Automatic Assessment of Global Craniofacial Differences between Crouzon mice and Wild-Type mice in terms of the Cephalic Index

Hildur Ólafsdóttir[1], Estanislao Oubel[2], Alejandro F. Frangi[2], Tron Darvann[3], Nuno V. Hermann[3,4], Sven Kreiborg[3,4,5], Per Larsen[1], Bjarne K. Ersbøll[1], Chad A. Perlyn[6], and Gillian Morriss-Kay[7]

[1] Informatics and Mathematical Modelling, Technical University of Denmark, ho@imm.dtu.dk
[2] Computational Imaging Lab, Department of Technology - D.326, Pompeu Fabra University, Barcelona
[3] 3D-Laboratory, School of Dentistry, University of Copenhagen; Copenhagen University Hospital
[4] Department of Pediatric Dentistry and Clinical Genetics, School of Dentistry, Faculty of Health Sciences, University of Copenhagen
[5] Department of Clinical Genetics, The Juliane Marie Centre, Copenhagen University Hospital, Copenhagen, Denmark
[6] Division of Plastic Surgery, Washington University School of Medicine
[7] Department of Physiology, Anatomy and Genetics, Oxford University, Oxford, UK

Abstract. This paper presents the automatic assessment of differences between Wild-Type mice and Crouzon mice based on high-resolution 3D Micro CT data. One factor used for the diagnosis of Crouzon syndrome in humans is the cephalic index, which is the skull width/length ratio. This index has traditionally been computed by time-consuming manual measurements that prevent large-scale populational studies. In this study, an automatic method to estimate cephalic index for this mouse model of Crouzon syndrome is presented. The method is based on constructing a craniofacial atlas of Wild-type mice and then registering each mouse to the atlas using affine transformations. The skull length and width are then measured on the atlas and propagated to all subjects to obtain automatic measurements of the cephalic index. The registration accuracy was estimated by RMS landmark errors. Even though the accuracy of landmark matching is limited using only affine transformations, the errors were considered acceptable. The automatic estimation of the cephalic index was in full agreement with the gold standard measurements. Discriminant analysis of the three scaling parameters resulted in a good classification of the mouse groups.

1 Introduction

Crouzon syndrome was first described nearly a century ago when calvarial deformities, facial anomalies, and abnormal protrusion of the eyeball were reported in

a mother and her son [1]. Later, the condition was characterized as a constellation of premature fusion of the cranial sutures (craniosynostosis), orbital deformity, maxillary hypoplasia, beaked nose, overcrowding of teeth, and high arched or cleft palate. Identification of heterozygous mutations in the gene encoding *fibroblast growth factor receptor type 2* (FGFR2) have been found responsible for Crouzon syndrome [2]. Recently a mouse model was created to study one of these mutations ($\mathrm{FGFR2^{Cys342Tyr}}$)[3].

The severity of Crouzon syndrome is of various stages and often it is difficult to estimate. One of the factors measured when diagnosing children with Crouzon syndrome is the cephalic index (CI). Currently, this is obtained by manual measurements of the skull. A step towards automating the diagnosis of children with Crouzon syndrome and quantifying the severity of the syndrome is presented here. In [4] it was shown that manual measurements of skull parameters on the mouse model were in correspondence with those obtained in humans. Further, they gave discrimination between Crouzon and Wild-type (WT) mice. This paper extends this work by presenting an automatic method for quantifying the differences in CI between mouse groups. This was achieved by constructing a craniofacial WT mouse atlas from 3D Micro CT data and then acquiring CIs by registering each subject to the atlas by an affine transformation. The consistency with manual measurements from [4] were then investigated. Additionally, a statistical analysis of the anisotropic scaling parameters was performed in order to further examine the differences between the two groups.

In addition to the automatisation of the diagnosis of Crouzon syndrome, it is anticipated that this study will contribute to our understanding of how premature fusion of sutures and synchrondroses affects craniofacial morphogenesis.

The outline of the paper is the following. In the next section, data acquisition and methodology will be discussed. Section 3 presents the experimental results in terms of registration accuracy, consistency with manual measurements of CI and statistical analysis of group differences. Section 4 provides a discussion of the results and conclusions.

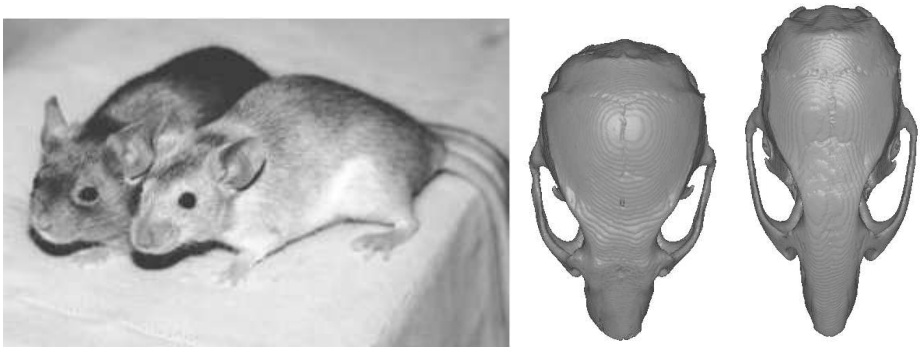
2 Methods and Materials

2.1 Data

Production of the $\mathrm{FGFR2}^{C342Y/+}$ and $\mathrm{FGFR2}^{C342Y/C342Y}$ mutant mouse (Crouzon mouse) has been previously described [3]. All procedures were carried out in agreement with the United Kingdom Animals (Scientific Procedures) Act, guidelines of the Home Office, and regulations of the University of Oxford. Mutant mice of breeding age were determined by phenotype.

For three-dimensional (3D) CT scanning, 10 WT and 10 $\mathrm{FGFR2^{C342Y/+}}$ specimens at six weeks of age (42 days) were sacrificed using Schedule I methods and fixed in 95% ethanol. They were sealed in conical tubes and shipped to the Micro CT imaging facility at the University of Utah. Images of the skull

were obtained at approximately 46μm $\times$ 46μm $\times$ 46μm resolution using a General Electric Medical Systems EVS-RS9 Micro CT scanner. Figure 1 shows an example of the mice and imaging data appearance.



(a) (b) (c)

Fig. 1. (a) Photo of a Crouzon mouse (left) and a WT mouse (right). Skulls Extracted from CT images of (b) a Crouzon mouse, (c) WT mouse.

2.2 Cephalic index

The cephalic index is defined as $CI = (W/L) \cdot 100$, where W is the skull width (the distance between the left- and rightmost lateral points (euryon) on the skull) and L is the skull length (the distance from nasale to to most distant point on occipital bone). This index contributes to quantify the severity of the Crouzon syndrome. For humans, a CI over 80 is connected to brachycephaly, seen in children with premature fusion of the coronal sutures bilaterally. This is one of the indicators for Crouzon syndrome.

2.3 Image registration

An image registration algorithm aims at optimising a similarity metric with respect to a transformation of a given type. In this study, Sum of Squared Differences (SSD) and Cross Correlation (CC) and Normalised Mutual Information (NMI) were tested as similarity metrics. In short, NMI outperformed the other two and was therefore used in the remaining experiments. The type of transformations used here was selected with the application in mind. Since the aim of the study was to investigate the CI, affine transformations with anisotropic scaling parameters were applied (i.e. nine degrees of freedom). For this study, the implementation of an affine registration algorithm by Rueckert[8] [5] was adopted.

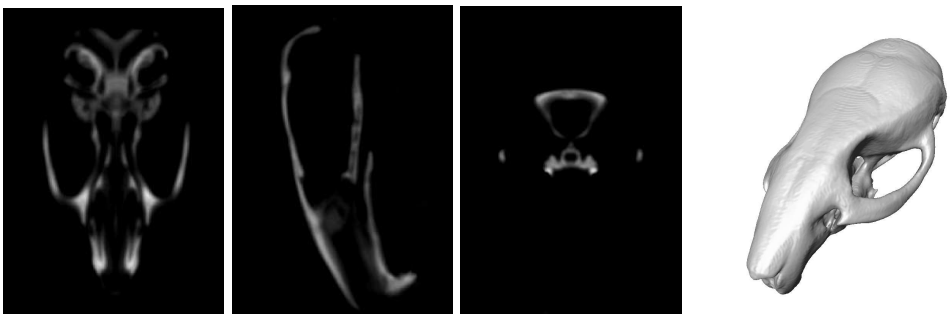
[8] http://www.doc.ic.ac.uk/∼dr/software/

2.4 Preprocessing

Prior to registration of the images, a few steps needed to be made. In order to initialise the registration close to the region of capture for NMI, four landmarks were used to align all mice with respect to their mid-sagittal planes (MSP). Secondly, the neck part was removed from all 20 cases since heads were removed at different positions prior to scanning. Additionally, the mice had different jaw positions, which could give problems in the registration due to a possible change in topology (closed vs. open mouth). The shape differences of the mandible (lower jaw) are not of interest in this study since they are considered a secondary consequence of the syndrome. Therefore, the mandible was manually segmented and masked out. The same applied to the hyoid bone. Furthermore, in order to speed up registrations and removing bias due to different orientations in background, a region of interest (ROI) was defined automatically in the images. The skull was segmented by thresholding and subsequently a sequence of dilations and erosions were carried out resulting in a ROI nicely body-fitted around the skull.

2.5 Atlas Construction

The craniofacial atlas was generated from the set of WT mice. Although the set includes few subjects (10 cases) the biological variation is small since they all had the same DNA and it was therefore considered reasonable to build an atlas from this set. The atlas was generated as follows. A reference subject was chosen from the set to represent the initial atlas. The remaining subjects were registered to the atlas using the affine transformation (9 degrees of freedom). In each iteration a new atlas was generated by averaging the transformed images, including only "active voxels" (i.e. voxels which had not been masked out). The procedure was repeated until the Mean Squared Error (MSE) between two successive atlases dropped below a given limit. In order to remove bias towards the choice of a reference subject, the atlas was transformed into its *affine* natural coordinates with a procedure similar to the one described in [6]. In our case, the natural coordinate system was defined in terms of an affine transformation, since we were to discriminate between the mouse groups in terms of scaling parameters. In this way, the atlas had aspect ratio and gross shape corresponding to the average of the population used in its construction as opposed to be biased to the mouse used as starting point for atlas building. The average scalings from all the registrations were calculated, and the inverses of these scalings were applied to the atlas to transform it into the *affine* natural coordinate system. After atlas construction, relabelling was carried out, i.e. mandible, hyoid bone and neck were segmented and masked out of the atlas in order to assure proper masks in the subsequent registration experiments. The resulting atlas is shown in Figure 2.



(a) (b) (c) (d)

Fig. 2. (a) Axial, (b) sagittal and (c) coronal views of the craniofacial WT mouse atlas. (d) 3D surface view of the atlas.

3 Experimental Results

3.1 Evaluation of Registration Accuracy

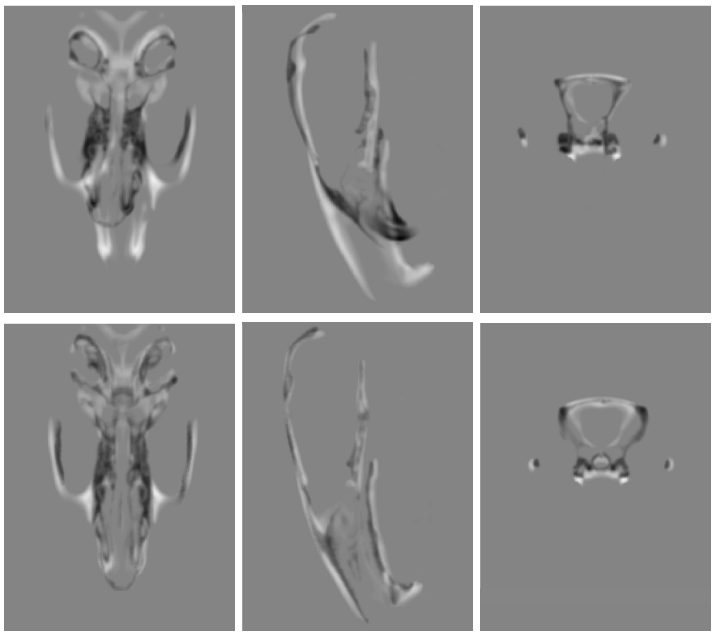
To give a qualitative impression of the registration accuracy, Figures 3 shows the difference images for registration of a Crouzon mouse to the atlas. For the quantitative evaluation of the registration accuracy, the average of two sets of manually annotated landmarks, set by two independent observers were used as a gold standard (GS). Landmarks were obtained in the atlas domain by propagating the GS landmarks from each subject to the atlas by using the transformations obtained by the registration. Landmark registration errors were calculated as the Euclidean distances of each of the transformed landmarks to the corresponding GS landmark on the atlas (point to point errors). From those, RMS errors were calculated. Table 1 presents the registration errors. An examination of the errors for Crouzon cases in terms of Euclidean distances of each landmark (point to point error) is presented in Figure 4.

Table 1. Registration errors in mm: Errors between an observer and GS and errors between the automatic method and the GS.

Errors [mm]	RMS	SD	Med	Max	Min
OBS - GS, WT	0.1845	0.0490	0.1827	0.2937	0.1318
Automatic - GS, WT	0.5435	0.07089	0.542	0.7081	0.4548
OBS - GS, Crouzon	0.1741	0.0269	0.1728	0.2175	0.1400
Automatic - GS, Crouzon	1.1370	0.1754	1.1761	1.4223	0.8713

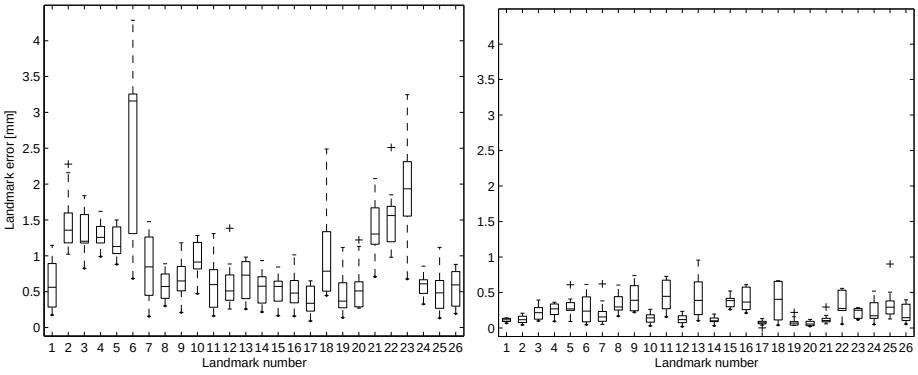
3.2 Consistency with previous results

In a previous study of the same dataset, manual measurements of seven skull parameters were performed [4]. Two of those were the skull width and length



(a) (b) (c)

Fig. 3. Difference images before (above) and after (below) registration of a Crouzon mouse to the atlas in (a) axial, (b) saggital and (c) coronal view.

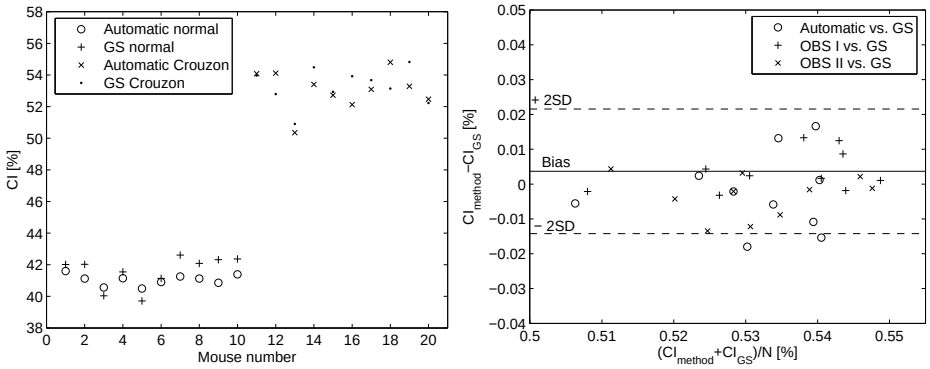


(a) (b)

Fig. 4. Errors for Crouzon mice. (a): Automatic vs. GS. (b): Observer vs. GS.

from which the CI was derived. These measurements were used as a GS. To automatically estimate the CI, landmarks defining the skull width and length were manually placed on the atlas. Subsequently, these were propagated to all subjects using the inverse of the transformations obtained previously by the registration. Figure 5 shows the CI plotted for the automatic approach and the

GS for both groups of mice. Further, the figure shows a Bland-Altman plot for the Crouzon mice to further evaluate the the accuracy of the automatic method with respect to the GS and to include an examination of inter-observer variability. In addition to the plots shown in Figure 5 a paired t-test was applied to compare the automatic method to the GS. The test stated that the difference between the two methods was statistically insignificant in terms of the CI.



(a) (b)

Fig. 5. (a) CI estimated by the automatic method compared to the GS, plotted for WT mice and Crouzon mice. (b) Bland-Altman plots for CI measured on Crouzon mice. The plot shows the bias with respect to the GS for three different approaches, the automatic method and the manual measurements from the two observers.

3.3 Statistical analysis of group differences

The differences between the two groups were analysed in terms of the scaling parameters, s_x, s_y and s_z required for each registration. For this purpose, a Fisher Linear Discriminant analysis was applied. Figure 6 shows a plot of the three scaling parameters with the Fisher discriminator plane included.

4 Discussion and Conclusions

The difference images shown in Figure 3 indicate that an affine registration is able to compensate for the gross craniofacial shape changes between a Crouzon mouse and the WT mouse atlas. The registration errors in Table 1 were not expected to outperform the observer error since differences between individual skulls can never be modelled completely by an affine transformation. Keeping that in mind we consider the registration errors for the WT mice acceptable. The errors for Crouzon cases are quite high. Judged from the left plot in Figure 4, the errors are highly dependent on landmark positions. Some landmarks have considerably

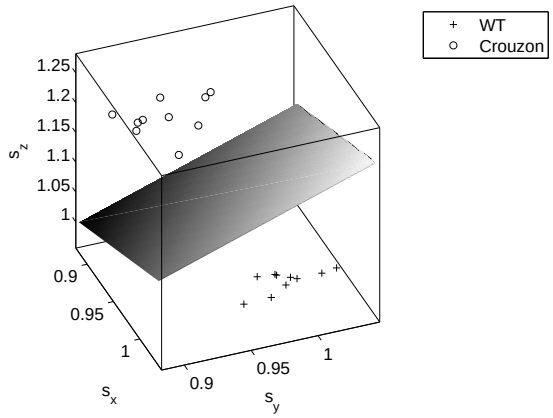


Fig. 6. Scaling parameters: s_x, s_y and s_z for registration of WT mice and Crouzon mice to atlas. Fisher linear discriminator is shown.

higher error than others. Due to the fact that an affine transformation is used, it is to be expected that some landmarks are more difficult to match than others. An inspection of the landmark positions revealed that the most problematic landmarks were placed at the skull sutures. The Crouzon syndrome has proven to cause early suture patency and the skull growth is known to be hindered and very complex at these regions. It is therefore not surprising that additional nonrigid warps are assumably required to match landmarks in this region and to describe in more detail the Crouzon phenotype. However, it is noted that most of the errors lie around 0.5 mm, which is close to the errors from the observer-GS plot to the right. In general the landmark errors are considered acceptable for the main topic of this paper, i.e. building a global morphometric tool to automate the discrimination of different groups of mice in terms of the CI and scaling parameters. From Figure 5, one can observe that the differences between the manual and automatic methods and the GS for estimating CI are very similar. This was confirmed by the paired t-test. This result suggests that it is possible to quantify the CI without actually measuring it manually. The plot in Figure 6 demonstrates that the skulls of WT mice and Crouzon mice differ in terms of scaling parameters of the registration. This especially applies to the z-direction, which is correlated to the difference in skull length. The Fisher discriminator confirms this, by a perfect classification between the two groups.

In summary, an automatic assessment of differences in Crouzon and Wild-Type mice in terms of the cephalic index and scaling parameters has been presented. Good consistency with manual measurements was obtained and it is concluded that automatic assessment of the cephalic index to quantify severity of Crouzon syndrome can replace manual measurements.

References

1. Crouzon, O.: De l'acrocephalosyndactylie. Bull Soc. Med. Paris **35** (1912) 1310
2. Reardon, W., Winter, R.M., Rutland, P., Pulleyn, L.J., Jones, B.M., Malcolm, S.: Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. Nat Genet **8** (1994) 98–103
3. Eswarakumar, V.P., Horowitz, M.C., Locklin, R., Morriss-Kay, G.M., Lonai, P.: A gain-of-function mutation of fgfr2c demonstrates the roles of this receptor variant in osteogenesis. Proc Natl Acad Sci, U.S.A. **101** (2004) 12555–60
4. Perlyn, C.A., DeLeon, V.B., Babbs, C., Govier, D., Burell, L., Darvann, T., Kreiborg, S., Morriss-Kay, G.: The craniofacial phenotype of the Crouzon mouse: Analysis of a model for syndromic craniosynostosis using 3D MicroCT. Cleft Palate Craniofac. J (2006, in press)
5. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. Medical Imaging, IEEE Trans. on **18**(8) (1999) 712–721
6. Frangi, A.F., Rueckert, D., Schnabel, J.A., Niessen, W.J.: Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. IEEE Transactions on Medical Imaging **21**(9) (2002) 1151–1166