AUTOMATIC SEGMENTATION OF HEAD STRUCTURES ON FETAL MRI

Jérémie Anquez, Elsa D. Angelini, Isabelle Bloch

Institut TELECOM - Télécom ParisTech - CNRS UMR 5141 LTCI, Paris, France.

ABSTRACT

Recent improvements of fetal MRI acquisitions now allow threedimensional segmentation of fetal structures, to extract biometrical measures for pregnancy follow-up. Automation of the segmentation process remains a difficult challenge, given the complexity of the fetal organs and their spatial organization. As a starting point, we propose in this paper a fully automated segmentation method to localize the eyes and segment the skull bone content (SBC). Priors, embedding contrast, morphological and biometrical information, are used to assist the segmentation process. A validation of the proposed segmentation method, on 24 MRI volumes of fetuses between 30 and 35 gestational weeks, demonstrated a high accuracy for eyes and SBC extraction.

Index Terms—Fetal imaging, MRI, eyes, skull, segmentation

1. INTRODUCTION

MRI constitutes a good adjunct to echography for pregnancy follow-up [1]. Higher image quality and shorter acquisition times have led to an increase in the number of MRI exams performed. However, very few works have been dedicated to automatic segmentation of in utero fetal structures, except for brain segmentation [2, 3, 4] on data of fetal eyes and skulls. Segmentation of other fetal structures are desirable to closely monitor fetal growth. As an example, a precise estimation of the fetus weight could be derived from biometrics measures, which is crucial when obstetricians need to plan for an optimal delivery procedure in case of fetal macrosomia (oversized fetuses). Using the Steady State Free Precession (SSFP) sequence, whole uterus MRI volumes can be acquired in less than 30 seconds, thus limiting imaging artefacts related to fetal motion. Moreover, the SSFP sequence provides adequate image quality for segmentation of fetal structures [5]. However, the MRI images contain complex information to process: (1) the SSFP sequence is highly sensitive to magnetic field inhomogeneity; (2) as a wide field of view is used, numerous maternal, uterine and fetal structures are visible; (3) the fetus orientation within the uterus (and hence with respect to the acquisition direction) is unknown. Our approach to process those images consists in identifying an initial structure in the fetal anatomy, to improve the knowledge of the fetal position and initialize the whole fetus segmentation process. Then, segmentation of neighboring organs can be performed sequentially in narrow regions of interest, using anatomical spatial relationships.

In this work, we use the eyes as initial structures, motivated by three factors: a fetal eye is small enough to be considered insensitive to image inhomogeneity, its spherical shape is rotationally invariant and it highly contrasts with its surroundings. We propose an eye-extraction scheme based on template matching. By explicitly using the eyes localization, the neighboring skull bone content (SBC) is then segmented using a graph cut approach. Shape, contrast and biometrical priors are used to guide both segmentation tasks.

2. FETAL IMAGING AND PROBLEM OVERVIEW

A database containing 24 MRI volumes of fetuses between 30 and 35 gestational weeks was gathered. All the images were acquired with the SSFP sequence on two 1.5 Tesla superconducting units (Magnetom, Siemens©, Erlangen, Germany and Avanto, General Electric©, Milwaukee (WI), United States). Typical acquisition parameters were: TR/TE = 4.2/1.8 ms, flip angle = 60°, FOV = 480, slice thickness/gap = 4/0 mm, matrix = 512×512, while voxel size were 0.94 x 0.94 x 4 mm³. An MRI slice is shown in Figure 1, with the structures of interest (fetal eye and SBC) outlined respectively in green and red. The SBC includes the brain and the cerebrospinal fluid and is delimited by the boundary between the cerebrospinal fluid and the skull bone. Due to the T2-weighting component of the SSFP sequence, voxels intensity is strikingly higher in aqueous tissues than in fetal soft tissues and fetal bones. This explains why (1) the eye, which contains fluid-filled chambers, is brighter than the surrounding muscles and (2) the cerebrospinal fluid, which water content is about 99%, appears brighter than the skull bone.

Fig. 1. 2D slice of an MRI volume $I_k$ of the database. The eye is outlined in green and the $SBC$ in red.

Figure 2 illustrates the different steps of the overall segmentation framework. (a) The eyes are detected in the whole image $I_k$ using an appearance model. (b) The midsagittal slice $M_k$ is reconstructed thanks to the eyes centers localization and the $SBC$ is segmented, taking advantage of the closeness between the $SBC$ and the eyes centers. (c) Orientation, position and scale information is extracted from the $SBC$ segmentation in $M_k$, to perform a precise 3D segmentation.

Simple appearance models are built to guide the segmentation of fetal eyes and $SBC$. Contrast priors between those structures and their surroundings are embedded in the segmentation process, as...
well as shape and biometrical priors. Those last two priors rely on information extracted through manual tracing on the image database. The eyes diameter \( d \) and the distance \( D \) between their centers \( E_1 \) and \( E_2 \) were measured, while the \( SBC \) was manually segmented to extract shape information.

Considering the limited number of cases available in our database, the different steps of the proposed segmentation framework were validated using a leave-one-out cross validation procedure. When an image \( I_k \) is processed, a training database \( DB_k \) is considered. \( DB_k \) corresponds to the whole database, leaving \( I_k \) out.

3. EYES DETECTION

3.1. Eyes appearance modeling

Let \( \mu_d \) and \( \sigma_d \) be the mean and the standard deviation of the eye diameter \( d \) on the database. We measured \( \mu_d = 15.03 \) mm and \( \sigma_d = 0.74 \) mm. Since \( \sigma_d \) is small (because the fetuses in our database are of similar gestational ages and do not exhibit any eye development pathology), the eye diameter is well modeled by using only \( \mu_d \). Considering the contrast and shape priors presented above, an eye template \( E \) is built, modeling the eye as a ball of diameter \( \mu_d \), brighter than the background. The typical image resolution in our database is \((r_x, r_y, r_z) = (0.94, 0.94, 4)\). \( r_x \) and \( r_y \) are of the same order of magnitude and the eye is affected by partial volume effect as shown in Figure 3 (a,d). We propose to take into account these visual artefacts.

A first isotropic template \( E' \) is considered at a resolution \( r' \) finer than \( r \) (Figure 3 (b,e)). Let \( v' \) be a voxel of \( E' \). \( E' \) is then downsampled at the resolution \( r \) to obtain \( E \) (Figure 3 (c,f)). Let \( v \) be a voxel of \( E \) and \( V \) be the set of voxels \( v' \) contained in \( v \). The partial volume effect being a mixture of multiple tissues inside a given voxel, we model the intensity of \( v \) as the mean intensity of the voxels belonging to \( V \). This is an approximation of the actual MRI acquisition sampling process, but this is sufficient to generate a realistic template.

3.2. Eyes localization

A set of eye candidates \( \{ E_{Ci} \} \) is computed using template matching. As a similarity measure, we use the normalized cross-correlation, which presents several advantages, such as being only sensitive to local contrast information. This is a key feature as the images obtained with the SSFP sequence are heterogeneous which can lead to great intensity variations in the eye and its surroundings from one image to another one. A normalized cross-correlation coefficients map \( N3C \) is computed using the template \( E \), and then thresholded at a value \( t \). Connected components are identified in the thresholded map and one eye candidate \( \{ E_{Ci} \} \) is extracted in each connected component, corresponding to the maximum of \( N3C \) in the component. The threshold value \( t \) is extracted from the \( N3C \) distribution considering all the images in the training database \( DB_k \). With \( \mu_{N3C} \) and \( \sigma_{N3C} \) the mean and standard deviation of \( N3C \), we use \( t = \mu_{N3C} + 2\sigma_{N3C} \).

![Fig. 2. Main steps of the proposed segmentation framework. (a) Eye detection in the image \( I_k \). (b) Extraction of the midsagittal slice \( M_k \) using the eyes centers and 2D segmentation of the \( SBC \). (c) 3D segmentation of the \( SBC \) using orientation, position and scale information extracted from 2D segmentation in \( M_k \).](image)

![Fig. 3. Eye appearance model: central slices of an image \( I_k \) across one eye (a,d), \( E' (b,c) \) and \( E (c,f) \). (a,b,c) and (d,e,f) are slices in \((X,Y)\) and \((X,Z)\) planes respectively. Image spatial resolutions are \( r' = (0.1,0.1,0.1) \), for \( E' \), and \( r = (0.94,0.94,4) \), for \( E \) and \( I \), in millimeters.](image)

![Fig. 4. Eye detection on an image \( I_k \): (a) connected components extracted from the thresholded \( N3C \) map overlayed on a slice of \( I_k \), (b) ten best \( E_{Ci} \), (c) membership function \( f(D_{ij}) \) and (d) ten best \( S_{ij} \). The blue color identifies the results corresponding to the eye centers.](image)

All couples \( (E_{Ci}, E_{Cj}) \) of eye center candidates are then considered. The distance \( D_{ij} \) between \( E_{Ci} \) and \( E_{Cj} \) is computed and a soft selection is proposed to discard couples separated by an inappropriate distance, building a trapezoid membership function \( f \) with keypoints \((k_1, k_2, k_3, k_4)\). Using the mean \( \mu_D \) and the standard deviation \( \sigma_D \) of the inter-eye distance \( D \) in \( DB_k \), keypoints are defined as \((k_1, k_2, k_3, k_4) = (\mu_D - 3\sigma_D, \mu_D - 2\sigma_D, \mu_D + 2\sigma_D, \mu_D + 3\sigma_D)\). Combining conjunctively \( f(D_{ij}) \) with the \( N3C \) values of \( E_{Ci} \) and \( E_{Cj} \) \((N3C_i \text{ and } N3C_j)\) leads to the scoring function \( S_{ij} = N3C_i, N3C_j \cdot f(D_{ij}) \), which is computed for every couple and identifies the couple with the highest score as the fetus eyes \((E'_{i1}, E'_{j2})\).

Figure 4 shows intermediate results when applying the algorithm to an MRI volume \( I_k \). On this image, \( t = 0.53 \) and 59 connected components were extracted from the thresholded \( N3C \) map using this value, which remains reasonable. In (a), the connected components are plotted in a slice containing an eye center (blue component). The ten best \( E_{Ci} \) are shown in (b), ordered by decreasing \( N3C_i \). The best two correspond to the eyes centers (blue circles).
The selection function $f$, built with $\mu_D = 33.2 \text{ mm}$ and $\sigma_D = 2.4 \text{ mm}$, is shown in black (c), along with the evaluation of the couples such that $D_{ij} < 100$. The couple corresponding to the eyes center returned 1 (blue circle) while 97.4% of the couples returned 0, demonstrating the high selectivity of $f$. Finally, (d) presents the ten best $(E_{ci}, E_{ci})$, ordered by decreasing $S_{ij}$, $(E_1^E, E_2^E)$, corresponding to the highest $S_{ij}$, indeed matches the fetal eyes in the image.

For two MRI datasets $I_k$, the eyes centers did not correspond to the best two $E_{ci}$. However, the use of $f$ discarded all false positives and the eyes centers were detected successfully on the whole database. The average discrepancy between the eyes centers $E_k^*_{ci}$ provided by the algorithm and those manually labeled $E_{ci}, i \in \{1, 2\}$, was 1.7 mm and a precise detection was hence performed.

4. SKULL BONE CONTENT SEGMENTATION

4.1. Skull bone content modeling

We now consider the SBC, defined as the volume delimited by the boundary between the skull bone and the cerebrospinal fluid. Several factors prevent the use of an appearance model, as was done for the eyes, for the SBC segmentation. First, important changes affect the fetal brain during the gyration phase, which occurs from the 28th to 35th gestational week and therefore concerns the fetuses in our database. During this phase, the SBC presents variable visual characteristics. Moreover, the fetal head surroundings can include different maternal and uterine structures, depending on the fetal position. Finally, given the SBC size, field inhomogeneities may make the intensity of the SBC structures variable and cannot be neglected. Hence, a different approach is adopted and an average shape model is built using the SBC manual segmentations. Leaving $I_k$ out, binary shaped shapes in $DB_k$ are registered using a similarity transform and averaged in order to derive a mean shape model of SBC denoted $SM$. Figure 5 illustrates one $SM$. It is important to point out here that the different $SM$ are very similar, showing a weak sensitivity to the $I_k$ left out of the training set. Therefore, any $SM$ could be used to process a new image added to the database.

**Fig. 5.** 3D mean shape model $SM$. (a) Right, (b) anterior and (c) inferior views.

$SM$ and a contrast prior between the cerebrospinal fluid and the skull bone are combined in a segmentation process involving two steps. Firstly, the midsagittal slice $M_k$ is extracted using $(E_1^E, E_2^E)$ and the SBC is segmented in this slice, to fully determine the fetal head orientation and initialize the 3D segmentation of the SBC. Indeed, $E_1^E$ and $E_2^E$ define the lateral orientation of the fetal head, but anterior-posterior and inferior-superior orientations are unknown. Secondly, the 2D segmentation of $SBC$ is used to precisely embed the shape model $SM$ in the 3D data, and perform a narrow-band 3D segmentation of the SBC.

4.2. Skull bone content segmentation in the midsagittal slice

Let $\bar{E}$ be the center of the $[E_1^E, E_2^E]$ segment. The midsagittal slice $M_k$, passing through $\bar{E}$ and orthogonal to $E_1^E E_2^E$, is reconstructed. Let $SM_{2D}$ denote the 2D mean shape model of SBC corresponding to the midsagittal slice of $SM$. For different locations and orientations of $SM_{2D}$, a segmentation is performed in a narrow-band, providing different SBC candidates. The best one is then selected as the most similar to $SM_{2D}$. The details of the method are as follows.

$SM_{2D}$ is translated in order to be at a proper distance from $\bar{E}$, and $L$ orientations of $SM_{2D}$ are considered, rotating $SM_{2D}$ around $\bar{E}$ with angles $\theta^l = (2\pi l)/L$, $l \in [0, L - 1]$. Let $SM'_{2D}$ be $SM_{2D}$ rotated with the angle $\theta^l$. $SM'_{2D}$ is eroded and dilated to provide two objects $\epsilon(SM'_{2D})$ and $\delta(SM'_{2D})$, which are used to define a narrow-band $NB^l = \delta(SM'_{2D}) \cap \epsilon(SM'_{2D})$. A graph cut segmentation is performed in $NB^l$ [6]. Oriented edges are created between neighboring pixels. Two special nodes, the source $S$ and the sink $T$, are added. Edges are created between $(1) S$ and the pixels of $NB^l$ adjacent to $\epsilon(SM'_{2D})$ and $(2) T$ and the pixels of $NB^l$ adjacent to $\delta(SM'_{2D})$. The minimal cut of the graph provides a binary segmentation corresponding to a minimal path in the region located between the source and the sink, considering the edge weights. Let $p$ and $q$ be two neighboring pixels and $I_p$ and $I_q$ their intensity. The weight of the edge linking $p$ to $q$ is defined as $w_{pq} = e^{-(I_p - I_q)^2/2\sigma^2}$, if $I_p > I_q$, and $w_{pq} = 1$, if $I_p \leq I_q$. This weighting includes contrast prior [7]. As the cerebrospinal fluid is brighter than the skull bone, edges linking cerebrospinal fluid pixels to skull bone pixels have a low weight. Thus, those edges are prone to belong to the minimal cut. The segmentation result provides the brain candidate $SBC_{2D}$ in the position $l$. The similarity between the brain candidates shape and the shape model is evaluated by computing the kappa measure $\kappa^l$ between $SBC_{2D}$ and $SM_{2D}$ for each orientation. The final $SBC_{2D}$ segmentation result, denoted by $SBC'_{2D}$, corresponds to $SBC_{2D}$ for which $\kappa^l$ is maximum.

![Fig. 6. SBC segmentation in 2D. (a) A midsagittal slice $M_k$. (b,c,d,e) The narrow-bands $NB^l$ and (f,g,h,i) the brain candidates $SBC_{2D}$ (orange) segmented in $NB^l$, overlayed on the rotated shape model $SM_{2D}$, for $\theta^l \in (0, \pi/2, \pi, 3\pi/2)$.](image)

We have used $L = 24$ rotation steps and $\sigma = 4$ in all the experiments. The erosion and dilation radius used to generate $NB^l$ ($R_e$ and $R_d$ respectively) are learned from $DB_k$, using manual segmentations of the $SBC_{2D}$, denoted $SBC_{2D}^{man}$. Let $r_e$ and $r_d$ be the minimum erosion and dilation radius guaranteeing that there exists an $l_0 \in L$ such that $\epsilon(SM'_{2D}) \subset SBC_{2D}^{man}$ and $SBC_{2D}^{man} \subset \delta(SM_{2D})$ for any element of $DB_k$. We ended up using $R_e = 0.5r_e$ and $R_d = 1.5r_d$, to tolerate $SBC_{2D}$ to segment to be smaller or bigger than all $SBC_{2D}$ represented in $DB_k$ and to ensure that boundaries are included in a given $NB^l$.

Figure 6 presents the midsagittal slice $M_k$ from one dataset and intermediate results for $\theta^l \in (0, \pi/2, \pi, 3\pi/2)$. For this image $R_e = 20$, $R_d = 18$. Figure 7 plots $\kappa^l$ values for all $\theta^l$. The highest $\kappa^l$ was obtained for $\theta^l = 3\pi/2$ and $SBC_{2D}^{man}$ shown in Figure 6 (i) corresponded to $SBC_{2D}^{man}$. For each element of the database, the mean distance, the kappa measure and the Hausdorff distance were computed between $SBC_{2D}^{man}$ and $SBC_{2D}^{man}$. The mean values on the
The mean distance between voxels is $SBC$. We set $w_{pq}$ if $I_q > I_p$ and $w_{pq} = 1 / \text{dist}(p, q)$ if $I_p \leq I_q$.

The mean kappa measure between the $SBC$ manual segmentation $SBC^{man}$ and $T(SM)$ was 93% on the database. This high accuracy of the registration process derives directly from the good quality of the $SBC$ segmentation in $M_k$. The erosion and dilation radiiues used to generate $NB$ are learned from $DB_k$ using the approach presented in Section 4.2. $T(SM)$ being close to the $SBC$ boundaries, small radiiues are considered. Hence, a very constrained $NB$ is defined which is crucial to obtain a correct segmentation of the $SBC$. As a matter of fact, partial volume effect degrades the quality of the algorithm. The mean value of the mean distance $\mu_{md}$, the kappa measure $\mu_\kappa$, and Hausdorff distance $\mu_{hd}$ were computed between the segmentation result and $SBC^{man}$ considering the whole database. High precision segmentations were obtained: $\mu_{md} = 0.5$ mm, $\mu_\kappa = 0.98$ and $\mu_{hd} = 3.4$ mm. Figure 8 presents the $SBC$ segmentation result in orthogonal views for one $I_k$.

### 5. CONCLUSION

A fully automatic method is proposed to segment the skull bone content on fetal MRI. The fetal eyes are first localized using a template combining contrast, shape and biometric priors. Based on the eyes localization, the midsagittal plane $M_k$ is reconstructed and the $SBC$ is segmented in $M_k$. This intermediate 2D segmentation result is used to embed a $SBC$ shape model in the 3D data, in order to segment the structure in a restricted region. The segmentation is performed using a graph cut approach on a graph exploiting contrast priors. Precise segmentation results were obtained on 24 fetal MRI cases, which can be used for morphological and biometrical analysis of the fetal $SBC$. The whole segmentation process takes less than 5 minutes on a 3.5 GHz PC running MATLAB (The MathWorks Inc.).

The method is applied on images on which the fetus orientation is unknown. The fetal eyes localization and the $SBC$ segmentation provide information on the fetus head orientation. Benefiting from the detection of the fetal spine root in the $SBC$, the segmentation of the fetus spine is now in progress, in order to increase the knowledge on the fetal trunk orientation and size. Further works will include a segmentation of the brain in the $SBC$ to extract morphological, biometrical and maturational information.

### 6. REFERENCES


