**Medical Image Analysis,** March 2000, 4(1): 31- 42
© Oxford University Press

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

Guido Gerig[1*], Daniel Welti[2], Charles R.G. Guttmann[3], Alan C.F. Colchester[4] and Gábor Székely[2]

[1]University of North Carolina, Department of Computer Science, Chapel Hill, NC, USA
[2]Swiss Federal Institute of Technology ETH, Zurich, Switzerland
[3]Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[4]University of Kent at Canterbury, Kent, UK

**Abstract**
This paper presents a new method for the automatic segmentation and characterization of object changes in time series of three-dimensional data sets. The technique was inspired by procedures developed for analysis of functional MRI data sets. After precise registration of serial volume data sets to 4-D data, we applied a 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 over the whole time series. Thus, static regions remain unchanged over time, whereas changes in tissue characteristics produce typical intensity variations in the voxel's time series. A set of features was derived from the time series, 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. The project was 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. The new approach replaces conventional segmentation of series of 3-D data sets by a 1-D processing of the temporal change at each voxel in the 4-D image data set. The new method has been applied to a total of 11 time series from different patient studies, covering time resolutions of 12 and 24 data sets over a period of about one year. The results demonstrate that time evolution is a highly sensitive feature for detection of fluctuating structures.

*Keywords:* time series analysis, combination of uncertainty, lesions in MRI, Multiple Sclerosis

*Received XX; revised July 20, 1999; accepted XX*

## 1. INTRODUCTION

Serial magnetic resonance imaging of patients has become increasingly attractive due to the non-invasive image acquisition, the shorter scanning and therefore decreased patient time, and the high spatial and tissue resolution. Time series scans reveal information about significant changes in diseased anatomical regions, about the effects of a drug or radiotherapy treatment, or about subtle morphological changes caused by neurological disease. The temporal sampling provides information about both morphological and functional changes.

A typical analysis of this type, which is routinely applied, is the analysis of functional MRI (fMRI) data sets. A patient is stimulated with a visual, auditory or motor activity for a specific time period. 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 with each pixel over time. Here, the signal processing aims at finding the best discrimination between noisy steady-state signals and signals correlated with the stimulus (Bandettini *et al.*, 1993). The processing most often assumes that a patient does not move during the examination,

*Corresponding author
(e-mail: gerig@cs.unc.edu)

although slight object motion due to breathing, pulsation of the heart or swallowing is unavoidable. It has been shown that a sub-voxel correction of 3-D motion (Hill *et al.*, 1995) can considerably improve the voxel-based time-series analysis.

Pharmacological studies and patient follow-up and monitoring differ from fMRI because 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 to months. In schizophrenia, temporal changes are studied by imaging a patient with yearly scans.

The development of a new segmentation technique is driven by the motivation of getting a better understanding of the disease process in multiple sclerosis (MS). Research in MS has already demonstrated the power of using serial imaging (Guttman *et al.*, 1995). Drug development for multiple sclerosis uses serial MRI as one of several diagnostic features to study 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 over a period of about one year. The significance of tests is increased by multi-center studies, collecting image data from various hospitals by using a standardized MR protocol. Image data are examined by radiologists, evaluating each scan in relation to the previous one to identify any new lesions. Quantitative analyses of total lesion load and of single lesions are performed by using interactive user-operated segmentation tools. A typical study may consist of a thousand 3-D data sets. The manual outlining of lesions in a 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 (Kikinis *et al.*, 1992, Evans *et al.*, 1996, Zijdenbos *et al.*, 1996, Kamber *et al.*, 1995). These systems consist of well-designed sequences of processing steps, including preprocessing, bias-field correction, feature-space clustering of multi-echo MRI data (Gerig *et al.*, 1992), and matching of a statistical anatomical atlas (Warfield *et al.*, 1996, Johnston *et al.*, 1996) to solve the ambiguities of statistical classification. The result is a significantly improved reproducibility with reduced inter- and intra-rater variabilities allowing efficient processing of a large amount of data.

Previous segmentation methods mostly intend to segment lesions from single data sets, not taking into account the significance of correlation over time. 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 over time was presented by (Metcalf *et al.*, 1992, Kikinis *et al.*, 1993) by proposing a 4-D connected component labeling on registered segmented images. The procedure serves as a 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 main aim was to improve 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. (Guttman *et al.*, 1995) presented a seminal paper on characterizing the evolution of lesions in serial MR data, suggesting the use of this valuable information for image processing. The present paper *explores the time domain information* inherently provided by serial MR data sets. The major question in research into 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 is a spatio-temporal analysis of the 4-D data sets, this paper only focuses on evaluating the discrimination power of time.

Besides exploring time as a new feature for segmentation, we are working toward 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 briefly 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 that allow the application of time-series analysis of single voxels.

**Registration:** The serial data sets obtained from Brigham and Women's Hospital in Boston (cf. section 4.1) have been registered by the INRIA research group using crest-line extraction and matching (Thirion, 1996). A second serial data set presented in this paper is processed by the KUL research group using the MIRIT registration software package (Maes *et al.*, 1997), which maximizes the mutual information between corresponding voxel intensities. Both registration methods perform a rigid transformation and work fully automatically after appropriate initialization. The transformation matrices are input to a geometric transformation which performs trilinear interpolation. The first 3-D image of each time series is declared as the reference image and all the other volumetric images are rigidly registered into this coordinate system.

**Image brightness normalization and bias correction:** 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 (Brechbühler *et al.*, 1996). We assume the original scene to be composed of tissue regions with homogeneous brightness only degraded by noise. 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 a series of volume images results in a combination of bias correction and brightness normalization. The marked artifactual changes in brightness on alternate slices in one of the data sets required a two step procedure by first correcting for brightness changes between individual slices as a 2-D processing and second for the 3-D bias field (Styner *et al.*, 1997) as a fully 3-D procedure. The brightness normalization and bias correction was calculated for each scan due to considerable differences over the observation period of one year.

**Brain mask:** MS lesion were analyzed only within brain white and gray matter. Brain tissue and fluid has been segmented by a well-established, highly automated segmentation system (Gerig *et al.*, 1992) which consists of supervised classification, erosion, connected component labeling and dilation. The mask has to be determined for only one dataset in the time series and takes about 10 minutes on a standard workstation.

**Result of Preprocessing:** 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 assessing its time course give a good impression of the quality of the preprocessing. We assume that the signal intensity of normal white matter should remain constant (figure 1b), whereas 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 result in 4-D $[x,y,z;t]$ data sets. The preprocessing yields a spatial and intensity-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 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 that can be explained by elastic tissue deformations, whereas some boundaries in the vicinity of lesions can demonstrate significant deformations due to a mass effect (see Fig. 2b lower middle). This effect has been recently explored by (Rey *et al*, 1999) between two rigidly registered 3-D data sets. 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 are used for discriminating between static tissue and active lesions.

**Brightness Difference:** The minimum and maximum brightness for each time series is obtained and the absolute difference $\triangle I = |I_{max} - I_{min}|$ calculated. This feature measures the maximum contrast change of a time series within the observed time period (Fig. 4a).

**Statistical measures:** Mean and variance of the time series at each voxel 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. Figures 4b,c,d display the standard deviation and variance to illustrate the lesion discrimination power of this feature but only the variance is used for evidence accumulation and classification.

**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 (3,4e,f,g).

**Time derivatives:** The gradient of the time function for each voxel 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 (4h). The attributed time will be further used for displaying temporal evolution (see results).

### 3.2. Evidence accumulation by combining uncertain measurements

The multiple features derived by signal processing provide probabilistic maps of the likelihood of the specific feature (Fig. 4a-h). Each of feature 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. The original Dempster-Shafer formalism (Shafer *et al.*, 1976, Gordon *et al.*, 1985) was computationally complex, but we obviated this by using binary frames of discernment (BFOD) as proposed by (Safranek *et al.*, 1990). More information about 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 of these functions were based on a comparison of the resulting feature maps with segmentation results produced by statistical classification followed by manual corrections, which was considered as a golden standard.

The Dempster's combination rule is associative and commutative, so that the final probability does not depend on the order in which evidence is combined (Fig. 5a).

The combined 3-D data set is again probabilistic, with a value range of $[0,\cdots,1]$ (Fig. 5b). A binary segmentation, for example for three-dimensional graphical visualization (Fig. 5c), 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 that a careful design of the cf-functions and bpa assignments had been carried out.

## 4. RESULTS

The new segmentation system has been applied to two time series from different patient studies. The first study was carried out at Brigham and Women's Hospital covers 40 patients with 24 brain scans and a fixed sequence of scanning intervals of one, two and 4 weeks. The second study was acquired at Guy's Hospital in London and analyzed in the European BIOMORPH project (BIOMORPH, 1996-1999), comprising 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 by 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) and half Fourier sampling (0.5 NEX), providing 54 slices, 3mm slice distance and thickness and 256x256 pixels, resulting in voxel dimensions of 0.96x0.96x3 $mm^3$. 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 MS study covers the serial scanning of 40 patients, and we selected one typical case for the temporal analysis. The study did not include the serial scanning of normal controls.

The 3-D visualizations (Fig. 9) 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. It is important to note that the procedure *only records the volume of tissue showing significant 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 were 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). Each scan consisted of 24 axial slices 256x256, with voxel dimensions 0.9x0.9x5.5$mm^3$. 12 scans were acquired over a period of 56 weeks: 11 scans with 4 weeks intervals and one scan with a 13 week interval. This unequally spaced time image series was converted into a regularly sampled sequence by linear interpolation. Totally, 40 patients have been scanned with 12 MRI each. The study did not include the serial scanning of normal controls. The processing described in this paper has been applied to 10 complete series of datasets.

Figures 10 and 11 illustrate the segmentation result, again color coding the time of appearance of the segmented lestions.

## 5. VALIDATION

Validation of the time series analysis method is not straightforward as it measures temporal changes rather than absolute lesion volumes. We 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 was replicated by the number of time steps, here 12. The limitations of the clinical 4-D data sets due to errors in spatial registration and the bias correction were therefore not subject to this validation. We confirmed visually that the dataset did not contain white matter lesions.

Lesions of varying size, contrast and lifetime (time between appearance and disappearance) were generated using a non-linear diffusion process with a source and a sink term. This decision was influenced by the visual analysis of realistic temporal patterns (Fig. 2). Lesions in magnetic resonance image data were always diffuse and most often showed a sharp increase and a smooth decrease of the size and brightness function.

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

Equation 1 produces time series of diffuse lesions with quick appearance and slightly slower disappearance (Figure 6**a**). The time series of the artificial lesions were merged with the time series of the head. Lesions were inserted in both the white matter and in the grey matter of the brain.

This 4-D data set, simulating lesion evolution, has then been corrupted by Gaussian noise with a variance equal to the variance of white matter (Fig 6**b**). The parameters for the temporal analysis were chosen to be 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 were detected (Fig 6**c**). No false positive lesions were detected which can be explained by the nature of the simulated 4-D dataset being composed of identical 3-D volumes.

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 low contrast lesions, 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. In addition to 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 time difference between the minimal and maximal gradients of the time course of this voxel and does not exactly match with the time points of appearance and disappearance of low-contrast lesion portions.

## 6. SUMMARY AND CONCLUSIONS

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

The paper describes the development of an image analysis technique 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 which discriminate static tissue from changes over time. 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 combination of different sources of evidence by considering the probability not only of the occurrence of a feature, but also the absence and the ignorance about the measurements. The design of confidence factor functions and the transformation of confidence factors into basic probabilities represent decisive steps comparable to supervised training in statistical classification. Tests 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 of processing time (SUN Ultra 1 with 128Mb). The results, visually compared with results from alternative segmentation methods, revealed a surprisingly good sensitivity and specifity to MS lesions. We so far applied the procedure to one serial dataset (23 MRI) from the Brigham and Women's Hospital and to 10 datasets (with 12 MRI each) from the European BIOMORPH study. However, our analysis so far is *only based on information of one MR echo over time*. We can expect an even higher sensitivity if multi-echo information could be embedded, and if we combined 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. 5). The quantitative validation and tests with simulated lesions show that all the lesions could be detected but that the lifetime of the detected lesion voxels is consistently shorter than that of the implanted lesion voxels. Our lesions were not binary blobs, however, but were generated by a nonlinear diffusion process, which starts with a very low contrast. Thereby, we tried to simulate lesions in real images that are often nearly invisible at early stages and which are not labeled as significant lesions by a manual observer.

We 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 because the goal is commonly evaluation of changes due to the disease process or a drug treatment. Besides detection of lesion voxels, our method reveals the time of appearance and disappearance as attributes of 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 possibly a new understanding of the disease mechanism.

Currently, we are extending the time-series analysis to spatial analysis in order 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. Confidence factor functions will be designed similarly to the single contrast application presented herein by including manual segmentation as a gold standard.

## ACKNOWLEDGMENT

The European Union and the Swiss funding agency BBW supported this research through the international project on Brain Morphometry (BIOMORPH, EU-BIOMED2 project number BMH4-CT96-0845). We thank Ron Kikinis and Ference 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). Jean-Philippe Thirion and Nicholas Ayache, INRIA Sophia Antipolis, provided 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 provided their expertise and made research funding available.

## A. DEMPSTER SHAFER (DS) THEORY

The DS theory (Shafer *et al.*, 1976, Gordon *et al.*, 1985) 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. 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 $\Theta = \{lesion, \neg lesion\}$ of mutually exclusive elements and represents the special case of a *binary frame of discernment* (BFOD) (Safranek *et al.*, 1990). The set of all subsets of $\Theta$ is called the power set $2^{|\Theta|}$ of $\Theta$ with the focal elements $(\{\emptyset\}, \{lesion\}, \{\neg lesion\}, \{\Theta\})$, where $\{\Theta\}$ stands for $\{lesion, \neg lesion\}$ or ignorance.

**Basic probability assignment (bpa):** A basic probability assignment 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 $0 \leq m(.) \leq 1, m(\emptyset) = 0, and \sum_{A \subseteq \Theta} m(A) = 1$, 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 basic probability assignments (bpa) $m_1$ and $m_2$, designated as $m_1 \oplus m_2$. The products of the assignments, 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.

$$m_1 \oplus m_2\{x\} = \frac{m_1\{x\}m_2\{x\} + m_1\{x,\neg x\}m_2\{x\} + m_1\{x\}m_1\{x,\neg x\}}{1-(m_1\{\neg x\}m_2\{x\} + m_1\{x\}m_2\{\neg x\})} \quad (2)$$

$$m_1 \oplus m_2\{\neg x\} = \frac{m_1\{\neg x\}m_2\{\neg x\} + m_1\{x,\neg x\}m_2\{\neg x\} + m_1\{\neg x\}m_1\{x,\neg x\}}{1-(m_1\{\neg x\}m_2\{x\} + m_1\{x\}m_2\{\neg x\})} \quad (3)$$

$$m_1 \oplus m_2(\{x,\neg x\}) = 1 - m_1 \oplus m_2(\{x\}) - m_1 \oplus m_2(\{\neg x\}) \quad (4)$$

**Confidence factor functions:** There remains the problem of transforming the output of a source (measurement) to an appropriate input for the bpa's. (Safranek *et al.*, 1990) introduced a confidence factor function $cf(v) : \Re \rightarrow [0,1]$ 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 basic probability (bpa) can be accomplished by defining appropriate transfer functions with parameters $A$ and $B$ (Eq. 5). Figure 7 displays typical functions for belief, disbelief and ignorance.

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

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 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 (Fig. 8 left).

**One-sided Gaussian function:** $y = e^{(x-\theta)^2\sigma}$, $\sigma$ is the scaling and $\theta$ a shift in x-direction (Fig. 8 right).

### A.1. Example of combination of evidence using Dempster Shafer's rule

A test example (Table 4 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. 5a).

## REFERENCES


Bandettini, P.A., Jesmanowicz, A., Wong, E.C. and Hyde,J.S. *Processing strategies for time-course data sets in functional MRI of the human brain*, Magnetic Resonance in Medicine, 30:161-173, 1993

Hill, D.L.G., Simmons, A., Studholme, C., Hawkes, D.J. and Williams, S.C.R. *Removal of stumulus corelated motion from echo plana fMRI studies*, In Proc. Soc. Magn. Res. 3rd annual meeting, page 840, 1995

Guttmann, Ch. R.G., *The Evolution of Multiple Sclerosis Lesions on Serial MR*, AJNR 16:1481-1491, Aug 1995

Kikinis, R., Shenton, M.E., Gerig, G., Martin, J., Anderson, M., Metcalf, D., Guttmann, Ch.R.G., McCarley, R.W., Lorensen, B., Cline, H. and Jolesz, F.A. *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

Evans, A.C., Frank, J.A., Antel, J. and Miller, D.H. *The role of MRI in clinical trials of multiple sclerosis: Comparison of image processing techniques*. Annals of Neurology, 1996. In press.

Zijdenbos, A., Evans, A., Riahi, F., Sled, J., Chui, H.-C. and Kollokian, V. *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

Kamber, M., Shinghal, R., Collins, D.L., Francis, G.S. and Evans, A.C. *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

Gerig, G., Martin, J., Kikinis, R., Kübler, O., Shenton M., and Jolesz, F.A. *Unsupervised tissue type segmentation of 3D dual-*

*echo MR head data*, image and vision computing, 10(6), pp. 349-360, July/August 1992

Warfield S., Dengler, J., Zaers, J., Gutmann, Ch.R.G., Wells, W.M., Ettinger, G.J., Hiller, J. and Kikinis, R. *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

Johnston et al., 1996, Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. IEEE TMI 15(2):154-169.

Metcalf, D. and Kikinis, R. and Guttmann, Ch. and Vaina L. and Jolesz, F., *4D connected component labelling applied to quantitative analysis of MS lesion temporal development*, Proceedings of the 14th annual internationa conference of the IEEE-EMBS, Paris, 1992; 3:945-946

Kikinis, R. and Metcalf, D. and Guttmann, Ch. and Martin, J. and Weiner, H. and Dengler, J. and Anderson, M. and Jolesz, F., *Quantitative Analysis of Temporal Development of MS Lesions*, Neuroradiology, 35(S72):108, 1993

Thirion, J-P. *New feature points based on geometric invariants for 3d image registration*, Int. Journal of Computer Vision, 18(2):121-137, May 1996

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. and Suetens, P. *Multi-Modality Image Registration by Maximization of Mutual Information*, IEEE Trans. on Medical Imaging, 16(2), pp. 187-198, April 1997

Brechbühler, Ch., Gerig, G. and Székely, G. *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

Styner M. and Gerig, G. *Evaluation of 2D/3D bias correction with 1+1ES-optimization*, Technical Report Image Science Lab, ETH Zurich, TR-179, 1997, submitted

Rey, D. and Subsol, G. and Delingette, H. and Ayache, N., *Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis*, Lecture Notes in Computer Science No. 1613, Information Processing in Medical Imaging, pp. 154-167

Shafer, Gl. *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press, 1976

Gordon, J. and Shortliffe, E.H. *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

Safranek, R. J., Gottschlich, S. and Kak, A. C. *Evidence Accumulation Using Binary Frames of Discernment for Verification Vision*, Actions on Robotics and Automation, Vol. 6, No. 4, August 1990

European project on Brain Morphometry, BIOMORPH, EU-BIOMED2 Project No. BMH4-CT96-0845, 1996-1999

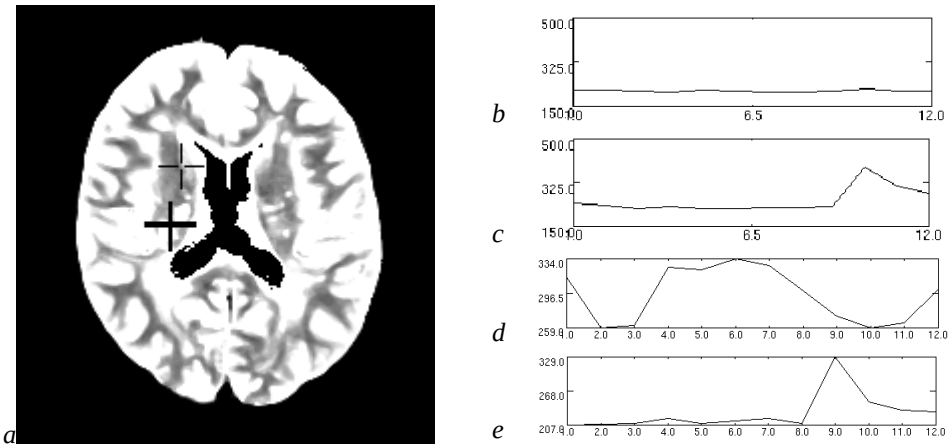



**Figure 1. Time series of voxels for healthy white matter (b) and typical lesions (c-e). Horizontal axis: time, vertical axis: MR intensity. Image (a) illustrates a typical MR slice presenting white matter lesions. The positions of the voxel generating the constant time series (b) is marked with a thin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

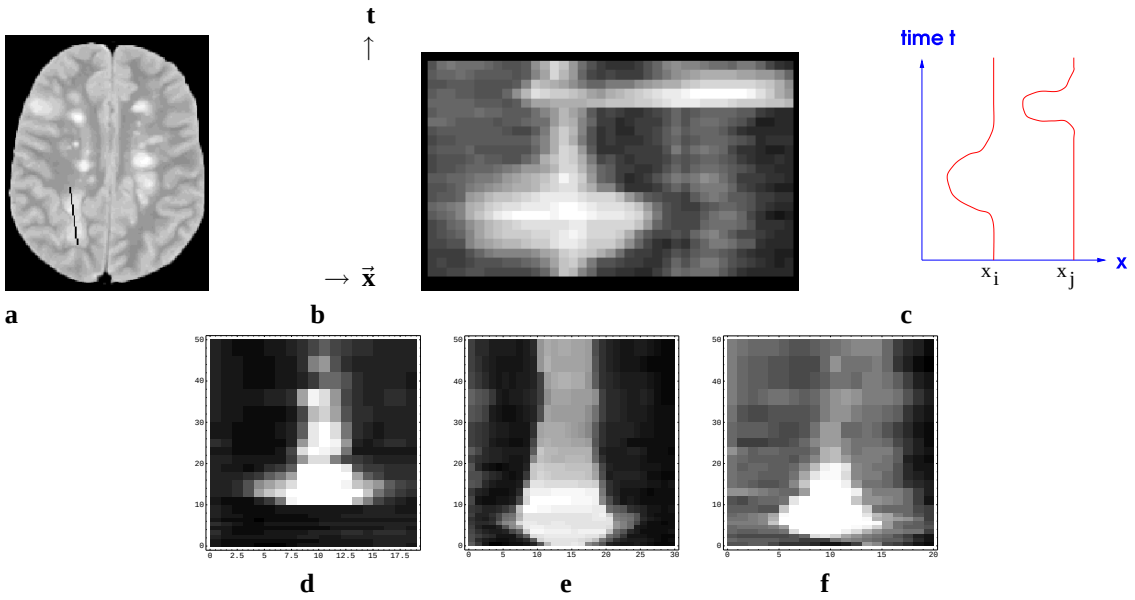



**Figure 2. Visualization of spatiotemporal lesion evolution. a Original image at time $t_i$ with profile, b space-time display (horizontal: spatial axis, vertical: time axis) and c sketch of typical time series at locations $x_i$ and $x_j$. Other typical lesion evolutions are displayed in d,e and f.**

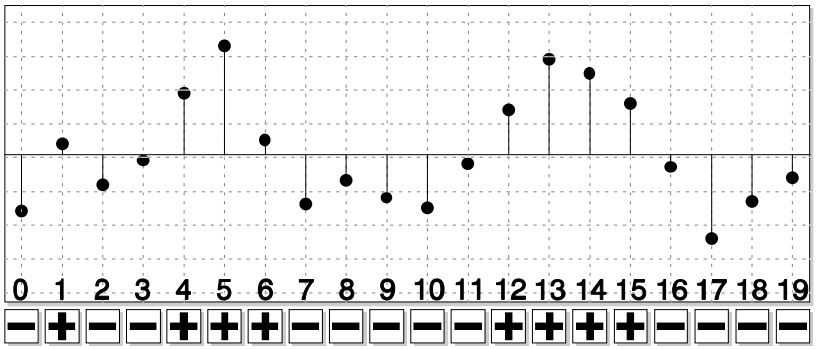



**Figure 3.** Evaluation of sign changes of zero-mean time series: Analysis of the sequence of signs: Number of "segments" (7), length of maximum segment (5), length of minimum segment (1) and average segment length (2.86). Observation: Low-frequency variations (lesions) result in large differences from the mean and generate small # of sign changes with large segments.

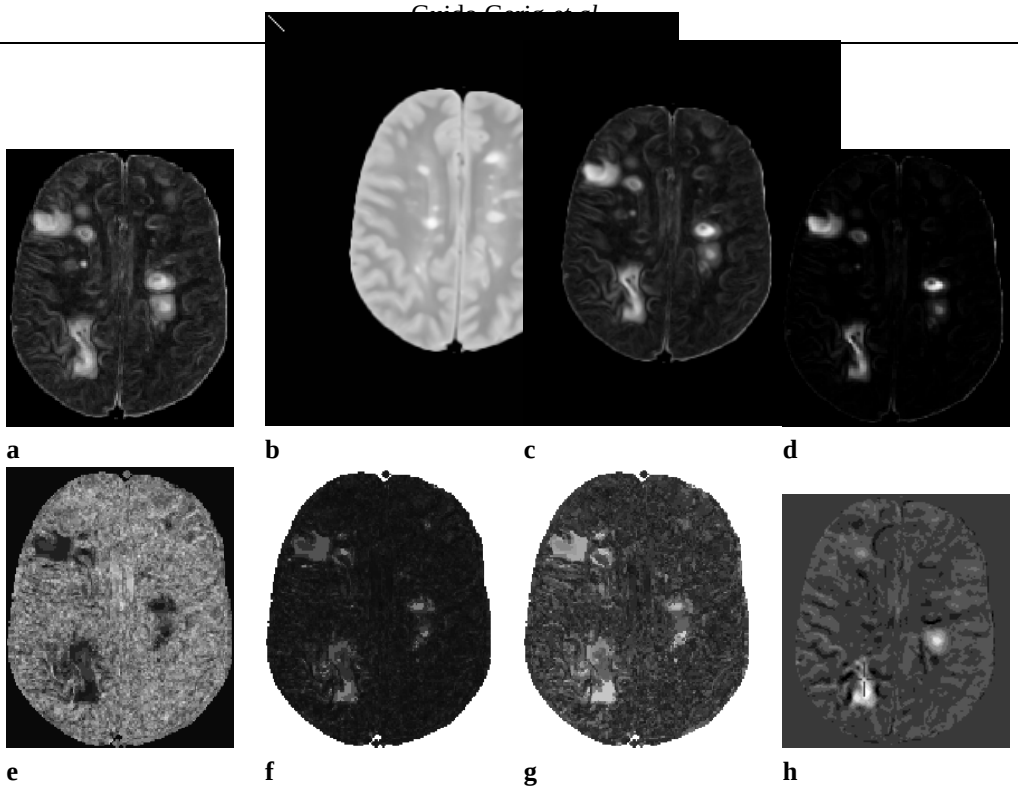



**Figure 4.** 2-D cuts of 3-D feature maps: Difference Max-Min (**a**), Mean (**b**), standard deviation (**c**) and variance (**d**), Nr. of zero-crossing segments (**e**), length of minimum segment (**f**), length of maximum segment (**g**), and maximum absolute time gradient (**h**)

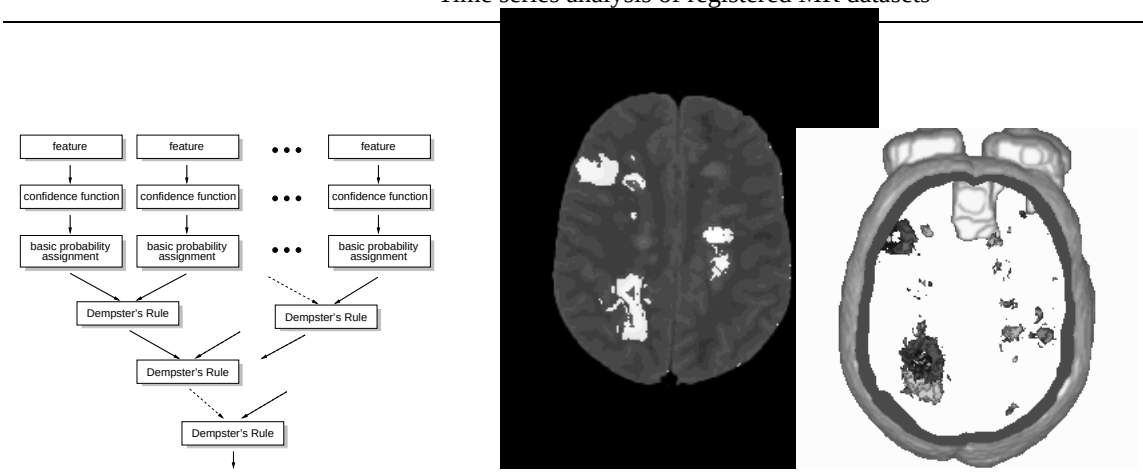
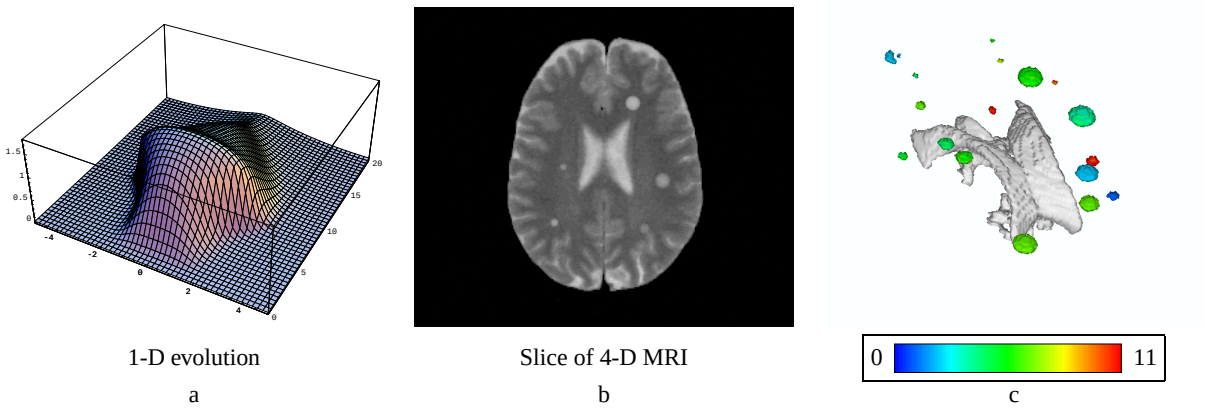



**Figure 5.** Combination of fuzzy features by Dempster's rule (a), segmented active lesions on 2-D slice (b) and as 3-D rendering (c).




1-D evolution
a

Slice of 4-D MRI
b

c

**Figure 6.** 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 color coded by the time of appearance.

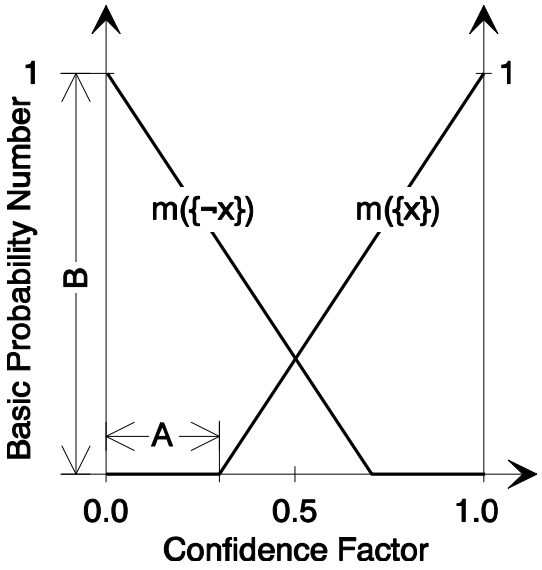



**Figure 7.** Transfer of confidence factors into basic probability values.

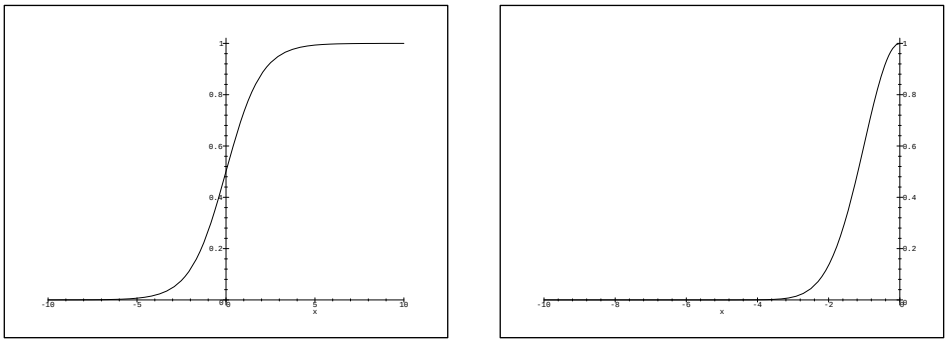



**Figure 8.** Choice of confidence functions: Sigmoid (left) and one-sided Gaussian (right).

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

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

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

**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 |

**Table 4.** Example of combination of evidence using Dempster Shafer's rule.

| 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 |
|---|---|---|---|---|---|
| $m_1 \oplus m_2\ \{x\}$ | 0.87 | $m_1 \oplus m_3\ \{x\}$ | 0.26 | $m_2 \oplus m_3\ \{x\}$ | 0.77 |
| $m_1 \oplus m_2\ \{\neg x\}$ | 0.02 | $m_1 \oplus m_3\ \{\neg x\}$ | 0.36 | $m_2 \oplus m_3\ \{\neg x\}$ | 0.17 |
| $m_1 \oplus m_2\ (\Theta)$ | 0.11 | $m_1 \oplus m_3\ (\Theta)$ | 0.38 | $m_2 \oplus m_3\ (\Theta)$ | 0.06 |
| contradiction $\kappa$ | 0.08 | contradiction $\kappa$ | 0.18 | contradiction $\kappa$ | 0.4 |

| combination | bpa |
|---|---|
| $m_1 \oplus m_2 \oplus m_3\ \{x\}$ | 0.84 |
| $m_1 \oplus m_2 \oplus m_3\ \{\neg x\}$ | 0.13 |
| $m_1 \oplus m_2 \oplus m_3\ (\Theta)$ | 0.03 |
| contradiction $\kappa$ | 0.145 |

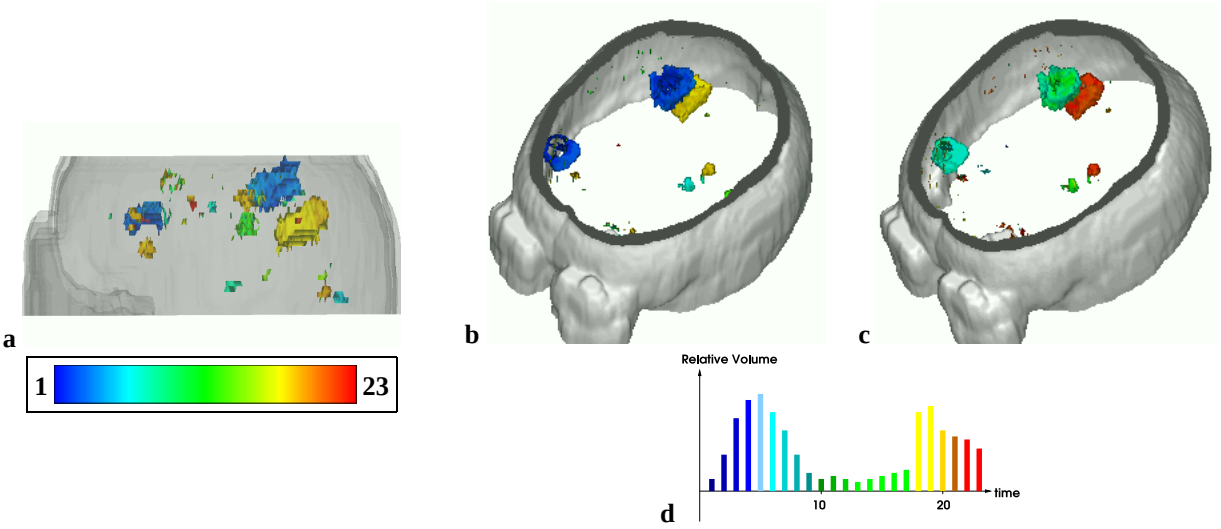



**Figure 9.** Three-dimensional display of lesions segmented from the Brigham and Women's data set. (a) Side-view with transparent intracranial cavity. (b) Time of appearance. (c) Time of disappearance. (d) Plot of total volume estimates versus time. Remember that the method analyzes only fluctuations and excludes static portions of lesions. The color represents the time of appearance or disappearance, respectively, coded from 1 to 23.

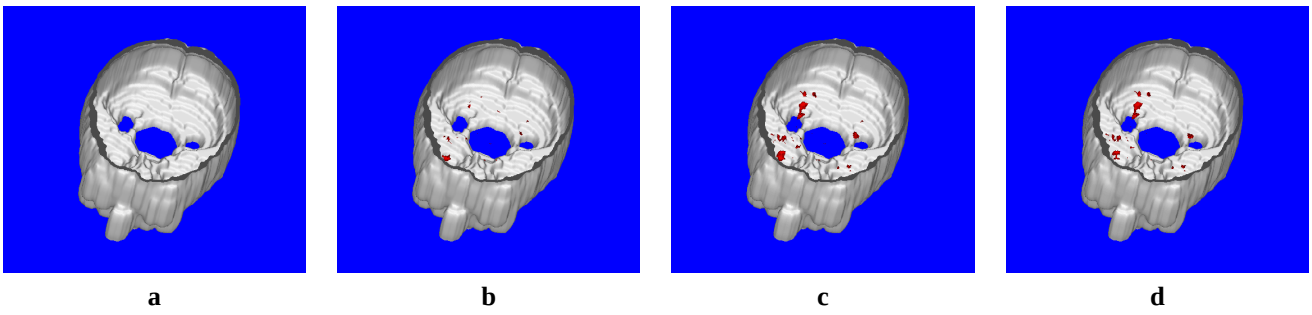



**Figure 10.** Three-dimensional renderings of time evolution resulting from the 4-D analysis of the BIOMORPH data set. Images (a-d) represent weeks 0, 28, 36 and 40.

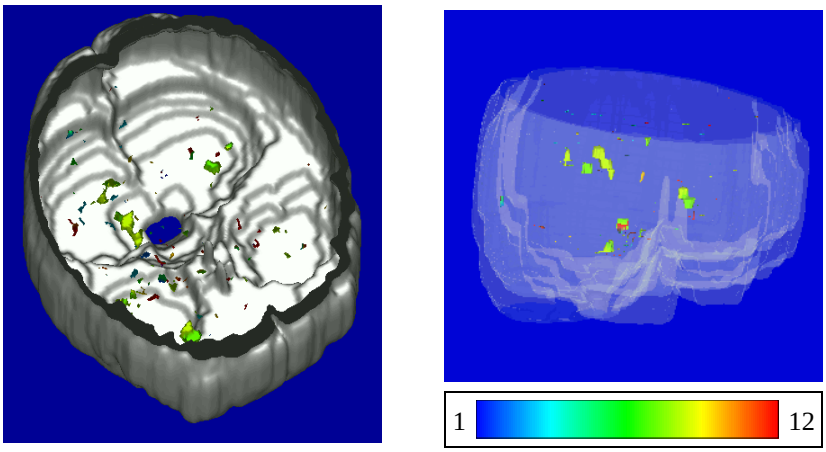



**Figure 11.** Three-dimensional displays of lesions segmented from the BIOMORPH data set, top and side view of lesions and intracranial cavity. The color represents the time of appearance, coded from 1 to 12.