

# Exploring the discrimination power of the time domain for segmentation and characterization of lesions in serial MR data

<sup>1</sup>Guido Gerig, <sup>2</sup>Daniel Welte, <sup>3</sup>Charles Guttman, <sup>4</sup>Alan Colchester and <sup>2</sup>Gábor Székely

<sup>1</sup>Computer Science, UNC Chapel Hill, NC

<sup>2</sup>Swiss Federal Institute of Technology ETHZ, Switzerland

<sup>3</sup>Brigham and Women's Hospital, Harvard Medical School, Boston

<sup>4</sup>University of Kent at Canterbury, Kent, England

gerig@cs.unc.edu, dwelte@vision.ee.ethz.ch

This is a submission to the MedIA journal, June 7, 1999

Correspondence Address:

Guido Gerig, Ph.D.  
Department of Computer Science  
University of North Carolina  
CB#3175, 219 Sitterson Hall  
Chapel Hill, NC 27599-3175, USA

Phone: (919) 960 0330

Fax: (919) 962 1799

e-mail: gerig@cs.unc.edu

www: <http://www.cs.unc.edu>

**Abstract.** This paper presents a new methodology for the automatic segmentation and characterization of object changes in time series of three-dimensional data sets. The purpose of the analysis is a detection and characterization of objects based on their dynamic changes. The technique was inspired by procedures developed for the analysis of functional MRI data sets. After precise registration of serial volume data sets to 4-D data, we applied a new time series analysis taking into account the characteristic time function of variable lesions. The images were preprocessed with a correction of image field inhomogeneities and a normalization of the brightness function over the whole time series. This leads to the hypothesis that static regions remain unchanged over time, whereas local changes in tissue characteristics cause typical functions in the voxel's time series. A set of features are derived from the time series and their derivatives, expressing probabilities for membership to the sought structures. These multiple sources of uncertain evidence were combined to a single evidence value using Dempster Shafer's theory. Individual processing of a series of 3-D data sets is therefore replaced by a fully 4-D processing. To explore the sensitivity of time information, active lesions are segmented solely based on time fluctuation, neglecting absolute intensity information. The project is driven by the objective of improving the segmentation and characterization of white matter lesions in serial MR data of multiple sclerosis patients. Pharmaceutical research and patient follow-up requires efficient and robust methods with high degree of automation. Further, an enhanced set of morphometric parameters might give a better insight into the course of the disease and therefore leads to a better understanding of the disease mechanism and of drug effects. The new method has been applied to two time series from different patient studies, covering time resolutions of 12 and 24 data sets over a period of roughly one year. The results demonstrate that time evolution is a highly sensitive feature to detect fluctuating structures.

## 1 Introduction

Serial magnetic resonance imaging of patients becomes attractive due to the minimal invasive image acquisition, the speed-up in scanning and therefore patient time, and the high spatial and tissue resolution. The time series reveal information about significant changes of diseased anatomical regions, about the changes as an effect of a drug or radiotherapy treatment, or about subtle morphological changes caused by a neurophysiological disease. The temporal sampling thus not only provides information about morphological but also functional changes.

A typical analysis of this type which is routinely applied is the analysis of functional MRI data sets. A patient is stimulated with a specific time pattern of visual, auditory or motor activity. Brightness changes due to local changes in the oxygenation state of blood are expected to show a similar time pattern and can be detected by a correlation of the stimulus function with the time series of each pixel. Here, the signal processing aims at finding the best discrimination between noisy steady state signals and signals correlated with the stimulus [1]. The processing most often assumes that a patient doesn't move during the examination, although slight object motion due to breathing, pulsation of the heart and swallowing is unavoidable. It has been shown that a sub-voxel correction of 3-D motion [2] can considerably improve the voxel-based time-series analysis.

Pharmacological studies or patient follow-up and monitoring, are different. Time frequency is not in the range of seconds, but can be days, months or even years. The study of a tumor change in relation to chemotherapy or radiotherapy, for example, typically requires time intervals of weeks till months. In schizophrenia, temporal changes are studied over long periods by imaging a patient with yearly scans.

The development of a new segmentation technique is driven by the motivation to get a better understanding of the disease process in multiple sclerosis (MS). Research in MS already demonstrated the power of using serial imaging [3]. Drug development for multiple

sclerosis uses serial MRI as one measurement among other diagnostic features to study the temporal changes of white matter lesions in the central nervous system. A series of patients is divided into two groups getting either placebo or the new drug. Patients are scanned in intervals of 1, 2 or 4 weeks during a period of about one year. The significance of tests is increased by multi-center studies, collecting image data from various hospitals using a standardized MR protocol. Image data are examined by radiologists, evaluating each scan in relation to the previous one to visually assess the occurrence of new lesions. A quantitative analysis of the total lesion load and of the single lesions is performed using interactive user operated segmentation tools. A typical study often comprises up to several thousands of 3-D data sets. The manual outlining of lesions in large number of series of 2-D slices is not only time consuming but also tedious and error prone. Errors for the segmentation of small structures are often in the range of the volume of the observed structures.

Automated image segmentation systems have been proposed by several groups [4–7]. They consist of well-designed sequences of processing steps, including preprocessing, bias-field correction, feature-space clustering of multi-echo MRI data [8], and a matching of a statistical anatomical atlas [9, 10] to solve ambiguities of statistical classification. As a result, they present a significantly improved reproducibility and therefore a reduced inter- and intra-rater variability and allow an efficient processing of large amount of data.

Previous segmentation methods mostly intend to segment lesions from single data sets, not taking into account the significance of correlation in the time domain. In radiological examination on the light-box, however, experts use previous scans of patients to decide about significant changes. An early attempt to consider the correlation in the time domain was presented by Metcalf et al. [11] by proposing a 4-D connected component labeling on registered segmented label images. The procedure serves as a postprocessing filter applied after individually segmenting the data sets, removing insignificant lesion candidates which appear only at one time point, or eliminating 4-D lesion patterns with volume below a predefined threshold. The aim still was an improved lesion segmentation, although the 4-D connectivity additionally could give access to time domain information.

So far, temporal changes in signal intensity patterns of multiple sclerosis lesions have not been used to improve and simplify the processing of time series. Guttmann [3] presented a seminal paper on characterizing the evolution of lesions in serial MR data, suggesting to use this valuable information for image processing. The present paper will *explore the time domain information* inherently given by serial MR data sets. The major question in research of disease mechanisms or drug studies is most often not a segmentation of static tissue or static lesions but of temporal changes. We claim that dynamic changes in lesion voxels can be detected by analyzing the time series of each voxel, assuming perfectly registered and normalized data sets. Although the ultimate goal will be a spatio-temporal analysis of the 4-D data sets, this paper only focuses on evaluating the discrimination power of the temporal domain.

Besides exploring time domain as a new feature for segmentation, we are working towards extracting a rich set of morphometric parameters. These include temporal information to analyze the time course of the disease, to understand time correlations of lesion groups and lesion patterns, to determine the lesion load versus time, and finally to combine the results with anatomic atlas information to describe major spatial categories

(periventricular, deep white matter, cortical) of lesions. Scientific visualisation of dynamic changes will be important to visually assess the disease course of individual patients.

The paper is organized as follows. Section two shortly describes the preprocessing including bias correction and image brightness normalization, and the combination of serial 3-D to 4-D data sets. The new time series analysis is explained in section three. Section four presents results obtained with data sets from different pharmaceutical studies. Details about the combination of multiple uncertain evidence sources are explained in the appendix.

## 2 Combination of serial 3-D data to 4-D data

Individual magnetic resonance volume data sets acquired in weekly to monthly time intervals can be combined to 4-D  $(x, y, z; t)$  data sets, which allows the application of time-series analysis of single voxels.

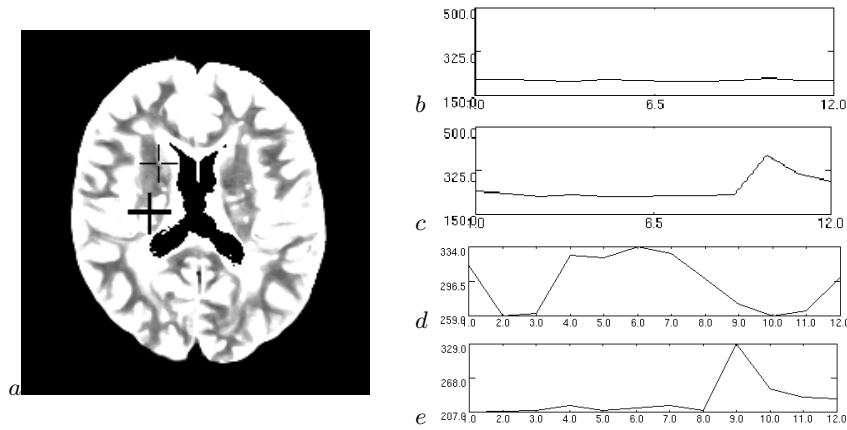
*Registration* The serial data sets obtained from the Brigham and Women’s Hospital Boston (cf. section 4.1) have been registered by the INRIA research group using crest-line extraction and matching [12]. A second serial data set presented in this paper is processed by the KUL research group using the MIRIT registration software package [13] which maximizes the mutual information between corresponding voxel intensities. Both registration methods work fully automatically. The transformation matrices are input to a geometric transformation which performs trilinear interpolation.

*Image brightness normalization and bias correction* The corruption of the image brightness values by a low-frequency bias field often occurs in MR imaging and impedes visual inspection and intensity-based segmentation. A mathematical model for bias correction using parametric bias field estimation was proposed in [14]. We assume the original scene to be composed of tissue regions with homogeneous brightness only degraded by noise. The estimation of the parametric bias field is formulated as a non-linear energy minimization problem. Input parameters are the statistics (mean, standard deviation) of expected categories. Using the same set of input parameters for each data set from series of volume images results in a combination of bias correction and brightness normalization. The presence of strong striping artifacts on one of the data sets required a two step procedure by first correcting for brightness changes between individual slices and then for the 3-D bias field [15].

*Result of Preprocessing* The normalization of brightness and correction of inhomogeneity artifacts results in sets of corrected 3-D data sets. After registration, they are combined to form 4-D data sets. Picking a voxel and visualizing its time course gives a good impression of the quality of the preprocessing. We assume that the signal intensity of white matter should remain constant (figure 1b), whereas voxels representing active lesions would show considerable changes (figure 1c-e).

## 3 Time series analysis to detect fluctuating lesions

Bias correction, image brightness normalization and spatial registration of serial 3-D image data results in 4-D  $[x, y, z; t]$  data sets. The preprocessing yields a spatial and intensity-



**Fig. 1.** Time series of voxels for healthy white matter (b) and typical lesions (c-e). **Horizontal axis: time, vertical axis: MR intensity.** Image (a) illustrates a typical MR slice presenting white matter lesions. The positions of the voxel generating the constant time series (b) is marked with a thin cross, the lesion time series (c) with the thick cross. Plots d and e represent time series of other lesions voxels.

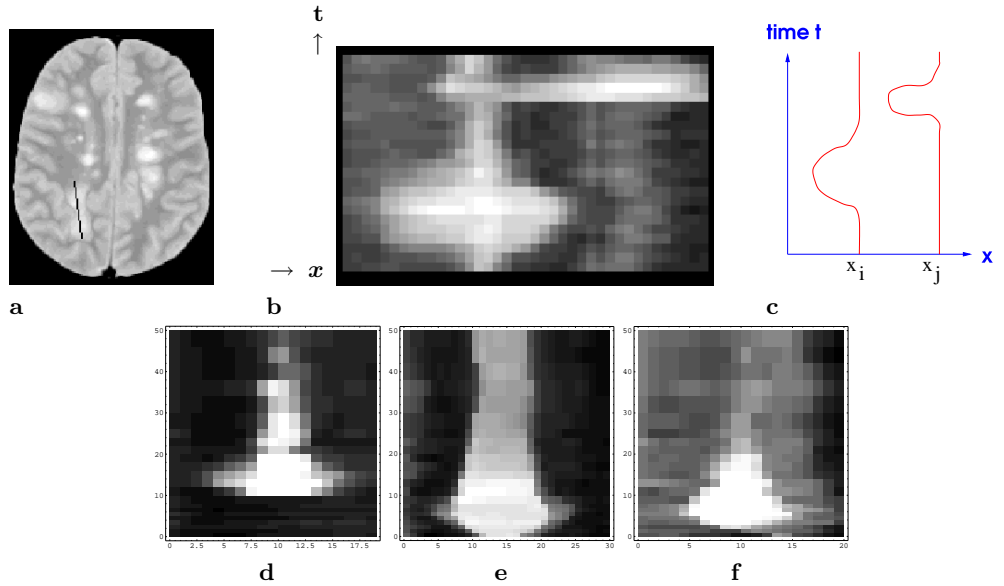
based normalization of the time series. Therefore, we can assume that static tissue will not change brightness over time, whereas voxels which are part of fluctuating lesions will depict typical variations. Each voxel can be considered as a time series, suggesting the application of methods for one-dimensional signal processing. The signal analysis shows similarities to the postprocessing of functional magnetic resonance data (fMRI), but there is one significant difference. Functional MRI is measured by a repetitive stimulation of a certain task, which allows a comparison of the stimulation function with the time series of each image pixel, most often using correlation techniques. The time course of MS lesion voxels, on the other hand, does not follow a fixed pattern and can only be characterized by a dynamic fluctuation of image brightness.

### 3.1 Visualization of brightness changes

The time course of lesion voxels can be studied by providing two-dimensional images of arbitrary profiles through 3-D image data versus time. The displays illustrate fluctuations of profiles over a typical time period of one year (Fig. 2). Tissue boundaries in general show very small spatial displacements which can be explained by elastic tissue deformations, whereas some boundaries in the vicinity of lesions can demonstrate larger deformations due to a mass effect (see Fig. 2b lower middle). A characteristic feature for lesion time series is a continuous fluctuation with time, presenting increasing and decreasing time changes or both.

Based on observations of typical time series of lesion voxels we developed features that describe fluctuations. The set of features will be used for discriminating between static tissue and active lesions.

**Brightness Difference:** A simple calculation determines the minimum and maximum brightness for each time series and calculates the absolute difference  $\Delta I = |I_{max} -$



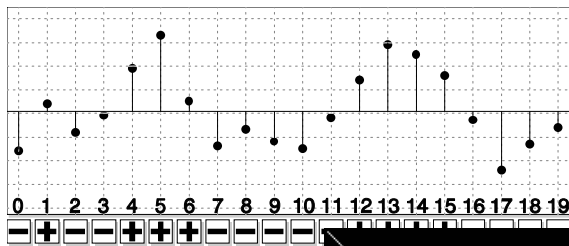
**Fig. 2. Visualization of spatiotemporal lesion evolution. a** Original image with profile, **b** space-time display (horizontal: spatial axis, vertical: time axis) and **c** sketch of typical time series at locations  $x_i$  and  $x_j$ . Other typical spatial lesion evolutions are displayed in **d**, **e** and **f**.

$I_{min}$ . This feature measures the maximum contrast change of a time series within the observed time period (Fig. 3a).

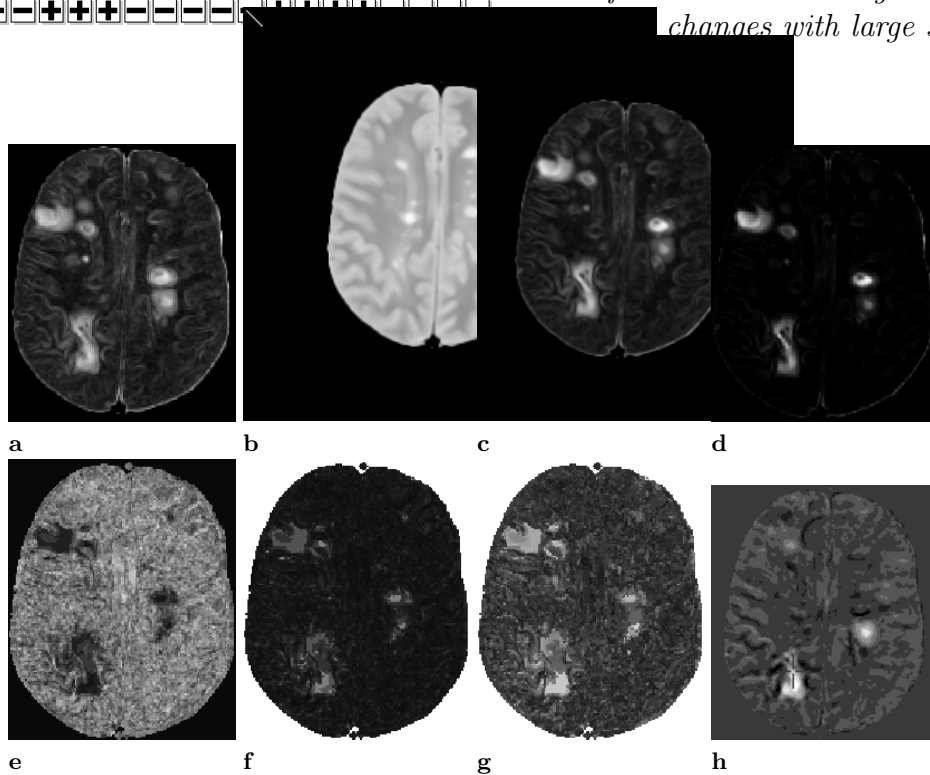
**Statistical measures:** Mean, standard deviation and variance form a set of statistical features expressing the temporal variation of brightness around the mean value. We expect much higher variance for lesion voxels than for static tissue (3b,c,d).

**Signs of fluctuation around mean:** The features discussed so far do not consider the temporal pattern or the frequency of fluctuations. We therefore determine the number of zero-crossings of the zero-mean time series and evaluate the time length of positive and negative segments. A noisy static signal will generate a large number of sign changes with small segments, whereas large fluctuations will generate a small number of long segments (3e,f,g).

**Time derivatives:** The gradient of the time function provides information about the rate of change, both for decreasing and increasing events. Fig. 2 illustrates that lesions often appear with a large brightness change. We used the minimum and maximum gradient as features for our lesion analysis (3h). The attributed time will be further used for displaying temporal evolution (see results).



*Evaluation of sign changes of zero-mean time series: Analysis of the sequence of signs: Nr. of “segments” (7), maximum (5), minimum (1) and average segment length (2.86). Observation: Low-frequency variations (lesions) result in large differences from the mean and generate a small changes with large segments.*



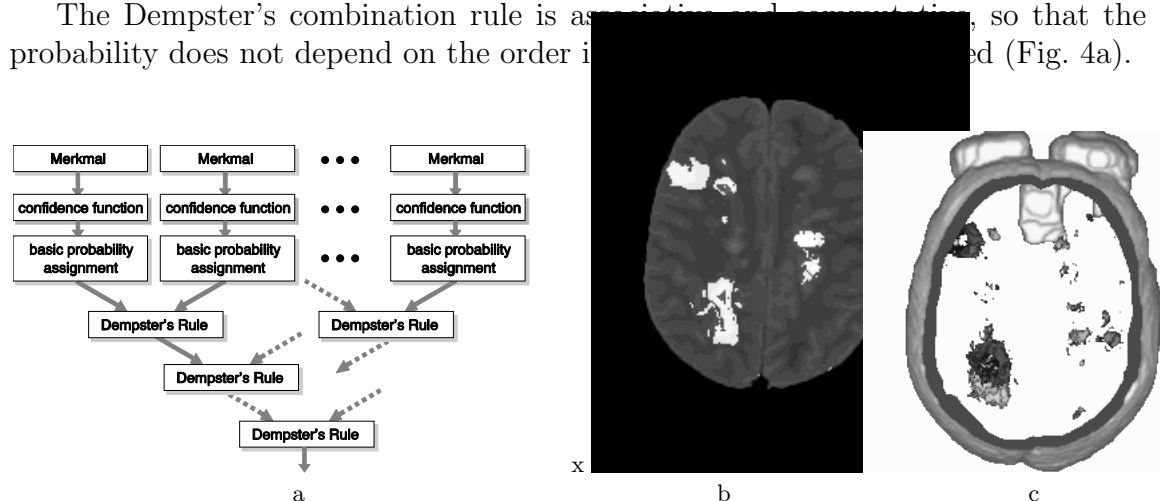
**Fig. 3.** 2-D cuts of 3-D feature maps: Difference Max-Min (a), Mean (b), standard deviation (c) and variance (d), Nr. of zero-crossing segments (e), length of minimum segment (f), length of maximum segment (g), and maximum absolute time gradient (h)

### 3.2 Evidence accumulation by combining uncertain measurements

The multiple features derived by signal processing provide probabilistic maps of the likelihood to characterize the sought structures (Fig. 3a-h). Each of these features is inherently uncertain, and they must somehow be combined to derive a measurement which incorporates different properties of the typical temporal pattern of a lesion. A pooling of evidence from different knowledge sources will strengthen beliefs in some cases and erode beliefs in others, even handling contradictory evidence. The following analysis assumes that the features are independent, although this might not be strictly true. A combination of probability measures can be accomplished by using Dempster-Shafer’s theory. To get around the computational complexity of the original DS method [16, 17], we used binary frames of discernment (BFOD) as proposed by [18]. Details describing the choice of confidence

factor functions (cf), basic probability assignments (cfa) and the combination rules can be found in the appendix. The design of these functions and probabilities (see Table 3) represents a crucial step. However, our tests with the analysis of very different serial data sets showed that only minor parameter adjustments were necessary. The initial design and training was based on a comparison of the resulting feature maps with segmentation results produced by statistical classification followed by manual corrections.

The Dempster’s combination rule is associative and commutative, so that the final probability does not depend on the order in which the features are combined (Fig. 4a).



**Fig. 4.** Combination of fuzzy features by Dempster’s rule (a), segmented active lesions on 2-D slice (b) and as 3-D rendering (c).

The combined 3-D data set is again probabilistic, with a value range of  $[0, \dots, 1]$  (Fig. 4b). A binary segmentation, for example for three-dimensional graphical visualization (Fig. 4c), is obtained by choosing an appropriate threshold either by visual inspection of overlay images or by comparing the segmentation output to hand-segmented training data. Tests with multiple data sets and visual inspection showed that the choice of the final threshold was not critical and revealed very similar results within a range of thresholds, provided a careful design of the cf-functions and bpa assignments.

## 4 Results

The new segmentation system has been applied to two time series from different patient studies. A first study carried out at the Brigham and Women’s hospital covers 40 patients with 24 brain scans, with a fixed sequence of scanning intervals of one, two and 4 weeks. Another study currently analyzed in the European BIOMORPH project [19] comprises 12 serial imaging sessions of 40 patients, each imaging session delivering multiple MR protocols (PD,T1,T2). The data sets are preprocessed as described in section 2 and analyzed using the signal processing methods described in section 3.

### 4.1 Brigham and Women’s Hospital data sets

The image data sets were acquired on a GE Signa 1.5 Tesla using a double echo spin echo pulse sequence (TR 3000ms, TE 30/80ms) half Fourier sampling (0.5 NEX). 54 slices



with 3mm slice distance and thickness and 256x256 pixels result in voxel dimensions of 0.96x0.96x3 mm<sup>3</sup>. The time series includes 24 data sets acquired over a period of 50 weeks with a specific time protocol: weekly scans for 10 weeks followed by every other week scans for 16 weeks and monthly scans for 24 weeks. We could use 23 out of the 24 scans for our analysis. The unequally spaced time image series was converted into a regularly sampled sequence by linear interpolation.

The 3-D visualizations (Fig. 6) display the time course of lesion evolution, coded as a color map ranging from 1 to 23. Additionally, the processing results in a quantification of the temporal evolution of the total lesion load, measured relative to the first image data set. Please remind that the procedure *only measures time changes* and excludes voxels that remain unchanged, thus providing information that is different from the conventional total lesion load over time.

## 4.2 BIOMORPH data sets

Image data are acquired on a Philips T5 magnetic resonance scanner, 1.5 Tesla, using a double echo spin-echo pulse sequence with TR 2816ms and TE 30/80ms). 24 axial slices 256x256 were measured, with voxel dimensions 0.9x0.9x5.5mm<sup>3</sup>. 12 scans are measured over a period of 56 weeks: 11 scans with approximately 4 weeks intervals and a last scan with a 13 week interval. This unequally spaced time image series was converted into a regularly sampled sequence by linear interpolation.

Figures 7 and 8 illustrate the segmentation result, again attributing the segmented lesions with the time of appearance.

## 5 Validation

Validation of the time series analysis method is not straightforward as it measures temporal changes rather than absolute lesion volumes. It has been decided to validate the method using simulated lesion patterns in 4-D data sets. We used a 3-D magnetic resonance data set of a healthy volunteer which has been replicated by the number of time steps, here 12. Remaining insufficiencies of the clinical 4-D data sets resulting from the spatial registration and the bias correction are therefore not subject to this validation. It has been visually verified that the dataset doesn't contain white matter lesions.

Lesions of varying size, contrast and lifetime have been produced with the help of a non-linear diffusion process with a source and a sink term. This decision has been influenced by the visual analysis of realistic temporal patterns (Fig. 2). Lesions in magnetic resonance image data are always diffuse and most often showed a sharp increase and a smooth decrease of the size and brightness function.

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = \text{div}(c(u(\mathbf{x}, t))\nabla(u(\mathbf{x}, t))) + q(\mathbf{x}, t) + r(u(\mathbf{x}, t)) \quad (1)$$

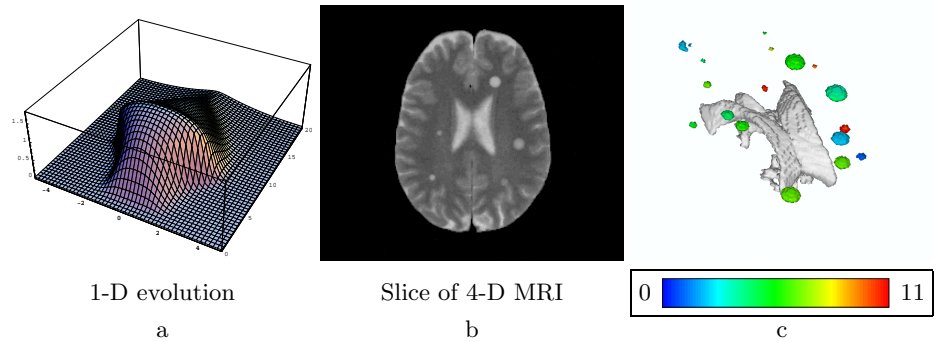
Equation 1 produces time series of diffuse lesions with quick appearance and slower disappearance (Figure 5a). The time series of the artificial lesions were merged with the time series of the head by taking the voxel with the maximal intensity of the sources. Due

to this operation, voxels which are only slightly brighter than the surrounding tissue are also counted as lesion voxels and therefore allowed to implant very diffuse lesions. The lesions were inserted in both the white matter and in the grey matter of the brain.

The 4-D data set (including the evolution of the lesions) has then been corrupted by Gaussian noise with a variance equally to the variance of white matter (Fig 5b). The parameters for the temporal analysis were chosen exactly the same as for the other segmentations. As a result, the number of detected lesion voxels per lesion and their time point of appearance and disappearance have been obtained. The 4D test data set consists of 12 time points and 18 lesions were implanted. All inserted lesions have been detected (Fig 5c).

Table 1 lists the maximal number of implanted/detected voxels during the time course for each lesion. For most generated lesions about 80% of the implanted voxels are detected, since very diffuse lesion voxels have been discarded by the time series analysis. For very small or weakly lesions of weak contrast the ratio of very diffuse lesion voxels to well contrasted lesion voxels is large and thus the detection rate in percent becomes worse.

Table 2 compares the number of implanted and detected voxels for different lesions at each time point. Beside the problems with the diffuse contours of the lesions it can also be seen that the estimate of the lifetime of the detected lesion voxels is often too short. The lifetime of a voxel is determined by calculating the minimal and maximal gradient of the time course of this voxel and does not exactly match with the time points of first appearance and final disappearance.



**Fig. 5.** Simulated lesions in 4-D MR data sets. Image **a** illustrates the non-linear diffusion with the one-dimensional profile as the horizontal axis time varying from front to back. Image **b** represents a cut through the 4-D data set with simulated lesions, and image **c** represents a 3-D display of the segmented lesions attributed by the time of appearance.

lesion number	implanted voxels	detected voxels	percentage of detected voxels
1	226	178	79%
2	600	463	77%
3	85	67	79%
4	450	352	78%
5	256	217	85%
6	31	15	48%
7	100	79	79%
8	67	48	72%
9	31	15	48%
10	31	18	58%
11	770	641	83%
12	797	649	81%
13	905	771	85%
14	254	133	52%
15	142	89	63%
16	69	44	64%
17	15	12	80%
18	160	113	70%

**Table 1.** Simulation experiment: Number of implanted versus detected voxels per 4-D lesion

lesion number	maximal intensity	implanted/detected	number of detected/implanted lesion voxels at each time point												
			1	2	3	4	5	6	7	8	9	10	11	12	
1	1300	implanted detected						39 3	192 178	226 178	78 1	6 1	1 1		2
2	900	implanted detected			119 2	600 10	565 463	134 394	1 1	1 1	1 1	1 1	1 1		1
3	1000	implanted detected						4 4	28 352	79 352	85 67	37 61	1 1		5 1
4	1100	implanted detected			1	1	1	97 4	387 352	450 352	176 1	26 1			
5	1200	implanted detected					1	38 1	79 35	192 177	256 217	114 7	44 1		1
6	1000	implanted detected								1 1	9 15	18 15	31 15		11
7	1100	implanted detected		4	28 14	80 79	100 79	61 1	16 1	4 1	1 1	1 1	1 1		1
8	900	implanted detected						7 11	29 48	64 42	67 42	33 1	12 1		4
9	1200	implanted detected									1 5	5 9	9 9		31 15
10	1200	implanted detected						5 1	31 16	31 18	9 1	1 1			
11	1200	implanted detected		1	1	2	2	2	2	163 8	671 641	770 641	375 7		79 2
12	1200	implanted detected		1	1	2	2	2	52 55	232 631	666 649	797 20	474 17	198	36 1
13	1300	implanted detected		18 3	120 3	301 120	681 639	905 771	731 310	456 24	249 2	84 2			9 1
14	900	implanted detected			60 15	254 133	196 95	38 2	2 2	1 1	1 1	1 1			1
15	1000	implanted detected		20 3	142 89	88 31	1 1								
16	800	implanted detected										4 4	25 18		69 44
17	900	implanted detected							5	15 12	14 11	5			
18	900	implanted detected										3 1	50 36		160 113

**Table 2.** Simulation experiment: Time course of implanted versus detected voxels per lesion. The intensity ranges of the MR image were  $\approx 500$ -610 for WM,  $\approx 650$ -800 for GM and  $\approx 1000$ -1300 for CSF.

## 6 Summary and Conclusions

We present a new image processing system for the time series analysis of serial data sets representing time series. The purpose of this project was the exploration of the discrimination power of the time axis, which is most often not directly used for the segmentation of structures. On purpose, we excluded any absolute scalar or multi-spectral information about structures as most often used for the voxel-wise segmentation of lesions from MRI by multi-dimensional thresholding and statistical clustering. Here, we exclusively analyzed the time series of each voxel to demonstrate the additional information obtained by taking into account the temporal evolution of brightness.

The paper describes the development of a the image analysis techniques for segmentation of fluctuating structures from 4-D data sets. Analyzing the time series of each voxel, we derive a set of statistical and structural features each of which discriminates static tissue from changes in the time function. The extraction of each feature creates a probability map for the presence of the sought structure. The multiple probabilities from the different evidence sources are combined using The Dempster-Shafer theory. We selected this technique because it allows to combine different sources of evidence by considering not only the probability of the occurrence of a feature, but also of the absence and of the ignorance about the measurements. The design of the confidence factor functions and the transformation of confidence factors into basic probabilities represent a decisive step which is comparable to supervised training in statistical classification. Test showed that once trained, these settings can be used for other data sets as well since measurements do not directly depend on absolute intensity values. Further, brightness and contrast of our data sets are normalized in the preprocessing step.

The analysis of normalized 4-D data sets is automatic and takes about 10 minutes processing time (SUN Ultra 1 with 128Mb). The results were visually compared with results from alternative segmentation methods and revealed a surprisingly good sensitivity and specificity to MS lesions. However, we have to keep in mind that our analysis so far is *only based on time series information of one MR echo*. We can expect an even higher sensitivity if multi-echo information could be embedded, and if we would combine the time series analysis with the segmentation of spatial structures. So far, data was inspected by clinical experts by evaluating overlays of the segmented lesions with the original MR scans (Fig. 4). The quantitative validation and tests with simulated lesions show that all the lesions could be detected, but that the lifetime of segmented is consistently shorter than the lifetime of the implanted lesions. However, our lesions were not binary blobs but were generated by a non-linear diffusion process which starts with a very weak contrast. Thereby, we tried to simulate lesions in real images which are often nearly invisible at early stages.

We can conclude that temporal changes represent a highly significant feature for the identification of active lesions and should be considered for future analysis. Further, temporal evolution and the detection of time changes are the most important features for pharmaceutical studies and research, as the goal most often is the evaluation of changes due the disease process or a drug treatment. Besides detection of lesion voxels, our method reveals the time of appearance and disappearance as attributes to each voxel. A dynamic visualization of this temporal information allows the detection of groups and patterns of

lesions which show a similar time course. If additionally combined with anatomical atlas information to link lesion positions to anatomy, we would get a new insight in the MS disease process and hopefully a new understanding of the disease mechanism.

Currently, we are extending the time-series analysis by spatial analysis to develop a spatio-temporal description of fluctuating lesion patterns. We will also include information from multiple spectral MR channels (PD, T1, T2, FLAIR) to replace the scalar image brightness by vector-valued measurements.

## Acknowledgment

The European Union and the Swiss funding agency BBW are acknowledged for supporting this research in the international project on Brain Morphometry (BIOMORPH, EU-BIOMED2 project no BMH4-CT96-0845). We would like to thank Ron Kikinis and Ferenc Jolesz, Brigham and Women’s Hospital, for providing a serial MR data set from a study supported by NIH (contract no. N01-NS-0-2397). Philip Thirion and Nicholas Ayache, INRIA Sophia Antipolis, are acknowledged for providing a registered time series. We thank Dirk Vandermeulen and Frederick Maes, KUL Belgium, for registering a serial data set as part of the BIOMORPH project. Dr. Kappos and Dr. Radü from the University Hospital Basel are acknowledged for providing their expertise and for making research funding available.

## A Dempster Shafer (DS) Theory

The DS theory [16, 17] makes inferences from incomplete and uncertain knowledge, provided by different independent knowledge sources. The theory allows the strengthening or erosion of beliefs by combining additional sources of evidence, even in the presence of partly contradictory evidence. It can be shown that the Dempster-Shafer theory contains the Bayesian theory of partial belief as a special case.

The problem of lesion detection allows a simplification, as the set of admitted answers, the so called *frame of discernment*, is simply the set  $\Theta = \{lesion, -lesion\}$  of mutually exclusive elements and represents the special case of a *binary frame of discernment* (BFOD) [18]. The set of all subsets of  $\Theta$  is called the power set  $2^{|\Theta|}$  of  $\Theta$  with the focal elements  $(\{\emptyset\}, \{lesion\}, \{-lesion\}, \{\Theta\})$ , where  $\{\Theta\}$  stands for  $\{lesion, -lesion\}$  or ignorance.

*Basic probability assignment (bpa)*: A bpa represents a belief in an elementary proposition or in a disjunction of several of them. Formally, a bpa is a function  $m : 2^\Theta \rightarrow [0, 1]$  with  $0 \leq m(\cdot) \leq 1$ ,  $m(\emptyset) = 0$ , and  $\sum_{A \subseteq \Theta} m(A) = 1$ , where  $\emptyset$  is the null proposition. A belief function  $Bel(A)$  over  $\Theta$  is defined by a sum of the bpa’s of all proper subsets of  $A$ ,  $Bel(A) = \sum_{B \subseteq A} m(B)$ , and calculates our total belief in a proposition  $A$ .

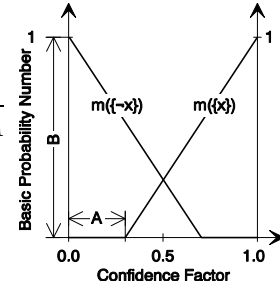
*Combination of evidence*: Dempster’s rule calculates a new evidence from two given basic probability assignments  $m_1$  and  $m_2$ , designated as  $m_1 \oplus m_2$ . The products of the bpa’s with focal elements with non-empty intersection are normalized with  $\frac{1}{1-\kappa}$ , where  $\kappa$  is the sum of the products with empty intersection.  $\kappa$  can be expressed as a measure for the *contradiction* or *inconsistency* of the combined evidence.

$$\begin{aligned}
m_1 \oplus m_2\{x\} &= \frac{m_1\{x\}m_2\{x\} + m_1\{x, \neg x\}m_2\{x\} + m_1\{x\}m_1\{x, \neg x\}}{1 - (m_1\{\neg x\}m_2\{x\} + m_1\{x\}m_2\{\neg x\})} \\
m_1 \oplus m_2\{\neg x\} &= \frac{m_1\{\neg x\}m_2\{\neg x\} + m_1\{x, \neg x\}m_2\{\neg x\} + m_1\{\neg x\}m_1\{x, \neg x\}}{1 - (m_1\{\neg x\}m_2\{x\} + m_1\{x\}m_2\{\neg x\})} \\
m_1 \oplus m_2(\{x, \neg x\}) &= 1 - m_1 \oplus m_2(\{x\}) - m_1 \oplus m_2(\{\neg x\})
\end{aligned}$$

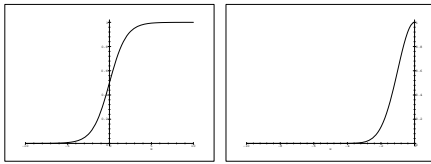
*Confidence factor functions:* It remains the problem of transforming the output of a source (measurement) to an appropriate input for the bpa's. [18] introduced a confidence factor function  $cf(v) : \mathfrak{R} \rightarrow [0, 1]$  that produces a confidence factor for the measurement  $v$  if it satisfies the following:

1.  $cf(v)$  is an increasing function
2.  $cf(v) = 1.0$  if the measurement  $v$  implies  $\{x\}$  with certainty
3.  $cf(v) = 0.0$  if the measurement  $v$  implies  $\{\neg x\}$  with certainty
4.  $cf(v) = 0.5$  if the measurement  $v$  favors neither  $\{x\}$  nor  $\{\neg x\}$

Once a confidence value is obtained, the transformation into a bpa can be accomplished by

$$\begin{aligned}
\text{beliefin}\{x\} : m(\{x\}) &= \frac{B}{1-A} * cf(v) - \frac{AB}{1-A} \\
\text{beliefin}\{\neg x\} : m(\{\neg x\}) &= \frac{-B}{1-A} * cf(v) + B \\
\text{ignorance} : m(\Theta) &= 1 - m(\{x\}) - m(\{\neg x\})
\end{aligned}$$


Tests with different confidence factor functions showed that the sigmoid function and the one-sided Gaussian were the most appropriate functions  $cf(v)$  for the combination of the measurements in our application.



**Sigmoid function**  $y = \frac{1}{1+e^{-k(x-\theta)}}$  with  $k = 1$  and  $\theta = 0$ ,  $k$  describes the steepness and  $\theta$  determines the offset on the x-axis.

**One-sided Gaussian function**  $y = e^{-(x-\theta)^2\sigma}$ ,  $\sigma$  is the scaling and  $\theta$  a shift in x-direction.

**Table 3.** Choice of confidence factor functions for evidence sources. The columns *invert*, *greatest lower bound*, and *sigmoid transient* define the function parameters

Evidence source	Confidence factor function	invert	greatest lower bound	sigmoid transient
standard deviation	One-sided Gaussian	no	0.3	-
nr. of segments	Sigmoid	yes	0.01	10.5
maximal length of segment	Sigmoid	no	0.01	13
maximum of gradient	Sigmoid	no	0.01	4.5
Minimum of Gradient	Sigmoid	yes	0.01	-6.4

## Example of combination of evidence using Dempster Shafer's rule

A test example illustrates the narrowing of the hypothesis by accumulation of evidence. Source 1, although with high ignorance 0.5, strengthens the bpa for  $\{x\}$  if combined with source 2. Source 3, although with higher belief in  $\{\neg x\}$  than in  $\{x\}$ , only slightly erodes the bpa for  $\{x\}$  if combined with source 2, whereas the the belief in  $\{\neg x\}$  is somewhat increased. The combination of all the three sources demonstrate the low ignorance  $\Theta = 0.03$  and the high belief in  $\{x\}$ . The combination rule is associative and commutative and consequently the belief function does not depend on the order in which the evidence was gathered (see Fig. 4a).

bpa	source 1	source 2	source 3
$m_i \{x\}$	0.4	0.8	0.3
$m_i \{\neg x\}$	0.1	0.0	0.5
$m_i (\Theta)$	0.5	0.2	0.2

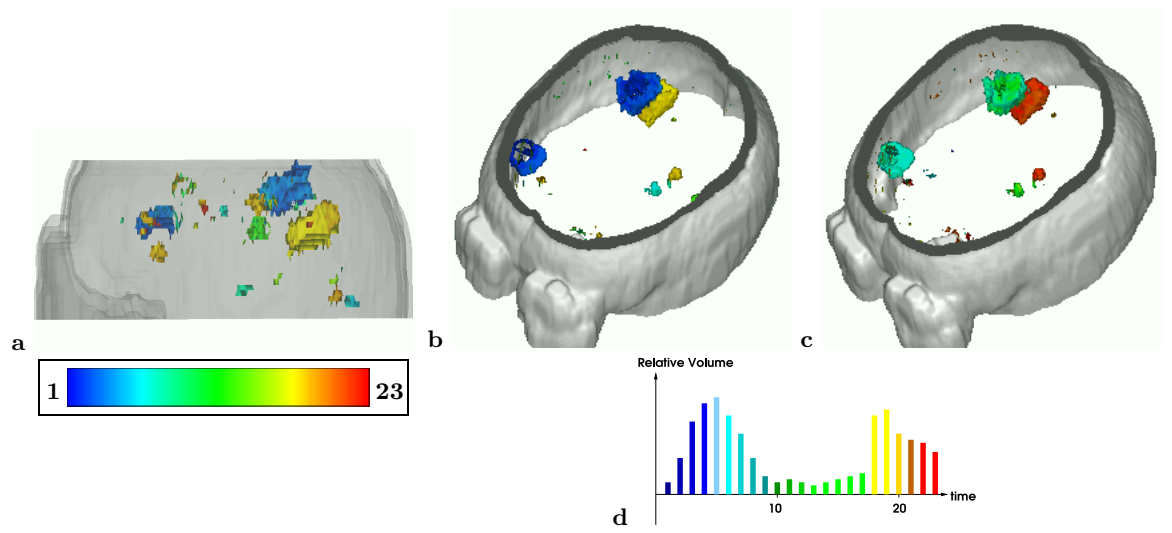
combination	bpa	combination	bpa	combination	bpa	combination	bpa
$m_1 \oplus m_2 \{x\}$	0.87	$m_1 \oplus m_3 \{x\}$	0.26	$m_2 \oplus m_3 \{x\}$	0.77	$m_1 \oplus m_2 \oplus m_3 \{x\}$	0.84
$m_1 \oplus m_2 \{\neg x\}$	0.02	$m_1 \oplus m_3 \{\neg x\}$	0.36	$m_2 \oplus m_3 \{\neg x\}$	0.17	$m_1 \oplus m_2 \oplus m_3 \{\neg x\}$	0.13
$m_1 \oplus m_2 (\Theta)$	0.11	$m_1 \oplus m_3 (\Theta)$	0.38	$m_2 \oplus m_3 (\Theta)$	0.06	$m_1 \oplus m_2 \oplus m_3 (\Theta)$	0.03
contradiction $\kappa$	0.08	contradiction $\kappa$	0.18	contradiction $\kappa$	0.4	contradiction $\kappa$	0.145

## References

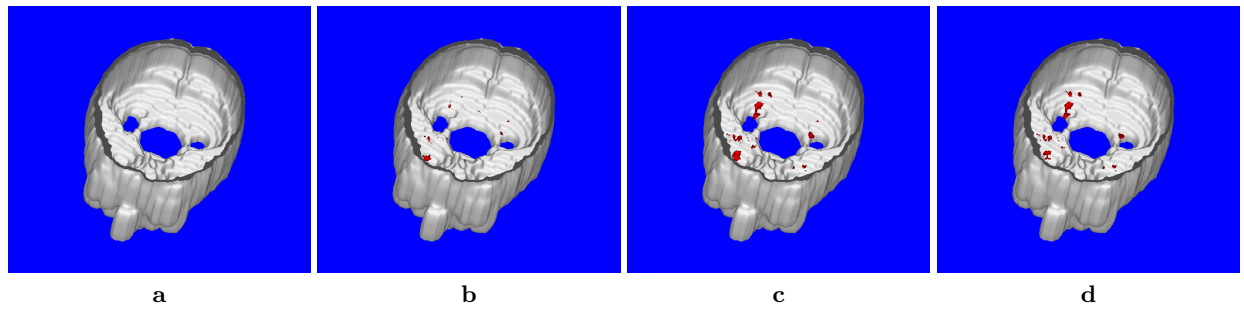
1. P.A. Bandettini, A Jesmanowicz, E.C. Wong, and J.S. Hyde, *Processing strategies for time-course data sets in functional MRI of the human brain*, Magnetic Resonance in Medicine, 30:161-173, 199
2. D.L.G. Hill, A. Simmons, C. Studholme, D.J. Hawkes, and S.C.R. Williams, *Removal of stimulus correlated motion from echo plana fmri studies*, In Proc. Soc. Magn. Res. 3rd annual meeting, page 840, 1995
3. Charles R.G. Guttmann et al., *The Evolution of Multiple Sclerosis Lesions on Serial MR*, AJNR 16:1481-1491, Aug 1995
4. R. Kikinis, M.E. Shenton, G. Gerig, J. Martin, M. Anderson, D. Metcalf, Ch.R.G. Guttmann, R.W. McCarley, B. Lorensen, H. Cline, F.A. Jolesz, *Routine Quantitative Analysis of Brain and Cerebrospinal Fluid Spaces with MR Imaging*, JMRI (Journal of Magnetic Resonance Imaging), Vol. 2 No. 6, pp. 619-629, Nov/Dec 1992
5. A.C. Evans, J.A. Frank, J. Antel, and D.H. Miller, *The role of MRI in clinical trials of multiple sclerosis: Comparison of image processing techniques*. Annals of Neurology, 1996. In press.
6. A. Zijdenbos, A. Evans, F. Riahi, J. Sled, H.-C. Chui, and V. Kollokian, *Automatic quantification of multiple sclerosis lesion volume using stereotaxic space*, In forth Int. Conf. on Visualization in Biomedical Computing (VBC), Hamburg, Germany, 1996, pp. 439-448
7. M. Kamber, R. Shinghal, D.L. Collins, G.S. Francis, and A.C. Evans, *Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images*. IEEE Transactions in Medical Imaging, 14(3):442-453, Sept. 1995
8. G. Gerig, J. Martin, R. Kikinis, O. Kübler, M. Shenton and F. A. Jolesz, *Unsupervised tissue type segmentation of 3D dual-echo MR head data*, image and vision computing, IPMI 1991 special issue, vol. 10 No. 6, pp. 349-360, July/August 1992
9. S. Warfield et al., *Automatic Identification of Gray Matter Structures from MRI to Improve the Segmentation of White Matter Lesions*, Journal of Image Guided Surgery, Vol. 1, No. 6, June 1996, pp. 326-338
10. Johnston et al., 1996, Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. IEEE TMI 15(2):154-169.
11. Labeling of 4D structures in registered 3-D segmentations (exact reference to be added)
12. J-P. Thirion, *New feature points based on geometric invariants for 3d image registration*, Int. Journal of Computer Vision, 18(2):121-137, May 1996
13. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, *Multi-Modality Image Registration by Maximization of Mutual Information*, IEEE Trans. on Medical Imaging, 16(2), pp. 187-198, April 1997
14. Christian Brechbühler, Guido Gerig, and Gabor Székely, *Compensation of spatial inhomogeneity in MRI based on a multi-valued image model and a parametric bias estimate*, Visualization in Biomedical Computing Proc. VBC'96, Lecture Notes in Computer Science, No. 1131, Springer, pp. 141-146, Sept. 1996



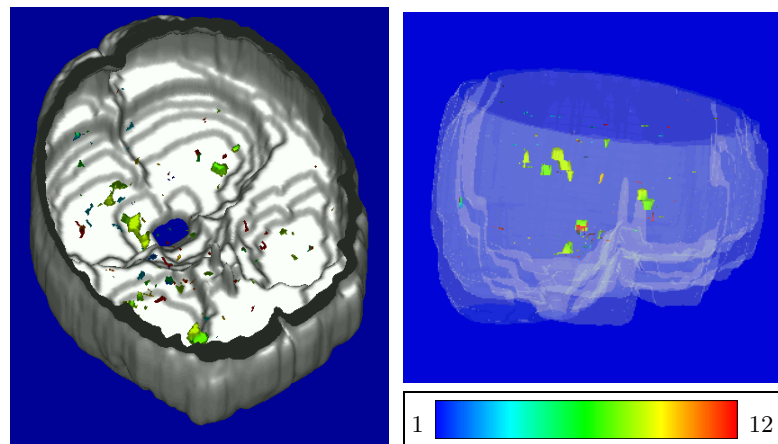
15. M. Styner and G. Gerig, *Evaluation of 2D/3D bias correction with 1+1ES-optimization*, Technical Report Image Science Lab, ETH Zurich, TR-179, 1997
16. Glenn Shafer, *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press, 1976
17. J. Gordon, and E.H. Shortliffe, *The Dempster-Shafer Theory of Evidence*, in: B.G. Buchanan and E.H. Shortliffe (Eds.), *Rule-Based Expert Systems*, pp. 272-292, Addison-Wesley, 1985
18. Robert J. Safranek, Susan Gottschlich, Avinash C. Kak, *Evidence Accumulation Using Binary Frames of Discernment for Verification Vision*, *Actions on Robotics and Automation*, Vol. 6, No. 4, August 1990
19. European project on Brain Morphometry (BIOMORPH, EU-BIOMED2 project no BMH4-CT96-0845, 1996-1998)



**Fig. 6.** Three-dimensional display of lesions segmented from the Brigham and Women's data set. (a) Side-view with transparent intracranial cavity. (b) Time of appearance, (c) time of disappearance. (d) Plot of total volume estimates versus time. Remember that the method analysis only fluctuations and excludes static portions of lesions. The color represents the time of appearance or disappearance, coded from 1 to 23.



**Fig. 7.** Three-dimensional renderings of time evolution resulting from the 4-D analysis of the BIOMORPH data set. Images (a-d) represent weeks 0, 28, 36 and 40.



**Fig. 8.** Three-dimensional displays of lesions segmented from the BIOMORPH data set, top and side view of lesions and intracranial cavity. The color represents the time of appearance, coded from 1 to 12).