

Graph Based Spatial Position Mapping of Low-Grade Gliomas

Sarah Parisot^{1,2,3}, Hugues Duffau⁴, Stéphane Chemouny³,
and Nikos Paragios^{1,2,*}

¹ Laboratoire MAS, Ecole Centrale Paris, Chatenay Malabry, France

² Equipe GALEN, INRIA Saclay - Ile de France, Orsay, France

³ Intrasense SAS, Montpellier, France

⁴ Département de Neurochirurgie, Hopital Gui de Chauliac, CHU Montpellier, France

Abstract. Low-grade gliomas (WHO grade II) are diffusively infiltrative brain tumors arising from glial cells. Spatial classification that is usually based on cerebral lobes lacks accuracy and is far from being able to provide some pattern or statistical interpretation of their appearance. In this paper, we propose a novel approach to understand and infer position of low-grade gliomas using a graphical model. The problem is formulated as a graph topology optimization problem. Graph nodes correspond to extracted tumors and graph connections to the spatial and content dependencies among them. The task of spatial position mapping is then expressed as an unsupervised clustering problem, where cluster centers correspond to centers with position appearance prior, and cluster samples to nodes with strong statistical dependencies on their position with respect to the cluster center. Promising results using leave-one-out cross-validation outperform conventional dimensionality reduction methods and seem to coincide with conclusions drawn in physiological studies regarding the expected tumor spatial distributions and interactions.

1 Introduction

Low-grade gliomas (WHO grade II) are diffusively infiltrative brain tumors that are generally revealed by seizures in young patients with a normal social and professional life [1]. Although surgery of low-grade gliomas has been a controversial subject for decades, it is now considered as the best therapeutic option as reported in the recent European Guidelines [2]. However, surgery can cause functional deficit, especially if the tumors are located near or within a functional area. For that matter, the tumor's location in the brain is of great importance.

Inferring spatial dependencies regarding tumor appearance in the brain is a research direction that has gained little attention [3,4]. The most natural approach to such a problem is through statistical modeling with respect to the position of low-grade gliomas. This could be achieved using dimensionality reduction techniques like principal [5] or independent [6] component analysis. This

* This work was supported by ANRT (grant 147/2010), Intrasense and the European Research Council Starting Grant Diocles (ERC-STG-259112).

can be easily achieved at the voxel level, but a huge number of observations is needed and the linearity assumption regarding the correlation between the spatial positions of tumors is imposed. Non-linear methods embed the observation space into a low-dimensional space and then seek for correlations in this space. Examples refer to isomap [7] or laplacian graphs [8]. Within the considered clinical setting, neither the expected rank of the reduced space is known nor the number of samples is sufficient to approximate the manifold.

Network connectivity analysis is an alternative to the above mentioned methods, a quite popular idea in functional imaging when targeting task-specific brain understanding through mutual activations of brain regions [9]. In such a context, the brain is parceled according to a certain criteria and considered to be a fully connected network, each node is associated with a multi-dimensional variable explaining the activations as a function of time/task. The aim is then to group parcels with important statistical correlations in terms of behavior [10].

In this paper, we are inspired from these methods but we amend them to deal with a static setting. While the brain parcels are known in functional imaging, in our case these parcels do depend on the observations and correspond to manually annotated low-grade gliomas. Our method first registers all samples to the same reference pose while explicitly taking into account the tumor position during the registration. The registered data are considered to define a statistical measure of coherence between different tumors based on their spatial position as well as their geometric form. This measure is used to determine a graph where messages are exchanged between nodes. The strength of the message depends on the statistical measure of coherence between tumors, while the overall capacity of a node depends on the total volume of messages being passed to it. The problem of understanding spatial dependencies between tumors appearance is then casted as an unsupervised clustering problem[11] where both the number of clusters, the cluster centers and the nodes assignments are to be determined.

The remainder of this paper is organized as follows: in section 2 we detail the preprocessing step and introduce the spatial position representation network. The optimization of the network is discussed in section 3 while experimental validation and comparisons with linear methods are presented in section 4. Discussion concludes the paper.

2 Network Representation of Low-Grade Gliomas Spatial Dependencies

Our data set consisted of 95 3D MRI images of 95 different patients with low-grade gliomas: 90 FLAIR T2 weighted, and 5 T2 weighted. We had a majority of male patients, age between 21 and 65 and tumor size between 3.5 and 123 cm^3 . The image size ranged from 256x256x24 to 512x512x33, and the pixel resolution from 0.4x0.4 to 0.9x0.9 mm^2 in the (x,y) plane and 5.3 to 6.4 mm in the z plane. Each image had been manually annotated by experts to indicate the position of the tumor. The atlas for registration consisted of a high resolution FLAIR image (size 256x256x24, resolution 0.9x0.9x5.45 mm^3) of a tumor free brain.

Let us consider without loss of generality n acquisitions $V_i()$ of brain volumes as well as the corresponding segmentations $S_i()$ binary maps indicating whether a pixel belongs or not to the tumor. Both, intensity volumes and binary maps are rigidly registered to the same pose. The first step towards spatial position mapping of low-grade gliomas consists in removing the brain local anatomical variability through deformable registration. This is achieved using the discrete optimization method presented in [12] that is amended to deal with the lack of visual correspondences in the brain tumor areas by not taking into account tumor voxels during registration. Using the optimal deformable registration parameters, let us consider now $S(\mathbf{x}) = S(T(\mathbf{x}))$ being the deformed segmentation map.

We then measure the proximity between two tumors by adopting as suggested in [13] the Mahalanobis distance. Let us consider the spatial coordinates of all pixels belonging to tumors S_i and S_j . Let $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{x}}_j$ be the center of mass of S_i and S_j . The Mahalanobis distance between the tumors is computed as:

$$d_M(S_i, S_j) = \sqrt{(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)^T \Sigma^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)} \quad \Sigma = \frac{(n_i - 1)\Sigma_i + (n_j - 1)\Sigma_j}{n_i + n_j - 2} \quad (1)$$

where Σ_i and Σ_j are the covariance matrices of pixels coordinates of S_i and S_j while n_i and n_j are the number of pixels in tumors S_i and S_j respectively. The complete network corresponding to our data-set is shown in [Fig. (1)] where the nodes correspond to the tumors and the strength and color of the edges correspond to the proximity of the observed tumors.

3 Network Optimization

Let us consider the set of distances $d(S_i, S_j)$ between all pairs of low-grade gliomas. Let us assume that the observations can be expressed through a compact sub-graph of k central nodes. Both the number and the position of the central nodes are to be determined while the remaining observations should be expressed from one of these nodes. In order to determine the quality of a central node, we adopt two measurements, a global and a local one. The global measurement assumes that a central node should have significant overlap with the remaining measurements. This can be quantified by the overall distance of the node from the rest of the data. The assignment of an observation to one of the central nodes (local measurement) should be based on their distance. The optimal network connectivity should minimize the local and the global criteria both in terms of central node selection and tumor assignments. This is an unsupervised clustering problem,

$$\min_k \min_{c_1, \dots, c_k} \min_{l_1, \dots, l_n} \left(\sum_{i=1}^n d(c_i, S_j) + \alpha \sum_{j=1}^k \delta(c_i - l_j) d(S_{l_j}, S_j) \right) \quad (2)$$

where l_j is the assignment of observation j and α is a constant coefficient balancing the contributions of the two terms. To recover the lowest potential of the

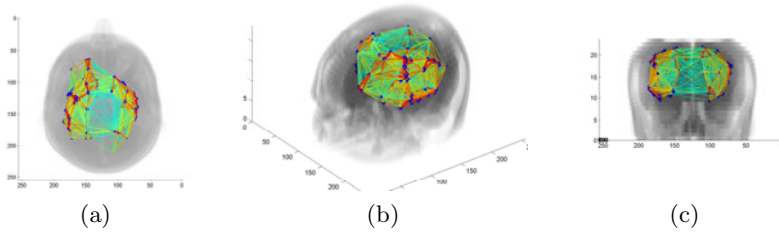


Fig. 1. Network representation of the data-set. Arcs corresponding to a distance greater than 5 are not displayed.

above functional, we will rely on a recently proposed clustering algorithm [11], which does not require any initialization and provides near-optimal clustering results both in terms of number of central nodes as well as in terms of remaining nodes assignments. The optimization results will heavily depend on the relative importance given between the two terms (*alpha* value). In order to determine the optimal number of central nodes, one should consider the quality of clusters being associated with them. We consider the following criteria to determine the compactness of clusters and the quality of the overall representation:

Dunn Index [14]: For a partition of *K* clusters, it is computed as

$$D = \min_{i \in [1:K]} \left\{ \min_{j \in [1:K], j \neq i} \left\{ \frac{d(c_i, c_j)}{d_{max}} \right\} \right\} \tag{3}$$

where d_{max} corresponds to the maximal distance of a sample to the center of the cluster it belongs to, and $d(c_i, c_j)$ is the distance between the centers c_i and c_j of clusters *i* and *j*. Intuitively, a good clustering will be characterized by a high Dunn index (see [Fig. (2b)]): compact and well separated clusters yield a low d_{max} and a high distance inter clusters.

Davies-Bouldin Index [15]: this index also identifies compact and well separated clusters. It computes the maximum similarity between clusters:

$$DB = \frac{1}{K} \sum_{i=1}^K \max_{j \in [1:K], j \neq i} \frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \tag{4}$$

where σ_i is the average distance of all points in cluster C_i to its center, *K* is the number of clusters and $d(c_i, c_j)$ is the distance between the centers of clusters *i* and *j*. A small DB value indicates little similarities between clusters and therefore, a better clustering ([Fig. (2c)]).

Global Silhouette Index [16]: For a given cluster $C_k = (S_1, \dots, S_n)$, the silhouette index assigns to each of its members a quality measure $s(S_i)$ ($i=1, \dots, n$) that is a confidence indicator on the membership of S_i in C_k . It is defined as

$$GS = \frac{1}{K} \sum_{j=1}^K \frac{1}{n_j} \sum_{i=1}^{n_j} s(S_i) \text{ with } s(S_i) = \frac{b(S_i) - a(S_i)}{\max(a(S_i), b(S_i))} \tag{5}$$

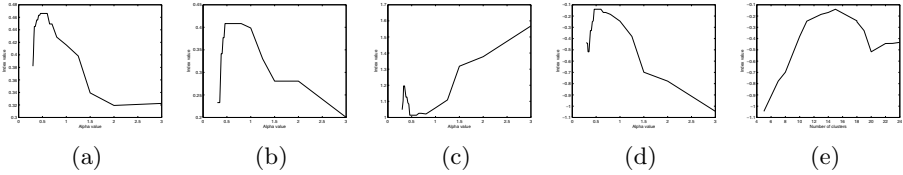


Fig. 2. Values of the 3 criteria in function of α (a,b,c,d) and of the number of clusters (e). (a) Global Silhouette index, (b) Dunn index, (c) Davies-Bouldin index, (d,e) Combination of the three indexes.

where n_j is the number of elements in cluster C_j , $a(S_i)$ is the average distance between S_i and all of the remaining elements in C_k , and $b(S_i)$ is the minimum average distance between S_i and all of the elements clustered in C_j , ($j = 1, \dots, K; j \neq k$). $s(S_i)$ takes values between -1 and 1. A value close to 1 indicates that S_i has been assigned to the appropriate cluster and a value close to -1 infers that S_i has been misclassified. A value close to zero suggests that S_i lies equally far away from 2 clusters. Since the largest silhouette index is indicative of the best clustering, we select the clustering that yields the maximum global silhouette index ([Fig. (2a)].

We have considered increasing relative importance between the global and the local term as shown in [Fig. (2)]. All the considered criteria reached their best value for the same 15 nodes graph-structure. The optimal network connectivity is shown in [Fig. (3)]. One can observe a notion of symmetry between the left and the right hemisphere. The distribution of the individual distance values before clustering is shown in [Fig. (4a)] while [Fig. ((4b)), and ((4c))] show the distribution per cluster of distances and individual silhouette indexes.

4 Experimental Validation

In order to evaluate the performance of the method we have considered a leave-one-out cross validation strategy. We have used $n - 1$ measurements to learn the topology of the graph and the remaining observation to predict its position in the network. We have performed this test $n = 95$ times.

We denote 3 different results that satisfy the 3 optimality criteria. In most cases, the number of central nodes has reached the same value as for the whole data-set, that is $K = 15$. In one case, the optimal number of clusters was 16. The removed sample is a large tumor (about 110 cm^3) that is the center of an important cluster (12 nodes) which splits in 2 without its center. The removal of 7 samples yielded 14 clusters networks. Those samples were assigned to small clusters (2 or 3 nodes) that were merged with a neighboring cluster when the sample is removed.

We have considered 2 criteria to determine the ability of the retained centers to express their respective populations. First the cluster members should be

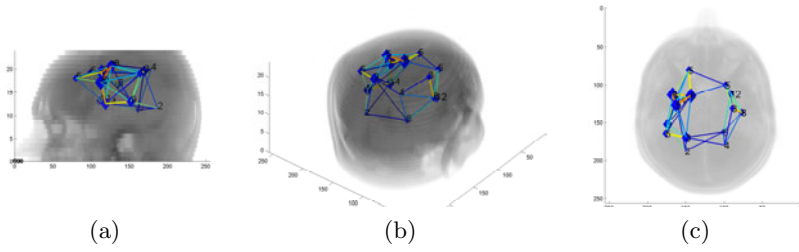


Fig. 3. 15 nodes network corresponding to the optimal criteria, connections corresponding to distances bigger than 4 are not displayed

overlapping with the center and their cluster's center c_k should be the closest, i.e. $d(S_i, c_k) = \min_{j \in [1:K]} d(S_i, c_j)$ and $d(S_i, c_k) \leq 2$. Secondly, we use the individual silhouette index $s(S_i)$ to evaluate the membership of a sample to a cluster. If the value is high, there is no doubt about the membership of the sample. If it is close to 0, we compare its value to the average silhouette index of the corresponding cluster : $s(S_i) \geq 0.3$ or $s(S_i) \geq \frac{1}{n_k} \sum_{j=1, j \neq i}^{n_k} s(S_j) - std(s_k)$, where $std(s_k)$ is the standard deviation of individual silhouette index values for cluster C_k , and n_k the number of elements in C_k . Such a configuration was able to properly assign 80% of the whole training example.

In order to evaluate the network prediction strength, we have considered the cluster that optimally represents the removed observation, and measured the quality of the cluster once this new sample has been added to it. Ideally, the quality of the clusters should remain the same if the network is able to express the variability of new samples. This was the case in 73% of the 95 cases. Failure can occur if the new sample is an outlier, equally close to 2 existing clusters or was the center of a cluster. We also estimate the effect of the removal of a sample on the clustered graph's structure by evaluating the quality of correspondences between the clustered graph obtained from the whole data-set G_0 , and the 15 nodes graphs obtained from cross validation experiments $G_k, (k = 1, \dots, 87)$. To this end, we seek a matching between the nodes of the graphs by using the algorithm proposed by [17]. We find complete match in 83% of the cases and only one node didn't correspond for the remaining graphs.

Finally, comparison with principal component analysis (PCA) was considered. PCA computes a new orthogonal coordinates system that regroups the maximum variance in a minimum vectors. It is a simple way to find correlation between data: the fewer vectors necessary to represent the data, the bigger the correlation. We performed PCA on the data-set at the voxel level. Each $[256, 256, 24]$ binary segmentation map was converted to a $[256 \times 256 \times 24, 1]$ vector, so that each voxel was a variable and each image was a sample. While our experiments using network representation suggest that 15 clusters represent 80% of the data, 15 PCA vectors regroup only 65% of it. Another drawback of PCA was that the obtained vectors did not correspond to specific preferential locations.

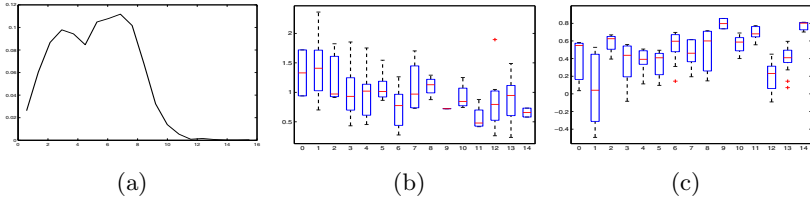


Fig. 4. Distance distribution between all the tumors (a), Box-plots of the distances (b) and individual Silhouette (c) values for each cluster

5 Discussion

Very few studies have dealt with spatial dependencies of gliomas appearances in the brain. In this paper, we proposed graph theory to represent our data-set spatial dependencies and clustering to regroup statistically dependent tumors. Cross-validation results on an important volume of clinical data support the idea that the complete graph could be reduced to a handful nodes, which indicates statistical preferential locations for low-grade gliomas in the brain that our clustered network enables to identify. There are very few tumors in or near the occipital and prefrontal lobes (7%). We find (symmetrically) the higher amount of tumors around the insula or the temporal lobe (27% in the right hemisphere and 33% in the left hemisphere). The remaining tumors are in the frontal and parietal lobes, mostly close to the motor areas. Those results are consistent with previous observations [4]. Furthermore, preliminary results indicate that the graph's structure remains unchanged while at the same time the method outperforms standard dimensionality reduction techniques.

Several open questions remained unanswered from this study. Despite the sizable validation set, increasing the number of cases considered in the study could further enhance the claims of the paper. On a more theoretical view-point, the impact of the registration process is critical since the selection of the reference pose introduce a strong bias on the results. The use of population registration methods [18] that simultaneously deform all the data while taking into account tumoral regions is a promising alternative. The estimation [19] and propagation of registration uncertainties is another mean of eliminating the bias while at the same time producing a qualitative interpretation of the results. This can also produce better means of measuring spatial coherence between tumors through high-dimensional embedding and Gaussian processes distance definition. Creating a statistical representation of tumor appearances can be of great interest and a valuable clinical tool for computer aided diagnosis. The network obtained through leave-one-out cross validation, through n-to-m graph matching and network topology optimization could lead to a unique representation for the spatial mapping of low-grade gliomas in the brain. Automatic segmentation is a task that could automatically benefit from the obtained graphical model since it

consists of a powerful position prior queue. Another possible outlook would be to integrate functional data and observe the consequences of a tumor on the functional brain and its plasticity.

References

1. Duffau, D.: Surgery of low-grade gliomas: towards a 'functional neurooncology'. *Current Opinion in Oncology* 21(6), 543–549 (2009)
2. Sofetti, R., Baumert, B., Bello, L., et al.: Guidelines on management of low-grade gliomas: report of an EFNS-EANO task force. *Eur. J. Neurol.* 17, 1124–1133 (2010)
3. Larjavaara, S., Mantyla, R., Salminen, T., et al.: Incidence of gliomas by anatomic location. *Neuro-oncology* 9, 319–325 (2007)
4. Duffau, H., Capelle, L.: Preferential brain locations of low-grade gliomas. *Cancer* 100, 2622–2626 (2004)
5. Shlens, J.: A tutorial on Principal Component Analysis. *Systems Neurobiology Laboratory, Salk Institute for Biological Studies* (2005)
6. Hyvarinen, A., Oja, E.: Independent component analysis: algorithms and applications. *Neural Netw.* 13, 411–430 (2000)
7. Tenenbaum, J.B., Silva, V., Langford, J.C.: A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290, 2319–2323 (2000)
8. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: *NIPS* (2001)
9. Zheng, X., Rajapakse, J.C.: Learning functional structure from fMR images. *Neuroimage* 31, 1601–1613 (2006)
10. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198 (2009)
11. Komodakis, N., Paragios, N., Tziritas, G.: Clustering via LP-based Stabilities. In: *NIPS*, pp. 865–872 (2008)
12. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through MRFs and efficient linear programming. *Medical Image Analysis* 12, 731–741 (2008)
13. Pokrajac, D., Megalookonomou, V., Lazarevic, A., Kontos, K., Obradovic, Z.: Applying spatial distribution analysis techniques to classification of 3D medical images. *Artificial Intelligence in Medicine* 33, 261–280 (2005)
14. Dunn, J.C.: Well separated clusters and optimal fuzzy partitions. *J. Cybern.* 4, 95–104 (1974)
15. Davies, D.L., Bouldin, D.W.: A cluster separation measure. *IEEE Trans. on PAMI* 1(4), 224–227 (1979)
16. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65 (1987)
17. Torresani, L., Kolmogorov, V., Rother, C.: Feature correspondence via graph matching: Models and global optimization. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008, Part II. LNCS*, vol. 5303, pp. 596–609. Springer, Heidelberg (2008)
18. Sotiras, A., Komodakis, N., Glocker, B., Deux, J.-F., Paragios, N.: Graphical models and deformable diffeomorphic population registration using global and local metrics. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009. LNCS*, vol. 5761, pp. 672–679. Springer, Heidelberg (2009)
19. Glocker, B., Paragios, N., Komodakis, N., Tziritas, G., Navab, N.: Optical Flow Estimation with Uncertainties through Dynamic MRFs. In: *CVPR* (2008)