PVR: Patch-to-Volume Reconstruction for Large Area Motion Correction of Fetal MRI

Amir Alansary, Martin Rajchl, Steven G. McDonagh, Maria Murgasova, Mellisa Damodaram, David F. A. Lloyd, Alice Davidson, Mary Rutherford, Joseph V. Hajnal, Daniel Rueckert, and Bernhard Kainz

***Abstract*— In this paper, we present a novel method for the correction of motion artifacts that are present in fetal magnetic resonance imaging (MRI) scans of the whole uterus. Contrary to current slice-to-volume registration (SVR) methods, requiring an inflexible anatomical enclosure of a *single* investigated organ, the proposed patch-to-volume reconstruction (PVR) approach is able to reconstruct a large field of view of non-rigidly deforming structures. It relaxes rigid motion assumptions by introducing a specific amount of redundant information that is exploited with parallelized patchwise optimization, super-resolution, and automatic outlier rejection. We further describe and provide an efficient parallel implementation of PVR allowing its execution within reasonable time on commercially available graphics processing units, enabling its use in the clinical practice. We evaluate PVR's computational overhead compared with standard methods and observe improved reconstruction accuracy in the presence of affine motion artifacts compared with conventional SVR in synthetic experiments. Furthermore, we have evaluated our method qualitatively and quantitatively on real fetal MRI data subject to maternal breathing and sudden fetal movements. We evaluate peak-signal-to-noise ratio, structural similarity index, and cross correlation with respect to the originally acquired data and provide a method for visual inspection of reconstruction uncertainty. We further evaluate the distance error for selected anatomical landmarks in the fetal head, as well as calculating the mean and maximum displacements resulting from automatic non-rigid registration to a motion-free ground truth image. These experiments demonstrate a successful application of PVR motion compensation to the whole fetal body, uterus, and placenta.**

***Index Terms*— Motion correction, fetal magnetic resonance imaging, GPU acceleration, image reconstruction, super-resolution.**

Manuscript received March 29, 2017; revised June 7, 2017 and July 19, 2017; accepted August 1, 2017. Date of publication September 1, 2017; date of current version September 29, 2017. This work was supported in part by the NIHR Biomedical Research Centre, Wellcome Trust/EPSRC under Award IEH 102431/NS/A000025/1, in part by ERC under Grant 319456, in part by the NVIDIA (GPU donations), and in part by EPSRC EP/N024494/1. The work of A. Alansary was supported by an Imperial College President's Ph.D. Scholarship. *(Corresponding author: Bernhard Kainz.)*

A. Alansary, M. Rajchl, S. G. McDonagh, D. Rueckert, and B. Kainz are with the Biomedical Image Analysis Group, Department of Computing, Imperial College London, London SW7 2AZ, U.K. (e-mail: b.kainz@imperial.ac.uk).

M. Murgasova, M. Damodaram, D. F. A. Lloyd, A. Davidson, M. Rutherford, and J. V. Hajnal are with the Department of Biomedical Engineering, King's College London, London WC2R 2LS, U.K.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2017.2737081

I. Introduction

THE advent of single shot fast spin echo (ssFSE) T2-weighted sequences has enabled spin echo image formation principles [1] in magnetic resonance imaging (MRI) to play an essential role in fetal diagnosis [2] and research [3]. In particular, cases for which ultrasound (US) fails to acquire conclusive image data benefit from fetal MRI [4], [5]. Recent advances in motion compensation for fetal MRI [6] facilitate advanced image-based diagnostics and lead to novel insights about the human development.

Fetal MRI enables an ability to distinguish between individual fetal structures such as brain, lung, kidney and liver, as well as pregnancy structures such as the placenta, umbilical cord and amniotic sac. It provides improved visualization and structural information of the fetal anatomy, which enables to study abnormalities during pregnancy such as neuro-developmental disorders [7], placental pathologies [8], fetuses with congenital lung masses [9], and conjoined twins [10]. MRI is considered to be safe after the first trimester [11] for 1.5T [12] and 3T [13] without the use of contrast agents, which may have teratogenic effects. Furthermore, this technology paves the way for researchers and clinicians to analyze correlations between childhood development and prenatal abnormalities.

During image acquisition the fetus is not sedated and moves freely as well as the mother breathes normally. As a result, movements are likely to corrupt the scans, hiding pathology and causing overlap between different anatomical regions. In order to limit these artifacts, fast scanning sequences such as ssFSE [14] allow for the rapid acquisition of single slices at high *in-plane* resolution in a large field of view and good tissue contrast of the uterus. However, when acquiring a 3D volume through a stack of slices, inter-slice artifacts in the *out-of-plane* views are highly likely. Consequently, this restricts reliable diagnostics to individual slices in the current clinical practice. Fig. 1 depicts a typical example of motion related artifacts in a fetal ssFSE scan. The observed motion (*c.f.* Fig. 1 b & c) is of unpredictable nature and consists of a combination of maternal respiration, bowel and fetal movements.

Slice-to-volume registration (SVR) combined with super-resolution image reconstruction techniques [15] can be applied to compensate motion between single slices by reconstructing a high-resolution (HR) image from multiple, overlapping low-resolution (LR) images, as shown in Fig. 2. To provide a sufficiently high number of samples for such an approach, multiple stacks of 2D-slices need to be acquired, ideally in

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/

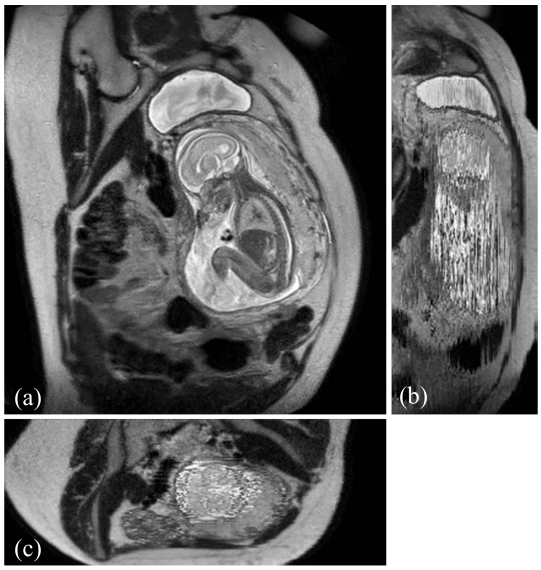


Fig. 1. Three view-planes for raw 3D data acquired through stacks of ssFSE images covering the whole uterus. The transverse (a) is the *in-plane* view, *i.e.*, native 2D slice scan orientation. Motion causes streaky artifacts for multi-planar reconstructions (MPR) in orthogonal views (b) and (c) caused by both maternal and fetal movements between the acquisition of individual slices.

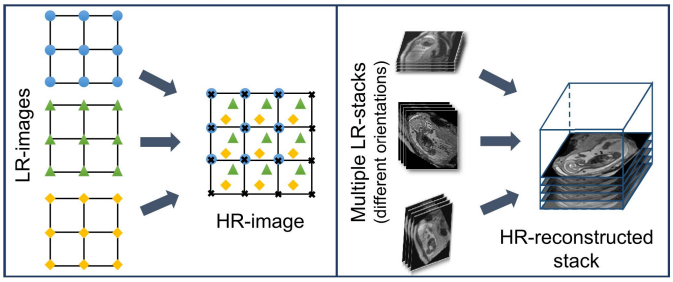


Fig. 2. Illustration of the basic ideas behind reconstruction [15]: A simplified example of a 2D 4×4 HR grid sampling from a 2D 3×3 LR grids (left) and a practical example of 3D fetal MRI using multiple overlapping stacks of slices, by reconstructing a 3D HR image with an isotropic voxel size from LR images with anisotropic voxel size.

orthogonal orientations. A simple LR $\rightarrow$ HR reconstruction model [15] can be formalized as:

$$Y_i = W_i X + n_i \quad \text{for} \quad 1 \leq i \leq N, \tag{1}$$

where Y_i denotes the i-th LR image of total N images, and X is the HR image. The matrix W_i combines motion, sub-sampling and degradation effects: $W_i = DBT_i$, where D is the sub-sampling matrix, B is the blurring matrix, and T_i is the transformation matrix of observation i. The noise of observation i is represented by n_i. LR images can be considered as a down-sampled, motion corrupted, blurred, and noisy version of the HR image. The resulting reconstruction can be divided into two main parts: motion correction (estimating W_i) and super-resolution (estimating X). Image registration can be used to estimate motion, interpolation to obtain a uniformly spaced HR image, and regularized super resolution with automatic outlier rejection to remove blur and noise. Volumetric fetal MR image reconstruction is more challenging than typical image reconstruction problems due to unconstrained random motion during slice acquisitions. Slice misalignments can lead to a loss of spatial coherence and typically introduce anisotropic voxel sizes and intensity inhomogeneities.

A. Practical Limitations

SVR methods have been successfully employed to address these problems in fetal MR and are typically applied to small regions and organs with rigid body characteristics that are identified by manual annotations [16]–[19] or less precise, automated segmentations [20]–[22]. Such approaches are prohibitive to whole body and uterus reconstruction because of the assumption of rigid motion in the 2D to 3D registration step of SVR. As a result, different areas in each slice that are likely to move in different directions will break this assumption, *e.g.*, the head and thorax. Further, an extension of 2D-3D registration to include non-rigid deformations is only well defined with each slice and not well-constrained in 3D. Current SVR approaches will fail in the presence of non-rigid deformations and unpredictable organ shapes. This restricts the application of SVR to regions that are manually or automatically annotated. Thus, most of the previous SVR methods have been limited to the fetal brain as the main region of interest for fetal reconstruction due to the high incidence of neuro-developmental disability in premature infants. Only recently, [23], [24] proposed a motion corrected 3D reconstruction of fetal thoracic structures from prenatal MRI. Moreover, SVR is computationally expensive due to the exponential increase of computation with the size of the target area. This leads to prohibitive post-processing times in the clinical practice. Parallelized implementations [25] can address run-time problems, however, methodologically SVR is still restricted to small, rigid body areas.

B. Reconstruction of Large-Scale Anatomy

MRI has further been shown to be useful for the evaluation of the Whole uterus and structures like the placenta. During both normal and high-risk pregnancies, the whole uterine appearance and the condition of the placenta are considered to be an indicator for fetal health after birth [26]. Placental functions affect the birth weight as it controls the transmission of nutrients from the maternal to the fetal circulation [27]. However, the whole fetal body and secondary uterine parts can be inherently inconsistent. Different fetal body parts can move independently from the uterus. This makes the application of SVR and 2D-3D registration to the full uterus impossible in the presence of fetal motion and maternal respiration.

Besides, multiple births is a case where classical SVR pipelines, based on preprocessing steps to identify consistent rigid regions, will likely fail. The presence of multiple instances of the same fetal structure is usually not considered in previous methods. Therefore, a fully automatic motion correction method for the whole uterus, as it is presented in this paper, is very desirable and will enable the application of standard 3D image analysis techniques, *e.g.*, [28], [29].

C. Related Work

Most motion compensation approaches for fetal MRI are based on SVR techniques that aim to obtain a motion-free and

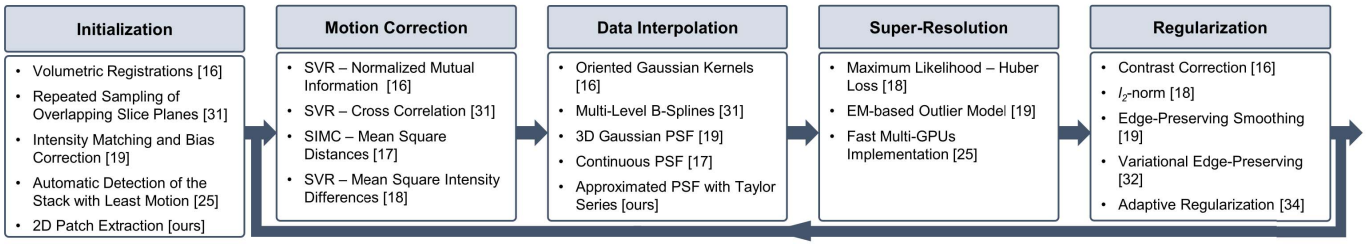


Fig. 3. Overview of the required modules of state-of-the-art SVR methods and main components introduced by previous work.

high resolution volume of a fetal target region. Registration of individual 2D slices with a higher resolution 3D volume [30] is the core approach of these algorithms. SVR methods assume that all acquired images are centered at a specific organ (*e.g.*, brain, thorax) and cover three orthogonal image directions. Fig. 3 shows the core elements of SVR and the contribution of previous frameworks from the literature.

The first SVR-based reconstruction framework for fetal MRI was introduced by Rousseau *et al.* [16]. It includes steps to correct 2D slice misalignments, intensity inhomogeneity distortions, and reconstructs an isotropic HR fetal brain from sets of LR images. Motion correction is done by applying a global 3D rigid alignment between the LR images using one image as a reference to define the global coordinate system. Then every slice is aligned to the initial reconstructed HR volume. Normalized mutual information is maximized using the gradient ascent method for both registration steps. A narrow Gaussian kernel is applied as a point spread function (PSF) for volume reconstruction and empty voxels are filled using the mean of the surrounding voxels. The image contrast is corrected using one LR image as a reference.

Jiang *et al.* [31] introduced the acquisition of many thin slices to provide sufficient sampling of the region of interest. Cross correlation is used as a cost function for the SVR steps assuming that the data have consistent contrast properties. After that, multilevel B-splines are applied to the volumetric reconstruction for data interpolation, which has the advantage of reducing blurring of the reconstructed image supported by including the thin slices.

Kim *et al.* [17] proposed a method for slice intersection motion correction (SIMC) of multi-slice MRI for 3D fetal brain image formation. The method is based on slice-to-slice registration using spatially weighted mean square intensity differences (MSD) of the signal between slices as an energy, assuming that the MRI contrasts are identical. Maternal tissues are excluded from the energy computations using a windowing function of a parametric ellipsoid model. Similar to [16], temporally adjacent slices are grouped together then divided into half iteratively. The splitting process is performed using discrete cosine basis functions.

Gholipour *et al.* [18] were the first to introduce a mathematical model for super-resolution (SR) volume reconstruction from slice acquisitions of fetal brains. The main difference to previous methods is that it includes knowledge of the slice acquisition model and the SR reconstruction is performed based on maximum likelihood and a robust M-estimation minimization for an error norm function. A Tikhonov regularization term is added to the cost function in order to enforce a solution when the number of acquired samples is not high enough for solving the reconstruction problem.

Rousseau *et al.* [32], [33] proposed to use a variational regularization including an approximation of Total Variation (TV) to better preserve edges. Later, Tourbier *et al.* [34] introduced an adaptive regularization by applying novel fast convex optimization techniques to design an efficient optimization algorithm for the super-resolution problem using edge-preserving TV regularization.

Murgasova *et al.* [19] were able to reconstruct the fetal brain using intensity matching and complete outlier removal. The main steps of their reconstruction method are: (*i*) 3D registration of the acquired stacks using a template stack; (*ii*) extracting region of interest (the fetal head) from all the stacks; (*iii*) intensity matching and bias correction between the slices based on an EM framework, where the differential bias fields and slice-dependent scaling factors are estimated during the reconstruction; (*iv*) motion correction using [16] based on the normalized cross correlation as a similarity measure and an approximated 3D Gaussian PSF similar to [31]. A posterior probability is used to define the inlier and outlier voxels within the EM framework in order to remove the motion-corrupted artifacts and misaligned data. Blurring in reconstructed images is reduced by integrating edge-preserving regularization based on anisotropic diffusion within the SR reconstruction framework.

Kainz *et al.* [25] developed a fast multi-GPU accelerated implementation for the method presented in [19], which is based on 2D-3D registration, SR with automatic outlier rejection and an optional intensity bias correction. They extended the reconstruction framework by automatically selecting the stack with least motion as the reference stack and using a fully flexible and accurate PSF instead of approximated functions. Using a multi-GPU framework enabled the SR reconstruction process to be approximately five to ten times faster than using a multi-CPU framework.

To our knowledge, modeling non-rigid transformations with multiple rigid transformations and without an initial registration target has only been preliminarily explored for fetal MRI in [35]. Other works regarding non-linear 3D-3D registration between an outlier-free, regularly sampled source and target volume, constrain parts of an image to move independently and rigidly (e.g. bones through manual segmentation and landmark annotation), while allowing other parts to deform [36].

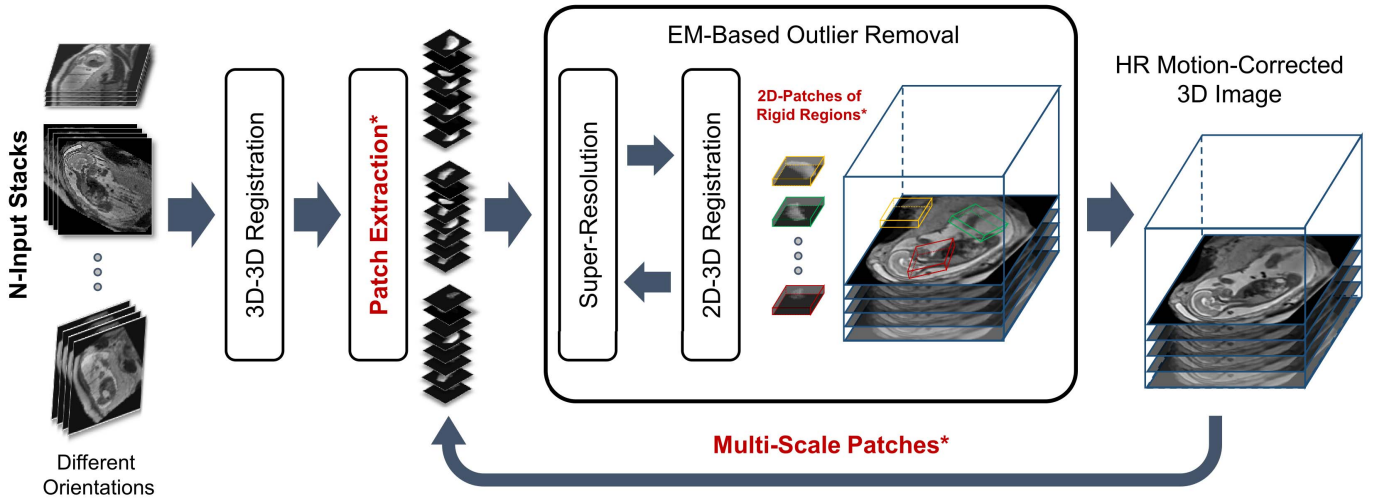


Fig. 4. A schematic and modular overview of the proposed patch-to-volume reconstruction (PVR) framework. The key parts are 3D-3D registration, patch extraction, 2D-3D registration, super-resolution, and EM-based outlier removal. Core contributions of PVR are written in red and marked with asterisk.

In 3D-3D settings globally continuous 3D deformation fields, which should not be assumed for 2D-3D fetal motion compensation problems, can be approximated by fusing multiple local rigid transformations as shown in [37] and [38].

D. Contributions

In this paper we propose and evaluate a new paradigm for motion correction based on SVR and flexible subdivision of the input space into overlapping, highly redundant and partly rigid image patches [35], thus solving the motion compensation problem for large field of view reconstructions. We split the input into small overlapping areas and find these, which contain rigid components. This allows to iteratively learn their consistency compared to a global reconstruction optimization volume. Corrupted and inconsistent data is automatically identified and excluded using robust statistics. Our approach facilitates the automatic reconstruction of whole collections of motion corrupted stacks without the need of corresponding image segmentations. By treating rigid image patches as piecewise constant segments of organs further allows limited correction of non-rigid tissue motion. We also evaluate the non-rigid deformations that may result from PVR using anatomical landmarks annotated in the fetal head and non-rigid registration. The presented patch-to-volume reconstruction (PVR) finds automatically rigidly connected areas, which can be used as segmentation prior for further refinement using conventional SVR in small regions of interest. In contrast to [25], we further introduce a multi-scale patch approach and thoroughly evaluate the reconstruction quality of the whole uterus including the fetal brain and placenta. We test the breaking points of SVR and variations of PVR on synthetically motion corrupted brain phantom data. The presented approach is the only currently available method that is able to reconstruct fetal organs and detailed 3D volumes of secondary, non-rigidly moving structures such as the placenta.

II. Method

SVR-based motion compensation methods make use of the assumption that rigid regions, *e.g.*, brain and thorax, of 2D input slices deforms rigidly, where a global 3D volume is reconstructed by iteratively registering these 2D input slices. We propose to increase the granularity of the input data by using 2D image patches of arbitrary shape instead of whole slices for SVR reconstruction. We explore square patches and dilated superpixels [39] for the definition of the patch shape. Superpixels provide a method to define semantically meaningful regions while reducing the required data redundancy and computational overhead.

PVR relies on the fact that certain regions of the scanned anatomy are rigid and can be reconstructed with SVR algorithms. However, unlike SVR, it is fully automatic and provides a full field of view reconstruction. Data consistency is obtained by oversampling a region of interest at different scan orientations. Robust statistics can be used to identify misregistered or heavily corrupted data [18], [19]. Fig. 4 depicts a schematic overview of the proposed PVR framework.

A. Input Data and Initialization

A template stack is either randomly or automatically chosen from available input stacks by detecting the stack with fewest motion artifacts [25]. Global intensity matching is applied to normalize intensity values of all input images followed by global 3D-3D alignments to spatially initialize the reconstruction target. Input data can be represented as stacks of 2D images (patches) consisting of

$$Y = \{y_s | s \in S\}, \tag{2}$$

where y_s is a patch of arbitrary 2D-shape and indexed by the location s. S is the set of all locations in all p stacks, $S = \{s_1, s_2, \ldots, s_M\}$, and M is total number of patches.

B. Patch Extraction

In the simplest, naïve case the shape of y_s is square defined via its edge length a and stride ω. This definition is generally applicable to any kind of oversampled motion corrupted data. If a and ω are fixed, no prior knowledge about the data is assumed. However, ideally each y_s corresponds

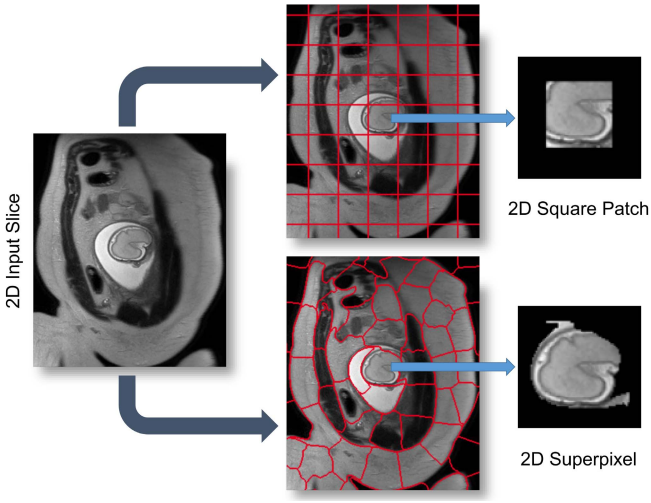


Fig. 5. An illustrative figure showing both square patches and superpixels methods for the patch extraction step. A 2D superpixel shows more flexibility than a square patch in extracting rigid regions or similar voxels. In practice, superpixels are dilated with few pixels to include some contextual information in order to increase the accuracy of the patch to volume registration step.

to a meaningful subregion of the volume in which motion can be characterized as rigid. Typically, square patches are overlapping to provide redundant representations of the same locations. The computational expense of such an approach increases with patch number and size a. Additional consideration must include the inherent trade-off between a and the assumption of it containing rigid motion.

An alternative to naïve shape definitions of y_s is to find correlation between voxel locations and their neighbors, which can be found by unsupervised image segmentation techniques such as *superpixels* (SP) [39]. These techniques allow similar-sized segments to be obtained from local intensity information (see Fig. 5) instead of employing dense sampling of overlapping patches, enabling the image reconstruction with fewer but more useful data blocks. Further, reducing the total amount of required data blocks for reconstruction lowers the computational overhead, positively impacting the overall runtime. Additionally, larger rigid areas require less computational effort for image registration and super-resolution, and more importantly less dependency on inherent image data parameters (*e.g.*, voxel spacing, organ size, subject size).

While there are several techniques for generating SP in the literature [39]–[41], a fast and efficient SP approach is desirable for the clinical practice. Simple linear iterative clustering (SLIC) [39] obtains regular SP based on minimizing the distance D between the centroids of SP with an initial step size a. D is defined as:

$$D = \sqrt{{d_c}^2 + \left(\frac{d_s}{a}\right)^2 t^2}, \tag{3}$$

where d_c and d_s are the intensity and spatial Euclidean distances that are controlled by the adaptive compactness parameter t for each SP. In our previous work [35], we use a defined number of SP centers $N_{sp} \approx k\sqrt{(n/2)}$, where n is the number of voxels in an image of size d_x, d_y and k a constant factor. Here, we use a regular grid for initialization with a constant distance a between SP centroids so that $N_{sp} = \frac{d_x}{a} + \frac{d_y}{a}$ and dilate by γ %. Thus, instead of providing a target number of SP, N_{sp} is defined by the image size and a controls the relative size of the SP.

C. Multi-Scale Patches

Although larger patch regions are less likely to include rigidly connected regions, they may perform better during 2D-3D registration due to the additional contextual information of each patch. In contrast, smaller patch sizes are more likely to represent rigidly deformed regions, but provide less contextual information, potentially affecting the 2D-3D registration. A good trade-off between the size of the patch region and the likelihood of rigid motion needs to be found. Here, we propose the use of multi-scale patches for reconstruction to exploit the advantages of different patch sizes. We represent input data as stacks of 2D patches:

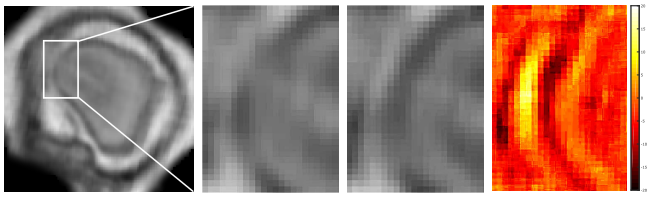
$$Y_i = \{y_s | s \in S_i\}, \tag{4}$$

where, instead of using the same Y as a unique input, a different scale of Y_i is used for each iteration i at γ % of its original size. Similar to Eq. 2, S_i is the set of all locations in all p stacks but with different size for each iteration i. This is done by re-calling the patch extraction module with a different patch size (see Fig. 4). Additionally, to increase contextual information for estimating the transformations, we compute overlapping y_s superpixel patches and dilate each by γ pixels using a flat structuring element b with a fixed neighborhood (26 pixels in our case), hence $\bar{y}_s = y_s \oplus b$. Smaller γ values result in a faster reconstruction. Ideally γ is $> 50\%$ of a, ensuring that every pixel is covered by multiple samples.

D. Patch to Volume Registration

An HR-image X is reconstructed from a number of motion corrupted patches y_s using 2D-3D registration-based super-resolution similar to [19] and [25], where an accurate PSF calculation is used to generate a gradually improving approximation of X and further employed to initialize the 2D-3D registration and computation of robust statistics. In [19] and [25], the PSF is a sinc function for the in-plane and is the slice profile for the through-plane, measured for the employed MRI sequence (ssFSE), according to [31].

We employ an implementation of the PSF function by applying a Taylor series for a better approximation of small values close to 0. We truncate the series after several terms and bound the remainder based on relative error ϵ. The Taylor series approximation of the sinc function is defined as $sinc(R) = 1 - \frac{R^2}{3!} + \frac{R^4}{5!} - \frac{R^6}{7!} + \cdots$. The proposed approximate PSF achieves qualitative improvement of the reconstructed image compared to the sinc implementation. An example from the first iteration of a fetal brain reconstruction is shown in Fig. 6. This is expected since Taylor series approximation is common for computations requiring a high level of accuracy close to the machine precision limits of floating point computations [42].



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

Fig. 6. Example for the observed differences in the first iteration of a fetal brain MRI reconstruction (a). (b) shows a magnified region using a sinc function for the PSF similar to [35] and (c) shows the result from using a Taylor series approximation of the sinc function as used in this work. Taylor series approximation allows a better approximating of small values close to zero. (d) shows the difference between both images.

During the optimization process, individual 2D patches are continuously rigidly registered to the current 3D reconstruction of X and reintegrated into X using iterative super-resolution with gradient descent optimization. Any similarity metric can be used as a cost function for the registration step such as mutual information [16], [19], cross correlation (CC) [25], [31], or mean square intensity differences [17], [18]. Choosing the best similarity metric for reconstruction depends on the input data. CC has been found to be effective for input data with similar intensity distribution [43]. In our experiments, we employ CC as the similarity metric for 2D-3D registration, after rescaling the intensities between the input stacks.

E. Super-Resolution

Given that the position and orientation of each patch y_s relative to X is known at any point, patches y_s^* can be simulated from sampling the current approximation of X using the PSF. X can subsequently be iteratively improved by minimizing the error between y_s and y_s^*. Gradient decent is used for the optimization. The optimization has to be regularized to prevent amplification of registration error, noise and under-sampling artifacts. Adaptive regularization is applied to reduce smoothing effects. Therefore the objective function is

$$\sum_i PSF(y_{s_i} - y_{s_i}^*)^2 + \lambda \sum_i \sum_d \phi \left(\frac{x_{i+d} - x_i}{\delta |d|} \right), \quad (5)$$

with $\phi(t) = 2\sqrt{1+t^2} - 2$, voxel locations $x \in X$, d a vector between the index of a voxel and one of its 26 neighbors, and δ a parameter that controls when the gradient between neighboring voxels is considered to be an edge. For our data a gradient magnitude between 100 and 400 yields good results. The regularization term represents edge preserving filtering [44] and is used similarly as presented in [25]. λ is decreased after each iteration. Following the experimental findings in [19] a suitable starting value for λ is $0.8\delta^2$, halved after every iteration.

F. EM-Based Outlier Removal

Correctly registered patches $\hat{y}_s$ should provide a higher contribution to the final reconstruction, presenting a low error e when compared to the original image data. [18] initially introduced an approach to account for outliers during super-resolution based on Huber function statistics. Similar to [19], we employ the expectation maximization algorithm for outlier removal by classifying $\hat{y}_s$ and the included pixels into an inlier and outlier class. A zero-mean Gaussian distribution $G_\sigma(e)$ with variance σ^2 is used for the inliers and a uniform distribution with constant density

$$m = \frac{1}{max(e) - min(e)} \quad (6)$$

for the outliers. This makes use of available, highly redundant information (*i.e.*, overlapping $\hat{y}_s$), to find partly matching patches and to depreciate or fully reject erroneous voxels. We aim to maximize the log-likelihood for each patch

$$y_s | log\, P(Y, \Phi) = \sum log\, P(e|\sigma, c) \quad (7)$$

to be part of a region of rigid motion. Φ is the current estimate of the reconstructed volume X, the variance σ^2 of the errors e, and the proportion of correctly matched voxels c. The posterior probability for a pixel $\in \hat{y}_s$ being identified as inlier is

$$p = \frac{G_\sigma(e)c}{G_\sigma(e)c + m(1-c)}. \quad (8)$$

We perform the updates of c and σ^2 according to [19]:

$$\hat{p} = \sqrt{(\sum_{\hat{y}_s} p^2)/N}, \quad (9)$$

where N is the number of pixels in $\hat{y}_s$. We further define an inlier and outlier probability for each $\hat{y}_s$ and exclude it from processing if classified as an outlier (*e.g.*, if it contains structures moving in opposite directions during scanning, such as the fetal head and thorax). Only if information in $\hat{y}_s$ is consistent with the originally acquired data, will the registered patch contribute to the SR reconstruction of X.

G. Identification of Rigid Regions and SVR Refinement

The rigidity of regions is measured by keeping track of the probability p of each pixel of every $\hat{y}_s$. This allows the identification of locations best fitting the rigid 2D-3D registration constraints. Candidate regions, that contain rigid motion components, can be identified by combining p and $\hat{p}$ into a 3D volume utilizing the same PSF as used in the reconstruction. This can further initialize the rigid SVR reconstruction and visualize the data uncertainty during reconstruction.

III. Implementation

A. Parallelization

The high data redundancy required for the proposed approach makes conventional single threaded implementation practically not feasible. Computational complexity of PVR is exponentially higher than SVR, depending on the employed patch overlap. For optimal performance we implemented our approach via General-Purpose Programming on Graphics Processing Units (GPGPU) using the Compute Unified Device Architecture (CUDA, NVIDIA, Santa Clara, CA) language [45], [46]. CUDA is a highly evolved single instruction multiple data (SIMD) programming language, which allows a large part of the proposed framework to be mapped onto GPU hardware. Currently, CUDA is the only high-level

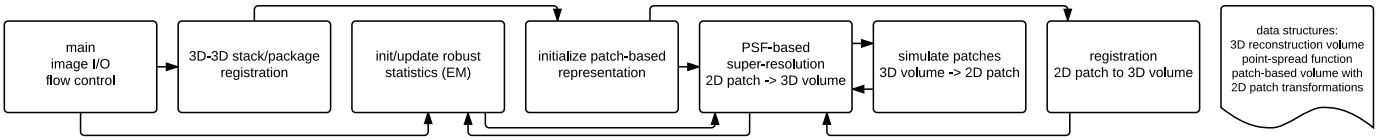


Fig. 7. The software modules defined for the implementation of the proposed approach. For implementation details, please refer to the provided source code.

general purpose GPU language that provides bi-directional texture access via surfaces in a kernel, which is essential for the efficient implementation of certain parts of our framework. In this section we discuss the key implementation details.

We use a modular design to allow experimentation with the separate components of the algorithm. An overview of this design is shown in Fig. 7. The modules are encapsulated in a CUDA library, which can be used independently from the instantiating framework. We employ the successor of IRTK[1] for interfacing with medical image data.

PVR is parallelized on three levels:

I. *Patch-level:* Individual patches are mapped to blocks of a CUDA computing grid and the contained voxels are mapped to individual threads. Depending on the used GPU hardware, patch processing can also be mapped directly to the computing grid, such that each thread works on a complete patch (limited by the employed patch size). The resulting thread divergence provides opportunities for advanced GPU scheduling strategies [47] and for a direct translation of optimization strategies for image registration, for example patch-wise gradient descent.

II. *Voxel-level:* For the parallelization of PSF-based super-resolution and robust statistics we follow a similar three-fold procedure definition as used in [25]. The voxels within each patch are processed using kernel level parallelization and parallel pixel-volume, volume-pixel, and volume-volume procedures are applied.

III. *Patch-batch:* PVR scales to multiple GPUs through distributing independent subsets of patches over the desired number of devices. Synchronization is done through averaging of the resulting sub reconstruction volumes on the master GPU. Initial 3D-3D registration is performed on a single master GPU, which allows optimal coalesced memory access.

B. Availability of Source Code

We provide the source code of a C++/CUDA implementation of the proposed method, including parallelization strategies, in a publicly available software repository.[2] The source code for the implementation of PVR is licensed under MIT license.

IV. Evaluation & Experiments

A. Evaluation of Global Image Quality

1) Evaluation of Adult Brain MRI Reconstruction: We evaluate the performance and limitations of PVR in terms of accuracy and robustness with synthetic deformations of *adult* brain data. Similar to [18], an isotropic 1 mm^3 T2-weighted adult brain phantom with no noise obtained from the *Brainweb* database [48] is used for this experiment. Three experiments with different deformations are performed, namely:

I. *Rigid Transformation:* Synthetic rigid motion artifacts are generated by translating the axial, sagittal and coronal images with $\pm d_x mm$ in the x-direction.

II. *Bulk Transformation:* Synthetic bulk motion artifacts are generated by rotating the upper half of the original image with $\pm\theta_z^\circ$ around the z-axis.

III. *Non-rigid deformation:* Synthetic non-rigid motion artifacts are generated by skewing (shearing) the axial, sagittal, coronal images using:

$$T_{S_{xyz}} = \begin{bmatrix} 1 & S_{xy} & S_{xz} & 0 \\ S_{yx} & 1 & S_{yz} & 0 \\ S_{zx} & S_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where we use one combined skewing value $S_{xyz} = tan(\pm\theta^\circ_{xyz})$ in the xyz-direction.

Following image deformation, a motion-corrupted 3D stack is constructed by sampling 2D images from both motion-free and motion-corrupted stacks in an interleaved manner similar to fetal MRI acquisition [7]. Where for the bulk motion experiment a 3D stack is constructed by only rotating the upper half of the brain. Three stacks are used for the reconstructions where each stack is sampled with a voxel size of $1.25 \times 1.25 \times 2.5\ mm^3$. We use standard axial, sagittal, and coronal orientations as shown in Fig. 8. An HR image with isotropic voxel size $1.25\ mm^3$ is reconstructed using SVR [25], square patch- and superpixel-based PVR.

2) Evaluation of Fetal Organ MRI Reconstructions: Evaluating the quality of reconstructed fetal MRI is challenging due to the absence of motion-free ground truth data. For this purpose, we introduce a novel approach for the evaluation problem based on the originally acquired slice images. Assuming that 2D in-plane patches extracted from the original stacks contain no motion artifacts, we use them as gold standard and compare with corresponding simulated patches from the reconstructed volume. Evaluation metrics (see Sec. IV-A.3) are computed between the reconstructed input stacks and the final motion corrected image and averaged over the whole volume. The fetal brain is typically used to assess the quality of reconstruction as it moves rigidly, fulfilling the rigid motion assumption for SVR-based methods in the 2D-3D registration step. However, soft tissue organs such as the placenta deform non-rigidly. For this reason, we additionally chose to reconstruct the placenta and whole uterus as challenging test cases for PVR and SVR.

[1] https://github.com/BioMedIA/MIRTK
[2] https://github.com/bkainz/fetalReconstruction

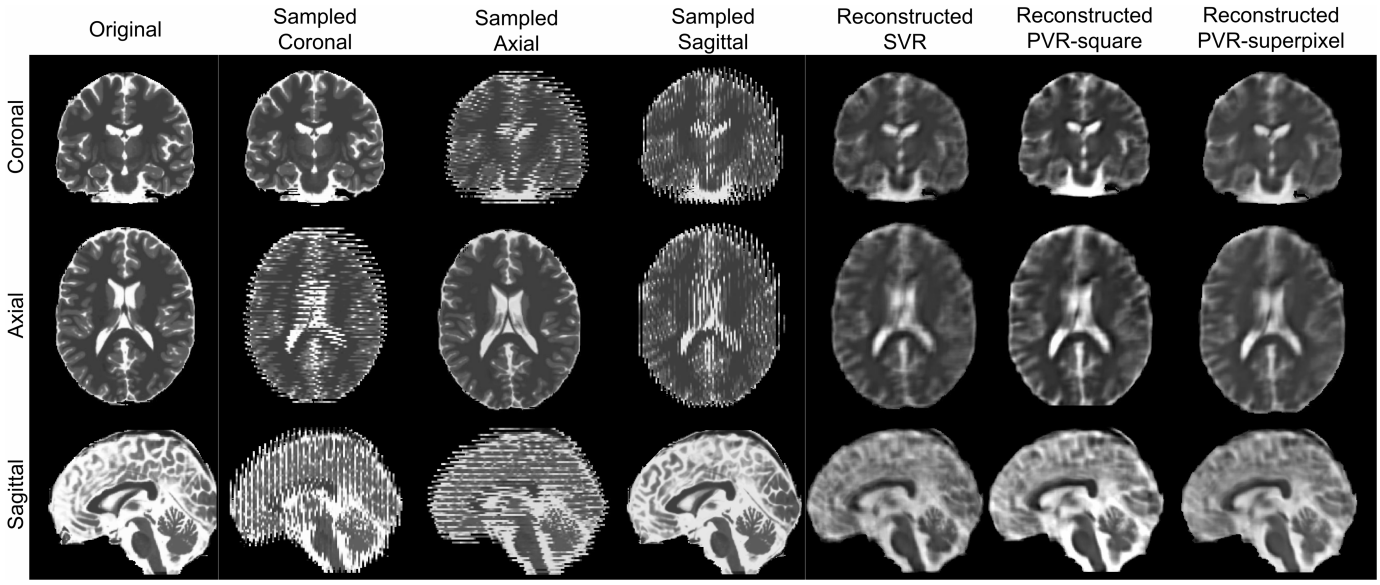


Fig. 8. Strong synthetic non-rigid motion artifacts caused by skewing an adult brain phantom with an angle of ($\theta^{\circ}_{xyz} = \pm 1$). Rows: MRI in standard orientations: coronal, axial, and sagittal. Columns: original scan ($1 \times 1 \times 1$ mm^3), and sampled ($1.25 \times 1.25 \times 2.5$ mm^3) axially, coronally and sagittally, and SVR, PVR-square patches and PVR-superpixel reconstructed isotropic images ($1.25 \times 1.25 \times 1.25$ mm^3).

3) Evaluation Metrics: We employ the following metrics for measuring the quality of the reconstructed image: Cross-correlation (CC) to measure the similarity between the intensities of input $I(i, j)$ and reconstructed image $\tilde{I}(i, j)$ at the location (i, j), which is defined as:

$$CC = \frac{1}{N \times M} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{(I(i,j) - I_{\mu})(\tilde{I}(i,j) - \tilde{I}_{\mu})}{\sigma_I \sigma_{\tilde{I}}} \quad (10)$$

where N and M are the dimensions of a 2D slice.

The peak signal-to-noise ratio (PSNR) is used to measure the error introduced by motion and is based on the mean squared error (MSE) between the original 2D in-plane patch and the reconstructed image. PSNR is defined as:

$$PSNR = 10 \log \frac{I_{max}^2}{MSE} \quad (11)$$

where I_{max} is the maximum intensity in the original image. An improved reconstruction quality usually results in higher PSNR. However, PSNR does not reflect well subjective human perception of image quality as it is mainly based on estimating absolute errors between individual pixels.

The structural similarity index (SSIM) accounts for image degradation as perceived changes in structural information [49]. It measures the structural similarity by comparing normalized local patterns of pixel intensities, which is similar to the human visual system's abilities to extract information based on structure. The SSIM is defined as:

$$SSIM = \frac{(2\mu_I \mu_{\tilde{I}} + c_1)(2\sigma_{I\tilde{I}} + c_2)}{(\mu_I^2 \mu_{\tilde{I}}^2 + c_1)(\sigma_I^2 + \sigma_{\tilde{I}}^2 + c_2)} \quad (12)$$

where μ_I, $\mu_{\tilde{I}}$, σ_I^2 and $\sigma_{\tilde{I}}^2$ are the average and variance of the intensities of the original 2D in-plane slice and the reconstructed slice respectively. $\sigma_{I\tilde{I}}$ is the covariance of I and $\tilde{I}$. c_1 and c_2 are defined as $(k_1 L)^2$ and $(k_2 L)^2$ in order to balance the division with weak denominator, where L is the dynamic range of the intensities in image I. Similar to [49], k_1 and k_2 are equal to 0.01 and 0.03 respectively.

Structural dissimilarity (DSSIM) heat maps are calculated in order to visualize the dissimilarities between original and reconstructed images. DSSIM is calculated as a distance metric derived from SSIM:

$$DSSIM = \frac{(1 - SSIM)}{2} \quad (13)$$

B. Evaluation of Local Deformations

Compared to SVR, PVR relaxes rigidity constraints and may therefore introduce deformation artifacts like shrinking of anatomical structures. Hence, in order to assess the geometric integrity of the 3D image reconstruction in terms of *mm* we conduct two experiments:

1) Evaluation With Anatomical Landmarks: We use a set of 3D anatomical landmarks that can be clearly defined in 3D reconstructions of real fetal MRI head scans to assess the accuracy of nonrigid deformations resulting from PVR compared to SVR. In particular, we have selected 10 clinically relevant, anatomical landmarks in the eyes and sub-cortical region of the fetal brain, which define stable anatomical structures of the corpus callosum and the lateral ventricles. A clinical expert has chosen a set of 11 fetal MRI scans with the best reconstruction quality of the fetal head using SVR. We have repeated the reconstruction of this set from the raw scan data using PVR with patches of size 32 and stride 16 pixels. Two different observers have annotated the selected anatomical landmarks independently across the 11 subjects. The accuracy is assessed by calculating the average 3D Euclidean distance errors between PVR and SVR (mean annotation points of the two observers), measured for each landmark for all 11 subjects.

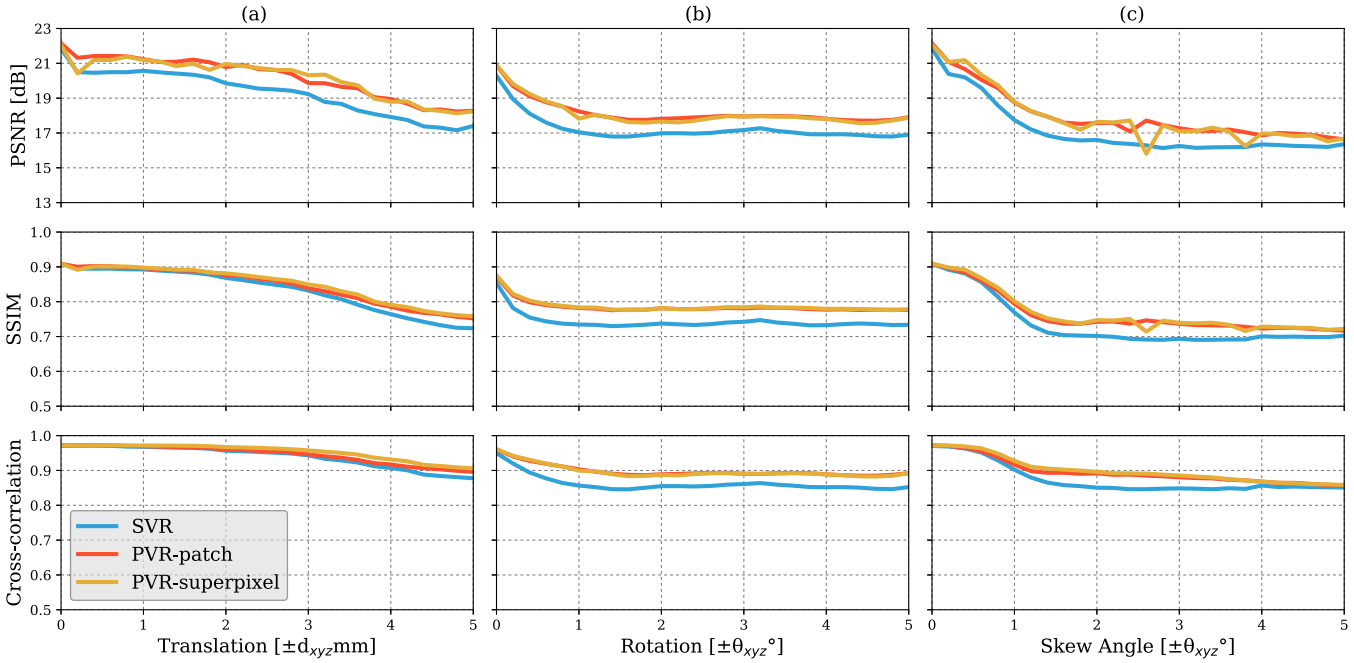


Fig. 9. Comparative reconstruction performance of SVR and PVR methods on synthetically corrupted Brainweb [48] data using rigid translational transformations (1st column), bulk transformations (2nd column) and non-rigid skewing deformations (3rd column). Left to right: PSNR, SSIM and CC over skew angle in degrees and translations in mm for SVR (blue), superpixel-based PVR ($a = 16, \gamma = 60\%$, yellow) and PVR using square patches ($a = 32, \omega = 16$, red).

2) Evaluation of Displacement: In this experiment we seek to evaluate the quality of PVR in cases of non-rigid motion between the scanned images using an approximated displacement error measured on the simulated skewing motion experiment of the brain phantom. The reconstructed images are first rigidly and then non-rigidly (b-spline [50]) registered to the motion-free ground truth image. The transformations resulting from non-rigid registration are used to transform a set of points sampled uniformly on a 3D grid (5% of the image size) inside the brain area. To ensure a fair comparison, we exclude reconstructed images from high motion skewing angles $> 1.5°$, which may affect registration to the reference image if the reconstructed image is blurred or distorted.

V. Results

A. Global Image Quality

1) Reconstruction of Adult Brain MRI: Experiments on adult brain MR data using the Brainweb database [48] includes introducing synthetic non-rigid motion artifacts as described in Sec. IV-A.1. Example slices of standard planes of original and corrupted data are depicted in Fig. 8. Comparative experimental results of SVR and PVR reconstruction methods are shown in Fig. 9 for PSNR, SSIM and CC. For all metrics, PVR shows an improved performance over SVR, particularly in the presence of deformations with higher skewing angles. Further, we observe that superpixel-based PVR achieves performance similar to PVR using arbitrary square patches, while requiring a lower number of input patches.

2) Reconstruction of Fetal Organs: Exemplary PVR and SVR reconstructions under motion introduced by kicking of the fetus are shown in Fig. 10. PVR reconstruction results show an improved visual appearance and less blurring in the region with severe motion artifacts (arrow).

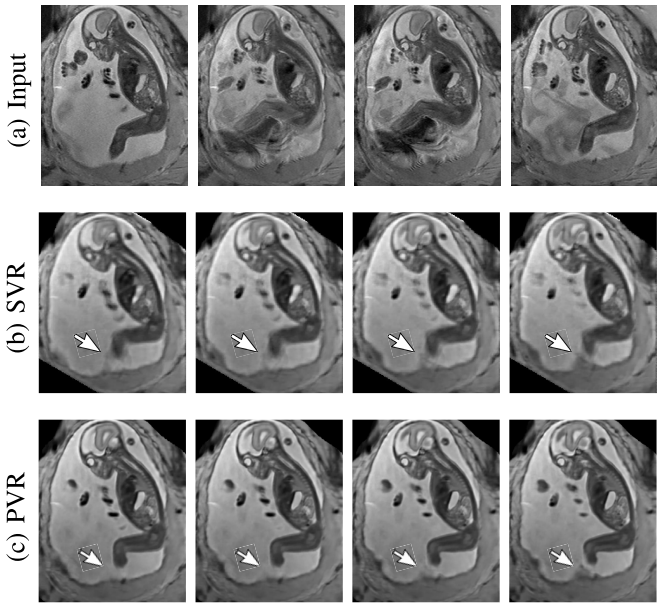


Fig. 10. Example reconstructions of consecutive MR scans of a moving fetus (kicking): input data (a) and corresponding cutting planes through an SVR- (b) and PVR-reconstructed (c) volumes. SVR produces blurry but readable results because of high data redundancy and outlier rejection through robust statistics. PVR with square patches of $a = 32$ and $\omega = 16$ appears visually superior. The arrow points at an area of substantial quality differences caused by independent rapid movements of the leg.

An example of a challenging clinical case with a kidney malformation in one of twin fetuses, is shown in Fig. 11. Our clinical partners confirmed that such complications are easier to examine and to quantify after PVR-based reconstruction.

TABLE I

AVERAGE (A) PSNR AND (B) SSIM RESULTS ($N = 32$) FOR THE INPUT STACK (BASELINE) AND PVR VARIANTS WITH FIXED (FS) AND MULTI-SCALE (MS) VARIANTS OF SQUARE PATCHES AND SUPERPIXELS. ALL MEAN DIFFERENCES OF PVR AGAINST BASELINE ARE STATISTICALLY SIGNIFICANT ($p < 0.05$). NAMES OF ALL STATISTICALLY SIGNIFICANTLY DIFFERENT PVR VARIANTS ARE STATED IN **BOLD**

(a)

Baseline	*Brain* 16.97 ± 3.77		*Placenta* 19.95 ± 4.40		*Uterus* 19.29 ± 3.82	
	Square Patches	Superpixels	**Square Patches**	**Superpixels**	Square Patches	Superpixels
FS	26.70 ± 1.45	26.60 ± 1.58	31.07 ± 1.35	31.00 ± 1.46	27.17 ± 1.43	27.35 ± 1.36
MS	27.03 ± 1.38	26.35 ± 1.54	30.85 ± 1.50	29.62 ± 1.30	26.87 ± 1.31	26.40 ± 1.24

(b)

Baseline	*Brain* 0.00 ± 0.01		*Placenta* 0.00 ± 0.02		*Uterus* 0.01 ± 0.03	
	Square Patches	**Superpixels**	**Square Patches**	**Superpixels**	**Square Patches**	**Superpixels**
FS	0.51 ± 0.03	0.51 ± 0.03	0.58 ± 0.03	0.58 ± 0.03	0.53 ± 0.03	0.50 ± 0.03
MS	0.48 ± 0.04	0.45 ± 0.05	0.54 ± 0.04	0.49 ± 0.05	0.50 ± 0.03	0.47 ± 0.04

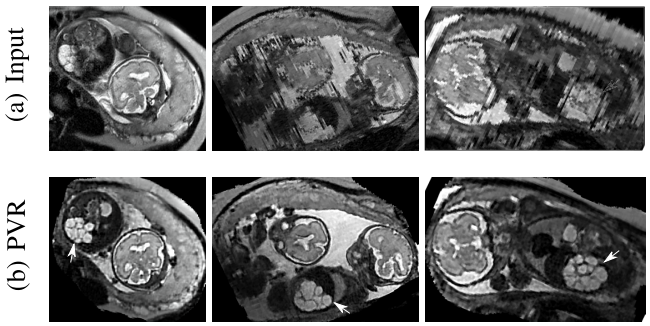


Fig. 11. Three viewing planes through the original motion corrupted scan of a moving twin with a gestational age of 28 weeks (a) and PVR reconstruction using multi-scale superpixels (b). For this dataset we masked the uterus to save unnecessary computation time in areas containing maternal tissue. The white arrow points at a unilateral multicystic kidney of one of the twins.

Comparative experiments of PVR variants were carried out on 32 fetal MR scans at gestational ages of approximately 20 weeks with voxel size around $1.00 \times 1.00 \times 2.50\ mm^3$ and $1.25 \times 1.25 \times 1.25\ mm^3$, constituting challenging image corruption samples. The ssFSE T2-weighted sequences data have been acquired on a Philips Achieva 1.5T, the mother lying 20° tilt on the left side to avoid pressure on the inferior vena cava and aligned to the main axes of the fetus. Table I (a) & (b) show numerical results of evaluating individual stacks before reconstruction (baseline), and the final reconstructed image from 3-6 stacks using square patches ($a = 32$ and $\omega = 16$), superpixels ($a = 16, \gamma = 60\%$) and multi-scale superpixels (with initial $a = 16, \gamma = 60\%$ and $\times 1.5$ scaling factor for every new iteration). Statistical testing between baseline and PVR variants was carried out using paired T-Tests and differences between using fixed or multi-scale and between square patches or superpixels were assessed via Two-factor ANOVA with repeated measures. In Table I (a) & (b) the names of PVR variants are marked in bold if statistically significant differences have been found during analysis, *i.e.*, **FS** and **MS** and/or **Square Patches** and **Superpixel** pairs are bold if the results between them differ significantly.

The evaluation of the reconstruction quality of a whole 3D image into a single-valued metric may not properly reflect the performance differences, as it is based on averaging values of all the pixels of all the input stacks. Furthermore, Table I indicates significant differences between variants of PVR but these differences have only minimal qualitative effect on reconstruction accuracy. Therefore, Fig. 12 evaluates the reconstruction quality of PVR additionally using dissimilarity heat maps based on the measured DSSIM (see Sec. IV-A.3). This approach allows further qualitative evaluation and allows for uncertainty visualization of PVR reconstructions.

3) Fetal Brain Reconstruction at Late Gestation: Examination of the brain is one of the primary uses for fetal MRI. While PVR's primary aim is to reconstruct large fields of view and non-rigid motion, we also compare how PVR performs under the conditions under which SVR is usually applied (accurate delineation of a brain mask followed by rigid full slice to volume registration). This experiment evaluates if PVR can efficiently substitute SVR in pipelines where only the reconstruction of a single, masked rigidly moving organ is required, *i.e.*, for compensation of pure bulk motion.

To compare PVR's performance, we collected a second dataset comprising of ten mature fetuses (GA $>$ 33 weeks) with little motion, expert brain segmentation, and validated state-of-the-art SVR reconstructions [25]. PVR is applied only in the delineated brain area and compared to clinical SVR reconstructions in terms of PSNR, SSIM, CC and computational overhead. In contrast to [35], where only a small area of the brain was evaluated, we evaluate the full brain mask in the current work. We chose patch parameters similar to [35] ($a = 32$, $\omega = 16$), which performed best for square patches according to the evaluation in [35] and used the multi-scale superpixel approach presented in this paper. Figure 14 shows that PVR provides similar reconstruction and motion correction quality for the brain as would be the case if a tight expert mask [20] for a region of interest would have been used for SVR. However, the computational overhead of processing patches is always larger than if SVR is used. In this experiment SVR is considered to be the baseline processing 100% of the available pixels. The required runtime of any PVR variant will always be longer than if SVR is used for bulk motion compensation of a defined region of interest.

4) Multi-Scale Reconstruction: We evaluate different scale parameters for the multi-scale superpixel reconstruction to further assess the influence of the number of scales on the reconstruction quality. Similar to the phantom brain experiment, we reconstruct a brain image from 3-stacks deformed with skewing angle $\theta_{xyz} = \pm 1°$. We use the ground truth image for evaluating quantitatively each reconstructed image with five different scales as shown in Table II. Multi-scale (MS) reconstruction is configured to increase the scale of the patch by a factor of $\times 1.5$, four times starting from $a = 16$ to 40 (2-iteration per each scale). Fixed-scale (FS) reconstruction is configured to fix the size of the patch for every reconstruction

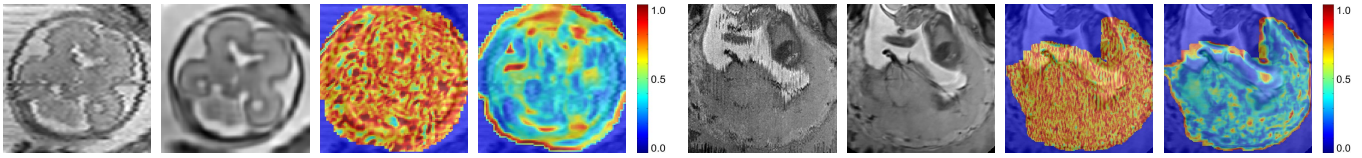


Fig. 12. A sample 2D cutting plane through a motion-corrupted fetal brain (a) and placenta (f), after PVR using square patches with $a = 32$ and $\omega = 16$ (b) and (g). The DSSIM heat map for a baseline before reconstruction (c) and (h), and after PVR (d) and (i). The average DSSIM of the fetal brain equals 0.497 (c) and 0.248 (d), while for the placenta equals to 0.491 (h) and 0.214 in (i).

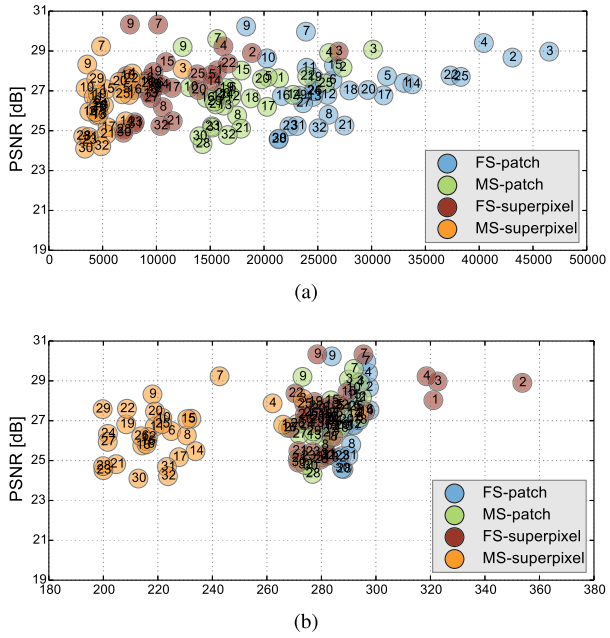


Fig. 13. (a) Number of generated patches and (b) necessary additional overhead pixels (%) of the different PVR variants versus their reconstruction PSNR quality of the whole uterus (see Table I-a). Optimal results are found in the upper left corner of the plots, *i.e.*, high reconstruction quality and low computational overhead. The subject number is highlighted inside each circle marker. Multi-scale superpixels (MS-superpixel) achieve similar reconstruction quality to fixed-size (FS-patch), multi-scale (MS-patch) square patches while clustering in the area of minimal computational overhead.

TABLE II
COMPARISON BETWEEN FIXED-SCALE (FS) AND MULTI-SCALE (MS) SUPERPIXELS RECONSTRUCTION USING 3 PHANTOM BRAIN INPUT IMAGES DEFORMED WITH SKEWING ANGLE $\theta_{xyz} = \pm 1°$

	FS-16	FS-24	FS-32	FS-40	FS-48	MS-16
PSNR	17.1	17.3	17.6	**17.9**	17.2	17.3
SSIM	0.441	**0.473**	0.464	0.435	0.391	0.420
CC	0.682	0.691	0.715	**0.721**	0.659	0.678

iteration. Figure 13 and Table II show that MS-PVR achieves similar performance as FS-PVR, while larger scale patches result in a lower number of input patches and overhead pixels.

B. Local Deformations

1) Landmarks: The results for each landmark are summarized in Table III. The inter-observer error is 0.92 ± 0.41 *mm*. The average distance between all landmarks in PVR reconstructions and the mean landmark location in SVR reconstructions between the two observers is 1.13 ± 0.87 *mm*. In order

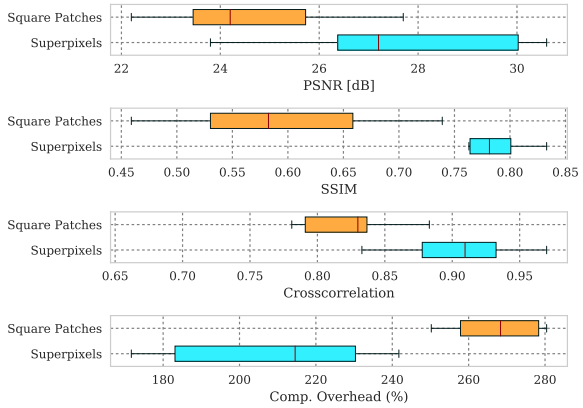


Fig. 14. Comparison of the best performing PVR parameters from [35] for square patches and multi-scale superpixels with SVR for motion compensation of the brain in mature (GA $>$ 33 weeks) fetuses with little motion. PVR's image quality is similar to the small-area SVR method currently used in the clinical practice. The necessary patch overlap for PVR requires more computation, which results generally in longer runtime for any PVR variant.

TABLE III
MOTION IN TERMS OF THE DISTANCE THE 3D LANDMARKS HAVE BEEN ESTIMATED (SVR) TO HAVE BEEN DISPLACED DURING IMAGING, AND THE DISPLACEMENT OF THESE LANDMARKS BETWEEN PVR AND CLINICALLY VALIDATED SVR RECONSTRUCTIONS

Landmarks	Slice-to-volume motion [*mm*]	Inter-observer error [*mm*]	PVR-SVR distance [*mm*]	*p*-value
Left center of the eye	15.10 ± 15.90	0.95 ± 0.29	1.00 ± 0.43	0.7983
Right center of the eye	11.87 ± 11.01	0.88 ± 0.40	0.96 ± 0.41	0.6381
Genu of corpus callosum outer aspect	13.86 ± 15.16	0.69 ± 0.36	0.83 ± 0.24	0.7670
Splenium of corpus callosum outer aspect	13.02 ± 15.06	1.08 ± 0.56	1.11 ± 0.75	0.8901
Left Frontal horn of lateral ventricle	14.93 ± 16.83	1.03 ± 0.40	1.35 ± 1.00	0.3296
Right Frontal horn of lateral ventricle	17.04 ± 18.83	0.90 ± 0.22	1.25 ± 0.89	0.2511
Left occipital horn of lateral ventricle	13.40 ± 14.59	1.04 ± 0.34	0.98 ± 0.48	0.7483
Right occipital horn of lateral ventricle	11.41 ± 13.20	1.05 ± 0.38	1.05 ± 0.56	0.9906
Left temporal horn of lateral ventricle	12.89 ± 15.88	0.69 ± 0.36	1.28 ± 1.43	0.2278
Right temporal horn of lateral ventricle	11.83 ± 12.76	0.80 ± 0.46	1.47 ± 1.33	0.1396
All landmarks	13.58 ± 15.23	0.92 ± 0.41	1.13 ± 0.87	0.0166

to provide an indication of the scale of rigid motion being addressed in the fetal brain landmark images, the estimated displacement distance of all 3D landmarks during imaging is measured by calculating the distance between each annotated landmark and its inverse-transformed point from the correspondent slice of an input stack, assuming rigid slice motion for SVR. The maximum of the calculated displacements equals to 79.85 *mm*. The total average equals to 13.58 ± 15.23 *mm* for all landmarks, see Table III. The p-values from a paired T-test show that the difference between the error introduced between observers and PVR can be considered to be not statistically significant ($p = 0.0166$). Overall, this evaluation demonstrates that PVR does not introduce notable structural distortion compared to rigid SVR.

TABLE IV
MEAN AND MAXIMUM DISPLACEMENT ERROR IN *mm* RESULTING FROM NON-RIGID REGISTRATION BETWEEN DIFFERENT RECONSTRUCTIONS AND THE MOTION-FREE ADULT BRAIN PHANTOM GROUND TRUTH

Displacement	SVR	PVR-Superpixel 16	PVR-Patch 64	PVR-Patch 32	PVR-Patch 16
Mean	1.83 ± 1.55	1.70 ± 1.10	1.58 ± 1.03	1.67 ± 1.07	2.40 ± 1.71
Max	4.91 ± 4.26	4.86 ± 3.54	4.87 ± 3.92	5.02 ± 3.92	5.72 ± 4.09

2) Displacement: The mean and maximum displacement of 9 reconstructed images are calculated for SVR and PVR with different parameters. The results in Table IV shows that PVR with patch size of 64 achieves the smallest mean error, while this error increases with the decrease of the patch size.

C. Performance Analysis

We further evaluate the computational performance of each PVR variant. Measuring the overall runtime is not meaningful because this would be highly machine specific and would include data transfer overhead and non optimized functions. The runtime varied between 2000–4000s on our testing machines, depending on the system configuration. Instead we analyze the computational overhead introduced by PVR in comparison to SVR. The overhead can be measured by counting the number of processed patches and the number of additionally processed voxels. We compare these values to the achieved reconstruction quality in Fig. 13. Multi-scale superpixels show significantly better performance than other PVR variants and introduce only minimal necessary overhead while gaining similar image quality to more naïve PVR variants. Multi-scale superpixels are potentially five times faster than other variants.

VI. DISCUSSION

PVR surpasses reconstruction performance of the state-of-the-art SVR method for cases with considerable non-rigid deformations and inconsistent bulk motion. We have evaluated different variants of PVR using fixed-size and multi-size square patches and fixed-size and multi-size superpixels. ANOVA analysis has shown significant differences between these approaches for different areas of the uterus. However, evaluation of motion compensation methods is difficult especially due to the lack of ground truth in fetal MRI. Mapping the reconstruction quality of a whole 3D volume into a single-valued metric may not properly reflect qualitative differences, if based on averaging all measured values of all input stacks. Therefore, we have additionally performed extensive qualitative analyses and present examples and evaluation based on structural dissimilarity (DSSIM) heat maps.

While an SVR reconstruction is constrained by using whole orthogonal slices assuming rigidity, PVR may introduce geometric deformations due to the relaxation of the rigidity assumption to smaller patches instead of whole slices. For this reason, we evaluated local deformations by selecting anatomical landmarks from the fetal head and compare PVR with SVR, which shows that the difference is not statistically significant for individual landmarks. We also employ dense, non-rigid registration to further validate the geometric accuracy of the reconstruction. Reconstructed images from SVR and PVR with different variants are non-rigidly aligned to a motion-free brain phantom. The mean and maximum displacements of the deformations are then calculated to measure the accuracy. Both experiments evaluate local deformations, although the latter may include errors introduced by the imperfect non-rigid registration process. Maximum displacement errors are mainly observed in feature-poor regions like the amniotic fluid, where automatic non-rigid registration does not perform well. Despite showing strong evidence that PVR works well for the whole uterus, it should be noted that our experiments to assess the reconstruction quality on brain anatomy do not necessarily extend to the achieved reconstruction accuracy of other areas (e.g. moving limbs). Focused validation of individual regions will be required for other, domain specific applications.

In addition to reconstruction and motion correction of the whole uterus, we have demonstrated that the proposed method works for multiple births cases with multiple fetal instances present in the volumetric image. These cases are more likely to have complications and to undergo MRI during pregnancy but would require extensive manual effort to be successfully reconstructed with state-of-the-art methods.

Although our method is able to reconstruct the whole uterus automatically, small parts like limbs that move rapidly between the acquisition of individual slices are more difficult to recover. This is especially problematic for very young fetuses that have more space to move inside the womb. In cases of rapid limb movements (>2 cm between individual slices) PVR is not able to find structural consensus between overlapping patches and blurry image regions will be reconstructed. This is a general problem of automatic intensity-based optimization methods and methods that are able to understand the semantic content of each patch will likely be required for future improvements.

PVR introduces a considerable computational overhead to the reconstruction stage of fetal MR image processing pipelines. We have evaluated the amount of necessary additional redundant information to give a general idea about the expected runtime of different PVR variants. Patches based on multi-scale superpixels are significantly more efficient than a naïve implementation of overlapping square patches, while maintaining a similar reconstruction accuracy. Quantitatively, square patches perform slightly better for the brain, which is most likely due to the rigid nature of the enclosing skull. Superpixel-based patches achieve better results for regions that are likely affected by non-rigid movements like the placenta and the whole uterus.

VII. CONCLUSION

We have introduced and evaluated the concept of patch-to-volume reconstruction (PVR) in order to compensate non-rigid motion artifacts from fetal MRI scans without requiring a defined region of interest. PVR splits the 3D input image into overlapping square patches and superpixels and employs automatic EM-based outlier rejection to find consistent data.

Our method is able to automatically reconstruct whole collections of motion corrupted stacks without the need for image segmentation and manual identification of rigid regions. We have shown that PVR can reconstruct the whole uterus, selected fetal organs, and secondary, non-rigidly moving pregnancy structures such as the placenta.

PVR's reconstruction quality and computational performance has been evaluated quantitatively and qualitatively on an adult phantom T2-weighted brain with synthetic non-rigid motion artifacts, as well as on the whole uterus from motion corrupted fetal MRI data including fetal brain, placenta and cases with multiple births.

ACKNOWLEDGMENT

The authors thank the volunteers, the radiographers J. Allsop and M. Fox, MITK [51], IRTK (CC public license from IXICO Ltd.). Data access only in line with the informed consent of the participants, subject to approval by the project ethics board and under a formal Data Sharing Agreement. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

[1] P. Mansfield and A. A. Maudsley, "Planar spin imaging by NMR," *J. Magn. Reson.*, vol. 27, no. 1, pp. 101–119, 1977.

[2] B. Ertl-Wagner, A. Lienemann, A. Strauss, and M. F. Reiser, "Fetal magnetic resonance imaging: Indications, technique, anatomical considerations and a review of fetal abnormalities," *Eur. Radiol.*, vol. 12, no. 8, pp. 1931–1940, 2002.

[3] D. Levine *et al.*, "Fetal fast MR imaging: Reproducibility, technical quality, and conspicuity of anatomy," *Radiology*, vol. 206, no. 2, pp. 549–554, 1998.

[4] B. Huppert, K. Brandt, K. Ramin, and B. King, "Single-shot fast spin-echo MR imaging of the fetus: A pictorial essay," *Radiographics*, vol. 19, no. 1, p. 10, 1999.

[5] C. Garel, "Imaging the fetus: When does MRI really help?" *Pediatric Radiol.*, vol. 38, pp. 467–470, Jun. 2008.

[6] C. Malamateniou *et al.*, "Motion-compensation techniques in neonatal and fetal MR imaging," *Amer. J. Neuroradiol.*, vol. 34, no. 6, pp. 1124–1136, 2013.

[7] O. A. Glenn and A. J. Barkovich, "Magnetic resonance imaging of the fetal brain and spine: An increasingly important tool in prenatal diagnosis, part 1," *Amer. J. Neuroradiol.*, vol. 27, no. 8, pp. 1604–1611, 2006.

[8] N. Linduska *et al.*, "Placental pathologies in fetal MRI with pathohistological correlation," *Placenta*, vol. 30, no. 6, pp. 555–559, 2009.

[9] I. J. Zamora *et al.*, "Fetal MRI lung volumes are predictive of perinatal outcomes in fetuses with congenital lung masses," *J. Pediatric Surg.*, vol. 49, no. 6, pp. 853–858, 2014.

[10] A. Pierro, E. M. Kiely, and L. Spitz, "Classification and clinical evaluation," *Seminars Pediatric Surg.*, vol. 24, no. 5, pp. 207–211, 2015.

[11] M. C. Frates, A. J. Kumar, C. B. Benson, V. L. Ward, and C. M. Tempany, "Fetal anomalies: Comparison of MR imaging and US for diagnosis," *Radiology*, vol. 232, no. 2, pp. 398–404, 2004.

[12] Y. Patenaude *et al.*, "The use of magnetic resonance imaging in the obstetric patient," *J. Obstetrics Gynaecol. Canada*, vol. 36, no. 4, pp. 349–363, 2014.

[13] M. M. Cannie *et al.*, "Potential heating effect in the gravid uterus by using 3-T MR imaging protocols: Experimental study in miniature pigs," *Radiology*, vol. 279, no. 3, pp. 754–761, 2015.

[14] D. Levine, "Fetal magnetic resonance imaging," *J. Maternal-Fetal Neonatal Med.*, vol. 15, no. 2, pp. 85–94, Feb. 2004.

[15] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[16] F. Rousseau *et al.*, "Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images," *Acad. Radiol.*, vol. 13, no. 9, pp. 1072–1081, 2006.

[17] K. Kim, P. A. Habas, F. Rousseau, O. A. Glenn, A. J. Barkovich, and C. Studholme, "Intersection based motion correction of multislice MRI for 3-D in utero fetal brain image formation," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 146–158, Jan. 2010.

[18] A. Gholipour, J. A. Estroff, and S. K. Warfield, "Robust super-resolution volume reconstruction from slice acquisitions: Application to fetal brain MRI," *IEEE Trans. Med. Imag.*, vol. 29, no. 10, pp. 1739–1758, Oct. 2010.

[19] M. Kuklisova-Murgasova, G. Quaghebeur, M. A. Rutherford, J. V. Hajnal, and J. A. Schnabel, "Reconstruction of fetal brain MRI with intensity matching and complete outlier removal," *Med. Image Anal.*, vol. 16, no. 8, pp. 1550–1560, 2012.

[20] K. Keraudren *et al.*, "Automated fetal brain segmentation from 2D MRI slices for motion correction," *NeuroImage*, vol. 101, pp. 633–643, Nov. 2014.

[21] A. Alansary *et al.*, "Automatic brain localization in fetal MRI using superpixel graphs," in *Machine Learning Meets Medical Imaging*. Cham, Switzerland: Springer, 2015, pp. 13–22. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-27929-9_2

[22] S. Tourbier *et al.*, "Automated template-based brain localization and extraction for fetal brain MRI reconstruction," *NeuroImage*, vol. 155, pp. 460–472, Jul. 2017.

[23] B. Kainz *et al.*, "Motion corrected 3D reconstruction of the fetal thorax from prenatal MRI," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Cham, Switzerland: Springer, 2014, pp. 284–291. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-10470-6_36

[24] D. F. A. Lloyd *et al.*, "Three-dimensional modelling of the fetal vasculature from prenatal MRI using motion corrected slice-to-volume registration," in *Proc. Int. Soc. Magn. Reson. Med.*, vol. 24. Sep. 2016, p. O413.

[25] B. Kainz *et al*, "Fast volume reconstruction from motion corrupted stacks of 2D slices," *IEEE Trans. Med. Imag.*, vol. 34, no. 9, pp. 1901–1913, Sep. 2015.

[26] E. Routledge, C. Malamateniou, and M. Rutherford, "Can MR imaging of the placenta demonstrate a distinct placental phenotype in normal and abnormal pregnancies," Bachelor Sci. Degree Imag. Sci., King's College London, Univ. London, London, U.K., Tech. Rep., Apr. 2015.

[27] C. Salafia, D. Misra, M. Yampolsky, A. Charles, and R. Miller, "Allometric metabolic scaling and fetal and placental weight," *Placenta*, vol. 30, no. 4, pp. 355–360, 2009.

[28] K. Keraudren *et al.*, "Automated localization of fetal organs in MRI using random forests with steerable features," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 9351. Cham, Switzerland: Springer, 2015, pp. 620–627. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-24574-4_74

[29] A. Alansary *et al.*, Fast fully automatic segmentation of the human placenta from motion corrupted MRI," in *Proc. 19th Int. Conf. Med. Image Comput. Comput.-Assist. Intervention—MICCAI*, Athens, Greece, Oct. 2016. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-46723-8_68

[30] B. Fei, J. L. Duerk, D. T. Boll, J. S. Lewin, and D. L. Wilson, "Slice-to-volume registration and its potential application to interventional MRI-guided radio-frequency thermal ablation of prostate cancer," *IEEE Trans. Med. Imag.*, vol. 22, no. 4, pp. 515–525, Apr. 2003.

[31] S. Jiang, H. Xue, A. Glover, M. Rutherford, D. Rueckert, and J. V. Hajnal, "MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): Application to fetal, neonatal, and adult brain studies," *IEEE Trans. Med. Imag.*, vol. 26, no. 7, pp. 967–980, Jul. 2007.

[32] F. Rousseau, K. Kim, C. Studholme, M. Koob, and J.-L. Dietemann, "On super-resolution for fetal brain MRI," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Berlin, Germany: Springer 2010, pp. 355–362. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-15745-5_44

[33] F. Rousseau *et al.*, "BTK: An open-source toolkit for fetal brain MR image processing," *Comput. Methods Programs Biomed.*, vol. 109, no. 1, pp. 65–73, 2013.

[34] S. Tourbier, X. Bresson, P. Hagmann, J.-P. Thiran, R. Meuli, and M. B. Cuadra, "An efficient total variation algorithm for super-resolution in fetal brain MRI with adaptive regularization," *NeuroImage*, vol. 118, pp. 584–597, Sep. 2015.

[35] B. Kainz *et al.*, "Flexible reconstruction and correction of unpredictable motion from stacks of 2D images," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 9350. Cham, Switzerland: Springer, 2015, pp. 555–562. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-24571-3_66

[36] J. A. Little, D. L. G. Hill, and D. J. Hawkes, "Deformations incorporating rigid structures [medical imaging]," in *Proc. IEEE Workshop Math. Methods Biomed. Image Anal.*, Jun. 1996, pp. 104–113.
[37] O. Commowick *et al.*, "An efficient locally affine framework for the smooth registration of anatomical structures," *Med. Image Anal.*, vol. 12, no. 4, pp. 427–441, 2008.
[38] V. Arsigny, O. Commowick, N. Ayache, and X. Pennec, "A fast and log-Euclidean polyaffine framework for locally linear registration," *J. Math. Imag. Vis.*, vol. 33, no. 2, pp. 222–238, 2009.
[39] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[40] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
[41] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in *Proc. ECCV*, 2008, pp. 705–718.
[42] D. Goldberg, "What every computer scientist should know about floating-point arithmetic," *ACM Comput. Surv.*, vol. 23, no. 1, pp. 5–48, Mar. 1991.
[43] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 586–595, Aug. 1998.
[44] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 298–311, Feb. 1997.
[45] J. Nickolls, I. Buck, M. Garland, and K. Skadron, "Scalable parallel programming with CUDA," *Queue*, vol. 6, no. 2, pp. 40–53, Mar. 2008.
[46] J. Sanders and E. Kandrot, *CUDA by Example: An Introduction to General-Purpose GPU Programming*, 1st ed. Boston, MA, USA: Addison-Wesley, 2010.
[47] M. Steinberger, B. Kainz, B. Kerbl, S. Hauswiesner, M. Kenzel, and D. Schmalstieg, "Softshell: Dynamic scheduling on GPUs," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 161:1–161:11, Nov. 2012.
[48] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "BrainWeb: Online interface to a 3D MRI simulated brain database," in *Proc. 3rd Int. Conf. Funct. Mapping Human Brain*, vol. 5. Copenhagen, Denmark, May 1997, p. S425.
[49] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[50] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.
[51] I. Wolf *et al.*, "The medical imaging interaction toolkit," *Med. Image Anal.*, vol. 9, no. 6, pp. 594–604, 2005.