Automatic Inference of Sulcus Patterns Using 3D Moment Invariants

Z.Y. Sun[1,2], D. Rivière[1,2], F. Poupon[1,2], J. Régis[3], and J.-F. Mangin[1,2]

[1] Neurospin, I2BM, CEA, France
zysun@cea.fr
[2] IFR 49, France
[3] Service de Neurochirurgie Fonctionnelle, CHU La Timone, Marseille, France

Abstract. The goal of this work is the automatic inference of frequent patterns of the cortical sulci, namely patterns that can be observed only for a subset of the population. The sulci are detected and identified using brainVISA open software. Then, each sulcus is represented by a set of shape descriptors called the 3D moment invariants. Unsupervised agglomerative clustering is performed to define the patterns. A ratio between compactness and contrast among clusters is used to select the best patterns. A pattern is considered significant when this ratio is statistically better than the ratios obtained for clouds of points following a Gaussian distribution. The patterns inferred for the left cingulate sulcus are consistent with the patterns described in the atlas of Ono.

1 Introduction

Human brain cortex folds to increase its surface area during development. It is intriguing to look at these folds. They are very complicated and variable, yet there is a certain consistency across brains [1]. Do they contain some information on the functional organization of the human brain? From the folds alone can we observe a pattern characteristic of a certain neurological disease? Thanks to recent advances in softwares dedicated to automatic recognition of cortical sulci [2,3,5,4], this kind of issues can now be tackled using large brain databases [6].

Each brain looks different and none of them looks exactly like the ones in the text books. Current studies of this variability focus on simple morphometric features like the length or the depth of the standard sulci or gyri. Unfortunately, the standard naming system cannot always account for the folding pattern variability. Hence some of the standard sulci can be difficult to define or to measure. This weakness of the nomenclature imposes difficulties on both morphometric studies and the pattern recognition softwares dedicated to automatic recognition of the sulci.

The most detailed description of the sulcus variability has been proposed in the atlas of Ono [7]. This atlas is not based on one single individual but on twenty different brains. For each sulcus, the authors propose a list of possible patterns and their frequencies. These patterns are defined for instance from the

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part I, LNCS 4791, pp. 515–522, 2007.

variability of the sulcus interruptions. In a way, the goal of the method proposed in this paper is to automate the work performed by Ono. We want to discover folding patterns that can be observed only for a subset of the population. For this purpose, once a sulcus has been defined in a population of brains, a non supervised clustering method provides subsets of brains with a characteristic trait. Each of these subsets is supposed to represent one of the patterns of interest.

In the following, we use two datasets of brains provided by the designers of brainVISA, an open software suite including a package dedicated to the study of cortical folds (http://brainvisa.info). The folds have been detected first using BrainVISA, then the sulci have been labeled either manually or automatically. The first dataset is made up of 36 brains, where each sulcus has been reliably labeled manually by a neuroanatomist. This dataset is used to train BrainVISA's sulcus recognition system. We also use another set of 150 brains, with the sulci automatically labeled. This database was provided by the International Consortium for Brain Mapping (ICBM) and acquired in the Montreal Neurological Institute of McGill University. The automatic recognition of the folds is less reliable but still gives reasonably good results [5].

The clustering of the sulci is based on 3D shape descriptors called moment invariants [6]. The first part of the paper describes several studies proving that these descriptors are well adapted to our purpose. The second part of the paper proposes a sketch of the agglomerative algorithm used to select interesting brain clusters [8]. Finally some results are shown for the cingulate sulcus, which provided the strongest patterns according to our criterion.

2 Shape Space and 3D Moment Invariants

The 3D moment invariants have been proposed as an interesting set of descriptors for the study of the shape of cortical sulci because they can be computed for any topology [6]. Hence they allow the management of the various sulcus interruptions. The construction of these descriptors is filtering out the influence of localization, orientation and scale from the 3D coordinate moments in order to obtain pure shape descriptors. While their theoretical derivation is complex, they can be computed in a simple and robust way from a black and white image defining an object. In the following, we use only the 12 invariants derived from the coordinate moments up to the power three. These 12 invariants denoted by $I1, I2, ..., I12$ are calculated from the software brainVISA and used as input to our clustering program.

Some investigations are carried out to verify that the set of moment invariants is a reasonably good shape representation to study the fold patterns. In order to confirm that similar shapes lead to similar representations, we verified first that a small shape variation leads to a small variation of the invariants. This is mandatory for our clustering purpose. Our experiments consist in creating series of shapes sampling a continuous shape transformation. An example of the resulting behaviour of the invariants is shown in Fig. 1.left. It is impossible to claim from these simple investigations that the invariants vary smoothly whatever the

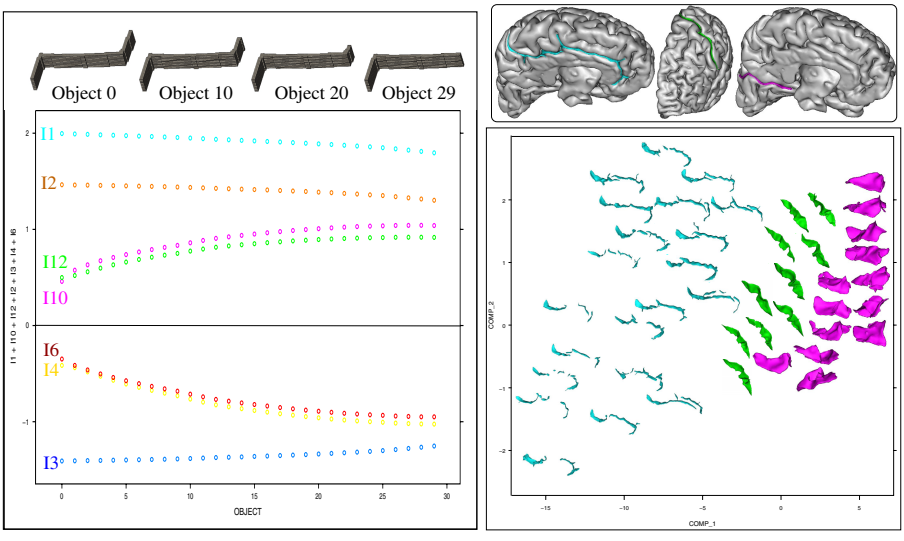

Fig. 1. Left: Variations of 7 moment invariants following continuous changes of a cylinder shape. **Right:** Two first axes of a PCA performed with 3 different sulci and 36 brains. Some of the sample points have been replaced by a snapshot of the corresponding sulcus in order to visualize the underlying shape. One can see gradual changes of the shapes, which shows that the invariant-based representation varies smoothly in the shape space.

underlying shape, and we will see further that we discovered some exceptions. Nevertheless, the behavior of these invariants seems to be continuous in general, except for two of them.

Studying the variability of the invariants across brains, we noticed that $I6$ and $I10$ were presenting bimodal distributions for some sulci. One mode was made up of positive values and the other one of negative values. There is no apparent correlation between the shape and the sign of $I6$ and $I10$. Furthermore, we managed to create slowly changing series of simulated shapes giving sign changes in $I6$ and $I10$. Such a series is illustrated in Fig. 2.left. This series evolves from a strong S cylinder towards a flat S by shortening both arms simultaneously. Notice that while most of the invariants behave smoothly all over the evolution, I6 and I10 fluctuate unexpectedly. They change sign three times very rapidly. To investigate this behavior further, we designed a new series using the finest-grain changes we could afford with our voxel-based representation (see Fig. 2.right). We discovered that adding only one single voxel could trigger the sign change. We do not know yet what kind of property would emerge if the shape space was sampled further with smaller voxels. The behaviour of the invariant could be continuous but very chaotic. Therefore, for further studies, we have chosen to discard $I6$ and $I10$ from our invariant-based representations.

It should be noted that our observation of the sign change of these two invariants has never been reported elsewhere. 3D moment invariants, indeed, have

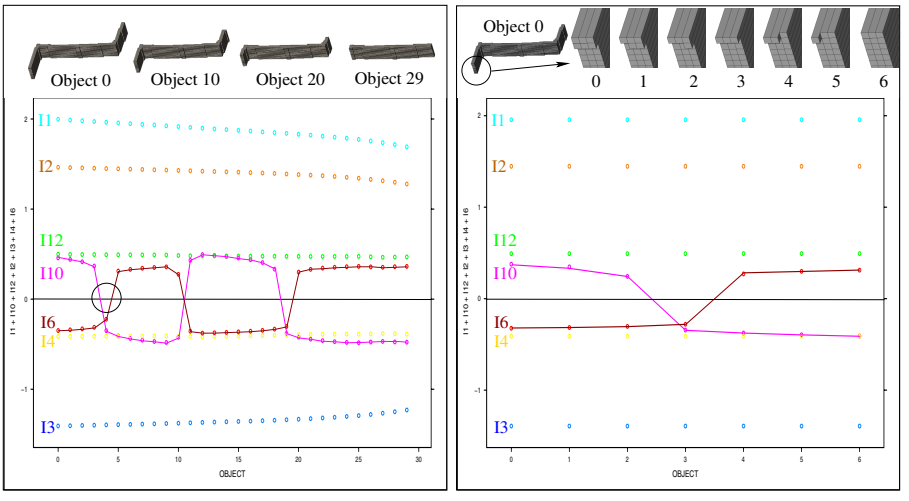

Fig. 2. Left: Variations of 7 moment invariants following continuous changes of a S-shaped cylinder. Note that I6 and I10 change signs abruptly several times. **Right:** A zoom on one of the sign change. Each step corresponds to the removal of one single voxel from the lower arm of the object.

mainly been considered as curiosities, because of the complexity of their derivation. Therefore, they were almost never used for actual applications. The invariants are made up of a sum of several hundreds of homogeneous polynomials of the central moments. This complexity is bound to hide some singularities. In fact we observed some sign change for a few other invariants, but for less than one percent of our total dataset. Therefore we decided to keep the ten remaining invariants as the basis of the representation used in this paper.

A second investigation aims at verifying that the information embedded in the invariants can distinguish the kind of patterns that characterize the cortical folds. For this purpose, we merge the datasets of several sulci, and we plot the resulting dataset using the two first axes of a principal component analysis. In most cases, the plot is made up of several clouds of points corresponding to the different sulci. An example is shown in Fig. 1.right. These clouds overlap more or less according to the shape of the sulci. The fact that each sulcus leads to a consistent cloud means that the invariant-based representations can be used to cluster groups of folds with similar shapes. To conclude, our different investigations have shown that the set of ten moment invariants can be considered as a good representation of the 3D shapes of the folds.

3 Clustering Sulci into Patterns

The cortical folding process can be considered as a chaotic phenomenon, in the sense that a slight difference among the factors that influence this process can

lead to a large difference in the folding patterns. The folding patterns are the result of the competition between a large number of forces influencing brain geometry. Some of these forces are for instance the tensions induced by the long fiber bundles trying to pull two different parts of the cortex as close as possible [9]. In our opinion, the huge variability observed at the level of the folding patterns results from the large number of attractors embedded in the dynamics of this folding process. Modeling this dynamics globally is at the present time largely beyond reach. In this paper, however, we focus only on local aspects: we try to infer automatically, sulcus by sulcus, some of the alternative patterns resulting from this multiplicity of attractors. The goal is not yet to perform an exhaustive enumeration of all the possible patterns but to detect a few very contrasted patterns. If such patterns can be defined, we hypothesize that their relative frequencies could be different in some patient populations compared to control subjects. Developmental pathologies, indeed, could modify the folding dynamics and favor some specific folding patterns for some of the sulci. Hence, the folding patterns could provide some signatures useful for diagnosis.

With this goal in mind, we designed a method looking for such patterns on a sulcus by sulcus basis. For each sulcus, the method is looking for reasonably large groups of brains that exhibit a similarity in shape. Each such group is representing a pattern. At least two patterns are required and the detected patterns have to be as different as possible. It is important to emphasize that the goal here is not to assign each sulcus to a pattern. Reproducible patterns may only characterize a subset of the population.

We have chosen to address our goal using unsupervised agglomerative clustering methods. Such hierarchical approaches to clustering, indeed, fit completely our need for finding compact clusters and discarding numerous outliers. Among the variants of agglomerative clustering algorithms, we have chosen the average-linkage method for its robustness and space conserving properties [8]. In the following, the algorithm is applied to the n different instances of a given sulcus. Each instance comes from a different brain and is represented by a vector of ten moment invariants. The clustering algorithm is building a tree by successive agglomeration of the closest clusters of sulci. At the initial stage, each sulcus is a singleton cluster. The two closest sulci are joined first, leaving us with $n-1$ clusters, one of which being the pair of joined sulci. In all succeeding steps, the two closest clusters are merged.

The complete agglomerative hierarchical tree has n levels. The lowest tree level corresponds to n singletons and the highest level to one single cluster gathering the whole set. To decide which level gives the best partition, we introduce a ratio defined as the average pairwise distance between cluster centers divided by the average cluster compactness. The compactness of a cluster is simply the average pairwise distance between the sulci. Note that the ratios are computed using only a subset of the sulci making up the kernel of each cluster. This kernel is defined as the t tightest sulci of the cluster, namely the t sulci that agglomerated the earliest in the hierarchy. Only these sulci will be considered as belonging to the putative patterns. Furthermore, only the clusters including at least t sulci

will generate a putative pattern. A good set of patterns should provide a very high ratio of distance versus compactness.

The optimal kernel size t is defined as the size providing the most reliable set of patterns. The null hypothesis is that the sulcus set follows a Gaussian distribution. If the null hypothesis is true, our sulcus set should embed only one single pattern. Therefore, whatever the choice for t, the ratio should be low. In order to evaluate the distribution of the ratio for each value of t, 1000 different random sets are generated using a multivariate Gaussian distribution based on the covariance matrix of the sulcus of interest. The hierarchical tree is built up for each of these random sets and the best ratio is obtained for each value of t. From these ratio distributions, we can test the null hypothesis: for a given t, we compute the p-value as the percentage of random sets providing a better ratio than the best ratio obtained with the actual data. Then, the best t is simply the one providing the best p-value. Finally, this p-value is used to evaluate the quality of the set of patterns associated with this best t.

4 Results

The method has been applied to the ten largest sulci of the left and right hemispheres using the database of 36 manually labelled brains. Among the 20 sulci, 3 provided a set of patterns endowed with a p-value lower than 0.01 (the left cingulate sulcus, the left inferior precentral sulcus, and the superior frontal sulcus). The sulcus providing the best p-value (0.001 for $t = 4$) is the left cingulate sulcus (see Fig. 3). For this sulcus, a first pattern is made up of sulci presenting a large anterior interruption, a second pattern is made up of sulci presenting a smaller and more posterior interruption, and a third pattern is made up of sulci appearing continuous. It should be noted that these patterns can not be inferred just from the number of connected components. Indeed, the sulci of the third pattern are only apparently continuous: some of them are made up of several connected components overlapping each other when the sulcus is viewed from above. In fact, the moment invariants are blind to connectivity. Therefore, these three patterns would be interpreted more reliably in terms of shape than in terms of interruption. For instance, the first pattern corresponds to sulci much deeper in the posterior part than in the middle, while the last pattern corresponds to sulci with more homogeneous depth.

It is not an accident that the cingulate sulcus provides the best p-value. This sulcus is one of the sulci with very varied shapes and many interruptions. According to Ono's atlas, around 60% of the instances of this sulcus have no interruption, around 24% have two segments with a posterior interruption or an anterior interruption, and around 16% are divided into three segments [7]. It should be noted that the small size of the manually labelled database used here (36 brains) prevents the detection of rare patterns. Therefore, much larger databases will be required to achieve a more exhaustive pattern enumeration.

To illustrate possible applications of our pattern inference process, we use the three patterns obtained with this database to mine the left cingulate sulci of

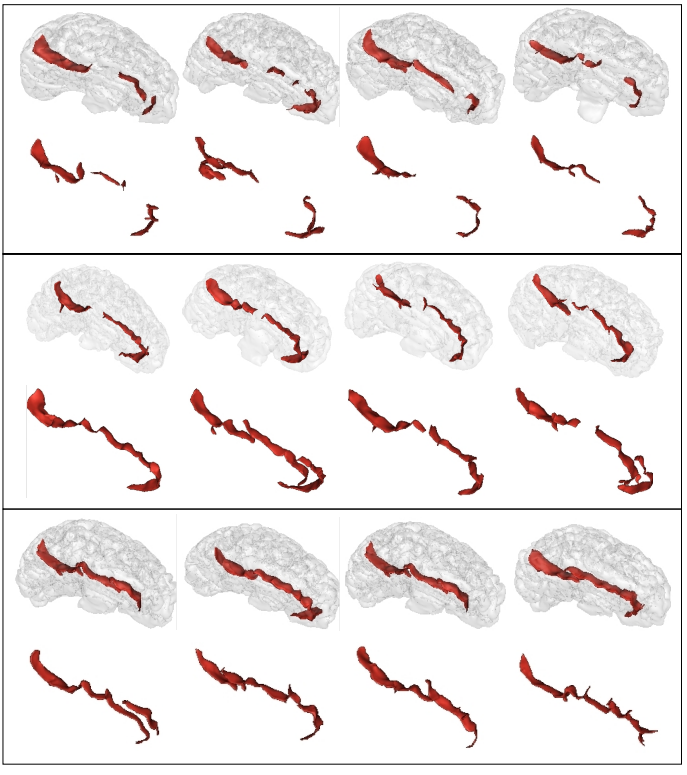

Fig. 3. The three patterns detected for the left cingulate sulcus. *Row 1,3,5:* the four tightest instances of each pattern in manually labelled database. *Row 2,4,6:* the four closest instances to the above pattern center in automatically labelled database.

another database, here the ICBM database. We selected in this database the closest samples to each of the three patterns. We observed that the shapes of these samples are consistent with the corresponding patterns (see Fig. 3). Note that when the anterior part of the sulcus is made up of two parallel folds (fourth row of Fig. 3), it is equivalent to a deeper sulcus for the moment invariants. To project the patterns from the first database onto the second database, we classify the sulci according to the closest distance to the pattern centers. This classification attributes 14 brains to the first pattern, 97 to the second and 35 to the third. It was found that the percentage of females increases gradually from 36% in the first class, to 41% in the second and to 49% in the third (global percentage of female is 42%).

5 Discussion

Compared with the traditional methods to characterize the folds by certain parameters such as the length, the number and position of interruptions, etc,

the moment invariants that are used in this study provide a more comprehensive description of the 3D shapes. The drawback of using the moment invariants is that the clusters found do not necessarily provide an easy physical explanation that can be readily observed. This presents a difficulty in the description of the patterns found. Furthermore, even when we found an explanation of the possible groupings, that explanation is not necessarily the reason that groups the folds by the moment invariants. However, the final goal is not necessarily to characterize the patterns physically using words or a set of measurements such as the length or number of interruptions. These patterns, indeed, can be directly described by their moment invariants. A future direction of research could consists in mixing different kinds of features, for instance simple morphometric parameters like length and depth with 3D moment invariants. This approach, however, requires large databases to overcome problems induced by the curse of dimensionality.

In this paper, the search for patterns has been applied to folds already labeled, either manually or automatically. A more ambitious project, that will be addressed in the future, will be the design of methods looking for patterns without the knowledge of the traditional nomenclature. Such an approach applied to large databases could reveal patterns beyond the reach of the first anatomists.

One possible use of the patterns we found is to compare the frequency of occurrence among normal and patient datasets. Similar comparisons can be carried out on other datasets for pure neuroscience questions: musicians versus athletes, kids with an early development on language versus an early development on motor-skills, etc. The hypothesis is that a certain developmental event or a certain strong training would leave an observable imprint on the folding patterns.

References

1. Welker, W.: Why does cerebral cortex fissure and fold? In: Cerebral Cortex, vol. 8B, pp. 3–136. Plenum Press, New York (1988)
2. Le Goualher, G., Procyk, E., Collins, D.L., Venugopal, R., Barillot, C., Evans, A.C.: Automated extraction and variability analysis of sulcal neuroanatomy. IEEE Trans. on Medical Imaging 18(3), 206–217 (1999)
3. Lohmann, G., von Cramon, D.Y.: Automatic labelling of the human cortical surface using sulcal basins. Medical Image Analysis 4(3), 179–188 (2000)
4. Fillard, P., Arsigny, V., Pennec, X., Hayashi, K.M., Thompson, P.M., Ayache, N.: Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. Neuroimage 34(2), 639–650 (2007)
5. Rivière, D., Mangin, J.F., Papadopoulos-Orfanos, D., Martinez, J.M., Frouin, V., Régis, J.: Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. Medical Image Analysis 6(2), 77–92 (2002)
6. Mangin, J.F., Poupon, F., Duchesnay, E., Rivière, D., Cachia, A., Collins, D.L., Evans, A.C., Régis, J.: Brain morphometry using 3D moment invariants. Medical Image Analysis 8, 187–196 (2004)
7. Ono, M., Kubik, S., Abernathey, C.D.: Atlas of the cerebral sulci. Thieme (1990)
8. Kaufman, L., Rousseuw, P.J.: Finding groups in data. Wiley series in probability and statistics (1990)
9. Van Essen, D.C.: A tension-based theory of morphogenesis and compact wiring in the central nervous system. Nature 385, 313–318 (1997)