Automatic strategy for extraction of anthropometric measurements for the diagnostic and evaluation of deformational plagiocephaly from infant's head models

Bruno Oliveira*[a,b], Helena R. Torres*[a,b], Fernando Veloso[a,b], Estela Vilhena[a], Nuno F. Rodrigues[a], Jaime C. Fonseca[b], Pedro Morais[a], João L.Vilaça[a]

[a]2Ai – Polytechnic Institute of Cávado and Ave, Barcelos, Portugal; [b]Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal

ABSTRACT

Deformational Plagiocephaly (DP) refers to an asymmetrical distortion of an infant's skull resulting from external forces applied over time. The diagnosis of this condition is performed using asymmetry indexes that are estimated from specific anatomical landmarks, whose are manually defined on head models acquired using laser scans. However, this manual identification is susceptible to intra-/inter-observer variability, being also time-consuming. Therefore, automatic strategies for the identification of the landmarks and, consequently, extraction of asymmetry indexes, are claimed. A novel pipeline to automatically identify these landmarks on 3D head models and to estimate the relevant cranial asymmetry indexes is proposed. Thus, a template database is created and then aligned with the unlabelled patient through an iterative closest point (ICP) strategy. Here, an initial rigid alignment followed by an affine one are applied to remove global misalignments between each template and the patient. Next, a non-rigid alignment is used to deform the template information to the patient-specific shape. The final position of each landmark is computed as a local weight average of all candidate results. From the identified landmarks, a head's coordinate system is automatically estimated and later used to estimate cranial asymmetry indexes. The proposed framework was evaluated in 15 synthetic infant head's model. Overall, the results demonstrated the accuracy of the identification strategy, with a mean average distance of 2.8±0.6 mm between the identified landmarks and the ground-truth. Moreover, for the estimation of cranial asymmetry indexes, a performance comparable to the inter-observer variability was achieved.

Keywords: ICP algorithm, landmark detection, anthropometric measurements, infant's head models, deformational plagiocephaly

1. INTRODUCTION

Deformational plagiocephaly (DP) is a condition characterized by a cranial asymmetry of an infant's skull caused by repeated external pressures, which result from restrictive intrauterine environment, unusual birth position, and supine sleeping [1]. Besides the DP-related cosmetic concerns, infants who present this condition can experience developmental delay [2]. Therefore, a correct diagnosis and evaluation of DP on infants is essential for the treatment outcome.

The evaluation and diagnosis of DP use anthropometric skull measurements to estimate clinical cranial asymmetry indexes [1]. To obtain the required measurements, 3D laser scanning is traditionally performed to construct a 3D patient head's model, being later manually processed to extract specific anatomical landmarks. These landmarks allow the estimation of the coordinate system of the head and further the extraction of the required anthropometric measurements [3]. However, manual identification of these landmarks is not straightforward, being this process time-consuming and susceptible to intra- and inter-observer variability. Therefore, strategies for the automatic identification of head model landmarks (for the extraction of anthropometric measurements) are claimed.

This paper aims to develop a fully automatic strategy for the computation of the anthropometric skull measurements and its respective cranial asymmetry indexes. Such tool can then be used to define a baseline for clinical documentation for the evaluation and diagnosis of DP [3]. This paper is structured as follows. Section 2 describes the methodology of the proposed framework, with implementation details presented in Section 3. Section 4 introduces the experiments, followed by results in Section 5. In Section 6, the performance of each module of the proposed framework is discussed. The conclusions are given in Section 7.
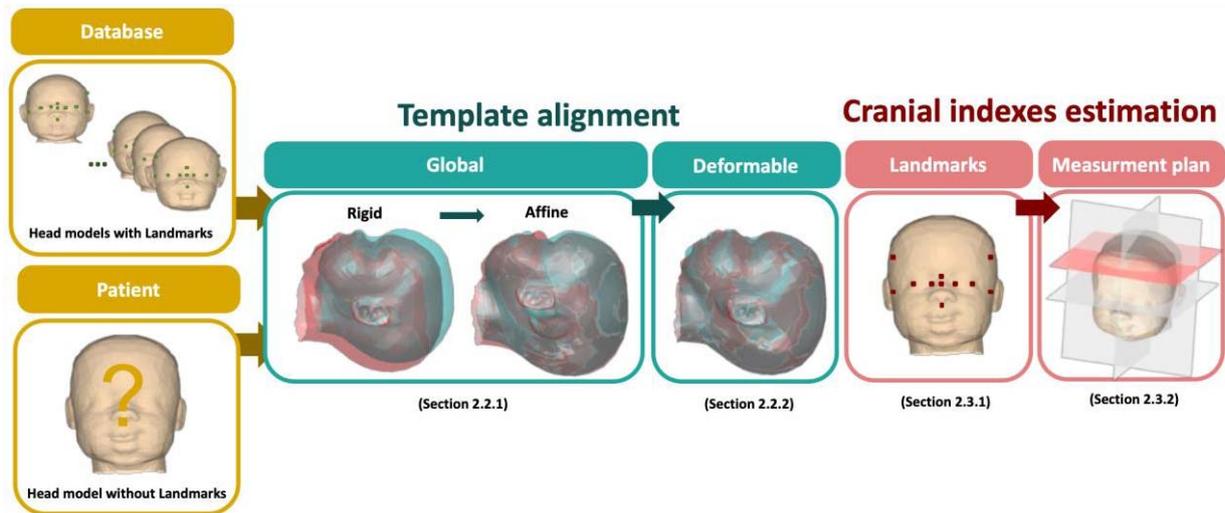
Figure 1– General overview of the proposed methodology.

2. METHODS

2.1 General overview

Figure 1 shows an overview of the proposed methodology, which relies on two conceptual blocks. The first relies on a template alignment approach and can be divided into two stages. Each template corresponds to a 3D head model with manual labelling of the relevant anthropometric landmarks. Initially, and to compensate high misalignments, rigid and affine transform models (section 2.2.1) are used to correct translation, rotation and scaling differences between each template and the target case. Then, deformable models (section 2.2.2) are applied to refine the template to the patient anatomy. This entire block is repeated for the set of cases available in the template database. The second block relies on the extraction of cranial asymmetry indexes and is divided into two stages. The first module applies a local weigh arithmetic mean to get, from all the template candidates, the final 3D landmarks positioning for the unlabelled patient (section 2.3.1). Next, in the second module, these landmarks are used to compute a head's coordinated system and further to extract the anthropometric measurements and cranial asymmetry indexes (section 2.3.2).

2.2 Template alignment

2.2.1 Global alignment

The ICP algorithm is the traditional strategy for effective point-cloud alignment/registration using global transformations [4]. However, the ICP is susceptible to local minima, requiring strong initialization. In this sense, to improve the robustness of the proposed pipeline to the initialization, a rigid alignment using the Global-ICP (Go-ICP) strategy was applied [5]. In detail, the Go-ICP searches for the optimal rigid alignment from the entire 3D motion space using a branch-and-bound scheme, which previously proved its high robustness in different situations [5]. In this sense, the optimal rigid alignment between each template and the target patient is potentiated, minimizing, therefore, the dependence to the initialization. However, possible global misalignments may still occur, due to the different sizes of the heads. Therefore, an affine transformation is later used to compensate size differences. Here, the traditional ICP algorithm using affine transformation models was used [6].

2.2.2 Deformable alignment

Typically, we must consider the anatomical difference among infants, namely the gender, age or breed [2] as well as the different facial features. Since the previous steps just allow rigid and affine transformations (globally adjusting the different surfaces), a deformable ICP alignment is afterwards used to refine the template information to the patient shape. Several strategies have been developed to extend the ICP algorithm to allow deformable alignments. Here we propose to use a non-rigid ICP strategy with local affine regularization, which minimizes the difference between transformations of neighbour vertexes [7]. Therefore, specific anatomical characteristics of individual templates may be preserved while the

remaining are aligned. With this strategy we may restrict the deformation when uncommon structures arise, maintaining the relative positioning of each landmark.

2.3 Cranial indexes estimation

2.3.1 Landmarks identification

After the alignment block, several candidate landmark positions (*i.e.* one per template) are obtained. To find the final optimal 3D landmark position, a strategy to combine all candidates is applied. Since the quality of the alignment may vary between cases, a local weigh strategy is implemented. Here, we start to define a region-of-interest (ROI) around each landmark, and then measure the alignment performance (of each template) at each region. Then, the quality of the alignment in each region is used as weight (*i.e.* computed based on the distance between the aligned template and the target case for each ROI), and the final position is estimated through a weighted arithmetic mean of all the candidate positions.

2.3.2 Measurement plan computation

After landmarks identification (Figure 2A), the head's coordinate system is automatically computed, as described in the original guidelines [8]. For it, the origin (*i.e.* midpoint between the right/left tragions) is initially calculated. The X-axis (*i.e.* line from the origin through the left tragion), Y-axis (*i.e.* line from the origin through the sellion), and Z-axis (*i.e.* line from the origin through the top of the head, such that it is perpendicular to the X- and Y-axis) are subsequently defined. Then, the opisthocranion, which is the most posterior point on the parallel plan to the y-axis at the glabella point, is found. Finally, the measurement plan is computed at the height of the opisthocranion, being parallel to the plan defined by the X and Y-axis. The measurement plan is used as a reference to calculate the anthropometric measures, namely oblique transcranial diagonals, cranial circumference, cranial length, and cranial width (Figure 2B). These measurements are used to estimate the cranial asymmetry indexes, namely cephalic index, diagonal difference, posterior cranial asymmetry, and interior cranial asymmetry. With these indexes, the physician can describe the infant' deformities and, thus, their condition [8].

3. IMPLEMENTATION DETAILS

Initially, a pre-processing stage was performed to smooth the templates and patient mesh (limiting it to 5000 vertices). Moreover, the template's size was normalized between -5 and 5 cm. Furthermore, to neglect templates with high anatomical differences to the unlabelled case, a template selection strategy, applied before the affine alignment, was used. For the template selection strategy, the point-to-point distance (P2P) between the template and the patient was measured. Then, the 25% of the atlases with lowest errors were selected for the remaining stages of the pipeline. Additionally, a ROI with the 25 closest vertices was defined around each landmark and the P2P was used to define the quality of the alignment of each template per landmark in section 2.3.1. All ICP parameters were experimentally selected to achieve a stable performance.
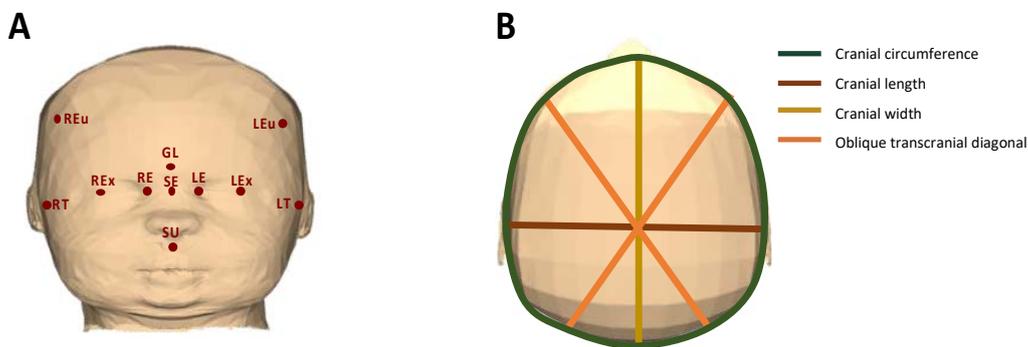
Figure 2 – Estimated anthropometric skull measurements from specific anatomical landmarks for DP evaluation: (A) Anatomical landmarks, namely glabella (GL), sellion (SE), subnasal (SU), right exocanthion (REx), left exocanthion (LEx), right endocanthion (RE), left endocanthion (LE), right tragion (RT), left tragion (LT), right eurion (REu), and left eurion (LEu); (B) Anthropometric skull measurements, namely oblique transcranial diagonals, cranial circumference, cranial length, and cranial width.

4. EXPERIMENTS

4.1 Data description

To evaluate the performance of the proposed methodology, a synthetic database with 65 infant head models (*i.e.* 50 for the template database and 15 as unknown patients' models, *i.e* target cases) was used. This database was created using the *MakeHuman* software and manually labelled by two different observers. Here, the facial features, the shape and the size of the head were randomly selected to get variability in the dataset. Moreover, these features were restricted to be representative of infant head models. In this work 11 landmarks were evaluated, namely glabella (GL), sellion (SE), subnasal (SU), right exocanthion (REx), left exocanthion (LEx), right endocanthion (RE), left endocanthion (LE), right tragion (RT), left tragion (LT), right eurion (REu), and left eurion (LEu) (Figure 2A).

4.2 Evaluation strategies

Initially, the performance of the landmarks' identification strategy at each module of the pipeline was evaluated using the average distance. The results for the alignment modules were produced after average all the candidate landmarks' positioning and the statistical significance of the results was tested using a Wilcoxon matched-pairs signed rank test (p < 0.05 being considered statistically significant). Additionally, a clinical validation was also performed by comparing the estimated cranial indexes with the ones obtained from the manual interaction of two observers through a Bland-Altman plot analysis using a two-tailed *t*-test against 0 and a two-tailed F-Test against the inter-observer ($p > 0.05$ being considered statistically not significant).

5. RESULTS

5.1 Landmarks' identification performance

The average distance of the entire database after global module, separately for the rigid and affine alignments (section 2.2.1), deformable (section 2.2.2), and landmarks' detection module (section 2.3.1) is presented in Figure 3. Looking for the boxplots, statistically significant improvements were observed between each module, therefore, validating the proposed pipeline. Moreover, Table 1 quantifies the performance of the proposed landmarks identification strategy and compare it against the inter-observer variability. Here, mean average distances of 2.8±0.6 and 4.6±1.8 mm were obtained for the proposed methodology and the inter-observer variability, respectively. Figure 4, shows the accuracy for each module of the proposed pipeline, using different distance thresholds (*i.e.* between 0-1 cm). Overall, the results showed the robustness of the proposed methodology, with average distances lower than the manual variability differences. In detail, the worst results were obtained for landmarks that are placed in smooth head regions with no specific particularities, namely REu and LEu.

5.2 Clinical Validation

The evaluation of relevant cranial indexes was performed through a Bland-Altman plot analysis (Figure 5). The inter-observer variability at these indexes is also presented. Overall, narrow limits-of-agreements (LOAs) were obtained by
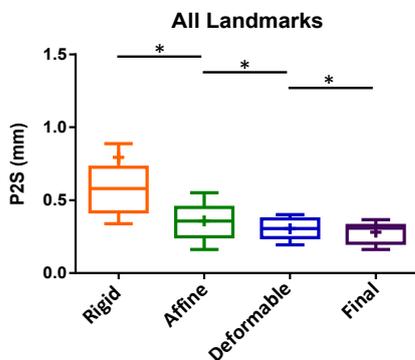
Figure 3 – Comparison of the performance of the proposed method throughout its different stages. ∗p < 0.05, Wilcoxon matched-pairs signed rank test
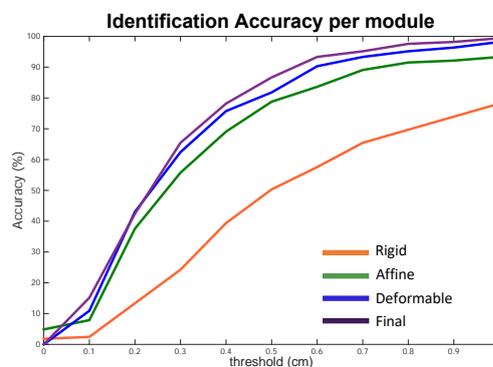
Figure 4 – Accuracy of the proposed method for different error threshold per method's module.

Table 1. Performance (in mm) of the proposed strategy per each landmark using the average distance ($\mu \pm \sigma$). Moreover. the inter-observer variability is also presented

	GL	SE	SU	REx	LEx	RE	LE	RT	LT	REu	LEu	Mean
Proposed	2.1±1.3	1.9±1.6	1.5±0.8	2.3±1.3	2.0±0.8	1.9±1.2	2.0±1.2	4.0±1.7	4.3±2.1	4.1±1.8	4.8±3.0	2.8±0.6
Manual Variability	4.5±2.5	2.8±1.3	2.4±1.0	7.0±1.5	5.9±3.2	2.1±0.7	2.4±1.1	2.5±1.5	2.3±1.2	10.5±3.0	8.3±3.2	4.6±1.8
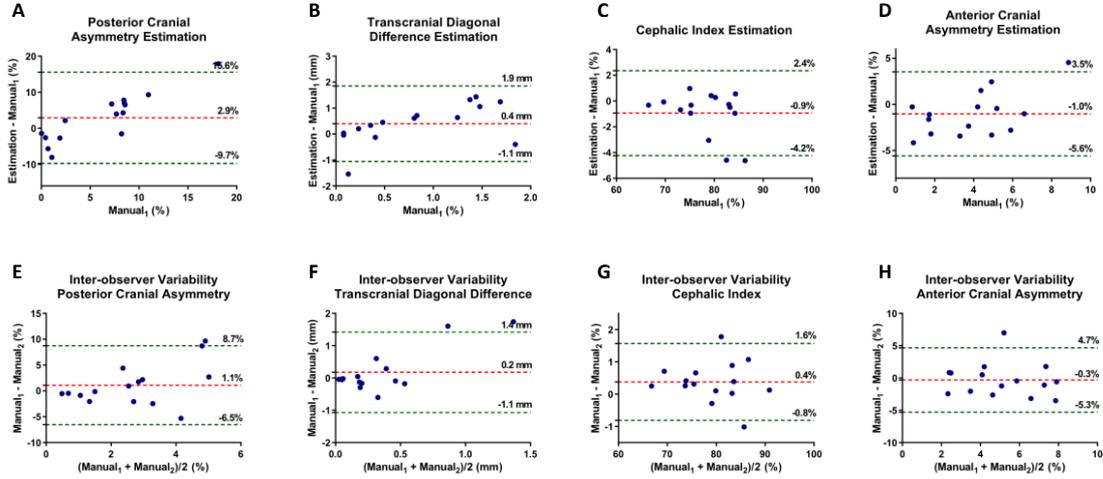
Figure 5 - Bland-Altman plots for the cranial indexes analysis between indexes estimated using automatic and manual approaches. (A),(B),(C),(D) represent the agreement between automatic and manual approaches for the different cranial indexes, while (E),(F),(G),(H) shows the agreement between different observers.

the automatic method with biases (*i.e.* mean error) not statistically significant. When compared the performance of the automatic method against the inter-observer variability, a similar performance was achieved. Statistically significant differences were only found for the cephalic index.

6. DISCUSSION

Deformational Plagiocephaly refers to cranial deformities occurring in infants that can result in developmental delay as well as cosmetic concerns over their life. Therefore, a correct diagnostic of this medical condition, which is normally performed using asymmetry indexes, is fundamental. In the current work a fully automatic strategy for the identification of head landmarks, and, consequently, extraction of cranial asymmetry indexes, was presented. The method was evaluated for the identification of 11 anatomical landmarks and estimation of 4 cranial indexes, and showed a performance comparable to the inter-observer variability.

Figure 3 and Figure 4 presents the error in each step of the landmarks' identification strategy showing statistically significant improvements between subsequent stages of the method, validating, therefore, the different stages of the proposed pipeline. Firstly, analyzing the boxplots for the rigid and affine stages, it is possible to verify that the low P2S distances obtained corroborate that the proposed strategy performed an accurate translation, rotation and scaling between the templates and the target head models. Secondly, the results of the deformable stage showed the capacity of the proposed strategy to refine the results obtained in the previous stages. Moreover, an accuracy of 86.7% and 99.3%, for a threshold of 0.5 and 1 cm respectively, was obtained for the final alignment, showing the robustness of the proposed strategy.

Looking for Table 1, it is possible to see that the proposed strategy showed better results than the manual variability for all landmarks, excepting for the left/right tragion, being a mean average distance of 2.8 mm obtained for all the landmarks. Indeed, most of the evaluated landmarks are located in anatomical places with specific anatomical features (i.e. shape, curvature), thus, potentiating the performance of the ICP methodology. In opposition, the identification of

landmarks located in specific regions with minimal anatomical features, hampering an optimal alignment, was sub-optimal. This can be seen for the left/right eurion, located in the central area of the parietal side, which is a smooth region of the head.

In addition to the assessment of the accuracy of the landmark identification, clinical validation of the cranial asymmetry indexes was also performed. When assessing the agreement between the cranial indexes estimated from automatic and manual landmarks' identification strategy, no statistically significant biases (red dashed line) were observed for the different measurements ($p > 0.05$ in a two-tailed t-test against 0). Moreover, narrow LOAs (green dashed lines) were obtained for all measurements, excepting for the cephalic index. In fact, for them, the LOAs proved to be statistically similar to the manual ones ($p > 0.05$ in a two-tailed F-Test against the intra-observer), which suggests the equivalence between the computed measurements obtained from manual and automatic landmarks' identification strategies. However, statistically significant differences were found for the cephalic index. Nevertheless, the results proved the added-value of the method and its potential to be used in clinical practice to diagnose and evaluate DP.

7. CONCLUSIONS

In this work, a novel pipeline for the automatic identification of 3D landmarks, from head models, and to estimate the relevant anthropometric measurements for DP evaluation is presented. The accuracy and robustness of this technique was initially validated in a synthetic database, demonstrating its potential for the current clinical practice. In the future, additional manual references of real datasets will be generated and, consequently, further validation of the proposed strategy will be performed.

REFERENCES

[1]   S. Robinson and M. Proctor, "Diagnosis and management of deformational plagiocephaly," *J. Neurosurg. Pediatr.*, vol. 3, no. 4, pp. 284–295, 2009.
[2]   A. L. C. Martiniuk, C. Vujovich-Dunn, M. Park, W. Yu, and B. R. Lucas, "Plagiocephaly and Developmental Delay: A Systematic Review," *Behav. Pediatr.*, vol. 38, no. 1, p. 12, 2017.
[3]   S. Nahles, M. Klein, A. Yacoub, and J. Neyer, "Evaluation of positional plagiocephaly: Conventional anthropometric measurement versus laser scanning method," *J. Cranio-Maxillofac. Surg.*, vol. 46, no. 1, pp. 11–21, Jan. 2018.
[4]   A. W. Fitzgibbon, "Robust registration of 2D and 3D point sets," *Image Vis. Comput.*, vol. 21, no. 13–14, pp. 1145–1153, Dec. 2003.
[5]   J. Yang, H. Li, D. Campbell, and Y. Jia, "Go-ICP: A Globally Optimal Solution to 3D ICP Point-Set Registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2241–2254, Nov. 2016.
[6]   Shaoyi Du, Nanning Zheng, Gaofeng Meng, and Zejian Yuan, "Affine Registration of Point Sets Using ICP and ICA," *IEEE Signal Process. Lett.*, vol. 15, pp. 689–692, 2008.
[7]   B. Amberg, S. Romdhani, and T. Vetter, "Optimal Step Nonrigid ICP Algorithms for Surface Registration," 2007, pp. 1–8.
[8]   H. Aarnivala *et al.*, "Accuracy of measurements used to quantify cranial asymmetry in deformational plagiocephaly," *J. Cranio-Maxillofac. Surg.*, vol. 45, no. 8, pp. 1349–1356, Aug. 2017.