CS 395T
Numerical Optimization for Graphics and AI (3D Vision)

Qixing Huang
August 29th 2018
3D Vision

• Understanding geometric relations
  – between images and the 3D world
  – between images

• Obtaining 3D information describing our 3D world
  – from images
  – From dedicated sensors

• Geometry understanding (new addition to 3D Vision)
  – Due to the emergence of big 3D data
  – Semantics for example
Virtual tourism

Photo Tourism
Exploring photo collections in 3D

Noah Snavely Steven M. Seitz Richard Szeliski
University of Washington Microsoft Research

SIGGRAPH 2006
3D Urban Modeling
Kinect Scanner
KinectFusion

SIGGRAPH Talks 2011

KinectFusion:
Real-Time Dynamic 3D Surface Reconstruction and Interaction

Shahram Izadi 1, Richard Newcombe 2, David Kim 1,3, Otmar Hilliges 1,
David Molyneaux 1,4, Pushmeet Kohli 1, Jamie Shotton 1,
Steve Hodges 1, Dustin Freeman 5, Andrew Davison 2, Andrew Fitzgibbon 1

1 Microsoft Research Cambridge  2 Imperial College London
3 Newcastle University           4 Lancaster University
5 University of Toronto
DynamicFusion

DynamicFusion:
Reconstruction & Tracking of Non-rigid Scenes in Real-Time

Richard Newcombe, Dieter Fox, Steve Seitz

Computer Science and Engineering,
University of Washington
Reconstructing of the Octagon Monument
Finding matching fragments is a tedious job
3D reconstruction of fragments
3D reconstruction and completion
Visual impact of the 3D restoration
3D ShapeNets

[Wu et al. 15]
PointNet

[Qi et al. 17]
Many other applications

- Augmented Reality
- Self-Driving Cars
- Human-machine interface
- Autonomous micro-helicopter navigation
- Performance capture
- Forensics
- Interactive 3D modeling
Image-based modeling

Modeling and Rendering Architecture from Photographs
a hybrid geometry- and image-based approach

Paul E. Debevec
Camillo J. Taylor
Jitendra Malik

University of California at Berkeley
January, 1996
Interdisciplinary aspect

- Vision
- Biology
- 3D Vision
- Architecture
- Medicine
- Robotics
- Geometry
- Processing
Topics to be Covered
Topic I: 3D Reconstruction

• Geometry
  – Epipolar geometry
  – Fundamental matrix
  – Extrinsic/Intrinsic camera parameters
  – Camera calibration
  – Vanishing points
  – Homogeneous coordinates
Topic I: 3D Reconstruction

• Algorithms
  – Feature extraction
  – Feature correspondences
  – Relative camera pose
  – Structure-from-motion
  – Multiview stereo
  – Bundle adjustment
  – ICP
Topic II: How to represent 3D Data

- Triangular mesh
- Implicit surface
- Light Field Representation
- Part-based models
- Point cloud
Conversion between different representations

• Implicit -> mesh (Marching Cube)
Conversion between different representations

- Point cloud -> triangular mesh (Delaunay triangulation)
Conversion between different representations

- Pointcloud -> Implicit -> Mesh

[Kazhdan et al. 06]
Two recommended books
Topic III: How to understand 3D Data

- Design algorithms to extract semantic information from one or a collection of shapes

  - Segmentation
  - Matching
  - Classification & Clustering

  [van Kaick et al. 11]
  [Karz and Tal 03]
  [Mitra et al. 06]
  [Funkhouser et al. 05]
Big 3D Data
Princeton Shape Benchmark

1800 models in 90 categories

Shilane et al. 04
Large-scale online repositories

3D Warehouse

Yobi3D

3M models in more than 4K categories
Image-based shape retrieval

20 years ago

10 years ago

now
Trends
Trends in Geometry Understanding

Single Shape/Scene Analysis

Joint Data Analysis (Unsupervised)  ML on 3D Data (Supervised)
Numerical Optimization
Optimization problems

We will see time and time again:

Translate

Real problem

into

Optimization problem

Examples in Vision/Robotics/NLP

/ML/Graphics/

Example of the contrary?

• This course: how to formulate $P$, how to solve $P$, and what are the guarantees

Many similar domains: Graphics/Robotics/Vision
Why bother how to solve P and what are the guarantees

- There are plenty of optimization softwares
  
  **Solvers:** CPLEX, Mosek, Gurobi, ECOS, Clp, Knitro, Ipopt,...

  **Modeling languages:** YALMIP, CVX, GAMS, AMPL, JuMP,...

- Almost all algorithms are data-dependent and can perform better or worse on different problems and data sets

- In many cases, studying P leads to new algorithms and... research papers
Categories of optimization models

• Linear vs. Nonlinear
• Convex vs. Nonconvex
• Continuous vs. Discrete
• Deterministic vs. Stochastic

We see all of them in 3D Vision tasks
Example I: Iterative Closest Point (or ICP)

- Multi-view
- Coordinate system alignment
- Optimization problem

$$\min_{\mathcal{T} \in SE(3)} d^2(P_T, \mathcal{T}(P_S))$$

- Performance
  - Local/global convergence
  - Convergence rate
  - Uncertainty

ICP [Besel and Mckay’ 92]
Example II: MRF Inference

\[ \arg \min_{\vec{w}} \sum_{n=1}^{N} U_n(w_n) + \sum_{(m,n) \in \mathcal{C}} P_{mn}(w_m, w_n), \]

**Unary terms**  
(compatibility of data with label \( w \))  
**Pairwise terms**  
(compatibility of neighboring labels)

The literature reflects almost all advances in optimization during the past decade:

Graphcut  
Coordinate descent

Linear programming relaxation  
Dual coordinate descent

Quadratic programming relaxation  
Stochastic gradient descent

Semidefinite programming relaxation  
Block-coordinate descent
Trends in optimization

Non-linear optimization

Convex optimization

Non-convex optimization
Trends in optimization

Second-order methods

First-order methods

Distributed optimization
Grading
Five assignments (70%)

• Theory + practice (Choose one of them)
• Theory
  – 3D geometry (for formulating numerical optimization problems)
  – Proof of local/global convergence

• Practice
  – Programming assignments of real computer vision tasks
Final Project (30%)

• Work in groups of 2-4 people

• Publishable results
  – ICCV/ICML/NIPS/SIGGRAPH/SIGGRAPH ASIA

• Written report
  – 8 pages
  – Introduction/related works/approach/results/conclusions

• Final presentation
Potential project topic I

- Potential topic I: Image-based modeling from internet images
  - A few images
  - Non-identical objects
  - Symmetries
  - Machine learning
Potential project topics

• Potential topic II: single-view reconstruction from real images
  – 3D Representation?
  – Network architecture
  – Domain adaptation
  – Depth/semantic labels
  – Human
  – Hand

[Song et al. 17]
Potential project topics

• Potential topic III: neural networks for 3D representations
  – Implicit surface
  – Multi-view
  – Point cloud
  – Mesh?
  – Arrangement?
  – Scene graph?
  – Hybrid?

[Wu et al. 16]
Potential project topics

• Potential topic IV: Geometry understanding
  – Task
    • Normal/Curvature
    • Feature extraction
    • Segmentation
    • Part/object decomposition
    • Correspondence
    • Affordance
  – Data
  – Thousands of categories
Potential project topics

• Potential topic V: Reconstruction from hybrid sensors
  – Low-res depth scan
    + High-res stereo
  – Depth + shading
  – Web cameras
    + internet images
  – 360 images
    + internet images
  – Street views
    + images from drones
Potential project topics

• Potential topic VI: Robot 3D Vision
  – Salient views
  – Next-best-view
  – Model-based View planning
  – Active vision
  – Policies for exploration
Potential project topics

• Potential topic VII: structure recovery via optimization
  – 2D human pose -> 3D human pose
  – Structure from symmetry
  – Structure from template
  – Shape from shading
  – Local/global convergence
  – Exact recovery condition

[Zhou et al. 16]
Potential project topics

• Potential topic VIII: Synchronization
  – Multiview structure from motion
    • Pose
    • Pose + Point cloud
  – Geometric alignment of point clouds
    • Pose
    • Pose + shape
  – Global/Local convergence
  – Exact recovery condition
Potential project topics

• Potential topic IX: uncertainties
  – Human pose estimation
  – Depth prediction
  – Camera poses in MVSFM
  – Geometry reconstruction
  – Geometry understanding
  – Quantification/Visualization
Logistics

• Office hour: Wednesdays 5:00 pm – 7:00 pm

• TA: Xiangru Huang
  – Office hour: TBD

• Homeworks
  – Six late days

• Final project
  – No late day
Questions?