Some books on linear algebra

Finite Dimensional Vector Spaces, Paul R. Halmos, 1947
Linear Algebra, Serge Lang, 2004
Linear Algebra and its Applications, Gilbert Strang, 1988
Matrix Computation, Gene H. Golub, Charles F. Van Loan, 1996
Last lecture

- 2-Frame Structure from Motion
- Multi-Frame Structure from Motion

\[ T = C' - C \]

\[ p'^T Ep = 0 \]

\[ x'^T Fx = 0 \]
Today

• Continue on Multi-Frame Structure from Motion:
• Multi-View Stereo

Unknown camera viewpoints
Structure from Motion by Factorization
Problem statement
SFM under orthographic projection

\[ q = \Pi (p - t) \]

For example, \[ \Pi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \]

In general, \[ \Pi = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \]

subject to \[ \Pi \Pi^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]
SFM under orthographic projection

\[ q_n = \Pi (p_n - t) \]

- Choose scene origin to be the centroid of the 3D points
  \[ \sum p_n = 0 \]
- Choose image origin to be the centroid of the 2D points
  \[ \sum q_n = 0 \]

\[ q = \Pi p \]
factorization (Tomasi & Kanade)

projection of \( n \) features in one image:

\[
\begin{bmatrix}
q_1 & q_2 & \cdots & q_n
\end{bmatrix} = \prod_{2 \times n} \begin{bmatrix}
p_1 & p_2 & \cdots & p_n
\end{bmatrix}_{2 \times 3}^{3 \times n}
\]

projection of \( n \) features in \( m \) images

\[
\begin{bmatrix}
q_{11} & q_{12} & \cdots & q_{1n} \\
q_{21} & q_{22} & \cdots & q_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
q_{m1} & q_{m2} & \cdots & q_{mn}
\end{bmatrix}_{2m \times n} = \prod_{m} \begin{bmatrix}
\Pi_1 \\
\Pi_2 \\
\vdots \\
\Pi_m
\end{bmatrix}_{2m \times 3}^{3 \times n}
\]

\( W \) measurement \hspace{1cm} \( M \) motion \hspace{1cm} \( S \) shape

Key Observation: \( \text{rank}(W) \leq 3 \)
Factorization

- **Factorization Technique**
  - W is at most rank 3 (assuming no noise)
  - We can use *singular value decomposition* to factor W:

\[
W = M' S'
\]

- S’ differs from S by a linear transformation A:

\[
W = M' S' = (MA^{-1})(AS)
\]

- Solve for A by enforcing *metric* constraints on M
Metric constraints

- Enforcing “Metric” Constraints
  - Compute $A$ such that rows of $M$ have these properties

$$
\begin{bmatrix}
\Pi_1 \\
\Pi_2 \\
\vdots \\
\Pi_m
\end{bmatrix} = M = M' \ A = 
\begin{bmatrix}
\Pi_1' \\
\Pi_2' \\
\vdots \\
\Pi_m'
\end{bmatrix} \ A
$$

**Trick** (not in original Tomasi/Kanade paper, but in followup work)

- Constraints are linear in $AA^T$:

$$
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} = \Pi_i \Pi_i^T = \Pi_i' A (\Pi_i'^T A)^T = \Pi_i' G \Pi_i'^T \quad \text{where} \ G = AA^T
$$

- Solve for $G$ first by writing equations for every $\Pi_i$ in $M$
- Then $G = AA^T$ by SVD
Results
 Extensions to factorization methods

- Paraperspective [Poelman & Kanade, PAMI 97]
- Sequential Factorization [Morita & Kanade, PAMI 97]
- Factorization under perspective [Christy & Horaud, PAMI 96] [Sturm & Triggs, ECCV 96]
- Factorization with Uncertainty [Anandan & Irani, IJCV 2002]
Perspective Bundle adjustment
Bundle Adjustment

\[ \hat{u}_{ij} = f(K, R_j, t_j, x_i) \]
\[ \hat{v}_{ij} = g(K, R_j, t_j, x_i) \]

• How to initialize?
  • 2 or 3 views at a time, add more iteratively [Hartley 00]

• What makes this non-linear minimization hard?
  • many more parameters: potentially slow
  • poorer conditioning (high correlation)
  • potentially lots of outliers
Lots of parameters: sparsity

\[
\begin{align*}
\hat{u}_{ij} &= f(K, R_j, t_j, x_i) \\
\hat{v}_{ij} &= g(K, R_j, t_j, x_i)
\end{align*}
\]

- Only a few entries in Jacobian are non-zero

\[
\frac{\partial \hat{u}_{ij}}{\partial K}, \quad \frac{\partial \hat{u}_{ij}}{\partial R_j}, \quad \frac{\partial \hat{u}_{ij}}{\partial t_j}, \quad \frac{\partial \hat{u}_{ij}}{\partial x_i}
\]
Structure from motion: limitations

- Very difficult to reliably estimate *metric* structure and motion unless:
  - large \((x\) or \(y\)) rotation  \textit{or}
  - large field of view and depth variation
- Camera calibration important for Euclidean reconstructions
- Need good feature tracker
- Lens distortion

[Diagram of a smiley face with blue, green, and red dots on it]
Track lifetime every 50th frame of a 800-frame sequence
lifetime of 3192 tracks from the previous sequence
Track lifetime

track length histogram
Nonlinear lens distortion
Nonlinear lens distortion

effect of lens distortion
Prior knowledge and scene constraints

add a constraint that several lines are parallel
Prior knowledge and scene constraints

add a constraint that it is a turntable sequence
Applications of Structure from Motion
Jurassic park
PhotoSynth

“What if your photo collection was an entry point into the world, like a wormhole that you could jump through and explore...”

http://labs.live.com/photosynth/
Multiview Stereo
Choosing the stereo baseline

What’s the optimal baseline?

- Too small: large depth error
- Too large: difficult search problem
The Effect of Baseline on Depth Estimation

Figure 2: An example scene. The grid pattern in the background has ambiguity of matching.
pixel matching score
Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

Fig. 6. Combining two stereo pairs with different baselines.

Fig. 7. Combining multiple baseline stereo pairs.
Multibaseline Stereo

Basic Approach

- Choose a reference view
- Use your favorite stereo algorithm BUT
  > replace two-view SSD with SSD over all baselines

Limitations

- Must choose a reference view (bad)
- Visibility!
MSR Image based Reality Project

http://research.microsoft.com/~larryz/videoviewinterpolation.htm
The visibility problem

Which points are visible in which images?

**Known Scene**

**Unknown Scene**

**Forward Visibility**

**Inverse Visibility**
Volumetric stereo

Scene Volume $V$

Input Images (Calibrated)

Goal: Determine occupancy, “color” of points in $V$
Discrete formulation: Voxel Coloring

Discretized Scene Volume

Input Images (Calibrated)

Goal: Assign RGBA values to voxels in \( V \) \textit{photo-consistent} with images
Complexity and computability

Discretized Scene Volume

$N^3$ voxels

$C$ colors

All Scenes ($C^{N^3}$)

True Scene

Photo-Consistent Scenes
Issues

Theoretical Questions
  • Identify class of *all* photo-consistent scenes

Practical Questions
  • How do we compute photo-consistent models?
Voxel coloring solutions

1. C=2 (shape from silhouettes)
   - Volume intersection [Baumgart 1974]
     > For more info: Rapid octree construction from image sequences. R. Szeliski, CVGIP: Image Understanding, 58(1):23-32, July 1993. (this paper is apparently not available online) or
     > W. Matusik, C. Buehler, R. Raskar, L. McMillan, and S. J. Gortler, Image-Based Visual Hulls, SIGGRAPH 2000 (pdf 1.6 MB)

2. C unconstrained, viewpoint constraints
   - Voxel coloring algorithm [Seitz & Dyer 97]

3. General Case
   - Space carving [Kutulakos & Seitz 98]
Reconstruction from Silhouettes (C = 2)

Approach:
- Backproject each silhouette
- Intersect backprojected volumes
Volume intersection

Reconstruction Contains the True Scene

- But is generally not the same
- In the limit (all views) get visual hull
  > Complement of all lines that don’t intersect S
Voxel algorithm for volume intersection

Color voxel black if on silhouette in every image

- $O(\ ? \ )$, for $M$ images, $N^3$ voxels
- Don’t have to search $2^{N^3}$ possible scenes!
Properties of Volume Intersection

Pros

- Easy to implement, fast
- Accelerated via octrees [Szeliski 1993] or interval techniques [Matusik 2000]

Cons

- No concavities
- Reconstruction is not photo-consistent
- Requires identification of silhouettes
Voxel Coloring Solutions

1. C=2 (silhouettes)
   • Volume intersection [Baumgart 1974]

2. C unconstrained, viewpoint constraints
   • Voxel coloring algorithm [Seitz & Dyer 97]

3. General Case
   • Space carving [Kutulakos & Seitz 98]
Voxel Coloring Approach

1. Choose voxel
2. Project and correlate
3. Color if consistent
   (standard deviation of pixel colors below threshold)

Visibility Problem: in which images is each voxel visible?
Depth Ordering: visit occluders first!

Condition: depth order is the same for all input views
Panoramic Depth Ordering

- Cameras oriented in many different directions
- Planar depth ordering does not apply
Panoramic Depth Ordering

Layers radiate outwards from cameras
Panoramic Layering

Layers radiate outwards from cameras
Panoramic Layering

Layers radiate outwards from cameras
Compatible Camera Configurations

Depth-Order Constraint

• Scene outside convex hull of camera centers

Inward-Looking

Outward-Looking
Calibrated Image Acquisition

Calibrated Turntable

Selected Dinosaur Images

Selected Flower Images
Voxel Coloring Results (Video)

Dinosaur Reconstruction
72 K voxels colored
7.6 M voxels tested
7 min. to compute
on a 250MHz SGI

Flower Reconstruction
70 K voxels colored
7.6 M voxels tested
7 min. to compute
on a 250MHz SGI
Limitations of Depth Ordering

A view-independent depth order may not exist

\[ \text{need more powerful general-case algorithms} \]

- Unconstrained camera positions
- Unconstrained scene geometry/topology
Voxel Coloring Solutions

1. C=2 (silhouettes)
   • Volume intersection [Baumgart 1974]

2. C unconstrained, viewpoint constraints
   • Voxel coloring algorithm [Seitz & Dyer 97]

3. General Case
   • Space carving [Kutulakos & Seitz 98]
Space Carving Algorithm

- Initialize to a volume $V$ containing the true scene
- Choose a voxel on the current surface
- Project to visible input images
- Carve if not photo-consistent
- Repeat until convergence
Which shape do you get?

The **Photo Hull** is the UNION of all photo-consistent scenes in $V$

- It is a photo-consistent scene reconstruction
- Tightest possible bound on the true scene
Space Carving Algorithm

The Basic Algorithm is Unwieldy

- Complex update procedure

Alternative: Multi-Pass Plane Sweep

- Efficient, can use texture-mapping hardware
- Converges quickly in practice
- Easy to implement
Multi-Pass Plane Sweep

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence
**Multi-Pass Plane Sweep**

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence
Multi-Pass Plane Sweep

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence
Multi-Pass Plane Sweep

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence
Multi-Pass Plane Sweep

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence
Multi-Pass Plane Sweep

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence
Space Carving Results: African Violet

Input Image (1 of 45)

Reconstruction

Reconstruction

Reconstruction

Reconstruction
Space Carving Results: Hand

Input Image (1 of 100)

Views of Reconstruction
Properties of Space Carving

Pros

- Voxel coloring version is easy to implement, fast
- Photo-consistent results
- No smoothness prior

Cons

- Bulging
- No smoothness prior
Alternatives to space carving

Optimizing space carving

- recent surveys
  - Slabaugh et al., 2001
  - Dyer et al., 2001
- many others...

Graph cuts

- Kolmogorov & Zabih

Level sets

- introduce smoothness term
- surface represented as an implicit function in 3D volume
- optimize by solving PDE’s
Alternatives to space carving

Optimizing space carving

• recent surveys
  > Slabaugh et al., 2001
  > Dyer et al., 2001
• many others...

Graph cuts

• Kolmogorov & Zabih

Level sets

• introduce smoothness term
• surface represented as an implicit function in 3D volume
• optimize by solving PDE’s
Level sets vs. space carving

Advantages of level sets
- optimizes consistency with images + smoothness term
- excellent results for smooth things
- does not require as many images

Advantages of space carving
- much simpler to implement
- runs faster (orders of magnitude)
- works better for thin structures, discontinuities

For more info on level set stereo:
- Renaud Keriven’s page:
  > http://cermics.enpc.fr/~keriven/stereo.html
References

Volume Intersection


Voxel Coloring and Space Carving


• Recent surveys

Other references from this talk