A discriminative feature selection approach for shape analysis: Application to fetal brain cortical folding

J. Pontabry [a,*], F. Rousseau [b], C. Studholme [c], M. Koob [d,e], J.-L. Dietemann [d,e]

[a] *Institute for Epigenetics and Stem cells, Helmholtz Zentrum München, Germany*
[b] *Institut Mines Télécom, Télécom Bretagne, INSERM, LaTIM U1101, Brest, France*
[c] *Biomedical Image Computing Group, Departments of Pediatrics, Bioengineering and Radiology, University of Washington, Seattle, USA*
[d] *Laboratoire ICube, UMR7357, CNRS-Université de Strasbourg, France*
[e] *Service de Radiopédiatrie, Hôpital de Hautepierre, Strasbourg, France*

A B S T R A C T

The development of post-processing reconstruction techniques has opened new possibilities for the study of in-utero fetal brain MRI data. Recent cortical surface analysis have led to the computation of quantitative maps characterizing brain folding of the developing brain.

In this paper, we describe a novel feature selection-based approach that is used to extract the most discriminative and sparse set of features of a given dataset. The proposed method is used to sparsely characterize cortical folding patterns of an in-utero fetal MR dataset, labeled with heterogeneous gestational age ranging from 26 weeks to 34 weeks.

The proposed algorithm is validated on a synthetic dataset with both linear and non-linear dynamics, supporting its ability to capture deformation patterns across the dataset within only a few features.

Results on the fetal brain dataset show that the temporal process of cortical folding related to brain maturation can be characterized by a very small set of points, located in anatomical regions changing across time. Quantitative measurements of growth against time are extracted from the set selected features to compare multiple brain regions (e.g. lobes and hemispheres) during the considered period of gestation.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Development of the central nervous system (CNS) is governed by a complex set of several spatio-temporal mechanisms. Understanding this sequence is fundamental for understanding perinatal neurology (Volpe, 2008). Many processes are involved in brain development at fetal and neonatal stages. These are asynchronous processes taking place at different locations in the brain. For instance, the main period of neural migration is 5 to 25 week postmenstrual age (PMA). Once neurons have been generated, they migrate through two different mechanisms (passive cell displacement and active cell migration). Axon and dendrite sprouting starts around 25 weeks PMA and declines around 1 year. The brain is also the scene of other major events such as synapse formation, glial cell proliferation, myelination, etc. (see de Graaf-Peters and Hadders-Algra, 2006 for a recent review of the ontogeny of the human CNS).

As highlighted by de Graaf-Peters and Hadders-Algra (2006), the knowledge on the exact timeline of ontogenetic events occurring during human brain development will provide new insights on the influence of injuries appearing at a specific point in time during this important period of brain building. The modeling of this process timeline from in vivo data is of great importance for improving, for instance, neonatal care services. It is clear that fetal and neonatal periods are key steps of brain development and a better understanding of the involved mechanisms is a corner stone in perinatal care.

Recent advances in ultrafast MRI sequences (such as half-Fourier turbo spin echo (HASTE) (Yamashita et al., 1997) or single shot fast spin echo (SSFSE) Busse et al., 2000) associated with the development of new techniques of image processing for retrospective motion correction (Rousseau et al., 2005) enabled to study in vivo fetal brain development. Since 2005, several registration-based reconstruction techniques have been proposed to estimate a 3D high resolution image from sets of scattered T2-weighted slices (Rousseau et al., 2006; Jiang et al., 2007; Limperopoulos and Clouchoux, 2009; Kim et al., 2010). These reconstruction techniques have been further enhanced by includ-

ing super-resolution framework (Rousseau et al., 2010; Gholipour et al., 2010; Kuklisova-Murgasova et al., 2012). Following these works on 3D fetal brain image reconstruction, several dedicated segmentation methods have been proposed (Habas et al., 2010b; Caldairou et al., 2011; Gholipour et al., 2012; Dittrich et al., 2014). See the review of Studholme (2011) for further reading on fetal MRI reconstruction and segmentation.

The development of these post-processing techniques had led to new ways to perform structural development studies of the fetal brain. Based on local volume changes using tensor-based morphometry techniques, maps of fetal brain growth patterns have been estimated leading to the detection of the emergence of sulci and gyri by differentiating between the changes in the cortical plate and the underlying cerebral mantle (Rajagopalan et al., 2011). This work has been extended to examine not only scalar expansion of tissue but the directional components of that expansion (Rajagopalan et al., 2012). The availability of high resolution 3D images of the fetal brain with tissue labeling enables also studies of tissue boundary shape changes. This has led to the computation of quantitative maps of brain folding further used for statistical detection of sulci and brain asymmetry emergence (Habas et al., 2012). Cortical folding analysis has been also used to predict physiological age (Wright et al., 2014). Such brain folding studies are of great importance to improve our understanding of malformation of the cortex.

In this work, we focus on the study of the evolution of brain folding during later stages of intrauterine life (third semester). Instead of computing local scalar features (e.g. cortical surface curvature), we have chosen to investigate a data-driven dimensionality reduction approach, allowing the extraction of global biomarkers of the brain development. The underlying question is: what is the smallest most discriminative set of features reflecting fetal brain folding? Following the work of Rajagopalan et al. (2012), a deformation based morphometry approach is adopted to capture directional growth information on the white matter / cortex interface. However, the voxel-based statistical testing method applied in Rajagopalan et al. (2012) does not lead to easily interpretable results of growth patterns. In this study, we have considered the use of feature selection techniques in order to extract the sparsest set of deformation fields describing the brain folding process. By studying brain development as a shape modeling problem, we are able to extract the most discriminative set of points related to brain folding during in-utero development.

2. Materials and methods

2.1. Subjects and fetal MRI acquisition

This study has been conducted on a population of 22 healthy fetus aged from 26 to 34 weeks of gestational age at scan time (normal MRI appearance checked by a radiologist, no issue during the first days after delivery). The data set includes 23 MRI T2 weighted images. Fetal MRI images have been obtained on a 1.5 T Siemens Avanto MRI Scanner (SIEMENS, Erlangen, Germany) using a 6-channel phased array coil combined to the spine array positioned around the mother abdomen. The resolution of the T2 weighted HASTE sequence (TE/TR = 147/3190 ms) is: $0.74 \times 0.74 \times 3.45$ mm. Every image acquisition took about 1 min; at least 3 min in total for each subject (acquisition of multiple orthogonal images).

2.2. Image reconstruction and segmentation

Raw dataset has been preprocessed using the "Baby Brain Toolkit" (BTK) in order to increase image quality without mod-

ifying the acquisition protocol used in routine (Rousseau et al., 2013). The retrospective motion correction method is based on a registration refined compounding of multiple sets of orthogonal fast 2D MRI slices to address the problem of fetal motion. This is achieved by first globally registering the low resolution images, and then applying an iterative slice alignment scheme which seeks to refine the 3D positioning of each slice to the current combined high resolution volume. This is driven by normalized mutual information to provide robustness to contrast variation induced by motion of the fetal brain with respect to the imaging coil in the magnet. Moreover, a super-resolution technique is applied in order to remove the effects of the blurring convolution and to increase the voxel grid density. The resolution of the reconstructed images is: $0.74 \times 0.74 \times 0.74$ mm. A topological based clustering technique is then applied on the motion-compensated high-resolution images to provide segmentation maps (ventricles, CSF, cortical plate, non-cortical plate –including white matter, subplate, intermediate zone and deep gray nuclei) (Caldairou et al., 2011).

2.3. Brain image normalization

Adopting a standard approach for deformation based morphometry to compare anatomies at different stages of brain development, a mapping is estimated to bring every subject's anatomy into correspondence within a common coordinate system. This common space corresponds to the average space of the subjects' anatomies. For this purpose, an iterative registration approach proposed by Guimond et al. (2000) has been used in this work. The procedure to estimate a common space is as following. First, a reference image is chosen among the dataset and the non linear transformations from this reference to the remaining images of the dataset are estimated using ANTS diffeomorphic registration technique (Avants and Gee, 2004). Then, the bias introduced by the choice of a reference is corrected by using the average of all transformations. Finally, this average transformation maps the chosen reference to a new space which corresponds to the average space of the subject's anatomies. The reference for the next iteration will be this new space. These steps are repeated iteratively, until convergence of the algorithm, which usually occurs within a few iterations (Guimond et al., 2000).

The choice of the first reference may impact the convergence of the normalization algorithm. A good option is to choose the closest subject to the average space, that is the subject sharing anatomy with any subject in the population. A natural choice for the first reference appears to be a middle age subject sharing all of the considered structures with both younger and older subjects. In this work, since the population age ranges from 26 to 34 (see the age histogram of population in Fig. 1), we empirically chose a subject aged of 30 weeks.

Parameters for ANTS registration are: cross correlation as similarity measure, Gaussian regularization with a symmetric diffeomorphic transformation model. To avoid any possibility of misregistration due to brain tissue contrast changes, the similarity criterion used during the registration process is a weighted combination of image intensities and tissue label maps (a similar strategy has been proposed by Habas et al. (2010a)). Weights have been set to 0.5 for both features.

To assess the quality of the atlas building step based on image registration, we have computed the average DICE coefficient of the gray matter maps between the estimated template and the population. This average DICE coefficient is equal to 0.85 ± 0.11, showing the validity of the approach.
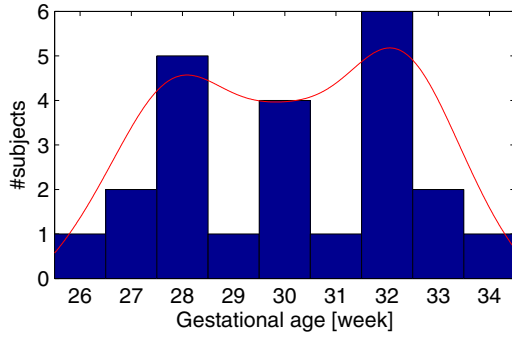
Fig. 1. Age histogram of the considered population.

Fig. 3. Dynamic of the two simulated sulci. A sigmoid function has been used to simulate a non-linear growth with parameters (A, λ, T) equal to (1.0, 1.0, 14.0) (red) and (0.7, 0.8, 7.0) (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Feature selection

Our objective concerns the selection of the most discriminative deformation-based features reflecting fetal brain development. Here we assume that the fetal brain development process implies heterogeneous deformation fields across the population (subjects have different gestational age). Please note that although we focus in this work on brain folding, the proposed approach is versatile and it could be used with any other kind of relevant features dedicated to brain anatomy study.

Let be $\mathcal{P} = \{I_1, I_2, \ldots, I_N\}$ a set of N images. Theses images are normalized in an average space $\bar{I}$ computed as described previously in Section 2.3 and that allows to capture shape changes across the population. The transformation $\mathcal{T}_i$ that maps the average space $\bar{I}$ to the image I_i is a composition of an affine transform and a non-linear deformation field. In order to capture local shape changes only and following the previous work of Aljabar et al. (2011), we only consider the non linear components of the mappings $\{\mathcal{T}_i\}_{i=1,\ldots,N}$. Indeed, since the volume change is captured by the affine component, we expect the deformation fields to be representative of the local deformation, such as the sulcal formation.

Each of the N non linear deformation fields that maps voxels from average space to population space consists of M vectors of $\mathbb{R}^3$ (where M is 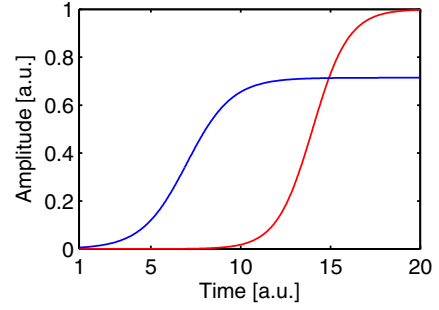the number of voxels considered in the region of interest). These 3D vectors are further arranged in a matrix $Y \in \mathcal{M}_{M \times N}(\mathbb{R}^3)$ where the rows and columns correspond respectively to the features and the samples:

$$Y = \begin{pmatrix} \vec{p}_{1,1} & \cdots & \vec{p}_{1,N} \\ \vdots & \ddots & \vdots \\ \vec{p}_{M,1} & \cdots & \vec{p}_{M,N} \end{pmatrix}, \tag{1}$$

where $\vec{p}_{i,j}$ is the 3D displacement vector of the voxel i in the deformation field of subject j. Since we focus in this study on brain folding, the analysis is restricted to the displacement vectors belonging to the cortical plate. The matrix Y is the full set of features from which a sparse representation of brain folding is estimated. Instead of adopting a voxel-based approach, that is to say applying a statistical testing method to detect the most significant folding patterns locally, we propose a dimensionality reduction approach aiming at extracting from the matrix Y a small set of discriminative features.

Such a dimension reduction issue can be tackled with either feature extraction or feature selection techniques. Feature extraction methods transform the high-dimensional data into a space of fewer dimensions. The main linear technique is the principal component analysis (PCA) which performs a linear mapping
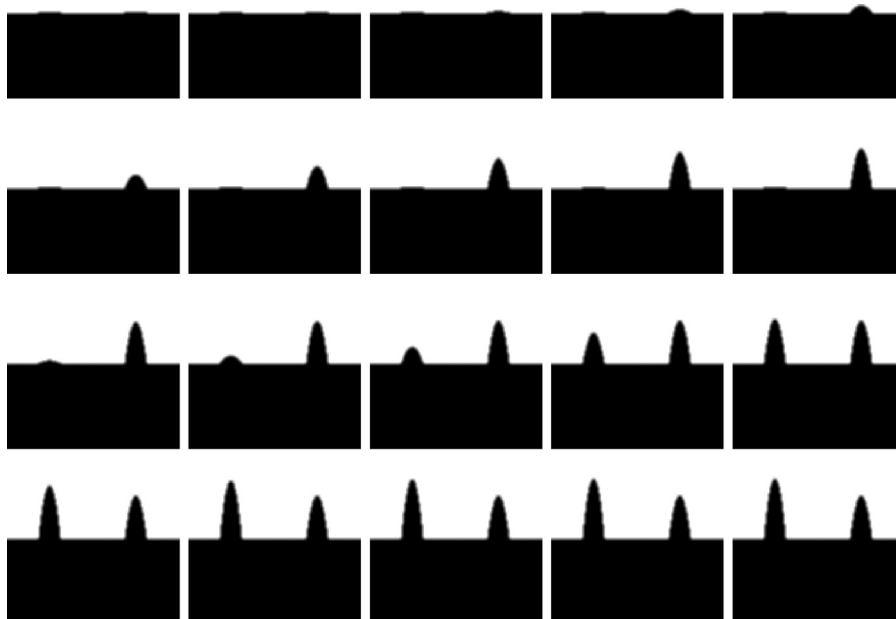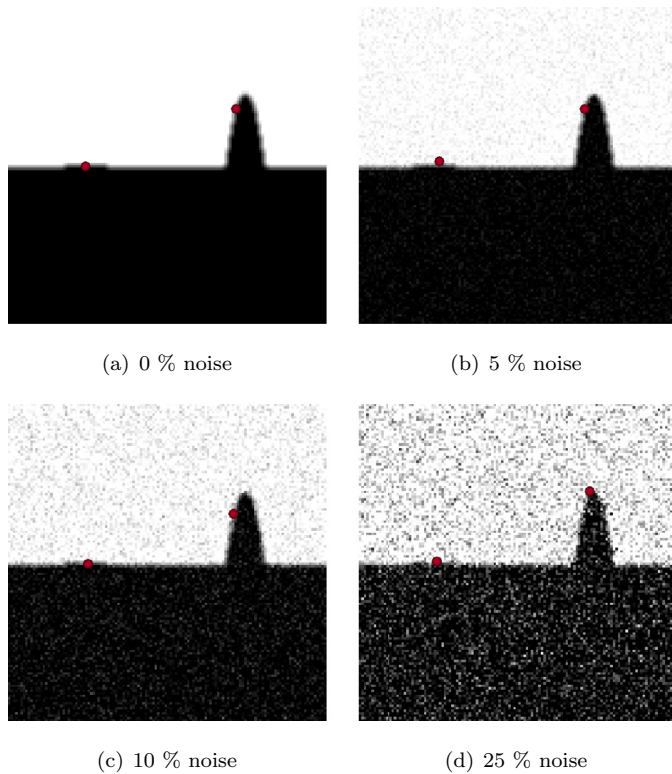

Fig. 2. Dataset used in the synthetic experiment, composed of 20 images of size 128 × 128 pixels. The 2D shapes have been created using a time-varying B-spline curve in order to simulate the folding cortical surface.

(a) 0 % noise

(b) 5 % noise

(c) 10 % noise

(d) 25 % noise

Fig. 4. Points selected by the proposed algorithm on the synthetic dataset with 50 points to select ($M' = 50$) for four different noise levels. The location of selected points, inside sulci, is not much affected by noise level.
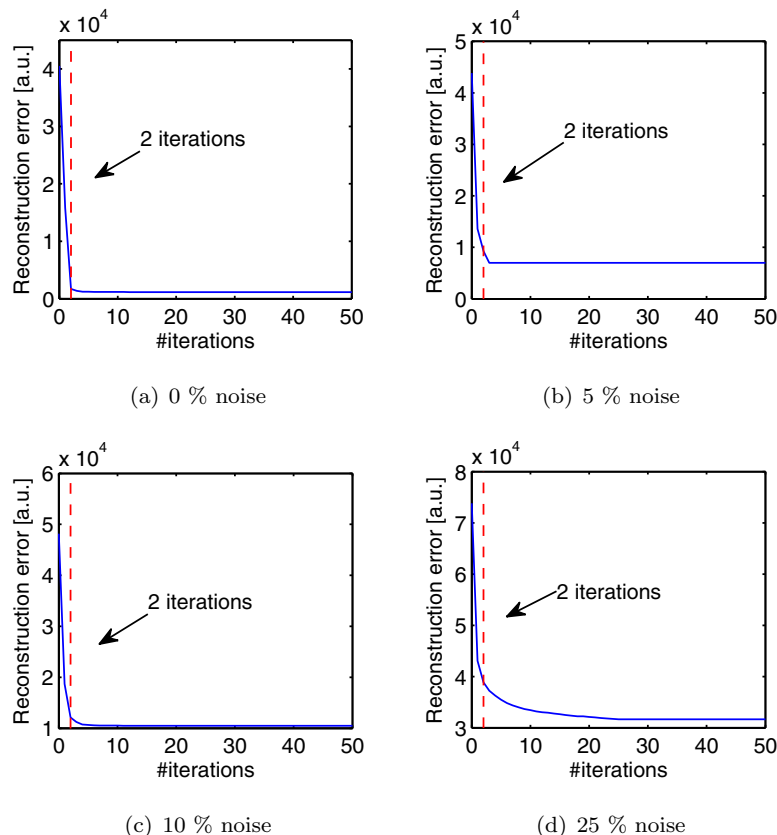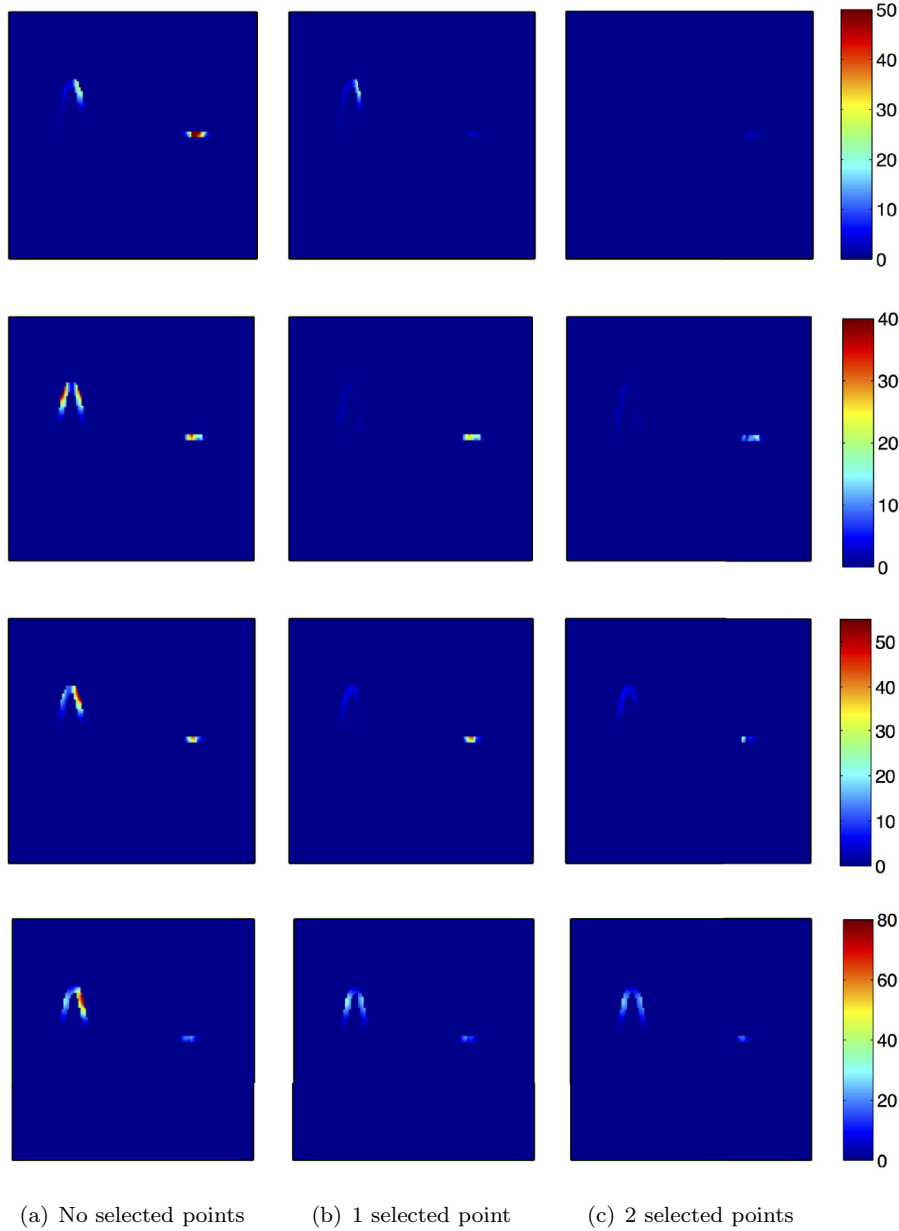
of the data to a lower-dimensional space in such a way that the variance in the low-dimensional representation is maximized and the co-variance is minimized. Although PCA may be applied to study affine deformation, the linear assumption makes it less adapted for non linear deformations. Non-linear feature extraction techniques have been proposed to preserve local data structures. An important example of such non linear techniques is Isomap (Tenenbaum et al., 2000) (see Gerber et al., 2010 for an example of Isomap application to manifold modeling for MRI dataset analysis). However, the lower-dimensional space may not be easily interpretable, meaning that in our context, the extracted patterns do not correspond to displacement fields anymore.

In this work, we investigate an alternate method by relying on a feature selection approach, aiming at computing a subset of the original variables. A general review of feature selection can be found in the paper of Guyon and Elisseeff (2003) and related selection techniques can be found in statistical learning books (Hastie et al., 2009; Elad, 2010). Within such framework, the sparsity problem is written as an optimization problem, described in the following paragraphs.

Let Y_j be the jth column of the matrix Y and X_j a subset vector of Y_j ($X_j \subset Y_j$) of size $M' \leq M$. The key assumption here is that the entire set of non-linear deformations can be reconstructed using only a very small set of 3D displacement vectors. Within such a sparse framework, Y_j and X_j are linked by a reconstruction function f:

$$Y_j = f(X_j). \tag{2}$$

In this work, we chose to formulate the reconstruction function f as a regression function because it is appropriate for a discrimination criterion. Since we do not want to put any prior on the shape or degree of freedom, we chose the Nadarya–Watson kernel

(a) 0 % noise

(b) 5 % noise

(c) 10 % noise

(d) 25 % noise

Fig. 5. Cost function evolution of the proposed algorithm on the synthetic dataset with 50 points to select ($M' = 50$) for four different noise levels. The cost functions decrease dramatically when selecting the first points and are almost constant after two selected points. As the noise level increases, the cost function becomes more smooth.

(a) No selected points (b) 1 selected point (c) 2 selected points

Fig. 6. Residuals of the reconstruction function with the considered synthetic dataset for 0, 1 and 2 selected parameter. From top to bottom: dataset with noise level of 0 %, 5 %, 10 % and 25 %. Each selected parameter allows to decrease dramatically the residuals in its area of influence. As the noise level increases, this observation becomes less clear.

function (Nadaraya, 1964):

$$\hat{f}(X, h) = \sum_{j=1}^{N} w_j(X, h) Y_j \tag{3}$$

where the weighting function w_j is written as

$$w_j(X, h) = \frac{K_h(X - X_j)}{\sum_{k=1}^{N} K_h(X - X_k)}. \tag{4}$$

The parameter h is the bandwidth of the kernel K_h acting on the subset X, which could be any of the X_j. The proposed sparsity-based approach relies on the assumption of redundant information in the original sample Y. Therefore, the relevant information is carried by a subset of parameters, i.e. a subset of rows in matrix $\mathbf{Y}$. Such a subset of M' parameters should minimize the following cri-

terion:

$$\{\hat{\gamma}, \hat{h}\} = \arg\min_{\gamma, h} J(\gamma, h) \text{ s.t. } \|\gamma\|_0 \leq M' \ , \tag{5}$$

where $\gamma \in \mathcal{M}_{M \times 1}(\{0, 1\})$ is a binary vector used to define the activated parameters and J is the reconstruction error defined as follows:

$$J(\gamma, h) = \sum_{j=1}^{N} \left\| Y_j - \hat{f}(\text{diag}(\gamma) \cdot Y_j, h) \right\|_2. \tag{6}$$

Optimization problems such as defined by Eq. (5) are NP-hard problems because of the use of L0 norm (Elad, 2010). An approximated solution of such problems can be obtained either by relaxing the L0 norm to the L1 norm or by using a heuristic algorithm. In this work, we used an alternate optimization scheme, i.e. the two parameters of interest $\{\gamma, h\}$ are estimated alternatively. The estimation of the bandwidth parameter h is performed using a

Algorithm 1: Feature selection algorithm.

Input:
Y Set of deformation fields
M' Maximum number of parameters to select
$\mathcal{E}$ Error threshold

Output:
X Selected features

$\gamma := (0 \cdots 0)^T$
Estimate parameter h
$J_0 := J(\gamma, \hat{h})$
$k := 1$

while $k \leq M'$ and $\bar{\mathcal{E}}(\gamma, h) \leq \mathcal{E}$ **do**
  $J_k := J_{k-1}$
  foreach *feature i such that* $\gamma(i) = 0$ **do**
    $\gamma(i) := 1$
    if $J(\gamma, \hat{h}) < J_k$ **then**
      $J_k := J(\gamma, \hat{h})$
      $i_k^{(+)} := i$
    end
    $\gamma(i) := 0$
  end
  if *one parameter to add is found* **then**
    $\gamma\left(i_k^{(+)}\right) := 1$
    Estimate parameter h
    $k := k + 1$
  else
    Stop the loop
  end
end

$X := \text{diag}(\gamma) \cdot Y$

---

gradient descent technique within a leave-one-out strategy (Wand and Jones, 1995). The estimation of the activation vector γ is obtained using a greedy algorithm (Pudil et al., 1994) with a complexity of $\mathcal{O}(M'M)$ cost function evaluations in the worst case, also within a leave-one-out strategy. The vector γ is first initialized to the null vector. Then, M' features at most are added iteratively. At each step, the feature added is the feature that maximizes the decrease of the reconstruction error. This procedure is detailed in Algorithm 1.

The following convergence criterion have been defined for the proposed algorithm:

1. The maximal number of parameter to select M' is reached;
2. No parameter can be added without increasing the cost function;
3. The reconstruction error is lower than a predefined error threshold $\mathcal{E}$.

Eq. (6) provides the global reconstruction error for which it could be difficult to set a convergence criterion. Therefore, we introduce the mean reconstruction error of parameters expressed in *mm* as:

$$\bar{\mathcal{E}}(\gamma, h) = \frac{1}{NM} \sum_{j=1}^{N} \sum_{i=1}^{M} \left\| Y_{i,j} - \hat{f}(\text{diag}(\gamma) \cdot Y_j, h)_i \right\|_2 , \quad (7)$$

where $\hat{f}(\text{diag}(\gamma) \cdot Y_j, h)_i$ is the reconstruction of the ith displacement vector of the jth subject. This last equation provides the average error in *mm* considering one parameter. Eq. (6) is used for the optimization process, whereas Eq. (7) is used as a convergence criterion only.

(a) 0 % noise (b) 5 % noise (c) 10 % noise (d) 25 % noise
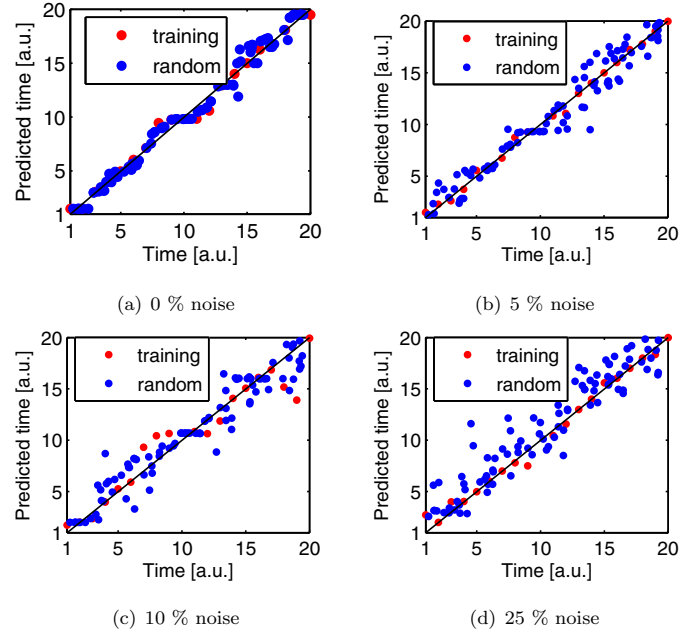
Fig. 7. Assessment of the proposed method to predict meta-data with synthetic dataset. The plots show time labels against predicted ones for the training dataset and a randomly generated dataset. On average, the prediction error is about 1 time unit. Even with 25 % of noise, the prediction error for randomly generated dataset is limited (1.67 ± 1.36 time unit). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Once a discriminative subset of features is sparsely selected, a quick and natural way to predict meta-data variables, such as gestational age for instance, is to take advantage of the weighted mean formulation of the reconstruction function (see Eq. (3)), replacing Y_j by meta-data variable v_j corresponding to subject j:

$$\hat{v}(X, h) = \sum_{j=1}^{N} w_j(X, h)v_j , \quad (8)$$

where X is the selected features of an eventually new subject and h is estimated as previously.

3. Results

Experiments have been conducted on both synthetic and real datasets. In the first part, the behavior and the performance of the algorithm are studied on synthetic datasets. Then, in a second part, the proposed method is applied on in vivo fetal brain dataset.

In each experiment, a Gaussian kernel is used within the reconstruction function (see Eq. (3)) and the maximum number of parameter is set to $M' = 50$. The mean parameter error threshold is set to $\mathcal{E} = 0.1$ mm for experiments on brain MR images. For synthetic experiment only, this threshold is set to 0 in order to study the convergence of the algorithm. Each cost function plot shows the full cost function up to M' iterations with a blue plain line and the iteration reached by the algorithm with the stopping criterion with a vertical dashed red line. Meta-data predictions (time and gestational age for synthetic and real datasets respectively) are provided for both training (initial) and randomly generated datasets.

3.1. Synthetic dataset experiments

The proposed algorithm has been tested on a synthetic dataset simulating the simultaneous growth of two sulci with non-linear dynamics. The images of the dataset have been processed in the
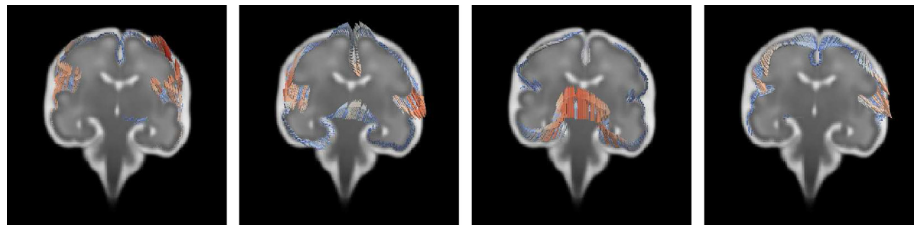
Fig. 8. Four randomly chosen deformation fields of the dataset displayed in the reference space. The color code corresponds to the magnitude of the vectors from blue (low magnitude) to red (high magnitude). The selection algorithm is applied on the cortical deformation fields mapping the reference space to population. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(a) $h = 0.5$ (b) $h = 1$ (c) $h = 1.5$ (d) $h = 2$ (e) $h = 3$
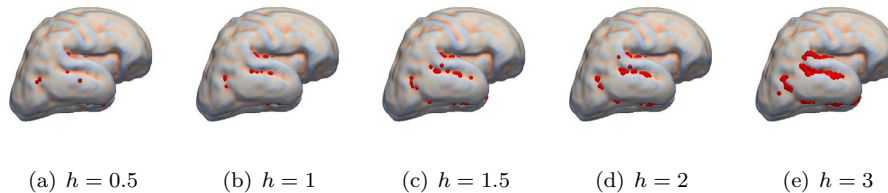
Fig. 9. Study of the influence of the choice of the bandwidth parameter of the proposed algorithm. As the bandwidth increases, the number of selected parameters increases as well. An important property is the location stability of the solution with various bandwidth values. For display purpose, the selected parameters are depicted on a surface reconstruction of the reference space.

same way as the fetal brain dataset (see the procedure described in Section 2.3). All the fitting procedures have been performed within MatLab's curve fitting application (non-linear least square function, with the Levenberg-Marquardt algorithm and the LAR robustness method).

3.1.1. Data generation

The synthetic dataset is composed of 20 two-dimensional images of size 128×128 pixels simulating the formation of two sulci (see Fig. 2). The first image of this dataset represents a flat cortical plate while the last one represents two sulci with different depth. The shapes have been created using a time-varying B-Spline curve to simulate the folding cortical surface (see Fig. 3). The growth dynamics have been modeled by a sigmoid function $x \mapsto \frac{A}{1+e^{-\lambda(x-T)}}$ with parameters (A, λ, T) equal to (0.7, 0.8, 7.0) for the first sulcus (right on Fig. 2 and blue curve on Fig. 3) and (1.0, 1.0, 14.0) for the second sulcus (left on Fig. 2 and red curve on Fig. 3).

In order to simulate a more realistic dataset, additive white Gaussian noise has been added to images. The noise level is described by the ratio between the standard deviation of the noise and the standard deviation of the image and is expressed as a percentage. Within this work, 5%, 10% and 25% of noise have been added to the clean dataset.

3.1.2. Selection

The results of the proposed algorithm on the synthetic dataset are depicted in Fig. 4. The algorithm allows to select only two points, which is consistent with the simulated dataset (two sulci). Moreover, as expected, the selected points are located inside the artificial sulci. As shown on the figure, the locations of the selected points are quite stable across noise level.

The evolution of cost functions are plotted in Fig. 5. Convergence is reached when only two features have been selected. Indeed, the cost functions decrease dramatically with the selection of the first features and are almost constant after the second iteration. The evolutions of the cost functions suggest that two selected points capture the main variability of the dataset.

While the locations of the selected are barely affected by the noise level, the cost function exhibits a smoother decrease as the noise level increases. It can be explained by the reconstruction error minimization used by the algorithm (see Section 2.4): the noisier the dataset is, the more points the proposed algorithm needs
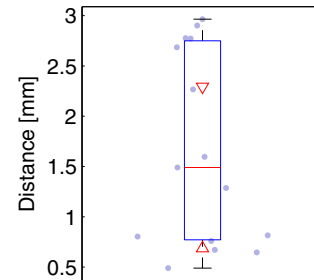
Fig. 10. Mean distances between clouds of selected points for several bandwidths (see Fig. 9). For two sets of selected points, the mean distance is computed as the mean distance of each point in the first set to the closest point in the second set. Since the measured error is about 1.5 mm on average (i.e. one voxel and a half), this quantification shows the anatomical stability of the selection with different bandwidths.

to select in order to achieve a proper reconstruction of the dataset. Moreover, it is interesting to note that the selected points are not necessarily located at the deepest points of the artificial sulci.

The residual errors (displayed in Fig. 6) decrease dramatically after the selection of the first and the second parameter and are mainly located inside the simulated sulci. Each selected parameter allows to decrease the residuals in its area of influence. As the noise level increases, the residuals increase as well.

3.1.3. Temporal pattern analysis

In this section, we study the ability of the proposed approach to estimate growth dynamic based on the selected points.

Sulci formation has been simulated using a sigmoid growth model using three parameter (A, λ, T). A is related to the maximum amplitude of each sulcus and then depends on selected point location. However, the dynamic of the growth change can be capture by estimating λ and T. The first sulcus has been simulated using $\lambda = 0.8$ and $T = 7$. Using a fitting procedure based on the extracted points, the estimated parameters (lambda,T) for noise level 0 %, 5 %, 10 % and 25 % are respectively (0.80, 6.80), (0.90, 6.64), (0.82, 6.38) and (0.90, 7.86). Concerning the second sulcus simulated using $\lambda = 1$ and $T = 14$, the estimated parameters (λ, T) for noise level 0 %, 5 %, 10 % and 25 % are respectively (1.27, 13.15), (1.80, 11.66), (1.59, 11.02) and (2.05, 11.64). In order to assess the
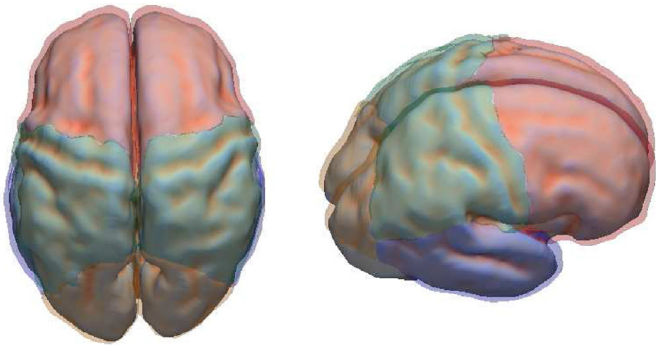
Fig. 11. The regions of interest used with the feature selection algorithm and depicted on a surface reconstruction of the reference space. These regions corresponds approximately to the frontal (red), temporal (blue), parietal (green) and occipital (yellow) lobes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5 %, 10 % and 25 % are respectively 0.40 ± 0.48 (0.58 ± 0.40), 0.27 ± 0.26 (1.09 ± 1.03), 1.01 ± 1.35 (1.16 ± 1.04) and 0.35 ± 0.53 (1.67 ± 1.36). On average, the prediction error is about 1 time unit. As noise level increases, the prediction error increases as well but is limited: even with 25 % of noise, the prediction error for randomly generated dataset is smaller than 2 time unit.

3.2. Application on fetal brain MRI dataset

Since our first objective is to study the geometrical changes of the cortical folding through in-utero brain maturation, we define as region of interest the cortical gray matter. Moreover, average deformation fields (computed in the orthogonal direction of brain surface) along the cortical gray matter have been used in order to reduce the initial parameter set to a smaller set of 3D displacement vectors (Rajagopalan et al., 2012). Fig. 8 shows five examples of deformation fields considered in this study.

3.2.1. Influence of the bandwidth parameter h

In this experiment, we evaluate the influence of the bandwidth parameter h over the selected points. Here, we applied the selection algorithm by setting manually this parameter $h \in \{0.5, 1, 1.5, 2, 3\}$. The results obtained applying the algorithm on the temporal lobe for different values for h are displayed in Fig. 9. First, it

quality of these estimates, we used the selected points for prediction.

Prediction error of time labels using Eq. (8) for the training and a randomly generated datasets are plotted in Fig. 7. The mean prediction errors of the training (random) dataset for noise level 0 %,
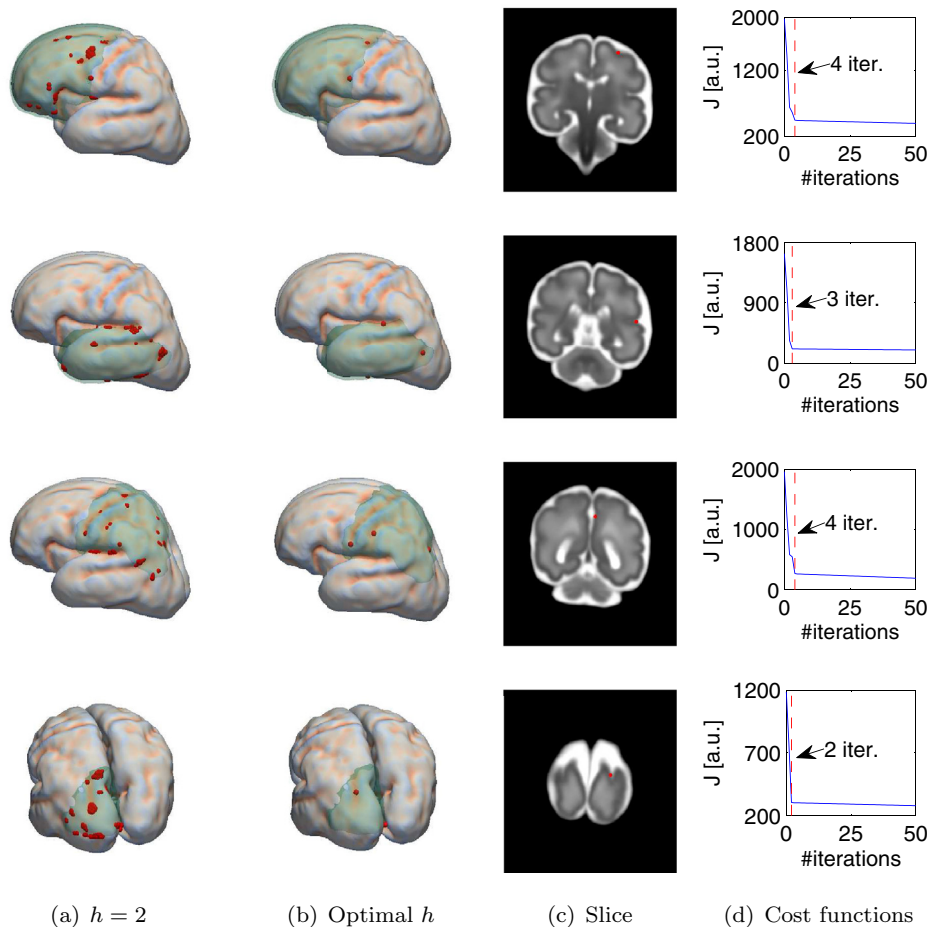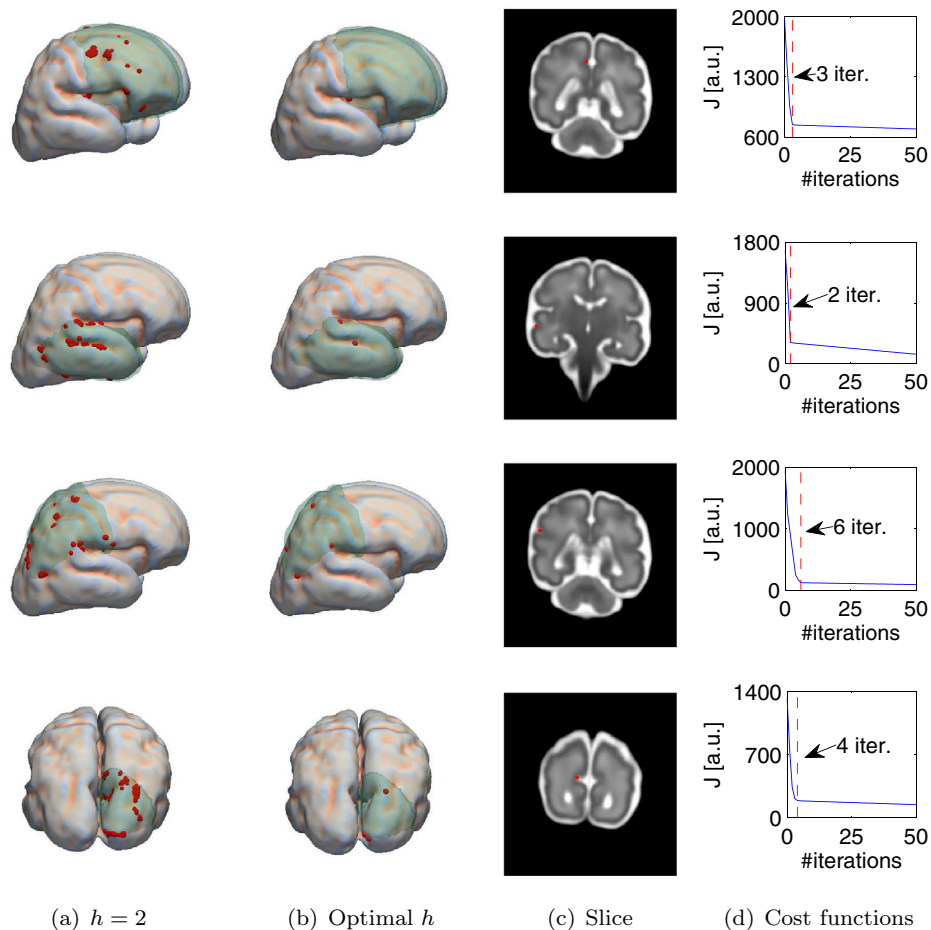

(a) $h = 2$ (b) Optimal h (c) Slice   (d) Cost functions

Fig. 12. Selection of cortical points on four lobe regions of the left hemisphere using the proposed algorithm depicted on a surface reconstruction of the reference space (for display purpose only). The left column shows the solution for a fixed bandwidth ($h = 2$). The other columns show the results for the optimal bandwidth. Top row: frontal lobe; second row: temporal lobe; third row: parietal lobe; last row: occipital lobe. The selected cortical points are depicted as red spheres. The color code in 3D mesh views maps the curvature of the cortical plate from blue color (positive curvature) to red color (negative curvature). The red dashed line in (d) locates the number of iterations reached by the algorithm when using stopping criterion (see Section 2.4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(a) $h = 2$ (b) Optimal h (c) Slice   (d) Cost functions

Fig. 13. Selection of cortical points of four lobe regions of the right hemisphere using the proposed algorithm depicted on a surface reconstruction of the reference space (for display purpose only). The left column shows the solution for a fixed bandwidth ($h = 2$). The other columns show the results for the optimal bandwidth. Top row: frontal lobe; second row: temporal lobe; third row: parietal lobe; last row: occipital lobe. The selected cortical points are depicted as red spheres. The color code in 3D mesh views maps the curvature of the cortical plate from blue color (positive curvature) to red color (negative curvature). The red dashed line in (d) locates the number of iterations reached by the algorithm when using stopping criterion (see Section 2.4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

can be easily seen that increasing the value of h makes the number of selected points increase. Please note that h is a parameter that controls the smoothness of the non parametric regression and therefore the value of h is related to the sparsity of the solution.

Second, this experiment also shows the anatomical stability of the selected features. For instance, see the distance error between sets of selected points for different bandwidths in Fig. 10. This error is computed as the mean distance of each point in the first set to the closest point in the second set. It is used as an indicator of anatomical stability of selected points across bandwidths. On average, this error is about 1.5 mm. Therefore, the average error between selected points across bandwidths is only about two pixels. It appears that the selected points are located in the main sulci (superior temporal sulcus, inferior temporal sulcus, lateral sulcus) in this region of interest, appearing during the considered period of age. This is a key point since it means that modifications of the bandwidth only influence the number of selected points, not their anatomical locations (i.e. the estimated growth pattern). In other words, while the value of the bandwidth h modifies the sparsity of the solution, it does not modify the spatial pattern of the most discriminative selected points.

Third, the experiment, restricted on the temporal lobe area, shows that growth variability can be captured by few points in the sulci of interest.

Please note that in the proposed algorithm, the bandwidth parameter h is estimated at each iteration using a leave-one-

out strategy in order to minimize the variance of the regression function. For the next experiments, h is then computed automatically.

3.2.2. Points selection on fetal brain lobes

In this section, we propose to apply the feature selection algorithm onto 8 regions of interest (ROI) corresponding approximately to the frontal, temporal, parietal and occipital lobes of each hemisphere (see Fig. 11) in order to characterize spatial pattern of cortical folding for the considered period of age (26 to 34 weeks). The purpose is 1) to reduce the computation time of the algorithm by pre-selecting ROI (common strategy for feature selection techniques) and 2) to provide statistics on the selected set of cortical points for each lobe separately. The cortical parcellation of the template has been obtained using a patch-based segmentation technique (Rousseau et al., 2011, 2013). Each of these ROI have been considered independently.

Experiments previously described have shown that the selected points tend to be in sulci valleys. Thus, the ROI defined here do not exactly match the anatomical lobes to avoid any bias of selected sulci. For instance, the central sulcus should define the boundary between the frontal and parietal lobes. However, we chose to associate the central sulcus with the parietal lobe, in order to give a chance to the algorithm to select cortical points of this sulcus.

The results of the selection algorithm on brain lobes are displayed in Figs. 12 and 13 respectively for lobes in left and right
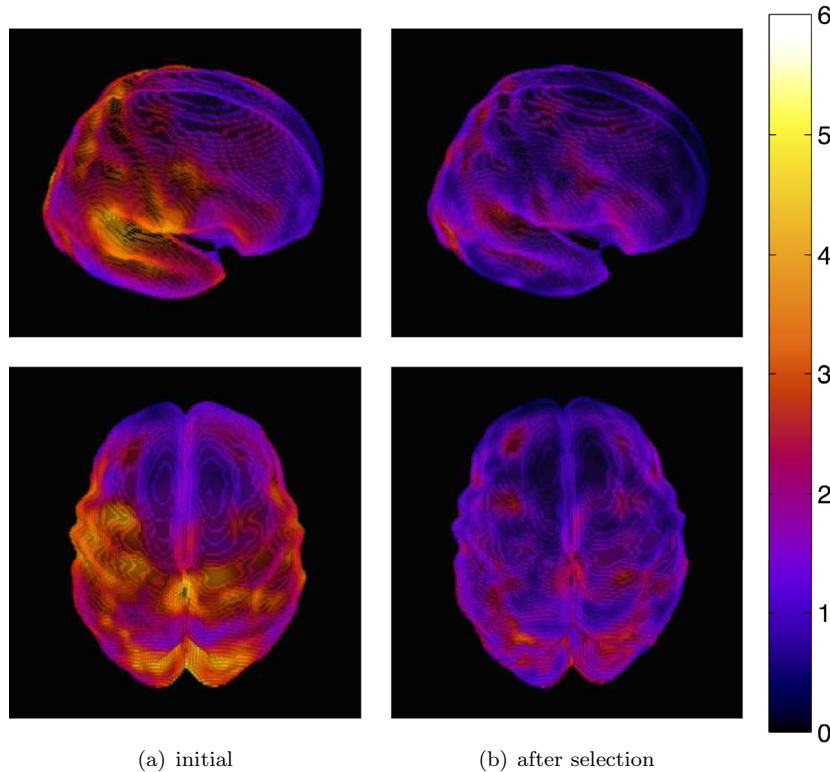
Fig. 14. Residuals of the cost function (6) before and after selection with automatic stopping criterion projected on population's average space. All lobes are displayed at the same time. This reconstruction experiment shows that within a few selected features (see Figs. 12 and 13), most of the variability of the dataset is captured.

hemispheres. For each considered ROI, four types of results are provided : (a) selected points when the bandwidth h is set to 2, (b) selected points when h is automatically estimated, (c) a coronal view showing the location of selected points and (d) the evolution of the cost function. While the first visualization (using $h = 2$) provides the main patterns of selected features, the second visualization (using automatic bandwidth computation) shows the sparsity of the reached solution of the algorithm. Only few features (up to 6) are required to discriminate the sets of images for each ROI. The coronal view confirms previous experiments (see Section 3.2.1): the selected points are located into sulci valleys.

The residuals of the reconstruction error (6) before ($M' = 0$) and after automatic selection (M' is found automatically using the stopping criterion described in Section 2.4) projected on population's average space are displayed in Fig. 14. While the initial residuals highlight regions with high variability across the population, the residuals after selection exhibit in comparison the ability of the proposed method to capture most of variability within a few selected features (less than 6 for each lobe).

As described in Section 2.4, it is possible to predict meta-data variables with Eq. (8) once a subset of cortical points has been selected with the proposed method. Predictions have been made for fetal brain dataset using this equation with a leave-one-out scheme in order to prevent overfitting. Fig. 15 shows the gestational age prediction errors for the lobe regions used independently as ROI within the proposed method. The average error for frontal, temporal, parietal and occipital left (respectively right) lobe regions are 0.74, 0.90, 0.75 and 1.10 (respectively 0.84, 0.87, 0.82 and 0.91) weeks. By merging all this predictions, the global prediction error is on average about 0.83 weeks.

3.2.3. Evolution of deformation vector magnitude

We examine in this section the temporal evolution of magnitudes of selected deformation vectors. Once the most discriminant

points are selected for each region of interest, we compute the set $\mathcal{V}$ of deformation fields between each subject and the mean image estimated at 26 weeks: $\mathcal{V} = \{\vec{v}_{\mathbf{x}}(t_i)\}$ where $\mathbf{x}$ is the location of each selected point and t_i is the time difference between the age of the subject i and 26 weeks. The evolution of the magnitudes of computed deformation fields at selected point locations is estimated using a temporal polynomial of degree 3 regression approach (see Fig. 16 for an instance of regression over time for one vector). We chose a polynomial regression model for the growth measurements because we want to make our quantifications comparable with previous work that often used such methods (see for instance the work of Habas et al. (2010a) and Habas et al. (2012)). Fig. 17 shows (a) the time evolution of the mean magnitude for the eight considered ROI (corresponding to the left-right lobes), (b) the mean magnitude for the four ROI by fusing data from left and right hemispheres, and (c) the mean magnitude by considering only left vs right. It can be seen that mean magnitude at occipital lobe has a different temporal evolution from the three other anatomical regions. Moreover, temporal differences in left and right temporal evolution are also observed. Statistical testing on larger dataset (i.e. using more fetal brain MRI images) has to be conducted in order to verify the significance of these observations.

4. Discussion

In this paper, we have shown that early cortical folding patterns occurring during fetal brain development can be expressed by a sparse representation using a discriminative feature selection approach. By using sparse non linear deformation fields, we have highlighted major cortical surface changes in the considered age interval. Such information extraction is important since sulci can be used to discriminate brain development stages. The period of age considered in this work (26 to 34 weeks) corresponds to a linear evolution of the surface curvature observed
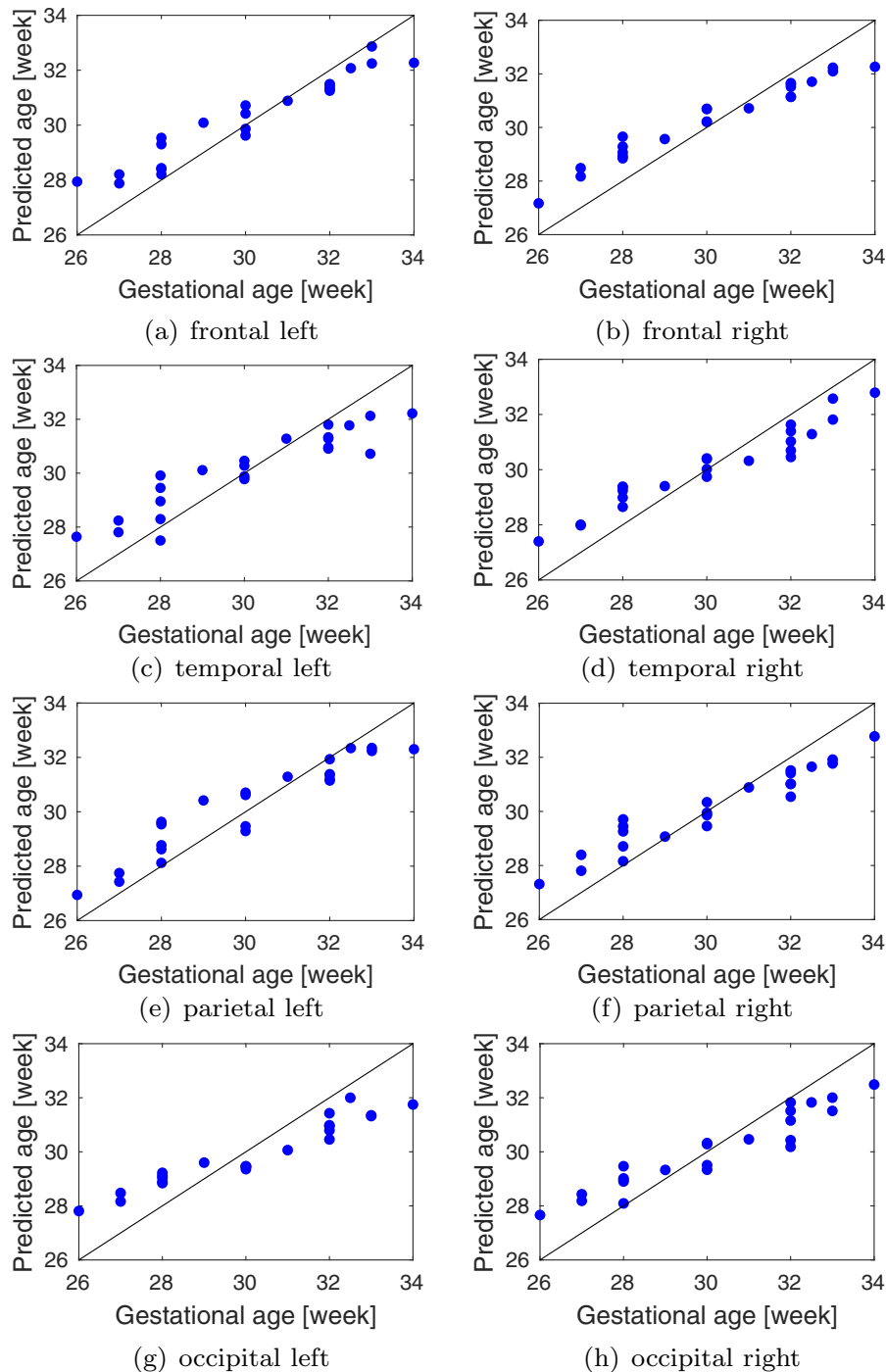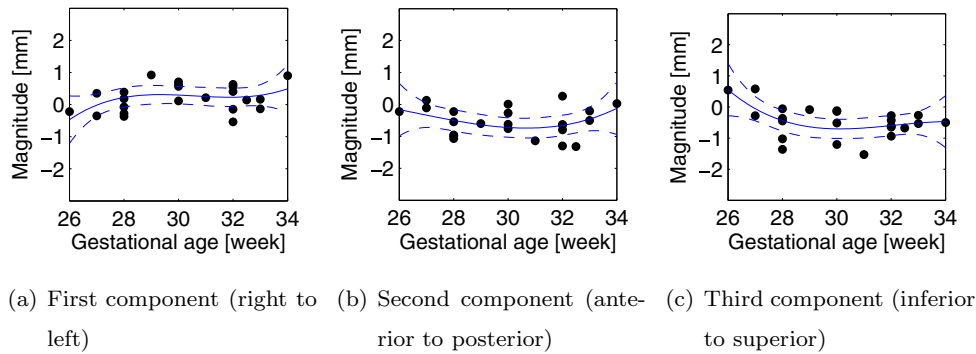
Fig. 15. Gestational age prediction error for each lobe region. The average error for frontal, temporal, parietal and occipital left (respectively right) lobe regions are 0.74, 0.90, 0.75 and 1.10 (respectively 0.84, 0.87, 0.82 and 0.91) weeks. By merging all this predictions, the global prediction error is on average about 0.83 weeks.

by Wright et al. (2014) (middle part of the sigmoid model used by Wright et al.)). Our experiments performed on mean magnitudes of deformation vectors tend to show the same trend.

In this work, we have considered non linear deformation fields providing voxel correspondence between each subject. Feature selection technique has been applied on these dense deformation fields in order to compute the smallest discriminative set of features characterizing the cortical folding process during fetal brain development. We found sparse fields by minimizing a reconstruction error subject to L0 norm regularization. An approximate solution has been estimated with a greedy algorithm. An alterna-

tive approach could have been to use an other sparsity inspired variable selection techniques like Lasso or Elastic-net methods. The main difference is the norm used in the regularization: L1 for the Lasso and weighted sum of L1 and L2 for the Elastic-net. While L1 and L2 norms take into account magnitudes of components, L0 norm does not. Indeed, L0 norm relates only the number of null vector, L1 norm penalizes vectors with non-null magnitudes and L2 norm penalizes mostly vectors with high magnitude. Often, L1 norm is regarded as a good approximation of L0 norm. Because of its convexity, the optimization problem would be easier to solve without resort to heuristic algorithm. A further work would be to

(a) First component (right to left)

(b) Second component (anterior to posterior)

(c) Third component (inferior to superior)

Fig. 16. Example of polynomial regression over time (degree 3) for a randomly selected vector from the right temporal lobe. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(a) All lobes

(b) 4 lobes

(c) Hemispheres
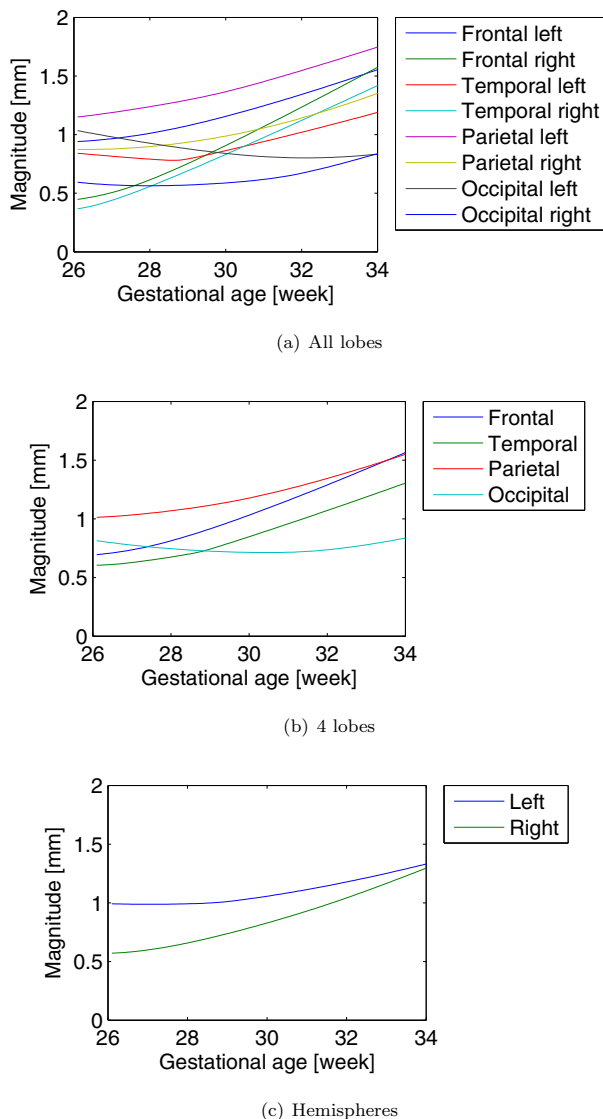
Fig. 17. Mean magnitude of selected deformation vectors of lobes over time.

compare the accuracy of the solutions provided by Lasso technique and the proposed method.

In order to be able to select features, the proposed method need to normalize the brain population in a common reference space. Usually, average space of the brain population is used as reference space. This normalization raises the question of how it will impact the selection output. In this work, we chose an it-

erative and unbiased method that is widely used in the community (Guimond et al., 2000). The principle is to initialize the normalization procedure with an empirically chosen image among the population and then correct for the bias iteratively by taking the average of deformations, until convergence of the algorithm to the average space of the population. The proposed method relies only on the deformation fields, mapping non-linearly a structure (e.g. sulci) in one subject to the same structure in the common reference space. Therefore, the output solution is dependent on the quality of the registration process that occur during the brain normalization.

It is a standard approach in morphometry to define an unbiased average anatomy as reference space. However, the specific case of fetal brain development analysis may imply to use alternative strategies. Indeed, thinking the brain development as a growing process with a beginning and an end, an average space might not be appropriate. A natural alternative would be to use for instance a smooth surface as the beginning of the developmental process and an adult brain as the end.

In this study, the registration step and the feature selection are independent. An alternative approach could have been to develop a registration technique incorporating a sparsity prior on the deformation fields. This kind of generative approach is related to the work of Durrleman et al. (2012) in which a new parametrization of deformations is proposed in order to capture the variability in image ensembles. The key idea in their work is to select the most relevant control points and to estimate their positions in a template domain. A major difference between these two works is that control points may not belong to the shape under study. Thus, measuring the variability through the distribution of control points may not be easily interpretable with respect to the studied temporal process. By separating the registration step and the shape variability study, we can focus on shape deformation locally related to cortical folding.

Our work is also related to the recent study of growth seeds performed by Lefèvre et al. (2009), where an analysis of depth maps of cortical surface based on the Helmholtz decomposition of the deformation fields. In the work of Lefèvre et al. (2009), which is also related to the works of Cuzol et al. (2005) and Grenander et al. (2007), the main idea relies in that a decomposition is used to make the physical interpretation of the deformation fields (capturing the brain maturation process) easier. Such decomposition has been recently investigated to discover critical regions of volume changes in Alzheimer disease (Lorenzi et al., 2015). Our approach is complementary to these techniques in the sense that the relevant (i.e. most discriminative) displacement vectors are selected using a non-parametric approach. Further work would consist in analyzing the impact of the deformation model on the selected features.

Due to the type of acquired data (i.e. fetal MR images of several subjects at different time points), the global variability of the dataset comes from both cross-subject and longitudinal variabilities. A strategy to disambiguate this limitation is to acquired longitudinal dataset (several subjects at multiple and different time points) and use an appropriate longitudinal trajectory estimation algorithm (see for instance the work of Hart et al., 2010 or Durrleman et al., 2009). So far, there is actually no longitudinal studies on fetal brain development with MRI data due to the difficulty of such data acquisition. In our work, despite we are not able to clearly differentiate both variabilities, we took the advantage of the discriminative property of the proposed method. Indeed, it will enforce the selected features to be the most different across the population, and therefore to locate anatomical areas with the most changing pattern across the population.

As shown in our experiments, the most discriminative points lie in deep sulcal regions. These specific locations could be used as consistent shape features (or biomarkers) in further studies. Future work will consist in analyzing these patterns with regard to previous related works on sulcal root (Régis et al., 2005) and sulcal pit (Lohmann et al., 2008). It has been shown that deepest parts of sulci generally show less inter-individual variability than superficial parts (Lohmann et al., 2008). The study of local fluctuations of these cortical patterns using a clustering approach such as in (Sun et al., 2009) could provide complementary insights on brain folding to gyrification index measurements (Clouchoux et al., 2011). The spatial distribution of deep sulcal landmarks (Im et al., 2010), especially over time (Meng et al., 2014), may help to set up new early markers of later functional development (Dubois et al., 2008) or particular cortical malformation (such as polymicrogyria) that might be difficult to detect and quantify visually. In this context, the proposed approach could be adapted to surface-based methodology that have been recently applied for the study of fetal cortex (Lefèvre et al., 2015; Auzias et al., 2015).

The stability of sparse solutions is a critical point within variable selection algorithms, especially when the number of features is greater than the number of subjects and when features exhibit strong correlations. Being related to the general issue of selecting a proper amount of regularization, this stability problem can be tackled using for instance a generic subsampling approach (Meinshausen and Buhlmann, 2010). In this work, we chose to use a greedy algorithm instead of a regularized estimator. Moreover, we chose to focus on anatomical stability, i.e. if the selected points are located in a same narrow area where changes across population take place. On one hand, the method is not stable in the sense that we won't have exactly the same set of selected features depending on the parameter h. On the other hand, the selected features are located in a relatively narrow anatomical area and are raised by the same phenomenon (apparition of the temporal sulcus). This effect sounds because the proposed method is seeking discriminative points. The voxel stability is related to sulcal roots and pits and could be a topic of further work.

5. Conclusion

In this paper, we have used a feature selection approach to characterize cortical folding patterns occurring during fetal brain development with a sparse representation.

The use of such a feature selection technique is not restricted to only fetal brain development studies. It could be used to define biomarkers, shape representation, diffusion / fiber tracts, etc.

Acknowledgment

References

Aljabar, P., Wolz, R., Srinivasan, L., Counsell, S., Rutherford, M., Edwards, A., Hajnal, J., Rueckert, D., 2011. A combined manifold learning analysis of shape and appearance to characterize neonatal brain development. IEEE Trans. Med. Imaging 30 (12), 2072–2086.

Auzias, G., De Guio, F., Pepe, A., Rousseau, F., Mangin, J.-F., Girard, N., Lefevre, J., Coulon, O., 2015. Model-driven parameterization of fetal cortical surfaces. HAL.

Avants, B., Gee, J., 2004. Geodesic estimation for large deformation anatomical shape averaging and interpolation. NeuroImage 23, 139–150 supplment 1.

Busse, R.F., Riederer, S.J., Fletcher, J.G., Bharucha, A.E., Brandt, K.R., 2000. Interactive fast spin-echo imaging.. Magn. Reson. Med. 44 (3), 339–348.

Caldairou, B., Passat, N., Habas, P.A., Studholme, C., Koob, M., Dietemann, J.-L., Rousseau, F., 2011. Segmentation of the cortex in fetal MRI using a topological model. In: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, pp. 2045–2048.

Clouchoux, C., Kudelski, D., Gholipour, A., Warfield, S.K., Viseur, S., Bouyssi-Kobar, M., Mari, J.-L., Evans, A.C., du Plessis, A.J., Limperopoulos, C., 2011. Quantitative in vivo mri measurement of cortical development in the fetus. Brain Struct. Funct. 217 (1), 127–139.

Dittrich, E., Riklin Raviv, T., Kasprian, G., Donner, R., Brugger, P.C., Prayer, D., Langs, G., 2014. A spatio-temporal latent atlas for semi-supervised learning of fetal brain segmentations and morphological age estimation.. Med. Image Anal. 18 (1), 9–21.

Dubois, J., Benders, M., Borradori-Tolsa, C., Cachia, A., Lazeyras, F., Ha-Vinh Leuchter, R., Sizonenko, S.V., Warfield, S.K., Mangin, J.F., Hüppi, P.S., 2008. Primary cortical folding in the human newborn: an early marker of later functional development. Brain 131 (8), 2028–2041.

Durrleman, S., Allassonnière, S., Joshi, S., 2012. Sparse adaptive parameterization of variability in image ensembles. Int. J. Comput. Vision 101 (1), 161–183.

Durrleman, S., Pennec, X., Trouv, A., Gerig, G., Ayache, N., 2009. Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009. Lecture Notes in Computer Science, Vol. 5761. Springer, pp. 297–304.

Elad, M., 2010. Sparse and redundant representations: from theory to applications in signal and image processing. Springer.

Gerber, S., Tasdizen, T., Fletcher, P., Joshi, S., Whitaker, R., 2010. Manifold modeling for brain population analysis. Med. Image Anal. 14 (5), 64–653.

Gholipour, A., Akhondi-Asl, A., Estroff, J.A., Warfield, S.K., 2012. Multi-atlas multi-shape segmentation of fetal brain MRi for volumetric and morphometric analysis of ventriculomegaly. NeuroImage 60 (3), 1819–1831.

Gholipour, A., Estroff, J.A., Warfield, S.K., 2010. Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI.. IEEE Trans. Med. Imaging 29 (10), 1739–1758.

de Graaf-Peters, V.B., Hadders-Algra, M., 2006. Ontogeny of the human central nervous system: what is happening when? Early Hum. Dev. 82 (4), 257–266.

Grenander, U., Srivastava, A., Saini, S., 2007. A pattern-theoretic characterization of biological growth. IEEE Trans. Med. Imaging 26 (5), 648–659.

Guimond, A., Meunier, J., Thirion, J.-P., 2000. Average brain models: a convergence study. Comput. Vision Image Underst. 77 (2), 192–210.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182.

Habas, P., Kim, K., Corbett-Detig, J., Rousseau, F., Glenn, O., Barkovich, A., Studholme, C., 2010a. A spatiotemporal atlas of MR intensity, tissue probability and shape of the fetal brain with application to segmentation. NeuroImage 53 (2), 460–470.

Habas, P.A., Kim, K., Rousseau, F., Glenn, O.A., Barkovich, A.J., Studholme, C., 2010b. Atlas-based segmentation of developing tissues in the human brain with quantitative validation in young fetuses. Hum. Brain Mapp. 31 (9), 1348–1358.

Habas, P.A., Scott, J.A., Roosta, A., Rajagopalan, V., Kim, K., Rousseau, F., Barkovich, A.J., Glenn, O.A., Studholme, C., 2012. Early folding patterns and asymmetries of the normal human brain detected from in utero MRI. Cereb. Cortex 22 (1), 13–25.

Hart, G., Shi, Y., Zhu, H., Sanchez, M., Styner, M., Niethammer, M., 2010. DTI longitudinal atlas construction as an average of growth models. MICCAI, International Workshop on Spatio-Temporal Image Analysis for Longitudinal and Time-Series Image Data.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning : Data Mining, Inference, and Prediction. Springer.

Im, K., Jo, H.J., Mangin, J.F., Evans, A.C., Kim, S.I., Lee, J.M., 2010. Spatial distribution of deep sulcal landmarks and hemispherical asymmetry on the cortical surface. Cereb. Cortex 20 (3), 602–611.

Jiang, S., Xue, H., Glover, A., Rutherford, M., Rueckert, D., Hajnal, J.V., 2007. MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): application to fetal, neonatal, and adult brain studies.. IEEE Trans. Med. Imaging 26 (7), 967–980.

Kim, K., Habas, P.A., Rousseau, F., Glenn, O.A., Barkovich, A.J., Studholme, C., 2010. Intersection based motion correction of multislice MRI for 3d in utero fetal brain image formation. IEEE Trans. Med. Imaging 29 (1), 146–158.

Kuklisova-Murgasova, M., Quaghebeur, G., Rutherford, M.A., Hajnal, J.V., Schnabel, J.A., 2012. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal.. Med. Image Anal. 16 (8), 1550–1564.

Lefèvre, J., Germanaud, D., Dubois, J., Rousseau, F., de Macedo Santos, I., Angleys, H., Mangin, J.-F., Hüppi, P.S., Girard, N., De Guio, F., 2015. Are developmental trajectories of cortical folding comparable between cross-sectional datasets of fetuses and preterm newborns? Cereb. Cortex 1–13.

Limperopoulos, C., Clouchoux, C., 2009. Advancing fetal brain MRI: targets for the future.. Sem. Perinatol. 33 (4), 289–298.

Lohmann, G., von Cramon, D.Y., Colchester, A.C.F., 2008. Deep sulcal landmarks provide an organizing framework for human cortical folding. Cereb. Cortex 18 (6), 1415–1420.

Lorenzi, M., Ayache, N., Pennec, X., Initiative, A.D.N., 2015. Regional flux analysis for discovering and quantifying anatomical changes: an application to the brain morphometry in Alzheimer's disease. NeuroImage 115 (15), 224–234.

Meinshausen, N., Buhlmann, P., 2010. Stability selection. J. R. Stat. Soc. Ser. B (Stat. Meth.) 72 (4), 417–473.

Meng, Y., Li, G., Lin, W., Gilmore, J.H., Shen, D., 2014. Spatial distribution and longitudinal development of deep cortical sulcal landmarks in infants. NeuroImage 100 (C), 206–218.

Nadaraya, E., 1964. On estimating regression. Theor Probab. Appl. 10, 186–190.

Pudil, P., Novovicova, J., Kittler, J., 1994. Floating search methods in feature selection. Pattern Recognit. Lett. 15 (11), 1119–1125.

Rajagopalan, V., Scott, J., Habas, P., Kim, K., Rousseau, F., Glenn, O., B., A., Studholme, C., 2012. Mapping directionality specific volume changes using tensor based morphometry: an application to the study of gyrogenesis and lateralization of the human fetal brain. NeuroImage 63 (2), 947–958.

Rajagopalan, V., Scott, J.A., Habas, P.A., Kim, K., Corbett-Detig, J.M., Rousseau, F., Barkovich, A.J., Glenn, O.A., Studholme, C., 2011. Local tissue growth patterns underlying normal fetal human brain gyrification quantified in utero. J. Neurosci. 31 (8), 2878–2887.

Régis, J., Mangin, J.-F., Ochiai, T., Frouin, V., Riviére, D., Cachia, A., Tamura, M., Samson, Y., 2005. "Sulcal root" generic model: a hypothesis to overcome the variability of the human cortex folding patterns. Neurol. Med. Chir. 45 (1), 1–17.

Rousseau, F., Glenn, O.A., Iordanova, B., Rodriguez-Carranza, C., Vigneron, D., Barkovich, A.J., Studholme, C., 2005. A novel approach to high resolution fetal brain MR imaging.. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: MICCAI, Pt 1, 8, pp. 548–555.

Rousseau, F., Glenn, O.A., Iordanova, B., Rodriguez-Carranza, C., Vigneron, D.B., Barkovich, A.J., Studholme, C., 2006. Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. Acad. Radiol. 13 (9), 1072–1081.

Rousseau, F., Habas, P., Studholme, C., 2011. A supervised patch-based approach for human brain labeling. IEEE Trans. Med. Imaging 30 (10), 1852–1862.

Rousseau, F., Kim, K., Studholme, C., Koob, M., Dietemann, J.-L., 2010. On super-resolution for fetal brain MRI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: MICCAI, Pt 2, 13, pp. 355–362.

Rousseau, F., Oubel, E., Pontabry, J., Schweitzer, M., Studholme, C., Koob, M., Dietemann, J.-L., 2013. BTK: an open-source toolkit for fetal brain MR image processing. Comput. Methods Programs Biomed. 109 (1), 65–73.

Studholme, C., 2011. Mapping fetal brain development in utero using magnetic resonance imaging: the big bang of brain mapping. Annu. Rev. Biomed. Eng. 13 (1), 345–368.

Sun, Z.Y., Perrot, M., Tucholka, A., Riviére, D., Mangin, J.-F., 2009. Constructing a dictionary of human brain folding patterns.. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: MICCAI, Pt 2, 12, pp. 117–124.

Tenenbaum, J., de Silva, V., Langford, J., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290 (5500), 2319–2323.

Volpe, J.J., 2008. Neurology of the Newborn. Elsevier Health Sciences.

Wand, M., Jones, M., 1995. Kernel Smoothing. of Monographs on Statistics & Applied Probability, Vol. 60. Chapman & Hall.

Wright, R., Kyriakopoulou, V., Ledig, C., Rutherford, M.A., Hajnal, J.V., Rueckert, D., Aljabar, P., 2014. Automatic quantification of normal cortical folding patterns from fetal brain MRI.. NeuroImage 91, 21–32.

Yamashita, Y., Namimoto, T., Abe, Y., Takahashi, M., Iwamasa, J., Miyazaki, K., Okamura, H., 1997. MR imaging of the fetus by a haste sequence.. Am. J. Roentgenol. 168 (2), 513–519.

Further reading

Cuzol, A., Hellier, P., Mémin, E., 2005. A novel parametric method for non-rigid image registration. In: Information Processing in Medical Imaging : Proceedings of the conference, 19, pp. 456–467.

Lefèvre, J., Leroy, F., Khan, S., Dubois, J., Hüppi, P.S., Baillet, S., Mangin, J.-F., 2009. Identification of growth seeds in the neonate brain through surfacic Helmholtz decomposition. In: Information Processing in Medical Imaging 21st International Conference, IPMI 2009, 21. Williamsburg, VA, USA, pp. 252–263. doi:10.1007/978-3-642-02498-6_21.