3D Fetal Face Reconstruction from Ultrasound Imaging

Antònia Alomar¹, Araceli Morales¹, Kilian Vellvé², Antonio R. Porras³,⁴,⁵, Fatima Crispi²,³, Marius George Linguraru³,⁶, Gemma Piella⁴,⁶ and Federico Sukno¹

¹Department of Information and Communications Technologies, Universitat Pompeu Fabra, Barcelona, Spain
²Fetal Medicine Research Center (BCNatal), Hospital Clínic and Hospital Sant Joan de Déu, Universitat de Barcelona, Barcelona, Spain
³Sheikh Zayed Institute for Pediatric Surgical Innovation, Children’s National Hospital, Washington, D.C., U.S.A.
⁴Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, U.S.A.
⁵Department of Pediatrics, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, U.S.A.
⁶Departments of Radiology and Pediatrics, School of Medicine and Health Sciences, George Washington University, Washington, D.C., U.S.A.

Keywords: Craniofacial Morphology, 3D Morphable Model, Facial Dysmorphology, Fetal Reconstruction, Prenatal Diagnosis.

Abstract: The fetal face contains essential information in the evaluation of congenital malformations and the fetal brain function, as its development is driven by genetic factors at early stages of embryogenesis. Three-dimensional ultrasound (3DUS) can provide information about the facial morphology of the fetus, but its use for prenatal diagnosis is challenging due to imaging noise, fetal movements, limited field-of-view, low soft-tissue contrast, and occlusions. In this paper, we propose a fetal face reconstruction algorithm from 3DUS images based on a novel statistical morphable model of newborn faces, the BabyFM. We test the feasibility of using newborn statistics to accurately reconstruct fetal faces by fitting the regularized morphable model to the noisy 3DUS images. The algorithm is capable of reconstructing the whole facial morphology of babies from one or several ultrasound scans to handle adverse conditions (e.g., missing parts, noisy data), and it has the potential to aid in-utero diagnosis for conditions that involve facial dysmorphology.

1 INTRODUCTION

Craniofacial malformations occur because of abnormal development comprise over one third of all congenital (i.e., birth) defects (Mossey and Catilla, 2001). These anomalies comprise a wide range of heterogeneous conditions and often have a multifactorial origin, including genetic and environmental factors (Sorop Florea et al., 2018). These malformations can impact swallowing, breathing, hearing, vision, speech, and cognitive development (on Government Affairs, 2020; EvansAnne et al., 2018), and they impose a large psychosocial, healthcare, and economic burden.

Early diagnosis is often crucial for the effective treatment of functional and developmental aspects (Learned-Miller et al., 2006; Tu et al., 2018; Tu et al., 2019). However, not all syndromes are easily identified, some of them having subtle physical manifestations; careful clinical assessment may be necessary to distinguish an isolated abnormality from an atypical or mildly manifested syndrome. Moreover, the identification of the specific syndrome is important for the overall care of the patient (EvansAnne et al., 2018). In this sense, the analysis of facial morphology can provide relevant information and serve as a pre-screening tool, facilitating the early detection of developmental disorders (Menezes et al., 2016; Merz and Welter, 2005). Efforts are being made to shift from diagnosis at birth, or during the first years of life, to prenatal diagnosis, which facilitates parents’ counselling and careful planning of delivery and postnatal treatment (Pooh and Kurjak, 2011). However, prenatal diag-

a https://orcid.org/0000-0003-3658-5832
b https://orcid.org/0000-0003-4930-6142
c https://orcid.org/0000-0002-2376-7664
d https://orcid.org/0000-0001-5989-2953
e https://orcid.org/0000-0002-7422-5240
f https://orcid.org/0000-0001-6175-8665
f https://orcid.org/0000-0001-5236-5819
h https://orcid.org/0000-0002-2029-1576

Copyright © 2021 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved
nosis of fetal syndromes is not easy, mainly because of the wide range of morphological features involved and the challenging nature of medical images.

Ultrasound is the primary imaging modality for fetal assessment. It is a noisy image modality, but it has the advantage of being widely available, cost-effective, non-ionizing, and it allows real-time examination. Three-dimensional ultrasound (3DUS) facilitates the evaluation of anatomical structures from the facial surface and can therefore aid diagnosis (Andresen et al., 2012; Werner et al., 2016; Merz and Abramowicz, 2012). A detailed 3D model of the fetus face could thus play a crucial role in prenatal diagnosis.

Little research has been done in 3D face reconstruction from fetal images, mainly due to the limitations of prenatal imaging itself. In (Dall’Asta et al., 2017), it was presented a statistical shape model constructed from 20 3DUS scans that were manually segmented and aligned, and statistically significant differences in face shape were found between normal and abnormal fetuses. There are some works on generating physical fetal models (although not always face specific) from 3DUS, magnetic resonance imaging, and computer tomography (Werner et al., 2010; Werner et al., 2015; Menezes et al., 2016; Speranza et al., 2017). However, they involve slice-by-slice manual segmentation and post-processing with proprietary software.

In this paper, we explore the feasibility of reconstructing the facial morphology before birth by analyzing 3DUS images of the fetus from routine scanning with the help of a recently proposed statistical model constructed from 3D scans of babies and newborns: the Baby Face Model (BabyFM) (Morales et al., 2020). Differently from previous works, we do not build our model directly from the noisy fetal images, but employ statistics from newborns to constrain the geometric reconstruction of the fetal face. In this way, we circumvent the difficulties associated with building accurate models from the noisy 3DUS images. Tests on a small set of fetal scans show promising results in both qualitative and quantitative terms, even in adverse conditions (e.g. missing parts, noisy data).

2 MATERIALS

2.1 3D Baby Face Morphable Model

A 3D morphable model (3DMM) is a tool for representing 3D shapes and textures. In the context of face analysis, the idea is to learn a general 3D face model that is able to encode the statistics of facial shape. A crucial aspect to consider when using a 3DMM is that the statistics encoded in the model must match those of the target population, e.g., in terms of ethnicity, gender and, especially important for this work, age. The latter has been an important obstacle for the application of pre-existing 3DMMs to fetal data, since all available 3DMMs were built from adults and, although sometimes they also included children, none of them included babies. However, very recently, Morales et al. (Morales et al., 2020) have published the Baby Face Model (BabyFM), which constitutes the first 3DMM built exclusively from babies, including an important proportion of data from newborns.

The BabyFM was built from 45 3D scans of baby faces (mean age 8.42 ± 6.45 months). Several ethnicities were included: Caucasian (47%), African American (24%), Hispanic (20%), and Asian (9%). Also, the data were roughly gender-balanced: 56% male and 44% female. The BabyFM covers the facial region that is delimited by the chin the forehead and the ears, all included. Additionally, the vertex indices for 23 anatomical landmarks (Figure 1) are provided, which are used to initialise the 3D facial reconstruction (see Section 3.1.1).

![Figure 1: Anatomical landmarks. Illustration of the 23 anatomical landmarks considered in this project on the mean baby morphable model face. Landmark abbreviations: enL/R = endocanthion Left/Right; n = nasion; exL/R = exocanthion Left/Right; aL/R = alare Left/Right; acL/R = alar crest Left/Right; prn = pronasale; sn = subnasale; chL/R = cheilion Left/Right; cphL/R = crista philtrum Left/Right; ls = labiale superius; li = labiale inferius; sl = sublabiale; pg = pogion; tL/R = tragion Left/Right; and oIL/R = otobasion inferior Left/Right.](image-url)

2.2 Test Database

To evaluate our methods, 10 3DUS scans from 4 fetuses were collected, i.e., there were multiple 3DUS images for each of fetus, corresponding to different
viewing directions. These fetuses had no relation to any of the babies used for the construction of the BabyFM. The 3DUS scans were obtained at Hospital Clinic and Hospital Sant Joan de Déu, Barcelona, according to its Ethical Research Committee and the current legislation. Images were acquired using a General Electric Voluson E6 (General Electric, IL, USA) US machine with a low-frequency probe (4-8 MHz).

Three-dimensional meshes were extracted using just a threshold segmentation. The meshes contained not only the face but also other parts of the fetus’ body, placenta, and noise. Using the Landmark software 3.6 1, we positioned a subset from the 23 targeted landmarks in the BabyFM (see Figure 1) on each fetal scan, according to their visibility. The identification of these anatomical landmarks in the fetal scan is challenging because of the occlusions (e.g., the baby may be positioned with the hand on the face) and the noisy nature of the data, and therefore not all of them could be positioned for each fetus.

Additionally to the 3DUS scans, for each of the babies we had three 2D postnatal photographs taken from different viewpoints by the parents with their mobile phones. These images were used to obtain a 2D-3D reconstruction of the baby face to which we fitted morphable face model and those of the US mesh is calculated as follows:

$$\alpha = (\Phi_r^T\Phi_r)^{-1}\Phi_r^T(x - \bar{x}), \quad (1)$$

where $\Phi_r$ is the reduced shape basis matrix (i.e., the rows of the eigenvector matrix that correspond to the landmarks), and then regularizing to ensure obtaining plausible faces. The shape parameters are assumed to follow a multivariate Gaussian distribution. Therefore, we constrain the shape parameter vectors to lie within a hyper-ellipsoid in the parameter space, the size of which is determined by the variances (the eigenvalues) of the data.

The two-stage landmark-based fitting is iterated 20 times to ensure convergence. Finally, the mean approximation error ($\bar{E}$) between the landmarks of the fitted morphable face model and those of the US mesh is calculated as follows:

$$\bar{E} = \frac{1}{m} \sum_{j=1}^{m} ||l_j - T(l_{\text{model},j})||, \quad (2)$$

where $l_j \in \mathbb{R}^3$ is the $j$-th landmark, $T$ is the transformation that maps the BabyFM mean to the US, and $m$ is the number of anatomical landmarks that were positioned in the US mesh.

### 3.1.2 Iterative Closest Point with Statistical Constraints

The fetal face reconstruction obtained from the landmark-based fitting is refined using an iterative closest point (ICP) algorithm. In every iteration, the ICP algorithm fits the face reconstruction to the 3DUS mesh and then recovers the model’s shape parameter $\alpha$, analogously to the landmark-based fitting (i.e., by Eq. 1 followed by regularization) but now using the whole surface, i.e., all rows of the shape basis $\Phi$ instead of just $\Phi_r$. To increase the robustness to input artifacts, the point-matching was applied under geometric and uniqueness constraints (to minimize the impact of outliers and ensure one-to-one mapping). The ICP with statistical constraints was repeated by alternating between the correspondence mapping and the model’s parameter update followed by statistical regularization, until the error difference between consecutive iterations was below a predefined threshold.

In this work, this convergence threshold was set to $10^{-2}$ mm.

---

1https://landmark2.software.informer.com/download/
3.2 Multiple Image Fitting

In order to quantitatively validate the fetal shapes estimated from the 3DUS scans, we reconstruct the newborn 3D face from a set of three 2D images (frontal, left, and right pose) taken shortly after birth. The BabyFM is used here to estimate the facial 3D geometry of the newborn. Once the 3D geometry is obtained, facial texture can also be added to obtain a photo-realistic 3D reconstruction.

We address the 3D-from-2D reconstruction problem using sparse geometric features (edges and landmarks). Our approach is based on the algorithm proposed in (Bas et al., 2016), but using multiple images rather than only a single image. We start by positioning the anatomical landmarks in the different images using a 2D landmarker and obtaining the edges by applying the Canny edge detector.

Then, an initialization of the 3D face is obtained using only the landmarks. The landmark fitting is then refined in an iterative closest point manner by finding the closest image edge to each model contour vertex. The model edge vertices can then be considered as landmarks with known 2D position, for which optimal pose and shape estimates can be easily computed under the assumption of a scaled orthographic projection. In particular, we obtain the optimal pose and shape parameter by minimizing an objective function that include landmark, edge, and prior terms:

$$ E(\alpha, R, t, s) = \frac{1}{2} E_{\text{land}}(\alpha, R, t, s) + \frac{1}{2} E_{\text{edge}}(\alpha, R, t, s) + \frac{1}{2} E_{\text{prior}}(\alpha), $$

where $\alpha$ is the shape parameters vector and $R, t, s$ the pose parameters (rotation, translation, and scale) assuming a scaled orthographic projection. The parameters $w_1, w_2, w_3$ correspond to the weights given to each error term, the sum of the three weights should be equal to one. The used values to perform the reconstructions were 0.25, 0.25, 0.50 respectively.

The landmark term penalizes differences between the actual landmarks positions on the images and the ones obtained by projecting the 3D model landmarks. The edge term compares the edges detected on the image with those induced by the model due to occluding boundaries. The prior term acts as a regularizer of the shape parameters based on the statistics encoded in the BabyFM.

4 RESULTS

4.1 US Fitting

We applied our reconstruction pipeline to each of the fetus scans. Figure 3 shows the US images obtained from the Voluson system (GE Healthcare), the US meshes obtained after the threshold segmentation, and the 3D reconstruction that we obtained. As can be observed in Figure 3, the input data is quite challenging. For example, in most of the 3DUS images, the ears are not present or are extremely noisy. Nevertheless, our method is able to estimate an approximate ear shape by exploiting the geometric correlations en-
Figure 3 shows US images (top), US meshes (middle), and their corresponding reconstructions (bottom). From left to right: case 1, case 2, case 3, and case 4.

coded in the BabyFM. Figures 4 displays the profile view of cases 1 and 2 to show the ears.

Figure 4: Example of reconstructions (left: case 1, right: case 2) in profile view.

4.1.1 Single US Fitting vs Multiple US Fitting

As we have multiple 3DUS scans from different views for each individual, we checked whether the reconstructions obtained independently for the same fetus were similar. Also, we investigated whether the simultaneous use of multiple views improves the obtained results. For this, the algorithm was adapted to merge the correspondences of each 3DUS view and then calculate a unique vector of shape parameters for the multiple scans.

Figure 5-6 show examples of reconstructions obtained independently for the same individual (cases 1 and 2, respectively) from two different 3DUS views. In Figure 5, we can appreciate that the fetus has plump lips (African ethnicity) and the model is not able to fully reconstruct them accurately, but it attempts to compensate this fact by opening the mouth. The used BabyFM was built including only a 24% of African American babies, which might explain why it is not able to perfectly reconstruct the lips. Nevertheless, the nose, eyes, and pose are well estimated, and the reconstructions are convincingly similar to the original renderings. More importantly, there are no substantial differences between the reconstructions obtained from the two 3DUS scans of this same fetus; the second one has wider cheeks, but the nose, mouth, and eyes are similar. In Figure 6 (which corresponds to a case of fetal growth restriction (Peleg et al., 1998; Albu et al., 2014)), the obtained reconstructions capture the skinny face. On the other hand, the second US has part of the mouth and chin occluded by an arm, which makes the reconstruction more challenging.

Next, we reconstruct the face considering simultaneously the multiple US scans for each fetus. This can
be especially useful to obtain an accurate complete reconstruction even when each scan provides only a partial view of the facial anatomy. Figure 7 illustrates this by reconstructing the face for case 3 from 4 different views.

### 4.2 Post-natal Fitting from Multiple Images

Since we wish to use the post-natal reconstructions to validate the geometries reconstructed from the 3DUS scans, it is important to assess the accuracy that we can expect in the post-natal reconstruction. To do so, we reconstructed the baby faces of an additional dataset, which contained three 2D images (frontal, right and left pose), as well as the 3D scan for each of the three babies from the Children’s National Hospital, Washington D.C. and one newborn from the Hospital Clinic and Hospital Sant Joan de Dèu, Barcelona. To evaluate the quality of the reconstructed 3D face, we compute the geometric error for each reconstruction, by first applying a transformation to align it with the ground-truth, which in this case is the 3D scan, and then computing its point-to-point distance.

The color error maps of the three babies in Figure 8 show that the 2D-3D fitting performed obtained acceptable results, as the mean error of the reconstructions is around 4 mm. In the three cases, the large errors in the forehead are caused by the lack of landmarks in this region and the black hat that the babies wear, as it interferes with the image edge detection resulting in wrong correspondences. However, notice that the error in the rest of the face, which corresponds to the main facial features, is considerably smaller. This, together with the fact of using uncalibrated input images, suggests that the accuracy of the reconstructions is satisfactory.

Once we have obtained the 3D reconstruction from multiple 2D images, we add texture to it to have more photo-realistic results to show to the parents. As expected, the textured meshes improve the visual resemblance with the images (bottom row of Figure 8).

### 4.3 Pre- and Post-natal Comparison

The reconstructions obtained by fitting the BabyFM to the 3DUS scan and to the set of 2D images after birth are compared quantitatively, using color error maps, as a proxy to evaluate the performance of the fetal face reconstruction. The color error maps are obtained computing the geometric error between the prenatal and the postnatal reconstructions.

We compared the reconstructions from the 3DUS fitting and the multiple image fitting for case 3, for which we had both the postnatal 2D images and the US scans, even if the difference between them was 16 weeks. Figure 9 shows the 2D frontal image of the baby, the baby face reconstruction using the multiple image fitting, the US fetal face obtained, and the color map error between the prenatal and the postnatal reconstruction. Although no age correction was applied to the 3DUS reconstruction, we can appreciate sufficient resemblance to support the statement these are two instances of the same baby, taken a few months apart from each other.

In quantitative terms, the overall mean reconstruction error is 1.73 mm. Main facial features such as eyes, nose, and mouth present the smallest reconstruc-
tion error (see Figure 9), whereas the forehead region - where no landmarks and few edges are available - shows the largest reconstruction error. It is expected that after applying some age correction mechanism, the resemblance will be even stronger.

5 DISCUSSION

In this study, our aim is to obtain accurate fetal face reconstructions from 3DUS images in the prenatal stage in order to assess craniofacial morphology as early as possible. In spite of the growing interest in the early assessment of craniofacial morphology, the 3D renderings provided by the standard clinical view software have some limitations, e.g., two images from different views seems to be two different babies, or the facial features cannot be distinguished due to the smoothing or the large amount of noise. For this reason, clinicians usually have to manually remove the noise to assess the baby’s morphology. Manual segmentation to remove the noise in the US is very time consuming, even for expert clinicians. The presented algorithm can help them to visualize the fetal face with more quality and without performing any manual segmentation. Therefore, the subjectivity introduced by manual segmentation is avoided.

Our algorithm overcomes the limitation of noisy data of the fetal face reconstruction algorithms proposed so far, as no restrictions to the US data are applied. In (Bonacina et al., 2016) and (Dall’Asta et al., 2017) studies, a strict inclusion criteria was applied to the US data, as their procedure is really susceptible to the quality of the input data. As a consequence, all the cases with missing facial parts, large amounts of noise, unsatisfactory definition of the facial borders or obstructed vision caused by the uterine wall/intervening limbs/cord loops, were excluded for those studies, limiting the applicability of their algorithms. In contrast, our approach exploits the facial geometry statistics encoded in the BabyFM, which allows us to reconstruct the 3D fetal face even with large amounts of noise, occlusions or missing parts. Also, it allows reconstructing a unique facial geometry from multiple US scans, so that the specific details available from each different view can be adequately combined to produce an accurate reconstruction of the whole facial geometry.

The reliability and accuracy of the reconstructions depend on the quality of the US and the fetal pose and expression. Higher levels of noise and/or occlusions inevitably result in less accurate reconstructions. Nevertheless, our algorithm demonstrates that reconstructions of the same individual from different US scans are very similar and the faces differ in small details. A large quantitative analysis was not possible, but the proposed 2D-3D fitting tests performed for those cases with ground-truth, show promising results that could be used as proxy ground-truth for the fetal reconstructions. Qualitative assessment of the results were largely convincing, highlighting the potential of the proposed technique to aid prenatal diagnosis.

The immediate next steps in our work will be to address a larger quantitative validation of the reconstruction algorithm, especially in terms of comparisons between the reconstructions before and after...
Figure 8: 3D reconstruction from 2D images. Top: 2D images (only the frontal view is shown). Middle: Error reconstruction (using the frontal, right, and left views) to the ground truth scans. Bottom: 3D reconstruction with texture.

Figure 9: Comparison between prenatal and postnatal reconstructions for case 3. From left to right: frontal 2D image, Us reconstruction, 3D from 2D reconstruction, and error between reconstructions.

After performing the quantitative analysis investigating whether the estimated morphology is significantly different between patients with intrauterine grow restriction and controls is also an interesting line of continuation of this work.

birth. The US scans analyzed so far, were acquired between weeks 20 and 30 of gestation and the postnatal images were taken once the baby was already a few weeks old. Thus, the time difference between the US and the photos used to get the 3D reconstructions was approximately about 20 weeks. This difference might have a clear impact in the face morphology of the baby. For this reason, we are also investigating an age correction mechanism to be applied to the mesh extracted from the 3D fetal faces to see if we can describe a realistic age progression of the face of the baby.

After performing the quantitative analysis investigating whether the estimated morphology is significantly different between patients with intrauterine grow restriction and controls is also an interesting line of continuation of this work.
6 CONCLUSIONS
To conclude, in this paper we proposed a fetal face reconstruction algorithm from 3DUS images. The approach differs from the existing ones proposed in the literature, as it is based on the fitting of a deformable BabyFM to 3DUS to remove the noise and to recover the whole baby facial morphology. It was demonstrated that our algorithm is able to reconstruct the whole facial morphology of the babies under different conditions (large amounts of noise, missing parts or multiple US scans), obtaining promising results. In the future, the presented technique could aid in the prenatal assessment and in-utero diagnosis of syndromes and diseases in which facial dysmorphology is an indicator of early craniofacial abnormalities.

ACKNOWLEDGEMENTS
This work is partly supported by the Spanish Ministry of Economy and Competitiveness under project grant TIN2017-90124-P, and the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502).

REFERENCES