To appear on ISBI-2002, Washington, DC.

BAYESIAN CLUSTERING METHODS FOR MORPHOLOGICAL ANALYSIS OF MR IMAGES

Hanchuan Peng, Edward Herskovits, and Christos Davatzikos

Center for Biomedical Image Computing,
Department of Radiology, School of Medicine, Johns Hopkins University,
601 N. Caroline St., JHOC # 3230, Baltimore, MD, 21287, USA.
Email: phc@cbmv.jhu.edu, ehh@braid.rad.jhu.edu, hristos@rad.jhu.edu

ABSTRACT

Determining the relationship between structure (i.e. morphology) and function is a fundamental problem in brain research. In this paper we present a new framework based on Bayesian clustering methods for the voxel-wise statistical morphology-function analysis of registered MR images. We construct a Bayesian network to automatically identify the significant associations between voxel-wise morphological variables and functional variables, such as cognitive performance. A Bayesian latent variable induction method is applied to locate the homogeneous association regions on registered maps of morphological variables. Experimental results on images with simulated atrophy confirm that the new method outperforms conventional statistical method, based on linear statistics.

1. INTRODUCTION

One of the most important tasks in medical imaging research is morphology-function analysis, which aims to identify the structures (from a medical-image collection of subjects) that possess statistical associations to the functionality (i.e. clinical variables) of the respective subjects. Many different approaches [1,3-7,10-13] have been proposed for generating statistical maps that identify groups of voxels for which significant correlations exist among morphological and clinical measurements. Given some types of the morphological measurement, e.g. density maps in a stereotaxic space [1,3,7,10], morphology-function analysis can be basically described as comparison of these morphological measurements (i.e. registered images of a group of subjects) and generation of Regions Of Interest (ROIs) that have statistical associations to the functional variables (i.e. clinical variables) of these subjects. For this purpose, our work in this paper is to present Bayesian clustering methods to generate the statistically optimal ROIs, as shown in Fig.1.

Fig.1 is a schematic framework of the morphology-function analysis, where the input data **D** include the morphological measurement maps of MR images and the corresponding clinical functionality-records of a collection of subjects. The outputs are the ROIs statistically associated with these functionality-records. For example, the data **D** can be described in terms of voxel-wise morphological variables and functional variables. In the experiments of this paper, we select the input image data as the RAVENS (Regional Analysis of Volumes Embedded in Stereotaxic Space) density maps [7] of MR images, in which the image brightness within any structure is proportional to the actual volume of the structure before registration. (Note: The RAVENS maps are generated via mass-preserving high-dimensional elastic transformations [7,10] that register individual images to a template. If a structure is compressed during this transformation, because it is larger in the individual compared to the template, then the tissue density in this structure will increase, so that the total amount of tissue is preserved. Accordingly, the brightness of RAVENS maps will increase.) These RAVENS maps can be overlaid and analyzed on a voxel-wise basis [3]. The clinical (i.e. functional) variables have states such as "normal", "abnormal", etc, depending on specific imaging studies.

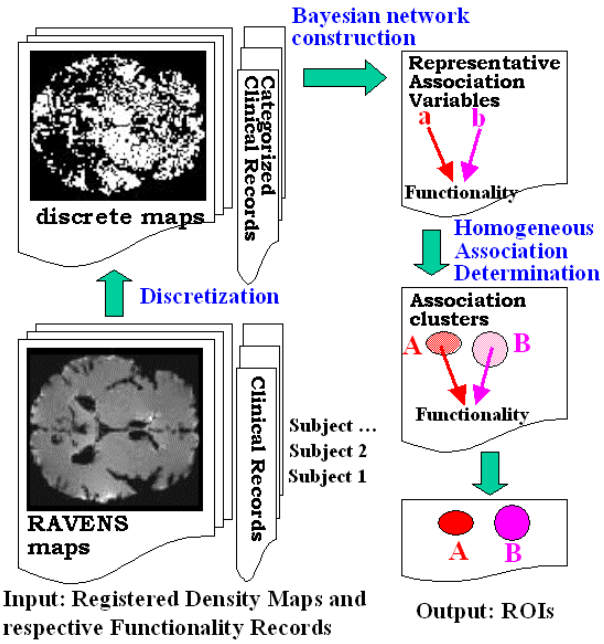


Fig. 1 Schematic illustration of the Bayesian morphology-function analysis. Top left shows a discrete map of the continuous RAVENS maps [3,7,10] at the bottom-left. Final result is a number of clusters (e.g. A and B), each associated with the functional (clinical) variable (e.g. a and b) in different ways.

As shown in Fig.1, our method is based on Bayesian clustering techniques to accomplish the middle steps. Essentially, we identify the association ROIs based on a clustering method of constructing Bayesian networks (BNs) from data; meanwhile a Bayesian latent variable induction method [2] is used to locate the homogeneous association regions in each association cluster. Basically, this framework is generally applicable to many different medical imaging applications; in this paper we only show an example with brain MR images.

Of note, conventionally this task has been accomplished via standard linear statistical tests, e.g. t-tests (TT), paired t-tests (PT), and ANOVAs, many of which suffer from a major drawback that a mechanism to locate nonlinear association patterns of the interaction of ROIs is not considered. Another drawback is that many statistical tests require the definition of a subjective confidence interval, or *p*-value threshold, which is not always easy to determine. As shown in this paper, our approach can overcome both difficulties.

This paper is arranged as follows. Section 2 introduces our methods concisely. Section 3 shows the experimental results. Section 4 gives a brief conclusion.

2. METHODS

We present Bayesian clustering methods for two steps in Fig.1, i.e. Bayesian network construction, and homogeneous association determination, as presented consequently.

2.1 Handle Continuous Voxel Variables: Bayesian Clustering for Data Discretization

The key part of our morphology-function analysis method is to identify the ROIs based on a Bayesian network construction algorithm (as explained in next subsection); for the purpose of nonlinear computation of Bayesian networks, the input data to the Bayesian network construction algorithm should be discrete [8]. Since the input data (e.g. RAVENS maps) are continuous, we must discretize the data. We use the z-score thresholding method in data discretization. That is, for any voxel-wise data, we calculate the mean value and compare each case to the mean value; the result label takes 1 when a case has a larger value than the mean value, otherwise takes 0.

2.2 Identify Association Clusters: Heuristic Bayesian Network Construction

The essence of our approach is to construct a Bayesian network that captures the statistical associations between the voxel variables and the functional variable(s). In this setting, both voxel variables and the functional variable(s) are nodes in a Bayesian network and associations among them are represented by edges. If an edge is favored by the input data, there is an association between corresponding pair-wise nodes. Given the fixed variable set, different combinations of edges actually mean different Bayesian network structures; hence the problem of discovering associations can be expressed as construction of the best Bayesian network structure **S** that represents the input data **D**. The Bayesian score [8], i.e. likelihood of the occurrence of the data given a specific structure, is used to approximate the posterior probability of this structure given the data, under the assumption of a uniform prior. Because of the hundreds of thousands of voxel variables in medical image data, it is impractical to compare all possible structures; the following heuristic algorithm is thus proposed to identify clusters of voxels that have similar associations with a functional variable.

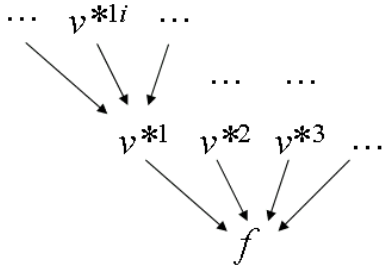


Fig. 2 The Bayesian network constructed by our algorithm.

Suppose that initially we have a voxel-wise defined morphological variable set **V** and a Bayesian network structure **S** with the functional variable *f*. As long as **V** is not empty, the search algorithm repeats the following 5-step loop: **First**, for each variable v_i in **V** the algorithm compares the pair-wise Bayesian networks $\{\mathbf{S} \cup [v_i \rightarrow f]\}$ and **S**, which stand for structures with and without the edge from v_i to *f*, respectively, to decide if v_i is associated with *f*. A voxel-function association is defined as that the structure $\{\mathbf{S} \cup [v_i \rightarrow f]\}$ has a larger Bayesian score than the structure **S**; otherwise we conclude that v_i is independent of *f*. **Second**, among all voxel variables with voxel-function associations we select the voxel variable v^*, which has the strongest voxel-function association, as the "representative association variable" of the current loop. The strength of a voxel-function association is indicated by the difference of the logarithm Bayesian scores of $\{\mathbf{S} \cup [v_i \rightarrow f]\}$ and **S**. **Third**, in **V** the algorithm finds the subset of voxel variables that are conditionally independent of *f* given the current representative association variable v^*. This voxel variable subset is called the "c-set". Here voxels are likely to be strongly associated with v^*, since they are conditionally independent of *f* given v^*. **Fourth**, the algorithm applies a latent-variable induction Bayesian clustering method (as explained in next subsection) to the c-set to isolate a subset (called the "e-set") in which voxel variables have approximately equal voxel-function associations (i.e. conditional probabilistic distributions) with the current representative association variable v^*. That is, in the e-set all voxels have homogeneous associations to the functional variable *f*. The union of the e-set and the current representative association variable v^* is regarded as the functional region of the current loop. **Finally**, the algorithm replaces **S** with $\{\mathbf{S} \cup [v^* \rightarrow f]\}$, and removes all variables in the e-set and the representative association variable v^* from the voxel variable set **V**.

In each loop, this algorithm outputs regions that have homogeneous associations to the functional variable(s). The final Bayesian network has a structure similar to that shown in Fig.2, where a few representative association variables v^{*1}, v^{*2}, … represent all other association voxel variables {…, v^{*1i}, …} and {…, v^{*2i}, …}, …, respectively. The functional region of the *j*th loop, i.e. the aggregate of v^{*j} and $\{v^{*ji}, i=1,2,\ldots\}$, is the cluster of association voxels for that loop. The corresponding ROIs are then obtained by observing the spatial groups of voxels in that aggregate.

In Fig.2 we see that the Bayesian network approach results in a hierarchy of functional ROIs; this is different from conventional morphology-function analysis methods (i.e. based on standard statistical tests). This hierarchical structure indicates that the importance of an ROI is evaluated based on the previously found ROIs. On the contrary standard statistical tests that evaluate the importance of each voxel in an isolated manner may not detect nonlinear associations induced by interaction among regions.

2.3 Determine Homogeneous Associations: Bayesian Latent Variable Induction

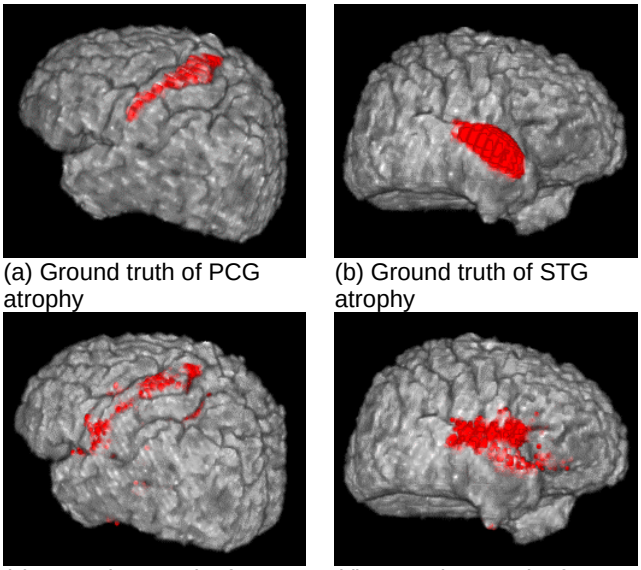
In the fourth step of the Bayesian network construction algorithm, we detect voxel variables that are "equivalent" to the current representative association variable, v^*. We refer to such "equivalence" as that two variables v and u have the same number of states, and there is $p(v|u) \approx p(u|v) \approx 1$ for every state. In our approach, the equivalent variables of v^{*j}, i.e. $\{v^{*ji}, i=1,2,\ldots\}$ as shown in the left top of Fig.2, are only searched in the corresponding c-set. We employ an approach similar to Chickering [2], which uses a Bayesian-network latent variable induction scheme to categorize cases of data, to detect variables with similar conditional-probability distributions (i.e. homogeneous associations). First, we transpose the data of the c-set variables as a pseudo data set; that is, the positions of variables and cases are exchanged. Then within the latent-variable induction scheme [2] a Monte Carlo approximation is used (because of its good accuracy) to estimate the posterior of data given the induced parameters. After the Monte Carlo iterations converge we find all variables with the same latent-variable state as v^*, and put these variables in the equivalent association voxel set, i.e. e-set.

2.4 Parameter Selection

As opposed to conventional statistical methods, in this Bayesian framework, there is no subjective parameter to be set; on the contrary, all parameters can be decided automatically. However, in the homogeneous association determination, the number of states of the latent variable can be specified in advance to reduce some computation. We can also adjust such a parameter, denoted as r_L, in the experiment for the purpose of comparison of different methods.

3. EXPERIMENTAL RESULTS

Due to the space limitation only partial results are shown in this paper. We used a simulated atrophy data set [3] of cerebral MR images (T1-weighted gradient-echo SPGR images; 3D sizes are 256×256×129 voxels; voxel resolution is 0.9375mm×0.9375mm ×1.5mm) of 11 normal elderly subjects. We manually selected two gyri, the right precentral gyrus (PCG) and the left superior temporal gyrus (STG), in all subjects and introduced a 30% uniform contraction into the labeled gyri, thereby creating 11 additional images with localized atrophy in these gyri. The labeled region of each subject is called the atrophy mask. Hence we have a group of 11 subjects without atrophy and another group of 11 subjects with atrophy. We defined the respective states of the functional variable as normal and abnormal. We obtained RAVENS density maps of these 22 images by using the STAR algorithm [7]. Because the STAR algorithm preserves the brain mass of each image, the atrophic regions in the abnormal group of images have smaller mean intensity than those for the normal group of images. As is customary in voxel-based morphometry, a Gaussian kernel (9mm diameter for the reported results) is utilized to smooth RAVENS maps and reduce registration error. Then we down-sampled each image by a factor of 2, and cropped it with the largest brain-region bounding box across all images to reduce the computational burden of the analysis. The final image size is 74×91×65.



(a) Ground truth of PCG atrophy (b) Ground truth of STG atrophy

(c) Detection result of PCG atrophy (d) Detection result of STG atrophy

Fig. 3 Comparison of the detection results versus the ground truth locations of atrophy.

We applied the z-score thresholding method to discretize the 22 RAVENS maps. The discrete data, 22 binary maps and respective 22 values for the functional variable, are input to the Bayesian network construction algorithm. The resulting ROIs of our Bayesian methods are visually compared with the ground truth atrophy-masks in Fig.3, where we overlapped the atrophy-masks and the result ROIs on a subject's brain for visualization. Our method successfully detected the atrophic regions (on both PCG and STG) that are associated with the functional states of subjects, although due to the complexity of brain structures and imperfectness of the image registration method, the ground truth and the detection results differ minimally.

Table 1. SDR and SNR results for different parameters with the non-smoothed atrophy-mask. BC is our Bayesian Clustering method; TT is t-test; t_{TT} is the confidence threshold of the *t*-map.

BC	r_L	2	3	4	5
	SDR	0.9488	0.9401	0.9041	0.9126
	SNR	0.0909	0.1359	0.1869	0.1741
TT	t_{TT}	3	5	10	15
	SDR	0.7822	0.4483	0.0903	0.0087
	SNR	0.1859	0.3640	1.4748	17

Theoretically a Receiver Operating Characteristic (ROC) curve analysis is equivalent to the SDR/SNR (Signal-Detection-Ratio and Signal-Noise-Ratio) analysis; for convenience the obtained SDRs and SNRs (using the standard definitions) with practical parameters in our experiments are shown in Table 1. The respective results of the standard t-test method are also compared. Our method outperforms the conventional t-test method in SDR. Notably, when both methods have similar SNRs (i.e. r_L=5 and t_{TT}=3), the SDR of our method is much better than that of t-test. Of note, in this study due to the complexity of data and the smoothing applied, there are surrounding areas detected around the ground-truth locations, which causes SNRs to be less than 1.

4. CONCLUSIONS

In this paper we have described a Bayesian framework for morphology-function analysis; particularly, we propose an algorithm for Bayesian network construction to identify significant morphology-function associations; for this purpose, a Bayesian latent variable induction method can be used to locate homogeneous associations for representative associations found by the Bayesian network construction algorithm. This method outperforms the t-test in our experiment on simulated data. We plan to extend our framework to effectively detect other morphology-function associations between image data and categorical variables.

ACKNOWLEDGEMENT

This work was supported in part by The Human Brain Project, National Institutes of Health grant R01 AG13743, which is funded by the National Institute of Aging, the National Institute of Mental Health, the National Aeronautics and Space Administration, and the National Cancer Institute. Dr. Herskovits was also supported by a Richard S. Ross Clinician Scientist Award.

REFERENCES

[1] Ashburner, J. and Friston, K.J., "Voxel-based morphometry: the methods," NeuroImage, vol.11, pp.805-821, 2000.

[2] Chickering, D.M., and Heckerman, D., "Efficient approximations for the marginal likelihood of Bayesian networks with hidden variables," Machine Learning, vol.29, pp.181-212, 1997.

[3] Davatzikos, C., Genc, A., Xu, D.R., and Resnick, R.M., "Voxel-based morphometry using RAVENS maps: methods and validation using simulated longitudinal atrophy," NeuroImage, vol.14, pp.1361-1369, 2001.

[4] Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., and Frackowiak, R.S.J., "Spatial registration and normalization of images," Human Brain Mapping, vol.2, pp.165-189, 1995.

[5] Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith C.D., and Frackowiak, R.S.J., "Statistical parametric maps in functional imaging: a general linear approach", Human Brain Mapping, vol.2, pp.189-210, 1995.

[6] Gaser, C., Volz, H.P., Kiebel, S., Riehemann, S., and Sauer, H., "Detecting structural changes in whole brain based on nonlinear deformations-application to schizophrenia research," NeuroImage, vol.10, pp.107-113, 1999.

[7] Goldszal, A.F., Davatzikos, C., Pham, D.L., Yan, X.H., Bryan, R.N., and Resnick, S.M, "An image processing system for qualitative and quantitative volumetric analysis of brain images," Journal of Computer Assisted Tomography, vol.22, no.5, pp.827-837, 1998.

[8] Herskovits, E.H., "Computer-based probabilistic-network construction," Doctoral Dissertation, Medical Informatics, Stanford University, 1991.

[9] Roberts, S.J., Husmeier, D., Rezek, I., and Penny, W., "Bayesian approaches to Gaussian mixture modeling," IEEE Trans. PAMI, vol.20, no.11, pp.1133 –1142, 1998.

[10] Shen, D.G., and Davatikos, C., "HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration", IEEE MMBIA 2001, Kauai, Hawaii, USA, Dec 9-10, 2001.

[11] Thompson, P.M., and Toga, A.W., "Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations," Medical Image Analysis, vol.1, no.4, pp.271-294, 1997.

[12] Thompson, P.M., MacDonald, D., Mega, M.S., Holmes, C.J., Evans, A.C., and Toga, A.W., "Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces," Journal of Computer Assisted Tomography, vol.21, pp.567-581, 1997.

[13] Woermann, F.G., Free, S.L., Koepp, M.J., Ashburner, J., and Duncan, J.S., "Voxel-by-voxel comparison of automatically segmented cerebral gray matter – a rater-independent comparison of structural MRI in patients with epilepsy," NeuroImage, vol.10, pp.373-384, 1999.