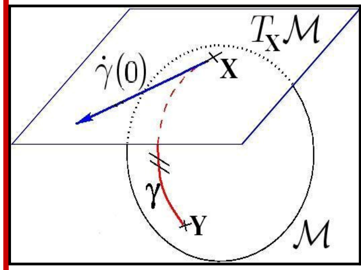
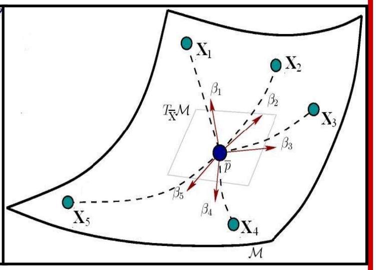
Background Modeling on Tensor Field for Foreground Segmentation

Rui Caseiro ♦ Pedro Martins ♦ Jorge Batista

Riemannian Geometry





▪ **Tensors** (**S+**) do not conform to Euclidean geometry, because the tensor space **is not a vector space**. Instead, tensors lies on a **Riemannian manifold** (differentiable manifold with a Riemannian metric). A manifold is a topological space which is locally similar to an Euclidean space.

▪ The **geodesic** between $\mathbf{X}$ and $\mathbf{Y}$ is defined as the minimum length curve γ(t) connecting these points. The **tangent space** $T_\mathbf{X}M$ is the vector space attached to $\mathbf{X}$**,** which contains the **tangent vectors** to all curves on M passing through $\mathbf{X}$. Given ,a tangent vector γ'(0) ∊ $T_\mathbf{X}M$ there exists a unique geodesic with γ(0) = $\mathbf{X}$, initial velocity γ'(0) and γ(1) = $\mathbf{Y}$.

▪ The **exponential map** $\exp_\mathbf{X}$: $T_\mathbf{X}M$→M maps the tangent vector γ'(0) at $\mathbf{X}$ = γ(0) to the point $\mathbf{Y}$ = γ(1) that is reached by the geodesic at time (t=ρ). The **logarithm map** $\log_\mathbf{X}$: M→$T_\mathbf{X}M$ maps any point $\mathbf{Y}$ to the unique tangent vector γ'(0) at $\mathbf{X}$ that is the initial velocity of the geodesic from $\mathbf{X}$ to $\mathbf{Y}$.

➢ **Euclidean:** the **distance** between two points $\mathbf{X}$,$\mathbf{Y}$ and the **distance gradient** are given as follows

$$\mathbf{D}_e\,(\mathbf{X},\mathbf{Y}) = |\,\mathbf{X}-\mathbf{Y}\,|_F = \sqrt{\mathrm{tr}((\mathbf{X}-\mathbf{Y})(\mathbf{X}-\mathbf{Y})^T)} \qquad \nabla_\mathbf{X}\mathbf{D}_e^2(\mathbf{X},\mathbf{Y}) = \mathbf{X}-\mathbf{Y}$$

➢ **Affine-Invariant:** the **geodesic** defined by the initial point γ(0) = $\mathbf{X}$ and the tangent vector γ'(0) is expressed as (t=1 → **exponential map**)

$$\gamma(t) = \exp_\mathbf{X}\left[t\dot{\gamma}(0)\right] = \mathbf{X}^{\frac{1}{2}}\exp\left[(t)\mathbf{X}^{-\frac{1}{2}}\dot{\gamma}(0)\mathbf{X}^{-\frac{1}{2}}\right]\mathbf{X}^{\frac{1}{2}}$$

The respective **logarithm map** is defined as

$$\dot{\gamma}(0) = \log_\mathbf{X}(\mathbf{Y}) = -\mathbf{X}\log\left(\mathbf{Y}^{-1}\mathbf{X}\right)$$

The **geodesic distance** between two points $\mathbf{X}$,$\mathbf{Y}$ induced by the Affine-Invariant metric, derived from the Fisher Information matrix is given as

$$\mathbf{D}_a\,(\mathbf{X},\mathbf{Y}) = \sqrt{\frac{1}{2}\mathrm{tr}\left(\log^2\left(\mathbf{X}^{-\frac{1}{2}}\mathbf{Y}\mathbf{X}^{-\frac{1}{2}}\right)\right)} \qquad \nabla_\mathbf{X}\mathbf{D}_a^2(\mathbf{X},\mathbf{Y}) = \mathbf{X}\log\left(\mathbf{Y}^{-1}\mathbf{X}\right)$$

The **distance gradient** is the negative of the initial velocity γ'(0).

➢ **Log-Euclidean:** based on specific properties of the matrix exponential on tensors, it is possible to define a vector space structure on tensors. Since under the matrix exponentiation, there is a one-to-one mapping between the tensor space and the vector space of symmetric matrices, one can transfer to tensors the standard algebraic operations with the matrix exponential.

The tensor vector space with this metric is in fact *isomorphic* and *isometric* with the corresponding Euclidean space of symmetric matrices. Results obtained on logarithms are mapped back to the tensor domain with the exponential. The **geodesic** is expressed as (t=1 → **exponential map**)

$$\gamma(t) = \exp_\mathbf{X}\left[t\dot{\gamma}(0)\right] = \exp\left[\log(\mathbf{X}) + \partial_\mathbf{X}\log.[t\dot{\gamma}(0)]\right]$$

The respective **logarithm map** is defined as

$$\dot{\gamma}(0) = \log_\mathbf{X}(\mathbf{Y}) = \partial_{\log(\mathbf{X})}\exp.\left[\log(\mathbf{Y}) - \log(\mathbf{X})\right]$$

Since the Log-Euclidean metrics corresponds to Euclidean metrics in the logarithms domain, the **interpolation** between tensors is simplified as

$$\gamma(t) = \exp\left[(1-t)\log(\mathbf{X}) + (t)\log(\mathbf{Y})\right]$$

The **geodesic distance** between two points $\mathbf{X}$,$\mathbf{Y}$ induced by the Log-Euclidean metric, is also extremely simplified as follows

$$\mathbf{D}_l\,(\mathbf{X},\mathbf{Y}) = \sqrt{\mathrm{tr}[(\log(\mathbf{Y}) - \log(\mathbf{X}))^2]} \qquad \nabla_\mathbf{X}\mathbf{D}_l^2(\mathbf{X},\mathbf{Y}) = -\dot{\gamma}(0)$$

The Log-Euclidean distance is much simpler than the Affine case where matrix multiplications, square roots, inverses are used. However, the **exponential** and **logarithm** mappings are complicated in the Log-Euclidean case by the use of the matrix differentials.
Using spectral properties of symmetric matrices, one can compute an explicit and efficiently closed-form expression for these differentials.

Background Modeling

➢ We propose a new method to background modeling using the tensor concept. The combination of color and texture features can improve segmentation performance. The **structure tensor** was used to convert the image into a more information rich form and is defined as $\mathbf{T} = K_\rho * (vv')$ where v = [$\mathbf{I_x}$; $\mathbf{I_y}$; $\mathbf{I_r}$; $\mathbf{I_g}$; $\mathbf{I_b}$] (K_ρ is a smoothing kernel).

➢ The tensor space does not form a vector space, thus linear statistical techniques do not apply. Taking into account the differential geometrical properties of the Riemannian manifold where tensors lie , we propose a novel approach for foreground detection on tensor field based on data modeling by means of GMM directly on tensor domain.

➢ We introduced a **K-means** approximation of the **EM** algorithm based on an **Affine-Invariant** metric. This metric has excellent theoretical properties but essentially due to the space curvature the computational burden is high. We propose a new K-means based on a new family of metrics, called **Log-Euclidean** , in order to speed up the process. Based on a novel vector space structure for tensors, the Log-Euclidean transforms computations on tensors into Euclidean computations on vectors in the logarithms domain.

➢From a practical point of view yield similar results, with an experimental computation time ratio of at least 2 and sometime more in favor of the Log-Euclidean.

Background Modeling

➢ We model the background with a **GMM** on tensor space. Based on the definition of a Gaussian law on this space, we can define a GMM as follows

$$p(\mathbf{T}_i\,|\,\Theta) = \sum_{k=1}^{K}\omega_k\frac{\exp\left(-(1/2)\varphi(\beta_{i,k})^T\Lambda_k^{-1}\varphi(\beta_{i,k})\right)}{\sqrt{(2\pi)^n\,|\,\Lambda_k\,|}} \qquad \beta_{i,k} = -\nabla_{\overline{\mathbf{T}}_k}\mathbf{D}^2\left(\overline{\mathbf{T}}_k,\mathbf{T}_i\right)$$

➢ The clustering of data lying on the **S+** is posed as a maximum likelihood estimation problem. An exact EM algorithm is a costly procedure. In order to speed up the process we propose a online **K-means** approximation of EM.

➢ **Kmeans (Euclidean):** the new mixture parameters combine the prior information with the observed sample. The model parameters are updated using an exponential decay scheme with learning rates (ρ) and (α).

$$\begin{cases}\overline{\mathbf{T}}_k^t = (1-\rho)\overline{\mathbf{T}}_k^{t-1} + \rho\mathbf{T}_i \\ \Lambda_k^t = (1-\rho)\Lambda_k^{t-1} + \rho\varphi(\beta_{i,k}^t)^T\varphi(\beta_{i,k}^t)\end{cases} \qquad \begin{cases}\rho = \alpha N(\mathbf{T}_i\,|\,\overline{\mathbf{T}}_k^{t-1},\Lambda_k^{t-1}) \\ \beta_{i,k}^t = -\nabla_{\overline{\mathbf{T}}_k^t}\mathbf{D}_e^2\left(\overline{\mathbf{T}}_k^t,\mathbf{T}_i\right)\end{cases}$$

➢ **Kmeans (Affine-Invariant):** the mean update equation presented previously can only be applied in the Euclidean case. To take into account the Riemannian geometry of the manifold, we proposed a method to update the mean, based on the concept of tensor interpolation. The point ($\mathbf{Z}$) that is reached by the geodesic at time (t=ρ) is estimated as

$$\mathbf{Z} = \overline{\mathbf{T}}_k^t \qquad \mathbf{X} = \overline{\mathbf{T}}_k^{t-1} = \gamma(0) \qquad \mathbf{Y} = \mathbf{T}_i = \gamma(1)$$

$$\begin{cases}\mathbf{Z} = \gamma(\rho) = \mathbf{X}^{\frac{1}{2}}\exp\left[(\rho)\mathbf{X}^{-\frac{1}{2}}\left[-\mathbf{X}\log\left(\mathbf{Y}^{-1}\mathbf{X}\right)\right]\mathbf{X}^{-\frac{1}{2}}\right]\mathbf{X}^{\frac{1}{2}} \\ \beta_{i,k}^t = -\nabla_{\overline{\mathbf{T}}_k^t}\mathbf{D}_a^2(\overline{\mathbf{T}}_k^t,\mathbf{T}_i) = -\nabla_\mathbf{Z}\mathbf{D}_a^2(\mathbf{Z},\mathbf{Y}) = -\mathbf{Z}\log(\mathbf{Y}^{-1}\mathbf{Z})\end{cases}$$

➢ **Kmeans (Log-Euclidean):** in this case, a closed-form and simple expression for interpolation between tensors exists. The point $\mathbf{Z}$ between $\mathbf{X}$ and $\mathbf{Y}$ that is reached by the geodesic γ(t) at time (t=ρ) is estimated as

$$\begin{cases}\mathbf{Z} = \gamma(\rho) = \exp[(1-\rho)\log(\mathbf{X}) + (\rho)\log(\mathbf{Y})] \\ \beta_{i,k}^t = -\nabla_{\overline{\mathbf{T}}_k^t}\mathbf{D}_l^2\left(\overline{\mathbf{T}}_k^t,\mathbf{T}_i\right) = -\nabla_\mathbf{Z}\mathbf{D}_l^2\left(\mathbf{Z},\mathbf{Y}\right) = \partial_{\log(\mathbf{Z})}\exp.[\log(\mathbf{Y}) - \log(\mathbf{Z})]\end{cases}$$

Results

Methods	Sequence 1 TPR	Sequence 1 FPR	Sequence 2 TPR	Sequence 2 FPR
GMM [Ir, Ig, Ib]	55.20	16.07	58.00	17.12
GMM [Ir, Ig, Ib, Ix, Iy]	61.30	13.50	63.20	14.65
KDE [Ir, Ig, Ib]	62.80	12.83	65.60	13.32
KDE [Ir, Ig, Ib, Ix, Iy]	69.10	10.95	71.50	11.26
GMM [Tensor] Euclidean	72.50	8.72	76.40	8.95
GMM [Tensor] Affine-Invariant	93.00	2.08	90.50	2.24
GMM [Tensor] Log-Euclidean	91.40	2.20	89.60	2.45

TPR = True positive ratio

FPR = False positive ratio

Top→Bottom: Original ; GMM(v) ; KDE(v) ; GMM(**T**)-Euclidean ; GMM(**T**)-AffineInvariant ; GMM(**T**)-LogEuclidean

Theoretic analysis/experimental evaluations demonstrate the promise/effectiveness of the proposed framework. It is stressed that no morphological operators were used.

This work was funded by BRISA, Auto-estradas de Portugal, S.A.

Institute of Systems and Robotics
University of Coimbra