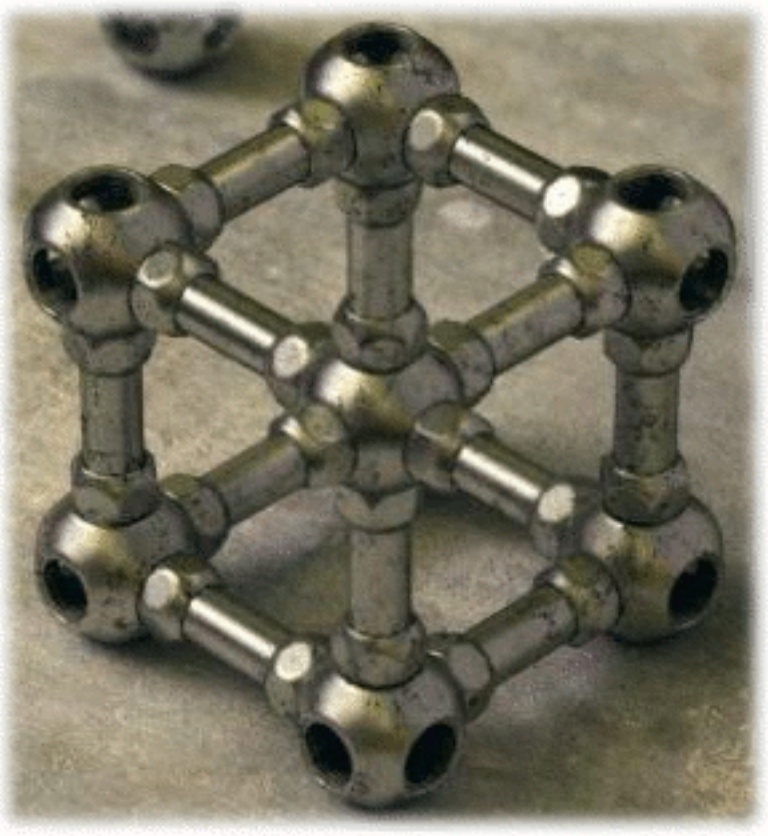


Exploring the knowledge contained in neuroimages: statistical discriminant analysis and automatic segmentation of the most significant changes [1]

Author(s):

Paulo Santos
Carlos Thomaza
Danilo dos Santosa
Rodolpho Freirea
João Sato
Mario Louzã
Paulo Sallet
Geraldo Busatto
Wagner Gattaz

[1]This work was supported by Fapesp Project LogProb, grant 2008/03995-5, São Paulo, Brazil.

Exploring the knowledge contained in neuroimages: statistical discriminant analysis and automatic segmentation of the most significant changes

Paulo Santos[*,a,1], Carlos Thomaz[a], Danilo dos Santos[a], Rodolpho Freire[a], João Sato[b], Mario Louzã[c], Paulo Sallet[c], Geraldo Busatto[c], Wagner Gattaz[c]

[a] *Electrical Engineering Department*
Centro Universitário da Fundação Educacional Inaciana (FEI)
Av. Humberto de A. Castelo Branco, SBC-SP, Brazil.
[b] *Universidade Federal do ABC,*
Rua Santa Adélia, 166. Bairro Bangu. Santo André - SP, Brazil
[c] *Instituto de Psiquiatria, Universidade de São Paulo,*
R. Dr. Ovídio Pires de Campos, 785 - São Paulo, Brazil

Summary

Objective: The aim of this article is to propose an integrated framework for extracting and describing patterns of disorders from medical images using a combination of linear discriminant analysis and active contour models.

Methods: A multivariate statistical methodology was first used to identify the most discriminating hyperplane separating two groups of images (from healthy controls and patients with schizophrenia) contained in the input data. After this, the present work makes explicit the differences found by the multivariate statistical method by subtracting the discriminant models of controls and patients, weighted by the pooled variance between the two groups. A variational level-set technique was used to segment clusters of these differences. We obtain a label of each anatomical change using the Talairach atlas.

Results: In this work all the data was analysed simultaneously rather than assuming *a priori* regions of interest. As a consequence of this, by using active contour models, we were able to obtain regions of interest that were emergent from the data. The results were evaluated using, as gold standard, well-known facts about the neuroanatomical changes related to schizophrenia. Most of the items in the gold standard was covered in our result set.

Conclusions: We argue that such investigation provides a suitable framework for characterising the high complexity of magnetic resonance images in schizophrenia as the results obtained indicate a high sensitivity rate with respect to the gold standard.

Key words: Multivariate statistical analysis, neuroimage, deformable models, schizophrenia research.

1. Introduction

Schizophrenia is a mental disorder characterised by symptoms of psychosis (*e.g.*, delusions and hallucinations), apathy and social withdrawal, as well as cognitive impairment [1]. Although the causes of schizophrenia are unknown, both genetic [2] and environmental factors (including biological - *e.g.*, prenatal infection and obstetric complications - and psychosocial factors) appear to play a role in its etiology. These factors are not sufficient to the emergence of schizophrenia, probably exerting their effect in a stress-vulnerability model of the disease [3].

The established illness is associated with structural and functional brain abnormalities, mainly in prefrontal and temporal lobes, findings being largely due to recent advances in *in vivo* Magnetic Resonance Imaging (MRI) techniques [4]. However, none of the brain abnormalities found in schizophrenia is characteristic of the disease, and no neuroanatomical finding alone has a diagnostic value for schizophrenia. It is conceivable that the abnormality in brain development is not restricted to a determined brain structure, being rather diffuse, affecting the different brain structures simultaneously. In an attempt to overcome these difficulties, in a previous study [5] we studied 12 Computed Tomography (CT) parameters in 30 schizophrenic patients and 30 sex- and age-matched controls, and evaluated the data simultaneously through multidimensional scaling (MDS). MDS offers a graphic representation in which subtle deviations in the different CT parameters can be detected, independently of predetermined criteria for the definition of abnormalities. MDS distinguished 13 patients from the controls as having deviant values in one or more CT parameters. Five of these patients were first-onset schizophrenics. Our results suggest that the use of multidimensional techniques may improve the sensitivity of neuroanatomical data to identify more precisely schizophrenic patients and to provide information of the possible influence of the structural brain abnormalities upon the course and prognosis of the disorder.

The purpose of this paper is to investigate multidimensional techniques on neuroanatomical data through the application of multivariate statistical methods to extract the most statistically significant differences between controls and patients and to segment automatically these findings. In order to report these developments, the present paper is organised in the following way: Section 2 describes some pieces of related research; Section 3 presents the statistical methods used to extract discriminant features that best differentiate neuroimages from healthy controls and patients with schizophrenia; Section 4 describes some results from the application of the methods discussed in Section 3 and shows how the obtained discriminant information can be automatically segmented using active contour models; Section 5 presents an automatic procedure for labelling the output obtained by the methods presented with the names of the related brain structure. The results of this procedure are evaluated according to the extent of which well-known findings about neuroanatomical changes related to schizophrenia were obtained automatically by the proposed methods (Section

*Corresponding author.
Email address: santosp@ieee.org (Paulo Santos)
[1]Phone and fax: +55 11 4353 2910 (Paulo Santos)

6). Discussions are drawn in Section 7 and Section 8 concludes this paper.

2. Related Research

The medical image analysis community has been mostly concerned with the segmentation of particular anatomic structures, even considering research on *bottom up* (i.e., autonomous and model-free) segmentation, as pointed out in various literature surveys [6–8]. From this literature we have to refer to some articles that propose similar steps to those discussed in the present paper. The work reported in [9] presents the results of the segmentation of the human lungs from MR images using a combination of neural networks and active contours, the perceptron is trained to classify a pixel as belonging to the boundary of the target structure or not. The resulting classification is then used as input to an active contour model. In [10] a combination of a C-means clustering algorithm and active contour is used to segment the thalamus from MRI scans, whereby C-means provides the input to the active contour segmentation. Adaptive clustering and active contours are also used in [11] to segment the hippocampus region in brain MRIs. Similarly, an expectation/maximisation segmentation is used in [12] to initialise an active model procedure that segments the brain tissue from MR images of the head.

However, to the best of our knowledge, the work reported here is the first to propose the combination of multivariate statistical analysis and active models to search for statistically significant regions (in contrast to particular anatomical structures) in the neuroimages that best differentiate MRI of schizophrenic patients from those of healthy controls.

3. Extracting discriminative information

The increasing resolution of 3D anatomical and functional images nowadays has allowed the visualisation of neuroanatomical structures of the human brain with impressive detail. For example, the widely used method of MRI gives good soft-tissue contrast with high resolution, that is, commonly less than 1mm [13]. However, depending on the brain abnormality and its progression, neuroanatomical changes may be too subtle, diffuse, or topologically complex to be detected by simple visual inspection [13]. Thus, in the last years, a considerable amount of effort has been devoted to the design of computational methods for morphological analysis of the human brain. Traditionally, such analysis of brain images has been based either on the definition of regions of interest given some *a priori* hypothesis or on voxel-wise measurements with little prior knowledge [14, 15]. In practice, these methodologies have shown yet limitations in their ability to identify previously unexplored relationships between control and patient groups.

In recent years, statistical pattern recognition methods have been proposed to extract and analyse morphological and anatomical structures of MR images between images of a reference group and the group under investigation [16, 15, 17–19]. Most of these techniques work in high-dimensional spaces of particular features such as shapes or statistical parametric maps and have overcome the difficulty of dealing with the inherent high dimensionality of medical data by analysing segmented structures individually or performing hypothesis tests on each feature separately. As a consequence, changes that are

relatively more widely distributed and involve simultaneously several structures of the pattern of interest might be difficult to detect, despite the possibility of some statistical learning methods [18, 19] of extracting multivariate differences between image samples of patients and controls.

The main goal of the methodology described in this section is to analyse all the data simultaneously rather than segmented versions separately or feature-by-feature. This approach has been specially designed for extracting discriminative information from high dimensional, small sample size problems [20–23]. In the next subsections, we discuss our general multivariate statistical approach to identify and analyse the most discriminating hyperplane separating two groups and explain how we have quantified the statistical significance of the multivariate group-differences found.

3.1. Multivariate Statistical Discriminant Model

In the generic discrimination problem, where the training sample consists of the class membership and observations for N samples, the outcome of interest falls into g classes and we wish to build a rule for predicting the class membership of an observation based on n variables or features. However, statistical discriminant methods are suitable not only for classification but also for characterisation of differences between a reference group of samples and the group under investigation. For example, in clinical diagnosis it might be helpful to determine on the original space of images the discriminant information captured (or used) by a statistical classifier to separate MRI samples of patients and controls.

However, before we can perform a multivariate pattern recognition analysis on the MR images, we need first to map all images into a common atlas coordinate system. This pre-processing step, called spatial normalisation or image registration, is essential because the construction of the multivariate statistical model relies on the correspondences of the image features when comparing patterns across samples. In MRI brain analysis, this procedure has essentially two goals [21]: (a) to reduce variability due to size, position and orientation of the brain shape [24] and (b) to reduce variability due to differences in the brain shape [24]. Each registered image can then be treated as a point in an n-dimensional space, where n is the total number of voxels. In this work, the coordinates of this point represent the intensity value of each voxel. For this feature representation to make sense in classification problems, we are making implicitly the assumption that two similar images correspond to two close points in the high dimensional original image space. I.e., the effectiveness of the extracting information technique would be determined by how well the samples from different classes can be separated.

The n-dimensional resulting images are then projected from the original vector space to a lower dimensional space using the well-known Principal Component Analysis (PCA) [25]. There are a number of reasons for using PCA to reduce the dimensionality of the original MR images. PCA is a linear transformation that is not only simple to compute and analytically tractable but also extracts a set of features that is optimal with respect to representing the data back into the original domain. Moreover, using PCA as an intermediate step reduces dramatically the computational and storage requirements for the subsequent linear discriminant covariance-based method. Since

in our application of interest the number of training samples N (or images) is much smaller than the number of features n (or voxels), it is possible to transform data in a way that the samples occupy regions that are as compact as possible in a lower dimensional feature space, with fewer degrees of freedom to estimate. Although much of the sample variability can be accounted for by a smaller number of principal components, that is, we could retain only the first two or three principal components, there is no guarantee that such additional dimensionality reduction will not add artifacts on the images when mapped back into the original image space. Since one of our main concerns here is to map the subsequent classification results back to the image domain for a further description of the information content of the findings, we must be certain that any modification on the images, such as blurring or subtle differences, is not related to the PCA dimensionality reduction. For example, Figure 1 [21] illustrates on the top a reference image (shown on the left) reconstructed using several principal components and on the bottom the corresponding differences between these reconstructions with respect to the original image. The values in parentheses represent, respectively, the number of principal components used and the corresponding total variance explained. We can see that, even when we use a set of principal components that represent more than 90% of the total sample variance we still have subtle differences between the reconstructed image and the original one. Therefore, in order to reproduce the total variability of the samples, we have composed the PCA transformation matrix by selecting all the principal components with non-zero eigenvalues, that is, the number m of retained principal components is $m = N - 1$.

A maximum uncertainty linear discriminant analysis (MLDA) approach [26] has been applied next to find the best linear discriminant features on that PCA subspace. The primary purpose of linear discriminant analysis (LDA) is to separate data samples of distinct groups by maximising their between-class separability while minimising their within-class variability. It is well known, however, that the performance of the standard LDA can be seriously degraded if there are only a limited number of total training observations N compared to the dimension of the feature space. Since the within-class scatter matrix S_w is a function of $(N - g)$, or fewer linearly independent vectors, where g is the number of groups, its rank is $(N - g)$ or less. In the current situation, where the number of training patterns is small with respect to the number of features, S_w might be singular or unstable and the standard LDA cannot be used to perform the task of the classification stage.

The main idea of the MLDA approach is to stabilise the within-class scatter matrix S_w with a multiple of the identity matrix. It is based on the maximum entropy covariance selection method that Thomaz et al. [27, 26] have developed to improve classification performance on limited sample size problems [20–23]. Since the estimation errors of the non-dominant or small eigenvalues are much greater than those of the dominant or large eigenvalues, the MLDA algorithm expands the smaller (less informative) eigenvalues of S_w and keeps most of its larger eigenvalues unchanged. It is a straightforward method that overcomes both, the singularity and the instability of the within-class scatter matrix when LDA is applied in limited sample, high dimensional

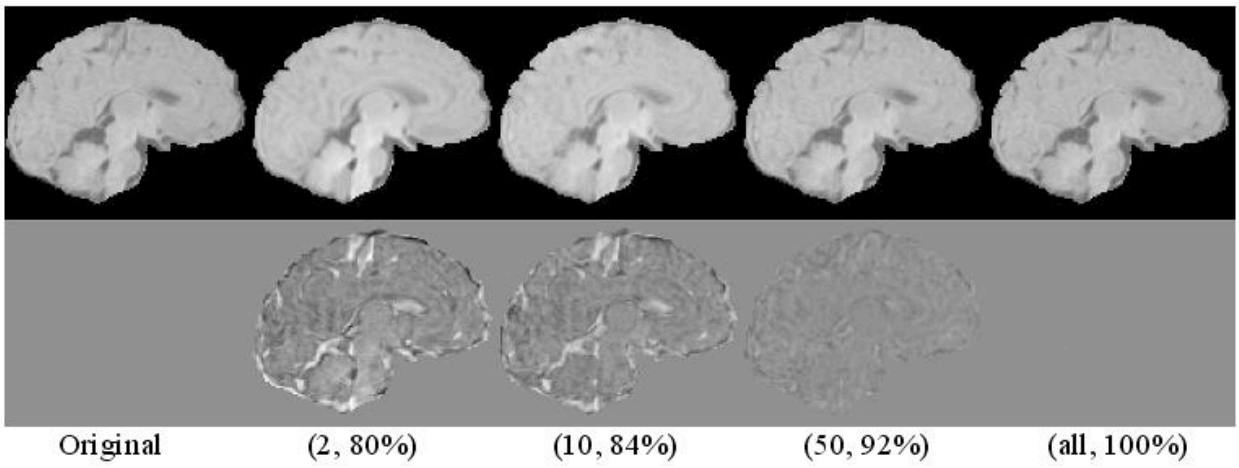


Figure 1: An example of the reconstruction of a reference image (shown on the left) using several principal components. The bottom row illustrates the corresponding differences between the reconstructions and the original image. The number of principal components used, and the corresponding total sample variance explained by each one of the sets of principal components, are shown in parentheses [21].

problems.

We can divide the design of the PCA+MLDA multivariate statistical discriminant model into two main tasks: classification (training and test stages) and visual analysis. In the classification task the principal components and the maximum uncertainty linear discriminant vector are generated. As illustrated in Algorithm 1 below, first, a training set is selected (T) and the average image vector of all the training images is calculated ($\bar{T}$) and subtracted from each pre-processed n-dimensional image vector (resulting on the matrix Z in Algorithm 1). Then the training matrix Z, composed of zero mean image vectors, is used as input to compute the PCA transformation matrix P. The columns of this transformation matrix are eigenvectors, not necessarily in eigenvalues descending order. Note that we have retained all the PCA eigenvectors with non-zero eigenvalues, that is, $m = N-1$, and P is a $n \times m$ matrix. The zero mean image vectors are projected on the principal components and reduced to m-dimensional vectors representing the most expressive features of each one of the pre-processed n-dimensional image vector (F). Afterwards, the $N \times m$ data matrix (F) is used as input to calculate the MLDA discriminant eigenvector (L). Since we are assuming only two classes to separate, that is, $g = 2$, there is only one MLDA discriminant eigenvector. The most discriminant feature (MF) of each one of the m-dimensional vectors is obtained by multiplying the $N \times m$ most expressive feature matrix by the $m \times 1$ MLDA linear discriminant eigenvector. Thus, the initial pre-processed training set consisting of N measurements on n variables, is reduced to a data set consisting of N measurements on only 1 most discriminant feature.

The other main task performed by this two-stage multivariate approach, and the one explored further in this work, is to visually analyse the most discriminant feature found by the PCA+MLDA statistical model. Any point on the most discriminant feature space can be converted to its

Algorithm 1 PCA+MLDA statistical discriminant method

N = number of images
n = number of voxels of each image
T = matrix (N x n) with training images, each row is an image vector

$A \Leftarrow \bar{T}$ {Average image (1xn)}
$Z \Leftarrow T - A$ {Zero mean sample set (Nxn)}
$P \Leftarrow PCA(Z)$ {Principal component of Z (nxm)}
$F \Leftarrow Z * P$ {Most expressive features (Nxm)}
$L \Leftarrow MLDA(F)$ {Linear discriminant eigenvector (mx1)}
$MF \Leftarrow L * F$ {Most discriminant features (Nx1)}

corresponding n-dimensional image vector by simply: (1) multiplying that particular point by the transpose of the linear discriminant vector previously computed (L^T); (2) multiplying its m most expressive features by the transpose of the principal components matrix (P^T); and (3) adding the average image (A) calculated in the training stage to the n-dimensional image vector. Therefore, assuming that the clouds of classes follow a multidimensional Gaussian joint distribution, and applying limits to the variance of each cloud, such as $\pm 3\sigma_i$, where σ_i is the standard deviation of each group $i \in \{1, 2\}$, we can move along this most discriminant feature and map the result back into the image domain. This mapping provides a sequence of images based on a statistical interpretation of the classification experiments and describes the discriminant information captured by the PCA+MLDA method to separate the groups.

To illustrate the performance of the knowledge extraction approach, we present in this subsection some results on facial expression analysis. Since this application does not require a specific knowledge to understand the differences between the groups, it seems a useful example to discuss the main idea of the multivariate statistical discriminant method. We have used frontal images of a face database maintained by the Department of Electrical Engineering of FEI to carry out the experiments. The FEI face database contains a set of face images taken between June 2005 and March 2006 at the Artificial Intelligence Laboratory in São Bernardo do Campo, São Paulo, Brazil, with 14 images for each of 200 individuals - a total of 2800 images. All images are colourful and taken against a white homogeneous background in an upright frontal position with profile rotation of up to about 180 degrees. Scale might vary about 10% and the original size of each image is 640x480 pixels. All faces are mainly represented by subjects between 19 and 40 years old with distinct appearance, hairstyle, and adorns[2].

To minimise image variations that are not necessarily related to differences between the faces, we first aligned all of the frontal face images to a common template so that the pixel-wise features extracted from the images correspond roughly to the same location across all subjects. In this manual alignment, we have randomly chosen the frontal image of a subject as template and the directions of the eyes and nose as a location reference. For implementation convenience, all the frontal images were then cropped to the size of 360x260 pixels and converted to 8-bit grey scale. Since the number of subjects is equal to 200 and each subject has two frontal images (one with a neutral or non-smiling expression and the other with a smiling facial expression), there are 400 images to perform the experiments. Therefore, $N = 400$, $n = 360 \times 260 = 93600$,

[2]This database is publicly available for download on the following site http://www.fei.edu.br/∼cet/facedatabase.html.

and the coordinates of each n-dimensional image sample represents a value in the range of 0 and 255.

Figure 2 shows the most discriminant features for the facial expression experiments. It displays the image regions captured by the multivariate knowledge extraction approach that change when we move from one side (left, group 1 of non-smiling samples) of the dividing hyper-plane to the other (right, group 2 of smiling samples), following limits to the standard deviation and mean of each sample group. As can be seen, the multivariate discriminant hyper-plane effectively extracts the subtle facial expression changes, showing exactly what we should expect intuitively from a face image when someone changes their expression from non-smiling to smiling. In fact, it is possible to note that the knowledge extraction approach has predicted a facial expression not necessarily present in our corresponding facial expression training set, that is, the "definitely non-smiling" or may be "angry" status and the "definitely smiling" or may be "happy" status represented respectively by the image models $-3\sigma_1$ and $+3\sigma_2$ in Figure 2.

The output of the PCA+MLDA extraction approach are models of images derived from the original image data set. The next subsection attempts to quantify the statistical significance of the models representing group samples. In order to simplify our terminology, next subsection uses the term *image model* to the output of the PCA+MLDA process and the term *original data images* to refer to the items in the original image dataset.

3.2. Statistical significance of the differences between groups

The PCA+MLDA multivariate discriminant approach provided a model of the most distinguishing features between two groups, however it does not quantify explicitly the change in each feature from any model in one group to any model in the other group, or how significantly relevant is this change with respect to the original data.

In order to quantify this level of change between two particular image models, we decided to subtract the extreme cases of group 1 and group 2 weighted by a factor that takes into account the pooled variance between the two groups. Therefore, given x_1 representing the "definitely group 1" image model, x_2 the "definitely group 2" image model (as image models $-3\sigma_1$ and $+3\sigma_2$ in Figure 2), N_1 the number of samples of group 1, N_2 the number of samples of group 2, and $\sigma^2 = \frac{(N_1-1)*\sigma_1^2+(N_2-1)*\sigma_2^2}{(N_1+N_2-2)}$ the pooled variance between the two groups, the differences are represented in the matrix *diff* obtained by the equation (1) below:

$$diff = \frac{x_1 - x_2}{\sqrt{\frac{\sigma^2}{N_1} + \frac{\sigma^2}{N_2}}}. \quad (1)$$

The extreme image models were chosen, instead of the mean models of each group, since the former represent the most differentiating features between the groups, as given by the PCA+MLDA multivariate approach. By making explicit the changes between the extremes of two groups we are able to verify the extent of which the results of the multivariate statistical approach (on neuroimages from patients and controls) agree with those presented in meta-analyses about anatomical changes related

Figure 2: Interpretation and reconstruction of the knowledge extraction approach for the facial expression experiments. From left (group 1 of non-smiling samples) to right (group 2 of smiling samples): $[-3\sigma_1, mean_1, +\sigma_1, boundary, -1\sigma_2, mean_2, +3\sigma_2]$.

to schizophrenia (as we shall see later in this paper).

To make explicit the most statistically significant changes we selected the features in *diff* that were greater than a threshold value. Considering the facial expression experiments described in the previous subsection, Figure 3 illustrates[3] the spatial distribution of the intensity changes when using the statistical extremes described by each sample group superimposed on a reference image. In this facial expression example, we considered a difference important if its absolute value exceeds at least one pooled standard deviation. In the picture, the colour-scale shows relative intensity change as a range of this thresholding. In yellow/red the marking on the image are brighter in the smiling samples compared to the non-smiling ones. Analogously, the areas in green/blue shows regions of relative darkness in the smiling samples compared to the non-smiling ones.

We can see clearly that, by exploring the separating hyper-plane found by the multivariate discriminant approach and quantifying its most significant changes, we are able to identify features that are most discriminant between the smiling and non-smiling

Figure 3: Effect size of the multivariate statistical differences comparing the intensity values described by the smiling and non-smiling image models.

[3]For colour images, please see the electronic version of the paper.

group samples, such as eyes, shadow, cheek, upper lip and mouth regions (as can be observed on Figure 3).

4. Multivariate statistical differences between controls and schizophrenia patients

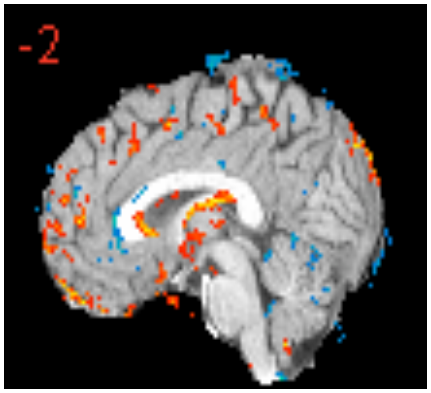
To analyse the most discriminant multivariate differences found on MRI samples of the adult human brain (considering controls and schizophrenic patients), we have used a data set that contains images of 43 patients with schizophrenia and 25 health controls. All these images were acquired using a 1.5T Philips Gyroscan S15-ACS MRI scanner (Philips Medical Systems, Eindhoven, The Netherlands), including a series of contiguous 1.2mm thick coronal images across the entire brain, using a T1-weighted fast field echo sequence (TE = 9ms, TR = 30ms, flip angle 30^o, field of view = 240mm, 256 x 256 matrix). All the images were reviewed by a MR neuro-radiologist. A previous study using this dataset was published in [28]. Ethical permission for this study was granted by the Ethics Committee of the *Hospital das Clínicas*, University of São Paulo Medical School, São Paulo, Brazil.

As stated in the previous section (more specifically in subsection 3.1), before we can perform a multivariate discriminant analysis on the MR images, we need to map all images into a common atlas coordinate system. The images were spatially normalised with the standard Statistical Parametric Mapping (SPM, version SPM2)[29] T1-MRI template [30], based on 152 health subjects from the Montreal Neurological Institute (MNI), using as matching criterion the residual sum of squared differences [31]. Such spatial normalisation step was restricted to linear 12-parameter affine transformations, in order to minimise confounding effects of the original data caused by different positions of the subjects' heads when acquiring the corresponding MR images. The normalised images were then resliced using tri-linear interpolation to a final voxel size of 2 x 2 x 2 mm^3 and final resolution of 91 x 109 x 91. As a final preprocessing step, an automated brain extraction procedure was performed in all images using the Steve Smith's method [32] available in the MRIcro software [33].

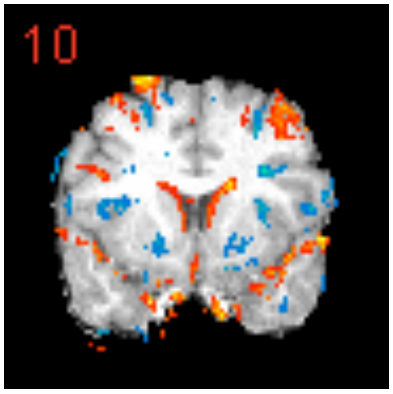
The statistical significant differences between the control and schizophrenia MRI samples captured by the multivariate discriminant approach are illustrated[4] in Figure 4. To determine these differences we have used the MR intensity values as input features and all the spatially normalised samples for training. In other words, $N = 43+25 = 67$ and $n = 91\times109\times91 = 902629$. As mentioned earlier, these changes correspond to the differences of one-dimensional sample group models ("definitely control" and "definitely patient" samples) on the PCA+MLDA space projected back into the image domain weighted by the pooled variance between the two sample groups. Analogously to the facial expression experiments described in the previous section (more specifically in subsection 3.2), these one-dimensional sample group models are represented by points in the PCA+MLDA space at 3 standard deviations from each corresponding sample group mean.

Figure 4 shows only three slices of the 3D MRI multivariate discriminant differences extracted, superimposed on a control brain image randomly selected. The numbers seen at the top of each slice rep-

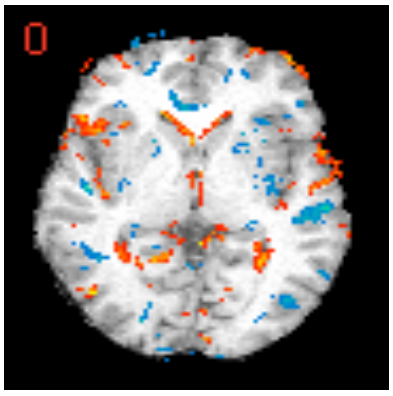
[4]For colour images, please see the electronic version of the paper.



(a) sagittal plane



(b) coronal plane



(c) axial plane

Figure 4: Effect size of the multivariate statistical differences comparing the intensity values described by the control and patient image models. The numbers seen at the top of each slice represent from left to right the standard coordinates in z (or axial), y (or coronal) and x (or sagittal) planes.

resent the standard coordinates in z (or axial), y (or coronal), and x (or sagittal) planes. We considered a difference important if its absolute value exceeds at least three pooled standard deviations. In the picture, the colour-scale shows relative intensity change as a range of this thresholding. In yellow/red the markings on the image are brighter in the control samples compared to the patient samples. Analogously, the areas in green/blue show regions of relative darkness in the control samples compared to the patients ones.

It is evident from Figure 4 that the differences found between the control and patient sample groups are not restricted to one specific neuroanatomic structure of the brain. In fact, there are clusters of multivariate differences that may be part of a single neuroanatomic structure or may fall in an intersection of various structures. In order to single out these discriminant regions, next subsection presents the automatic segmentation of the multivariate differences found (such as those shown in Figure 4).

Automatic segmentation of the most statistically significant differences

The segmentation of the cluster of points representing the most statistically significant differences is accomplished in two stages. First, we apply a dilation morphological operator on the matrix *diff*; second, the results of this dilation process are used to initialise an active contour model procedure that runs on the original matrix *diff*. The dilation operator is applied so that the active contours can merge clusters of points that might be closer together. However this is only used to generate the contour, whereas the statistically significant changes detected are kept unchanged.

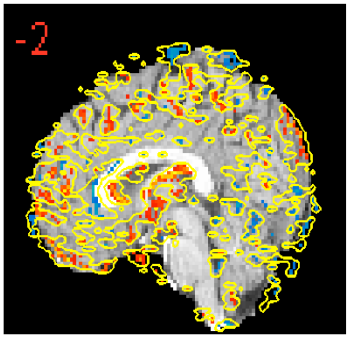
Our main interest with the segmentation procedure is to automatically generate multivariate discriminant regions in neuroimages of schizophrenia based on raw data, in contrast to the regions of interest (ROI) usually assumed as hypotheses beforehand and detected separately [15].

In this work we used the active contour technique known as level sets. In brief, level-set models (or geometric models) [34, 35] are continuous deformable models whose curves evolve using only geometric computation (i.e. without an explicit parametrisation of the contour). These models are capable of dealing with complex topology, including splitting or merging shapes. Its capacity to handle complex shapes is the reason for the recent popularisation of this procedure in medical image analysis [36].

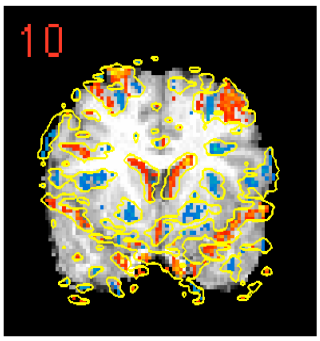
In the present paper we use an off-the-shelf variational level-set technique proposed in [37], whose evolution of the level-set function is derived from the minimisation of an energy functional. This energy functional is composed of an internal energy (which modulates the degree of dispersion of the level-set function with respect to a boundary) and an external energy term (that forces the level-set function towards image features). The motivation for using this technique is two fold: first, due to the internal energy, there is no need for reinitialisation; and second, the initialisation of the level-set functions can be done from arbitrary regions in the image domain (more details in [37]).

The results of the level set segmentation procedure on the images in Figure 4 are shown in Figure 5.

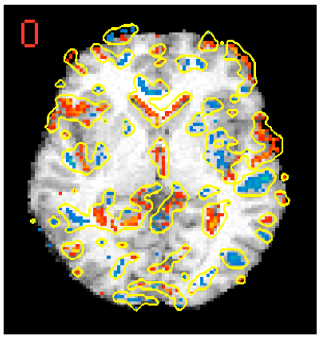
Figure 5 shows the contours built by the level-set procedure in yellow, and the clusters of most statistically significant points



(a) sagittal plane



(b) coronal plane



(c) axial plane

Figure 5: Results of the automatic segmentation using level sets on the images in Figure 4.

in red and blue (as discussed in Section 4). We can see that the level-set contours managed to merge, into unique regions, points from distinct anatomical structures (such as the contour surrounding part of the lateral ventricles in Figure 5(c)) or discard points that fell outside the contours.

Therefore, the combination of procedures proposed so far in this paper provided a number of discriminant regions of interest for schizophrenia research that were emergent from the data. Some of these discriminant regions may indeed point to relevant changes related to neuroimages of patients in comparison to those of controls.

The results shown in Figures 4 and 5 are samples of our result set, since the procedure described was applied to all 91 slices of the extreme models, each of which included an average number of 20 regions (containing around ten points each in average).

Some of the regions emergent from the data may not have been explored before in schizophrenia research. However it may also be the case that they were due to image artifacts, by-products of the image analysis itself. As illustrated in Figure 6, the PCA+MLDA linear statistical model has achieved a relatively high false positive rate (approximately 40%) with a true positive rate of around 70%, showing that both control and patient small sample groups overlap and have non-zero linear classification error.

In order to evaluate the results obtained, and to provide a ranking of the most informative regions obtained from the data, we confront the most statistically significant points found with evidences of neuroanatomical changes, related to schizophrenia, as presented in the literature. This process starts by labelling the neuroimage results according to an atlas, as discussed in the next section.

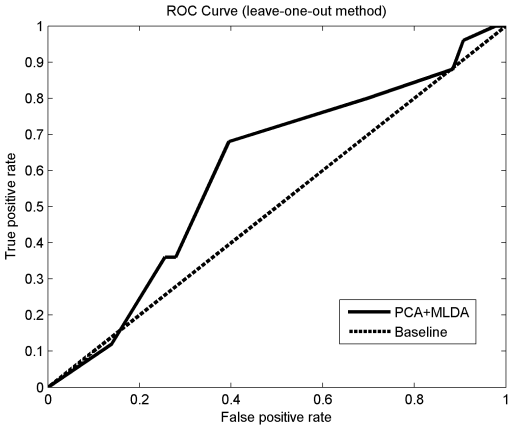


Figure 6: ROC curve based on the leave-one-out method of the PCA+MLDA statistical discriminant model.

5. Labelling the findings

The methods described in the previous sections are very powerful tools for discovering the most important discriminating features between two sets of input images, and singling them out into distinct regions, however they are incapable of providing a qualitative description of what these features represent within a certain context. For instance, in [38, 21], MLDA discovered a set of contraction and expansion regions that best classified neuroimages of preterm from control groups of babies. However, a paediatrician was needed to point out which of these findings were actual differentiating features between these two groups and which were possibly originated from other sources (image artifacts for instance). Our goal is to automate at least part of this process.

In order to bridge the findings presented above to a more conceptual level, first, the points grouped by level-set contours are mapped onto an atlas. As a result we obtain, for each group of points, in every im-

age slice, the names of the neuroanatomical structures related to the point's location. These descriptions are then contrasted with findings in the literature in order to rank the most relevant clusters, according to the number of matches between anatomical structures pointed out within the clusters with the expected results (as given in the medical literature). We discuss this procedure as follows.

Mapping points to an atlas

The points found by the procedures discussed in Sections 4 were mapped onto the automated Talairach atlas [39], which is a freely available web application that provides the anatomical labels related to 3D coordinates of locations in the human brain[5]. However, as the images in this work were normalised using the MNI brain, we had to apply the coordinate transformation discussed in [40], summarised in Algorithm 2 below.

Algorithm 2 Coordinate Transformation from MNI to Talairach spaces

x, y, z = MNI coordinates

if $z >= 0$ **then**
  $x_{Talairach} \Leftarrow 0.9900x$
  $y_{Talairach} \Leftarrow 0.9688y + 0.0460z$
  $z_{Talairach} \Leftarrow -0.0485y + 0.9189z$
else if $z < 0$ **then**
  $x_{Talairach} \Leftarrow 0.9900x$
  $y_{Talairach} \Leftarrow 0.9688y + 0.0420z$
  $z_{Talairach} \Leftarrow -0.0485y + 0.839z$
end if

As an example of the labelling procedure, Table 1 shows the labels related to the points shown in Figure 5(c). It is worth pointing out that the Talairach atlas labels coordinates into five distinct levels: *Hemisphere*, *Lobe*, *Gyrus*, *Tissue* and *Cell* (as shown in the top row of the table).

The task now is to verify whether these descriptions were already reported in the medical literature. Next subsection discusses this issue.

6. Evaluation

In order to evaluate the results obtained from the procedures discussed above we consider, as *gold standard*, literature reviews of meta-analyses relating schizophrenia with neuroanatomical alterations. Table 2 shows, in the first column, the neuroanatomical structure affected and, in the second, the related reviews.

We then contrasted every description of the most significant changes (such as those presented in Table 1), with the anatomical structures related to schizophrenia as provided in the literature (Table 2). In this context we understand as *true positives* (TP), descriptions that agree with the gold standard, whereas *false positives* (FP) are descriptions not present in the gold standard. In this work, true and false negatives are not defined as the gold standard only represents positive facts.

The results were evaluated according to their sensitivity and selectivity. Sensitivity was obtained as the rate between the number of distinct true positives (DTP), i.e. the number of items in the gold standard present in the descriptions, over the size of the gold-standard set (GS): DTP/GS (according to Table 2: GS = 14). Selectivity measures the rate of true findings (i.e. number of true positives) over the total amount descriptions produced by the framework proposed in this work: $TP/(TP + FP)$.

[5] In this work we used the Talairach Client available in http://www.talairach.org/, accessed in 27/02/2009.

Hemisphere	Lobe	Gyrus	Tissue	Cell
Left Cerebrum	Limbic Lobe	PG	GM	BA 27
Left Cerebrum	Sub-lobar	Caudate	GM	Caudate Head
Left Cerebrum	Sub-lobar	Thalamus	GM	MB
Left Cerebrum	Sub-lobar	Third Ventricle	CSF	*
Right Cerebrum	Limbic Lobe	PG	GM	BA 30
Right Cerebrum	Sub-lobar	Caudate	GM	Caudate Head
Right Cerebrum	Sub-lobar	Claustrum	GM	*
Right Cerebrum	Sub-lobar	Lateral Ventricle	CSF	*
Right Cerebrum	Temporal Lobe	Extra-Nuclear	WM	*
Right Cerebrum	Temporal Lobe	STG	GM	BA 22

Table 1: Some of the anatomic labels of points from Figure 5(c) as given by the Talairach atlas. The abbreviation PG stands for parahippocampal gyrus; GM is gray matter; CSF is cerebro-spinal fluid; STG is superior temporal gyrus; MB is mammillary body and BA is Brodmann area. The symbol "*" refers to coordinates without a related label in the atlas.

Affected Structure	Literature reviews
Lateral Ventricles	[4, 41–46]
Third Ventricles	[4, 43]
Fourth Ventricles	[4]
Temporal Lobe	[4, 41, 42]
Medial Temporal Lobe	[4]
Amygdala	[41, 42]
Hippocampus	[41–43]
PG	[41–43]
STG	[4]
Thalamus	[4, 41–43]
Basal Ganglia	[45]
Frontal Lobe	[4, 41, 42]
Occipital Lobe	[4, 41–43]
Parietal Lobe	[4, 41, 42]

Table 2: Reviews of anatomical changes related to schizophrenia.

Informally, sensitivity measures the rate of "correct" answers provided by our framework, whereas selectivity measures how objective the framework is. We obtained sensitivity and selectivity values for the data before and after the level-set segmentation. Table 3 summarises the results obtained.

	TP	DTP	FP
Before segm.	14,692	10	37,104
After segm.	14,472	10	36,582

Table 3: Results related to the procedure before and after level-set segmentation. In this table, TP means *true positives*, DTP *distinct true positives* and FP *false positives*

We can observe on Table 3 that the application of level-set segmentation kept almost all the original points, as the difference between TP and FP before and after its application is not significant. Therefore, in both cases we obtained around 0.28 of selectivity and 0.71 of sensitivity, meaning that 71% of the gold standard was covered by the results obtained and that 28% of the points generated by the procedures discussed in this paper were relevant with respect to the assumed gold standard. However, these results do not mean that the points that did not find a match in the gold standard should be rejected as noise, but that they should be considered in further research either for verifying an increase in selectivity in future extensions of this work, or as indicators of new regions of interest to be considered in future schizophrenia research. On this latter point we could use the individual sensitivity values for each region segmented by the level-set procedure to point out the most promising candidates of *Discriminant Regions*, as obtained from the data. Table 4 shows the ten highest sensitivity values of segmented regions, the corresponding image slices containing these regions are shown in Figure 7.

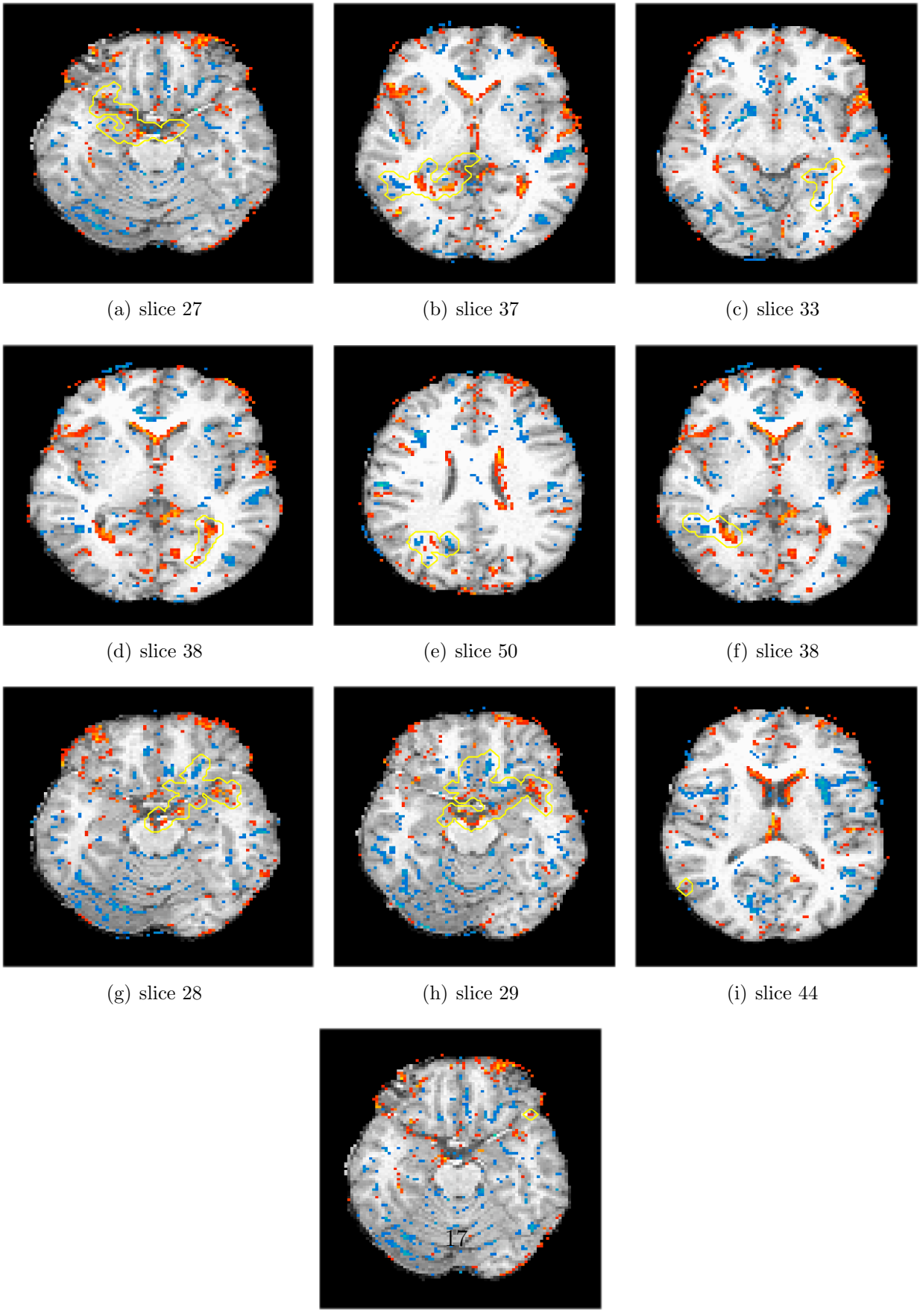
Rank	slice	Sensitivity
1	27	0.36
2	37	0.36
3	33	0.29
4	38	0.29
5	50	0.29
6	38	0.29
7	28	0.29
8	29	0.29
9	44	0.22
10	27	0.22

Table 4: Segmented regions with the highest sensitivity value.

7. Discussion

This work proposed an integrated framework for classifying and interpreting patterns of the schizophrenia disorder from 3D MR images using a combination of statistical discriminant analysis and active contour models. In the following paragraphs, we discuss some points that have emerged from this study which might be relevant in other similar investigations.

It is important to remark that the construction of the multivariate statistical model (PCA+MLDA) for extracting the most discriminant features between two groups relies on the quality of the inter-subject correspondences calculated by either affine or non-rigid registration algorithms. In other words, when we use PCA as a feature extraction technique we must have in mind that PCA outputs projection directions that maximise the total scatter composed of all images of all classes. As a consequence, when we retain all the PCA eigenvectors and choose such projection without any previous alignment of



(a) slice 27 (b) slice 37 (c) slice 33

(d) slice 38 (e) slice 50 (f) slice 38

(g) slice 28 (h) slice 29 (i) slice 44

(j) slice 27

Figure 7: Slices containing the regions with the highest sensitivity values, according to Figure 4.

the images, PCA might describe unwanted variations inherent to any image acquisition process, such as differences in rotation, translation, scaling, and shape [21]. In order to minimise those variations (that are not necessarily related to anatomical differences between the images) and to transform data in a way that the images belonging to distinct classes occupy as compact and disjoint regions in a lower dimensional feature space as possible, the spatial normalisation of the images provided by the image registration algorithms is a fundamental preprocessing stage for the success of the multivariate statistical model [21].

The output of the multivariate statistical model was analysed by subtracting the extreme models between the two groups ($-3\sigma_1$ and $+3\sigma_2$, cf. Section 3.2) weighted by a factor related to the pooled variance between the groups. Taking into account the implicit hypothesis that the two groups are normally distributed, the extreme models are on the opposite ends of the Gaussians associated with each group. In order to take this into account, we then selected the differences that were higher than a threshold of three standard deviations. It is left for future investigations the analysis of whether other statistical tests, or a different threshold value, would have any clinical significance.

This method of selecting points according to a threshold on standard deviations is known as *effect-size pruning* [47], which is an objective measure of interestingness that evaluates the degree of diversity in a compact description (i.e. the image models) of the data. Usually a threshold on the group variances is used as such measure in these cases, but (as discussed in Section 3.2) we preferred to use the effect size of the matrix *diff* (Formula 1) as our goal was to identify the most differentiating features between controls and patients. These features are more explicit when comparing extreme models of the two image groups rather than their means, which are smooth by definition.

In Section 6 we present a ranking of the segmented regions according to the highest degree of sensitivity value. If this ranking was made on the other way around (i.e. ranking the regions with the lowest sensitivity values) it could be used as a measure of *novelty* [47]. However, we prefer not to use the term *novelty* due to the nature of our domain: stating that our system *found* a novel structure, and that this is indeed related to schizophrenia, warrants research on the direction of a meticulous medical evaluation of the findings (which is outside the scope of this paper).

It might be argued that the dilation before the level-set segmentation could interfere with the data observed. However, we have to emphasise that the contours were made to converge on the non-dilated data. It is possible to initialise the active contours without the previous application of the morphological operation but, in this case, some contours end up collapsing and there would be no merging of various cluster of points into one region.

We did not obtain a significant distinction between sensitivity and selectivity values for our results before and after the application of the level sets. This means that the level-sets were used largely to merge cluster of points and not to filter isolated points in the image (which could indicate outliers). However, the active contours allowed the definition of discriminant regions that were emergent from the data. The sensitivity values for these regions were used to rank them, indicating discriminant candidates to be con-

sidered in future research.

The images submitted to the level-set segmentation were automatically generated by the multivariate statistical method and, therefore, there is no ground truth directly associated to them. In fact, there is no ground truth associate to image segmentation related to brain regions linked to schizophrenia, since neuroimaging for schizophrenia diagnostics is still a matter of debate [48]. Evaluating the results obtained with a gold standard based on literature findings (as done in Section 6) was the best that could be done under these conditions.

We could have evaluated the level-set procedure, however, using the *Internet Brain Segmentation Repository* (IBSR)[6] (or any other MRI repository available [8]), but this would lead to an orthogonal research path to that taken in this paper for two reasons: first, the ground truth segmentation usually given in such repositories is on raw data (not on statistically significant changes as provided by a linear discriminant analysis method); second, the purpose of the submitted paper is not to prove that the statistical discriminant analysis, or the segmentation method, were better than other competing procedures in terms of some evaluation procedure, but to show that it is possible to extract and interpret statistically significant regions representing changes between neuroimages from controls and patients that can be meaningful for schizophrenia research.

Future directions

This article extended a general multivariate linear framework [21] to extract statistical differences of 3D MR brain images of adult subjects suffering from schizophrenia compared to a healthy control group. The framework is not restricted to a specific linear discriminant analysis model, other separating hyper-plane methods can be used, such as Support Vector Machines (SVM) [49, 50]. Although the multivariate linear approach has been demonstrated in two-class problems, it is extensible to several classes. Since the brain changes found in schizophrenia are not exclusively characteristic of this disease, a multi-class analysis involving a number of brain disorders and controls could provide a comprehensive understanding of abnormalities in brain development.

Future work should also consider the most sensitive regions obtained by the level-set procedure (as presented in Section 6) as prior hypothesis on schizophrenia research. This will provide an experimental evaluation of the methods discussed in this work, as well as a possible new standpoint on the neuroanatomical structures related to the disease.

In Section 6, neuroanatomical structures (commonly linked to schizophrenia in the literature) were compared with the changes in neuroimages obtained in the present work. This was done in order to evaluate the present findings. We are currently investigating the construction of an ontology about neuroanatomy that could also encompass the *knowledge about* a domain, and not only the domain itself, in order to allow the organisation and correlation of what is known about schizophrenia (e.g. in literature reviews) and the neuroanatomical structures involved [51]. However, *findings about* schizophrenia falls within the epistemology (and not ontology) umbrella. A complete solution of combining ontologies

[6] http://www.cma.mgh.harvard.edu/ibsr/, last accessed on 5/11/09

with epistemology is a matter for future research. Nevertheless we can envisage that the results presented in this work could serve as input knowledge for this formalisation, providing new epistemological classes about the input data.

8. Concluding remarks

This article discussed our current research on methods for automatically finding discriminative features and anatomical regions that characterise schizophrenia from neuroimages.

In the present work, we discussed the main issues involving the construction of an integrated framework for classifying and analysing patterns of disorders from medical images using a combination of image registration, multivariate statistics, and active contour models. Our central goal was achieved: in this work all the data was analysed simultaneously rather than assuming *a priori* regions of interest. As a consequence of this, by using active contours models, we were able to obtain regions of interest that were emergent from the data. The results were evaluated using, as gold standard, well-known facts about neuroanatomical changes related to schizophrenia. Most of the items in the gold standard was covered in our result set, implying on a high sensitivity rate.

Acknowledgement

We would like to thank Luciana Cristina dos Santos (IPq-USP) for helping us with the normalisation procedures and Dr. Debora P. Bassit for obtaining the original MRI images. We are also in debt with the three anonymous reviewers for their very sharp comments that helped improving this paper greatly.

This work is partially funded by FAPESP Project: LogProb 08/03995-5, and CNPq, Brazil.

References

[1] K. T. Mueser, S. R. McGurk, Schizophrenia, Lancet 363 (2004) 2063–72.

[2] P. J. Harrison, M. J. Owen, Genes for schizophrenia: recent findings and their pathophysiological implications, Lancet 361 (2003) 417–19.

[3] K. H. Nuechterlein, M. E. Dawson, A heuristic vulnerability/stress model of schizophrenic episodes, Schizophrenia Bulletin 10 (2) (1984) 300–12.

[4] M. E. Shenton, C. C. Dickey, M. Frumin, R. W. McCarley, A review of MRI findings in schizophrenia, Schizophrenia Research 49 (1–2) (2001) 1–52.

[5] W. F. Gattaz, W. Rost, K. Kohlmeyer, K. Bauer, C. Hubner, T. Gasser, CT scans and neuroleptic response in schizophrenia: a multidimensional approach, Psychiatry Research 26 (3) (1988) 293–303.

[6] J. Duncan, N. Ayache, Medical Image Analysis: Progress over two decades and the challenges ahead, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (1) (2000) 85–106.

[7] T. Mcinerney, D. Terzopoulos, Deformable models in medical image analysis: A survey, Medical Image Analysis 1 (2) (1996) 91–108.

[8] D. Withey, Z. Koles, A review of medical image segmentation: methods and available software, International Journal of Bioelectromagnetism 10 (3) (2008) 125–148.

[9] I. Middleton, R. Damper, Segmentation of magnetic resonance images using a combination of neural networks and active contour models, Medical engineering & physics 26 (1) (2004) 71–86.

[10] L. Amini, H. Soltanian-Zadeh, C. Lucas, M. Gity, Automatic segmentation of thalamus from brain mri integrating fuzzy clustering and dynamic contours, IEEE transactions on biomedical engineering 51 (5) (2004) 800–11.

[11] E. Ashton, M. Berg, K. Parker, J. Weisberg, C. Chen, L. Ketonen, Segmentation and feature extraction techniques, with applications

to mri head studies, Magnetic resonance in medicine 33 (5) (1995) 670–677.

[12] T. Kapur, W. Grimson, W. Wells, R. Kikinis, Segmentation of brain tissue from magnetic resonance images, Medical image analysis 1 (2) (1996) 109–27.

[13] J. Ashburner, J. G. Csernansky, C. Davatzikos, N. C. Fox, G. B. Frisoni, P. M. Thompson, Computer-assisted imaging to assess brain structure in healthy and diseased brains, Lancet neurology 2 (2) (2003) 79–88.

[14] E. R. Sowell, P. M. Thompson, C. J. Holmes, R. Batth, T. L. Jernigan, A. W. Toga, Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping, Neuroimage 9 (6) (1999) 587–597.

[15] J. Ashburner, K. J. Friston, Voxel-based morphometry - the methods, Neuroimage 11 (6) (2000) 805–821.

[16] K. J. Friston, A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith, R. S. J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach, Human Brain Mapping 2 (4) (1995) 189–210.

[17] C. Davatzikos, Why voxel-based morphometric analysis should be used with great caution when characterizing group differences, Neuroimage 23 (1) (2004) 17 – 20.

[18] Z. Lao, D. Shen, Z. Xue, B. Karacali, S. Resnick, C. Davatzikos, Morphological classification of brains via high-dimensional shape transformations and machine learning methods, Neuroimage 21 (1) (2004) 46–57.

[19] P. Golland, W. Grimson, M. Shenton, R. Kikinis, Detection and analysis of statistical differences in anatomical shape, Medical Image Analysis 9 (1) (2005) 69–86.

[20] C. E. Thomaz, F. L. S. Duran, G. F. Busatto, D. F. Gillies, D. Rueckert, Multivariate statistical differences of MRI samples of the human brain, Journal of Mathematical Imaging and Vision 29 (2-3) (2007) 95–106.

[21] C. E. Thomaz, J. P. Boardman, S. Counsell, D. L. G. Hill, J. V. Hajnal, A. D. Edwards, M. A. Rutherford, D. F. Gillies, D. Rueckert, A multivariate statistical analysis of the developing human brain in preterm infants, Image and Vision Computing 25 (6) (2007) 981–994.

[22] J. R. Sato, C. E. Thomaz, E. F. Cardoso, A. Fujita, M. G. Morais-Martin, E. A. Junior, Hyperplane navigation: a method to set individual scores in fMRI group datasets, Neuroimage 42 (4) (2008) 1473–1480.

[23] A. Fujita, L. R. Gomes, J. R. Sato, R. Yamaguchi, C. E. Thomaz, M. C. Sogayar, S. Miyano, Multivariate gene expression analysis reveals functional connectivity changes between normal/tumoral prostates, BMC Systems Biology 2 (2008) 106.

[24] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, D. J. Hawkes, Non-rigid registration using free-form deformations: Application to breast MR images, IEEE Transactions on Medical Imaging 18 (8) (1999) 712–721.

[25] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd Edition, Academic Press, Boston, MA, 1990.

[26] C. E. Thomaz, E. C. Kitani, D. F. Gillies, A maximum uncertainty LDA-based approach for limited sample size problems - with application to face recognition, Journal of the Brazilian Computer Society 12 (2) (2006) 7–18.

[27] C. E. Thomaz, Maximum entropy covariance estimate for statistical pattern recognition, Ph.D. thesis, Department of Computing, Imperial College London (2004).

[28] D. P. Bassitt, M. R. L. Neto, C. C. de Castro, G. F. Busatto, Insight and regional brain volumes in schizophrenia, European Archives of Psychiatry and Clinical Neuroscience 257 (1) (2007) 58–62.

[29] K. J. Friston, A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith, R. S. J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach, Human Brain Mapping 2 (4) (1995) 189 – 210.

[30] J. C. Mazziotta, A. W. Toga, A. Evans, P. Fox, J. Lancaster, A probabilistic atlas of the human brain: Theory and rationale for its development, NeuroImage 2 (2) (1995) 89–101.

[31] C. D. Good, I. S. Johnsrued, J. Ashburner, R. N. Henson, K. J. Friston, R. S. Frackowiak, A voxel-based morphometric study of ageing in 465 normal adult human brains, Neuroimage 14 (1) (2001) 21 – 36.

[32] S. M. Smith, Fast robust automated brain extraction, Human Brain Mapping 17 (3) (2002) 143 – 155.

[33] C. Rorden, M. Brett, Stereotaxic display of brain lesions, Behavioural Neurology 12 (4)

(2000) 191 – 200.

[34] V. Caselles, F. Catte, T. Coll, F. Dibos, A geometric model for active contours, Numerische Mathematik 66 (1) (1993) 1–31.

[35] J. A. Sethian, Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, 2nd Edition, Cambridge Monographs on Applied and Computational Mathematics (n. 3), Cambridge University Press, Cambridge, UK, 1999.

[36] E. Angelini, Y. Jin, A. Laine, Handbook of Biomedical Image Analysis, Topics in Biomedical Engineering International Book Series, Springer, New York, NY, USA, 2007, Ch. State of the Art of Level Set Methods in Segmentation and Registration of Medical Imaging Modalities, pp. 47–101.

[37] C. Li, C. Xu, C. Gui, M. D. Fox, Level-set evolution without re-initialization: A new variational formulation, in: Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society Press, Silver Spring, Md., 2005, pp. 430–436.

[38] C. E. Thomaz, J. P. Boardman, D. L. G. Hill, J. V. Hajnal, A. D. Edwards, M. A. Rutherford, D. F. Gillies, D. Rueckert, Whole brain voxel-based analysis using registration and multivariate statistics, in: Proc. of the 8th Medical Image Understanding and Analysis MIUA'04, BMVA Press, London, UK, 2004, pp. 73–76.

[39] J. Lancaster, M. Woldorff, L. Parsons, M. Liotti, C. Freitas, L. Rainey, P. Kochunov, D. Nickerson, S. Mikiten, P. Fox, Automated Talairach atlas labels for functional brain mapping, Human Brain Mapping 10 (3) (2000) 120–31.

[40] M. Brett, I. S. Johnsrude, A. M. Owen, The problem of functional localization in the human brain, Nature Reviews Neuroscience 3 (3) (2002) 243–9.

[41] L. Howard, Comprehensiveness of systematic review – update, The British Journal of Psychiatry 176 (2000) 396–401.

[42] S. M. Lawrie, S. Abukmeil, Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies, The British Journal of Psychiatry 172 (1998) 110–120.

[43] P. Sallet, As esquizofrenias segundo a classificação das psicoses endogenas de Karl Leonhard and sua correlação com imagens cerebrais por meio de ressonância magnetica (MRI), Phd thesis, Medical School, Universidade de São Paulo, São Paulo, Brazil (2002).

[44] A. Vita, M. Dieci, C. Silenzi, F. Tenconi, G. M. Giobbio, G. Invernizzi, Cerebral ventricular enlargement as a generalized feature of schizophrenia: a distribution analysis on 502 subjects, Schizophrenia Research 44 (1) (2000) 25–34.

[45] P. J. Harrison, The neuropathology of schizophrenia: A critical review of the data and their interpretation, Schizophrenia Research 44 (1) (2000) 25–34.

[46] H. Elkis, L. Kimura, L. M. Nita, M. C. R. G. Tissot, Neuroimagem estrutural e psicopatologia: sintomas positivos e negativos e dilatação ventricular na esquizofrenia, Revista Brasileira de Psiquiatria 23 (2) (2001) 19–23.

[47] L. Geng, H. J. Hamilton, Interestingness measures for data mining: A survey, ACM Computing Surveys 38 (3) (2006) 37–68.

[48] J. Woolley, P. McGuire, Neuroimaging in schizophrenia: what does it tell the clinician?, Advances in Psychiatric Treatment 11 (2005) 195–202.

[49] V. N. Vapnik, The Nature of Statistical Learning Theory, 2nd Edition, Springer,NY, 1999.

[50] J. R. Sato, A. Fujita, C. E. Thomaz, M. G. Morais-Martin, J. Mourao-Miranda, M. J. Brammer, E. A. Junior, Evaluating SVM and MLDA in the extraction of discriminant regions for mental state prediction, NeuroImage 46 (1) (2009) 105–114.

[51] P. Santos, R. Freire, D. N. dos Santos, C. Thomaz, P. Sallet, M. L. a, A region-based ontology of the brain ventricular system and its relation to schizophrenia, in: Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions, IGI Global, USA, 2010, to appear.