18CSE390T
Computer Vision

S4-Exploiting Sparsity

Exploiting Sparsity

- Large bundle adjustment problems, such as those involving 3D scenes from thousands of Internet photographs can require solving non-linear least square problems with millions of measurements
- Structure from motion is *bipartite* problem in structure and motion.
- Each feature point $\boldsymbol{x_{ij}}$ in a given image depends on one 3D point position $\boldsymbol{p_i}$ and 3D camera pose $\boldsymbol{(R_j,c_j)}$.

Exploiting Sparsity (cont).

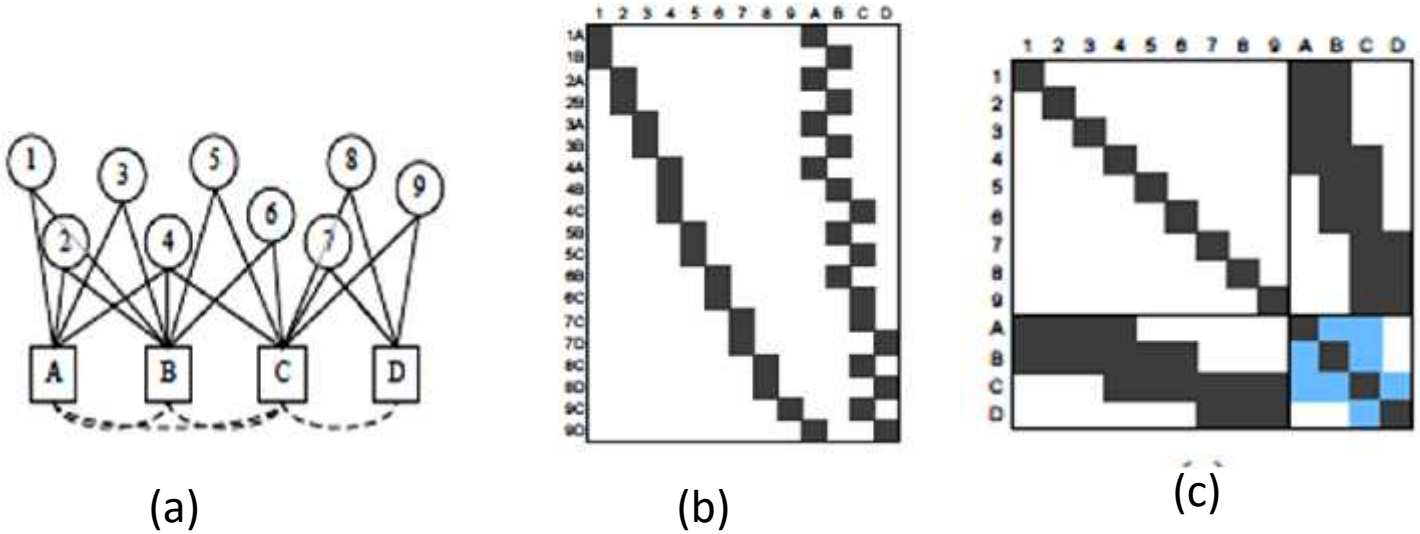

Figure: (a) Bipartite graph for a toy structure from motion problem and (b) its associated *Jacobian J* and (c) *Hessian A*. Numbers indicate cameras. The dashed arcs and light blue squares indicate the fill-in that occurs when the structure (point) variables are eliminated.

Uncertainty and Ambiguity

- Structure from motion involves the estimation of so many highly coupled parameters, often with no known "*ground truth*" components.
- The estimates produces by structure from motion algorithm can often exhibit large amounts of uncertainty .
- Example: *bas-relief ambiguity*, which makes it hard to simultaneously estimate 3D depth of scene and the amount of camera motion.

Uncertainty and Ambiguity (cont).

- A unique coordinate frame and scale for a reconstructed scene can not be recovered from monocular visual measurements alone.
- This seven-degrees-of-freedom gauge ambiguity makes it tricky to compute the variance matrix associated with a 3D reconstruction.
- To compute a convex matrix that ignores gauge freedom is to throw away the seven smallest eigenvalues of the information matrix, whose values are equivalent to the problem Hessian A up to noise scaling.

Reconstruction from Internet Photos

- Widely used application of structure from motion: the reconstruction of 3D objects and scenes from video sequences and collection of images.
- Before structure from motion comparison can begin, it is first necessary to establish sparse correspondences between different pairs of images and to then link such correspondences into *feature track,* which associates individual 2D image feature with global 3D points .

Reconstruction from Internet Photos (cont).

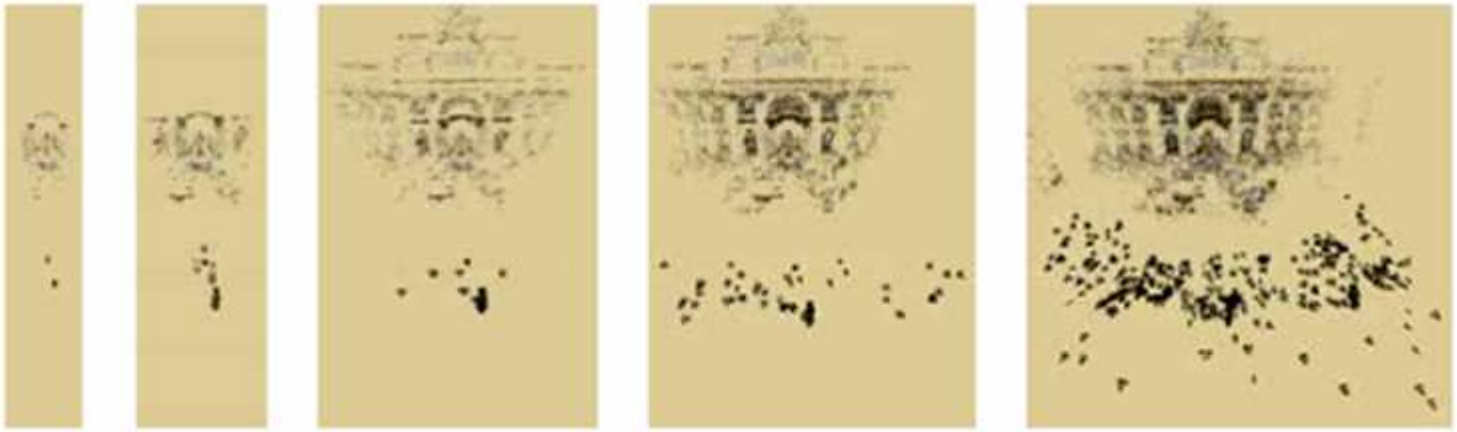

Figure: Incremental structure from motion: Starting with an initial two-frame reconstruction of Trevi Fountain, batches of images are added using pose estimation, and their positions (along with 3D model) are refined using bundle adjustment

Reconstruction from Internet Photos (cont).

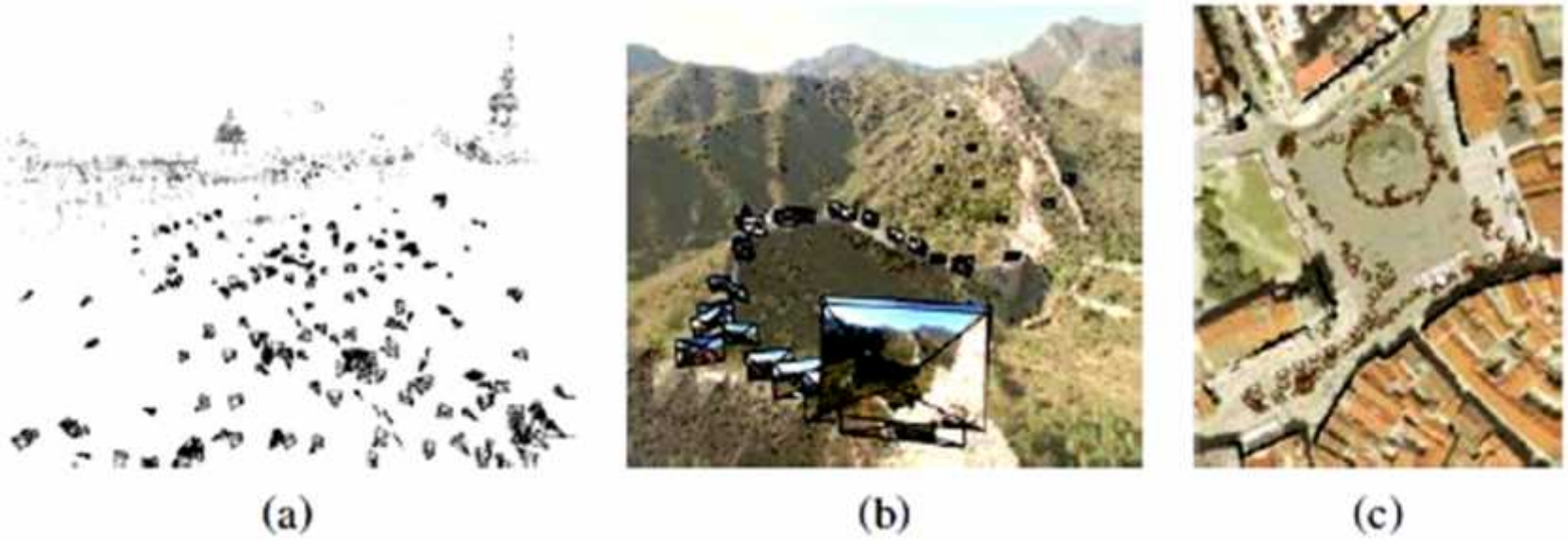

Figure: 3D reconstruction produced by the incremental structure from motion algorithm. (a) cameras and point cloud from Trafalgar Square; (b) cameras and points overlaid on an image from the Great Wall of China.; (c) overhead view of reconstruction of Old Town Square in Prague registered to an aerial photograph.

Reconstruction from Internet Photos (cont).

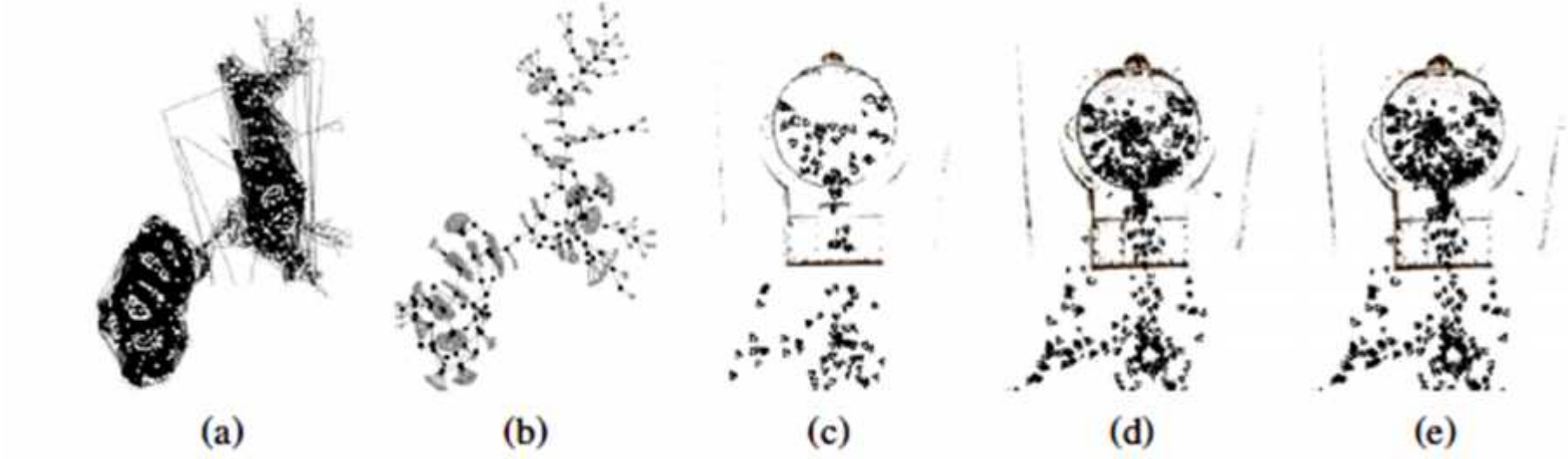

Figure: Large scale structure from motion using skeletal sets: (a) original match graph for 784 images; (b) skeletal set containing 101 images; (c) top-down view of scene (Pantheon) reconstructed from the skeletal set; (d) reconstruction after adding in the remaining images using pose estimation; (e) final bundle adjusted reconstruction, which is almost identical.