Consensus Convolutional Sparse Coding

Biswarup Choudhury[1], Robin Swanson[2], Felix Heide[3], Gordon Wetzstein[3], Wolfgang Heidrich[1]

1. King Abdullah University of Science and Technology (KAUST)

2. University of Toronto 3. Stanford University

Abstract

Convolutional sparse coding (CSC) is a promising direction for unsupervised learning in computer vision. In contrast to recent supervised methods, CSC allows for convolutional image representations to be learned that are equally useful for high-level vision tasks and low-level image reconstruction and can be applied to a wide range of tasks without problem-specific retraining. Due to their extreme memory requirements, however, existing CSC solvers have so far been limited to low-dimensional problems and datasets using a handful of low-resolution example images at a time.

In this paper, we propose a new approach to solving CSC as a consensus optimization problem, which lifts these limitations. By learning CSC features from large-scale image datasets for the first time, we achieve significant quality improvements in a number of imaging tasks. Moreover, the proposed method enables new applications in high dimensional feature learning that has been intractable using existing CSC methods. This is demonstrated for a variety of reconstruction problems across diverse problem domains, including 3D multispectral demosaicking and 4D light field view synthesis.

Mathematical Framework

Classical CSC learning

$$\underset{d,z}{\text{argmin}} \quad \frac{1}{2}\sum_{j=1}^{n}(||b_j - DZ_j||_2^2 + \beta \sum_{w=1}^{W} ||z_{j_w}||_1) + \sum_{w=1}^{W} ind_c(d_w)$$

- b_j is our dataset used for learning ($j \in [1, n]$)
- z_{j_w} are the w sparse feature maps corresponding to element j
- For good feature learning, n is often very large!

- Partition data matrix $\mathbf{b}$ and feature maps Z as

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_j \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_N \end{bmatrix}, \quad Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_i \\ \vdots \\ Z_N \end{bmatrix}$$

- b_i: i^{th} block of data with respective Z_i ($i \in [1, N], N \leq n$)

- Can each block be handled independently? D?
- To facilitate independence of each block:

$$\underset{d_i, Z_i}{\text{argmin}} \quad \sum_{i=1}^{n} \frac{1}{2}||b_i - Z_i D_i||_2^2 + ||Z_i||_1 + ind_c(D_i)$$

- Multiple kernels D_i?

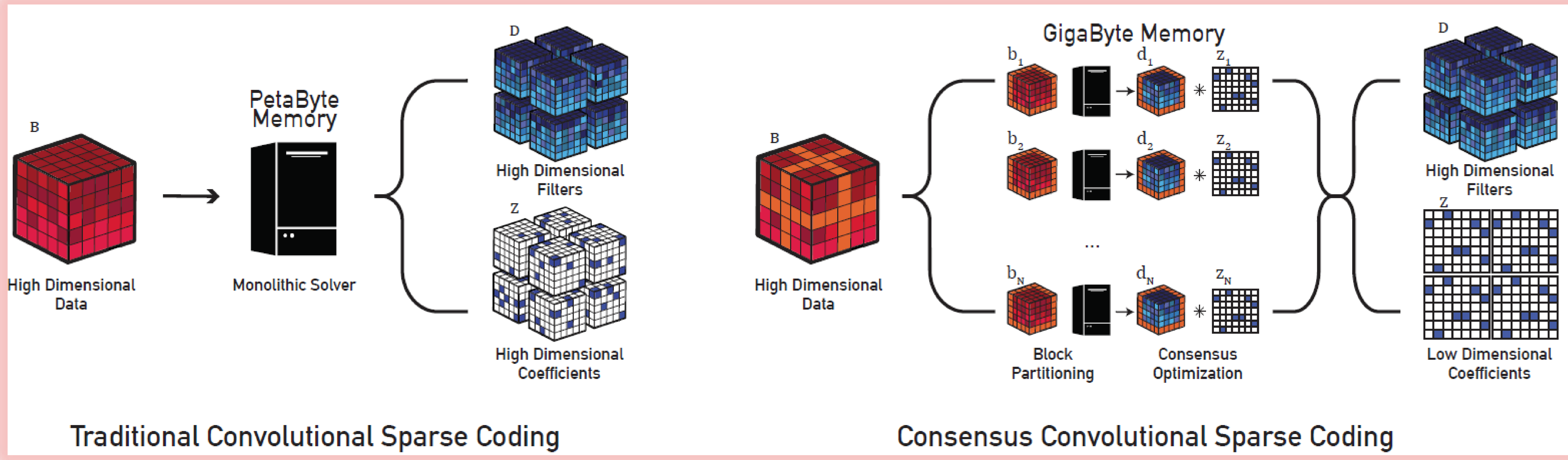
- Desire single kernel for entire data matrix $\mathbf{b}$
- Add equality constraint on the D_i

$$\underset{D_i}{\text{argmin}} \quad \frac{1}{2}||b_i - Z_i D_i||_2^2 + ||Z_i||_1 + ind_c(y)$$

$$\text{subject to} \quad D_i - y = 0, i = 1, \ldots, N$$

Global Consensus Form

- Consensus among local variables D_i to be equal
- Each equal to global variable y



Traditional Convolutional Sparse Coding

Consensus Convolutional Sparse Coding

Consensus CSC learning

ADMM Algorithm for d^{opt}

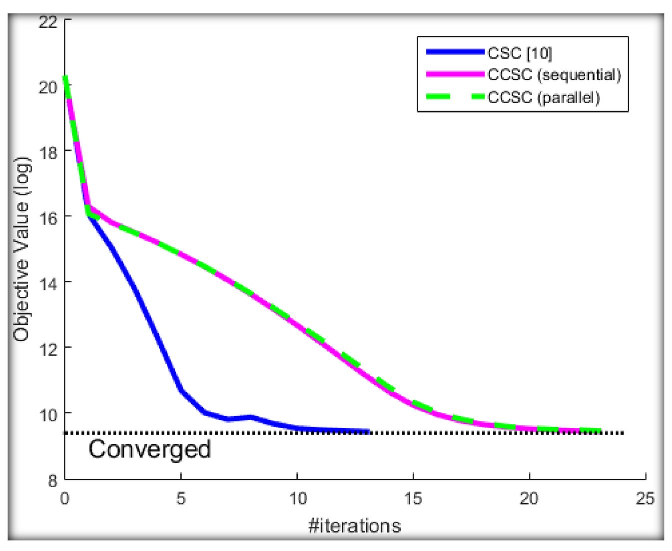
for all $k = 1$ to V **do**
  for all $i = 1$ to N **do**
    $D_i^{k+1} := \underset{D_i}{\text{argmin}}\Big((\frac{1}{2})\|\|b_i - Z_i D_i\|\|_2^2 + (\rho/2)\|\|D_i - y^k + u_i^k\|\|_2^2\Big)$
  end for
  $\overline{D}^{k+1} = \frac{1}{N}\sum_{i=1}^{N} D_i^{k+1}$.
  $\overline{u}^k = \frac{1}{N}\sum_{i=1}^{N} u_i^k$
  $y^{k+1} := \underset{y}{\text{argmin}}\Big(ind_c(y) + (N\rho/2)\|\|y - \overline{d}^{k+1} - \overline{u}^k\|\|_2^2\Big)$
  for all $i = 1$ to N **do**
    $u_i^{k+1} := u_i^k + D_i^{k+1} - y^{k+1}$
  end for
end for
$D^{opt} = y^{k+1}$

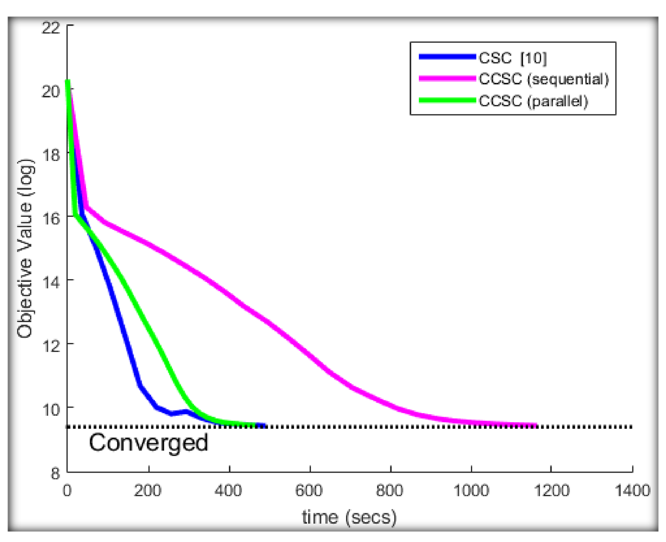
ADMM Algorithm for Z_i^{opt}

for all $i = 1$ to N **do**
  for all $k = 1$ to V **do**
    $Z_i^{k+1} := \underset{Z_i}{\text{argmin}}\Big((\frac{1}{2})\|\|b_i - D_i Z_i\|\|_2^2 + (\rho/2)\|\|Z_i - y_i^k + u_i^k\|\|_2^2\Big)$
    $y_i^{k+1} := \underset{y_i}{\text{argmin}}\Big(\beta\|y_i\|_1 + (\rho/2)\|\|y_i - Z_i^{k+1} - u_i^k\|\|_2^2\Big)$
    $u_i^{k+1} := u_i^k + Z_i^{k+1} - y_i^{k+1}$
  end for
  $Z_i^{opt} = Z_i^{k+1}$
end for

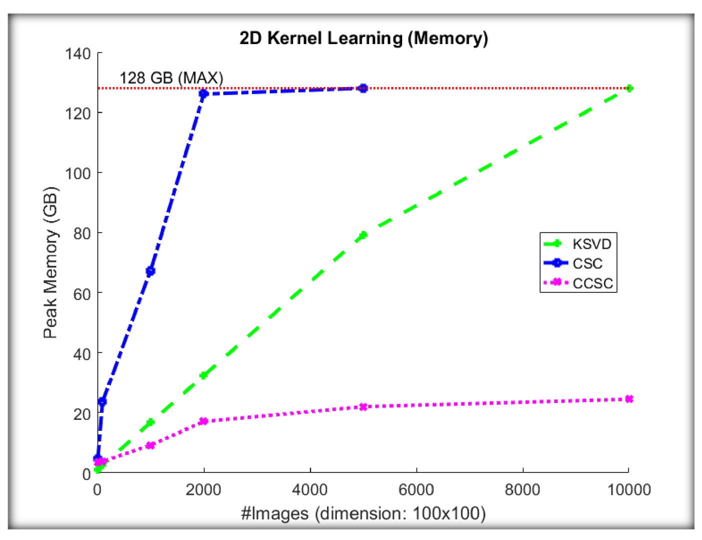
Evaluation

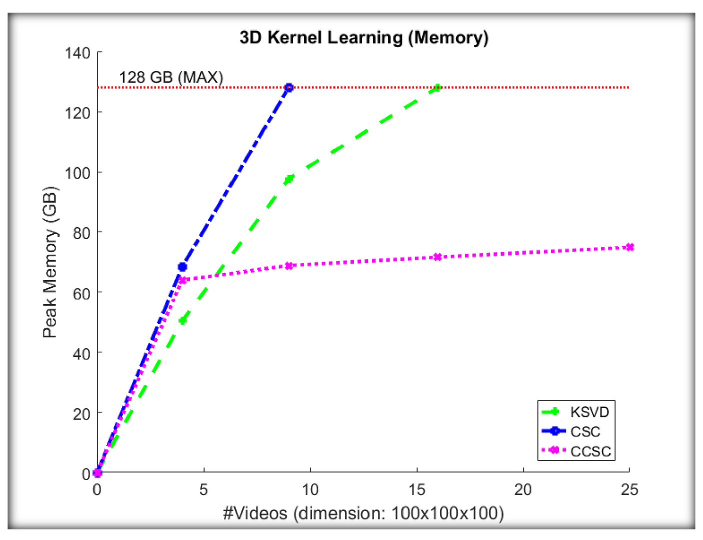
Convergence Analysis





Memory Analysis

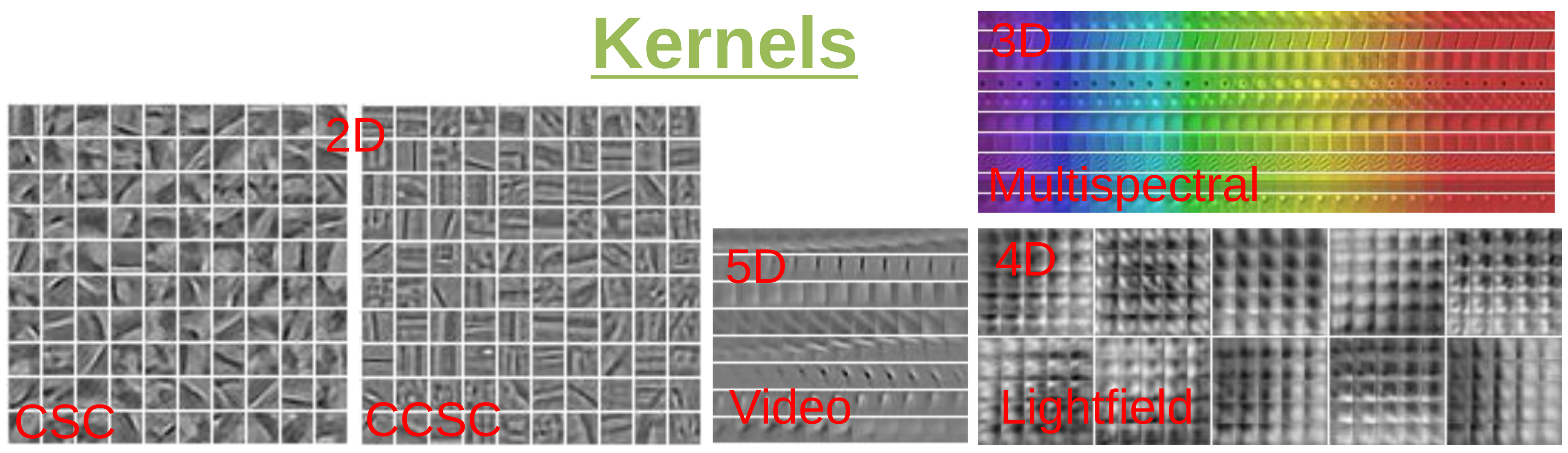




Runtime Analysis

2D Dataset Size	*Classical CSC (sec.)*	*CCSC (U = #PCs) (sec.)*		
		U=5	U=10	U=50
100	203.56	35.35	25.69	**25.20**
500	1530.71	259.30	82.69	**28.57**
1000	Out of Memory!	387.68	255.38	**35.63**

Kernels



Results

Lightfield Novel View Synthesis

In-Painting

Poisson Deconvolution