Multiscale Geometric Dictionaries for Point-cloud Data

William K. Allard[1], Guangliang Chen[1], Mauro Maggioni[1,2]

Departments of Mathematics[1] and Computer Science[2], Duke University, PO BOX 90320, Durham, NC 27708

Abstract

We develop a novel geometric multiresolution analysis for analyzing intrinsically low dimensional point clouds in high-dimensional spaces, modeled as samples from a d-dimensional set $\mathcal{M}$ (in particular, a manifold) embedded in $\mathbb{R}^D$, in the regime $d \ll D$. This type of situation has been recognized as important in various applications, such as the analysis of sounds, images, and gene arrays. In this paper we construct data-dependent multiscale dictionaries that aim at efficient encoding and manipulating of the data. Unlike existing constructions, our construction is fast, and so are the algorithms that map data points to dictionary coefficients and vice versa. In addition, data points have a guaranteed sparsity in terms of the dictionary.

Keywords: *Wavelets. Multiscale Analysis. Sparse Coding. Point Clouds.*

Introduction

Data sets are often modeled as point clouds in $\mathbb{R}^D$, for D large, but having some interesting low-dimensional structure, for example that of a d-dimensional manifold $\mathcal{M}$, with $d \ll D$. When $\mathcal{M}$ is simply a linear subspace, one may exploit this assumption for encoding efficiently the data by projecting onto a dictionary of d vectors in $\mathbb{R}^D$ (for example found by SVD), at a cost $(n+D)d$ for n data points. When $\mathcal{M}$ is nonlinear, there are no "explicit" constructions of dictionaries that achieve a similar efficiency: typically one uses either random dictionaries, or dictionaries obtained by black-box optimization. This type of situation has been recognized as important in various applications, and has been at the center of much investigation in the applied mathematics and machine learning communities during the past several years.

We formalize this approach by requesting to find a dictionary Φ of size I, based on training data, such that every point (at least from the training data set) may be represented, up to a certain precision ϵ, by at most m elements of the dictionary. This requirement of sparsity of the representation is very natural from the viewpoints of statistics, signal processing, and interpretation of the representation. Of course, the smaller I and m are, for a given ϵ, the better the dictionary.

Current constructions of these dictionaries such as K-SVD [1], k-flats [5] and Bayesian methods [6] have several deficiencies. First, they cast the requirements above as an optimization problem, with many local minima, and for iterative algorithms little is known about their computational complexity. Second, no guarantees are provided about the size of I, m (as a function of ϵ). Lastly, the dictionaries found in this way are in general highly over-complete and unstructured. As a consequence, there is no fast algorithm for computing the coefficients of a data point w.r.t. the dictionary, thus requiring appropriate sparsity-seeking algorithms.

In this paper we construct data-dependent dictionaries based on a *geometric multiresolution analysis (GMRA)* of the data, inspired by multiscale techniques in geometric measure theory [4]. These dictionaries are structured in a multiscale fashion; the expansion of a data point on the dictionary elements is guaranteed to have a certain degree of sparsity m; both the dictionary elements and the coefficients may be computed by a fast algorithm; the growth of the number of dictionary elements I as a function of ϵ is controlled theoretically, and easy to estimate in practice. We call the elements of these dictionaries *geometric wavelets*, since in some aspects they generalize wavelets from vectors that analyze functions to affine vectors that analyze point clouds.

Geometric Wavelets

Let $(\mathcal{M}, g)$ be a d-dimensional compact Riemannian manifold isometrically embedded in $\mathbb{R}^D$, with $d \ll D$. Assume we have n samples drawn i.i.d. from $\mathcal{M}$, according to the natural volume measure $d$vol on $\mathcal{M}$. We use such training data to present how to construct geometric wavelets, though our construction easily extends to any point-cloud data, by using locally adaptive dimensions $d_{j,k}$ (rather than a fixed d).

Multiscale decomposition. We start by constructing a multiscale nested partition of $\mathcal{M}$ into dyadic cells $\{C_{j,k}\}_{k\in\Gamma_j, 0\le j\le J}$ that satisfy the usual properties of dyadic cubes in $\mathbb{R}^D$. There is a natural tree $\mathcal{T}$ associated to the family: For any $j \in \mathbb{Z}$ and $k \in \Gamma_j$, we let children$(j,k) = \{k' \in \Gamma_{j+1} : C_{j+1,k'} \subseteq C_{j,k}\}$. Also, for $x \in \mathcal{M}$, we denote by $C_{j,x}$ the unique cell at scale j that contains x (similar notation like $P_{j,x}, \Phi_{j,x}, \Psi_{j,x}$ associated to $C_{j,x}$ are used later).

Multiscale SVD. For every $C_{j,k}$ we define the mean (in $\mathbb{R}^D$) by $\overline{c}_{j,k} := \mathbb{E}[x|x \in C_{j,k}] = \frac{1}{\mathrm{vol}(C_{j,k})}\int_{C_{j,k}} x\, d\mathrm{vol}(x)$ and the covariance by $\mathrm{cov}_{j,k} = \mathbb{E}[(x-\overline{c}_{j,k})(x-\overline{c}_{j,k})^*|x\in C_{j,k}]$. Let the rank-$d$ Singular Value Decomposition (SVD) of $\mathrm{cov}_{j,k}$ be $\mathrm{cov}_{j,k} = \Phi_{j,k}\Sigma_{j,k}\Phi_{j,k}^*$, where $\Phi_{j,k}$ is orthonormal and Σ is diagonal.The subspace spanned by the columns of $\Phi_{j,k}$, and then translated to pass through $\overline{c}_{j,k}$, $\langle\Phi_{j,k}\rangle + \overline{c}_{j,k}$, is an approximate tangent space to $\mathcal{M}$ at location $\overline{c}_{j,k}$ and scale 2^{-j}. We think of $\{\Phi_{j,k}\}_{k\in\Gamma_j}$ as a family of *geometric scaling functions* at scale j. Let $P_{j,k}$ be the associated affine projection

$$P_{j,k}(x) = \Phi_{j,k}\Phi_{j,k}^*(x-\overline{c}_{j,k}) + \overline{c}_{j,k}, \quad \forall\, x \in C_{j,k} \tag{1}$$

We define the coarse approximations, at scale j, to the manifold $\mathcal{M}$ and to any point $\forall x \in \mathcal{M}$, as follows:

$$\mathcal{M}_j := \cup_{k\in\Gamma_j} P_{j,k}(C_{j,k}), \quad x_j \equiv P_{\mathcal{M}_j}(x) := P_{j,x}(x). \tag{2}$$

Multiscale geometric wavelets. We introduce our wavelet encoding of the difference between $\mathcal{M}_j$ and $\mathcal{M}_{j+1}$, for $j < J$. Fix $x \in C_{j+1,k'} \subset C_{j,k}$. The difference $x_{j+1} - x_j$ is a high-dimensional vector in $\mathbb{R}^D$, however it may be decomposed into a sum of vectors in certain well-chosen low-dimensional spaces, shared across multiple points, in a multiscale fashion. We proceed as follows:

$$\begin{aligned} Q_{\mathcal{M}_{j+1}}(x) &:= P_{\mathcal{M}_{j+1}}(x) - P_{\mathcal{M}_j}(x) \\ &= x_{j+1} - P_{j,k}(x_{j+1}) + P_{j,k}(x_{j+1}) - P_{j,k}(x) \\ &= (I - \Phi_{j,k}\Phi_{j,k}^*)(x_{j+1} - \overline{c}_{j,k}) + \Phi_{j,k}\Phi_{j,k}^*(x_{j+1} - x). \end{aligned} \tag{3}$$

Define the wavelet subspace and translation as

$$W_{j+1,k'} := (I - \Phi_{j,k}\Phi_{j,k}^*)\,\langle\Phi_{j+1,k'}\rangle; \tag{4}$$

$$w_{j+1,k'} := (I - \Phi_{j,k}\Phi_{j,k}^*)(\overline{c}_{j+1,k'} - \overline{c}_{j,k}). \tag{5}$$

Clearly $\dim W_{j+1,k'} \le d$. Let $\Psi_{j+1,k'}$ be an orthonormal basis for $W_{j+1,k'}$ which we call a *geometric wavelet basis*. Then we may rewrite (3) as

$$\begin{aligned} Q_{\mathcal{M}_{j+1}}(x) = \Psi_{j+1,k'}\Psi_{j+1,k'}^*(x_{j+1} - \overline{c}_{j+1,k'}) + w_{j+1,k'} \\ - \Phi_{j,k}\Phi_{j,k}^*(x - x_{j+1}). \end{aligned} \tag{6}$$

The last term $x - x_{j+1}$ can be closely approximated by $x_J - x_{j+1} = \sum_{l=j+1}^{J-1} Q_{\mathcal{M}_{l+1}}(x)$ as the finest scale $J \to +\infty$, under general conditions. These operators are "detail" operators analogous to the wavelet projections in wavelet theory, and satisfy, by construction, the crucial multiscale relationship

$$P_{\mathcal{M}_{j+1}}(x) = P_{\mathcal{M}_j}(x) + Q_{\mathcal{M}_{j+1}}(x), \quad \forall\, x \in \mathcal{M}. \tag{7}$$

We have therefore constructed a multiscale family of projection operators $P_{\mathcal{M}_j}$ (one for each node $C_{j,k}$) onto approximate local tangent planes and detail projection operators $Q_{\mathcal{M}_{j+1}}$ (one for each edge) encoding the differences, collectively referred to as a GMRA structure. The cost of encoding the GMRA structure is dominated by that of the scaling functions $\{\Phi_{j,k}\}$, which is $\mathcal{O}(dD\epsilon^{-\frac{d}{2}})$, and the time complexity of the algorithm is $\mathcal{O}(Dn\log(n))$ [2].

Theorem 1 (Geometric Wavelet Decomposition)**.** *Let $(\mathcal{M}, g)$ be a $\mathcal{C}^2$ manifold of dimension d in $\mathbb{R}^D$, and $\{P_{\mathcal{M}_j}, Q_{\mathcal{M}_{j+1}}\}$ a GMRA. For any $x \in \mathcal{M}$, there exists a constant $C = C(x)$ and a scale $j_0 = j_0(\mathrm{reach}(\mathbb{B}_x(1)))$, such that for any $j \ge j_0$,*

$$\|x - P_{\mathcal{M}_{j_0}}(x) - \sum_{l=j_0+1}^{j} Q_{\mathcal{M}_l}(x)\| \le C \cdot 2^{-2j}. \tag{8}$$

Geometric Wavelet Transforms (GWT). Given a GMRA structure, we may compute a discrete Forward GWT for a point $x \in \mathcal{M}$ that maps it to a sequence of wavelet coefficient vectors:

$$q_x = (q_{J,x}, q_{J-1,x}, \ldots, q_{1,x}, q_{0,x}) \in \mathbb{R}^{d+\sum_{j=1}^J d_{j,x}^w} \tag{9}$$

where $q_{j,x} := \Psi_{j,x}^*(x_j - c_{j,x})$, and $d_{j,x}^w := \dim W_{j,x} \le d$. Note that, for a fixed precision $\epsilon > 0$, q_x has a maximum possible length $(1 + \frac{1}{2}\log_2\frac{1}{\epsilon})d$, which is independent of D and nearly optimal in d [3]. On the other hand, we may easily compute the wavelets at all scales using the GMRA structure and the wavelet coefficients, by a discrete Inverse GWT:

$$Q_{\mathcal{M}_J}(x) = \Psi_{J,x}q_{J,x} + w_{J,x}; \tag{10}$$

$$Q_{\mathcal{M}_j}(x) = \Psi_{j,x}q_{j,x} + w_{j,x} - \Phi_{j-1,x}\Phi_{j-1,x}^* \sum_{\ell>j} Q_{\mathcal{M}_\ell}(x) \tag{11}$$

for j gradually decreasing from $J-1$ to 1. The finest approximation of x is $\hat{x}_J = x_0 + \sum_{j\ge1} Q_{\mathcal{M}_j}(x)$.

Experiments

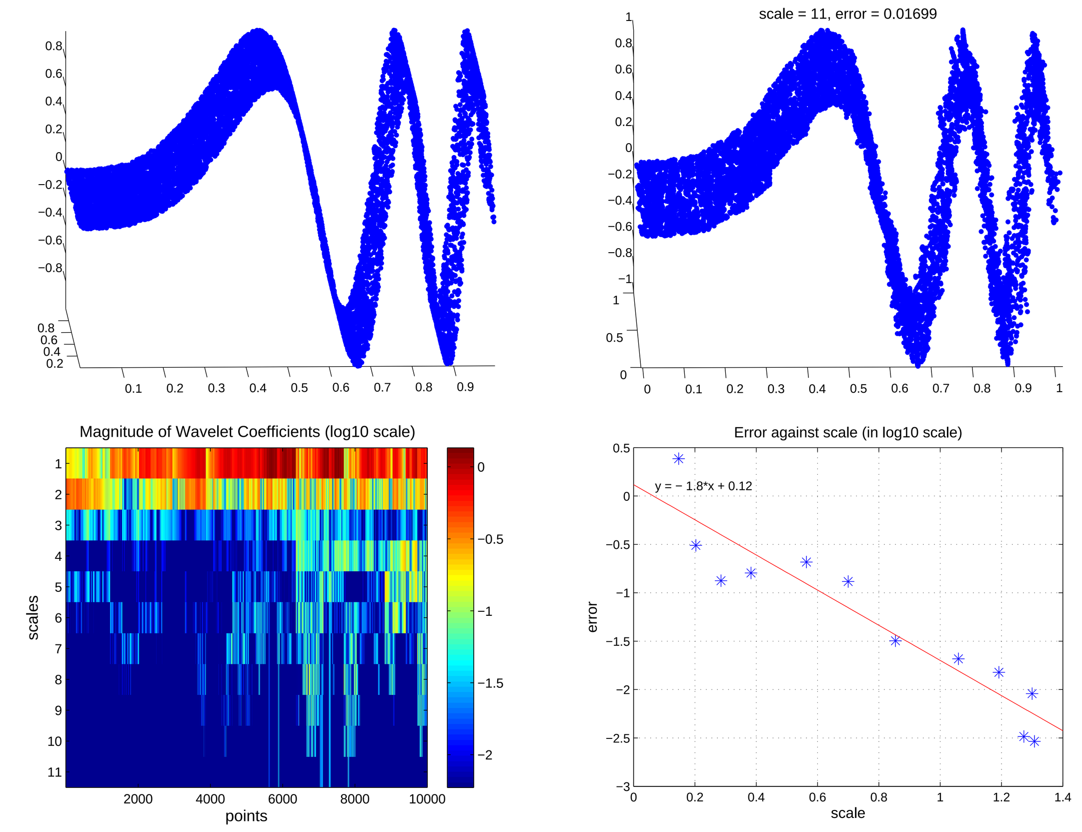

Figure 1. We sample 10,000 points from a 2-D *Oscillating2DWave* embedded in $\mathbb{R}^{50}$ and compute the GWT of the data. Bottom left figure shows the wavelet coefficients arranged into the natural tree: The x-axis indexes the points, and the y axis indexes the scales from coarsest (1) to finest (11)

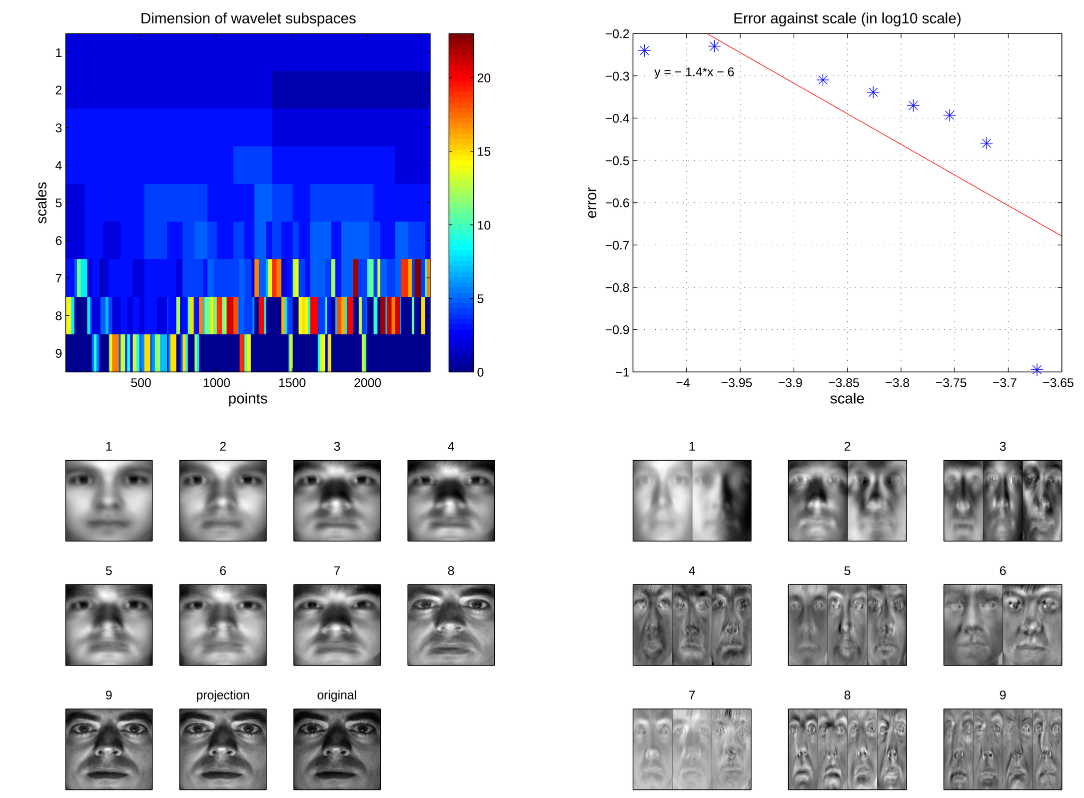

Figure 2. We apply the GMRA to 2414 (cropped) face images from 38 human subjects in fixed frontal pose under varying illumination angles. Bottom row shows approximations of a fixed point (image) and corresponding dictionary elements used

Acknowledgements. MM is grateful for partial support from NSF (DMS 0650413, CCF 0808847, IIS 0803293), ONR N00014-07-1-0625, the Sloan Foundation and Duke. GC was partially supported by ONR N00014-07-1-0625 and NSF CCF 0808847.

References.

[1] M. Aharon, M. Elad, and A. M. Bruckstein. The K-SVD algorithm. In *SPARSE'05.*

[2] W. K. Allard, G. Chen, and M. Maggioni. Multiscale geometric methods for data sets II: Geometric wavelets. In preparation.

[3] G. Chen and M. Maggioni. Multiscale geometric wavelets for the analysis of point clouds. In *CISS '10.*

[4] G. David. *Wavelets and Singular Integrals on Curves and Surfaces.* Springer, 1991.

[5] A. Szlam and G. Sapiro. Discriminative k-metrics. In *ICML '09.*

[6] M. Zhou, H. Chen, J. Paisley, L. Ren, G. Sapiro, and L. Carin. Non-parametric Bayesian dictionary learning for sparse image representations. In *NIPS '09.*