Image-Based Localisation

Roberto Cipolla

Duncan Robertson and Ben Tordoff
Department of Engineering
Background
User takes a picture of a nearby building. System tells you what you are looking at and exactly where you are on a map.
The problem
Why difficult?

- Extreme perspective distortion
- Differences in colour / lighting conditions
- Occlusion
Unconstrained matching

326 matches (score 57.2) 373 matches (score 51.2)
Constrained matching

- Building façades are roughly planar
- They contain many horizontal and vertical features
- We can use this to get a "front view" (rectified image)
- Front-views are related by translation and scale only
Constrained matching
Constrained matching
Constrained matching
Constrained matching
Constrained matching
Overview of solution

1. vanishing point detection
2. image rectification
3. database search
4. viewpoint determination
Rectification
Rectification
Detection of straight lines

Detect straight lines:
Finding vanishing points
Find vertical and horizontal lines

Allocate all lines as vertical, horizontal or “clutter”
Rectification by homography
Align horizon

Only difference is now scale + x translation
Matching
Matching

Input
- Image
- Corner features 1
- Corner features 2

Process
- Pyramid
- Harris detector
- Patch descriptors
- Constrained correlation matching
- RANSAC

Output
- Corner features
- Alignment transform
Matching
Matching

With only 2 params \((s, t_x)\), can search rather than RANSAC.
Matching

With only 2 params \((s, t_x)\), can search rather than RANSAC.
Matching

With only 2 params \((s, t_x)\), can search rather than RANSAC.
Examples over wide baselines
Summary of matching
Camera pose estimation - localisation
Localisation

Input

Rectified image
Map image
Alignment transform
Rectifying transform
Map transform
Camera calibration

Process

4 click registration
Combine transforms
Find viewpoint

Output

Map transform
User Location
Register database view

First align database view to map
Localisation

Knowing the rectifying homography ($H_{\perp}$), the alignment ($H_A$), and the database view registration, can work backwards to find user:

Rectifying rotation $R_{\perp}$ gives the angle from perpendicular and focal length the distance to camera.
Localisation of query view
Localisation

Summary:

- Using geometric information generic matching is reduced to a 2 DOF search problem
- We are also able to find the camera (ie user) position and orientation
Evaluation
Evaluation
Evaluation
Image-based localisation
Image-based localisation
Image-based localisation
Image-based localisation
Image-based localisation
Conclusions

- Effective wide baseline matching and image registration
- Mobile phone localisation:
  - Where am I?
  - What am I looking at?
- Scaling up to real applications?
- Technology is ripe for adaptation and exploitation
Perspective projection
Perspective projection
1. 3D shape from uncalibrated images
3D model acquisition

Photorealistic models from uncalibrated images of architectural scenes
Ambiguity in a single view

\[
\begin{bmatrix}
\lambda_u \\
\lambda_v \\
\lambda
\end{bmatrix}
= K \begin{bmatrix} R & T \end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]
Stereo vision

\[
\begin{bmatrix}
\lambda_u \\
\lambda_v \\
\lambda
\end{bmatrix} = K \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} X \\
Y \\
Z \\
1
\end{bmatrix}
\]

\[
\begin{bmatrix}
\lambda_{u'} \\
\lambda_{v'} \\
\lambda
\end{bmatrix} = K' \begin{bmatrix} R' & T' \end{bmatrix} \begin{bmatrix} X \\
Y \\
Z \\
1
\end{bmatrix}
\]
Epipolar geometry

\[
\begin{bmatrix}
    u' & v' & 1
\end{bmatrix}
\begin{bmatrix}
    F & 0 & \vdots \\
    0 & F & \vdots \\
    \vdots & \vdots & \vdots
\end{bmatrix}
\begin{bmatrix}
    u \\
    v \\
    1
\end{bmatrix} = 0
\]
Trumpington Street Data
Camera pose determination
3D reconstruction
Reconstruction texture mapped
Conclusions

- Wide baseline matching and image registration
- Mobile phone localisation
- Technology is ripe for adaptation and exploitation