Computing Sparse Representations in *O(NlogN)* time

May 3, 2013

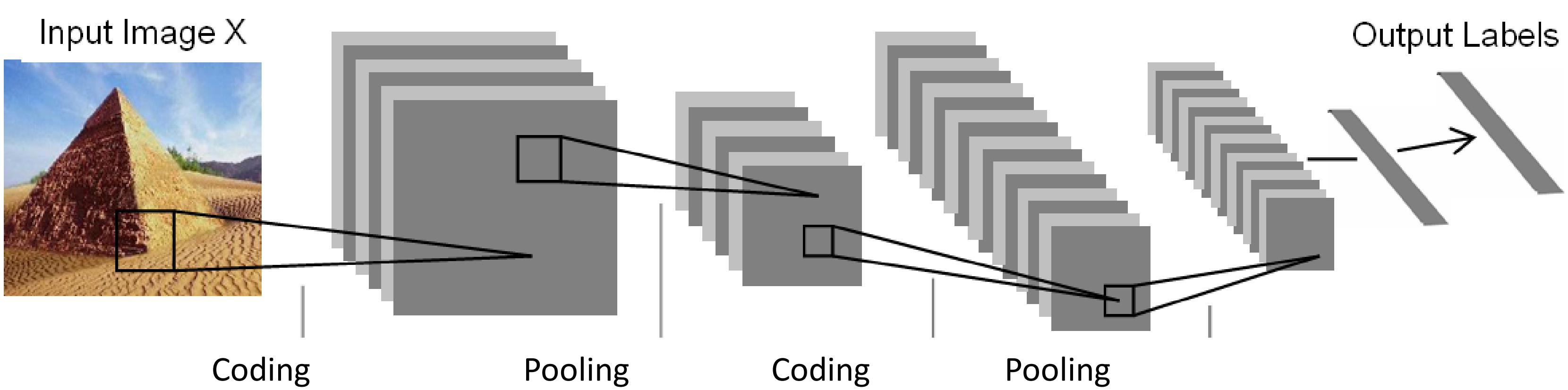
Tsung-Han Lin and H.T. Kung

Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS 2013), July 2013

HARVARD
School of Engineering and Applied Sciences

Hierarchical Feature Extraction

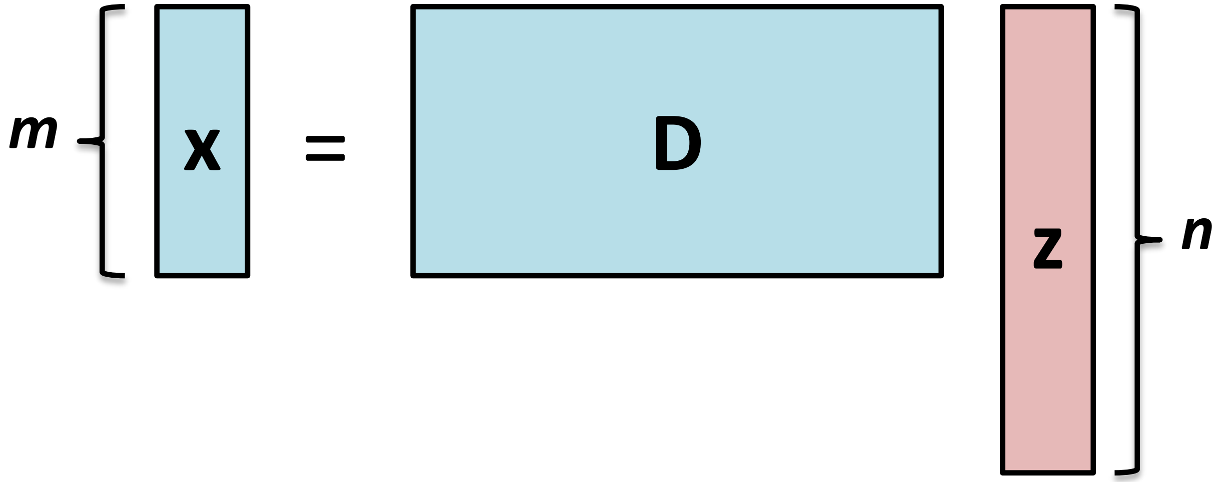
- Deep learning
  - Stack multiple feature extraction layers in hierarchy
  - Layer 1: find sparse representations of image patches
  - Layer 2: find sparse representations of layer-1 output



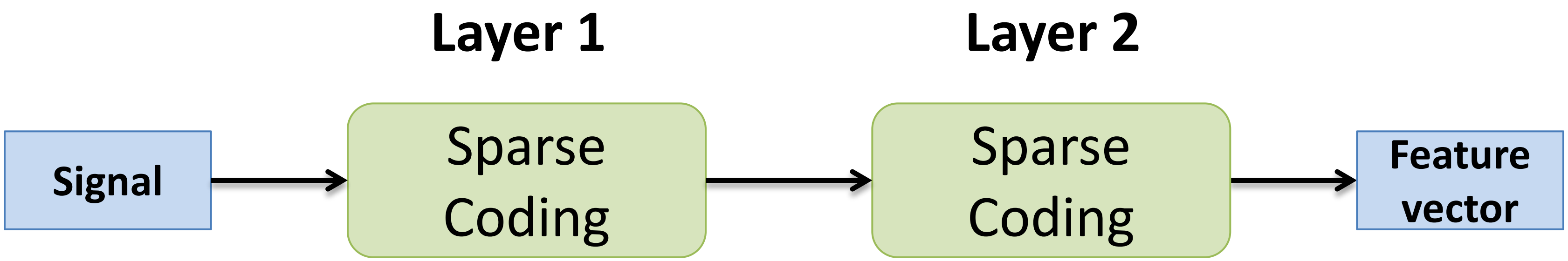
[Yu 2012]

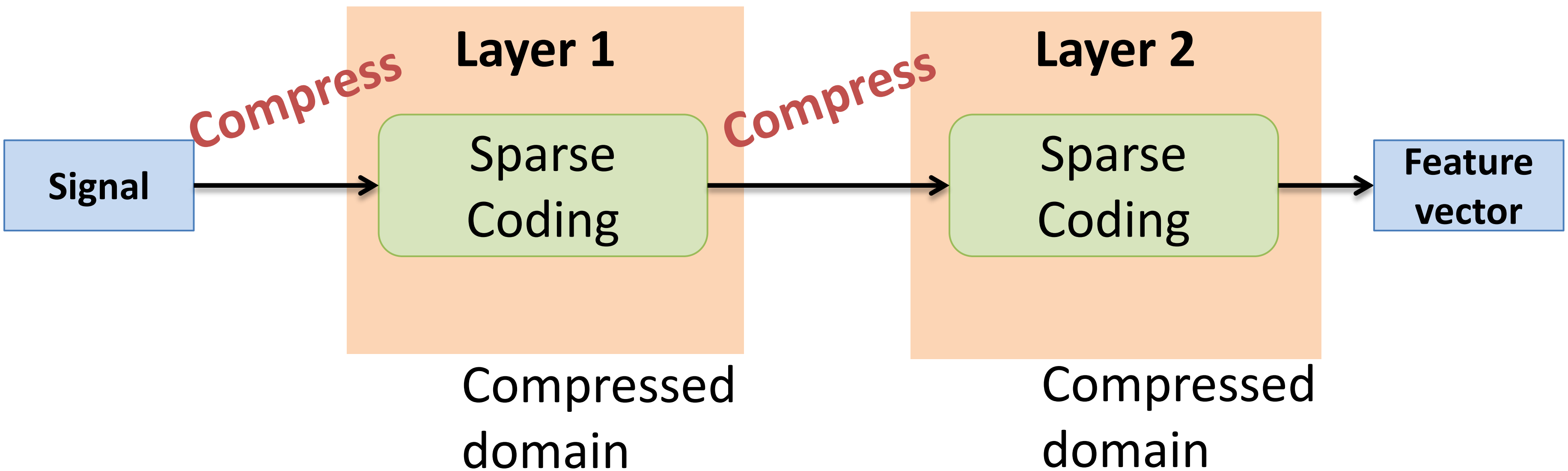
Computation Cost at a Feature Extraction Layer

- Complexity is ***O*(*mn*)**
  - *mx1* input signal **x** and *nx1* sparse code **z**
- *m* depends on the output code length in the previous layer, can be large in deeper layer
- *n* depends on dictionary size, governed by the machine learning task



Move Computations to Compressed Domain, i.e., Reducing *m*





How Much Can We Compress?

Compression by random projections make dictionary atoms less distinguishable
→ Compression ratio depends on the machine learning task, i.e., the dictionary size ***n***

Theorem. For a dictionary **D** that has *n* atoms, the input signal length *m* can be reduced to as small as $O(\log n/e^2)$, as long as **D** is sufficiently incoherent, or, the coherence *u* of the dictionary satisfies:

$$u < 1/(2K-1) - e$$

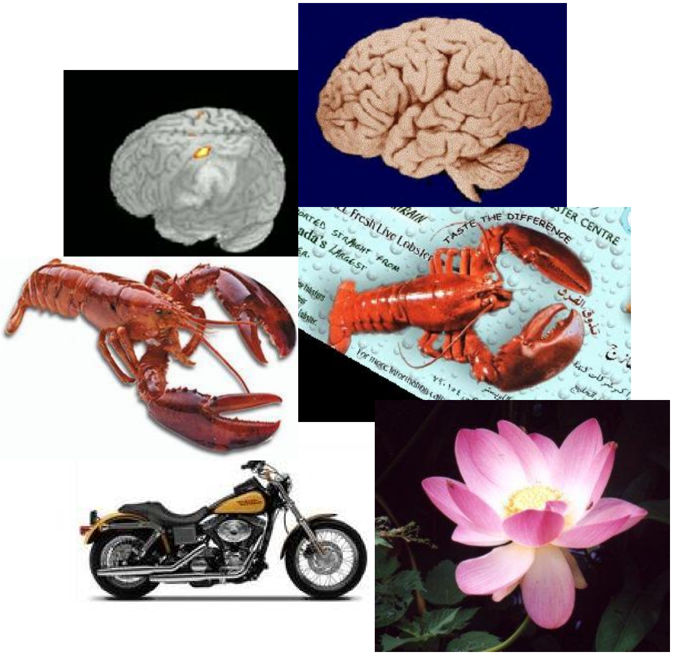
where *e* is a small positive number and *K* is the sparsity.

Experiments on Object Recognition

Recognition accuracy and run time

No compression D: 2268 x 1000	**2x** compression D: 1134 x 1000	**10x** compression D: 226 x 1000
59.9%	59.3%	56.7%
75.4 sec	40.3 sec	8.0 sec

- Two-layer sparse coding, compress second layer dictionary
- Test on Caltech-101, 101 object classes, 2945 images



Conclusion and Future Work

- The computations of deep learning can be performed in a low dimensional space
- Savings in # operations, meaning savings in energy and time
- Future work
  - Learning in the compressed domain
  - Novelty detection (afternoon)