# Riemannian Coding and Dictionary Learning: Kernels to the Rescue

Mehrtash Harandi, Mathieu Salzmann
Australian National University & NICTA

While sparse coding on non-flat Riemannian manifolds has recently become increasingly popular, existing solutions either are dedicated to specific manifolds, or rely on optimization problems that are difficult to solve, especially when it comes to dictionary learning. In this paper, we propose to make use of kernels to perform coding and dictionary learning on Riemannian manifolds. To this end, we introduce a general Riemannian coding framework with its kernel-based counterpart. This lets us **(i)** generalize beyond the special case of sparse coding; **(ii)** introduce efficient solutions to two coding schemes; **(iii)** learn the kernel parameters; **(iv)** perform unsupervised and supervised dictionary learning in a much simpler manner than previous Riemannian coding approaches.

More specifically, let $\mathcal{D} = \{d_i\}_{i=1}^N$, $d_i \in \mathcal{M}$, be a dictionary on a Riemannian manifold $\mathcal{M}$, and $x \in \mathcal{M}$ be a query point on the manifold. We define a general Riemannian coding formulation as

$$\begin{aligned} &\min_{\alpha} \ \delta^2\Big(x, \biguplus_{j=1}^N \alpha_j d_j\Big) + \lambda\gamma(\alpha; x, \mathcal{D}) \\ &\text{s.t.} \quad \alpha \in \mathcal{C}, \end{aligned} \tag{1}$$

where $\delta : \mathcal{M} \times \mathcal{M} \to \mathbb{R}^+$ is a metric on $\mathcal{M}$, $\alpha \in \mathbb{R}^N$ is the vector of Riemannian codes, $\gamma$ is a prior on the codes $\alpha$ and $\mathcal{C}$ is a set of constraints on $\alpha$. Moreover, $\biguplus : \mathcal{M} \times \cdots \times \mathcal{M} \times \mathbb{R} \times \mathbb{R} \cdots \times \mathbb{R} \to \mathcal{M}$ is an operator that combines multiple dictionary atoms $\{d_j \in \mathcal{M}\}$ with weights $\{\alpha_j\}$ and generates a point $\hat{x}$ on $\mathcal{M}$. This general formulation encapsulates intrinsic sparse coding [2, 5], but also lets us derive and intrinsic version of Locality-constrained Linear Coding [10]. Such intrinsic formulations, however, depend on the logarithm map, which may be highly nonlinear, or not even have an analytic solution.

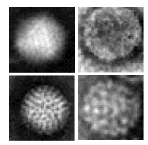
To overcome these weaknesses and obtain a general formulation of Riemannian coding, we propose to perform coding in RKHS. This has the twofold advantage of yielding simple solutions to several popular coding techniques and of resulting in a potentially better representation than standard coding techniques due to the nonlinearity of the approach. To this end, let $\phi : \mathcal{M} \to \mathcal{H}$ be a mapping to an RKHS induced by the kernel $k(x,y) = \phi(x)^T\phi(y)$. Coding in $\mathcal{H}$ can then be formulated as

$$\begin{aligned} &\min_{\alpha} \Big\| \phi(x) - \sum\nolimits_{j=1}^N \alpha_j \phi(d_j) \Big\|_2^2 + \lambda\gamma(\alpha; \phi(x), \phi(\mathcal{D})) \\ &\text{s.t.} \quad \alpha \in \mathcal{C}. \end{aligned} \tag{2}$$

As shown in the paper, the reconstruction term in (2) can be kernelized. More importantly, after kernelization, this term remains quadratic, convex and similar to its counterpart in Euclidean space. This lets us derive efficient solutions to two coding schemes: kernel Sparse Coding (kSC) and kernel Locality Constrained Coding (kLCC).

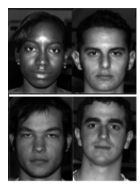
In many cases, it is beneficial not only to compute the codes for a given dictionary, but also to optimize the dictionary to best suit the problem at hand. Given training data, and for fixed codes, we then show that, by relying on the *Representer theorem* [8], the dictionary update has an analytic form. Furthermore, we introduce an approach to supervised dictionary learning, which, given labeled data, jointly learns the dictionary and a classifier acting on the codes. The resulting supervised coding schemes are referred to as kSSC and kSLCC.

We demonstrate the effectiveness of our approach on three different types of non-flat manifolds, as well as illustrate its generality by also applying it to Euclidean space, which simply is a special type of Riemannian manifold. In particular, we evaluated our different techniques on two challenging classification datasets where the images are represented with region covariance descriptors (RCovDs) [9], which lie on SPD manifolds. To this end, we considered two classification tasks: virus classification using the Kylberg dataset and material categorization using the KTH-TIPS2b dataset. Furthermore, to illustrate the effectiveness of our coding schemes in Euclidean space, we made use of the extended YALE-B dataset and of Caltech101. The results given in Tables 1 and 2 evidence the benefits of our kernel coding schemes over existing baselines, with kSLCC generally arising as the best performer. Additional results confirming this trend for Grassmann manifolds and for the shape manifold are provided in the paper and in supplementary material.




| Method | Virus | KTH-TIPS2-b |
|---|---|---|
| CDL [11] | 69.5% | 76.3% |
| logEuc-SC | 68.3% | 67.8% |
| logEuc-LCC | 72.3% | 75.9% |
| kSC | 78.5% | 78.8% |
| kLCC | 79.4% | 79.8% |
| kSSC | 81.7% | 79.9% |
| kSLCC | **82.0%** | **81.2%** |

Table 1: **Coding on SPD manifolds. Left panel.** Samples from the Virus dataset [7]. **Middle panel.** Samples from the KTH-TIPS2b texture dataset [1]. **Right panel.** Recognition accuracies.




| Method | YALE-B | Caltech101 |
|---|---|---|
| SRC [12] | 80.5% | 70.7% |
| LC-KSVD [6] | 95.0% | 73.6% |
| kSC | 96.9% | 75.1% |
| kLCC | 97.2% | 75.4% |
| kSSC | 98.2% | 75.7% |
| kSLCC | **98.4%** | **76.2%** |

Table 2: **Coding in Euclidean space. Left panel.** Samples from extended YALE-B [4]. **Middle panel.** Samples from Caltech101 [3]. **Right panel.** Recognition accuracies.

[1] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *Proc. Int. Conference on Computer Vision (ICCV)*, volume 2, pages 1597–1604, 2005.

[2] H.E. Cetingul, M.J. Wright, P.M. Thompson, and R. Vidal. Segmentation of high angular resolution diffusion mri using sparse riemannian manifold clustering. *IEEE Transactions on Medical Imaging*, 33(2):301–317, Feb 2014.

[3] Li Fei-Fei, Rob Fergus, and Pietro Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding (CVIU)*, 106(1): 59–70, 2007.

[4] Athinodoros S. Georghiades, Peter N. Belhumeur, and David Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.

[5] Jeffrey Ho, Yuchen Xie, and Baba Vemuri. On a nonlinear generalization of sparse coding and dictionary learning. In *Proc. Int. Conference on Machine Learning (ICML)*, pages 1480–1488, 2013.

[6] Zhuolin Jiang, Zhe Lin, and L.S. Davis. Label consistent k-svd: Learning a discriminative dictionary for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2651–2664, 2013.

[7] Gustaf Kylberg and Ida-Maria Sintorn. Evaluation of noise robustness for local binary pattern descriptors in texture classification. *EURASIP Journal on Image and Video Processing*, 2013(1):1–20, 2013.

[8] Bernhard Schölkopf, Ralf Herbrich, and Alex J Smola. A generalized representer theorem. In *Computational learning theory*, pages 416–426. Springer, 2001.

[9] Oncel Tuzel, Fatih Porikli, and Peter Meer. Region covariance: A fast descriptor for detection and classification. In *Proc. European Conference on Computer Vision (ECCV)*, pages 589–600. Springer, 2006.

[10] Jinjun Wang, Jianchao Yang, Kai Yu, Fengjun Lv, Thomas Huang, and Yihong Gong. Locality-constrained linear coding for image classification. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3360–3367, 2010.

[11] Ruiping Wang, Huimin Guo, Larry S Davis, and Qionghai Dai. Covariance discriminative learning: A natural and efficient approach to image set classification. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2496–2503, 2012.

[12] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma. Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(2):210–227, 2009.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.