

A Spatio-temporal Multi-modal Data Management and Analysis Environment for Tracking MS Lesions

Tilman Steinberg¹, Yuhang Wang¹, Fillia Makedon¹, Li Shen¹,
Andrew Saykin², Heather Wishart²

¹ *Dartmouth Experimental Visualization Laboratory (DEVLAB),
Department of Computer Science*

² *Brain Imaging Laboratory, Dartmouth Medical School
Dartmouth College, Hanover, NH 03755*

{tilmann, wyh, makedon, li}@cs.dartmouth.edu, {saykin, wishart}@dartmouth.edu

Abstract

We describe the development of a system that automates data collection, metadata extraction and analysis of spatio-temporal multi-modal data, combining data management and data analysis to provide an efficient resource for clinicians. Though the system is extensible to many applications, the current focus is on managing Multiple Sclerosis (MS) lesion data, which are disparate streams of image, numeric, and text data. In order to discover patterns of MS pathology and plan early and effective treatment, multispectral magnetic resonance (MR) image streams collected over time need to be correlated efficiently with each other and with patient performance and clinical data streams.

1. Introduction

Magnetic Resonance Imaging (MRI) is typically used in conjunction with performance tests to monitor the development and treatment of Multiple Sclerosis (MS), a brain disease affecting about 1 in 1,000 people in the USA. The discovery of common patterns and correlations between image data (showing pathology) and clinical data (showing effects) is paramount for finding effective monitoring strategies and treatments for the disease. Most previous quantitative lesion time series analysis approaches have used only single MRI modalities [1-3]. Currently, there are many similarity search algorithms on multi-attribute time series data [4]. In 2D, a lesion indexing approach using Force Histograms has been introduced [5].

In our approach, different MRI modalities are used to reveal lesion pathology. Currently, heterogeneous data streams are collected and analyzed manually. Our approach provides: (a) a common representation of heterogeneous data; (b) standardized metadata extraction to represent changes in lesions at different time points; (c) integration of data streams (scans, performance tests, and treatments); and (d) analysis tools for pattern extraction.

2. Data Collection & Management

We have the following input data streams for image data: scan images from n subjects (patients or controls) acquired in m sessions (ideally equidistant in time), each with l modalities. Scans are traced or segmented by k experts or tools to identify lesions from which basic features such as location and volume are extracted. At the same date the scans are generated, clinical test data are collected.

From these streams, we extract feature information to create vector sequences, translating the problem of finding similar medical cases to that of finding similar paths in an n -dimensional space (a solution to this kind of problem is described in [6]). In the case of MS lesions, this requires quantifying lesion changes in and applying data fusion techniques to multiple imaging modalities, and data stream correlation analysis for the discovery of drug effects on lesions. These techniques are being implemented in the system as research components to help improve prognosis, disease tracking, and evaluation of response to treatment in MS.

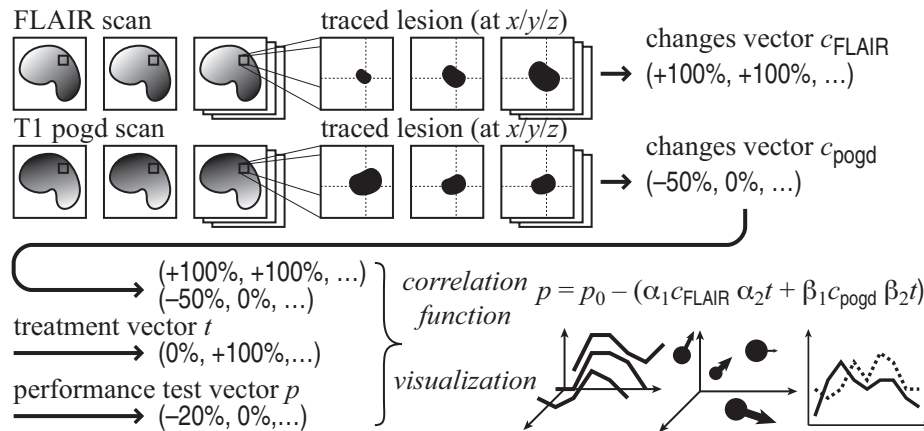


Figure 1. Tracking pathology by combining multiple modalities. MS lesions are captured and traced, and the changes in their features between temporal data points are encoded as “change vectors”. These are combined with vectors calculated from other data streams such as dosage of treatment or clinical tests to yield results in form of a correlation function or visualizations.

3. Extracting Features and Finding Patterns

The overall goal is to correlate effectiveness of a treatment or drug by tracking lesion development and clinical data from performance tests. To run analysis tools, we need a complete set for n_1 subjects, m_1 sessions, and l_1 modalities before applying a given tool. From this set, features are extracted that describe MS lesions, such as but not limited to location (of center of mass, relative to brain structures), size, average/maximum diameter, and surface area (in 3D).

We then calculate changes from time t_i to t_{i+1} for each feature, using feature-specific difference functions, and compare these *change vectors* across lesion objects from the same or different subjects (see Figure 1 for an example combining two imaging modalities with two non-image data streams). In particular, we are interested in finding *patterns of changes* in a given set of features that can be reproduced across subjects, so that a new subject can be placed along a timeline of the condition based on a library of existing subject data and relatively little data from the new subject, and to allow correlation of patterns with other, non-image based factors.

4. Conclusion

We have outlined a system to collect, manage, and analyze heterogeneous data streams for research in MS, with the goal to allow better evaluation of subjects' progress and correlation of image data with clinical data. The aim is to allow correlation of very different data streams by focussing on patterns of changes in features.

References

- [1] G. Gerig, D. Welti, C. R. G. Guttman, A. C. F. Colchester, and G. Szekely, “Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial MR data,” *Medical Image Analysis*, vol. 4, pp. 31-42, 2000.
- [2] R. Kikinis, C. R. G. Guttman, D. Metcalf, W. M. Wells, G. J. Ettinger, H. L. Weiner, and F. A. Jolesz, “Quantitative follow-up of patients with multiple sclerosis using MRI: Technical aspects,” *Jmri-Journal of Magnetic Resonance Imaging*, vol. 9, pp. 519-530, 1999.
- [3] S. Prima, N. Ayache, A. Janke, S. J. Francis, D. L. Arnold, and D. L. Collins, “Statistical Analysis of Longitudinal MRI Data: Applications for Detection of Disease Activity in MS,” *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Tokyo, Japan, 2002.
- [4] T. Kahveci, A. Singh, and A. Gurel, “Similarity searching for multi-attribute sequences,” *Proceedings of the 14th International Conference on Scientific and Statistical Database Management*, Edinburgh, Scotland, 2002.
- [5] C.-R. Shyu and P. Matsakis, “Spatial lesion indexing for medical image databases using force histograms,” *Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii, 2001.
- [6] M. Vlachos, G. Kollios, and D. Gunopulos, “Discovering similar multi-dimensional trajectories”, *Proceedings of the 18th IEEE International Conference on Data Engineering*, San Jose, California, 2002.