Learning Orthogonal Bases for k-Sparse Representations

Henry Schütze, Erhardt Barth, Thomas Martinetz

Institute for Neuro- and Bioinformatics, University of Lübeck
Ratzeburger Allee 160, 23562 Lübeck, Germany
schuetze@inb.uni-luebeck.de

Sparse Coding aims at finding a dictionary for a given data set, such that each sample can be represented by a linear combination of only few dictionary atoms. Generally, sparse coding dictionaries are overcomplete and not orthogonal. Thus, the processing substep to determine the optimal \( k \)-sparse representation of a given sample by the current dictionary is \( NP \)-hard. Usually, the solution is approximated by a greedy algorithm or by \( l_1 \) convex relaxation. With an orthogonal dictionary, however, an optimal \( k \)-sparse representation can not only be efficiently, but exactly computed, because a corresponding \( k \)-sparse coefficient vector is given by the \( k \) largest absolute projections.

In this paper, we present the novel online learning algorithm Orthogonal Sparse Coding (OSC), that is designed to find an orthogonal basis \( U = (u_1, \ldots, u_d) \) for a given data set \( X \in \mathbb{R}^{d \times L} \), such that for any \( k \in \{1, \ldots, d\} \), the optimal \( k \)-sparse coefficient vectors \( A \in \mathbb{R}^{d \times L} \) minimize the average representation error

\[
E = \frac{1}{dL} \| X - UA \|_F^2.
\]

At each learning step \( t \), OSC randomly selects a sample \( x \) from \( X \) and determines an index sequence \( h_1, \ldots, h_d \) of decreasing overlaps \( |u_i^T x| \) between \( x \) and the basis vectors in \( U \). In the order of that sequence, each

(a) Learned basis from 1,000 synthetic image patches of size 16×16 pixel.

(b) Learned basis from 20,000 natural image patches of size 16×16 pixel.

Fig. 1: Basis patches learned with OSC.
basis vector $u_h$ is updated by the Hebbian learning rule $\Delta u_h = \varepsilon_t (u_h^T x)x$ with a subsequent unit length normalization. After each basis vector update, $x$ and the next basis vector $u_{h+1}$ to be adapted are projected onto the orthogonal complement $\text{span}(\{u_1, ..., u_h\})^\perp$ wherein the next update takes place.

We applied OSC to (i) 1,000 synthetic ($k=50$)-sparse patches of size 16$\times$16 pixel, randomly generated with a 2D Haar basis, and (ii) 20,000 natural image patches of size 16$\times$16 pixel, that were randomly sampled from the first image set of the nature scene collection [1] (308 images of nature scenes containing no man-made objects or people). The basis patches learned by OSC are shown in Figure 1 and demonstrate that OSC reliably recovers the generating basis from synthetic data (see Figure 1a). Figure 1b illustrates that the OSC basis learned on the natural image patches resembles a wavelet decomposition, and is distinct from PCA, DCT, and Haar bases.

In Figure 2, the average $k$-term approximation performance of the OSC basis is compared with PCA, DCT, Haar and JPEG 2000 wavelets on the natural image patch data set. For this data set, OSC yields a consistently better $k$-term approximation performance than any of the alternative methods.

![Fig. 2: Average $k$-term approximation performance of 20,000 natural image patches of size 16$\times$16 pixel.](image)

**Acknowledgement.**

The research is funded by the DFG Priority Programme SPP 1527, grant number MA 2401/2-1.

**References**