Graph-based whole body segmentation in fetal MR images

T. Zhang[1], J. Matthew[1], M. Lohezic[1], A. Davidson[1], M. Rutherford[1], D. Rueckert[2], J.V. Hajnal[1], and P. Aljabar[1]

[1] Departments of Biomedical Engineering & Perinatal Imaging and Health, King's College London, London, UK
[2] Department of Computing, Imperial College London, London, UK

Abstract. Fetal body segmentation in MR images has an important role in fetal weight estimation, growth monitoring and anomaly detection. Significant fetal motion during MR image acquisition, especially around mid-gestation, makes automatic segmentation of fetal body very challenging. In this work, we propose an automatic graph-based framework to segment uterine images into amniotic fluid (AF), fetal body and background. First, the image is over-segmented into a set of supervoxels, which are then viewed as nodes in a weighted graph model. The AF region is then defined by clustering the supervoxels based on intensities and connectivity. The detected AF region is used to define a region of interest (ROI) within which we randomly select a set of edge nodes to act as queries. All graph nodes are then ranked based on similarities to the queries using manifold ranking. The fetal body is then segmented iteratively with the updated query node(s) and volume constraints that derived from manual delineations. The proposed method has been tested on 10 uterine MR images with comparable US measurements and a promising comparison with manual segmentation results has been carried out.

Keywords: fetal imaging, magnetic resonance, graph labelling, image segmentation, manifold ranking

1 Introduction

Compared with ultasonography (US), which is routinely used for antenatal screening, magnetic resonance (MR) imaging has a number of advantages, such as extended field-of-view (FOV) and superior soft-tissue contrast. In particular, the limited FOV of fetal US after the first trimester makes it difficult to visualise and examine the whole fetus, placenta and even large organs. MR imaging (MRI) has gained popularity for prenatal examinations in recent years, especially when abnormalities are not clearly defined in US imaging or when they are detected in US but require further investigations [6].

Segmentation of fetal body and computing fetal body volume (FBV) have been extensively studied in a range of antenatal research projects and clinical

studies and in a variety of applications [2–5, 10, 12, 13]. For instance, Cannie et al. [5] showed that fetal total lung volume is more strongly correlated with FBV than gestational age (GA). Seed et al. [12] investigated correlations among fetal blood flow, gestational age, and fetal weight, where fetal weight is calculated from fetal density and FBV with the regression formula $weight\,(g) = 0.12 + 1.031 \times FBV\,(ml)$ derived by Baker et al. [4] More recent studies [10, 12] have reported that, for fetal weight estimation, more accurate results are obtained from MR than US. Based on comparisons with actual birth weight, Kacem et al. [10] estimated significant differences in the error rates of MRI and US weight estimations (1.1%, and over 10%, respectively).

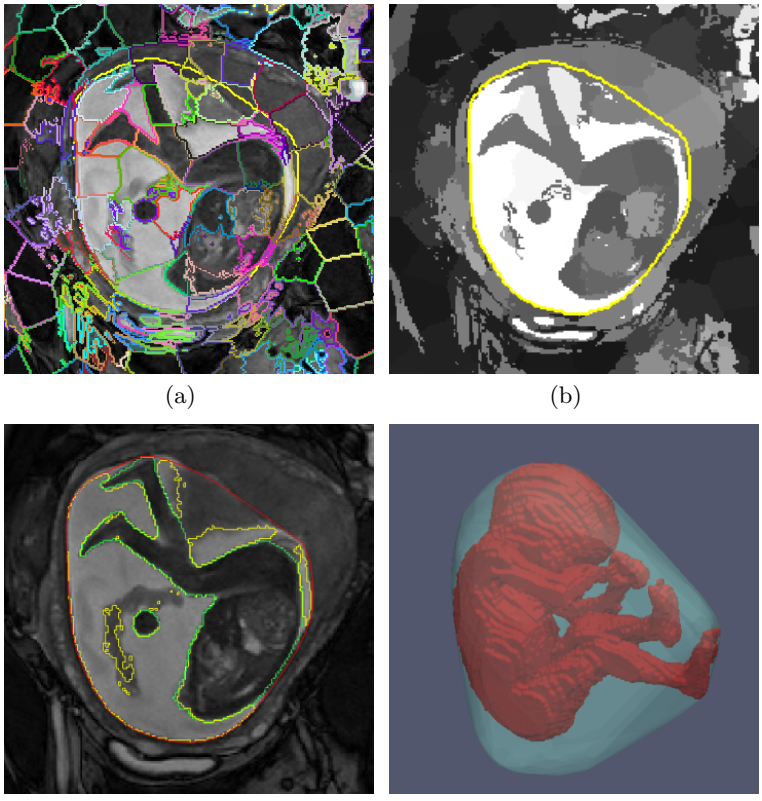


Fig. 1. Examples of fetal MRI and the segmented 3D volumes: (a) uterus MRI scan with the boundaries of a supervoxel segmentation; (b) supervoxels showing with mean intensities and a defined ROI; (c) segmented fetal body, AF and ROI; (d) 3D rendering of the fetus and ROI

Typically, the fetal body is manually delineated by experts, which is time-consuming and subjective process and motivates the development of more automated methods although motion corruption and other artefacts in fetal MRI

scans make automatic segmentation of the fetal body very challenging. Anquez et al. [3] developed a semi-automatic method for segmenting the fetus between 30 to 35 weeks GA applied to a steady state free precession (SSFP) MRI sequence. They first defined a fetal skeleton based on a set of landmarks, and then registered a fetal image to the skeleton using the landmarks. Finally, the registration result is used to annotate the fetal body and segmentation is achieved with a graph-cut approach. While this method does not require manual segmentation, landmark annotation may also entail significant time and effort.

Effective monitoring of fetal growth and timely detection of disease requires imaging and analysis of fetal images across the gestational range. In particular, accurate estimation of FBV and fetal weight at mid-term is important for early diagnosis of anomalies. Compared to later in gestation, mid-term MR images of the fetus are more motion corrupted, which necessitates a fast imaging protocol along with robust methods to compensate and correct for motion. For imaging the whole uterus, we employ a fast balanced SSFP (bSSFP) sequence [7].

In this work, we present an automatic fetal body segmentation method for mid-term subjects in the age range $20-24$ weeks GA. bSSFP MRI scans contain maternal tissue, amniotic fluid (AF), placenta and the whole fetus. In order to break the problem down, we propose a two-stage graph-based segmentation framework which can segment the 3D fetal image into AF, fetal body and 'other'. In the first stage, the image is over-segmented into a set of supervoxels, which are used to construct a weighted graph model. Based on the graph connectivity and intensity, the AF region is derived and then used to restrict the region of interest (ROI). The second stage focuses on the sub-graph within the ROI. We randomly select a node(s) at the edge of the ROI as a query, and rank all graph nodes based on their similarities to the query via manifold ranking [14, 15]. Finally, the updated query node(s) and volume constraints derived from manual delineations are used to segment the fetal body in an iterative procedure.

2 Methodology

2.1 Overview

The MRI data used in this work have an in-plane resolution of $1.25 \times 1.25\ mm$ and $5\ mm$ slice spacing with a $2.5\ mm$ overlap. Each slice can be acquired with $500ms$, which freezes nearly all in-plane fetal movement. An example of bSSFP fetal image is shown in Fig 1.

The proposed approach for the automatic segmentation of the fetal body consists of two main stages as shown in Fig 2. In the first stage, the ROI that contains AF, the whole fetal body and a partial placenta is identified. The volume priors of AF, fetal body and 'other' within the ROI are obtained from a set of manual segmentations beforehand. The second stage iteratively segments the fetal body via graph-based manifold ranking and the volume constraints.

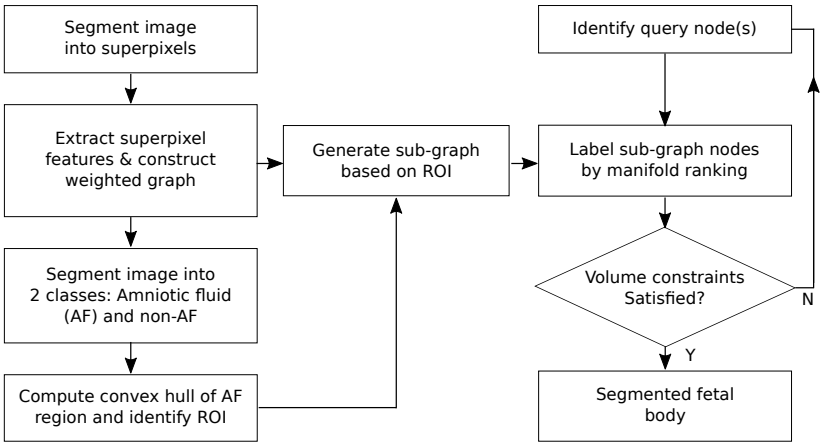


Fig. 2. The workflow of the proposed fetal body segmentation method.

2.2 Stage I: ROI Identification

With improved computation and boundary adherence properties, superpixel/supervoxel based segmentation methods have been extensively researched for a variety of 2D/3D image processing tasks [1, 9, 14]. In this paper, we use the 3D version of the popular Simple Linear Iterative Clustering (SLIC) algorithm [1] to generate an over-segmentation with a set of supervoxels. Based on the spatial localisations and k-means clustering, the SLIC algorithm is well recognised for its efficiency for segmenting images into homogeneous regions. Denote $X = \{x_1, x_2, ..., x_n\}$ as a 3D fetal MR image, where x_i represents the i-th supervoxel.

We subsequently define a graph $G = (X, E)$ on the 3D image X, where each supervoxel x_i represents a graph node. The nodes with shared bounderies are defined as connected and a weighted graph is then constructed in which the weights are defined by the Kullback-Leibler (KL) divergence of the intensity distributions of adjacent supervoxels.

We carry out segmentation of the AF where, for simplicity, we use the mean value to represent each supervoxel. In conjunction with the graph connectivities, this is generally straightforward and successful in characterising the AF as the intensities of supervoxels in the AF region tend to be significantly high as illustrated in Fig 1(b). In this work, we first define a coarse ROI which is a cuboid centred in the 3D volume with half the side lengths of the whole field of view. After combination of the ROI with high mean valued nodes, the largest connected component is then defined as the AF region. It should be noted that partial nodes within brain region are usually included during this clustering pro-

cess, as they also have generally high intensities and it is possible for supervoxel containing brain tissue to be connected to AF supervoxels.

As shown in Fig 1, the fetus is surrounded by AF. Intuitively, the ROI for fetal body segmentation can be derived from the convex hull of the AF region. The convex hull represents the smallest convex region that contains both AF and the whole fetal body. In this study, the ROI for fetal body segmentation is defined as the convex hull of the AF region $ROI = hull(AF)$.

2.3 Stage II: Graph-based Fetal Segmentation

In the second stage, the fetal segmentation is turned into a graph labelling problem. We define a set of query node(s) as those in the largest connected component of the nodes that are on the boundary of the ROI and that have not previously been defined as AF. These nodes typically represent placental tissue. Given this initial query, all the graph nodes are then ranked based on their similarity to the query via manifold ranking [14]. High ranked nodes are more likely to be the same category as query nodes. The volume priors (see below) of AF and fetal body are derived from the manual segmentations beforehand and are used as a stopping criteria for the segmentation process.

Manifold Ranking is a distance metric learning tool [14, 15] that estimates a low dimensional manifold from a graph of the data which connects samples by their proximity. Distances among the features are computed by diffusing relevance values on the graph, typically using an iterative process. Manifold ranking has been shown to significantly improve accuracy for 3D model retrieval compared to a fixed 'simple' distance such as Euclidean or L_1-norm [11].

Let $X = \{x_1, x_2, ..., x_q, x_{q+1}, ..., x_n\}$ denote a set of graph nodes, where the first q nodes are the queries and the rest are to be labelled. Define a function $\mathbf{f} : X \to R^n$, which assigns a ranking value f_i to each node x_i, and $\mathbf{f} = [f_1, ..., f_n]^T$. Denote $\mathbf{y} = [y_1, y_2 ..., y_n]^T$ as an indicator vector, where $y_i = 1$ if x_i is query, and $y_i = 0$ otherwise. Define a sub-graph $G[V]$, where $V \subset X$ denotes a subset of $m = |V|$ nodes of X derived from the ROI. For the weighted subgraph $G = (V, W)$, $W = [\omega_{ij}]_{m \times m}$ is the symmetric weighting matrix, where ω_{ij} denotes the weight between nodes v_i and v_j. Define $D = \mathrm{diag}(d_{11}, ..., d_{nn})$ as the degree matrix, where $d_{ii} = \sum_j \omega_{ij}$. Manifold ranking seeks to optimise the following equation

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} \frac{1}{2} \sum_{i,j=1}^{n} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 + \mu \sum_{i=1}^{n} \|f_i - y_i\|^2$$

where μ is a balancing coefficient that controls the data fidelity to the query node(s). During each iteration, similarly to spectral clustering, the optimal ranking vector $\mathbf{f}^*$ can be obtained from unnormalised Laplacian matrix:

$$\mathbf{f}^* = (D - \alpha W)^{-1} y$$

where $\alpha = 1/(1+\mu)$. Here, we fix the percentage of the ranked nodes that are assigned to the placenta class based on empirical experience according to the number observed in previous data sets.

3 Experiments and Results

Datasets: Research ethics committee approval and informed consent was obtained from all pregnant women prior to the scans. MR examinations were performed on a 1.5 Tesla MR scanner (Ingenia; Philips Medical Systems, Best, the Netherlands) using a combination of posterior and anterior coil arrays (28 channels in total). Ten volunteers, with uneventful pregnancies, were included in the study at gestational ages ranging from 20 weeks and 2 days to 24 weeks and 6 days. One multislice bSSFP sequence covering the whole gravid uterus was acquired in free breathing along the subject's coronal orientation, with $TR = (3.8 - 4.0)\,ms$, $TE = (1.89 - 1.98)\,ms$, filp angle $\alpha = 90^0$, and field-of-view $420 \times 420\ mm^2$. The total acquisition time was less than 2 minutes for all the volunteers.

Measurements and evaluation: Since the US based fomula is commonly used for fetal weight estimation, we also include it for comparison. US Measurements were obtained by two sonographers and, according to Hadlock's US weight estimation formula [8], abdominal circumference(AC), femur length(FL), head circumference(HC), and biparietal diameter(BPD) are measured for computing the estimated fetal weight (EFW):

$$\begin{aligned}\log_{10}(EFW) = 1.3596 - 0.00386\,AC\,FL + 0.0064\,HC \\ +0.00061\,BPD\,AC + 0.0424\,AC + 0.174\,FL\end{aligned}$$

It is not practical to directly measure the FBV from US scans after the first trimester of pregnancy. Based on the estimated fetal weight, Baker's formula can be used to estimate the fetal volume (EFV) by $EFV = (EFW - 0.12)/1.031$

The relative errors of the estimated fetal volumes were used to evaluate the approach using the volume of the manually delineated MR fetal volume MR_{FV} as a reference:

$$Error = \frac{100 \times |EFV - MR_{FV}|}{MR_{FV}} \tag{1}$$

The Dice similarity coefficient (DC) was also used to evaluate the accuracy of the fetal body segmentations as defined by

$$DC = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{2}$$

Fetal Body Segmentation results: In order to examine the performance of the proposed method, the segmentation algorithm was applied to 10 fetal MR images. In Fig. 3, the automatic segmentation results in three orthogonal views are shown in the first column, in the second and third columns are two expert manual delineations. It can be seen that, for this example, the automatic

segmentation results are close to those from human experts. The main difference is that the automatic segmentation results cover a larger area, but the main ROI has been identified correctly. Expert raters require about 45 mins to one hour to delineate each image. In comparison, the automated algorithm takes less than five minutes for all cases. In some cases, the automatic segmentation did not succeed, producing a less accurate segmentation of the body (Figure 4(b)) compared with human raters (4(a), 4(c)). We suspect that this may be due to a number of reasons including a relatively small volume of AF compared with other subject and the distribution of AF being predominantly on one side of the fetal body rather than surrounding it. For the segmentations provided by each of the manual raters, the mean Dice coefficient of the automated segmentation was 0.69 (S.D. 0.08). The average relative errors of the estimate fetal volume based on the automated segmentation of the MR data and from the US measurements are given in Table 1. These show that both approaches are comparable and both have a tendency to over-estimate the volume.

As shown in Fig 5, the FBV estimates derived from the proposed method are comparable with those estimated from the US scans. In particular, compared to the measurements obtained from manual delineations of the MR data, eight out of ten automated segmentations are overestimated, while the other two are underestimated. The red line indicates the linear regression of the mean FBVs from two groups of MR manual segmentations, while the blue line is the counterpart from the US measurements. It is worth noting that the EFV values derived from US measurements are consistently larger than the measurements obtained from manual segmentations of MR images. These results may indicate that either Baker's [4] or Hadlock's [8] weight estimation formulas may need to be adapted specifically to provide estimates of volume for mid-term fetues.

Table 1. The average relative errors of the EFV, using the MR manual segmentations as a reference for the proposed method and estimates derived from annotations of US scans.

Error Rate	Proposed	US1	US2
Mean	12.01%	13.82%	13.5%
S.D.	11.90%	9.02%	8.54%

4 Conclusion

In this work, we present a graph-based framework for fetal body segmentation in MRI data, which integrates supervoxel extraction and manifold ranking techniques. These preliminary experiments are based on 10 mid-term fetal MR scans using a bSSFP sequence. We fix the percentage of the ranked nodes assigned to the placenta class empirically, in future work we aim to define the number assigned based on learned shape or appearance features that may constrain the selection of nodes for the placenta. To further investigate the statistics of the

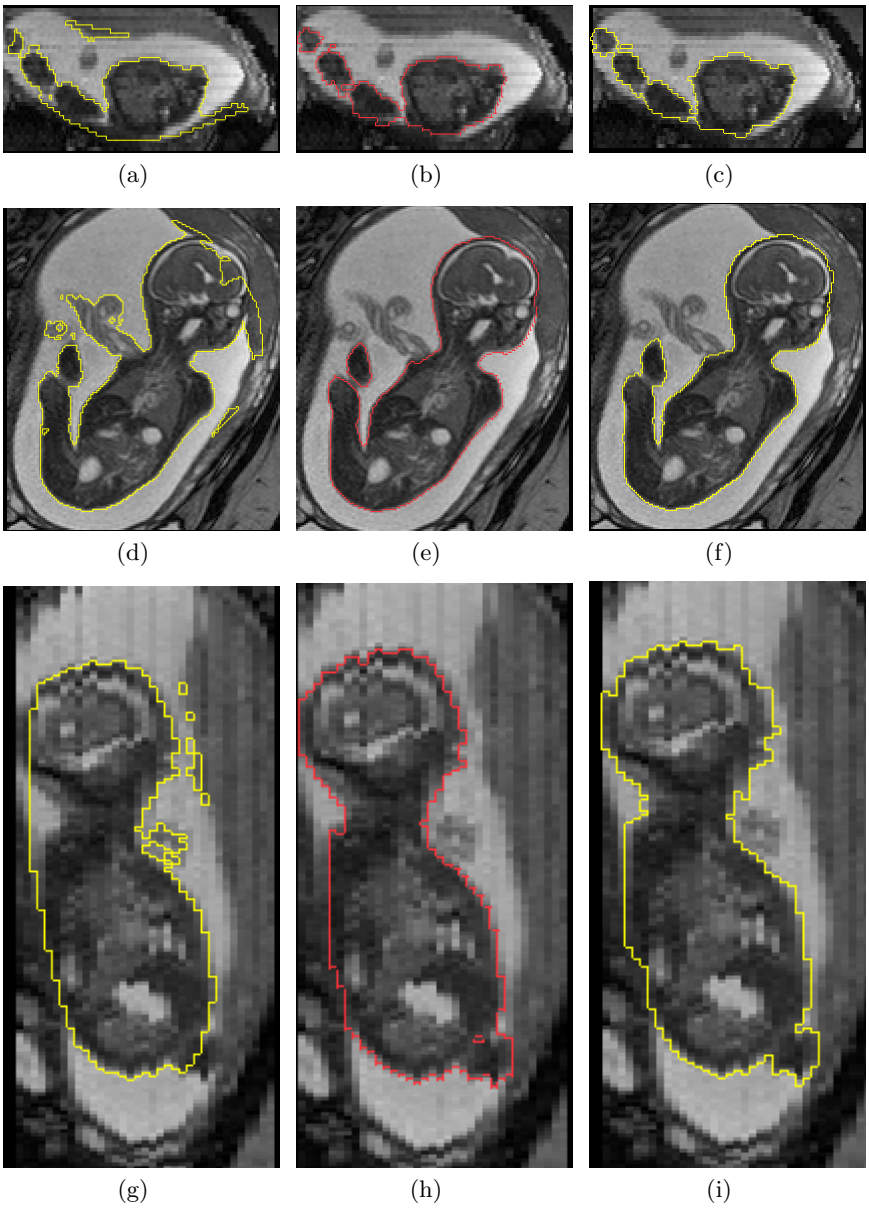
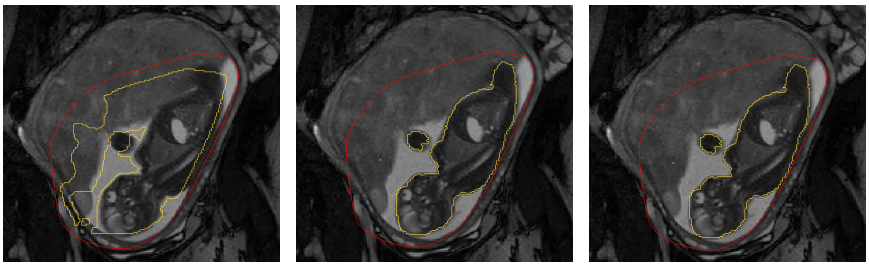


Fig. 3. Comparison of the segmentation results and two manual delineations in axial, coronal, and sagittal views: the left column shows the segmentation results from the proposed method; the middle and right columns are manual delineations from expert I and II.



(a) (b) (c)

Fig. 4. Example of a challenging subject. Red line denotes automated identified ROI and yellow line represents the segmentation derived from: (a) the proposed method; (b) manual delineation from expert I; (c) manual delineation from expert II.

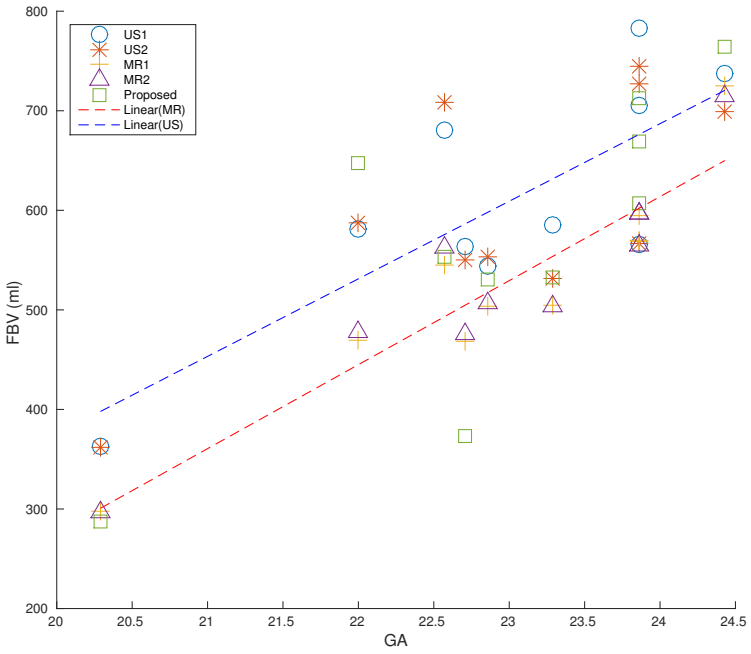


Fig. 5. Comparison of FBV (ml)

segmentation results, we will conduct more experiments on a variety range of MR images, especially those with wider GA range.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence 34(11), 2274–2281 (2012)
2. Anquez, J., et al.: Interest of the steady state free precession (SSFP) sequence for 3D modeling of the whole fetus. In: International Conference of the IEEE Engineering in Medicine and Biology Society.771–4 (2007)
3. Anquez, J., Bibin, L., Angelini, E.D., Bloch, I.: Segmentation of the fetal envelope on ante-natal MRI. In: 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 896-899 (2010)
4. Baker, P., Johnson, I., Gowland, P., Hykin, J., Harvey, P., Freeman, A., Adams, V., Mansfield, P., Worthington, B.: Fetal weight estimation by echo-planar magnetic resonance imaging. The Lancet 343(8898), 644–645 (1994)
5. Cannie, M.M., Jani, J.C., Van Kerkhove, F., Meerschaert, J., De Keyzer, F., Lewi, L., Deprest, J.A., Dymarkowski, S.: Fetal body volume at MR imaging to quantify total fetal lung volume: normal ranges. Radiology 247(1), 197-203 (2008)
6. DEWILDE, J., RIVERS, A., PRICE, D.: A review of the current use of magnetic resonance imaging in pregnancy and safety implications for the fetus. Progress in Biophysics and Molecular Biology 87(2-3), 335-353 (apr 2005)
7. Gholipour, A., et al.: Fetal MRI: A technical update with educational aspirations. Concepts in Magnetic Resonance Part A 43(6), 237-266 (nov 2014)
8. Hadlock, F.P., Harrist, R.B., Sharman, R.S., Deter, R.L., Park, S.K.: Estimation of fetal weight with the use of head, body, and femur measurements–a prospective study. American journal of obstetrics and gynecology 151(3), 333-7 (feb 1985)
9. Heinrich, M.P., et al.: Edge- and detail-preserving sparse image representations for deformable registration of chest MRI and CT volumes. Lecture Notes in Computer Science 7917, 463-474 (2013)
10. Kacem, Y., Cannie, M.M., Kadji, C., Dobrescu, O., Lo Zito, L., Ziane, S., Strizek, B., Evrard, A.S., Gubana, F., Gucciardo, L., Staelens, R., Jani, J.C.: Fetal weight estimation: comparison of two-dimensional US and MR imaging assessments. Radiology 267(3), 902–10 (2013)
11. Ohbuchi, R., Furuya, T.: Distance metric learning and feature combination for shape-based 3D model retrieval. In: Proceedings of the ACM workshop on 3D object retrieval - 3DOR '10. p. 63. ACM Press, New York, New York, USA (2010)
12. Seed, M., et al.: Feasibility of quantification of the distribution of blood flow in the normal human fetal circulation using CMR: a cross-sectional study. Journal of Cardiovascular Magnetic Resonance 14(1), 79 (2012)
13. Weidner, M., et al.: MRI-based ratio of fetal lung volume to fetal body volume as a new prognostic marker in congenital diaphragmatic hernia. American Journal of Roentgenology 202(6), 1330-1336 (2014)
14. Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.H.: Saliency detection via graph-based manifold ranking. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition pp. 3166-3173 (2013)
15. Zhou, D., Weston, J., Gretton, A., Bousquet, O., Schölkopf, B.: Ranking on Data Manifolds. Advances in Neural Information Processing Systems 16(c), 177-186 (2003)