Automatic 3D foetal face model extraction from ultrasonography through histogram processing

Original
Automatic 3D foetal face model extraction from ultrasonography through histogram processing / Bonacina, Luca; Froio, Antonio; Conti, Daniele; Marcolin, Federica; Vezzetti, Enrico. - In: JOURNAL OF MEDICAL ULTRASOUND. - ISSN 0929-6441. - STAMPA. - 24:4(2016), pp. 142-149.

Availability:
This version is available at: 11583/2648638 since: 2018-03-23T12:41:24Z

Published
Elsevier

DOI:10.1016/j.jmu.2016.08.003

Terms of use:
oneAccess
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)
Automatic 3D Foetal Face Model Extraction From Ultrasonography Through Histogram Processing

L. Bonacina\textsuperscript{a}, A. Froio\textsuperscript{b}, D. Conti\textsuperscript{c}, F. Marcolin\textsuperscript{d}, E. Vezzetti\textsuperscript{d,*}

\textsuperscript{a}Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
\textsuperscript{b}Dipartimento Energia, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
\textsuperscript{c}Dipartimento Scienza Applicata e Tecnologia, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
\textsuperscript{d}Dipartimento di Ingegneria Gestionale e della Produzione, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Abstract

Ultrasound is by far the most adopted method for safe screening and diagnosis in the prenatal phase, thanks to its non-harmful nature with respect to radiation-based imaging techniques. The main drawback of ultrasound imaging is its sensitivity to scattering noise, which makes automatic tissues segmentation a tricky task, limiting the possible range of applications.

An algorithm for automatically extracting facial surface is presented here. The method provides a comprehensive segmentation process and does not require any human intervention or training procedure, leading from the output of the scanner directly to the 3D mesh describing the face. The proposed segmentation technique is based on a two-steps statistical process that relies on both volumetric histogram processing and 2D segmentation. The completely unattended nature of such a procedure makes it possible to rapidly populate a large database of 3D point clouds describing healthy and unhealthy faces, enhancing the diagnosis of rare syndromes through statistical analyses.

Keywords: DICOM, 3D ultrasound, Image segmentation, 3D histogram processing

1. Introduction

The popularity of 3D ultrasound (US) scans (figure 1) has rapidly increased during the last years [1]. However, even though it could provide useful hints to the physicians, its importance as a diagnostic tool is somewhat limited in prenatal imaging, as nowadays it is mainly used as a visualization tool. On the other hand, it is self-evident that a detailed 3D model of the foetus could play a crucial role in prenatal diagnosis. For instance, the

\textsuperscript{*}Corresponding author

Email addresses: bonacina.l@gmail.com (L. Bonacina), antonio.froio@polito.it (A. Froio), danielo.conti@polito.it (D. Conti), federica.marcolin@polito.it (F. Marcolin), enrico.vezzetti@polito.it (E. Vezzetti)

Figure 1: Three examples of 3D ultrasound images exported from 4DView by General Electric [15]. Given that these are rendered 2D images, although image quality is among the best achievable today, they cannot be 3D-processed.

Figure 2: Scheme of the algorithm. The “Volumetric Processing” and “Slice Processing” blocks are repeated three times, extracting the slices from the 3D volume along three perpendicular directions.

detection of dysmorphic features could be tackled by exploiting biometrics and landmark identification on a 3D model of the foetus [2–9]. Algorithms and software exist for 3D US segmentation and volume extraction, mainly based on region growing [10, 11] or deformable models [12, 13], but human intervention is requested, e.g. to choose the seeds or for pre/post-processing. The algorithm by Feng et al. [14] is automatic, but it requires a training stage before it can be applied; moreover, its output is a coarse mesh.

The method here proposed automatically extracts relevant 3D information of foetus’s face from a set of planar 2D US slices, giving as output the a high resolution mesh of the facial surface; being not based on machine learning techniques, it does not require any training step. To achieve the maximum interoperability, the whole procedure has been designed and tested with DICOM [16] files as input and a triangular mesh in PLY [17] files as output. Although the algorithm has been implemented using MATLAB®, it is easily portable to any other language.

The main phases of the procedure are shown in figure 2.
2. Methods

2.1. Preprocessing

The DICOM is initially imported as a stack of 2D planar slices. This can be done in any of the three coordinate directions, as the algorithm does not depend on the foetus’s position. Input stack $I$ is a now an N-dimensional matrix (3D in ultrasonography case) such that:

$$I(v) = g,$$

where $v = (r, c, s)$ is a voxel, $r, c, s$ are respectively row, column, and slice indexes, and $g$ is an integer value ($0 \leq g \leq (2^N - 1)$, where $N$ is the image bit depth) representing grey-scale.

US wave scattering damages the images by introducing a type of noise called speckle [18]. In order to enhance segmentation performance, in the first phase noise is removed through filtering. Current state-of-the-art algorithms performing this task are based on statistical analysis, inverse PSF filtering, and log separation [19]. In our algorithm, a standard 3D low-pass Gaussian filtering (or, alternatively, a Laplacian of Gaussian to better preserve the edges) resulted accurate enough for a preliminary speckles removal. Nonetheless, the most effective image cleaning is done after the extraction of the surface. After this preprocessing step, the image improves in terms of quality and helps enhancing the performance, although it is not clean enough for achieving complete surface extraction yet.

2.2. Surface reconstruction through 3D histogram processing

The most widely adopted method to extract a volume from a point cloud relies on the marching cube algorithm [20]. To extrapolate the facial surface, the detection (per each DICOM slice) of the edges of the face itself would be desirable.

Both volume and surface reconstruction are demanding tasks, due to the coexistence of different tissues in the same image, such as mom’s and foetus’ ones. Edge extraction is even trickier when the foetus’s head lays directly on mother’s uterus; in this scenario, the two tissues are virtually merged from the US imaging standpoint. To face this situation, a novel algorithm has been developed, based on a statistical analysis of local and 3D histograms that combines slice and volumetric processing.

2.2.1. Volumetric processing

As a first step, the algorithm processes the entire stack as a whole, applying a star-shaped cubic kernel to each voxel of the 3D array higher than a threshold grey-scale value $t$.

The kernel support is a subset of the stack. In particular, for each voxel $v = (r, c, s)$, considering a side $2l + 1$ (with $l$ chosen by the user, in order to finely tune the result) and the cube $C^v(l) = \left\{ [r-l, r+l] \times [c-l, c+l] \times [s-l, s+l] \right\}$ centred in $v$, the kernel $K^v_l$ will be:

$$K^v_l(i, j, k) = \begin{cases} 
I(i, j, k) & (i, j, k) \in S^v(l) \subset C^v(l) \\
0 & (i, j, k) \in C^v(l) \setminus S^v(l) 
\end{cases}$$

(2)
with:

\[ S^v(l) = \left\{ (i, j, k) \in C^v(l) : \right. \]
\[ \left. \left( |i - r| = |j - c| \land k = s \right) \lor \left( |i - r| = |k - s| \land j = c \right) \lor \right. \]
\[ \left. \lor \left( |j - c| = |k - s| \land i = r \right) \lor \left( |i - r| = |j - c| = |k - s| \right) \lor \right. \]
\[ \left. \lor \left[ |i - r| + |j - c| + |k - s| = \max(|i - r|, |j - c|, |k - s|) \right] \right\} \]  \( (3) \)

In other words, this is equivalent to extracting the voxels belonging to the main diagonals of \( C^v(l) \), as well as those on the diagonals of the three mid-planes of the cube (i.e., the three coordinate planes passing through \( v \)) and on the segments passing through \( v \) on the three coordinate axes. The \( 0 \)-centred kernel, with the default value of \( l = 10 \), is plotted in figure 3.

The grey-scale histogram \( H^v \) of the kernel \( K^v \) will be computed using the standard iterative formula:

\[ H^v(K^v(i, j, k)) = H^v(K^v(i, j, k)) + 1, \quad (i, j, k) \in S^v(l) \]  \( (4) \)

starting from \( H^v = 0 \). A local (per voxel) threshold value \( L(v) \) is set to the \((100 - p)\)-th percentile of such histogram, with \( p \) being an input parameter to be set by the final user. Thus, as an intermediate output, a new 3D array is produced, containing in each cell the local threshold value.

### 2.2.2. Slice processing

A second statistical analysis is then performed on a per-slice basis, i.e. the volume segmentation is obtained via individual segmentation of each 2D slice by means of a new global thresholding technique [21]. All the local threshold values lower than \( t \) are set to
0 and the slice histogram is computed. The global slice threshold is set to $T(s) = M - \left\lfloor \sigma \left(2^N - 1\right) \right\rfloor$, where $M$ is the brightest pixel appearing at least once in the slice, $\sigma$ is the standard deviation of the computed histogram, $N$ is the image bit depth, and $\left\lfloor \cdot \right\rfloor$ represents the nearest integer function. The volume to be extracted is then mapped into a new output binary slice $B(\mathbf{v})$, defined as

$$B(\mathbf{v}) = \begin{cases} \text{true} & I(\mathbf{v}) \geq T(s) \\ \text{false} & \text{otherwise} \end{cases} \quad \forall \mathbf{v} = (r, c, s).$$

Each slice of the obtained binary volume represents a “filled” level set of the facial surface (figures 6c to 10c). The algorithm is reported as pseudo-code in AppendixA.

2.3. Data fusion

Potentially, the model extracted after the second step is robust and complete enough to recognize the shape of the foetus’s face, but to enhance detection and improve the accuracy we have added a data fusion [22] step to the workflow. We repeat the previous 3D shape extraction from all the three directions by permuting the dimensions of the DICOM stack, and combine the extracted surfaces evaluating the logical AND of the three binary arrays. This operation increases the robustness of the algorithm, as it is not likely that the artefacts generated by the processing from a viewpoint would be the same generated changing the point of view. This is why the data fusion step generates further noise removal and increased veracity. Moreover, this kind of procedures heavily reduces the number of false positives, i.e. edges that do not actually belong to the desired surface.

2.4. Post-processing

After the data fusion, the output is fairly “clean”, but can still be refined via post-processing. Three main operations take place in this phase:

2.4.1. Edge detection

As a first stage, the Marr–Hildreth laplacian of Gaussian 2D edge detection algorithm [23] is applied to each slice of the fused stack to obtain another binary stack with a more precise definition of the facial surface.

2.4.2. Island removal

The whole 3D image is further “cleaned” by removing the isolated edges which are not connected to the main surface. To perform this task, the algorithm searches for all the 3D $c$-connected components, together with their bounding boxes. $c$ is an input parameter chosen by the user, with a default value of 26. Finally, only the connected components belonging to the $r$ bounding boxes with the largest volumes are kept. The algorithm for island removal is reported as pseudo-code in AppendixB.
2.4.3. Point cloud and mesh generation

In this phase, each true value in the 3D matrix represents a point of the surface. Thus, the final 3D point cloud can be obtained by simply looping over the 3D structure and assigning the Cartesian coordinates to each non-zero element, starting from an arbitrary origin. Finally, the surface is reconstructed via the Poisson surface reconstruction\(^{[24]}\) algorithm, exploiting the publicly Available pointCloud2rawMesh function \(^{[25]}\). The resulting vertices and faces are written in the PLY file.

3. Results and Discussion

From September 2012 to January 2013, 30 three-dimensional volumes of 30 foetuses at 22–32 weeks’ gestation were acquired. Written informed consent was obtained from the parents for publication of clinical details, clinical images, and videos. Principles outlined in the Declaration of Helsinki have been followed.

Among these acquisitions, five were selected and processed for the purposes of the study. The remaining ones were excluded due to high noise or practical “inconveniences” such as baby’s hands on face, or simply too much inaccuracy in the scanning process.

The ultrasound equipment was a Voluson system (GE Healthcare, Wauwatosa, WI, USA), with a RAB 4–8 (real time 4D convex transducer probe). The GE RAB 4–8 has a frequency range of 4 to 8 MHz and is used for OB applications (Footprint \(63.6 \times 37.8\) mm, FOV \(70^\circ\), V \(85^\circ \times 70^\circ\)).

Indeed, the algorithm has been tested upon five 3D US scans. A rendered example of the resulting 3D model is shown in figure 4. To show the clearness of the output, this model has also been 3D printed, as shown in figure 5, proving also how this technology might enable alternative business scenarios and applications, e.g. sales of 3D printed foetus face. The intermediate steps are shown in figures 6 to 10 (the rendered example corresponds to that of figure 6), reporting a 2D slice in each phase for each of the five 3D US source scans. In all the cases, the 2D slice corresponds to a transverse plane cutting the face at eye level, as shown by the dash-dotted line in figure 5. The regions of interest are in the upper part of the surfaces where the nose is clearly visible; the rear part of the foetus’ heads are instead full of noise (visible in the lower part of the images), and thus cannot be processed. However, please note that this does not affect the validity of the proposed method (e.g., it is visible also in the rendered examples shown in figure 1), as it is due to the US scattering noise. Moreover, the surfaces of interest for the present work are those on the front part of the head (i.e. the face), as the final aim is to produce a large set of foetus face models for the statistical pre-birth detection of face dysmorphic features \(^{[9]}\).

Figure 6 shows the segmentation process; the effectiveness of all the steps is well-rendered. The validity of the proposed segmentation technique pointedly emerges in figure 6c. The data fusion step removes noise from the regions of interest, i.e. in the back side, and the small island in the lower right corner of the slice is easily removed during post-processing.

In particular, the effectiveness of the data fusion step is self-evident in the example of figure 7. Here the main volume is properly extracted after the local 3D histogram procedure, but some of mother’s tissues are still visible on the right side in figures 7a to 7c. On the
Figure 4: Rendering (produced with MeshLab [26]) of the triangular mesh obtained by the algorithm. The input US is the same as figures 5 and 6 below.

Figure 5: 3D printed version of the model shown in figure 4 above. The dash-dotted line represents the plane where the slices shown in the next figures are extracted.
Figure 6: Example of the output of the different steps. The segmentation of the surface of interest is excellent in this case.
Figure 7: Example of successful application of the algorithm. The efficacy of the data fusion step is self-evident.

contrary, the surface reconstruction from the other two perpendicular directions, together with the data fusion step, results in a complete elimination of all the unwanted tissues (figure 7d).

The example in figure 8 shows the importance of the island removal. Even though their dimensions are strongly reduced, two portions of unwanted tissue survive to the data fusion (in the upper left corner and in the centre of the frame, as shown in figure 8d). These two islands disappear after the last step; the suppression of the upper one is of utter importance, as it stands above the surface of interest. The only small island left, in the lower left corner in figure 8f, survives to this phase because it is connected to the foetus in other slices. In this sense, it is important to underline the strong connection between the data fusion and the island removal. As seen in this example, the reduction of the noisy effects obtained by the former might be crucial for the success of the latter.

In figure 9, instead, the algorithm does not perform as expected. In this case the face is completely laid on the mother’s tissue, as can be seen in figure 9a, and the algorithm is not able to cope with this situation. In fact, the mother’s tissue is segmented together with the desired one and, even though the profile is still evident, the extracted surface is less meaningful. However, with a reduced manual post-processing, the actual surface might still be extracted. In addition, it can be noticed that in this case also the island removal fails, although an island is distinctly visible in figure 9e; this is due to the connection of that portion of mother’s tissue with the foetus’s ones in the bottom slices. Finally, in figure 10 another example of successful segmentation is reported.

The procedure is completely automatic, enabling fully unattended segmentation. This leads to the possibility of creating a large database of 3D facial point clouds, like the ones required for morphological studies through statistical analyses, without the necessity of human intervention. This kind of work has been already started [9]. Our tool is applied to build the database for clustering faces into healthy and unhealthy.
Figure 8: Example of successful application of the algorithm. The efficacy of the island removal is mainly evident.

Figure 9: Example of the output of the different steps. In this case, the foetus’s face is completely laid on the mother’s tissue (clearly visible on the right part of (a)), and the segmentation does not work as expected.
Figure 10: Example of successful application of the algorithm.

The algorithm works with standard protocols and file format (DICOM as input and any mesh file format as output), so it is completely platform-independent.

The algorithm is preliminary and might be refined, for instance by acting on the kernel shape or on the post-cleaning techniques (e.g. to obtain efficient cleaning also in situations like the one in figure 9d). The same idea may also be extended to internal tissues segmentation to reach a complete 3D model of foetus’s face and of other organs (e.g. heart). Furthermore, it can be extended to other applications of US imaging and for working with different 3D imaging techniques (CT or MRI), taking advantage, in this case, of the absence of the speckle noise effect.

4. Conclusions

An algorithm for facial automatic segmentation from 3D foetal ultrasound scans has been developed. The stack of 2D images is initially pre-processed, to remove speckle effects for a more assisted segmentation. This is achieved through a standard 3D Gaussian low-pass filtering.

The second step consists in the actual segmentation, which is performed with a new two-steps statistical processing, based on both 3D (volumetric) histogram processing and 2D (per-slice) global thresholding. In the first stage, an intermediate 3D structure is created, containing a local threshold chosen on the basis of the histogram of a 3D kernel. The histogram of each slice of this structure is then computed to find a global grey-scale threshold per each slice. The whole procedure is carried out three times, “cutting” the input slices among the three coordinate axes; the three binary masks are then fused together to reduce the number of false positives.

The surface point cloud is extracted from the fused binary stack through edge detection. Then, it is further “cleaned” by removing all the isolated edges not belonging to the tissue of interest. Finally, the cloud is meshed using the Poisson surface extraction procedure and saved in any standard mesh format (.ply in our case).
The whole procedure can be executed unattendedly, allowing the segmentation of large batches of 3D US scans.

Appendix A. Volume extraction algorithm

**input:** $I$ (3D image array), $l$ (proportional to kernel dimension), $s$ (threshold), $p$ (percentile for threshold selection)

1: for each slice $k$ do
2: for each $i$ voxel in the $k$-th slice do \(\%\) volumetric processing
3: \quad if $I(i) \geq s$ then \(\%\) voxel is "bright enough"
4: \quad apply the cubic kernel of side $l$ centred in voxel $i$, as described in section 2.2.1;
5: \quad extract the voxels belonging to the main diagonals of the kernel;
6: \quad compute the histogram of this subset of voxels;
7: \quad identify the $(100 - p)$-th percentile of the histogram;
8: \quad set the $i$-th local threshold equal to the percentile above;
9: end if
10: end for
11: \(\%\) slice processing
12: set all local thresholds $< s$ equal to 0;
13: compute the histogram of the $k$-th thresholds slice;
14: estimate the histogram standard deviation $\sigma$;
15: identify the highest value $M$ appearing at least once;
16: set the $k$-th global threshold as $M - \left\lfloor \sigma \left(2^N - 1\right) \right\rfloor$;
17: create the binary output slice, putting ones in the pixels above threshold and zeros otherwise;
18: end for

Appendix B. Island removal algorithm

**input:** $I$ (3D binary fused stack), $r$ (number of connected components to keep), $c$ (connectivity of the components to keep)

1: find all the 3D $c$-connected components in $I$;
2: for each connected component $CC$ do
3: \quad find the bounding box $BB(CC)$;
4: \quad compute the volume of $BB(CC)$;
5: end for
6: set to false all the voxels not belonging to the $r$ connected components with the largest bounding boxes;

References

URL http://www.volusonclub.net/emea/4dview
URL http://paulbourke.net/dataformats/ply/
[26] P. Cignoni, M. Corsini, G. Ranzuglia, MeshLab: an open-source 3D mesh processing system, Ercim
URL http://meshlab.sourceforge.net