Modeling MS Lesion Objects for Medical Information Exchange

Tilmann Steinberg[1], Zhengyi Le[1], Yi Ouyang[1], Zhifeng Wang[1], Wei Zheng[1], Fillia Makedon[1], Heather Wishart[2], Andrew Saykin[2], Euripides Loukis[3]

[1] The Dartmouth Experimental Visualization Laboratory (DEVLAB), Dept. of Computer Science, Dartmouth College, Hanoverppsala University, Medical Informatics and Engineering, University Hospital, Uppsala, Sweden

[2] Brain Imaging Laboratory, Dartmouth Medical School, Hanover

[3] University of the Aegean, Samos, Greece

Abstract

Current medical technologies allow for capturing patient brain pathology in scans of several different types at the same visit, yielding multiple modality image sequences from which objects of interest, such as multiple sclerosis (MS) lesions, can be extracted. It is of interest to model the development of these objects (e.g., lesions, nodules or tumors) over time and in different modalities, in order to characterize and group them. Grouping lesions of different behavior patterns allows a better understanding of the development of the disease. Information exchange between clinicians and computational experts is possible through a common collection platform that takes into account disparity of data and need for integration as well as security. This paper focuses on the issues related to collaboration with regulated data. We describe an approach based on converting private data into an abstracted form that unprivileged researchers can use.

J. Qual. Life Res. Vol. 3, Issue 2, 2005

INTRODUCTION

Brain disease analysis is complex, rich in imaging media, and in need of analytical tools. Cooperation between medical and computational groups has yielded good results in the past, but faces challenges related to complying with new federal regulations on privacy, and limits in local data set size. Frequently, the physical separation between computational and medical collaborators presents a practical limitation on how closely groups can work together. The problem, then, is to develop methods for processing data to enhance collaboration and seamless exchange of data.

In this paper, we describe an infrastructure to facilitate exchange of medical imaging data amongst physically distributed medical and computational collaborators, using MS neuroimaging as an example. We provide the medical team with data collection and management tools to produce clean data that can then be used by outside collaborators to build advanced analysis tools. Our framework identifies patterns of change in MS lesions based on abstractions of the objects' location, shape, mass distribution, and relative positioning, and assigns extracted objects from new patient scans to the most appropriate group of objects with similar multi-modal features. Augmenting this data modeling framework are ontologies that are built through information extracted from the lesions. Ontologies are used to model different phases of lesions and classify new scans into learned categories. We show how this framework can be utilized under current privacy regulations and in telemedicine.

RELATED WORK

Previous work in analyzing MS lesion data has focused on development in individual modalities. One approach has modeled intensity evolutions on individual voxels to produce values characteristic for normal and pathological areas (Rey, et al., 2001), and used displacement fields in time series data to perform segmentation in new scans (Rey, et al., 2002). Another approach employs a volumetric analysis technique using nonrigid deformation computation and flow-field analysis to detect and measure changes in lesions over time (Thirion and Calmon, 1999). One of the fundamental problems that a model should solve is change detection. Depending on number of patient visits, there are two categories: one for small number of visits, another for a series of visits. Models working

for a series create 4-D data by precise registration of serial volume data sets. Statistical model is used then for time series analysis to monitor changes (Gerig et al., 2000, Meier and Guttmann, 2003). Models working with few visits focus on identifying localized changes in either large structures or small ones (Bosc et al., 2003). Miller et al. (2004) investigated the benefits and the development requirements of web-accessible ontologies for discrete-event simulation and modeling. Soo et al. (2003) described algorithms for creating an automated semantic annotation by parsing natural language content descriptions. Hyvonen et al. (2002) described the structure of an ontology used for both images and queries that supplies similar images to users. Moenne-Loccoz et al. (2004) proposed a framework to accommodate developments in content-based multimedia retrieval and methods to handling an overwhelming volume of temporal data. Soo et al. (2004) investigated the feasibilities and advantages of semantic retrieval and automated ontology acquisition from semantically annotated poems. Khan et al. (2004) proposed a concept-based model using domain-dependent ontologies and an automatic mechanism for selecting appropriate concepts that describe media documents as well as user requests.

MS Lesions

Basic project idea. Our primary goal is to capture multiple scan types at same time to get a more comprehensive view of lesions and their development. Each scan type contributes particular aspects of each lesion; by combining these aspects, we can see different stages of lesion development. Table 1 shows examples of medical implications of different combinations of lesion manifestations on imaging.

Table 1. Interpretation of Neurobiology based on Lesion Appearance in Scan Types.

Flair	Post Gad	SPGR	Interpretation
1	1	1	severe, currently active pathology
1	0	0	chronic, nonactive pathology
0	1	0	active pathology

(Source: Brain Imaging, Laboratory, Dartmouth Medical School).

The development of analysis tools is complicated (on the computational side) by privacy regulations that require limited access to the collected data: only appropriate local clinical users are allowed access to all data. For others, all private data have to be stripped out first. As we still want to include outsiders in developing new solutions, we have to come up with methods to perform the data cleaning in an efficient way that does not unduly tax the clinical users. We have two kinds of data sets: (a) data registered for same-patient analysis, to track individual development of lesions; and (b) data registered for cross-patient analysis, to compare developments over multiple subjects. While the latter meets privacy regulations, as the brain and skull volume is resampled into a standard space, the former is more desireable as cross-subject registration is likely to introduce distortions.

Data and Development Modeling

Goals. Abstract data modeling aims to provide the computational group with real data with which to work, so the medical group needs to have a tool to extract data from the scans and other input. The medical group also needs to specify appropriate problems to be solved (e.g., to what extent do active lesions predict the subsequent evolution of local tissue destruction); and develop measures of quality for solutions.

Data modeling and change detection. Challenges in MS lesion segmentation include the fact that intensity values within lesions may be similar to those of noise or other elements in the scan. Certain MS lesion segmentation approaches assume prior knowledge, such as that MS lesions typically occur in a periventricular distribution within the white matter of the brain on a given type of scan. MS lesion segmentation remains a challenging task. Change detection can be realized in both 2D and 3D. In 2D we assume pixel model while in 3D we assume voxel model. Pixel model is the result of image capturing (MRI) and image segmentation. Voxel model is an extension to pixel model. In addition to the tasks required by pixel model, connectivity analysis between adjacent slides is usually required. We will employ this connectivity as a means of countering effects of noise, since random noises are usually not clustered.

Image registration is required for change detection. While rigid transformation modeling has been extensively studied and utilized, brain tissues are not rigid, thus deformable models may be more suitable. We will assess a simple approach employing both landmarks (along the borders of brain tissues) and thin plates. In certain cases, it is beneficial or necessary to map brain tissues to a standard reference system, such as the Montreal or Talairach atlas (e.g., for crosssubject analyses). If we have a set of registered images over time, we can calculate time intensity curves (TIC) for each individual point of interest. TIC can be used as a way to distinguish changes. For certain area of interest, we can get several TICs (blue curves in Figure 1). We compute the average TIC (red curve) and analyze it to decide whether there are changes.

Patterns of individual 3D shapes can be captured by shape descriptors: shape distributions, reflective symmetry descriptors, and spherical harmonics. Each approach may contribute unique information. Given the hetero-

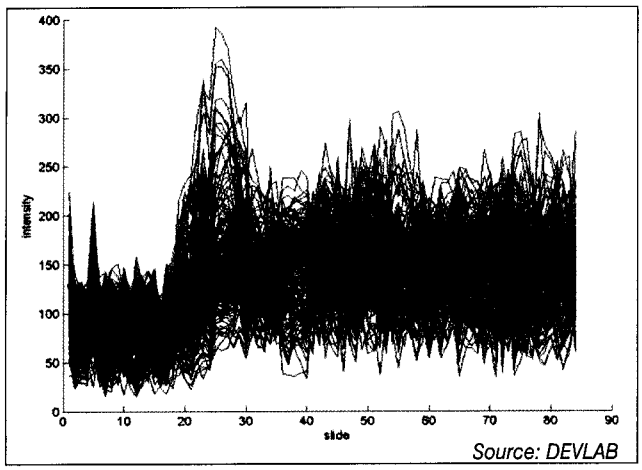

Figure 1. Time Intensity Curves for a Monitored Region and their Average.

geneity in presentation and progression of MS lesions, an approach using both shapes and their relationships seems most appropriate. 2D spatial relationship can be partial modeled using the R-histogram approach. We extend R-histogram by adding the 3rd dimension. We add "view points" (plus the origin) on the lattice of xyz coordinate system and convert the xyz coordinate system into spherical coordinate system (see Figure 2; red dots represent view points). Each voxel in the target object can be viewed from each view point and recorded as distance, longitude and latitude tuple. These tuples form a space that represents the whole spatially distributed system. In this representation both shapes and their relationships are modeled. Apparently this representation is very sensitive to origin and coordinates, but assuming registration is fairly accurate, it could serve as a foundation for change detection.

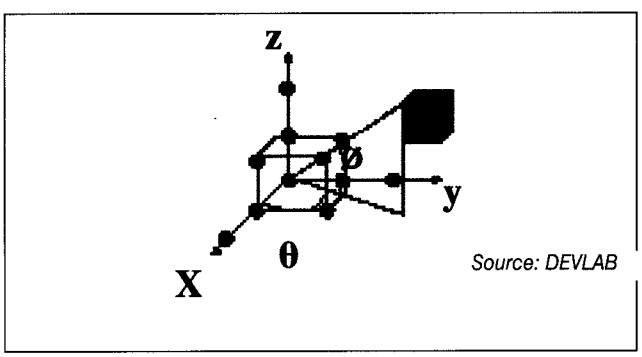

Figure 2. Spatial Relationship between Objects in 3D as expressed by Polar Coordinates.

Basic approach for tracking lesion development. For each scan, we obtain expert-generated segmentation, from which individual lesion objects are taken. Based on criteria for proximity and overlap, we group these individual objects into sets of objects that represent the same physical lesion. For the objects within each of these sets, we (the computational group) can describe differences and changes in lesions using the methods described above. There are two basic tasks, classification and similarity search (simple classification: by present modalities), to answer the interesting question for clinicians: what classes does a lesion go through over the course of its tracked development? Especially when coupled with information about treatment, this is helpful in understanding the disease and its consequences, and how it can be addressed.

Using a combination of multiple approaches. Spherical harmonics are good for identifying similar objects with invariance to rotation, but not as good for directional differences; bitmap division or histograms along major axes can identify specific differences but requires that objects are aligned, e.g. in the case of MS lesions, within the same subject, and with multiple scans aligned to each other (which is relatively easy since there are only translational and rotational variances). Our approach uses a combination of these techniques: we apply spherical harmonics to bitmap differences to create signatures of "change objects" (i.e., a signature for the differences between two objects), as well as simplifications of the bitmap differences. How data model ensures privacy of original subject. The data shared with the computational group consists of specific lesion objects only, with only limited information about the surrounding brain or skull tissue. For example, segmented binary lesion maps yield insufficient information to reconstruct original skull/face features from which participants could be identified, and therefore protect patient privacy consistent with federal regulations.

MS Ontologies

An ontology describes a set of concepts and their relationships in a formal structure. It is a specification of an abstract, simplified view of the world, either domain-dependent or generic. As such, the ontology approach offers an alternative method of representing information about MS lesions than the previously described modeling approaches, while at the same time offering a way to clearly distinct between private and public data.

MS Ontologies. After MS images are annotated, the meta-data can be extracted from the images and put into an ontology. As a result, the related concepts about MS lesions can be constructed into a taxonomy, which can be used as a reference when computing the similarities between objects such as different lesions located in different areas with different shapes. Figure 3(a) is an example of an ontology representing locations and types of MS lesions. The "Brain" is divided into several different areas, within which there may be several different types of lesions (e.g. based on immunopathological properties). MS

lesions of different types may have different attributes in terms of shape and size. Figure 3(b) shows an example of attributes of concepts for a given type of lesion which has "shape", "size", and other attributes. The values of "size", "shape", and other additional attributes are knowledge about that particular type of lesion that we can extract from annotated MS lesion images. The attributes of a concept can be used to compute similarity between a concept and a real object in order to determine which concept is the proper one for the target object. In this example, we can estimate which type the given MS lesion is by comparing its values of different attributes to different concepts. The values of attributes of a stage can be learned from previous scans that have been evaluated by experts. Thus, these ontologies give us a way to store knowledge of MS lesion types and may facilitate evaluation of new scans.

We use C to denote a concept. *Ai* denotes the *i-th* attribute of concept *C*. The similarity between a concept and an object is

$$SIM(C,O) = \frac{1}{q}\sum_{i=1}^{q} S(A_i)$$

where *S(Ai)* denotes the satisfaction value of attribute *Ai*. It depends on the value domain of attribute *Ai*. Given a target lesion and its location, the area of this lesion can be determined first. Then by using this formula, the similarity value between target lesion and different concepts representing different stages of lesions can be computed. The stage with maximum similarity value is the estimated current stage of target lesion.

Sharing Data

Advantages of sharing data. Converting data into abstract form can improve workflow between the cooperating groups, as newly created data sets become available for inclusion in analysis tools as soon as they are converted. Manually uploading or downloading data sets is no longer necessary, eliminating communication bottlenecks (assuming that previously agreed upon access rights have been set up). In return, data owners immediately benefit from updated analysis results.

Trigger mechanisms for notification or processing of new shared data. As data becomes available, triggers set to detect new files or database entries are used to either notify human collaborators or automatically start processing jobs for the new data.

Conclusions and Future Work

Contributions. We discussed the issue of collaboration with regulated data; and proposed a solution in which we convert private data into abstracted form that unprivileged researchers can use for development.

Future work. Further work is required to develop additional forms of abstracted data to facilitate collaboration, to incorporate other scan types, such as diffusion tensor imaging (a structural imaging modality with new attributes) and functional MRI (a non-structural modality), and to incorporate additional biological (e.g., genetic, immunologic) assays of relevance.

Acknowledgements

This material is based in part upon work supported by the National Science Foundation under award numbers IDM 0308229 and ITR 0312629. Any opinions, findings,

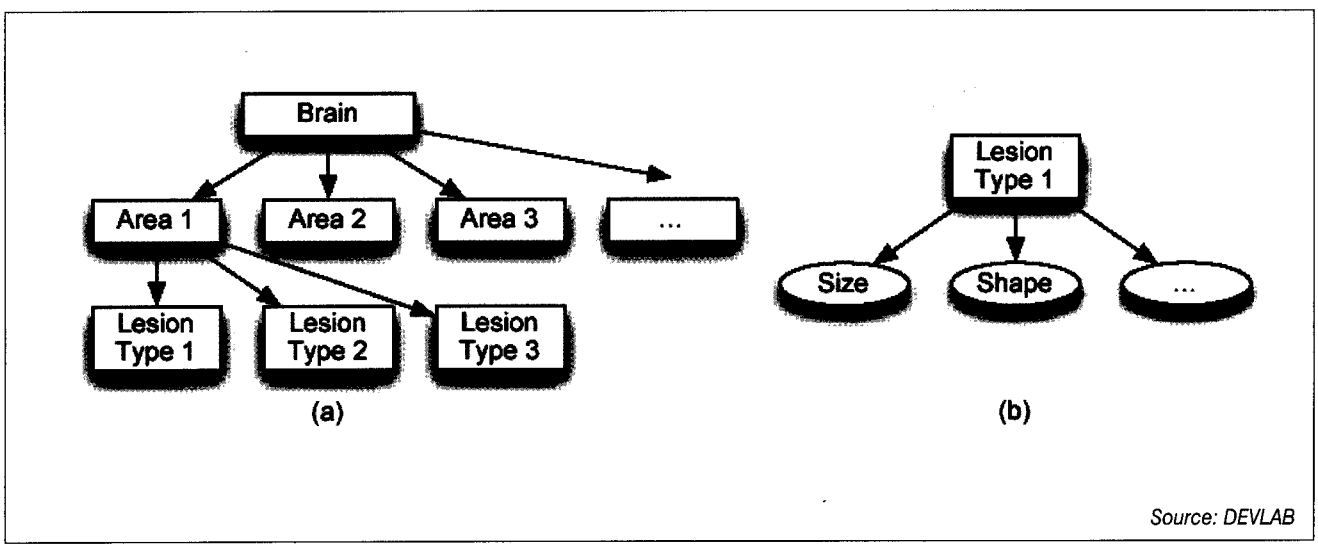

Figure 3. An Ontology Example of MS Lesions **(a)**, and Attributes of One Concept **(b)**.

and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Rey, D., Stoeckel, J., Malandain, G., Ayache, N. (2001). A spatio-temporal model-based statistical approach to detect evolving multiple sclerosis lesions. *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis* (pp. 105-112). Kauai, Hawaii.
2. Rey, D., Subsol, G., Delingette, H., Ayache, N. (2002). Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Medical Image Analysis 6* (2002), 163-179.
3. Thirion, J.-P., Calmon, G. (1999). Deformation analysis to detect and quantify active lesions in threedimensional medical image sequences. *IEEE Transactions on Medical Imaging,* 18 (5), 429-441.
4. Gerig, G., Welti, D., Guttmann, C. R. G., Colchester, A. C. F., Szekely, G., (2000). Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial MR data. *Medical Image Analysis* 4, 31-42.
5. Bosc, M., Heitz, F., Armspach, J.-P., Namer, I., Gounot, D., Rumbach, L., (2003). Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution. *Neuroimage* 20 (2), 643-656.
6. Meier, D. S., Guttmann, C. R. G., (2003). Time-series analysis of MRI intensity patterns in multiple sclerosis. *Neuroimage* 20 (2), 1193-1209.
7. Miller, J., Baramidze, G., Sheth, A., Fishwick, P., (2004). Investigating ontologies for simulation modeling. *Proceedings of the 37th Annual Symposium on Simulation*
8. Hyvönen, E., Styrman, A., Saarela, S., (2002). Ontology-based image retrieval. *HIIT Publications*, 2002-03, pp. 15-27. Helsinki Institute for Information Technology (HIIT), Helsinki, Finland.
9. Soo, V., Lee, C. , Li, C., Chen, S., (2003). Automated semantic annotation and retrieval based on sharable ontology and case-based learning techniques. *Proceedings of the third ACM/IEEE-CS Joint Conference on Digital libraries*. pp. 61-72.
10. Moënne-Loccoz, N., Janvier, B., Marchand-Maillet, S., Bruno, E., (2004). Managing video collections at large. *Proceedings of the 1st International Workshop on Computer Vision Meets Databases*
11. Soo, W., Yang , Y., Chen , L., Fu, T., (2004). Ontology acquisition and semantic retrieval from semantic annotated chinese poetry. *Proceedings of the 2004 joint ACM/IEEE conference on Digital Libraries*, Tuscon, AZ, USA
12. Khan, L., McLeod, D., Hovy, E., (2004). Retrieval effectiveness of an ontology-based model for information selection. *The International Journal on Very Large Data Bases 13*