IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

BAYESIAN TECHNIQUE (MEDICAL IMAGE SEGMENTATION) FOR IMAGE CLASSIFYING REGISTRATION

Sumangla Pawar*, Shashank Maane
Digital Communication, Balaji Institute of Technology & Management, Betul, India

ABSTRACT

In this paper, we propose a new model in which the similarity criterion is adapted locally to images by classification of image intensity dependencies. Defined in a Bayesian framework, the similarity criterion is a mixture of probability distributions describing dependencies on two classes. The model also includes a class map which locates pixels of the two classes and weights the two mixture components. The registration problem is formulated both as an energy minimization problem and as a Maximum A Posteriori (MAP) estimation problem. It is solved using a gradient descent algorithm. In the problem formulation and resolution, the image deformation and the class map are estimated at the same time, leading to an original combination of registration and classification that we call image classifying registration. Finally, we illustrate the interest of our model on two real applications from medical imaging: template-based segmentation of contrast-enhanced images and lesion detection in mammograms. We also conduct an evaluation of our model on simulated medical data and show its ability to take into account spatial variations of intensity dependencies while keeping good registration accuracy.

KEYWORDS— (Image registration, image segmentation, image classification)

INTRODUCTION (BAYESIAN TECHNIQUE)

The selectivity introduced in the previous section shows that we do not need to update each class's posterior for each descriptor. This section shows that by focusing on a much smaller, local neighbourhood (rather than on a particular log odds threshold), we can use an alternate search strategy to speed up the algorithm, and also achieve better classification performance by ignoring the distances to classes far from the query descriptor. Instead of performing a search for a query descriptor's nearest neighbour in each of the classes' reference sets, we search for only the nearest few neighbours in a single, merged dataset comprising all the features from all Labelle training data from all classes. Doing one approximate k-nearest-neighbour search in this large index is much faster than querying each of the classes'

- **Magnetic resonance imaging (MRI)**

*T*_ 2 , Transversal relaxation time incorporating *B*0 which are determined for each experiment in order to obtain separation or depending on what part of the body that is being examined.

*T*1 **and** *T*2

*T*1 and *T*2 are time constants, where *T*1 describes the time to reduce the difference between the longitudinal magnetization, *Mz*, and its equilibrium value and *T*2 is the time it takes for the transversal net magnetization, *Mxy*, to return to its equilibrium. When the magnetization vector lies along the direction of the applied magnetic field, *B*0, is the magnetization at equilibrium and is then called equilibrium magnetization, *M*0.

*T*1 **recovery**

The recovery of *T*1, also known as the spin-lattice relaxation time, arises by interaction between protons and the electromagnetic fields in the surrounding structures. The time constant *T*1 represents the time it takes for *Mz* to reach 63 % of the longitudinal magnetizations maximum value, see Figure
2.4. The *T*1 relaxation time is constant for a particular tissue and occurs between one excitation pulse and the next or the *TR* (repetition time).

*T*2 **relaxation**

Interactions between the magnetic fields of neighbouring spins cause *T*2 decay and are called spin-spin relaxation time. It is a result of the intrinsic magnetic fields of the nuclei interacting with each other. This process produces a loss of phase coherence or dephasing and results in a decay of the net magnetization vector in the transverse plane. The time it takes for *Mxy* to fall back to 37 % of its original level is represented by *T*2, see Figure 2.5.

The relaxation rates (*R*) are described by the inversed time constants respectively:

*R*1 = 1/*T*1and

$R2 = 1/T2$

Dynamic contrast-enhanced MRI

Dynamic contrast-enhanced MRI (DCE-MRI) is an imaging method where MR images are obtained after an intravenous injection of a contrast agent to describe the dynamic effect of it. DCE-MRI can for example be used to examine the liver non-invasively. The contrast in an image is obtained by the different $T1$ relaxation times in the tissues in the images. If the

Difference in $T1$ relaxation times between different biological tissues is not large enough to get a good enough contrast in the image; a contrast agent that enhances the difference can be used. An example of a contrast agent

- **Medical Image Segmentation.**

To be able to separate data in the data segmentation of the MR images is needed. Segmentation could be automatic, semi-automatic or manual. An automatic segmentation method is k means, which is a widely used and simple unsupervised clustering method, as for example in. The same applies for the supervised classification method k-nearest neighbor, a simple and often used classification tool, as in. However, to my knowledge they have not been used together to automatically segment images, which probably would be an easy first attempt towards image segmentation in this case. The large data sets and the pre-classified small areas in the images do these two methods appropriate. A semi-automatic segmentation is a method which involves human interaction as described in this gives an accurate segmentation, however the method is more time consuming than an automatic segmentation but more rapid then a manual segmentation. In the following chapters the methods used in this thesis explained.

The field of medical imaging has undergone revolutionary advances over of the past two decades. New medical imaging technologies have provided physician's powerful, non-invasive techniques to probe the structure, function and pathology of the human body. A few years ago, only a small number of non-invasive techniques were available to radiologists. Besides much experience/practice, deep insight, even intuition, was required for clinical diagnostic imaging. In recent years, the improvement and the development of many image acquisition techniques, the enhancement of the general quality of the acquired images, advances in image processing and development of large computational capacities, have considerably eased this task. Acquisition of medical images in two (2D), three (3D), or higher dimensions, has become a routine task for clinical and research applications. Image acquisition techniques include magnetic resonance imaging (MRI), magneto encephalography (MEG), 3D ultra sound imaging, computed tomography (CT), positron emission tomography (PET), single photon emission computed tomography (SPECT), functional MRI (fMRI), and diffusion weighted imaging (DWI). This increasingly vast and detailed amount of information needs to be interpreted in a timely and accurate manner in order to provide better diagnosis and treatment options for a family of clinical applications. It requires significant innovation in all aspects of image processing, such as image segmentation, image registration, visualization, compression and communication, among others. In medical image processing, the automated recognition of meaningful image components, anatomical structures, and other regions of interest, is a fundamental task commonly referred to as image segmentation. Image segmentation greatly facilitates visualization and manipulation of specific structures. It is a critical step that often dictates the outcome of the entire clinical or research analysis. One approach to image segmentation is to have a trained anatomist or technician manually label some regions of interests. However, manual approaches are considerably time consuming. For instance, the labeling of some or all the structures in the brain can take up to a week for high-resolution images. Also, manual or interactive segmentations, which are often restricted to 2D slice-wise processing, often suffer from inconsistency across segmented slices. Finally, studies have shown a large amount of variance among manual segmentations, an effect which seems to increase the risks related to inter- and intraobserver reliability. Quantitative analysis of medical images requires reproducible, accurate and efficient segmentation methods. In medical imaging studies, the segmentation of a large number of images is often necessary for obtaining meaningful (i.e. statistically significant) results. Therefore, automated segmentation is desirable. However, the challenge is that images are usually corrupted by several artifacts, such as image noise, image intensity in homogeneity or non-uniformity, and partial volume averaging effect. In recent years, many segmentation algorithms have been proposed and designed to account for such unwanted artifacts. While these techniques produce repeatable and accurate results, few of them guarantee that the obtained segmentations respect the true anatomy of the structures. Too often, segmentations contain small geometric inaccuracies that alter the true anatomy of the modeled structures. For instance, cortical segmentations often include handles that incorrectly connect different regions of the cortex. In medical imaging, the overall shape of a region of interest is prescribed by medical knowledge; it is usually known a priori. Segmentation techniques should be able to produce results that reflect the anatomy of the structures several clinical and research applications (e.g. visualization, surgical planning and surface-based processing of functional data, surface based atlas, and intersubject registration.) depend critically on the accuracy and correctness of the representations. However, accurate segmentation under anatomical consistency is challenging. In this paper we adapted some of the existing segmentation algorithms using Bayesian classifier.

- **Analysis of Gated Myocardial Perfusion Spect Image Using Computational Image Registration Techniques.**

a) **Motivation & Goals**

Modern medicine has been widely using imaging as a fundamental tool to aid in diagnosis procedures, monitoring the development of disease and planning treatment or even surgeries. Thus, it became a key element to support medical decision making through non-invasive procedures. In the last years, deep researches and developments have been increasing lead to a higher number of information types that can be acquired from such diagnostic tool.

Image registration computer techniques enable the integration of different medical image modalities and the easier detection of changes between images acquired from different points of view, different acquisition times or even with subject atlas to attain prior anatomic or functional information. It stresses changes in size, shape or image intensity over time, and relates both preoperative image and surgical plans to the physical reality of patients during intervention and aligns patient's anatomy to a standardized atlas. Techniques of image registration aim the establishment of spatial correspondence with the goal of find the optimal transformation that best aligns the structures of interest in the input images. The minimization of an error measure, or of a cost function, is the goal of these techniques. Additionally, an optimization algorithm is needed to find the most suitable transformation, and an interpolator is employed to resample the features into the new registered space.

- **Brain segmentation**

Efficient Multilevel Brain Tumor Segmentation with Integrated Bayesian Model Classification A key problem in medical imaging is automatically segmenting an image into its constituent heterogeneous processes. Automatic segmentation has the potential to positively impact clinical medicine by freeing physicians from the burden of manual labeling and by providing robust, quantitative measurements to aid in diagnosis and disease modeling. One such problem in clinical medicine is the automatic segmentation and quantification of brain tumors. We consider the GBM tumor because it is the most common primary tumor of the central nervous system, accounting for approximately 40% of brain tumors across patients of all ages, and the median postoperative survival time is extremely short (8 months) with a 5-year recurrence-free survival rate of nearly zero .Quantifying the volume of a brain tumor is the key indicator of tumor progression. However, like most segmentation problems, automatic detection and quantification of a brain 2 tumor is very difficult. In general, it is impossible to segment a GBM tumor by simple thresholding techniques. Brain tumors are highly varying in size, have a variety of shape and appearance properties, and often deform other nearby structures in the brain. In the current clinic, the tumor volume is approximated by the area of the maximal cross-section, which is often further approximated to an ellipse. Such a rough approximation is used because the time cost to compute a more accurate manual volume estimate is too high. Liu et al. Present an interactive system for computing the volume that reduces the cost of manual annotation and shows promise in volume estimates on a small number of cases. However, no completely automatic segmentation algorithm has yet been adopted in the clinic. In Table I we present a concise review of the prior art in automatic tumor segmentation.

Both GBM and non-GBM methods are given in the table for completeness. Fuzzy clustering methods (voxelbased) across all tumor types appear to be the most popular approach. Philips et al. give an early proof-of-concept fuzzy clustering for brain tumor by operating on the raw multi-sequence data. They visually demonstrated that even with multi-sequence data the intensity distributions for tumor and normal tissue overlap. This led future researchers to incorporate additional knowledge into the feature vectors being clustered. Clark et al. integrate knowledge-based techniques and multi-spectral histogram analysis to segment GBM tumors in a multichannel feature space. Fletcher-Heathet al. takes a knowledge-based fuzzy clustering approach to the segmentation followed by 3D connected components to build the tumor shape. Prastawa et al also present a knowledge-based detection/segmentation algorithm based on learning voxel-intensity distributions for normal brain matter and detecting outlier voxels, which are considered tumor. The distributions are learned with kernel-based density estimation methods, and the initial outlier detection is followed by a region competition algorithm. Voxel-based statistical classification methods include. Kaus et al.use the adaptive template-moderated classification algorithm to segment the MR image into five different tissue classes: background, skin, brain, ventricles, and tumor. Their technique proceeds as an iterative sequence of spatially varying classification and non-linear registration. Prastawa et al. Define a parametric distribution across multiple channels of tumor as a mixture of Gamma and Gaussian components. They use the Expectation-Maximization algorithm to perform segmentation and iteratively adapt the model parameters to the case at hand. These two sets of methods are limited by their extreme degree of locality, i.e., they are voxel-based and do not take local or global context into account. While they have had some success in segmenting low-grade gliomas and meningiomas (relatively homogeneous) on a good-sized data set, their success is limited in the more relevant GBM (heterogeneous) segmentation examples. Furthermore, it is not clear this limited success will scale to the more difficult inevitable cases arising in larger data-sets (like the one used in this paper). There have been few attempts at solving this problem of local ambiguity.

GRADIENT DECENT ALGORITHM

Gradient descent is a first-order optimization algorithm. To find a local minimum of a function using gradient descent, one takes steps proportional to the *negative* of the gradient(or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the *positive* of the gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent. Gradient descent is also known as steepest descent, or the method of steepest descent. However, gradient descent should not be confused with the method of steepest descent for approximating integrals.

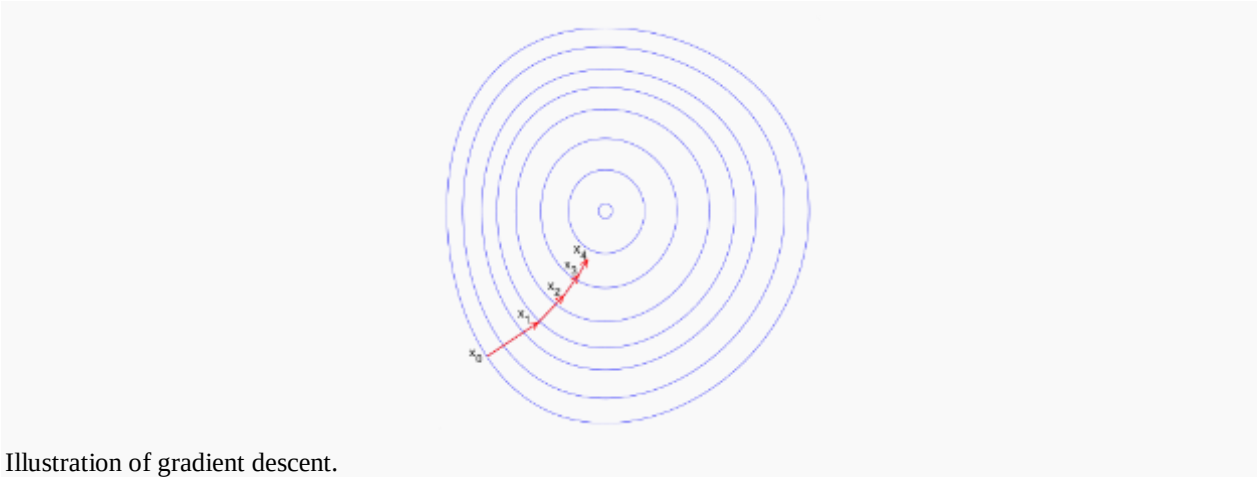


Illustration of gradient descent.

Gradient descent is based on the observation that if the multivariable

Function $F(\mathbf{x})$ is defined and differentiable in a neighborhoods of a point $\mathbf{a}$,then $F(\mathbf{x})$ decreases *fastest* if one goes from $\mathbf{a}$ in the direction of the negative gradient of

- F at $\mathbf{a}$, $-\nabla F(\mathbf{a})$.
- F It follows that,
- F If
- F $\mathbf{b} = \mathbf{a} - \gamma \nabla F(\mathbf{a})$
- F for γ small enough, then

$F(\mathbf{a}) \geq F(\mathbf{b})$.

With this observation in mind, one starts with a guess $\mathbf{x}_0$ for a local minimum of F, and considers the sequence $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \cdots$ such that

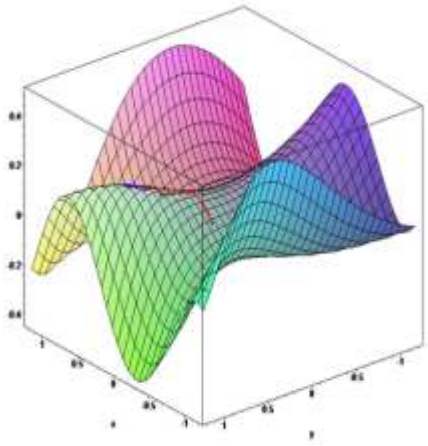
$$\mathbf{x}_{n+1} = \mathbf{x}_n - \gamma_n \nabla F(\mathbf{x}_n), \ n \geq 0.$$

We have $F(\mathbf{x}_0) \geq F(\mathbf{x}_1) \geq F(\mathbf{x}_2) \geq \cdots$, so hopefully the sequence $(\mathbf{x}_n)$ converges to the desired local minimum. Note that the value of the *step size* γ is allowed to change at every iteration. With certain assumptions on the function F (for example, F convex and ∇F Lipchitz) and particular choices of γ (e.g., chosen via a line search that satisfies the Wolfe conditions), convergence to a local minimum can be guaranteed. When the function F is convex, all local minima are also global minima, so in this case gradient descent can converge to the global solution. This process is illustrated in the picture to the right. Here F is assumed to be defined on the plane, and that its graph has a bowl shape. The blue curves are the contour lines, that is, the regions on which the value of F is constant. A red arrow originating at a point shows the direction of the negative gradient at that point. Note that the (negative) gradient at a point is orthogonal to the contour line going through that point. We see that gradient *descent* leads us to the bottom of the bowl, that is, to the point where the value of the function F is minimal. Gradient descent has problems with pathological functions such as the Rosenbrock function shown here.

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2.$$

The Rosenbrock function has a narrow curved valley which contains the minimum. The bottom of the valley is very flat. Because of the curved flat valley the optimization is zigzagging slowly with small step sizes towards the minimum.The "Zigzagging" nature of the method is also evident below, where the gradient descent method is applied

For some of the above examples, gradient descent is relatively slow close to the minimum: technically, its asymptotic rate of convergence is inferior to many other methods. For poorly conditioned convex problems, gradient descent increasingly 'zigzags' as the gradients point nearly orthogonally to the shortest direction to a minimum point. For more details, see the comments below. For non-differentiable functions, gradient methods are ill-defined. For locally Lipchitz problems and especially for convex minimization problems, bundle methods of descent are well-defined. Non-descent methods, like sub gradient projection methods, may also be used. These methods are typically slower than gradient descent. Another alternative for non-differentiable functions is to "smooth" the function, or bound the function by a smooth function. In this approach, the smooth problem is solved in the hope that the answer is close to the answer for the non-smooth problem (occasionally, this can be made rigorous).



$$F(x,y) = \sin\left(\frac{1}{2}x^2 - \frac{1}{4}y^2 + 3\right)\cos(2x + 1 - e^y).$$

Gradient descent can be used to solve a system of linear equations, reformulated as a quadratic minimization problem, e.g., using linear least squares. The solution of

$$A\mathbf{x} - \mathbf{b} = 0$$

In the sense of linear least squares is defined as minimizing the function

$$F(x) = \|A\mathbf{x} - \mathbf{b}\|^2.$$

In traditional linear least squares for real A and $\mathbf{b}$ the Euclidean norm is used, in which case.

$$\nabla F(\mathbf{x}) = 2A^T(A\mathbf{x} - \mathbf{b}).$$

In this case, the line search minimization, finding the locally optimal step size γ on every iteration, can be performed analytically, and explicit formulas for the locally optimal γ are known.[2]For solving linear equations, gradient descent is rarely used, with the conjugate gradient method being one of the most popular alternatives. The speed of convergence of gradient descent depends on the maximal and minimal eigenvalues of A, while the speed of convergence of conjugate gradients has a more complex dependence on the eigenvalues, and can benefit from preconditioning. Gradient descent also benefits from preconditioning, but this is not done as commonly. Gradient descent can also be used to solve a system of nonlinear equations. Below is an example that shows how to use the gradient descent to solve for three unknown variables, x_1, x_2, and x_3. This example shows one iteration of the gradient descent. This chapter introduces gradient descent and the conjugate gradient methods used in the project. It also describes how gradient descent works in finding a search direction, deciding on a step search to figure out how much to move in that direction, explicit and induced derivatives, and the complexity involved in the problem. Gradient descent, which is sometimes referred to as the method of steepest descent, is an optimization algorithm that helps you pick a search direction when trying to minimize a function. The equation to find the search direction is

$\sim zk = -\sim g + \beta k \sim zk-1$, (2.1)

Where $\sim z$ is the search direction, $-\sim g$ is the negative of the gradient, and β is a real valued coefficient representing an added momentum in the old search direction. For gradient descent, the value for β is zero. However, this value is used in the conjugate gradient methods discussed in the following section. Gradient descent can be used to find minima for everything from multi-dimensional functions to very large systems of equations. It can be thought of as a ball rolling down a hill and finding equilibrium at a minimum. An example is shown in Figure 1. In a system of equations that consists of many equations and

unknowns where not much else is known about the problem, it becomes one of the fastest options available for finding a numerical solution. One of the major competitors to gradient 5 6 descent is the multi-dimensional Newton's method. However, it becomes impractical for large problems, as it requires inverting large matrices at each iteration. This computation often leads to large rounding errors and can take a very long time to complete. Gradient descent is also very useful in problems where you have a general idea of where the solution should be located. Since gradient descent relies heavily on a starting guess, providing a guess that is close to a solution will usually lead to quick convergence to the solution.

CONCLUSION

In the paper a series of steps combining image analysis, image reconstruction and analysis, was analyzed. The output was compared to literature values and seemed to be in right interval, however since this only is a first version further analyses and development are needed in order to validate and confirm the accuracy with this method. All steps needs to be controlled and validated to ensure a correct visualization and quantification of regional liver function, however a first proposal was able to be produced in this paper.

REFERENCES

1. Björnerud A. The physics of magnetic resonance imaging in fys-kjm 4740. Department of Physics, 2008.
2. Geller David A, Goss John A, and Tsung Allan. Schwartz's principles of surgery9e. chapter 31. liver. E-book, 2011.http://www.accessmedicine.com/content.aspx?aID=5018142.
3. Westbrook C. MRI at a glance. John Wiley & Sons, Chichester, West Sussex, United Kingdom, 2010.
4. MacKay D. Chapter 20. an example inference task: Clustering. Information theory, inference and learning algorithms.
5. Nils Dahlström. Quantitative Evaluation of Contrast Agent Dynamics in Liver MRI. Phd thesis, Linköping University, 2010.
6. Levitt DG. The pharmacokinetics of the interstitial space in humans. BMC Clin Pharmacol, 3:3, 2003.
7. Larose DT. Discovering Knowledge in Data: An Introduction to Data Mining. John Wiley & Sons, NJ, USA, 2005.
8. Malmberg F and Lindblad J and Nyström I. Sub-pixel segmentation with the image foresting transform. In In Proceedings of the 13th International Workshop on Combinatorial Image Analysis (IWCIA'09), Playa del Carmen., 2009. www.cb.uu.se/ filip/publications.html.
9. Mikael Forsgren. Human whole body pharmacokinetic minimal model for the liver specific contrast agent gd-eob-dtpa, Jun 2011. 51 52 Bibliography.
10. Cedersund G and Roll J. Systems biology: Model based evaluation and comparison of potential explanations for given biological data. FEBS J, 276:903-922, 2009.
11. Liney G. MRI from A to Z. Springer-Verlag, London, Great Britian, 2005.
12. Weinmann HJ, Brasch RC, Press WR, and Wesbey GE. Characteristics of gadolinium-dtpa complex: A potential nmr contrast agent. Radiographics, 29:1725-1748, 2009.
13. Olof Dahlqvist Leinhard. Quantitative Magnetic Resonance in Diffuse Neurological and Liver Disease. Phd thesis, Linköping University, 2010.
14. Ghany M and Hoofnagle JH. Harrison's principles of internal medicine 17e. chapter 295. approach to the patient with liver disease. E-book, 2011. http://www.accessmedicine.com/content.aspx?aID=2873408.
15. Narita M, Hatano E, Arizono S, Miyagawa-Hayashino A, Isoda H, Kitamura K, Taura K, Yasuchika K, Nitta T, Ikai I, and Uemoto S. Expression of oatp1b3 determines uptake of gd-ebo-dtpa in hepatocellualar carcinoma. J Gastroenterol, 44:793-798, 2009