34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Non-parametric Bayesian Estimation of Apparent Diffusion Coefficient from Diffusion-Weighted Magnetic Resonance Imaging Data

Andrew Cameron, Jeffrey Glaister, Alexander Wong, Masoom A. Haider

Abstract— **A promising approach to prostate cancer diagnosis is multi-parametric MRI. One of the key modalities used in multi-parametric MRI is diffusion weighted MRI. Using multiple diffusion weighted MR acquisitions taken with different magnetic gradient strengths, the apparent diffusion coefficient (ADC) is calculated and can be used to identify tumors in the prostate. Current algorithms used to calculate ADC assume a parametric measurement model, but this assumption is not true due to the presence of additional phenomena during the acquisition process. A novel Non-parametric Estimated ADC (NEstA) algorithm is proposed which uses a Monte Carlo strategy to learn the inherent measurement distribution model based on the underlying statistical behavior of the DWI measurements to estimate the ADC values. The proposed algorithm is compared to the results of the commonly used least-squares (LS) estimation algorithm for computing ADC values. Nine test patient cases with visible tumors in the prostate gland were processed using both algorithms and compared visually. It was found that NEstA produced ADC data with reduced artifacts while preserving structure. Quantitatively, Fisher's criterion measuring the separability of the healthy prostate and tumor tissues was computed for the nine patient cases, comparing the NEstA and LS methods. It was found that Fisher's criterion increased with the NEstA method, meaning the separation of classes was more pronounced.**

I. Introduction

Prostate cancer is the second most common type of cancer in males, with 913,000 new diagnoses globally in 2008 [1]. It is also the sixth leading cause of death from cancer in males, with an estimated 258,000 deaths globally that same year [1]. If the cancer is detected early, the prognosis is excellent with a relative 5-year survival rate of 100% for cancer in the local stage [2]. This due not only to better treatments available, but also from better tests than can detect cancer earlier. One common test looks for an increased presence of prostate-specific antigen (PSA) in blood serum tests, which can indicate cancer in the prostate [3]. However, PSA tests can also indicate benign prostatic disease. Instead, a positive PSA test suggests that additional tests should be used to image the prostate and detect cancer.

Ultrasound is commonly used to image tumors in the prostate [4]. In conventional ultrasound techniques, the cancer should appear as a hypoechoic lesion. However, in many cases, the cancer is isoechoic and is not visible using ultrasound. Furthermore, prostatitis and benign prostatic hyperplasia can also appear as a hypoechoic lesion in ultrasound images, making it difficult to identify the prostate cancer.

Magnetic resonance imaging (MRI) is an alternative to ultrasound for imaging the prostate. In particular, multi-parametric MRI, which makes use of MRI modalities such as T2-weighted and diffusion-weighted MRI, has been used to locate prostate cancer [5], [6]. T2-weighted images detect the anatomical structures of the prostate. Diffusion-weighted MRI images (DWI) can detect prostate cancer from differences in the diffusion of water molecules of the normal and tumor tissues. A diffusion gradient is applied to measure the diffusion characteristics of the tissue and this gradient is quantified as the b-value. A particularly important parameter derived from diffusion-weighted MRI is the apparent diffusion coefficient (ADC), which is computed based on DWI acquisitions at different b-values [5]. The ADC value associated with each imaged tissue is the parameter of interest used to locate the cancer.

The ADC value associated with a specific tissue type is estimated from a set of diffusion-weighted images (DWI) and with different b-values. The relationship between these parameters is given in Eq. 1, where A is the ADC value, b_i is the b-value associated with signal intensity S_i, and S_α is the reference signal intensity associated with b_α.

$$S_i = S_\alpha e^{-(b_i - b_\alpha)A} \tag{1}$$

In practice, DWI measurements S are taken using different b-values, which allows for a more robust estimation of the ADC value. Previous estimation approaches rely exclusively on fixed, parametric measurement distribution models for the DWI measurements when estimating the ADC to account for measurement deviations (e.g., Rician measurement distribution model). Such approaches include the most common and widely used least squares (LS) estimation algorithm and parametric maximum likelihood approaches [7]. However, they do not perform well in situations where the fixed measurement distribution models employed may not fit well with the underlying DWI measurements, such as in the presence of additional phenomena during the acquisition process that is unaccounted for by the parametric models. Furthermore, the use of such models increase the complexity for fitting the signal decay model for DWI.

More recent approaches include minimizing total variation of ADC values, proposed by Chen et al [8]. While this approach produces low ADC variance inside the prostate

This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Ontario Ministry of Economic Development and Innovation.

Jeffrey Glaister, Andrew Cameron, and Alexander Wong are with Department of Systems Design Engineering, University of Waterloo, Ontario, Canada, N2L 3G1. {a4camero,jlglaist,a28wong}@uwaterloo.ca

Masoom A. Haider is with the the Department of Medical Imaging, Princess Margaret Hospital and the University Health Network, University of Toronto, Canada, M5S 1A1. m.haider@utoronto.ca

and tumor, it results in piece-wise smooth images with a noticeable loss in detail. Walker-Samuel, Orton, Boult and Robinson propose a Bayesian Adaptive Smoothing (BAS) based approach [9]. They propose using a Markov Chain Monte Carlo algorithm to learn parameters to a fixed parametric Bayesian Adaptive Smoothing model. A limitation to this approach is that a specific parametric model must be assumed, which is less adaptive to the underlying probabilistic behavior of the DWI measurements. Furthermore, due to the number of iterations required to learn the parameters, the algorithm runtime can be large.

In this paper, a Non-parametric Estimated ADC (NEstA) strategy is proposed to determine the ADC values. The strategy employs a Monte Carlo strategy to learn the inherent measurement distribution model directly based on the underlying statistical behavior of the DWI measurements. This allows the NEstA strategy to account for the probabilistic behavior of the DWI measurements in an adaptive and dynamic manner, hence resulting in more robust ADC estimation process while at the same time allowing for simpler assumptions to be made when fitting to the signal decay model for DWI.

The rest of the paper is organized as follows: in Section II, the methodologies of the proposed algorithm are presented. In Section III, quantitative and qualitative results are presented that demonstrate the performance of the proposed algorithm and compare it to existing algorithms. Finally, conclusions are given in Section IV.

II. Methodology

The proposed method approaches ADC estimation as the Bayesian estimation problem,

$$\hat{A} = \arg\max_{A} P(E(S|M)|A). \tag{2}$$

Here, M is the collection of DWI measurements, S is the collection of true DWI signals, A is the ADC, $E(S|M)$ is the conditional mean of S given M, and $P(\cdot|A)$ is the conditional probability given A. By assuming the measurements comprising M are statistically independent, the probability in Eq. 2 can be expressed as

$$P(E(S|M)|A) = \prod_{i} P(E_i(S|M)|A). \tag{3}$$

Since we can only observe S through our measurements, M, it is in M that random noise, as well as other unknown processes, which we will characterize collectively as η, are introduced. These contaminations of the signal are modeled as

$$M = S + \eta. \tag{4}$$

Given Eq. 4, the conditional mean $E(S|M)$ can be expressed as

$$E(S|M) = \int S\, p(S|M)\, dS\,, \tag{5}$$

where $p(S|M)$ is the posterior distribution of S given M. Most other approaches use a parametric model in order to estimate $p(S|M)$. However, parametric models do not work well when the signal contaminations do not fit the model assumptions. In order to better model a general unknown process η, we propose a non-parametric approach which learns the posterior directly from the data. We have implemented a Monte Carlo method for its ability to directly learn the underlying posterior in an adaptive, non-parametric manner [10].

Finally, since the arbitrary unknown processes η have been accounted for with $E(S|M)$, we can assume that

$$P(\cdot|A) \overset{\text{iid}}{\sim} \mathcal{N}(S_\alpha e^{-(b_i - b_\alpha)A}, \sigma^2) \tag{6}$$

in order to account for estimation error. For the purposes of our implementation, the number of samples used to estimate a single value in the ADC estimate is on the order of $100 \times m$, where m is the number of b-values used to estimate the ADC. For the datasets in Table I, on the order of 300 samples were used per ADC value, for those in Table II, on the order of 400 samples were used per ADC value, and for those in Table III, on the order of 700 samples were used per ADC value.

TABLE I: Summary of patients with DWI taken with 3 b-values $\{0, 100, 1000$ s/mm$^2\}$.

	Age	DFOV (cm^2)	Resolution (mm^3)	TE (ms)	TR (ms)
1	57	24×24	$1.36 \times 1.36 \times 4$	67	3336
2	56	24×24	$1.36 \times 1.36 \times 4$	67	3336
3	76	24×24	$1.36 \times 1.36 \times 4$	67	4876

TABLE II: Summary of patients with DWI taken with 4 b-values $\{0, 100, 400, 1000$ s/mm$^2\}$.

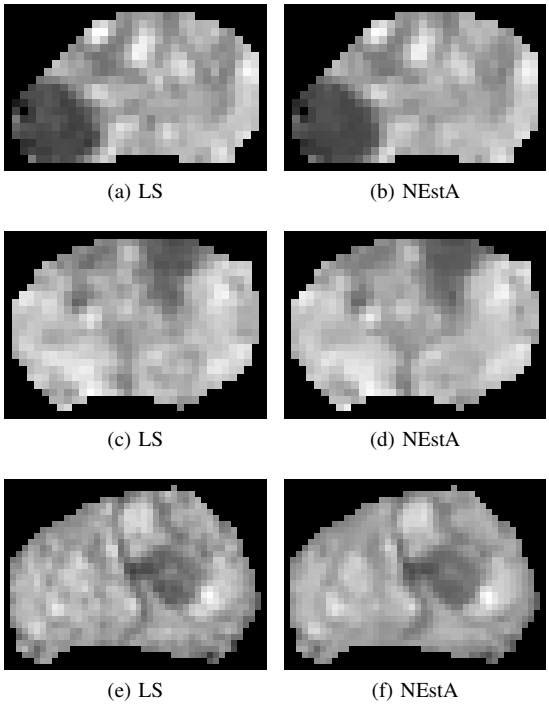
	Age	DFOV (cm^2)	Resolution (mm^3)	TE (ms)	TR (ms)
4	61	24×24	$1.66 \times 1.66 \times 3$	61	6153
5	76	20×20	$1.56 \times 1.56 \times 3$	61	6178
6	58	24×24	$1.66 \times 1.66 \times 3$	61	6153

III. Experimental Results

The experimental results comparing the proposed ADC estimation algorithm to using the widely used least squares (LS) algorithm to estimate the ADC values. Nine different patient cases with visible tumors in the prostate gland were used. Three of the patients had diffusion-weighted imaging performed using three different b-values, three of the patients had imaging performed for seven different b-values, and three patients had imaging performed for four different b-values. All MRI acquisitions were obtained using a Philips Achieva 3.0T machine at Sunnybrook Health Sciences Centre, Toronto, Canada. Information about the patient cases used is summarized in Tables I, II, and III, including Displayed Field of View (DFOV), echo time (TE), and repetition time (TR). The NEstA algorithm was performed within the ProCanVAS (Prostate Cancer Visual Analysis System) platform developed at the University of Waterloo Vision and Image Processing research group.

TABLE III: Summary of patients with DWI taken with 7 b-values $\{0, 10, 20, 50, 75, 100, 1000$ s/mm$^2\}$.

	Age	DFOV (cm^2)	Resolution (mm^3)	TE (ms)	TR (ms)
7	78	24×24	$1.66 \times 1.66 \times 3.5$	61	4890
8	66	24×24	$1.66 \times 1.66 \times 3.5$	61	4118
9	76	24×24	$1.66 \times 1.66 \times 3.5$	61	4890



(a) LS (b) NEstA

(c) LS (d) NEstA

(e) LS (f) NEstA

Fig. 1: Example ADC slices of prostates containing tumors derived from 3-b data.

A. Visual Comparison Results

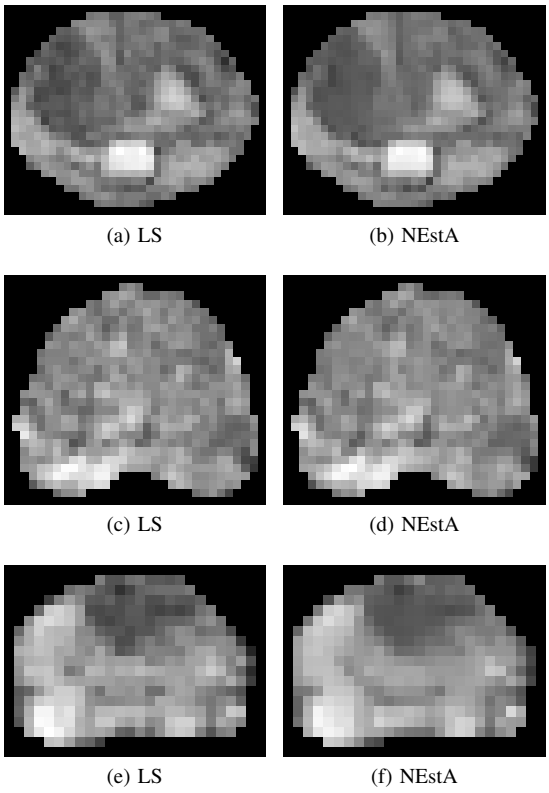
The NEstA approach achieves better homogeneity, both in the prostate region as well as in tumors, as shown in Figs. 1, 2, and 3; however, local structure is preserved. These figures also demonstrate better separation of the prostate from surrounding tissues, which will facilitate automatic prostate segmentation. We also observe from Fig. 4 that the contrast between the prostate and the tumor is improved by using the NEstA method. Furthermore, all figures demonstrate reduced artifacts.

B. Fisher's Criterion Results

For quantitative evaluation of the NEstA algorithm for the purpose of computerized cancer analysis, Fishers criterion was chosen for comparing separability of cancerous tissue from healthy tissue. Fisher's criterion is a measure of class separability, defined as

$$J(A) = \frac{|m_1 - m_2|^2}{s_1^2 + s_2^2}, \tag{7}$$

in the one-dimensional case, where m_1 and m_2 are the class means, and s_1^2 and s_2^2 are the class sample variances. A higher value of $J(A)$ indicates that the two classes (in this case,

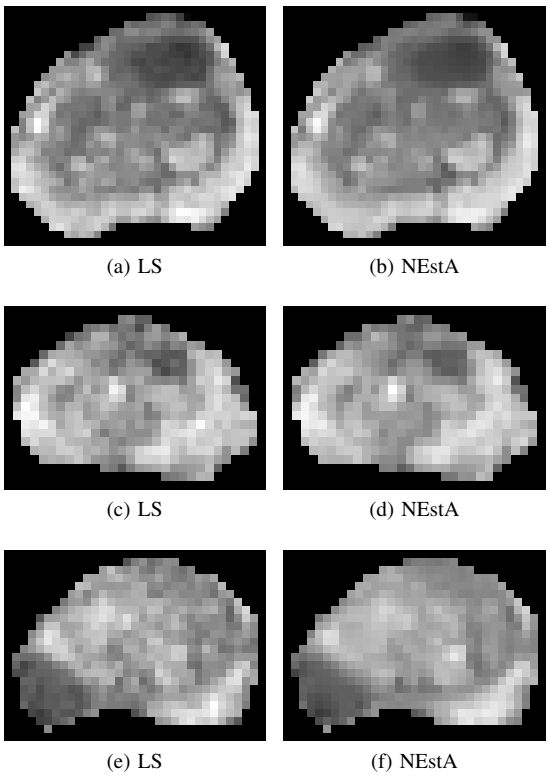


(a) LS (b) NEstA

(c) LS (d) NEstA

(e) LS (f) NEstA

Fig. 2: Example ADC slices of prostates containing tumors derived from 4-b data.

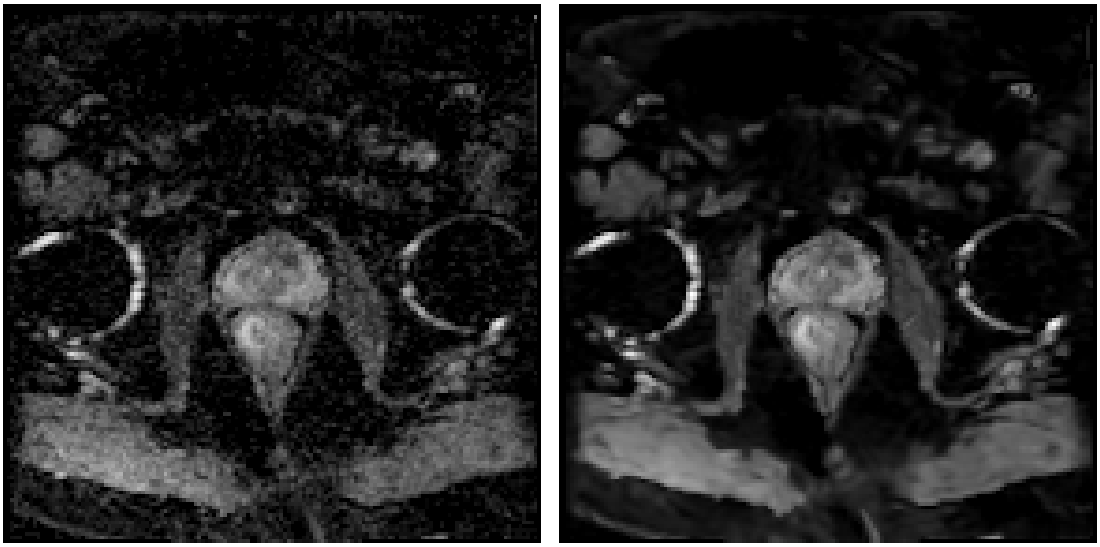


(a) LS (b) NEstA

(c) LS (d) NEstA

(e) LS (f) NEstA

Fig. 3: Example ADC slices of prostates containing tumors derived from 7-b data.



(a) LS (b) NEstA

Fig. 4: Full-size comparison of LS and NEstA estimations showing improved contrast between prostate and surrounding tissue in (b).

TABLE IV: Fisher's criterion for separability of cancerous and healthy tissues.

Patient #	LS	NEstA	% improvement
1	5.168	5.614	8.63
2	5.939	6.310	6.26
3	2.714	3.051	12.41
4	1.215	1.294	6.56
5	1.566	1.655	5.72
6	2.385	2.652	11.19
7	2.214	2.311	4.41
8	3.439	3.561	3.53
9	2.997	3.358	12.05

cancerous tissue vs. healthy tissue) are more easily separable [11].

Table IV compares prostate/tumor separability using Fisher's criterion for ADC estimates prepared using a least squares (LS) estimation method vs. the NEstA approach. These experimental results indicate that cancerous tissue is better separated from healthy tissue in all cases, with an average improvement over LS of 7.86%. Improved separability translates directly into better performance of classifiers designed to detect prostate cancer using ADC.

IV. CONCLUSIONS

A Non-parametric Estimated ADC (NEstA) strategy was presented for the estimation of ADC from DWI data. It departs from previous methods which use parametric techniques in that the NEstA strategy accounts for random phenomena in a direct non-parametric manner when estimating the ADC based on the probabilistic behavior of the underlying DWI data. The NEstA method is advantageous for its ability to adapt to unknown sources of error by learning the posterior directly from the data. Experimental results confirm that the separability of cancerous tissue from healthy prostate tissue, measured using Fisher's criterion, improved. Further improvements to the NEstA method presented herein are anticipated to improve the method such that it offers a better ADC estimate for the purpose of developing a very high b-value computed DWI system, which has the potential to greatly improve the identification and segmentation of cancerous tissue.

REFERENCES

[1] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, *Global Cancer Statistics*, CA: A Cancer Journal for Clinicians, vol. 61, iss. 2, pp. 69-90, March 2011.

[2] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, T. Murray, and M. J. Thun, *Cancer Statistics, 2008*, CA: A Cancer Journal for Clinicians, vol. 58, iss. 2, pp. 71-96, March 2008.

[3] U. Stenman, J. Leinonen, W. Zhang, and P. Finne, *Prostate-specific antigen*, Seminars in Cancer Biology, vol. 9, iss. 2, pp. 83-93, 1999.

[4] E. J. Halpern, *Contrast-Enhanced Ultrasound Imaging of Prostate Cancer*, Reviews in Urology, vol. 8, suppl. 1, pp. S29-S37.

[5] M. A. Haider, T. H. van der Kwast, J. Tanguay, A. J. Evans, A. Hashmi, G. Lockwood, and J. Trachtenberg, *Combined T2-weighted and diffusion-weighted MRI for localization of prostate cancer*, American Journal of Roentgenology, vol. 189, iss. 2, pp. 323-328, August 2007.

[6] D. L. Langer, T. H. van der Kwast, A. J. Evans, J. Trachtenberg, B. C. Wilson, and M. A. Haider, *Prostate Cancer Detection With Multi-Parametric MRI: Logistic Regression Analysis of Quantitative T2, Diffusion-Weighted Image, and Dynamic Contrast-Enhanced MRI*, Journal of Magnetic Resonance Imaging, vol. 30, iss. 2, pp. 327-334, August 2009.

[7] S. Walker-Samuel, M. Orton, L. D. McPhail, and S. P. Robinson, *Robust Estimation of the Apparent Diffusion Coefficient (ADC) in Heterogeneous Solid Tumors*, Magnetic Resonance in Medicine, vol. 62, iss. 2, pp. 420-429, August 2009.

[8] Y. Chen, et al., *Estimation, Smoothing, and Characterization of Apparent Diffusion Coefficient Profiles from High Angular Resolution DWI*, Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 588-593, July 2004.

[9] S. Walker-Samuel, M. Orton, J. K. R. Boult, and S. P. Robinson, *Improving Apparent Diffusion Coefficient Estimates and Elucidating Tumor Heterogeneity Using Bayesian Adaptive Smoothing*, Magnetic Resonance in Medicine, vol. 65, iss. 2, pp. 438-447, February 2011.

[10] W. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, vol. 57, iss. 1, pp. 97-109, April 1970.

[11] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973, pp. 114-129, 221-225.