Flexible Reconstruction and Correction of Unpredictable Motion from Stacks of 2D Images

Bernhard Kainz[1,2], Amir Alansary[1], Christina Malamateniou[2], Kevin Keraudren[1], Mary Rutherford[2], Joseph V. Hajnal[2], and Daniel Rueckert[1]

Department of Computing, Imperial College London, UK
{b.kainz,kevin.keraudren10,d.rueckert}@imperial.ac.uk
King's College London, Division of Imaging Sciences, London, UK
{christina.malamateniou,mary.rutherford,jo.hajnal}@kcl.ac.uk

Abstract. We present a method to correct motion in fetal in-utero scan sequences. The proposed approach avoids previously necessary manual segmentation of a region of interest. We solve the problem of non-rigid motion by splitting motion corrupted slices into overlapping patches of finite size. In these patches the assumption of rigid motion approximately holds and they can thus be used to perform a slice-to-volume-based (SVR) reconstruction during which their consistency with the other patches is learned. The learned information is used to reject patches that are not conform with the motion corrected reconstruction in their local areas. We evaluate rectangular and evenly distributed patches for the reconstruction as well as patches that have been derived from super-pixels. Both approaches achieve on 29 subjects aged between 22–37 weeks a sufficient reconstruction quality and facilitate following 3D segmentation of fetal organs and the placenta.

1 Introduction

Evaluation of fetal organs and the placenta is an important diagnostic tool during prenatal screening and is considered to be an indicator for fetal health after birth. Fetal Magnetic Resonance Imaging (MRI) allows to acquire high resolution slices from the fetus at a large field of view and with good tissue contrast [9]. However, the fetus is not sedated during these scans and may move freely inside the uterus. Because of a scan time of up to 500 ms per slice, motion artefacts are likely to corrupt volumetric scans. Therefore, several (usually 3–12) orthogonal stacks of slices are acquired and reconstructed using approaches based on slice-to-volume registration (SVR) to obtain an artefact free, high resolution volume of a fetal target region [8,5]. So far, this process has been applied only to small regions and organs with rigid body characteristic such as the fetal brain. Usually, these areas have to be identified by manual labor intensive segmentation methods. Such approaches cannot be applied to the whole fetal body and uterus because of the assumption of rigid motion in the 2D to 3D registration step of SVR methods.

© Springer International Publishing Switzerland 2015
N. Navab et al. (Eds.): MICCAI 2015, Part II, LNCS 9350, pp. 555–562, 2015.
DOI: 10.1007/978-3-319-24571-3_66

Different areas in each slice that are likely to move in different directions will break this assumption. Because an extension of 2D-3D registration to non-rigid deformations is not well-constrained, current SVR approaches will fail for non-rigid deformations and movements.

Contribution: We solve the motion compensation problem for large field of views in stacks of 2D slices. 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 volume and to exclude corrupted data automatically. This approach paves the way to fully automatically reconstruct whole collections of motion corrupted stacks without the need for manually segmented input. The method also finds rigidly connected areas automatically, which can be used for further refinement or as segmentation prior.

Related Work: Fetal MRI is increasingly used as a complementary diagnostic tool to ultrasound sonography. Currently, the brain [8], thorax [6], and the appearance of the whole fetus [10] are qualitatively examined using MRI in the clinical practice. Fetal motion and its unpredictable nature make the acquisition of 3D MR sequences very challenging. Fast MR sequences such as single shot fast spin echo (ssFSE) [9] are often used in order to freeze motion within a single 2D image. Using several overlapping stacks of 2D images provides coverage of a 3D volume of a target region of interest. However, the resulting stacks are usually corrupted by relative motion between individual slices. Often six to twelve stacks need to be acquired to sufficiently oversample a 3D volume. Segmentation and localization of selected organs can be automated, however, the available approaches provide either a very rough segmentation of the central slices of a stack [7] or require less motion corrupted stacks [4]. Furthermore, they are only applicable to a specificity trained region, *e.g.*, the fetal brain.

2 Method

The proposed method is based on the fact that certain regions of a scanned anatomy are rigid and can be reconstructed with SVR super-resolution algorithms. These methods usually use manually defined rigid regions of the 2D input images (slices) and register them to an iteratively improving global 3D reconstruction volume. Robust statistics can be used to identify mis-registered or heavily corrupted data [8,11]. Data consistency is reached by oversampling a region of interest with different scan orientations. We propose to reduce the granularity of the input data by using 2D data *patches* of arbitrary shape instead of whole slices for SVR reconstruction. That way, multiple, large motion

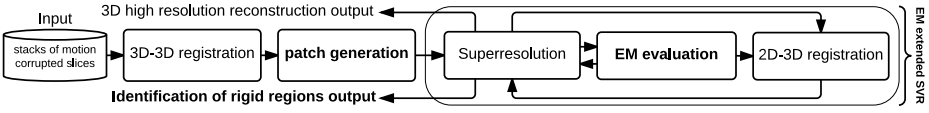


Fig. 1. An overview over our approach. Bold parts are extensions to SVR.

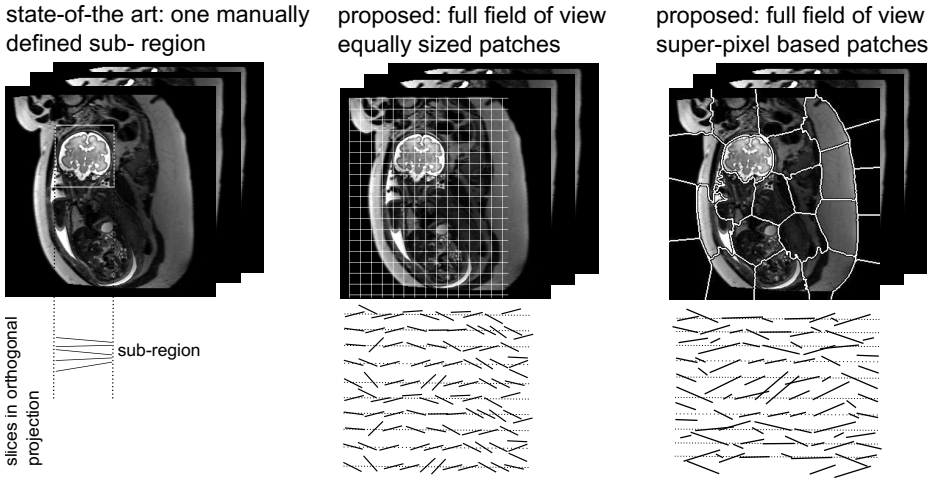


Fig. 2. Comparison of the proposed reconstruction methods to the state-of-the-art.

corrupted field of views can be reconstructed and regions with rigid motion can be found automatically. Fig. 1 gives a schematic overview over our approach and Fig. 2 shows a high level overviews over the current state-of-the-art manual segmentation-based reconstruction paradigm compared to the here presented fully automatic, full field of view reconstruction method.

The input data can be represented as stacks of 2D images consisting of $Y = \{y_s | s \in S\}$, where y_s is a 2D patch indexed by the location s and S is the set of all locations in all p stacks, $S = \{s_1, s_2, ...s_M\}$. y_s can have arbitrary (2D) shape. In this work we explore using overlapping square patches as a general application of our novel reconstruction method and Simple Linear Iterative Clustering (SLIC) super-pixels [1] as a method to reduce the required data redundancy.

SVR: We can reconstruct a high resolution image X from a number of motion corrupted y_s using 2D-3D registration-based super-resolution [8,5]. After initial 3D-3D alignment a gradually improving approximation of X (super-resolution with the measured point spread function of the used MRI sequence) is used to initialize and perform 2D-3D registration and robust statistics. To provide enough structural information for rigid registration of y_s to X we dilate each y_s by γ pixels using a flat structuring element b with a fixed (26 in our case) pixels neighbourhood, hence $\bar{y}_s = y_s \oplus b$.

Patch Generation: The shape of y_s can be square with similar edge sizes in the simplest naïve case. These patches can be defined by their edge length a and stride ω. While this definition is likely to be generally applicable to any kind of oversampled motion corrupted data, it does not assume any knowledge about the data and a and ω are likely to depend on for example the gestational age. Ideally, each y_s corresponds to a subregion of the volume in which the motion can be characterized as rigid. A good trade-off between the size of the patch region and the likelihood of rigid motion has to be found. The larger the cho-

sen patch regions are the less likely they will cover rigidly moving areas. An alternative to naïve shape definitions of y_s is to use correlation between each pixel and its neighbors. Such correlations can be found by popular unsupervised image segmentation techniques like *super-pixels*. Super-pixels clusters the image into areas of pixels with local correlations. Correlated regions may define rigidly connected areas, which can support the image reconstruction step with less but more useful data blocks. In the literature, there are different techniques for generating super-pixels. We aim for clinical applicability of our method. Therefore, we selected a super-pixel algorithm, which is fast to compute, i.e., we use Simple Linear Iterative Clustering (SLIC) [1]. SLIC allows to segment 2D image slices into compact and uniform super-pixels as shown in Fig. 2. This approach is also computationally more efficient for image reconstruction because larger rigid areas require less redundant image registration and super-resolution effort, independent from data parameters like gestational age. To determine the number of super-pixels N_{sp} for each 2D slice, we use the rule of thumb proposed by [3], which is based on the total number of pixels n. To handle the high variability of the size and shape of our data, we have weighted this generation rule with a constant factor k, where $k \in \mathbb{R}_{>0}$ and is chosen depending on the resolution of the input data and thus $N_{sp} \approx k \cdot \sqrt{(n/2)}$.

EM Evaluation: We aim to use only voxels from $\bar{y}_s$ that can be well registered and that have a minimal error e when compared to the originally scanned data. To achieve this we propose to classify $\bar{y}_s$ and the included pixels into an inlier and an outlier class using an expectation maximization (EM) framework. Inspired by [2,8], we use a zero-mean Gaussian distribution $G_\sigma(e)$ with variance σ^2 for the inliers and a uniform distribution with constant density $m = \frac{1}{max(e)-min(e)}$ for the outliers. This allows us to use redundant information, i.e., overlapping $\bar{y}_s$, to find partly matching patches and to depreciate or fully reject erroneous voxels of each $\bar{y}_s$. For these regions we try to maximize the log-likelihood for each patch $y_s|logP(Y, \Phi) = \sum logP(e|\sigma, c)$ to be part of an area that undergoes rigid motion. Φ contains 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 $\in \bar{y}_s$ being identified as inlier is $p = \frac{G_\sigma(e)c}{G_\sigma(e)c+m(1-c)}$. We perform the updates of c and σ^2 similar to [8]. Using $\bar{p} = \sqrt{(\sum_{\bar{y}_s} p^2)/N}$ (with N the number of pixels in $\bar{y}_s$) we can also define an inlier and outlier probability for each patch $\bar{y}_s$ and stop processing this patch if it gets classified as outlier.

Identification of Rigid Regions: Keeping track of the probability p of each pixel of every $\bar{y}_s$ allows to identify areas that best fit the rigid 2D-3D registration constraint of SVR methods. Integrating p and $\bar{p}$ into a separate probability volume P using the same slice to volume integration scheme as used by SVR can be used to identify candidate regions, which contain only rigid motion components. This can be useful to apply the classic SVR reconstruction approach at a higher level of detail only in these regions. As shown by [7], tight region of interest masks can lead to a higher reconstruction quality for rigid regions like the fetal brain. In practice we can identify such rigid regions by blurring P using

a 3D Gaussian filter with σ related to the size of the desired regions followed by blob detection. These regions can subsequently be reconstructed and motion corrected in the high-resolution volume and used for automatic classification, *e.g.*, in organ classes with machine learning. Fig. 5(g) shows cross section views through P, which was generated using super-pixels at $k = 0.2$.

3 Evaluation and Results

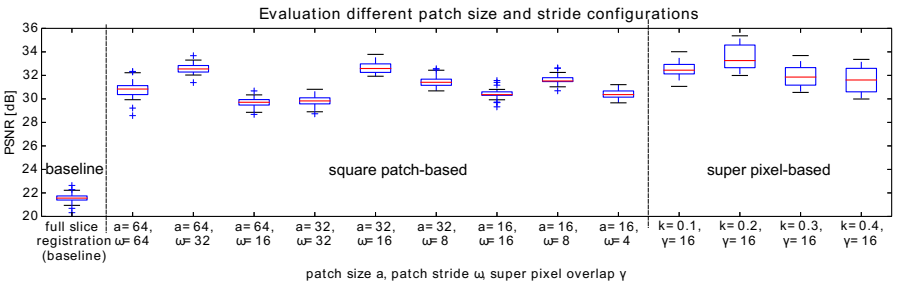
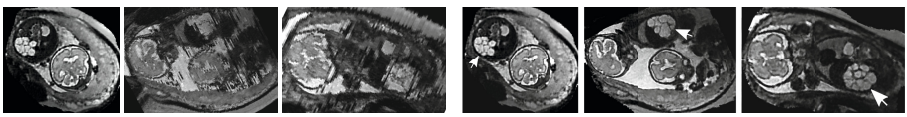


Fig. 3. PSNR comparison in the brain region of 29 subjects between using the slices of the input stacks directly for full field of view SVR reconstruction (baseline, most left whisker-box) to different regular patch sizes a with varying stride ω. The last four whisker-box plots are using super pixels with varying k and an overlap $\gamma = 16$ pixels, which yielded good results during our experiments. A PSNR above 30 dB shows that all the proposed configurations produce a result, which is very similar to a reconstruction when using a tight manually defined mask. Small square patches with $a = 32$, $\omega = 16$ and super pixels with $k = 0.2$, $\gamma = 16$ produced the best results during this experiment.

We evaluate our method with experiments using 29 data sets from fetuses with gestational ages between 22–37 weeks. To the best of our knowledge our method



(a) original (b) reconstruction

Fig. 4. Three viewing planes through the originally scanned (a) and the reconstruction (b) of a motion corrupted scan from moving twins with a gestational age of 28 weeks using super-pixels $k = 0.2$ $\gamma = 16$. For this dataset we used a mask of the uterus so save unnecessary computation time in areas containing maternal tissue. The white arrow points at a unilateral multicystic kidney of one of the twins.

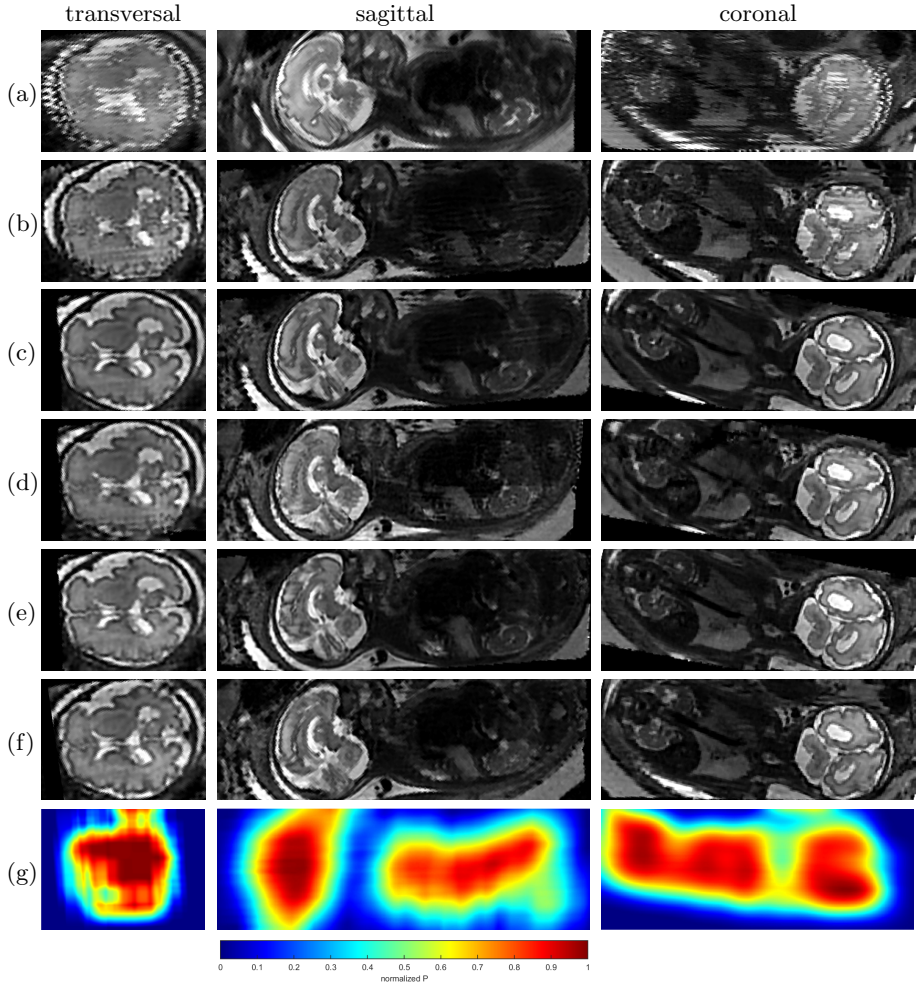


Fig. 5. Visual comparison between the input data (a) and different configurations for full fetal body reconstruction of a motion corrupted 3T MRI dataset with gestational age of 33 weeks. The categorization into best and worst has been made by an expert. (b): reconstruction with full slices and [8]; (c): best: $s = 32 \times 16, \omega = 16$; (d): worst $s = 64 \times 64, \omega = 64$; (e): best: super-pixels $k = 0.2$ $\gamma = 16$, (f): worst: $k = 0.05$ super-pixels/slice $\gamma = 16$. (g) indicates which regions move as a rigid body as a function of P ($\sigma = 15$ with super-pixels $k = 0.2$, $\gamma = 16$) for subsequent automatic identification of rigid regions.

provides the first approach to reconstruct other areas than the brain or the lung, hence there is no ground truth for the full fetal body to compare with. However, we can compare the reconstruction quality with a well researched organ: the fetal brain. Our hypothesis is that our method provides similar reconstruction and motion correction quality for the brain as it would be the case if a tight mask [7] for a region of interest would have been used for SVR. We expect the

quality of our results to be close to the results from the state-of-the-art SVR approach [5] for rigid regions. Fig. 3 shows peak signal-to-noise ratio (PSNR) comparisons with a defined region in the fetal brain for different patch sizes (a) and super-pixel sizes (b) with different strides and overlaps compared to a full field of view reconstruction using the slices directly without masking or splitting into patches (baseline). Fig. 5 shows an expert quality assessment of the results from different image parcellations applied to a full fetal body 3T ssFSE dataset.

Our method allows for the first time the reconstruction and motion correction of scans of the whole uterus with more than one fetus. Up to now only selected regions could be reconstructed and malformations in multiple births as shown in the kidney of one twin in Fig. 4 were difficult to examine.

Runtime: Super-pixels can be generated for all slices of motion corrupted stacks within a few seconds ($\sim$ 800 2D images/examination). Our approach becomes slower the more patches/super-pixels and the more overlap is used (approx. quadratically). We use a parallelized and hardware accelerated SVR reconstruction method based on [5]. A full field of view reconstruction of 8 input stacks at $288 \times 288 \times 100$ voxels takes up to $1 - 2$ hours using a small patch size (*e.g.*, $a = 32$, $\omega = 16$) on a multi GPU System (Intel Xeon E5-2630 2.60GHz system with 16 GB RAM, an Nvidia Tesla K40 and a Geforce 780). Using large ($k = 0.1$) overlapping super-pixels reduces this time to approximately $45min$ for a full field-of-view volume, while maintaining a comparable result to the best configuration of overlapping square patches.

4 Discussion and Conclusion

We have presented a method to fully automatically reconstruct the full overlapping field of view of multiple motion corrupted stacks of 2D slices. This method is generally applicable to motion corrupted scan protocols and especially useful for fetal MRI. We discuss how data patches can be used to tackle the problem of locally rigid body movements between different body parts. The presented method can be used to provide a segmentation prior for rigid and stable regions like the fetal brain and the thorax and provides a motion corrected overview over the whole uterus in 3D including multiple births and the placenta. For certain configurations our method might produce a slightly less accurate reconstruction of specific body parts as it would be the case when using tight masks of these regions. However, this can be solved by using standard slice-based SVR reconstruction as a subsequent step, applied only to the automatically detected rigid areas or an automatically derived mask. In fact our method provides an excellent starting point to apply state-of-the-art but motion sensitive 3D volume analysis and segmentation methods directly to the motion corrected result. We introduce the use of super-pixels as an alternative to naïve overlapping square patches. While the parameter configuration for naïve patches is likely to be data dependent (*e.g.*, gestational age), super-pixels provide a framework, which is invariant towards such variations in the data. In future work we will investigate potential improvements of the super-resolution step in SVR methods by using

consistent patches for dictionary learning methods to provide a sparse representation of the desired high-resolution volume.

Acknowledgements. We used MITK [12]. This work was supported by Wellcome Trust and EPSRC IEH award [102431] for the iFIND project.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Analysis and Machine Intelligence 34(11), 2274–2282 (2012)
2. Gholipour, A., Estroff, J.A., Warfield, S.K.: Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI. IEEE Trans. Med. Imaging 29(10), 1739–1758 (2010)
3. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate data analysis, vol. 6. Pearson Prentice Hall, Upper Saddle River (2006)
4. Kainz, B., Keraudren, K., Kyriakopoulou, V., Rutherford, M., Hajnal, J.V., Rueckert, D.: Fast Fully Automatic Brain Detection in Foetal MRI Using Dense Rotation Invariant Image Descriptors. In: IEEE ISBI 2014, pp. 1230–1233 (2014)
5. Kainz, B., Steinberger, M., Wein, W., Murgasova, M., Malamateniou, C., Keraudren, K., Torsney-Weir, T.K., Aljabar, P., Rutherford, M., Hajnal, J., Rueckert, D.: Fast Volume Reconstruction from Motion Corrupted Stacks of 2D Slices. IEEE Trans. Med. Imaging 34(9), 1901–1913 (2015)
6. Kainz, B., Malamateniou, C., Murgasova, M., Keraudren, K., Rutherford, M., Hajnal, J.V., Rueckert, D.: Motion Corrected 3D Reconstruction of the Fetal Thorax from Prenatal MRI. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 284–291. Springer, Heidelberg (2014)
7. Keraudren, K., Murgasova, M., Kyriakopoulou, V., Malamateniou, C., Rutherford, M., Kainz, B., Hajnal, J., Rueckert, D.: Automated Fetal Brain Segmentation from 2D MRI Slices for Motion Correction. NeuroImage 101, 633–643 (2014)
8. Kuklisova-Murgasova, M., Quaghebeur, G., Rutherford, M.A., Hajnal, J.V., Schnabel, J.A.: Reconstruction of Fetal Brain MRI with Intensity Matching and Complete Outlier Removal. Medical Image Analysis 16(8), 1550–1560 (2012)
9. Levine, D.: Fetal Magnetic Resonance Imaging. J. Matern. Fetal Neonatal Med. 15(2), 85–94 (2004)
10. Prayer, D., Brugger, P., Prayer, L.: Fetal MRI: techniques and protocols. Pediatric Radiology 34(9), 685–693 (2004)
11. Rousseau, F., Glenn, O.A., Iordanova, B., Rodriguez-Carranza, C., Vigneron, D.B., Barkovich, J.A., Studholme, C.: Registration-Based Approach for Reconstruction of High-Resolution In Utero Fetal MR Brain Images. Academic Radiology 13(9), 1072–1081 (2006)
12. Wolf, I., Vetter, M., Wegner, I., Böttger, T., Nolden, M., Schöbinger, M., Hastenteufel, M., Kunert, T., Meinze, H.P.: The medical imaging interaction toolkit. Medical Image Analysis 9(6), 594–604 (2005)