HHS Public Access

Author manuscript

Proc IEEE Int Symp Biomed Imaging. Author manuscript; available in PMC 2020 January 14.

Published in final edited form as:

Proc IEEE Int Symp Biomed Imaging. 2019 April ; 2019: 396–399. doi:10.1109/ISBI.2019.8759429.

CORTICAL FOLDINGPRINTS FOR INFANT IDENTIFICATION

Dingna Duan[1,2], **Shunren Xia**[1], **Zhengwang Wu**[2], **Fan Wang**[2], **Li Wang**[2], **Weili Lin**[2], **John H Gilmore**[3], **Dinggang Shen**[2], **Gang Li**[2]

[1]Key Laboratory of Biomedical Engineering of Ministry of Education, Zhejiang University, China

[2]Department of Radiology and BRIC, University of North Carolina at Chapel Hill, USA

[3]Department of Psychiatry, University of North Carolina at Chapel Hill, USA

Abstract

Cortical folding of the adult brain is highly convoluted and encodes inter-subject variable characteristics. Recent studies suggest that it is useful for individual identification in adults. However, little is known about whether the infant cortical folding, which undergoes dynamic postnatal development, can be used for individual identification. To fill this gap, we propose to explore cortical folding patterns for infant subject identification. This study thus aims to address two important questions in neuroscience: 1) whether the infant cortical folding is unique for individual identification; and 2) considering the region-specific inter-subject variability, which cortical regions are more distinct and reliable for infant identification. To this end, we propose a novel discriminative descriptor of regional cortical folding based on multi-scale analysis of curvature maps via spherical wavelets, called *FoldingPrint*. Experiments are carried out on a large longitudinal dataset with 1,141 MRI scans from 472 infants. Despite the dramatic development in the first two years, successful identification of 1-year-olds and 2-year-olds using their neonatal cortical folding (with accuracy > 98%) indicates the effectiveness of the proposed method. Moreover, we reveal that regions with high identification accuracy and large inter-subject variability mainly distribute in high-order association cortices.

Keywords

cortical folding; infant; individual identification; multi-scale curvatures

1. INTRODUCTION

The human cerebral cortex is highly convoluted and exhibits remarkable inter-subject variability. Current studies have found that the adult brain cortex, when characterized by structural or functional features from MRI, is unique and reliable for individual identification [1, 2]. For example, a shape descriptor of brain morphology based on Laplace-Beltrami operator, called BrainPrint [1], was developed to characterize the cortical and subcortical structures and is effective for adult identification. Several other studies found the inter-individual variability of the functional connectivity and proposed the functional connectome fingerprinting for individual identification [2]. However, all existing individual identification studies based on the brain MRI are carried out on adult datasets, in which the

brain change is typically subtle between different scans. It is still an open question that whether the infant brain MRI is reliable for individual identification during infancy.

Compared to the adult identification studies, the infant identification studies are more challenging. The difficulties are mainly in two aspects: first, the infant cortex undergoes regionally-heterogeneously dynamic growth, especially in the first two years, which is the most dynamic phase of postnatal brain development [3]. Thus, the features developed for infant identification should be able to capture the unique individual-specific information, which is robust to the dynamic early brain development. The existing features for adult brain identification are not applicable for infant identification. For instance, BrainPrint [1], which is only appropriate for the simple isometric changes, cannot deal with the region-variable cortical expansion and the emerging tertiary cortical folds during infancy. Second, compared to the wide availabilities of adult MRI datasets and processing tools, large-scale longitudinal infant datasets and dedicated computational tools are rare and precious. This is because it is difficult to collect the quality motion-free infant MR images and their follow-up scans, and also difficult to process infant MR images, typically exhibiting extremely low tissue contrast and dynamic appearance.

To this end, in this work, we propose a novel highdimensional cortical folding descriptor based on multi-scale decomposition of curvature maps via spherical wavelets, called *FoldingPrint*, for infant identification. The motivation of using cortical folding is that, although the infant brain develops dynamically in the first years, all the primary and secondary cortical folds forming the major cortical folding patterns are established at term birth and are largely preserved during brain development [4], as shown in Fig. 1. Experiments are carried out on a large longitudinal infant dataset with 1,141 MRI scans during the first two years from 472 infants. Based on our results, we address two important neuroscientific questions: 1) whether the cortical folding of each infant is unique for individual identification; and 2) which cortical regions are more variable across individuals and more reliable for infant identification.

2. METHODS

2.1. Data Acquisition and Cortical Surface Mapping

Longitudinal brain MR images from 472 healthy infants were acquired at birth, 1 and 2 years of age, using a 3T Siemens scanner. In total, 1,141 longitudinal scans were obtained, including 472 neonates, 387 1-year-olds, and 282 2-year-olds. T1-weighted images were acquired by using the following imaging parameters: TR / TE = 1900 / 4.38 ms, and resolution = $1\times1\times1$ mm^3. T2-weighted images were acquired by using the following settings: TR / TE = 7380 / 119 ms, and resolution = $1.25\times1.25\times1.95$ mm^3.

All MR images were processed using an infant-dedicated computational pipeline [5]. The preprocessing includes the following main steps: skull stripping, cerebellum removal, intensity inhomogeneity correction, tissue segmentation, hemisphere separation, and topology correction. For each hemisphere of each scan, the inner cortical surface was reconstructed by using a deformable surface method and mapped onto a spherical space [6]. Each spherical cortical surface was aligned onto the age-matched template in the UNC 4D

infant cortical surface atlas [5, 7] using Spherical Demons [8]. For each cortical surface, a parcellation map with 34 regions of interest (ROIs) [9] was obtained by propagation of the parcellation map in the atlas.

2.2. Computing Cortical FoldingPrints

To identify infants based on the rapidly developing cortical folding, we need to construct discriminative cortical folding features, which are able to capture subject-specific distinct folding patterns and meanwhile are robust to the emerging tertiary cortical folds during the early brain development. To this end, we first decomposed the mean curvature map on the cortical surface into multiple complementary spatial-frequency scales using over-complete spherical wavelets [10], thus comprehensively characterizing the cortical folding in multi-scales. Then, for each cortical region, we created a high-dimensional multivariate folding descriptor, i.e., FoldingPrint, based on the joint probability distribution of the decomposed multi-scale curvature maps.

Multi-scale Decomposition of Curvature Maps: First, the mean curvature map of each inner cortical surface (with a spherical topology) was calculated for each scan. Then, the mean curvature map was decomposed into multiple spatial-frequency scales, via the over-complete spherical wavelet transform [10], where each wavelet scale is sufficiently sampled, thus free from sampling aliasing. Let $\boldsymbol{C}$ be the input spherical curvature map and $\{\tilde{\boldsymbol{h}}_n\}_{n=1}^{N}$ be the spherical analysis filters at $N = 7$ frequency levels, a series of wavelet coefficient maps $\boldsymbol{W}_n$ can thus be obtained at multiple spatial-frequency scales by convolving them in the spherical domain as:

$$\boldsymbol{W}_n = \boldsymbol{C} * \tilde{\boldsymbol{h}}_n \tag{1}$$

Herein, $\tilde{\boldsymbol{h}}_n$ is defined as: $\tilde{\boldsymbol{h}}_n = Q_n\boldsymbol{\psi}$, where n is the frequency level (with a larger n corresponding to a narrower filter), and $\boldsymbol{\psi}$ is the mother wavelet, and $Q_n\boldsymbol{\psi}$ represents the dilations of $\boldsymbol{\psi}$. The Laplacian-of-Gaussian was adopted as the mother wavelet $\boldsymbol{\psi}$ as in [10]. As the underlying wavelet basis functions have local supports in both space and frequency, the multi-scale wavelet coefficient maps can thus encode rich information of cortical folding at different levels. As we can see from Fig. 2(a), at coarser levels, the wavelet coefficients encode the larger scale folding information; while at finer levels, the wavelet coefficients capture the smaller scale folding information. Thus, the decomposed curvature maps lead to multi-scale folding characteristics.

Region-based Folding Descriptor: To create a comprehensive folding descriptor, we computed a highdimensional joint probability distribution of three most informative curvature scales, i.e., a 3D joint histogram, in each ROI as its FoldingPrint. The proposed regional FoldingPrint is able to couple the robustness of region-based analysis with the richness of multi-scale curvature maps.

Given a cortical surface and its decomposed curvature maps $\{\boldsymbol{W}_n\}n=1\ldots7$, we adopted three most informative scales of the decomposed curvature maps, in consideration of the

exponential growth of the feature dimension when adding all scales of curvature. For a given ROI r with a collection of vertices $\{v\}_{v=1}\ldots V_r$, its FoldingPrint $\boldsymbol{H}_r$ is a 3D joint histogram for capturing multi-scale structural information. Of note, we can also adopt the kernel density estimation (KDE) to improve the smoothness of the descriptor. Herein, the histogram was calculated through the following steps: 1) Linearly dividing the 3D space of the selected decomposed curvature maps $\boldsymbol{W}_i$, $\boldsymbol{W}_j$, and $\boldsymbol{W}_k$ into $X\times Y\times Z$ bins; 2) Inserting all vertices in $\{v\}_{v=1}\ldots V_r$ into $\boldsymbol{H}_r$: $\forall 1 \le x \le X$, $\forall 1 \le y \le Y$, $\forall 1 \le z \le Z$. $\boldsymbol{H}_r$ is thus defined as:

$$\boldsymbol{H}_r(x, y, z) = Count\left\{v \in \boldsymbol{H}_r : \left(\boldsymbol{W}_i(v), \boldsymbol{W}_j(v), \boldsymbol{W}_k(v)\right) \in B(x, y, z)\right\} \tag{2}$$

where $\boldsymbol{W}_i(v)$, $\boldsymbol{W}_j(v)$, and $\boldsymbol{W}_k(v)$ are the decomposed curvature values of vertex v at levels i, j, and k, respectively. $B(x, y, z)$ denotes the bin corresponding to the x-th interval of w_i, y-th interval of W_j, and z-th interval of w_k; 3) Normalizing $\boldsymbol{H}_r$ by dividing the total vertex number in the ROI r. Herein, we chose $X = Y = Z = 10$, thus obtaining a 1000 dimensional descriptor for each ROI. Of note, $\boldsymbol{H}_r$ is *independent* of the resolution of the cortical surface mesh, thus is robust to different surface reconstruction tools. Moreover, $\boldsymbol{H}_r$ is *independent* from surface registration.

In our application, we selected the decomposed curvature maps $\boldsymbol{W}_2$, $\boldsymbol{W}_3$, and $\boldsymbol{W}_4$, i.e., levels 2 to 4, as shown in Fig. 2(a). We abandoned other scales due to the following reasons. On one hand, the coarsest level (level 1) mainly captures very large scale folding information, which is very similar across subjects and thus not discriminative for infant identification. On the other hand, the finer levels (levels 5 to 7) mainly embed small scale folding information, which is typically sensitive to the tertiary cortical folds emerging during postnatal development and noises, and thus not reliable for infant identification. In contrast, the middle levels (levels 2 to 4) are robust to cortical development and thus are discriminative for infant identification.

In this way, for each subject, we can obtain a series of high-dimensional regional folding descriptors $\{\boldsymbol{H}_r\}_{r=1\ldots R}$, together forming FoldingPrints. In this study, $R = 68$ is the total number of ROIs in two hemispheres. Fig. 2(b) shows the graphical representations of the FoldingPrint of a region (i.e., the left caudal middle frontal gyrus) from 2 subjects at their 0 and 2 years of age. Herein, each 3D FoldingPrint is projected onto three 2D planes for better visualization. It can be seen that the FoldingPrint is able to capture the stable subject-specific folding pattern during development.

2.3. Infant Identification based on FoldingPrints

With the FoldingPrint of each region, we can simply compare all regions' FoldingPrints to identify infants. To measure the difference between two scans p and q in r-th ROI with FoldingPrints $\boldsymbol{H}_r^p$ and $\boldsymbol{H}_r^q$, we chose the widely-used Chi-squared histogram distance:

$$D_r(p, q) = \sum_{x=1}^{X} \sum_{y=1}^{Y} \sum_{z=1}^{Z} \frac{\left(\boldsymbol{H}_r^p(x, y, z) - \boldsymbol{H}_r^q(x, y, z)\right)^2}{\boldsymbol{H}_r^p(x, y, z) + \boldsymbol{H}_r^q(x, y, z)} \tag{4}$$

Of note, other histogram distance measures, e.g., earth mover's distance, could also be alternatively adopted here.

Given a new coming infant to be identified, in each cortical ROI, the distance of FoldingPrints between this infant and each subject in a specific age group in the database was calculated. Then, all distances were sorted in an increasing order, and the subject with the smallest distance in the database was regarded as a potentially-identified subject. Based on all 68 ROIs, for each subject to be identified, 68 identification results were obtained. Then through a simple majority voting, *without* training any classification model, the subject with the highest frequency among the potentially-identified subjects was considered as the finally-identified subject. Of note, if more than one subjects show the highest frequency, the one with smaller distance was regarded as the identified subject.

3. RESULTS

The proposed method was applied to a large longitudinal infant MRI dataset with 1,141 scans at 0, 1, and 2 years of age. Three sets of experimental tasks were carried out to evaluate the identification accuracy of our method, i.e., 1) using the scans at birth to identify their corresponding scans at year 1; 2) using the scans at birth to identify their corresponding scans at year 2; and 3) using the scans at year 1 to identify their corresponding scans at year 2. Since cortical folding develops more rapidly in the first year than the second year [3], the first two tasks are more challenging.

Table 1 reports the accuracies of infant identification on three tasks using our proposed FoldingPrints and other feature combinations. Compared to different 2D or 3D joint distributions of the decomposed curvatures at levels 2–5, we found that the 3D joint distribution with the decomposed curvatures at levels 2, 3 and 4 achieved the best performance. In comparison, we have also extensively tested different joint distributions of the traditional cortical folding features, i.e., mean curvature, sulcal depth, and average convexity. As we can see, these high-dimensional features still failed to show comparable performance as the proposed FoldingPrints, especially in the first two tasks. The underlying reason may be that multi-scale decomposed curvature maps contain complementary information. Thus, their combinations capture more complete information of the cortical folding. In contrast, the relationship among those traditional cortical folding features are unclear. Thus, their combinations would not necessarily improve the performance. Besides, by setting the numbers of the histogram bins to 10, 15, and 20, we found that the performance of FoldingPrints does not strongly rely on this number, since they maintain better performance in all cases, compared to other feature combinations. The best performance is found with 10 bins for the above three tasks, with the accuracies of 98.97%, 98.58%, and 100.00%, respectively. These results suggest that our proposed FoldingPrints are effective for infant identification, even for the challenging tasks involving neonates.

We further evaluated the identification accuracy of each ROI in the first two tasks and mapped the results onto the cortical surface, as shown in Fig. 3. The regions with high accuracy are more reliable for infant identification, and thus their folding patterns should be more distinctive and variable across individuals. As can be seen, overall most of the high-

order association cortices, i.e., the prefrontal cortex (including the caudal middle frontal gyrus, rostral middle frontal gyrus and superior frontal gyrus), middle temporal gyrus, and inferior parietal gyrus, as well as the precentral gyrus show higher accuracy in infant identification, indicating their distinct folding patterns across subjects. While the auditory, visual, and insular regions show lower accuracy, suggesting their less variable folding patterns across individuals. These region-specific patterns are similar in both tasks and also are relatively symmetric across left and right hemispheres. These findings are supported by the current knowledge on inter-individual variability [11].

4. CONCLUSION

For the first time, we explored infant identification based on cortical folding in a large longitudinal dataset. Our contribution is mainly in two aspects. *First*, we proposed a regional cortical folding descriptor, i.e., FoldingPrint, for infant identification, which is capable of comprehensively characterizing the invariance of developing cortical folding and achieves promising identification accuracy. Thus, once brain MRI scans become cheap and convenient to acquire, FoldingPrints could be used as reliable biometric traits for infant identification. *Second*, leveraging FoldingPrints, we found that the cortical regions with high identification accuracy are distributed mostly in the high-order association cortices, indicating their high inter-subject variability. This study sheds some light on the individual identification and the inter-subject cortical variability during infancy.

Acknowledgments

This work was partially supported by NIH grants (MH100217, MH107815, MH108914, MH109773, MH110274, MH116225, MH117943, MH070890, MH064065, and HD053000).

5. REFERENCES

[1]. Wachinger C, Golland P, Kremen W, et al., “BrainPrint: a discriminative characterization of brain morphology,” Neuroimage, vol. 109, pp. 232–248, 2015. [PubMed: 25613439]

[2]. Finn ES, Shen X, Scheinost D, et al., “Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity,” Nat Neurosci, vol. 18, pp. 1664–1670, 2015. [PubMed: 26457551]

[3]. Li G, Wang L, Shi F, et al., “Mapping longitudinal development of local cortical gyrification in infants from birth to 2 years of age,” JNeurosci, vol. 34, pp. 4228–4238, 2014. [PubMed: 24647943]

[4]. Duan D, Xia S, Rekik I, et al., “Exploring folding patterns of infant cerebral cortex based on multi-view curvature features: Methods and applications,” Neuroimage, pp. 575–592, 2019.

[5]. Li G, Wang L, Shi F, et al., “Construction of 4D high-definition cortical surface atlases of infants: Methods and applications,” Med Image Anal, vol. 25, pp. 22–36, 2015. [PubMed: 25980388]

[6]. Fischl B, Sereno MI, Dale AM, “Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system,” Neuroimage, vol. 9, pp. 195–207, 1999. [PubMed: 9931269]

[7]. https://www.nitrc.org/projects/infantsurfatlas/.

[8]. Yeo BTT, Sabuncu MR, Vercauteren T, et al., “Spherical demons: fast diffeomorphic landmark-free surface registration”, IEEE Trans Med Imaging, vol. 29, pp. 650–668, 2010. [PubMed: 19709963]

[9]. Desikan RS, Ségonne F, Fischl B, et al., “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest,” Neuroimage, vol. 31, pp. 968–980, 2006. [PubMed: 16530430]

[10]. Yeo BTT, Yu P, Grant PE, et al., “Shape analysis with overcomplete spherical wavelets,” in MICCAI, pp. 468–476, 2008.

[11]. Hill J, Dierker D, Neil J, et al., “A surface-based analysis of hemispheric asymmetries and folding of cerebral cortex in term-born human infants,” J Neurosci, vol. 30, pp. 2268–2276, 2010. [PubMed: 20147553]

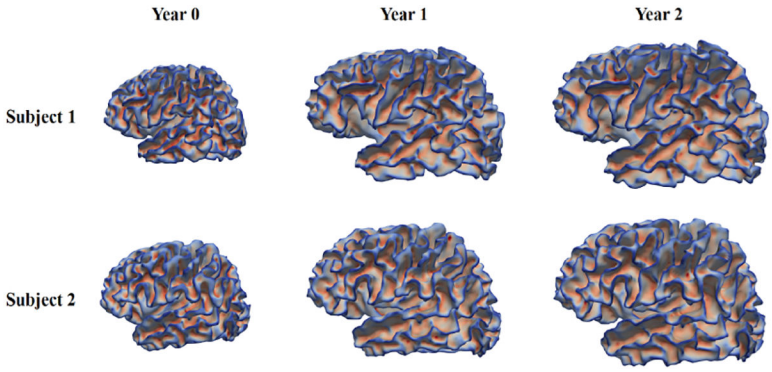

Fig. 1.

Longitudinal cortical surfaces at 0, 1, and 2 years of age of two subjects. Cortical surfaces are color-coded by mean curvature.

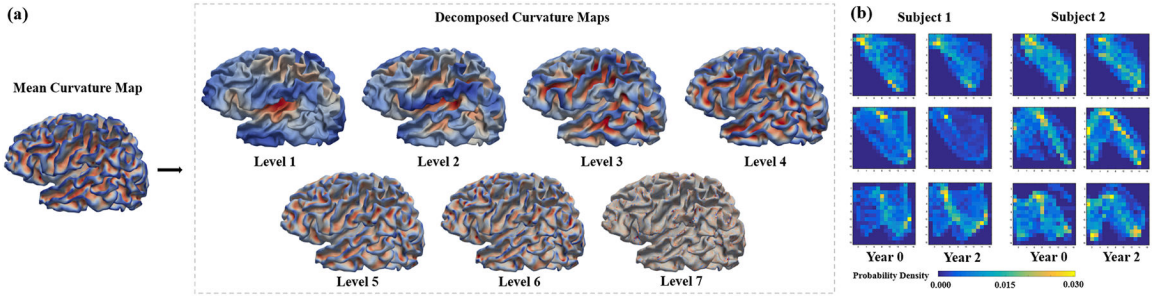

Fig. 2.

(a) The mean curvature map (left panel) and the decomposed curvature maps at levels 1–7 by over-complete spherical wavelets (right panel). (b) Graphical representations of the FoldingPrint of the left caudal middle frontal gyrus from two subjects at their 0 and 2 years of age. For visualization, each 3D FoldingPrint is projected onto three 2D planes.

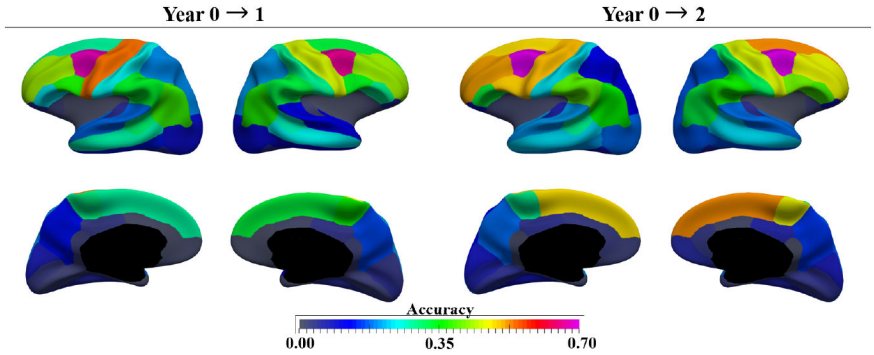

Fig. 3.

Region-specific accuracies in infant identification.

Table 1.

Evaluation of the identification accuracies using FoldingPrints (3D histogram based on $\boldsymbol{W_2}$-$\boldsymbol{W_3}$-$\boldsymbol{W_4}$(**10**)) and related feature combinations. ($\boldsymbol{W_n}$: Decomposed wavelet curvature at scale n; Cur: mean curvature; Con: average convexity; Dep: sulcal depth)

Hist. (bins)	W_2-W_3 **(30)**	W_2-W_4 **(30)**	W_2-W_5 **(30)**	W_3-W_4 **(30)**	W_3-W_5 **(30)**	**Cur-Dep (30)**	**Cur-Con (30)**
Year 0→1	54.26	13.95	4.39	74.16	51.42	1.29	1.29
Year 0→2	72.70	18.44	6.74	83.33	58.87	1.77	0.71
Year 1→2	100.00	78.17	59.39	99.49	98.98	53.81	61.42
Hist. (bins)	**Dep-Con (30)**	**Cur-Dep-Con (10)**	W_2-W_3-W_5 **(10)**	W_3-W_4-W_5 **(10)**	W_2-W_3-W_4 **(20)**	W_2-W_3-W_4 **(15)**	W_2-W_3-W_4 **(10)**
Year 0→1	4.39	4.91	94.06	78.29	96.90	97.93	**98.97**
Year 0→2	4.96	2.13	93.97	87.23	95.74	98.94	**98.58**
Year 1→2	94.42	94.92	100.00	99.49	100.00	100.00	**100.00**