UNIVERSITY OF CAMBRIDGE

Image-Based Localisation

Roberto Cipolla

Duncan Robertson and Ben Tordoff

Department of Engineering

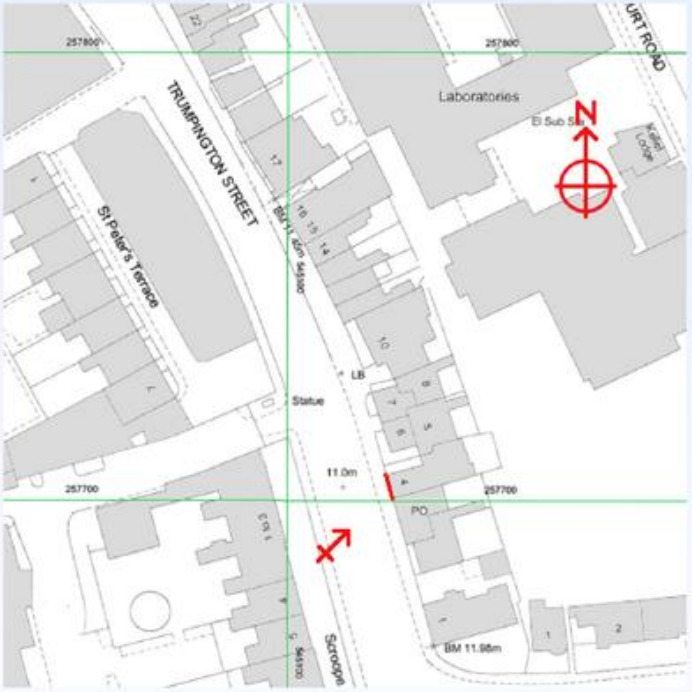
UNIVERSITY OF CAMBRIDGE

Background

The goal – where am I?

UNIVERSITY OF CAMBRIDGE

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

UNIVERSITY OF CAMBRIDGE

?

Why difficult?

UNIVERSITY OF CAMBRIDGE

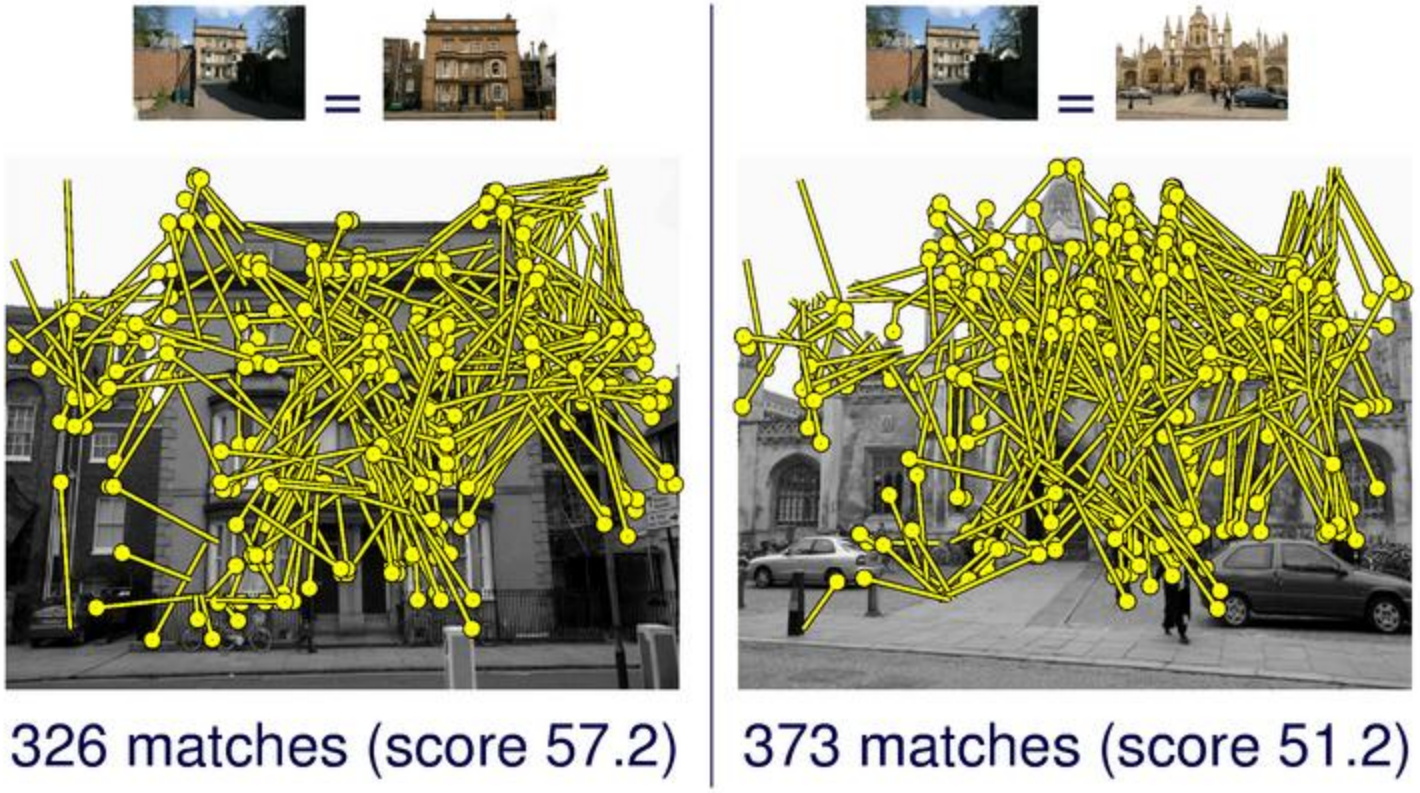
Extreme perspective distortion

Differences in colour / lighting conditions

Occlusion

Unconstrained matching

UNIVERSITY OF CAMBRIDGE



326 matches (score 57.2)

373 matches (score 51.2)

Constrained matching

UNIVERSITY OF CAMBRIDGE

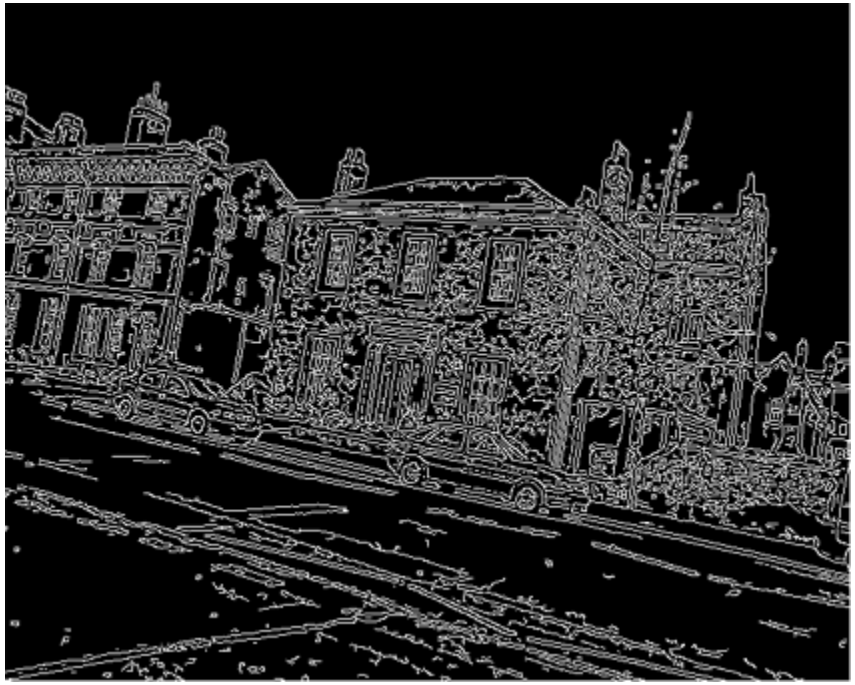
- Building façades are roughly planar
- They contain many horizontal and vertical features
- We can use this to get a “front view” (rectified image)
- Front-views are related by translation and scale only

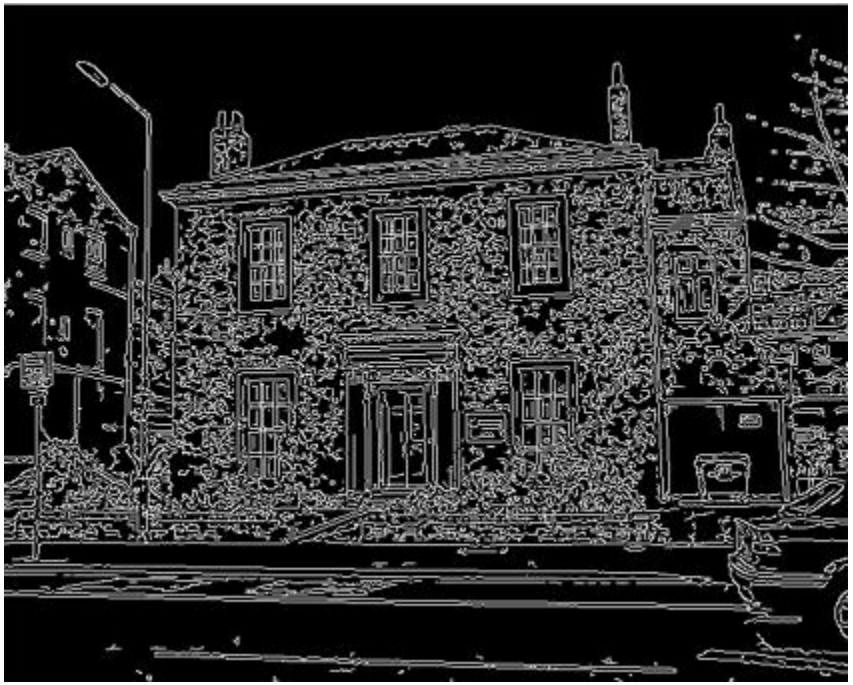
Constrained matching

UNIVERSITY OF CAMBRIDGE

Constrained matching

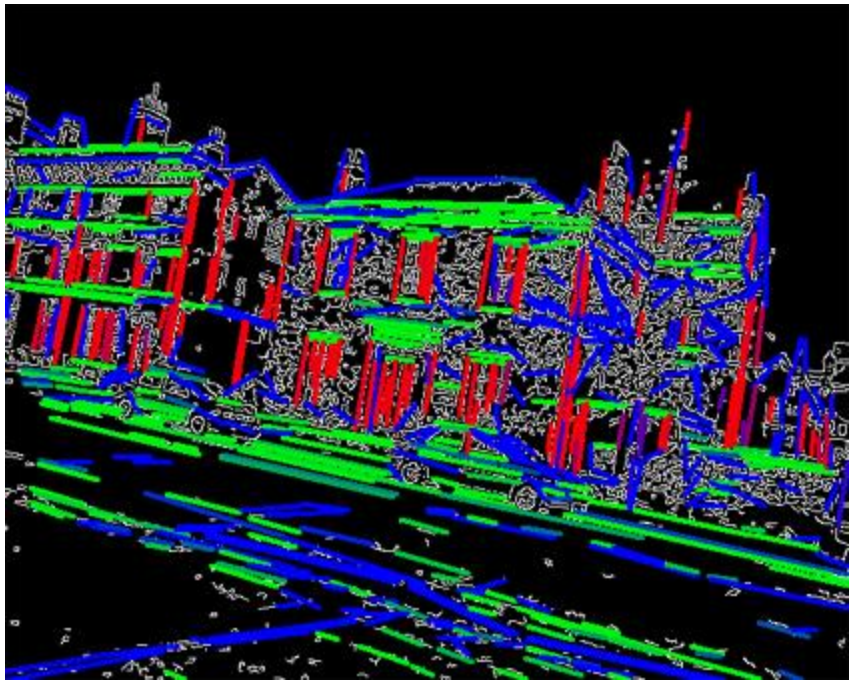
UNIVERSITY OF CAMBRIDGE

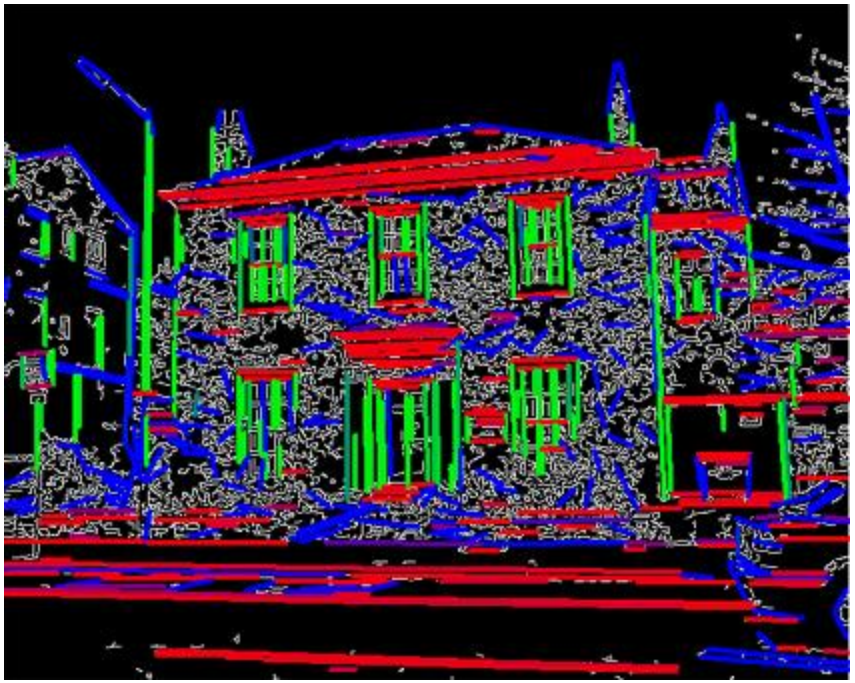




Constrained matching

UNIVERSITY OF CAMBRIDGE





Constrained matching

UNIVERSITY OF CAMBRIDGE

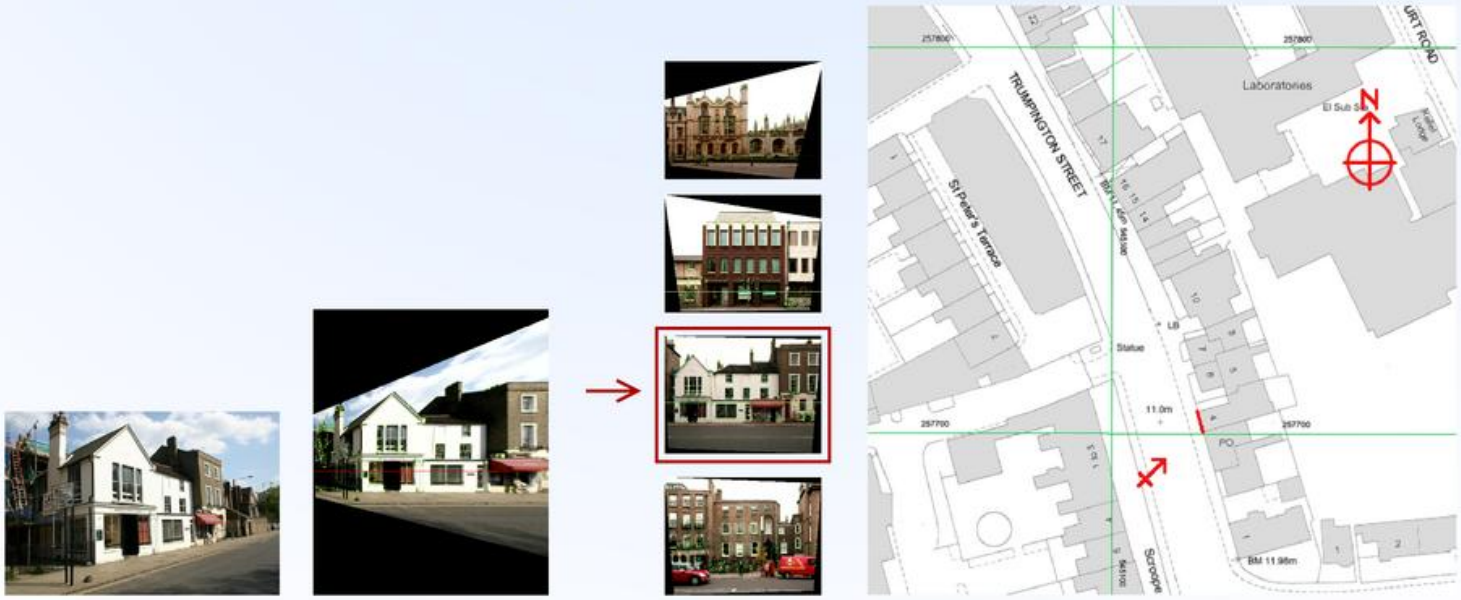
Constrained matching

UNIVERSITY OF CAMBRIDGE

Overview of solution

UNIVERSITY OF CAMBRIDGE

1 vanishing point detection
2 image rectification
3 database search
4 viewpoint determination

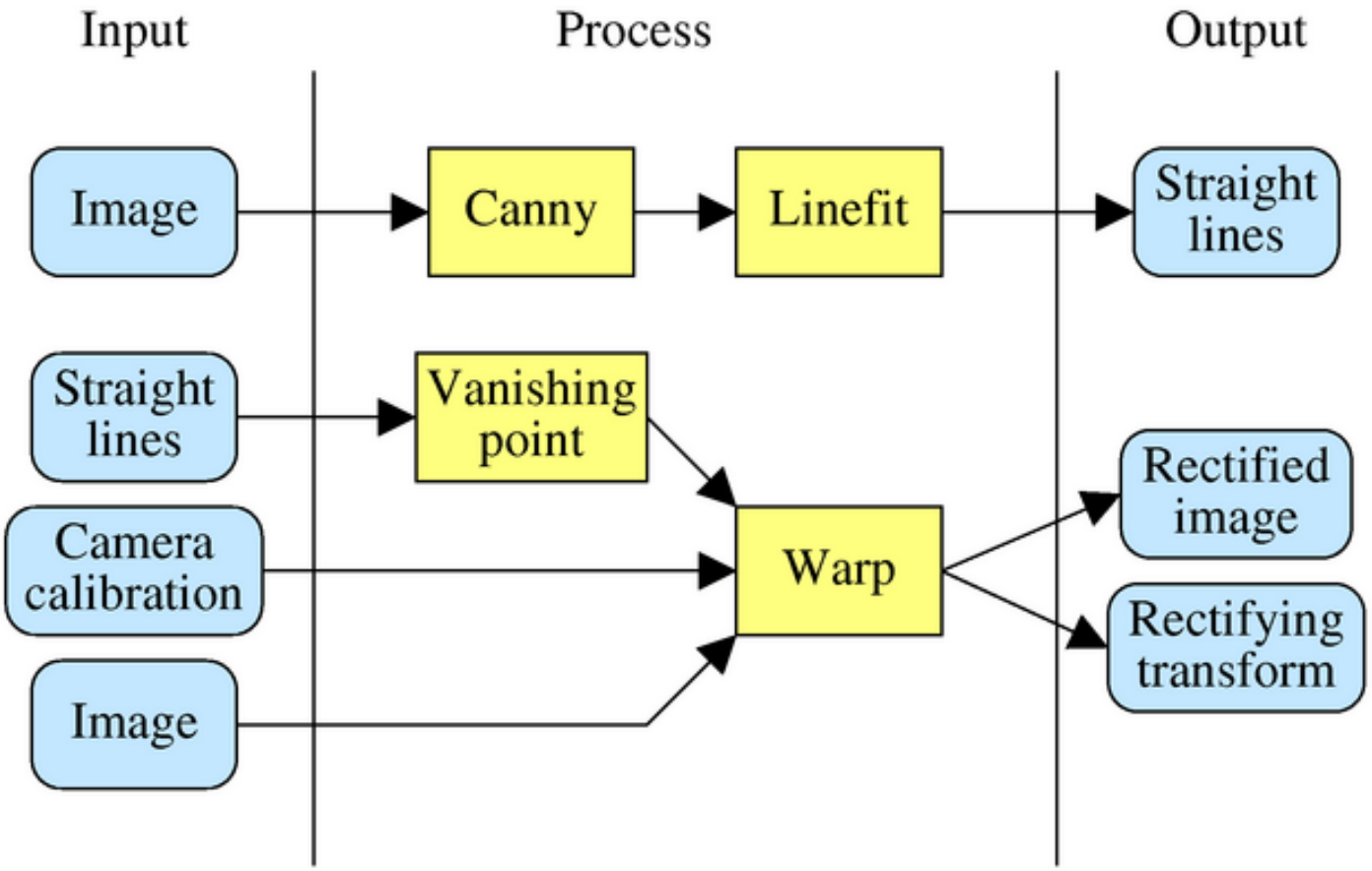


UNIVERSITY OF CAMBRIDGE

Rectification

Rectification

UNIVERSITY OF CAMBRIDGE



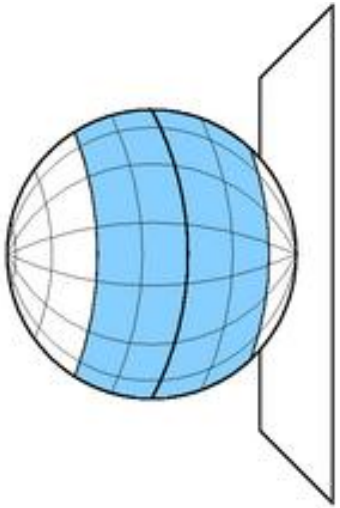
Detection of straight lines

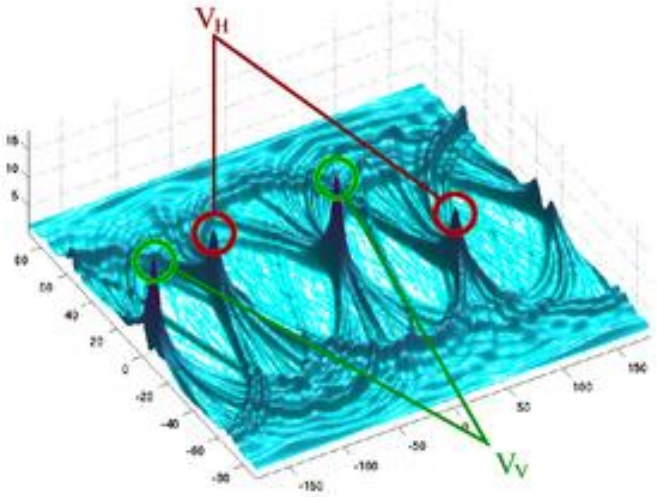
UNIVERSITY OF CAMBRIDGE

Detect straight lines:

Finding vanishing points

UNIVERSITY OF CAMBRIDGE





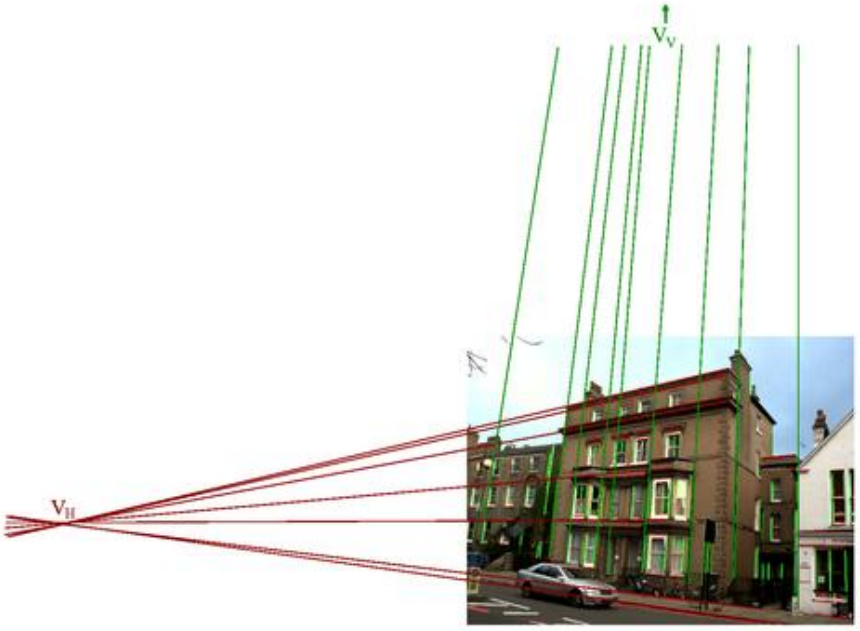
Find vertical and horizontal lines

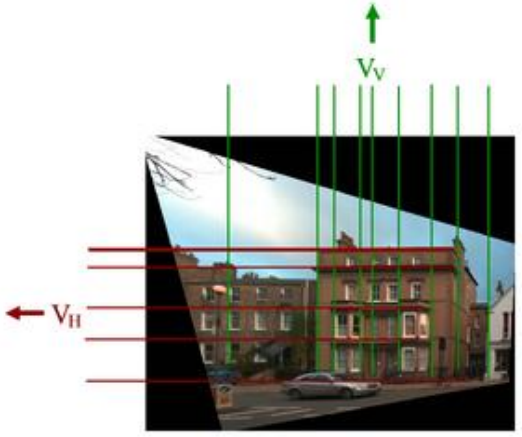
UNIVERSITY OF CAMBRIDGE

Allocate all lines as vertical, horizontal or “clutter”

Rectification by homography

UNIVERSITY OF CAMBRIDGE





Align horizon

UNIVERSITY OF CAMBRIDGE

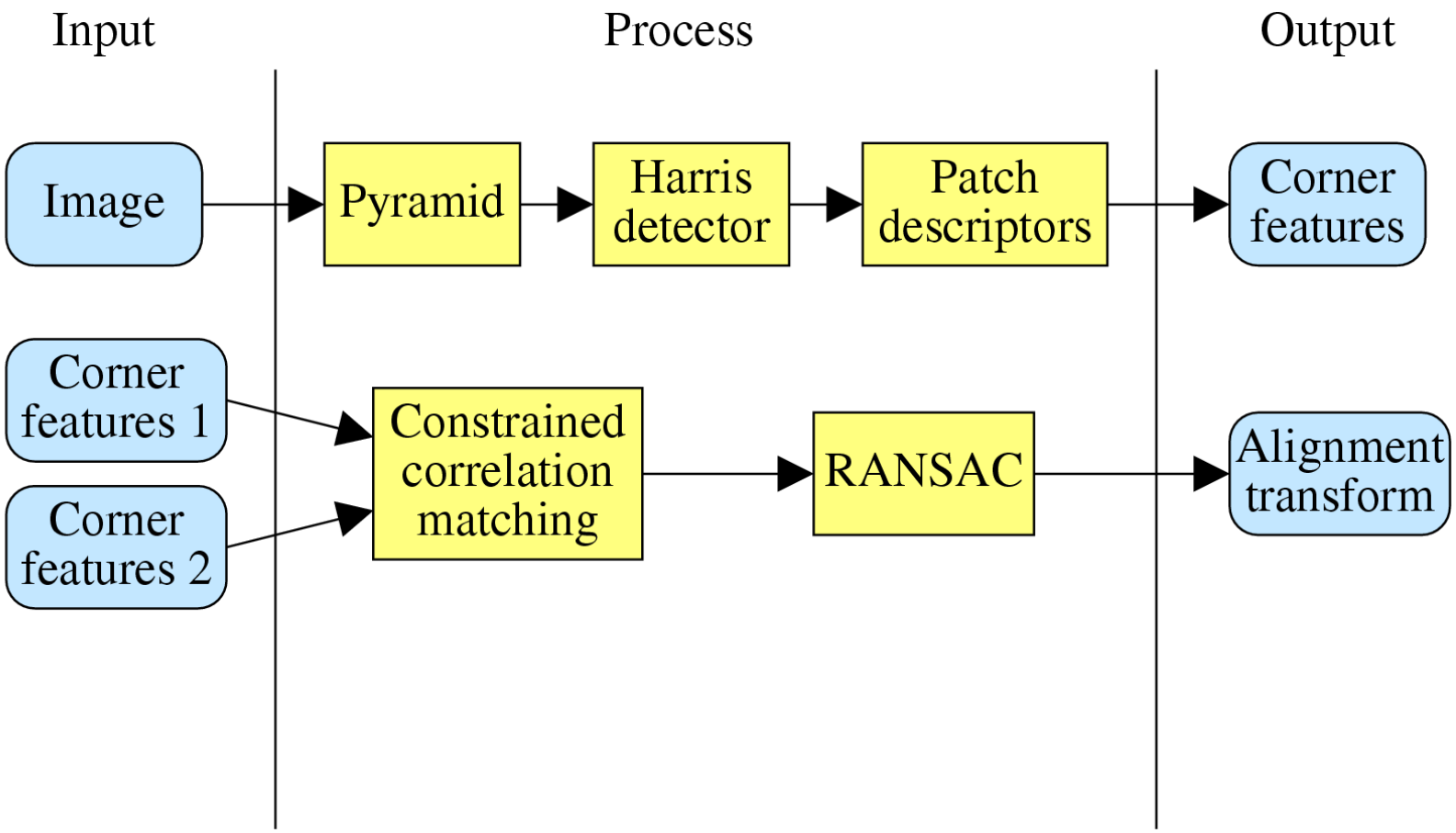
Only difference is now scale + x translation

UNIVERSITY OF CAMBRIDGE

Matching

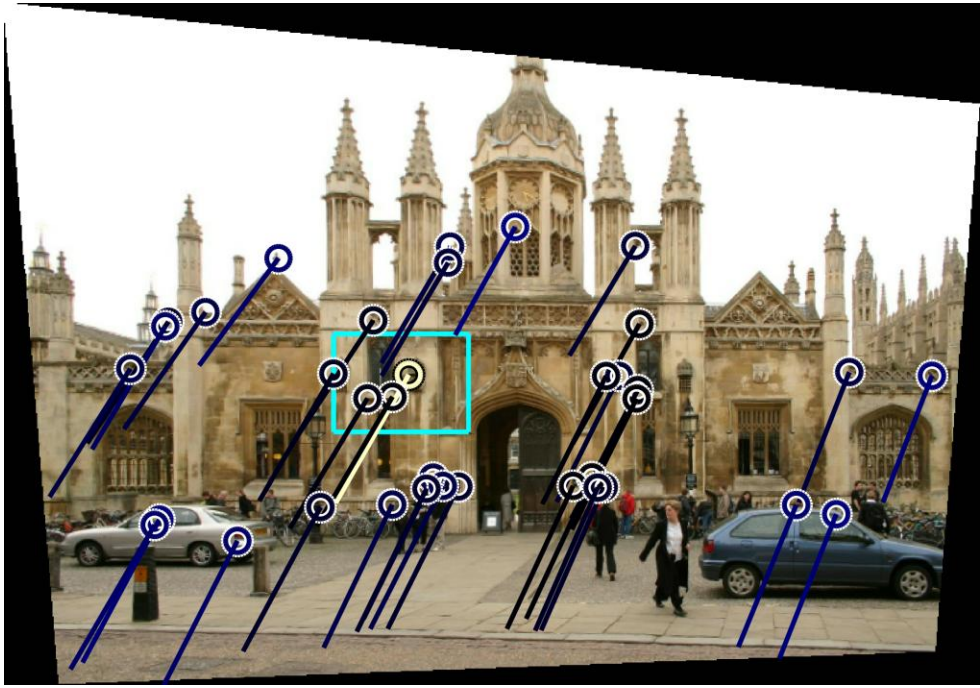
Matching

UNIVERSITY OF CAMBRIDGE



Matching

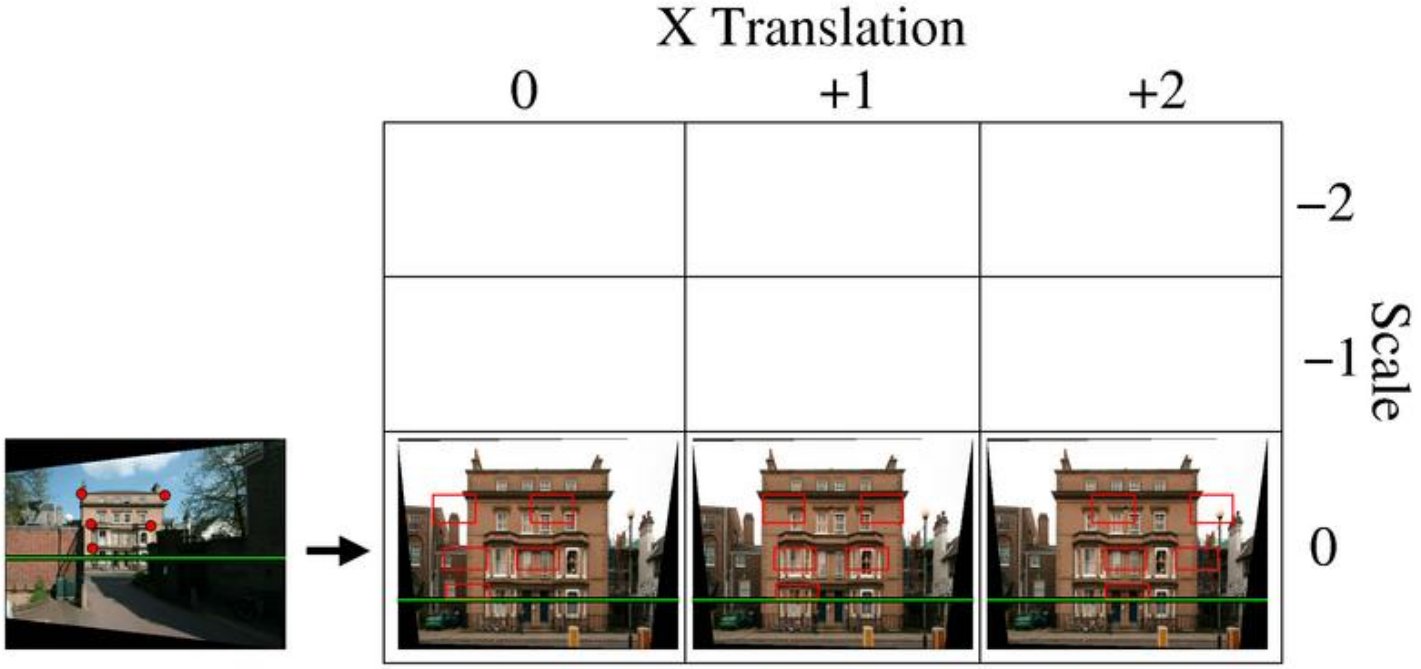
UNIVERSITY OF CAMBRIDGE



Matching

UNIVERSITY OF CAMBRIDGE

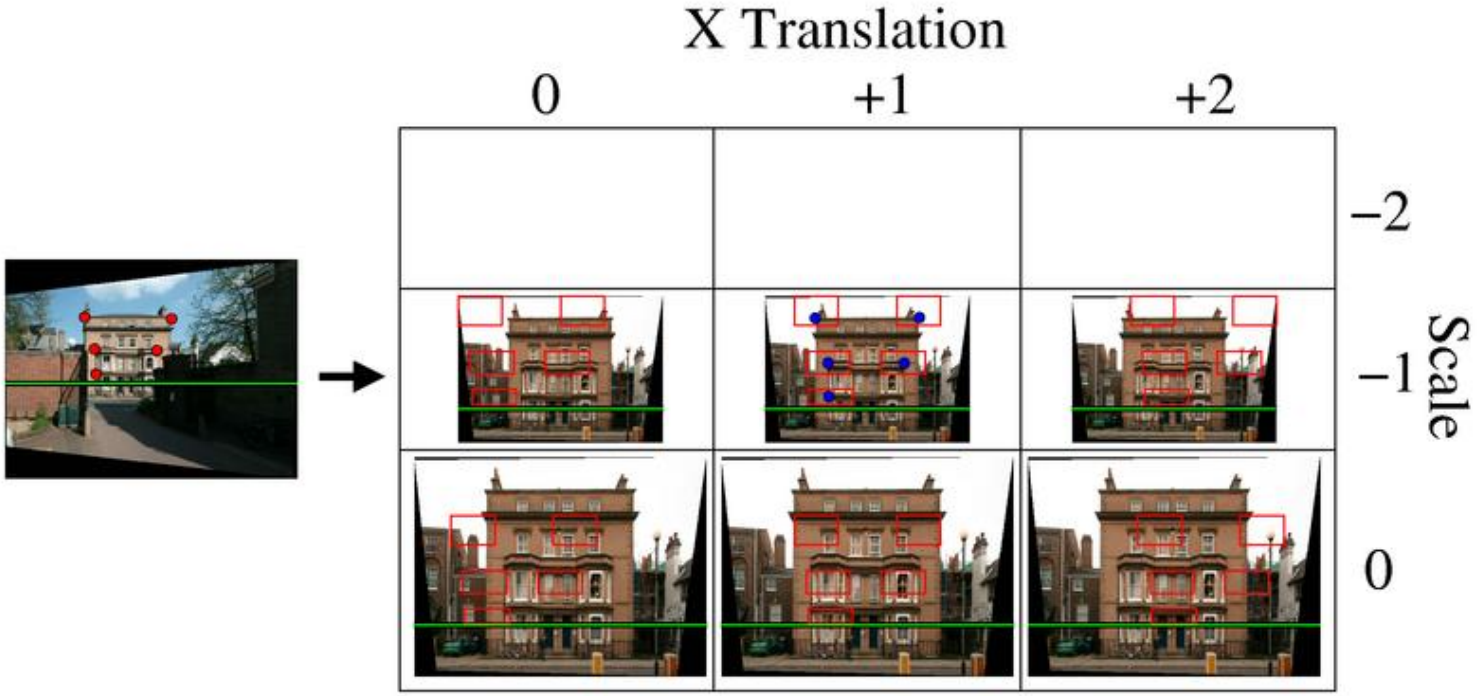
With only 2 params (s,t_x), can search rather than RANSAC.



Matching

UNIVERSITY OF CAMBRIDGE

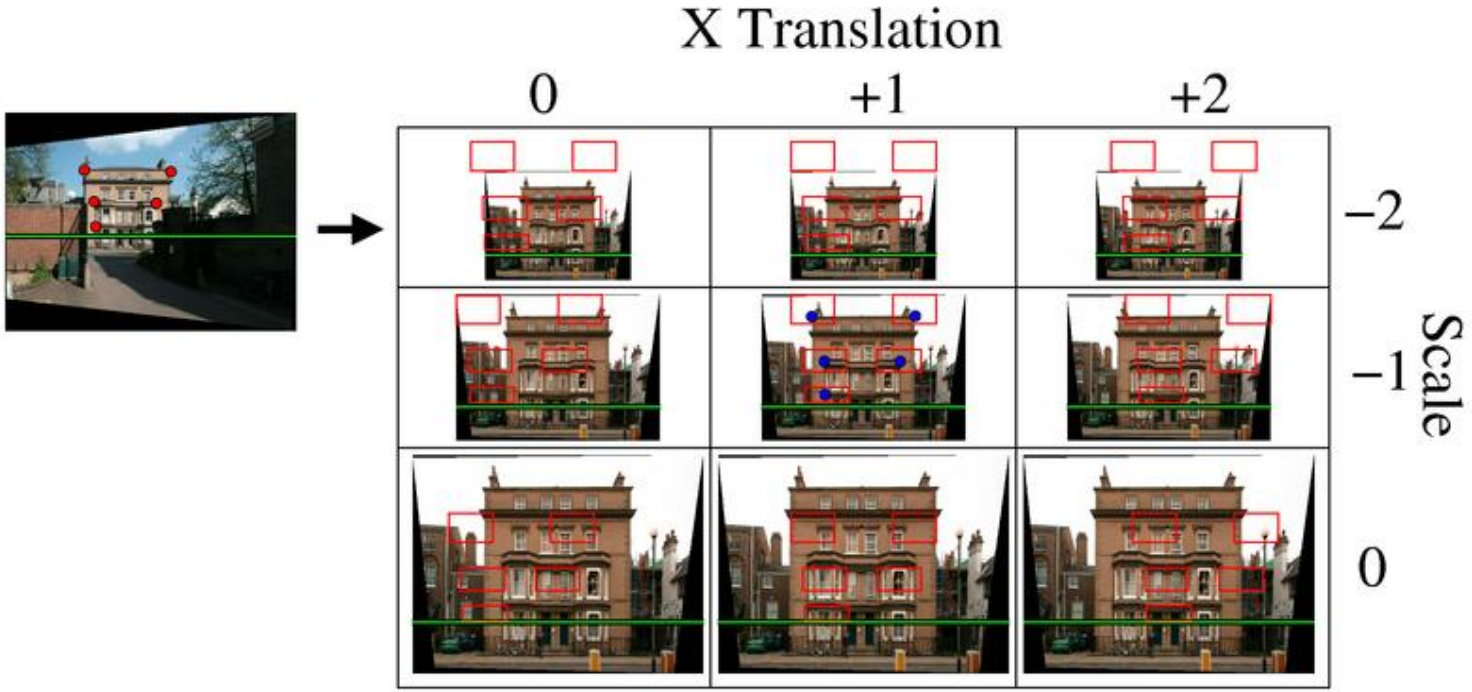
With only 2 params (s,t_x), can search rather than RANSAC.



Matching

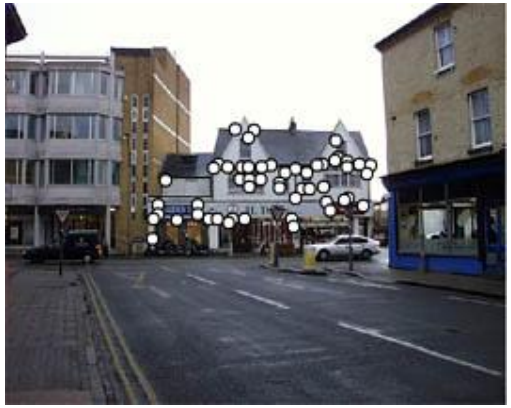
UNIVERSITY OF CAMBRIDGE

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



Examples over wide baselines

UNIVERSITY OF CAMBRIDGE



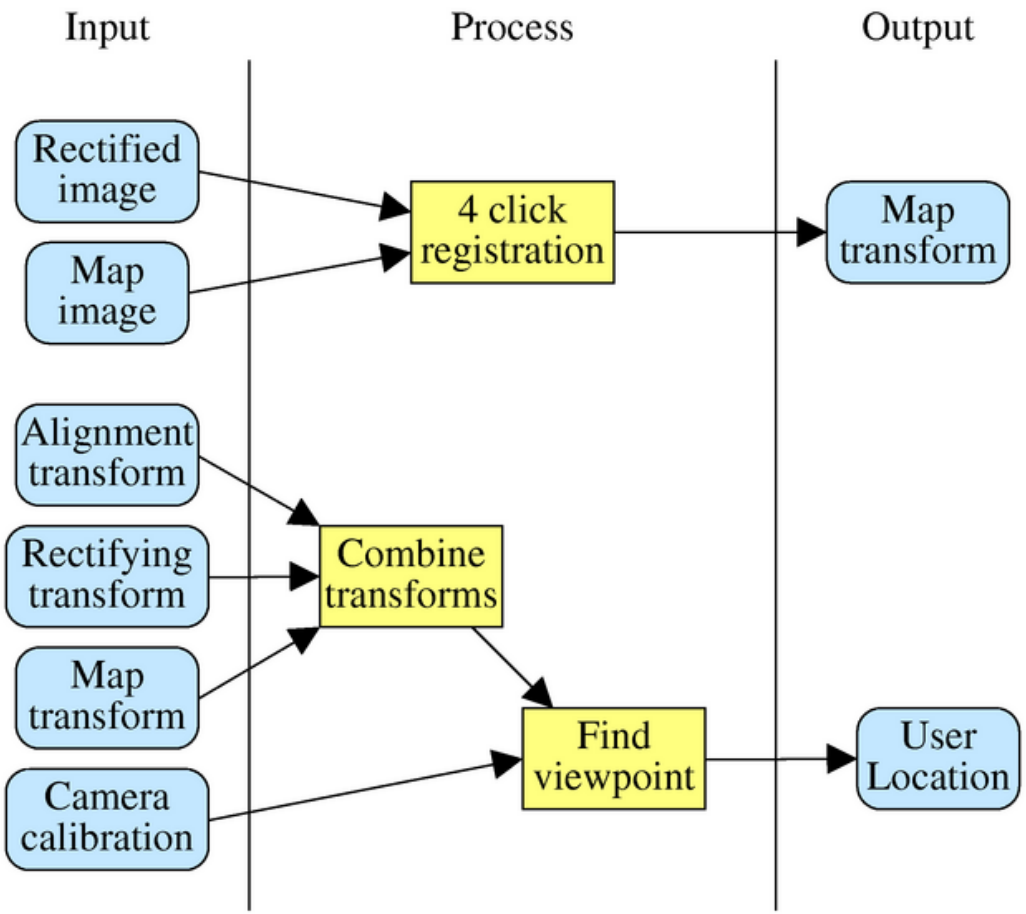
Summary of matching

UNIVERSITY OF CAMBRIDGE

Camera pose estimation - localisation

Localisation

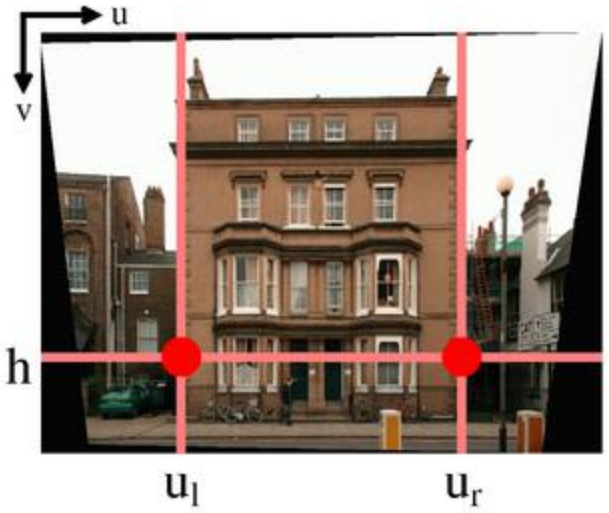
UNIVERSITY OF CAMBRIDGE

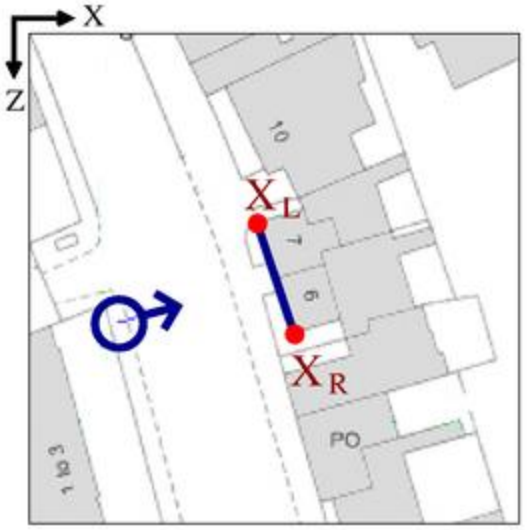


Register database view

UNIVERSITY OF CAMBRIDGE

First align database view to map





Localisation

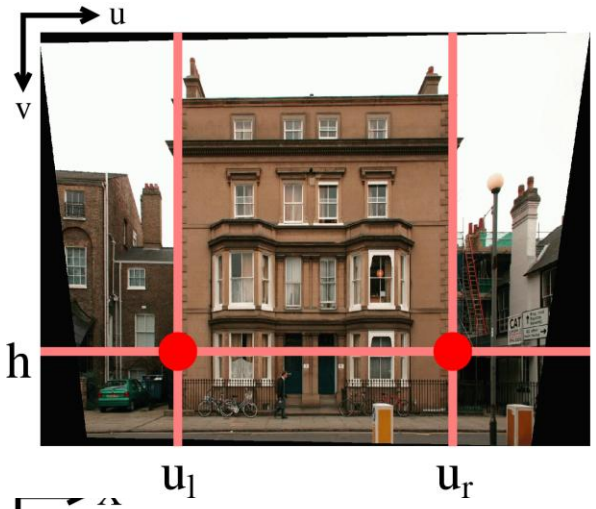
UNIVERSITY OF CAMBRIDGE

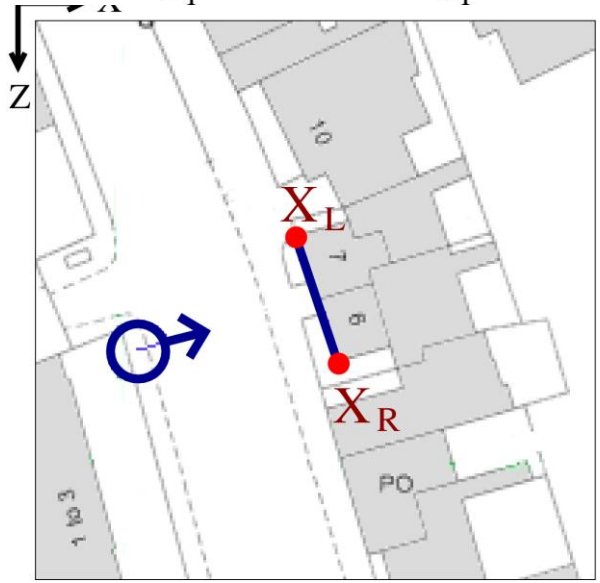
Knowing the rectifying homography ($H_\perp$), the alignment (H_A), and the database view registration, can work backwards to find user:

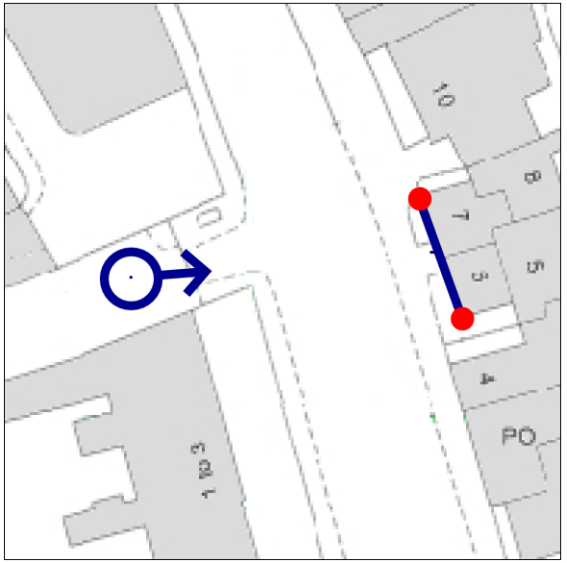
Rectifying rotation $R_\perp$ gives the angle from perpendicular and focal length the distance to camera.

Localisation of query view

UNIVERSITY OF CAMBRIDGE







Localisation

UNIVERSITY OF CAMBRIDGE

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

UNIVERSITY OF CAMBRIDGE

Evaluation

UNIVERSITY OF CAMBRIDGE

Image-based localisation

UNIVERSITY OF CAMBRIDGE

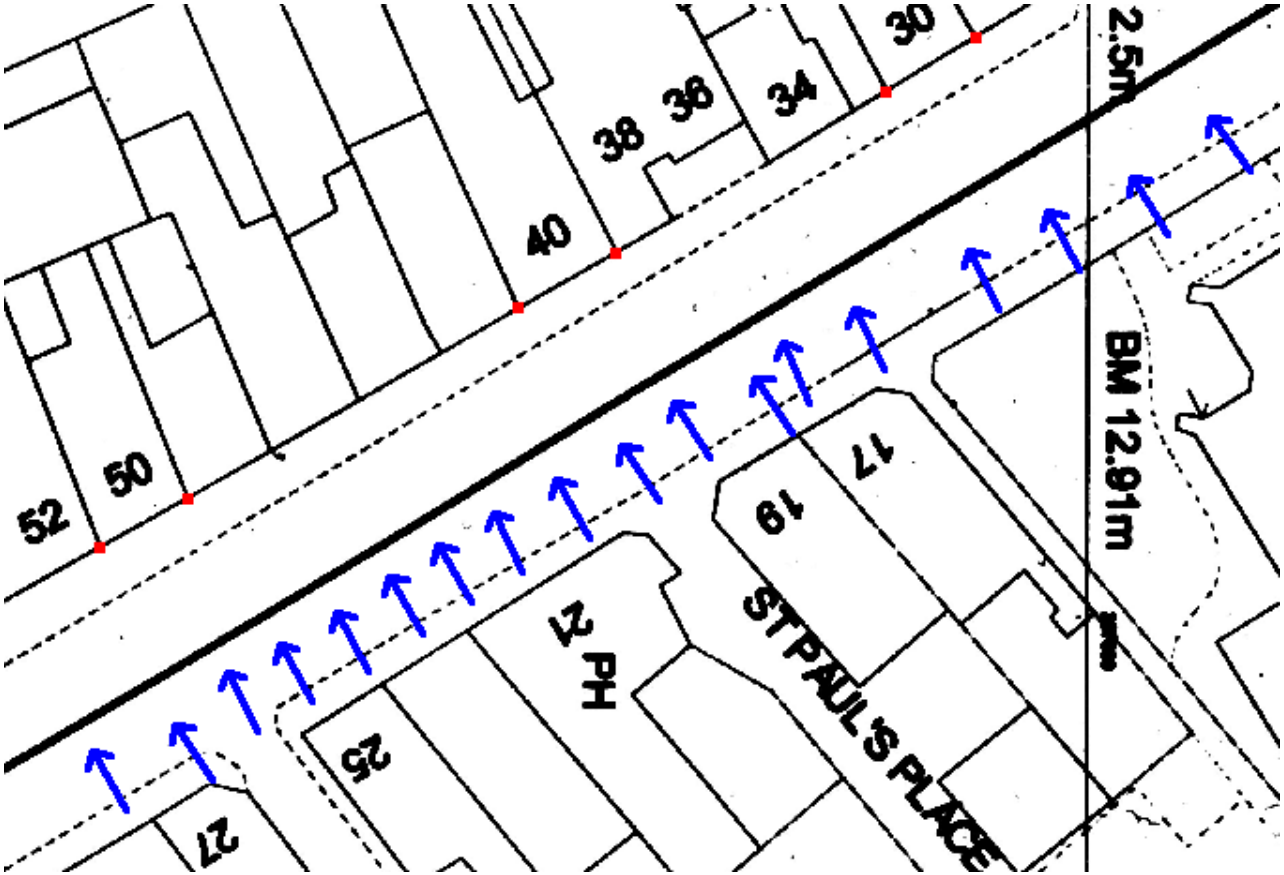


Image-based localisation

UNIVERSITY OF CAMBRIDGE

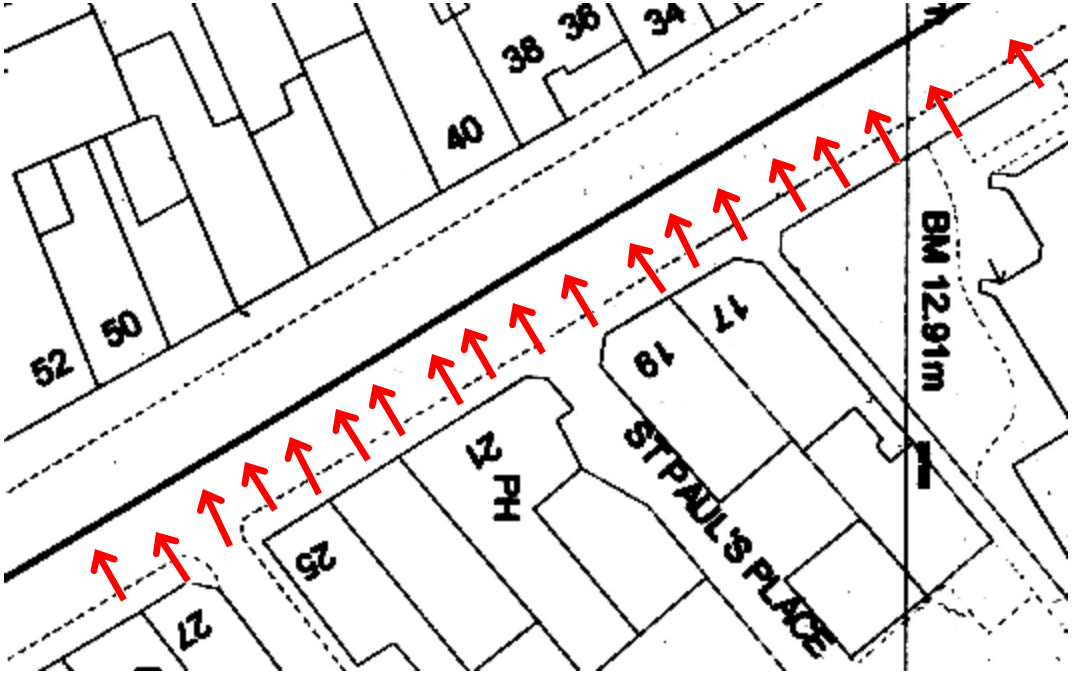


Image-based localisation

UNIVERSITY OF CAMBRIDGE

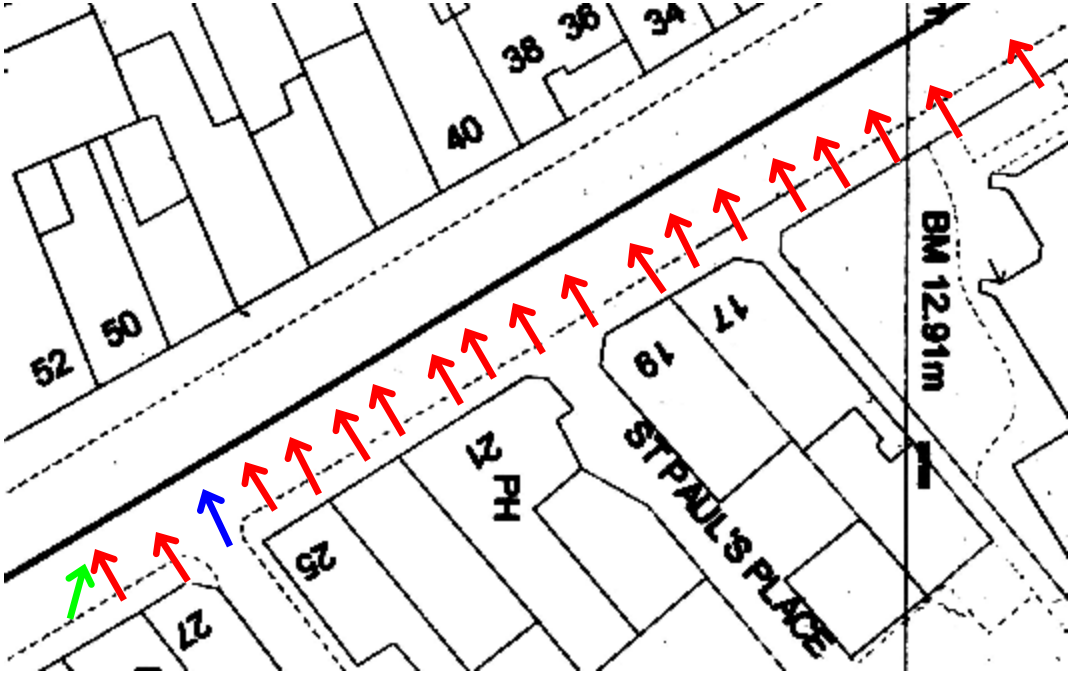


Image-based localisation

UNIVERSITY OF CAMBRIDGE

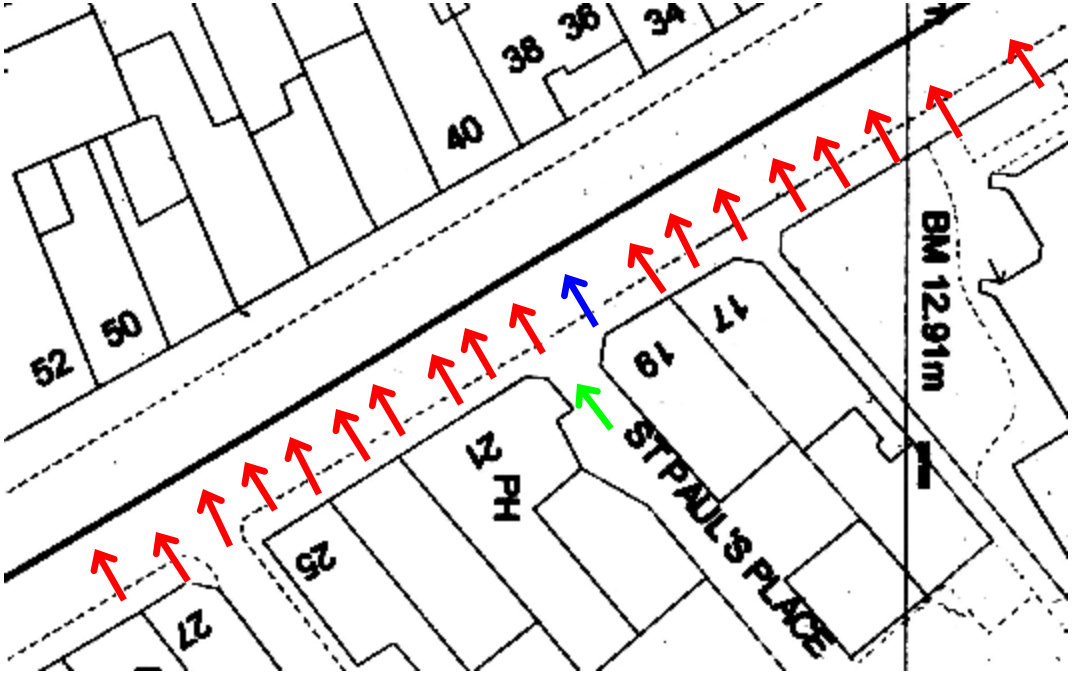
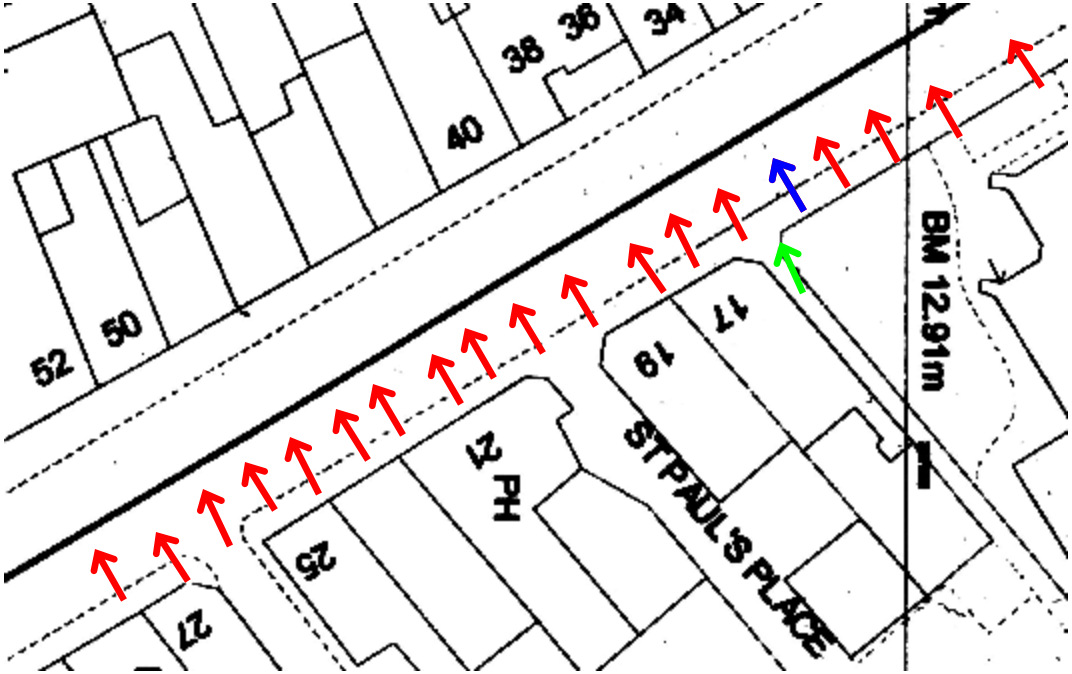


Image-based localisation

UNIVERSITY OF CAMBRIDGE



Conclusions

UNIVERSITY OF CAMBRIDGE

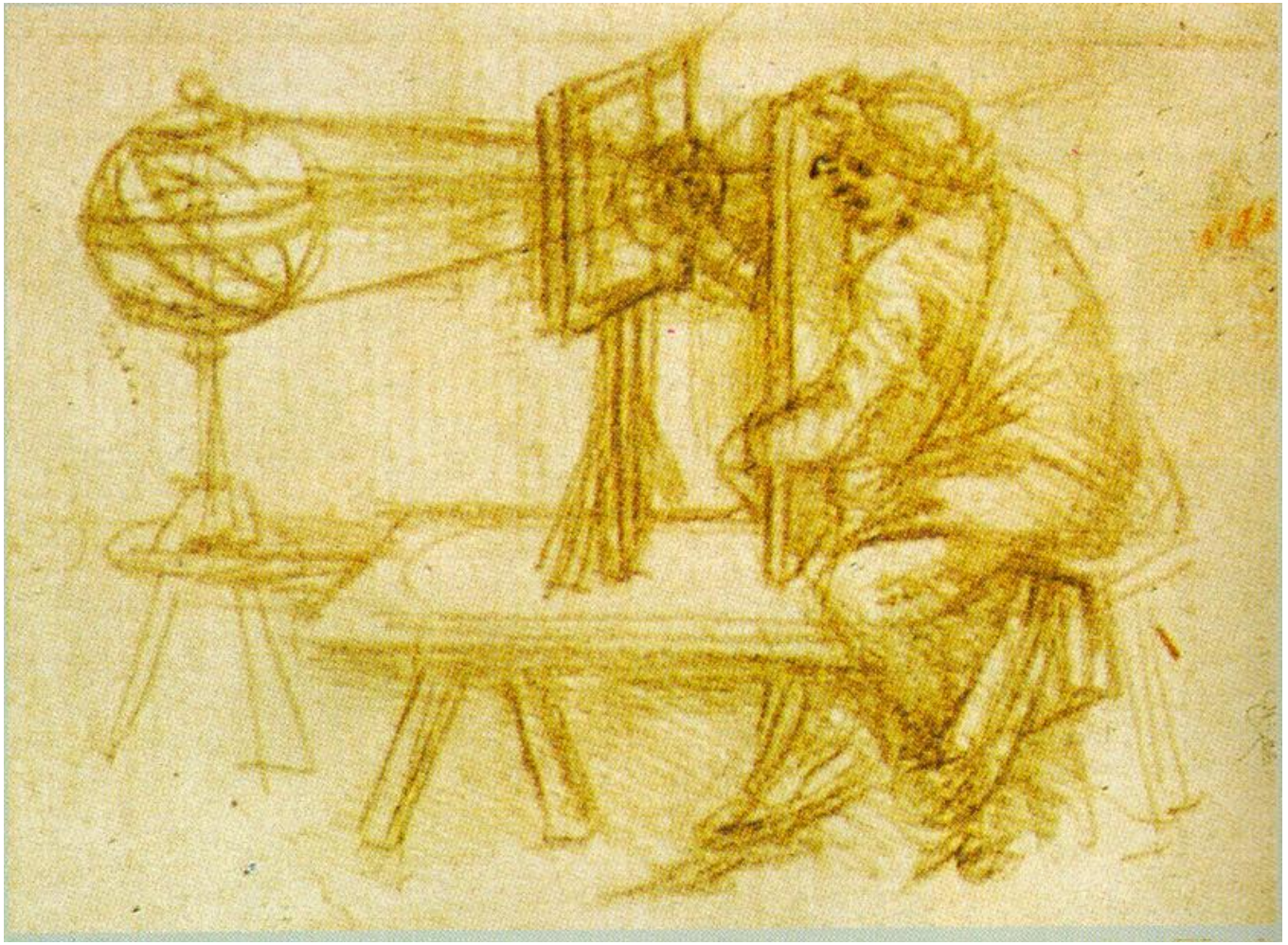
- Effective wide baseline matching and image registration
- Mobile phone localisation:
  - Where am I?
  - What am I looking at?
- Scaling up to real applications?
- Technology is ripe for adaptation and exploitation

UNIVERSITY OF CAMBRIDGE

Perspective projection

Perspective projection

UNIVERSITY OF CAMBRIDGE



UNIVERSITY OF CAMBRIDGE

1. 3D shape from uncalibrated images

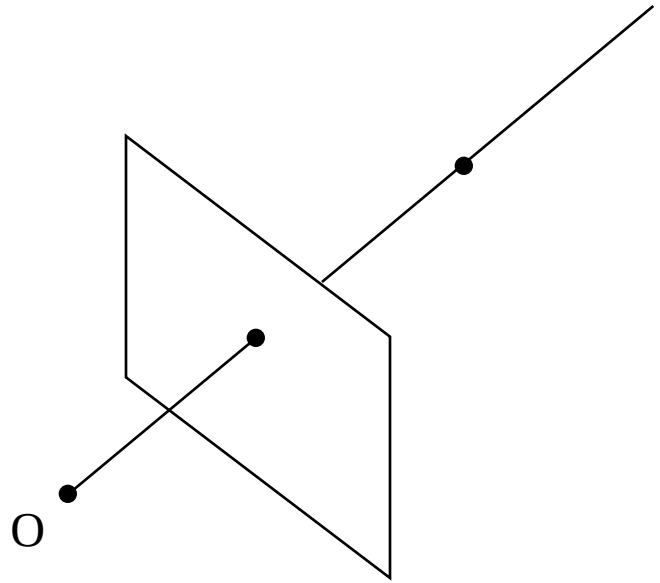
3D model acquisition

UNIVERSITY OF CAMBRIDGE

Photorealistic models from uncalibrated images of architectural scenes

Ambiguity in a single view

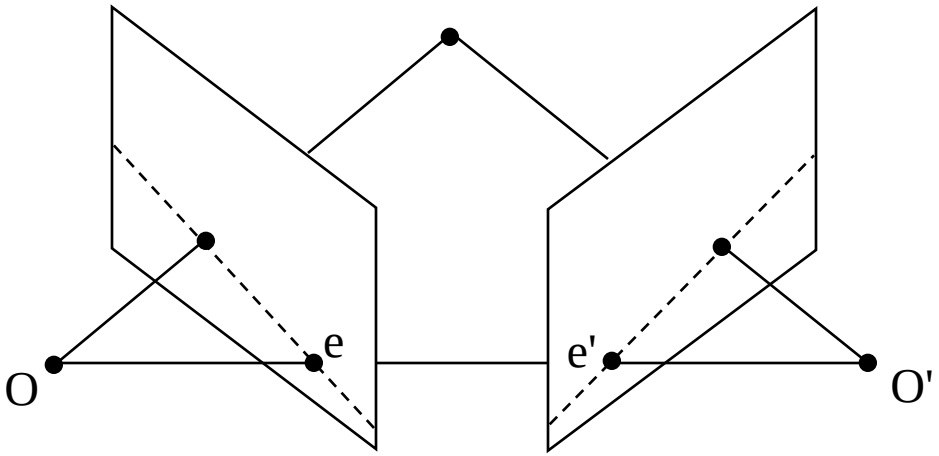
UNIVERSITY OF CAMBRIDGE



$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \mathrm{K}\begin{bmatrix} \mathrm{R} & \mathrm{T} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Stereo vision

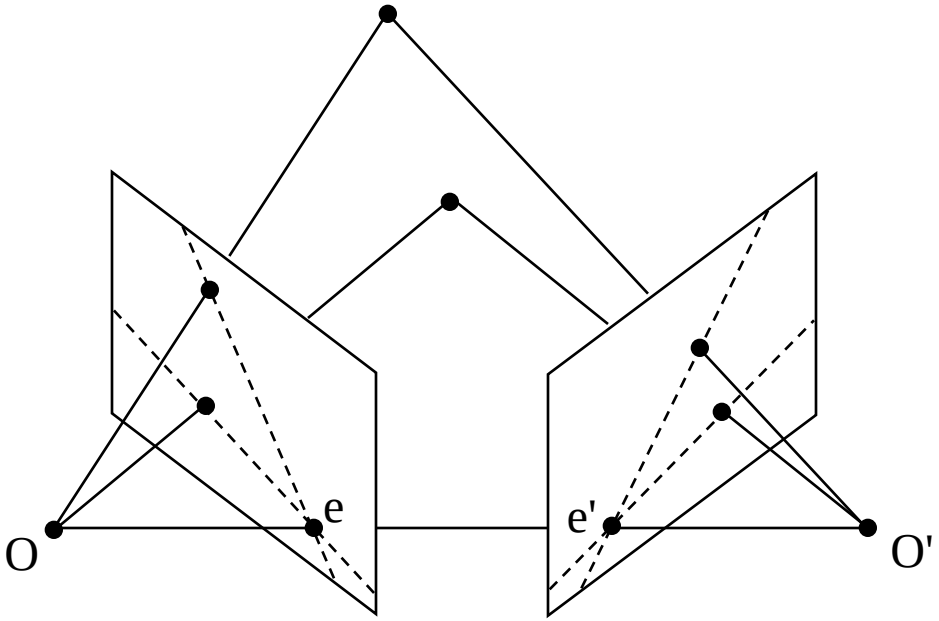
UNIVERSITY OF CAMBRIDGE



$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \mathrm{K}\begin{bmatrix} \mathrm{R} & \mathrm{T} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \begin{bmatrix} \lambda u' \\ \lambda v' \\ \lambda \end{bmatrix} = \mathrm{K}'\begin{bmatrix} \mathrm{R}' & \mathrm{T}' \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Epipolar geometry

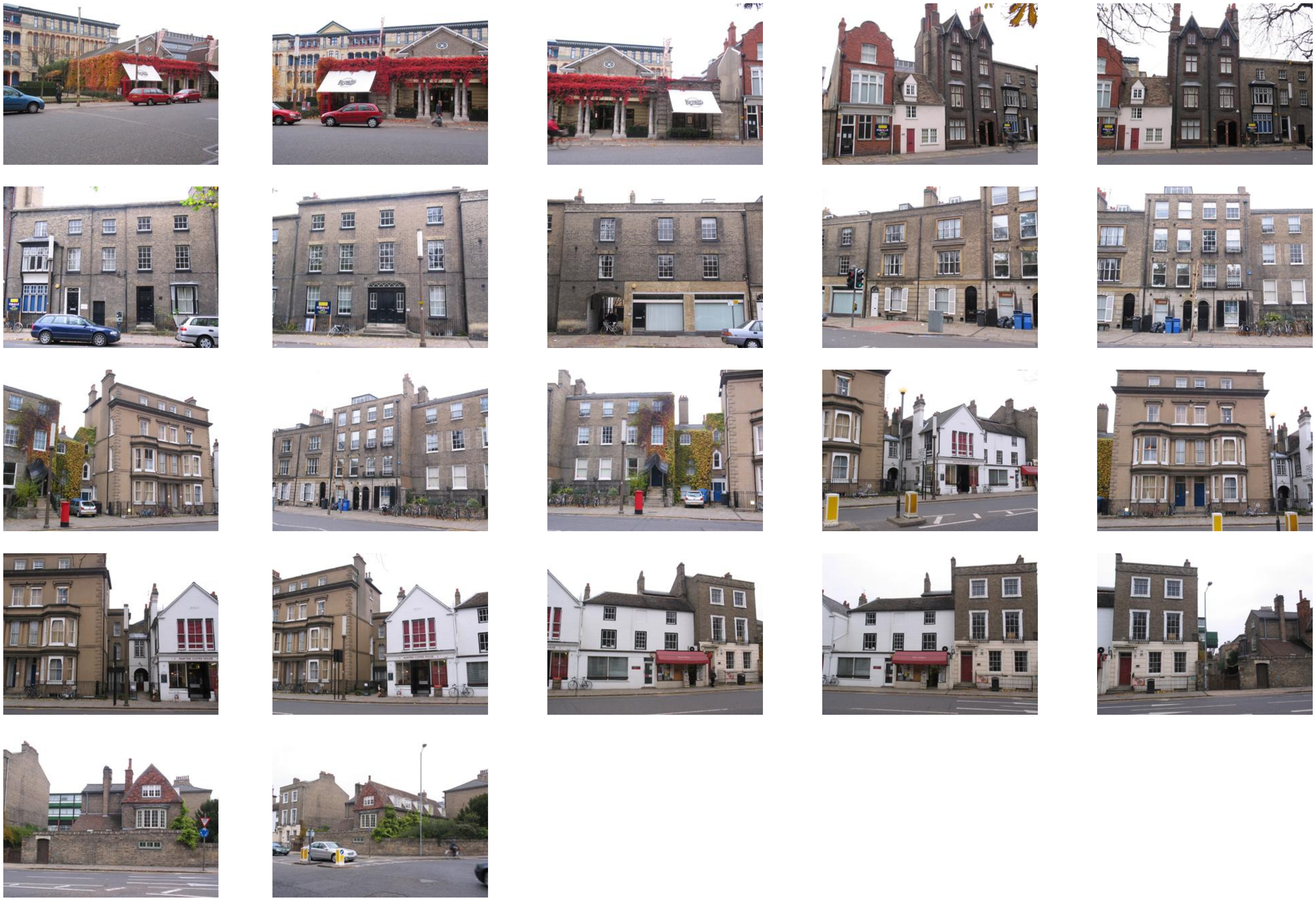
UNIVERSITY OF CAMBRIDGE



$$\begin{bmatrix} u' & v' & 1 \end{bmatrix} \begin{bmatrix} & \mathrm{F} & \\ & & \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0$$

UNIVERSITY OF CAMBRIDGE

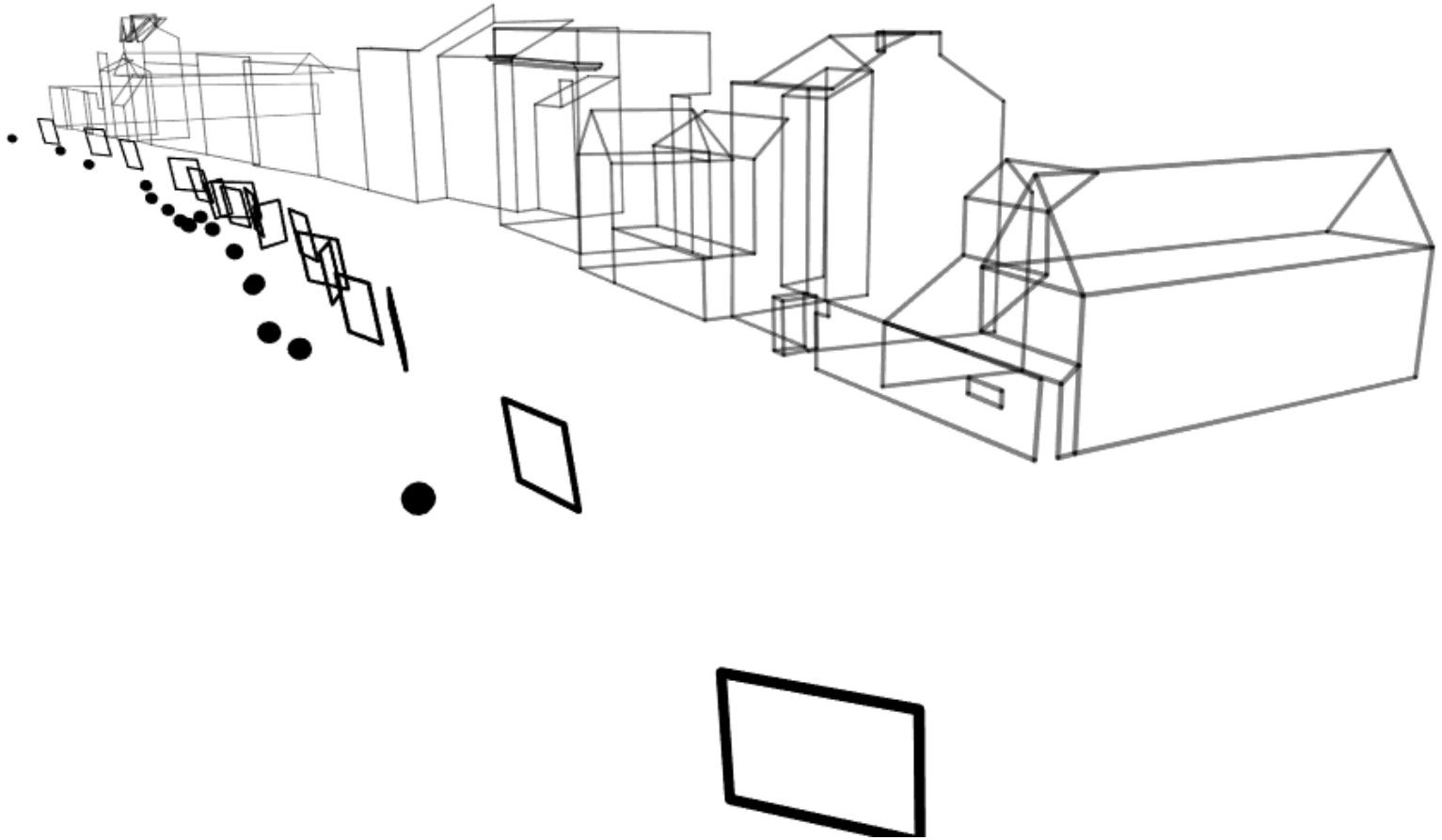
Trumpington Street Data



Camera pose determination

3D reconstruction

UNIVERSITY OF CAMBRIDGE



Reconstruction texture mapped

UNIVERSITY OF CAMBRIDGE

Conclusions

UNIVERSITY OF CAMBRIDGE

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