation differences as confounding variables to determine true differences in form, defined as outliers, with a given configuration  $C$ . Note that  $\hat{C}$  is the norm-equivalent configuration of  $C$  after ML-estimation. The full set of transformation model parameters  $\theta$  to estimate are three Euler angles, three translation directions, a scaling factor and  $D$  PDM parameters. Without favoring any rotations, translations or scaling the transformation model regularization is completely dependent on the PDM parameters only and the eigenvalues  $\alpha$  of the eigenvectors and can be defined as:

$$\|L(\theta)\|^2 = \frac{1}{2} \sum_{k=1}^D \frac{c_k^2}{\alpha_k^2} \quad (14)$$

This is the Mahalanobis distance of a PDM parameter setting favoring more plausible configurations within the population sample and statistics. Stated differently, it restricts the PDM from diverging too far away from the consensus configuration taking into account the within-population variation.

It is interesting to note that when using a PDM the covariance matrix of the within-population variation is moved from being used as the superimposition metric to being part of the superimposition transformation model. This allows for an additional superimposition metric to be defined like the modelling of i.i.d AWGN in the practical extended Procrustes ML-estimator. Doing so, results in an analysis that remains multivariate in nature whilst being able to determine local regions of discordancy. They are then defined as being significantly bigger than the estimated noise-level after compensation for within-population variation in the norm-equivalent besides the other confounding variables.

### Appendix C: Augmented Superimposition

#### Likelihood formulations

Given the observed landmarks a joint estimation is presented where besides the transformation parameters  $\theta$  also the outlier map of latent variables or unobserved data  $Z = \{z_j | j = 1, \dots, K\}$  is to be estimated. Following an Expectation-Maximization (EM) based estimation strategy, the likelihood of transformation parameters can be re-written in terms of the latent variables as:

$$\Pr(D|\theta) = \sum_z \Pr(D, Z|\theta) \quad (15)$$

EM produces a sequence of parameter updates  $\{\hat{\theta}^{(t)} | t = 0, 1, \dots\}$  by alternating two steps: the E-step and the M-Step. In the E-step, the latent variables are estimated given the observed landmarks and the current update of the transformation model parameters. This is achieved using the conditional expectation generating the so-called Q-function:

$$Q^{(t+1)} = \sum_Z \Pr(Z|D, \hat{\theta}^{(t)}) \log \Pr(D, Z|\theta) = E_{\Pr(Z|D, \hat{\theta}^{(t)})} [\log \Pr(D, Z|\theta)] \quad (16)$$

In the M-step, the Q-function is maximized generating a new update for the transformation parameters:

$$\hat{\theta}^{(t+1)} = \operatorname{argmax}_{\theta} Q^{(t+1)} \quad (17)$$

If the latent variable is modeled as a stochastic variable the outlier map  $Z$  becomes a random map with an associated prior-distribution  $\Pr(Z)$ . Under the assumption that the outlier map is independent from the transformation model parameters:

$$\Pr(D, Z|\theta) = \Pr(D|Z, \theta)\Pr(Z) \quad (18)$$

*Focusing on the second factor in (18):* Let  $P$  be the prior probability of having an inlier (for example the fraction of form differences thought to be generated by the inlier-process) and let  $1 - P$  be the prior probability of having an outlier. Then,

assuming  $Z$  to be i.i.d we can specify  $Pr(Z)$  as a product of i.i.d. Bernoulli distributed variables  $z_j$  :

$$Pr(Z) = \prod_{j=1}^K P^{z_j} (1 - P)^{(1-z_j)} \quad (19)$$

Note that extra constraints on the outliers can be incorporated here by choosing an alternative prior distribution  $Pr(Z)$ . For example, in [33] spatially coherent outliers are modeled by considering a binary Markov-Random-Field with associated Gibbs prior.

*Focusing on the first factor in (18):* The complete likelihood under the assumption of independent landmarks can be specified by conditioning individual residual likelihoods on the state of the binary-valued latent variable  $z_j$  signaling whether a local form difference was generated by the inlier-distribution  $Pr_i(\cdot)$  or the outlier-distribution  $Pr_o(\cdot)$ :

$$Pr(D|Z, \theta) = \prod_{j=1}^K Pr(d_j|z_j, \theta) \quad (20)$$

with

$$Pr(d_j|z_j, \theta) = \begin{cases} Pr_i(d_j|\theta), & \text{if } z_j = 1 \\ Pr_o(d_j), & \text{if } z_j = 0 \end{cases} \quad (21)$$

or, equivalently using a mixture model notation:

$$Pr(d_j|z_j, \theta) = Pr_i(d_j|\theta)^{z_j} Pr_o(d_j)^{(1-z_j)} \quad (22)$$

The extension of a negative log-likelihood assuming a perturbation model only into a complete negative log-likelihood assuming an outlier-process as well is obtained by taking the negative logarithm of (18) using (19-22):

$$l(\theta) = - \sum_{j=1}^K z_j (\log Pr_i(d_j|\theta) + P) + \sum_{j=1}^K (1 - z_j) (\log Pr_o(d_j|\theta) + 1 - P) \quad (23)$$

In the E-step the Q-function of (23) is obtained by dropping any constant terms independent from the parameters and by replacing the values of  $z_j$  with their expected conditioned values:

$$E_{Pr(Z|D, \hat{\theta}^{(t)})}[z_j] = \sum_{z_j \in \{0,1\}} z_j Pr(z_j|d_j, \hat{\theta}^{(t)}) = Pr(z_j = 1|d_j, \hat{\theta}^{(t)}) = b_j \quad (24)$$

Using the Bayes Rule we can write:

$$b_j = \frac{Pr(d_j|z_j = 1, \hat{\theta}^{(t)}) Pr(z_j = 1)}{\sum_{x \in \{0,1\}} Pr(d_j|z_j = x, \hat{\theta}^{(t)}) Pr(z_j = x)} \quad (25)$$

Based on (19) and (22) we then obtain:

$$b_j = \frac{Pr_i(d_j|\hat{\theta}^{(t)})P}{Pr_i(d_j|\hat{\theta}^{(t)})P + Pr_o(d_j|\hat{\theta}^{(t)})(1 - P)} \quad (26)$$

#### A practical extended Procrustes ML-estimator

As an example ML-estimator extension, we start from the Procrustes ML-estimator assuming a zero-mean Gaussian perturbation model with i.i.d landmark displacements (using the negative log-likelihood in (19)). Hence, the model metric reflects additive white Gaussian noise (AWGN) with a single noise parameter  $\sigma$ . This constitutes the inlier-distribution  $Pr_i(x)$ . For the outlier-distribution  $Pr_o(x)$  several choices can be considered. If some knowledge about the associated distribution of the outlier generating process is given, than this can be used. In most cases, however, this knowledge is not known and could be estimated as well [33,36]. However, we simply assume the outlier-distribution to be uniform such that:

$$Pr_o(x) = \delta \text{ with } 0 < \delta < 1 \quad (27)$$

In this simplified case, the negative Q-function, written down from an M-estimator point of view, has a  $\rho$ -function equal to:

$$\rho(d_j) = \frac{b_j}{2\sigma^2}d_j^2 + b_j \log \sqrt{2\pi}\sigma - (1 - b_j) \log \delta \quad (28)$$

with

$$b_j = \frac{Pr_i(d_j|\hat{\theta}^{(t)})P}{Pr_i(d_j|\hat{\theta}^{(t)})P + \delta(1 - P)} \quad (29)$$

From (28) we see that for inlier landmarks ( $b_j = 1$ ) the extended Procrustes ML-estimator is equal to the original Procrustes ML-estimator. Furthermore, an outlier landmark ( $b_j = 0$ ) does not contribute to the estimation of the transformation parameters  $\theta$  or model metric parameter  $\sigma$  because  $\log \delta$  is independent of both. From (29), we see that the estimation  $\sigma$  of the latent variables during superimposition is dependent on the estimated of the inlier-distribution and also on the prior outlier distribution parameters  $P$  and  $\delta$ . In order to simplify the choice of prior values, we combine these inter-dependent parameters into a single parameter  $\lambda = \delta(1 - P)/P$  such that:

$$b_j = \frac{Pr_i(d_j|\hat{\theta}^{(t)})}{Pr_i(d_j|\hat{\theta}^{(t)}) + \lambda} \quad (30)$$

Furthermore,  $\log \delta$  can be substituted with  $\log \lambda$  in (28) without any effect to the parameter estimation. Changing the parameter  $\lambda$  changes the amount of outliers versus inliers. A proper choice or fine-tuning of this parameter is therefore required, but can prove to be challenging. This choice is made easier and, more importantly, statistically relevant following [37]. Equation (30) reflects the posterior probability of a local

residual to belong to the inlier-distribution, called the inlier-belief. The outlier-belief can then be defined similarly as:

$$1 - b_j = \frac{\lambda}{Pr_i(d_j|\hat{\theta}^{(t)}) + \lambda} \quad (31)$$

The outlier-belief exceeds the inlier-belief if  $b_j < 0.5$  or  $Pr_i(d_j|\hat{\theta}^{(t)}) < \lambda$  which is equivalent to  $MD_j^2 > -2 \log \lambda \sqrt{2\pi} \sigma$  with  $MD_j = d_j/\sigma$  the Mahalanobis distance. In statistics, an observation is called abnormal or atypical with respect to a given normal distribution if its (Mahalanobis) distance exceeds a predefined threshold. Because of its dependence on  $\sigma$  in, the Mahalanobis distance threshold above which a local form difference is considered abnormal changes when  $\sigma$  changes or is updated. Furthermore, observations are more easily rejected from classes with a broad distribution (e.g. overall big landmark displacements) than from classes with a narrow one, making the choice of  $\lambda$  dependent on the current superimposition problem, leading to a different  $\lambda$  setting for every different configuration to superimpose. Because of these problems it is not clear how  $\lambda$  should be chosen. Ideally, a spatial displacement should be considered abnormal if  $MD_j > \kappa$ , where  $\kappa > 0$  is an explicit Mahalanobis distance threshold that is equal for all normal inlier-distributions alike. Therefore, taking into account the dependence on  $\sigma$ ,  $\lambda$  is replaced by:

$$\lambda = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\kappa^2\right) \quad (32)$$

$\lambda$  is now inlier-distribution dependent and re-parameterized using a more interpretable prior outlier-parameter  $\kappa$ . The actual choice of  $\kappa$  is equivalent to the choice of a statistical significance level above which local form differences are considered atypical (outlier) compared to a typical Gaussian (inlier-) distribution.

#### Appendix D: Computational results

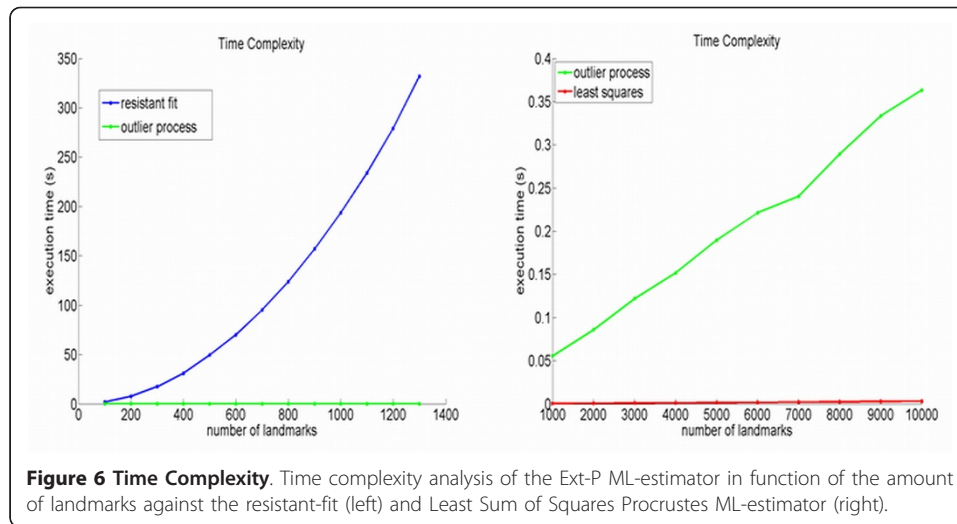
In this section the results obtained using an outlier process as formulated in the Ext-P ML-estimator are numerically compared to two commonly used alternative estimators: the popular Procrustes least sum of squares (LLS) and the resistant-fit or repeated median estimator.

Using increasing amounts of landmarks from the surgical intervention superimposition problem depicted in Figure 4(a) and 4(c), a time-complexity analysis is given in Figure 6. It is obvious that the resistant-fit suffers from an exponential time-complexity, whereas the LLS as well as the Ext-P ML estimator remain computationally practical even up to 10.000 landmarks.

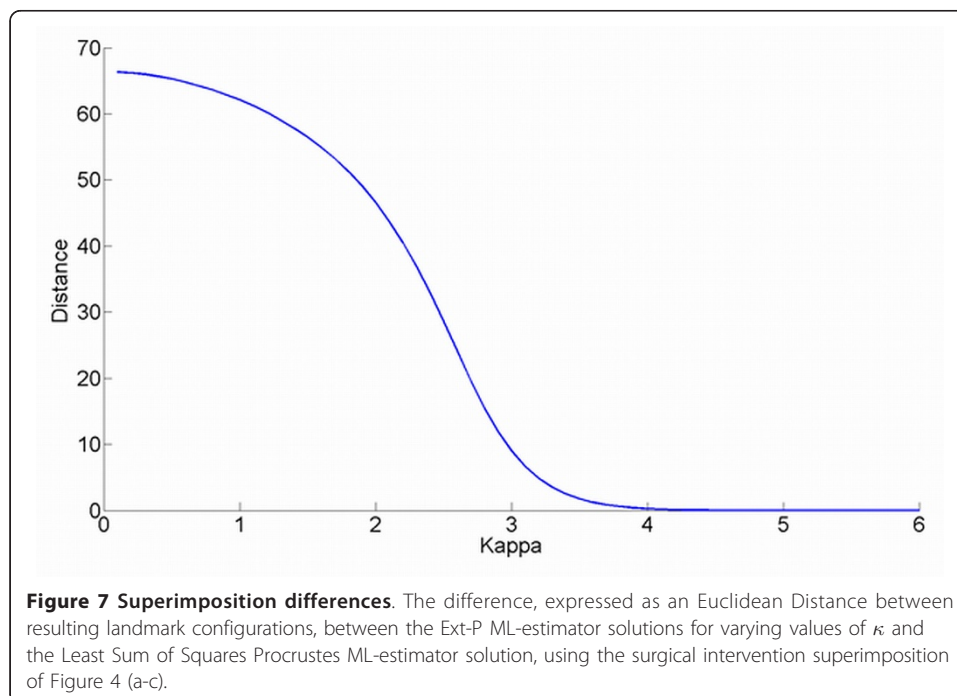
Using the same surgical intervention superimposition problem, the Euclidean distance between landmark configurations after superimposition as a function of  $\kappa$  (Kappa) are given in Figure 7. For lower values of  $\kappa$ , the Ext-P solution differs from the LSS solution. A clear bending point is observed for  $\kappa$  in-between 2 to 3, corresponding to the typically used p-values of 0.05 and 0.001 for significance assessment. As expected, for higher values of  $\kappa$ , the Ext-P solution equates the LSS solution.

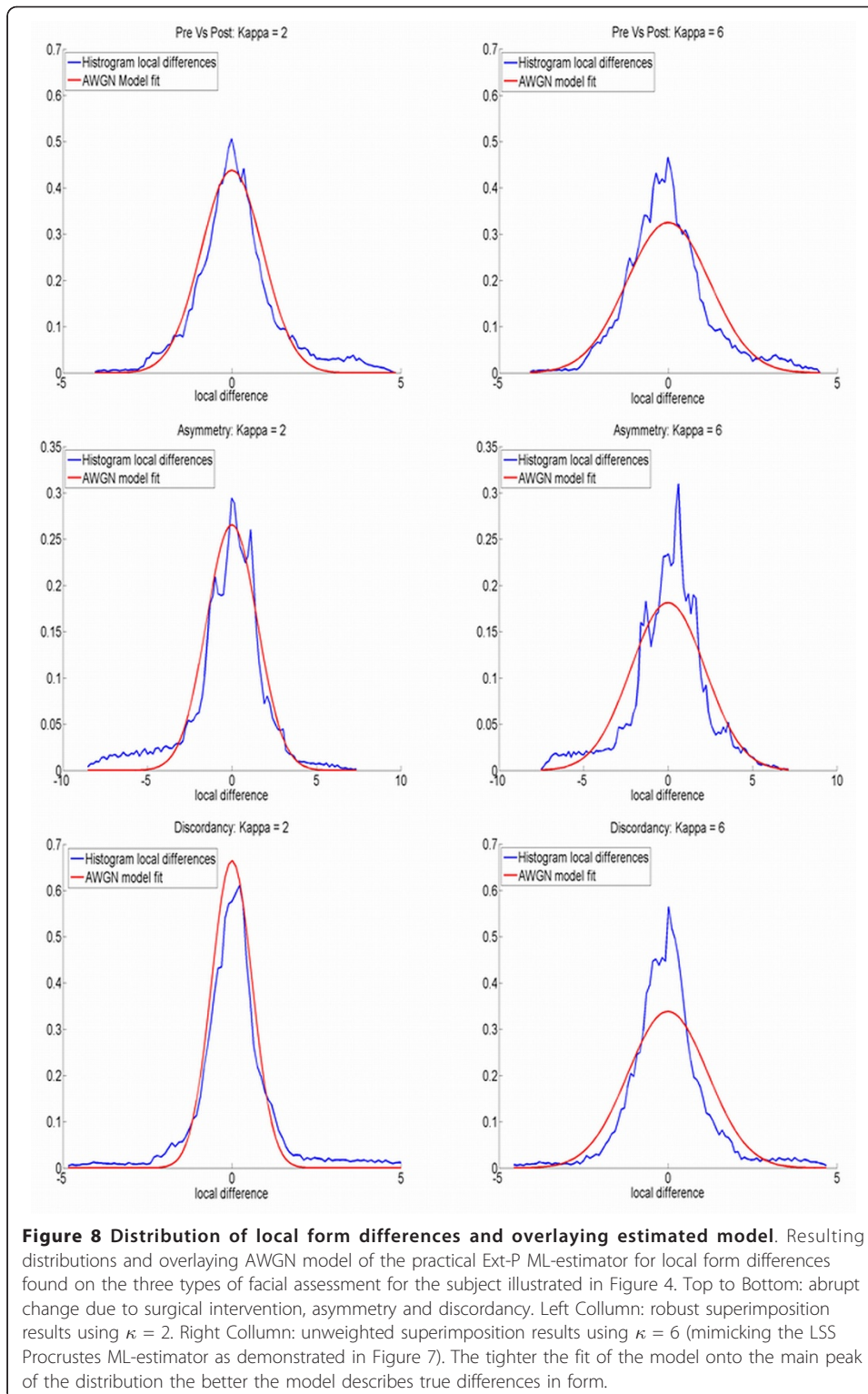
Finally, for the three types of assessment (abrupt changes, asymmetry and discordancy) in Figure 4, the Ext-P ML-estimator is used with two different values of Kappa;





$\kappa = 2$  (robust superimposition) and  $\kappa = 6$  (un-weighted superimposition, mimicking the Procrustes LSS superimposition (see Figure 7)). Histograms of pooled X, Y, and Z directed local form differences after superimposition are visualized in Figure 8, with the respective estimated AWGN models overlaid. The tighter the model-fit onto the histogram, the better the respective model assumption and therefore the better the superimposition. For all three assessments, the robust superimposition with  $\kappa = 2$ , provides a better model in the case of a presented facial abnormality.





#### Acknowledgements

This work is supported by the Australian Research Council (ARC) [grant number 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 like to thank, Dr Paul Sillifant, Dr David Gillett, Dr Matthew Hansen and Dr Gareth Baynam for the input on the data used in this work.

#### Author details

<sup>1</sup>Department of Electrical Engineering - processing of speech and images - medical image computing, KU Leuven, Herestraat 49 Bus 7003, Leuven, 3000, Belgium. <sup>2</sup>Cranio-Maxillo-Facial Unit, Princess Margaret Hospital for Children, Perth, Australia. <sup>3</sup>Melbourne Dental School, The University of Melbourne, Melbourne, Australia. <sup>4</sup>IBBT-KU.Leuven Future Health Department, KU Leuven, Leuven, Belgium.

#### Authors' contributions

PC, DV and PS devised the technical methods and details. PC, MW and JC explored the biomedical applications, re-worked and applied the methodology on all the cases. MW collected the clinical and required facial data to build up a population norm. PC, DV and JC situated everything within morphometrics and defined the concept of dysmorphometrics. KD, as part of her master's dissertation, generated the computational results. PC wrote the manuscript with input and revisions from KD, MW, JC, DV and PS. All authors have read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 26 November 2011 Accepted: 6 February 2012 Published: 6 February 2012

#### References

1. Blackith R, Reyment R: *Multivariate Morphometrics* New York: Academic Press; 1971.
2. Penrose L: **Distance, size and shape**. *Annals of Eugenics* 1954, **18**:337-343.
3. Lestrel P: *Morphometrics for the life sciences Volume 7 of Recent advances in human biology* World Scientific Publishing Co. Pte. Ltd; 2000.
4. Zelditch M, Swiderski D, Sheets H, Fink W: *Geometric morphometrics for biologists: A primer* New York and London: Elsevier Academic Press; 2004.
5. Richtsmeier JT, Deleon VB: **Morphological integration of the skull in craniofacial anomalies**. *Orthodontics and Craniofacial Research* 2009, **12**:149-158.
6. Chandola V: **Anomaly detection: A survey**. *ACM computing surveys* 2009, **41**(3).
7. Adams D, Rohlf F, Slice D: **Geometric Morphometrics: ten years of progress following the revolution**. *Italian Journal of Zoology* 2004, **71**:5-16.
8. Lestrel P: *Fourier Descriptors and their applications in biology* Cambridge University Press; 1997.
9. Slice D: **Geometric Morphometrics**. *The annual review of Anthropology* 2007, **36**:261-281.
10. Dryden I, Mardia K: *Statistical Shape Analysis* John Wiley and Sons; 1998.
11. Richtsmeier JT, Deleon VB, Lele SR: **The Promise Of Geometric Morphometrics**. *American Journal of Physical Anthropology* 2002, **45**:63-91.
12. Rohlf F, Slice D: **Extensions of the Procrustes method for the optimal superimposition of landmarks**. *Systematic Zoology* 1990, **39**:40-59.
13. Siegel A, Benson R: **A robust comparison of biological shapes**. *Biometrics* 1982, **38**:341-350.
14. Bookstein F: *Morphometric tools for landmark data: geometry and biology* Cambridge: Cambridge University Press; 1991.
15. Cheverud J, Lewis J, Bachrach W, Lew W: **The measurement of form and variation in form: an application of three-dimensional quantitative morphology by finite-element methods**. *American Journal of Physical Anthropology* 1983, **62**:151-165.
16. Farkas L: *Anthropometry of the Head and Face in Medicine* New York: Elsevier; 1981.
17. Thomson D: *On growth and form: the complete revised edition* New York: Dover; 1992.
18. Lele SR, Richtsmeier JT: *An invariant approach to the statistical analysis of Shapes* Boca Raton: Chapman & Hall/CRC; 2001.
19. Smeets D, Claes P, Vandermeulen D, Clement JG: **Objective 3D face recognition: Evolution, approaches and challenges**. *Forensic Science International* 2010, **201**:125-132.
20. Thomas CDL: *Three-Dimensional Quantification of Facial Shape* London: Elsevier Academic Press; 2005.
21. Andresen PR, Bookstein F, Conradsen K, Ersboll BK, Marsh J, Kreiborg S: **Surface-bounded growth modeling applied to human mandibles**. *IEEE Transactions on Medical Imaging* 2000, **19**(11):1053-1063.
22. Claes P, Walters M, Vandermeulen D, Clement JG: **Spatially-dense 3D facial asymmetry assessment in both typical and disordered growth**. *Journal of Anatomy* 2011, **219**:444-455.
23. Claes P, Walters M, Clement JG: **Improved Facial Outcome Assessment using a 3D Anthropometric Mask**. *International journal of oral and maxillofacial surgery* 2011, DOI: 10.1016/j.ijom.2011.10.019.
24. Wei R, Claes P, Walters M, Wholley C, Clement JG: **Augmentation of linear facial anthropometrics through modern morphometrics: A facial convexity example**. *Australian Dental Journal* 2011, **56**:141-147.
25. Claes P: **A robust statistical surface registration framework using implicit function representations: Application in craniofacial reconstruction**. *Phd thesis* KULeuven, Belgium, Department of Electrical Engineering; 2007.
26. Theobald D, Wuttke D: **Empirical Bayes hierarchical models for regularizing maximum likelihood estimation matrix Gaussian Procrustes problem**. *Proceedings of the national Academy of science of the United States of America* 2006, **103**(49):18521-18527.
27. Goodall C: **Procrustes Methods in the Statistical Analysis of Shape**. *Journal of the Royal Statistical Society. Series B (Methodological)* 1991, **53**(2):285-339.
28. Bookstein F: **Biometrics, biomathematics and the morphometric synthesis**. *Bulletin of Mathematical Biology* 1996, **58**(2):313-365.

29. Hampel F, Ronchetti E, Rousseeuw P, Stahel W: **Robust statistics: The approach based on influence functions.** *Wiley series in Probability and Statistics* New York: John Wiley and sons; 2005.
30. Rousseeuw P: **Least median of squares regression.** *Journal of American Statistical Association*, 79:871-880 1984, 871-880.
31. Mechelke M, Habeck M: **Robust probabilistic superposition and comparison of protein structures.** *BMC Bioinformatics* 2010, **11**:13.
32. Black M, Rangarajan A: **The outlier process: unifying line processes and robust statistics.** *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '94* 1994, 15-22.
33. Fransens R, Strecha C, Van Gool L: **Robust estimation in the presence of spatially coherent outliers.** *Conference on Computer Vision and Pattern Recognition Workshop, CVPRW 2006*, 102-102.
34. Chuong BD, Serafim B: **What is the expectation maximization algorithm?** *Nature Biotechnology* 2008, **26**(8):897-899.
35. Dempster A, Laird N, Rubin D: **Maximum Likelihood from Incomplete Data via the EM Algorithm.** *Journal of the royal statistical society* 1977, **39**:1-38.
36. Hasler D, Sbiaz L, Susstrunk S, Vetterli M: **Outlier Modeling in Image Matching.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2003, **25**(3):301-315.
37. Van Leemput K, Maes F, Vandermeulen D, Colchester A, Suetens P: **Automated segmentation of multiple sclerosis lesions by model outlier detection.** *IEEE Transactions on Medical Imaging* 2001, **20**(8):677-688.
38. Ferrario V, Sforza C, Miani A, Serrao G: **A three-dimensional evaluation of human facial asymmetry.** *Journal of Anatomy* 1995, **186**:103-110.
39. Yu Z, Mu X, Feng S, Han J, Chang T: **Flip-Registration procedure of Three-dimensional laser surface scanning images on quantitative evaluation of facial asymmetries.** *The journal of craniofacial surgery* 2009, **20**:157-160.
40. Klingenberg C, Barluenga M, Meyer A: **Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry.** *Evolution* 2002, **56**(10):1909-1920.
41. Mardia K, Bookstein F, Moreton IJ: **Statistical assessment of bilateral symmetry of shapes.** *Biometrika* 2000, **87**(2):285-300.
42. Walters M, Claes P, Kakulas E, Clement JG: **3D Facial Asymmetry Assessment in Hemimandibular Hyperplasia and Hemimandibular Elongation Anomalies.** *International journal of oral and maxillofacial surgery* 2012, accepted 23/12/2011.
43. Hammond P, Hutton TJ, Allanson JE, Buxton B, Campbell LE, Clayton-Smith J, Donnai D, Karmilo-Smith A, Metcalfe K, Murphy KC, Patton M, Pober B, Prescott K, Scambler P, Shaw A, Smith ACM, Stevens AF, Temple IK, Hennekam R, Tassabehji M: **Discriminating power of localized three-dimensional facial morphology.** *American Journal of Human Genetics* 2005, **77**(6):999-1010.
44. William J, Berger T, Elston D: *Andrews' diseases of the skin: clinical dermatology Saunders*, 10 2005.
45. Winchester B, Vellodi A, Young E: **The molecular basis of lysosomal storage diseases and their treatment.** *Biochemical Society Transactions* 2000, **28**(2):150-154.
46. Cory R, Clayman D, Faillace W, McKee S, Gama C: **Clinical and radiologic findings in progressive facial hemiatrophy (Parry-Romberg syndrome).** *American Journal of Neuroradiology* 1997, **18**(4):751-757.
47. Baynam G, Claes P, Craig JM, Goldblatt J, Kung S, Le Souef P, Walters M: **Intersections of epigenetics, twinning and developmental asymmetries: Insights into monogenic and complex diseases and role for 3D Facial Analysis.** *Twin Research and Human Genetics* 2011, **14**(4):305-315.
48. Baynam G, Walters M, Claes P, Le Souef P: **3D facial analysis can investigate vaccine responses.** *Medical hypotheses* 2012, accepted 11/01/2012.
49. Collet D: *Modeling binary data.* 2 edition. Chapman and Hall/CRC; 2003.
50. Hajeer MY, Mao Z, Millett DT, Ayoub AF, Siebert JP: **A new three-dimensional method of assessing facial volumetric changes after orthognatic treatment.** *Cleft Palate-Craniofacial Journal* 2005, **42**(2):113-120.
51. Tang X, Chen J, Moon Y: **Accurate 3D face registration based on the symmetry plane analysis on nose regions.** *16th European Signal processing conference (EUSIPCO)* 2008.
52. Barry SJE, Bowman AW: **Linear mixed models for longitudinal shape data with applications to facial modeling.** *Biostatistics* 2008, **9**(3):555-565.
53. Benz M, Laboureaux X, Maier T, Nkenke E, Seeger S, Neukam F, Häusler G: **The symmetry of faces.** *Vision Modeling and Visualization* 2002, 43-50.
54. Shaveesh AI, Clement JG, Thomas CDL, Bankier A: **Construction and use of facial archetypes in anthropology and syndrome diagnosis.** *Forensic Science International* 2006, **159**:175-185.
55. Klingenberg C, Monteiro R: **Distances and Directions in Multidimensional Shape Spaces: Implications for morphometric applications.** *Systematic Biology* 2005, **54**(4):678-688.
56. Mardia K, Kent JT, Bibby JM: *Multivariate analysis* London: Academic Press; 1979.
57. Taylor C, Cooper DH, Graham J: **Training models of shape from sets of examples.** *Proc BMVC92, Springer-Verlag* 1992, 9-18.

doi:10.1186/1742-4682-9-5

**Cite this article as:** Claes et al.: Dysmorphometrics: the modelling of morphological abnormalities. *Theoretical Biology and Medical Modelling* 2012 **9**:5.