General Imaging
Design, Modeling and Applications

Peter Sturm, INRIA Grenoble - Rhône-Alpes, France

http://perception.inrialpes.fr/people/Sturm

With contributions from:

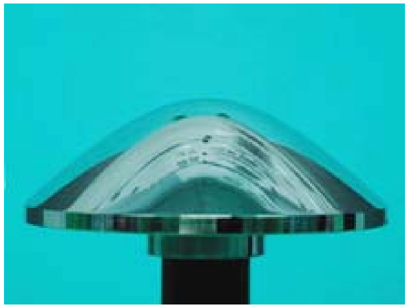
Srikumar Ramalingam	INRIA Grenoble - Rhône-Alpes and UC Santa Cruz
Jean-Philippe Tardif	Université de Montréal

Contents

- Introduction
- General imaging models
- Non-parametric calibration and distortion correction
- Non-parametric self-calibration
- Structure-from-motion

Introduction

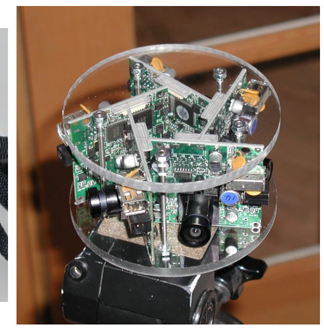
There exist lots of camera designs:

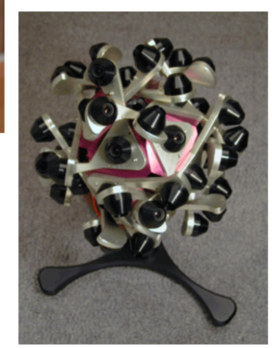


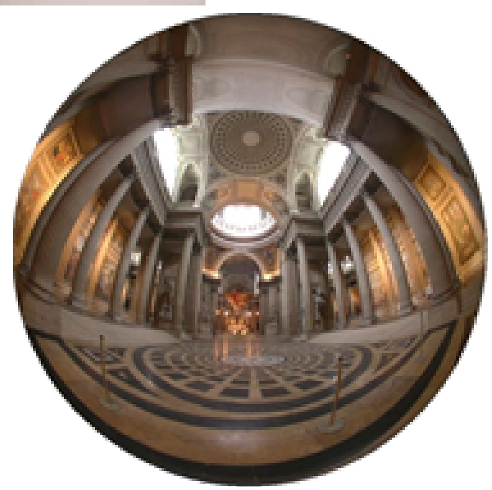











Introduction

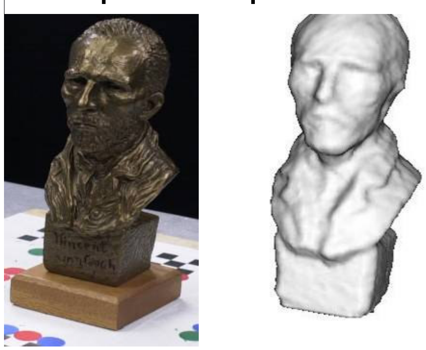
Some applications:

Automatic Vehicle Navigation

Aerial Mosaics

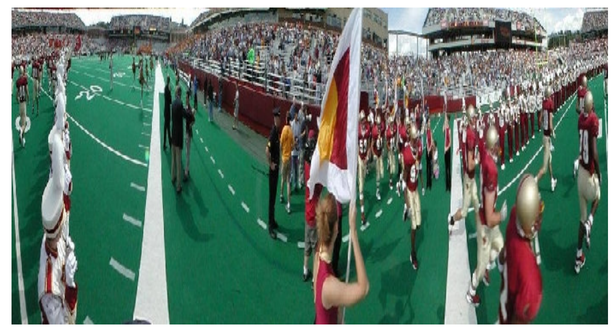
3D Video Conferencing

Shape Computation



Surveillance

Panoramic Imaging



- Many applications require/benefit from a specific type of imaging system
- Work underlying this talk started by considering omnidirectional systems (large field of view)

Introduction

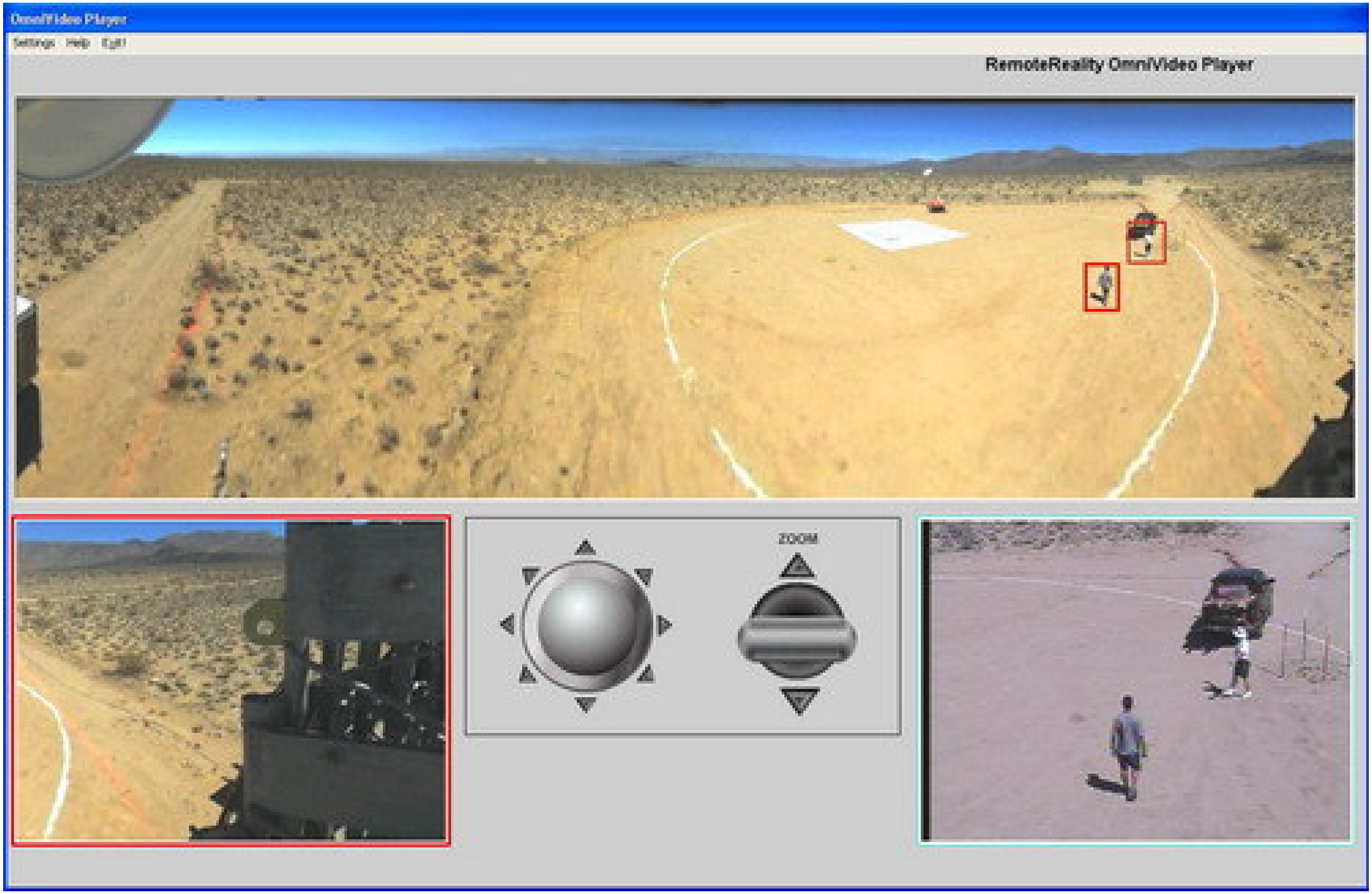
Videoconferencing:

Introduction

Surveillance:

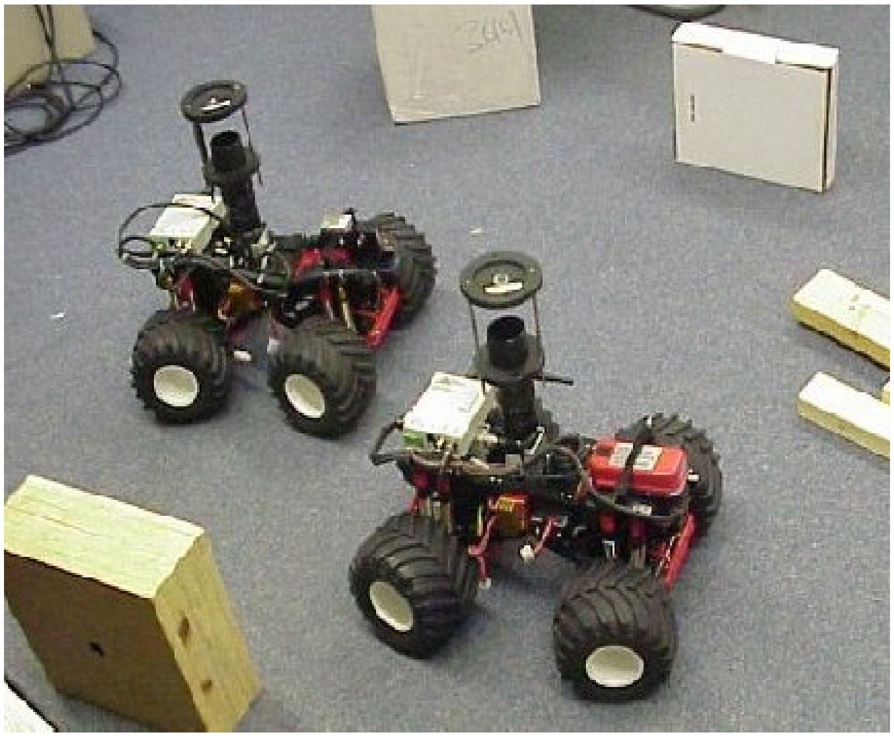
Introduction

Surveillance:



Introduction

Robot navigation (including obstacle avoidance):



Taylor et al. – GRASP

Santos Victor et al. – ISR/IST

Introduction

Panoramic imaging, here mosaicing:

Problematic for dynamic scenes:

Introduction

Panoramic imaging with omnidirectional cameras:

Fisheye

Catadioptric

Introduction

Design of tailor-made imaging systems:

Usual:

Desired:

[Swaminathan et al.]

Introduction

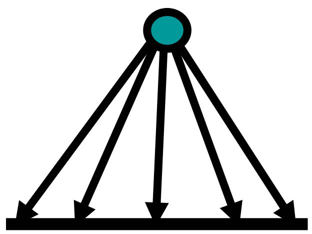
Design of tailor-made imaging systems:

By Julian Beever

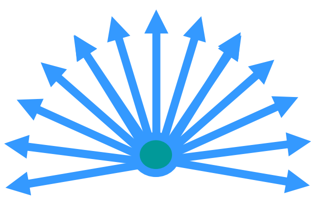
Introduction

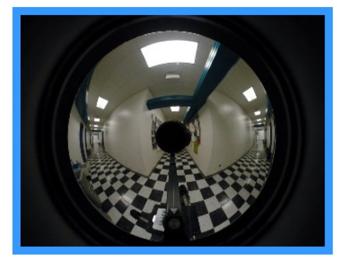
Different cameras “sample light rays” in different ways:

Perspective cameras:

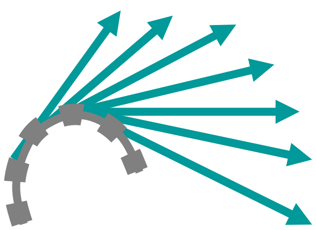
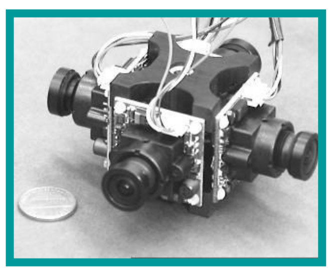


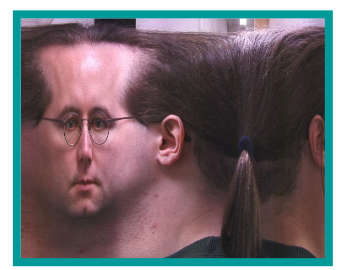

Single viewpoint cameras:



Non-single viewpoint cameras:



Introduction

Each camera type comes with a particular model and often, particular calibration and structure-from-motion algorithms

Main motivations for my related works:

- Propose generic camera models and calibration algorithms
- Highlight common principles underlying structure-from-motion algorithms for different camera models
- Generalize (parts of) the structure-from-motion theory, e.g. multi-view geometry (epipolar, trifocal and quadrifocal geometry)

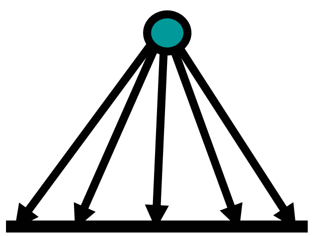
Contents

- Introduction
- General imaging models
- Non-parametric calibration and distortion correction
- Non-parametric self-calibration
- Structure-from-motion

Imaging Models

Perspective cameras:

- Imaging model well-known…
- Calibration information (intrinsic parameters)
  allows to perform **projection**: 3D points → image points)
  and **back-projection**: image points → projection rays (lines of sight)



Imaging Models

Single viewpoint cameras:

- Perspective projection plus radial or decentering distortion
  - imaging model well-known…
  - again, calibration (intrinsic parameters) allows to perform **projection** and **back-projection**
  - calibration approaches:
    - plumbline calibration: use images of straight line patterns to estimate “non-perspective” parameters
    - calibration with control points: compute all parameters of the model using bundle adjustment

Imaging Models

Single viewpoint cameras:

- Fisheyes
  - several models have been proposed (ad hoc or derived from actual lens designs)
  - e.g. equi-angular model (existence of distortion center and optical axis such that distance of image point to distortion center is proportional to angle between projection ray and optical axis)

Imaging Models

Catadioptric systems (camera + mirror):

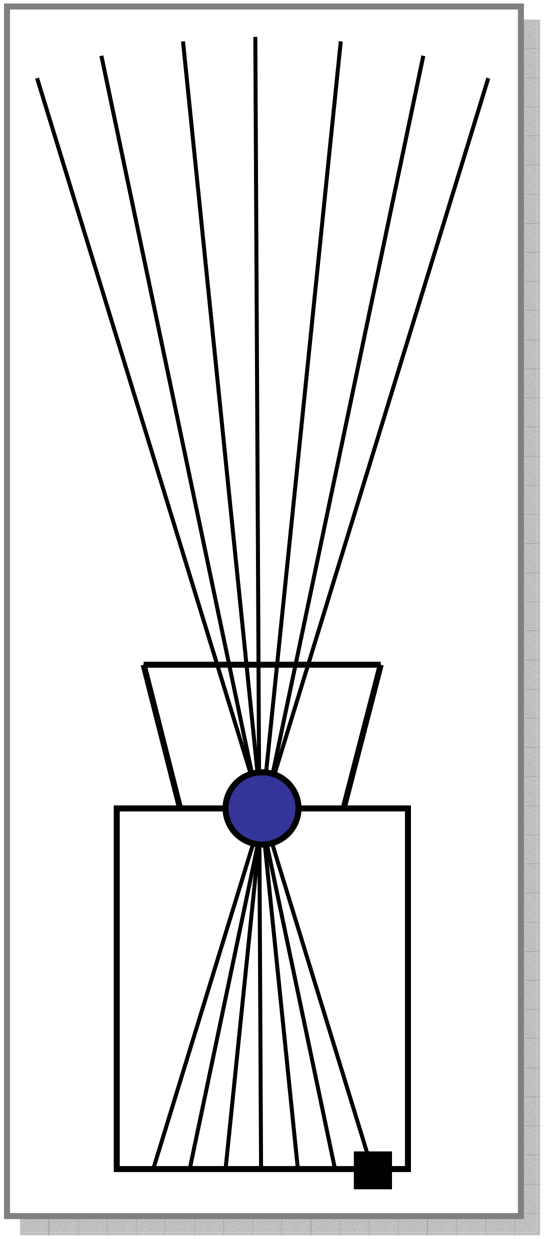
- Knowledge of mirror shape and position relative to camera, together with calibration of camera, allows to perform back-projection

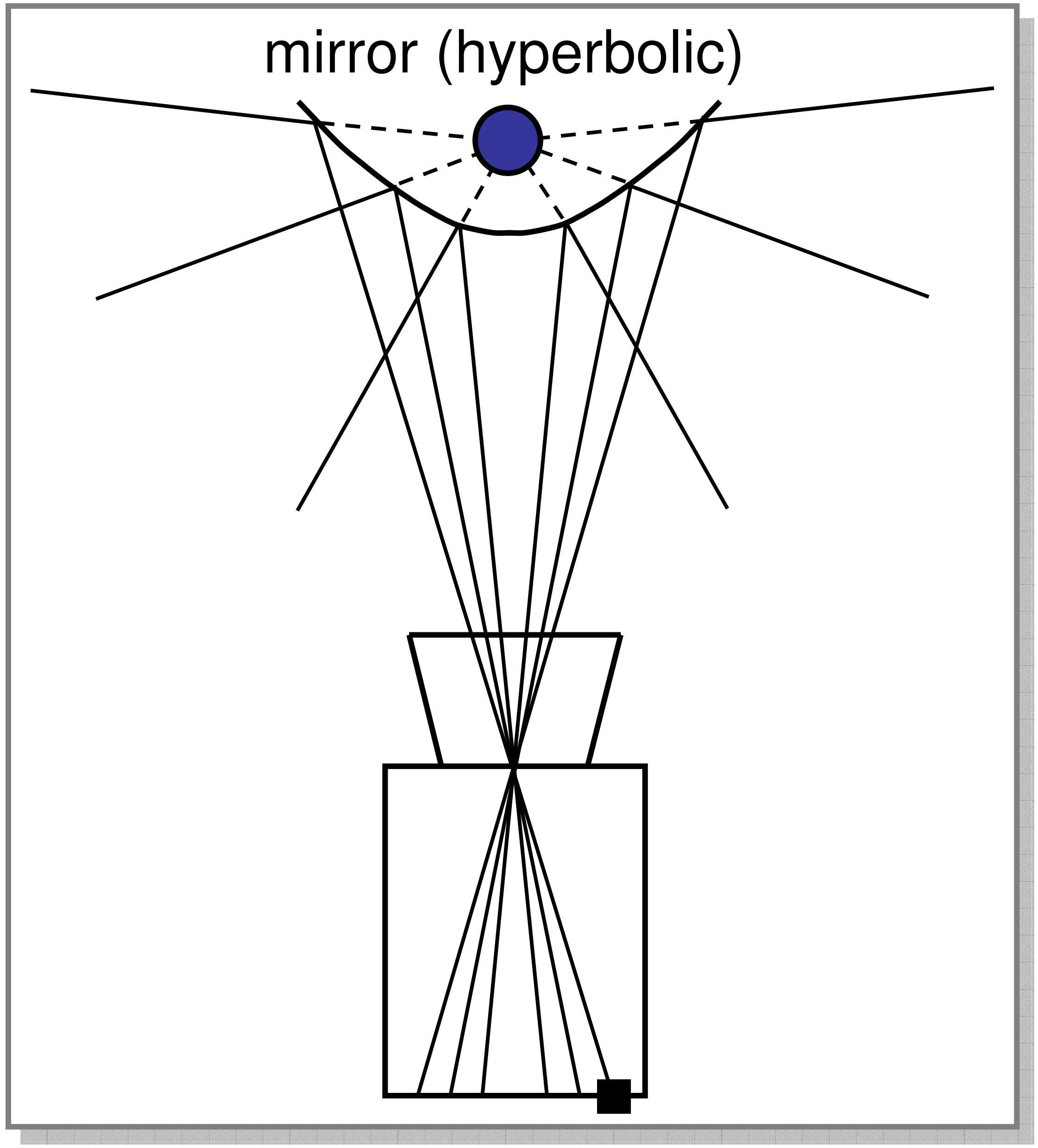
Imaging Models

Back to single viewpoint cameras:

- **Central** catadioptric systems
  - with appropriate mirror shape and position, system has a single effective viewpoint (cf. next slide)
  - practically relevant: parabolic mirror + orthographic camera, hyperbolic mirror + perspective camera
  - various imaging models have been proposed:
    - models whose parameters represent correlations between mirror shape/position and calibration of camera
    - unifying models for all types of central catadioptric cameras
  - calibration approaches:
    - plumbline approaches (sometimes with closed-form solutions)
    - calibration with control points: compute all parameters of the model using bundle adjustment

Imaging Models





Imaging Models

Single viewpoint cameras:

- Central catadioptric system using **multiple** planar mirrors and cameras (so-called Nalwa pyramid)

  - perspective camera + planar mirror
    ≡ perspective camera with effective optical center on the other side of the plane

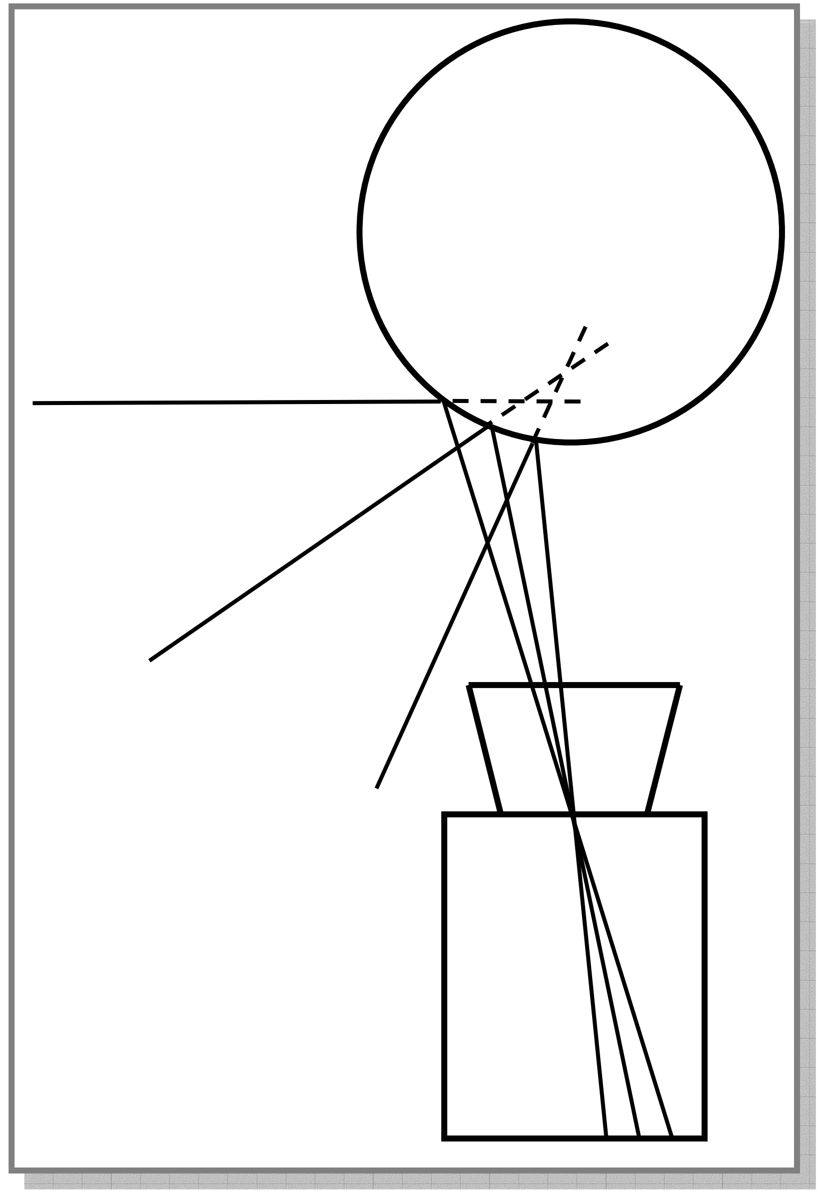
  - Nalwa pyramid: assemble pairs (camera, mirror) such that effective optical centers coincide

→ possibility to construct a high-resolution panoramic image

Imaging Models

Non-single viewpoint cameras:

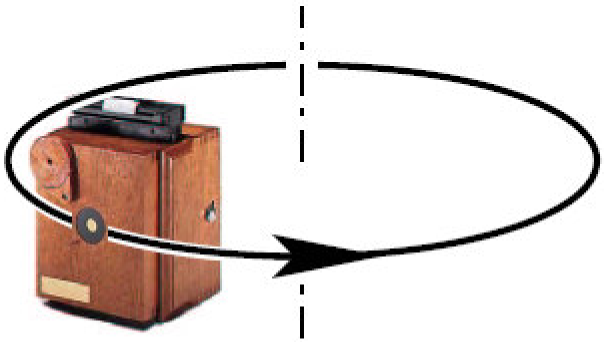
- **Non-central** catadioptric systems
  - spheres, cones or any non-quadric mirrors give non-central system: projection rays do not intersect in a single point
  - calibration approaches have been developed for individual systems
  - example:
    - mirror that leads to equi-angular imaging model

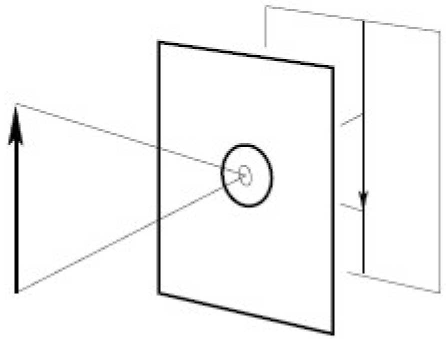


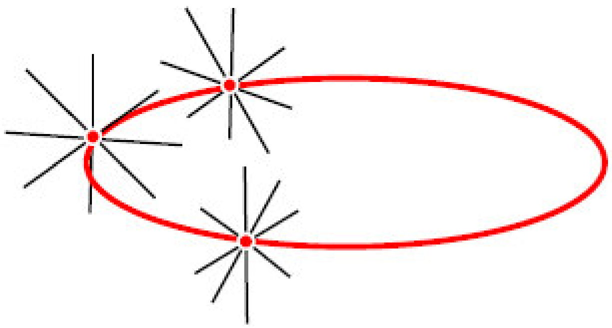
Imaging Models

Other non-single viewpoint cameras:

- Pushbroom cameras
  - Moving linear camera acquires 1D images that are stitched together to a 2D image (motion is usually a lateral translation)
- So-called non-central mosaics
  - Acquired by a camera rotating about an axis not containing the optical center (from each image, take one or several columns of pixels and stitch them all together)







Imaging Models

Other non-single viewpoint cameras:

- So-called multi-perspective images
  - Acquired like a non-central mosaic but with camera looking inwards

Imaging Models

All above imaging models are subsumed by the following
generic imaging model:

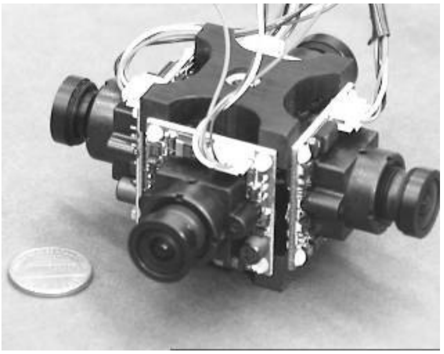
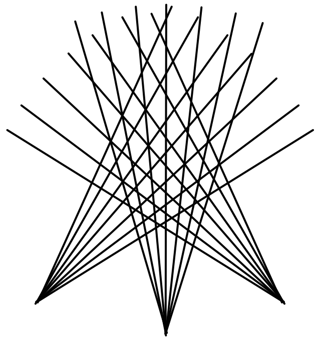
- A pixel “watches along” one viewing ray
- Camera model is lookup table, containing for each pixel the coordinates of the associated ray

Calibration = computation of all these rays

Imaging Models

Comments on the generic imaging model:

- is idealized (in reality, a pixel sees more than a line)
- more complete model, including radiometric properties, is used by Grossberg and Nayar (ICCV 2001)
- other sampling than pixel-wise is possible (e.g. sub-pixel)
- conceptually, allows to consider a stereo or multi-camera system as a single camera: union of their pixels and associated rays

Contents

- Introduction
- General imaging models
- Non-parametric calibration and distortion correction
- Non-parametric self-calibration
- Structure-from-motion

Non-parametric calibration — Basic idea

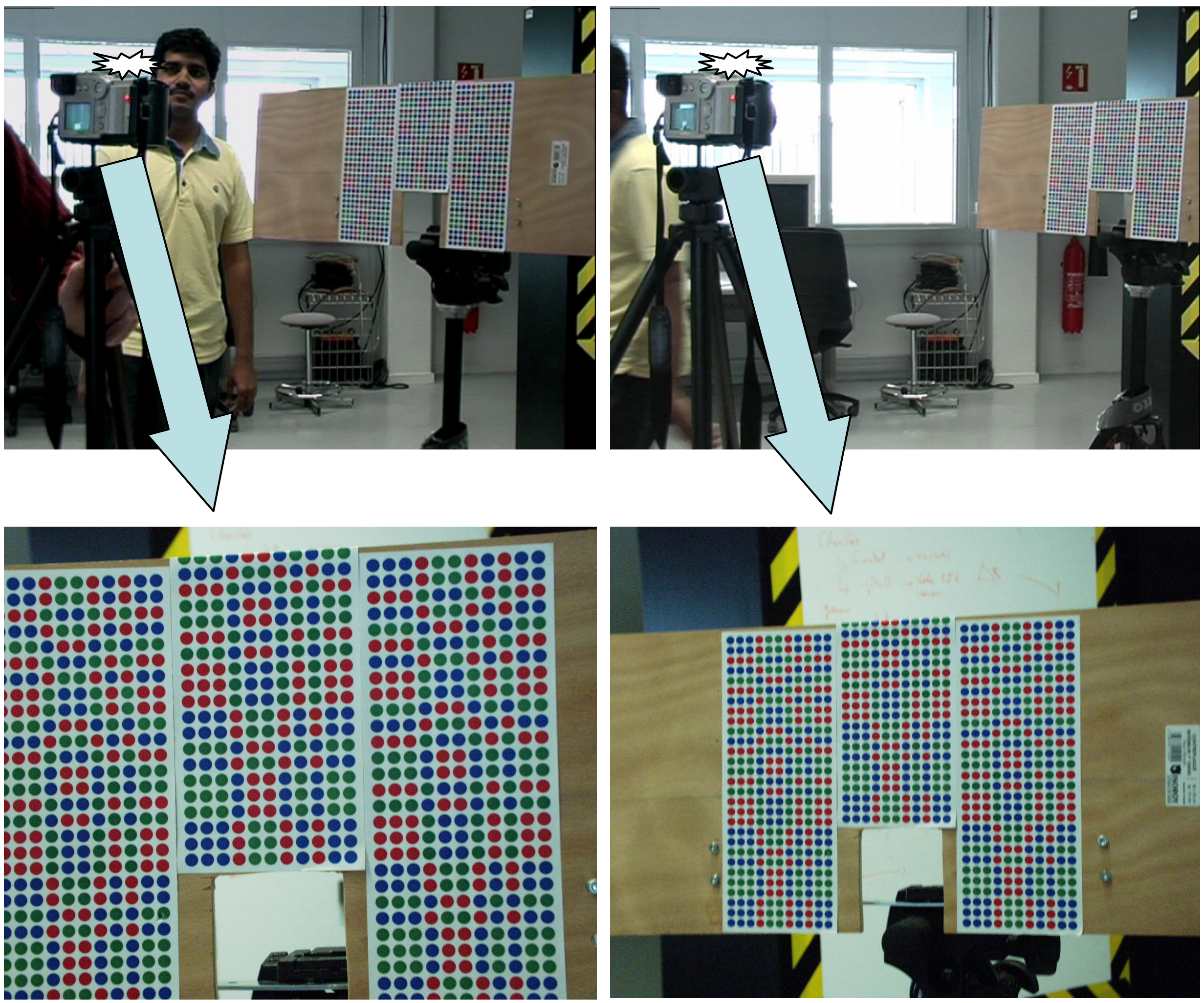
Input: images of calibration objects

Goal: compute projection ray for each pixel, in some 3D coordinate system

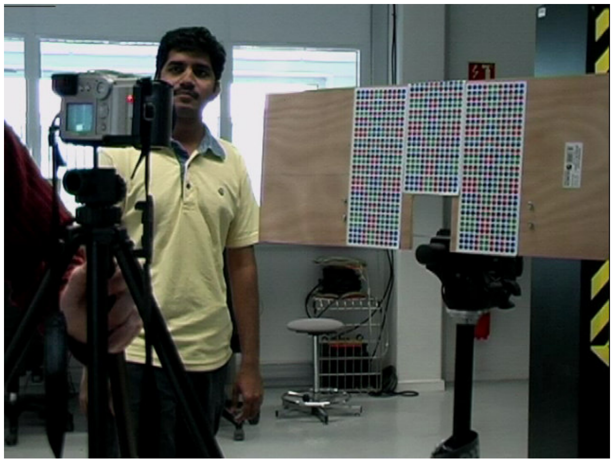
- General approach applicable for non-central cameras
- Variants for special cases (central and axial cameras)

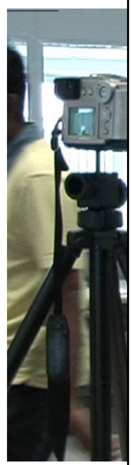
Non-parametric calibration — Basic idea

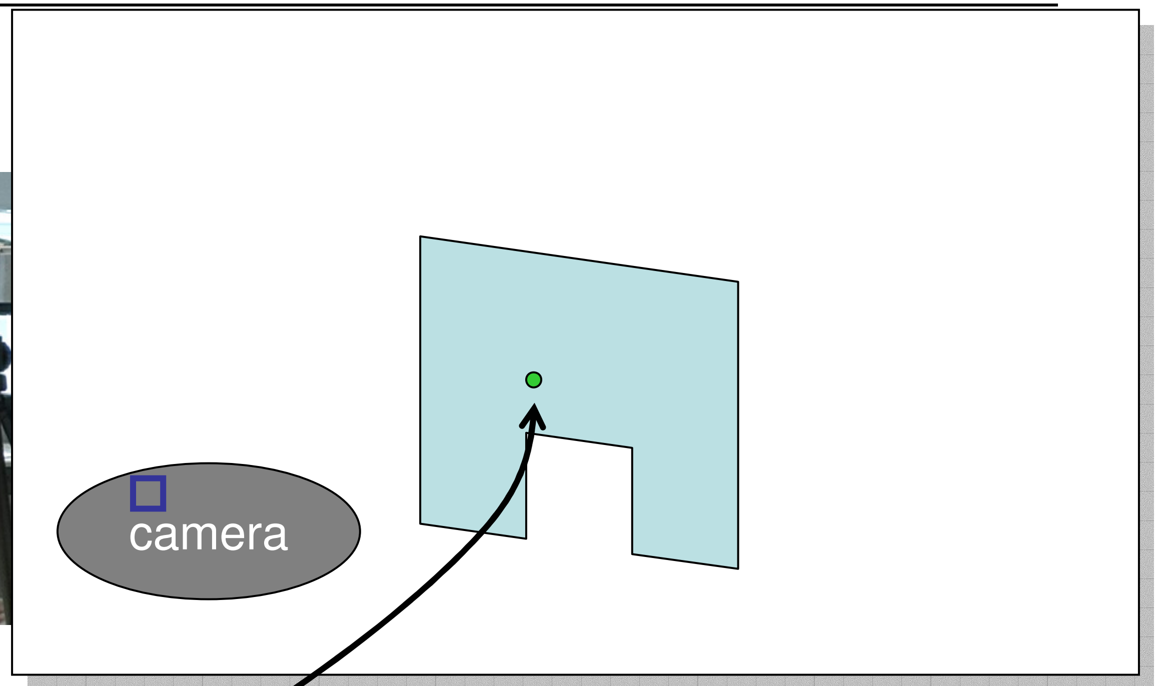
Approach using known motion: [Gremban-etal-ICRA'88,Champleboux-etal-ICRA'92, Grossberg-Nayar-ICCV'01]

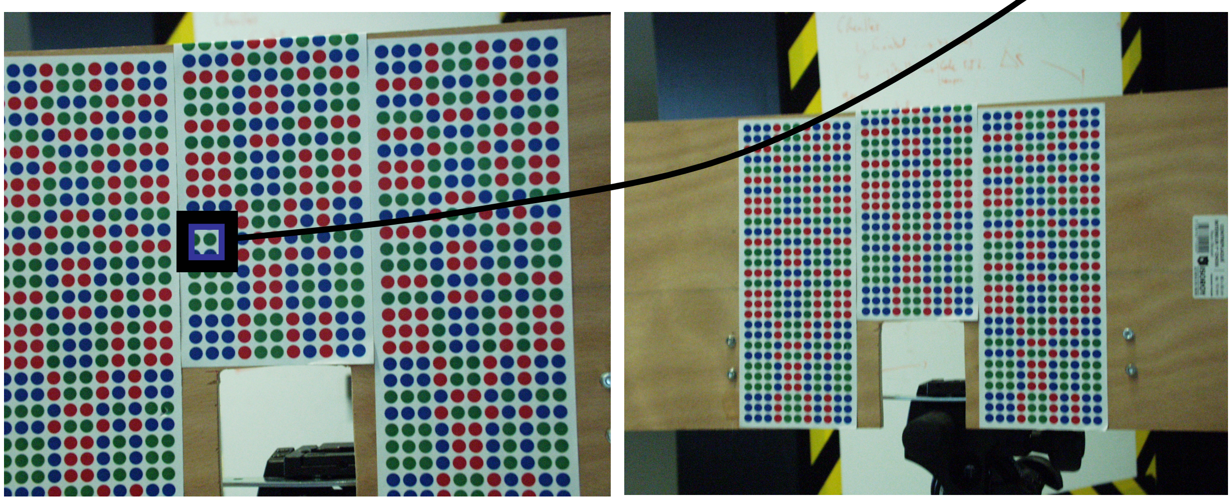


Approach using known motion:

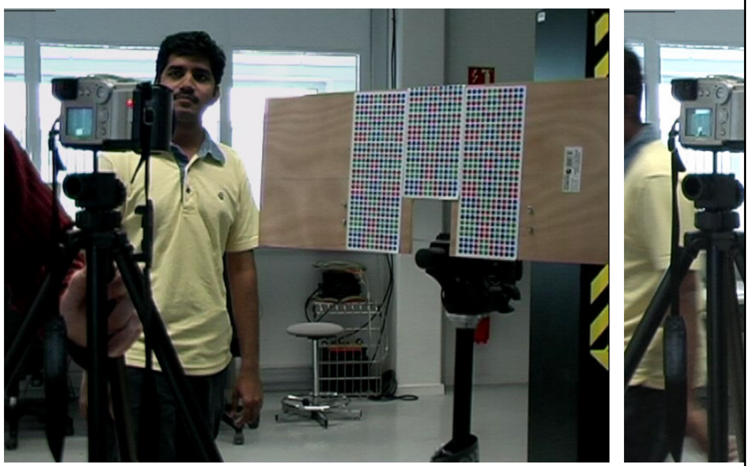


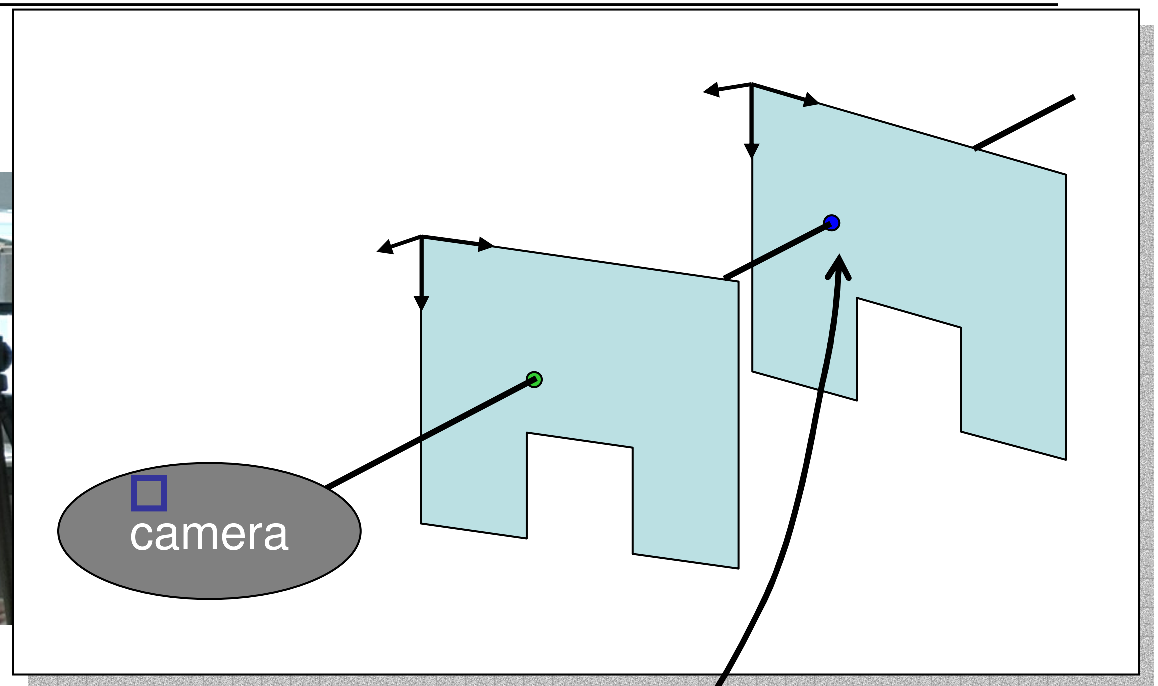


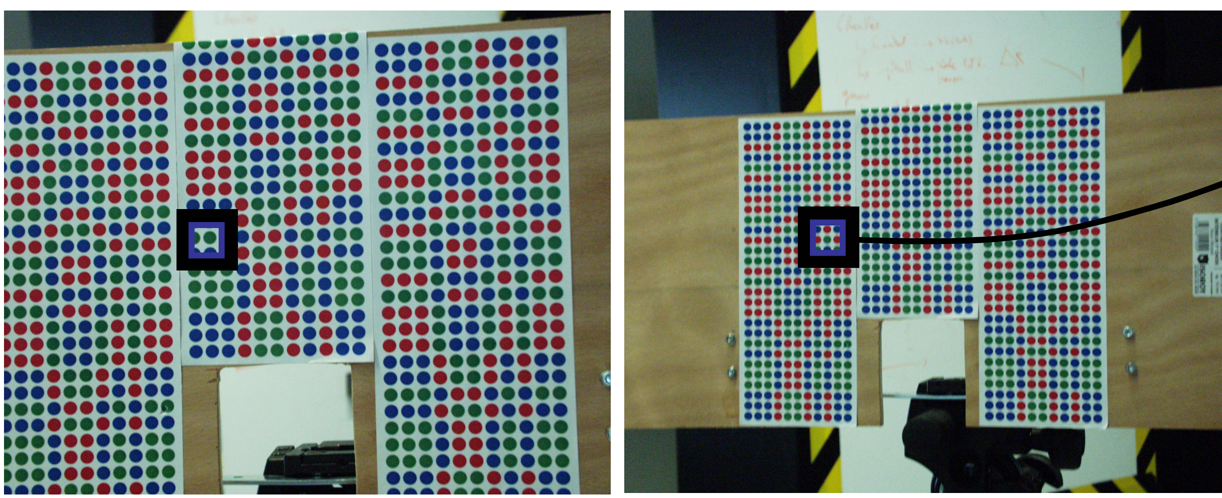




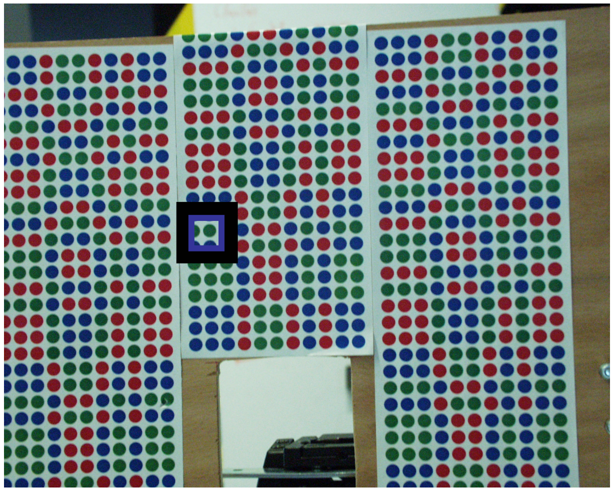
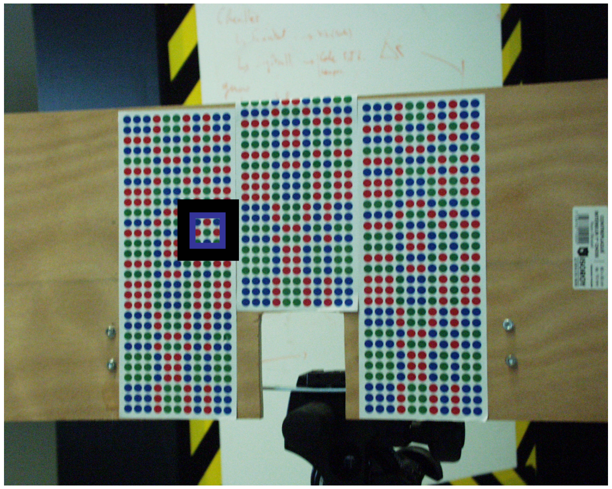
Approach using known motion:







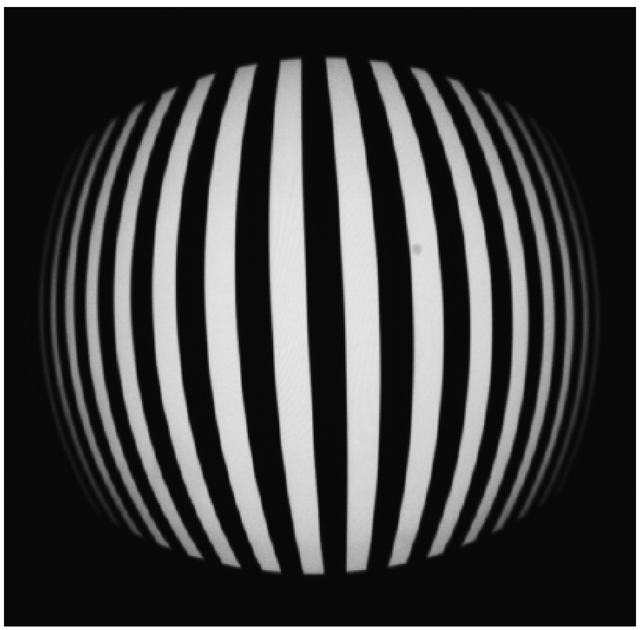
Using color coded grid:

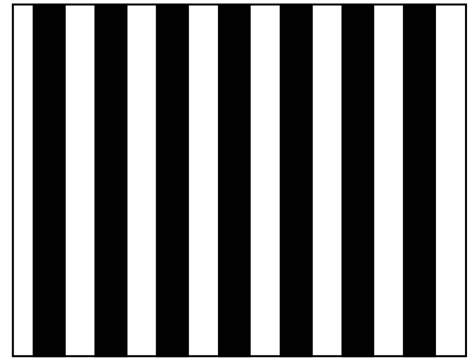



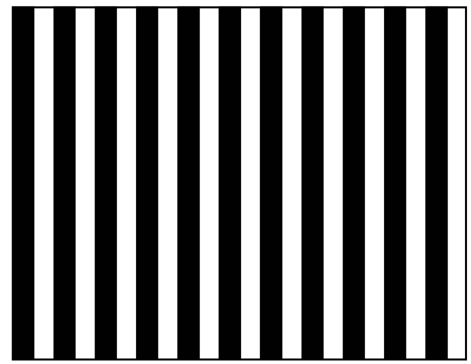
- Sparse matches, only for center pixels of circular targets
- We interpolate, for example using an homography:
  - for a pixel *p*, determine 4 closest pixels that have a match
  - compute 2D homography between these 4 image points and the matched points on the planar grid
  - apply this homography to compute point on grid that matches *p*

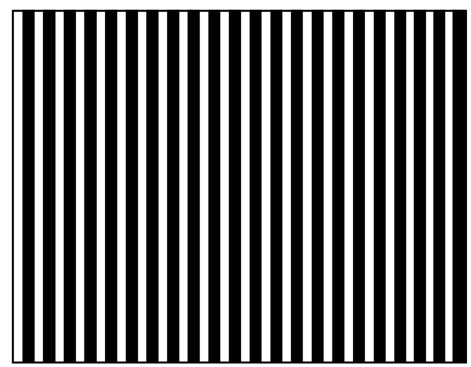
Better: structured light, e.g. acquiring images of a flat screen displaying a series of Gray code images (series of vertical and horizontal stripe patterns)

- Each screen pixel has its own unique sequence of black-white successions
- Dense matching between image and calibration grid (screen)



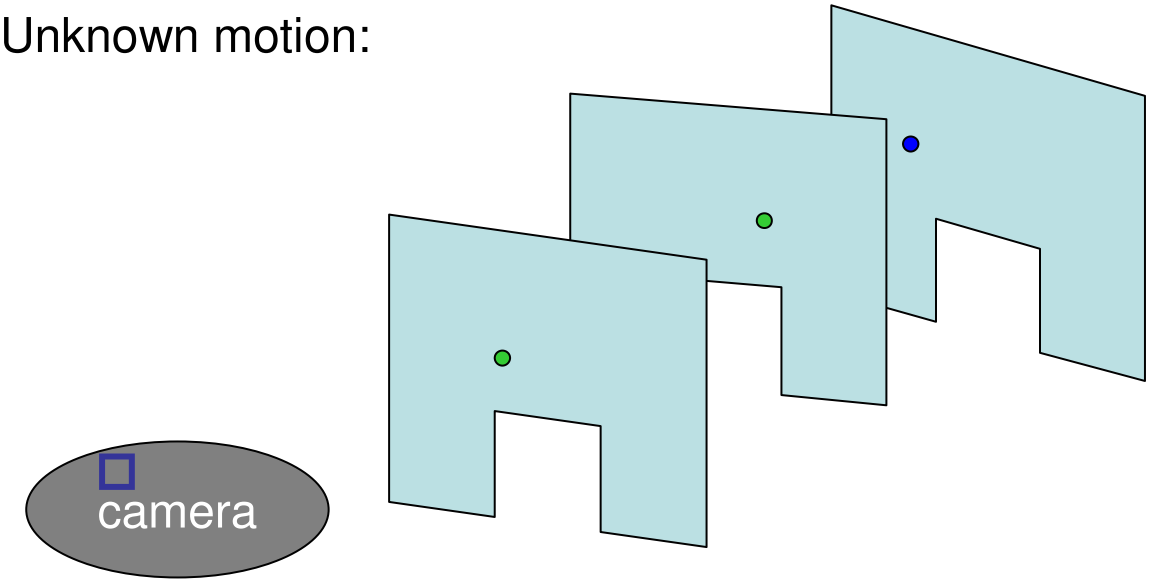


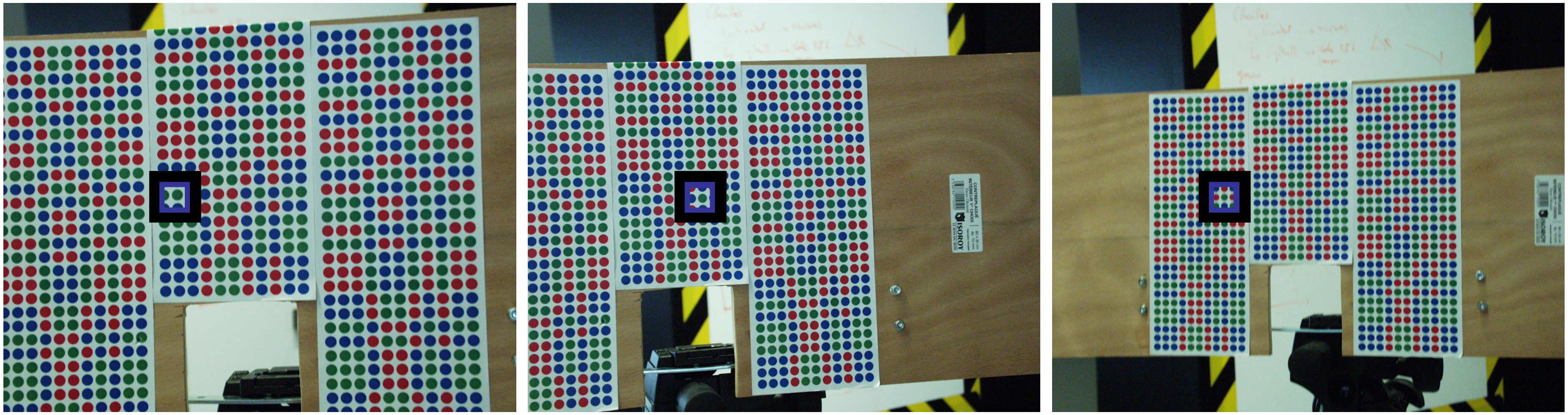




[Sturm-Ramalingam-ECCV'04]

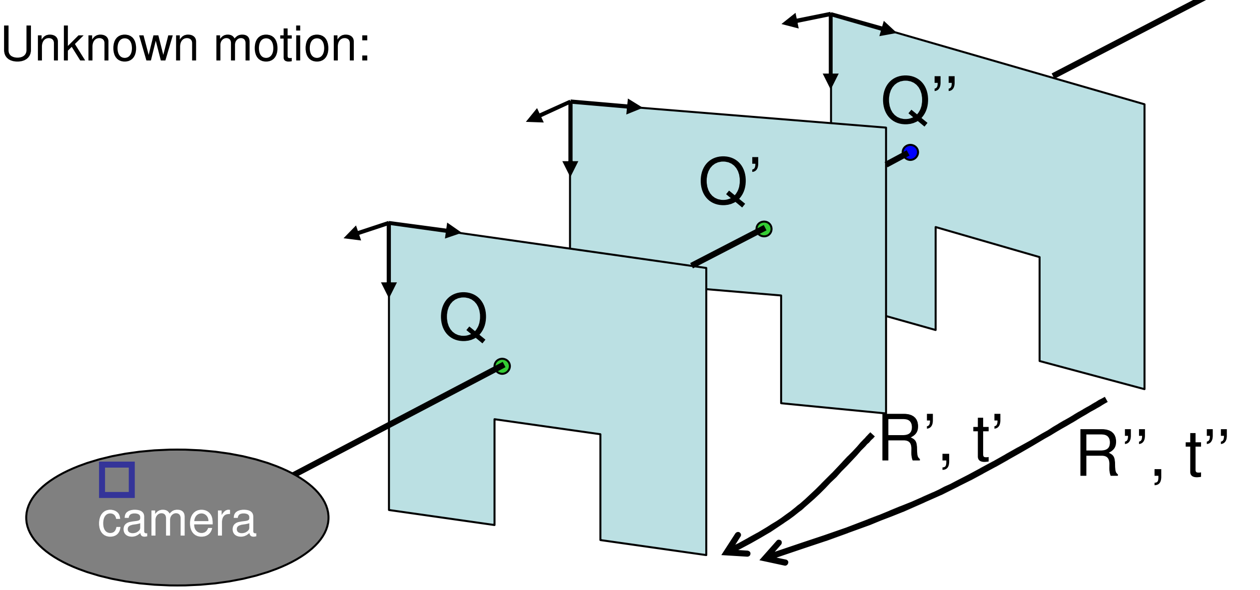
Unknown motion:



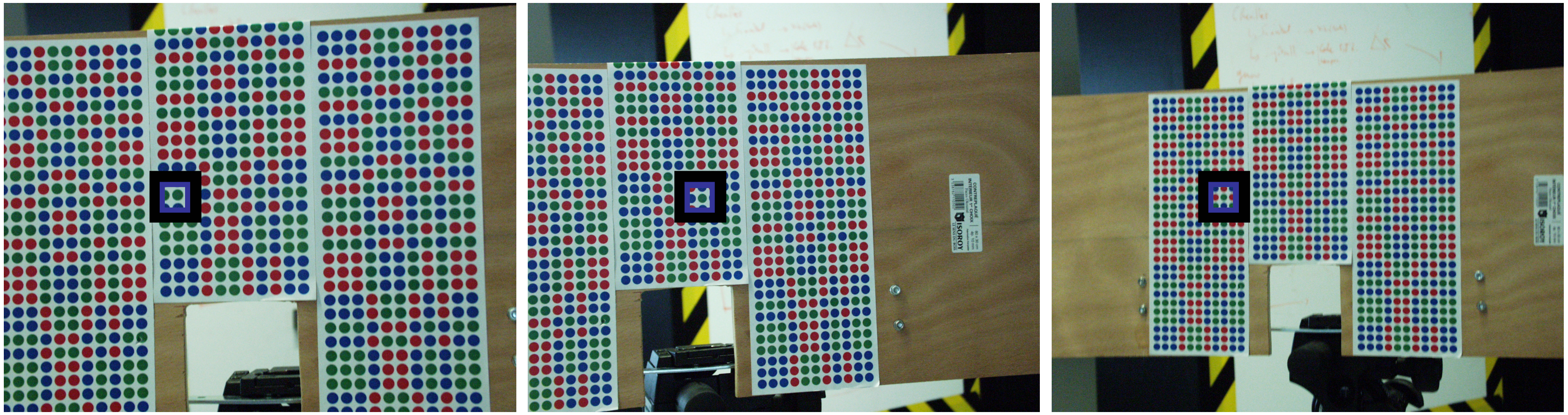


[Sturm-Ramalingam-ECCV'04]

Unknown motion:

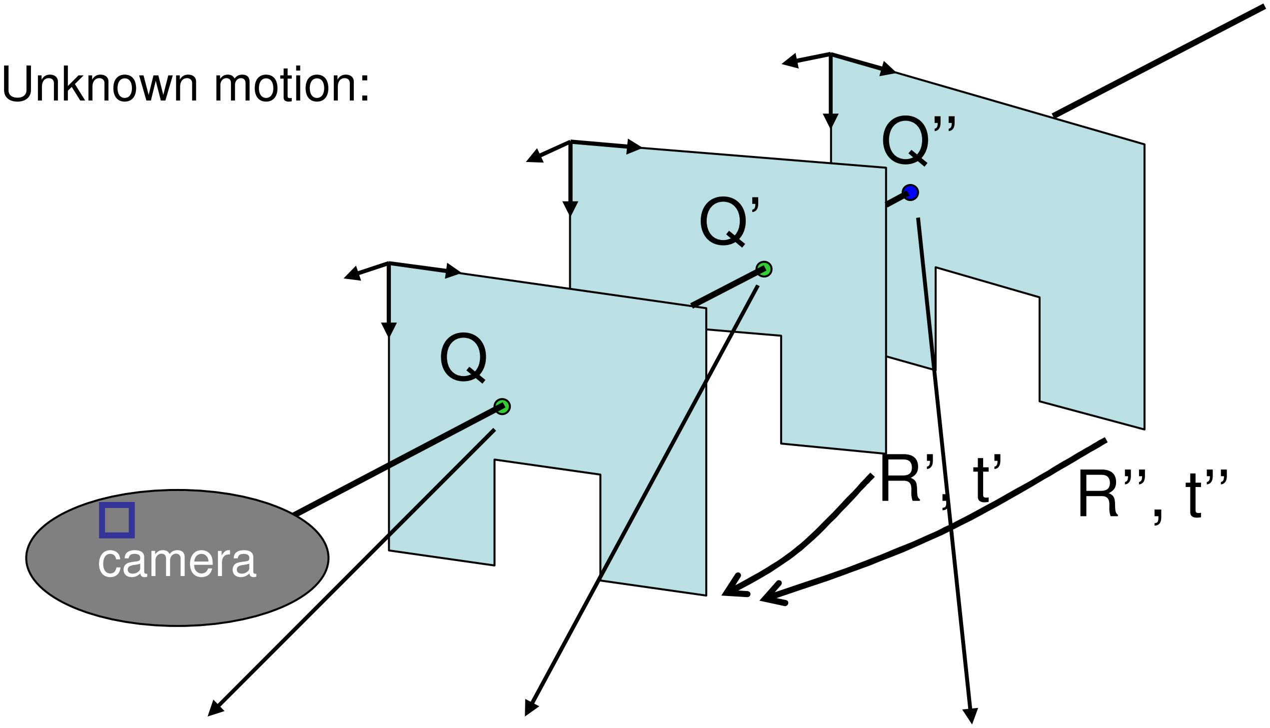


Estimate motions that make points collinear



[Sturm-Ramalingam-ECCV'04]

Unknown motion:

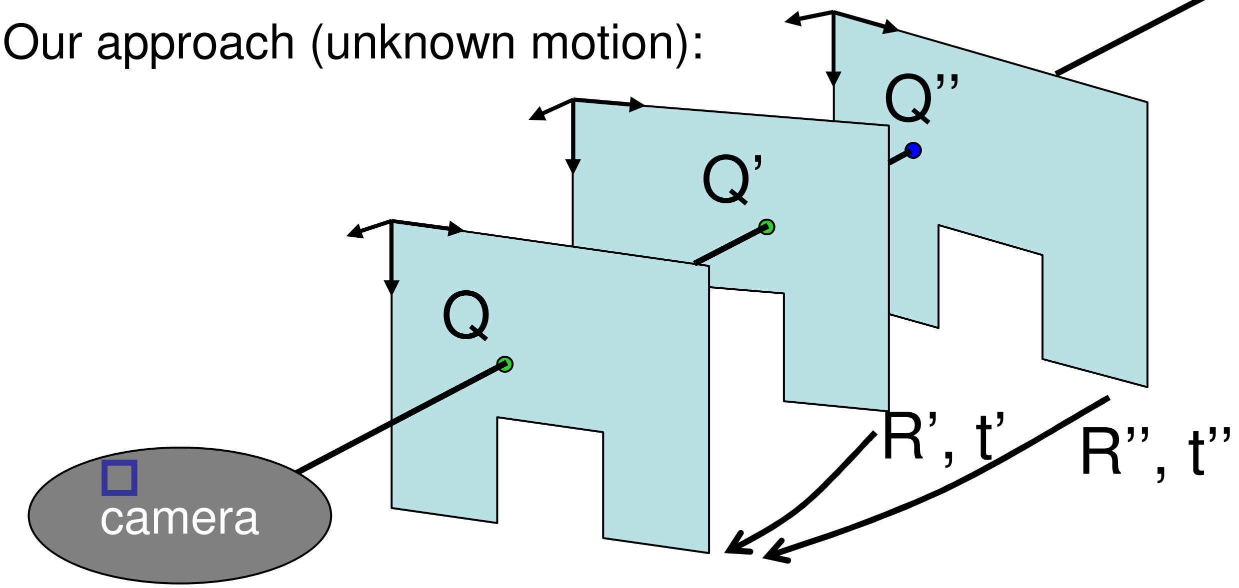


Estimate motions that make points collinear

$$\begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \end{pmatrix} \quad \begin{pmatrix} R' \begin{pmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{pmatrix} + t' Q'_4 \\ Q'_4 \end{pmatrix} \quad \begin{pmatrix} R'' \begin{pmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{pmatrix} + t'' Q''_4 \\ Q''_4 \end{pmatrix}$$

[Sturm-Ramalingam-ECCV'04]

Our approach (unknown motion):



Estimate motions that make points collinear

$$\begin{pmatrix} \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \end{matrix} & R'\begin{pmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{pmatrix} + t'\,Q'_4 & R''\begin{pmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{pmatrix} + t''\,Q''_4 \\ Q_4 & Q'_4 & Q''_4 \end{pmatrix}_{4\text{x}3} \longrightarrow \text{rank} < 3$$

$$\begin{pmatrix} Q_1 & R'\begin{bmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{bmatrix} + t' Q'_4 & R''\begin{bmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{bmatrix} + t'' Q''_4 \\ Q_2 & & \\ Q_3 & & \\ Q_4 & Q'_4 & Q''_4 \end{pmatrix}_{4x3} \longrightarrow \text{rank} < 3$$

Non-parametric calibration

$$\begin{pmatrix} \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \end{matrix} & \begin{matrix} R' \begin{bmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{bmatrix} + t' Q'_4 \\ Q'_4 \end{matrix} & \begin{matrix} R'' \begin{bmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{bmatrix} + t'' Q''_4 \\ Q''_4 \end{matrix} \end{pmatrix}_{4\text{x}3} \longrightarrow \text{rank} < 3$$

$$\Downarrow$$

$$\begin{pmatrix} \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \\ \cancel{Q_4} \end{matrix} & \begin{matrix} R' \begin{bmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{bmatrix} + t' Q'_4 \\ \cancel{Q'_4} \end{matrix} & \begin{matrix} R'' \begin{bmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{bmatrix} + t'' Q''_4 \\ \cancel{Q''_4} \end{matrix} \end{pmatrix} \longrightarrow \det = 0$$

$$\Downarrow$$

$$\det = \sum_{i,j,k=1}^{4} Q_i \, Q'_j \, Q''_k \, T_{i,j,k} = 0$$

a trifocal tensor

$$\begin{pmatrix} \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \end{matrix} & \begin{matrix} R' \begin{pmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{pmatrix} + t' Q'_4 \\ Q'_4 \end{matrix} & \begin{matrix} R'' \begin{pmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{pmatrix} + t'' Q''_4 \\ Q''_4 \end{matrix} \end{pmatrix}_{4\text{x}3} \longrightarrow \text{rank} < 3$$

$$\Downarrow$$

$$\begin{pmatrix} \begin{matrix} Q_1 \\ Q_2 \\ Q_3 \\ \cancel{Q_4} \end{matrix} & \begin{matrix} R' \begin{pmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \end{pmatrix} + t' Q'_4 \\ \cancel{Q'_4} \end{matrix} & \begin{matrix} R'' \begin{pmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{pmatrix} + t'' Q''_4 \\ \cancel{Q''_4} \end{matrix} \end{pmatrix} \longrightarrow \det = 0$$

$$\Downarrow$$

$$\det = \sum_{i,j,k=1}^{4} Q_i \, Q'_j \, Q''_k \, \underbrace{T_{i,j,k}}_{\text{a trifocal tensor}} = 0$$

Non-parametric calibration — General approach

4 such tensors exist, striking out one row in turn:

$$\begin{pmatrix} Q_1 & & \\ Q_2 & R'\begin{bmatrix}Q'_1\\ Q'_2\\ Q'_3\end{bmatrix} + t'Q'_4 & R''\begin{bmatrix}Q''_1\\ Q''_2\\ Q''_3\end{bmatrix} + t''Q''_4 \\ Q_3 & & \\ \cancel{Q_4} & \cancel{Q'_4} & \cancel{Q''_4} \end{pmatrix}_{4\mathrm{x}3} \longrightarrow \det = 0$$

Each one has a particular structure, see the following slide for two examples

Non-parametric calibration — General approach

i	C_i	V_i	W_i
1	$Q_1Q'_1Q''_4$	0	R'_{31}
2	$Q_1Q'_2Q''_4$	0	R'_{32}
3	$Q_1Q'_3Q''_4$	0	R'_{33}
4	$Q_1Q'_4Q''_1$	0	$-R''_{31}$
5	$Q_1Q'_4Q''_2$	0	$-R''_{32}$
6	$Q_1Q'_4Q''_3$	0	$-R''_{33}$
7	$Q_1Q'_4Q''_4$	0	$t'_3 - t''_3$
8	$Q_2Q'_1Q''_4$	R'_{31}	0
9	$Q_2Q'_2Q''_4$	R'_{32}	0
10	$Q_2Q'_3Q''_4$	R'_{33}	0
11	$Q_2Q'_4Q''_1$	$-R''_{31}$	0
12	$Q_2Q'_4Q''_2$	$-R''_{32}$	0
13	$Q_2Q'_4Q''_3$	$-R''_{33}$	0
14	$Q_2Q'_4Q''_4$	$t'_3 - t''_3$	0
15	$Q_3Q'_1Q''_4$	$-R'_{21}$	$-R'_{11}$
16	$Q_3Q'_2Q''_4$	$-R'_{22}$	$-R'_{12}$
17	$Q_3Q'_3Q''_4$	$-R'_{23}$	$-R'_{13}$
18	$Q_3Q'_4Q''_1$	R''_{21}	R''_{11}
19	$Q_3Q'_4Q''_2$	R''_{22}	R''_{12}
20	$Q_3Q'_4Q''_3$	R''_{23}	R''_{13}
21	$Q_3Q'_4Q''_4$	$t''_2 - t'_2$	$t''_1 - t'_1$
22	$Q_4Q'_1Q''_1$	$R'_{21}R''_{31} - R''_{21}R'_{31}$	$R'_{11}R''_{31} - R''_{11}R'_{31}$
23	$Q_4Q'_1Q''_2$	$R'_{21}R''_{32} - R''_{22}R'_{31}$	$R'_{11}R''_{32} - R''_{12}R'_{31}$
24	$Q_4Q'_1Q''_3$	$R'_{21}R''_{33} - R''_{23}R'_{31}$	$R'_{11}R''_{33} - R''_{13}R'_{31}$
25	$Q_4Q'_1Q''_4$	$R'_{21}t''_3 - R'_{31}t''_2$	$R'_{11}t''_3 - R'_{31}t''_1$
26	$Q_4Q'_2Q''_1$	$R'_{22}R''_{31} - R''_{21}R'_{32}$	$R'_{12}R''_{31} - R''_{11}R'_{32}$
27	$Q_4Q'_2Q''_2$	$R'_{22}R''_{32} - R''_{22}R'_{32}$	$R'_{12}R''_{32} - R''_{12}R'_{32}$
28	$Q_4Q'_2Q''_3$	$R'_{22}R''_{33} - R''_{23}R'_{32}$	$R'_{12}R''_{33} - R''_{13}R'_{32}$
29	$Q_4Q'_2Q''_4$	$R'_{22}t''_3 - R'_{32}t''_2$	$R'_{12}t''_3 - R_{32}t''_1$
30	$Q_4Q'_3Q''_1$	$R'_{23}R''_{31} - R''_{21}R'_{33}$	$R'_{13}R''_{31} - R''_{11}R'_{33}$
31	$Q_4Q'_3Q''_2$	$R'_{23}R''_{32} - R''_{22}R'_{33}$	$R'_{13}R''_{32} - R''_{12}R'_{33}$
32	$Q_4Q'_3Q''_3$	$R'_{23}R''_{33} - R''_{23}R'_{33}$	$R'_{13}R''_{33} - R''_{13}R'_{33}$
33	$Q_4Q'_3Q''_4$	$R'_{23}t''_3 - R'_{33}t''_2$	$R'_{13}t''_3 - R'_{33}t''_1$
34	$Q_4Q'_4Q''_1$	$R''_{31}t'_2 - R''_{21}t'_3$	$R''_{31}t'_1 - R''_{11}t'_3$
35	$Q_4Q'_4Q''_2$	$R''_{32}t'_2 - R''_{22}t'_3$	$R''_{32}t'_1 - R''_{12}t'_3$
36	$Q_4Q'_4Q''_3$	$R''_{33}t'_2 - R''_{23}t'_3$	$R''_{33}t'_1 - R''_{13}t'_3$
37	$Q_4Q'_4Q''_4$	$t'_2t''_3 - t'_3t''_2$	$t'_1t''_3 - t''_1t'_3$

Calibration algorithm:

(1) Take images of calibration object in different poses

(2) 2D-3D matching (pixels to points on object)

(3) Estimation of tensors, based on linear equations

$$\sum_{i,j,k=1}^{4} Q_i \, Q'_j \, Q''_k \, T_{i,j,k} = 0$$

and taking into account the tensors' structure (e.g. coefficients that are zero)

(4) Extraction of motion parameters from tensors:

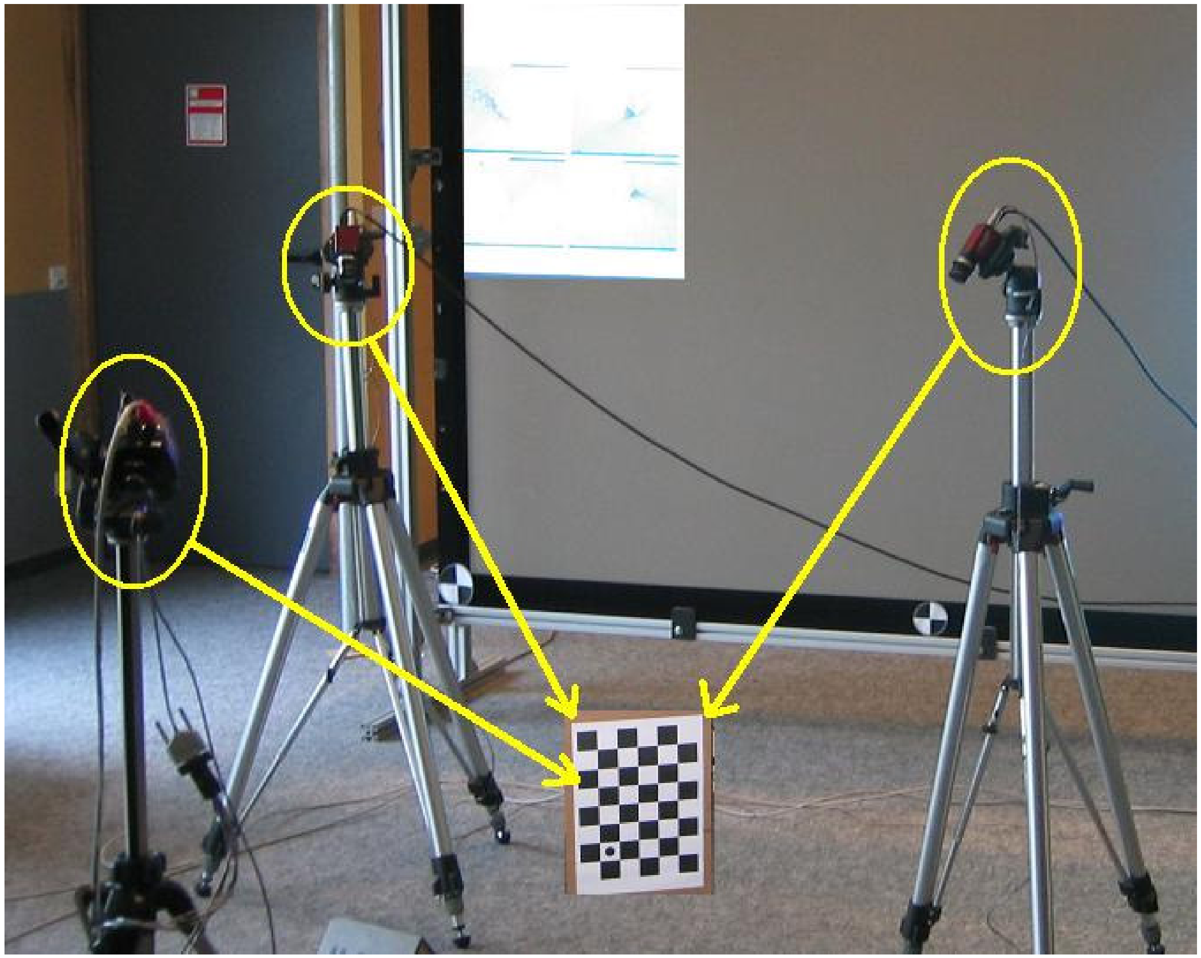
- some can be directly read off (some rotation coefficients, cf. previous slide)
- others can be computed using orthonormality constraints on R' and R''

(5) Put calibration grids in same 3D coordinate system

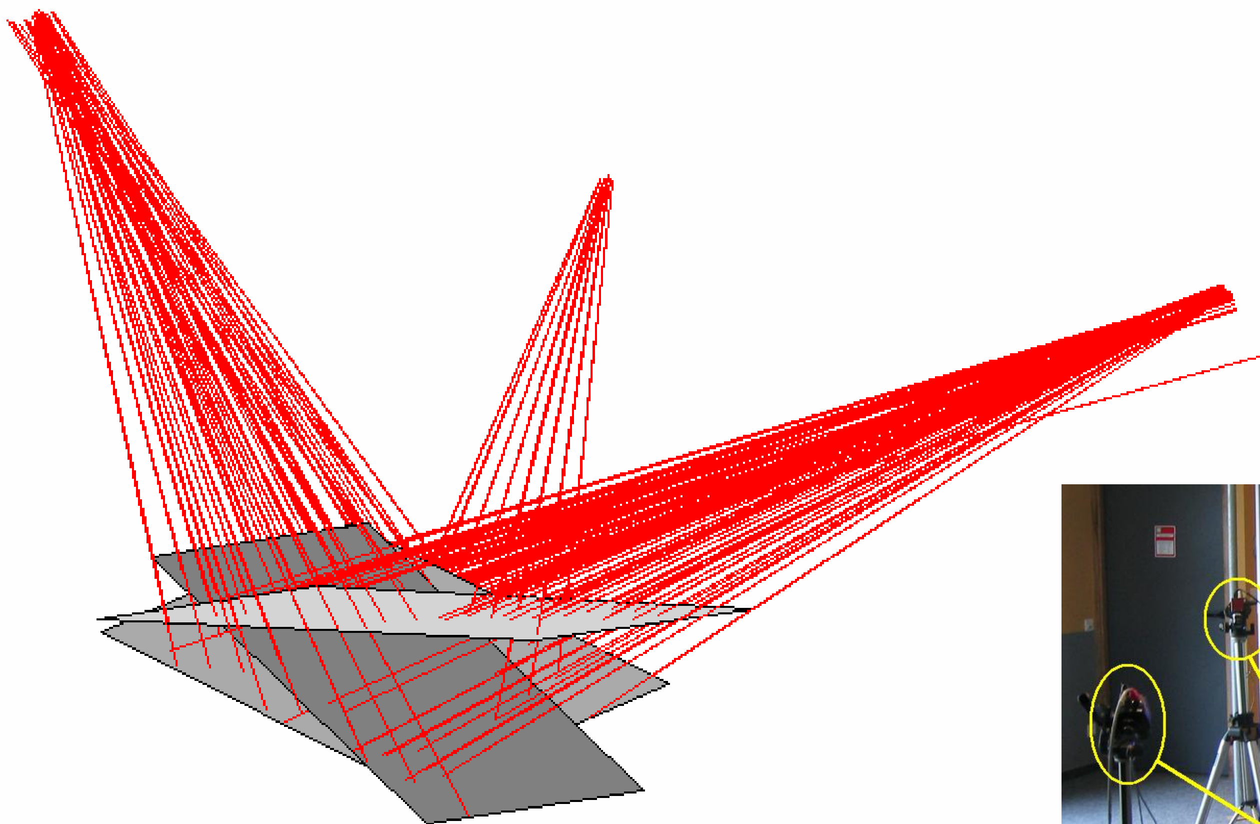
(6) Compute projection rays: for each pixel join the associated calibration points

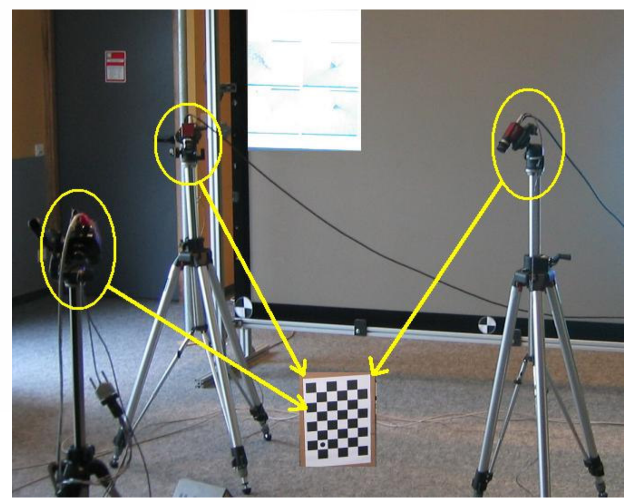
(7) Bundle adjustment

Results for non-central camera
(multi-camera system, considered as single non-central camera):

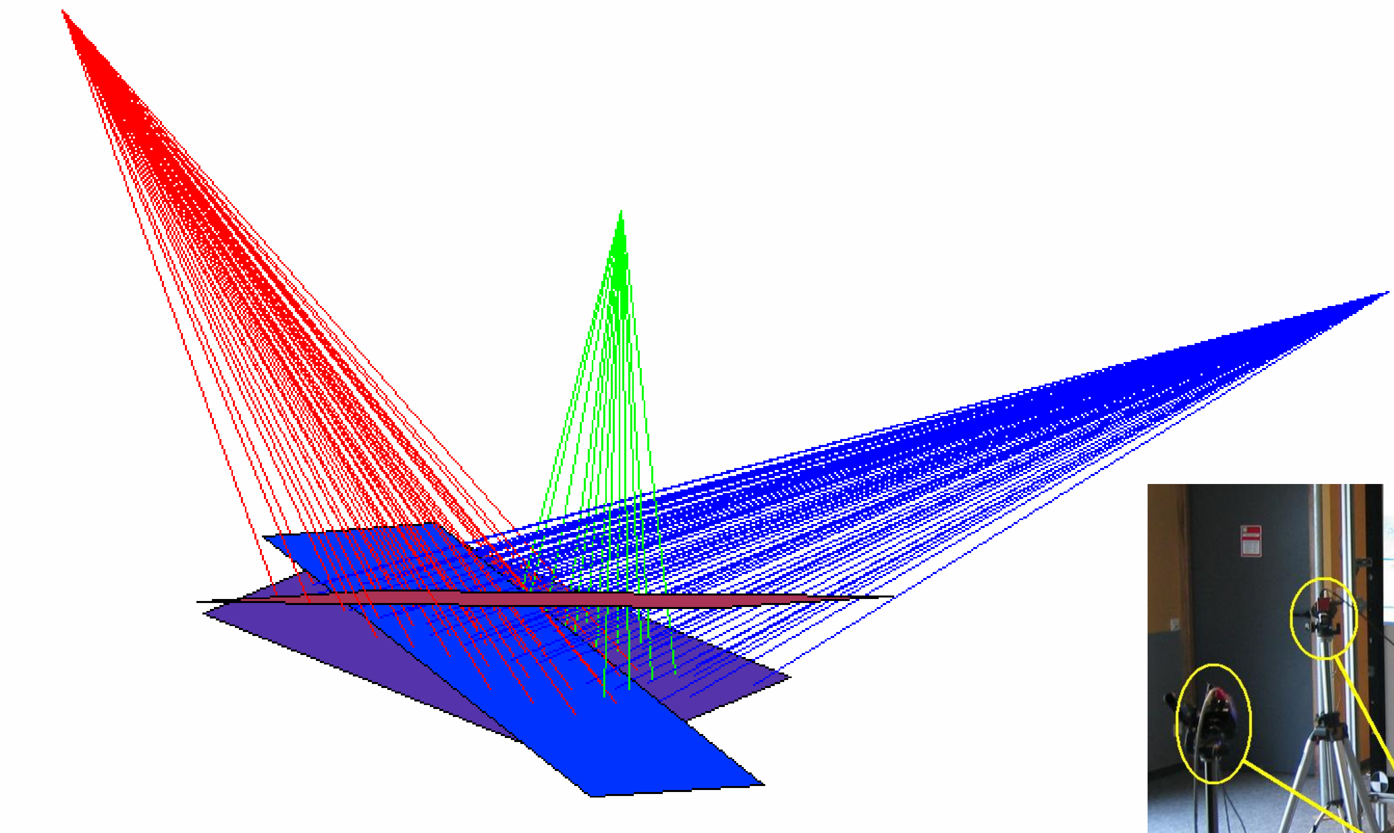


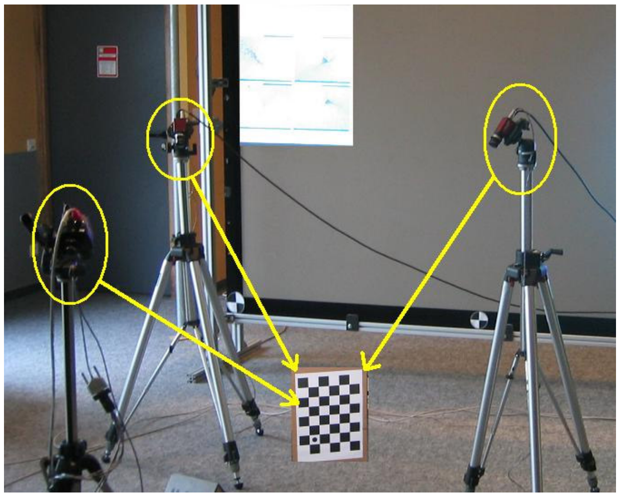
Results for non-central camera:





Results for non-central camera: after constraining rays into central clusters





Non-parametric calibration — Discussion

Intermediate discussion:

- the approach is designed for 3D calibration objects

  → ***variant for using*** **planar** ***calibration objects***

- this approach uses *exactly* 3 images
  - only pixels covered by all 3 images of the calibration grid are calibrated
    → especially with large field of view, difficult to calibrate whole image
  - results may not be highly accurate

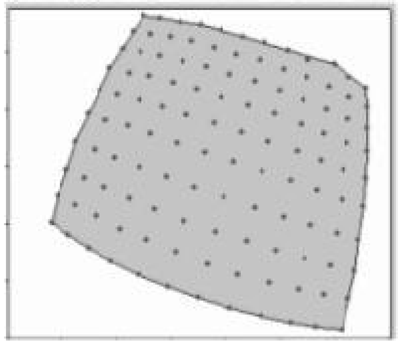
  → ***methods for using multiple images***

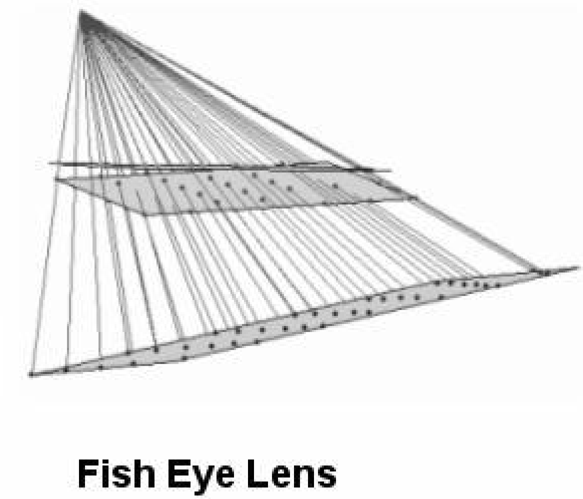
- the approach allows to calibrate non-central cameras!
- BUT: if used with images acquired by central camera
  - tensors are not computed uniquely (linear equation system of too low rank)
    → calibration fails

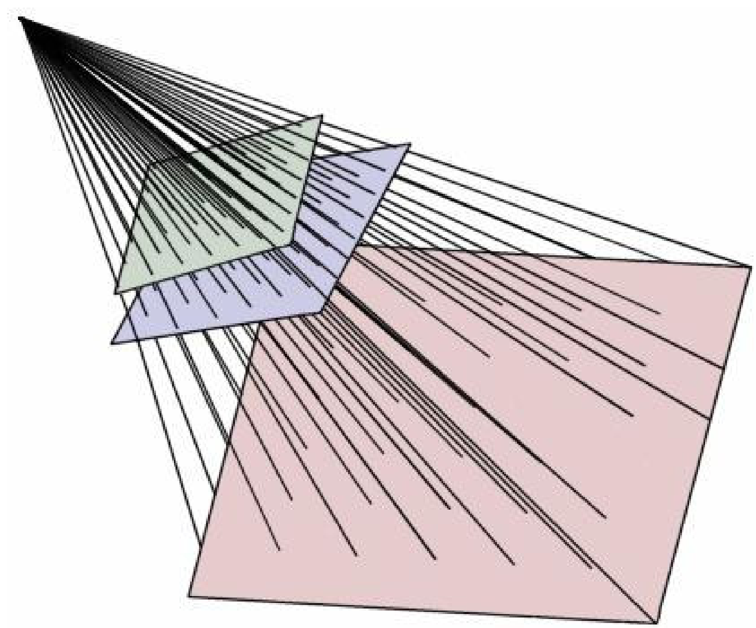
  → ***variant of the approach for central cameras and other special cases***

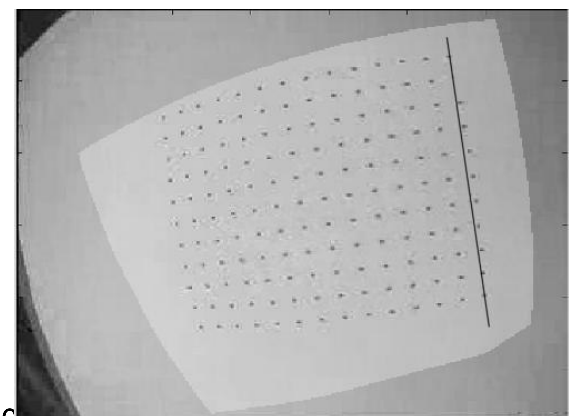
Non-parametric calibration — Approach for central model

Results for fisheye camera

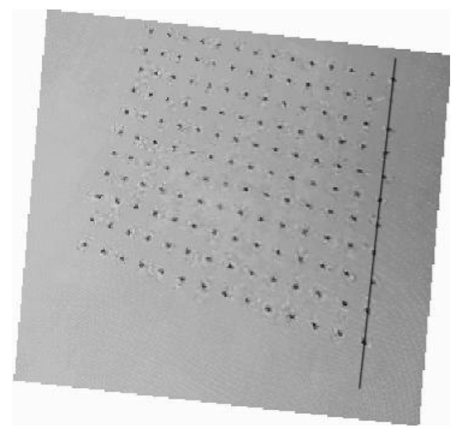






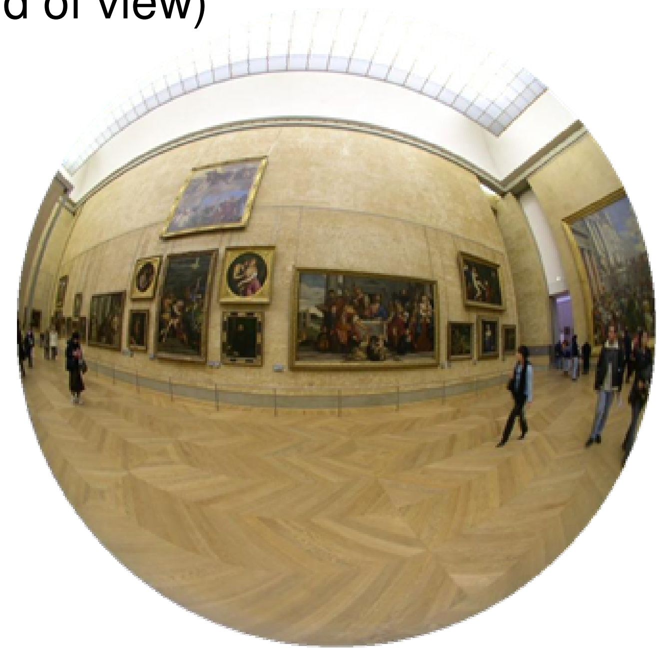


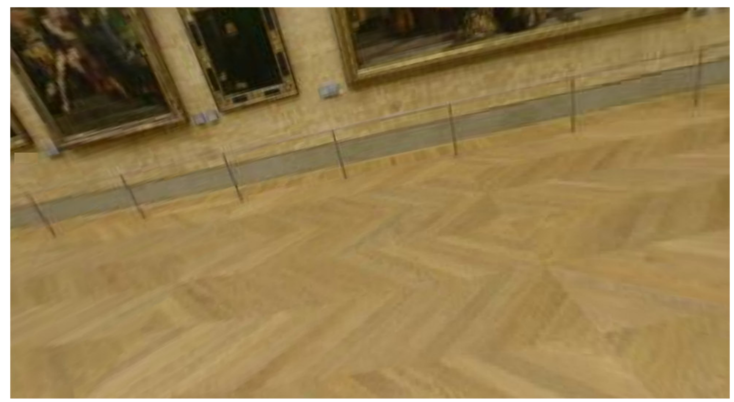
Distortion correction



Non-parametric calibration — Approach for central model

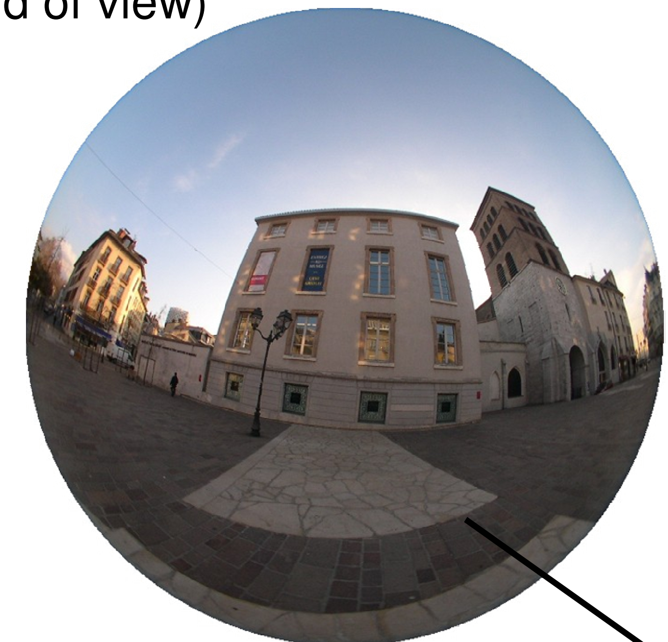
Results for fisheye camera
(183° field of view)

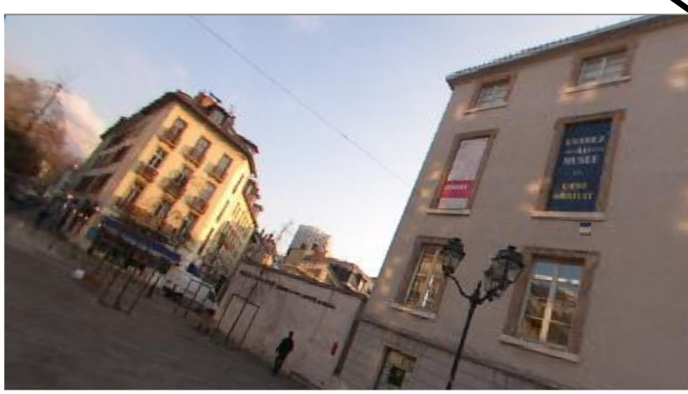




Non-parametric calibration   Approach for central model

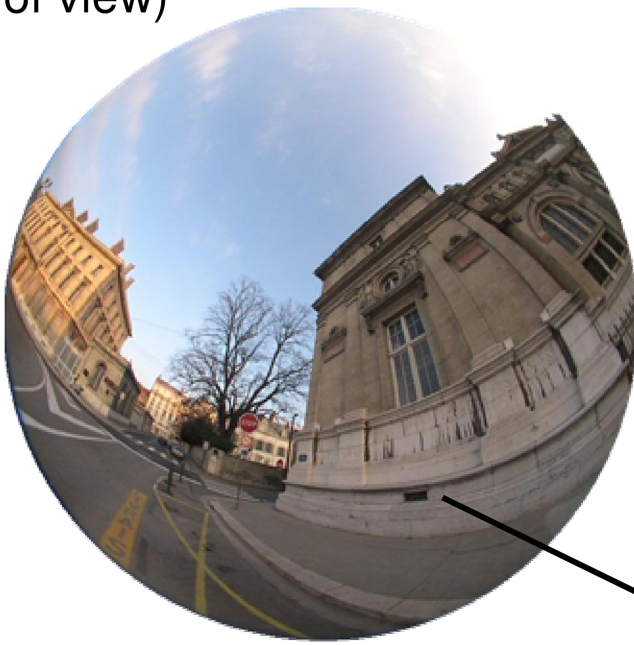
Results for fisheye camera
(183° field of view)

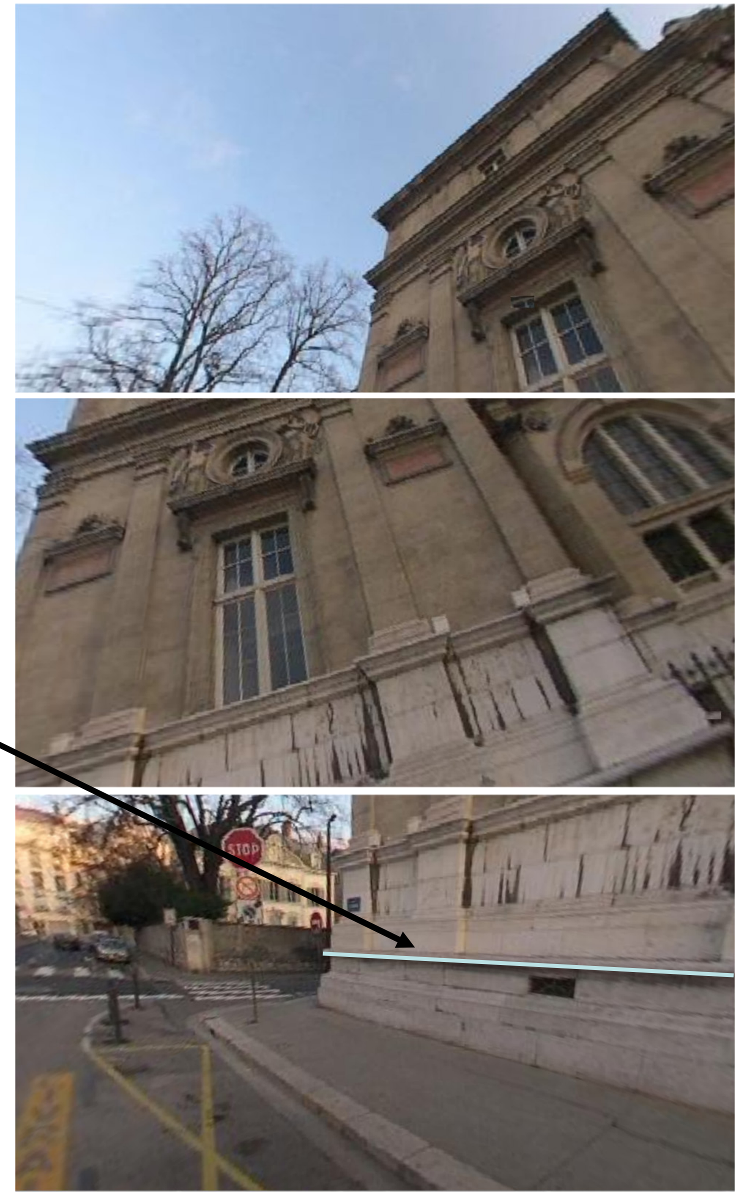




Non-parametric calibration   Approach for central model

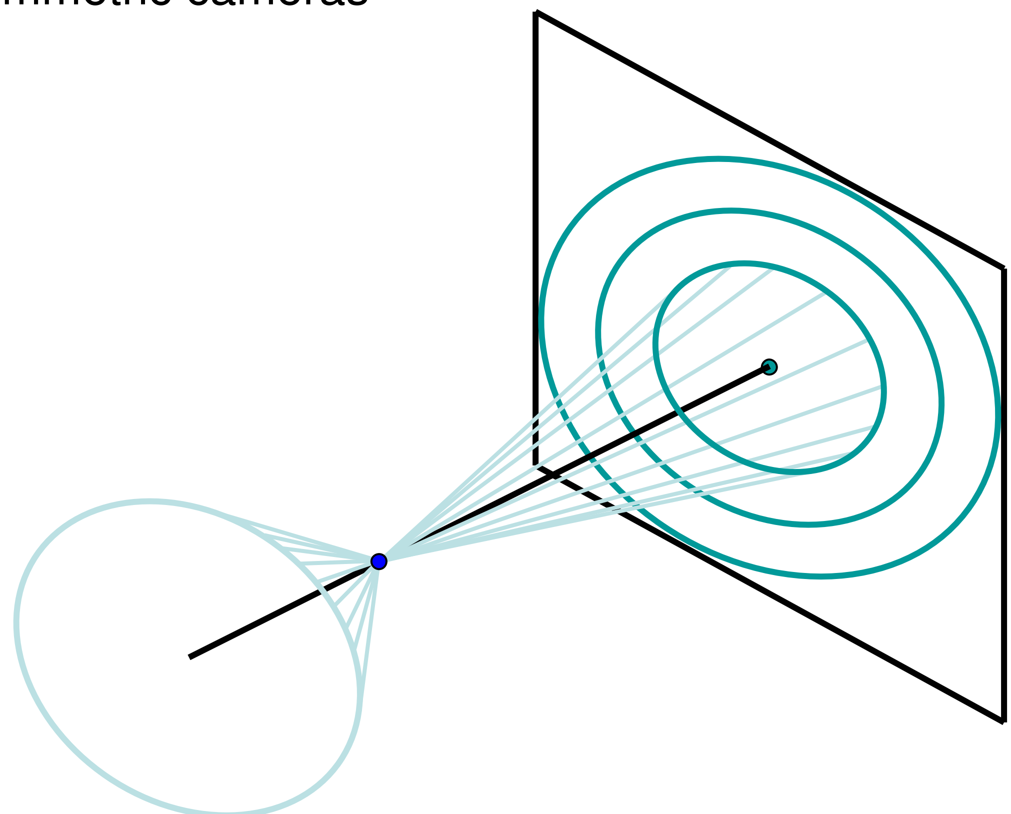
Results for fisheye camera
(183° field of view)





Non-parametric calibration Radially symmetric cameras

Interesting special case: radially symmetric cameras



Calibration:

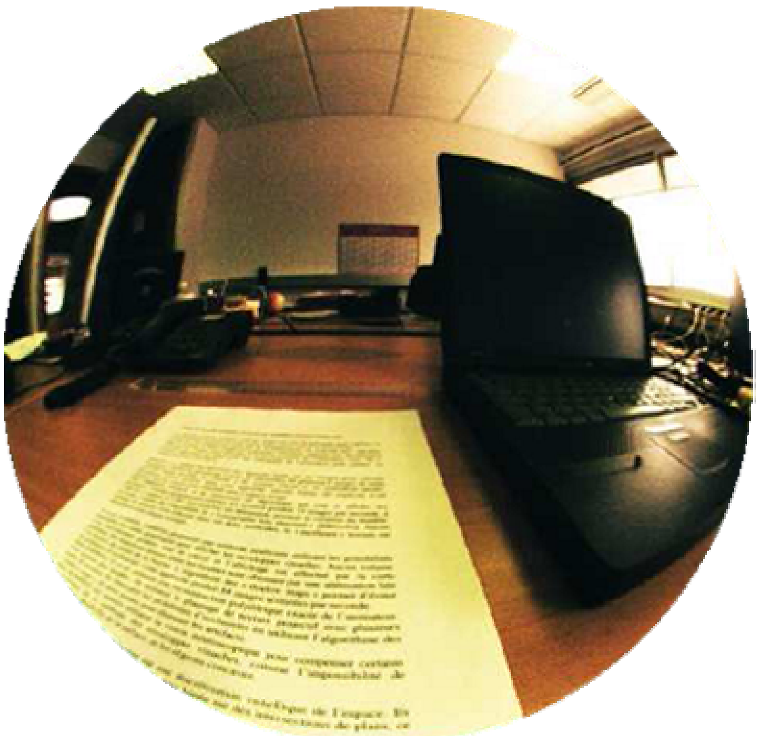
Computation of distortion center and
distortion function: *radius → view angle / focal length*

Note: each distortion circle ≡ perspective camera

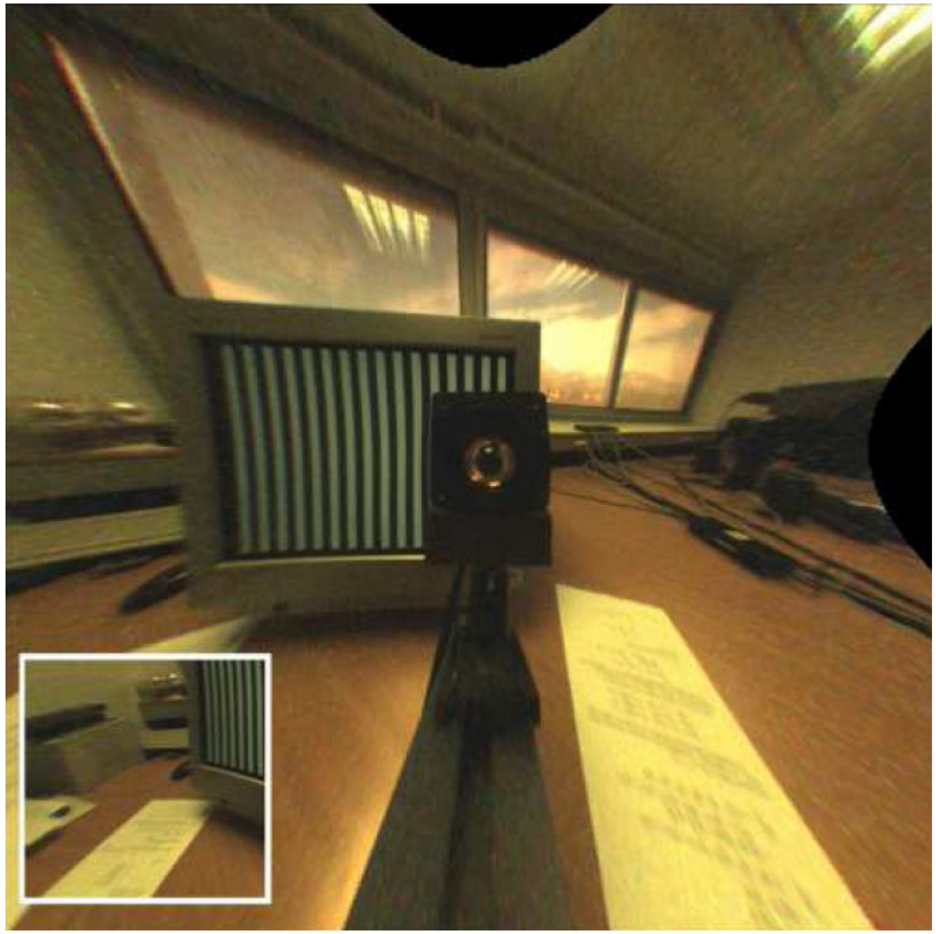
[Tardif-Sturm-OMNIVIS'05]

Non-parametric calibration Radially symmetric cameras

Result of distortion correction for fisheye



Result for homemade "Christmas camera"



Non-parametric calibration — Discussion

- General approach that allows to calibrate any camera
- Variants for central and axial camera modes
- Variants for using planar or 3D calibration objects

- How about stability?
  - Possible overfitting when calibrating “not very non-central cameras” with the general approach (result may be worse than with the central approach)
  - Stability depends on:
    - amount of “non-centrality”
    - number of images
    - accuracy of matches
  - If unstable:
    use more images, regularization, assumption of radial symmetry, …

Non-parametric calibration — Discussion

- Here, pixel-wise discretization of camera model
- Any other discretization (sub-pixel or super-pixel) is possible
- Trade-off between
  - potential accuracy of calibration (the finer the discretization, the better)
  - potential instability (the finer the discretization, the more unknowns…)

Contents

- Introduction
- General imaging models
- Non-parametric calibration and distortion correction
- Non-parametric self-calibration
- Structure-from-motion

Non-parametric self-calibration

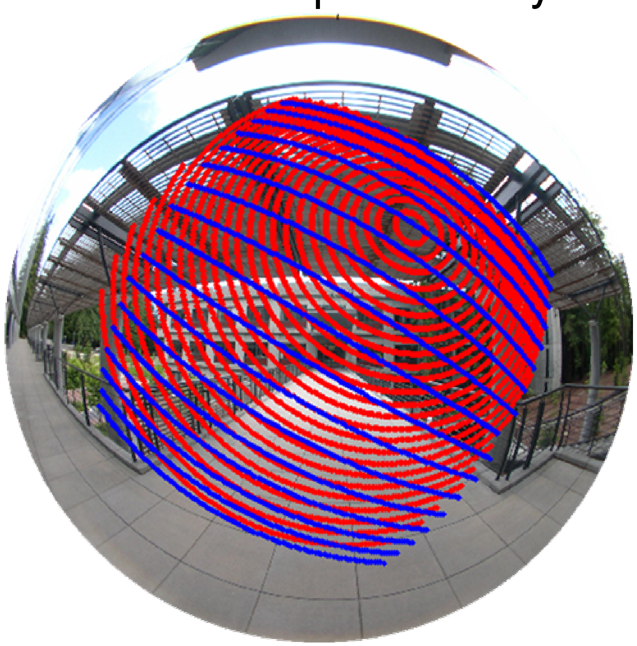
Self-calibration from several translational motions:

- Ray directions can be computed up to a projective transformation

  → amount of calibration knowledge is now equivalent to that of an uncalibrated perspective camera

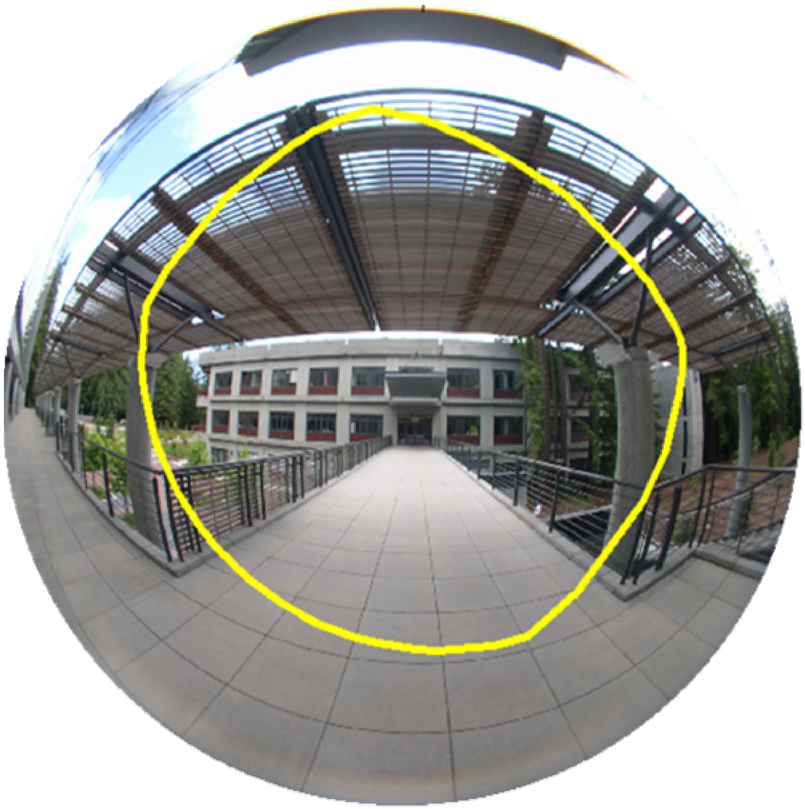
  → any self-calibration method for perspective cameras can be applied to complete the self-calibration

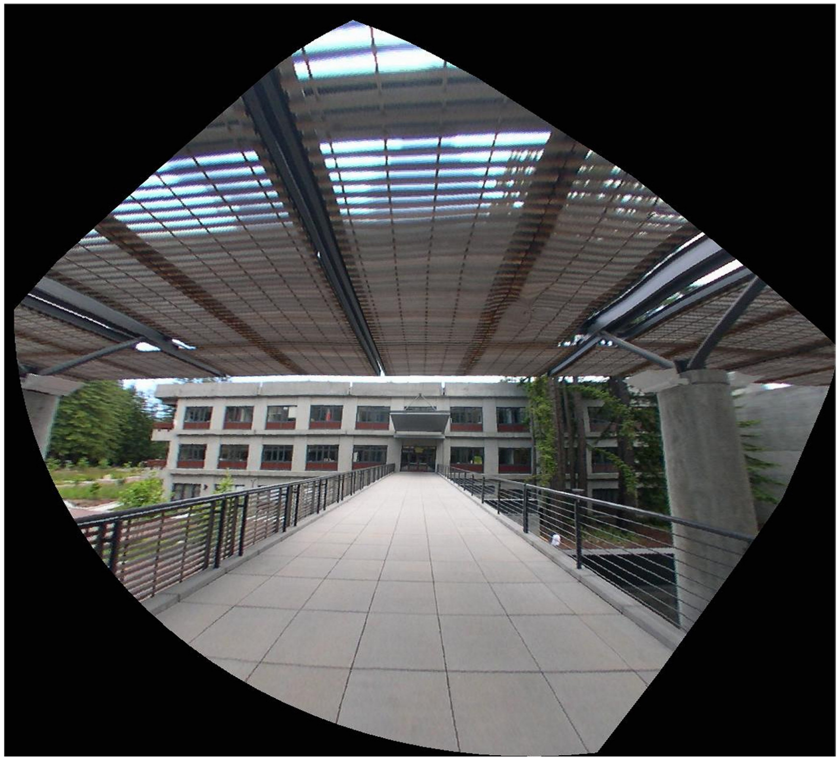
Complete self-calibration is possible by doing translational and rotational motions



Non-parametric self-calibration

Result of distortion correction using self-calibration result:





Non-parametric self-calibration

Result of distortion correction using self-calibration result:

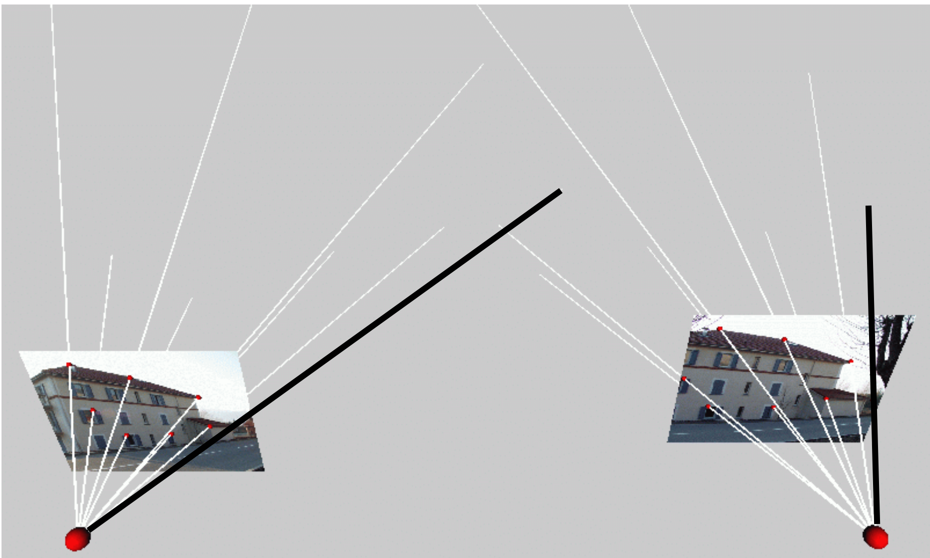
Contents

- Introduction
- General imaging models
- Non-parametric calibration and distortion correction
- Non-parametric self-calibration
- Structure-from-motion

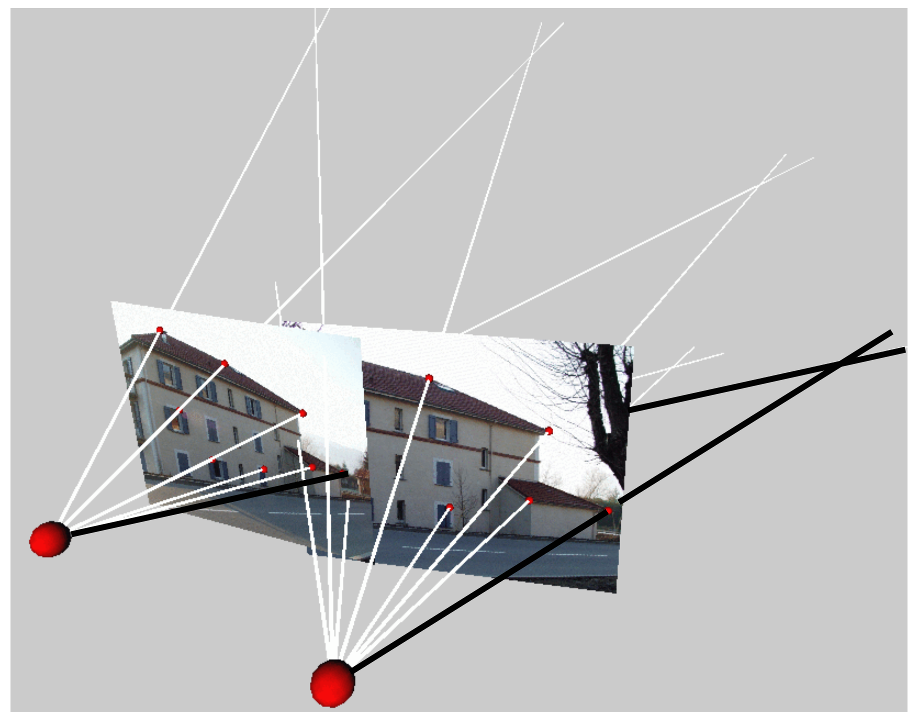
Motivation:

- Many different SfM algorithms (pose, motion, triangulation, …) exist, for different camera types
- But, in principle, if calibrated cameras are considered, one single approach for each SfM problem is sufficient, for all camera types

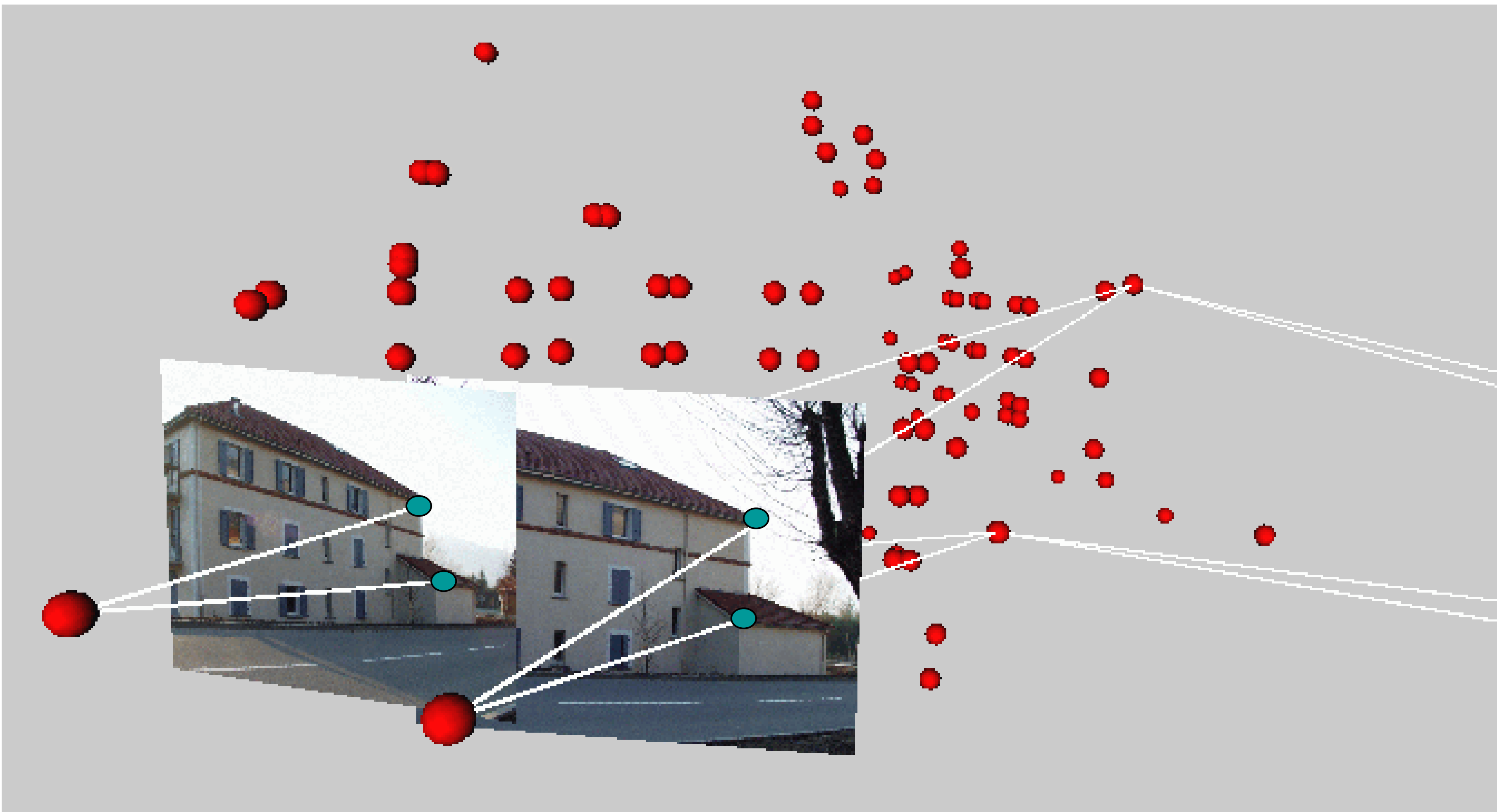
Calibration: determine, for each pixel, the corresponding line of sight ("projection ray")



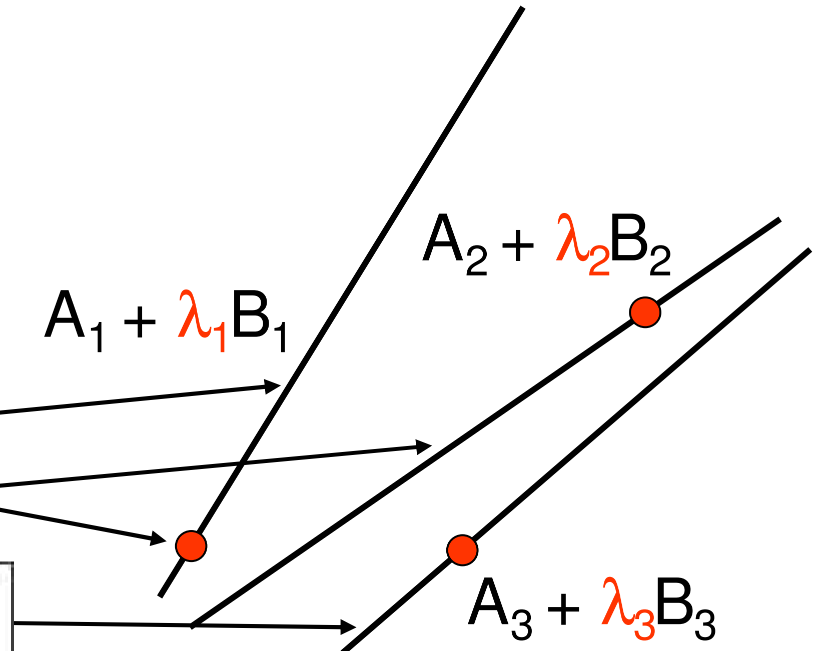
Motion estimation: compute motion such that matching rays intersect



Triangulation / 3D Reconstruction



Pose estimation of known object

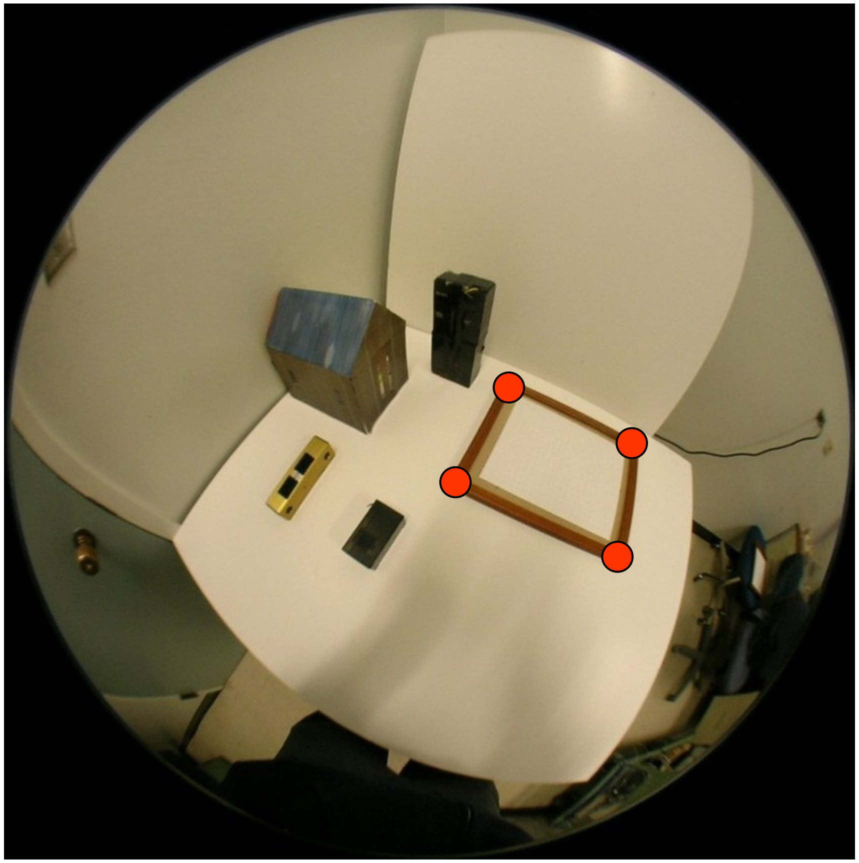


$$\mathrm{dist}^2(A_1 + \lambda_1 B_1 \,,\, A_2 + \lambda_2 B_2) = d_{12}^2$$

$$\mathrm{dist}^2(A_1 + \lambda_1 B_1 \,,\, A_3 + \lambda_3 B_3) = d_{13}^2$$

$$\mathrm{dist}^2(A_2 + \lambda_2 B_2 \,,\, A_3 + \lambda_3 B_3) = d_{23}^2$$

Pose estimation of known object



- 3 quadratic equations:
  up to 8 solutions
- Central camera: solutions come in
  mirrored pairs (for a solution in front
  of the camera, another one behind
  exists too)
- Non-central camera: no such simple
  symmetry exists
- With 4 points, unique solution
  in general

[Chen-Chang-PAMI'04,Nistér-CVPR'04,Ramalingam-etal-OMNIVIS'04]

Motion estimation

Motion estimation: unknown scene

- Pixel matches gives rise to ray matches
- Represent rays using Plücker coordinates
- Displacement for Plücker coordinates:

$$L'_1 = \begin{pmatrix} R & 0 \\ -[t]_x R & R \end{pmatrix}_{6\times 6} L_1$$

- Rays intersect if

$$L_2^T \begin{pmatrix} 0 & Id \\ Id & 0 \end{pmatrix}_{6\times 6} L'_1 = 0$$

$\rightarrow$

Essential matrix

$$E = \begin{pmatrix} -[t]_x R & R \\ R & 0 \end{pmatrix}_{6\times 6}$$

$$L_2^T \, E \, L_1 = 0$$

Motion estimation:

(1) Estimation of E (possible using linear equations: minimum 17 matches)

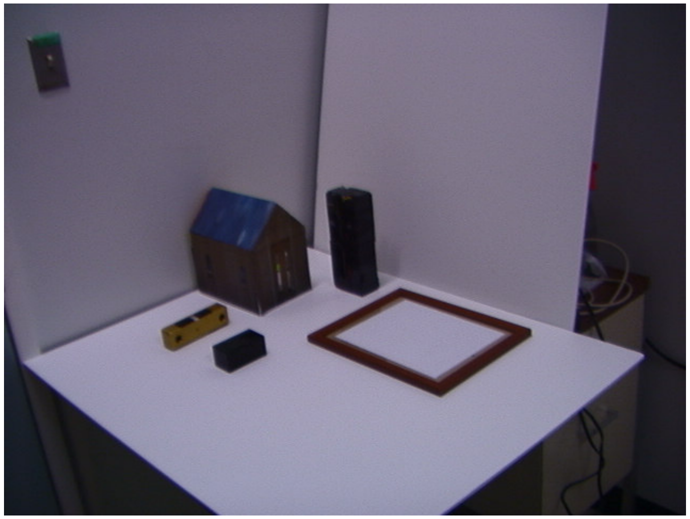
(2) Extraction of R and t from E (simple)

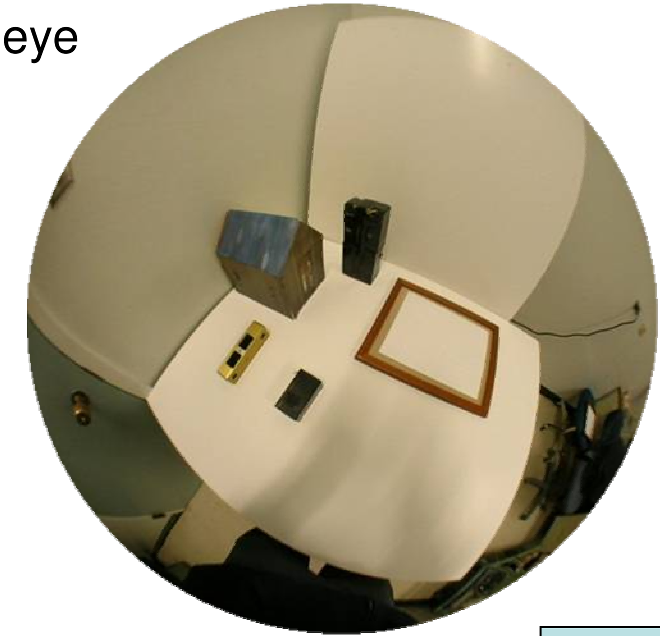
Note: scale of motion can be estimated if non-central cameras!
(but may be unreliable if cameras not very non-central)

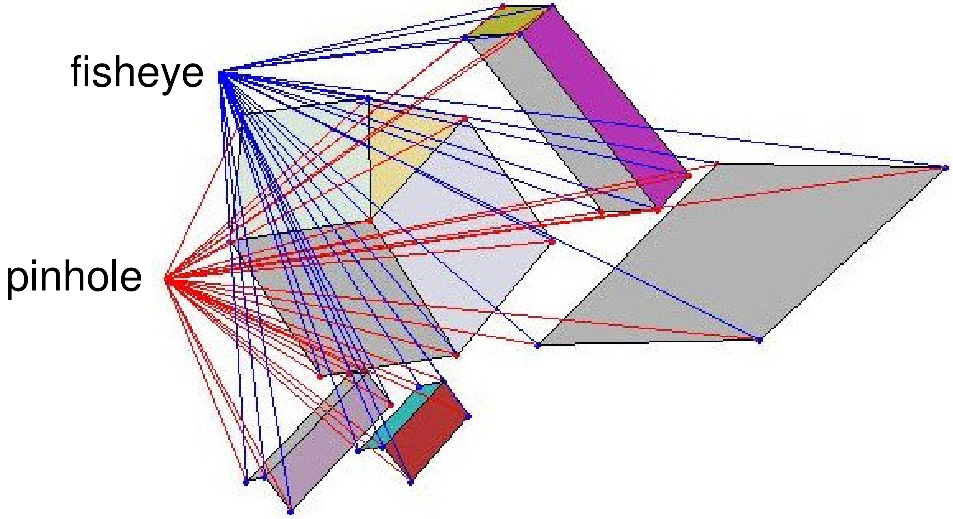
Variants for: axial, x-slit, central cameras

[Pless-CVPR'03,Sturm-etal-Bookchapter'06]

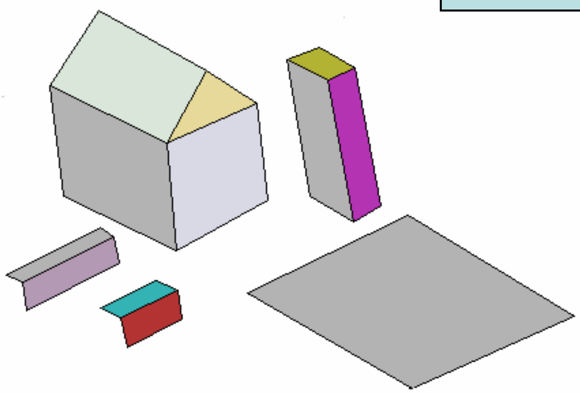
Motion estimation and 3D from pinhole+fisheye



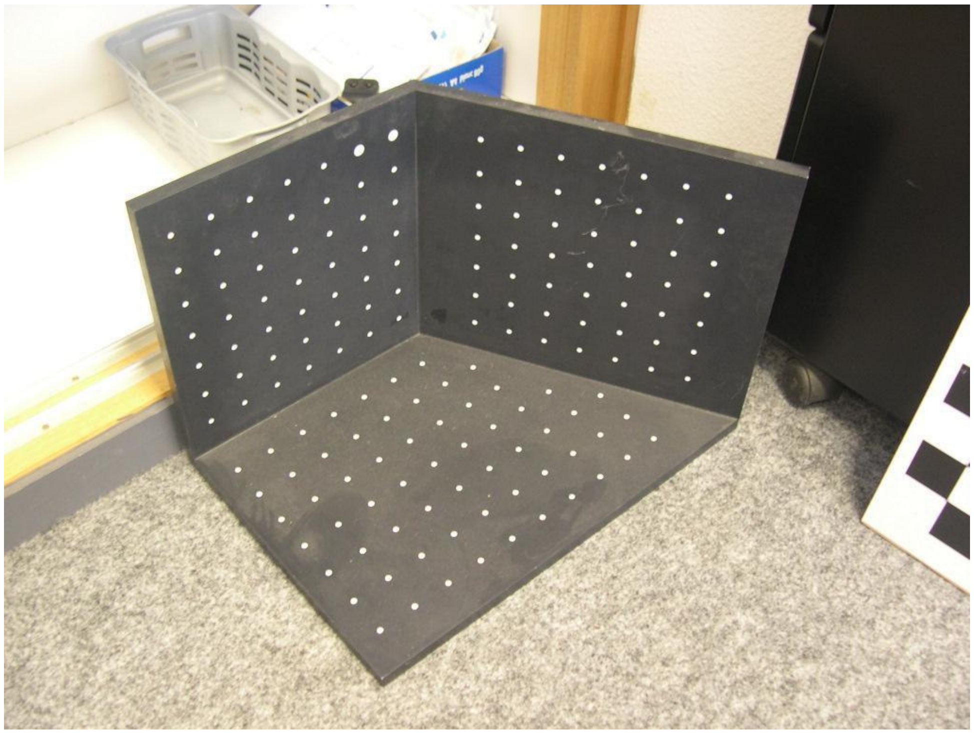


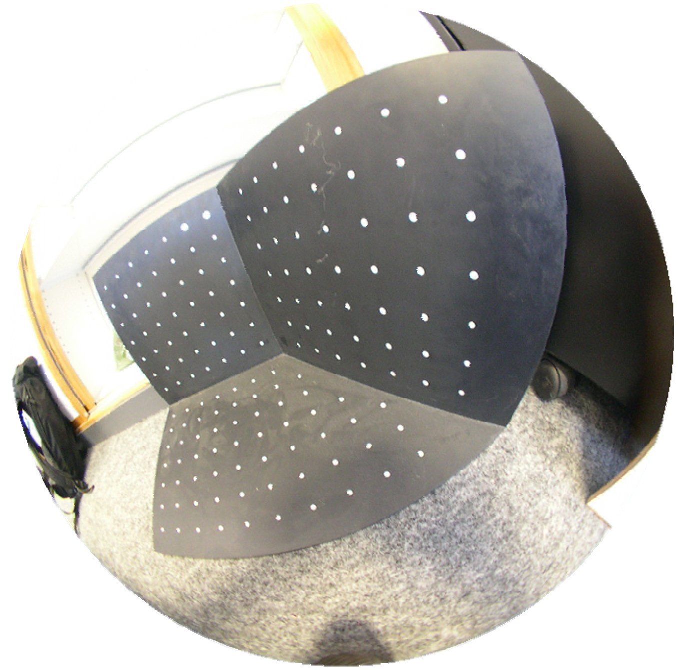


3D Model



Motion estimation and 3D from pinhole+fisheye





3D Model

Perspective epipolar geometry:

- Epipolar line of a pixel *p* computed via the fundamental matrix: $v{=}Fp$

Such a parametric epipolar geometry exists for some omnidirectional cameras, e.g. para-catadioptric ones

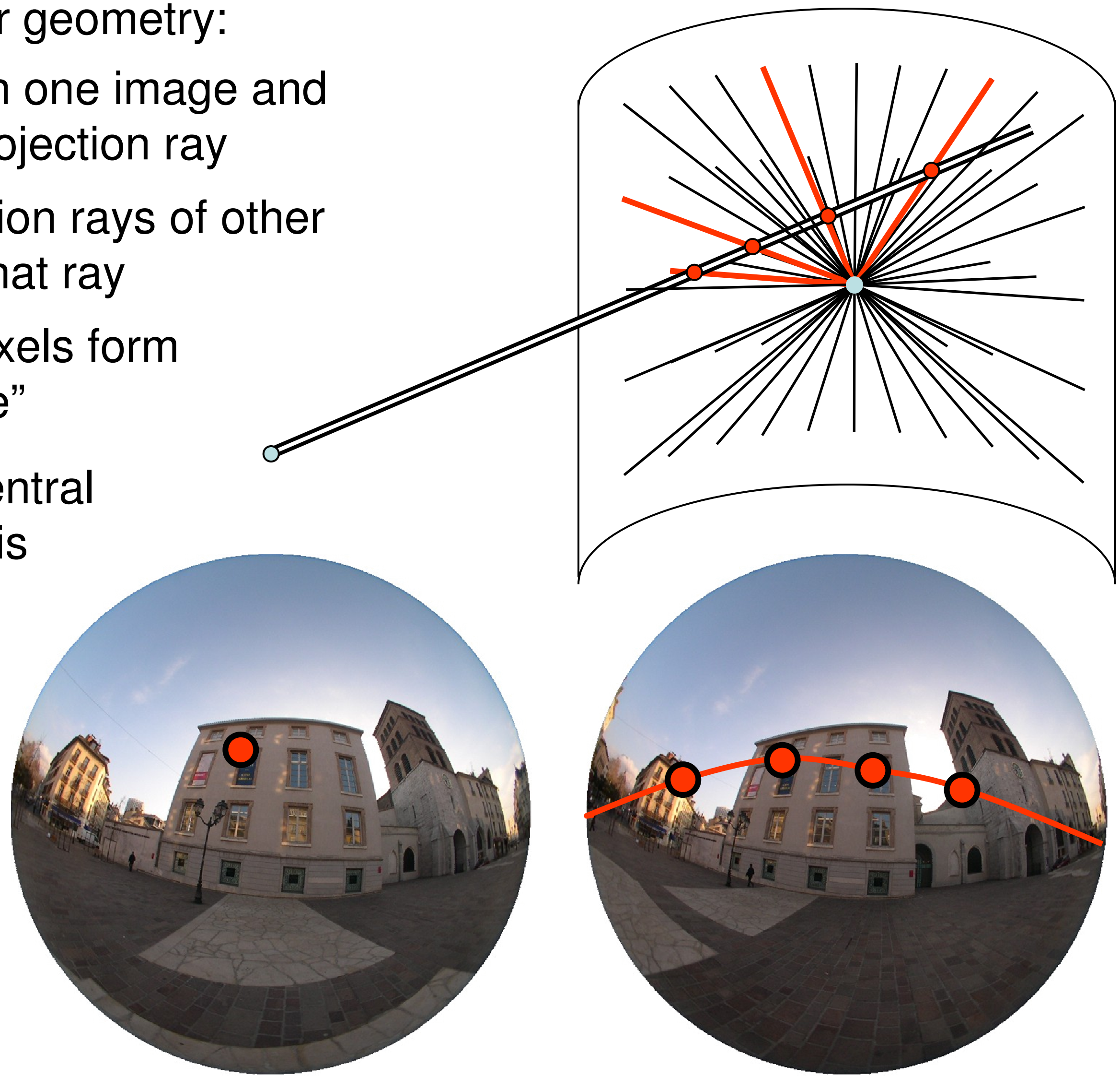
It also exists between cameras of different types, e.g. a stereo pair consisting of a perspective and a para-catadioptric camera

[Svoboda-etal-ECCV'98,Feldman-et-al-ICCV'05,Sturm-OMNIVIS'02]

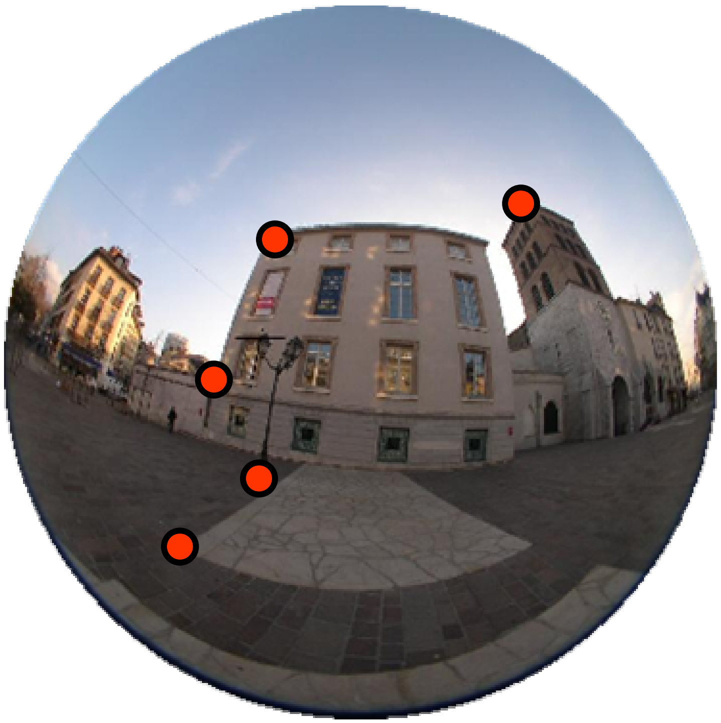
Non-parametric epipolar geometry:

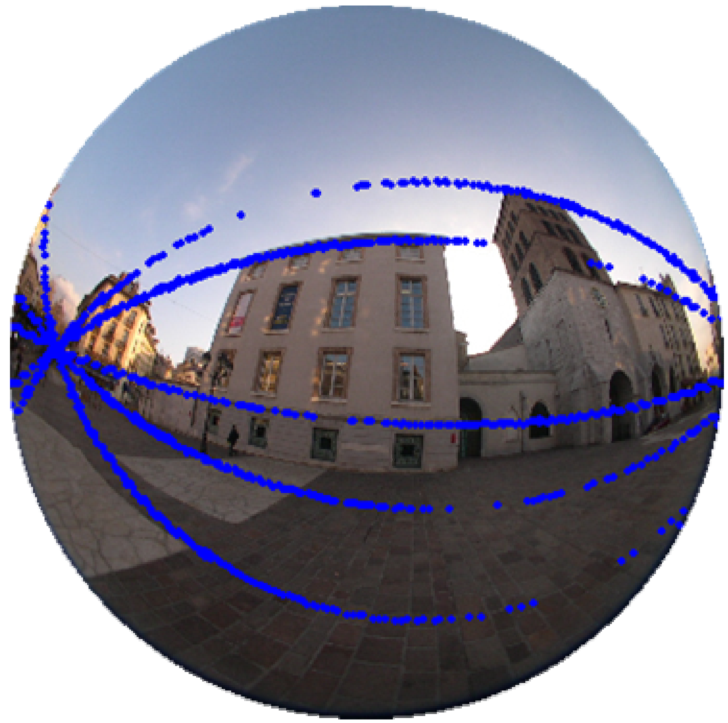
- Consider a pixel in one image and the associated projection ray
- Determine projection rays of other camera that cut that ray
- The associated pixels form an "epipolar curve"

Here: illustration with central cameras, but concept is applicable to whatever camera, i.e. also non-central ones



Non-parametric epipolar geometry:





Multi-view geometry for perspective images:

- Consider points (or other features) in images
- Which geometric constraints exist that tell if points are potential matches?
  - 2 images: epipolar geometry (fundamental/essential matrix)

$$\mathrm{q}_2^{\mathrm{T}} \mathbf{E}\, \mathrm{q}_1 = 0$$

  - 3 or 4 images: trifocal and quadrifocal tensors

$$\sum_{i_1=1}^{3} \sum_{i_2=1}^{3} \cdots \sum_{i_n=1}^{3} q_{1,i_1} q_{2,i_2} \cdots q_{n,i_n} T_{i_1,i_2,\cdots,i_n} = 0$$

Multi-view geometry for generic imaging model:

- Constraints between *projection rays*

$$\sum_{i_1=1}^{6} \sum_{i_2=1}^{6} \cdots \sum_{i_n=1}^{6} L_{1,i_1} L_{2,i_2} \cdots L_{n,i_n} T_{i_1,i_2,\cdots,i_n} = 0$$

Perspective multi-view geometry:

- Consider points $\mathbf{q}_i$ in *n* images with projection matrices P_i
- They are potential matches if scalars λ_i and a 3D point $\mathbf{Q}$ exist with:

$$\lambda_i \mathbf{q}_i = \mathsf{P}_i \mathbf{Q}, \quad \forall i = 1 \cdots n$$

- This can be written as:

$$\underbrace{\begin{pmatrix} \mathsf{P}_1 & \mathbf{q}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{q}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathsf{P}_n & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{q}_n \end{pmatrix}}_{\mathsf{M}} \begin{pmatrix} \mathbf{Q} \\ -\lambda_1 \\ -\lambda_2 \\ \vdots \\ -\lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- Existence of null-vector implies rank-deficiency of M
- M is of size 3n × 4+n
  → all submatrices (4+n) × (4+n) have zero determinant

$$\underbrace{\begin{pmatrix} \mathsf{P}_1 & \mathbf{q}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{q}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathsf{P}_n & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{q}_n \end{pmatrix}}_{\mathsf{M}}$$

- Determinants of submatrices can be written as:

$$\sum_{i_1=1}^{3} \sum_{i_2=1}^{3} \cdots \sum_{i_n=1}^{3} q_{1,i_1} q_{2,i_2} \cdots q_{n,i_n} T_{i_1,i_2,\cdots,i_n} = 0$$

- where: matching tensors $\mathbf{T}$ depend exactly on the projection matrices P_i

  n = 2: fundamental (essential) matrix

  n = 3: trifocal tensors

  n = 4: quadrifocal tensors

- Uses of matching tensors:
  - Matching constraints
  - Useful for motion estimation from image correspondences

Multi-view geometry for generic imaging model:

- Projection rays are represented by Plücker coordinates:
  - let $\mathbf{A}$ and $\mathbf{B}$ be any 2 points on a 3D line
  - Plücker coordinates can be defined as:

$$\mathbf{L} = \begin{pmatrix} A_4B_1 - A_1B_4 \\ A_4B_2 - A_2B_4 \\ A_4B_3 - A_3B_4 \\ A_3B_2 - A_2B_3 \\ A_1B_3 - A_3B_1 \\ A_2B_1 - A_1B_2 \end{pmatrix}$$

  - they are independent of the choice of $\mathbf{A}$ and $\mathbf{B}$

[Sturm-CVPR'05]

- Consider projection rays $\mathbf{L}_i$ for *n* calibrated cameras
- For the moment, parameterize rays by two points $\mathbf{A}_i$ and $\mathbf{B}_i$ each.
- Pose of cameras is parameterized as

$$\mathsf{P}_i = \begin{pmatrix} \mathsf{R}_i & \mathbf{t}_i \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix}$$

- Rays are potential matches if scalars λ_i and μ_i and a 3D point $\mathbf{Q}$ exist with:

$$\lambda_i \mathbf{A}_i + \mu_i \mathbf{B}_i = \mathsf{P}_i \mathbf{Q}, \quad \forall i = 1 \cdots n$$

- Rays are potential matches if scalars λ_i and μ_i and a 3D point $\mathbf{Q}$ exist with:

$$\lambda_i \mathbf{A}_i + \mu_i \mathbf{B}_i = \mathsf{P}_i \mathbf{Q}, \quad \forall i = 1 \cdots n$$

- This can be written as:

$$\underbrace{\begin{pmatrix} \mathsf{P}_1 & \mathbf{A}_1 & \mathbf{B}_1 & \cdots & \mathbf{0} & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathsf{P}_n & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_n & \mathbf{B}_n \end{pmatrix}}_{\mathsf{M}} \begin{pmatrix} \mathbf{Q} \\ -\lambda_1 \\ -\mu_1 \\ \vdots \\ -\lambda_n \\ -\mu_n \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$$

- Existence of null-vector implies rank-deficiency of M
- M is of size 4n × 4+2n
  → all submatrices (4+2n) × (4+2n) have zero determinant

$$\underbrace{\begin{pmatrix} \mathsf{P}_1 & \mathbf{A}_1 & \mathbf{B}_1 & \cdots & \mathbf{0} & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathsf{P}_n & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_n & \mathbf{B}_n \end{pmatrix}}_{\mathsf{M}}$$

- When developping determinants of submatrices, coordinates of points $\mathbf{A}_i$ and $\mathbf{B}_i$ appear in terms of this form:

$$A_{i,j} B_{i,k} - A_{i,k} B_{i,j}$$

$\rightarrow$ Plücker coordinates of $\mathbf{L}_i$

- We obtain matching constraints of the form:

$$\sum_{i_1=1}^{6} \sum_{i_2=1}^{6} \cdots \sum_{i_n=1}^{6} L_{1,i_1} L_{2,i_2} \cdots L_{n,i_n} T_{i_1,i_2,\cdots,i_n} = 0$$

- Matching tensors $\mathbf{T}$ depend on pose matrices P_i

- Like for perspective images, matching tensors exist for 2, 3, and 4 cameras
- Example: two views

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & A_{1,1} & B_{1,1} & 0 & 0 \\ 0 & 1 & 0 & 0 & A_{1,2} & B_{1,2} & 0 & 0 \\ 0 & 0 & 1 & 0 & A_{1,3} & B_{1,3} & 0 & 0 \\ 0 & 0 & 0 & 1 & A_{1,4} & B_{1,4} & 0 & 0 \\ R_{11} & R_{12} & R_{13} & t_1 & 0 & 0 & A_{2,1} & B_{2,1} \\ R_{21} & R_{22} & R_{23} & t_2 & 0 & 0 & A_{2,2} & B_{2,2} \\ R_{31} & R_{32} & R_{33} & t_3 & 0 & 0 & A_{2,3} & B_{2,3} \\ 0 & 0 & 0 & 1 & 0 & 0 & A_{2,4} & B_{2,4} \end{pmatrix}$$

of size 8x8

M is rank-deficient, thus singular

$\rightarrow$ matching constraint is: $$\det M = L_2^T \underbrace{\begin{pmatrix} -[t]_x R & R \\ R & 0 \end{pmatrix}}_{\text{essential matrix}} L_1 = 0$$

- Matching tensors for non-central cameras are of size $6 \times 6 \times \ldots$
- Reduced parameterizations exist:
  - Axial cameras: $5 \times 5 \times \ldots$
  - X-slit cameras: $4 \times 4 \times \ldots$
  - Central cameras: $3 \times 3 \times \ldots$
- Matching tensors between cameras of different types are straightforward, e.g.:
  - Essential matrix of a non-central and a central camera: 6×3

Summary for structure-from-motion:

- When calibrated cameras are considered, an SfM problem (pose, motion, …) can be solved with one and the same algorithm, whatever the type of camera
- But: results are not optimal (e.g. in the sense of reprojection errors)
  → methods are useful for embedding in RANSAC, but should be followed by bundle adjustment if good accuracy required
- Extension of structure-from-motion theory from perspective to general camera model
- Some missing pieces, e.g. matching tensors for line images

Contents

- Introduction
- General imaging models
- Non-parametric calibration and distortion correction
- Non-parametric self-calibration
- Structure-from-motion

Conclusions

- Generic camera model
- Generic approaches for calibration and structure-from-motion tasks

General Imaging

Design, Modeling and Applications

Peter Sturm, INRIA Grenoble - Rhône-Alpes, France

http://perception.inrialpes.fr/people/Sturm

With contributions from:

Srikumar Ramalingam	INRIA Grenoble - Rhône-Alpes and UC Santa Cruz
Jean-Philippe Tardif	Université de Montréal