Measurement of folding in surfaces of arbitrary size in human brain development

Claudia Rodriguez-Carranza, P. Mukherjee, Dan Vigneron, James A. Barkovich, C. Studholme

Measurment of folding in surfaces of arbitrary size in human brain development

C. Rodriguez-Carranza[1], P. Mukherjee[2], D. Vigneron[2], J. Barkovich[2], and C. Studholme[1,2]

[1] NCIRE/VAMC, San Francisco, CA
[2] Department of Radiology, University of California, San Francisco

Abstract. This paper describes a novel approach to in-vivo measurement of brain surface folding in clinically acquired neonatal MR image data. Specifically it addresses the problem of comparing folding of surfaces of arbitrary area. Most of the current measures of folding are not independent of the area of the surface from which they are derived. Therefore, applying them to whole brains or subregions of different sizes result in differences which may or may not reflect true differences in folding. In this paper we describe alternative approaches to deriving area independent measures. The measures were applied to twelve premature infants (age 28-37 weeks) from which cortical gray and white surfaces were extracted. Experimental results show that previous folding measures are sensitive to the size of the surface patch, and that the area independent measures proposed here provide significant improvements. Such a system provides a tool to allow the study of structural development in the neonatal brain within specific functional subregions, which may be critical in identifying later neurological impairment.

1 Introduction

The percent of infants born preterm, or at less than 37 completed weeks of gestation, has increased significantly in the last two decades. The latest report [6] in the USA indicates that almost one in every eight births is preterm. There is growing evidence that premature birth can result in structural and functional alterations of the brain, which are related to adverse neurodevelopmental outcome later in life [7, 8]. Some of the challenges that preterm infants face range from spastic motor deficits (cerebral palsy) [20], impaired academic achievement [3,5], and behavioral disorders [9,12]. However, the conditions that cause the cerebral abnormalities that underlie these common and serious developmental disabilities are not entirely understood [8]. The wider availability of clinical in vivo magnetic resonance imaging of neonatal brain anatomy, provided by systems which make use of an MRI compatible incubator, creates a new opportunity to quantify brain development. In this work we are particularly interested in the study of the cortical folding, or gyrification, in preterm infants, because it may reflect underlying functional organization.

In our earlier work [16] we applied previously proposed global average measures of folding to surfaces extracted from neonatal MRI, and showed their use in tracking global age changes. The main limitation was that the previously proposed measures were heavily dependent on the size of the surface being examined. Since brain surface area increases dramatically with brain development, previous measures cannot probe whether cortical folding is following a normal pattern, independent of size. Additionally these cannot be calculated on functional sub-regions of the cortex which may have a different surface area in different individuals. To address the problem, in this paper we describe and analyze measures of global folding that are independent of the size of the surface of analysis. We apply the measures to surfaces extracted from MR images of twelve premature infants with ages from 28-37 weeks. The folding of whole surfaces, as well as their left and right hemispheres, was studied.

2 Theory

Several global 3D brain surface measures of gyrification have previously emerged. Some are based on surface principal curvatures (k_1, k_2), mean curvature $(H = \frac{1}{2}(k_1 + k_2))$, or Gaussian curvature $(K = k_1 k_2)$. Examples are folding index (FI) [4], the intrinsic curvature index (ICI) [4], L^2 norm of the mean curvature (MLN) [1], L^2 norm of Gaussian curvature (GLN) [1], and average curvature [13]; global shape index (GS) and global curvedness (GC) were defined in our earlier work [16] based on the local shape descriptors curvedness (c) [10] and shape index (s) [10]. Examples of non-curvature based measures are the gyrification index [21], calculated as the ratio of the entire cortical contour of the brain to the superficially exposed contour, and roundness[1] (Rn), based on surface area (A) and volume (V). A list of current global 3D measures of folding is shown on the left column of Table 1. These expressions are normalized to yield the unit value for a sphere [16].

2.1 Size-independent measures of surface folding

Examination of the form of the current global measures in Table 1 reveals a critical dependency on surface area, with the exception of global shape index and gyrification index. This can be better illustrated with the following example, using MLN to measure folding. Take a whole sphere of radius Ro, half a sphere of radius Ro, and a whole sphere of half that radius. The three objects have the same surface complexity, hence a measure of global folding should yield identical results. By definition, $H = 1/R$ at each point in a sphere. The whole sphere with radius Ro yields $MLN = \frac{1}{4\pi}\sum_A H^2 = \frac{1}{4\pi} \times 4\pi Ro^2 \times (1/Ro)^2 = 1$, and so does the whole sphere with radius $Ro/2$: $MLN = \frac{1}{4\pi} \times (4\pi\frac{Ro}{2})^2 \times (2/Ro)^2 = 1$. On the other hand, MLN for the half sphere of radius Ro is half that value: $MLN = \frac{1}{4\pi} \times \frac{1}{2} 4\pi Ro^2 \times (1/Ro)^2 = \frac{1}{2}$. This demonstrates the dependency of the

[1] The terms surface complexity [13] and isoperimetric ratio [1] have also been used.

measure on the size of the surface of analysis. Normalizing MLN with the surface area does not alleviate the problem: $MLN = \frac{1}{A}\sum_A H^2 = \frac{1}{4\pi Ro^2}(4\pi Ro^2)(1/Ro)^2 = \frac{1}{Ro^2}$ (whole sphere of radius Ro). We propose two normalization factors: 1) $T = 3V/A$ and 2) $|\overline{H}| = \frac{1}{A}|\sum_A H|$. The resulting T-normalized and H-normalized expressions (T-measures and H-measures, respectively) are shown in the right column of Table 1. For a sphere, $T = Ro$, $|\overline{H}| = 1/Ro$, and it can be verified that $MLN_T = MLN_H = 1$ for the three sphere cases. Similarly for the rest of the newly defined area-independent measures.

Table 1. Measures of surface folding

Current measures	**New measures**
Area-dependent	Area-independent
$MLN=\frac{1}{4\pi}\sum_A H^2$ [1]	$MLN_T=\frac{T^2}{A}\sum_A H^2$
$GLN=\frac{1}{4\pi}\sqrt{A*\sum_A K^2}$ [1]	$GLN_T=T\sqrt[4]{\frac{1}{2}\sum_A K^2}$
$ICI=\frac{1}{4\pi}\sum_A K^+$ [4]	$ICI_T=T\sqrt{\frac{1}{A}\sum_A K^+}$
$FI=\frac{1}{4\pi}\sum_A ak$ [4]	$FI_T=T\sqrt{\frac{1}{A}\sum_A ak}$
$GC=\frac{1}{\sqrt{A4\pi}}\sum_A c$ [16]	$GC_T=T\sqrt{\frac{1}{A}\sum_A c}$
$\overline{U}=\frac{1}{A}\sum_A U$ [13]	$\overline{U_T}=\frac{T}{A_U}\sum_A U$
$\overline{V}=\frac{1}{A}\sum_A V$	$\overline{V_T}=T\sqrt{\frac{1}{A_V}\sum_A V}$
$Rn=\frac{A}{\sqrt[3]{36\pi V^2}}$ [1]	$SH2SH=T\sum_A H^2/\sum_A H$
	$SK2SK=T\sqrt{\sum_A K^2/\sum_A K}$
Area-independent	
$GS=\frac{1}{A}\sum_A s$ [16]	$MLN_H=\frac{1}{\lvert\overline{H}\rvert}\sqrt{\sum_A H^2/A}$
Gyrification Index [21]	$GC_H=\frac{1}{A\lvert\overline{H}\rvert}\sum_A c$
	$AF_i=A_i/A,\ i\in\{H^+,K^+\}$

Notation: $c = \sqrt{\frac{1}{2}(k_1^2+k_2^2)}$ (curvedness [10]), $s = \frac{2}{\pi}\arctan\frac{k_2+k_1}{k_2-k_1}$ (shape index [10]), $ak = |k_1|(|k_1| - |k_2|)$, $T = 3\frac{V}{A}$, and $|\overline{H}| = \frac{1}{A}|\sum_A H|$. $U \in [H^+, H^-]$, $V \in [K^+, K^-]$, and A_{H+} (or A_{K+}) is the area of the surface with positive H (or K) curvature.

In addition, we defined three new area-independent folding measures: $SH2SH$, $SK2SK$, and $AF_{\{H^+,K^-\}}$. The first two were based on the rationale that measures composed of ratios of local curvature factors would intrinsically eliminate the dependence on area. The latter was based on the idea that a reasonable characterization of the degree of folding in a brain surface is the fraction of the surface which contains convex folds (gyri). This can be mathematically characterized by looking at the relative portion of the surface which has positive mean or Gaussian curvature. For an undeveloped brain, the fraction would be

large, and it would progressively decrease as the concave folds (sulci) appear. The definition of the three expressions is shown in Table 1.

3 Method

The sphere example was useful to identify global folding measures that have the potential to consistently evaluate surface folding, independent of the size of surface of analysis. A more formal verification of this area-independence property is in progress. In this work we present an experimental comparison of the measures on two datasets. The first dataset consisted of 15 neonatal brains. The surfaces of this brains not only increase in size with age, but they also become more complex. So, folding varies proportionally to size. The second dataset consisted of iso-surfaces for which surface folding varied inversely to size. They were generated from one neonatal brain at various percentage occupancy thresholds (see Section 3.3). A larger threshold produces a smaller brain and larger separation between sulci walls, which translates to increased folding. For both datasets, gyrification was measured on left and right hemispheres, in addition to the whole brains. The infants in this study are normal, therefore gyrification is expected to be similar on both sides. Area-independent measures are expected to yield similar values for each hemisphere and the full-brain, but not so the area-dependent measures.

3.1 Data

Human Neonatal Images. High resolution ($0.703 \times 0.703 \times [1.5 - 2.2]$ mm) 3D T1 weighted SPGR images were acquired on premature infants using a 1.5T GE MRI scanner with an MRI compatible incubator. The gestational ages (GA) of the twelve premature infants in this study ranged between 24-31 weeks. The postmenstrual ages (gestational age + postnatal age) at the time of acquisition were 28-37 weeks. A subsequent scan for three infants was available, hence a total of 15 brains were processed.

3.2 Image Segmentation

The outer gray matter and the gray-white matter interface[2] surfaces of the premature brains were extracted semi-automatically as described in [16] using the rview software package [17]. Cortical gray matter was segmented for fifteen brains and gray-white matter interface for thirteen brains. A value of 1000 was assigned to brain voxels and 0 to background voxels.

[2] In the remaining of the text we will refer to the gray-white matter interface simply as white matter.

3.3 Computation of Brain Surface Folding

Each binary segmented brain was supersampled using voxel replication to prevent loss of fine scale features ($2 \times 2 \times 4$). A voxelwize approach to surface curvature estimation from iso-surfaces, derived from that developed by Rieger [14], was then employed. It avoids the need of a parametric model. A summary of the sequence of steps (described in detail in [16]) is as follows: 1) computation of the image gradient $g = \nabla f(x, y, z)$ of the replicated binary volume; 2) computation of the gradient structure tensor (GST) [14] defined as $T = \overline{gg^t}$; 3) calculation of the eigenvectors v_1, v_2, v_3 of T and the mapping $M(v_1) = \frac{v_1 v_1^t}{||v_1||}$; 4) on the 50% occupancy iso-surface, computation of the principal curvatures from $|k_{1,2}| = \frac{1}{\sqrt{2}}||\nabla_{v_{2,3}} M(v_1)||_F$ (Fröbenius norm); the sign of k_1 and k_2 is determined from the Hessian matrix of the image and the eigenvectors v_2 and v_3. The curvatures k_1 and k_2 were then used to compute mean (H) and Gaussian (K) curvature and the folding measures of Table 1. Each step involves image smoothing or differentiation. For all images, the 3D filters used at the ith step had a full-width at half maximum (f_i) of: $f_1 = 2.1$mm, $f_2 = 1$mm, $f_3 = 2$mm, $f_4 = 1$mm. The rationale behind the choice of values was to create a smooth surface that would yield smooth variations in curvature. Normalized Gaussian derivatives [11] were used for all computations.

The input to this algorithm is the binary replicated segmented brain, for which foreground voxels have a value of 1000 and background voxels a value of 0. Near the border of the brain the kernel incorporates brain and background voxels, hence the raw effect of the convolution on the image is that the resulting volume has voxel values in the range [0,1000]. The resulting voxel values are interpreted as the partial occupancy by brain tissue. The larger the number indicates that more foreground voxels were present in the kernel. For each brain, the surface on which curvature and the global measures of folding were calculated was taken to be the iso-surface with at least 50% occupancy (i.e. voxel values of at least 500) and satisfying 6-connectivity with the background.

3.4 Cortical Partitioning using Spatial normalization

Full-brains and their left and right hemispheres were processed with the algorithm just described. The procedure used to identify the left and right hemispheres in each neonatal brain was the following. A reference neonatal anatomy was manually partitioned into left and right hemisphere. A multi-resolution B-Spline based spatial normalization, adapted from that described in [18, 19] and as used in [2], was then applied to estimate a spatial transformation mapping from this reference to each subject MRI being studied. This transformation was then numerically inverted to allow the assignment of the nearest reference label to a given subject image voxel, which brought the voxelwize partitioning of the reference anatomy to the same space as the surface extracted from each infant. The partitioning was then used to constrain the voxelwize evaluation of the folding measures.

4 Results

We applied the measures to surfaces from the 15 MRI studies to investigate their response to folding with age. The relationship between age and folding was analyzed with regression and for this work a linear model was used. Due to space limitations results from only seven measures are shown in Figure 1. Top two rows correspond to gray matter results, and bottom two rows to white matter. The first observation is that measures with large goodness-of-fit for gray matter, had very small goodness-of-fit for white matter, and viceversa. This was true for the whole brain and both hemispheres. The same was observed in our results [15] on 10 MRI studies. The measures with greater goodness-of-fit for gray matter were the T-normalized measures: ICI_T, FI_T, $\overline{H^+}$, $\overline{K^-}$, $\overline{K^+}$; for white matter, GS (global shape index) and the $|\overline{H}|$-normalized measures: MLN_H and GC_H. The goodness-of-fit for gray matter and white matter (in that order) on whole surfaces (depicted as $*$ in the figure) was: CI_T (**0.85**, 0.004), FI_T (**0.85**, 0.19), GLN_T (0.80, 0.21), ICI_T (0.83, 0.002), MLN_T (0.80, 0.07), GS (0.61, **0.95**), $SK2SK$ (0.67, 0.46), $SH2SH$ (0.77, 0.25), MLN_H (0.48, **0.96**), GC_H (0.52, **0.95**), $\overline{H^-}$ (0.73, 0.07), $\overline{H^+}$ (**0.85**, 0.12), $\overline{K^-}$ (**0.85**, 0.27), $\overline{K^+}$ (same as ICI_T), AF_{H^+}=(0.62, 0.87), AF_{K^+}=(0.45, 0.13). The measures had similar goodness-of-fit scores for the left and right hemispheres. Area, volume, and T varied linearly in all cases.

Next we analyzed the dataset for which smaller brains were more folded. The iso-surfaces were created from a single brain (age 28 weeks) at percentage occupancy thresholds in the range 5-95% (in increments of 5%). An example of one coronal slice from five iso-surfaces obtained at five different thresholds is shown in Figure 2. The results for five measures on whole surfaces and their left and right hemispheres are shown in Figure 3. In contrast with the first dataset, the folding measures had similar goodness-of-fit for both gray and white matter (either both large or both small). Another difference was that FI_T presented a small goodness-of-fit. For this iso-surface dataset the area of gray matter varied in a non-linear fashion (plot not shown). This behavior is explained by the separation of the gray matter volume into two hemispheres, for larger thresholds (see Figure 2). The increase in surface area due to this separation is dominant over the relative decrease in area due to the larger threshold.

It can be observed that in the first dataset (Figure 1) the slopes of T-normalized measures have different tendencies for gray matter (positive) and white matter (negative). This difference is explained by the slope of T itself. For gray matter the rate of change of volume, with age, was faster than that of area, therefore $T = 3V/A$ had a positive slope. The inverse occurs for white matter. For the second dataset (Figure 3), the slope of T was negative for both gray and white matter, therefore the slope of all T-normalized measures is also negative. Once this slope difference is considered, it is clear that the new measures (T- and $|\overline{H}|$-normalized) change consistently with folding, independent of the size of the surface.

Results on both datasets for previously defined measures corroborate their dependency on the size of the surface of analysis. For each hemisphere, the

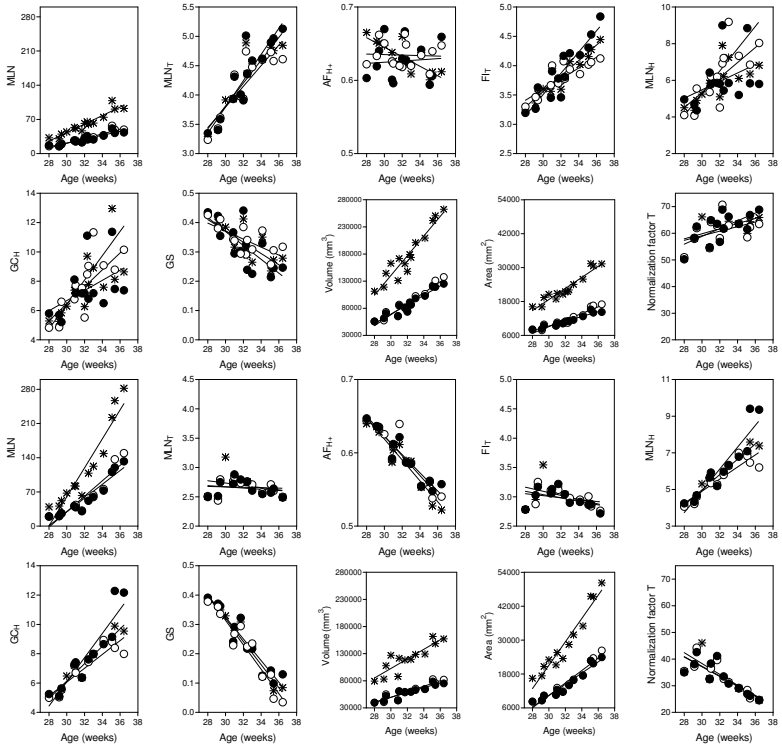

Fig. 1. Plots of seven folding measures applied to gray matter (top two rows) and white matter (bottom two rows). Filled shapes correspond to results for left hemisphere, and empty ones for right hemisphere. Results for the whole surface ($\star$) are also included.

magnitude of previously defined measures was at least half the value than for the whote-brain. In contrast, for the newly defined measures, their magnitude for whole-brain was in between that for left hemisphere and right hemisphere. As an example, results for MLN and GLN are shown in Figures 1 and 3, respectively.

Assessment of intra-observer segmentation variability (preliminary stage). The white matter of two brains, from the 15 available, was segmented and analyzed with an approach similar to the one in [1]. One brain was 29.1 weeks and was segmented four times; the other brain was 32.4 weeks and was segmented twice. Measures were computed on all the segmented surfaces. Given a set of segmentations for one brain, the segmentation variability was assessed as the maximum percentage variation of a measure from the mean value in the set. This was compared to the percentage change of the mean value of a measure between the 29.1 week brain and the 32.4 week brain. This number was taken as an indication of the extent of the variability in segmentations. An example of the results is given for MLN_H, GS, and AF_{H^+}. In that order, the maximum percentage variation for each measure on the 29.1 week old brain was: 2.81%,

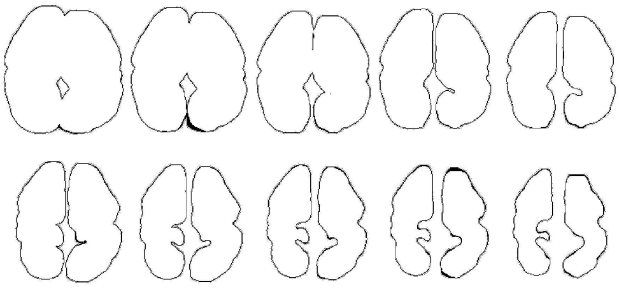

Fig. 2. A coronal slice each from five gray matter (top) and white matter (bottom) iso-surfaces extracted at percentage occupancy thresholds: 10%,30%,50%,70%,90%. All slices come from one brain.

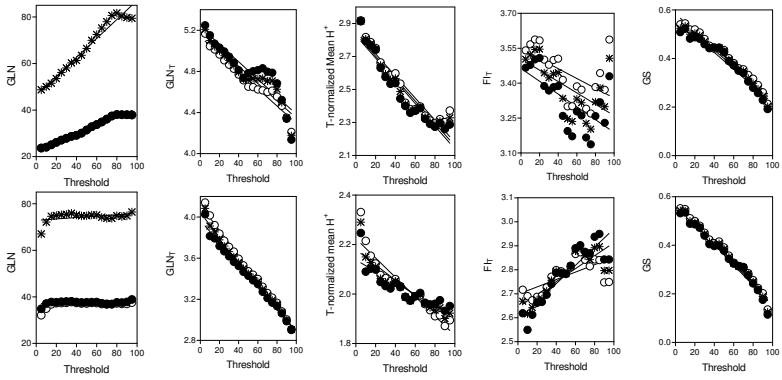

Fig. 3. Plots of five folding measures applied to iso-surfaces of a single brain at various occupancy thresholds. Top row shows results for gray matter and bottom row for white matter. Filled shapes correspond to left hemisphere and empty ones to right hemisphere. Results for the whole surface ($\star$) are also included.

2.57%, and 0.55%. The percentage difference between the 29.1 week brain and the 32.4 week brain was: 35.4%, 41.9%, 8.1%.

5 Discussion

An understanding of the cortical folding process in the development of premature infants may be important in explaining and predicting abnormal neurological outcome. The use of formal mathematical descriptions provides a more quantitative tool to study the folding process than is available with simple visual evaluation of MRI scans. The long term goal of our research is to create a model to track preterm *in vivo* neonatal brain cortical development that will help characterize normal gyrification and departures from it.

In this work we have shown that most previously proposed folding measures [1, 4] are dependent of the size of the surface patch on which they are calculated. To alleviate this problem we examined two approaches to area normalization of the measures and proposed new forms of measure, which satisfy the requirement of area-independence. These were evaluated on neonatal brain surfaces. We evaluated the measures ability to detect change in normal development by computing the linear regression of each measure with age. Finding the relationship between folding measures and age is an important part of the goal of understanding brain development. We will explore higher order models when more data is available.

In the dataset of 15 neonatal brains, folding varied proportionally to size. White matter folding tracked best with age, while no single measure rated consistently well for both gray and white matter surfaces. For the iso-surfaces dataset, where folding varied inversely to size, most measures rated well, and similarly for both gray and white matter. Despite differences in the behaviour of the normalization factor T and in the linearity of area on both datasets, most newly defined measures appeared to consistently assess folding. Recommended measures for gray matter are ICI_T, $\overline{H^+}$, $\overline{K^-}$; and for white matter GS, MLN_H, GC_H, and AF_{H^+}. The proposed normalization factors T and $|\overline{H}|$ opened the possibility to adapt previously defined measures to quantify gyrification in subregions of the brain. The slope of T appears to determine that of the T-normalized measures. Awareness of this is important for the correct interpretation of results. Alternative normalization factors need to be explored.

Preliminary results indicate that the effect of segmentation variability on the measures seems to be minimal: the differences in the measures with segmentation seems to be at least 10 times smaller than the differences with gestational age. We need to further investigate the sensitivity and validity of the measures in a larger cohort, containing both normal and pathological cases. In conclusion, the proposed new and normalized measures provide area-independent assessment of folding which provide the ability to study local gyrification. Such measures will provide a new tool in assessing global and local surface folding which is independent of the overall surface area, and are hence applicable to developing a model that tracks development in premature infants

Acknowledgments

This work was funded by the NIH grant R01 MH65392.

References

1. P. G. Batchelor, A. D. Castellano Smith, D. L. G. Hill, D. J. Hawkes, T. C. S. Cox, and A. F. Dean. Measures of folding applied to the development of the human fetal brain. *IEEE Transactions on Medical Imaging*, 21(8):953–965, August 2002.
2. V. Cardenas, L.L. Chao, R. Blumenfeld, and *et al.* Using automated morphometry to detect association between ERP latency and structural brain MRI in normal adults. *Human Brain Mapping*, 25(3):317–327, 2005.

3. R. W. I. Cooke, L. Foulder-Hughes, D. Newsham, and D. Clarke. Ophtalmic impairment at 7 years of age in children born very preterm. *Archives of Disease in Childhood Fetal and Neonatal Edition*, 89:249–253, 2004.
4. D.C. Van Essen and H. A. Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *The Journal of Neuroscience*, 17(18):7079–7102, 1997.
5. M. Hack, D. J. Flannery, and *et al.* Outcomes in young adulthood for very-low-birth weight infants. *New England Journal of Medicine*, 346(3):149–157, 2002.
6. B. E. Hamilton, J. A. Martin, and *et al.* Births: Preliminary data for 2004. *National Vital Statistics Reports*, 54(8), 2005.
7. P. S. Hüppi, B. Schuknecht, C. Boesch, and *et al.* Structural and neurobehavioral delay in postnatal brain development of preterm infants. *Pediatric Research*, 39(5):895–901, 1996.
8. T. E. Inder, S. K. Warfield, H. Wang, P. S. Hüppi, and J. J. Volpe. Abnormal cerebral structure is present at term in premature infants. *Pediatrics*, 115(2):286–294, 2005.
9. T. E. Inder, S. J. Wells, N. B. Mogridge, C. Spencer, and J. J. Volpe. Defining the nature of the cerebral abnormalities in the premature infant: a qualitative magnetic resonance imaging study. *Journal of Pediatry*, 143:171–179, 2003.
10. J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, 10(8):557–564, 1992.
11. T. Lindeberg. On scale selection for differential operators. In *Proc. 8th Scandinavian Conference on Image Analysis*, pages 857–866, 1993.
12. M. Luciana. Cognitive development in children born preterm: implications for theories of brain plasticity following early injury. *Development and Psychopathology*, 15:1017–1047, 2003.
13. V. A. Magnotta, N. C. Andreasen, and S. K. Schultz *et al.* Quantitative *in vivo* measurement of gyrification in the human brain: changes associated with aging. *Cerebral Cortex*, 9:151–160, 1999.
14. B. Rieger, F. J. Timmermans, L. J. van Vilet, and P. W. Verbeek. On curvature estimation of iso surfaces in 3D gray-value images and the computation of shape descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8):1088–1094, August 2004.
15. C. Rodriguez-Carranza, P. Mukherjee, D. Vigneron, J. Barkovich, and C. Studholme. A system for measuring regional surface folding of the neonatal brain from MRI. In *9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2006)*, 2006.
16. C. Rodriguez-Carranza, F. Rousseau, B. Iordanova, O. Glenn, D. Vigneron, J. Barkovich, and C. Studholme. An iso-surface folding analysis method applied to premature neonatal brain development. In *Proc. SPIE Medical Imaging: Image Processing*, volume 6144, pages 529–539, 2006.
17. C. Studholme. http://rview.colin-studholme.net/.
18. C. Studholme, V. Cardenas, and *et al.* Accurate template-based correction of brain MRI intensity distortion with application to dementia and aging. *IEEE Transactions on Medical Imaging*, 23(1):99–110, 2004.
19. C. Studholme, V. Cardenas, and *et al.* Deformation tensor morphometry of semantic dementia with quantitative validation. *Neuroimage*, 21(4):1387–1398, 2004.
20. J. J. Volpe. Neurologic outcome of prematurity. *Archives of Neurology*, 55:297–300, 1998.
21. K. Zilles, G. Schlaug, M. Matelli, G. Luppino, and et al. The human pattern of gyrification in the cerebral cortex. *Anatomy and Embriology*, 179:173–179, 1988.