Inference of a dictionnary of cortical folding

Zhongyi Sun

UNIVERSITE PARIS-SUD XI
Faculté des Sciences d'Orsay

THÈSE DE DOCTORAT

SPÉCIALITÉ : PHYSIQUE

École Doctorale « Sciences et Technologies de l'Information des Télécommunications et des Systèmes »

PRÉSENTÉE PAR : Zhongyi SUN

SUJET: Inférence d'un dictionnaire des motifs des plissements corticaux

Soutenance prévu 7 Juin 2011, devant les membres du jury :

Cachia Arnaud	Rapporteur
Colliot Olivier	Examinateur
Durand Emmanuel	Examinateur
Malandain Grégoire	Rapporteur
Mangin Jean-François	Directeur de thèse
Régis Jean	Examinateur

Résumé

Cette thèse vise à faire émerger de nouvelles descriptions de la variabilité des plissements du cortex humain en s'appuyant sur des techniques de fouilles de données. L'objectif principal est la conception d'algorithmes permettant de découvrir des motifs de plissement spécifiques à une sous-population d'individus. Le but final est de réaliser un dictionnaire de ces motifs et de les associer à des particularités cognitives ou architecturales, voire à des pathologies. Deux stratégies de « clustering » sont proposées pour mettre en évidence de tels motifs. La première repose sur des descripteurs de formes globaux correspondant aux invariants de moment 3D, la seconde repose sur l'estimation d'une matrice de distances entre chaque paire d'individus. Un algorithme de clustering dédié est conçu pour détecter les motifs les plus fréquents de manière robuste. Une technique de réduction de dimension est utilisée pour mettre en évidence les transitions entre motifs au sein de la population. Les méthodes algorithmiques proposées sont utilisées pour étudier la forme du cortex sensori-moteur d'une population de gauchers contrariés. Des résultats originaux sur le lien entre la forme du sillon central et la latéralité manuelle sont mis en évidence. Les méthodes développées sont ensuite utilisées pour construire le premier dictionnaire des motifs observés dans les plissements corticaux issu d'une approche algorithmique.

Abstract

This thesis aims at proposing new descriptions of the variability of the folding of the human cortex using data mining. The main objective is the design of algorithms detecting folding patterns specific to a sub-population. The long term goal is the constitution of an exhaustive dictionary of all the folding patterns enriched with links to cognitive or architectural specificities, or to pathologies. Two clustering strategies are proposed to detect such patterns. The first one is based on global shape descriptors called the 3D moment invariants, the second one implies the computation of a pairwise distance matrix. A dedicated clustering algorithm is designed for robust detection of the most frequent patterns. A dimension reduction strategy is proposed to study the transition from one pattern to another across the population. The proposed framework is applied to the study of the shape of the sensori-motor cortex of a population of left-handers forced to write with the right hand. Original discoveries relating the shape of the central sulcus to handedness are achieved. The framework is finally used to build the first computerized dictionary of the cortical folding patterns.

Table of Contents

List of Figures and Illustrations

Epigraph

A perfection of means, and confusion of aims, seems to be our main problem.

- Albert Einstein

Chapter One: Cortical Folding and Cortical Morphology

1.1 Summary

The primary goal of the thesis work is applying computational methods to analyze the cortical folding patterns. In this chapter, a general overview is given to the fundamental subjects of concern: the phenomenon of cortical folding and the computational morphometry of cortical folding. The computational morphometry provides the general framework where the thesis work fits in. Within this framework, the cortical folding patterns are studied applying computational methods using magnetic resonance imaging data.

The rest of the thesis is organized as follows: Chapter One gives an introduction to the subjects the most relevant to this work; Chapter Two introduces the approach of using clustering algorithms on selected morphometrical features of cortical folding; Chapter Three further explores various approaches of finding comprehensive information in cortical folding; Chapter Four applies the methods developed to a real dataset, illustrating the exciting potentials of such analysis; Chapter Five presents the summary of the detailed dictionary of cortical folding patterns and Chapter Six concludes the thesis work.

1.2 Introduction

As of all the work combining different fields, in this case, neuroscience and computer science, it is of pivotal importance to gain sufficient understanding of the fields involved. The questions asked need to be well understood before the right solutions can be found. Consequently, this thesis work begins by understanding the neuroscience issues behind. The computational methods then need to be selected or designed to answer these specific neuroscience questions. In this work, the journey starts with the cortical folding process.

The cortex of the brain folds during development. The folding process, mysterious and fascinating, is not yet well understood. Does the cortical folding 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? There exists a wealth of knowledge on the subject. Thanks to recent advances in software dedicated to automatic recognition of cortical sulci (Fillard et al., 2007; Le Goualher et al., 1999; Lohmann and von Cramon, 2000; Riviere et al., 2002; Welker, 1988) issues regarding cortical folding can now be tackled using large brain databases (Mangin et al., 2004b).

Each brain looks different and none of them looks exactly like the ones in the text books. Refer to **Fig 1.1**, examples of individual folding patterns are shown. It can be observed that huge amount of folding variability exists.

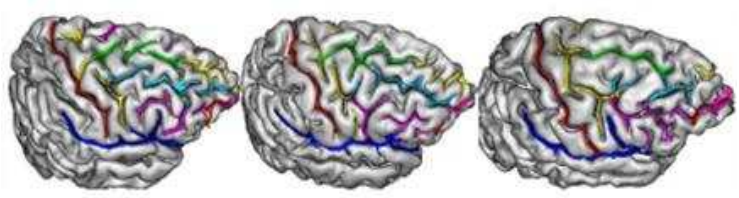


Fig 1.1. The variability of cortical folding

Right hemispheres of three individuals, the cortical folds of the frontal lobe are highlighted according to the traditional nomenclature.

Current studies of this variability focus mainly on simple morphometric features, such as 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 software dedicated to automatic recognition of the sulci.

The most detailed description of the sulcus variability has been proposed in the atlas of Ono (Ono et al., 1990). 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 variability of the sulcus interruptions. In a way, the initial goal of the thesis work is to automate the work performed by Ono. We want to discover folding patterns that can be observed for a subset of the population. Furthermore, we want to find links between folding patterns and function.

An overview of the study of brain cortical folding patterns is discussed in the next section. The field of computational cortical morphometry is discussed further in section 1.4, the focus is on the analysis of cortical folding. Note that it is impossible to cover in every detail these two vast domains; let alone many related subjects not directly concerned in the thesis work. The goal of this chapter is to give enough background knowledge to pave the way for the chapters that follow.

1.3 The study of cortical folding patterns

Different conceptual approaches are undertaken to map the brain: the functional (initiated by Broca, followed by Jackson), the cytoarchitectural (initiated by Baillarger, followed by Ramon y Cajal and later Brodman) and the study of sulci and gyri. In this work, the third approach, the study of sulci and gyri, is taken.

It is fascinating to look at the brain folding patterns. As stated by Welker: “the most striking, interesting, yet poorly understood gross morphological features of the cerebral hemispheres in mammals are the diverse and complex arrangements of their cortical gyri and sulci” (Welker, 1988).

Indeed, while the nuclei in the spinal cord and the brain stem of the mammals are quite similar, the cerebrum and cerebellum exhibit great variation in size, shape and convolutional complexity. The size and amount of convolution of the mammalian brain

has been linked to intelligence. It is believed that those with bigger and more convoluted brains possess more behavioral complexity and higher intelligence. Comparative studies revealed that cortical thickness and columnar architecture differ relatively little in different mammals. The huge difference lies in cortical surface area. Compared to macaque monkeys, the surface area of the human brain is approximately 10 times greater, whereas the thickness of the human cortex is only twofold greater (Barondes et al., 1997). The larger brains have more gyri and sulci, the variability of the three-dimensional convolutional patterns increases as well (Jerison, 1973), see **Fig 1.2** for examples.

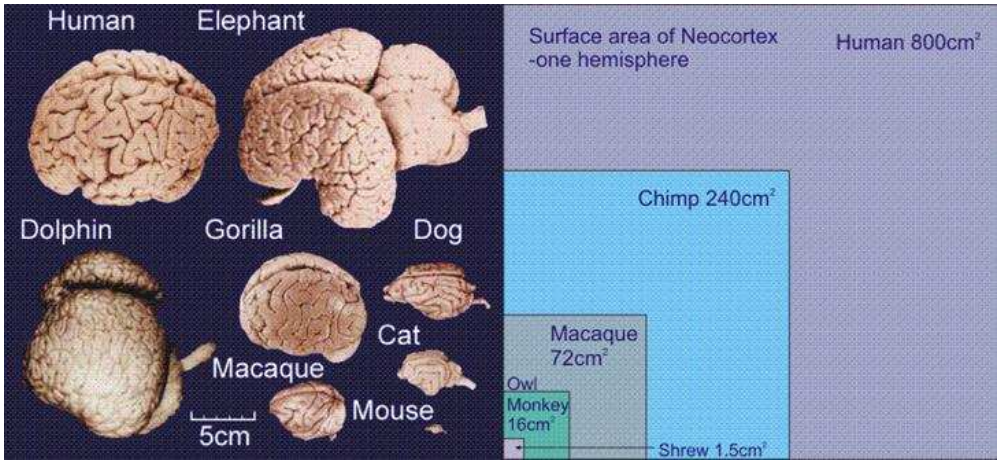


Fig 1.2: The brain of different species

It is believed that during evolution, a greater number and diversity of brain functions were achieved by increasing the surface area of cerebral neocortex. The increase in surface area is achieved by mechanical buckling, infolding, and fissuring. The folding of the cortex might provide more functional modules (Jerison, 1973).

Historically, the importance of folding patterns of the cortex was questioned: “similar to the loops of small intestines that seemed to lie in no particular order” (Edwin Clarke, 1973). In the early 1800s attention was drawn to the possibility of localizing specific mental faculties to specific gyri by Gall and his followers (Clarke, 1968). Phrenology is history now, but this initiative leads to many descriptive studies of the cerebral convolutions (Welker, 1988). These studies include humans as well as primates. As

advances were made in experimental methods and technology, this interest in localizing specific functions to specific convolutions diminished. More studies focused on the microscopic structure and architecture of the cerebral cortex. However, there starts to be more structural, connectional, and functional studies that reveal correlates of gyri and sulci to brain function (Watson et al., 1993; Welker, 1988). These evidences suggest that brain morphology of the folding patterns might be linked to brain connection and function. It is thus interesting to explore further the folding patterns with more modern and automated computational methods, and try to link these patterns to functions or behaviors.

1.3.1 Gyri and Sulci

The folding patterns are already very varied when observed from the brain surface. When the cortex is viewed by dissection or serial sections, the gyri and sulci appear to be much more complex (Welker, 1988). Traditionally, a nomenclature system is set up to describe the convolution of the surface of the human brain (Anatomica, 1983; Clemente, 1985; Ono et al., 1990), refer to **Fig 1.3** for an example. Due to the enormous variability that exists in folding patterns, achieving a unified conception for the description of these sulci and gyri is extremely difficult. One interesting example of such attempt for a systematic understanding of the sulci and gyri is the sulcal root model (Regis et al., 2005). Indeed, a unified scheme cross different species can be confirmed only when the neurological data concerning developmental, architectural, connectional, physiological and chemical features are available.

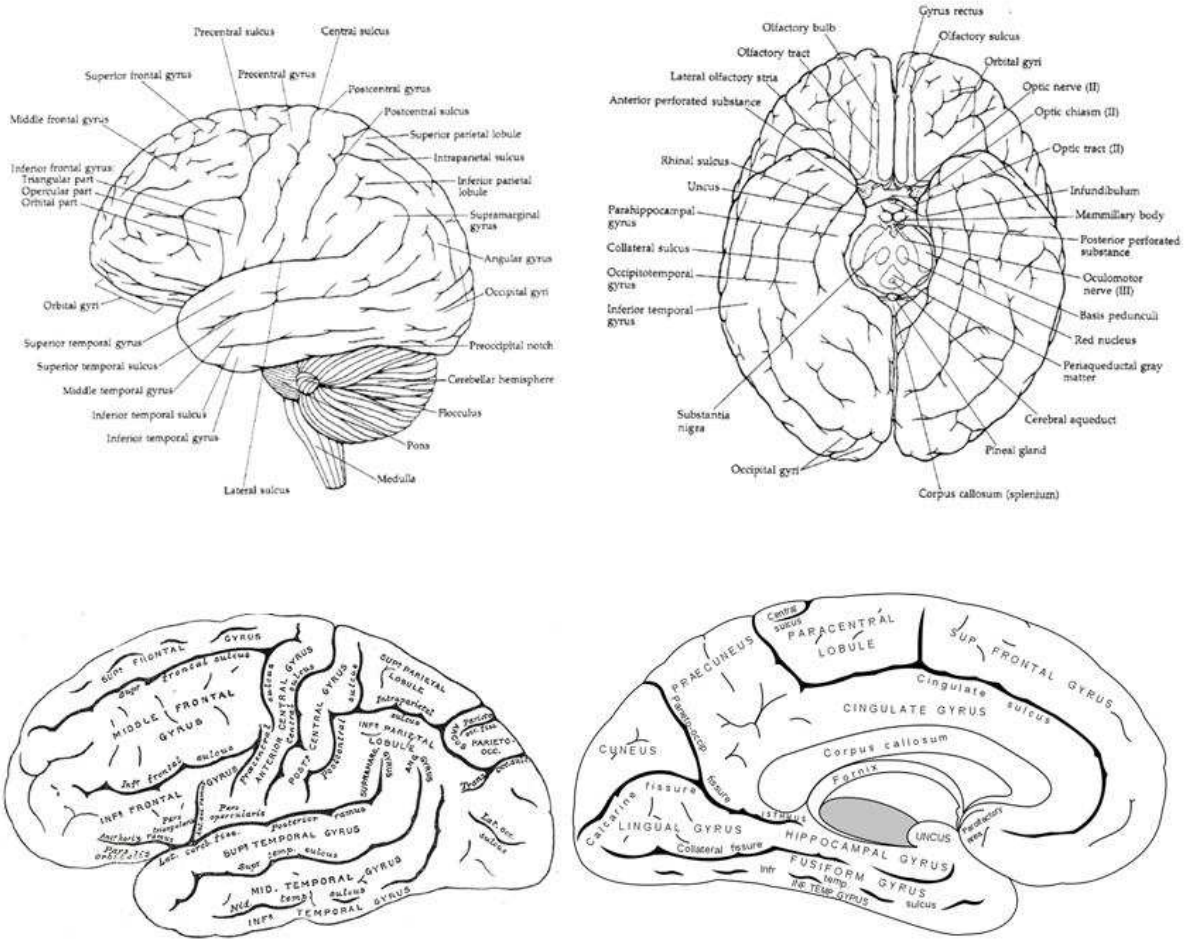


Fig 1.3 Gyri and Sulci, an example of the text book presentation of nomenclature.

Gray's Anatomy: The Anatomical Basis of Medicine and Surgery (British Edition. 38th Ed)

1.3.2 Development of the cerebral cortex and gyrogenesis

Human cerebral cortex development may be divided into three, partly overlapping periods (Marin-Padilla, 1990): embryonic period; intermediate, fetal or migration period and the perinatal period. The perinatal period starts about the 24th week of gestation.

Disorders of neuronal migration are likely to occur in the fetal period so defined. On the other hand, abnormalities of the structural organization of the cerebral cortex are

common in the perinatal period. At about 22 weeks of gestation, the first thalamocortical synapses are formed within the cortical plate, cortical interneurons and dendrites of layers III and V pyramidal neurons are developing simultaneously. In the period of 22–26 weeks of gestation, dense granularity is observed in the position of the future layer VI. The six-layered adult laminar pattern gradually appears after 28 weeks of gestation. At 1-2 years after birth, major outgrowth of dendrites occurs, both for pyramidal and for nonpyramidal neurons. Mature dendritic extension has been reached at 2–4 years of age. Finally, the mature level of outgrowth is reached at about 3–4 years (Uylings, 2001). Certain cortical areas such as the frontal and parietal cortices keep increasing until the age of 12–13 years (Hori, 2006).

In humans, gyrogenesis usually starts during fetal development. In ferrets, this process starts shortly after birth (Neal et al., 2007). While the gyral tissue continues to expand, some areas (the sulcal roots) remain in a relatively stable position (Regis et al., 2005). Gyrification in humans reaches adult values around age of 10 years (Armstrong et al., 1995). The primary effect of the folding process is an increase of surface area relative to volume, which correlates with an increased number of neurons (Panizzon et al., 2009). This increase is presumed to enhance the computational capacities of the cortex with some metabolic and connectivity limits (Wen and Chklovskii, 2008).

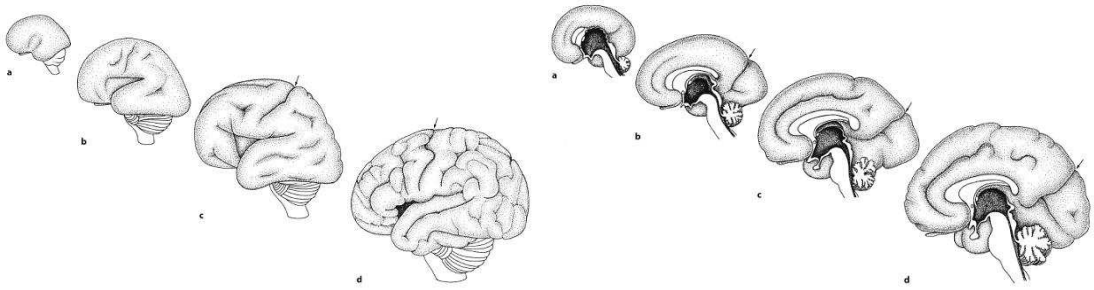


Fig 1.4 Development of the human brain

Lateral and medial views of the developing human brain in the fourth (**a**), sixth (**b**) and eighth (**c**) gestational months, and in a neonate (**d**).The *arrows* indicate the position of the central sulcus (Hori, 2006)

Gyrogenesis is composed of an intriguingly complex series of events. Without going into the details of each, the processes hypothesized during gyrogenesis include: neuronal differentiation and dendrogenesis, neuronal orientation, afferent arrival, penetration, fasciculation, and arborization, synaptogenesis, glial proliferation and ensheathment, laminar aggreration and segregation, rearrangement of cell adhesion molecules and related membrane structures, and the differential development of gyri and lobules. From the perspective of external morphology, differential development of different gyri and lobules affects their relative width, height, shape, orientation, and spatial pattern. The gyral crowns, sulcal walls, and fundi are constructed differently and according to different developmental timetables (Welker, 1988).

The cortical folding is a sequential process. Refer to **Fig 1.4** for the external morphology of the cortical folds during development. The Sylvian fissure and insula can be recognized at the 14th gestational week as a shallow indentation on the lateral surface of the cerebrum. Cerebral sulcus formation begins around the 16th gestational week with the appearance of the parieto-occipital and cingulate sulci. Central sulcus formation is seen in the 20th to the 21st week (Armstrong et al., 1995; Feess-Higgins and Larroche, 1987).

The classification of sulci into primary, secondary and tertiary sulci has been adopted. Definitions are proposed but the precise classification remains controversial. The definition based on comparative anatomy defines the primary sulci as those that can be found in all gyrencephalic primates. The ontogenetic approach defines the primary sulci as those generally appearing before 30th week of gestation (Feess-Higgins and Larroche, 1987). The secondary and tertiary sulci are those that give the cortex the adult appearance (Tamraz and Comair, 2006). The definition concerning the orientation of the sulci is as follows: primary sulci are oriented perpendicular to the neuraxis, whereas secondary sulcus formation is parallel to the neuraxis, the tertiary sulci bind the primary and secondary sulci (Hori, 2006).

Normal gyration and sulcal pattern has been studied prenatally and in preterm and term neonates with MRI. Brain maturation was found to start in the central area and to proceed

towards the parietooccipital cortex (Ruoss et al., 2001). The frontal cortex develops last. As more knowledge concerning early brain development becomes available, the classification of sulci could be converged.

It should be kept in mind that gyrogenesis is not an isolated process. The central nervous system as a whole is developing as well, together with the development of other tissues and organs such as the eye, the heart, the teeth etc. **Fig 1.5** gives a timing of these events. More specific to the cortex, while the gyri and the sulci are forming, the six-layered adult laminar pattern is forming simultaneously (refer to **Fig 1.6** as an example). While it is interesting to observe the cortical folding process, it is important to put cortical folding into the context of the human development as a whole.

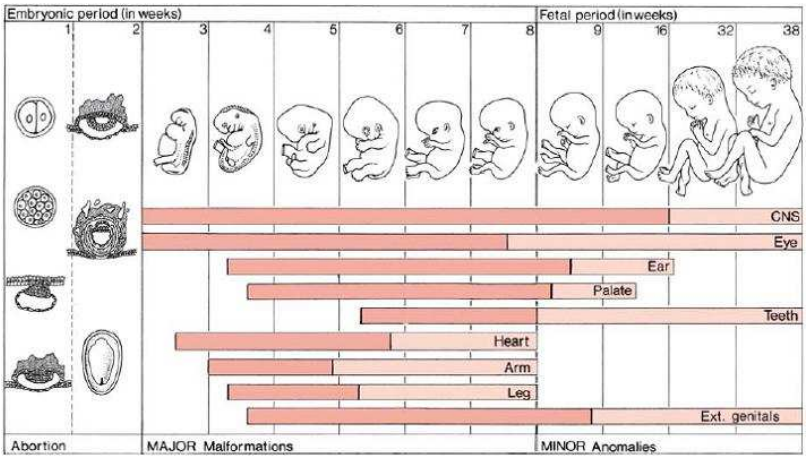


Fig 1.5 Critical periods of human development

Critical periods in human development. In the *horizontal columns*, the period of major complications is shown in *red*, that of minor anomalies in *light red*. (Hori, 2006)

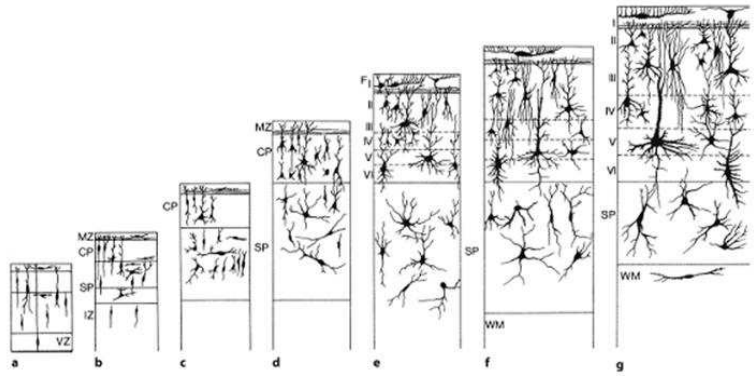


Fig 1.6 Prenatal development of neurons in the prefrontal cortex

Golgi-stained sections at 10.5(a), 13.5(b), 17(c), 19-25(d), 26-29(e) and 32-34(f) weeks of gestation and a neonate(g). CP: cortical plate, FI-VI: fetal cortical layers, IZ: intermediate zone, MZ: marginal zone, SP: subplate, VZ: ventricular zone, WM: white matter. (Hori, 2006)

1.3.3 Relation of architectonic and function to gyrus

The parcellation of neocortex into structurally different cytoarchitectonic areas has a long history; the criteria used to differentiate the areas differ from one author to another. It is generally accepted that architectonic distinctions alone do not provide an adequate or accurate view of areal differences in cortical organization (Welker, 1988).

Different architectonic fields often occupy different, but adjacent, gyri. When the border of the transitional zones lies at the fundus of the interposed sulcus, such sulci have been called *limiting sulci*. Examples of limiting sulci are the central, cingulate and sylvian sulci. However, many sulci lie within a single architectonic field, within-field sulci are called *axial sulci.* Examples are the superior and inferior precentral sulci and the calcarine sulcus. In some cases a single gyrus contains two or more architectonic fields; an example is the postcentral gyrus of primates. Refer to **Fig 1.7** for the Brodman's areas relative to the gyrus and sulcus.

One of the most striking evidence for a precise correlation of functional areas with gyral and sulcal features came from mapping studies of somatosensory cerebral cortex in

raccoons (Welker, 1988). The functional areas of the forepaw correlate well to the cortical areas separated by the folding. The crown cortex has larger or more densely activated neuronal populations, compared to gyral walls and fundi. Additional studies of the raccoon found that not only the ventrobasal thalamus, but also the dorsal column nuclei, were subdivided into as many subnuclei as there were gyral crowns in somatosensory cortex (Welker, 1988).

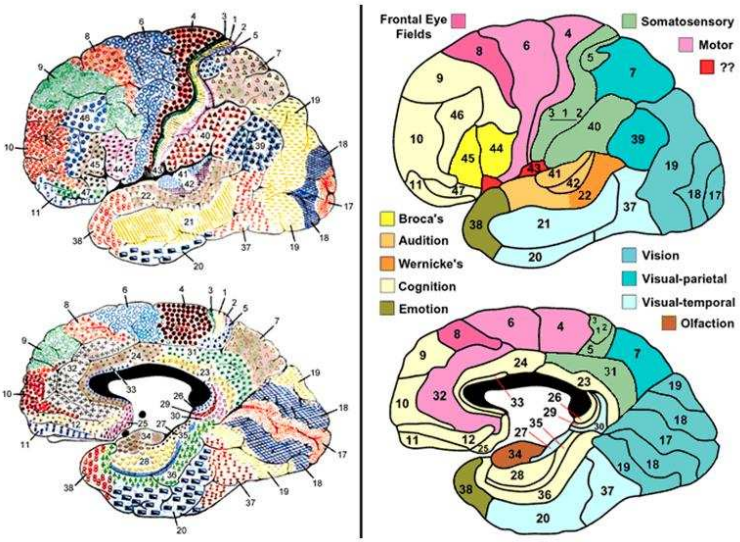


Fig 1.7 Map of the cerebral cortex in man (Brodmann, 1909).

1.3.4 Cortical Connections

To have a better understanding of cortical folding, the underlying fiber connections cannot be overlooked. Large and well-organized thalamic nuclei send projections to, and receive reciprocal connections from, different specific cortical gyri or gyral groups. *Limiting sulci* which lie at the borders or cortical areas receive projections from different thalamic nuclei or nulear complexes. There are also cases where adjacent gyri separated by *axial sulci* within a single cortical field are interconnected with different but adjacent thalamic *sub*nuclei. A well-studied example of this is found in the connections of the

subnuclei of the raccoon's ventrobasal complex with the several cortical subgyri within somatosensorry cortex (Welker, 1988).

Other than cortical-thalamic connections, numerous cortical-cortical connections exist. In human, the superior and inferior longitudinal fascicule, the uncinate fasciculus, and the cingulum are some of the largest fiber bundles which interconnect major lobes of the cerebral hemisphere (Clemente, 1985). The intergyral "U" fiber connections are also demonstrated (Krieg, 1966). It is shown that adjacent gyral crowns are richly interconnected, whereas fundic cortex is sparsely interconnected. Regarding interhemispheric connections, experiments reveal that most topographically homologous gyri of the two hemispheres have reciprocal connections that are symmetric. Many gyral crowns project to gyral crowns, and sulcal walls to sulcal walls (Welker, 1988). However, some cortical gyri send projections primarily to the walls and fundi of certain sulci in the opposite hemisphere. It is interesting to notice that in several mammals, gyri that contain specialized sensory or motor representations are devoid of reciprocal interhemispheric connections. The examples include the somatosensory hand cortex, the visual foveal cortex and the primary auditory cortex in some mammals. Local circuits exist between different parts of a single cytoarchitectonic field located on one gyrus, as well as between different adjacent subfields on the same gyrus (Welker, 1988).

Finally, it should be kept in mind that many gyral regions send descending projections not only to specific thalamic nuclei, but also to specific basal ganglia, brain-stem, cerebellar, medullary and spinal cord nuclei.

1.3.5 Pathological gyral and sulcal patterns

Abnormal gyral and sulcal patterns have been observed. There are terms describing the anomalies of cortical development in humans and animals, which include: agyria (lacking gyri), pachygyria (broad gyri with thin cortex), ulegyria (narrow, distorted, and scarred

gyri), microgyria (abnormally narrow gyri), polymicrogyria (supernumerary tiny gyri), polygyria (an unusually large number of gyral formations), schizogyria (gyri with disrupted continuity), ectopic gyri (gyri that occur in unusual places), and cortical warts (small innervated cellular protrusions of cortex) (Welker, 1988).

Abnormalities of cortical development produce alternations in form and pattern of gyri and sulci, many of these are also associated with sensory, motor, cognitive, and motivational disorders. The interactions during development are complex. The difficulty in deciphering the folding mechanism based on abnormalities is that these abnormalities usually affect many levels of development, and little is known of the time and nature of the cause (Welker, 1988).

One interesting example illustrating the link between the underlying connections and the sulcal pattern is the case of brains without corpus callosum (callosal agenesis). In these cases, the bundles of Probst become longitudinal instead of crossing the hemispheres as corpus callosum. Typically, the cingulate region shows more radial arrangement. Refer to **Fig 1.8** for the abnormal bundles and abnormal sulcal pattern. The other example is the Williams syndrome, a rare disorder characterized by dissociation between language, face processing and spatial cognition. Overly social behavior is observed that is opposite to that seen in autism (Bellugi et al., 1999). It is found that the central sulcus of the patients are shorter and does not become opercularized in the interhemispheric fissure (Galaburda et al., 2001). The central sulci are also observed to be separated by unusual gyral convolutions (Hori, 2006).

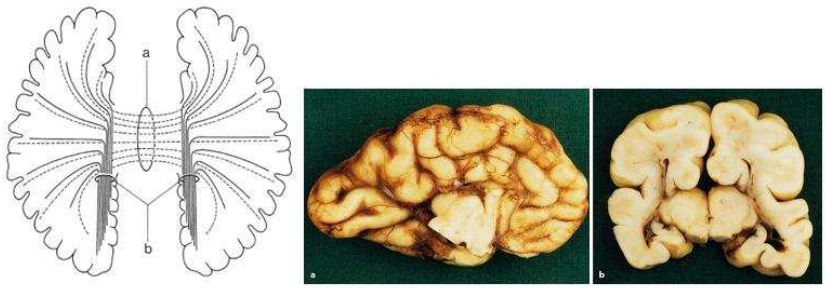


Fig 1.8 Callosal agenesis

Left: Development of the bundles of Probst. The commissural fibres of the corpus callosum in a normal brain (*a*) are shown by *broken lines*; the abnormal longitudinal bundles of Probst (*b*) which fail to cross are shown in *solid lines*. (Hori, 2006) Right: Photographs show a fetal case of callosal agenesis: **a** medial view of the brain; note radial arrangement of gyri; **b** coronal section of the brain

The examples above implies that even though folding abnormalities are most likely consequences of multiple functional abnormalities, it nevertheless links altered form to altered function in the cortex, and can shed some light on the underlying developmental mechanism.

1.3.6 Comparative and developmental studies

Comparative studies among different species provide many interesting insights into brain evolution and development. For example, it is observed that the size of the localized areas of the cortex is different among mammals. Relatively large cortical somatosensory hand area is observed in raccoons, larger rhinarial areas in pigs, larger lip area in llamas, and larger tail area in spider monkeys (Welker, 1988).

More specific to gyrogenesis, historical comparative studies reveal that gyrification is not associated with brain size or body size across species. Smaller brains can be more convoluted than bigger ones. An example is the least weasel, the smallest living carnivore, which has a highly convoluted brain smaller than that of a smooth-brained rodent (Jerison, 1973). Another well documented fact is that the cortical thickness of cerebral cortex varies relatively little (1-4mm) among brains of different mammals over a wide range of brain and body sizes, from mouse to elephant (Braitenberg, 2001).

More recent studies confirm that as the brain size increases, the cortical thickness increases only slightly, while the degree of sulcal convolutions increases dramatically (Im et al., 2008). Furthermore, the relation of gyrification to brain size might follow region-specific patterns within species. Toro et al (Toro et al., 2008) found that as the brain size

increases in humans, the cortical folding is increased more specific to the prefrontal cortex. The major implications are: the processing operations of cerebral cortex are probably the same everywhere; increasing the number of processing modules is mainly by the increase in cortical surface area.

It is interesting to note that shapes and orientations of most gyri differ in predictable ways in different species, as well as in different cortical regions in any one species. Convolutions and sulci do not appear randomly in different mammalian groups, but tend to occur in taxon-specific patterns. Within each group, there are greater similarities, despite wide variations in brain size and gyral and sulcal complexity. The study of the somatosensory cortex in raccoons by Welker (Welker, 1988) reveals that minor interanimal differences were associated with variations in the deployment of specific peripheral somatosensory projections to cerebral cortex.

The development of sulcal and gyral patterns is strongly influenced by genetic processes (Piao et al., 2004), yet studies of monozygotic twins reveal considerable differences in their surface morphology (Thompson et al., 2001; White et al., 2002). This could be due to environmental influences during early development. It is found in twin studies that the deeper and developmentally earlier sulci of the brain (i.e., the central sulcus or the sylvian fissure) are more highly correlated than the superficial or tertiary sulci, which develop mainly after birth, and appear to be more affected by non-genetic influences (Lohmann et al., 1999).

1.3.7 Explanation of gyrification and fissuration

Gyrus building is considered to be consisting of numerous constructional processes. The gyral crowns, sulcal walls, and fundi (bottom) are different in architectural, connectional and functional features (Welker, 1988). Various models have been proposed to account for the phenomena of convolutions. The non-isotropic forces might be due to the complex

interaction of differential growth of the cortical layers, cell migration, cell myelination, cortical-cortical and cortical-thalamic connectivity, synaptic pruning, brain size and metabolism (White et al., 2010).

It was first proposed that sulcation is due to cortex expansion constrained by the skull and the basal ganglia (LeGrossClark, 1945). It is later found out that removal of large amount of cortical and subcortical structure in sheep brain results in normal sulcal size and organization (Barron, 1950). So cortical folding is likely not due to constrained growth solely. Van Essen (Van Essen, 1997) proposes that neuronal connections that develop during the second trimester produce localized fiber tension which draws densely interconnected regions closer together. Tension along axon in the white matter is suggested to be the primary driving force of cortical folding. As regions of greater connectivity move closer together in an enclosed and rapidly growing brain, they form gyri (outward fold). The more sparsely connected regions drift apart and the sulci (inward fold) form. The tension, although very small for an individual axon, is summed by the very large number of neurons. The characteristic pattern of the convolutions can thus be explained by the highly specific organization of the underlying connectivity. This model suggests that differential growth of different cortical layers is a consequence rather than a cause of cortical folding.

Such a theory links brain surface morphology with regional neuronal connectivity, in a developmental framework (White and Hilgetag, 2008). The link between gyrification and axonal tension has been supported by experimental findings in the primate brain (Hilgetag and Barbas, 2006).

Alternate folding theories emphasize on mechanical factors such as abutting cortical plates (Richman et al., 1975), or the differential growth with mechanical constrains (Todd, 1982). Genetic factors likely play a crucial role in cortical folding. The genetic control of cortical development is proposed (Rakic, 1988); links between cortical folding and cytoarchitecture is confirmed (Fischl et al., 2008). Cortical folding has also been

linked to genetic factors by studying the abnormal folding of diseases with known genetic link such as William's syndrome (Gaser et al., 2006; Kippenhan et al., 2005). Indeed, the fiber development, the differential growth in cortical layers and in different sub-regions, could be a complex concerted process determined by genetics.

The major mechanical factors that contribute to the folding process, fiber pulling or differential growth under constrain, remain to be clarified in the future. Some interesting computer based models of cortical folding are proposed, which will be discussed in detail in the next section.

1.4 Computational morphometry of cortical folding

The general framework of cortical folding is introduced above. Next, we discuss some computational methods applied to the study of cortical folding. The conventional naming system seems inadequate for describing the folds; this is partly due to the variability that exists, partly due to the use of external morphological criteria alone in the description of folding patterns. One of the goals of this work is to understand better the variability in cortical folding patterns.

The recent advancement in brain imaging such as the Magnetic Resonance (MR) techniques can provide valuable information. Multiple subjects can be followed consistently in longitudinal studies. Applying modern computational methods, massive data can be analysed automatically. Computational morphometry is the field concerned with the qualification of anatomical features and changes in individual brains or brain populations (Mietchen and Gaser, 2009). This thesis work is an application of computational morphometry. In the following, only topics relevant to the thesis work are introduced. Subjects less relevant are not discussed; the aim is to introduce subjects that our work can contribute in advancing the current understanding.

1.4.1 Cortical morphology and the modeling of cortical folding

The voxel-based-morphometry (VBM) approach is used in many brain morphometry studies (Ashburner and Friston, 2000). In VBM studies, three-dimensional spatial alignment is applied; voxel-wise comparison of the local concentration of grey and white matter across populations is then carried out. When the cortex is of interest in particular, the cortical thickness analysis is often used after alignment of the cortical surfaces (Ashburner, 2009; Fischl and Dale, 2000). More related to our work of cortical folding, a group of computational methods utilizing the sulci/gyri are developed, which will be discussed here.

Regarding the cortical folding process, as discussed in section 1.3, numerous factors are involved. Different hypothesis regarding cortical folding are proposed, emphasis are put on fiber tension (Van Essen, 1997), differential growth of the sub-layers (Richman et al., 1975), and differential growth of different sub-regions of the brain (Welker, 1988). The fold formation could be mainly due to genetic factors or mechanical factors. With the advancement in MRI and computer algorithms, these models can be tested.

Computer simulation of the morphogenetic model (Toro and Burnod, 2005) has been proposed to clarify the importance of mechanical and genetic factors in cortical folding. The results suggest that convolutions could be a natural consequence of cortical growth. Such a model can produce primary, secondary and tertiary folds.

The causation of fiber tension for gyrification hence remains controversial. In another recent study, cerebral cortical folding has been modeled combining structural and diffusion tensor MRI in sheep. Finite element modeling is combined with explicit growth mechanisms to be tested. The growth mechanisms tested are white matter tension or tangential cortical growth that drives cortical folding (Geng et al., 2009). It is found that tangential cortical growth is a plausible biomechanism of sulcal root formation and hence

cortical folding. It is also shown in the developing ferret brain that, even though the axons are verified to be under considerable tension, the tension likely does not drive folding. Using computational models, it is shown that deferential cortical growth accompanied by remodeling of the subplate leads to outward folds and stress fields consistent with microdissection experiments. This result supports a mechanism involving differential growth of the layers (Xu et al., 2010).

Another interesting approach to explain the consistency and variability of the cortical folding pattern proposes a phenomenological model (Lefevre and Mangin, 2010). This model is based on reaction diffusion mechanisms, where the Turing morphogens are responsible for the differential growth of the sulci and the gyri. This model mimics the progressive folding of the cortical surface; it can generate reproducible yet variable patterns using sulcal roots (Regis et al., 2005). The study suggests that interactions between growth factors may be sufficient in the formation of consistent yet variable folding patterns. Such a model would give more emphasis to genetic factors, which subsequently determines the timing and amount of the generation of growth factors.

The mystery of cortical folding and the importance of various factors contributing to this phenomenon remain to be deciphered. The studies discussed above illustrate the exciting potential of computational methods in testing hypothesis of complex phenomenon such as cortical folding.

The cortical folding analysis is of growing interest due to the potential connection of cortical folding to white matter connections underneath (Van Essen, 1997). Recent links has also been found between cortical folding and cytoarchitectony (Fischl et al., 2008). The analysis of cortical folding would thus be valuable for the study of development and pathology (White et al., 2010). Abnormal developments possibly leave traces that can be observed in abnormal folding patterns.

Next, the different approaches in the study of cortical folding are discussed.

1.4.2 Approaches in the study of cortical folding

One way to allow comparison of cortical folding and morphology across subjects is to first align the surfaces. Many approaches can be used for such two-dimensional spatial normalization. The alignment based on folding depth and curvature on the stable sulci can be carried out. These stable sulci can be defined manually or automatically (Fischl et al., 1999; MacDonald et al., 2000). In cortical thickness and gyral surface analysis (Fischl et al., 2004), the alignment is carried out first.

Beyond cortical thickness and cortical surface analysis, the nature of cortical folding can be studied. An index is used to quantify the extent of folding, the gyrification index (GI). The GI is first defined as the ratio between the lengths of coronal outlines for the brain including and excluding the sulcal regions (Zilles et al., 1988). This approach leads to interesting findings. For example, it is found that the GI increases dramatically in the third trimester of development, then remains more or less constant throughout life (Armstrong et al., 1995).

The limitation to this method of measuring the GI is that it is obtained in two-dimensions, the measurements could be biased. The possibility of defining three-dimensional GI is being explored. The 3D GI can be defined locally from the geometry of the cortical surface (Schaer et al., 2008; Toro et al., 2008); it can also be defined globally (Rogers et al., 2010) or both at the global level and the sulcus level (Cachia et al., 2008). In the study of premature infants, specific GI is designed (Dubois et al., 2008).

Other than the thickness and gyrification index, other features have been used to study cortical folding. The easier to access features are the length, the depth and the surface area of the sulci (Mangin et al., 2004b). Another interesting feature useful to monitor

aging is the opening of the folds, defined as the distance between the two walls of a particular sulcus (Kochunov et al., 2005).

Our methods developed in this work are using a more comprehensive measure, the 3D shape of the folds. The shape needs to be detected and further represented reliably for pattern analysis. The details of the representation of the sulcal shape can be found in Chapter Two and Chapter Three. The sulcal identification and extraction is done by using the software BrainVISA (Mangin et al., 2004b). The image processing algorithms in BrainVISA obtain the shapes of the sulci in three steps. First, a hollow object made up of gray matter and cerebrospinal fluid is extracted from the T1-weighted image. Second, the object from step one is skeletonized to obtain the hemisphere hull and the numerous medial surfaces of the cortical folds. Third, the skeleton obtained in step two is split to separate the folds from each other and from the hemisphere hull. The naming of the sulci can be carried out automatically (Mangin et al., 2004b; Perrot et al., 2009b).

Sulcal extraction and analysis is carried out by other teams as well. In the work of Le Goualher (Le Goualher et al., 1999), the active ribbon method is used to extract the superior and the fundus trace of the sulci; these are then extended to the surface of the whole sulcus. The labeling is semi-automated, in the sense that for the labeling of each sulcus, the user is given a list of choices as the most likely candidates. This list is based on the priors for the expected sulcal spatial distribution. The sulcal shape can then be compared statistically to investigate such issues as the link of sulcal shape to genetics in twin studies (Le Goualher et al., 2000).

Various other methods are used to extract the fundi of the sulcus (Kao et al., 2007; Li et al., ; Seong et al., 2010; Shi et al., 2009) or the surface sulci lines (Fillard et al., 2007; Tu et al., 2007). The results of the analysis using the top or bottom of the sulci would provide different but interesting information regarding the sulcal pattern and cortical morphometry in general.

The normalization process discussed before would possibly create difficulties in further variability analysis. When studying the shape variability of specific sulci, an interesting approach is to create a co-ordinate system based on the sulcal depth profile (Cykowski et al., 2008). In the next section, some of the findings using the methods discussed above are briefly discussed.

1.4.3 Applications of cortical morphology

The study of cortical morphology can shed some light on many important issues. In particular, the findings in the four domains would be discussed: brain abnormalities, brain development and aging, brain plasticity and the relation of genetics to brain development and cortical folding.

1.4.3.1 Brain abnormalities

Since cortical folding pattern is very stable through-out life in normal population (Armstrong et al., 1995), a deviation from the normal gyrification rates or gyrification patterns has thus a high probability to indicate brain malfunction. Global and regional abnormal gyrification is found in a variety of disorders. Using the 2D normalization approach discussed above, abnormal symmetry of cortical folding has been found in William's syndrome (Van Essen et al., 2006) and Schizophrenia (Csernansky et al., 2008). Sulcal depth difference is found in autism (Nordahl et al., 2007).

The GI study reveals many interesting results as well. Using the 2D GI, it is found that gyrification pattern in monozygotic twin pairs is different for autism. Increased folding is found in the right parietal lobe, and the increase in folding is associated with more symptoms of autism (Kates et al., 2009). Decreased folding is found in the prefrontal lobe in patients with obsessive-compulsive disorder (OCD) (Wobrock et al.), while increased

folding in the same region is found to be linked to the high risk group for schizophrenia (Harris et al., 2007). Many other regions of abnormal folding are linked to pathology. For example, abnormal folding in the anterior cingulate cortex is linked to bipolar disorder (Fornito et al., 2007) and OCD (Shim et al., 2009). Measuring gyrification index at both the global and local sulcal level of language-related cortex (Cachia et al., 2008), it is found that the schizophrenia patients with auditive hallucination have a decrease in sulcal index in the superior temporal sulcus, the middle frontal sulcus and the diagonal branch of the Sylvian valley, the region defining the Broca's area.

Specific sulcal shape has been linked to pathology as well. Change in the frequency of specific folding patterns of the collateral sulcus is linked to temporal lobe epilepsy (Kim et al., 2008). The shape of the temporo-parietal junction and the superior temporal sulcus is linked to inner or outer space hallucinations (Plaze et al., 2009).

1.4.3.2 Brain development and aging

The study of cortical folding can provide many insights of brain development and aging. The computational approaches provide exciting potential of studying cortical folding in newborns (Dubois et al., 2008). Early structural measurements such as cortical folding can be useful in the follow-up study of highly premature infants.

In the study of brain volume related to GI, it is found that there is a disproportionate increase in cortical surface, especially in the prefrontal area (Toro et al., 2008). More general traits such as brain asymmetry can also be studied. The position and surface area asymmetry can be studied in adults (Lyttelton et al., 2009) and infants (Hill et al., 2010). More specific to sulci, the asymmetry of the depth of the central sulcus is studied (Cykowski et al., 2008). It is found that the central sulcus is asymmetrical in surface area and length (Kloppel et al., 2010; Mangin et al., 2004a).

Cortical analysis can also shed some light on the aging process. The age-related change in sulcal width and depth are studied to show that sulcal structures in multimodal cortical areas have more profound age-related changes than those in unimodal areas (Kochunov et al., 2005). The sulcal depth and surface area change is found in brain atrophy such as in the case of Cerebral Autosomal Dominant Arteriolopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL) (Jouvent et al., 2009).

1.4.3.3 Brain plasticity

Plasticity and the effect of learning is a subject of enduring interest. A series of VBM studies found links between brain morphometry and function in terms of proficiency in various performances. Brain has the potential for changes in structure and function throughout life. For example, a bilateral gray matter expansion in the medial temporal visual area (also known as V5) is found in juggling novices (Draganski et al., 2004). Furthermore, such changes can be detected after just 7 days of juggling practice (Driemeyer et al., 2008).

Plasticity in terms of cortical folding is less easy to be studied, probably due to the fact that folding changes take much longer than gray matter changes to be established. Some studies do report for example the central sulcus of amputees may eventually lose its characteristic shape (Dettmers et al., 1999). Furthermore, such flattened central sulcus is not observed in amputees of the arms that used their foot for sophisticated activities such as painting or sculpting (Yu et al., 2006). In the study of hand converters, the natural left-handers who are forced to write with the right hand, it is found that the asymmetry of the central sulcus surface area is changed compared to left-handers (Kloppel et al., 2010). In Chapter Four of this thesis, this work on the analysis of the central sulcus of hand-converters is continued with interesting results.

1.4.3.4 Brain development and genetics

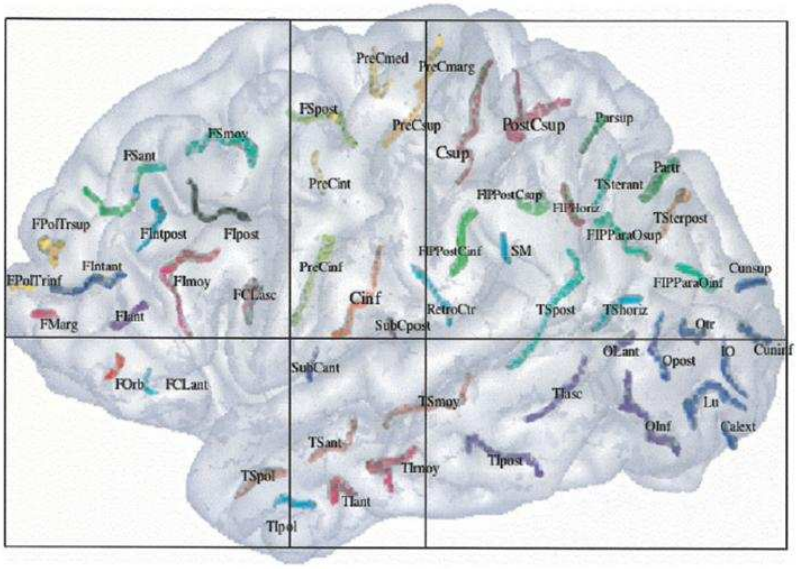
The genetic influence on brain structure is found in twin studies (Peper et al., 2007; Schmitt et al., 2007; Thompson et al., 2001). The thickness and surface area of the cortex is likely influenced by different and distinct genetic factors (Panizzon et al., 2009).

Evolutionary studies can provide information complementary to those of clinical studies, since many biological mechanisms behind development; aging, learning and disease are shared between a wide range of organisms (Carroll, 2005).

1.4.4 Organizational framework of cortical folding and future work

Some organizational framework has been proposed regarding the cortical folding patterns. In the sulcal root model (Regis et al., 2005), the variability observed in cortical folding is explained based on the gyri buried inside the sulci (the plis de passage). In rare cases the “plis de passage” is too buried to be observed in adult brains. The sulcal roots are the units corresponding to the first folding locations during fetal development. The units of sulcal roots are organized in a system of meridians and parallels of the cortical surface. The locations of sulcal roots are relatively stable across individuals, the variability occurs in later development during the folding process. These relatively stable entities across subjects could be very useful for sulcal labeling and spatial normalization before anatomical or functional analysis.

The sulcal roots are obtained in three steps: i) each main sulcus is split into portions according to interruptions observed in literature, mainly from Ono and Kubik atlas (Ono et al., 1990); ii) split based on the “plis de passage”; iii) split based on embryological literature when available. It is assumed that the depth of the “plis de passage” is correlated with the date of apparition.



Generic description of cerebral cortex anatomy. The Talairach grid system represents precursors of the sulcal roots in the right hemisphere. The 57 sulcal roots of the lateral aspect are displayed schematically. Calext: calcarinalis externalis, Cinf: centralis inferior, Csup: centralis superior, Cuninf: cunealis inferior, Cunsup: cunealis superior, FCLant: fissura cerebri lateralis anterior, FCLasc: fissura cerebri lateralis ascendens, FIntant: frontalis intermedialis anterior, FIntpost: frontalis intermedialis posterior, FImoy: frontalis inferior intermedius, FIPHoriz: fissura intraparietalis horizontalis, FIant: frontalis inferior anterior, FIpost: frontalis inferior posterior, FIPParaOinf: fissura intraparietalis paraoccipitalis inferior, FIPParaOsup: fissura intraparietalis paraoccipitalis superior, FIPPostCinf: fissura intraparietalis postcentralis inferior, FIPPostCsup: fissura intraparietalis postcentralis superior, FMarg: frontalis marginalis, FOrb: fronto-orbitalis, FPO: fissura parieto-occipitalis, FPolTrinf: fronto-polar transversalis inferior, FPolTrsup: fronto-polar transversalis superior, FSant: frontalis superior anterior, FSmoy: frontalis superior intermedius, FSpost: frontalis superior posterior, int 1: 1st ramus intermedius FIP, int 2: 2nd ramus intermedius FIP, IO: interoccipitalis, Lu: lunatus, Oinf: occipitalis inferior, OLant: occipitalis lateralis anterior, Opost: occipitalis posterior, Otr: occipitalis transversalis, Parsup: parietalis superior, Partr: parietalis transversalis, PostCsup: postcentralis superior, PreCinf: precentralis inferior, PreCint: precentralis internalis, PreCmarg: precentralis marginalis, PreCmed: precentralis medialis, PreCsup: precentralis superior, RetroCtr: retrocentralis transversalis, SM: sulcus marginalis, SubCant: subcentralis anterior, SubCpost: subcentralis posterior, TImoy: temporalis inferior intermedius, TIant: temporalis inferior anterior, TIasc: temporalis inferior ascendens, TIpol: temporalis inferior polaris, TIpost: temporalis inferior posterior, TSant: temporalis superior anterior, TShoriz: temporalis superior horizontalis, TSmoy: temporalis superior intermedialis, TSpol: temporalis superior polaris, TSpost: temporalis superior posterior, TSterant: temporalis superior terminalis anterior, TSterpost: temporalis superior terminalis anterior.

Fig 1.9 The sulcal roots (as fig 5 of the paper on sulcal roots (Regis et al., 2005))

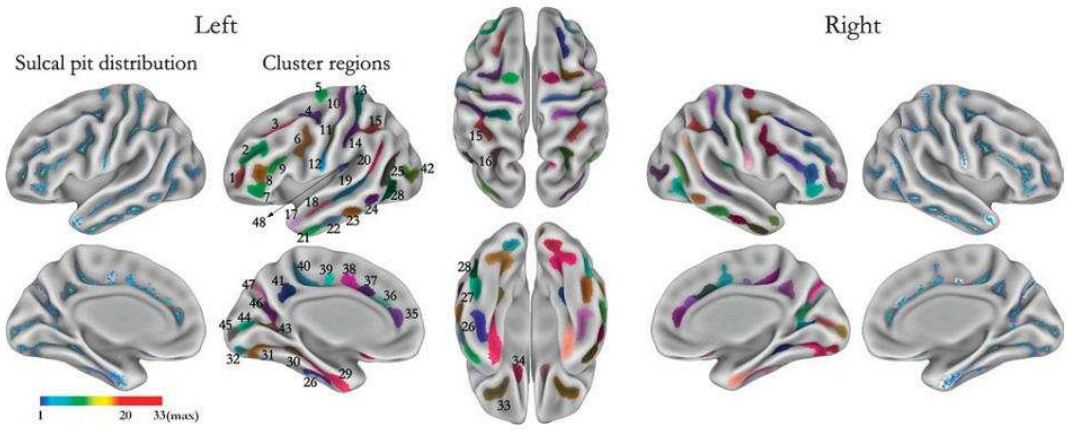


Fig 1.10 The sulcal pits (as presented in (Im et al., 2010))

s.: sulcus, 1 middle frontal s. a, 2 middle frontal s. b, 3 superior frontal s., 4 junction between superior frontal s. and precentral s., 5 precentral s., 6 junction between precentral s. and inferior frontal s., 7 inferior frontal s.a, 8 inferior frontal s. b, 9 inferior frontal s. c, 10 central s. a, 11 central s. b, 12 central s. c, 13 postcentral s. a, 14 postcentral s. b, 15 intraparietal s. a, 16 intraparietal s. b, 17 superior temporal s. a, 18 superior temporal s. b, 19 superior temporal s. c, 20 superior temporal s. d, 21 inferior temporal s. a, 22 inferior temporal s. b, 23 inferior temporal s. c, 24 inferior temporal s. d, 25 inferior temporal s. e, 26 occipito-temporal s. a, 27 occipito-temporal s. b, 28 occipito-temporal s. c, 29 collateral s. a, 30 collateral s. b, 31 collateral s. c, 32 collateral s. d, 33 orbital s., 34 olfactory s., 35 cingulate s. a, 36 cingulate s. b, 37 cingulate s. c, 38 cingulate s. d, 39 cingulate s. e, 40 cingulate s. f, 41 subparietal s., 42 lateral occipital s., 43 calcarine s. a, 44 calcarine s. b, 45 calcarine s. c, 46 parieto-occipital s. a, 47 parieto-occipital

Another approach which obtained very similar map as that of the sulcal roots is the sulcal pits model (Lohmann et al., 2008). The sulcal depth is used; the sulcal pit is the zone of the sulcus where the depth is maximal. These sulcal pits are hypothesized to be under genetic control based on sulcal analysis of monozygotic twins.

The sulcal roots and sulcal pits results are displayed below in **Fig 1.9** and **Fig 1.10**, taken from the original papers (Im et al., 2010; Regis et al., 2005).

To understand better the mechanism of cortical folding, large scale developmental studies (longitudinal) on folding would need to be carried out. Functional analysis would need to be combined with folding pattern analysis to understand better the implications of specific folding patterns. Mathematical models such as the reaction-diffusion model

(Lefevre and Mangin, 2010) can be used to provide more insights into the process of folding.

On the other hand, systematic folding pattern analysis such as that of the work of Ono (Ono et al., 1990) is still important. Such work can be extended with the knowledge obtained from the above approaches. The knowledge gained can likely provide new insights to the organizational framework of cortical folding discussed above.

1.5 Discussion

A brief tour has been given to the subjects of concern to the thesis work: the biological process of cortical folding and cortical morphology. It should be emphasized that this thesis work is an interdisciplinary effort to understand brain development and cortical folding. The methods designed and chosen need to suit the specific neuroscience questions. The evaluation of the validity and quality of the results should be based on neuroscience as well. It is our hope that the knowledge gained through this work may add to the existing knowledge base of neuroanatomy and neuroscience.

The methods developed and used, together with some example results can be found in Chapter Two and Three. An application of such analysis to the understanding of handedness can be found in Chapter Four. Chapter Five present the dictionary and Chapter Six gives a summary of the knowledge gained in terms of cortical folding patterns through this work.

Chapter Two: Clusters of folds

2.1 Summary

In this chapter, the concept of using clustering algorithms to analyze cortical folding is introduced. The effort of selecting the suitable shape descriptor and the development of a clustering algorithm specific to cortical folding analysis is described. Some interesting results obtained are presented and discussed.

2.2 Introduction to the analysis of cortical folding

Human brain cortex folds to increase its surface area during development. To better understand the nature and the degree of variability, some real brains are presented in **Fig 2.1** (as Fig 1.1, shown here for easier inspection). In the first brain, the superior frontal sulcus (the green fold) is broken into three pieces, while in the second and the third brain it is continuous as shown in most textbooks. Refer to the intermediate frontal sulcus (the cyan fold): in the first brain it is broken into three pieces, in the second brain there exists less discontinuity, while in the last brain it is extremely discontinuous to the extent that it is hard to be labeled in the conventional naming framework.

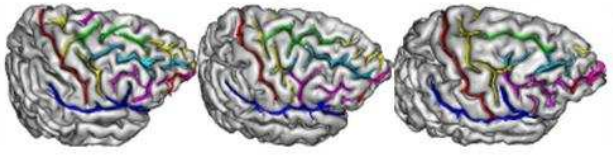


Fig 2.1 The variability in brain folding patterns

Indeed, the traditional naming system cannot account for the huge variability in brain folding patterns. Furthermore, the naming system is based on external morphological

criteria alone, the three-dimensional complexity of the folds is hidden and thus cannot be accounted for.

This leads to the goal of our current effort, try to understand better the variability of sulcal folding patterns in a comprehensive manner. This includes the description of the various folding patterns. The possible link of certain characteristic folding patterns to neurological interpretations such as pathology or behaviour will be explored as well. From the folds alone can we observe patterns characteristic of certain neurological diseases? The knowledge of folding patterns could also be added to the traditional naming system to better characterize the brain morphology. Eventually, this information may be used by the artificial vision system such as BrainVISA for improving automatic fold recognition.

The most detailed description of the sulcus variability has been proposed in the atlas of Ono (Ono et al., 1990). Here we try to approach the analysis of folding patterns in a more systematic way. In one sense, we attempt to automate the work performed by Ono. With the help of computer algorithms, a large amount of subjects can be analyzed in a more consistent manner than visual inspection by human eyes. It is important to keep in mind, however, that this type of cortical folding analysis up till now can only be performed by experienced neuroanatomists. Due to the huge variability that exists, this type of analysis is challenging even for an expert. So to automate this type of analysis, even partially, is not a trivial task.

What are these cortical folding patterns that we are looking for? In the preliminary study these patterns are loosely defined as a group of brains that show a characteristic trait which distinguishes them from the other brains. This characteristic could be based on the 3D shape of the fold, the surface area of the fold, the degree of curvature, the number and position of interruptions etc. Since this type of unsupervised learning on the shape of the folds has never been carried out systematically before, we do not have a clear definition of how to define the resulting patterns to start with. It is also very likely that no single

finite set of parameters (angles, length, depth etc) can be defined to characterize the patterns found. For now, we simply define a pattern to be a certain characteristic that can distinguish one group of folds from another.

We expect to observe characteristic folding patterns in only a subset of the population. Refer back to **Fig 2.1**, suppose we found two patterns for the superior frontal sulcus (the green fold): one being broken into three pieces, the other being continuous. It is important to keep in mind that, in a given dataset, it is likely that only a subset of the subjects would have one of the two patterns. The rest of the subjects would not fit these two pattern descriptions. The computer algorithm used for clustering should be able to choose the subset that contains interesting patterns.

2.2.1 Clustering algorithms

Now the general goal is defined, we discuss the method that can be used to tackle such a problem of sulcal pattern discovery. Clustering analysis would need to be carried out. Clustering is the unsupervised classification of patterns (observations, data items, or feature vectors) into groups, or clusters. It is distinguished from the supervised learning by the fact that there are no training examples to teach some a priori output. Clustering analysis is performed often when little prior information is available about the data; it is useful for exploration of the interrelationships among the data points, to make an assessment of their structure (Jain et al., 1999). Clustering is an essential component of data mining, a process of exploring and analyzing large amounts of data in order to discover useful information (Berry and Linoff, 2000). While there are still debate toward the ultimate definition of clustering, a rough definition is nonetheless possible: for a given set of data points and a similarity measure, we regroup the data such that objects in the same cluster are similar and objects in different clusters are distinct (Jain and Dube, 1988).

Large number of clustering algorithms exists in literature from various domains, ranging from pattern recognition, artificial intelligence, image processing, statistics, and applied mathematics to marketing, psychology, and biology. In each field, a set of algorithms tend to dominate for historical and practical reasons. For example, the hierarchical-based approach is more used in the artificial intelligence community, while the model-based approach is more used in the statistical community. While exciting new advancement has been made in clustering algorithm development, challenges still remain. Part of the reason is that domain specific problems often require the use of specifically designed algorithms, general clustering algorithms are often not sufficient. The behavior of real-life situations is always complex, unpredictable challenges rise, demanding more sophisticated or more specific solutions.

There is no clustering technique that is universally applicable for uncovering the variety of structures present in multi-dimensional datasets. Not all clustering techniques can uncover all the clusters present with equal facility, because clustering algorithms often contain implicit assumptions about cluster shapes. In reality, data hardly follow the "ideal" structures such as being hyperspherical or linear. Very often, a dedicated clustering algorithm performs slightly better than the existing ones on a specific distribution of patterns (Jain et al., 1999). Furthermore, it is difficult for humans to obtain an intuitive interpretation of the clustering results, especially for data in high-dimensional space.

Clustering analysis is intensively used in Bioinformatics, especially in gene expression analysis and Microarray analysis. The application of clustering algorithms to the analysis of brain folding patterns is new, the behavior of existing algorithms would need to be studied, new algorithms suitable for this specific domain would need to be designed if necessary and consequently validated.

In terms of the types of clusters, Lorr (Lorr and Maurice, 1983) suggested that there appear to be two types of clusters, compact clusters and chained clusters. For a compact

cluster, objects of the same cluster have high mutual similarity; usually a compact cluster can be presented by a representative point or center. A chained cluster is a set of data points in which every member is more similar to other members in the cluster. Any two data points in a chained cluster are reachable through a path. Refer to **Fig 2.2** to have an intuitive understanding of these two types of clusters.

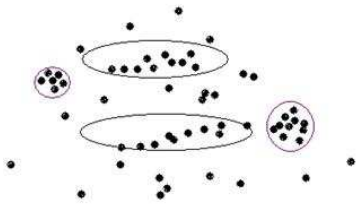


Fig 2.2 Compact and Chained clusters

The chained clusters are typically elongated clusters (black) due to the chaining effect. The compact clusters are typically more dense and round-shaped (purple).

In the application of clustering analysis to the discovery of cortical folding patterns, it is important to define the type of clusters that we are looking for. In the context of cortical folding, the compact clusters would be the groups of subjects whose sulcal patterns are very similar to each other. The chained clusters would be the groups of subjects that are not as highly similar as in the compact clusters, but there exist a higher similarity among subjects within the group than subjects outside the group. These two types of clusters could both be very interesting and convey important yet different information regarding the nature of the dataset.

Clustering algorithms can be divided into two categories: hard or crisp clustering and fuzzy or soft clustering. In hard clustering, each data point belongs to one single cluster; in fuzzy clustering, each data point belongs to two or more clusters with certain probabilities. Hard clustering algorithms are categorized into two kinds, hierarchical and partitional algorithms. Hierarchical algorithm can be further divided into agglomerative and divisive approaches (Gan et al., 2007). A sketch of the general structure of clustering algorithms is given, (see **Fig 2.3**).

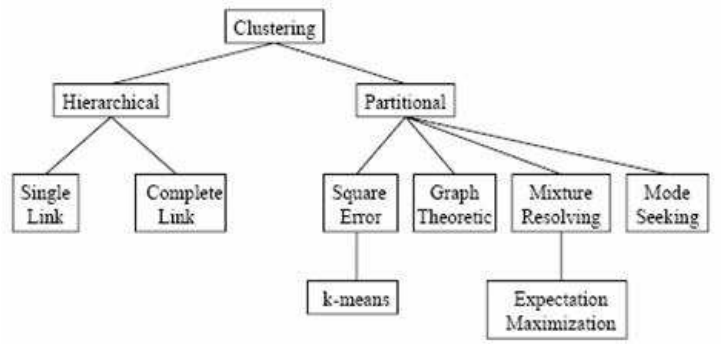


Fig 2.3 Different types of clustering methods (as in (Jain et al., 1999)

A *partitional* method constructs N clusters. That is, it classifies the data into N groups, which together satisfy the requirements of a partition. For this type of algorithm, each group must contain at least one object; and each object must belong to exactly one group. Partitional methods are applied if one wants to classify the objects into N clusters; N is usually given by the user and fixed. In general, the algorithm tries to find a "good" partition in the sense that objects of the same cluster should be close or related to each other, whereas objects of different clusters should be far apart or very different. **Fig 2.4** illustrates the principle of a partitional algorithm. When N given to the algorithm is 3, the data points are divided into three partitions or three classes by the algorithm.

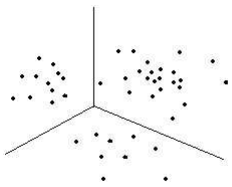


Fig 2.4 Cluster data using partitional algorithm

Hypothetical data divided into three clusters by a partitional algorithm.

Hierarchical algorithms, on the other hand, do not construct a single partition with N clusters; instead they deal with all values of N in the same run. There are two kinds of

hierarchical techniques: the agglomerative (bottom-up) and the divisive (top-down). Agglomerative algorithms begin with each element as a separate cluster and merge them in successively larger clusters. Divisive algorithms begin with the whole set and proceed to divide it into successively smaller clusters.

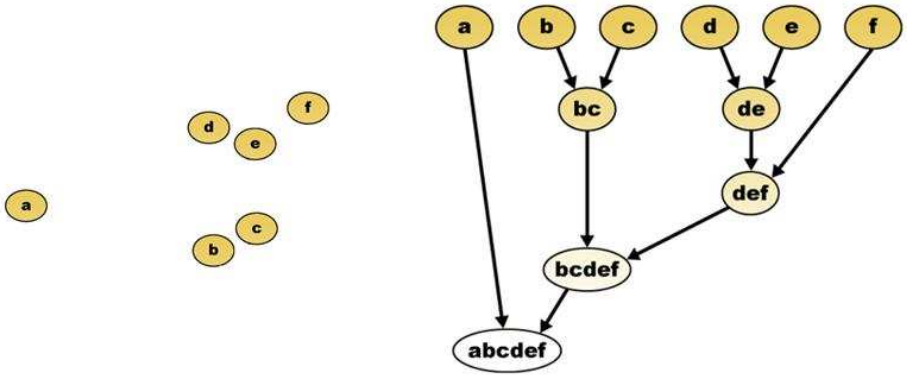


Fig 2.5 Sample data distribution and the formation of agglomerative clustering tree

The principle of agglomerative clustering is illustrated in **Fig 2.5**, where a diagram of the agglomerative process is shown. Groups are formed consecutively at each iteration, until at the last step all elements are in the same group. In divisive clustering the arrow of dataflow is reversed. Notice that the hierarchical tree can provide information on the data structure itself. From this tree we can deduce various information: points “b” and “c”, and points “d” and “e” are the closest pairs ; “f” is closer to the “de“ group than to the “bc” group ; “a” is an outlier point that was joined to the rest of the points only at the very end.

2.2.2 Clustering analysis on cortical folding

Which type of clustering algorithm is more suitable for cortical folding analysis, partitional or hierarchical? To make a decision we need to go back to the goal of our data analysis: finding patterns of folds. Brain folding is a chaotic phenomenon. As we have seen earlier, there is a huge variability in the patterns of folding. Not every brain in the data set need to belong to a particular pattern. We are looking for reasonably large groups

of brains that exhibit a similarity in folding. Hence, we are not trying to classify each brain in terms of its folding pattern. Practically, we want to be able to visually inspect the clusters found and validate them. This means we want to “fish out” the tightest elements of each of the clusters. These elements are expected to be the closest in shape, and the best representatives of a particular pattern of folding.

To achieve this goal, we need an algorithm that can find the cluster “centers”, the tightest elements of a given cluster. More importantly, we want an algorithm that can discard the outliers. Therefore, the goal is not to divide the elements into N clusters. Agglomerative methods can help us achieve such a goal. This algorithm groups the tightest elements first; it guarantees that the outliers would only be joined at the end of the process. So if we only consider the clusters formed at the beginning of the clustering process, these outliers would be discarded automatically.

The main drawback of agglomerative algorithm and of hierarchical algorithm in general, is that it is computationally expensive, and it cannot correct a possibly wrong grouping at a later step. However, for the purpose of a preliminary study, it is sufficient to provide some insight into the folding patterns. For a more complete study in the future, algorithm that is less expensive and performs reasonably well for discarding outliers could be further explored. Refer to Chapter Six for a more detailed discussion.

In the rest of this chapter, the selection of descriptors for clustering analysis, and the development of a clustering algorithm dedicated to sulcal pattern analysis are discussed in detail. Validations of the method are then presented, followed by some interesting patterns found.

2.3 The shape descriptor

The clustering analysis starts by the selection of shape descriptors. How to represent the shapes is an important choice which would consequently decide the nature and quality of

the results. In the following, the initial effort is described first to illustrate the specific problems and difficulties encountered during the first try of finding patterns. Shape descriptors more suitable for such studies are further explored, using more sophisticated algorithms.

2.3.1 The initial effort

As a first trial of the discovery of patterns, we started by selecting a set of folds; some morphometric features of each fold were selected to characterize their shapes and relations. The central sulcus and the Sylvian fissure were selected due to the fact that they are among the biggest and the most stable folds (see **Fig 2.6**). The selected features are: the surface area of the two sulci, the shortest distance between the bottom of the central sulcus and the Sylvian fissure, and the angle between the two sulci.

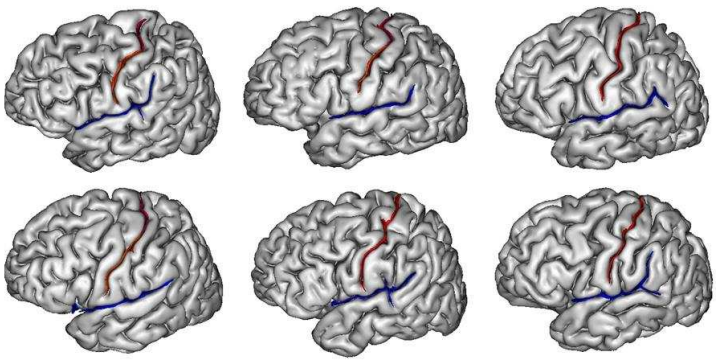


Fig 2.6 The Central Sulcus (red) and the Sylvian fissure (blue), highlighted on brain surfaces.

The simple clustering algorithm K-means was used to cluster the data as a first try. K-means is a greedy algorithm for partitioning the n samples into K clusters, so as to minimize the sum of the squared distances to the cluster centers (MacQueen, 1967). This clustering algorithm is simple, intuitive and fast, weaknesses exist however. In particular, it is not robust, the initialization influences the result, and the result depends on the value of K chosen. There is no simple solution to these problems. Furthermore, because the

folds are complicated 3D objects, it is very difficult to compare their forms to detect similarities. The result of the k-means clustering was inspected and no strong evidence was found to support a good clustering: no obvious pattern could be detected from the resulting clusters.

To be able to obtain convincing clustering results, it became evident that an algorithm that can select the most similar folds is needed. In other words, the clusters found need to be compact, the corresponding sulci in the cluster would be expected to be very similar in shape. The identification of such highly similar subset of folds provides a better chance for a describable 3D folding pattern. This algorithm should also be robust and outlier-proof. In terms of the information we use as the input to the clustering algorithm, we need some descriptors that can capture more precisely and more comprehensively the information of the 3D forms. It is difficult to choose interesting features by hand; also as the number of features increases we encounter the problem of the curse of dimensionality (Bellman, 1961).

This very first try illustrates the importance of selecting a good shape descriptor, where comprehensive information on the cortical folds should be reliably coded. It also shows the special challenges of clustering on very noisy datasets, due to the huge variability that exists in folding patterns. This motivates the selection of the moment invariant as a shape descriptor, which is discussed next. This first try also motivates the design of a dedicated clustering algorithm which is discussed in the section after.

2.1.1 The 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 (Mangin et al., 2004a). Hence they allow the management of various sulcus interruptions. The construction of these descriptors filters 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.

Here we give a brief insight into the computation of the 3D moment invariants. The 3D moments of order n = p+q+r, n∈N of a 3D density function ρ(x,y,z) are defined by

$$m_{pqr} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q z^r \rho(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

For our purpose ρ(x,y,z) is equal to 1 inside the object of interest and 0 elsewhere, because we deal with objects defined by binary images. The moments of order higher than 3 are not considered in this report; but the derivation of moment invariants is theoretically possible for any order. By discarding moments of order higher than 3, a small set of global descriptors are obtained which embed simple shape information, such as bending, tapering, pinching etc. The derivation of the invariants aims at filtering out the influence of localization, orientation and scale on the 3D moments in order to obtain "pure shape" descriptors.

This derivation is done in three steps. First, translation invariance is obtained using the centroid of the object as the origin of the coordinate system, leading to the definition of the central moments denoted by Mpqr. Second, a new set of moments invariant to scale, denoted by μpqr, is obtained by normalizing central moments with the suitable power of the volume M000:

$$\mu_{pqr} = \frac{M_{pqr}}{M_{000}^{\frac{p+q+r}{3}+1}}.$$

Finally, the invariance to 3D rotations is derived from sophisticated group theory techniques usual in quantum mechanics. These techniques combine decomposition into harmonic polynomials and tensor calculus beyond the scope of the thesis work (Lo and Don, 1989). The resulting invariants turn out to be homogeneous polynomials of the central moments made up of several hundreds of terms. Because of various symmetries, we get only 12 invariants denoted by I1, I2, …, I12 in the following.

For the work reported here, the moment invariant data are calculated using the software brainVISA (http://brainvisa.info). This data is then used as input to the clustering program. The invariance to scale and rotation of the descriptors provided by this implementation was checked elsewhere through the resampling of a couple of objects with 28 different orientations and several scales (Mangin et al., 2004a). The variability of the invariant estimation resulting from this resampling was always less than 5% (it should be noted that perfect invariance would be achieved only for continuous objects).

2.3.2 3D Moment invariants as sulcal shape descriptors

Some investigations are carried out to verify that the set of moment invariants is a reasonably good shape representation to study the folding patterns.

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 behavior of the invariants is shown in **Fig. 2.7**. It is impossible to claim from these simple investigations that the invariants vary smoothly whatever the 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. 2.8**. 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. 2.9**). 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 behavior 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 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 work.

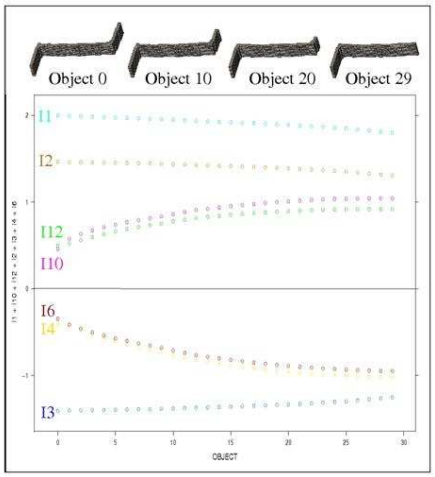


Fig 2.7 The variation of moment invariants corresponding to the change in 3D shape.

I1, I2, I12, I10, I6, I4 and I3 are the seven moment invariants illustrated. As the shapes vary, following a trend shown by the four 3D shapes on top, these invariants vary following smooth curves.

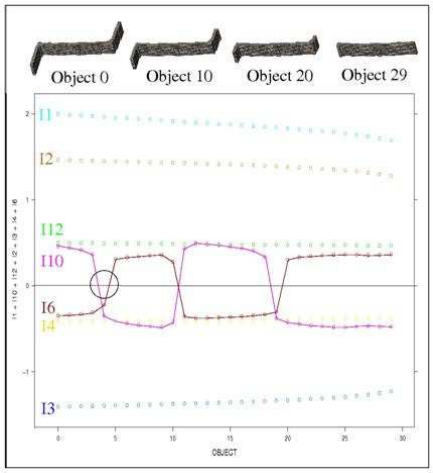


Fig 2.8. Variation of moment invariants corresponding to the change in the 3D shape of the objects.

On top, the trend of the change in shape is depicted; the upper and the lower arm of the object are shortened gradually. Seven out of the twelve moment invariants are drawn on the graph, notice that while the rest of them are smoothly changing in value when the shape change gradually, I6 and I10 change signs abruptly.

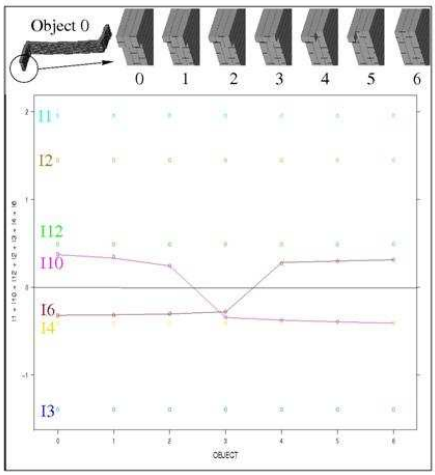


Fig 2.9. The variation of the moment invariants corresponding to the fine change in the 3D shape of the objects.

On top, the trend of the change in shape is depicted; one voxel is removed at a time, from the lower arm of the object. Seven out of the twelve moment invariants are drawn on the graph, notice that while the rest of them are smoothly changing in value when the shape change gradually, I6 and I10 change signs abruptly.

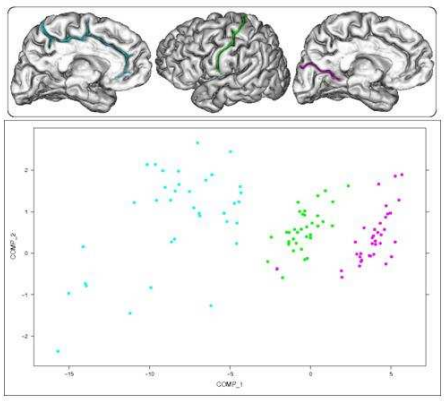


Fig 2.10 The distribution of different sulci.

The data of the three sulci, the cingulate sulcus (cyan), the central sulcus (green) and the parieto-occipital sulcus (magenta) are plotted, using the first two axes of PCA as the X and Y coordinates. Dataset of 36 manually labeled brains is used, respective locations of the three sulci are shown on top.

A second investigation aims at verifying that the information on the shape 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 a lot of cases, the plot is

made up of several clouds of points corresponding to the different sulci. These clouds overlap more or less according to the choice of 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. The fact that some of the clouds overlap would simply mean that some sulci have similar shapes. This is not a problem for achieving our objective.

An example is shown (**Fig 2.10**) using three different sulci: the cingulate sulcus (cyan), the central sulcus (green) and the parieto-occipital sulcus (magenta). A dataset of 36 manually labeled and normalized brains is used. Using the two first PCA axes, the three sulci are almost perfectly distinguished. Notice also the large variability of the cingulate sulcus distribution (cyan) compared to the two other sulci. The central sulcus (green) leads to the tightest cloud, which is consistent with the fact that this sulcus is one of the most stable ones. **Fig 2.11** is showing that one of the sources of variability of the cingulate sulcus pattern is its frequent interruptions.

In **Fig2.11** sibling to **Fig 2.10**, some of the dots have been replaced by a snapshot of the corresponding sulcus. The points of view chosen for these snapshots correspond to the 3D renderings of the brain shown in the box at the top. This graph further confirms that the moment invariants provide reliable representations of the 3D shapes of the folds that vary smoothly across the shape space. It is possible, indeed, to see gradual changes of the shape of the folds. For example, consider the cingulate sulcus: at the bottom left we see sulci more fragmented, while towards the top right corner, the shapes become more continuous. To conclude, the moment invariants can be considered as good descriptors of the 3D shapes of the folds.

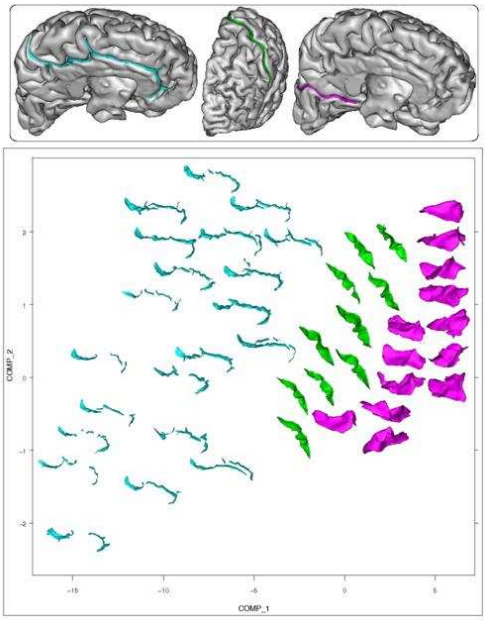


Fig. 2.11 The distribution of different sulci, with the sample shapes.

The data of the three sulci, the cingulate sulcus (cyan), the central sulcus (green) and the parieto-occipital sulcus (magenta) are plotted, using the first two axes of PCA as the X and Y coordinates. A normalized dataset of 36 manually labeled brains is used. The shapes of some samples randomly chosen are plotted. The corresponding locations of the three sulci are shown on top.

2.4 The clustering algorithm for sulcal pattern discovery

2.4.1 Agglomerative hierarchical clustering

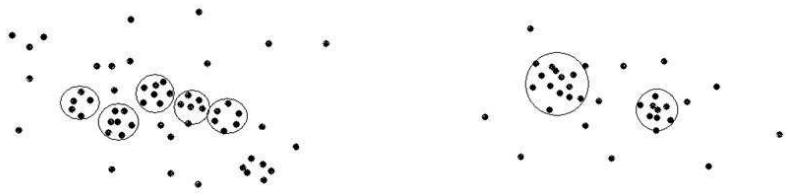
Once the sulcal form descriptor is determined, the next step is to design an algorithm dedicated to sulcal pattern analysis. Following the discussion on clustering algorithms, the hierarchical approach to clustering is considered suitable for finding compact clusters and discarding numerous outliers. This method is thus chosen for the preliminary study.

There exist many agglomerative algorithms, which only differ in their definition of between-cluster dissimilarity (Kaufman and Rousseeuw, 1990). The most common ones are the nearest neighbor method (single-linkage), the furthest neighbor method

(complete-linkage) and the unweighted pair-group average method (average linkage). In the nearest neighbor method, during agglomeration, the dissimilarity between the cluster R and the cluster Q is the smallest dissimilarity between an object of R and an object of Q. On the other hand, the furthest neighbor rule uses the largest dissimilarity between an object of R and an object of Q. In average-linkage, the dissimilarity between clusters R and Q is taken to be the average of all dissimilarities d(i,j), where i is any object of R and j is any object of Q.

The nearest neighbor rule is not always appropriate. Whenever both clusters come too close to each other, even when this happens at just one point, the clusters immediately stick together. Notice that they cannot be separated at later steps. This is called the chaining effect because many objects may be chained together resulting in a drawn-out cluster, some members of which are very far from each other. This algorithm tends to give elongated clusters because of the chaining effect. What we are more interested in here are the more round shaped clusters.

The furthest neighbor rule possesses the opposite property. It tends to produce very compact clusters. Every member of such a cluster must be close to every other member of the same cluster, outlying points will not be incorporated. The resulting clusters are not necessarily well shaped, because clusters will not be joined when they contain at least one pair of too distant points. **Fig 2.12A** illustrates the typical clusters formed by using this method.



A **B**

Figure 2.12 Typical clustering result of complete-linkage and average-linkage.

Whereas single linkage usually leads to too few clusters which are drawn out, complete linkage often yields the opposite effect: many clusters with small within-cluster dissimilarities. As a consequence, relatively similar objects will often stay in different clusters for a long time, hence complete linkage is sometimes said to be space dilating. Single linkage will often bring rather different objects into the same cluster due to the chaining effect, and therefore said to be space contracting. The necessity to compromise between these two extremes has lead to group average linkage and other methods, which are space conserving (Kaufman and Rousseeuw, 1990). The group average technique is aimed at finding roughly ball-shaped clusters. Being rather robust, this method can even deal with more potato-shaped clusters, (see **Fig 2.12B**).

In our data analysis, we want to use a robust method that is space conserving, and we want to find clusters with the ball-shape. So the group average method is used for the clustering.

2.4.2 Merging of clusters: tight-head join

Regarding the detail of the algorithm, we have N clusters at the beginning. We proceed by successive fusions until a single cluster is obtained containing all the objects. We start by constructing a dissimilarity matrix, which records the dissimilarity coefficients of the samples pair-wisely. Dissimilarity coefficients between objects are obtained here by the computation of distances.

At the first step, the two closest or most similar *objects* are joined. In the second and all the subsequent steps, we will want to merge the two closest *clusters.* The dissimilarity d(R,Q) between clusters R and Q is defined as the average of all dissimilarities d(i,j), where i is any object of R and j is any object of Q. We update the dissimilarity matrix each time a new join occurs.

In order to extract a partitioning from a hierarchy, we have to choose an appropriate level. Various stopping rules were proposed to select a suitable number of clusters based on the distribution of clustering criteria (Mojena, 1977). Two graphs are provided to illustrate the idea of different stopping rule and their corresponding clusters, (see **Fig 2.13**). In **Fig 2.13A**, two stopping rules are illustrated (the earlier stopping rule in green, the later stopping rule in purple). In **Fig 2.13B**, the two resulting cluster groups are presented. The green cluster groups are discarding "f" and "a", the purple cluster groups are discarding only "a" instead. The cluster groups we obtain from the agglomerative clustering are different depending on when we stop the process.

In the clustering algorithm, as discussed earlier, we want to be able to extract out the center elements of a given cluster. Each time a new element is added to a cluster, we want to guarantee that the center elements remain at the head of the cluster list. To achieve this goal, we need a different join algorithm, so that after each join, we reorder the elements of the cluster in terms of tightness. Note that this algorithm does not guarantee that the center elements are the tightest elements of the cluster as well. We consider the clusters that are formed earlier in the hierarchical process to be more important. For evaluation by visual inspection, we can then simply take the n elements at the head of the clusters.

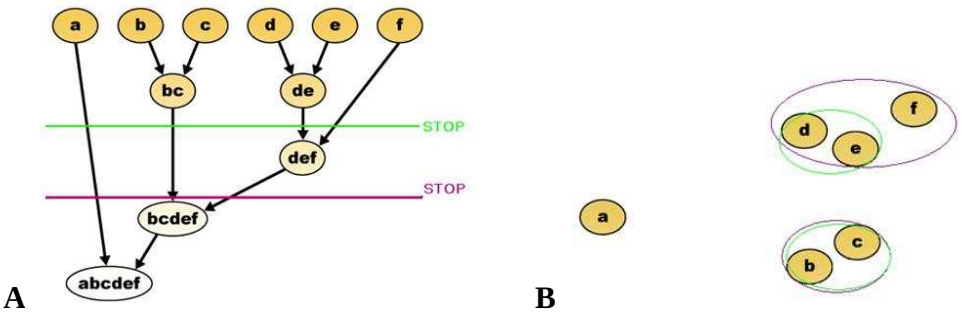


Fig 2.13. Two different stopping rules.

Depending on the stopping rule, the iteration where we stop the agglomerative or divisive process is different. **A**: Shown in green and purple are two different stopping rules. **B**: The clusters in green are the ones resulting from the green stopping rule; the purple clusters are resulting from the purple stopping rule

The goal is then to keep the n objects at the center of a given cluster (the objects that we are interested in) at the head of each cluster list. First, we define a “core” cluster to mean a cluster with at least C elements. We are only interested in the big enough core clusters. How C is chosen will be discussed later in this chapter. For now, it is enough to describe these core clusters to be the tight clusters of brains with a similar pattern that could not have occurred purely by chance.

Here is how we join two clusters; we call it “tightHeadJoin”. The head is defined to be the *C elements* at the beginning of the cluster list that we are interested in. If both clusters have less than C elements, the elements are ordered from the most tight to the least tight; if both clusters have more than C elements, the cluster with the tighter head (the first C elements) is put in front; if one cluster has more than C elements, the other has less than C elements, the one with more than C elements is put in front.

To reorder the head of the clusters, we start by picking out the tightest pair of the elements from the merged group of two clusters. These two elements are set as the new head. For the rest of the steps, we compare distance from the mean of the head to each element left in the merged group; the closest element is picked to be added to the end of the head to form a new head. This process is repeated until all the elements are added to the head.

Note that in a situation depicted in **Fig 2.14**, this algorithm will not give the tightest group of elements as the head. The algorithm will give “a”, “b” and “c” as the three-element head, while the tightest group with three elements is “d”, “c” and “e”. Nonetheless, the resulting head elements given by this algorithm are useful and interesting in the sense that the elements grouped together earlier during the hierarchical process are given more importance and put in front. The objects grouped earlier during the agglomerative process are tighter groups compared to the ones grouped later in the process. The subjects grouped earlier in the agglomerative process are expected to be interesting when we try to observe the similarity in 3D shapes of these objects. Keep in

mind that only the “head” of each cluster is used for further processing and evaluation, so only the ordering of the “head” needs to be guaranteed.

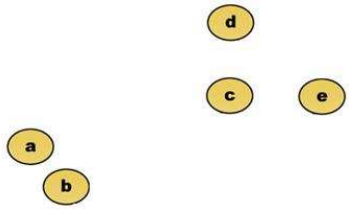


Fig 2.14. An example distribution.

A hypothetical data distribution to illustrate the characteristic of the “tightHeadJoin”.

2.4.3 The stopping rule based on competition

So how to pick out the most “interesting” folds? How to find the corresponding clusters of brains for these folds? The goal is to pick out the sulci that are most likely to exhibit possible patterns of interest, we are not attempting to compose an exhaustive list. Because the results are difficult to analyze (as discussed in section 2.3.1), the idea we use to carry out this selection process is a competition among the different sulci, and a competition among the different clusters formed at each step of the agglomerative clustering process.

A **B**

Fig 2.15. Sample distribution A and B

To have a more intuitive understanding, refer to the two distributions depicted in **Fig 2.15**. To “fish out” patterns, **Fig 2.15A** is considered to be less interesting than **Fig**

2.15B. The reason is that from the distribution depicted in **Fig 2.15B**, we can observe two possible local distribution centers that could be corresponding to two patterns. In terms of choosing the clusters for one particular sulcus, there exist many possibilities depending on the stopping rule. The clusters obtained change depending on when we stop the agglomeration process. We call the clusters found at a particular iteration of the agglomerative process a ***cluster group***.

The "interesting" cluster groups are defined to be the ones with tight elements within the cluster; and with large dissimilarity among clusters. A ratio is used to characterize the clusters, and different clusters from different folds compete based on this ratio. The folds whose clusters have the highest ratios are selected. This competition is really a selection based on the characteristics of the clusters; and these characteristics are described by the ratios.

We first define the *distance* between two clusters to be the distance between the centers of each cluster. When there are more than two clusters, the distance among the clusters is calculated as the average pair-wise distance between cluster centers. The *compactness* of a cluster is defined as the average pair-wise distance to the median. The ratio is defined as the average distance divided by the average cluster compactness. The ratio allows us to find cluster groups whose clusters are relatively tight within a given cluster, yet the clusters are far from each other. The ratio **r** is computed for clusters formed at each iteration of the agglomerative clustering process. The iteration that wins at the end is the one that has the highest ratio **r**. Notice that there could be more than one iteration winning. This competition is carried out among different sulci and combination of sulci. The sulci are then ranked according to the ratio. The winning cluster groups of the winning sulci are expected to have some strong patterns, and would be evaluated further.

Now that the quality of the clusters found can be evaluated by the ratio introduced, we go back to the issue of choosing C, the size of the "head" of a cluster. C is the size of a core cluster as discussed in the previous section. To find this number, random sets from

Gaussian distribution using the covariance matrix of the original data sets are generated. The same competition using the random sets is then carried out. For a given C, a cluster group obtained with real data is considered significant only if its ratio is better than the best ratio obtained with the random set. This procedure is using the parametric sampling process (Good, 2004), the p-value of the clusters formed at each step can be estimated, clusters with the best p-values can be studied further.

In the clustering literature, Ray and Turi (Ray and Turi, 1999) introduced a compactness-separation-based validity measure to determine the number of clusters in k-means clustering. ***Mintra***, the intracluster distance is defined as the mean square distance to the cluster center. ***Minter***, the intercluster distance is defined as the square distance between the centers. A good clustering result should have a small Mintra and a large Minter. The validity measure **V** is defined as ***Mintra/Minter***. V is to be minimized. The ratio used in the designed algorithm is the same as this compactness-separation-based validity measure. In our algorithm, the validity measure is used for the selection of the best stopping rule, instead of the selection of the best number of clusters.

To determine the optimal number of clusters, C=2 up to C=number of subjects in the sample can be tried, the validity measure V can be calculated for each C, that gives the smallest V is then chosen to be the optimal cluster number. The number of clusters is automatically decided when the stopping rule is found.

2.4.4 Bagging and final clusters

The clustering algorithm described is not very stable, in the sense that a small change in the sample leads to large variations in the result. **Fig 2.16** illustrates this problem. Relatively small variations of the sample set leads to large enough differences in the resulting clusters. To add stability to the clusters found by using the above algorithm, bagging is used. This method creates a set of clusters from a given dataset. Each group of

clusters is extracted from a bootstrap of the dataset. Note that the bagging technique improves the estimate only if the learning algorithm is unstable. It degrades the estimate if the algorithm is stable. As it is demonstrated that our original algorithm has stability issues, bagging is thus suitable to resolve this problem.

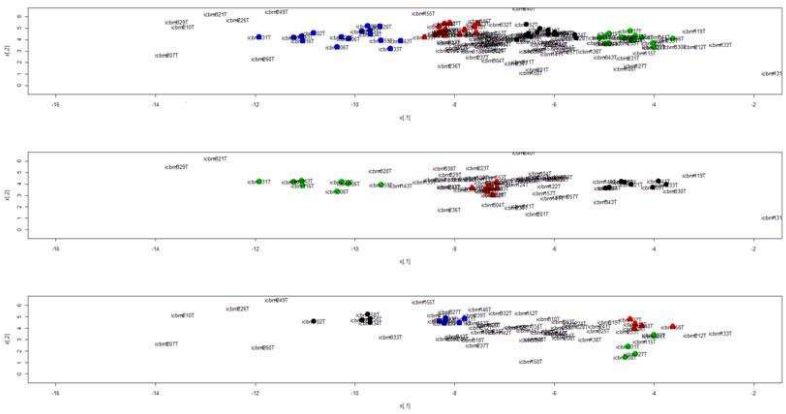


Fig 2.16 Resulting clusters on bootstrap samples

The three images show three different datasets generated from the original dataset by bootstrap. The clusters obtained are marked in different colours (black, red, blue and green). PCA is used for visualization. It can be observed that a slight variation of the sample leads to large enough differences in the resulting clusters. This difference in the clustering results is both in terms of the number of clusters found and the location of the clusters found.

In the method validation section, it is shown that the bagging approach greatly improves the resulting clusters. Thus the bagging technique is integrated into the clustering algorithm. The resulting algorithm is summarized in the next section.

Once the salient points are found from each of the bootstrap samples, they are gathered together as the new data points for the final clustering step. Since the salient points found in the bootstrap samples are very selective, for this work a simple K-means or aggomerative kind of algorithm is considered sufficient to find the final cluster centers. In the work presented in this chapter, a K-medoid algorithm is used, which is a partitional clustering algorithm related to the K-means algorithm. In contrast to the K-means algorithm, K-medoid algorithm choses real data points as centers. In this implementation,

Partioning Around Medoid (PAM) is used as a K-medoid clustering algorithm (Kaufman and Rousseeuw, 1990). The characterization of clusters in PAM is by the representatives of the clusters, the medoids. Compare with K-means, PAM is more robust by minimizing a sum of dissimilarities instead of a sum of squared Euclidean distances. A medoid is defined as an object of the cluster, whose average dissimilarity to all the objects in the cluster is minimal. A measure called Silhouette can be used to select the number of clusters. The Silhouette method is as follows:

if **a(i)** is defined as the average dissimilarity of **i** with all other data points within the cluster, and **b(i)** is defined as the average dissimilarity of **i** with all other data points in another cluster which has the lowest average dissimilarity to **i**, the Silhouette measure **s(i)** can then be defined as

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \qquad s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases}$$

a(i) measures how i is dissimilar to its cluster; **b(i**) indicates how i is matched to its neighbouring cluster; **s(i**) is a number within the range from -1 to 1, a value close to one indicates that i is properly clustered, a value close to -1 would indicate that i is better clustered to its neighbouring cluster. The average **s(i)** of the whole dataset thus measures the quality of the clustering. Using this measure, the appropriate number of final clusters can be estimated.

2.4.5 The PCBB algorithm

Adding all the previous steps together, we summarize here the algorithms designed for cortical folding analysis, named PCBB, which stands for Partial Clustering by Bootstrap sampling and Bagging. The steps of the algorithms are described below and summarized in two flow charts (**Fig 2.17A** and **Fig 2.17B**).

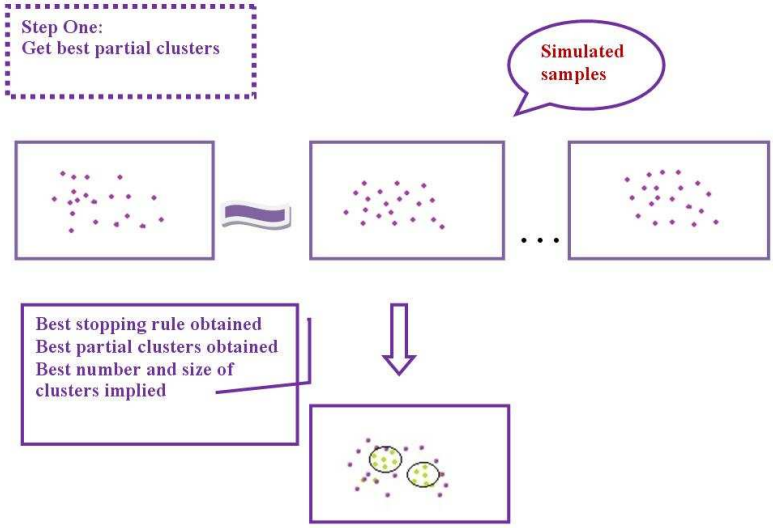


Fig 2.17A Obtaining the best partial clusters

The simulated samples are obtained by simulating a Gaussian distribution using the covariance matrix of the real dataset. The hierarchical clustering is run on the real dataset and all the simulated datasets, the step of the agglomeration where the best ratio is obtained is considered the step where the strongest clusters can be found. The quality of the clusters found at this step is verified by calculating the p-value. This value is calculated by counting the number of times that the simulated dataset performed better than the real dataset.

Step 1: (Fig 2.17A)

Agglomerative hierarchical clustering is performed; the agglomeration process is guided by an objective function:

$$R = \sum \text{compactness of the clusters formed} / \sum \text{distance among the clusters formed}$$

In each step of the agglomeration process, R is calculated. The p-value of the clusters formed at each step is then estimated by a parametric sampling process. Simulated distributions are generated using the covariance matrix of the real data, the same clustering process is applied to these simulated datasets and p-value is estimated by counting the number of times the simulated data have a better R score than the real data. Finally the clusters with the best p-value are chosen as the salient points for the next step of the algorithm.

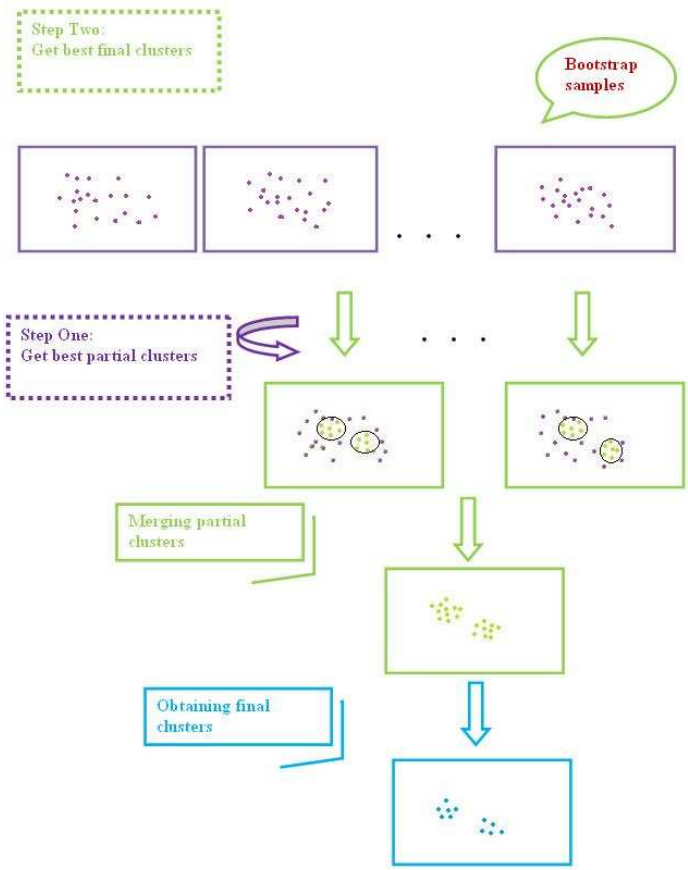


Fig 2.17B Obtaining the best final clusters

Step 2: (Fig 2.17B)

The process described in step 1 is performed many times on the bootstrap datasets of the original data. A repertoire of salient points is identified to form the new dataset. A simple K-medoid algorithm (PAM) is then used to find the location of the final clusters.

The goal of step one is to estimate the number of clusters and their size automatically. Notice that the clustering is “partial”, not all data points are assigned to clusters. The goal here is to extract the most interesting sample points that might contain strong and significant patterns. We are not trying to assign each point to a pattern. Note also that the p-values estimated here are only used in the ranking system to pick out the most “interesting” clusters; they are not used to perform statistical tests.

Step two is using the idea of bagging, the goal is to overcome the instability of the clusters found in the first step. The assumption is that the first step of the clustering on the bootstrap samples gives the strongest clusters and eliminates most of the noise. So in this step a relatively simple clustering algorithm is sufficient to identify the final clusters.

The strong point of the PCBB algorithm is that it is very robust, also it makes no assumption that the clusters span the whole data space, as all the division-based clustering algorithms do. When we think about our problem of finding cortical folding patterns here, we are aware that the cortical folding process is very complicated. The final folding pattern of the brain is the end product of numerous chemical and mechanical forces acting on the brain, very well orchestrated throughout the time of brain development. Here we are not trying to model the folding process and explain all the variability we observe in folding patterns. Instead we are trying to identify some typical patterns that might exist in only a part of the population, but are significant and can give us some insight into the folding process and certain pathologies. So partial clustering is more relevant for this particular problem.

2.5 Validation of the algorithm

2.5.1 Simulated datasets

To evaluate the performance of the PCBB algorithm, some experiments are performed on simulated datasets. The procedure and results are presented below. The dataset we use as a model for generating the simulated dataset is a real dataset made up of 36 brains, where each sulcus has been reliably labelled manually by a neuroanatomist. This dataset was used to train the sulcus recognition system of brainVISA. We chose the moment invariant data of the ten biggest sulci of each hemisphere for further analysis.

The simulated datasets are generated as follows: take the mean and covariance matrix of any pair of sulci, generate a new dataset using these same parameters. This gives a dataset with two known clusters. Then a series of noisy datasets are generated by adding 10, 20, 30, 40 50, 60, 70, 80, 90, 100% noise to the new dataset. The noise added follows a Poisson distribution, within the min and max value of the original dataset in their respective dimensions. The mean and covariance matrix of the real sulci are used to keep the simulated data close to the distribution of the real data.

Both the Gaussian mixture modelling algorithm (GMM) and the PCBB algorithm are run on these simulated datasets. The results are evaluated in terms of the number of clusters found and the distance from the cluster centres found to the real centres. GMM involves first fitting a mixture model, usually by the expectation-maximization (EM) algorithm (Duda et al., 2000). Some success has been shown using the Bayesian Information Criterion (BIC) to choose the right number of components. However, in general, equating a component of GMM with a cluster is questionable (Ray and Lindsay, 2005). In our experiments, we use the Mclust toolbox from R to run the GMM algorithm. Mclust is a state-of-the-art mixture-model-based clustering tool (Fraley and Raftery, 2002, 2006). We did two GMM runs for each dataset. The first run allows the algorithm to optimally select the structure of the covariance matrices using BIC, but without the initialization of the proportion of noise as a prior. In the second run the real proportion of noise in the dataset is given as a prior to the algorithm.

Two comparisons are made to evaluate the quality of the clustering. First, the distance of the cluster centres found to the real centres are measured. When there are more cluster centres found by the algorithm than the real centres, only the two clusters closest to the real centres are taken into consideration. Second, the numbers of cluster centres found by the algorithm are compared for each simulated dataset.

The result is shown in **Fig 2.18** below:

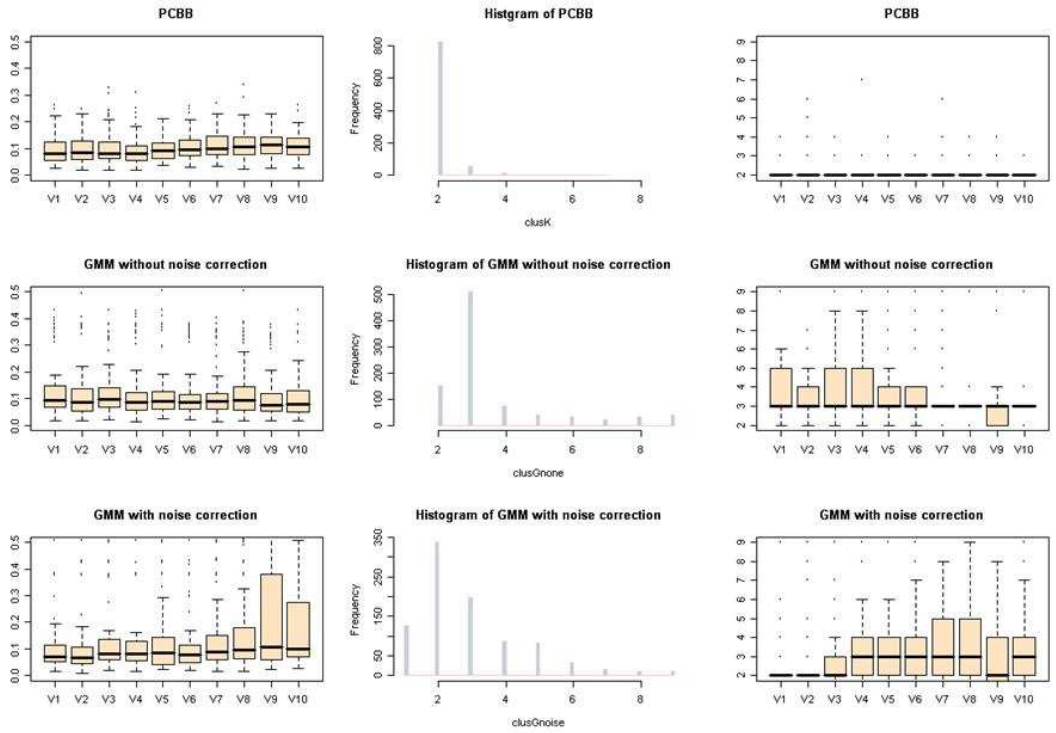


Fig 2.18 The comparison of performances

The first column shows the boxplot of the distance of the two closest cluster centres found by the algorithm to the real centres. (In the box plot, within the box is the data from the first to the third quartile, the dark line inside the box represents the median. Below the box shows the line of the minimum, above the box the line of the maximum, the outliers are shown as dots.) The x-axis shows the ten simulated datasets, averages across the pairs of sulci, with the percentage of noise from 10 to 100 percent. The results of the PCBB method is shown on the first row, the results of GMM without noise correction are shown on the second row, and the results of GMM with noise correction are shown on the third row.
The second column shows the histogram of the distributions of the number of centres by the three algorithms, the third column shows the boxplot of the number of centres for the ten different datasets from 10% to 100% noise.

Results show that the PCBB algorithm is comparable to the GMM algorithm in terms of locating the centres of clusters. However, in terms of estimating the number of clusters, PCBB is more accurate and stable than the GMM algorithm, with increasing number of noise in the data. Feeding the GMM a percentage of noise during initialization does not seem to help the performance in this case. The result shows that PCBB is more robust than GMM for finding clusters in this particular problem.

2.5.2 Real datasets

The clustering experiment is run on the 36 brain dataset used for generating the simulated data. The sample is normalized as follows:

Normalized sample = (sample -median) / median of the absolute deviation from the median

The clusters found and their corresponding forms are observed, one example on the analysis of the cingulate sulcus is given here to illustrate the stability of the algorithm. In **Fig 2.19**, we observe the forms of the patterns found. The first cluster has a pattern with an anterior interruption, the second cluster has a pattern with a posterior interruption, and the third cluster appears to be continuous. These patterns found are consistent with those described in the atlas of Ono, which stated that around 60% of the instances of the cingulated 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. It should be noted that the size of the database used prevents the detection of rare patterns. Therefore, larger and more comprehensive databases will be required to achieve a more exhaustive pattern search.

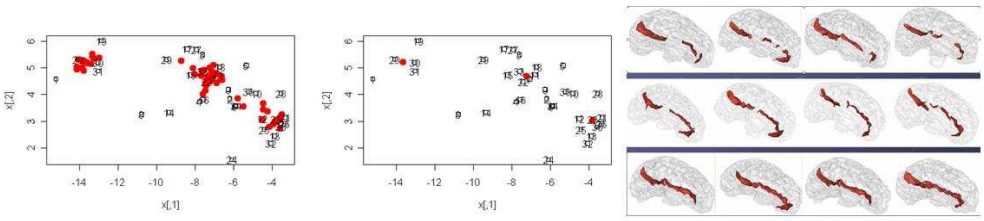


Fig 2.19 The clusters of the cingulate sulcus

The image to the left shows the salient points found with 100 bootstrap samples. The image in the middle shows the final cluster centres found by the clustering algorithm. The X and Y-axis are the first and second dimensions of the PCA. The third image shows the shapes of the cingulate sulcus of the three corresponding clusters.

Next the PCBB algorithm is performed on three real and different datasets. Moment invariant data of the left cingulate sulcus is used. The clusters found are shown in **Fig 2.20** below.

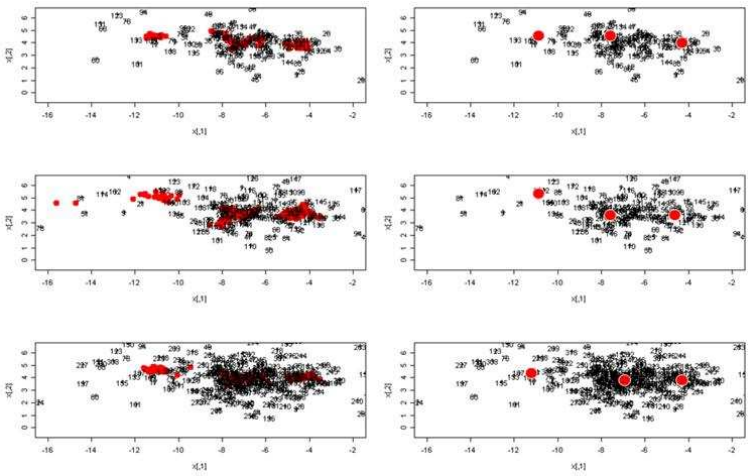


Fig 2.20 The clusters found on three different datasets

The dataset is shown using the first two axes of PCA, as the X and Y axis. The first row shows one dataset, the second row shows another dataset, the third row shows the dataset composed by mixing the data of the first and the second datasets. The first column shows the salient points found by the first part of PCBB, using 100 bootstrap samples. The second column shows the final cluster centres found by the second part of the PCBB.

Here two large real datasets are used, each containing 150 brains. The sulci are automatically identified and labelled by brainVISA. We observe that the PCBB algorithm is stable over different datasets. The bagging procedure used in the second part of the algorithm helps to achieve greater stability.

2.6 Folding patterns

Some of the strongest patterns found are shown and discussed next. The method has been applied to the ten largest sulci (**Table 2.1**) of the left and right hemispheres using the database of 36 manually labelled brains. From these 10 sulci, all the combinations of two and three sulci are taken, leading to 45 pairs and 120 triplets. This constructive approach

allows us to look for patterns at different levels of details or at different degrees of resolution. The motivation for building any pair or triplet of sulci stems from the idea that long range interactions could occur during the folding process, either because of long fibre bundles or because of correlated development of different brain areas. This constructive process could be pushed to the extreme, where we would study the pattern made up by all the sulci of the brain. It should be noted, however, that extending the approach to bigger groups than triplets would lead to intractable combinatorial explosion. Therefore, in the future, this multi-resolution strategy should be more selective, for example, larger groups can be built only from sulci close to each other in the brain.

Among the 20 sulci (of the two hemispheres), three sulci 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 left superior frontal sulcus). The sulcus providing the best p-value (0.001 for $t = 4$) is the left cingulate sulcus. Results based on the three sulci where the strongest patterns are found are presented below. Among the pairs and triplets, some patterns are found as well, one of the examples is shown below.

We also use another set of 150 brains, with the sulci automatically labelled. 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 (Riviere et al., 2002).

Number	Name	Shorthand BrainVISA
1	calloso-marginal fissure	F.C.M.
2	calcarine fissure and calcarine scissure	F.Cal.ant.-Sc.Cal.
3	collateral fissure	F.Coll
4	inter-pariatal fissure	F.I.P.
5	pariatal-occipital fissure	F.P.O.
6	central sulcus	S.C.
7	inferior frontal sulcus	S.F.inf.
8	superior frontal sulcus	S.F.sup.
9	inferior precentral sulcus	S.Pe.C.inf.
10	superior temporal sulcus	S.T.S

Table 2.1 The ten sulci for folding analysis

2.6.1 The cingulate sulcus

The cingulate sulcus ranks the best in p-value. Three patterns are found (refer to **Fig 2.21**). The 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. 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 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 (**Fig. 2.21**). Note that when the anterior part of the sulcus is made up of two parallel folds (fourth row of **Fig. 2.21**), it is equivalent to a deeper sulcus for the moment invariants.

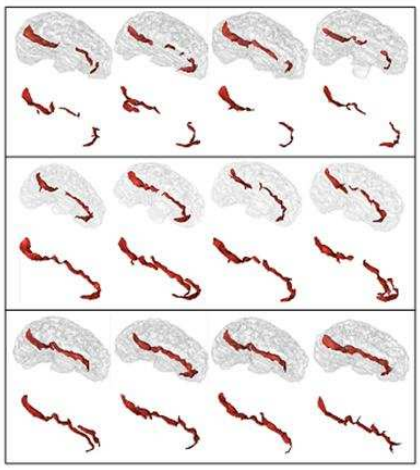


Fig 2.21 The three patterns detected for the left cingulate sulcus. Row 1,3,5: the four tightest instances of each pattern in manually labelled database. Row 2,4,6: the four closest instances to the above pattern centre in automatically labelled database.

To project the patterns from the first database onto the second database, we classify the sulci according to the closest distance to the pattern centres. 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%).

2.6.2 The inferior precentral sulcus

The sulcus that has the second best p-value is the precentral sulcus. Refer to **Fig 2.22**, two patterns are found, depicted on the first and the second row, each pattern contains three subjects. The precentral sulci (in green) are plot on the 3D rendering of the cortex for the six subjects. The central sulci are plot as a reference.

The first pattern (first row on **Fig 2.22**) is more frequent, comparing with the second pattern; the sulci here are more elongated. Notice however that even though all of the folds are long, they are not always in the same orientation. The folds of the first and the third subjects have different orientation compare with the second. When observing the second row pattern, it can be seen that they contain the upper portion of the first pattern, the lower portion is likely either missing or not being identified as precentral sulcus.

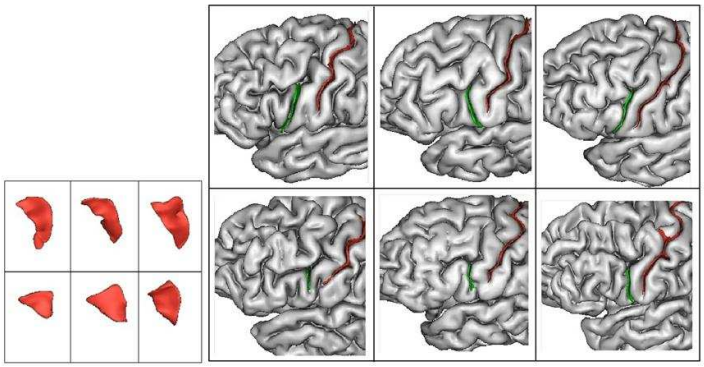


Fig 2.22 The patterns of the inferior precentral sulcus

Indeed, in anatomical literature, a small fissure originating from the Sylvian fissure is seen that is called “sulcus subcentralis anterior” (Eberstaller, 1890). This sulcus sometimes unite with the lower end of the central sulcus, sometimes unites with the precentral sulcus or the diagonal ramus of the Sylvian Fissure. This information can be

useful to the automatic sulcal recognition system. The anterior subcentral sulcus is thought to be rarely present and not used by the previous versions of the sulcal recognition system of BrainVISA. The result of this pattern analysis inspired the addition of this sulcus to the naming system of the new version (Perrot et al., 2009a).

2.6.3 The calcarine, collateral and central sulcus

This example illustrates the possibility of finding patterns in sulci groups. Here the three sulci concerned are: the calcarine sulcus, the collateral sulcus and the central sulcus. Two patterns are found (**Fig 2.23**), which are illustrated in colour red and green, each pattern contains three subjects. The overall differences between the patterns can be easier to observe when they are plot together. The green group has a bigger calcarine sulcus, and the red group has a bigger and longer collateral sulcus. With respect to the central sulcus, the green group has the central sulcus closer to the other two sulci. It would be interesting to further explore the meaning of such differences.

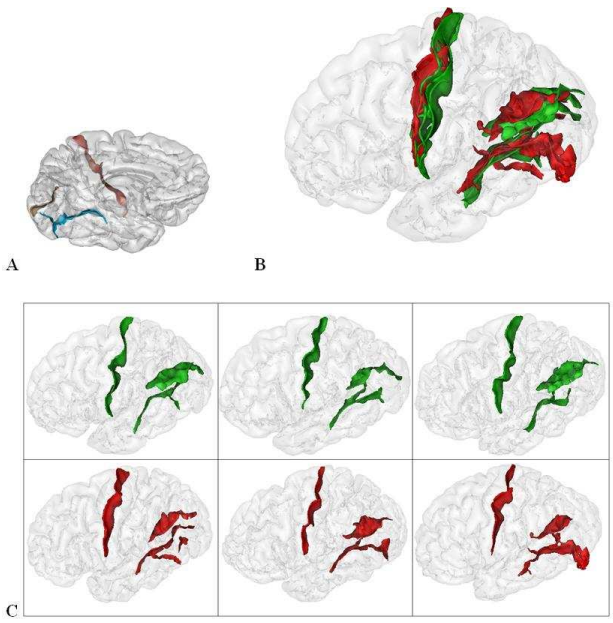


Fig 2.23 The patterns of the calcarine, collateral and central sulcus

A: The location of the calcarine sulcus, the collateral sulcus and the central sulcus. The calcarine sulcus is in brown, the collateral sulcus is in blue, the central sulcus is in red. **B:** The subjects from the two patterns plot together **C:** The two patterns in green and red, each pattern contain 3 subjects.

2.7 Discussion

The preliminary cortical pattern analysis produces interesting results. The moment invariants provide stable and comprehensive description of the shapes compare with any specific set of measurements such as length or depth.

One possible use of the patterns we found is to compare the frequency of the occurrence of these patterns among normal and patient datasets, to verify if we can see any significant difference. 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 to verify is that a certain developmental event or a certain strong specific training would leave an observable imprint on the cortex folding patterns.

Chapter Three: Patterns of folds

3.1 Summary

In this chapter, some new approaches are explored in the study of folding patterns. In terms of the similarity measure among the folds, new descriptors are used; in terms of the algorithms to mine the information, the type of algorithm that can handle high dimensional data is experimented. New information obtained, combined with that obtained from using the methods described in the previous chapter, allows a more comprehensive description of the sulcal shape. The change in the shape of the folds across the population can be observed, together with descriptions of the more frequent and characteristic patterns of the population.

3.2 Introduction

The work described in the previous chapter consists of finding the most frequent patterns in terms of cortical folding. Further study of the variability of sulci involves three important choices: which sulci to study or where to look in the brain; which similarity measure to apply; and which algorithm to use for mining the patterns.

Regarding the choice of where to look among the cortical folds, new directions can be explored. In the work presented before, the biggest folds are chosen for pattern detection. The biggest folds are likely the more stable; consequently they can be more reliably detected. Bigger sulci are more likely to exist in any given individual, while the smaller folds may or may not exist in a given individual due to the high variability of folding patterns. The drawback or limitation of such an approach is that the important patterns might include the smaller folds. A more sensible approach may target the regions of the cortex, not restricted to the prominent sulci that are easier to be labelled according to traditional nomenclature. There are regions linked to functions, from the bigger regions such as the prefrontal region, to the smaller regions such as the Broca's area and

Wernicke's area linked to language. Indeed, the sulci of the brain are the folds that separate the gyri, and the gyri are where the neuronal activities are located. So to include the notion of a functional area, it makes sense to include at least two sulci, the sulci anterior and posterior to a given gyrus. As a simple example, the central sulcus and pre-central sulcus can be grouped together, to give a better description of the pre-central gyrus region. When a bigger region is of interest, a group of sulci within this functional region can be studied. Based on this reasoning, in further exploration of cortical folding patterns, we decided to include a group of sulci within a chosen region of interest.

Regarding the choice of the similarity measure, more possibilities can be explored. We used a 3D shape descriptor, the 3D moment invariant in the previous studies. While this descriptor can capture the form of the folds, the information is nonetheless limited since only ten numbers are used to describe each fold. Another interesting and more intuitive approach is the use of the similarity of the folds among the subjects. Each fold of a given subject can be characterized as the similarity to the same fold of all the other subjects in the data set. The characterization of the fold is much more comprehensive, since it would include the whole dataset. This direction is further explored in the work described in this chapter.

Regarding the algorithm to use for pattern analysis, many new directions can be explored. It is important to realize, however, when the nature of the input data is changed, the clustering algorithm might need to be changed as well. In particular, if the approach of similarity measure discussed above is adopted, the dimension of the input would be very high. When the data dimension is high, the distance between any two data points becomes almost the same (Beyer et al., 1999). In such case, it becomes difficult to differentiate similar data points from dissimilar ones. In the mean time, clusters are embedded in the subspaces of high-dimensional data space; different clusters may exist in different subspaces (Agrawal et al., 1998). Because of these reasons, almost all conventional clustering algorithms fail to work well for high-dimensional data sets. Consequently, the algorithm used should either be able to reliably reduce the dimension,

or feature selection techniques should be used to select only certain features to be used in the final clustering. The other alternative is to use algorithms specifically designed for high-dimensional datasets. In this chapter, we describe how we choose the algorithm based on the nature of the shape description.

3.3 Searching for a framework of patterns

In the following, the approach of describing the cortical folds is presented, together with the algorithms that best suit the similarity measure. This approach is compared with the approach described in the previous chapter, to illustrate the type of information that can be discovered when applying different approaches. Some interesting results are presented and discussed.

3.3.1 The Similarity measure: ICP

Sophisticated shape descriptors based on 3D moment invariants have been proposed in the previous chapter. This approach might be limited in terms of the representation of all the complexity in folding patterns. Very different shapes can sometimes lead to similar descriptors which can disturb the clustering process. In order to overcome this weakness, a different approach is explored. A given sulcal shape is described by a vector of distance of this shape to a large number of similar shapes. This approach has been proven to be very efficient to compare shapes in large dimension spaces (Besl and McKay, 1992). Hence the representation of the sulcal set of one subject is consisting of the distances to the same sulcal set in all the other subjects. Each pair-wise distance is computed using the simple Iterated Closest Point (ICP) algorithm after affine global spatial normalization of the brains (Jain et al., 1999). Note that performing a global normalisation removes non-interesting patterns induced by global differences in brain size. Our ICP implementation is providing the minimal distance obtained whatever the rotation between the two shapes (Kaufman and Rousseeuw, 1990).

The idea of representing each sulcus by its distance to all the other sulci in the dataset is illustrated in **Fig 3.1**. The simplest ICP algorithm is used, in which a form (data) is matched to another form (model) iteratively by rotation and translation (Besl and McKay, 1992). The algorithm stops when the distance cannot be significantly improved or a given number of iteration has been reached. It should be noted that this algorithm has been extensively studied and many variations of the algorithm provide more efficient performance. In this preliminary study, the simplest version of the algorithm is used. It is however by choice that we use only rigid motion for matching in this study. It is reasoned that in the application of cortical pattern analysis, even the small shape differences are important. Consequently, the distance should faithfully reflect this difference. However, an affine matching could be very interesting as well, in the sense that smaller details in folding difference could be ignored, so that we can concentrate on more prominent differences. It should be kept in mind that the tuning of ICP is not only related to nonlinear alignment, it is also related to robustness. The estimation of the distance between dissimilar shapes could be less accurate compare with the distance between more similar shapes. However, our method does not require accurate distance estimation for dissimilar shapes because only the ordering is important in the algorithm used here.

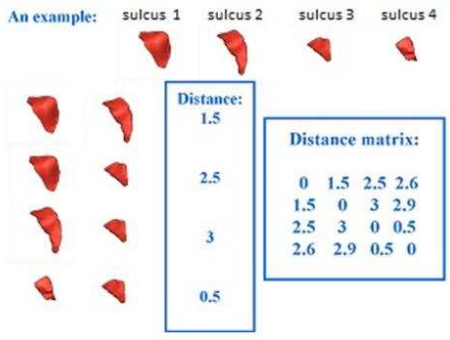


Fig 3.1: An example of the similarity measure

As an example to illustrate the concept of the similarity measure, four sulci are used (in red) named sulcus 1 to 4. The distance (using ICP) can be calculated for each pair of the set. For each sulcus, the distance to the other three sulci can be calculated. The final distance matrix can be obtained.

3.3.2 Algorithm to handle high dimensional data: Isomap

The resulting similarity matrix obtained using the ICP algorithm could have a very high dimensionality, since the number of dimension is equal to the number of subjects in a given dataset. The curse of dimensionality is a well-known problem occurring in such situations (Duda et al., 2000). The dimension of data needs to be reduced before applying the clustering algorithms.

The assumption of dimension reduction is that data points may lie on a lower dimensional manifold, which is embedded in the original space of higher dimension. The goal is to find such a lower dimensional embedding of the data, while keeping the original geometry of the data as much as possible.

The techniques of dimension reduction could be roughly divided into two categories, the linear and nonlinear techniques (Van der Maaten and al, 2008). Examples of linear dimension reduction techniques are Principle Components Analysis (PCA) and Linear Discriminant Analysis (LDA). Nonlinear techniques can be further divided into three main types: techniques trying to preserve global properties of the original data; techniques attempting to preserve local properties of the original data; and techniques performing global alignment using some linear models. Examples of techniques preserving global properties are MultiDimensional Scaling (MDS), Isomap, Diffusion maps and Kernal PCA. Local Linear Embedding (LLE) is an example of the technique preserving local properties, where multi-linear PCA is used for local representation. Local Linear Coordination (LLC) is one of the techniques using local linear models for global alignment.

The Isomap algorithm is chosen for this preliminary study. This algorithm has the computational efficiency and global optimality of Principal Component Analysis (PCA) and Multi-Dimensional Scaling (MDS), it also has the flexibility to learn a broad class of

non-linear manifolds (Tenenbaum et al., 2000). The distance between any two data points is measured by using the geodesic distance between the two points over the manifold. Despite some weaknesses that will be further discussed below, this algorithm has been successfully applied in artificial vision and the visualization of biomedical data (Gerber and al, 2010).

The input of the Isomap is the distance matrix among the subjects. Linking each point to its K nearest neighbours, a graph is created that is supposed to describe a low dimensional manifold. In the Isomap algorithm, the geodesic distance is first estimated between points (here the sulci of each subject). The distance between two points is the length of the shortest path between these two points in the graph constructed. It is important to choose an appropriate neighbourhood size. When the neighbourhood is too large, too many "short-circuit" edges would be created; when the neighbourhood is too small, the graph becomes too sparse to approximate the geodesic paths. To our knowledge, there is no consensual general way to choose K whatever the problem. This is the main weakness of the Isomap approach. It is topologically unstable, erroneous connections could be created in the neighbourhood graph, as the "short-circuiting" effect when the K is too large. Some methods were proposed to overcome such problems such as removing nearest neighbours that violate local linearity of the graph (Van der Maaten and al, 2008). Once a matrix of geodesic distance has been computed, a simple dimension reduction algorithm such as MultiDimensional Scaling (MDS) can then be applied.

There are different choices concerning the MDS algorithm. The classical MDS algorithm of Torgerson chooses a N-dimensional configuration minimizing the stress defined by: $\Sigma(gij - dij)^2/\Sigma dij^2$, where gij denotes distances in the original dimension, and dij denotes pair wise distance in the low dimensional space (Borg and Groenen, 2005). Many variants of the classical MDS exist. They can be divided in general to two categories, the metric and non-metric MDS. The metric MDS generalizes the optimization procedure to a variety of loss functions and input matrices with different weights. It tries to find an embedding in a lower dimension such that distances are preserved. The non-metric MDS

instead try to find a non-parametric monotonic relationship between dissimilarities from the high to the low dimensional spaces. It is an estimation problem under ordering restrictions, the rank order of entries in the data is assumed to contain significant information. When the input data is “distance-like” but not actual Euclidean distances, or when we are interested in finding the patterns in the input, non-metric MDS may prove to be an interesting approach (Cox and Cox, 2001).

In this work, the non-metric MDS (IsoMDS) as well as metric MDS are tested. Considering that the input similarity is an estimation of the “likeliness” of two forms, not an actual Euclidean distance, the ordering of the folds may be more important than the absolute distance. The non-metric approach may reveal important properties such as intrinsic patterns of distribution.

While the Isomap approach has found some success, some limitations exist. It could be trapped in local optima. The non-Euclidean description based on Shared Nearest Neighbour (SNN) could be an interesting approach, where the strength of the link between two data points is based on the number of their shared neighbours (Ertöz et al., 2002). Other algorithms emphasizing the preservation of local properties such as LLE (Roweis and Saul, 2000) could be investigated in the future. The approaches that can find global optimum can be further explored. One example is the Sammon mapping, an alternative to the classical MDS, where the scaling cost function is adapted by weighting on the inverse of the pair wise distance. Thus larger pair-wise distances will not be emphasized over smaller pair wise distances. It is reasoned that the pairs with smaller distances are important to the local geometry of the data (Van der Maaten and al, 2008).

3.4 Validation of the ICP-Isomap approach

The approach of ICP and Isomap for cortical folding analysis would need to be validated. The efficiency of the clustering algorithm PCBB presented in Chapter Two needs to be

verified as well based on the new approach. Below the validation is done on both simulated and real datasets.

3.4.1 The suitability of ICP-Isomap for cortical pattern analysis: real dataset

The efficiency of the Isomap as a dimension reduction tool is validated first, in the context of the study of the cortical folding patterns. Three sulci are chosen for the validation (**Fig 3.2A**): the central sulcus, the superior temporal sulcus and the cingulate sulcus.

A dataset of $3*62$ shapes is generated combining the datasets of the three different sulci from our most recent manually labelled database (Perrot et al., 2009a). Dimension reduction of the ICP-based distance matrix is performed with three alternative classical approaches: Isomap, classical MDS and PCA. The results are shown in **Fig 3.2B**. Referring to the central sulcus (in black), one of the most stable sulcus: Isomap gives a dense presentation, the representation of classical MDS is less dense, and the result from PCA is the least dense. The same trend exists for the superior temporal sulcus (in green). Concerning the cingulate sulcus (in red), this trend is less obvious; Isomap organized the distribution into a distinctive shape. The significance in terms of real shape distribution is studied and presented in the section below. Considering the three sulci together, the Isomap algorithm separates the sulcal distribution more successfully than the other two methods. In conclusion, Isomap outperforms the other methods for sulcal analysis.

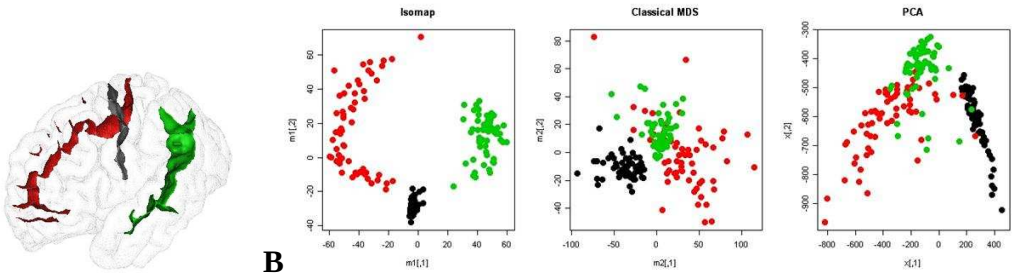


Fig 3.2 The comparison of dimension reduction methods

A: The central sulcus in red, the superior temporal sulcus in green and the cingulate sulcus in red. **B**: From left to right: the distribution using Isomap, classical MDS and PCA.

3.4.2 The suitability of ICP-Isomap for cortical pattern analysis: simulated dataset

The performance of the clustering algorithm is also evaluated with simulations. For each simulation, three subjects are picked randomly from the original database. Six random variations are generated for each of them. Each variation results from a random transformation applied to the original sulcus. This transformation is an affine transformation endowed with a diagonal of 1 and with 6 random numbers sampled from a Gaussian distribution. An example is provided in **Fig. 3.3**. The database of 62 central sulci (Perrot et al., 2009a) is used for the generation of simulated datasets. Each of them is composed of 3 simulated tight clusters of 7 sulci plus 41 original central sulci, leading to a total of 62 sulci.

Additional subjects are picked randomly among the 62 minus 3 other subjects to complete the dataset. Ten different sets of three subjects are picked, and five different standard deviations ranging from 0.11 to 0.15 are used for generating the deformation. A total of 50 simulated datasets are obtained. For each simulation, the ICP-based distance matrix is computed, Isomap is used for dimension reduction (d=4, K=6).

Three different clustering methods are applied: PCBB, Gaussian mixture modelling (GMM) and Partitioning Around Medoids (PAM). PAM is an algorithm similar to K-means including estimation of the optimal number of clusters. In our experiments, R cluster toolbox is used (Crawley, 2007). The estimation of the number of clusters performed by PAM in this context has been shown to be reliable (refer to Chapter Two). GMM involves first fitting a mixture model by expectation-maximization and computation of posterior probabilities (Duda et al., 2000). The Bayesian Information

Criterion (BIC) provides the number of components. The state-of-the-art Mclust toolbox from R is used to run GMM (Fraley and Raftery, 2006).

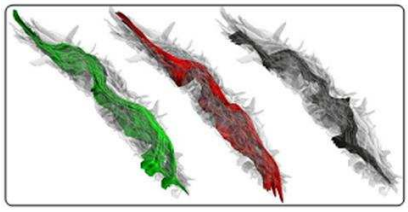


Fig 3.3 The simulated central sulcus

Shown in grey are all the real shape of the central sulcus from the dataset superimposed. Shown in green, red and black are three simulated clusters. It can be seen that the variability within each cluster is different, however, the shapes of all three clusters follow closely the real shape of the central sulcus. During the alignment for the visualization of the sulcal shapes, the most neutral shape is chosen for better visualization.

When the clustering using PCBB is carried out using the dimension-reduced similarity matrix of the three methods described above, it is found that Isomap provides the best clusters: the clusters that are tight within the group and distant among the groups. Refer to **Fig 3.4** for the clusters found. The members of the clusters are shown in black, red and green respectively. Isomap algorithm consistently finds tight clusters compare with the other two methods.

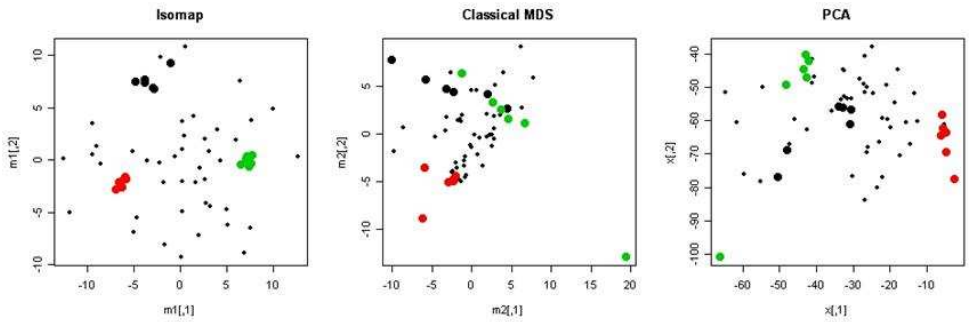


Fig 3.4 The comparison of dimension reduction methods on final clusters

Left to right: the distribution using Isomap, classical MDS and PCA. The clusters shown are those found by running the PCBB algorithm.

Fig 3.5 shows the performance statistics in terms of the number of clusters found. The simulation has three arbitrary clusters, most of the clustering algorithms tested here found one or more clusters. The distribution of the number of clusters found can give an indication of the quality of the results, even though the data might hide other valid clusters in addition to the three simulated ones. It can be observed that the PCBB algorithm has the highest incidence of finding three clusters, PAM finds slightly less, GMM still less. Overall, GMM may find none, one, two or three clusters. PAM and GMM finds at least one cluster, their incidences of finding one of two clusters are less than GMM. PAM performs reasonably well compared to PCBB, however, it is less likely than PCBB to find all three simulated clusters. **Fig 3.6** presents one of the examples of the clustering results highlighting the real clusters found. The quality of clustering is evaluated on multiple runs. It is found that PCBB outperforms the two other methods. This is not so surprising when we consider the nature of these different clustering algorithms. PAM and GMM aim at providing a complete partitioning of the dataset. This goal is not always compatible with the detection of tight clusters. PCBB, on the other hand, is designed specifically to detect dense regions and pick out subjects from the dense regions.

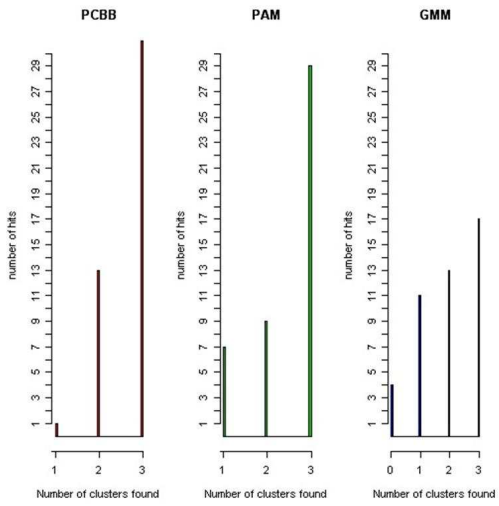


Fig 3.5 Comparison of algorithms: effect on the estimation of the number of clusters

The three algorithms compared are: PCBB, PAM and GMM. The upper image presents the result on the number of clusters found (the simulated data has three clusters).

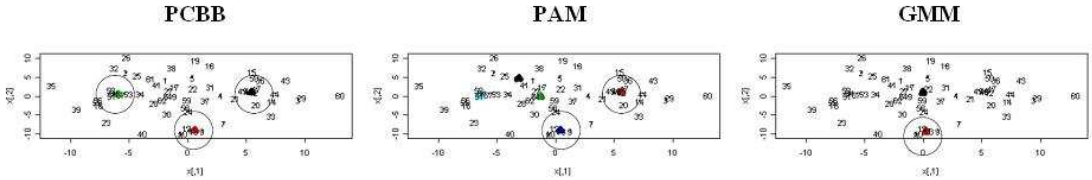


Fig 3.6 Comparison of algorithms: clusters found

The three algorithms compared are: PCBB, PAM and GMM. An example of the clustering results is shown applying the three methods. The clusters found are circled. The simulated data has three clusters.

3.5 How to look: the clusters of sulci

Once the dimension reduced similarity matrix is obtained, the clustering can be carried out using the clustering algorithm PCBB discussed in the previous chapter. The characteristic patterns of each cluster can then be visualized.

3.6 Where to look: the selection of folds

As discussed in introduction, we try the approach of selecting a group of folds or sulci that define loosely the functional regions of the cortex. The simplest groups would consist of single folds; the most complex group would be consisting of all the folds in the cortex. There are all the possible combinations in between the two extreme cases. In the case of a single fold, the advantage is that the resulting patterns are easier to be visualized and interpreted; the disadvantage is that some interesting patterns could be overlooked. In the case of the folds of the whole brain, the advantage is that the information is the most comprehensive and complete, the obvious disadvantage is that the difficulty in visual interpretation of the results. Here we try to find a compromise between the two extreme cases. We attempted to define regions of neurological significance, instead of arbitrarily selecting regions of a certain size.

3.7 Results

The database of 62 central sulci (Perrot et al., 2009a) is used for the clustering analysis. Once the lower dimension representation similarity matrices are obtained, clustering can be carried out using PCBB as described in the previous Chapter. Clustering was applied to 32 regions (sulcal sets) using PCBB. Group of clusters with scores below 0.05 were collected for the dictionary. 13 sets of the left hemisphere and 12 sets of the right hemisphere passed the threshold.

In the following sections, some of the most representative results are briefly presented and discussed. Refer to Chapter Five for a more complete representation and discussion of the results.

3.7.1 The Cingulate region

The cingulate region is found to have the strongest clusters. In this study, the strategy is changed from using the single sulcus to the Cingulate region, which includes the cingulate sulcus and the smaller sulci around it (paracingulate sulcus, intralimbic sulcus, and superior and inferior rostral sulcus). Essentially all the sulci around the corpus callosum ventral and dorsal to the cingulate sulcus are taken into account. Compared with the clustering on the cingulate sulcus alone, the clustering on the region can reveal different and more comprehensive information on variability. Some of the potential labelling ambiguities can be further investigated, for example, the labelling of the intralimbic sulcus, the paracingulate sulcus, and the superior and inferior rostral sulcus.

The resulting clusters of the cingulate region are shown in **Fig 3.7**. Five patterns are found. Notice that the purple pattern has a parallel segment on top, which may correspond to the paracingulate sulcus. Across the patterns, this paracingulate sulcus takes a varied appearance. At the posterior end, the curvature appears to be slightly different for the five patterns. The cyan pattern is relatively curved; while the blue pattern is more flat. Compared with the results obtained from Chapter Two on the cingulate

sulcus, the position of the interruption is not as clear on the regional clustering results. The interruptions might be filled up by another fold not labelled as the cingulate sulcus. It is thus interesting to approach the clustering analysis using both of the method from chapter Two and from this chapter.

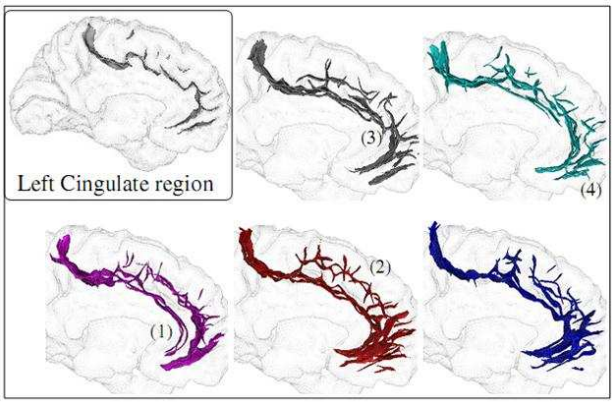


Fig 3.7 The clusters found on the Cingulate region

One example of the sulcal set is shown first. Then for each pattern, three aligned subjects are superimposed in order to highlight the areas of stability making up the patterns. The **left cingulate region** is highly variable. The key features are (1) the development of the intracingulate sulcus (long shallow fold at the bottom of violet pattern), (2) the development of the paracingulate sulcus (series of small folds at the top of red pattern), (3) the interruptions of the cingulate sulcus and (4) the shape of the anterior part of the region. The posterior part is relatively stable.

Such clustering analysis promotes the further investigation of certain issues. Refer to **Fig 3.7**, the pattern 1 (the purple pattern) shows a possible intracingulate sulcus ventral of a less “heavy” cingulate. This observation also leads to the question: is it possible that when paracingulate is prominent, the cingulate becomes intracingulate? This subject will be further discussed in Chapter Five. In pattern 2 (the red pattern), it can be observed that, compare with pattern 1, the paracingulate is more broken and more likely becoming vertical to the cingulate sulcus. The anterior region tends to be heavier as well. In pattern 3 (the grey pattern), the interruptions of the paracingulate sulcus can be observed. The pattern 4 (the cyan pattern) may have a more curved posterior.

3.7.2 The Central Sulcus

The central sulcus is discussed here as an example of the sulci or sulcal regions where clusters cannot be found due to the homogeneity of the variability. Using both approaches, the 3D moment invariant or the ICP-Isomap approach, when clustering with PCBB, no clusters can be found.

It is realized that from this approach of clustering alone it is not straightforward to deduce some simple characteristics. When observing the shape of the central sulcus, it is clear that even though the shape variability is less than the cingulate sulcus, variability still exist. In such cases when clusters cannot be found, it would be interesting to observe the variability in a more systematic fashion.

3.7.3 The inferior frontal region

The result of the inferior frontal region is shown here to illustrate the potential and limitation of such a regional approach. In the regions of the cortex where the folding is highly variable, the results could be difficult to interpret by visual inspection.

The clusters of the inferior frontal region are shown in **Fig 3.8**. The Inf-Inter frontal gyrus is made up of the intermediate precentral and intermediate, marginal, orbitary and inferior frontal sulci. Three patterns are found. The main difference among the three patterns shown here lies in the different configurations of the intermediate frontal sulcus. The violet and the cyan patterns show small and split intermediate frontal sulcus, the red pattern shows a large and transverse pattern instead. Notice that the red pattern shows a configuration more parallel to the Sylvian valley compare with the other two patterns, which are more perpendicular to the Sylvian valley. This observation is in agreement with the hypothesis of sulcal roots in terms of the orientations of the sulci, where the units of sulcal roots are organized in a system of meridians and parallels of the cortical

surface (Regis et al., 2005). While these patterns are interesting, the interpretation of the patterns could become very subjective. When defining the region for study, it is thus important to keep the region relatively small in highly variable regions.

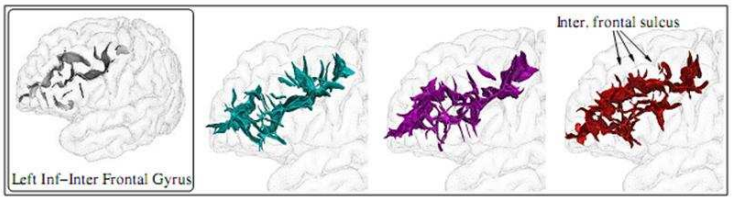


Fig 3.8 The inferior frontal region

The main difference among the three patterns shown here lies in the different configurations of the intermediate frontal sulcus (small and split: violet, large: red, large and transverse: cyan).

3.8 Discussions

New dimensions have been explored in the following three directions: the selection of cortical region for pattern search, data presentation and the algorithm for analysis. Regions of interest are defined where all the sulci within the region can be studied, so the analysis is not restricted to a single sulcus but a group of sulci. Each sulcus is presented as its distance to all the other sulci in the dataset, and the algorithm to reduce the dimension of the data is explored. Validation is done to ensure that this new approach is suitable to the study of cortical folding. It is also validated that the clustering algorithm PCBB developed in Chapter Two is suitable for pattern search using this new approach.

The results are adding new patterns in addition to those found in Chapter Two. The definition of region however, should be carefully selected. The highly variable region should not be large, so that local variations can be studied in detail. When the region is large and the folding very variable, the folding patterns found could be very difficult to be reliably described. Note that the results presented in this chapter are preliminary, the main goal is to validate the methods chosen for sulcal analysis.

In the study presented in this chapter and the previous one, we noticed that a huge difference exists in terms of the degree of variability among different folds. For example, the cingulate sulcus is very different across individuals, while the central sulcus is very stable and quite similar across the population. Indeed, the strongest clusters are found when studying the cingulate sulcus, while no strong clusters are found when studying the central sulcus. This leads to the question that, regarding the more stable folds, is there a way to describe their shape variability? In the next chapter, an interesting alternative is found when no strong clusters exist. In the attempt to give a more comprehensive description of the folding distribution, we try to establish a framework where a more general description can be given. This would complement the description based on clusters alone.

Chapter Four: Handedness in the folds

4.1 Summary

In this chapter, the methods developed in the previous chapters are applied to a neuroscience issue with interesting discoveries. In this study, we investigate the possible manifestation of handedness in cortical folding patterns. Three populations are studied: the right-handers, the left-handers and the hand-converters: lefthanders who were forced to use their right hand for writing. The investigation focuses on the central sulcus, one of the most stable and prominent sulcus that separate the motor and somatosensory areas of the cortex (Penfield and Boldrey, 1937). Anatomical MRI images are used; brain folds are extracted automatically using the BrainVisa software platform (Mangin et al., 2004b). The difference in terms of shapes among the central sulcus of all subjects are calculated, this information is further reduced in data dimension using Isomap.

A strikingly simple morphological trait of the central sulcus was discovered which can be described as the more or less centred position of the hand "notch" or "knob", where hand knob corresponds to the well-established functional motor hand area (Yousry et al., 1997). The notch position can shift along the central sulcus, the location can be from around the middle of the fold moving upwards (dorsally). Also observed is a second "knob" below the hand knob. When the hand knob position moves upwards, the second lower knob is becoming more prominent and moves upwards as well.

The two hemispheres are found to be different in terms of the shape of the central sulcus. The left central sulcus of the three populations (the left-handers, the right-handers and the hand-converters) has a characteristic "two-knob" pattern, i.e. the hand knob and a second knob below. The right central sulcus, on the other hand, has a characteristic "one-knob" pattern, containing a prominent hand knob; the second lower knob is very weak or missing.

The shape of the left central sulcus for both the left-handers and the hand-converters are found to be statistically different to the right-handers (right-hander vs. left-hander: p=0.002; right-hander vs. hand-converter: p=0.02). On the right hemisphere, the central sulcus of the converters has a tendency to be different to the left-handers (p=0.06), the right-handers has a tendency to be different to the left-handers as well (p=0.15). The hand knob position of the left-handers is more centred than that of the right-handers for both hemispheres.

The study provides insights into handedness and the effect of converting hand use. The results also reveal the exciting potential of such algorithmic approaches in the analysis of cortical folding patterns. In the following, the study of handedness is introduced, methods are described, and results are then presented. The chapter concludes with discussion and future work.

4.2 Introduction

Brain asymmetry is a fascinating topic with a huge amount of research behind. Humans, dogs and chickens alike are not perfectly symmetrical, in terms of anatomy as well as behaviour. It is shown that dogs wag their tails to the left when facing an unfamiliar dominant dog, while wagging their tails to the right when facing their owners. Marmoset monkeys, which communicate using bird-like calls, opened the left side of the mouth wider when expressing fear, and opened the right side of the mouth wider when making social contacts. Similarly in humans, the left side of the mouth is more prominent for emotional expression, while the right side of the mouth more for speech (Corballis, 2009).

Handedness and language are two of the most prominent examples of brain asymmetry or laterality in humans, with much research being done. Language laterality can be measured reliably, while the cause is still unclear. In terms of handedness, exactly how and why the brains of the left-handers are different from that of the majority right-

handers is an intriguing question that many are trying to decode and understand (Sun et al., 2006). Various researches are conducted to investigate the behavioural, physiological, and anatomical differences between the left-handers and the right-handers. The left-handedness is found to be linked to a spectrum of conditions, ranging from schizophrenia to special talents, passing through birth defects, shorter life span, dyslexia and autism, to mention just a few (Coren, 1990). The consequence of being left-handed is not well understood, neither is the cause. Many hypotheses are proposed, from purely genetic (Annett, 1972; McManus, 1985) to mainly environmental (Provins, 1997). While it is more accepted today that there are genetic factors involved in handedness, it is not clear how important is the role of the environment.

Not that long ago the left-handers were trained to switch the hand use in many cultures, especially for writing, due to the belief that left-handedness is awkward and abnormal. A Newspaper story on the 20th November 1922 under the title "*Left-handedness is cured among pupils*" reported: An intensive campaign to cure left-handers among pupils in local schools here [Elizabeth, New Jersey, USA] has resulted in a reduction from 250 to 66 since 1919 (Coren, 1990).

While it is astonishing to read the story today, the converters, the left-handers that were forced to change their hand use, make up an interesting population to study the consequences of a specific learning experience on the brain, both in terms of anatomy and in terms of behaviour. It should be kept in mind that the majority of the individuals who have a past history of shifting hands do not result in a change in handedness classification. The action of switching hands is found to lead to either an increase of the use of both hands, or to a tolerance for the use of the contralateral limb (Porac et al., 1990).

Studying the anatomical difference of different populations is one of the ways to further our understanding of handedness and brain asymmetry in general. In this quest, the brain shape and folding patterns represent a unique opportunity. Brain folds to accommodate

the growth while the fibre connections are maturing (Van Essen, 1997; Welker, 1988). Brain folding patterns are incredibly variable from one individual to another. For a more detailed introduction to cortical folding, refer to Chapter One. The variability in cortical folding might be partly due to genetics, and partly due to the fact that each individual has a unique environmental growing experience from the first day of embryo development. Many studies are conducted on brain anatomy and folding morphology. For example, the difference between the left-handers and the right-handers in terms of the hemispheric size, the size and shape of the corpus callosum, the folding patterns of the regions around the Sylvian fissure, the size of the motor and somatosensory areas, the cerebellum and the basal ganglia are studied, just to name a few (Jäncke and Steinmetz, 2003).

Among the many sulci of the human cortex, the central sulcus has been a subject of many interesting investigations (Amunts et al., 1996; Yousry et al., 1997). It marks the well-known separation of the motor and the somatosensory regions anterior and posterior of the central sulcus, where the famous homunculus is mapped (Penfield and Boldrey, 1937) refer to **Fig 4.1a** for an example. Penfield also observed and hypothesized that the motor sensory strips were arranged in horizontal strips extending from precentral to postcentral sulci, through the central sulcus (Penfield and Jasper, 1954). This observation is confirmed in animal experiments (Murphy et al., 1978). Among the huge variability of sulcal folding patterns that exist, the central sulcus is one of the most stable and one of the easiest to identify. The central sulcus can be observed appearing around 20-21 weeks of gestation, and it attains its more definitive curvature in the eighth month of gestation (Ono et al., 1990).

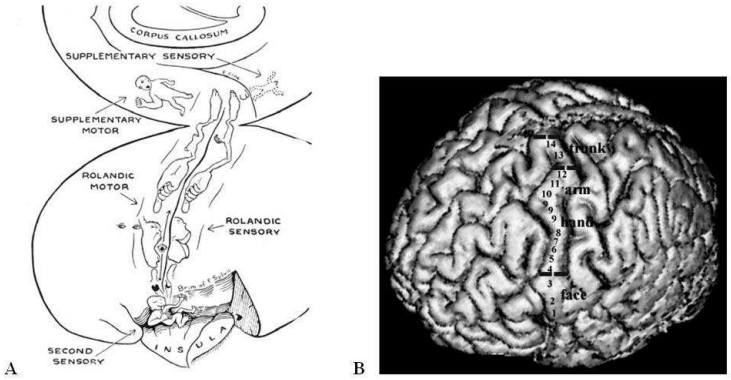


Fig 4.1a The functional organization of the central region

A The homunculus of Penfield: figurines drawn on the left hemisphere (Penfield and Jasper, 1954)
B The functional subdivision of the precentral gyrus *1*, Tongue; *2*, lips; *3*, face; *4*, thumb; *5*, index finger; *6*, middle finger; *7*, ring finger; *8*, little finger; *9*, hand; *10*, wrist; *11*, elbow; *12*, shoulder; *13*, trunk; *14*, proximal leg (Tamraz and Comair, 2006)

In anatomical literature, the central sulcus is described as a two-piece structure by Broca, and a three-piece structure by Testut-Laterjet and Paturet. The two-piece structure of Broca consists of the superior and the inferior genu. The superior knee is directed anteriorly, and that of the inferior knee posteriorly. In the three-piece model, the upper and lower knees (genou supérieur and inférieur) are anteriorly directed, and the middle knee (genou moyen) is posteriorly directed. The upper knee is always well-defined, while the lower knee is more variable and less well-defined (Ono et al., 1990). Refer to **Fig 4.1b** about the two/three-piece structure. Talairach and Tournoux described the form of the central sulcus as “step-like” or a “zigzag” (Talairach and Tournoux, 1993). At the level of the Sylvian fissure, anterior to the central sulcus, a small sulcus called the sulcus subcentralis anterior (anterior subcentral sulcus) can be seen in some cases. This small sulcus may unite with the lower end of the central sulcus in some cases, and frequently it may unite with the precentral sulcus. In some cases a posterior subcentral sulcus can be observed posterior to the central sulcus (Eberstaller, 1890).

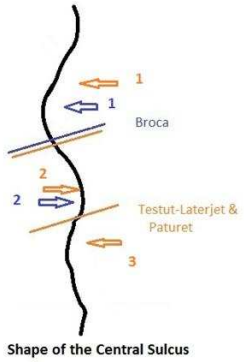


Fig 4.1b The structure model of the central sulcus.

In the two-piece model of Broca, the superior knee is directed anteriorly, and the inferior knee is directed posteriorly. In the three-piece model of Testut-Laterjet and Paturet, the upper and lower knees are anteriorly directed, the middle knee is posteriorly directed (Ono et al., 1990).

Functional MRI studies reveal that the hand motor area is located at the middle genu (number 2 in **Fig 4.1b**) of the central sulcus in a portion of the precentral gyrus that displays a characteristic “knob” or “knuckle”. This “knob” which has either an omega shape (90%) or an epsilon shape (10%) can be reliably observed in the axial plane (Yousry et al., 1997). It has also been shown that the cortical representation of the sensory hand area is located along the anterior bank of the postcentral gyrus at a characteristic curve of the central sulcus immediately posterior to the motor hand area (Boling et al., 2008).

The precentral gyrus in between the central sulcus and the precentral sulcus is the motor region. The functional subdivision of the precentral gyrus can be defined in four parts: the inferior face region, the middle hand-arm region, the superior trunk region and the paracentral leg-foot region. It is interesting to note that the depth of the central sulcus corresponding to these four regions are different: the face level (first 3cm) is averaging 15mm, the hand-arm region roughly to its midpoint is averaging 17mm, the trunk region averages 12mm due to the annectant gyrus (the PPFPM) that reduces the depth, the interhemispheric leg portion reaches at most 13mm (Tamraz and Comair, 2006). The hand-arm region is thus the deepest section of the central sulcus. The profile of the depth could be corresponding to the timing of apparition of the sulcus, the deeper being earlier. The depth could be affected during later differential growth as well.

It should be kept in mind that despite the fact that the central sulcus is relatively simple and stable, the actual shape from one individual to another is still extremely variable, refer to **Fig 4.2** for some examples. With respect to the converters, the action of converting the hand use for a prolonged period of time is expected to leave certain anatomical trace around the hand motor and sensory areas of the brain. A related recent study reveals that the converters have anatomical changes both on the surface area of the central sulcus and in the deep structure such as the putamen (Kloppel et al., 2010).

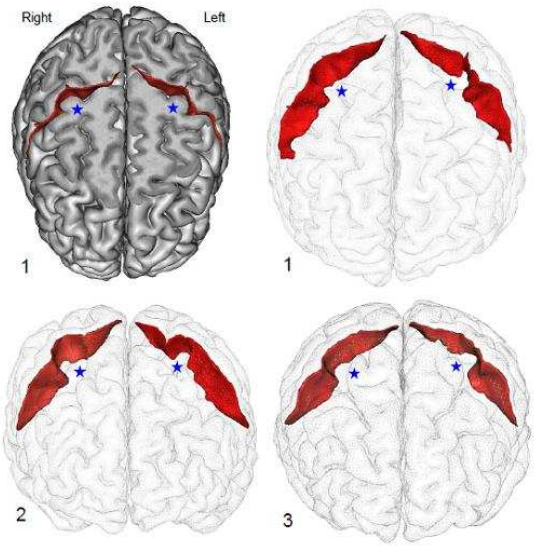


Fig 4.2 A negative cast of the central sulcus of three subjects

The left and right central sulci of the three subjects are highlighted in red, the hand knob position is marked in blue stars.

Tools such as magnetic resonance imaging (MRI) allow us to study the anatomy of the brain noninvasively, refer to chapter One for a more detailed introduction. Our contribution in this work is to use computer algorithms to discover new information in terms of the 3D shape of the central sulcus.

4.3 Material and Methods

In this section, the datasets used are described in detail, followed by the methods used for the measure of similarity among the subjects. The algorithm used for dimension reduction of the similarity matrix is then presented; finally the visualization algorithm used for obtaining the summary of sulcal forms is discussed.

4.3.1 Datasets

Two datasets are used. In the first dataset, 31 'converted' left-handers *(mean age 40, range 24-56 years; 22m, 12f)* are compared with 19 consistent age and sex matched right-handed (*mean age 34, range 22-59 years; 17 males, 6 females*) and 16 similar left-handed subjects (*mean age 36, range 25-56 years; 12 males, 6 females*). Individuals were only labelled as a 'converted' left-hander, if the subjects and their parents clearly recollected that writing commenced with the left hand at school, but was switch to the right. Data from these participants were reported in a previous study (Kloppel et al., 2010), which lists all inclusion and exclusion criteria. All subjects gave written informed consent and the local Ethics Committee approved the experimental procedures. The Edinburgh handedness inventory (Oldfield, 1971) was used in which a score of –100 reflects extreme left-handedness and +100 extreme right-handedness. The second dataset contains the subjects from the training base of the Brainvisa software (Perrot et al., 2009a).

Additional information regarding the family history of the hand converters is available for the hand converters (Kloppel et al., 2010). The impact of family history on sulcal forms is investigated.

For the first database, high-resolution structural MRI was performed on a 3T system (TRIO; Siemens, Erlangen, Germany) with a T1-weighted FLASH 3D sequence (TR = 15 ms, TE = 4.92 ms, flip angle 25°, 192 slices, slice thickness = 1 mm, matrix: 256 x 256 mm). For the second database, the 62 healthy subjects come from 6 different scanners and research protocols (Perrot et al., 2009a). An expert confirmed that the

software correctly identified the central sulcus. To control for the influence of variable brain size the central sulcal representations are all normalized to the linear standard Talairach reference frame (by a 9 parameter affine transformation). The right central sulci were flipped relative to the inter-hemispheric plane to allow asymmetry studies.

The information concerning the number of left-handers, right-handers and hand converters in each of the two datasets are listed in **Table 4.1**. Since the first dataset has much less right-handers than the left-handers and hand converters, in some analysis the subjects from the second dataset are added to the first to balance the number of left-handers and right-handers. The detail is further explained in 4.3.3. Note that some subjects of the base set are of unknown handedness, so a general 10% left-hander is estimated. We are interested in understanding the effect of input data size and data composition (in terms of percentage of each population in the input data) on the final analysis results. The second dataset (mostly right-handed individuals) provides an opportunity to vary the percentage of right-handers in the dataset. It provides also a bigger dataset for analysis.

The central sulci of all subjects of the first dataset are extracted and automatically labelled by BrainVisa. For the second dataset, all central sulcus are manually labelled by an expert. The central sulcus is described as a list of 3D points.

Dataset	Total # subjects	# left-handed	# right-handed	# converters
1. Left-handed set	66	16	19	31
2. Base set	62	Estimated 5 (8-10%), 3 known	Estimated 57, 40 known	0

Table 4.1 Handedness Information of the Two Datasets

The distances among the central sulcus of different subjects are calculated using the Iterated Closest Point (ICP) algorithm. This efficient algorithm iteratively aligns one

sulcus relative to another through successive rotations and translations. The iterations are driven by minimization of the average distance between shifted and target sulci. For each pair of sulci, the ICP algorithm is applied twice before and after sulcal role switches. The smallest residual distance across the two alignments is used as a similarity measure. A distance matrix is obtained. The Isomap algorithm combined with isoMDS (refer to Chapter Three for detail) is then used to reduce the dimension of the distance matrix. Statistical tests are performed on the dimension reduced distance matrix. Weighted SPAM is used to visualize the shape of the central sulcus. In the following, each step is discussed in more detail.

4.3.2 Sulcal similarity measure

The ICP algorithm is applied for calculating the distance or similarity of two sulci. Many variants of the classical ICP algorithm exists, here the simple rigid alignment is used for the estimation of similarity. This choice is made since the simple alignment is estimated to capture the most prominent variation of two forms. Based on the fact that little prior knowledge is known in terms of how the forms of the two sulci involved would vary, a rigid alignment for this preliminary study is deemed most appropriate.

Using the ICP algorithm, a distance matrix is obtained. The dimension of this matrix is twice the number of subjects in the dataset, since both the left hemisphere and the flipped right hemisphere are used. For each subject, there are N numbers specifying its distance to the N sulci in the dataset.

The resulting distance matrix is not symmetrical. This matrix needs to be processed to be made symmetrical for further analysis. Some simple approaches to obtain a symmetrical distance matrix experimented are: take the sum of the distance from A to B and the distance from B to A; take the maximum of the distances A to B and B to A; or take the

minimum of the distances A to B and B to A. These approaches provide different information and will be explored in the following sections.

4.3.3 Dimension reduction with Isomap

Isomap is a powerful algorithm for the analysis of high-dimensional datasets (refer to Chapter Three for detail). Here, the algorithm is used to reduce the dimension for further analysis and visualization. In the Isomap algorithm, two parameters need to be specified: the dimension and the neighbourhood size. In this work, the final Isomap-reduced similarity matrices of dimension one are used for statistical analysis. The dimension two distance matrices are also obtained for visual inspection in terms of the forms.

4.3.4 Visualization of the forms using weighted SPAM

We would like to have a reliable method to study the folding variability as well. In this study, we chose to explore further the information hidden in the Isomap reduced lower dimension representation. When the dimension is reduced to one or two, the folds can be easily ordered and visualized. Visual inspection is important in this type of exploratory analysis to ensure the quality of the algorithm. For example, using Isomap with data dimension reduction to one, each sulcus is represented by one number. Usually, moving from one extreme to the other of the axis, the corresponding shape of the individual sulcus is changed gradually. Visual inspection can provide important information regarding this gradual change of form.

To ensure that the hypothesis in terms of sulcal shape based on visual inspection is objective, we developed a technique highlighting the specific shape variability underlying this feature of the axis. For this purpose, we adapted the classical Statistical Probability Anatomical Map (SPAM) strategy consisting in averaging images after alignment (Evans

and D.L.Collins, 1997). Several averaged sulci can be computed at regularly spaced intervals along the Isomap axis, we call the strategy the weighted SPAM. For any specific location, the closer a sulcus is to this location, the more its shape contributes to the average image. The weight of this contribution follows an exponential decay relative to the square of the distance to the location. Hence, each average image provides a good representation of the shape of the sulci around the corresponding location in the axis.

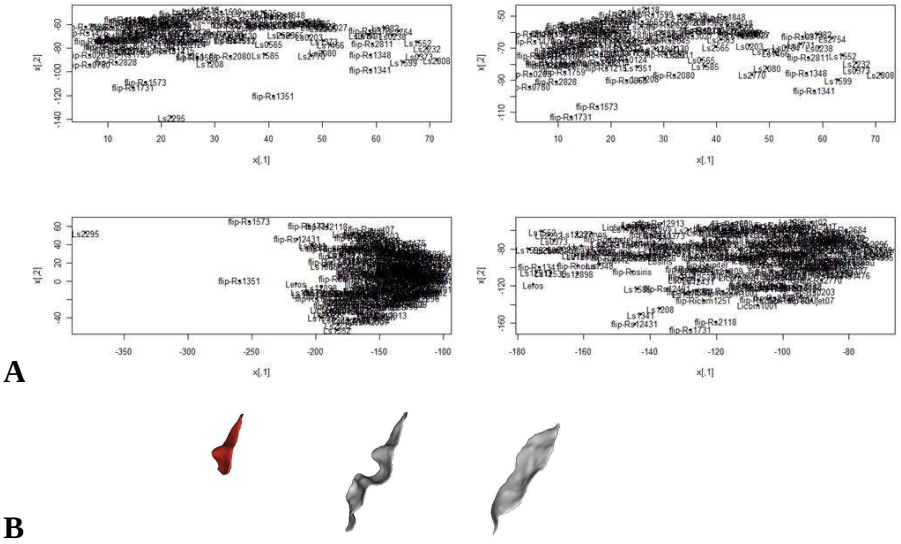
The technique of weighted SPAM can be used to visualize one dimensional as well as two dimensional similarities. In the first dimension, more prominent characteristics in terms of the folds are expected. While in the second or higher dimensions, progressively more details are expected. In a comprehensive description of the folds, different dimensions are expected to reveal different information. The sulcal forms obtained using weighted SPAM are summarized in the section below.

4.4 Results

The results are presented in six different sections. Section one presents the distance matrix representation; section two presents the outlier selection process and results; section three presents the algorithmic parameter selection results; section four presents the analysis of individual sulcal forms; section five presents the analysis and visualization of statistically summarized sulcal forms; section six concludes with the results of statistical analysis.

4.4.1 Distance matrix composition

Two approaches of composing the distance matrix discussed in the method section are compared: the minimum distance and the maximum distance. The resulting distributions of the original distance matrix for the two approaches are shown in **Fig 4.3A**.



A

B

Fig 4.3 Data distribution before and after outlier removal and correction

A: First row: the minimum distance before (left) and after (right) outlier removal. Subjects "Ls2295" and "flip-Rs1351" are considered as outliers. Second row: the maximum distance before (left) and after (right) outlier removal. PCA is used for plotting the sample points. Subjects "Ls2295", "flip-Rs1351" and "flip-Rs1573" are considered as outliers.
B: From left to right: Ls2295 before correction; Ls2295 after correction; Rs1351

It is difficult to judge from the PCA distribution which method is better. It is likely that these approaches provide different information regarding the similarities among the sulci. Minimum distance contains less outlier compared to the maximum distance approach. This is likely due to the fact that there are frequently branches and even interruptions on the central sulcus. The minimum difference approach can in a sense overlook these details and match the main form. The maximum approach, on the other hand, would emphasize these differences. The minimum approach is used in further analysis in this study, to focus on the study of the main shape difference of the central sulcus.

4.4.2 Outlier selection

It is important to remove outliers once the similarity matrix is constructed, before Isomap analysis is carried out. The presence of outliers might skew the final result, since this

method is sensitive to noise. Outlier subjects are selected by choosing sample points that are outside the three times the standard deviation range, from the average distance to the whole set. The distribution of the dataset after the outliers are removed is displayed in **Fig 4.3A**. The distributions are more homogenous after the removal of the outliers. The shapes of the outliers are verified to ensure the ICP and PCA combined method of identifying outliers is effective, which are displayed in **Fig 4.3B**. In the final analysis, the minimum distance is used, subjects Ls2295 is corrected and Rs1351 removed as outlier.

4.4.3 The choice of K

Here dimension one is used for preliminary study and visual comparison of results. The neighbourhood size k is chosen to be 7 and 14 as 10% of the whole datasets One and the combination of One and Two. The effect of k on the final statistics is discussed in section 4.4.6.5.

4.4.4 Results in terms of the position of the hand knob

The forms of all central sulci from dataset One and Two (refer to **Table 4.1**) are presented in **Fig 4.4**, using the axis specified by Isomap. Note that at the two extremities, interrupted sulci are found. At the left extremity the interruption is more ventrally located, while at the right extremity it is more dorsally located. Three parallel lines are drawn to aid in visualization. Around the middle line, it can be observed that the hand knob has a tendency to move dorsally from the left to the right of the Isomap axis. This tendency is further clarified using weighted SPAM.

Results are obtained using only Dataset One and the combination of Dataset One and Two are investigated. In both cases, the real sulci and the weighted SPAMs are shown (see **Fig 4.5** and **Fig 4.6**). The gradual change in shape from one extreme to the other

along the axis can be clearly seen in both cases. The weighted SPAM helps in the visualization of the location of the knob. The hand knob moves up while a lower knob appears, moving from the left to the right of the axis. Notice that in the case of the combination of datasets (**Fig 4.6**), the right extreme of the Isomap axis is more populated, which provides a smoother form transition in terms of the SPAM visualization. The characteristic of the location of the hand knob moving upwards from the left to the right of the axis is consistent. Moreover, as the hand knob moves up, a ventral second knob becomes more evident as well.

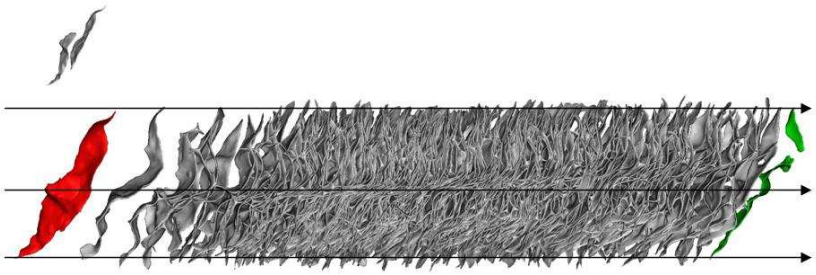


Fig 4.4 The central sulci aligned along the Isomap axis

The two datasets are mixed to generate the Isomap. Plot here is the central sulci of all subjects along the axis defined by Isomap in one dimension. The two extremity sulci are highlighted in red and green respectively. The red sulcus is depicted at top left in another orientation. The red sulcus has a ventral interruption, the green sulci has a dorsal interruption.

The exact functional role of the lower knob is unknown. Further investigation is needed also to investigate the correlation between the positions of the upper and the lower knob. Dataset Two contains mostly right-handers; it can be observed that the right extreme of the Isomap axis is more populated in the presence of this dataset. It is likely that the right-handers exhibit a more double-knob configuration, with the first knob pushed more upwards, and the second knob becomes more prominent. The results presented on the next section when the handedness groups are studied separately would reveal more detail.

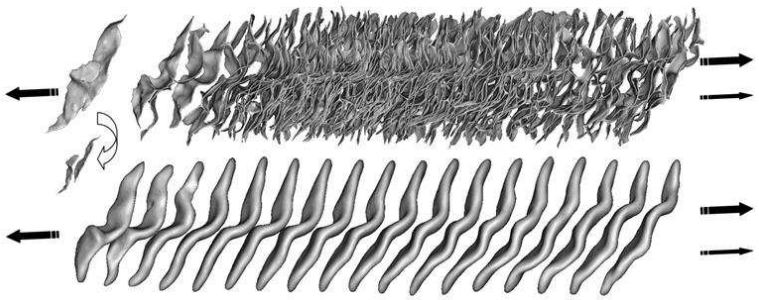


Fig 4.5 Hand knob position moving dorsally from left to right along the Isomap axis.

(***Top***) First, all the sulci were aligned to a template sulcus in order to obtain similar orientations across the whole set. Then each sulcus was translated along the antero-posterior brain axis proportionally to its Isomap coordinate. Note the outlier sulcus located at the extreme left of the axis corresponding to an atypical interrupted central sulcus (this sulcus is also presented from a different viewpoint below to make this clearer). (***Lower***) Local averages of the central sulci were computed at regularly spaced positions to clarify shape variability coded on the Isomap axis.

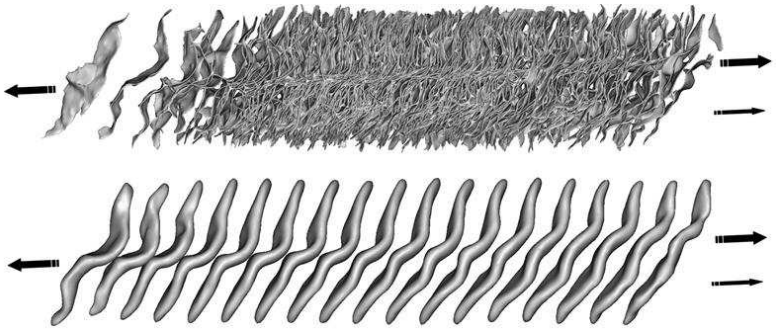


Fig 4.6 Hand knob position of Dataset One combined with Dataset Two

Analogous to Fig 4.5, the combination of dataset One and Two are used instead. Note that a second interrupted sulcus has been pushed to the extreme right of the axis.

4.4.5 Results in terms of hemispheric asymmetry and handedness

The sulci of different groups are then plot separately, refer to **Fig 4.7**. Dataset One is analyzed; the sulci of different handedness groups on the two hemispheres are plot along the Isomap axis. The shift in sulcal shape from one hemisphere to another for each handedness group can be compared easily. The SPAM of the median of the form of different handedness groups on the two hemispheres are plot also, which aid in

summarizing the characteristic shape of each sub-group (**Fig 4.7**, last row). In **Fig 4.8**, the same information is plot on the combination of dataset One and Two.

It can be observed that the SPAM of the right hemisphere has a more characteristic single knob, while the left hemisphere has a more double knob characteristic. The median shape of the dominant hemisphere for hand (right hemisphere for left-handers and left hemisphere for right-handers) are more different among the three groups, and the median shape located more away from the centre of the Isomap axis. This can be more easily observed on the SPAM images when the median shapes of different subgroups are plot on the same axes (**Fig 4.7**, last row). The right central sulcus of the left-handers (cyan) is located more towards the left extremity of the Isomap axis, the left central sulcus (green) of the right-handers is more towards the right extremity of the axis. The left central sulcus of the left-handers (red) and the right central sulcus of the right-handers (purple) cannot be distinguished by their median shape, notice that on the SPAM summaries of **Fig 4.7** and **Fig 4.8**, these two forms are at the same location, towards the middle of the axis. When the hand converters are concerned, it can be observed that their right central sulcus (dark blue) is shifted to be towards the centre of the axis, the median form quite similar to the right-handers (purple) on the right hemisphere. The left central sulcus of the converters (orange), on the other hand, is more shifted towards the right extreme of the axis, where the average from of the left central sulcus of the right-handers (green) is located.

It is interesting to observe that the average form of the non-dominant central sulcus (purple, red) tend to be located towards the centre of the Isomap axis; the dominant ones (green, cyan) tend to be shifted towards the extremities. However, depending on handedness or hemisphere, the direction in shape shift is not the same. For the right-handers on the left hemisphere (green), the shift is towards the right, where a characteristic “double knob” configuration is observed, as discussed in the previous section. For the left-handers on the right hemisphere (cyan), the shift is towards the left of the axis where the “single knob” configuration is observed. Considering the hand

converters, their left central sulcus is moved away from the centre towards the right of the axis (to be more like the right-handers); their right central sulcus is more like that of the right-handers as well. This shows the evidence of a “neutral zone” towards the centre of the Isomap axis, which represents a “neutral form”. More use of the dominant hand (either due to genetics or environment or both) induces a shift away from this “neutral shape”. More on this observation will be explored in the discussion.

Another observation is that the spread of the sulcal forms are different for the dominant and non-dominant hemispheres. Consider the right-handers, the form of the right central sulcus (purple) is more evenly spread over the Isomap axis, while the left central sulcus (green) is found more dense towards the right of the axis. For the left-handers, the form distribution of the dominant hemisphere (cyan) is denser towards the left of the axis, compared to the non-dominant hemisphere (red). Such a shift in density cannot be clearly observed in the case of the hand converters.

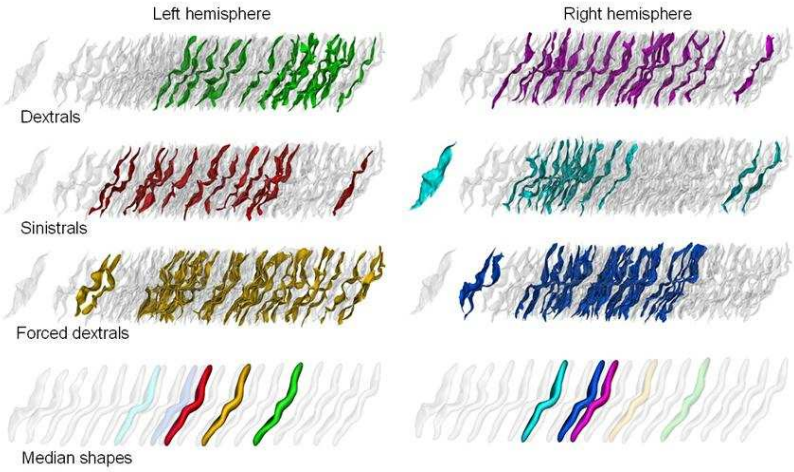


Fig 4.7 Population localization along the Isomap axis

Fig 4.5 is plotted as a transparent background. Note the fact that the forced dextral group is double the size of the natural handedness groups ***Lower*****:** Local average sulcus computed for the median location of each population. Colour code: Cyan: left-handed, right hemisphere; Blue: converters, right hemisphere; Purple: right-handed, right hemisphere; Red: left-handed, left hemisphere; Yellow: converters, left hemisphere; Green: right-handed, left hemisphere

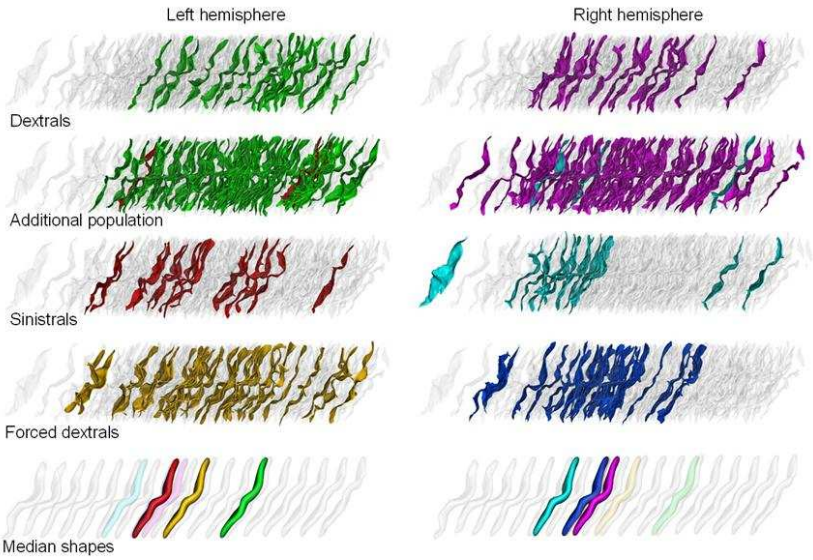


Fig 4.8 Population localization along the Isomap axis for dataset One and Two combined

The additional database charts include right-handed and left-handed subjects, which are colour-coded correspondingly. Colour code: Cyan: left-handed, right hemisphere; Blue: converters, right hemisphere; Purple: right-handed, right hemisphere; Red: left-handed, left hemisphere; Yellow: converters, left hemisphere; Green: right-handed, left hemisphere

The SPAMs of the Isomap of dimension two are shown in **Fig 4.9**. It can be observed that in the dimension from left to right, the same trend exists as that of the 1D Isomap. Concerning the other dimension from top to bottom, another trend can be implied. The hand knob is changed from smaller to be more prominant, while the position of it remains the same, and not sliding up or down. Another interesting observation is that while the hand knob becomes more profound, the sulci form is becoming more flat, especially regarding the two extremities of the form.

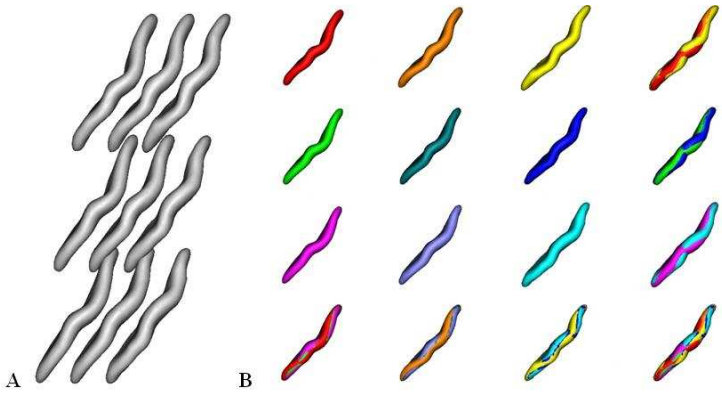


Fig 4.9 The 2D weighted SPAM

A: the 2D weighted SPAM **B**: The first three rows and columns correspond to the weighted SPAM at different coordinates as in **A**, the fourth row and the fourth column correspond to the first three rows and columns merged together for easy comparison.

4.4.6 Statistical analysis

4.4.6.1 The difference of the three groups

Statistical analyses are carried out to study the difference of the three groups, the left-handers, the right-handers and the hand converters. Looking at the left hemisphere, it is discovered that the shape of the central sulcus are significantly different between the left-handers and the right-handers ($p=0.002$). The right-handers, compared to the left-handers, have a more prominent “two-knob” structure, with the hand-knob shifting upwards and the lower knob more profound. The “two-knob” configuration might reflect the fact that the left hemisphere is the dominant hemisphere for both language and hand for the right-handers. The converters have the shape pattern statistically similar to the left-handers ($p=0.22$), and statistically different from the right-handers ($p=0.02$).

On the right hemisphere, it is discovered that the shape of the central sulcus is not significantly different for the left-handers and the right-handers, even though the tendency exist ($p=0.15$). However, there also is a strong tendency for the converters to be different from the left-handers ($p=0.06$), making them very similar to the right-handers

(p=0.56). Refer to **Fig 4.10** for a comparison of the sulcal forms of the three groups on the two hemispheres.

The asymmetry index of the form is simply calculated as (r-l). The right-handers are found to be asymmetrical in form (p=0.0002). The converters are found to be asymmetrical in form (p=0.03). The left-handers are not asymmetrical.

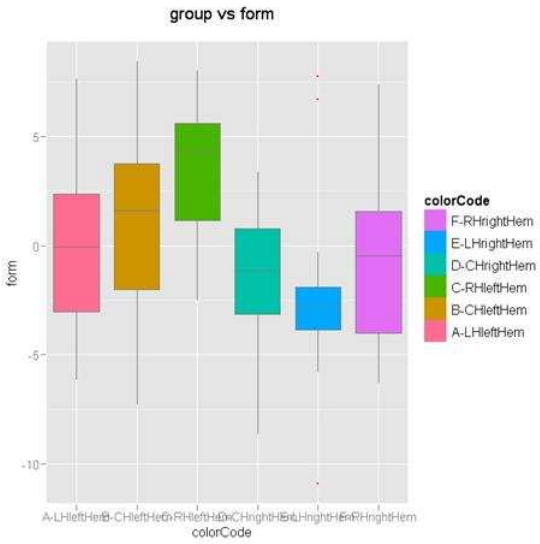


Fig 4.10 The analysis of form of the central sulcus of the three groups on the two hemispheres

The box plot of sulcal form of the three handedness groups on the two hemispheres. The colour-coding is the same as Fig 4.5, namely: Cyan: left-handed, right hemisphere; Blue: converters, right hemisphere; Purple: right-handed, right hemisphere; Red: left-handed, left hemisphere; Yellow: converters, left hemisphere; Green: right-handed, left hemisphere

4.4.6.2 The correlation of various measures to shape

The correlation of sulcal form with various measures such as the length, the surface area, the depth and the thickness of the central sulcus are studied. The thickness here refers to thickness of the cortical mantle on both sides of the sulcus. The form is found to be not correlated to surface, depth and thickness on both hemispheres. The length is found to be correlated to form (p=0.05) on the left hemisphere, but not correlated on the right

hemisphere. This correlation on the left hemisphere is anticipated since the right-handers are known to have a longer and larger left central sulcus.

It is verified that the algorithm indeed captured a unique shape characteristic, not a by-product of other parameters such as the length or the surface area.

4.4.6.3 Handedness correlates in length

The stability of left central sulcus shape of converters contrasts with the flipped asymmetry of surface area observed in the previous study with the same population (Kloppel S et al., 2010). To complement Klöppel et al. study, we similarly tested handedness correlates for central sulcus length and depth asymmetries and found no effect with depth (we used the index (2*(r-l)/(r+l)). However, the length asymmetry indices replicated the surface area flipped effect. The length asymmetry index distributions of the natural handedness populations are different ($p=0.002$), confirming a previous study (Mangin JF et al., 2004). The length indices of forced and natural right-handers cannot be distinguished ($p>0.5$) however converters differ significantly from left-handers (Wilcoxon, $n=31/16$, $p=0.02$). In order to obtain greater insight, we also tested whether the absolute length of the central sulcus is different between populations after affine spatial normalization into a standardised (Talairach) space. The only differences in length found were in the left hemisphere in the two naturally handed populations ($p=0.01$), and also on the left, between converters and lefthanders ($p=0.06$). Natural left-handers have a shorter left hemispheric sulcus. No significant differences were found in terms of absolute sulcal surface area or depth.

4.4.6.4 The influence of family history

Additional information regarding family history of handedness for the hand converters is provided. The analysis of the influence of family history on the shape of the central

sulcus is carried out. Compared to the right-handers on the left hemisphere, converters without family history (p=0.08) are less different than converters with family history (p=0.01) in sulcal form. It is also found that the asymmetry index is stronger (p=0.05) for converters without family history than those with family history (p=0.9).

On the left hemisphere the converters with family history (median:-1.8) tend to be more similar to left-handers (median:-0.06). The median of converters without family history is 1.7, the median of right-handers is at 4.35 respectively. The corresponding medians on the right hemisphere are as follows: -2.83 for left-handers, -1.4 for converters with family history, -0.8 for converters without family history, and -0.47 for right-handers (**Table 4.2**).

leftHem/rightHem	median	Difference to other groups (pVal)		Difference within group (pVal)		Asymmetry index	
		lefthanded	righthanded	history	noHistory	mean	pVal
history	-1.8/-1.4	0.8/0.3	0.01/0.7	--	0.2/0.7	-0.3	0.9
noHistory	1.7/-0.8	0.2/0.1	0.08/0.7	0.2/0.7	--	-2.8	0.05

Table 4.2 Statistics of converter with or without history

4.4.6.5 Data distribution and parameter selection

The distribution pattern of the dataset is studied to decide the statistical tests to use. This provides an opportunity to observe the behaviour of parameter selection related to statistical results. The distribution of the data is deviated from Gaussian (Kolmogorov-Smirnov test, p=0.66), the Wilcoxon test results were consequently used for analysis.

The effect of varying the neighbourhood parameter K is explored (K = 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35). The goal is to study the behaviour of the Isomap using various neighbourhood sizes. It is also important as a verification of the impact of neighbourhood

size on the resulting p-values. Does changing the neighbourhood size affect the resulting p-value on the difference of form among the handedness groups? To what extent is this variation?

The resulting p-values using various neighbourhood sizes are displayed graphically in **Fig 4.11**. For these experiments, the dataset with the whole Set One is used. Both the classical multidimensional scaling (CMDS) and the Isometric MDS (isoMDS) methods are used; the statistical results of running the Wilcoxon test are shown. The groups studied are the left-handers and right-handers on the left and right hemisphere.

The results show that varying the neighbourhood size has little impact on the conclusion for the tests that are already statistically different. For the difference on form between the right-handers and left-handers on the left hemisphere, the p-value remains significant with various neighbourhood sizes, with the minimum around k=7. For the difference on form between the right-handers and left-handers on the right hemisphere, the p-value remains not significant with various neighbourhood sizes. The isoMDS and CMDS approach produce similar trends in variation of p-values. The isoMDS approach, however, provides lower p-values in general.

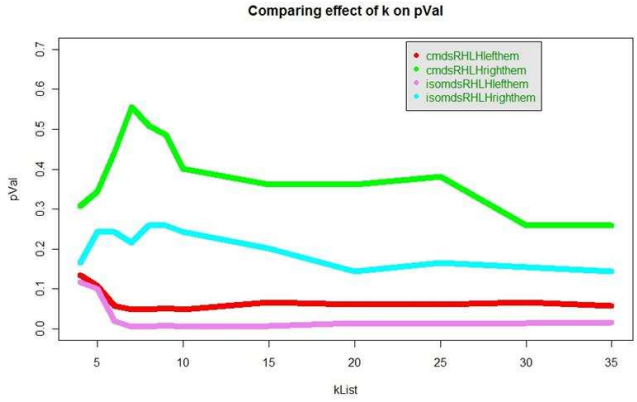


Figure 4.11 Neighbourhood size and p-value

The lines present the evolution of p-value while varying the neighbourhood size k. The red and violet lines present the Isomap using classical and isometric MDS for dimension reduction, both on the difference between the form of left-handers and right-handers on the left hemisphere. The green and cyan lines present the Isomap using classical and isometric MDS for dimension reduction, both on the difference between the

form of left-handers and right-handers on the right hemisphere. The minimum for the violet and the red line are both at k=8.

4.4.6.6 Influence of data composition on the results

The central sulcus of the left-handers, the right-handers and the hand converters are expected to be not completely identical. The percentage of each group in the dataset could have an impact on the resulting Isomap reduced similarity measures. When the majority of subjects in the dataset are from one group, the Isomap may focus on the difference within the group. Some experiments are performed to better understand the behaviour of Isomap with different percentage of right-handers in the whole dataset. The two original datasets (**Table 4.1**) are mixed (Refer to **Table 4.3)** for the study of the impact of data composition on the statistical results.

Dataset	Total # subjects	# left-handed + converters	# right-handed	% right-handed
Whole Set 1	66	47	19	29%
Set1 + half of Set2	97	Estimated 49	Estimated 48	Estimated 50%
Set 1 + whole Set 2	128	53	75	58%

Table 4.3 Handedness Information of Mixing the Two Datasets

Row 1 and 3: the dataset One and the combination of set One and Two. Row 2: the new experimental dataset with more right-handers added to increase the total percentage of right-handers

The three datasets form interesting samples with different percentage of left-handers in the set. It is likely that when there are equal amount of left-handers and right-handers in the dataset, the Isomap can capture the most difference between these two groups. When one of the handedness groups outnumbers the other one, the focus could be partly shifted to the difference within the majority handedness group. To further investigate, statistics is done on comparing the form of the central sulcus of the lefthanders and the right-handers on the left hemisphere, where known significant difference is found between these two groups. The resulting p-value could be an indication of the focus of Isomap on differentiating the forms of different handedness groups. It can be observed that the p-

value is the most significant (0.0009) for set1 combined with set 2, where there are roughly equal amount of left-handers and right-handers. When the percentage of the right-handers increased to 58%, the p-value increased slightly (0.006). When the percentage of right-handers is 30%, the p-value is 0.002.

This experiment confirms the influence of data composition on the results. Choosing a balanced dataset with equal amount of the subjects from different groups could give more focus on discovering the differences among the groups.

4.5 Discussion

The discussion is divided into three sections due to the interdisciplinary nature of the work. The first section focuses on the insights gained on handedness and cortical development, in the second section some computational and algorithmic issues are discussed, the third section concludes by discussing the future directions.

4.5.1 Insights in Neuroscience

The intriguing neuroscience issue we tried to investigate here is: if and how are the hand converters changed following the hand conversion experience? Cortical folding is used as a vehicle to study these questions; in particular the study focuses on the central sulcus due to its functional regional specificity. Can the event of switching hands leave a trace on gross brain anatomy? Twin study reveals that while the variability in tertiary folds is mainly due to environmental factors, primary folds such as the central sulcus are determined more genetically (Lohmann et al., 1999). Indeed, gyrification (cortical folding) is a rather stable property suitable for comparisons across long time spans (Francis et al., 2006). Such study may provide further insights on brain development. Are certain anatomical traits more responsive to learning? How normal development can be affected by early intensive learning? Regarding brain asymmetry, do the two hemispheres react differently facing the same learning experience?

Regarding the effect of handedness on the anatomy of the central sulcus, some differences are revealed in terms of the surface and the depth of the central sulcus, where the central sulcus is found to be bigger on the dominant hemisphere (Amunts et al., 1996; Csernansky et al., 2008; Mangin et al., 2004a; White et al., 1994), Other studies found no difference in terms of depth (White et al., 1997). In the study of the depth, it should be kept in mind that an observer-independent depth measure needs to be used such as in the work of Cykowski et al (Cykowski et al., 2008). When the depth is estimated based on the measures directly from the 2D slices, this measure could become arbitrary since the depth depends on the direction of the tissue slicing.

In this study, it is revealed that the shape of the central sulcus of the left-handers is consistently different from the right-handers; this difference is not affected by later prolonged training as in the case of converting the hand use. A trend of shape change for the converters can be observed on both hemispheres, but such change does not reach statistical significance. The characteristic shape of the left central sulcus can be described as a “two-knob” configuration, while the right central sulcus can be described more as a “one-knob” configuration.

This differential configuration might be due to the fact that language centre in the human brain is lateralized. In most of the individuals, the language is localized to the left hemisphere. The position of the second knob of the central sulcus may be related to language lateralization.

Concerning the shape of the central sulcus of the converters, the results suggest that developmental events leave different yet observable traces on the two hemispheres. On the left-hemisphere, where the central sulcus of the right-handers and that of the left-handers are found to be statistically different ($p=0.002$), the shape of the central sulcus of the converters remains similar to the left-handers and different from the right-handers. On the right-hemisphere, where the central sulcus of the right-handers and that of the left-

handers are found to be not statistically different (p=0.15), the shape of the central sulcus of the converters has a tendency to be different to the left-handers compared to the right-handers (p=0.06) This difference is not only in terms of the shape change, but also in terms of the change in the variability of shape within the group.

The results reveal interesting insights with respect to plasticity. The change of human brain can be observed at the scale from hours to years. Grey matter change can be observed several days to several months after consistent training such as juggling (Draganski et al., 2004); and central sulcus folding pattern is observed to be flattened years after amputation (Dettmers et al., 1999). The cortical folding pattern is an example of such a more stable trait. The stability of such traits may prove to be useful in the diagnosis of problems of early abnormal development. In a more general sense, such traits can be used to study how genetic interact with environmental factors such as learning in shaping our anatomy and behaviour.

The converters were forced to write in school with their non-preferred right hand, the results here show that inborn mechanisms sculpt the shape of the human central sulcus, while environmental mechanisms modify only its length, on the left hemisphere. Forcing the conversion of hand use does not modify "hand knob" location along the left central sulcus, but results in a more right-handed pattern of sulcus length asymmetry. Thus, the shape of the central sulcus, once established, is resistant to mechanisms of use-dependent plasticity during childhood. However, behavioural constraints during the critical period of learning to write modify sulcus length. In this specific case the sulcus shape reflects early developmental mechanisms and the sulcus size later environmental effects. Characterising normal variation of cortical morphology provides a means of systematically correlating behaviour with cortical development.

It can be observed that for both the right-handers and the left-handers, the distribution of the sulcal shape is more clustered on their dominant hemisphere for hand use. The non-dominant hemisphere exhibits greater spread (**Fig 4.7, 4.8**). When the converters are

concerned, they have a different distribution pattern compared to the left-handers on both hemispheres. If we assume that the shape of the central sulcus on the non-dominant hemisphere is less influenced by the impact of hand use, the shape on the non-dominant hemisphere could be then considered more ‘in-born” and less influenced by environmental variations such as hand use. Future genetic analysis could be carried out correlating with the sulcal forms of the non-dominant hemisphere.

The higher variability in sulcal form might be more “in-born”, the hand usage changes this distribution from more uniform to be more clustered. This might explain the difference of shape distribution observed on the two hemispheres of the converters. On both hemispheres, the distribution of form has a tendency (not significant) to be shifted towards the pattern of the right-hander (**Fig 4.7, 4.8**). The experience of converting hand use then is most likely to be a more complex activity beyond simply change the writing hand from the left to the right.

Concerning family history, the converters with family history tend to differ more than those without family history when compared to the right-handers on the left hemisphere. On the right hemisphere, the converters with family history tend to be more similar to the left-handers than those without family history. In summary, the converters with family history seem to resist the change in the shape of the central sulcus on both hemispheres. The results suggest that left-handedness has degrees, which is shown in the degree of resistance to change of shape. Concerning the nature of handedness, Woo and Pearson suggested that handedness might be a continuous variable (Woo and Pearson, 1927) instead of a categorical one consisting of left-handers, right-handers and ambidextrous individuals. There are also some functional studies supporting this hypothesis.

Concerning the anatomy of the central sulcus, it is interesting that the algorithm placed two interrupted central sulcus at the left and right extremities (**Fig 4.4**). Furthermore, the positions of the interruption are very different. At the left extreme, the interruption is at the inferior part of the sulcus; at the right extreme, the interruption is at the superior part

of the sulcus. This change in the position of interruption may be related to the change of the location of the hand knob.

The interruptions are generally very rare, in about 1% of the cases only (Ono et al., 1990). The interruption is believed to be caused by a non-operculated deep convolution. Broca observed that there are three bridges connecting the precentral and the postcentral gyri (Broca and Pozzi, 1888), and he termed them the pli de passage frontopariétal supérieur (PPFPS), the pli de passage frontopariétal moyen (PPFPM) and the pli de passage frontopariétal inférieur (PPFPI). The PPFPS is found at the interhemispheric fissure, corresponding to the paracentral lobule. The PPFPM is described as a bulge into the central sulcus at the level of the middle knee of the central sulcus. The PPFPI separates the central sulcus from the Sylvian fissure, corresponding to the subcentral gyrus. Cunningham confirmed the observation of Broca by describing a deep annectant gyrus between the pre- and postcentral gyrus. In rare cases, the PPFPM arises completely to the cortical surface, cutting the central sulcus into two separate parts (Cunningham, 1892).

The PPFPM is functionally linked to the hand knob region, both in hand motor activation (Boling et al., 1999), and hand sensory functions (Alkadhi and Kollias, 2004; Boling et al., 2008), the functional implication of the PPFPI is less clear. Even though the “classic” relatively straight lower knee is described (Rademacher et al., 2001), Fesl (Fesl et al., 2003) reports the existence of from two to four additional curves in this region, most commonly (69%) two additional curves. Boling et al (Boling et al., 2002) demonstrated that the tongue sensory area is within the triangular region situated at the base of the postcentral gyrus, superior to the sylvan fissure.

The two cases of interruption observed in this study might be the result of the lower (ventral) or upper (dorsal) extremities of the Pli de Passages Fronto-Pariétal Moyen (PPFPM) arising to the cortical surface. Historically, Wagner first described such interruption of the Rolandic fissure in the right hemisphere, in the brain of a celebrated

physician (Wagner, 1862). Broca found it in that of “an idiot, who was known to suffer from a large number of severe anomalies” (Broca and Pozzi, 1888). More recently, a case of a normal male with PPFPM rising to the surface on the left hemisphere is studied (**Fig 4.12**). The PPFPM is found to completely segregate the primary motor (M1) finger from the M1 elbow representation, the M1 wrist representation was consistently split by the PPFPM into a medial and lateral activation cluster (Alkadhi and Kollias, 2004). The interruptions might reveal developmental events (Regis et al., 2005), thus being an important characteristic of the central sulcus in terms of hemispheric asymmetry and function.

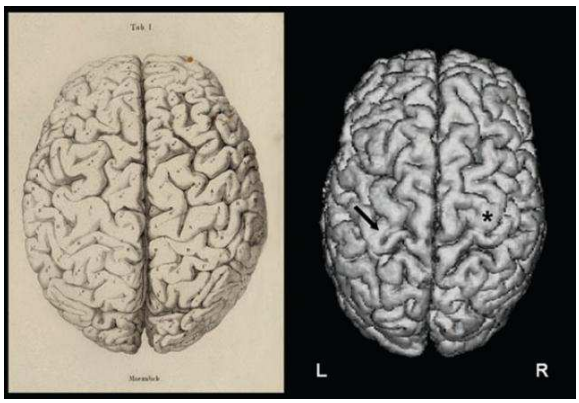


Fig 4.12 Examples of central sulcus interruptions, from (Alkadhi and Kollias, 2004)

The left image: the image of pli de passage fronto-pariétal moyen provided by Wagner (Wagner, 1862); the right image: the image of the two hemispheres of the subject studied (Alkadhi and Kollias, 2004). The connecting gyrus is marked by an arrow; the opposite side is marked by an asterisk.

From **Fig 4.4**, it can be observed that the positions of the upper and lower extremes of the hand knob are corresponding to the positions of the two interruptions. However, the hand knob and the PPFPM may be closely related, especially based on the fact that functional studies reveal consistent mapping of hand motor (Boling et al., 1999) and sensory (Alkadhi and Kollias, 2004; Boling et al., 2008) functions to the PPFPM.

Assuming that the cortical folding patterns are related to functionality (the relevance of cortical folding to functional regions is discussed in Chapter One), the anatomical information found in this study could be useful for functional analysis. For example,

based on the knowledge that the hand knob could move along the central sulcus, more precise functional localization of the hand motor region could be possibly achieved by taking into account the individual location of the hand knob.

To conclude, it is confirmed in this study that the converters are differentially changed from the left-handers by prolonged training. On the left hemisphere, while the characteristic shape of the left-handers is preserved, the consistent use of the right hand might be the cause of the elongation of the left central sulcus. On the right hemisphere, however, the impact of the lack of use or change of use of the left hand on the sulcal shape is much less profound. There is a not significant but possible weak change in the shape of the sulcus towards the “neutral shape” of the non-dominant hemisphere found at the middle of the Isomap axis, while the sulcal length remains the same. In one sense, the “more use” of the right hand and the “less use” of the left hand leave different traces on the left and right hemispheres of the brain.

This is the first time that a behavioural trait is associated with a shape-based mesoscopic feature of specific cortical folds. The sulcal shape analysis in hand converters shows that certain developmental processes are stable and unaffected by mechanisms of late, use-dependent plasticity. It should be possible to establish a dictionary of stable and variable shape-based features associated with each cortical fold, using the Isomap or similar technique, to provide a reference for comparison with brain developmental diseases. A normal morphometric description at a mesoscopic level would then lead to a series of endophenotypes that could provide stable markers of maldevelopment for gene association studies. Additionally, the cortical folding concomitants of local functional cortical changes associated with environment-dependent behavioural variability lend themselves to much more precise definition with potential prognostic implications.

4.5.2 Insights on approaches and methods

Through this analysis, insights are gained on some important issues: the choice of the similarity measure, the outlier identification, and the percentage of different populations in the data input.

The choice of the similarity measure influences the resulting similarity matrix. In this analysis, the minimum measure is chosen considering the difficulty of the reliable matching of branches in central sulcus. When a distance is calculated, it is desirable that the main piece of the sulcus is given the priority, while the details such as branches affect less the result. It is also observed that the maximum measure finds more outliers. In the future studies, the maximum measure can be used to find outliers, while the minimum measure can be used for the discovery of patterns.

It is confirmed in this analysis that ICP distance is efficient in identifying outliers in terms of sulcal shape. The identified outliers can be confirmed by further visual inspection. When corrections are needed in terms of cortical labelling, the corrections can be further verified by the updated ICP distance distribution.

Regarding the influence of the input data composition on the final results, using the Isomap approach, it is found that when a particular population out-numbers the others, more details within this population are revealed; while the other minor populations would be given less focus. The term population here can be corresponding to handedness groups such as left-handers and right-handers; it can also be corresponding to the majority patterns and the rarer patterns of a given sulcus. It is observed that for comparison between two populations, the number of subjects from each population affects the results. When possible, it would be interesting to “zoom-in” to a minor population to be able to explore it in more detail. One limiting factor is the size of the input data; the minor population must reach a certain size for reliable data analysis.

The weighted SPAM proves to be a very useful tool for the interpretation of the results. While the statistical tests give quantitative measures of the difference between groups, it

does not provide visual information on the forms. Using weighted SPAM, the forms can be perceived directly; the results of statistical tests can be further verified; visual interpretations of folding patterns can be carried out more easily as well.

In this study, the combination of algorithms proves to be a powerful tool for the analysis of 3D sulcal shapes. This approach provides a unique opportunity to analyze the central sulcus of multiple subjects simultaneously. Interesting characteristics difficult to detect even by human experts can be proposed using this approach. This is illustrated in the finding of the new “one-knob” vs. “two-knob” patterns that separates the left from the right hemisphere. It should be noted that the change from a “one-knob” to a “two-knob” pattern is gradual, the discussion of patterns of the two hemispheres are based on statistical probability distributions. For a given individual, the possibility to have a “two-knob” configuration on the right hemisphere is not excluded.

4.5.3 Future work

Further cytoarchitectonic studies and the study of underlying fibre bundles could be carried out in the future to clarify if and how the gross anatomical changes are related to the physiological and connectional differences. Further functional MRI studies based on previous work (Kim et al., 1993) could be designed to further verify the relation between the positioning of the hand notch on the 3D form of the central sulcus. The relation of the existence of the second notch to language, arithmetic and sensori-motor activities could be further explored.

The difference in terms of asymmetry between the left-handers and the right-handers could be due to both genetic and environmental factors. Young children might learn to be right-handed by imitation and living in a right-handed environment, where toys and tools are designed for the right-handed (Coren, 1990). Study of the central sulci of infants would shed some light on the importance of later environmental factors related to living

in a “right-handed world”. It would be interesting to study the shape of the central sulci of infants to understand better when the asymmetry started. The pericentral cortex region is already well developed at birth (Ono et al., 1990). It would be also valuable to study other species such as the chimpanzees to know if this type of asymmetry is unique to humans.

In addition to the central sulcus, studies can be carried out focusing on different parts of the brain, for example, the language areas, the corpus callosum and the Perisylvian region where interesting results have been found (Jäncke and Steinmetz, 2003). Different patient populations can be compared based on their anatomy; this will be of particular interest for the better understanding of certain pathologies.

The results obtained in this study invite many interesting further investigations. The current study uses mainly 1D Isomaps, more dimensions can be explored. Observing the forms of the central sulcus, it can be deduced that some individuals have a simpler and smoother **stair** shape, while others have a more **wavy** shape, there are still others with more branches than usual. Application of more dimensions of Isomap, SPAM and eventually different clustering algorithms could unveil more traits with further biological or pathological implications.

Finally, this work is an example of what “computer vision” can contribute to the analysis of cortical folding patterns. The term “computer vision” is used here in the sense that instead of using the human eye, the computer “eye” is used to explore the shape space attempting to understand the difference among the three populations. For the human eye, it is hard to go beyond some relatively simple shape descriptors. Manually, solely based on visual analysis by human experts, these characteristics are difficult to be detected. Even when such characteristics are found, they are hard to be characterized and to be analyzed systematically and quantitatively. Moreover, here it is demonstrated that computer algorithms have the power to analyze hundreds of brains simultaneously and infer interesting characteristics without any prior domain-specific knowledge. The same

task is overwhelming for human, even for experienced neuroanatomists. This approach open doors to further collaboration among the domains of neuroanatomy, and neuroscience in general, with the domains of datamining and computer vision.

Chapter Five: Sample Dictionary of Folds

5.1 Summary

In this chapter, the example dictionary is presented, using various approaches and methods detailed in Chapters Two and Three. Some selected regions are studied; interesting variability patterns found are discussed. The implications and usage of such dictionary and the potential applications are discussed in Chapter Six.

5.2 Dictionary of patterns

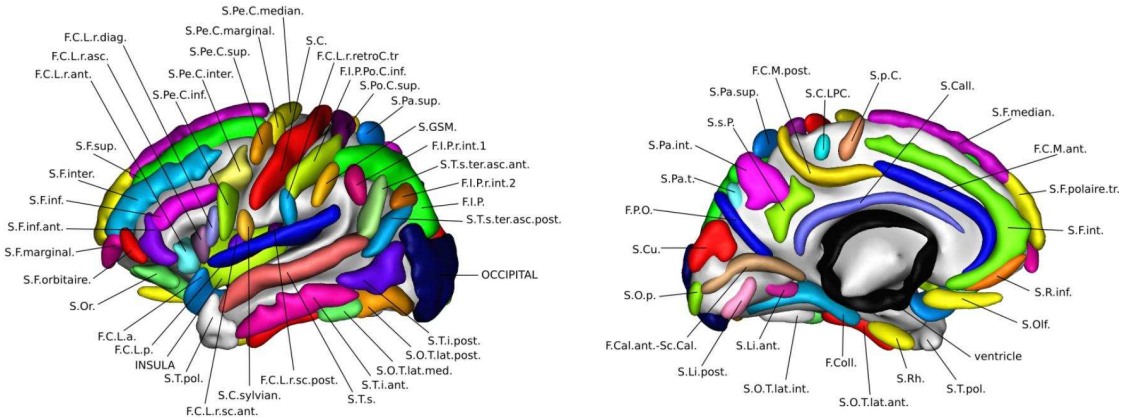
Several regions are chosen for this preliminary analysis of folding patterns. Namely: the region of the central sulcus, the region of the cingulate sulcus, the region of the superior temporal sulcus, the Broca's area, and the superior frontal region. Finally, some combinations of regions are explored. This is a preliminary systematic study of folding patterns, based on computerized automatic analysis developed in this thesis work. The list of regions can be expanded in future work. In the following, the method used for this analysis is summarized; the regions and their patterns are then introduced.

5.3 Analysis and clustering methods

The approach introduced in Chapters Two and Chapter Three are used in this analysis. The dataset consists of 62 brains manually labelled (Perrot et al., 2009a), which serve as the training base of BrainVISA. The definition of the sulci is illustrated in **Fig 5.1**. Both hemispheres are used; the right hemisphere is flipped to match the orientation of the left hemisphere for analysis. The similarity measure among the folds is calculated using the ICP algorithm. In terms of the distance measure in this analysis, both the maximum and the minimum distance are used. As discussed in Chapter Four, these two approaches could reveal different information regarding the folds. Combination of Isomap and IsoMDS is then applied to reduce the dimension of the distance matrix to two for

clustering analysis using PCBB; the dimension is also reduced to one for visualization using SPAM. The dimension two is chosen for the clustering analysis in this preliminary study so that the quality of the clustering results can be assessed by visual inspection. We target the patterns that can be observed visually in this work. It is reasoned that higher dimension images might hide details not observable in 2D visualization. Future targeted analysis can further explore the complexity of higher dimensions.

It should be noted that the SPAM images are produced given a threshold, so that the resulting images do not have holes regarding the central sulcus, refer to Chapter Four for detail. When a certain fold does not appear in the SPAM image, it could be due to the fact that the folds are not aligned. For this analysis the same threshold is used for all the regions analyzed. When a specific sub-region is absent in the SPAM image, it could be that the sub-region is absent of folds, it could also mean that the folds are not aligned at the sub-region. To prevent false interpretation of the SPAM images, the images of real sulci corresponding to the sub-region of the SPAMs should be consulted before drawing conclusions.



F.C.L.a.	anterior lateral fissure
F.C.L.r.ant.	anterior ramus of the lateral fissure
F.C.L.r.diag.	diagonal ramus of the lateral fissure
F.C.L.r.sc.ant.	anterior sub-central ramus of the lateral fissure
F.C.M.ant.	calloso-marginal anterior fissure
F.Cal.ant.-Sc.Cal.	calcarine fissure
F.I.P.Po.C.inf.	superior postcentral intraparietal superior sulcus
F.I.P.r.int.1	primary intermediate ramus of the intraparietal sulcus
F.P.O.	parieto-occipital fissure
OCCIPITAL	lobe occipital
S.C.	central sulcus
S.Call.	subcallosal sulcus
S.F.inf.	inferior frontal sulcus
S.F.int.	internal frontal sulcus
S.F.marginal.	marginal frontal sulcus
S.F.orbitaire.	orbital frontal sulcus
S.F.sup.	superior frontal sulcus
S.Li.ant.	anterior intralingual sulcus
S.O.T.lat.ant.	anterior occipito-temporal lateral sulcus
S.O.T.lat.med.	median occipito-temporal lateral sulcus
S.O.p.	occipito-polar sulcus
S.Or.	orbital sulcus
S.Pa.sup.	superior parietal sulcus
S.Pe.C.inf.	inferior precentral sulcus
S.Pe.C.marginal.	marginal precentral sulcus
S.Pe.C.sup.	superior precentral sulcus
S.R.inf.	inferior rostral sulcus
S.T.i.ant.	anterior inferior temporal sulcus
S.T.pol.	polar temporal sulcus
S.T.s.ter.asc.ant.	anterior terminal ascending branch of the superior temporal sulcus
S.p.C.	paracentral sulcus
ventricle	ventricle

F.C.L.p.	posterior lateral fissure
F.C.L.r.asc.	ascending ramus of the lateral fissure
F.C.L.r.retroC.tr.	retro central transverse ramus of the lateral fissure
F.C.L.r.sc.post.	posterior sub-central ramus of the lateral fissure
F.C.M.post.	calloso-marginal posterior fissure
F.Coll.	collateral fissure
F.I.P.	intraparietal sulcus
F.I.P.r.int.2	secondary intermediate ramus of the intraparietal sulcus
INSULA	insula
S.C.LPC.	paracentral lobule central sulcus
S.C.sylvian.	central sylvian sulcus
S.Cu.	cuneal sulcus
S.F.inf.ant.	anterior inferior frontal sulcus
S.F.inter.	intermediate frontal sulcus
S.F.median.	median frontal sulcus
S.F.polaire.tr.	polar frontal sulcus
S.GSM.	sulcus of the supra-marginal gyrus
S.Li.post.	posterior intra-lingual sulcus
S.O.T.lat.int.	internal occipito-temporal lateral sulcus
S.O.T.lat.post.	posterior occipito-temporal lateral sulcus
S.Olf.	olfactory sulcus
S.Pa.int.	internal parietal sulcus
S.Pa.t.	transverse parietal sulcus
S.Pe.C.inter.	intermediate precentral sulcus
S.Pe.C.median.	median precentral sulcus
S.Po.C.sup.	superior postcentral sulcus
S.Rh.	rhinal sulcus
S.T.i.post.	posterior inferior temporal sulcus
S.T.s.	superior temporal sulcus
S.T.s.ter.asc.post.	posterior terminal ascending branch of the superior temporal sulcus
S.s.P.	sub-parietal sulcus
unknown	unknown cortical folds

Fig 5.1 The nomenclature of the sulci (Perrot et al., 2009a)

5.4 The region of the central sulcus

The anatomy of the region of the central sulcus is discussed in detail in Chapter Four, where anatomical traits are linked to behaviour. It was found that there exist two typical configurations, the “single knob” and the “double knob” configurations. Furthermore, on the left hemisphere, the right-handers tend to have the double-knob configuration, while the left-handers tend to have the single knob configuration. Such form characteristics is relatively stable and is not changed by later learning events such as the switching of hand use in the hand converters.

The definition of the central sulcus in this study uses the combination of central sulcus (S.C.) and sylvian central sulcus (S.C. sylvian) in the version of the current automatic naming system of BrainVISA. The interruptions towards the bottom of the central sulcus is documented to be rare (around 4%) (Eberstaller, 1890; Ono et al., 1990). Refer to **Fig 5.1** for the definitions of these two sulci. Note that this image shows the SPAMs used for the automatic recognition system of BrainVISA, obtained using a different approach compared to the current work (Perrot et al., 2009a). To conform to the classical definition

of the central sulcus, here the continuous piece is used, including the portion of S.C. sylvian. The goal is to study the sulcal anatomy of the region and its characteristic variations. It should be noted that such study can give further hint to the system of nomenclature used for automatic recognition. For example, when the definition of the central sulcus is including two pieces (S.C. and S.C.sylvian), as in the automatic sulcus recognition model, the SPAM image of the (**Fig 5.1**) shows two separate pieces. How often in a given dataset that such interruption exists? What is the most typical configuration of the region of the central sulcus? The results of the current analysis may provide additional information.

Refer to **Fig 5.2**, the SPAM images are shown, it can be seen that the sulci of different subjects are relatively evenly distributed along the Isomap axis. The trend of evolution from the left to the right of the Isomap axis goes from the “single knob” to “double-knob” configuration as discussed in Chapter Four. Observe the real forms at each Isomap coordinates (**Fig 5.2 E, G**), it can be seen that even though the central sulcus is a relatively stable sulcus, huge variability in form still exists. Without the aid of the Isomap organization and the SPAM images, the evolution of the position of the hand knob cannot necessarily be detected.

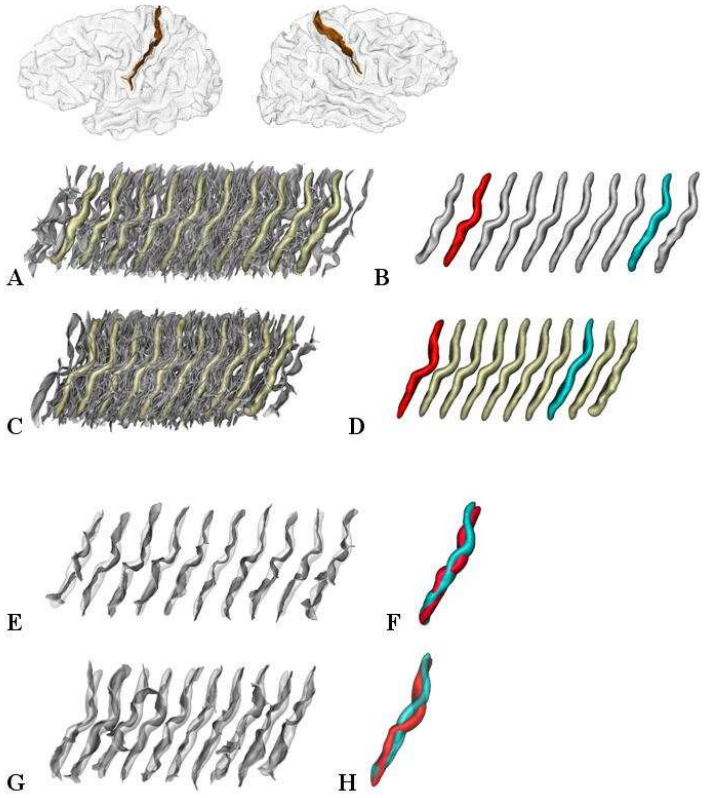


Fig 5.2 The Isomap of the central sulcus

The position of the central sulcus on the whole brain is illustrated in the first row. **A/C**: The isomap of the central sulcus using maximum/minimum distance. The sulci are superimposed on the SPAM, according to their relative positions. **B/D**: The SPAM of the central sulcus using maximum/minimum distance. The forms at the two extremities are coloured blue and red respectively. **E/G**: The SPAM using maximum/minimum distance is plot as a transparent background; superimposed on top of each SPAM of sulcus is the real sulcus with its coordinate the closest to the SPAM coordinate. **F/H**: the SPAM of **B/D** at the left (red) and right (blue) extremities of the Isomap are superimposed on each other.

The pericentral region is usually stable in morphology; the variability observed is much lower than in the other regions of the brain. In such case of a more homogeneous sulcus, finding clustering patterns is relatively difficult. In the clustering analysis introduced in Chapter Two using moment invariants, no clusters could be found. No clusters could be found neither using the approach introduced in Chapter Three. However, the clustering analysis of Chapter Three is based on analyzing the central sulcus of the left hemisphere

alone. Here, when both the left and the right central sulci are included, some clusters could finally be revealed.

Clusters are found using the minimum distance approach; the maximum distance approach gives no clusters. The two clusters found are likely representing the typical forms of the two hemispheres (**Fig 5.3**). The SPAMs of the clusters are obtained by weighting on the distance to the centre of the cluster. It is similar to the method of obtaining the 1D SPAMs, except that 2D data are used. The final result is plot at the cluster center in the 1D SPAM for visualization. The red cluster has a more characteristic single knob configuration, while the black cluster has a more characteristic double knob configuration. The relative positions of the hand knob on the two clusters can be compared when the two cluster forms are superimposed on each other (**Fig 5.3 D**). As discussed in Chapter Four, for the right-handers (estimated 95% in the dataset used for this analysis), the left hemisphere has a more typical "double-knob" pattern (the black pattern), and the right hemisphere has a more typical "single-knob" pattern (the red pattern). The plot of the data distribution with the information of the location of the clusters found conveys further information concerning the clusters (**Fig 5.3 B**). The black and the red clusters are located at the two extremities of the distribution. These clusters are not very strong, the black cluster is found mainly due to the low probability of finding a group of subjects relatively similar and not at the centre of the distribution. Keep in mind that the PCBB algorithm aims at detecting the regions where the distribution is unlikely dense.

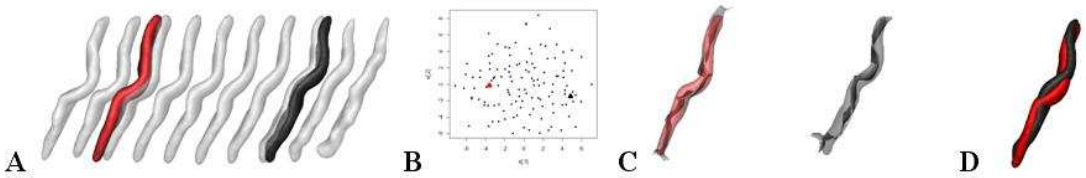


Fig 5.3 The clusters of the central sulcus

A: The locations of the clusters found are indicated along the Isomap axis. **B**: The locations of the clusters are plot on the distribution of the whole dataset. **C**: the real sulcal shapes of the subject at the centre of the clusters are plot, superimposed with the SPAM forms as in **A**. **D**: the SPAM forms are superimposed for easier comparison.

The SPAM of the central sulcus of the whole dataset can be generated as well (where each sulcus is given the same weight), which convey information of the most common form of the sulcus. Refer to **Fig 5.4**, the SPAM forms at various locations of the Isomap can be used for comparison with the average form. It can be seen that the average form (**Fig 5.4 A**) is in the middle of the form variations (plot in transparent). Notice that the variation of single or double knob can be observed when the central sulcus is viewed in a standard angle (**Fig 5.4 B**). When the angle is changed to facilitate the observation of the depth (**Fig 5.4 C**), it can be observed that the variation of the depth profile is much less prominent compared to the variability of the curvatures.

As discussed in Chapter Four, there exist three pli de passage through the central sulcus, the pli de passage frontopariétal superior, moyen and inferior (Broca and Pozzi, 1888). The superior and inferior genoux of the central sulcus is likely corresponding to the superior and inferior frontal sulci (Dejerine, 1895). Further detailed studies need to be carried out relating the pli de passage and the position of the hand knob. If the shallowest point in the depth profile is corresponding to the PPFPM (Broca and Pozzi, 1888), it is likely that the PPFPM and the position of the hand knob are two separate features of the central sulcus. While it is observed here that the hand knob position can move up and down along the central sulcus, the position of the pli de passage frontopariétal moyen (PPFPM) is supposed to be very stable (White et al., 2002). The PPFPM is thus likely corresponding to the very stable bottom of the hand knob, while towards the top of the central sulcus the variability increases, this variability manifests in terms of the position of the hand knob observed towards the surface. This is in agreement with the sulcal roots theory of the more primitive and more stable bottom or “roots” of the sulci. The variability in terms of the shape or curvature, on the other hand, conveys additional information on the accumulative genetic and environmental influence during development till adulthood.

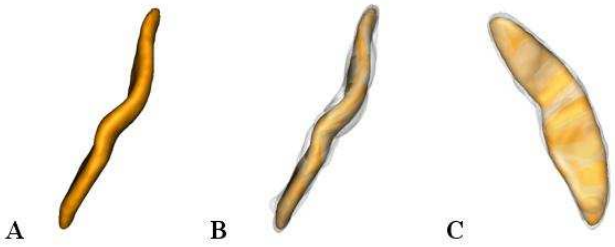


Fig 5.4 The SPAM of the whole dataset

A: The SPAM of the whole dataset is plot. **B**: the forms along the Isomap axis are plot as transparent, superimposed on the whole dataset SPAM as in **A**. **C**: as in B with a different angle

As the analysis carried out in Chapter Four, the forms of the left and the right hemispheres can be plot. Refer to **Fig 5.5**, the left hemisphere is plot in green, and the right hemisphere plot in red. It is confirmed that the left hemisphere has a tendency towards the right extreme of the Isomap axis, while the right hemisphere exhibit the tendency towards the left extreme of the Isomap axis.

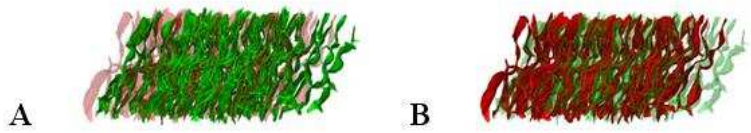


Fig 5.5 The comparison of the two hemispheres of the central sulcus

A: the central sulci of the left hemisphere **B**: the central sulci of the right hemisphere

Since the precentral gyrus is the functional motor region, the patterns of the combination of the central sulcus and the precentral sulcus are studied next. This could reveal the pattern of the precentral gyrus, as well as the sulcal variability of the precentral region.

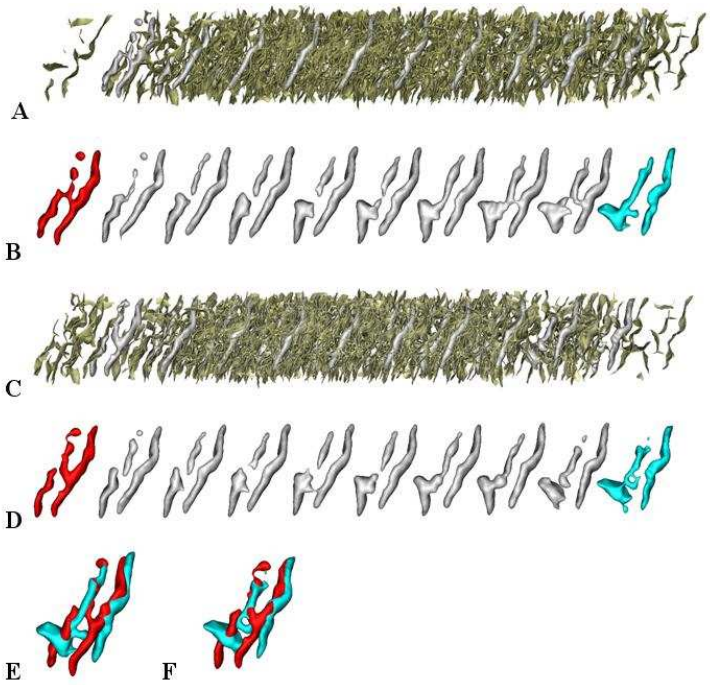


Fig 5.6 The SPAM and fold distribution of the precentral gyrus

A/C: The isomap of the precentral gyrus using maximum/minimum distance. The sulci are superimposed on the SPAM, according to their relative positions. Notice that at the two extremities there are much less sulci compared to the centre. **B/D**: The SPAM of the precentral gyrus using maximum/minimum distance. The forms at the two extremities are coloured blue and red respectively. **E/F**: the highlighted SPAMs at the extremes of **B/D** are superimposed

The analysis using both the maximum and the minimum distances provide consistent results. The SPAM images along the Isomap axis are produced. Refer to **Fig 5.6 B** (maximum distance used), in terms of the form of the precentral sulcus, from left to right of the axis, the intermediate precentral sulcus changes orientation from more parallel to more perpendicular with respect to the central sulcus. Such change in orientation observed might be related to the underlying architecture and the orientation of the fibre bundles. It is also important to notice the distribution of sulcal forms along the Isomap axis. The majority of the forms are located from the fourth to the seventh of the coordinates, while the forms towards the two extremities are rarer.

The trend described is less obvious using the minimum distance (**Fig 5.6D**). While the trend of the intermediate precentral sulcus is evident, that of the superior region is not. Notice the trend of single to double knob observed on the central sulcus cannot be deduced here. The strength and orientation of the intermediate precentral sulcus appear to be a dominating feature when the region is concerned.

To investigate the sulcal shape in more detail, individual folds along the Isomap axis can be studied. Refer to **Fig 5.7**, when the sulci located at the coordinates of the SPAM are plot, more details of the shape evolution can be studied. Based on individual variations, it can be confirmed that the intermediate precentral sulcus changes orientation from the left to the right of the axis. The superior region of the precentral sulcus is more variable in terms of shape and interruptions, which is harder to be captured by Isomap of 1D. Nonetheless, from left to right of the axis, the superior portion of the precentral sulcus is not only more continuous but also more profound or "heavier" using maximum distance (**Fig 5.7 A**). This evolution of the trait of the superior precentral sulcus is confirmed related to the change of position and orientation of the intermediate precentral sulcus.

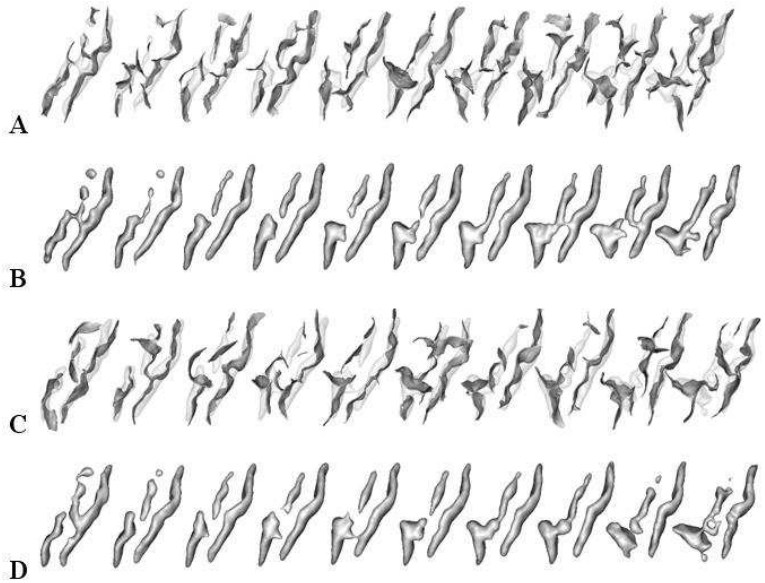


Fig 5.7 The SPAM of the precentral region with the corresponding sulcal forms

A/C: The SPAM using maximum/minimum distance is plot as a transparent background, superimposed on top of each SPAM of sulcus is the real sulcus with its coordinate the closest to the SPAM coordinate. **B/D**: The SPAM using maximum/minimum distance, plot here for comparison.

Clustering is carried out on the maximum and minimum based distance. The results are illustrated in **Fig 5.8**. Two clusters are found using the minimum distance, and two different clusters are found using the maximum distance. The black cluster is stronger than the red cluster statistically, i.e. it is more likely that the black group forms a cluster compared with the red group. Overall, the trend of change of orientation of the intermediate precentral sulcus can be observed.

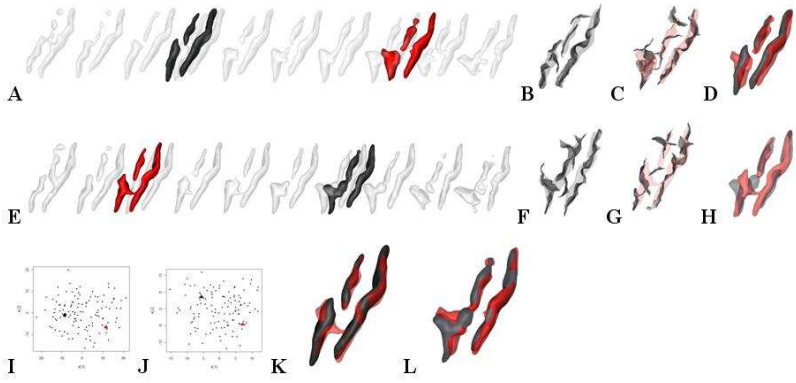


Fig 5.8 The clusters of the precentral gyrus

A/E: The locations of the clusters found are indicated along the Isomap axis, maximum/minimum distance is used. **B/F**: the real sulcal shapes of the subject at the centre of the black cluster in **A/E** are plot, superimposed with the SPAM of the black cluster. **C/G**: the real sulcal shapes of the subject at the centre of the red cluster in **A/E** are plot, superimposed with the SPAM of the red cluster. **D/H**: the SPAM forms of the clusters in **A/E** are superimposed for easier comparison. **I/J**: The locations of the clusters in **A/E** are plot on the distribution of the whole dataset. **K/L**: the clusters of **B** and **G/ C** and **F** are superimposed for comparison.

In **Fig 5.9** the average form of the precentral region as a whole is plot, it can be observed that the inferior portion of the precentral sulcus is not as thin as the superior portion. This indicates a greater variability in the inferior portion of the sulcus. This effect is likely due to the variability in terms of orientation of the intermediate precentral sulcus discussed

above. When comparing the average form with the forms at the two extremes of the Isomap (**Fig 5.9 B**), this variability in the orientation becomes more evident.

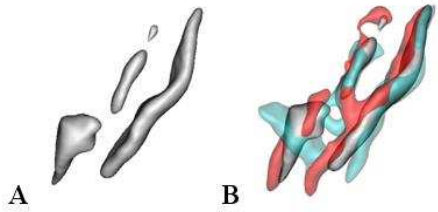


Fig 5.9 Global form of the precentral gyrus

A: the average from of the precentral region. **B**: the two extremes of the Isomap SPAM forms in red and blue are superimposed on the average form.

Another interesting direction for pattern analysis is the study of asymmetry by comparing the folding patterns of the two hemispheres. In **Fig 5.10**, the samples from the two different hemispheres are plot in different colours, using minimum and maximum distances. In **Fig 5.10 D and H**, the average forms of the two hemispheres are plot together for comparison. It can be observed that on the left hemisphere (in green), compared to the right hemisphere (in red), the intermediate precentral sulcus is more parallel in orientation to the central sulcus, and the superior precentral region is more likely to be lighter and interrupted. Such asymmetry may be related to hemisphere specific functionalities such as language and handedness. Further investigation taking into account functional and fibre tracking information would be needed to further the understanding of such interesting anatomical variability.

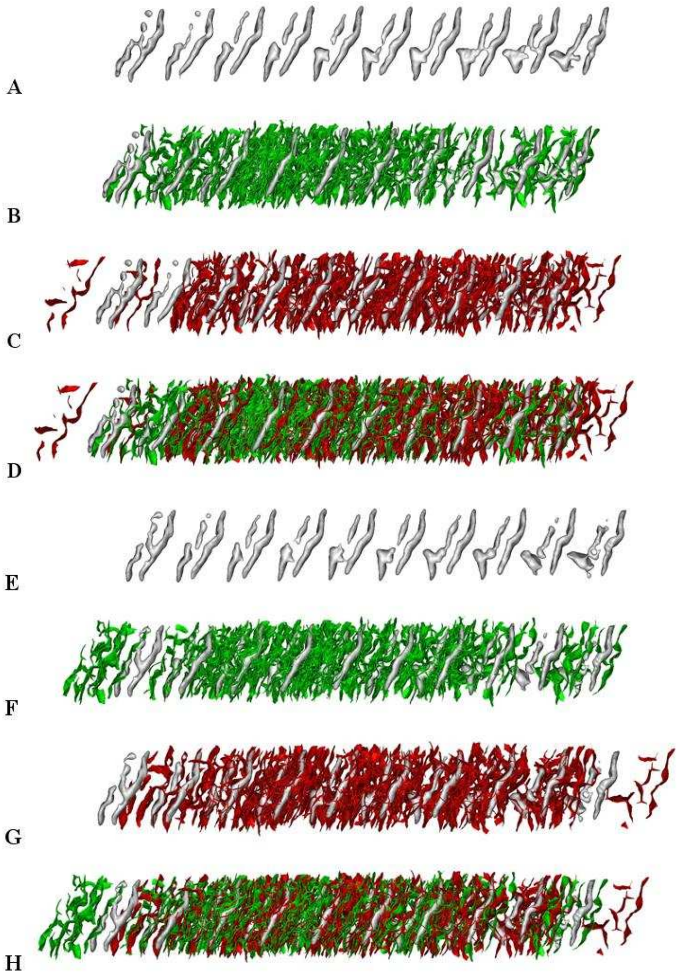


Fig 5.10 The two hemispheres of the precentral gyrus

A/E: the SPAM of the precentral region using maximum/minimum distance **B/F**: the sulci of the left hemisphere are plot in green against a grey background of SPAM using maximum/minimum distance; **C/G**: the sulci of the right hemisphere are plot in green against a grey background of SPAM using maximum/minimum distance; **D/H**: the sulci of both hemispheres are plot against a grey background of SPAM using maximum/minimum distance.

One of the possible applications of the automated analysis introduced in this work is to deduce cortical folding models based on folding variability. The Isomap approach proposes a possible framework for the organization of sulci, the clustering results provide information on the frequent folding patterns observed which are unlikely due to chance. Clinicians and researchers without extensive knowledge in anatomy of cortical folding can gain insights into the folding variability utilising such analysis tools. The folding

pattern analysis presented in this work provides a starting point for further understanding of the folding variability and the underlying causes of such variability.

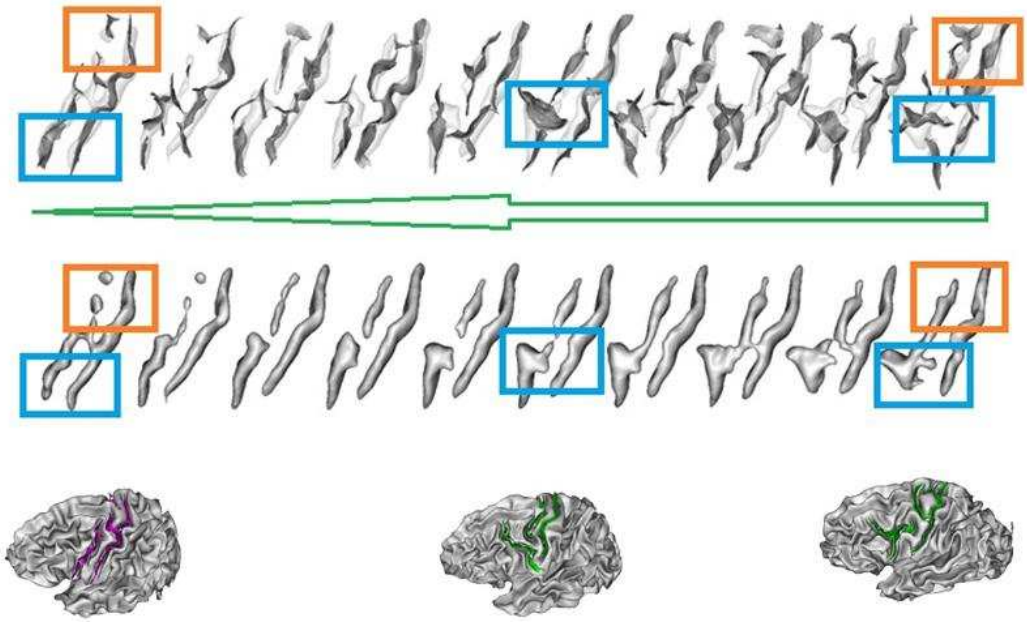


Fig 5.11 The variability model of the precentral region

The green arrow indicates the direction of change, the blue and orange squares indicate the regions of interest, for the observation of change in sulcal patterns. Three brains are given as examples of the typical patterns, the brain surface is plot together with the sulci to facilitate the localization of the sulci of interest.

Fig 5.11 presents an example of such folding variability model of the precentral sulcus. From the observations of the sulcal patterns along the Isomap axis discussed above, it can be deduced that when the intermediate precentral sulcus becomes more perpendicular to the central sulcus, the superior portion of the precentral sulcus tends to become more prominent and continuous. This is an example of a more systematic understanding of folding variability, beyond a simple categorization based on interruptions and characteristic shapes, such as that introduced in Ono. Furthermore, since such understanding is gained through the analysis of 3D sulcal forms, the depth information is implicitly included, so it is not solely based on surface morphology.

The main difference between such deduced variability model and the development-related model such as the sulcal roots model is that this approach aims at summarizing the *outcome* of the gyrification dynamics, not at deducing the *origins* of sulcal formation. Combining these two types of models leads to important and interesting insights concerning the folding dynamics. Such variability model is deduced here not as a definitive hypothesis, but as an example to illustrate the possibilities being opened using systematic analysis of the variability of 3D folding patterns, where a large number of subjects can be analyzed simultaneously and automatically. Such model can be (and is expected to be) modified based on the number and nature of the input data, as well as the algorithms used for analysis.

5.5 The region of the cingulate sulcus

Switching from the lateral to the medial surface, the Cingulate region is analyzed, which includes the cingulate sulcus (or callosomarginal fissure) and the smaller sulci surrounding it: the paracingulate sulcus, the intralimbic (or intracingulate) sulcus, the superior and inferior rostral sulcus. The focus of many anatomical and functional studies is on the cingulate and paracingulate sulci; here the whole region is studied, in the hope of finding not only the patterns concerning the cingulate and paracingulate structure alone, but also how the surrounding smaller sulci change with respect to the cingulate sulcus.

The cingulate sulcus is a primary sulcus that is present in both hemispheres in normal subjects. It appears around 16-19 weeks of gestation (Hori, 2006), together with the parieto-occipital and the calcarine sulcus, before the appearance of the central sulcus. Historically, the study of the sulcal pattern of the cingulate sulcus (CinS) is carried out by the early anatomists such as Eberstaller (Eberstaller, 1890). The caudal end of the CinS (the marginal ramus) is located just behind the medial portion of the central sulcus (**Fig 5.12**). The CinS follows the corpus callosum; at the rostral end the fusion with the superior rostral sulcus (SRS) is possible. This fusion with the SRS could occur rostral or

ventral to the cingulate sulcus (Paus, 1996). Interruption is one of the important features of the CinS. 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 (Ono et al., 1990). One or two branches could be extended from the CinS, a caudal branch right after the marginal ramus and a rostral branch. Tertiary sulci can also be observed: the paracingulate sulcus (PCS) and the intralimbic sulcus. The PCS occurs in 30-60% of the individuals, while the occurrence of the intralimbic sulcus is very rare (6% on the right side and 4% on the left side) (Paus, 1996).

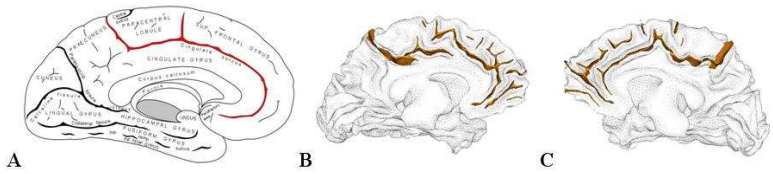


Fig 5.12 The cingulate sulcus and the surrounding functional organizations

A: the traditional nomenclature of the cingulate region (Grey's Anatomy) **B/C**: examples of the regional sulci used in this analysis (the combination of cingulate, paracingulate, rostral and intralimbic sulcus)

In terms of the left-right asymmetry, the left CinS is found to have interruptions less frequently than the right CinS. The left CinS has fewer branches as well. Concerning the paracingulate sulcus, it is found to be present more frequently in the left hemisphere in the work of Paus (Paus, 1996). This asymmetry is also presented in Weinberg (1905), but it is not observed in Ono. The cases of prominent and absent PCS are found to be more frequent in females (Paus, 1996).

The more developed PCS in the left hemisphere is hypothesized to be related to language lateralization, the left paracingulate cortex is activated in fMRI studies of word generation in humans (Paus, 1996). More recent study of the anterior cingulate and the paracingulate region shows that the PCS morphology is related to executive functions (Fornito, 2004). It is found that a leftward asymmetry, compared with rightward asymmetric patterns, is related to both verbal and spatial task engaging executive

cognitive processes. It is thus possible that the leftward PCS asymmetry represents an efficient configuration for executive cognitive processes. This leftward asymmetry is less frequent in neuropsychiatric populations (Fornito et al., 2008) and early-onset male with schizophrenia (Provost and al, 2003). These interesting studies revealed that the anatomical variations of the cingulate region are related to functional variations. The accurate mapping of function to this region is likely dependent on the accurate description of the underlying anatomical variability.

Fig 5.12 illustrates the traditional nomenclature of the region, two examples of the definition of the sulci included in this analysis is given. Refer to **Fig 5.13** for the summary of the shape analysis at the cingulate region. Using maximum and minimum distance measures, different SPAM images are obtained. Despite local differences, the overall shape evolution is similar. At one extreme, the paracingulate structure is much more prominent than at the other extreme. Moreover, as the paracingulate structure becomes more prominent, the anterior and rostral region of the cingulate is becoming heavier as well, this is more evident regarding the maximum distance (**Fig 5.13 B, E**). As the paracingulate structure becomes heavier, the anterior and rostral cingulate structure is pushed more towards the corpus callosum as well.

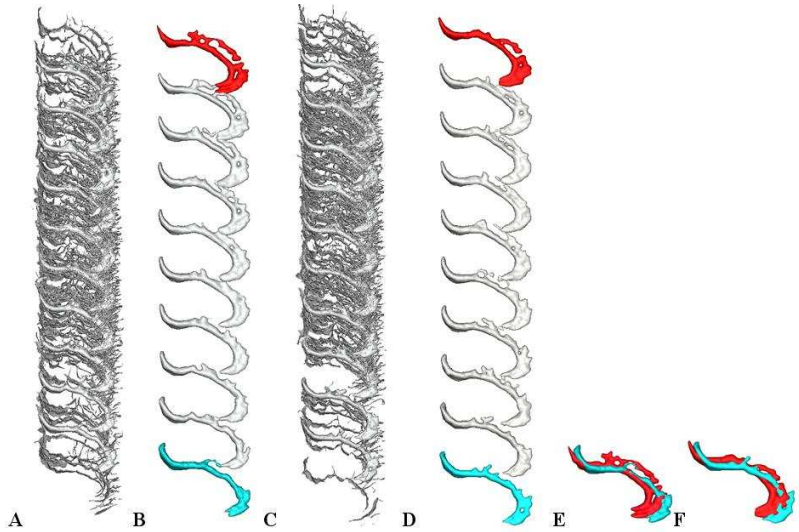


Fig 5.13 The Isomap of the cingulate region

A/C: The SPAM of the cingulate region using maximum/minimum distance. The sulci are superimposed on the SPAM, according to their relative positions. Notice that at the two extremities there are much less sulci compared with the centre. **B/D**: The SPAM of the cingulate region using maximum/minimum distance. The forms at the two extremities are coloured blue and red respectively. **E/F**: The two extremities of B/D are superimposed for easier comparison.

For a better understanding of the actual sulcal shapes at various positions on the Isomap axis, refer to **Fig 5.14**. From the top to the bottom of the axis, the evolution from more to less prominent paracingulate structure can be confirmed. The correlation between the more prominent paracingulate structure and the more prominent anterior and rostral structure can be observed as well. The change in number and position of interruptions of the cingulate sulcus cannot be reliably observed here. The main trend captured here appears to be the relative "heaviness" between the cingulate and paracingulate sulci. When the paracingulate is very prominent, the cingulate sulcus becomes weaker. When the paracingulate sulcus is missing or consisting of small vertical pieces, the cingulate sulcus becomes heavier. As the paracingulate structure gets more prominent (regardless of orientation and continuity), the presence of the superior and the inferior rostral sulcus becomes more likely, the cingulate sulcus or the rostral sulcus appear to be more advanced towards the parolfactory area as well.

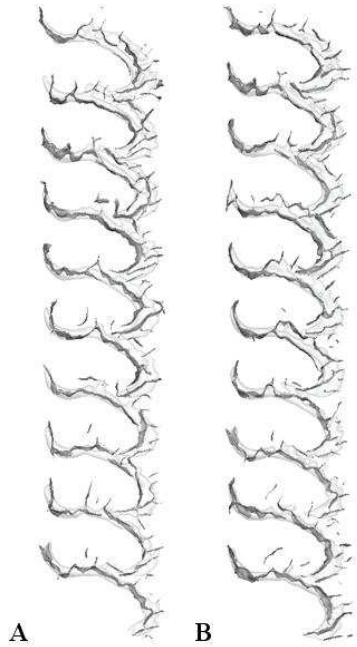


Fig 5.14 The SPAM of the cingulate region with the corresponding sulcal forms

A/B: The SPAM using maximum/minimum distance is plot as a transparent background, superimposed on top of each SPAM of sulcus is the real sulcus with its coordinate the closest to the SPAM coordinate. The form of the real sulcus thus provides more detailed information of the sulcal form at each coordinates.

The clustering analysis using PCBB method described in Chapter Two is used. Three clusters are found using the maximum distance measure (**Fig 5.15**). Notice that the clusters span the Isomap axis. The black cluster has a heavier configuration on the anterior region; overall its paracingulate structure is more prominent as well.

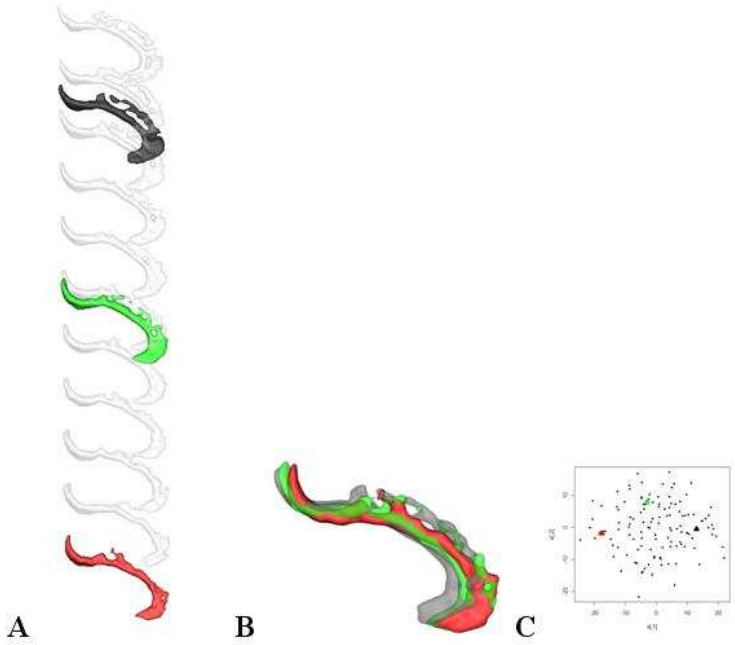


Fig 5.15 The clusters of the cingulate region

The three clusters found applying PCBB clustering algorithm to maximum distance. **A**: The SPAM image is plot as a transparent background, the three clusters found are coloured orange, purple and blue respectively. **B**: The three clusters found are superimposed to facilitate comparison among the shape of the three clusters **C**: the location of the three clusters related to the whole dataset distribution

Based on the SPAM analysis and the cluster analysis, a variability model can be deduced emphasizing the relative position and heaviness of the cingulate and paracingulate sulcus

(**Fig 5.16**). While the paracingulate region is prominent (towards the top of the axis), the cingulate region is very weak; this characteristic is reversed towards the bottom of the axis. When the paracingulate region becomes prominent, the paracingulate sulcus tends to become continuous. Some examples of the sulci of the cingulate region superimposed on brain surfaces are shown at the right of **Fig 5.16**. The detail regarding the interruption and orientation of the paracingulate structure cannot be reached in the current analysis. One possible reason is the limitation in Isomap dimension. It should be emphasized, however, that the current analysis takes into account the whole 3D shape of the sulci, including sulcal depth, while the interruption pattern described by Ono is mainly based on surface anatomy. The interruptions are not necessarily linked to sulcal depth or "heaviness" of the folds, thus it is not surprising that the current analysis does not find variability patterns based on interruptions.

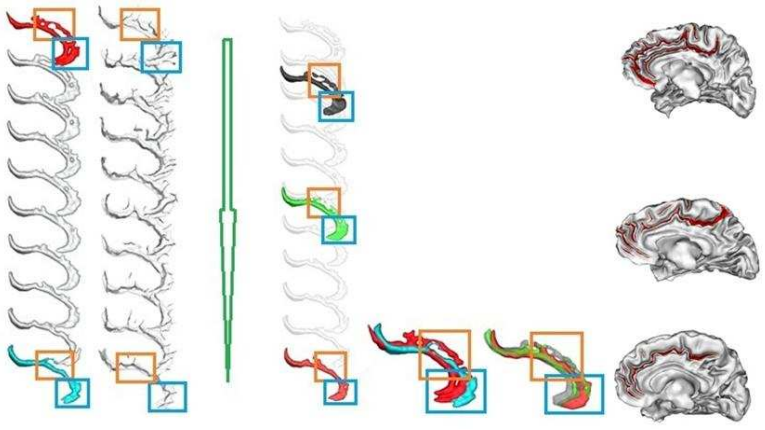


Fig 5.16 The variability model for the cingulate region

The green arrow indicates the direction of change. For the top to the bottom of the Isomap axis, the orange boxes emphasize the "heaviness" change in the paracingulate region; the blue boxes emphasize the changes in the anterior cingulate region in terms of both "heaviness" and curvature. At the right end three examples are given illustrating the relative change in "heaviness" of the paracingulate sulcus relative to the cingulate sulcus.

One possible region of variation in labelling is that of the intralimbic, the cingulate and the paracingulate sulci. The intralimbic (or intracingulate) sulcus is defined as a sulcus starting at the anterior part of the corpus callosum and joining the middle part of the

cingulate sulcus (Paus, 1996). Following the analysis above, when paracingulate is prominent, the cingulate becomes weaker. Refer to **Fig 5.16**, at the top extreme of the Isomap axis, the cingulate could become very weak and shallow. Does a weak cingulate structure corresponds to the intracingulate sulcus? In other words, when paracingulate is unusually prominent, should the cingulate be labelled as intracingulate and the paracingulate as the cingulate? Indeed, if the cingulate sulcus is the main sulcus separating the Brodmann areas 32 and 24, when a parallel configuration exists where the paracingulate is much more prominent than the cingulate, the paracingulate might be properly labelled as the cingulate (Regis, 1994). Our variability model of the cingulate region in this work assumes the dorsal sulcus in a parallel configuration to be the paracingulate, regardless the strength of the ventral counterpart. The paracingulate region, by definition, corresponds to the dorsal portion of the anterior cingulate (AC) cortex where the Brodmann areas 24b', 24c' and 32' are located, it is a relative expansion of the limbic and the paralimbic anterior AC cortex (Brodmann areas 24 and 32) (Fornito, 2004). Refer to **Fig 5.17 B and C** for the definition and the Brodmann Areas of the region. Refer to **Fig 5.17 B**, area 32 occupies the gyrus between the cingulate and the paracingulate sulcus. Areas 6aα and 6aβ are above the paracingulate sulcus. The area 6aβ corresponds to pre-SMA (Supplementary Motor Area), while the area 6aα corresponds to SMA proper (Crosson and al, 1999). A portion of medial area 8 may be present as well.

To investigate further this question regarding the labelling of the cingulate, paracingulate and the intracingulate sulcus, two subjects are chosen, one with prominent paracingulate structure, the other with prominent cingulate structure. The sulci of the two subjects are superimposed for easier comparison (**Fig 5.17A**). There is a possibility that the intracingulate structure corresponds to a weak cingulate, when no additional branch towards the corpus callosum can be observed. However, the automatic alignment used here is not necessarily optimal, so no reliable conclusions can be drawn. Further study related to the relative positions of the sulci in Talairach is needed to clarify this issue, combining with further architectonic and functional activation information.

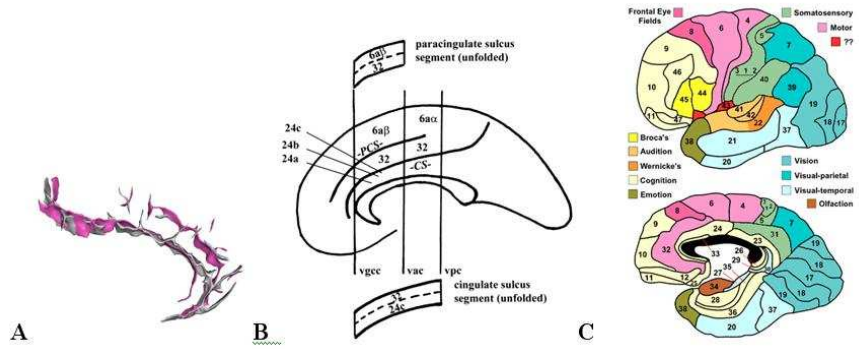


Fig 5.17 The cingulate regions of two subjects superimposed.

A: The cingulate regions of two subjects are chosen, the purple case corresponds to a prominent paracingulate structure with a weak cingulate structure; the grey case consists of a prominent cingulate structure with a weak paracingulate structure. **B**: The paracingulate region (Crosson and al, 1999) **C**: The Brodmann areas (Brodmann, 1909)

The configuration of the two hemispheres are compared (**Fig 5.18**), no clear asymmetry regarding the whole cingulate region can be observed. Clusters are found only on the left hemisphere, the results are illustrated in **Fig 5.19**. The general evolution in terms of the relative heaviness of the cingulate and the paracingulate regions can be observed as well among the clusters. The fact that clusters are found only on the left hemisphere but not on the right infers a difference in sulcal form distribution among the two hemispheres. The hemispheric asymmetry is thus likely to be related to the variation in sulcal form density distribution, not as simple as a shift in form, as observed for the region of the central sulcus.

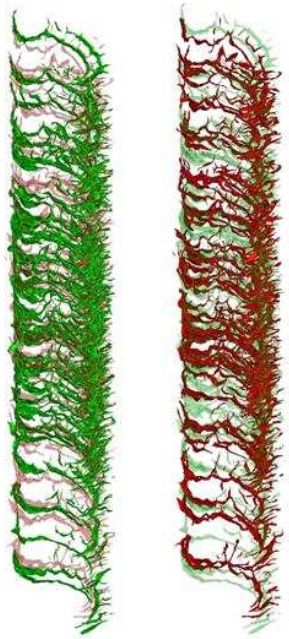


Fig 5.18 The cingulate region by hemisphere

The left hemisphere is plot in green, and the right hemisphere plot in red. The opposite hemisphere is superimposed in transparent for easier comparison.

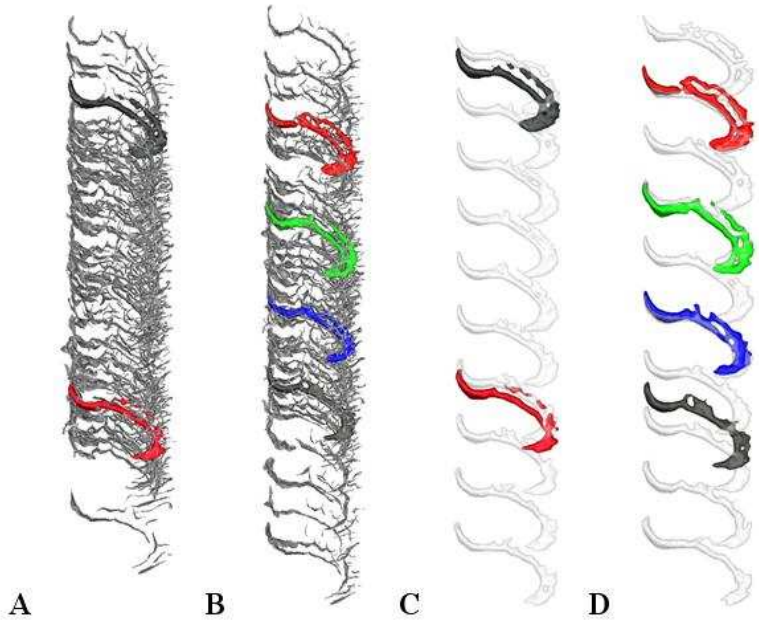


Fig 5.19 Clusters found on the left hemisphere

A/B: clusters found on the left hemisphere using the minimum/maximum distance; **C/D**: the SPAM of the clusters corresponding to A/B.

5.6 The region of the superior temporal sulcus

The region of the superior temporal sulcus (STS) is highly variable. It is appearing around 20-23 weeks of gestation, together with the central sulcus (Hori, 2006). Interruptions are one of the features described in (Ono et al., 1990). Compared to the left STS, the right STS is found to be more likely continuous (36% continuous on right STS vs. 28% continuous on left STS). The left STS is more likely to be broken into four segments (24%) compared to the right STS (0%). The anterior end of the STS is found to be extending lateral, medial, at or far posterior to the temporal pole. The left and right hemispheres are found to be distinctive in this characteristic as well. Various connections are found, such as the connection with the Sylvian fissure, the intraparietal sulcus, and the inferior temporal sulcus.

Some systematic anatomical analysis of STS was carried out (Ochiai et al., 2004) based on the sulcal root model. The sulcal roots of the STS are illustrated in **Fig 5.20 A**. Different sulcal roots (STs. Ant, STs.mid, STs. Post, STs horizontal, STs.ter.asc.ant and STs.ter.asc.post) corresponding to the "plis de passage" in this region are described. It is found that the generic model proposed using the "pli de passage" is consistent, and further asymmetry between the two hemispheres is observed (Ochiai et al., 2004).

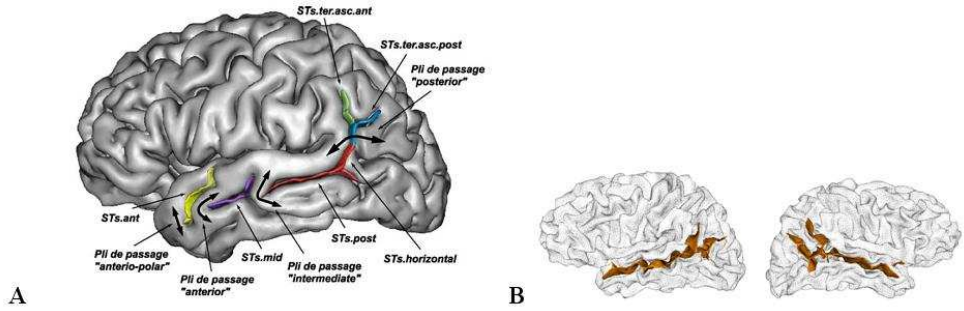


Fig 5.20 The sulcal roots and "plis de passage" of the STS complex

A: the pli de passage in the STS region, from T. Ochiai et al. (Ochiai et al., 2004) **B**: two examples of the STS used in the current analysis

The superior temporal sulcus is a region extensively studied for functional activations. The anterior STS is found to be related to voice and sentence processing, spatial awareness and biological motion processing; the middle part related to word comprehension; and the posterior part related to spatial and motion processing, face processing, social perception and the Theory of Mind activities (Hein and Knight, 2008). The functional activity in the STS region is linked to various systems such as the visual, auditory and the limbic systems, connections to and from the frontal lobe, the parietal lobe and the deeper structures such as the amygdale and striatum have been observed. Abnormalities in this region have been linked to pathologies such as autism (Redcay, 2008; Zilbovicius and al, 2006). Anatomical studies linking the form of the STS to symptoms has also been demonstrated as discussed in Chapter One (Plaze et al., 2009), where the shape of the STS is linked to the nature of auditory hallucination in schizophrenia.

Refer to **Fig 5.20 B** for examples of STS analyzed, and refer to **Fig 5.21** for the Isomap analysis of the STS region. The prominent feature found by Isomap is a gradual opening of the posterior "fork" of the STS. At the top of the Isomap axis, the posterior region of the STS (anterior and posterior terminal ascending sulci of STS) forms a wide "fork", either the anterior or the posterior terminal ascending portion can be broken off the main STS branch. At the bottom of the axis, these two pieces are much closer to each other. When observing the two extremities of the SPAM superimposed on each other, another interesting characteristic can be observed at the anterior end: as the "fork" closes, the external (surface) of the anterior STS is rotated dorsally compared to the bottom of the sulci. This trend observed can be seen more clearly on **Fig 5.21 E, F, G and H**, where the corresponding sulci of each SPAM coordinate along the Isomap axis are plot. Notice that from the top to the bottom of the axis, the posterior opening is diminishing, while the anterior portion is also getting longer and heavier, with the external part of the sulci rotated more dorsally. The change of the angle of the anterior end is more easily observed on the SPAM based on minimum distance in this case (**Fig 5.21 F**).

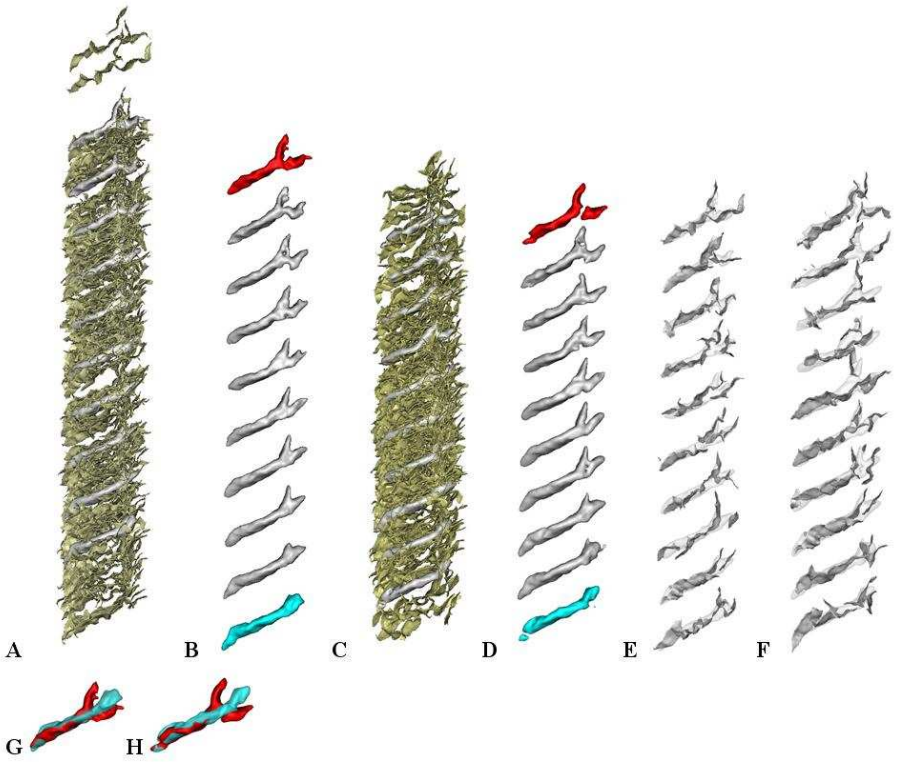


Fig 5.21 The isomap of the superior temporal sulcus

A/C: The isomap of the superior temporal sulcus (STS) using maximum/minimum distance. The sulci are superimposed on the SPAM, according to their relative positions. Notice that at the two extremities there are much less sulci compare with the centre. **B/D**: The SPAM of the STS using maximum/minimum distance. The forms at the two extremities are coloured blue and red respectively. **E/F**: The SPAM using maximum/minimum distance is plot as a transparent background; superimposed on top of each SPAM is the real sulcus with its coordinate the closest to the SPAM coordinate. The form of the real sulcus thus provides more detailed information of the sulcal form at each coordinates. **G/H**: The two extremities of **B/D** are superimposed for comparison.

Next, the clustering analysis is carried out. Refer to **Fig 5.22**, three clusters are found using both maximum and minimum distances. The three clusters follow the trend observed above, especially the degree of "opening" of the "fork" at the posterior of the STS. The anterior change of angle can be clearly observed in **Fig 5.22 F and J**. While the anterior angle changes, the anterior piece tends to be broken off the main branch as well, this interruption is likely due to the operculation of the anterior pli de passage (**Fig 5.20**). Compare with the anterior part of the STS, the posterior part is much more variable.

Consider for instance **Fig 5.22 G, H and K**. Even when the "opening" of the "fork" is similar in degree according to the SPAM images, the actual folds can be very different in configuration (consider the red cluster in **Fig 5.22 G and K**).

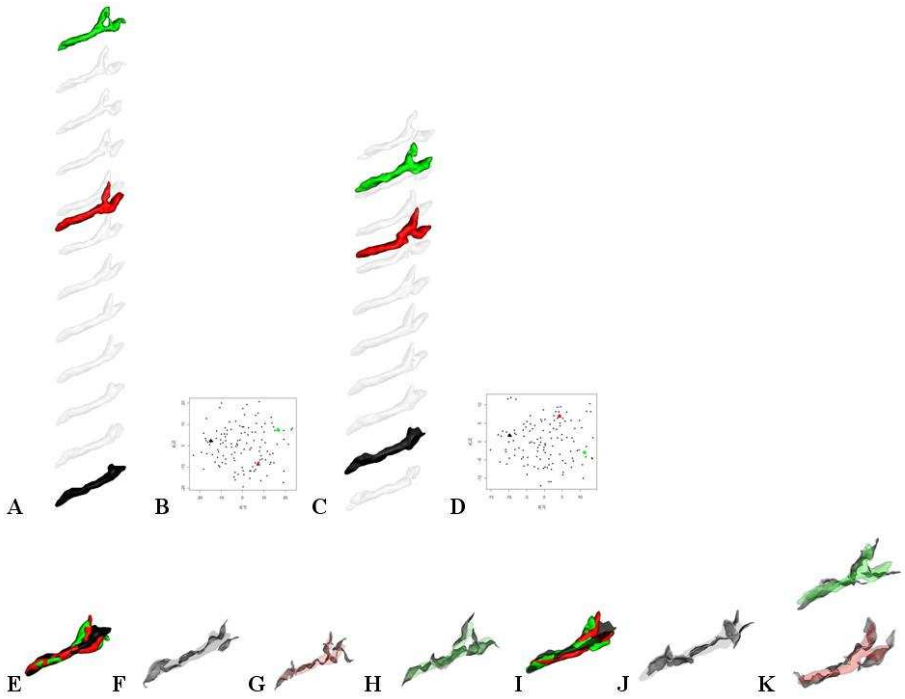


Fig 5.22 The clusters of the superior temporal sulcus

A/C: the clusters found using maximum/minimum distance **B/D**: the location of the three clusters related to the whole dataset distribution using the maximum/minimum distance **E/I**: the clusters of **A/C** superimposed together for easier comparison **F,G,H**: the three clusters of **A** with the real sulcal shape superimposed **J,K**: the three clusters of **C** with the real sulcal shape superimposed

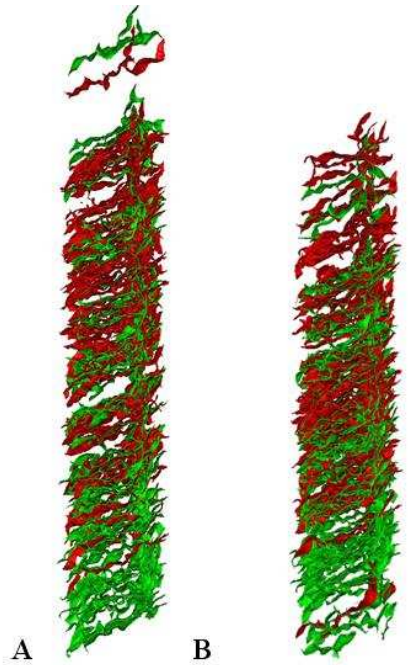


Fig 5.23 The left and right hemisphere of the superior temporal sulcus

A/B: the left (green) and right (red) hemisphere using maximum/minimum distance measure, the hemispheres are superimposed along the Isomap axis for easier comparison

Finally, the hemispheric asymmetry of the STS is studied (**Fig 5.23**). A strong asymmetry is observed, the left hemisphere is located more towards the bottom of the axis where the "fork" is closing; the right hemisphere is located more towards the top of the axis where the fork is opening wider. It is interesting to notice that the region between the upper fork of STS and the Sylvian fissure is corresponding to the Planum Temporale (PT), part of the Broadmann area 22 where the Wernicke's area is located. Hemispheric asymmetry has been observed in PT, the region is larger on the left hemisphere. Refer to Chapter One or **Fig 5.24B** for the corresponding Brodmann areas. Refer to **Fig 5.23**, the left hemisphere STS (green) shows a closing of the "fork", this also indicates an enlargement of the region directly above the upper "fork", where PT is located. This analysis also shows that when the PT region is enlarged on the left hemisphere, the anterior end is also elongated and rotated, following the Isomap analysis on shape (**Fig 5.21**). This anterior piece may correspond to the relative configuration of Brodmann area 38 related to 22, where BA 38 is now known to contain at least 7 subspaces according to cytoarchitectonic

studies. This area is among the earliest affected by Alzheimer's disease and one of the earliest involved at the start of temporal lobe seizures (Ding and al, 2009).

Based on the above analysis, a variability model of the STS can be proposed which is illustrated in **Fig 5.24A**. The model emphasizes on two regions of interest, at the anterior and the posterior end respectively. The variability at the anterior end is related to the angle of rotation and the length (the blue region), the variability at the posterior end is related to the opening of the ascending "fork", the anterior and posterior terminal ascending sulci of the STS (the orange region). The variability of the closing of the angle of the "fork" can also be interpreted as the possible enlargement of the region directly above the upper fork, approximately corresponding to the ventral border of Planum Temporale. To be more certain of this aspect of enlargement of the region another analysis including both the STS and the Sylvian fissure should be carried out. The "fork" itself could be related to BA 40, 39 and 37, the temporo-parieto-occipital area. BA 40 is involved in meaning and phonology of reading, and damage to BA 39 plays a role in semantic aphasia (Stoeckel and al, 2009). Assuming the significance of morphology in brain function, further studies linking such sulcal variability of STS to function should be very interesting.

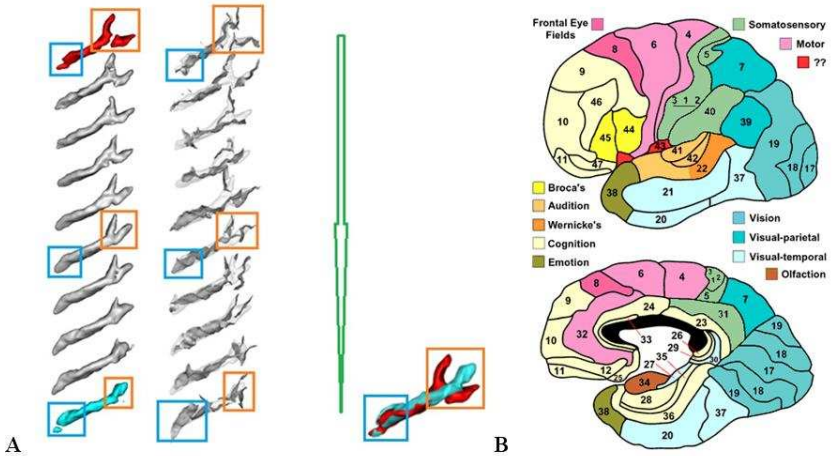


Fig 5.24 The variability model of the superior temporal sulcus

A: the variability model of STS **B**: The definition of the Brodmann areas are illustrated (detail refer to Chapter One)

5.7 The Broca's area

The Broca's area is explored; the sulci included in this analysis are the diagonal ramus, the ascendant ramus and the anterior ramus of the Sylvian fissure, the anterior inferior frontal sulcus and the inferior frontal sulcus (F.C.L.r.diag, F.C.L.r.asc, F.C.L.r.ant, S.F.inf.ant and S.F.inf). It is a region of wide interest because of its connection to language production. The definition of the region by a group of sulci is illustrated in **Fig 5.25**, with an example of the region.

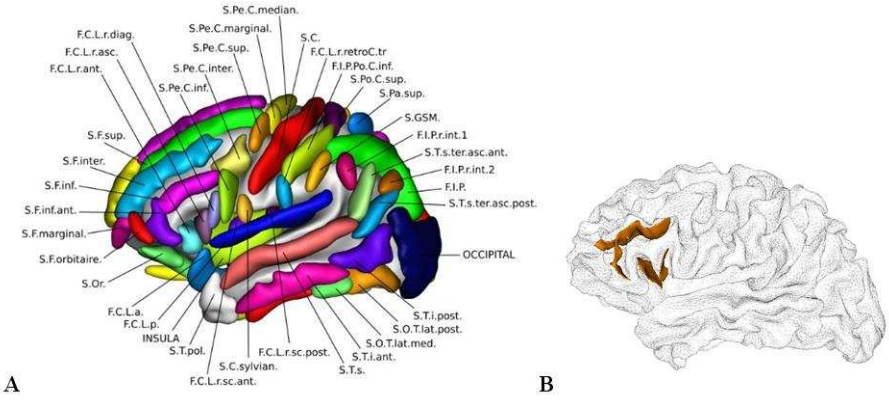


Fig 5.25 The definition of the sulci of the Broca's area

A: the definition of the sulci (as Fig 5.1); **B**: the Broca's area defined by the sulci. The definition of the Broca's area by a group of sulci: F.C.L.r.diag, F.C.L.r.asc, F.C.L.r.ant, S.F.inf and S.F.inf.ant.

The Broca's area is consisted of Brodmann Areas (BA) 44 and 45 (**Fig 5.24 B**). The F.C.L.r.asc generally separates BA 44 and 45. BA 44 is more involved in phonological and syntactic processing, also music perception. BA 45 is the triangular area (or pars triangularis) of the inferior frontal gyrus, it surrounds the F.C.L.r.ant and bounded caudally by the F.C.L.r.asc. The Broca's area is traditionally viewed as receiving afferents mainly from Wernicke's area through the arcuate fasciculus (part of the superior longitudinal fasciculus). Recent evidences suggest that the arcuate fasciculus connects the

posterior brain areas with the Broca's area through a relay-station in the premotor/motor areas (Bernal and Ardila, 2009).

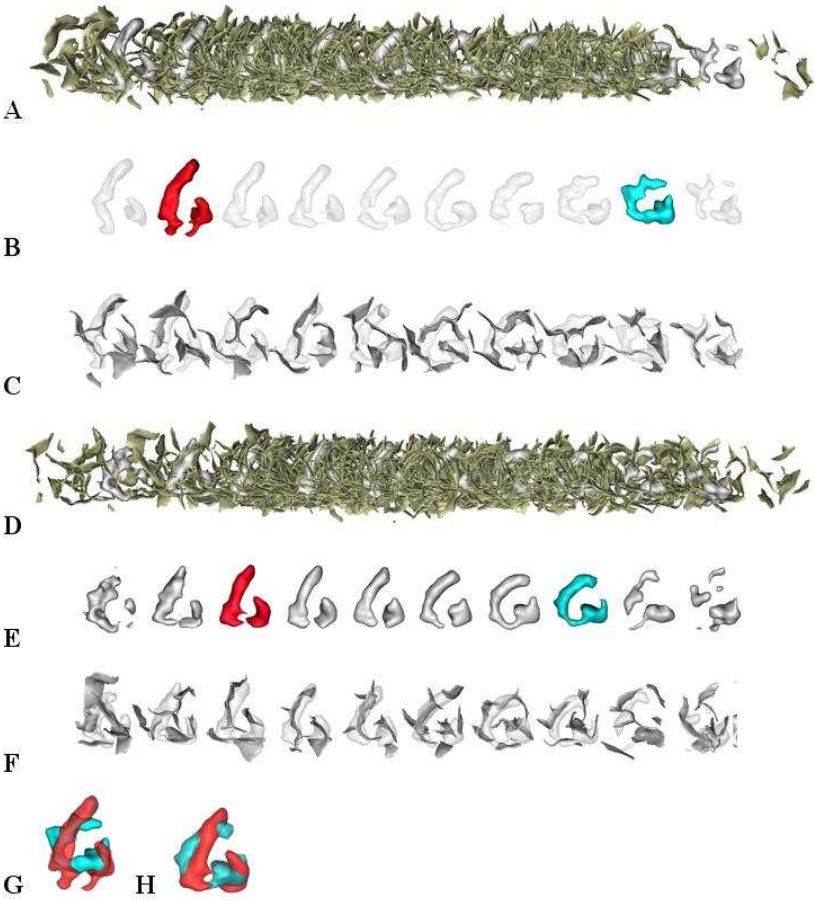


Fig 5.26 The analysis of the sulci around the region of Broca's area

A/D: The isomap of the Broca's area using maximum/minimum distance. The sulci are superimposed on the SPAM, according to their relative positions. Notice that at the two extremities there are much less sulci compare with the centre. **B/E**: The SPAM of the Broca's area using maximum/minimum distance. The forms at the two extremities are coloured blue and red respectively. **C/F**: The SPAM using maximum/minimum distance is plot as a transparent background; superimposed on top of each SPAM is the real sulcus with its coordinate the closest to the SPAM coordinate. The form of the real sulcus thus provides more detailed information of the sulcal form at each coordinates. **G/H**: The two extremities of **B/E** are superimposed for comparison.

The same methods are applied to study this region with results presented in **Fig 5.26**.

The complexity of this region increases compared to the regions analyzed before, the deduction of variability models becomes difficult. From the Isomaps, certain trends can nonetheless be observed. Refer to **Fig 5.26**, from the left to the right of the axis, the inferior frontal sulcus (S.F.inf) becomes less "heavy" or prominent and its orientation appears to be changing. Notice that the SPAM images show this trend, even though the image is not a precise summary of the shape. From the left to the right of the axis, as the relative "heaviness" and orientation of S.F.inf changes, the folds of the region become less "loose" and more compact as well. Further functional analysis with respect to the trend found would be very interesting.

The results of the clustering are shown in **Fig 5.27**. Two clusters are found. The red cluster shows a longer S.F.inf compare to the black cluster.

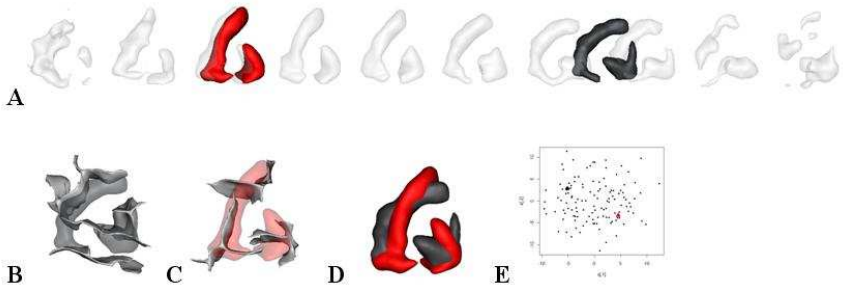


Fig 5.27 The clusters of the sulci of the Broca's area

A: The locations of the clusters found are indicated along the Isomap axis, minimum distance is used. **B**: the real sulcal shapes of the subject at the centre of the black cluster in A is plot, superimposed with the SPAM of the black cluster. **C**: the real sulcal shapes of the subject at the centre of the red cluster in A are plot, superimposed with the SPAM of the red cluster. **D**: the SPAM forms of the clusters in A are superimposed for easier comparison. **E**: The locations of the clusters in A are plot on the distribution of the whole dataset.

While these first results with Broca's area worth further exploration, we realize the possibility that because of the wide inter-individual variability, the behaviour of the ICP was more questionable than for the simpler groups of folds mentioned before. One of the

consequences is that the global alignment of the whole population toward the most neutral subject is not reliable. We need to explore further the consequences on the Isomap organization described here. Our future research program to overcome this difficulty is described in the last chapter.

5.8 The prefrontal region

The folds of the prefrontal region are analyzed. To aid in analysis, two sets of sulci are analyzed, called the smaller and larger superior frontal region. The smaller region contains the superior precentral sulcus, the marginal precentral sulcus, the superior frontal sulcus, the median frontal sulcus, and the transverse frontopolar sulcus. The larger region contains all the sulci of the smaller region, plus the intermediate frontal sulcus. The definitions of the regions are illustrated in **Fig 5.28**.

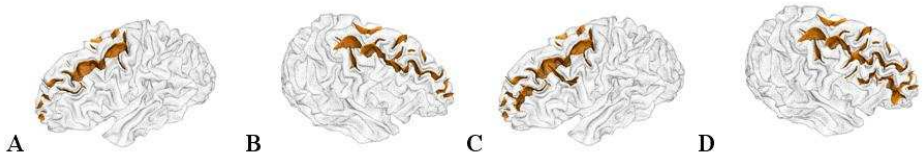


Fig 5.28 The definition of the smaller and larger superior frontal region

A/B: examples of the smaller superior frontal region; **C/D**: examples of the larger superior frontal region. The smaller region contains the superior precentral sulcus, the marginal precentral sulcus, the superior frontal sulcus, the median frontal sulcus, and the transverse frontopolar sulcus. The larger region contains all the sulci of the smaller region, plus the intermediate frontal sulcus.

5.8.1 The smaller superior frontal region

The results of the smaller region are shown in **Fig 5.29.** It can be observed that the complexity and variability of this region is higher compared to those analyzed above; finding a trend in sulcal shape becomes extremely difficult for human eye. A dominant trend that can still be observed is concerning the more or less “heavy” posterior portion,

around the superior precentral sulcus. This relative heaviness of the posterior region is also related to the total extent of sulcal complexity and heaviness of the region as a whole (**Fig 5.29 B and C**). In summary, from the top towards the bottom of the Isomap axis, the total heaviness and complexity increases, the superior precentral sulcus appears to elongate and is oriented more towards the frontal region, away from the central sulcus.

The clusters found in this region are illustrated in **Fig 5.30**. From the top to the bottom of the Isomap axis, the typical form of the clusters become heavier in general, the elongation of the superior precentral sulcus (the vertical posterior piece) can be observed as well. It should be noted that labelling error could exist regarding the superior and intermediate frontal sulcus, due to the complexity in this region. The introspection of each brain could clarify this issue.

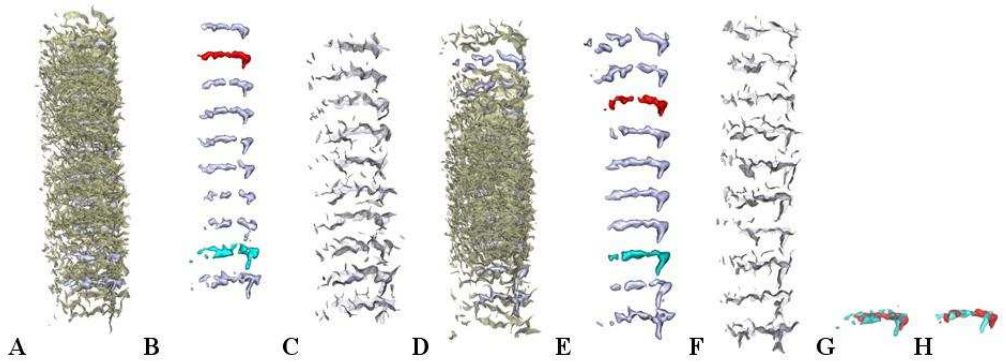


Fig 5.29 The analysis of the sulci around the region of the superior frontal area

A/D: The isomap of the superior frontal area using maximum/minimum distance. The sulci are superimposed on the SPAM, according to their relative positions. Notice that at the two extremities there are much less sulci compare with the centre. **B/E**: The SPAM of the superior frontal area using maximum/minimum distance. The forms at the two extremities are coloured blue and red respectively. **C/F**: The SPAM using maximum/minimum distance is plot as a transparent background; superimposed on top of each SPAM of sulcus is the real sulcus with its coordinate the closest to the SPAM coordinate. The form of the real sulcus thus provides more detailed information of the sulcal form at each coordinates. **G/H**: The two extremities of **B/E** are superimposed for comparison.

Fig 5.30 The clusters of the superior frontal area

Three clusters are found which are coloured green, yellow and purple respectively on the Isomap SPAM map. The more precise shape at the cluster centres are then plot, the subject is chosen as that the closest to the centre of the clusters found.

5.8.2 The larger superior frontal region

The results of the larger superior frontal region are illustrated in **Fig 5.31**. Using the maximum distance (**Fig 5.31 A, B, E and F**), an interesting trend can be observed. In contrast to the results obtained using the smaller superior frontal region, here from the top to the bottom of the Isomap axis, the overall complexity and "heaviness" appear to be decreasing, while the superior precentral sulcus appears to be somehow elongating. The reduction in complexity from the top to the bottom of the axis appears to be partly due to the less complex intermediate frontal sulcus.

Refer to **Fig 5.32** for the variability model of the superior frontal region. Here it is demonstrated that the definition of the region for shape analysis would influence the results. When the larger superior frontal region is studied, the dominant variability factor found is the "heaviness" of the intermediate frontal sulcus. When a smaller region is investigated, where the intermediate frontal sulcus is excluded, the "heaviness" of the superior precentral sulcus became the main factor for variability. In a sense, when we zoom into a region, the variability can be analyzed in more detail.

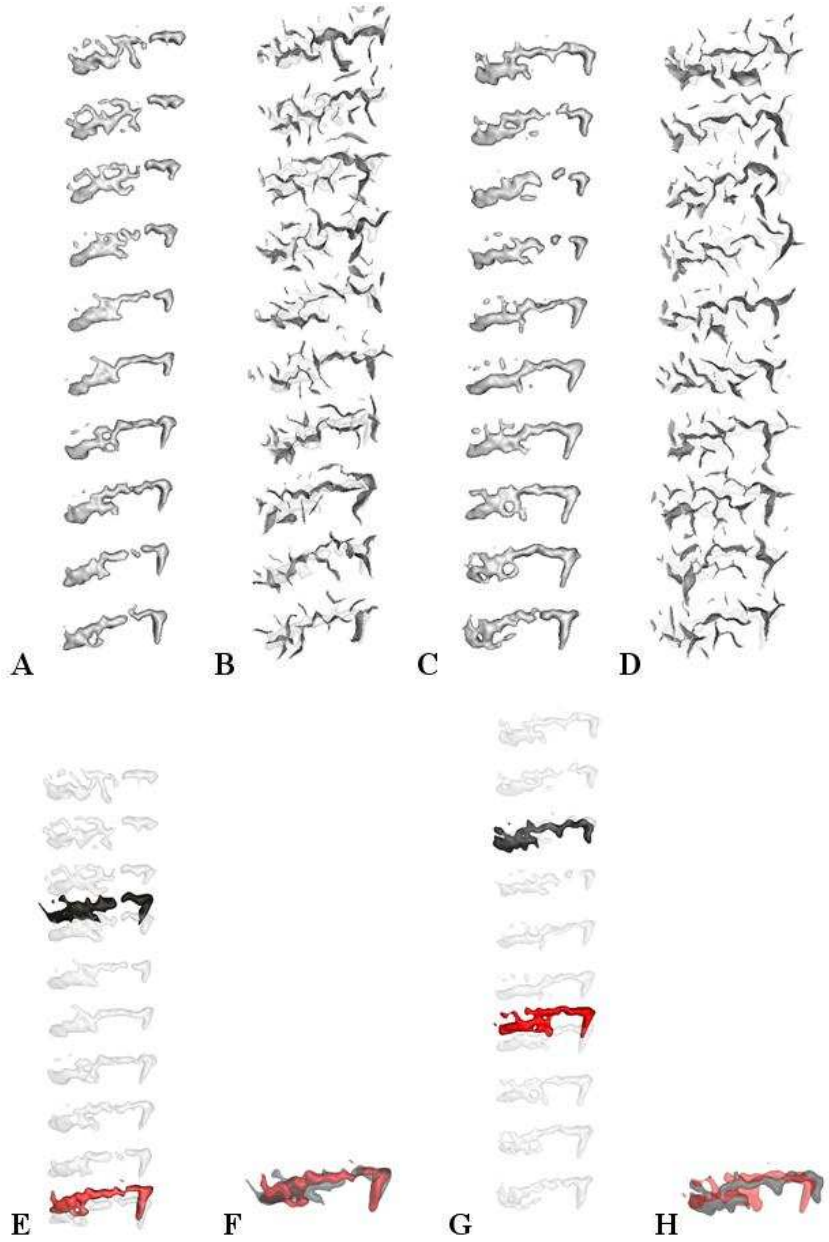
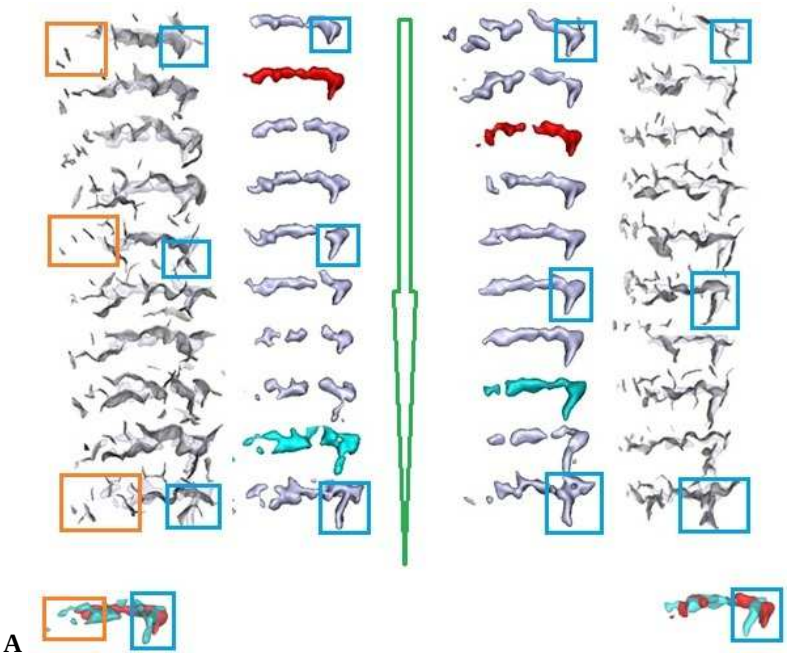
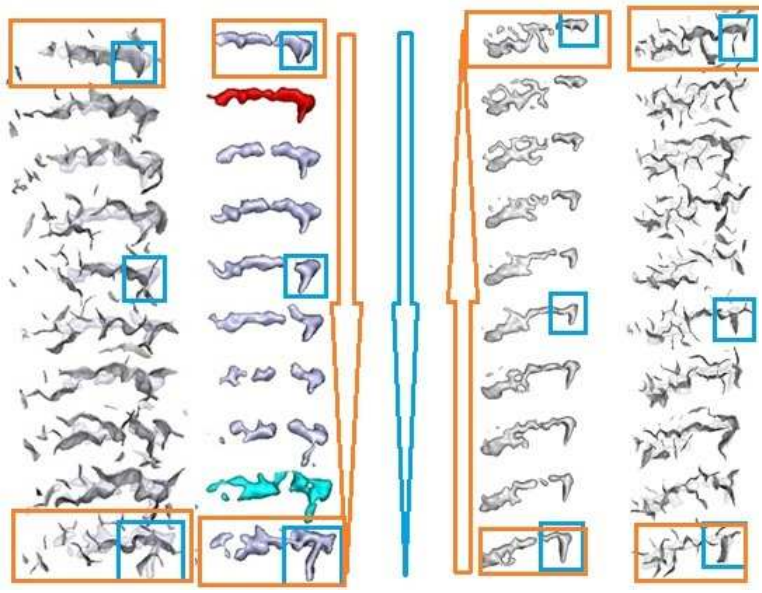


Fig 5.31 The analysis of the bigger superior fontal region

A/C: the isomap of the region using maximum/minimum distance **B/D**: the SPAM as in **A/C** with the real folds superimposed in the SPAM image **E/G**: the clusters found using the maximum/minimum distance **F/H**: the clusters of E/G plot together for easier comparison



A



B

Fig 5.32 Variability models for the superior frontal region

A/B: The variability model of the smaller/larger superior frontal region

5.9 The combinations: CS and cingulate

Finally, a combination of the sulci is analyzed to illustrate the potential of more comprehensive analysis. The advantage of such analysis is that relative positions of the sulci relatively far apart can be studied, the complexity of such analysis is beyond simple visual inspection. As an example, the combination of the cingulate region and the central sulcus is analyzed, the definition of the region is illustrated in **Fig 5.33**.

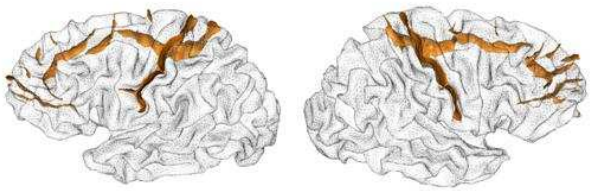


Fig 5.33 The combination of central sulcus and the cingulate region

The results are presented in **Fig 5.34** and **Fig 5.35**. Similar trend related to the heaviness of the paracingulate region is found here, from the left to the right of the Isomap axis, the paracingulate structure is more prominent. Regarding the central sulcus, it is interesting to notice that from the left to the right of the axis, the angle between the central sulcus and the cingulate sulcus is diminishing. In other words, at the left extreme of the axis, the central sulcus is more perpendicular to the cingulate sulcus; towards the right extreme, the central sulcus is more parallel to the cingulate sulcus. In one sense, when the paracingulate structure becomes heavier, the cingulate is pushed towards the corpus callosum, which explains why the angle between the central sulcus and the cingulate sulcus is changed. This trend can be observed also when the two extremities are superimposed together (**Fig 5.34 F**).

The clustering results are presented in **Fig 5.35 B and C**. The hemispheric asymmetry in the region is plot in **Fig 5.35 E**, no clear asymmetry can be observed.

The variability model is summarized in **Fig 5.36**. Two regions of interest are the orange one representing the change in angle of the central sulcus with respect to the cingulate sulcus; the blue region draws attention to the change in heaviness of the paracingulate structure.

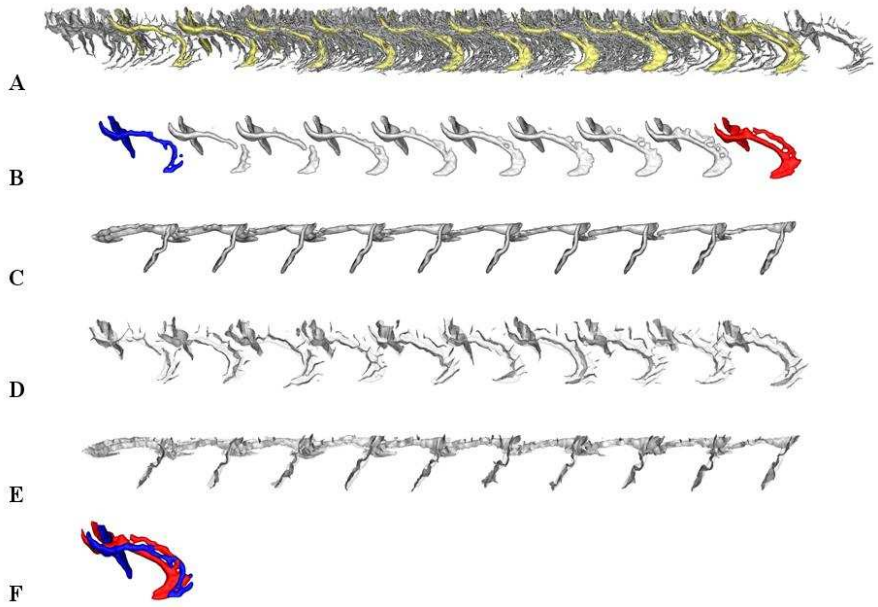


Fig 5.34 The Isomap analysis of the combination of the central sulcus and the cingulate region

A: the distribution of the sulci superimposed on the SPAM images, minimum distance is used here, similar results are obtained using maximum distance. **B**: the SPAM images with the two extremes highlighted **C**: the same as **B** with a change of point of view to facilitate the observation of the central sulcus **D/E**: same as **B/C**, with the real sulci superimposed on the SPAM images **F**: the two extremities of the Isomap axis as in **B** are superimposed to facilitate comparison

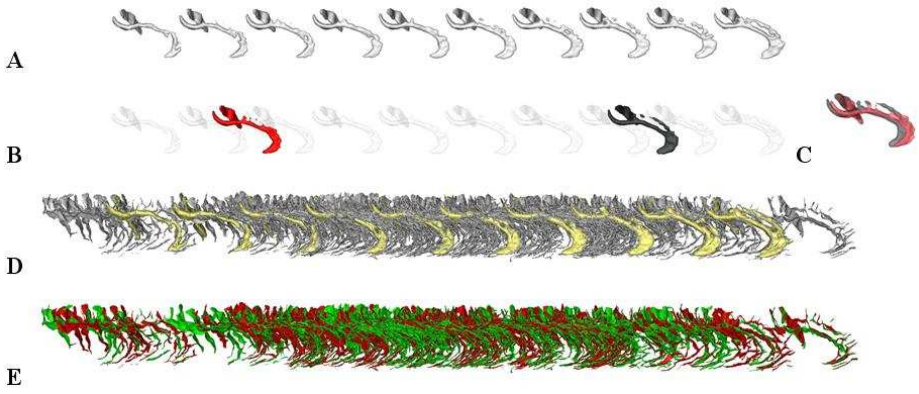


Fig 5.35 The clustering and asymmetry analysis of the combination of the central sulcus and the cingulate region

A: the SPAM images using maximum distance **B**: the clusters found using maximum distance, in black and red **C**: The clusters of **B** superimposed for comparison **D**: the real sulci plot on the SPAM images **E**: the sulci of the left (green) and right (red) hemispheres superimposed

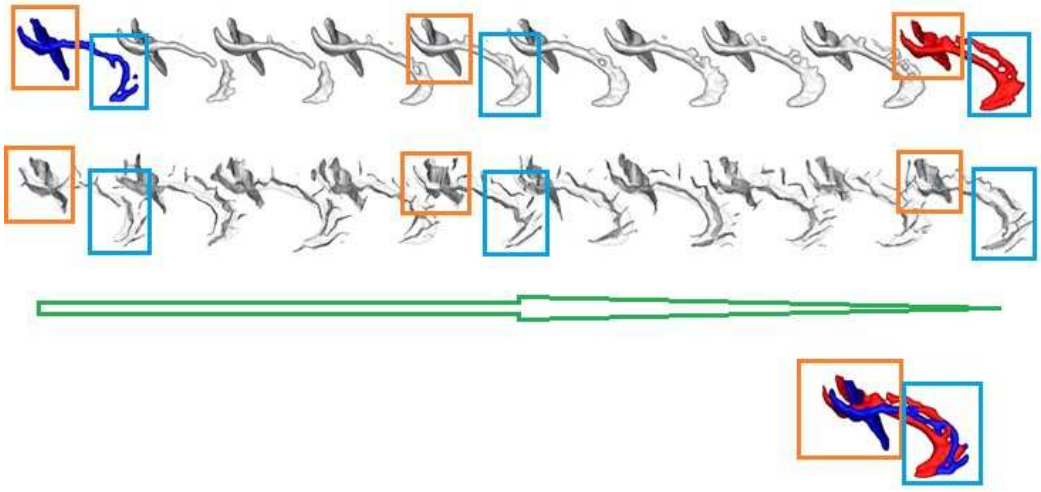


Fig 5.36 The variability model of the central sulcus with the cingulate region

Chapter Six: Discussion on Folding Analysis

6.1 Summary

In this chapter, the potentials of cortical pattern analysis, new data representation and interesting directions to further explore in the future is discussed. The example dictionary is presented in Chapter Five, using various approaches and methods detailed in Chapters Two and Three. The implications and usage of such dictionary and the potential applications are discussed in this chapter.

6.2 Introduction

Brain folding patterns are explored in this work, using algorithmic approaches. As discussed in Chapter One, the nature of this work is multi-disciplinary. Refer to the simple diagram in **Fig 6.1**, computer algorithms are used to solve specific problems in brain cortical folding. The results of such study can provide insight into the brain folding process; the results can also consequently provide insight into algorithm development.

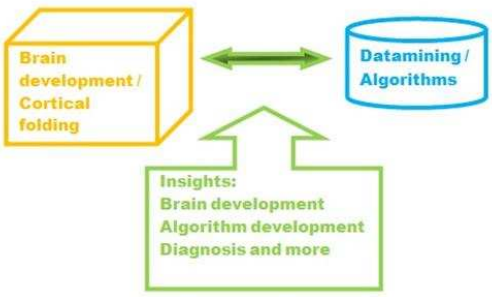


Fig 6.1 The diagram of the interaction among domains

Many factors play important roles in the analysis of cortical folding. In particular, the choice of the sulci or region of interest, the shape descriptor, the similarity measure, the clustering algorithm and the dataset used for analysis all impact the final results. These

five factors will be discussed in more detail below. The chapter concludes by the discussion of the use of such a dictionary and the future of cortical folding analysis in general.

6.3 The implications of sulcal analysis

The implication of the sulcal analysis is two-fold. On one hand, the pattern of cortical folding is studied systematically. Many of the knowledge gained through this process can be further used in other applications, such as automatic sulcal recognition or diagnosis of certain pathologies. On the other hand, the process of designing algorithms for cortical analysis provides useful insights to the design of algorithms in general.

6.3.1 Sulcal analysis and brain development

The results of cortical folding analysis can provide insights to the automatic naming system, for example, in the case of the inferior precentral sulcus (refer to **Fig 2.26** of Chapter Two). A pattern is consistently observed that consists of a shorter precentral sulcus, with the lower portion sometimes connected to another sulci (examples are the anterior subcentral sulcus, the diagonal ramus of the Sylvian Fissure). The automatic naming system may take into account this type of knowledge in terms of frequent patterns of the folds. This approach may lead to improved performance in automatic fold recognition. A more detailed analysis is presented below on the section of the analysis of the precentral gyrus.

More generally, the sulcal analysis can be used to verify existing hypothesis. For example, the sulcal roots theory (Regis et al., 2005) provides a systematic framework to study folding variability. Using the approach of brain folding analysis described in this work, such framework can be further explored. One possible direction for further study is the variability related to the sulcal roots model. At certain regions such as the pericentral region, the two-piece configuration of the central sulcus as the configuration of the sulcal

roots (the central sulcus with an interruption) can be rarely observed in adults (around 1%) (Ono et al., 1990). At other regions such as the prefrontal regions however, the interruptions are very frequent in adult brains, so that the sulcal roots pieces can be observed. Such analysis of the difference between adult folding patterns and the sulcal roots model provides information on region-specific variability. Why some regions resemble more the configuration of the sulcal roots with interruptions, while other folds become more typically continuous with few interruptions?

The folds formed earlier are more likely to be continuous than those formed later in development. The local resemblance to the sulcal roots, or the local degree of variability may correspond to the relative timing of the appearance of the sulcal roots. It would be interesting to explore further such region specific difference in variability, which may give insights into the timing and nature of brain development in general.

One possible use of the results of such systematic analysis is the deduction of possible sulcal variability models based on the analysis by Isomap, combined with the knowledge of the clusters. With the aid of the pattern dictionary, the very complex cortical folding can be analyzed by not only experienced neuroanatomists. Researchers and clinicians with less experience in sulcal anatomy can deduce hypothesis on sulcal patterns, and further test them. Some examples of the sulcal variability models are presented in Chapter Five.

Related to the timing of the appearance of the sulcal roots, the study of adult folding patterns can be extended to add the patterns of foetus or children at different age, to further understand the onset and nature of variability. The advantage of the method of cortical pattern analysis described in this work is that it is observer-independent and automated. It provides the possibility to detect patterns sometimes beyond the comprehension of human inspection alone, as in the case of the analysis of central sulcus presented in Chapter Four.

Such cortical pattern analysis can also provide possible landmark for normal versus abnormal brain development. When a certain pattern found rarely in normal populations becomes exceedingly frequent in certain pathologies, such an incidence could help with diagnosis. In a similar sense, if a gradual change in certain folding pattern can be observed, such as in the case of handedness and the form and length of the central sulcus discussed in Chapter Four, such information can be further used to monitor change during learning or treatment in the case of pathology.

6.3.2 Algorithmic development and knowledge discovery

Switching from neuroscience to computer science, it is demonstrated through this work that algorithmic development is not a stand-alone process. Based on the specific questions being investigated, the algorithms need to be changed to adapt to the question. In the first part of the work, we try to find dense clusters that would become the representative patterns of a given sulcus or a group of sulci. An algorithm is designed specifically for this purpose. It is later realized that there exist a varied degree of variability among the sulci. Certain sulci are very stable, so that no dense clusters can be found in these cases. To give a summary of the forms of such stable sulci, the algorithm needs to be changed. Eventually, the information regarding the sulci produced by different algorithmic approaches can be combined to give a more comprehensive description of the cortical folding patterns. This is an ongoing and dynamic process.

6.4 The main factors in the analysis of cortical folding

6.4.1 Folding groups and region of interest

Two different approaches towards sulcal analysis are experimented in this work. In the first approach, the biggest sulci are chosen (refer to Chapter Two for detail). These sulci also correspond to the deepest or mostly primary sulci in the sense that they appear earlier in brain development. These sulci are chosen because they are among the most

reliably labelled sulci and they can be found in the great majority of individuals, unlike the case of the smaller sulci. The combination of two or three of these sulci is also analyzed, following the reasoning that there would likely be fibre bundles connecting these sulci.

The major drawback of such an approach is that it does not emphasize the gyrus, which is the functional entity of the cortical convolution. Furthermore, the exhaustive combination of two or three of the major sulci is still doable, but this approach quickly leads to an explosion in computation when smaller sulci are concerned. More specific strategies need to be used. This leads to the second approach introduced in Chapter Three. Instead of choosing the most interesting sulci, the most interesting regions of the cortex are chosen. These regions can be as small as one single sulcus or as large as the whole brain. Most of the regions chosen are referring to the gyrus instead of the sulcus. The number of sulci is inferred by the region chosen.

These two approaches can find different patterns. There exist many further possibilities. For example, the sulci can be grouped by timing (from primary to tertiary), by depth, by surface area or by length. It would be interesting to verify if the depth corresponds to timing, this is expected to be generally true, but exceptions might exist. In the work on sulcal roots (Regis et al., 2005), the emphasis is on developmental timing, sulcal structure, variability and stability, the work of sulcal pits (Lohmann and von Cramon, 2000), on the other hand, is more based on sulcal depth. These two approaches yield similar maps. With the pattern analysis tools developed in this thesis work, the interconnections and differences of these two approaches can be further explored.

The regions of known functional correlations (such as the Broca's area and the Wernike's area) or known anatomical connectivity through fibre bundles can be explored together. In the cases when known link in specific pathologies exist, these regions can be combined as well. One such example is the cingulate region and orbital frontal region known to be linked in the obsessive compulsory disorders (Shim and al, 2009). Finally,

of course, when no constrain in computational power and time exists, a pure exploratory approach can be taken, when exhaustive combinations of regions can be explored in sequence without any prior knowledge.

Another approach worth investigation in the future is clustering analysis on unlabelled sulci. This approach would break the limitations of the traditional naming system. Interesting patterns concerning pieces of different sulci can be found, which may carry biological or neurological significance. Such an approach applied to large databases could reveal patterns beyond the reach of the anatomists using the classical nomenclature. This objective will have to deal with combinatorial explosion. In brainVISA, each brain, indeed, is made up of more than 500 elementary folds. These folds are different for each brain. A sulcus of the traditional nomenclature is usually made up by several of these elementary folds, some representing the stem of the sulcus and the others representing the branches. The number of elementary folds for a particular sulcus varies from one brain to another, because one given sulcus can be broken or interrupted in various ways. Therefore, inferring some reproducible patterns across a large set of folds will be a challenge.

Note that such analysis can be carried out on the raw data as well (Toews et al., 2010). In feature-based morphometry, folding patterns can be discovered using volumetric imagery.

6.4.2 Shape descriptor and Similarity definition

To describe the cortical folding, the first and more traditional approach is to use a shape descriptor. The simplest descriptors that can be used include the depth, length, the surface area and the number of connected components of the folds. The three-dimensional moments invariant to translation, rotation and scaling used in this work is originally developed to analyse simple shapes such as that of the thalamus of the brain (Poupon,

1999). The application of 3D moment invariant to the analysis of complex cortical folding shapes is validated in this study.

In the second part of the study, a new way to describe the shape is introduced. Instead of using sophisticated shape descriptors for each shape, the similarity among all the shapes in the population is calculated. This different approach goes beyond the description of each single shape, the group closeness in shape is used which may embed more comprehensive information concerning the individual related to the whole population. The results using the two approaches can both bring useful information to cortical folding analysis.

As discussed in Chapter One, other methods exist that can extract the top or the bottom of the cortical folding for further analysis. In our approach, we used the three-dimensional form of the folds. However, the top (external) or bottom of the sulci could embed important and different information. Our algorithmic approach can be used to carry out pattern analysis on the top or bottom of the sulci only.

In addition to the sulci line at the top or bottom of the fold, indeed, lines at different depth can be taken out for systematic pattern analysis. For example, when comparing the results of sulcal roots and sulcal pits models, an interesting question to ask is: is there an optimal depth which corresponds to a maximal fit between the two models?

Other than depth, another important factor in cortical folding is the sulcal direction. In the sulcal roots model, it is further proposed that the folding follows a grid system (Regis et al., 2005). The sulcal direction related to sulcal patterns can be further explored. In the study of folding orientation, the simplest approach is to use the most stable folds (for example the cingulate and central sulcus) as reference directions for the calculation of a direction index. This index could be calculated globally, locally, or in a more region specific sense (Clouchoux et al., 2010).

As discussed in Chapter One, the gyrification index (GI) is useful in the study of cortical folding and its relation to certain pathologies. In a similar sense, the extent of folding can be measured as to which extent a certain sulcus is stretched flat or being distorted or "wrinkled". Two folds with the same surface area may be more or less smooth. The Isomap study introduced in Chapter Four hinted the importance of such a measure. The extent of such folding can be observed in the Isomap analysis in two dimensions presented in Chapter Four (**Fig 4.9**). A further index dedicating the extent of folding can be added to the pattern description. The higher dimensions of Isomap can be systematically studied as well, to reveal possible additional information.

For patterns concerning more than one sulcus, the surface area of the gyri in between can be calculated. This adds another parameter to the pattern definition. Systematic comparison of asymmetry of folding patterns between the two hemispheres can be carried out. Other possible systematic comparisons include the analysis on gender and age. Such comparison can be expanded when other information are available.

While moment invariants discussed in Chapter Two turned out to be a good description of the complex shapes made up by the folds, they cannot describe certain details like branches that only weakly contribute to the coordinate moments. Therefore, in the future, a complementary approach based on non linear registration of images could be developed, where the main difference is removed, the focus can be on details such as the variability of the sulcal branches.

6.4.3 Clustering algorithms and approaches

In this section, the more general framework of clustering algorithm development is discussed. The algorithms developed and used are put into this general framework. The close link between cortical folding phenomenon and the "right" algorithm or algorithms to study such a phenomenon is further discussed.

6.4.3.1 Clustering algorithms

The definition of data clustering is itself vague; one of the definitions is as follows: “for a given set of data points and similarity measure, we regroup the data such that objects in the same clusters are similar, and objects in different clusters are distinct” (Jain et al., 1999). More intuitively, clustering is an exploratory process, a way to summarize the information in the dataset, so that an improved understanding can be achieved.

The major difficulties or challenges of clustering are mainly in three domains: how to find clusters with differing size, shape and densities; how to handle noise and outliers; and how to determine the number of clusters. These three issues are not independent, many times a given clustering algorithm excels in resolving one of the issues while fails in resolving the others. Indeed, there is a trade-off among the three issues. The “best” algorithm is ultimately based on an understanding of the specific real-life question(s) we are trying to solve by clustering analysis.

Clustering algorithms can be roughly put into six different categories: centre-based, density-based, grid-based, graph-based, search-based, and model-based (Gan et al., 2007). This is a conceptual categorization; a given clustering algorithm can belong to more than one category. These approaches are discussed very briefly below, the advantage and disadvantage based on the three issues mentioned above are discussed.

In the centre-based approach, the goal is to find clusters each represented by its centre. Examples of such algorithms are K-means and Expectation Maximization (EM) as a generalization of the K-means algorithm (Duda et al., 2000). This type of algorithms can find clusters of variable sizes, but cannot deal with arbitrary shape clusters easily. The PCBB algorithm can be put into this category, a cluster is represented by its centre.

In the density-based approach, the goal is to find dense regions. By definition, it is difficult for such an approach to find clusters of different densities. However, this approach can find clusters of arbitrary shapes, it is usually robust and the number of clusters is automatically decided by defining the dense regions. The algorithm DBSCAN belongs to this category, where the density of a point is obtained by counting the number of points in a region of specified radius (Gan et al., 2007). BRIDGE, an algorithm combining DBSCAN with K-means is another example (Gan et al., 2007). The PCBB algorithm designed in this work can be put in this category as well since the goal is to find local dense regions.

The grid-based approach is related to the density-based approach. A grid structure is created to partition the data into finite number of cells for further density-based clustering. Such an approach allows reduction in computational complexity. The advantages and disadvantages are similar to that of the density-based approach, the number of cluster is determined automatically, the clusters can be of arbitrary shapes, but the density of the clusters cannot be different.

In the graph-based approach, graph or hyper-graph is constructed; the data points are the graph nodes linked in a specific manner. The goal of the clustering algorithm becomes that of graph partitioning. The algorithm Chameleon can be grouped into this category where the K-nearest neighbour graph is generated and partitioned, and then a hierarchical clustering schema is used to combine sub-clusters (Gan et al., 2007). Compared to the density-based approaches, the definition of cluster can be more comprehensive; examples include clusters as dense regions, or homogenous regions. The definition of a cluster and how it will affect the clustering results will be discussed further below.

In the search-based approach, clustering is considered as an optimization problem. While algorithms such as k-means or fuzzy k-means can find local optimums, algorithms such as simulated annealing, Tabu search or genetic algorithms can go beyond local optimum (Michalewicz and Fogel, 2004).

Compared to the model-based approach, the other categories (centre-based, density-based, grid-based, graph-based and search-based) are considered as heuristics. In the model-based approach, data is considered to be generated by a finite mixture of underlying probability distributions. The models are used for clustering to optimize the fit between the data and the model. The clustering problem becomes the estimation of parameters of the assumed mixture model. In the end, both the model-based approach and heuristics face the same challenge: how to find a "good" clustering algorithm in the heuristics approach; or how to perform a "good" model selection in the probability framework.

In a more general sense, clustering algorithm development could take two routes: rule-based or model-based. Today, these two approaches are very often combined to solve specific problems. However, a somehow more philosophical separation exists between these two approaches. Does Nature follow mathematical models or very elementary rule-based computer programs such as that of the Turing Machine (a theoretical machine that can manipulate one input symbol on an infinite tape at a given time step) (Turing, 1948)? Simple models such as cellular automata can be used for such exploration of natural phenomenon (Wolfram, 2002). Schmidhuber asked the question: "is God a Mathematician or a Programmer?"(Schmidhuber, 1997). In the understanding of nature, both approaches reach interesting findings such as the morphogenesis of natural patterns. Some authors (for example (Wolfram, 2002)) argue, however, that a mathematical approach cannot explain the whole complexity of natural systems while rule-based computer heuristics (such as the genetic algorithms and the cellular automata approach) may have a better chance.

In our study, we tried to combine both approaches, although the emphasis is on heuristics. In the future, both approaches separately, or the combination of the two, can be used for more discoveries. Next, the approach of applying clustering algorithms to the study of cortical folding is discussed.

6.4.3.2 Clustering algorithms on cortical folding

The first and arguably the most important step of the clustering process is the definition of a cluster. By doing this, an assumption is made on the kind of clusters that can be found on a dataset. Indeed, the definition of the kind of cluster that we try to find can be very different. For example, the clusters can be defined as: compact groups with high mutual similarity within the group; or chained clusters where the group members can be just loosely linked. The clusters can be defined as regions of Gaussian distributions in noisy data, or can be defined as regions of homogeneous density in sparsely populated background. The definition of a cluster decides the desired outcome. The algorithm can consequently be designed or selected based on the definition of the cluster.

The definition of the type of cluster should be based on the specificity of the real data upon which the clustering analysis would be carried out. The data in this work, the cortical folding, would surely be different from the distribution of bacteria culture or gene expression data. Blindly applying clustering algorithms regardless the domain specificity leads to failure. So the first step of selecting a clustering algorithm in this work is the careful observation of the distribution patterns of different sulci. Such study is done as discussed in Chapter Two, it is found that the cortical folding distribution is very varied. For the less homogenous sulci such as the cingulate sulcus, the distribution resembles random homogeneous distribution with dense islands, for the less variable sulci such as the central sulcus, the distribution resembles random homogeneous distribution without any dense islands. In the case of very noisy datasets such as these, a particularly robust algorithm is needed. It is also likely more sensible to define clusters as the dense islands of arbitrary shape instead of regions of Gaussian distributions. The PCBB algorithm is consequently designed to find such dense regions in noisy data.

An interesting question to ask is: what does such data distribution pattern tell us about the biological nature of cortical folding? Indeed, distribution as mixture of Gaussians is

observed in many Biological phenomenons, such as the size of different bacteria species in a cell culture dish. On the other hand, many more complex systems such as geological systems exhibit non-Gaussian distributions. Why a mixture of cortical folds does not exhibit a mixture of Gaussian distribution pattern? In the 1952 paper of Turing (Turing, 1952), it is proposed that pattern might grow from an initially nearly homogeneous state due to instability. The result of small initial instability could result in the emergence of patterns as a consequence of the breakdown of symmetry and homogeneity. The observed final distribution pattern could be close to or very different from homogeneous.

Linking data distribution with genetics, a Gaussian distribution pattern may hint the existence of simple genetic control for expression, such as the case when one gene is involved in determining the size of a given bacteria population. When complex irregular distributions are observed, it is possible that there are more genetic or environmental factors involved that interact with each other in gene expression. In the case of cortical folding, a huge variability in shape and pattern can be observed. This variability is likely due to interactions of multiple genetic and environmental factors. In the Turing morphogen model (Turing, 1952), the reaction-diffusion of two factors, the activator and the inhibitor morphogens, can cause pattern formation. An interesting study exploring the cortical pattern formation using Turing morphogens is discussed in Chapter One (Lefevre and Mangin, 2010). In real life, a group of such morphogens are likely to be involved in sulcal pattern formation. In addition, such morphogens can be genetic or environmental, further adding complexity to the process.

It is possible that in certain pathologies where genetic abnormalities are known, the cortical folding pattern distribution would change from that of the normal population. The folding abnormality can be difficult to detect visually; the abnormal folding distribution of the dataset as a whole may be observed more easily. Such change in distribution can be reflected in the shift of locations of the dense regions. Another possible change is in the nature of distribution, for instance a shift from non-Gaussian to Gaussian pattern of distribution would hint a shift in the dynamics of folding.

To conclude, the definition of clusters should be determined by careful observation of the distribution of the cortical folding patterns. Different clustering algorithms can be used based on the kind of clusters we are trying to find. A combination of clustering algorithms can be used to give a more comprehensive description of the data.

6.4.4 ICP

We recently realized that because of the amount of variability observed with a group of sulcus like Broca's area, a better control of the quality of the ICP alignment was mandatory. For instance, combining the actual sulci used in **Fig. 5.26** with the complete brains, we realized a shift in orientation from the left to the right of the Isomap. Furthermore, some of the individual alignments with the template brain were spurious because either ICP yielded a local minimum or the two sulcus groups are too different to be aligned correctly. Future work will aim at overcoming these difficulties following alternative research directions:

- Adding a low variability sulcus like central sulcus to the group to impose the global orientation;
- Using local alignments to compute the local SPAM rather than the alignment to the template (we think that ICP alignment is more robust when dealing with similar shapes);
- Using more robust ICP algorithms;
- Controlling the matrix of pair wise alignments globally to introduce some regularization.

6.4.5 Datasets

The datasets used in part determines the patterns that would be found. Many factors are important in the selection and preparation of datasets, and would consequently determine the outcome. The size and the type of the datasets (patient versus normal, animal versus human etc) are some of the important factors. In the future, many more datasets should be used in pattern analysis to get a more comprehensive dictionary of patterns. The datasets can be mixed to increase the statistical power; however, caution should be taken before mixing to ensure that the datasets are compatible. When different populations are used for pattern analysis, the patterns obtained would be different. Another issue is the method to reliably compare the patterns or clusters obtained.

6.5 Future directions

In the future, different directions can be taken. For instance, other datasets could be used for the same type of analysis. It is important to confirm the results on other datasets, the nature of the input may change the results as well. In **Fig 6.2**, another dataset containing 486 brains is analyzed. This database was processed by brainVISA first, and then the branches of the central sulci were unselected manually. The resulting Isomap confirms our results, the hand knob moves upwards from the left to the right of the axis. With this higher number of brains, the Isomap organization looks almost perfect (a neighbourhood of 10 brains is used to define the nearest neighbour graph). This database is part of a pedigree study (collaboration with P. Kochunov) that will allow addressing the genetics underlying the hand knob feature. Preliminary results tend to show that the hand knob location is heritable.

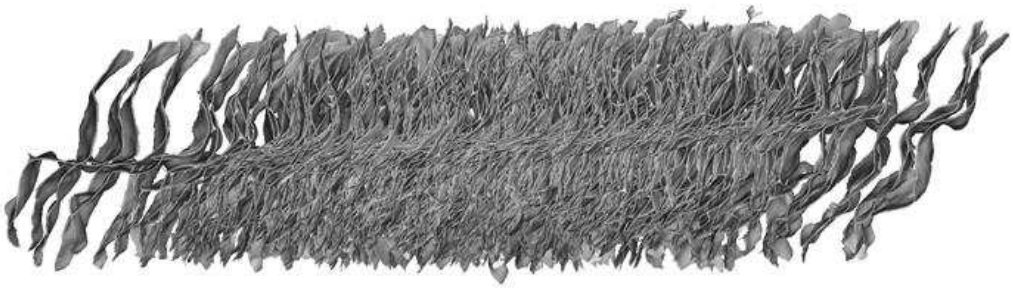


Fig 6.2 The analysis of the central sulcus of a dataset containing 972 adult sulci (courtesy P. Kochunov)

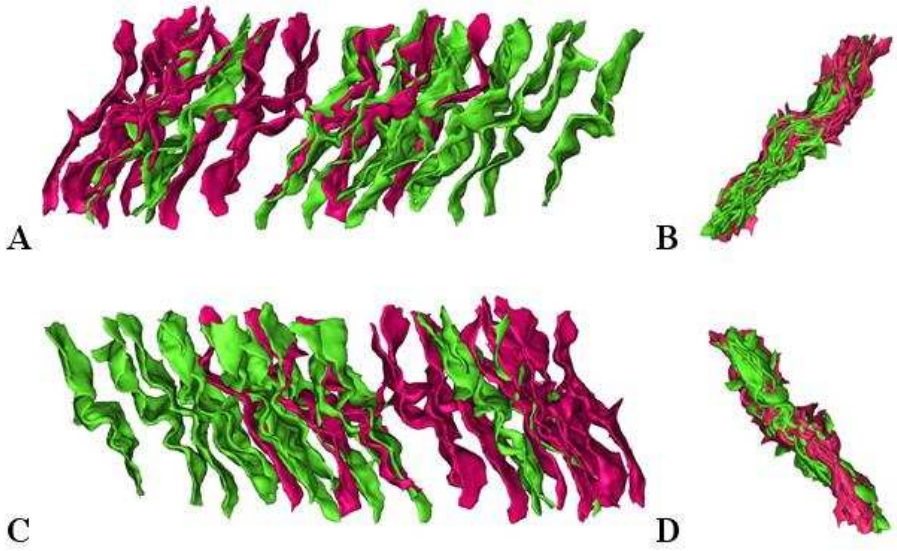


Fig 6.3 The isomap of the central sulcus of a dataset of infants (courtesy F. Leroy, J. Dubois, L. Hertz-Pannier and G. Dehaene-Lambertz)

A: the left (magenta) and right (green) hemispheres superimposed **B**: all the sulci plot together **C/D**: similar to A/B, the point of view is changed to the bottom instead of the surface of the brain

It is also interesting to study the sulcal forms during development. **Fig 6.3** illustrates the preliminary images of the analysis of the central sulcus of infants, aged from 1 to 6 months. The higher position of the hand knob can be observed (**Fig 6.3 B**) from the surface, but not from the bottom (**Fig 6.3 D**). This is likely due to the stability at the bottom of the sulci as discussed in Chapter Four and Chapter Five concerning the central

sulcus. The investigation could be extended to children of various ages and especially backward in time: recent advances provide similar central sulcus representations in highly premature babies (courtesy J. Dubois and P. Hüppi). Finally some comparative studies performed in collaborations with the groups of W. Hopkins and P. Kochunov lead to similar observations with chimps and other primates.

We also tested to which extent our method is robust enough to provide the same qualitative results without any manual correction. 250 brains from the Localizer database of Neurospin were processed by A. Moreno and P. Pinel, from the group of S. Dehaene. They applied the automatic sulcus recognition of brainVISA. We processed the resulting 500 central sulci, 10 of them being discarded by the outlier detection mechanism. The SPAM of the resulting one dimensional Isomap can be visualized below (**Fig. 6.4**) and fit perfectly the previous ones.

Ideally, such sulcal form analysis should be combined with functional data and fibre bundle data. **Fig 6.4** and **Fig 6.5** illustrate preliminary such results. The functional SPAM highlighted in **Fig.6.4** proves that the hand knob keeps its status of landmark for the hand motor area whatever its location along the central sulcus. The language activation extent seems correlated to the size of the second lower knob that we supposed associated with the language system. Further analyses are required to confirm this association at the individual level. Indeed, a bias could be introduced by differences in the quality of alignment from one side to the other side of the Isomap. More in depth work is needed, these images are only shown here to illustrate the many exciting directions that can be taken using the automatic sulcal pattern analysis methods developed in this work.

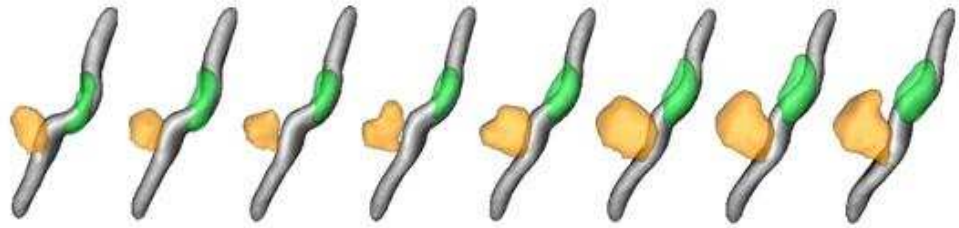


Fig 6.4 Functional activation SPAM superimposed on central sulcus SPAM along the Isomap axis of the central sulcus (250 brains, localizer protocol, courtesy P. Pinel and S. Dehaene).

The green SPAM is related to the right hand motor activation, the orange SPAM is related to silent reading involving part of the motor language system, maybe an area controlling the larynx. The functional SPAMs are obtained using the isomap coordinate of the left central sulcus as individual weight when averaging individual activations and the individual alignment used for the left central sulcus relative to the template sulcus. Therefore the functional SPAM should be interpreted as group studies for population with similar left central sulcus shape.

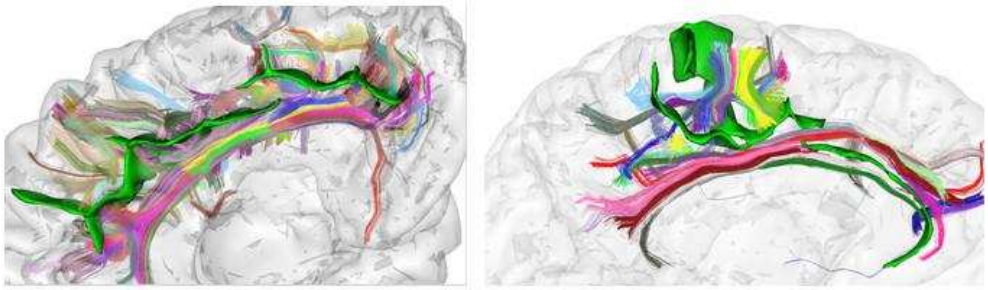


Fig 6.5 The fibre bundles surrounding the cingulate sulcus of two subjects

It can be seen that even though in both cases interruptions can be observed. A closer look at the underlying fibre connection reveals that the interruptions are likely due to different configurations of the underlying fibre architecture (courtesy P. Guevara and C. Poupon).

6.6 Conclusion

As illustrated in the sample dictionary (details refer to Chapter Five), description of the variability can be more systematic, and results on large datasets can be obtained automatically. Such analysis adds to the valuable work such as that of Ono (Ono et al., 1990), using modern imaging and computational techniques.

Based on the expanded pattern dictionary, hypothesis of folding mechanism can be further verified. It might inspire new hypothesis based on the expanded information. As our understanding of cortical folding variability gets more complete, more insights can be gained, models can be refined and expanded.

Genetic factors (when patterns are linked to genetic diseases, twins, families) in the folding process can be further tested. Landmark for development and change (in the study of plasticity, development and aging) can be established. Such pattern analysis may help with diagnosis and the monitoring of treatment. Finally, such cortical folding analysis can be combined with the information of fibre bundles beneath, functional data and information from behavioural experiments. Development can be systematically studied in longitudinal studies to reveal the timing of cortical pattern formations.

Indeed, the cortical folding patterns can be landmarks for the past (developmental abnormality, in-born traits), the present (benchmark for treatment and change) and the future (in early diagnosis).

References

Agrawal, R., Johannes, G., Dimitrios, G., Raghavan, P., 1998. Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications. Proc. of the ACM SIGMOD Int'l Conference on Management of Data.
Alkadhi, H., Kollias, S.S., 2004. Pli de passage fronto-parietal moyen of broca separates the motor homunculus. AJNR. American Journal of Neuroradiology 25, 809-812.
Amunts, K., Schlaug, G., Schleicher, A., Steinmetz, H., Dabringhaus, A., Roland, P.E., Zilles, K., 1996. Asymmetry in the human motor cortex and handedness. NeuroImage 4, 216-222.
Anatomica, N., 1983. Nomina Anatomica

Armstrong, E., Schleicher, A., Omran, H., Curtis, M., Zilles, K., 1995. The ontogeny of human gyrification. Cereb Cortex 5, 56-63.
Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry--the methods. NeuroImage 11, 805-821.
Barondes, S.H., Alberts, B.M., Andreasen, N.C., Bargmann, C., Benes, F., Goldman-Rakic, P., Gottesman, I., Heinemann, S.F., Jones, E.G., Kirschner, M., Lewis, D., Raff, M., Roses, A., Rubenstein, J., Snyder, S., Watson, S.J., Weinberger, D.R., Yolken, R.H., 1997. Workshop on schizophrenia. Proceedings of the National Academy of Sciences of the United States of America 94, 1612-1614.
Barron, D.H., 1950. An experimental analysis of some factors involved in the development of the fissure pattern of the cerebral cortex. Journal of Experimental Zoology 113, 553-581.
Bellman, R.E., 1961. Adaptive Control Processes. Princeton University Press, Princeton.
Bellugi, U., Lichtenberger, L., Mills, D., Galaburda, A., Korenberg, J.R., 1999. Bridging cognition, the brain and molecular genetics: evidence from Williams syndrome. Trends in Neurosciences 22, 197-207.
Bernal, B., Ardila, A., 2009. The role of the arcuate fasciculus in conduction aphasia. Brain 132, 2309-2316.
Berry, M., Linoff, G., 2000. Mastering Data Mining.
Besl, P.J., McKay, N.D., 1992. A Method for Registration of 3-D Shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14, 239-256.
Beyer, K., Goldstein, J., Ramakrishnan, R., Shaft, U., 1999. When Is "Nearest Neighbor" Meaningful? Int. Conf. on Database Theory, 217-235.
Boling, W., Olivier, A., Bittar, R.G., Reutens, D., 1999. Localization of hand motor activation in Broca's pli de passage moyen. J Neurosurg 91, 903-910.
Boling, W., Parsons, M., Kraszpulski, M., Cantrell, C., Puce, A., 2008. Whole-hand sensorimotor area: cortical stimulation localization and correlation with functional magnetic resonance imaging. J Neurosurg 108, 491-500.
Boling, W., Reutens, D.C., Olivier, A., 2002. Functional topography of the low postcentral area. J Neurosurg 97, 388-395.
Borg, I., Groenen, P.J.F., 2005. Modern Multidimensional Scaling. Springer.

Braitenberg, V., 2001. Brain size and number of neurons: an exercise in synthetic neuroanatomy. Journal of Computational Neuroscience 10, 71-77.
Broca, P., Pozzi, S., 1888. Mémoires sur le cerveau de l'homme et des primates. Reinwald, Paris.
Brodmann, 1909. Vergleichende Lokalisationslehre der Grosshirnrinde. Barth-Verlag, Leipzig.
Cachia, A., Paillere-Martinot, M.L., Galinowski, A., Januel, D., de Beaurepaire, R., Bellivier, F., Artiges, E., Andoh, J., Bartres-Faz, D., Duchesnay, E., Riviere, D., Plaze, M., Mangin, J.F., Martinot, J.L., 2008. Cortical folding abnormalities in schizophrenia patients with resistant auditory hallucinations. NeuroImage 39, 927-935.
Carroll, S.B., 2005. Evolution at two levels: on genes and form. PLoS Biol 3, e245.
Clarke, O.M., 1968. The Human Brain and Spinal Cord. A Historical Study Illustrated by Writings from Antiquity to the Twentieth Century.
Clemente, C.D.e., 1985. Anatomy of the Human Body by Henry Gray.
Clouchoux, C., Riviere, D., Mangin, J.F., Operto, G., Regis, J., Coulon, O., 2010. Model-driven parameterization of the cortical surface for localization and inter-subject matching. NeuroImage 50, 552-566.
Corballis, M.C., 2009. The evolution and genetics of cerebral asymmetry. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 364, 867-879.
Coren, S. (Ed.), 1990. Left-handedness: Behavioral Implications and Anomalies. North-Holland.
Cox, T.F., Cox, M.A.A., 2001. Multidimensional Scaling. Chapman & Hall.
Crawley, M., 2007. The R Book. John Wiley & Sons, Ltd.
Crosson, B., al, e., 1999. Activity in the Paracingulate and Cingulate sulci during word Generation: An fMRI Study of Functional Anatomy Cerebral Cortex 9, 307-316.
Csernansky, J.G., Gillespie, S.K., Dierker, D.L., Anticevic, A., Wang, L., Barch, D.M., Van Essen, D.C., 2008. Symmetric abnormalities in sulcal patterning in schizophrenia. NeuroImage 43, 440-446.
Cunningham, D.J., 1892. Contribution to the surface anatomy of the derebral hemispheres. Royal Irish Academy, Dublin.
Cykowski, M.D., Coulon, O., Kochunov, P.V., Amunts, K., Lancaster, J.L., Laird, A.R., Glahn, D.C., Fox, P.T., 2008. The central sulcus: an observer-independent characterization of sulcal landmarks and depth asymmetry. Cereb Cortex 18, 1999-2009.
Dejerine, J., 1895. Anatomie des centres nerveux (Tome I). Masson, 238-312.
Dettmers, C., Liepert, J., Adler, T., Rzanny, R., Rijntjes, M., van Schayck, R., Kaiser, W., Bruckner, L., Weiller, C., 1999. Abnormal motor cortex organization contralateral to early upper limb amputation in humans. Neuroscience Letters 263, 41-44.
Ding, S., al, e., 2009. Parcellation of human temporal polar cortex: a combined analysis of multiple cytoarchitectonic, chemoarchitectonic, and pathological markers. Journal of Comparative Neurology 514, 595-623.
Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., May, A., 2004. Neuroplasticity: changes in grey matter induced by training. Nature 427, 311-312.
Driemeyer, J., Boyke, J., Gaser, C., Buchel, C., May, A., 2008. Changes in gray matter induced by learning--revisited. PLoS One 3, e2669.

Dubois, J., Benders, M., Cachia, A., Lazeyras, F., Ha-Vinh Leuchter, R., Sizonenko, S.V., Borradori-Tolsa, C., Mangin, J.F., Huppi, P.S., 2008. Mapping the early cortical folding process in the preterm newborn brain. Cereb Cortex 18, 1444-1454.
Duda, R.O., Hart, P.E., Stork, D.G., 2000. Pattern classification.
Eberstaller, O., 1890. Das Stirnhirn. Wien and Leipzig: Urban and Schwarzenberg.
Edwin Clarke, K.D., 1973. An Illustrated History of Brain Function. 154 pp.
Ertöz, L., Steinbach, M., Kumar, V., 2002. A New Shared Nearest Neighbor Clustering Algorithm and its Applications. SIAM International Conference on Data Mining.
Evans, A.C., D.L.Collins, C.H.,
T.Paus,D.MacDonald,A.Zijdenbos,A.Toga,P.Fox,J.Mazziota, 1997. A 3D probalistic atlas of normal human neuroanatomy. Human Brain Mapping 4, S349.
Feess-Higgins, Larroche, 1987. Development du cerveau foetal humain: atlas anatomique. Masson, Paris, France.
Fesl, G., Moriggl, B., Schmid, U.D., Naidich, T.P., Herholz, K., Yousry, T.A., 2003. Inferior central sulcus: variations of anatomy and function on the example of the motor tongue area. NeuroImage 20, 601-610.
Fillard, P., Arsigny, V., Pennec, X., Hayashi, K.M., Thompson, P.M., Ayache, N., 2007. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. NeuroImage 34, 639-650.
Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B.T., Mohlberg, H., Amunts, K., Zilles, K., 2008. Cortical folding patterns and predicting cytoarchitecture. Cereb Cortex 18, 1973-1980.
Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage 9, 195-207.
Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb Cortex 14, 11-22.
Fornito, A., Malhi, G.S., Lagopoulos, J., Ivanovski, B., Wood, S.J., Saling, M.M., Pantelis, C., Yucel, M., 2008. Anatomical abnormalities of the anterior cingulate and paracingulate cortex in patients with bipolar I disorder. Psychiatry Research 162, 123-132.
Fornito, A.e.a., 2004. Individual differences in anterior cingulate/paracingulate morphology are related to executive functions in healthy males. Cereb Cortex 14, 424-431.
Fraley, C., Raftery, A.E., 2002. Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association 97, 611-631.
Fraley, C., Raftery, A.E., 2006. Mclust version 3 for R: Normal mixture modelling and model-based clustering. Technical Report / Brookhaven National Laboratory, Department of Statistics, University of Washington 504.
Francis, F., Meyer, G., Fallet-Bianco, C., Moreno, S., Kappeler, C., Socorro, A.C., Tuy, F.P., Beldjord, C., Chelly, J., 2006. Human disorders of cortical development: from past to present. European Journal of Neuroscience 23, 877-893.
Galaburda, A.M., Schmitt, J.E., Atlas, S.W., Eliez, S., Bellugi, U., Reiss, A.L., 2001. Dorsal forebrain anomaly in Williams syndrome. Archives of Neurology 58, 1865-1869.

Gan, G., Ma, C., Wu, J., 2007. Data Clustering: Theory, Algorithms, and Applications.
Gaser, C., Luders, E., Thompson, P.M., Lee, A.D., Dutton, R.A., Geaga, J.A., Hayashi, K.M., Bellugi, U., Galaburda, A.M., Korenberg, J.R., Mills, D.L., Toga, A.W., Reiss, A.L., 2006. Increased local gyrification mapped in Williams syndrome. NeuroImage 33, 46-54.
Gerber, S., al, e., 2010. Manifold modeling for brain population analysis. Medical Image Analysis 14, 643-653.
Good, P., 2004. Permutation, Parametric, and Bootstrap Tests of Hypotheses.
Hein, G., Knight, R.T., 2008. Superior Temporal Sulcus - It's My Area: Or Is It? Journal of Cognitive Neuroscience 20, 2125-2136.
Hilgetag, C.C., Barbas, H., 2006. Role of mechanical factors in the morphology of the primate cerebral cortex. PLoS Comput Biol 2, e22.
Hill, J., Dierker, D., Neil, J., Inder, T., Knutsen, A., Harwell, J., Coalson, T., Van Essen, D., 2010. A surface-based analysis of hemispheric asymmetries and folding of cerebral cortex in term-born human infants. Journal of Neuroscience 30, 2268-2276.
Hori, D.L., 2006. Clinical Neuroembryology: Development and Developmental Disorders of the Human Nervous System.
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, 602-611.
Im, K., Lee, J.M., Lyttelton, O., Kim, S.H., Evans, A.C., Kim, S.I., 2008. Brain size and cortical structure in the adult human brain. Cereb Cortex 18, 2181-2191.
Jain, A., Dube, R., 1988. Algorithms for Clustering Data. Prentice Hall, New Jersey.
Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data Clustering: A Review. ACM Computing Survays 31.
Jäncke, L., Steinmetz, H., 2003. Anatomical Brain Asymmetry and Their Relevance for Functional Asymmetries. In: Hugdahl, K., Davidson, R. (Eds.), The Asymmetrical Brain. A Bradford Book, The MIT Press Cambridge, Massachusetts
London, Engliand.
Jerison, H.J., 1973. Evolution of the Brain and Intelligence. Academic Press, New York, 482 pp.
Jouvent, E., Viswanathan, A., Chabriat, H., 2009. Cerebral atrophy in cerebrovascular disorders. Journal of Neuroimaging 20, 213-218.
Kates, W.R., Ikuta, I., Burnette, C.P., 2009. Gyrification patterns in monozygotic twin pairs varying in discordance for autism. Autism Res 2, 267-278.
Kaufman, L., Rousseeuw, P.J., 1990. Finding groups in Data.
Kim, H., Bernasconi, N., Bernhardt, B., Colliot, O., Bernasconi, A., 2008. Basal temporal sulcal morphology in healthy controls and patients with temporal lobe epilepsy. Neurology 70, 2159-2165.
Kim, S.G., Ashe, J., Hendrich, K., Ellermann, J.M., Merkle, H., Ugurbil, K., Georgopoulos, A.P., 1993. Functional magnetic resonance imaging of motor cortex: hemispheric asymmetry and handedness. Science 261, 615-617.
Kippenhan, J.S., Olsen, R.K., Mervis, C.B., Morris, C.A., Kohn, P., Meyer-Lindenberg, A., Berman, K.F., 2005. Genetic contributions to human gyrification: sulcal morphometry in Williams syndrome. Journal of Neuroscience 25, 7840-7846.

Kloppel, S., Mangin, J.F., Vongerichten, A., Frackowiak, R.S., Siebner, H.R., 2010. Nurture versus nature: long-term impact of forced right-handedness on structure of pericentral cortex and basal ganglia. Journal of Neuroscience 30, 3271-3275.
Kochunov, P., Mangin, J.F., Coyle, T., Lancaster, J., Thompson, P., Riviere, D., Cointepas, Y., Regis, J., Schlosser, A., Royall, D.R., Zilles, K., Mazziotta, J., Toga, A., Fox, P.T., 2005. Age-related morphology trends of cortical sulci. Human Brain Mapping 26, 210-220.
Krieg, W., 1966. Functional neuroanatomy. Brain Books, Evanston.
Le Goualher, G., Argenti, A.M., Duyme, M., Baare, W.F., Hulshoff Pol, H.E., Boomsma, D.I., Zouaoui, A., Barillot, C., Evans, A.C., 2000. Statistical sulcal shape comparisons: application to the detection of genetic encoding of the central sulcus shape. NeuroImage 11, 564-574.
Le Goualher, G., Procyk, E., Collins, D.L., Venugopal, R., Barillot, C., Evans, A.C., 1999. Automated extraction and variability analysis of sulcal neuroanatomy. IEEE Transactions on Medical Imaging 18, 206-217.
Lefevre, J., Mangin, J.F., 2010. A reaction-diffusion model of human brain development. PLoS Comput Biol 6, e1000749.
LeGrossClark, W.E., 1945. Deformation patterns on the derebral cortex. Essays on growth and form, 1-23.
Lo, C.H., Don, H.S., 1989. 3D moment forms: their construction and application to object identification and positioning. IEEE PAMI 11 (October), 1053-1064.
Lohmann, G., von Cramon, D.Y., 2000. Automatic labelling of the human cortical surface using sulcal basins. Medical Image Analysis 4, 179-188.
Lohmann, G., von Cramon, D.Y., Colchester, A.C., 2008. Deep sulcal landmarks provide an organizing framework for human cortical folding. Cereb Cortex 18, 1415-1420.
Lohmann, G., von Cramon, D.Y., Steinmetz, H., 1999. Sulcal variability of twins. Cereb Cortex 9, 754-763.
Lorr, Maurice, 1983. Jossey-Bass social and behavioral science series, San Francisco.
Lyttelton, O.C., Karama, S., Ad-Dab'bagh, Y., Zatorre, R.J., Carbonell, F., Worsley, K., Evans, A.C., 2009. Positional and surface area asymmetry of the human cerebral cortex. NeuroImage 46, 895-903.
MacDonald, D., Kabani, N., Avis, D., Evans, A.C., 2000. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. NeuroImage 12, 340-356.
MacQueen, J.B., 1967. Some Methods for classification and Analysis of Multivariate Observations. University of California Press, Berkeley.
Mangin, J.F., Poupon, F., Duchesnay, E., Riviere, D., Cachia, A., Collins, D.L., Evans, A.C., Regis, J., 2004a. Brain morphometry using 3D moment invariants. Medical Image Analysis 8, 187-196.
Mangin, J.F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Regis, J., 2004b. A framework to study the cortical folding patterns. NeuroImage 23 Suppl 1, S129-138.
Marin-Padilla, M., 1990. Origin, formation, and prenatal maturation of the human derebral cortex: An overview. Journal of Craniofacial Genetics and Developmental Biology 10, 137-146.
Michalewicz, Z., Fogel, D.B., 2004. How to solve it: Modern Heuristics. Springer.

Mietchen, D., Gaser, C., 2009. Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution. Front Neuroinformatics 3, 25.
Mojena, R., 1977. Hierarchical grouping methods and stopping rules: An evaluation. Computer J. 20, 359-363.
Murphy, J.T., Kwan, H.C., MacKay, W.A., Wong, Y.C., 1978. Spatial organization of precentral cortex in awake primates. III. Input-output coupling. Journal of Neurophysiology 41, 1132-1139.
Nordahl, C.W., Dierker, D., Mostafavi, I., Schumann, C.M., Rivera, S.M., Amaral, D.G., Van Essen, D.C., 2007. Cortical folding abnormalities in autism revealed by surface-based morphometry. Journal of Neuroscience 27, 11725-11735.
Ochiai, T., Grimault, S., Scavarda, D., Roch, G., Hori, T., Riviere, D., Mangin, J.F., Regis, J., 2004. Sulcal pattern and morphology of the superior temporal sulcus. NeuroImage 22, 706-719.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97-113.
Ono, M., Kubik, S., Abarnathey, C.D., 1990. Atlas of the Cerebral Sulci. Georg Thieme, New York.
Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., Kremen, W.S., 2009. Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex 19, 2728-2735.
Paus, T.e.a., 1996. Human Cingulate and Paracingulate Sulci: Pattern, Variability, Asymmetry, and Probalistic Map. Cerebral Cortex 6 (2), 207-214.
Penfield, W., Boldrey, E., 1937. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain 60, 389-443.
Penfield, W., Jasper, H., 1954. Epilespsy and the Functional Anatomy of the Human Brain. Little, Brown & Co., Boston.
Peper, J.S., Brouwer, R.M., Boomsma, D.I., Kahn, R.S., Hulshoff Pol, H.E., 2007. Genetic influences on human brain structure: a review of brain imaging studies in twins. Human Brain Mapping 28, 464-473.
Perrot, M., Riviere, D., Tucholka, A., Mangin, J., 2009a. Joint Bayesian cortical sulci recognition and spatial normalization. Inf Process Med Imaging 21, 176-187.
Perrot, M., Riviere, D., Tucholka, A., Mangin, J.F., 2009b. Joint Bayesian cortical sulci recognition and spatial normalization. Inf Process Med Imaging 21, 176-187.
Plaze, M., Paillere-Martinot, M.L., Penttila, J., Januel, D., de Beaurepaire, R., Bellivier, F., Andoh, J., Galinowski, A., Gallarda, T., Artiges, E., Olie, J.P., Mangin, J.F., Martinot, J.L., Cachia, A., 2009. "Where Do Auditory Hallucinations Come From?"--A Brain Morphometry Study of Schizophrenia Patients With Inner or Outer Space Hallucinations. Schizophrenia Bulletin.
Porac, C., Rees, L., Buller, T., 1990. Switching Hands: A Place for Left Hand Use in a Right Hand World. In: Coren, S. (Ed.), Left-Handedness: Behavioral Implications and Anomalies. North-Holland.
Poupon, F., 1999. Parcellisation systématique du cerveau en volumes d'intérêt. Le cas des structures profondes, INSA Lyon.

Provins, K.A., 1997. Handedness and speech: a critical reappraisal of the role of genetic and environmental factors in the cerebral lateralization of function. Psychological Review 104, 554-571.
Provost, L., al, e., 2003. Paracingulate sulcus morphology in men with early-onset schizophrenia. The British Journal of Psychiatry 182, 228-232.
Rademacher, J., Burgel, U., Geyer, S., Schormann, T., Schleicher, A., Freund, H.J., Zilles, K., 2001. Variability and asymmetry in the human precentral motor system. A cytoarchitectonic and myeloarchitectonic brain mapping study. Brain 124, 2232-2258.
Rakic, P., 1988. Specification of cerebral cortical areas. Science 241, 170-176.
Ray, S., Lindsay, B.G., 2005. The topography of multivariate normal mixtures. Annals of Statistics 33, 2042-2065.
Ray, S., Turi, H., 1999. Determination of number of clusters in K-means clustering and application in colour segmentation. The 4th International Conference on Advances in Pattern Recognition and Digital Techniques, 137-143.
Redcay, E., 2008. The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. Neuroscience & Biobehavioral Reviews 32, 123-142.
Regis, J., 1994. Anatomie Sulcale Profonde et Cartographie Fonctionelle du Cortex Cerebral. Marseille.
Regis, J., Mangin, J.F., Ochiai, T., Frouin, V., Riviere, D., Cachia, A., Tamura, M., Samson, Y., 2005. "Sulcal root" generic model: a hypothesis to overcome the variability of the human cortex folding patterns. Neurologia Medico-Chirurgica 45, 1-17.
Richman, D.P., Stewart, R.M., Hutchinson, J.W., Caviness, V.S., 1975. Mechanical model of brain convolutional development. Science 189(4196), 18-21.
Riviere, D., Mangin, J.F., Papadopoulos-Orfanos, D., Martinez, J.M., Frouin, V., Regis, J., 2002. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. Medical Image Analysis 6, 77-92.
Rogers, J., Kochunov, P., Zilles, K., Shelledy, W., Lancaster, J., Thompson, P., Duggirala, R., Blangero, J., Fox, P.T., Glahn, D.C., 2010. On the genetic architecture of cortical folding and brain volume in primates. NeuroImage 53, 1103-1108.
Roweis, S., Saul, L., 2000. Nonlinear dimensinality reduction by locally linear embedding. Science 290, 2323-2326.
Ruoss, K., Lovblad, K., Schroth, G., Moessinger, A.C., Fusch, C., 2001. Brain development (sulci and gyri) as assessed by early postnatal MR imaging in preterm and term newborn infants. Neuropediatrics 32, 69-74.
Schaer, M., Cuadra, M.B., Tamarit, L., Lazeyras, F., Eliez, S., Thiran, J.P., 2008. A surface-based approach to quantify local cortical gyrification. IEEE Transactions on Medical Imaging 27, 161-170.
Schmidhuber, J., 1997. A Computer Scientist's View of Life, the Universe, and Everything. Foundations of Computer Science, Lecture Notes in Computer Science. Springer, pp. 201-208.
Schmitt, J.E., Eyler, L.T., Giedd, J.N., Kremen, W.S., Kendler, K.S., Neale, M.C., 2007. Review of twin and family studies on neuroanatomic phenotypes and typical neurodevelopment. Twin Res Hum Genet 10, 683-694.

Shim, G., al, e., 2009. Reduced cortical folding of the anterior cingulate cortex in obsessive-compulsive disorder. Journal of Psychiatry and Neuroscience 34, 443-449.
Shim, G., Jung, W.H., Choi, J.S., Jung, M.H., Jang, J.H., Park, J.Y., Choi, C.H., Kang, D.H., Kwon, J.S., 2009. Reduced cortical folding of the anterior cingulate cortex in obsessive-compulsive disorder. Journal of Psychiatry and Neuroscience 34, 443-449.
Stoeckel, C., al, e., 2009. Supramarginal gyrus involvement in visual word recognition. Cortex 45, 1091-1096.
Sun, T., Collura, R.V., Ruvolo, M., Walsh, C.A., 2006. Genomic and evolutionary analyses of asymmetrically expressed genes in human fetal left and right cerebral cortex. Cereb Cortex 16 Suppl 1, i18-25.
Talairach, J., Tournoux, P., 1993. Referentially oriented cerebral MRI anatomy. New York: Georg Thieme Verlag.
Tamraz, Comair, 2006. Atlas of Regional Anatomy of the Brain Using MRI: with Functional Correlations. Springer.
Tenenbaum, J.B., de Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290, 2319-2323.
Thompson, P.M., Cannon, T.D., Narr, K.L., van Erp, T., Poutanen, V.P., Huttunen, M., Lonnqvist, J., Standertskjold-Nordenstam, C.G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C.I., Toga, A.W., 2001. Genetic influences on brain structure. Nature Neuroscience 4, 1253-1258.
Todd, P.H., 1982. A geometric model for the cortical folding pattern of simple folded brains. Journal of Theoretical Biology 97, 529-538.
Toews, M., Wells, W., D.L.Collins, Arbel, T., 2010. Feature-based morphometry: discovering group-related anatomical patterns. NeuroImage 49, 2318-2327.
Toro, R., Burnod, Y., 2005. A morphogenetic model for the development of cortical convolutions. Cereb Cortex 15, 1900-1913.
Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., Paus, T., 2008. Brain size and folding of the human cerebral cortex. Cereb Cortex 18, 2352-2357.
Tu, Z., Zheng, S., Yuille, A.L., Reiss, A.L., Dutton, R.A., Lee, A.D., Galaburda, A.M., Dinov, I., Thompson, P.M., Toga, A.W., 2007. Automated extraction of the cortical sulci based on a supervised learning approach. IEEE Transactions on Medical Imaging 26, 541-552.
Turing, A.M., 1948. Intelligent Machinery. Collected Works of A.M. Turing: Mechanical Intelligence, Amsterdam, pp. 87-106.
Turing, A.M., 1952. The Chemical Basis of Morphogenesis. Phil. Trans. Roy. Soc., 37-72.
Uylings, H., 2001. The human cerebral cortex in development. Handbook of Brain and Behaviour in Human Development, 63-80.
Van der Maaten, L.J.P., al, e., 2008. Dimensinality reduction: A comparative review. MICC, Maastricht University, Tech Report.
Van Essen, D.C., 1997. A tension-based theory of morphogenesis and compact wiring in the central nervous system. Nature 385, 313-318.
Van Essen, D.C., Dierker, D., Snyder, A.Z., Raichle, M.E., Reiss, A.L., Korenberg, J., 2006. Symmetry of cortical folding abnormalities in Williams syndrome revealed by surface-based analyses. Journal of Neuroscience 26, 5470-5483.

Wagner, R., 1862. Vorstudien zu Einer Wissenschaftlichen Morphologie und Physiologie des Menschlichen Gehirns als Seelenorgan. Göttingen:Verlag der Dieterichschen Buchhandlung 2, 1-105.
Watson, J.D., Myers, R., Frackowiak, R.S., Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S., Zeki, S., 1993. Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. Cereb Cortex 3, 79-94.
Welker, W., 1988. Why does cerebral cortex fissure and fold? Cereb Cortex 8B, 3-135.
Wen, Q., Chklovskii, D.B., 2008. A cost-benefit analysis of neuronal morphology. Journal of Neurophysiology 99, 2320-2328.
White, Hilgetag (Eds.), 2008. Gyrification of the human brain MIT Press, Cambridge, MA.
White, L.E., Andrews, T.J., Hulette, C., Richards, A., Groelle, M., Paydarfar, J., Purves, D., 1997. Structure of the human sensorimotor system. I: Morphology and cytoarchitecture of the central sulcus. Cereb Cortex 7, 18-30.
White, T., Andreasen, N.C., Nopoulos, P., 2002. Brain volumes and surface morphology in monozygotic twins. Cereb Cortex 12, 486-493.
White, T., Su, S., Schmidt, M., Kao, C.Y., Sapiro, G., 2010. The development of gyrification in childhood and adolescence. Brain and Cognition 72, 36-45.
Wobrock, T., Gruber, O., McIntosh, A.M., Kraft, S., Klinghardt, A., Scherk, H., Reith, W., Schneider-Axmann, T., Lawrie, S.M., Falkai, P., Moorhead, T.W., Reduced prefrontal gyrification in obsessive-compulsive disorder. European Archives of Psychiatry and Clinical Neuroscience 260, 455-464.
Wolfram, S., 2002. A New Kind of Science. Wolfram Media.
Woo, T.L., Pearson, K., 1927. Dextrality and sinistrality of hand and eye. Biometrika 19.
Xu, G., Knutsen, A.K., Dikranian, K., Kroenke, C.D., Bayly, P.V., Taber, L.A., 2010. Axons pull on the brain, but tension does not drive cortical folding. Journal of Biomechanical Engineering 132, 071013.
Yousry, T.A., Schmid, U.D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., Winkler, P., 1997. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. Brain 120 (Pt 1), 141-157.
Yu, X., Zhang, S., Liu, H., Chen, Y., 2006. The activation of the cortical hand area by toe tapping in two bilateral upper-extremities amputees with extraordinary foot movement skill. Magnetic Resonance Imaging 24, 45-50.
Zilbovicius, M., al, e., 2006. Autism, the superior temporal sulcus and social perception. Trends in Neurosciences 29, 359-366.
Zilles, K., Armstrong, E., Schleicher, A., Kretschmann, H.J., 1988. The human pattern of gyrification in the cerebral cortex. Anat Embryol (Berl) 179, 173-179.

APPENDIX A: RÉSUMÉ

Trouver des motifs dans les plissements corticaux

Résumé

Le contexte de cette thèse est l'étude de la variabilité des plissements du cortex. L'objectif principal est la conception d'algorithmes permettant de découvrir des motifs spécifiques à une sous-population d'individus. Le but final est de réaliser un dictionnaire de ces motifs et de les associer à des particularités cognitives ou architecturales, voire à des pathologies. Deux stratégies de clustering sont proposées pour mettre en évidence de tels motifs. La première repose sur des descripteurs de formes globaux correspondant aux invariants de moment 3D, la seconde repose sur l'estimation d'une matrice de distances entre chaque paire d'individus. Un algorithme de clustering dédié est conçu pour détecter les motifs les plus fréquents de manière robuste. Une technique de réduction de dimension est utilisée pour mettre en évidence les transitions entre motifs au sein de la population. Les méthodes algorithmiques proposées sont utilisées pour étudier la forme du cortex sensori-moteur d'une population de gauchers contrariés. Des résultats originaux sur le lien entre la forme du sillon central et la latéralité manuelle sont mis en évidence. Les méthodes développées sont ensuite utilisées pour construire le premier dictionnaire des motifs observés dans les plissements corticaux issu d'une approche algorithmique.

Le reste du résumé est organisé comme suit: La première section propose une introduction aux domaines les plus pertinents pour ce travail; la seconde section introduit une première approche de clustering fondée sur des descripteurs de forme; la troisième section explore les approches fondées sur le calcul de distances entre paires de sillons; La section 4 applique les méthodes mises au point à l'étude des corrélats de la latéralité manuelle sur la forme du sillon central, illustrant le potentiel des méthodes proposées; la section 5 comprend la discussion des divers aspects de l'analyse de la morphologie du

cortex et la section 6 présente les premières pages du dictionnaire que nous souhaitons construire à plus long terme.

1. Introduction

Dans cette section, un aperçu général est donné aux domaines de préoccupation fondamentale pour cette thèse: le phénomène du plissement cortical et la morphométrie informatisée de ces plis à partir de données d'imagerie par résonance magnétique.

1.1 Plissement cortical du cerveau

Le voyage commence avec le processus de plissement cortical. Ce processus de plissement, mystérieux et fascinant, n'est pas encore bien compris. Est-ce que les plissements corticaux contiennent des informations sur l'organisation fonctionnelle du cerveau humain? Dans la forme des plis, peut-on observer un motif caractéristique de certaines maladies neurologiques? Il existe une foule de connaissances éparses sur le sujet accumulées par les études détaillées de générations de neuroanatomistes. Grâce aux progrès récents des logiciels dédiés à la reconnaissance automatique des sillons corticaux, les questions concernant les plissements corticaux peuvent maintenant être revisitées à partir de bases de données incluant un grand nombre de cerveaux (Mangin et al., 2004c).

Il est fascinant de regarder les représentations 3D des plissements du cerveau. Comme l'a dit Welker: "le plus frappant, intéressant, encore mal compris des caractères morphologiques des hémisphères cérébraux chez les mammifères correspond aux dispositions diverses et complexes des circonvolutions corticales et des sillons" (Welker, 1988). Chaque cerveau est différent et aucun d'entre eux ne ressemble exactement à ceux des livres d’anatomie (des exemples sont présentés dans la figure 1). L’énorme variabilité entre ces plissements est manifeste.

Fig. 1. La variabilité des plissements du cortex

Hémisphères droits de trois personnes, les plis corticaux du lobe frontal sont mis en évidence selon la nomenclature traditionnelle.

Historiquement, certains ont beaucoup douté de la signification des formes des plissements du cortex : «semblable à la boucle de l'intestin grêle qui semblait ne présenter aucun ordre particulier» (Edwin Clarke, 1973). Dans le début des années 1800 l'attention a été attirée par Gall et ses disciples sur l'idée de localiser certaines facultés mentales dans des gyri spécifiques (Clarke, 1968). La phrénologie est maintenant du passé, mais cette initiative a conduit à de nombreuses études descriptives des circonvolutions cérébrales. De nombreuses observations suggèrent que la morphologie des plissements peut être liée à la connectivité et aux fonctions cérébrales (Welker, 1988). Il est donc intéressant d'explorer davantage les modèles de plissement extraits des images IRM avec des méthodes de calcul plus modernes et automatisées, et d'essayer de lier ces modèles à des fonctions, des comportements ou des pathologies.

Les études actuelles de la variabilité des plis corticaux s'appuient principalement sur des caractéristiques morphométriques simples, telles que la longueur ou la profondeur des sillons ou des gyri standards. La description la plus détaillée de la variabilité des sillons a été proposée dans l'atlas de Ono (Ono et al., 1990). Cet atlas ne repose pas sur un seul individu, mais sur vingt cerveaux différents. Pour chaque sillon, les auteurs proposent une liste de motifs observés et leurs fréquences. Ces modèles sont définis par exemple en tenant compte de la variabilité des interruptions des sillons. D'une certaine manière, l'objectif initial du travail de thèse est d'automatiser le travail effectué par Ono. Nous voulons découvrir des motifs stables pour un sous-ensemble de la population. En outre, nous voulons trouver des liens entre ces motifs et l'organisation fonctionnelle du cortex.

Un système de nomenclature standard a été mis en place pour décrire les circonvolutions de la surface du cerveau humain (Anatomica, 1983; Clemente, 1985; Ono et al, 1990.), (fig. 2 pour un exemple). Malheureusement, en raison de l'énorme variabilité qui existe entre les individus, la réalisation d'une conception unifiée de la description des sillons et des circonvolutions est extrêmement difficile. Un exemple intéressant d'une tentative de compréhension systématique des sillons et des circonvolutions est le modèle des racines

sulcales, des entités élémentaires correspondant aux premières ébauches des plis supposées stables à travers les sujets (Regis et coll., 2005). Ce modèle générique permet des analogies entre les motifs de l'ensemble des individus, voire même des études comparatives à travers les espèces. Il nécessite néanmoins encore des validations croisant des données issues de l'étude du développement cérébral, de l'architecture corticale intrinsèque au manteau cortical et de sa connectivité.

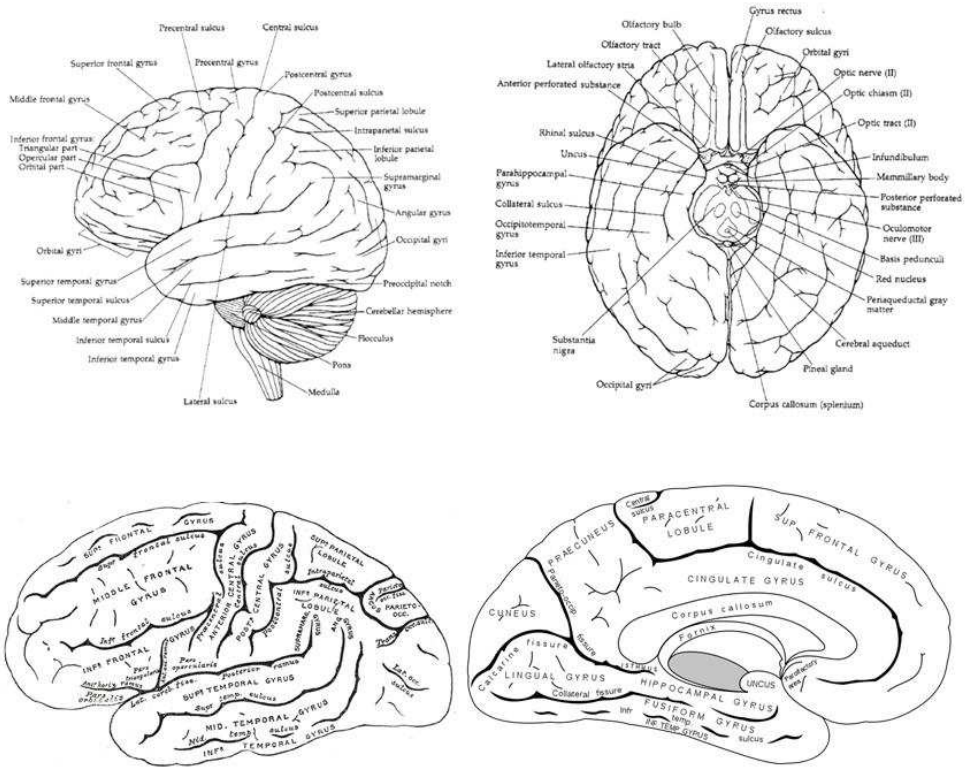


Fig 2 gyri et sillons, un exemple de la présentation du livre du texte de la nomenclature.
Gray's Anatomy: les bases anatomiques de Médecine et Chirurgie (édition britannique 38e éd.)

1.2 Les théories sur le mécanisme de plissement cortical

Il a d'abord été proposé que la sulcation est due à l'expansion du cortex limitée par le crâne et les ganglions de la base (LeGrossClark, 1945). Il fut découvert par la suite que l'élimination d'une grande quantité des structures corticale et sous-corticale du cerveau d'un mouton n'affecte pas la taille et l'organisation des sillons (Barron, 1950), invalidant cette hypothèse. Van Essen (Van Essen, 1997) propose que les connexions neuronales qui

se développent pendant le deuxième trimestre de grossesse produisent des tensions qui attirent fortement les régions interconnectées juqu'à les rapprocher grâce aux plis. D'autres théories du plissement mettent l'accent sur des facteurs mécaniques tels que la croissance différentielle des couches corticales (Todd, 1982). Les facteurs génétiques jouent probablement un rôle crucial dans la forme des plissements du cortex. Un contrôle génétique du développement cortical a ainsi été proposé (Rakic, 1988) et des liens importants entre les plissements du cortex et la cytoarchitecture ont été confirmés (Fischl et al, 2008.).

1.3 Approches informatiques pour étudier le plissement cortical

Les progrès récents de l'imagerie du cerveau à partir de la résonance magnétique (IRM) fournissent de précieuses informations concernant le plissement cortical. Par exemple, il est possible d'étudier la dynamique du plissement chez un sujet impliqué dans une étude longitudinale. Cette dynamique peut également être simulée : la simulation informatique d'un modèle morphogénétique (Toro et Burnod, 2005) a été proposée pour tenter de préciser l'importance relative des facteurs mécaniques et génétiques dans les plissements du cortex. Une autre approche intéressante pour expliquer la cohérence et la variabilité du plissement cortical s'appuie sur un modèle phénoménologique (Lefèvre et Mangin, 2010). L'étude suggère que des interactions non linéaires régies par un modèle de type reaction/diffusion suffisent à expliquer à la fois la stabilité de l'organisation générale des plissements et la variabilité issues de l'interruption de certains sillons.
Grâce à des approches algorithmiques, plusieurs cadres organisationnels génériques ont maintenant été proposés en ce qui concerne les schémas de plissement cortical. Dans le modèle des racines sulcales inspiré du développement (Regis et coll., 2005), la variabilité observée dans le plissement cortical adulte est expliquée sur la base de gyri plus ou moins enfouis dans les sillons (les plis de passage). De ce fait, une autre approche qui repose sur les maxima de profondeur des plissements principalement induits par ces plis de passage conduit au modèle des « sulcals pits » qui est très similaire au précédent (Lohmann et al., 2008 ; Im et al., 2010).

Une façon de permettre la comparaison des plissements du cortex et de la morphologie à travers les sujets est d'abord d'aligner les surfaces corticales. Cette approche permet de comparer l'épaisseur corticale et la surface des structures lobaires ou gyrales, mais aussi la géométrie des plissements du cortex. Un index est souvent utilisé pour quantifier l'ampleur du plissement, l'indice de gyrification (IG). L'IG a d'abord été défini comme le rapport entre les longueurs des contours coronaux 2D du cerveau incluant et excluant les sillons (Zilles et al., 1988). Des versions 3D globales ou locales sont aujourd'hui utilisées en routine. Une multitude d'autres paramètres morphologique peuvent être utilisés lorsque les sillons ont été identifiés au sein du plissement : la longueur, la profondeur, la surface et l'ouverture des sillons (Mangin et al., 2004b, Kochunov et al., 2005).

L'étude de la morphologie corticale apporte de la lumière sur de nombreuses questions importantes. En particulier, des résultats pertinents ont été mis en évidence dans quatre domaines: les anomalie du développement, le vieillissement, la plasticité, et la relation de la génétique au développement des plissements. Dans cette thèse, un lien intéressant concernant le développement du cortex et sa plasticité sera mis en évidence.

2. Clusters de plis

Dans cette section, le concept de l'aide d'algorithmes de clustering pour analyser les plissements du cortex est introduit. L'effort de sélection du descripteur de forme approprié et le développement d'un algorithme de clustering spécifique à l'analyse des plissements du cortex sont décrits.

2.1 Définition

Tout d'abord, nous avons besoin de définir les motifs de plissement cortical que nous recherchons. Dans une étude préliminaire, ces motifs sont vaguement définis comme un

groupe de cerveaux qui montrent un trait caractéristique qui les distingue des autres cerveaux. Cette caractéristique pourrait être basée sur la surface du pli, le degré de courbure, le nombre et la position des interruptions, etc. Puisque ce type d'apprentissage non supervisé sur la forme des plis n'a jamais été réalisé systématiquement avant, nous n'avons pas une définition claire. Il est également très probable qu'aucun ensemble restreint de paramètres (angle, longueur, profondeur, topologie…) ne permette de caractériser la totalité des motifs intéressants.

2.2 Méthodes de clustering

Nous discutons maintenant de la méthode qui peut être utilisée pour découvrir des motifs répondant à la définition ci-dessus. Il s'agit a priori d'un problème de clustering, c'est-à-dire
un problème de classification non supervisée. Cette situation se distingue de l'apprentissage supervisé par le fait qu'il n'y a pas d'exemples a priori à apprendre pour guider la méthode. Ce type d'analyse est souvent effectué lorsque peu d'information a priori est disponible sur les données. Le clustering est une composante essentielle du monde de l'exploration de données qui peut parfois permettre de découvrir des informations utiles (Berry et Linoff, 2000).

Un grand nombre d'algorithmes de classification existe dans la littérature de différents domaines : la reconnaissance des formes, l'intelligence artificielle, le traitement d'images, les statistiques et mathématiques appliquées, la psychologie et la biologie. Dans chaque domaine, un ensemble d'algorithmes ont tendance à dominer pour des raisons historiques et pratiques. Par exemple, l'approche hiérarchique est plus utilisée dans la communauté de l'intelligence artificielle, alors que l'approche basée sur un modèle est plus utilisée dans la communauté statistique. Bien que de nouvelles avancées soient régulièrement présentées dans le développement d'algorithmes de clustering, les défis demeurent. Une partie de la raison est que les problèmes spécifiques à un domaine nécessitent souvent l'utilisation d'algorithmes spécialement conçus, les algorithmes de clustering généraux ne sont en effet souvent pas suffisants. Le comportement des situations de la vie réelle est

toujours complexe et imprévisible, et exige des solutions plus sophistiquées ou plus spécifiques.

Il n'y a pas de technique de regroupement universellement applicable pour découvrir la variété des structures présentes dans les ensembles de données multidimensionnelles. Une technique donnée ne met en évidence qu'un type de structures, car les algorithmes de regroupement contiennent souvent des hypothèses implicites ou explicites sur la forme des clusters. Par exemple, il y a des clusters compacts et des clusters chaînés. Les membres d'un cluster compact ont deux à deux une grande similarité mutuelle; habituellement un cluster compact peut être représenté par un centre. En revanche, un cluster chaîné est constitué d'une série de points dans laquelle chaque membre n'est très semblable qu'à quelques autres points. Ces deux types de clusters sont intéressants même s'ils mettent en évidence des informations différentes.

De nombreux algorithmes de classification existent. Une taxonomie usuelle consiste à distinguer les approches hiérarchiques des approches partitionnelles (Gan et al., 2007). Les méthodes partitionnelles sont appliquées lorsque l'on veut classer les objets en N groupes. Les algorithmes hiérarchiques, quant à eux, ne construisent pas une seule partition mais des partitions emboîtées pour tous les N possibles.

Quel type d'algorithme de clustering est le plus adapté à l'analyse des plissements du cortex ? Pour prendre une décision, nous devons revenir sur notre objectif initial: trouver des motifs de plis. Nous recherchons des groupes de cerveaux de taille raisonnable qui présentent une similitude dans leurs plissements. Par conséquent, nous ne cherchons pas à classer chaque cerveau. En outre, notre idée intuitive d'un motif correspond a priori à un cluster compact. Finalement, au regard de l'incroyable variabilité des plissements, il semble impératif de disposer d'une technique très robuste au bruit. Il est probable que les motifs les plus clairs que nous allons chercher à mettre en évidence ne s'appliquent qu'à une minorité de cerveaux. Notre idée initiale était en effet que des motifs purs correspondent à des situations biologiques extrêmes, où le nombre de phénomènes

impliqués dans le plissement local est faible, et qui ne sont valables que pour un nombre restreint d'individus.

2.3 Les descripteurs de forme

Les moments invariants 3D ont été proposés comme des descripteurs intéressants pour l'étude de la forme des sillons corticaux, car ils peuvent être calculés pour n'importe quelle topologie (Mangin et al., 2004). Par conséquent, ils permettent la gestion des fréquentes interruptions des sillons. La construction de ces descripteurs les rend insensibles à la localisation, à l'orientation et à l'échelle des objets. On peut donc les considérer comme des descripteurs de la forme « pure ». Bien que leur dérivation théorique soit complexe, ils peuvent être calculés de manière simple et robuste à partir d'une image en noir et blanc définissant un objet. Dans la suite, nous utilisons seulement les 12 invariants calculés à partir des moments de coordonnées d'ordre inférieur ou égal à trois. Notez néanmoins que le calcul des moments invariants est théoriquement possible pour n'importe quel ordre. Il est malheureusement relativement difficile de se faire une idée intuitive des informations de formes codées par ces invariants car ils s'agit de polynomes complexes des moments. Pour le travail présenté ici, les invariants d'un sillon ou d'un groupe de sillons sont calculés en utilisant BrainVISA (http://brainvisa.info) pour chaque individu avant d'être utilisés pour le clustering.

Des études préalables ont été menées pour vérifier que l'ensemble des 12 moments invariants utilisé est une représentation raisonable de la forme d'un sillon. Il était en particulier important de s'assurer qu'une petite variation de la forme ne conduit pas à une grande modification des invariants, car cette propriété n'est pas garantie par leur construction. Notre approche très pragmatique du problème a consisté à créer des séries de formes échantillonées correspondant à une déformation régulière d'une forme simple. De multiples expériences de ce type ont montré que les invariants varient lentement pour les formes explorées, mis à part deux d'entre eux qui peuvent changer de signe lors de l'ajout d'un seul voxel et présentent une distribution bimodale pour certains sillons. Ils

ont donc été supprimés de notre représentation. Ce type de phénomène existe peut-être pour des formes rares pour les autres invariants, mais il n'était pas problématique pour notre objectif qui vise essentiellement à découvrir quelques motifs afin de tester leur signification biologique.

Une deuxième étude visait à vérifier que les informations sur la forme pure intégrées dans les invariants peuvent distinguer les formes qui caractérisent différents sillons corticaux. Pour ce but, nous avons réuni à de nombreuses reprises plusieurs grands sillons pour un grand groupe d'individus, et nous avons réalisé à chaque fois une analyse en composante principale (Fig. 3). Dans beaucoup de cas, les premiers axes de l'ACP montrent que les invariants séparent les sillons. On observe parfois des recouvrements mais ce n'est pas étonnant car certains sillons ont des formes similaires. Le fait que chaque sillon conduise à un nuage relativement compact implique que les représentations à base d'invariants peuvent être utilisées pour rechercher des motifs.

2.4 L'algorithme de clustering pour la découverte de motifs sulcaux

Au vu des impératifs mentionnés ci-dessus, il semble assez évident que l'approche hiérarchique du clustering est la plus adaptée à notre volonté de trouver des clusters compacts noyés au sein de nuages de points très hétérogènes. Cette méthode est donc choisie dans la phase qui suit.

Pour trouver la règle permettant d'arrêter le processus d'agglomération hiérarchique au niveau le plus intéressant, une méthode spécifique a été conçue. Elle s'appuie sur la génération d'un grand nombre de nuages de points aléatoires engendrés par une distribution gaussienne estimée à partir des données réelles. L'idée sous jacente est que la découverte d'un ensemble de motifs intéressant doit invalider l'hypothèse que la distribution des points résulte d'une simple gausienne. Un ratio classique entre la compacité des clusters et leur éloignement permet de caractériser un ensemble de motifs. Les distributions de ce ratio sur les ensembles aléatoires sont utilisées pour sélectionner

le niveau de l'arbre procurant la meilleure configuration, c'est-à-dire celle qui s'éloigne le plus des distributions observées dans les ensembles aléatoires.

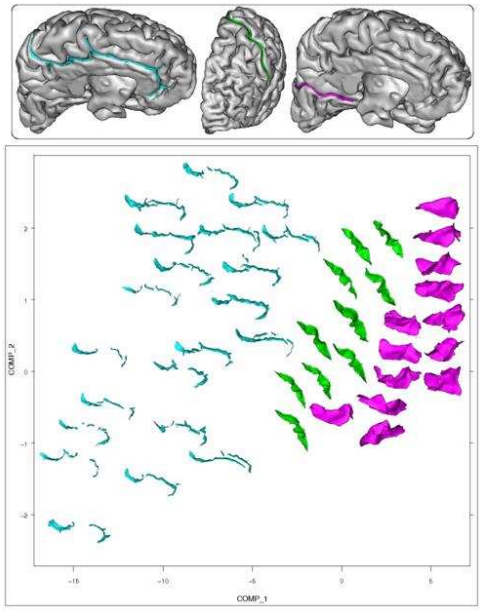


Fig. 3 Les invariants de moment ont un pouvoir discriminant entre les formes des sillons.

Les données de trois sillons, le sillon cingulaire (cyan), le sillon central (vert) et le sillon pariéto-occipital (magenta) sont plottées en utilisant les deux premiers axes d'une ACP appliquée au vecteur d'invariants normalisé (36 cerveaux étiquetés manuellement sont utilisés). Les formes de certains sillons choisis au hasard sont superposées à l'ACP. Un exemple de chaque sillon est proposé pour un cerveau.

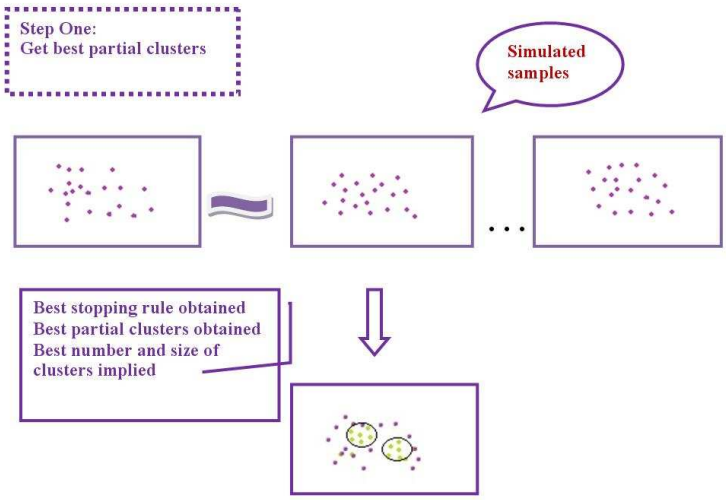


Fig. 4A Obtenir les meilleurs clusters partiels

Les échantillons simulés sont obtenus en échantillonant une distribution gaussienne estimée à partir de la matrice de covariance des données réelles. La classification hiérarchique est exécutée sur l'ensemble de

données réelles et tous les ensembles de données simulées. L'étape de l'agglomération dotée du meilleur rapport compacité/éloignement obtenu est séléctionnée. La qualité des clusters trouvée à cette étape est définie par le calcul d'une p-valeur. Cette valeur est calculée en comptant le nombre de fois que l'ensemble de données simulées donne de meilleurs résultats que les données réelles.

La résistance finale des clusters trouvés est fondée sur une p-valeur calculée en comptant le nombre de fois où les jeux aléatoires sont plus performants que l'ensemble de données réelles. Ces groupes sont appelés clusters partiels (figure 4A). Un algorithmes de bootstrap est ensuite utilisé pour augmenter la stabilité de l'algorithme (cf figure 4B). Le procédé décrit dans l'étape 1 est réalisé à plusieurs reprises sur des sous-ensembles des données d'origine. Les centres des clusters obtenus sont conservés pour procéder à un second niveau de clustering relativement aisé généralement réalisé avec l'algorithme K-medoid associé à la méthode PAM pour estimer le nombre de clusters.

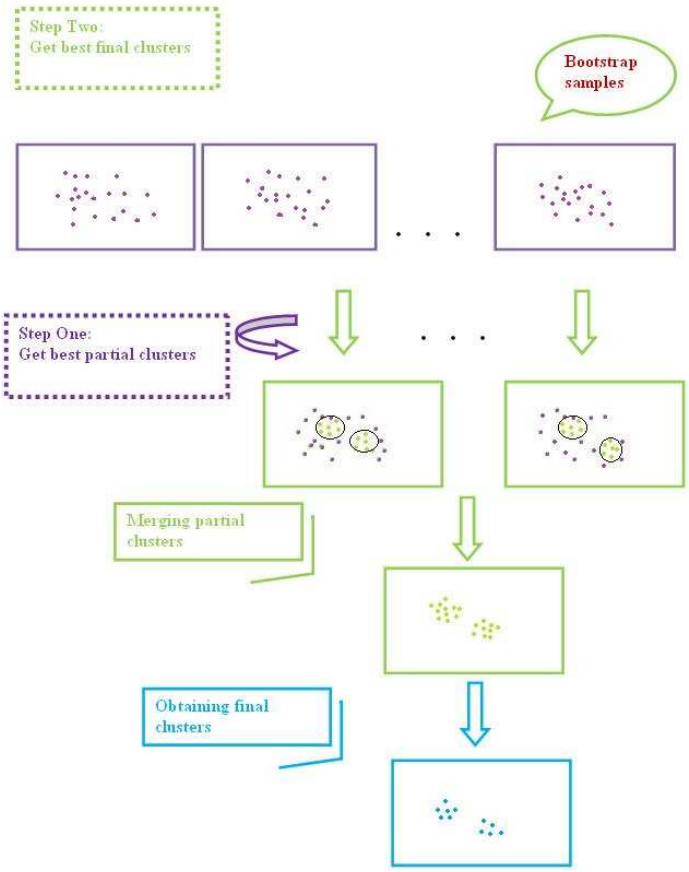


Fig. 4B Robustification à l'aide d'un bootstrap

2.5 Validation de l'algorithme de clustering

Pour évaluer la performance de l'algorithme PCBB, plusieurs expériences ont été effectuées sur des ensembles de données simulées. Les ensembles de données simulées sont générés comme suit. Choisir une paire de sillons de façon aléatoire. Générer pour chaque sillon un cluster tiré d'une distribution Gaussienne estimée à partir de la distribution de ce sillon. Cela donne un ensemble de données avec deux clusters connus. Ajouter du bruit pour créer une suite de données simulées de plus en plus bruitées. Le bruit ajouté suit une distribution de Poisson, dont les valeurs minimale et maximale correspondent au jeu de données d'origine.
Deux algorithmes sont exécutés sur ces ensembles de données simulées : une méthode classique de Modèle de Mélange Gaussiens (GMM, Duda et al., 2000) et l'algorithme PCBB. Les résultats sont évalués en fonction du nombre de clusters trouvés et la distance entre les centres trouvés et les véritables centres. Le nombre de gaussienne du GMM est estimé à partir du critère d'information bayésien (BIC). Pour nos expériences, nous utilisons la boîte à outils Mclust de R pour évaluer le GMM (Fraley et Raftery, 2002, 2006). Nous testons en fait deux variantes de GMM. Une approche classique est une variante où la méthode est informée de la proportion de bruit ajoutée aux données.

Deux comparaisons sont effectuées pour évaluer la qualité d'analyse. Tout d'abord, la distance entres les centres trouvés et les véritables centres est calculée. Quand il y a plus de centres trouvés par l'algorithme que de centres réels, seuls les deux groupes les plus proches des centres réels sont pris en considération. Deuxièmement, les nombres de clusters trouvés par chaque algorithme sont comparés au valeurs réelles.

Le résultat est montré dans la figure 5 ci-dessous. Les résultats montrent que l'algorithme PCBB est comparable à l'algorithme de GMM en termes de localisation des centres des clusters. Toutefois, en termes d'estimation du nombre de clusters, PCBB est plus précis et

plus stable que l'algorithme de GMM lorsque le bruit augmente. Informer le GMM du pourcentage de bruit lors de l'initialisation ne semble pas améliorer les performances. Le résultat montre que PCBB est plus robuste que GMM pour ce problème particulier.

2.6 Résultats en utilisant l'algorithme PCBB

La recherche de motifs est réalisée sur l’ensemble de 36 cerveaux déjà utilisé pour générer les données simulées pour la validation. Un deuxième ensemble de données composé de 150 cerveaux est aussi utilisé pour vérifier la reproductibilité des résultats. Un exemple de clustering obtenu pour le sillon cingulaire est représenté (Fig. 6).

Un premier motif est constitué de sillons présentant une grande interruption antérieure, un second motif est constitué de sillons présentant une interruption plus petite et plus postérieure, et un troisième modèle est constitué de sillons continus. Il convient de noter que ces motifs ne peuvent être déduits simplement du nombre de composantes connexes. En effet, les sillons du troisième motif ne sont continues qu’en apparence: certains d'entre eux sont constitués de plusieurs composantes connexes qui se chevauchent les uns les autres lorsque le sillon est vu de dessus. En fait, les moments invariants sont aveugles à la topologie. Par conséquent, ces trois motifs seraient interprétées de manière plus fiable en termes de forme que sur le plan de l'interruption. Par exemple, le premier motif correspond à des sillons beaucoup plus profonds dans la partie postérieure que dans le milieu, tandis que le dernier motif correspond à une profondeur des sillons plus homogène.

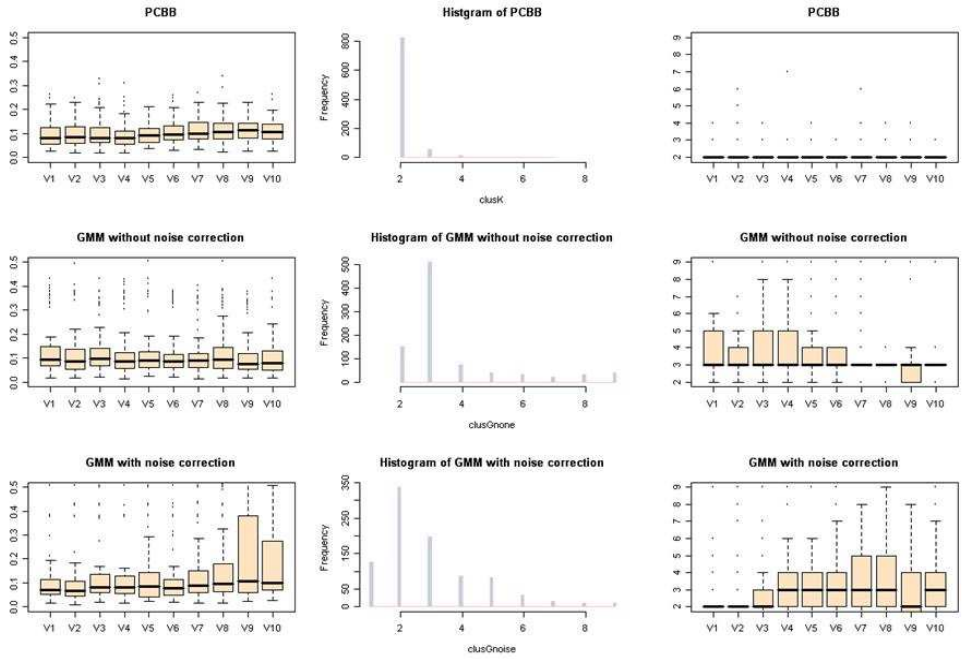


Figure 5 Performance de plusieurs algorithmes de clustering sur une simulation

La première colonne indique le boxplot de la distance entre les deux centres les plus proches trouvés par l'algorithme et les vrais centres. (La boîte à moustaches contient les données du second et du troisième quartile, la ligne sombre à l'intérieur de la boîte représente la médiane, les moustaches indiquent les minima et les maxima, et les valeurs aberrantes sont représentées par des points.) L'axe des abscisses correspond à l'augmentation du pourcentage de bruit de 10 à 100 pour cent. Les résultats de la méthode PCBB sont indiqués sur la première ligne, les résultats de GMM sans correction du bruit sont indiqués sur la deuxième ligne, et les résultats de GMM avec correction du bruit, sont indiqués sur la troisième ligne. La deuxième colonne montre l'histogramme de la distribution du nombre de centres pour les trois algorithmes, la troisième colonne indique le boxplot du nombre de centres.

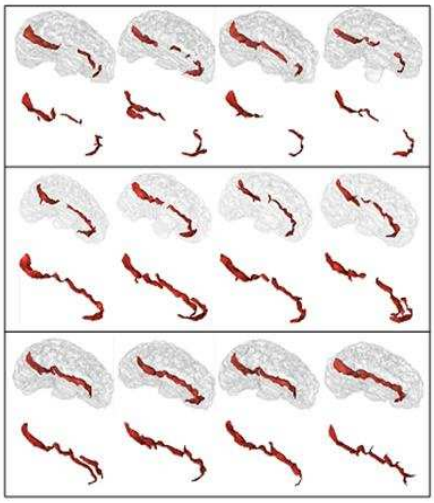


Fig 6. Les trois motifs détectés pour le sillon cingulaire gauche.

Ligne 1,3,5: les quatre cas les plus centraux du motif dans la base de données manuellement étiquetées.

Ligne 2,4,6: les quatre cas les plus proches dans une base de données automatiquement étiquetées.

3. Des modèles de la variabilité des plis

Dans cette section, de nouvelles approches sont explorées pour l'étude de la variabilité des plissements. En particulier, de nouveaux descripteurs de formes sont utilisés. Ils correspondent simplement à des mesures de similarité fondées sur une distance entre paires de sillons. Cette approche est couplée avec une méthode de réduction de dimension, « l'isomap », qui permet d'appréhender les modes de variabilité les plus importants pour un sillon donné. Cette démarche complète celle qui consiste à faire émerger des motifs fréquents et permet d'aboutir à une description très riche de la variabilité des formes d'un sillon. En effet cette description permet de comprendre les transitions progressives qui permettent de passer d'un motif à un autre.

3.1 Passer des motifs à des modèles du plissement cortical

Les travaux décrits dans la section précédente consistent à trouver les motifs les plus fréquents en termes de plissements du cortex. Ces travaux ont montré qu'une étude de la variabilité des sillons implique trois choix importants: quels sillons étudier, quelle mesure

de similarité utiliser pour comparer deux sillons, et quel algorithme utiliser pour extraire une structure intéressante à partir de ces mesures de similarité.

En ce qui concerne le choix de l'endroit où chercher parmi les sillons corticaux, deux nouveaux angles peuvent être explorés. Dans le travail présenté dans la première partie, les plus grands plissements avaient été choisis pour la détection de motifs. Ce sont probablement les plus stables, et par conséquent ils peuvent être définis de manière plus fiable. Les plus grands sillons sont en effet supposés exister chez tous les individus, tandis que les plis plus petits peuvent ne pas exister chez un individu donné en raison de la forte variabilité du processus de plissement. L'inconvénient ou la limitation d'une telle approche est que les petits plis secondaires recèlent probablement une mine d'information du fait de leur grande variabilité. Une approche qui ciblerait non pas un sillon mais une région complète du cortex pourrait donc s'avérer plus intéressante. En outre, elle permettrait parfois de lever un certain nombre d'ambiguïtés au cours de la phase d'étiquetage des sillons qui viennent perturber l'interprétation des motifs mis en évidence.
La vision actuelle de l'organisation du cortex associe à chaque fonction un réseau de localisations. Chacune de ces localisations correspond à une région spécifique du cortex traversée par un certain nombre de plis. Les associations de ce type les plus anciennes sont les aires de Broca et de Wernicke associées au language. En général, ces régions ne sont pas associées à un sillon mais plutôt à un gyrus considéré comme un module fonctionnel siège de l'activité cérébrale. Il nous a donc semblé tentant d'étudier la forme de régions du cortex à travers la forme de l'ensemble des sillons les délimitant. Par exemple, le sillon central et le sillon précentral peuvent être regroupés pour constituer une description de la forme du gyrus précentral associé aux structures motrices. Sur la base de ce raisonnement, nous avons décidé d'étudier la forme de groupes de sillons de ce type.

En ce qui concerne le choix de la mesure de similarité, d'autres possibilités peuvent être explorées. Nous avons juqu'à présent utilisé des descripteurs de forme 3D, les invariants

de moment 3D. Bien que ces descripteurs puissent capturer une partie de la forme des plis, l'information codée était toutefois limitée car nous n'avons utilisé que dix invariants. Une autre approche intéressante et plus intuitive est l'utilisation de similarités calculées directement entre sillons. Chaque sillon peut être caractérisé par ses similarités plus ou moins importantes avec les sillons de tous les autres sujets. La caractérisation du sillon est beaucoup plus complète qu'avec des descripteurs de forme génériques, car elle s'appuie sur une base de formes similaires. Elle peut donc contenir des subtilités difficiles à coder autrement.

En ce qui concerne l'algorithme à utiliser pour l'analyse, de nombreuses nouvelles orientations peuvent être explorées. Il est important de réaliser, cependant, que lorsque la nature des données d'entrée est modifiée, l'algorithme de clustering peut avoir besoin d'être changé. En particulier, si l'approche du vecteur de mesures de similarité discutée ci-dessus était adoptée, la dimension de l'entrée serait très élevée. Quand la dimension des données est élevée, la distance entre deux points de données quelconques devient presque constante (Beyer et al., 1999). Dans ce cas, il est difficile de faire émerger des clusters de manière robuste (Agrawal et al, 1998.). Pour ces raisons, presque tous les algorithmes de clustering classiques ne fonctionnent pas bien pour les ensembles de données de grande dimension. Par conséquent, l'algorithme utilisé doit être en mesure soit de réduire la dimension globalement, soit de procéder à une sélection des dimensions les plus intéressantes pour le clustering. Une alternative est néanmoins d'utiliser un des rares algorithmes spécialement conçus pour les ensembles de données de grande dimension.

3.2 La mesure de similarité

Nous avons déjà évoqué que la description fondée sur les 10 invariants 3D est trop limitée pour bien représenter toute la complexité des plissements. Des formes très différentes peuvent générer des descripteurs presques similaires, ce qui perturbe la recherche de motifs stables. Afin de surmonter cette faiblesse, une approche différente est explorée ici. Une forme de sillon est décrite par un vecteur de distances de cette forme à

un grand nombre de formes semblables. Chaque paire de distance est calculée en utilisant l'algorithme « Iterated Closest Point » (ICP), après normalisation spatiale affine globale du cerveau. Notez que le but de cette normalisation globale est de supprimer d'éventuels motifs induits par des différences globales de taille du cerveau. Notre mise en œuvre de l'ICP calcule la distance moyenne quadratique minimale obtenue quelle que soit la rotation appliquée pour tenter d'aligner les deux formes.

L'idée de représenter chaque sillon par sa distance à l'ensemble des autres sillons dans le jeu de données est illustrée dans la figure 7. Le plus simple algorithme ICP est utilisé : une forme (donnée) est adaptée à une autre forme (modèle) de façon itérative par rotation et translation (Besl et McKay, 1992). L'algorithme s'arrête lorsque la distance moyenne de la forme mobile au modèle ne peut plus être considérablement diminuée ou lorsqu'un nombre maximum d'itérations a été atteint. Il convient de noter que cet algorithme a été largement étudié et que de nombreuses améliorations en ont été proposées que ce soit pour gagner en efficacité de calcul ou en robustesse, ou pour l'utiliser avec des transformations non rigides. Dans notre étude, nous nous sommes contentés de la version la plus simple pour plusieurs raisons.

Notre but initial étant de défricher l'intérêt potentiel d'une analyse des motifs corticaux, nous ne cherchons pas à atteindre l'exhaustivité dans leur cartographie. L'approche fondée sur un recalage rigide nous a paru suffisante pour montrer l'intérêt de la démarche si elle met en évidence quelques motifs intéressants. Nous avons préféré cette approche à une approche intégrant des degrés de liberté affines du fait de la normalisation globale utilisée comme préalable. Des différences de tailles perdurant après la normalisation globale pourraient être particulièrement intéressantes d'un point de vue neurosciences. Nous avons renoncé dans un premier temps à la tentation des approches non linéaires essentiellement pour des raisons de temps de calcul, mais il est clair que dans le futur nous définirons la notion de motif sulcal par rapport à un ensemble de transformations possibles pour recaler les sillons, dans l'esprit usuel utilisé pour définir un invariant. Le travail décrit par la suite considère seulement l'ensemble des transformations rigides.

Nous n’avons également pas jugé nécessaire de mettre en œuvre les raffinements proposés dans la littérature pour améliorer la robustesse de l’ICP vis-à-vis des minima locaux. En effet, nous verrons par la suite que les seules distances véritablement utilisées par notre méthode sont les distances faibles obtenues entre formes voisines. L’ICP s’avère naturellement robuste lorsqu’on l’applique avec deux formes très similaires, les difficultés ne surviennent en effet que lorsque deux formes diffèrent de façon importante. Mais dans ce dernier cas notre seul besoin est d’obtenir une distance élevée, ce qui arrive quel que soit le minimum local.

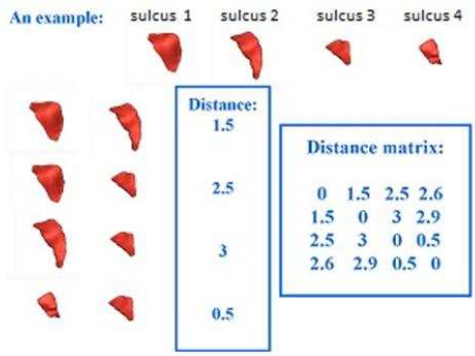


Figure 7: Un exemple de la mesure de similarité

A titre d'exemple pour illustrer le concept de la mesure de similarité, quatre sillons sont utilisés (en rouge) appelé modèle 1 à 4. La distance (par ICP) peut être calculée pour chaque paire de l'ensemble. Pour chaque motif, la distance aux trois autres modèles peut être calculée. Elle donne ici une idée sur l’existence de deux groupes.

3.3 Algorithme pour gérer des données de grande dimension

La matrice de distance obtenue en utilisant l'algorithme ICP peut avoir une dimension très élevée, puisque le nombre de dimensions est égal au nombre de sujets. La « malédiction de la dimensionnalité » est un problème bien connu survenant dans ce genre de situations (Duda et al., 2000). La dimension de la matrice doit donc être réduite avant d'appliquer les algorithmes de clustering standard.

La légitimité de la réduction de dimension est que les points de données évoluent probablement dans un espace (ou une variété) de dimension bien inférieure à la

dimension des vecteurs de distances qui décrivent ces points. L'objectif est donc de construire un tel sous espace préservant au maximum la répartition spatiale initiale des données. L'algorithme Isomap a été choisi pour cette étude préliminaire, mais de nombreuses alternatives pourraient être explorées à l'avenir. L'algorithme Isomap a l'efficacité de calcul et l'optimalité globale de l'analyse en composantes principales (ACP) et du Multi-Dimensional Scaling (MDS). Il a également la possibilité d'apprendre une large classe de variétés non-linéaire (Tenenbaum et al., 2000). Cet algorithme utilise le MDS appliqué à des distances géodésiques entre les points. Ces distances sont estimées comme les plus courts chemins dans un graphe des plus proches voisins pour la similarité initiale. En dépit de quelques faiblesses, cet algorithme a été appliqué avec succès dans la vision artificielle et la visualisation de données biomédicales (avis 2008).
Un des intérêts des approches de type isomap réside dans le peu de paramètres à régler. En ce qui concerne l'isomap, il y en a essentiellement un, le nombre de plus proches voisins utilisé pour construire le graphe dont découlent les distances géodésiques. Lorsque le voisinage est trop grand, des « court-circuits » risquent d'être créés ; lorsque le voisinage est trop petit, les chemins dans le graphe sont trop rares pour procurer une bonne approximation de l'espace sous-jacent. À notre connaissance, il n'existe pas de manière consensuelle de choisir K quel que soit le problème. C'est la principale faiblesse de l'approche Isomap. Cet algorithme est « topologiquement instable » : des connexions erronées peuvent être créés dans le graphe de voisinage. Certaines méthodes ont néanmoins été proposées pour surmonter ces problèmes comme la suppression des voisins les plus proches qui violent la linéarité locale du graphe (Van der Maaten and al, 2008), mais nous n'avons pas eu le temps de les considérer.

Une fois la matrice de distance géodésique calculée, un algorithme de réduction de dimension simple telle que le Multidimensional Scaling (MDS) peut être utilisé. Nous avons en fait surtout utilisé le MDS non-métrique (IsoMDS) qui préserve l'ordre du voisinage plutôt que les distances, dans la mesure où notre mesure de similarité n'est pas une véritable distance.

3.4 Validation de l'approche ICP-Isomap

Combiner l'ICP et l'Isomap pour l'analyse des plissements du cortex nécessite une validation. L'efficacité de l'algorithme de clustering PCBB présenté dans la deuxième section a été vérifiée. La validation est effectuée sur des ensembles de données simulés et réels.

3.4.1 La pertinence de l'ICP-Isomap pour l'analyse du cortex: ensemble de données simulées

La performance de l'algorithme de clustering est d'abord évaluée à partir de simulations générées en utilisant le sillon central. Pour chaque simulation, trois sujets sont choisis au hasard dans la base de données originale. Six variations aléatoires sont générées pour chacun d'eux par transformation affine. Un exemple est fourni dans la figure. 8. La base de données centrale de 62 sillons (Perrot et al., 2009) est utilisée pour la génération de ces ensembles de données simulées. Chacun d'eux est composé de 3 clusters simulés compacts incluant chacun 7 sillons auxquels sont ajoutés 41 autres sillons centraux de la base, conduisant à un total de 62 sillons.

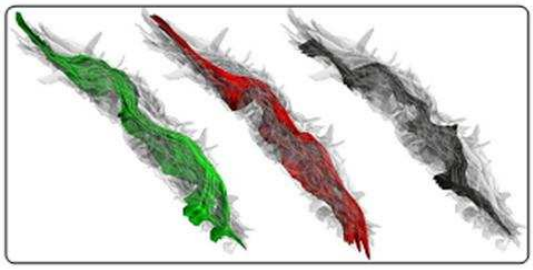


Figure 8 Les sillons centraux simulés : Chaque cluster est indiqué par une couleur et est superposé au reste de l'ensemble. On peut voir que la variabilité au sein de chaque cluster est différente, mais que les formes des trois groupes suivent de près la forme du sillon central graine. Pour ces visualisations, tous les sillons ont été alignés avec le sillon le plus neutre, celui qui minimise la distance moyenne à l'ensemble.

Lorsque le clustering en utilisant PCBB est réalisé après une ACP, un MDS ou l'Isomap, il est constaté qu'Isomap offre les meilleurs clusters: des clusters compacts éloignés les uns des autres (Fig. 9 pour les clusters trouvés).

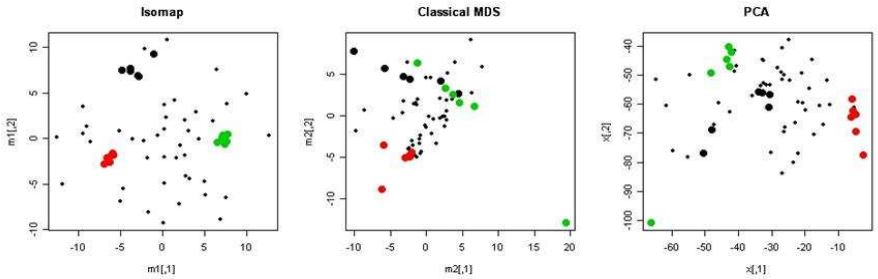


Figure 9 La comparaison des méthodes de réduction de dimension sur les clusters simulés. De gauche à droite: la distribution en utilisant Isomap, MDS classique et ACP. Les groupes représentés sont ceux trouvés en exécutant l'algorithme PCBB. La couleur des sillons correspond à celle de la figure 8.

Les résultats sont évalués en fonction du nombre de clusters trouvés. Un cluster détecté est considéré comme un succès si la distance de son centre au centre du cluster simulé le plus proche est dans le rayon de ce cluster. Le rayon est défini comme la médiane des distances au centre. Les clusters supplémentaires constatés ne sont pas pénalisés, car il est possible que les données réelles contiennent des clusters. Les trois algorithmes comparés sont: PCBB, PAM (K-medoid) et GMM (se reporter à la section 2 pour plus de détails). Il est constaté que PCBB surpasse les deux autres méthodes. Ce n'est pas si surprenant quand on considère la nature de ces différents algorithmes de clustering. L'objectif de PAM et de GMM est de fournir une partition complète de l'ensemble de données. Cet objectif n'est pas toujours compatible avec la détection des clusters compacts. PCBB, d'autre part, est conçu spécifiquement pour détecter ce type de clusters.

3.4.2 La pertinence de l'ICP-Isomap pour l'analyse du cortex: données réelles

Une combinaison de trois sillons est choisie pour une étude de l'efficacité de l'approche ICP-Isomap. L'algorithme Isomap sépare mieux les distributions de ces sillons que les

deux autres méthodes. En outre, le sillon le moins variable, le sillon central, est représenté comme un nuage compact.

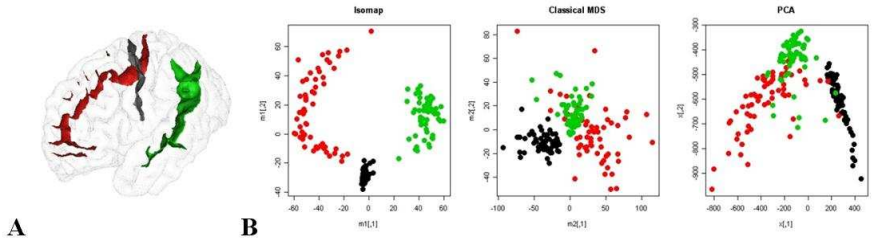


Figure 10 La comparaison des méthodes de réduction de dimension

A: Le sillon central en noir, le sillon temporal supérieur en vert et le sillon cingulaire en rouge

B: De gauche à droite: la distribution en utilisant Isomap, MDS classique et de l'PCA.

3.4.3 Un exemple de clustering sur une région

La Fig. 11 illustre les clusters trouvés sur la région cingulaire en utilisant l'approche enchaînant ICP, isomap et PCBB. Le groupe de sillons utilisé vise à élucider la variabilité de la partie frontale de la face interne du cortex qui cause beaucoup de soucis d'interprétation.

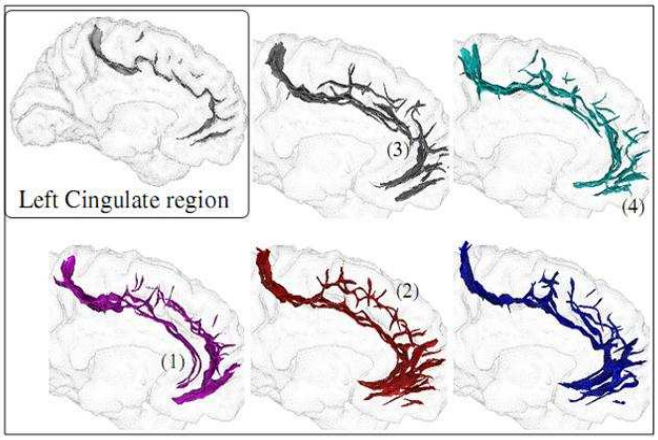


Figure 11 Les clusters trouvés sur la région cingulaire

Un exemple de la région s est affiché en premier. Ensuite, pour chaque motif, trois sujets alignés se superposent afin de mettre en évidence les domaines de stabilité. La région cingulaire gauche est très variable. Les principales caractéristiques permettant d'interpréter ces motifs sont (1) le développement important du sillon appelé « intracingulaire » (peu profond en bas dans le motif violet), (2) le développement important du sillon « paracingulaire » (série de petits plis en haut dans le motif rouge), (3)

les interruptions du sillon cingulaire et (4) la forme de la partie antérieure de la région. La partie postérieure est relativement stable.

3.5 Visualisation des formes à l'aide de SPAMs

Quand il n'y a pas de motifs évidents dans un ensemble de données, ce qui est en fait fréquent, nous aimerions disposer d'une alternative pour étudier la variabilité des plissements individuels. Quand la dimension de l'isomap est réduite à un ou deux, l'organisation des plis peut facilement être visualisée et souvent interprétée. En outre, cette inspection visuelle est importante dans une analyse exploratoire comme la notre pour s'assurer de la qualité des résultats. Par exemple, en utilisant une isomap de dimension un, chaque sillon est représenté par un nombre unique, une coordonnée. En général, lorsqu'on passe d'une extrémité à l'autre de l'axe de l'isomap, la forme des sillons change progressivement. L'inspection visuelle peut souvent suffir à deviner la nature de ce changement graduel de la forme.

Afin de s'assurer que les hypothèse alors formulées sont objectives, nous avons développé une technique permettant de mettre en évidence les changements de forme qui prévalent dans l'organisation de l'isomap. À cette fin, nous avons adapté la stratégie classique de la communauté qui consiste à moyenner des images à travers les individus après les avoirs alignées. On parle de Cartes Anatomiques Statistiques Paramétriques (SPAM, Evans et DL Collins, 1997). L'alignement réalisé pour nos moyennes est celui calculé par l'ICP relativement au sillon le plus neutre de l'ensemble, c'est-à-dire celui qui minimise la distance moyenne à l'ensemble des sillons. De tels sillons « moyens » peuvent être calculés localement en les répartissant régulièrement le long de l'axe Isomap. Pour une position donnée, on calcule une moyenne pondérée de sorte qu'un sillon contribue plus ou moins à la moyenne en fonction de sa distance à la position sélectionnée dans l'isomap. Le poids de cette contribution correspond à une décroissance exponentielle par rapport au carré de la distance à l'emplacement. Un paramètre d'échelle est réglé de sorte que chaque sillon ne contribue essentiellement qu'aux deux SPAM situées directement à sa droite et à sa gauche dans l'isomap. Ainsi, chaque image

moyenne fournit une bonne représentation de la forme des sillons autour de l'emplacement correspondant sur l'axe.

4. Un exemple d'application en neurosciences

Dans cette section, les méthodes développées dans les sections précédentes sont appliquées à un problème de neurosciences : les conséquences éventuelles de la latéralité manuelle dans la forme du sillon central. Trois populations sont étudiées: les droitiers, les gauchers et les gauchers contrariés : des gauchers qui ont été contraints d'utiliser leur main droite pour écrire. Le sillon central est un des sillons les plus stables. Il sépare le cortex moteur du cortex somatosensoriel (Penfield et Boldrey, 1937), (fig. 12A). Il est extrait automatiquement dans l'ensemble des IRM à l'aide du logiciel BrainVisa (Mangin et al, 2004b.). Les algorithmes décrits précédemment permette de résumer la variabilité de la forme du sillon central de la population à travers un axe de dimension 1 fourni par l'isomap.

Deux ensembles de données sont utilisés. Dans le premier ensemble de données, 31 gauchers contrariés sont comparés à 19 droitiers et 16 gauchers. Les individus sont considérés comme des gauchers contrariés si les sujets et leurs parents se rappellent clairement que l'écriture a commencé avec la main gauche à l'école, mais que le sujet a ensuite été forcé d'utiliser la main droite. Les détails sur les participants ont été décrits dans une étude précédente (Klöppel et al., 2010), qui répertorie tous les critères d'inclusion et d'exclusion. Pour corriger un éventuel biais lié au faible nombre de droitiers dans cette première base de données, elle est ensuite combinée avec une seconde base de données contenant essentiellement des droitiers. Cette deuxième série de données est la base d'apprentissage du système actuel de reconnaissance des sillons de BrainVISA (Perrot et al., 2011).

Nous avons d'abord découvert un trait morphologique frappant et très simple du sillon

central, qui peut être décrit comme la position plus ou moins centrée de la « bosse » de main. Cette bosse de la main correspond de manière bien établie au siège de la motricité de la main (Yousry et al ., 1997). (La Fig. 12B illustre la définition de la bosse de la main, la fig13 montre comment la position de l'encoche se déplace le long du sillon central dans l'isomap). Nous avons également observé une deuxième "bosse" sous la bosse de main. Lorsque la position de la bosse de la main se déplace vers le haut, la deuxième bosse inférieure devient plus imposante. Des résultats très similaires sont obtenus lorsque le premier jeu de données est combiné avec la deuxième série de données, comme illustré sur la fig14.

Les deux hémisphères sont différents en termes de forme du sillon central. Le sillon central gauche des trois populations (les gauchers, les droitiers et les gauchers contrariés) présente généralement le motif à « deux bosses », c'est à dire la bosse de la main plus une deuxième bosse en dessous. Le sillon central droit, en revanche, correspond plutôt au motif « une-bosse », contenant une bosse de la main de premier plan, la deuxième bosse inférieure étant quasi inexistante. (cf fig15, pour l'analyse sur le premier jeu de données, et fig16 pour l'analyse sur le jeu de données combiné).

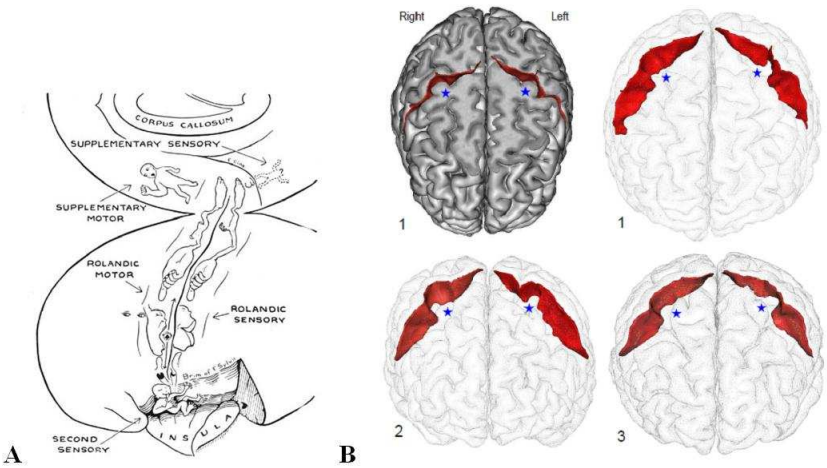


Figure 12 L'homonculus et la bosse de la main

A: L'homonculus de Penfield: figurines dessinées sur l'hémisphère gauche (Penfield et Jasper, 1954)

B: Les sillons centraux gauche et droit de trois sujets sont surlignés en rouge ; la position de la bosse de la main est marquée avec des étoiles bleues.

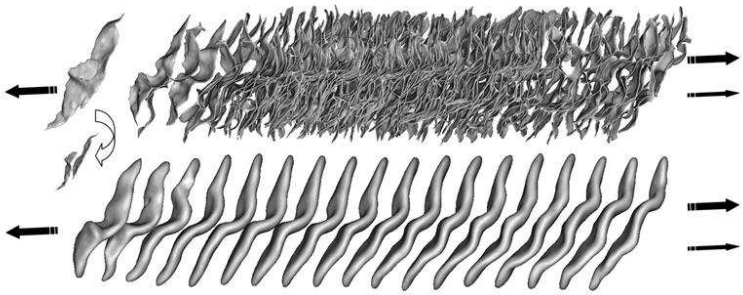


Figure 13 positions de la bosse de main selon l'axe Isomap.

(HAut) Tout d'abord, tous les sillons ont été alignés sur un sillon modèle afin d'obtenir des orientations similaires dans l'ensemble. Ensuite, chaque sillon a été translaté dans l'axe antéro-postérieure du cerveau proportionnellement à sa coordonnée dans l'Isomap. Notez le sillon atypique situé à l'extrême gauche de l'axe correspondant à un sillon central ainterrompu (ce sillon est également représenté à partir d'un point de vue différent ci-dessous pour rendre l'interruption plus claire). (Bas) Des moyennes locales des sillons centraux ont été calculées à des positions régulièrement espacées de manière à préciser la variabilité de forme codée par l'axe Isomap.

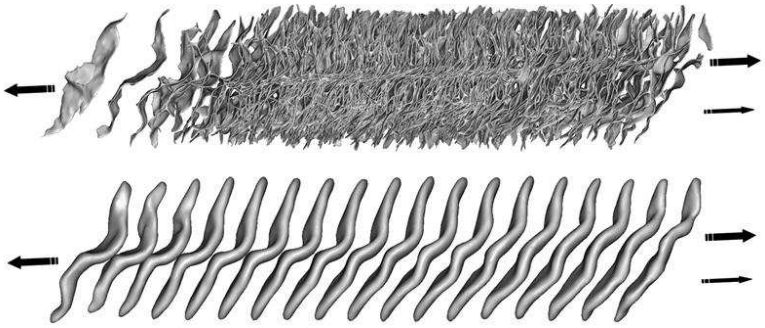


Figure 14 positions de la bosse de main d'un ensemble de données combinée

Analogue à la figure précédente avec le jeu de données combiné. Notez qu'un second sillon interrompu a été poussé à l'extrême droite de l'axe.

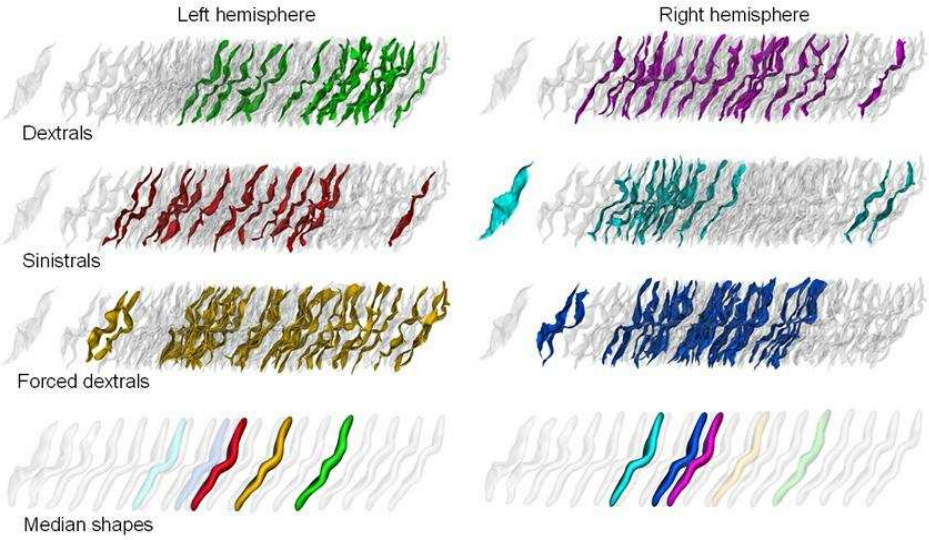


Fig 15 localisation de chaque population le long de l'axe Isomap

La Fig 13 est représentée en transparence. En bas: Les moyennes locales calculées pour les positions médianes de chaque population. Code couleur: Cyan: gaucher, hémisphère droit; Bleu: gaucher contrarié, hémisphère droit; Violet: droitier, hémisphère droit; Rouge: gaucher, hémisphère gauche; Jaune: gaucher contrarié, hémisphère gauche; Vert: droitier, hémisphère gauche

Que l'ont considère les gauchers ou les gauchers contrariés, la forme de leur sillon gauche diffère significativement de celle des droitiers (droitiers vs gauchers: $p = 0,002$; droitier vs gauchers contrariés: $p = 0,02$). En ce qui concerne l'hémisphère droit, le sillon central des gauchers contrariés a tendance à être différent de celui des gauchers ($p = 0,06$), et les droitiers ont tendance à être différent des gauchers ($p = 0,15$) . On peut également noter que le sillon de l'hémisphère dominant des gauchers et des droitiers s'écarte de la zone neutre centrale dans deux directions opposées ($p=0,0002$).

Des études antérieures avaient que des asymétries de la taille du sillon central permettaient de distinguer les gauchers des droitiers. Toutefois, personne ne savait si cette tendance devait être attribuée à des facteurs innés ou à la plasticité induite par les expériences et les influences environnementales. Dans notre étude, nous avons montré que les mécanismes innés sculptent la forme du sillon central, alors que l'expérience ne fait que modifier sa longueur. Chez les gauchers contrariés, en effet, l'asymétrie de la longueur du sillon central que nous avons calculée est typique des droitiers. En revanche,

les résultats évoqués ci-dessus montrent que l'emplacement de la «bosse de main » dans le sillon gauche, qui s’est avérée corrélée avec la latéralité manuelle, ne change pas chez les gauchers contrariés. Nous avons même mis en évidence que cet stabilité de la forme était encore plus forte chez les huit gauchers contrariés dont au moins un des parents était gaucher. Ainsi, la morphologie du cortex chez l'adulte combine des informations sur l’inné et sur l’acquis.

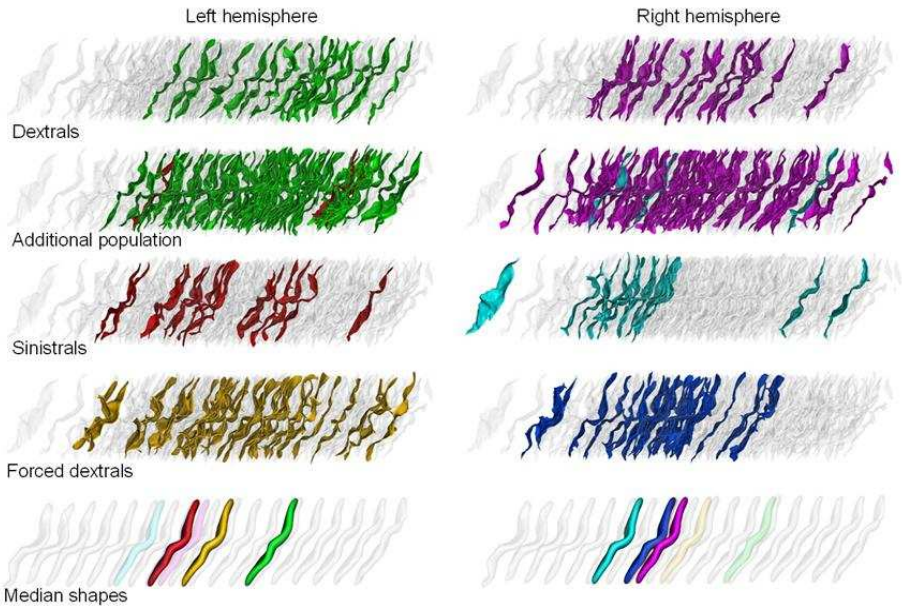


Fig 16 localisation des populations le long de l'axe Isomap (pour les deux bases combinées)
Le code de couleur utilisé pour la base supplémentaire indique les gauchers et les droitiers. Code couleur: Cyan: gaucher, hémisphère droit; Bleu: gaucher contrarié, hémisphère droit; Violet: droitier, hémisphère droit; Rouge: gaucher, hémisphère gauche; Jaune: gaucher contrarié, hémisphère gauche; Vert: droitier, hémisphère gauche

5. Discussion sur l'analyse des plissements

Dans cette section, nous évoquons le potentiel de l’analyse des motifs des sillons et des nouvelles représentations des données fournies par l’isomap. Nous indiquons également

quelques perspectives. Quelques pages de notre futur dictionnaire sont présentées pour finir.

De nombreux facteurs jouent un rôle important dans l'analyse des plissements du cortex. En particulier, le choix des sillons ou des régions d'intérêt, les descripteurs de forme, la mesure de similarité, l'algorithme de clustering et l'ensemble de données utilisé pour l'analyse. Il faudra encore du temps pour peaufiner chacun de ces choix. La limpidité des résultats complètement inattendus mis en évidence pour le sillon central met néanmoins en évidence le potentiel de notre démarche. Il faut noter que l'équipe de Neurospin a exploré l'impact de la latéralité sur le sillon central pendant des années en passant complètement à coté des différences de forme que nous avons découvertes. L'assymétrie mise en évidence est elle aussi complètement inédite. Les quelques pages du dictionnaire présentées par la suite mette en évidence d'autres caractéristiques de formes de nature très similaires à celle découverte sur le sillon central. Nous espérons donc qu'elles pourront également être associées à des particularités cognitives voire à des pathologies développementales.

Outre ce type d'applications aux neurosciences, nous avons également fait émerger que les descriptions de la variabilité très synthétiques fournies par nos motifs et les premières dimensions de l'isomap permettent d'avancer sur les aspects propres à la nomenclature des sillons. Nous avons à plusieurs reprise observé que certains motifs étaient induits par des faiblesses de la nomenclature actuelle, mais aussi des erreurs manifestes lors de l'utilisation de cette nomenclature pour étiqueter des configurations atypiques. Certains motifs résultent quant à eux d'insuffisances du système de sur-segmentation des plissements utilisé par brainVISA avant l'étiquetage des sillons. L'ensemble de ces observations va donc permettre d'améliorer la nomenclature actuelle. Il est également probable que la description synthétique de la variabilité fournie par les premières dimensions de l'isomap puisse être injectée dans l'a priori Bayésien utilisé dans le dernier modèle conçu pour reconnaître automatiquement les sillons (Perrot et al., Med. Image Analysis, 2011).

6. Exemple de dictionnaire des motifs des plissements

Nous illustrons ici le type d'information que nous pensons utiliser pour réaliser notre dictionnaire, en nous focalisant sur quelques régions exemples : le gyrus précentral, la région du sillon cingulaire, la région du sillon temporal supérieur, l'aire de Broca, et des régions frontales supérieures. Nous évoquons également la possibilité de combiner plusieurs régions.
Il s'agit bien sur d'une étude préliminaire qui pourra être systématisée pour l'ensemble du cortex.

6.1 La région précentrale

La fig17 montre que l'isomap révèle que le premier mode de variabilité de la forme du gyrus précentral réside dans l'orientation du sillon précentral intermédiaire.

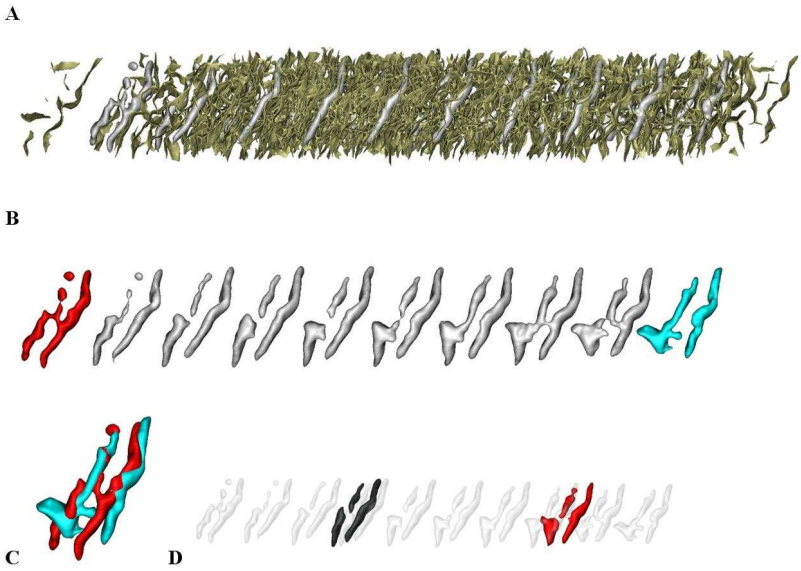


Figure 17 Variabilité des plissements du gyrus précentral

A: L'isomap du gyrus précentral. Les sillons et les SPAM locales sont associés afin d'indiquer la localisation de ces dernières dans la distribution. Notez que la distribution est beaucoup moins dense

lorsqu'on s'éloigne de la partie centrale, la forme des SPAMs les plus extrêmes est donc à considérer avec précaution.

B. Les SPAMs locales Les formes aux deux extrémités sont de couleur bleu et rouge respectivement. C: Les SPAMs extrêmes sont superposées. D: Les SPAMs des deux motifs détectés par PCBB localisés sur l'axe Isomap

6.2 La région cingulaire

Les résultats de l'analyse de cette région sont présentés dans la figure 18. Le premier mode de variabilité est capturé par l'Isomap. Il s'agit selon nous de l'ampleur relative des développements des sillons cingulaires et paracingulaires. Les SPAMs des trois motifs trouvés par PCBB sont également représentés.

6.3 Le sillon temporal supérieur (STS)

Le premier mode de variation capturé correspond à «l'ouverture» des deux branches postérieures du STS (fig19). Les implications de cette variabilité en termes de fonction et le comportement mériteraient d'être approfondies.

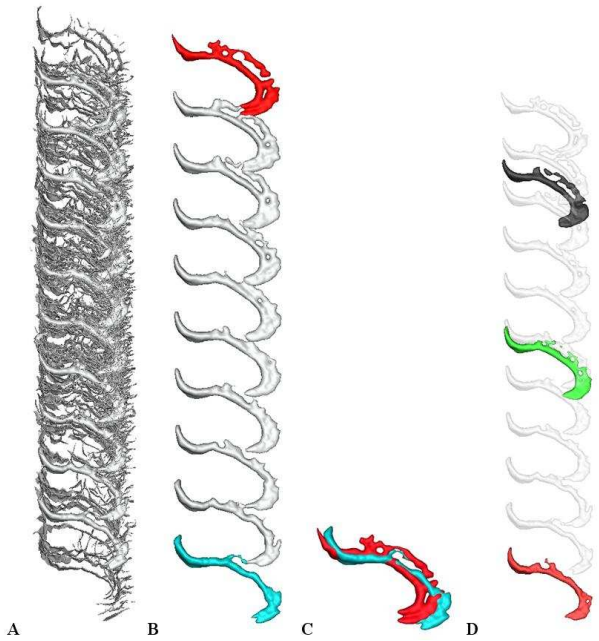


Figure 18 Variabilité des plissements de la région cingulaire (cette figure correspond au mode de présentation introduit dans la figure précédente)

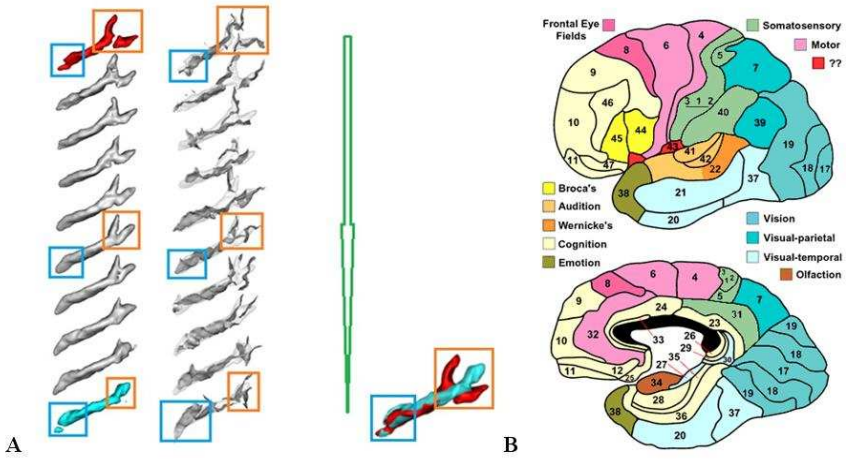


Figure 19 Variabilité des plissement du sillon temporal supérieur (STS)

7. Conclusion

Comme l'illustre notre étude, la description de la variabilité des formes des plissements du cortex peut être systématisée. Notre démarche permet d'obtenir des résultats sur de grands ensembles de données. Elle va permettre de prolonger le travail initié par Ono (ref) à une échelle susceptible d'engendrer une description exhaustive des modes de variabilités importants. Nous disposons déjà à Neurospin de plusieurs milliers de cerveaux utilisables avec cet objectif. La multiplication des bases de données publiques devrait rapidement encore étendre cette base.

Les facteurs génétiques (lorsque des motifs sont liés à des maladies génétiques, des jumeaux, des familles) dans le processus de plissement vont pouvoir être évalués. Réduire la complexité de la variabilité de la forme d'un sillon à quelques dimensions va considérablement clarifier les études de la plasticité, du développement et du vieillissement. Nous espérons faire émerger de la forme des sillons un nouveau potentiel en termes de biomarqueurs. Enfin, nous espérons intégrer dans l'analyse des plissements

du cortex des informations issues des faisceaux de fibres et des données. Le développement cérébral pourra être étudié systématiquement dans les études longitudinales et révéler la chronologie associée à la différentiation des motifs.

APPENDIX B: ETAT DES PUBLICATIONS

1. Articles dans des revues avec comité de lecture

Soumises :

Sun, Z.Y., Klöppel, S., Riviere, D., Perrot, M., Frackowiak, R., Siebner, H., Mangin, J.F., 2011. Innate and experience-dependent brain markers of left-handedness. Cerebral Cortex submitted for review.

2. Communications avec actes

2 communications à MICCAI, la conférence la plus sélective du domaine de l'imagerie médicale (1/3 des soumissions acceptées) :

Sun, Z.Y., Perrot, M., Tucholka, A., Riviere, D., Mangin, J.F., 2009. Constructing a dictionary of human brain folding patterns. MICCAI 12, Londres, pp 117-124, LNCS, Springer Verlag

Sun, Z.Y., Riviere, D., Poupon, F., Regis, J., Mangin, J.F., 2007. Automatic inference of sulcus patterns using 3D moment invariants. MICCAI 10, Brisbane, pp 515-522, LNCS, Springer Verlag

Présentation orale, conférence IEEE International Symposium on Biomedical Imaging

Sun, Z.Y., Riviere, D., Duchesnay, E., Thirion, B., Poupon, F., Mangin, J.F., 2008. Defining cortical sulcus patterns using partial clustering based on bootstrap and bagging. Proc. IEEE ISBI 5, Paris, 1629-1632.

3. Communications sans actes

Sun, Z.Y., Riviere, D., Poupon, F., Regis, J., Mangin, J.F., 2007, Automatic Inference of Sulcus Patterns, Human Brain Mapping, Chicago

Sun, Z.Y., Riviere, D., Poupon, F., Regis, J., Mangin, J.F., 2007, Mining Sulcal Folding Patterns, Human Brain Mapping, San Francisco