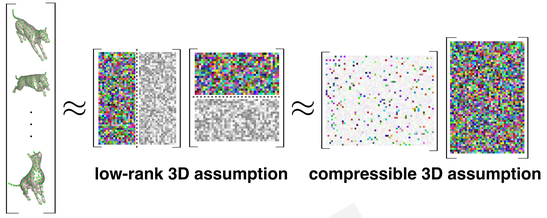
Prior-Less Compressible Structure from Motion

Chen Kong, Simon Lucey
{chenk, slucey}@cs.cmu.edu
Carnegie Mellon University

Introduction

- Re-interpret compressible SfM as a Block Sparse Dictionary Learning (BSDL) problem
- Theoretically characterize the *uniqueness* of BSDL
- Utilize the uniqueness of BSDL to recover the camera motion and 3D structures



NRSfM = BSDL?

$$\mathbf{S}^\sharp = \begin{bmatrix} x_{11} & y_{11} & z_{11} & \dots & x_{1P} & y_{1P} & z_{1P} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{F1} & y_{F1} & z_{F1} & \dots & x_{FP} & y_{FP} & z_{FP} \end{bmatrix} = \mathbf{C}^\sharp \, \mathbf{B}^\sharp$$

$$\mathbf{S} = \begin{bmatrix} x_{11} & \dots & x_{1P} \\ y_{11} & \dots & y_{1P} \\ z_{11} & \dots & z_{1P} \\ \vdots & \vdots & \vdots \\ x_{F1} & \dots & x_{FP} \\ y_{1P} & \dots & y_{FP} \\ z_{F1} & \dots & z_{FP} \end{bmatrix} = \left(\mathbf{C}^\sharp \otimes \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \right) \mathbf{B}$$

3-by-3 block sparse

dictionary

coefficients

$$\mathbf{W} = \begin{bmatrix} u_{11} & \dots & u_{1P} \\ v_{11} & \dots & v_{1P} \\ \vdots & \dots & \vdots \\ u_{F1} & \dots & u_{FP} \\ v_{F1} & \dots & v_{FP} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1 & & \\ & \ddots & \\ & & \mathbf{M}_F \end{bmatrix} \left[\mathbf{C}^\sharp \otimes \mathbf{I}_3 \right] \mathbf{B} = \left[\mathbf{\Pi} \right] \mathbf{B}$$

2-by-3 ($\mathbf{M}_1$) … 2-by-3 ($\mathbf{M}_F$); 3-by-3 block sparse; 2-by-3 block sparse

By assuming the 3D structure $\mathbf{S}^\sharp$ is compressible, we found that the 2D measurement $\mathbf{W}$ is 2-by-3 block compressible.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1526033. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.

The Uniqueness of BSDL

We show that if the dictionary $\mathbf{B}$ satisfies the block spark condition*, then any feasible block sparse factorization $\{\hat{\mathbf{\Pi}}, \hat{\mathbf{B}}\}$ of $\mathbf{W}$ satisfies that

$$\hat{\mathbf{\Pi}} = \mathbf{\Pi} \left(\text{Permutation Matrix} \otimes \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \right) \text{Block Diagonal Matrix}, \quad \hat{\mathbf{B}} = \text{Block Diagonal Matrix} \left(\text{Permutation Matrix} \otimes \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \right) \mathbf{B}$$

corrective matrix

A mix inside atoms

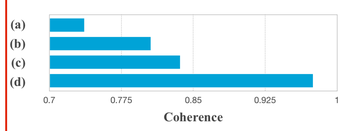
Change the ordering of atoms

*Check our paper for definitions and the the complete proof.

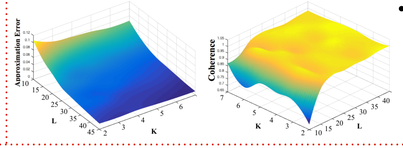
NRSfM

- Initialize BSDL by solving relaxed block ℓ_1-norm problem using Alternating Direction Method of Multipliers (ADMM).
- Solve BSDL by block K-SVD, block OMP, and block FOCUSS.
- Estimate corrective matrix.
- Estimate camera motion by enforcing camera consistency.
- Estimate 3D structure.
- Predict the confidence of the reconstruction by the coherence of the dictionary

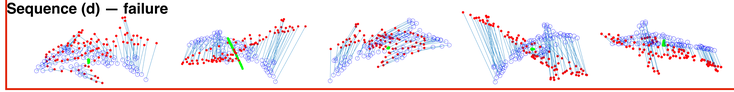
Incoherence



Control incoherence

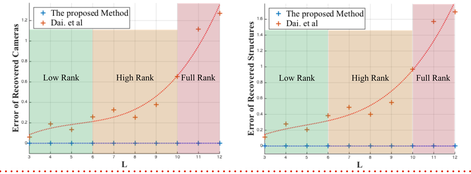


Sequence (d) — failure

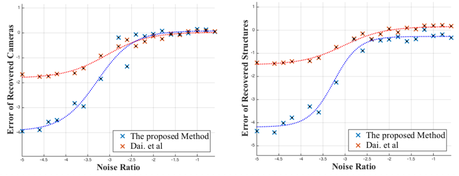


Experiments

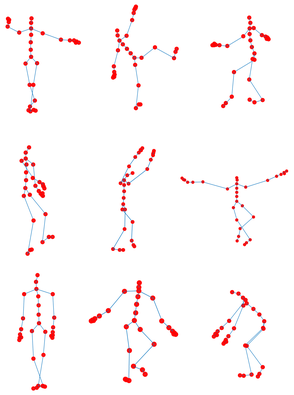
Performance vs. Rank of 3D structure



Performance vs. Noise



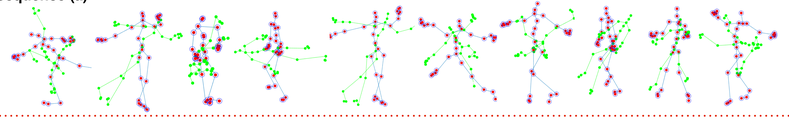
Selected recovered bases



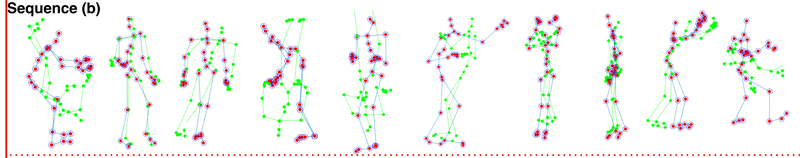
Visual performance

ground truth — The proposed — Dai. et.al

Sequence (a)



Sequence (b)



Sequence (c)