Image Registration and 3D Reconstruction in Computer Vision

Adrien Bartoli et al.
Clermont Université, France
Image Understanding

Object detection

Object recognition

View morphing

Photogrammetry

Rendering

Augmented reality

Special effects, motion capture, ...
“Vision is the act of knowing what is where by looking.”
- Aristotle
Computer Vision

Does not study biological vision

Uses simple models

Handles various types of camera

From [Baker et al., CVPR’01]
3D Computer Vision: Shape-from-X

- Shading
- Focus
- Shadows
- Occlusions, silhouettes, ...
- Texture
Techniques with Multiple Views

Image formation

Photometric stereo

Structure-from-Motion - SfM

From [Higo et al. ICCV’09]
SfM – Structure-from-Motion
SfM has Two Main Steps

Image registration

3D reconstruction
Rigid vs Deformable SfM

Image registration 3D reconstruction

Image registration 3D reconstruction
Why is Registration Difficult?
Rigid SfM: A Mature Paradigm

Inputs: a set of images

1 – registration

2 – projective reconstruction

3 – camera self-calibration

4 – surface reconstruction

Outputs: camera parameters and scene structure
Keypoint Detection and Matching

- Feature-based methods: abstract the images by features
- Keypoints lie at local maxima in intensity variations
- They should be stable under change of viewpoint
- They are matched by comparing invariant descriptors
Robust Estimation Methods

• Objective: estimate the model parameters and classify each datum as valid or erroneous
• This is a chicken-and-egg problem
• RANSAC (Random Sample Consensus)
  – Handles more than 50% outliers in the data
  – Basic idea: sample minimal data sets and maximize the support

Inliers
Outliers
Homography based panorama
Rigid SfM: A Mature Paradigm

Inputs: a set of images

1 – registration

2 – projective reconstruction

3 – camera self-calibration

4 – surface reconstruction

Outputs: camera parameters and scene structure
The Projective Camera

Image point

Centre of projection

Retina point

Image point

3D point

Centre of projection

Intrinsics

Extrinsics

3D point

Rational in euclidean coordinates

Linear in homogeneous coordinates
Projective Geometry
3D Reconstruction

• The basic constraint: intersection of the viewing rays for matching points
Projective Reconstruction

• Nonlinear optimization problem
• Initialization
  – Split the image set
  – Reconstruct from 2 – 4 views
  – Merge the reconstructions
• Refinement: bundle adjustment
Camera Self-Calibration

- Uses constraints on the intrinsics, such as
  - Constant intrinsics for all cameras
  - Known intrinsics, except the (varying) focal length
- Linear or nonlinear least squares
- Why using camera self-calibration?
  - Flexibility: no prior knowledge on structure and extrinsics
  - Simplicity: projective is easier than metric
- Bundle adjustment
Bundle Adjustment

- Minimization of the reprojection error
- Thousands of unknowns
- Exploits the problem structure
Rigid SfM: A Mature Paradigm

Inputs: a set of images

1 – registration

2 – projective reconstruction

3 – camera self-calibration

4 – surface reconstruction

Outputs: camera parameters and scene structure
Dense 3D Reconstruction

• Based on geometric primitives (planes, cubes, etc)

Keypoint detection and matching

3D reconstruction

Piecewise planar segmentation
Dense 3D Reconstruction

- Stereopsis, voxel colouring
- Based on matching colours between the images

Results from [Boykov et al, PAMI’04]

Results from [Furukawa et al, CVPR’07]
Augmentation
Deformable Structure-from-Motion

- The viewing rays for matching points do not generally meet
- The problem is generally ill-posed
Deformable Structure-from-Motion

• Priors on the structure
  – Single object
  – Empirical: smoothness, \textit{etc}
  – Physical: inextensibility, developability, mechanical models, \textit{etc}
  – Statistical: face, body, \textit{etc}

Image registration $\rightarrow$ 3D reconstruction
Deformable Surface Registration

- Find transformations that make the input images similar
- Geometric transformation: changes pixel locations
- Photometric transformation: changes pixel colours
- Sequential registration
Deformable Surface Registration

Current image

Template

Transformed image

Difference image

Unknown transformation

$\rightarrow$ minimization over the transformation parameters
Explicit Photometric Transformation

Template

Current image

Videos of the difference image

No photometric model

Affine photometric model
Other Registration Results and Methods

Curve-based registration

Retexturing
Other Registration Results and Methods

- Sea breeze tracking
- Computation of mechanical constants
- Fluid flow registration
Retargetting
Deformable Surface 3D Reconstruction

- Highly dependent on the surface (cloth, paper, skin, etc)
- Example algorithm for isometric surfaces
1. Compute the upper bounds for each pair of points
2. Keep the tightest upper bounds
3. Recompute the bounds and loop to 2
4. Interpolate the upper bounds to get a surface
Summary for Rigid Structure-from-Motion

- Mature for textured environments
- Camera and ‘sparse’ structure recovery, self-calibration
- Companies: 2d3 (University of Oxford then Vicon), RealViz (INRIA then Autodesk), etc
- Textbooks: [Hartley et al, 03 ; Faugeras et al, 01 ; Forsyth et al, 03]
Summary for Deformable Structure-from-Motion

• Some results in the ‘surface’ case for registration and 3D reconstruction
• Image registration for fluid flow
• The general case is completely open
• Prior knowledge on the environment is required
Bibliography

• Keypoint detection, description and matching: [Lowe, IJCV’04 ; Mikolajczyk et al, IJCV’05]
• Robust estimation (RANSAC): [Fischler et al, CVGIP’81]
• Projective reconstruction: [Nister, PhD-Thesis’00]
• Bundle adjustment: [Triggs et al, VA’00]
• Self-calibration: [Gurdjos et al, ICCV’09]
• (Rigid) dense stereo: [Boykov et al, PAMI’04 ; Strecha et al, CVPR’08]
• Deformable surface registration: [Gay-Bellile et al, PAMI’10; Pilet et al, IJCV’07]
• Deformable surface reconstruction: [Perriollat et al, IJCV’10; Salzmann et al, ICCV’09]
Image Registration and 3D Reconstruction in Computer Vision

Adrien Bartoli et al.
Clermont Université, France