Efficient Recursive Algorithms for Computing the Mean Diffusion Tensor and Applications to DTI Segmentation*

Guang Cheng, Hesamoddin Salehian, and Baba C. Vemuri**

Department of Computer and Information Science and Engineering,
University of Florida, Gainesville, FL
{gcheng,salehian,vemuri}@cise.ufl.edu

Abstract. Computation of the mean of a collection of symmetric positive definite (SPD) matrices is a fundamental ingredient of many algorithms in diffusion tensor image (DTI) processing. For instance, in DTI segmentation, clustering, etc. In this paper, we present novel recursive algorithms for computing the mean of a set of diffusion tensors using several distance/divergence measures commonly used in DTI segmentation and clustering such as the Riemannian distance and symmetrized Kullback-Leibler divergence. To the best of our knowledge, to date, there are no recursive algorithms for computing the mean using these measures in literature. Recursive algorithms lead to a gain in computation time of several orders in magnitude over existing non-recursive algorithms. The key contributions of this paper are: (i) we present novel theoretical results on a recursive estimator for Karcher expectation in the space of SPD matrices, which in effect is a proof of the law of large numbers (with some restrictions) for the manifold of SPD matrices. (ii) We also present a recursive version of the symmetrized KL-divergence for computing the mean of a collection of SPD matrices. (iii) We present comparative timing results for computing the mean of a group of SPD matrices (diffusion tensors) depicting the gains in compute time using the proposed recursive algorithms over existing non-recursive counter parts. Finally, we also show results on gains in compute times obtained by applying these recursive algorithms to the task of DTI segmentation.

1 Introduction

Finding the mean of a population of symmetric positive definite (SPD) matrices/tensors is an often encountered problem in medical image analysis and computer vision, specifically in diffusion MRI processing, tensor-based morphometry, texture analysis using the structure tensor etc. The mean tensor can be used to represent a population of structure tensors in texture analysis, diffusion tensors in diffusion tensor image (DTI) segmentation or for interpolation of diffusion

* This research was funded by the NIH grant NS066340 to BCV.
** Corresponding author.

A. Fitzgibbon et al. (Eds.): ECCV 2012, Part VII, LNCS 7578, pp. 390–401, 2012.

tensors or in clustering applications. It is also useful in atlas construction where the atlas is usually defined as the mean of a population of DTI data. It is well known that computation of the mean can be posed as a minimization problem in which one minimizes the sum squared distances between the unknown mean and the members of the set whose mean is being sought. Mathematically speaking, we want to find, $\mu^* = \min_\mu \sum_i^n d^2(x_i, \mu)$, where, d is the chosen distance, x_i are the data samples whose mean is being sought and μ^* is the mean. Depending on the definition of distance d, one gets different kinds of means. For example, if we choose the Euclidean distance for d, we get the arithmetic mean, where as if we choose the geodesic distance in the domain of x_i, we get the Karcher mean and so on. Also, if we chose the L_1-norm instead of the L_2 norm in the above formula, we get the median.

Tomes of research has been published on finding the mean tensor using different kinds of distances/divergences and has been applied to DTI as well as structure tensor field segmentation, interpolation and clustering. In [1], authors generalized the geometric active contour model-based piece-wise constant segmentation [2,3] to segmentation of DTIs using the Euclidean distance to measure the distance between two SPD tensors. Authors in [4], present a geometric active contour [5,6] based approach for tensor field segmentation that used information from the diffusion tensors to construct the so called structure tensor which is a sum of structure tensors formed from each component of the diffusion tensor. A Riemannian metric on the manifold of SPD matrices was used in [7,8,9] and [10,11,12] for DTI segmentation and for computing the mean interpolant of diffusion tensors respectively. In [13,14,15] and [10] the symmetrized KL-divergence was used for DTI segmentation and interpolation respectively. The *Log*-Euclidean distance was introduced to simplify the computations on the manifold of SPD matrices and this was achieved by using the principal *Log*-map from the manifold to its tangent space and then using the Euclidean metric on the *Log*-mapped matrices in the tangent space at the identity [16]. More recently, in [9], a statistically robust measure called the total Bregman divergence (tBD) family was introduced and used for interpolation as well as DTI segmentation.

None of the above methods for computing the mean of SPD matrices which are used within the segmentation algorithms or in their own right for interpolation purposes are in recursive form with the exception of the Log-Euclidean metric based scheme for which a recursive form was proposed in [17]. A recursive formulation would be more desirable as it would yield a computationally efficient algorithm for computing the means of regions in the segmentation application. Also, in many applications such as DTI segmentation, clustering and atlas construction, data are incrementally supplied to the algorithm for classification or assimilation to update the mean and an algorithm that recursively updates the mean rather than one that recomputes the mean in a batch mode would be much more efficient and desirable. In this paper, we pursue this very task of recursive mean computation. The key contributions of this paper are: (i) first, we present novel theoretical results proving the probabilistic convergence of the recursive intrinsic Karcher expectation computation of a set of SPD matrices to the true

Karcher expectation. This is a significant result namely, a restricted form of the *law of large numbers for the space of SPD matrices*. (ii) We present recursive formulations for computing the mean using commonly used distance/divergence measures mentioned above and present experiments that depict significant gains in compute time over their non-recursive counterparts. (iii) We present synthetic and real data experiments depicting gains in compute time for DTI segmentation using these recursive algorithms.

The rest of the paper is organized as follows: in Section 2 we present novel theoretical results leading to the intrinsic Karcher expectation computation algorithm. In addition, we present the recursive formulations for the commonly used symmetrized KL-divergence based mean computation as well as the Log-Euclidean distance based mean. Section 3 contains synthetic and real data experiments depicting the improvements in computation time of the DTI segmentation task. Finally, we present the conclusions in Section 4.

2 Methods

2.1 The Recursive Karcher Expectation Estimator

Let P_n denote the space of $n \times n$ SPD matrices, which is not a vector space, since it is not closed under vector operations. Instead, the General Linear Group ($GL(n)$ denoting the group of $n \times n$ non-singular matrices) is the natural group action on P_n, where the group action can be defined as $\forall \mathbf{g} \in GL(n), \forall \mathbf{M} \in P_n$, $\mathbf{M}[\mathbf{g}] = \mathbf{g}\mathbf{M}\mathbf{g}^{\mathbf{t}}$. Let $\mathbf{U}, \mathbf{V} \in T_{\mathbf{M}}P_n$, where $T_{\mathbf{M}}P_n$ denotes the tangent space of P_n at point $\mathbf{M}$. The GL invariant metric (a metric invariant to GL group action) in P_n can then be uniquely defined as

$$< \mathbf{U}, \mathbf{V} >_{\mathbf{M}} = trace(\mathbf{M}^{-1}\mathbf{U}\mathbf{M}^{-1}\mathbf{V}) \tag{1}$$

With this metric, the curve length on the manifold can then be evaluated and the geodesic distance between any given points $\mathbf{M}, \mathbf{N} \in P_n$ can be computed in a closed form

$$dist(\mathbf{M}, \mathbf{N})^2 = trace(Log(\mathbf{M}^{-1}\mathbf{N})^2) \tag{2}$$

where Log is the matrix logarithm. As the distance is defined, given m points $\mathbf{M}_1, \mathbf{M}_2, ...\mathbf{M}_m \in P_n$ the center of the mass on P_n can then be defined as the minimizer of the sum of squared geodesic distance $\mu^* = argmin_\mu \sum_{i=1}^{m} dist(\mu, \mathbf{M}_i)^2$, which is also known as Karcher mean[18]. Unlike the arithmetic mean in the Euclidean space, the Karcher mean on P_n has no existing analytic (closed form) solution in P_n for $m > 2$. A gradient based optimization is needed whenever we need to compute this Karcher mean [10], which makes it computationally expensive.

The Karcher mean can be viewed as an extension of the arithmetic mean from the Euclidean space to the Riemannian manifold. Similarly, the expectation and

the variance can also be extended. Given a random variable $\mathbf{M} \in P_n$ with a probability density $P(\mathbf{M})$

$$E(\mathbf{M}) = argmin_{\mu} \int_{P_n} dist(\mu, \mathbf{X})^2 P(\mathbf{X})[d\mathbf{X}] \tag{3}$$

$$Var(\mathbf{M}) = \int_{P_n} dist(E(\mathbf{M}), \mathbf{X})^2 P(\mathbf{X})[d\mathbf{X}] \tag{4}$$

where $[d\mathbf{X}]$ denotes the unique GL invariant measure on P_n.

We now develop an estimator for the intrinsic (Karcher) expectation that can be used to represent a set of data points in P_n (the space of diffusion tensors) and can be computed recursively. This recursive computation property is a very important property especially for online problems where the data points are provided sequentially. This is very pertinent to applications such as DTI and structure tensor field segmentation, diffusion/structure tensor clustering etc.

Let $\mathbf{X}_k \in P_n$, $k = 1, 2, ...$ be iid samples in P_n from probability measure $P(\mathbf{M})$. The recursive Karcher expectation estimator can be defined as:

$$\mathbf{M}_1 = \mathbf{X}_1 \tag{5}$$

$$\mathbf{M}_{k+1}(w_{k+1}) = \mathbf{M}_k^{\frac{1}{2}} (\mathbf{M}_k^{-\frac{1}{2}} \mathbf{X}_{k+1} \mathbf{M}_k^{-\frac{1}{2}})^{w_{k+1}} \mathbf{M}_k^{\frac{1}{2}} \tag{6}$$

Here we set $w_{k+1} = \frac{1}{k+1}$. We now prove the following properties of the recursive Karcher expectation estimator presented here in the form of theorems with their proofs.

Theorem 1. Let iid samples $\mathbf{X}_k$ be generated from a density $P(\mathbf{X}; \mu)$ that is symmetric w.r.t. to its expectation μ, then $\mathbf{M}_k$ is a unbiased estimator. By symmetry we mean that $\forall \mathbf{X} \in P_n \; P(\mathbf{X}; \mu) = P(\mu \mathbf{X}^{-1} \mu; \mu)$, note that $\mathbf{X}, \mu, \mu \mathbf{X}^{-1} \mu$ are on the same geodesic and $dist(\mu, \mathbf{X}) = dist(\mu, \mu \mathbf{X}^{-1} \mu)$

Proof. Without loss generality we assume that $\mu = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. We can now prove the theorem by induction. For $k = 1$,

$$E(\mathbf{M_1}) = E(\mathbf{X_1}) = \mathbf{I}$$

where E denotes the Karcher expectation. $P_{m_1}(\mathbf{M}_1; \mathbf{I})$ is obviously symmetric.

Assuming $E(\mathbf{M_k}) = \mathbf{I}$, and the the posterior $P_{m_k}(\mathbf{M}_k)$ is symmetric. Then,

$$P_{m_{k+1}}(\mathbf{M}_{k+1}) =$$

$$\int_{P_n} P_x(\mathbf{X}_{k+1}) P_{m_k}(\mathbf{X}_{k+1}^{\frac{1}{2}} (\mathbf{X}_{k+1}^{-\frac{1}{2}} \mathbf{M}_{k+1} \mathbf{X}_{k+1}^{-\frac{1}{2}})^{w_{k+1}-1} \mathbf{X}_{k+1}^{\frac{1}{2}})[d\mathbf{X}_{k+1}] =$$

$$P_{m_{k+1}}(\mathbf{M}_{k+1}^{-1})$$

since P_x, P_{m_k} are symmetric and

$$(\mathbf{X}_{k+1}^{\frac{1}{2}} (\mathbf{X}_{k+1}^{-\frac{1}{2}} \mathbf{M}_{k+1} \mathbf{X}_{k+1}^{-\frac{1}{2}})^{w_{k+1}-1} \mathbf{X}_{k+1}^{\frac{1}{2}})^{-1} = \mathbf{X}_{k+1}^{-\frac{1}{2}} (\mathbf{X}_{k+1}^{\frac{1}{2}} \mathbf{M}_{k+1}^{-1} \mathbf{X}_{k+1}^{\frac{1}{2}})^{w_{k+1}-1} \mathbf{X}_{k+1}^{-\frac{1}{2}}$$

Thus, $P_{m_{k+1}}$ is symmetric with respect to $\mathbf{I}$, and $E(\mathbf{M}_{k+1}) = I = \mu$ since

$$\int_{P_n} Log(\mathbf{M}) P_{m_{k+1}}(\mathbf{M})[d\mathbf{M}] = \int_{P_n} Log(\mathbf{N}^{-1}) P_{m_{k+1}}(\mathbf{N}^{-1})[d\mathbf{N}^{-1}] = 0$$

Theorem 2. $\forall \mathbf{A}, \mathbf{B} \in P_n$, and $w \in [0,1]$

$$||Log(\mathbf{A}^{\frac{1}{2}}(\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{A}^{-\frac{1}{2}})^w\mathbf{A}^{\frac{1}{2}})||^2 \leq tr(((1-w)Log(\mathbf{A}) + wLog(\mathbf{B}))^2) \quad (7)$$

Note that the left side of the inequality is the square distance between the identity matrix and a geodesic interpolation between $\mathbf{A}$ and $\mathbf{B}$, and the right side of the inequality is the square distance between the identity and the log linear interpolation. This inequality is true based on the fact that P_n is a space with non-positive sectional curvature.

Proof. Let $\gamma(w) = \mathbf{A}^{\frac{1}{2}}(\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{A}^{-\frac{1}{2}})^w\mathbf{A}^{\frac{1}{2}}$. Then $\gamma(w)$ is a geodesic between $\mathbf{A}, \mathbf{B}$ Since P_n is a Hadamard space, based on Lurie's notes on Hadamard space [19], we know that

$$lhs = dist(\mathbf{I}, \gamma(w))^2 \leq (1-w)dist(\mathbf{I}, \mathbf{A})^2 + wdist(\mathbf{I}, \mathbf{B})^2 - w(1-w)dist(\mathbf{A}, \mathbf{B})^2 \quad (8)$$

Also

$$\begin{aligned} rhs - ((1-w)dist(\mathbf{I}, \mathbf{A})^2 + wdist(\mathbf{I}, \mathbf{B})^2 - w(1-w)dist(\mathbf{A}, \mathbf{B})^2) \\ = w(1-w)(dist(\mathbf{A}, \mathbf{B})^2 - tr(Log\mathbf{A} - Log\mathbf{B})^2) \geq 0 \end{aligned} \quad (9)$$

where the last inequality is based on the Cosine inequality in [20]. Thus, we have proved that $lhs \leq rhs$.

Theorem 3. Let $w_k = \frac{1}{k}$, we then have, $Var(\mathbf{M}_k) \leq \frac{1}{k}u^2$, where $u^2 = Var(\mathbf{X}_i)$, $i = 1, 2....$

Proof. We will use induction to prove the theorem. We still assume that $E(\mathbf{X}_k) = \mathbf{I}$. When $k = 1$, $Var(\mathbf{M}_1) = Var(\mathbf{X}_1) = u^2$.

Assume that the claim is true for $k = i$, that is $Var(\mathbf{M}_i) \leq \frac{1}{i}u^2$. When $k = i + 1$, using the lemma above, we know that,

$$\begin{aligned} Var(\mathbf{M}_{i+1}(w)) \leq \\ \int_{P_n}\int_{P_n} ||(1-w)Log(\mathbf{M_i}) + wLog(\mathbf{X_{i+1}})||^2 P(\mathbf{M}_i)P(\mathbf{X}_{i+1})[d\mathbf{M}_i][d\mathbf{X}_{i+1}] \\ = (1-w)^2 Var(\mathbf{M_i}) + w^2 Var(\mathbf{X}_{i+1}) \\ \leq (1-w)^2\frac{1}{i}u^2 + w^2u^2 \end{aligned} \quad (10)$$

$(1-w)^2\frac{1}{i}u^2 + w^2u^2 = \frac{1}{i+1}u^2$ when $w = \frac{1}{i+1}$.

From the theorems above, we can find that the recursive Karcher expectation estimator is an unbiased estimator for the Karcher expectation when the samples are drawn from a symmetric distribution on P_n. And it weakly converges to the expectation. *Note that what we have proved is a restricted form of the law of large numbers*, where the restriction is that the samples be drawn from a symmetric distribution on P_n. Of course, this recursive Karcher expectation estimator can be viewed as an approximation of the Karcher sample mean. However,

in our experiments, we find that it actually has similar accuracy as the Karcher sample mean. Also, because it is a recursive estimator, it would be far more computationally efficient to use our estimator than the standard non-recursive Karcher mean algorithm when the diffusion tensors are input sequentially to the estimation algorithm as in all the aforementioned applications.

2.2 Recursive Form of the Symmetrized KL-Divergence Mean

We now present a recursive formulation for computing the symmetrized KL-divergence based mean. Let's recall, the symmetrized KL divergence also called the J-divergence, is defined by $\mathbf{J}(p,q) = \frac{1}{2}(\mathbf{KL}(p||q)+\mathbf{KL}(q||p))$ Using the square root of J, one can define a divergence between two given positive definite tensors. The symmetrized KL (KL_s) divergence based mean of a set of SPD tensors is the minimizer of the sum of squared KL divergences. This minimization problem has a closed form solution as shown in Wang et al. [13] and regurgitated here for convenience,

$$\mathbf{M_{KL}} = \sqrt{\mathbf{B^{-1}}}[\sqrt{\sqrt{\mathbf{B}}\mathbf{A}\sqrt{\mathbf{B}}}]\sqrt{\mathbf{B^{-1}}} \tag{11}$$

where $\mathbf{A} = \frac{1}{N}\sum_i \mathbf{T_i}$ is the arithmetic mean, $\mathbf{B} = \frac{1}{N}\sum_i \mathbf{T_i^{-1}}$ is the harmonic mean, $\mathbf{T} = \{\mathbf{T_i}\}$ is the given tensor field and N is the total number of tensors. The closed form equation 11 can be computed in an recursive manner as follows. Let the arithmetic and harmonic means at iteration n be denoted by (A_n) and (B_n), respectively. When a new $*(n+1)^{st}$ tensor, $\mathbf{T_{n+1}}$ augments the data set, the quantities A_n and B_n are recursively updated via the following simple equations,

$$\mathbf{A_{n+1}} = \frac{n}{n+1}\mathbf{A_n} + \frac{1}{n+1}\mathbf{T_{n+1}} \tag{12}$$

$$\mathbf{B_{n+1}} = \frac{n}{n+1}\mathbf{B_n} + \frac{1}{n+1}\mathbf{T_{n+1}^{-1}}. \tag{13}$$

Using the above recursive form of the arithmetic and harmonic means of a set of tensors and the closed form expression 11, we can recursively compute the KL_s mean of a set of SPD tensors.

2.3 Recursive Mean for the Log-Euclidean Metric

We now present the recursive form of the Log-Euclidean (LE) based mean. It is well known that P_n can be diffeomorphically mapped to the Euclidean space using the matrix *Log* function, which makes it possible to directly induce the Euclidean metric on P_n called the *Log*-Euclidean metric [16]. Then, the *Log*-Euclidean distance can be defined as,

$$\mathbf{D_{LE}}(\mathbf{T_1}, \mathbf{T_2}) = ||Log(\mathbf{T_1}) - Log(\mathbf{T_2})|| \tag{14}$$

where ||.|| is the Euclidean norm. The LE-mean on a set of SPD matrices, is obtained by minimizing the sum of the squared LE distances which leads to a closed form solution

$$\mathbf{M_{LE}} = Exp(\sum_{i=1}^{n} Log(\mathbf{T_i})) \tag{15}$$

This closed form expression can be rewritten in a recursive form for more efficient computation. Let $\mathbf{M_n}$ be the Log-Euclidean mean in the n^{th} iteration. When the $(n+1)^{st}$ tensor, say $\mathbf{T_{n+1}}$ is added, the current mean can be recursively updated using the following equation,

$$\mathbf{M_{n+1}} = Exp(\frac{n}{n+1} Log(\mathbf{M_n}) + \frac{1}{n+1} Log(\mathbf{T_{n+1}})). \tag{16}$$

We were made aware of this recursive form having appeared in [17] by one of the reviewers which was not known to us at the time of this submission.

3 Experiments

3.1 Performance of the Recursive Estimators

To justify the performance of the recursive estimators, we first generate iid samples from the Log-normal distribution [21] on P_3 with the expectation at the identity matrix. Then, we input the 100 random samples sequentially to all estimators including the recursive karcher expectation estimator ($RKEE$), Karcher mean (KM), recursive KL_s mean ($RKLS$), non-recursive KL_s (KLS) mean, recursive *Log*-Euclidean mean ($RLEM$) and the non-recursive *Log*-Euclidean mean (LEM) respectively. To compare the accuracy of $RKEE$ and KM, we evaluate the error of the estimator using the squared distance in Equation 2 between the ground truth (the identity matrix) and the computed estimate. *The accuracy test of the remaining algorithms is not included because for KL_s and Log-Euclidean "metrics" , the recursive and non-recursive algorithm will generate the exact same results.* Also, the computation time for each step (each sample) is recorded. For comparison, we have the same settings for all the mean computation algorithms. We run the experiment 20 times and plot the average error and the average computation time at each step in Figure 1. In Figure 1 (a), we see that the accuracy of computed mean is nearly the same for both the non-recursive Karcher mean and the recursive Karcher expectation estimators after they are given 10 samples. The computation time (in CPU seconds on an I-7, $2.8GHZ$ processor) for the Karcher mean however increases linearly with the number of steps, while that for the recursive Karcher expectation estimator is nearly a constant and far less than the non-recursive case. This means that the recursive Karcher expectation estimator is computationally far superior especially for large size problems where data is input incrementally, for example in algorithms for segmentation, clustering, classification and atlas construction. Similar conclusions can also be drawn in Figure 1 (c) and (d), where for sequentially input data the recursive mean algorithm for KL_s divergence and *Log*-Euclidean mean are much more efficient than their own batch versions.

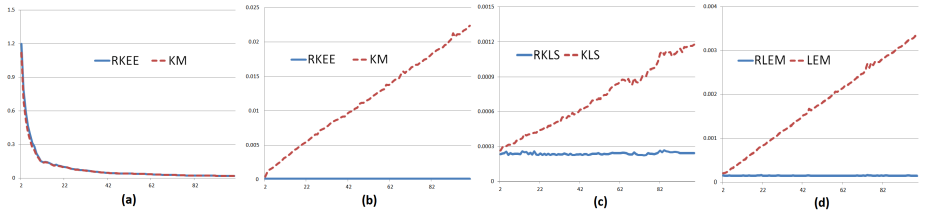

Fig. 1. Accuracy and speed comparisons of the recursive versus non-recursive mean computation algorithms for data on P_3. Figure (a) is the mean error of the Karcher mean (red dashed line) and the recursive Karcher expectation estimator (blue solid line) for each step. Figures (b) (c) (d) are the comparisons of computation time (in seconds) between the recursive (red dashed line) and non-recursive (blue solid line) mean computation algorithms for different "metrics" respectively. Results for the Riemannian metric are shown in Figure (b), KL_s in (c), *Log*-Euclidean in (d) respectively.

3.2 Application to DTI Segmentation

In this section, we present results of applying our recursive algorithms to the DTI segmentation problem. In [13], the classical levelset based segmentation algorithm [2] was extended to the field of diffusion tensors. In this algorithm, based on a piecewise constant model, the segmentation procedure became a EM like algorithm, where at each iteration, the mean tensor is computed over each region and the region boundary is then evolved based on the mean tensor. In this section, we use this algorithm to segment DTIs, and plug in different tensor field mean computation techniques for comparison.

Firstly, experiments on DTI segmentation of synthetic datasets is presented here. We manually generated an image region of size (64×64) which contains two different kind of tensors (differing in orientation, one vertical and another horizontal). Then DW-MR signal is generated based on [22] with 5 different level of Riccian noised added to the DW-MR signal $\sigma = 0.1, 0.15, 0.2, 0.25, 0.3$, where σ^2 is variance of the Gaussian noise added to the real and image part of the DW-MR signal. DTIs are constructed by using the technique in [23]. Exact same dataset and same setting are used for all six methods. The initialization curve overlaid on a noisy dataset is depicted in Figure 2 (a). To evaluate the segmentation result, the dice coefficient between the ground truth segmentation and the estimated segmentation are computed. These results are shown in Figure 2 with Figure (b) depicting the dice coefficients and Figure (c) showing the comparison of the running times. From the Figure (b) we can see that the segmentation accuracies are very similar for the recursive and non-recursive methods for the same "distance metric". For different "distance metric"s, result of the Riemannian (GL invariant) metric is the most accurate, since the GL invariant metric is the natural metric on P_n. In Figure 2, we can find that segmentation using the KM takes significantly longer time than other methods, this is because there is no closed form computation formula for Karcher mean on P_n, and hence the Karcher mean computation is very time consuming which can also be seen in

Table 1. The recursive Karcher expectation estimator is about 2 times faster and has the similar accuracy. For KL_s and Log-Euclidean "metrics", the time saved by the recursive method is not so significant as for the GL-invariant metric. This is because, although the mean computation time for the recursive method is at least one tenth of the non-recursive method (0.01 versus 0.1 in Table 1), the CPU time used for curve evolution is about $1 \sim 4$ seconds which makes the savings in total segmentation time not significant. From these results we can find that the RKEE is the most attractive from an accuracy and efficiency viewpoint.

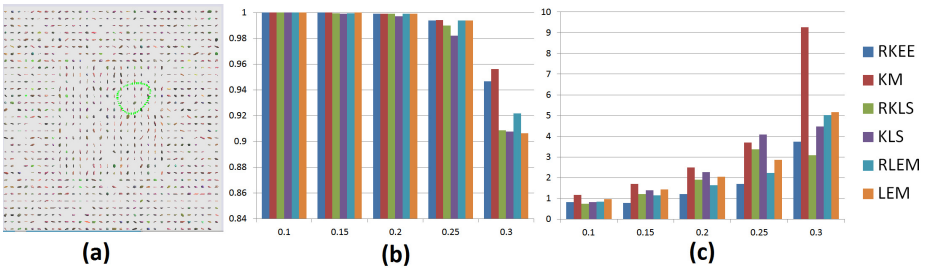

Fig. 2. Results for the DTI segmentation experiments on the synthetic dataset. Figure (a) is the initialization overlaid on the synthetic dataset at one of the noise levels used in the experiments. Figure (b) is the segmentation accuracy evaluated by the dice coefficient of the segmentations from all the methods at different noise levels. Figure (c) is the total segmentation time (in seconds) for all methods at different noise levels.

Table 1. Time (in seconds) for mean computation in the DTI segmentation on synthetic dataset

Noise Level	$RKEE$	KM	$RKLS$	KLS	$RLEM$	LEM
0.1	0.01	0.75	0.01	0.12	0.01	0.13
0.15	0.007	1.02	0.006	0.23	0.01	0.15
0.2	0.02	1.55	0.01	0.42	0.005	0.22
0.25	0.01	2.22	0.01	0.63	0.006	0.28
0.3	0.01	4.52	0.008	0.55	0.01	0.45

For the real data experiment, the DTI are estimated [13] from a DW-MR scan of a rat spinal cord. The DW-MR data were acquired using a PGSE with $TR = 1.5s$, $TE = 28.3ms$, bandwidth $= 35$ Khz, 21 diffusion weighted images with a b-value of $1250s/mm^2$ were collected. The image size is $128 \times 128 \times 10$. We used the same initialization for each segmentation. We applied all of the six methods (recursive and non-recursive for each of the three "distance" measures) to perform this experiment. In order to compare the time efficiency, we report the whole segmentation running time, including the total time required to compute the means. Table 2 shows the result of this comparison, from which we can find that it is much more efficient to use the recursive mean estimator in the segmentation than using the batch mean estimator. Especially, in the case of the

Karcher mean, which has no closed form formula and takes nearly half of the total reported segmentation time, whereas, using the recursive Karcher expectation estimator makes the computation much faster, and also significantly reduces the total segmentation time. The segmentation results are depicted in Figure 3 for each method. Each $(3, 3)$ diffusion tensor in the DTI data are illustrated as an ellipsoid whose axis directions and lengths correspond to the eigen-vectors and eigen-values respectively. From the figure we can see that the segmentation results are visually similar to each other, while our recursive Karcher expectation based method takes much less time which would be very useful in practice.

Table 2. Timing in seconds for segmentation of grey matter in a rat spinal cord

Segmentation Method	$RKEE$	KM	$RKLS$	KLS	$RLEM$	LEM
Mean computation time	0.02	3.56	0.01	0.56	0.01	0.4
Total segmentation time	5.09	8.13	3.41	4.41	5.45	5.82

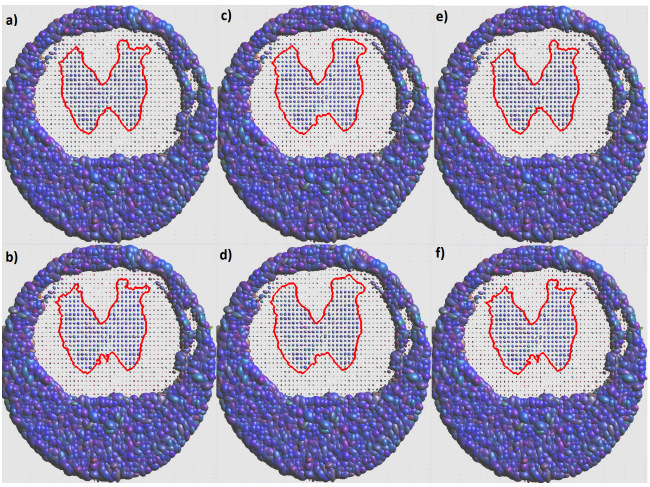

Fig. 3. Segmentation results of grey matter in a rat spinal cord for 6 different methods. Figure (a) is $RKEE$ based segmentation. Figure (b) is segmentation using the Karcher mean KM. Figure (c) and (d) are results for the recursive and non-recursive KL_s mean estimators respectively. Figure (e) and (f) are results for the recursive and non-recursive Log-Euclidean mean respectively.

A second real data set from an isolated rat-hippocampus was used to test the segmentation algorithms. Figure 4 depicts the segmentation of the molecular layer in the rat hippocampus. For the sake of space, we present only the segmentation results from the recursive algorithms presented in this paper and not the non-recursive counterparts as their results are visually similar and the key difference is in the time savings.

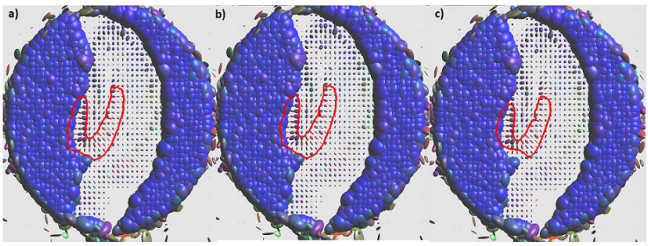

Fig. 4. Segmentation results of the molecular layer in a rat hippocampus for 3 different methods. Figure (a) $RKEE$ based segmentation. (b) Recursive KL_s based segmentation (c) Recursive Log-Euclidean based segmentation.

4 Conclusion

In this paper, we presented a novel recursive Karcher expectation estimator (RKEE) on P_n which can be used to recursively estimate the expectation of the data distribution by taking data input sequentially. We proved that under the symmetric distribution assumption, the RKEE is an unbiased estimator and it converges (weakly) to the Karcher expectation, which to our knowledge has never been proved before. In effect, we proved the Law of Large Numbers for P_n for samples drawn from symmetric distributions. Recursive algorithms to compute the KL_s mean and Log-Euclidean mean are also presented in the paper. The synthetic and real data experiments with comparison demonstrated the efficiency of our proposed method.

References

1. Wang, Z., Vemuri, B.C.: Tensor Field Segmentation Using Region Based Active Contour Model. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3024, pp. 304–315. Springer, Heidelberg (2004)
2. Chan, T., Vese, L.: Active contours without edges. IEEE Trans. on Image Proc. 10(2), 266–277 (2001)
3. Tsai, A., Yezzi, A.J., Willsky, A.: Curve Evolution Implementation of the Mumford-Shah Functional for Image Segmentation, Denoising, Interpolation, and Magnification. IEEE Trans. on Image Proc. 10(8), 1169–1186 (2001)
4. Feddern, C., Weickert, J., Burgeth, B.: Level-set Methods for Tensor-valued Images. In: Proc. 2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Comp. Vis., pp. 65–72 (2003)
5. Malladi, R., Sethian, J., Vemuri, B.C.: Shape Modeling with Front Propagation: A Level Set Approach. IEEE Trans. on PAMI 17(2), 158–175 (1995)
6. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. Intl. Journ. of Compu. Vision 22(1), 61–79 (1997)
7. Lenglet, C., Rousson, M., Deriche, R.: Dti segmentation by statistical surface evolution. IEEE Trans. on Medical Imaging 25(6), 685–700 (2006)

8. Goh, A., Vidal, R.: Segmenting Fiber Bundles in Diffusion Tensor Images. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part III. LNCS, vol. 5304, pp. 238–250. Springer, Heidelberg (2008)
9. Vemuri, B., Liu, M., Amari, S., Nielsen, F.: Total bregman divergence and its applications to dti analysis. IEEE Trans. on Medical Imaging 30(2), 475–483 (2011)
10. Moakher, M., Batchelor, P.G.: Symmetric Positive-Definite Matrices: From Geometry to Applications and Visualization. Visualization and Processing of Tensor Fields. Springer (2006)
11. Pennec, X., Fillard, P., Ayache, N.: A riemannian framework for tensor computing. International Journal of Computer Vision 66(1), 41–66 (2006)
12. Barmpoutis, A., Vemuri, B.C., Shepherd, T.M., Forder, J.R.: Tensor splines for interpolation and approximation of DT-MRI with applications to segmentation of isolated rat hippocampi. IEEE Trans. Med. Imag. 26 (2007)
13. Wang, Z., Vemuri, B.: Dti segmentation using an information theoretic tensor dissimilarity measure. IEEE Trans. on Medical Imaging 24(10), 1267–1277 (2005)
14. Ziyan, U., Tuch, D., Westin, C.-F.: Segmentation of Thalamic Nuclei from DTI Using Spectral Clustering. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 807–814. Springer, Heidelberg (2006)
15. Weldeselassie, Y., Hamarneh, G.: Dt-mri segmentation using graph cuts. In: SPIE Medical Imaging, vol. 6512 (2007)
16. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magn. Reson. Med. 56, 411–421 (2006)
17. Wu, Y., Wang, J., Lu, H.: Real-time visual tracking via incremental covariance model update on log-euclidean riemannian manifold. In: Proc. IEEE Chinese Conference on Pattern Recognition, CCPR (2009)
18. Karcher, H.: Riemannian center of mass and mollifier smoothing. Comm. Pure Appl. Math. 30, 509–541 (1977)
19. Lurie, J.: Lecture notes on the theory of hadamard spaces (metric spaces of nonpositive curvature), `http://www.math.harvard.edu/~lurie/papers/hadamard.pdf`
20. Ballmann, W.: Manifolds of non positive curvature. In: Hirzebruch, F., Schwermer, J., Suter, S. (eds.) Arbeitstagung Bonn 1984. Lecture Notes in Mathematics, vol. 1111, pp. 261–268. Springer, Heidelberg (1985) 10.1007/BFb0084594
21. Schwartzman, A.: Random ellipsoids and false discovery rates: Statistics for diffusion tensor imaging data. PhD thesis, Stanford University (2006)
22. Söderman, O., Jönsson, B.: Restricted diffusion in cylindrical geometry. Journal of Magnetic Resonance. Series A 117(1), 94–97 (1995)
23. Barmpoutis, A., Vemuri, B.C.: A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. In: Proceedings of ISBI 2010: IEEE International Symposium on Biomedical Imaging, pp. 1385–1388 (2010)