Semi-intrinsic Mean Shift on Riemannian Manifolds

Rui Caseiro, João F. Henriques, Pedro Martins, and Jorge Batista

Institute of Systems and Robotics - University of Coimbra, Portugal
{ruicaseiro,henriques,pedromartins,batista}@isr.uc.pt

Abstract. The original mean shift algorithm [1] on Euclidean spaces (MS) was extended in [2] to operate on general Riemannian manifolds. This extension is *extrinsic* (Ext-MS) since the mode seeking is performed on the tangent spaces [3], where the underlying curvature is not fully considered (tangent spaces are only valid in a small neighborhood). In [3] was proposed an *intrinsic* mean shift designed to operate on two particular Riemannian manifolds (IntGS-MS), i.e. Grassmann and Stiefel manifolds (using manifold-dedicated density kernels). It is then natural to ask whether mean shift could be intrinsically extended to work on a large class of manifolds. We propose a novel paradigm to intrinsically reformulate the mean shift on general Riemannian manifolds. This is accomplished by embedding the Riemannian manifold into a Reproducing Kernel Hilbert Space (RKHS) by using a general and mathematically well-founded Riemannian kernel function, i.e. *heat kernel* [4]. The key issue is that when the data is implicitly mapped to the Hilbert space, the curvature of the manifold is taken into account (i.e. exploits the underlying information of the data). The inherent optimization is then performed on the embedded space. Theoretic analysis and experimental results demonstrate the promise and effectiveness of this novel paradigm.

1 Introduction

Mean shift (MS) is a popular nonparametric and unsupervised clustering algorithm which does not assume a fixed number of clusters and their shape [1]. It allows to analyze complex multimodal feature spaces from data points, which are assumed to be sampled from an unknown distribution on a Euclidean space. This conjecture is not always true which poses a challenge for some computer vision applications where data often lies in complex manifolds, namely in Riemannian manifolds, i.e. a nonlinear, curved yet smooth, metric space (e.g. motion/pose/epipolar segmentation, multi-body factorization [2, 3], object recognition/classification [5, 6], foreground segmentation [7], diffusion tensor processing [8], activity recognition, text categorization, shape analysis [9]. It is mandatory to take into account the Riemannian structure of the space in order to extract the underlying data information. From this point of view, it is natural to see several attempts in the recently literature for devise mean shift algorithms to operate on the Riemannian manifolds [2, 3, 10–12].

A. Fitzgibbon et al. (Eds.): ECCV 2012, Part I, LNCS 7572, pp. 342–355, 2012.
© Springer-Verlag Berlin Heidelberg 2012

Prior Work. Tuzel *et al.* [10] proposed a mean shift specifically to operate on Lie groups (it was applied to multiple rigid motion estimation). Subbarao *et al.* [2, 13] extended the MS to operate on general Riemannian manifolds. This extension is *extrinsic* (Ext-MS) since the mode seeking is performed on the tangent spaces [3]. The mean shift is computed as a weighted sum of tangent vectors (logarithmic map) and the resulting vector is mapped back to the manifold using the exponential map at each iteration. It was employed [2] to camera pose based segmentation, affine image motion, camera pose estimation (Lie Groups), epipolar segmentation (Essential manifold), translation viewed by a calibrated camera, multi-body factorization, chromatic noise filtering (Grassmann manifold) and diffusion tensor filtering (Tensor manifold). The underlying curvature is not fully considered since the tangent spaces only provide an approximation of the manifold in a small neighborhood. This important constraint induces a loss of accuracy in the clustering results. Following the above extensions, the medoid shift [11] and the quick shift [12] methods were proposed to cluster data on non-Euclidean spaces and applied for image categorization and segmentation [3]. Ertan *et al.* [3] derived an *intrinsic* mean shift designed to operate on two particular Riemannian manifolds (IntGS-MS), i.e. Stiefel/Grassmann (using manifold-dedicated density kernels). The Ertan's paradigm [3] cannot be generalized to other Riemannian manifolds due to the specificity of the inherent kernel density function. It is then natural to ask whether mean shift could be intrinsically extended to work on general Riemannian manifolds.

Learning problems on Riemannian manifolds are generally solved by fattening the manifold via local diffeomorphisms (tangent spaces), i.e. the manifold is locally embedded into a Euclidean space. However, embedding the manifold using those local diffeomorphisms leads to some problems. The *exponential map* is *onto* but only *one-to-one* in a neighborhood of a point. Therefore, the inverse mapping (*logarithmic map*) is uniquely defined only around a small neighborhood of that point. Those constraints are restrictive in the sense that the intrinsic structure and curvature of the space are not fully considered.

Mercer kernels [14] have been widely used to devise several well-known statistical learning techniques (e.g. support vector machines), particularly in order to convert linear classification and regression algorithms into nonlinear counterparts [15]. A key assumption of the most commonly used Mercer kernels (e.g. the radial basis function kernel) is that the data points needs to be represented as vectors in an Euclidean space [15]. This issue is often solved in an *ad hoc* manner since there is little theoretical knowledge on how the representation of data as real-valued (Euclidean) feature vectors should be carried out [15]. Recently a new paradigm emerged [16, 17]. This new paradigm suggests to use specific Grassmann kernel functions in order to embed a particular Riemannian manifold, i.e. Grassmann manifold, into a Reproducing Kernel Hilbert Space [14].

Contributions. We proposed a novel paradigm to intrinsically reformulate the mean shift on general Riemannian manifolds. This is accomplished by embedding the Riemannian manifold into a Reproducing Kernel Hilbert Space by using a general and mathematically well-founded Riemannian kernel function, i.e. *heat* *kernel* [4]. The novelty is that when the data is implicitly mapped from the manifold to the Hilbert space, the structure and curvature of the manifold is taken into account (i.e. exploits the underlying information of the data). This is the reason for the expression *semi-intrinsic* - SInt-MS). The inherent optimization is then performed on the embedded space.

As proved in the mathematics literature [4] a Riemannian manifold can be embedded into a Hilbert space using the *heat kernel* (HK) paradigm. The heat kernel on Riemannian manifolds is based on expansions involving eigenfunctions of a Laplace operator on the manifold. Considering that different Laplace operators give different heat kernels, the use of the HK in this context involves two important challenges : firstly, one must prove that the heat kernel used is a Mercer kernel [14] ; secondly, for most geometries there is not a closed form and tractable solution to compute the HK. The heat kernel is a concept that proved to be a powerful tool in physics, geometric analysis and has been a important subject of research in mathematical and physical literature. In physics, this study is motivated, in particular, by the fact that it gives a framework for investigating the quantum field theories [18]. In fact we tackle the above challenges by seeking inspiration from physics. We adopted a framework applicable for generic Riemannian manifolds proposed by Avramidi [18–20], which gives a tractable solution to compute the heat kernel and defines a Mercer kernel (to our knowledge this is the first time that the Avramidi's work is used outside the field of theoretical Physics/Mathematics and related areas).

Considering the specificity of the density-kernels used by Ertan *et al.* [3], to the best of our knowledge this is the first work that proposes an intrinsic reformulation of the mean shift for general Riemannian manifolds. Consequently, we believe this mean shift can be used to solve many different problems in the same way that the original MS has so widely been used in Euclidean cases.

2 Differential Geometry

In this section we briefly review and define some elements of differential geometry crucial to understand the proposed work. For the sake of brevity, our treatment will not be complete. For more details, please refer to [18, 20–22]. A *manifold* is a topological space locally similar to an Euclidean space. A *Riemannian manifold* is a differentiable manifold $\mathcal{M}$ endowed with a Riemannian *metric* g. Let $(\mathcal{M}, g)$ be a smooth, compact and complete n-dimensional Riemannian manifold. Let $T\mathcal{M}$ and $T^*\mathcal{M}$ be the *tangent* and *cotangent bundles* of $\mathcal{M}$. We denote a smooth vector *bundle* over a Riemannian manifold as $\mathcal{V}$, and the respective *dual* as $\mathcal{V}^*$. The $\mathrm{End}(\mathcal{V}) \cong \mathcal{V} \otimes \mathcal{V}^*$ is the bundle of all smooth *endomorphisms* of $\mathcal{V}$ [18, 20]. The space of smooth real-valued *functions* on $\mathcal{M}$ is denoted as $C^\infty(\mathcal{M})$. The $C^\infty(\mathcal{M}, \mathcal{V})$ and $C^\infty(\mathcal{M}, \mathrm{End}(\mathcal{V}))$ are the spaces of all smooth *sections* of the bundles $\mathcal{V}$ and $\mathrm{End}(\mathcal{V})$, respectively [18, 20]. The vector bundle $\mathcal{V}$ is assumed to be equipped with a Hermitian metric, identifing the dual bundle $\mathcal{V}^*$ with $\mathcal{V}$, and defines L^2 inner product. The Hilbert space $L^2(\mathcal{M}, \mathcal{V})$ of square integrable sections is defined to be the completion of $C^\infty(\mathcal{M}, \mathcal{V})$ in this norm [18, 20]. Let ∇^{LC} be the *canonical connection* (*Levi-Civita*) on the tangent bundle $T\mathcal{M}$. Let, further, $\nabla^{\mathcal{V}}: C^\infty(\mathcal{M},\mathcal{V}) \to C^\infty(\mathcal{M}, T^*\mathcal{M}\otimes\mathcal{V})$ be a *connection*, on the vector bundle $\mathcal{V}$. Using ∇^{LC} together with $\nabla^{\mathcal{V}}$, result in connections on all bundles in the tensor algebra over $\mathcal{V}, \mathcal{V}^*, T\mathcal{M}$ and $T^*\mathcal{M}$ [18, 20]. Let $\nabla^*: C^\infty(T^*\mathcal{M}) \to C^\infty(\mathcal{M})$ be the *formal adjoint* of $\nabla: C^\infty(\mathcal{M}) \to C^\infty(T^*\mathcal{M})$, and let $\mathcal{Q} \in C^\infty(\mathcal{M}, \mathrm{End}(\mathcal{V}))$ be a smooth Hermitian *section* of the endomorphism bundle $\mathrm{End}(\mathcal{V})$ [18, 20]. Let x^μ, $(\mu = 1, 2, ..., n)$, be a system of local coordinates. Greek indices, $\mu, \nu, ...$, label the components with respect to local coordinate frames $x = (x^\mu)$ on $\mathcal{M}$. Let ∂_μ be a local coordinate basis (frames) for the tangent space $T\mathcal{M}$ at some point $\in \mathcal{M}$ and dx^μ be dual basis for the cotangent space $T^*\mathcal{M}$. Let $g_{\mu\nu} = (\partial_\mu, \partial_\nu)$ and $g^{\mu\nu} = (dx^\mu, dx^\nu)$ be the *metric* on the tangent/cotangent bundle [18, 20].

3 Kernel Mean Shift

In this section we describe a kernel-based mean shift technique proposed by Tuzel *et al.* [14, 23] which will be the basis for our semi-intrinsic mean shift on general Riemannian manifolds (SInt-MS). In order to study the Riemannian structure of the space, the data is implicitly mapped to an enlarged feature space $\mathcal{H}$ (Hilbert space) by using an appropriate Riemannian kernel function (a Riemannian manifold can be embedded into a Hilbert space using the heat kernel [4]). Consider $\mathcal{M}$ as the input space and the N data points given by $\mathbf{Z}_i \in \mathcal{M}$, $i = 1, ..., N$. Let $K: \mathcal{M}\times\mathcal{M} \to \Re$ be a *positive definite kernel function* : $K(\mathbf{Z}, \mathbf{Z}') = \phi(\mathbf{Z})^T\phi(\mathbf{Z}')$ for all $\mathbf{Z}, \mathbf{Z}' \in \mathcal{M}$. The input space $\mathcal{M}$ is projected into the d-dimensional feature space $\mathcal{H}$ by the mapping function ϕ (e.g. $\mathbf{z} = \phi(\mathbf{Z})$ with $\mathbf{z} \in \mathcal{H}$). Let's now present the derivation of the mean shift on the space $\mathcal{H}$ in function of the mapping ϕ [23]. In the feature space $\mathcal{H}$, the density estimator at some point sample $\mathbf{z} \in \mathcal{H}$ is defined as [23]

$$f_{\mathcal{H}}(\mathbf{z}) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_i^d}k\left(\|\frac{\mathbf{z}-\phi(\mathbf{Z}_i)}{h_i}\|^2\right) \tag{1}$$

where h_i is the bandwidth and $k(\cdot)$ is a multivariate normal profile. The stationary points of the function $f_{\mathcal{H}}$ are obtained computing the gradient of Eq. 1 with respect to the mapping ϕ [23], i.e. they satisfy the condition

$$\frac{2}{N}\sum_{i=1}^{N}\frac{1}{h_i^{d+2}}(\phi(\mathbf{Z_i})-\mathbf{z})g\left(\|\frac{\mathbf{z}-\phi(\mathbf{Z}_i)}{h_i}\|^2\right) = 0 \tag{2}$$

where $g(\cdot) = -k'(\cdot)$. This problem is solved iteratively [23] similarly to the standard mean shift [1]

$$\bar{\mathbf{z}} = \left(\sum_{i=1}^{N}\frac{\phi(\mathbf{Z}_i)}{h_i^{d+2}}g\left(\|\frac{\mathbf{z}-\phi(\mathbf{Z}_i)}{h_i}\|^2\right)\right)\left(\sum_{i=1}^{N}\frac{1}{h_i^{d+2}}g\left(\|\frac{\mathbf{z}-\phi(\mathbf{Z}_i)}{h_i}\|^2\right)\right)^{-1} \tag{3}$$

Let the matrix $\mathbf{K} = \mathbf{\Phi}^T\mathbf{\Phi}$ be the $N\times N$ Kernel (Gram) matrix, where $\mathbf{\Phi} = [\phi(\mathbf{Z}_1)\ \ \phi(\mathbf{Z}_2)\ \ ...\ \ \phi(\mathbf{Z}_N)]$ is the $d\times N$ matrix of the feature points. Given the

Input: 1 - $\mathbf{K}$ computed using Heat Kernel on Riemannian manifolds (Section 4)
2 - Bandwidth selection parameter k

Calculate the bandwidths h_i as the k^{th} smallest distance from the point using Eq. 4 with $\mathbf{K}$ and $d = \mathrm{rank}(\mathbf{K})$

for *All data points* $i = 1, ..., N$ **do**

a) - Let $\alpha_{\mathbf{z}_i} = \mathbf{e}_i$

b) - Repeat until convergence (with $D' = \alpha_{\mathbf{z}_i}^T \mathbf{K} \alpha_{\mathbf{z}_i} + \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j - 2\alpha_{\mathbf{z}_i}^T \mathbf{K} \mathbf{e}_j$)

$$\bar{\alpha}_i = \left(\sum_{j=1}^{N} \frac{\mathbf{e}_j}{h_j^{d+2}} g\left(\frac{D'}{h_j^2}\right) \right) \left(\sum_{j=1}^{N} \frac{1}{h_j^{d+2}} g\left(\frac{D'}{h_j^2}\right) \right)^{-1} \tag{5}$$

c) - Group the points $\bar{\alpha}_{\mathbf{z}_i}$ and $\bar{\alpha}_{\mathbf{z}_j}$, $i, j = 1, ..., N$ satisfying $\bar{\alpha}_i^T \mathbf{K} \bar{\alpha}_i + \bar{\alpha}_j^T \mathbf{K} \bar{\alpha}_j - 2\bar{\alpha}_i^T \mathbf{K} \bar{\alpha}_j = 0$.

Algorithm 1. Semi-Intrinsic Meanshift on Riemannian Manifolds. The bandwidth h_i is computed as the k^{th} smallest distance from the point i to all the N data points on the feature space, where k is computed as a fraction of N [23].

feature matrix $\mathbf{\Phi}$, the solution $\bar{\mathbf{z}}$ at each iteration of the mean shift algorithm (Eq. 3) lies in the column space of $\mathbf{\Phi}$ [23]. Considering the subspace spanned by the columns of the matrix $\mathbf{\Phi}$, a point $\mathbf{z}$ is defined as $\mathbf{z} = \mathbf{\Phi}\alpha_{\mathbf{z}}$ ($\alpha_{\mathbf{z}}$ is an N-dimensional vector of weights) [23]. All the computation involving data points can be expressed in terms of inner products, e.g. the distances can be formulated in the form of the inner product of the points and $\alpha_{\mathbf{z}}$ can be iteratively updated. In this case the distance $D = \|\mathbf{z} - \mathbf{z}'\|^2$ between $\mathbf{z}$ and $\mathbf{z}'$ is given by [23]

$$D = \|\mathbf{\Phi}\alpha_{\mathbf{z}} - \mathbf{\Phi}\alpha_{\mathbf{z}'}\|^2 = \alpha_{\mathbf{z}}^T \mathbf{K} \alpha_{\mathbf{z}} + \alpha_{\mathbf{z}'}^T \mathbf{K} \alpha_{\mathbf{z}'} - 2\alpha_{\mathbf{z}}^T \mathbf{K} \alpha_{\mathbf{z}'} \tag{4}$$

We can therefore use the well-known *kernel trick* [14], i.e. it is possible to compute the distances between $\mathbf{z} = \phi(\mathbf{Z})$ and $\mathbf{z}' = \phi(\mathbf{Z}')$ in the mapped space without the explicit knowledge of the mapping of $\mathbf{Z}$ and $\mathbf{Z}'$ to $\phi(\mathbf{Z})$ and $\phi(\mathbf{Z}')$ respectively. The mappings are replaced by the dot product $K(\mathbf{Z}, \mathbf{Z}') = \phi(\mathbf{Z})^T \phi(\mathbf{Z}')$. The input of the algorithm is the kernel matrix $\mathbf{K}$ of inner products. It can be demonstrated that a positive semi-definite matrix $\mathbf{K}$ can be seen as a kernel matrix [14, 24]. Consider $\phi(\mathbf{Z}_i) = \mathbf{\Phi}\mathbf{e}_i$, where $\mathbf{e}_i$ represent the i-th canonical basis for $\Re^N$. Substituting Eq. 4 into Eq. 3 the solution $\bar{\alpha}_{\mathbf{z}}$ is obtained as follows [23].

$$\bar{\alpha}_{\mathbf{z}} = \left(\sum_{i=1}^{N} \frac{\mathbf{e}_i}{h_i^{d+2}} g\left(\frac{D'}{h_i^2}\right) \right) \left(\sum_{i=1}^{N} \frac{1}{h_i^{d+2}} g\left(\frac{D'}{h_i^2}\right) \right)^{-1} \tag{6}$$

where $D' = \alpha_{\mathbf{z}}^T \mathbf{K} \alpha_{\mathbf{z}} + \mathbf{e}_i^T \mathbf{K} \mathbf{e}_i - 2\alpha_{\mathbf{z}}^T \mathbf{K} \mathbf{e}_i$. Since any positive semi-definite matrix $\mathbf{K}$ is a kernel for some feature space [25], the original proof of convergence in the Euclidean case [1] is sufficient to prove the convergence of this kernel-based mean shift algorithm (as argued in [23]). The vectors of weights at the initial of the iterative procedure are defined as $\alpha_{\mathbf{z}_i} = \mathbf{e}_i$, which results in $\mathbf{z}_i = \mathbf{\Phi}\alpha_{\mathbf{z}_i} = \phi(\mathbf{Z}_i), i = 1, ..., N$. The final mode is represented as $\mathbf{\Phi}\bar{\alpha}_{\mathbf{z}_i}$. For more details about this kernel-based mean shift, please refer to [23]. The matrix $\mathbf{K}$ is computed using the heat kernel on Riemannian manifolds, which will be presented in Section 4.

4 Mercer Kernel on Riemannian Manifolds

Our goal is to define a class of positive definite kernels (Mercer kernels [14]) on general Riemannian manifolds ($\mathcal{M}$), and thus both prove that the problem presented in Section 3 is well-posed and define tools for solving it. The functions must be devised in order to take into account the intrinsic geometry of the manifold ($\mathcal{M}$). These objectives will be accomplished by considering expansions involving eigenfunctions of the Laplace operator on $\mathcal{M}$ (i.e. heat kernel [4]).

A continuous, complex-valued kernel $K(\cdot\,,\cdot\,) \in C(\mathcal{M}\times\mathcal{M})$ is a *positive definite* function on $\mathcal{M}$ if $\bar{K}(\mathbf{Z}', \mathbf{Z}) = K(\mathbf{Z}, \mathbf{Z}')$ with $\mathbf{Z}, \mathbf{Z}' \in \mathcal{M}$ and if for all finite set of points $\mathcal{C} = \{\mathbf{Z}_1, ..., \mathbf{Z}_N\} \in \mathcal{M}$, the self-adjoint, $N \times N$ matrix with entries $K(\mathbf{Z}_j, \mathbf{Z}_k)$ is *positive semi-definite* [26, 27]. We will be particulary interested in C^∞ positive definite kernels. This positivity can be seen in terms of distributions. The concept of a distribution can be represented on $\mathcal{M}$; typically it is a linear functional defined on $C^\infty(\mathcal{M})$. Let $\mathcal{D}'(\mathcal{M})$ be the set of all distributions on $\mathcal{M}$ [26, 27].

If $\delta_{\mathbf{Z}}$ is the Dirac delta function located at $\mathbf{Z}$ and if $u = (\sum_{j=1}^{N} c_j \delta_{\mathbf{Z}_j}) \in \mathcal{D}'(\mathcal{M})$, then the matrix $\mathbf{K} = [K(\mathbf{Z}_j, \mathbf{Z}_k)]$ being positive semi-definite is equivalent to the quadratic form $(\bar{u} \otimes u, K) \geq 0$ for arbitrary c_j, where $\otimes$ corresponds to the tensor product between distributions (Theorem 2.1 - [26]).

Theorem 1. *Let $K(\mathbf{Z}, \mathbf{Z}') \in C^\infty(\mathcal{M} \times \mathcal{M})$ have the eigenfunction expansion. $K(\mathbf{Z}, \mathbf{Z}') = \sum_{l\in\mathcal{A}} = a_l \varphi_l(\mathbf{Z})\bar{\varphi}_l(\mathbf{Z}')$ where $\mathcal{A}$ is some countable index set, and the φ_l are the eigen-functions of the Laplace operator on $\mathcal{M}$. Then K is positive definite on $\mathcal{M}$ if and only if $a_l \geq 0$ for all $l \in \mathcal{A}$. Moreover, K is strictly positive definite if and only if $a_l > 0$ for all $l \in \mathcal{A}$ [26].*

Corollary 1. *For some countable index set $\mathcal{A}$, let $\{\varphi_l\}_{l\in\mathcal{A}}$ be an orthogonal basis for $L^2(\mathcal{M}, g)$ comprising eigenfunctions of the Laplace operator on $\mathcal{M}$. In addition, let $K \in C^\infty(\mathcal{M} \times \mathcal{M})$ be such that $(\bar{\varphi}_l \otimes \varphi_r, K) = a_l\delta_{lr}$. Then, K is positive definite on $\mathcal{M}$ if and only if $a_l \geq 0$ for all $l \in \mathcal{A}$. Moreover, K is strictly positive definite if and only if $a_l > 0$ for all $l \in \mathcal{A}$ [26].*

On a Riemannian manifold $\mathcal{M}$, a Laplace type operator $\mathcal{L}$: $C^\infty(\mathcal{M}, \mathcal{V}) \to C^\infty(\mathcal{M}, \mathcal{V})$ can be defined as a second-order partial differential operator of the form [18, 20]

$$\mathcal{L} = \nabla^*\nabla + \mathcal{Q} = -g^{\mu\nu}\nabla_\mu\nabla_\nu + \mathcal{Q} \tag{7}$$

where $\nabla^*\nabla$ is the generalized Laplacian and $\mathcal{Q}$ is an endomorphism (Section 2). The Laplace operator $\mathcal{L}$ is *elliptic* and has a *positive leading symbol.* $\mathcal{L}$ is *symmetric* with respect to the natural L^2 inner product and is *self-adjoint*, i.e. its closure is self-adjoint, meaning that there is a unique self-adjoint extension $\bar{\mathcal{L}}$ of the operator $\mathcal{L}$ [18, 20].

Gilkey [28] defined an important theorem about the spectrum of a self-adjoint, elliptic differential operator $\mathcal{L}$ with a positive definite principal symbol, operating over a compact manifold $\mathcal{M}$ (on smooth sections of a vector bundle $\mathcal{V}$) , $\mathcal{L}\colon C^\infty(\mathcal{M},\mathcal{V}) \to C^\infty(\mathcal{M},\mathcal{V})$, [18, 20] : 1 - the spectrum of the operator $\mathcal{L}$ is constituted by a sequence of discrete real nondecreasing eigenvalues $\{\lambda_l\}_{l=1}^{\infty}$; 2 - the eigenvectors $\{\varphi_l\}_{l=1}^{\infty}$ form a complete orthonormal basis in $L^2(\mathcal{M},\mathcal{V})$ (are smooth sections of the vector bundle $\mathcal{V}$) ; and the eigenspaces are finite-dimensional ; 3 - as $l \to \infty$ the eigenvalues increase as $\lambda_l \sim \beta l^2$ as $l \to \infty$, with some positive constant β. Hence, the operator $U_{\mathcal{L}}(t) = \exp(t\mathcal{L})$ for $t > 0$ is well defined as a *bounded* operator (form a semi-group of *bounded* operators) on the Hilbert space $L^2(\mathcal{M},\mathcal{V})$ of square integrable sections of the bundle $\mathcal{V}$ [18, 20].

The kernel $U_{\mathcal{L}}(t|\mathbf{Z},\mathbf{Z}')$ of that operator satisfies the heat equation, $(\partial_t + \mathcal{L})U_{\mathcal{L}}(t|\mathbf{Z},\mathbf{Z}') = 0$ and is called the *heat kernel*, being defined as follows [18, 20]

$$K_t(t|\mathbf{Z},\mathbf{Z}') = U_{\mathcal{L}}(t|\mathbf{Z},\mathbf{Z}') = \sum_{l=1}^{\infty} e^{-t\lambda_l}\varphi_l(\mathbf{Z}) \otimes \bar{\varphi}_l(\mathbf{Z}') \tag{8}$$

The heat kernel can be seen as a smooth section of the *external tensor product* of the vector bundles $(\mathcal{V} \boxtimes \mathcal{V}^*)$ over the tensor product manifold $\mathcal{M} \times \mathcal{M}$: $K_t(t|\mathbf{Z},\mathbf{Z}') \in C^\infty(\Re_+ \times \mathcal{M} \times \mathcal{M}, \mathcal{V} \boxtimes \mathcal{V}^*)$, i.e. is an endomorphism from the fiber of $\mathcal{V}$ over $\mathbf{Z}'$ to the fiber of $\mathcal{V}$ over $\mathbf{Z}$ [18, 20, 21].

Since all of the coefficients of $K_t(t|\mathbf{Z},\mathbf{Z}')$ given by Eq. 8 (derived from the operator $\mathcal{L}$) are positive, Theorem 1 and Corollary 1 yield the following result :

Corollary 2. *The heat kernel $K_t(\mathbf{Z},\mathbf{Z}') \in C^\infty(\mathcal{M} \times \mathcal{M})$ is a strictly positive definite kernel on $\mathcal{M}$.*

From the Corollary 2 we conclude that the heat kernel can define a Mercer Kernel and therefore it is a suitable kernel for represent the similarity between data points $\in \mathcal{M}$, while respecting the intrinsic geometry of the Riemannian space. In the current literature, there is not a closed form and tractable solution to estimate the heat kernel on general Riemann manifolds. However, the asymptotic solution of the heat kernel in the short time is a well-studied problem in theoretical physics and mathematics. In mathematics - corresponding to the determination of the spectrum of the Laplacian - can give topological information and in physics, it gives the solution, on a fixed spacetime background, of the Euclidean Schrödinger equation. Avramidi [18–21] proposed an asymptotic expansion of the heat kernel for second-order elliptic partial differential operators acting on sections of vector bundles over a Riemannian manifold, defined as

$$K_t(t|\mathbf{Z},\mathbf{Z}') = (4\pi t)^{-\frac{n}{2}} \Delta(\mathbf{Z},\mathbf{Z}')^{\frac{1}{2}} \exp\left(-\frac{W(\mathbf{Z},\mathbf{Z}')}{2t}\right) \mathcal{P}(\mathbf{Z},\mathbf{Z}') \Omega(t|\mathbf{Z},\mathbf{Z}') \tag{9}$$

We consider $t = \sigma^2$ by analogy with the typical variance σ^2 of the standard gaussian kernel in Euclidean spaces. Let $W = W(\mathbf{Z},\mathbf{Z}')$ be the geodetic interval, also called *world function*, defined as one half the square of the length of the geodesic (geodesic distance D_g) connecting the points $\mathbf{Z}$ and $\mathbf{Z}'$, i.e. $W(\mathbf{Z},\mathbf{Z}') =$

$1/2D_g((\mathbf{Z},\mathbf{Z}')$. The first derivatives of this function with respect to $\mathbf{Z}$ and $\mathbf{Z}'$ define tangent vector fields to the geodesic at the points $\mathbf{Z}$ and $\mathbf{Z}'$ [18–21]

$$u^{\mu} = u^{\mu}(\mathbf{Z},\mathbf{Z}') = g^{\mu\nu}\nabla_{\nu}W \qquad u^{\mu'} = u^{\mu'}(\mathbf{Z},\mathbf{Z}') = g^{\mu'\nu'}\nabla'_{\nu'}W \tag{10}$$

and the determinant of the mixed second derivatives defines the so-called *Van Vleck-Morette determinant* [18–21]

$$\Delta = \Delta(\mathbf{Z},\mathbf{Z}') = |g(\mathbf{Z})|^{-\frac{1}{2}}|g(\mathbf{Z}')|^{-\frac{1}{2}}det(-\nabla_{\mu}\nabla'_{\nu'}W) \tag{11}$$

Let, $\mathcal{P} = \mathcal{P}(\mathbf{Z},\mathbf{Z}')$ denote the *parallel transport operator* of sections of the vector bundle $\mathcal{V}$ along the geodesic from the point $\mathbf{Z}'$ to the point $\mathbf{Z}$ [18–21]. It is an endomorphism from the fiber of $\mathcal{V}$ over $\mathbf{Z}'$ to the fiber of $\mathcal{V}$ over $\mathbf{Z}$ (or a section of the external tensor product $\mathcal{V}\boxtimes\mathcal{V}^*$ over $\mathcal{M}\times\mathcal{M}$). The function $\Omega(t) = \Omega(t|\mathbf{Z},\mathbf{Z}')$, called the *transport function*, is a section of the endomorphism bundle $\mathrm{End}(\mathcal{V})$ over the point $\mathbf{Z}'$ and satisfies the transport equation [18–21]

$$\left(\partial_t + \frac{1}{t}D + L\right)\Omega(t|\mathbf{Z},\mathbf{Z}') = 0 \qquad \Omega(t|\mathbf{Z},\mathbf{Z}') \sim \sum_{k=0}^{\infty}\frac{(-t)^k}{k!}a_k(\mathbf{Z},\mathbf{Z}') \tag{12}$$

with the intial condition $\Omega(0|\mathbf{Z},\mathbf{Z}') = I$, where $\mathbf{I}$ is the identity endomorphism of the bundle $\mathcal{V}$ over $\mathbf{Z}'$, D and L are operators defined as [18–21]

$$D = u^{\mu}\nabla_{\mu} \qquad L = \mathcal{P}^{-1}\Delta^{-1/2}F\Delta^{1/2}\mathcal{P} \tag{13}$$

where D is the radial vector field, i.e. operator of differentiation along the geodesic and L is a second-order differential operator [18–21].

The transport function $\Omega(t)$ can be defined using an asymptotic expansion in terms of the coefficients $a_k = a_k(\mathbf{Z},\mathbf{Z}')$, called *Hadamard-Minakshisundaram-DeWitt-Seeley coefficients* (HMDS) - Eq. 12 [18–21]. Taking into account that the calculation of the HMDS coefficients in the general case offers a complex theorical and technical problem, we will use a trade-off solution by considering only the two lowest order terms, i.e. $k = 0, 1$ (which is in pratical a very good solution). The so-called *off-diagonal* HMDS coefficients (i.e. $a_k = a_k(\mathbf{Z},\mathbf{Z}')$, with $\mathbf{Z} \neq \mathbf{Z}'$) are determined by a differential recursive system given by [18–21]

$$\left(1 + \frac{1}{k}D\right)a_k = La_{k-1} \qquad a_k = D_k^{-1}LD_{k-1}^{-1}L...D_1^{-1}LI \tag{14}$$

with $a_o = I$ and $D_k = 1 + \frac{1}{k}D$. A close solution (explicit solution) to compute the so-called *diagonal* HMDS coefficients (i.e. $a_k^{diag} = a_k(\mathbf{Z},\mathbf{Z}')$, with $\mathbf{Z} = \mathbf{Z}'$) has been a important subject of research in recent years in mathematical and physical literature [18–21]. Recently a_k^{diag} has been computed up to order $k = 8$. These formulas become exponentially more complicated as k increases. For example, the formula for a_6^{diag} has 46 terms. However, the explicit formula for the first two coefficients ($k = 0, 1$) are fairly easy to compute and are given by $a_0^{diag} = I$ and

$a_1^{diag} = Q - 1/6R$ where R is the *scalar curvature*. [18–21]. This paradigm is applicable to general Riemannian manifolds $\mathcal{M}$ and it is very algorithmic. Considering that there are a number of usual algebraic operations on symmetric tensors that are easily pre-programmed, this technique is appropriate to automatic computation, i.e. symbolic computations easily lead to the components of the heat kernel.

5 Experimental Results

The clustering accuracy of our method is evaluated on synthetic data (Section 5.1) as well as on real data (Section 5.2). Specifically, we compare the proposed *semi-intrinsic* mean shift (SInt-MS) with the *extrinsic* counterpart (Ext-MS) [2, 13]. The experimental evaluation, serve mainly as a proof of concept, which is reasonable given the novelty of the method. We present results for the following manifolds : Lie Group (Special Orthogonal Group - SO_3), Tensor manifold (Symmetric Positive-Definite Matrices - $\mathcal{S}^+$), Grassmann ($\mathcal{G}$) and Stiefel ($\mathcal{V}$) manifolds. However, our paradigm is general and applicable to a large number of different Riemannian manifolds and applications. We remark that the goal is not to compare our *semi-intrinsic* meanshift (SInt-MS) with the *intrinsic* version (Int-MS) proposed by Ertan *et al.* [3]. The Ertan's method is not generalizable due to the specificity of the inherent kernel density function. The proof of concept (extrinsic general vs semi-extrinsic general) is not affected since unlike Ertan [3] we have looked to the Grassmman and Stiefel manifolds as general Riemannian manifolds and not as Riemannian manifolds with specific kernel density functions. Let $\mathcal{C}$ be the number of classes and $\mathcal{P}$ be the number of points per class.

5.1 Simulations on Synthetic Data

We conduct our synthetic experiments using the following data configurations for clustering : (**A** = 4 classes |100 points/class) ; (**B** = 4 classes |200 points/class); (**C** = 8 classes |100 points/class); (**D** = 8 classes |200 points/class). Table 1 shows the clustering rates of the Ext-MS and the SInt-MS on synthetic data.

Grassmann ($\mathcal{G}_{k,m-k}$) and Stiefel ($\mathcal{V}_{k,m}$) Manifolds
Regarding the synthetic tests, we generate pseudo-random matrices on the Grassmann ($\mathcal{G}_{k,m-k}$) and Stiefel ($\mathcal{V}_{k,m}$) manifolds using the procedures presented in [3, 29]. It is possible to represent a orthogonal matrix $\mathcal{S} \in \mathcal{O}(m)$ as a product of $(0.5m^2 - 0.5m)$ orthogonal matrices of the form $R_m^{\nu}(\theta)$, for $1 \leq \nu \leq m-1$ (please refer to [3]). The orthogonal matrix $S \in \mathcal{O}(m)$ can be defined as $S = \prod_{\nu=1}^{m-1} S_m^{\nu}$, with $S_m^{\nu} = \prod_{j=\nu}^{m-1} R_m^j(\theta_{\nu,j})$. For each one of the $\mathcal{C}$ classes, we generated $(0.5m^2 - 0.5m)$ angles $\{\theta_{\nu,j}\}$. Each angle $\{\theta_{\nu,j}\}$ is randomly drawn from one of $(0.5m^2 - 0.5m)$ bins in $[0, \pi]$ [3]. Then the angles are corrupted with random noise. Given the set of matrices $\{S\}$ for each class, we form the matrices $\mathbf{X} \in \mathcal{V}_{k,m}$ by taking the first k orthonormal columns of each S. Regarding the Grassmann manifold, the matrices are computed as $\mathbf{P} = \mathbf{X}\mathbf{X}^T \in \mathcal{G}_{k,m-k}$ [3].

Table 1. Clustering rates (%) of extrinsic (Ext-MS) and semi-intrinsic (SInt-MS) meanshift on Riemannian manifolds in the case of simulations on synthetic data

$\mathcal{G}$		**Performance**		$\mathcal{V}$		**Performance**		SO_3		**Performance**		$\mathcal{S}^+$		**Performance**	
m	k	Ext-Ms	SInt-Ms	m	k	Ext-Ms	SInt-Ms	r	-	Ext-Ms	SInt-Ms	d	d	Ext-Ms	SInt-Ms
A $\mapsto$ **Classes** = 4 | **Points per Class** = 100															
3	1	51.23	63.42	3	1	64.83	84.05	1	-	55.53	78.23	3	3	57.68	85.64
5	3	67.41	78.50	3	3	71.15	86.25	2	-	66.78	86.86	4	4	66.46	91.04
5	4	46.80	61.91	5	3	77.10	91.76	3	-	59.45	81.35	5	5	65.71	92.56
10	4	53.27	65.34	10	1	74.91	90.35	4	-	61.59	82.34	6	6	65.75	90.84
20	4	50.05	64.08	50	1	70.03	87.50	5	-	57.50	80.29	7	7	61.28	88.39
B $\mapsto$ **Classes** = 4 | **Points per Class** = 200															
3	1	49.51	72.25	3	1	60.25	84.24	1	-	52.38	82.74	3	3	53.81	87.95
5	3	64.28	82.39	3	3	66.12	87.05	2	-	62.70	89.20	4	4	61.91	92.63
5	4	45.23	69.72	5	3	69.86	91.41	3	-	55.04	85.06	5	5	59.72	92.73
10	4	51.08	73.01	10	1	67.05	90.25	4	-	56.41	86.13	6	6	59.30	92.68
20	4	49.16	72.89	50	1	65.41	88.95	5	-	54.78	85.42	7	7	57.59	91.68
C $\mapsto$ **Classes** = 8 | **Points per Class** = 100															
3	1	46.85	69.85	3	1	57.85	80.53	1	-	49.85	79.69	3	3	51.35	84.61
5	3	60.05	78.05	3	3	60.25	86.01	2	-	57.92	86.53	4	4	56.45	90.47
5	4	42.91	67.91	5	3	65.12	84.25	3	-	51.31	80.57	5	5	55.81	86.91
10	4	47.35	69.35	10	1	63.51	82.42	4	-	52.93	80.34	6	6	55.72	85.90
20	4	45.19	69.19	50	1	60.39	83.05	5	-	50.29	80.62	7	7	52.84	86.35
D $\mapsto$ **Classes** = 8 | **Points per Class** = 200															
3	1	44.05	71.05	3	1	53.83	80.05	1	-	46.44	80.05	3	3	47.63	84.51
5	3	57.35	78.35	3	3	55.15	83.10	2	-	53.75	85.25	4	4	51.95	88.60
5	4	41.03	70.03	5	3	58.10	81.31	3	-	47.06	80.17	5	5	50.08	85.24
10	4	45.15	67.15	10	1	59.91	85.65	4	-	50.03	80.93	6	6	52.47	87.95
20	4	44.10	70.10	50	1	54.03	79.19	5	-	46.82	79.14	7	7	47.79	83.56
Average Improvement : Δ = (SInt-MS) - (Ext-MS)															
$\Delta\mathcal{G}$ = + 20 %				$\Delta\mathcal{V}$ = + 22 %				ΔSO_3 = + 28 %				$\Delta\mathcal{S}^+$ = + 32 %			

Lie Group (SO_3)

The Lie Group used in the synthetic tests was the special orthogonal group SO_3, which is the group of rotations in 3D. The set of 3×3 skew-symmetric matrices $\mathbf{\Omega}$ forms the Lie algebra so_3 (tangent space to the identity element of the group). The skew-symmetric matrices can be defined in vector form $\omega = (\omega_x, \omega_y, \omega_z)$ [10]. From a geometrical perspective $\mathbf{\Omega}$ can be seen as a rotation of an angle $||\omega||$ around the axis $\omega/||\omega||$. Let $x \in so(3)$ be an element on the Lie algebra and $\mathbf{X} = \exp(x) \in SO(3)$ be its exponential mapping to the manifold and consider that the tangent space is a vectorial space. Using these facts we generated the synthetic points of the several classes directly in the Lie algebra. We choose $\mathcal{C}$ points in the Lie algebra defined in the vectorial form as $\omega^i = (\omega_x^i, \omega_y^i, \omega_z^i)$ for $i = 1, ..., \mathcal{C}$. Each of those $\mathcal{C}$ points corresponds to the ground-truth center of a class/cluster. Let r be the radius of a 2-sphere centered in each of the $\mathcal{C}$ clusters. We randomly drawn $\mathcal{P}$ points inside of each one of the $\mathcal{C}$ spheres. We then corrupted all the $\mathcal{CP}$ points with random noise. Finally, we map back to the manifold all the $\mathcal{CP}$ points using the exponential map and we obtain $\mathcal{CP}$ rotation matrices distributed in $\mathcal{C}$ known clusters/classes.

Tensor Manifold ($\mathcal{S}_d^+$)

The ($\mathcal{S}_d^+$) manifold corresponds to the Riemannian manifold of the $d \times d$ symmetric positive-definite matrices. We generated the synthetic points directly in a tangent space of the manifold $\mathcal{S}_d^+$. There exist two well-founded Riemannian metrics for ($\mathcal{S}_d^+$) e.g. Affine-Invariant and Log-Euclidean. When the ($\mathcal{S}_d^+$) manifold is endowed

Table 2. Clustering rates (%) of extrinsic (Ext-MS) and semi-intrinsic (SInt-MS) meanshift on Riemannian manifolds on object categorization for selected objects on the data set ETH-80[30]

$\mathcal{G}$				**Performance**		$\mathcal{V}$				**Performance**		$\mathcal{S}^+$		**Performance**	
Scales	Bins	m	k	Ext-Ms	SInt-Ms	Scales	Bins	m	k	Ext-Ms	SInt-Ms	d	d	Ext-Ms	SInt-Ms
E $\mapsto$ **Classes** = 3 | **Points per Class** = 250															
2	32	32	4	68.31	88.46	2	32	128	1	69.83	91.05	5	5	63.75	85.44
2	64	64	4	76.95	95.40	2	64	256	1	76.15	93.22	7	7	62.02	84.22
3	32	32	6	72.28	91.26	3	32	192	1	81.10	97.55	9	9	58.50	82.45
F $\mapsto$ **Classes** = 3 | **Points per Class** = 300															
2	32	32	4	63.98	91.55	2	32	128	1	65.25	91.24	5	5	55.12	91.84
2	64	64	4	73.55	97.53	2	64	256	1	71.12	94.05	7	7	53.29	89.50
3	32	32	6	67.35	94.68	3	32	192	1	73.86	91.41	9	9	58.16	86.11
G $\mapsto$ **Classes** = 4 | **Points per Class** = 250															
2	32	32	4	61.47	88.42	2	32	128	1	62.85	87.53	5	5	52.72	83.65
2	64	64	4	68.60	95.09	2	64	256	1	65.25	93.01	7	7	54.97	80.21
3	32	32	6	63.84	89.62	3	32	192	1	70.12	91.25	9	9	52.03	81.42
H $\mapsto$ **Classes** = 4 | **Points per Class** = 300															
2	32	32	4	57.98	88.67	2	32	128	1	58.83	87.05	5	5	48.61	82.18
2	64	64	4	64.80	93.58	2	64	256	1	60.15	90.10	7	7	49.24	86.24
3	32	32	6	61.08	88.63	3	32	192	1	63.10	88.31	9	9	45.28	81.66
Average Improvement : Δ = (SInt-MS) - (Ext-MS)															
$\Delta\mathcal{G}$ = + 25 %						$\Delta\mathcal{V}$ = + 23 %						$\Delta\mathcal{S}^+$ = + 30 %			

with the Log-Euclidean metric turns into a space with a null curvature, i.e. we can map all the points $\in \mathcal{S}_d^+$ for the tangent space centered at $\mathbf{I}_d$ (identity matrix) using the ordinary matrix logarithm (log) and map back to the manifold with the matrix exponential (exp). The tangent space corresponds to the space of $d \times d$ symmetric matrices $\mathcal{S}_d$ with $m = d(d+1)/2$ independent elements. We choose $\mathcal{C}$ points in $\Re^m$. Each of those $\mathcal{C}$ points corresponds to the ground-truth center of a class/cluster. Let r be the radius of a sphere S^{m-1} centered in each of the $\mathcal{C}$ clusters. We randomly drawn $\mathcal{P}$ points inside of each one of the $\mathcal{C}$ spheres. We then corrupted all the $\mathcal{CP}$ points with random noise. Finally, we construct the corresponding $\mathcal{C}$ symmetric matrices $d \times d$ and map back to the manifold all the $\mathcal{CP}$ points using the matrix exponential (exp). We obtain $\mathcal{CP}$ symmetric positive-definite matrices (tensors) distributed in $\mathcal{C}$ known clusters/classes. We used Log-Euclidean metric to generate the synthetic tensors, but we endowed the ($\mathcal{S}_d^+$) manifold with the Affine-Invariant metric for the clustering task, because with the Log-Euclidean metric the tensor space turns into a null curvature space.

5.2 Real Data

In our real experiments we select two well-known data sets, ETH-80 [30] and CIFAR10 [31], typically used in visual object categorization tasks (grouping similar objects of the same class). We conduct our real experiments using the following data configurations for clustering : (**E** = 3 classes |250 points/class); (**F** = 3 classes |300 points/class) ; (**G** = 4 classes |250 points/class) ; (**H** = 4 classes |300 points/class). Tables 2 and 3 show the clustering rates of the Ext-MS and the SInt-MS on real data. The process to extract features on the different Riemannian manifolds will be described next.

Table 3. Clustering rates (%) of extrinsic (Ext-MS) and semi-intrinsic (SInt-MS) meanshift on Riemannian manifolds on object categorization for selected objects on the data set CIFAR10 [31]

$\mathcal{G}$				**Performance**		$\mathcal{V}$				**Performance**		$\mathcal{S}^+$		**Performance**	
Scales	Bins	m	k	Ext-Ms	SInt-Ms	Scales	Bins	m	k	Ext-Ms	SInt-Ms	d	d	Ext-Ms	SInt-Ms
E $\mapsto$ **Classes** = 3 | **Points per Class** = 250															
3	32	32	6	69.49	92.35	3	32	192	1	66.12	92.68	5	5	63.08	89.60
4	32	32	8	74.20	94.94	4	32	256	1	73.60	97.18	7	7	66.83	92.68
5	32	32	10	73.25	94.41	5	32	320	1	73.55	97.38	9	9	62.63	89.61
F $\mapsto$ **Classes** = 3 | **Points per Class** = 300															
3	32	32	6	63.95	95.54	3	32	192	1	61.66	94.26	5	5	56.31	94.26
4	32	32	8	68.55	97.70	4	32	256	1	69.48	98.55	7	7	60.29	96.40
5	32	32	10	69.01	94.66	5	32	320	1	67.35	95.81	9	9	59.53	93.16
G $\mapsto$ **Classes** = 4 | **Points per Class** = 250															
3	32	32	6	61.48	90.53	3	32	192	1	59.21	90.84	5	5	54.09	88.60
4	32	32	8	65.50	93.43	4	32	256	1	63.97	96.92	7	7	58.83	90.52
5	32	32	10	64.90	91.44	5	32	320	1	64.03	93.32	9	9	54.98	88.41
H $\mapsto$ **Classes** = 4 | **Points per Class** = 300															
3	32	32	6	57.64	89.96	3	32	192	1	55.30	90.73	5	5	50.35	88.29
4	32	32	8	60.48	93.97	4	32	256	1	59.40	94.71	7	7	53.94	92.78
5	32	32	10	58.98	90.30	5	32	320	1	59.13	91.49	9	9	50.22	88.16
Average Improvement : Δ = (SInt-MS) - (Ext-MS)															
$\Delta\mathcal{G}$ = + 28 %						$\Delta\mathcal{V}$ = + 30 %						$\Delta\mathcal{S}^+$ = + 34 %			

Stiefel ($\mathcal{V}_{k,m}$) Manifold

Regarding the real tests with the Stiefel manifold ($\mathcal{V}_{k,m}$) we follow the procedures presented in [3]. As refered in [3], directional data in the form of unit-norm feature vectors can be extracted from an object image. The authors in [3] proposed to obtain this type of feature vector from the magnitude of the image gradient and the Laplacian at three different scales. In our experiments a maximum of five different scales and two types of histograms were used, that is, $\upsilon_1 = \{1, 2, 4, 6, 8\}$ where υ_1 is the variance of the Gaussian filter and $\upsilon_2 = \{32, 64\}$ where υ_2 is the number of bins of the histogram. Let s be the number of scales used, for each of the $2s$ images a υ_2-bin histogram is computed and concatenated as a feature vector of length $m = 2s\upsilon_2$ (and then normalized). Therefore, the problem is posed as a clustering problem of points on the Stiefel manifold $\mathcal{V}_{1,m} \equiv S^{m-1}$ [3].

Grassmann ($\mathcal{G}_{k,m-k}$) Manifold

Regarding the real tests with the Grassmann manifold ($\mathcal{G}_{k,m-k}$) we follow the procedures presented in [3]. As refered in [3] the normalization process of the feature vector (used on the Stiefel manfiold case) may corrupt the underlying information of the individual histograms and then decrease the class separability. To tackle the above problem we assume that the l_1-norm of the each histogram is 1 and then we take the square root of each entry to make their l_2-norms equal to 1 [3]. Next, we form a feature matrix by stacking the $2s$ aforementioned υ_2-bin histograms as columns and then we take the SVD of the resulting $\upsilon_2 \times 2s$ matrix [3]. Its singular vectors span a subspace of dimension $k = 2s$ in $\Re^{m=\upsilon_2}$ [3]. Therefore, it is obtained a new representation of the feature as a point on G_{2s,υ_2-2s} [3].

Tensor Manifold ($\mathcal{S}_d^+$)

We used the well-known concept of region covariance matrix (RCM) [32] in order to test the proposed method in the tensor manifold using real data. We constructed three different types of $d \times d$ RCM's with $d = \{5, 7, 9\}$. The respective features extracted for each one of the configurations are ($d = 5 \rightarrow [x\ y\ \mathrm{I}\ |\mathrm{I}_x|\ |\mathrm{I}_y|]$) ; ($d = 7 \rightarrow [x\ y\ \mathrm{R\ G\ B}\ |\mathrm{I}_x|\ |\mathrm{I}_y|]$) ; ($d = 9 \rightarrow [x\ y\ \mathrm{R\ G\ B}\ |\mathrm{I}_x|\ |\mathrm{I}_y|\ |I_{xx}|\ |_{Iyy}|]$) where (x, y) is the pixel image position, $\mathrm{I}_x, \mathrm{I}_y, \mathrm{I}_{xx}, \mathrm{I}_{yy}$ are the gradients and $\mathrm{I}, \mathrm{R}, \mathrm{G}, \mathrm{B}$ are the gray/color components.

6 Conclusions

To the best of our knowledge this is the first work that proposes an intrinsic reformulation of the mean shift algorithm for general Riemannian manifolds. Experimental results on synthetic data as well as on real data clearly demonstrate that significant improvements in clustering accuracy can be achieved by employing this novel semi-intrinsic mean shift (SInt-MS) over the extrinsic counterpart (Ext-MS). We conclude that : the consequent usage of the intrinsic Riemannian structure of the space, in conjunction with an embedding of the Riemannian manifold into a Reproducing Kernel Hilbert Space (RKHS) by using a general and mathematically well-founded Riemannian kernel function (i.e. *heat kernel*), yields the most accurate and reliable approach presented so far to extend the well-known mean shift algorithm to general Riemannian manifolds. This allows us to extend mean shift based clustering and filtering techniques to a large class of frequently occurring Riemannian manifolds in vision and related areas, paving the way for other researchers.

Acknowledgments. This work was supported by the Portuguese Science Foundation (FCT) under the project Differential Geometry for Computer Vision and Pattern Recognition (PTDC/EEA-CRO/122812/2010). Rui Caseiro, João F. Henriques and Pedro Martins acknowledge the FCT through the grants SFRH/ BD74152/2010, SFRH/BD/75459/2010 and SFRH/BD/45178/2008, respectively.

References

1. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. TPAMI 24, 603–619 (2002)
2. Subbarao, R., Meer, P.: Nonlinear mean shift over riemannian manifolds. IJCV 84, 1–20 (2009)
3. Ceting, H., Vidal, R.: Intrinsic mean shift for clustering on stiefel and grassmann manifolds. In: CVPR (2009)
4. Berard, P., Besson, G., Gallot, S.: Embedding riemannian manifolds by their heat kernel. Geometric And Functional Analysis 4, 373–398 (1994)
5. Carreira, J., Caseiro, R., Batista, J., Sminchisescu, C.: Semantic Segmentation with Second-Order Pooling. In: Fitzgibbon, A., Lazebnik, S., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part VII. LNCS, vol. 7578, pp. 430–443. Springer, Heidelberg (2012)
6. Turaga, P., Veeraraghavan, A., Srivastava, A., Chellappa, R.: Statistical computations on grassmann and stiefel manifolds for image and video-based recognition. TPAMI 33, 2273–2286 (2011)
7. Caseiro, R., Henriques, J.F., Martins, P., Batista, J.: A nonparametric riemannian framework on tensor field application to foreground segmentation. In: ICCV (2011)

8. Pennec, X.: Statistical Computing on Manifolds for Computational Anatomy. In: L'Habilitation Diriger des Recherche (2006)
9. Turaga, P., Veeraraghavan, A., Srivastava, A., Chellappa, R.: Statistical Analysis on Manifolds and Its Applications to Video Analysis. In: Schonfeld, D., Shan, C., Tao, D., Wang, L. (eds.) Video Search and Mining. SCI, vol. 287, pp. 115–144. Springer, Heidelberg (2010)
10. Tuzel, O., Porikli, F., Meer, P.: Simultaneous multiple 3d motion estimation via mode finding on lie groups. In: ICCV (2005)
11. Sheikh, Y., Khan, E., Kanade, T.: Mode-seeking by medoidshifts. In: ICCV (2007)
12. Vedaldi, A., Soatto, S.: Quick Shift and Kernel Methods for Mode Seeking. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 705–718. Springer, Heidelberg (2008)
13. Subbarao, R., Meer, P.: Nonlinear mean shift for clustering over analytic manifolds. In: CVPR (2006)
14. Schpolkopf, B., Smola, A.J.: Learning with Kernels Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press (2002)
15. Lafferty, J., Lebanon, G.: Diffusion kernels on statistical manifolds. Journal of Machine Learning Research 6, 129–163 (2005)
16. Hamm, J., Lee, D.D.: Grassmann discriminant analysis: a unifying view on subspace-based learning. In: International Conference on Machine Learning (2008)
17. Harandi, M.T., Sanderson, C., Lovell, B.C.: Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching. In: CVPR (2011)
18. Avramidi, I.: Heat kernel approach in quantum field theory. Nuclear Physics B 104, 3–32 (2002)
19. Avramidi, I.: The covariant technique for calculation of the heat kernel asymptotic expansion. Physics Letters B 238, 92–97 (1990)
20. Avramidi, I.G.: Covariant techniques for computation of the heat kernel. Reviews in Mathematical Physics 11, 947–980 (1999)
21. Avramidi, I.G.: Analytic and Geometric Methods for Heat Kernel Applications in Finance. New Mexico Institute of Mining and Technology, NM 87801, USA (2007)
22. O'Neill, B.: Semi-Riemannian Manifolds: With Applications to Relativity. Academic Press (1983)
23. Tuzel, O., Porikli, F., Meer, P.: Kernel methods for weakly supervised mean shift clustering. In: ICCV (2009)
24. Kulis, B., Basu, S., Dhillon, I., Mooney, R.: Semi-supervised graph clustering: A kernel approach. In: International Conference on Machine Learning (2005)
25. Cristianini, N., Shawe-Taylor, J.: Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press (2000)
26. Narcowich, F.: Generalized hermite interpolation and positive definite kernels on a riemannian manifolds. Journal of Mathematical Analysis 190, 165–193 (1995)
27. Dyn, N., Narcowich, F., Ward, J.: A framework for interpolation and approximation on riemannian manifolds. In: Approximation Theory and Optimization: Tributes to M. J. D. Powell, pp. 133–144. Cambridge University Press (1997)
28. Gilkey, P.B.: Invariance Theory, the Heat Equation and the Atiyah-Singer Index Theorem. Publish or Perish (1984)
29. Chikuse, Y.: Statistics on Special Manifolds. Springer (2003)
30. Leibe, B., Schiele, B.: Analyzing appearance and contour based methods for object categorization. In: CVPR (2003)
31. Krizhevsky, A.: Learning multiple features from tiny images. T. Report (2009)
32. Tuzel, O., Porikli, F., Meer, P.: Region Covariance: A Fast Descriptor for Detection and Classification. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part II. LNCS, vol. 3952, pp. 589–600. Springer, Heidelberg (2006)