



# Photo-based Augmented Reality

## Registration And Navigation Methods

*Application to Discrepancy Check*

*Imperial College London*

*11/09/09*

*Pierre Georgel*

---

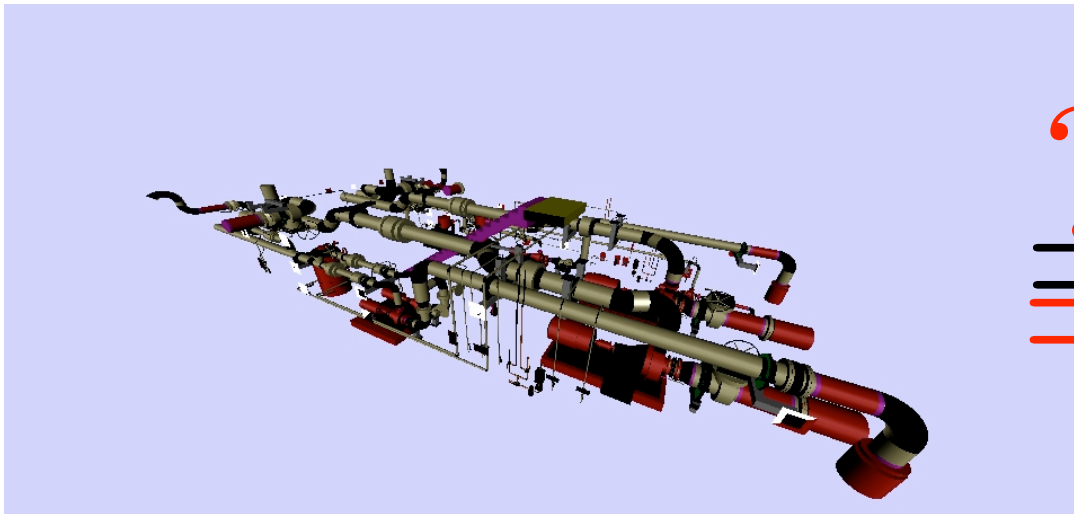


## Agenda

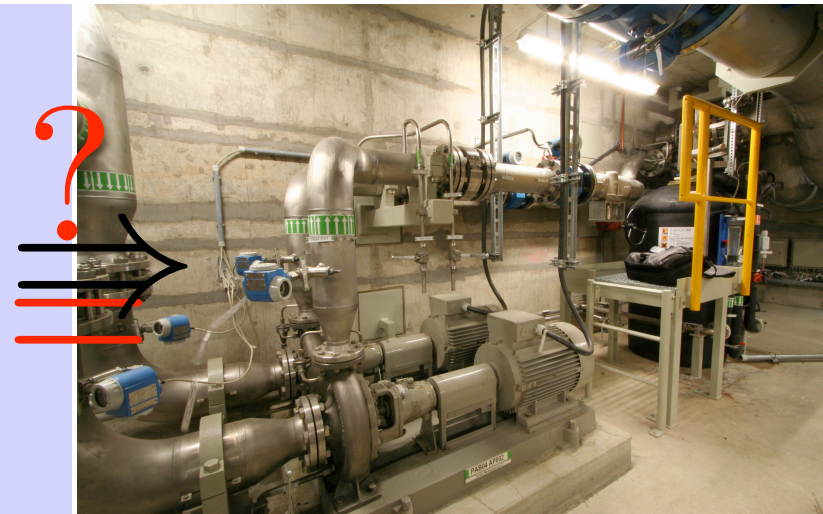
- Problem statement and application
- Software features
- Recovering the Full Pose from a Single Keyframe
- Outline of future work

## Application Overview

- Discrepancy check between CAD data and built items



*3D Model used as Planning for Construction*



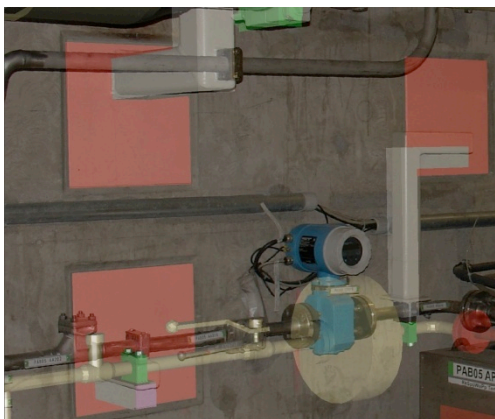
*Finished Construction*

## Application Overview

- Support engineer in charge of this verification task
- Project started in February 2006
- Consortium of CAMP – Siemens CT – Areva NP



*"[...] senior project manager at Siemens, estimates that the software will reduce the cost of constructing a typical medium-sized coal-fired power plant by more than \$1m." The economist 2007*



## Photo-based Augmented Reality



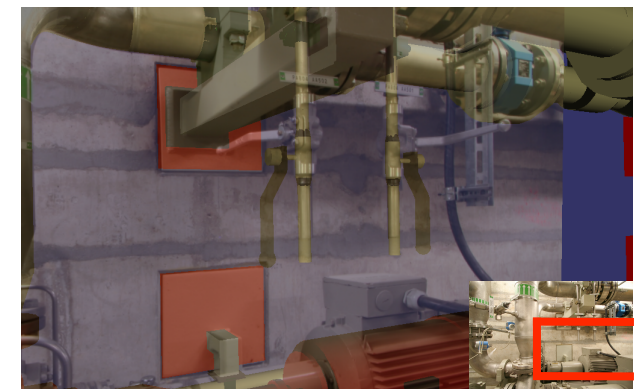
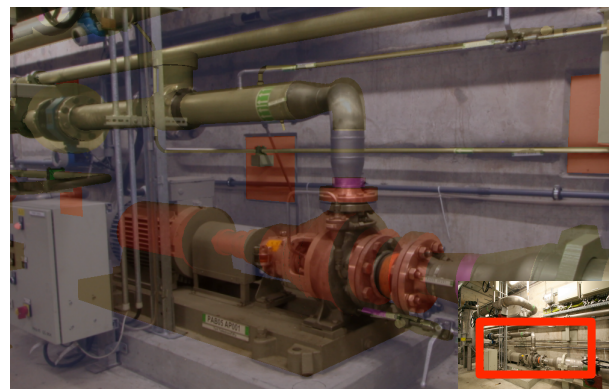
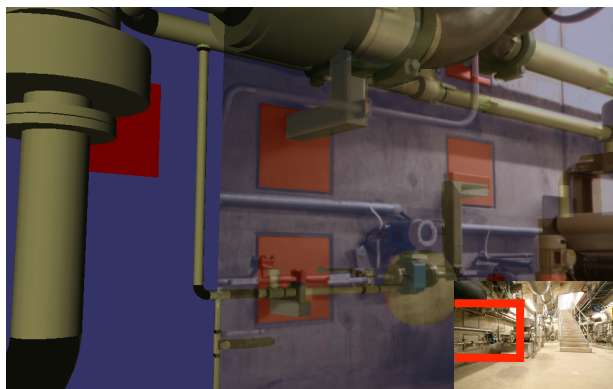
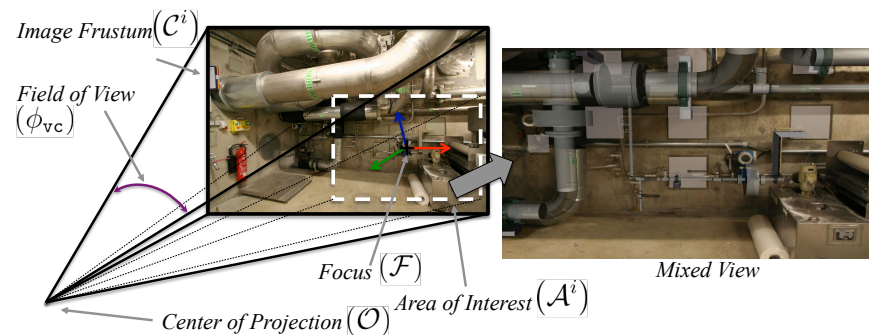
*Single frame augmentation*

## Requirements

1. Interaction
2. Registration

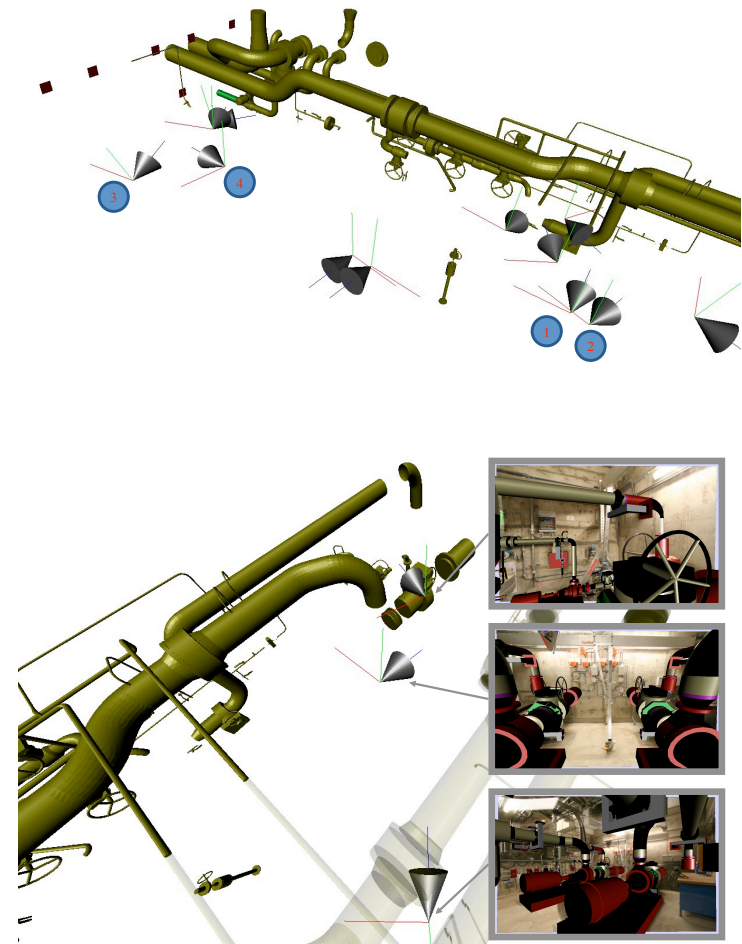
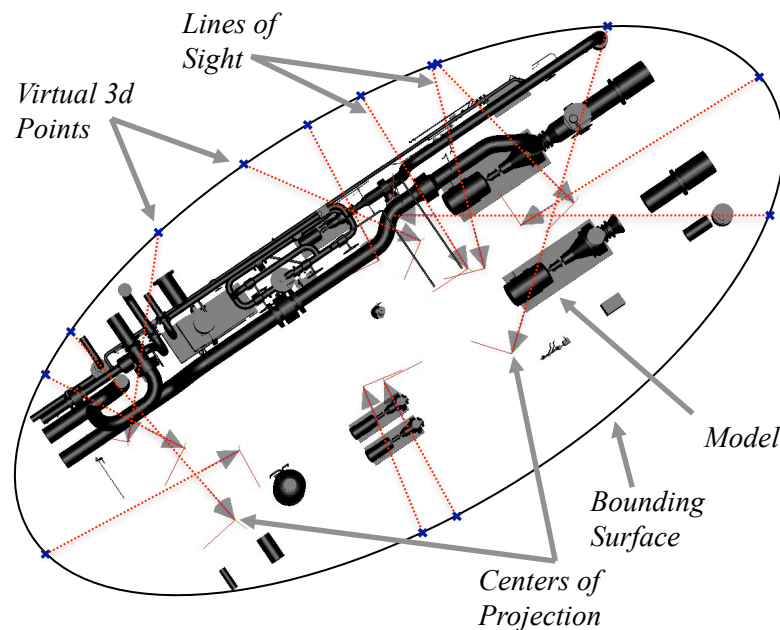
## 3D User Interface for an Augmented CAD Software

- New paradigm requires new interactions
  - Extension of 2D zoom/pan interface to Mixed views



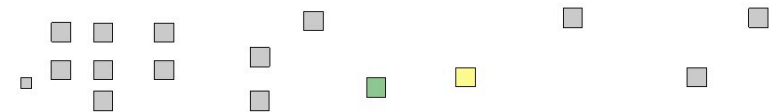
## 3D User Interface for an Augmented CAD Software

- New paradigm requires new interactions
  - Intuitive 3D image navigation
    - No points track are required
    - No Model is considered



## Landmark-based Registration

- Goal
  - Register images to the CAD coordinate system
- Approach
  - Extract landmarks (Anchor-Plates) from images
  - Find corresponding 3D landmarks
  - Compute the Pose (Rotation, Translation)

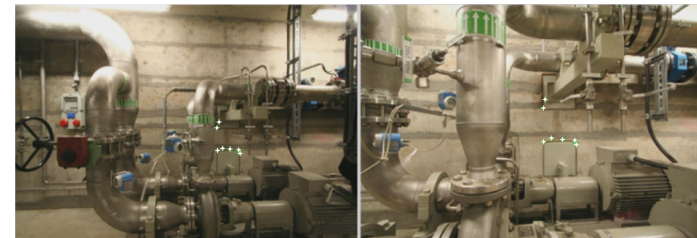


- Limitation
  - Need of 2d visible Landmarks

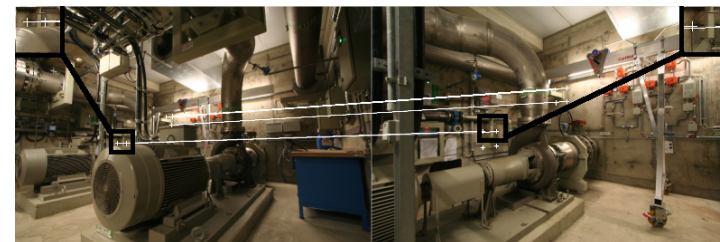


## Automatic 6 DoF Stereo Registration

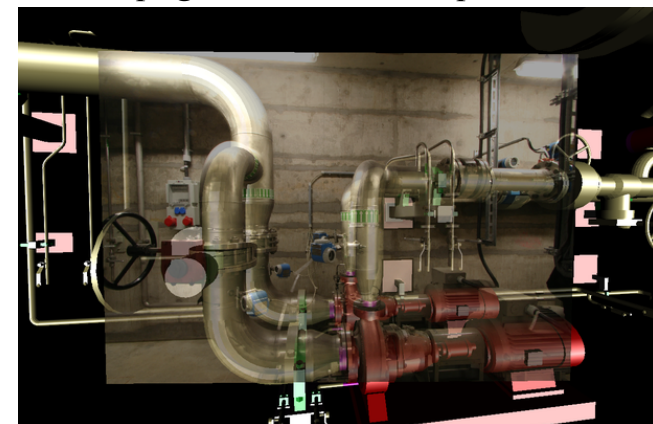
- Goal
  - Register images using a unique keyframe
- Approach
  - Compute epipolar geometry using keypoints
  - Extend the relative pose to a full pose using
    - Planar structure
    - 2d-3d correspondences
  - Non-linear estimation
    - Bundle adjustment with planarity constraint
    - Hybrid pose estimation that includes intensities



*Extracted planar structure*



*Propagated 2d-3d correspondences*





# Recovering the Full Pose from a Single Keyframe

WACV

12/2009

*Pierre Fite Georgel, Selim Benhimane, Juergen Sotke and Nassir Navab*

---



# Agenda

- Registration
- Registration with single keyframe
  - Challenges
  - Initial scale estimates
  - Non-linear refinement
- Results
  - Synthetic experiments
  - Plant inspection images
- Closing statement

## Registration

### *Find Geometric Transform between CAD and Image*

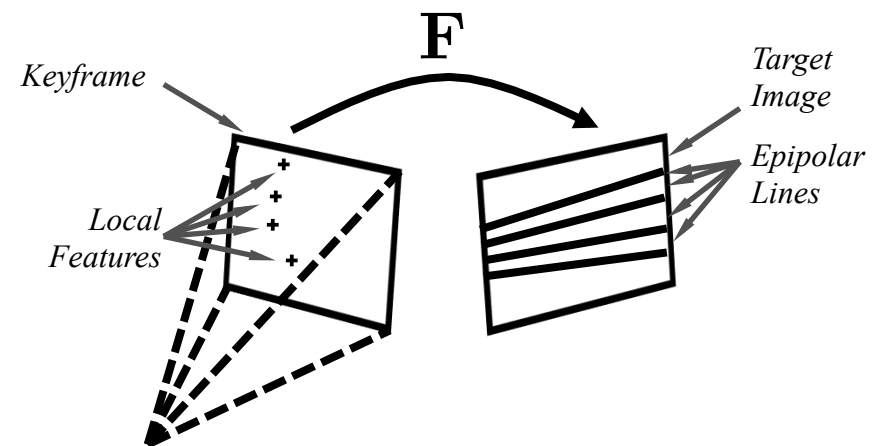
- Marker based [1,2]
- External tracking system
  - Magnetic [3]
  - Optical [4]
  - GPS + Compass [5]
- Model based approach
  - Edges [6,7]
  - Keyframes (Multiple [8], Unique with model [9, 10])

- [1] Goose *et al.* Speech-enabled augmented reality supporting mobile industrial maintenance. Pervasive Computing 2003.
- [2] Pentenrieder *et al.* Augmented Reality-based factory planning - an application tailored to industrial needs. ISMAR, 2007.
- [3] Webster *et al.* Architectural Anatomy. Presence, 1995.
- [4] Schoenfelder & Schmalstieg. Augmented Reality for Industrial Building Acceptance. IEEE VR, 2008.
- [5] Schall *et al.* Virtual redlining for civil engineering in real environments. ISMAR, 2008.
- [6] Lowe. Fitting parameterized three-dimensional models to images. IEEE Trans. PAMI, 1991.
- [7] Drummond & Cipolla. Real-time tracking of complex structures with on-line camera calibration. BMVC, 1999.
- [8] Chia *et al.* Online 6 dof augmented reality registration from natural features. ISMAR, 2002.
- [9] Vacchetti *et al.* Stable real-time 3d tracking using online and offline information. IEEE Trans. PAMI, 2004.
- [10] Platonov *et al.* A mobile markerless AR system for maintenance and repair. ISMAR, 2006.

## Full Pose from Single Keyframe *Challenge*

- Compute fundamental matrix using keypoints  $\langle \mathbf{p}_i, \mathbf{q}_i \rangle$

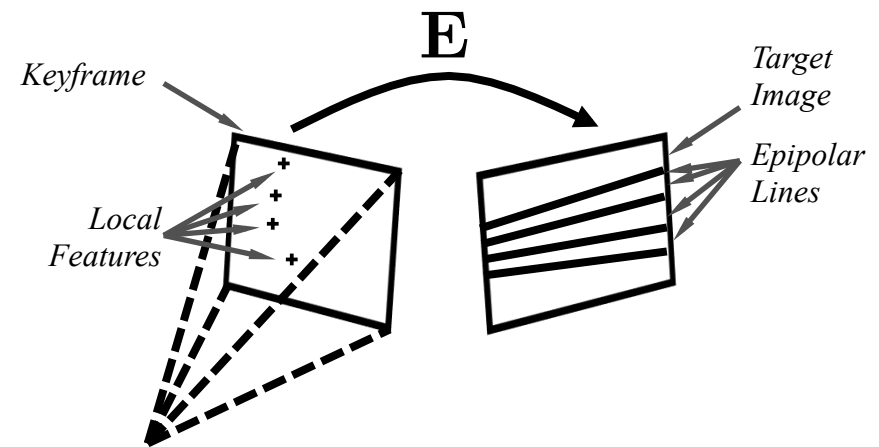
$$\mathbf{q}_i^\top \mathbf{F} \mathbf{p}_i = 0$$



## Full Pose from Single Keyframe *Challenge*

- Compute fundamental matrix using keypoints  $\langle \mathbf{p}_i, \mathbf{q}_i \rangle$ 
$$\mathbf{q}_i^\top \mathbf{F} \mathbf{p}_i = 0$$
- Derivation of essential for calibrated cameras [11]

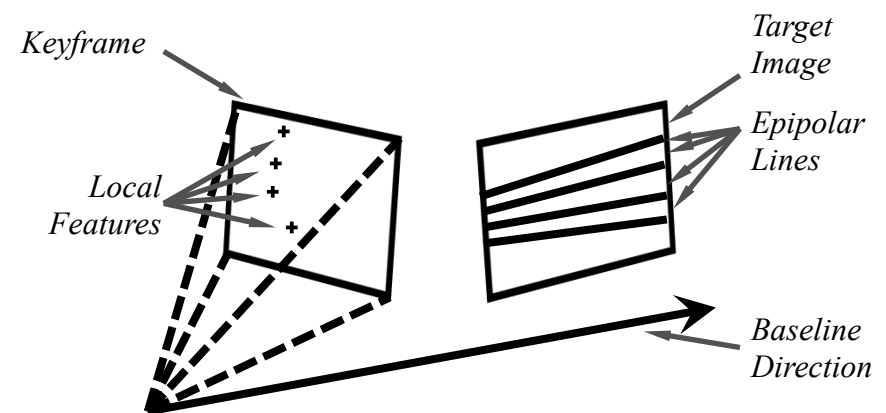
$$\mathbf{E} = \mathbf{K}_T^\top \mathbf{F} \mathbf{K}_S$$



## Full Pose from Single Keyframe *Challenge*

- Compute fundamental matrix using keypoints  $\langle \mathbf{p}_i, \mathbf{q}_i \rangle$
- Derivation of essential for calibrated cameras [11]
- Essential matrix decomposition

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$$

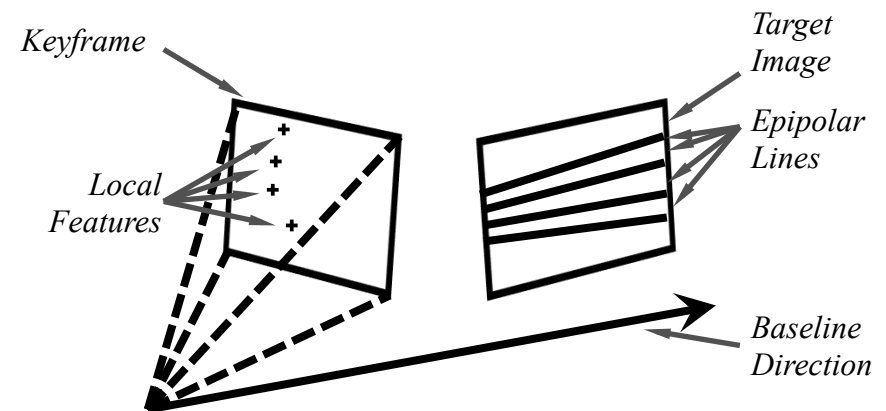


## Full Pose from Single Keyframe *Challenge*

- Compute fundamental matrix using keypoints  $\langle \mathbf{p}_i, \mathbf{q}_i \rangle$
- Derivation of essential for calibrated cameras [11]
- Essential matrix decomposition

$$\mathbf{E} = \boxed{[\mathbf{t}]_{\times}} \mathbf{R}$$

*Unknown translation norm*





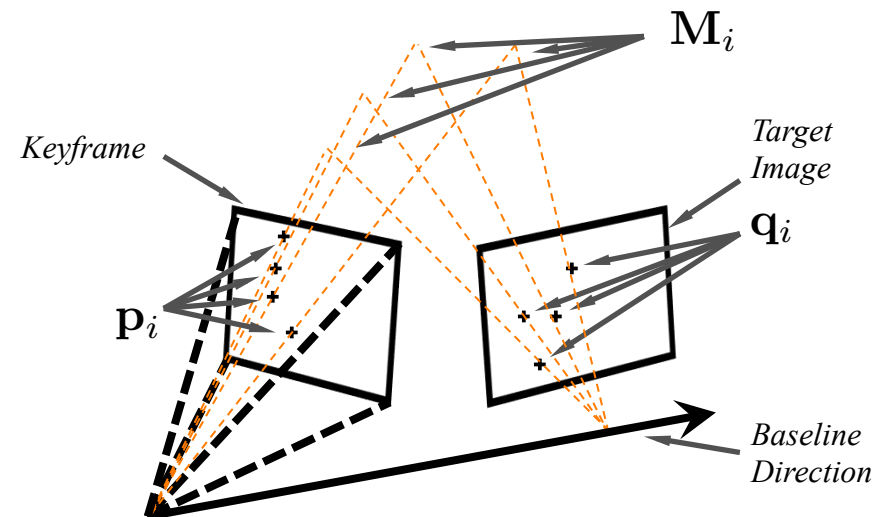
## Full Pose from Single Keyframe Challenge

- Compute fundamental matrix using keypoints  $\langle \mathbf{p}_i, \mathbf{q}_i \rangle$
- Derivation of essential for calibrated cameras [11]
- Essential matrix decomposition

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$$

- Bundle adjustment cost

$$\mathcal{C}_{\mathcal{G}}(\mathbf{M}_i, \mathbf{R}, \mathbf{t}) = \sum_{i=1}^n + \frac{\|\mathbf{K}_s \mathbf{w}(\mathbf{M}_i) - \mathbf{p}_i\|^2}{\|\mathbf{K}_t \mathbf{w}(\mathbf{R}\mathbf{M}_i + \mathbf{t}) - \mathbf{q}_i\|^2}$$



## Full Pose from Single Keyframe Challenge

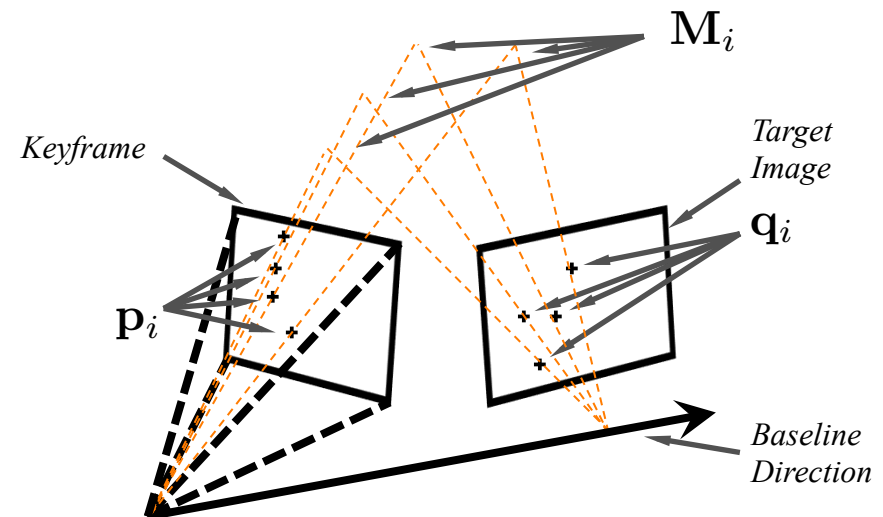
- Compute fundamental matrix using keypoints  $\langle \mathbf{p}_i, \mathbf{q}_i \rangle$
- Derivation of essential for calibrated cameras [11]
- Essential matrix decomposition

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$$

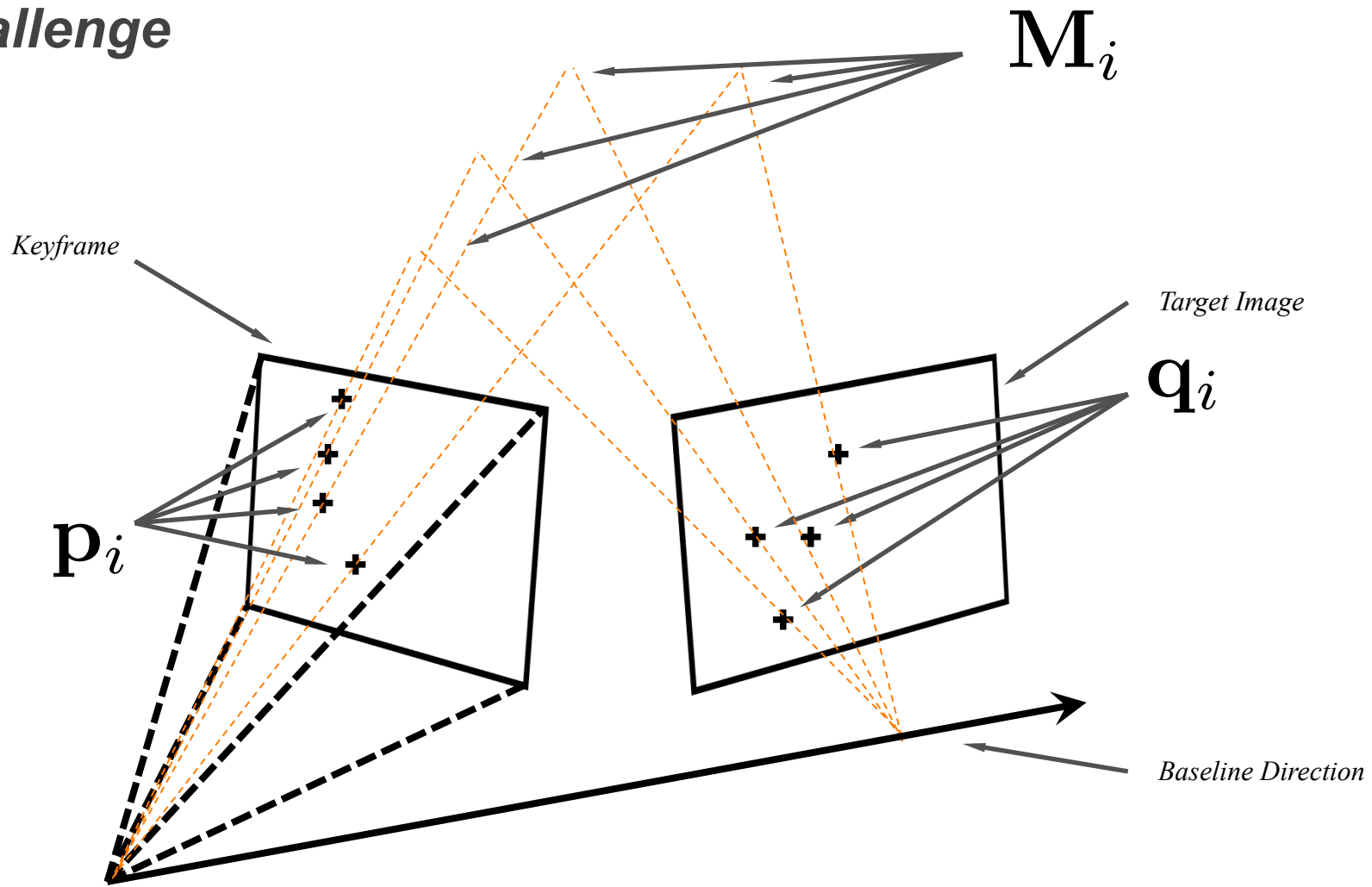
- Bundle adjustment cost

$$\mathcal{C}_{\mathcal{G}}(\mathbf{M}_i, \mathbf{R}, \mathbf{t}) = \sum_{i=1}^n \left( \|\mathbf{K}_s \mathbf{w}(\mathbf{M}_i) - \mathbf{p}_i\|^2 + \|\mathbf{K}_t \mathbf{w}(\mathbf{R}\mathbf{M}_i + \mathbf{t}) - \mathbf{q}_i\|^2 \right)$$

$$\forall s \neq 0, \quad \mathcal{C}_{\mathcal{G}}(s\mathbf{M}_i, \mathbf{R}, s\mathbf{t}) = \mathcal{C}_{\mathcal{G}}(\mathbf{M}_i, \mathbf{R}, \mathbf{t})$$

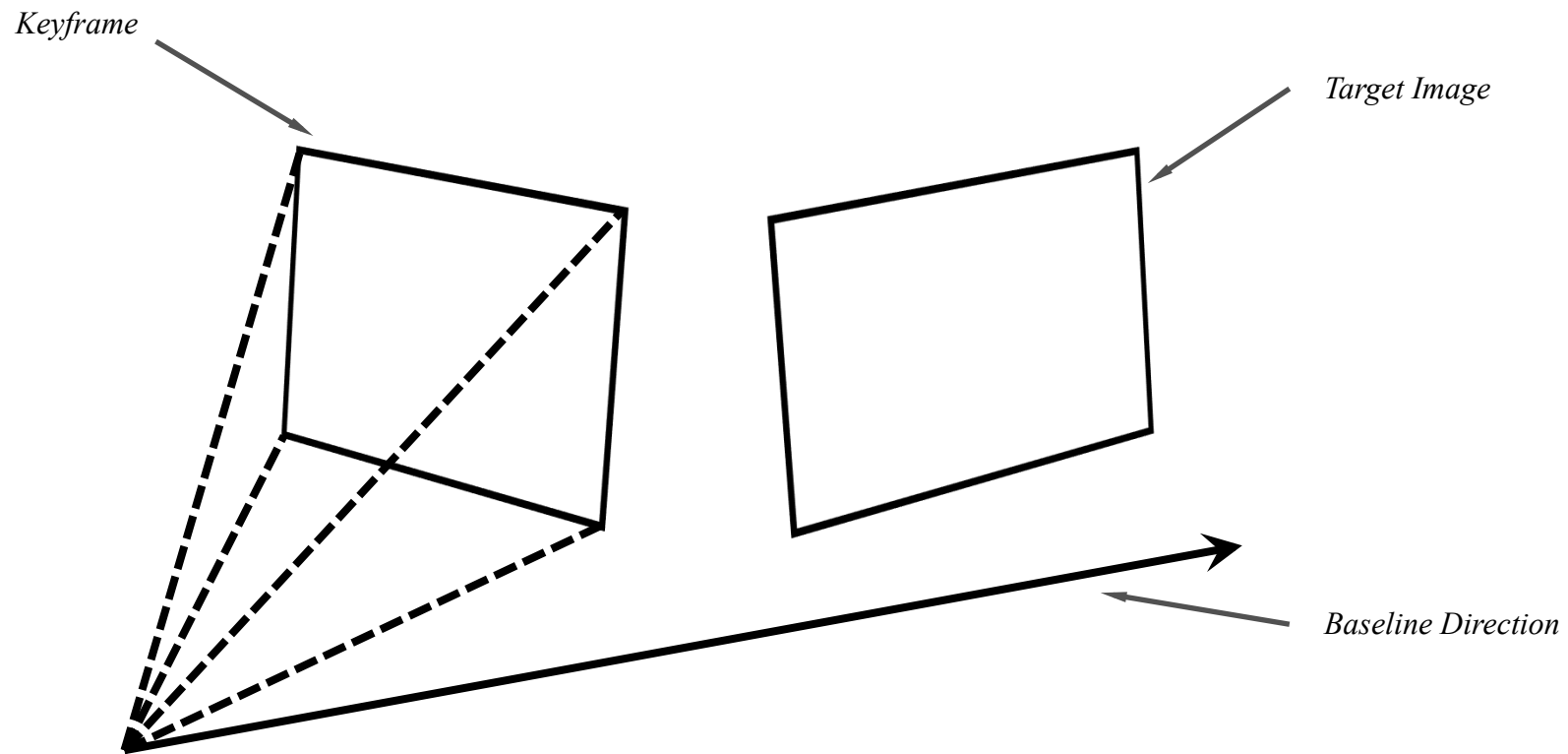


# Full Pose from Single Keyframe *Challenge*



## Full Pose from Single Keyframe Challenge

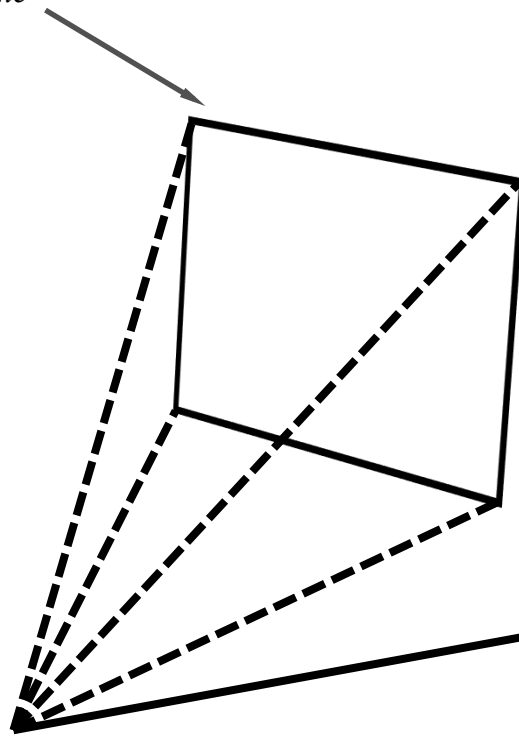
- We suppose that we know  $\mathbf{R}$  and  $\mathbf{t}$  ( $\|\mathbf{t}\| = 1$ )



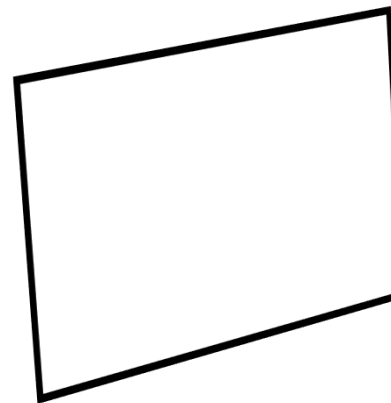
## Full Pose from Single Keyframe *Challenge*

- We suppose that we know  $\mathbf{R}$  and  $\mathbf{t}$  ( $\|\mathbf{t}\| = 1$ )
- We search for  $\mathcal{S}$

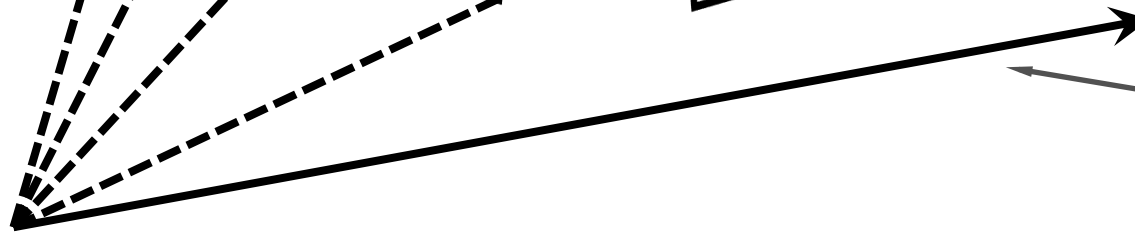
*Keyframe*



*Target Image*

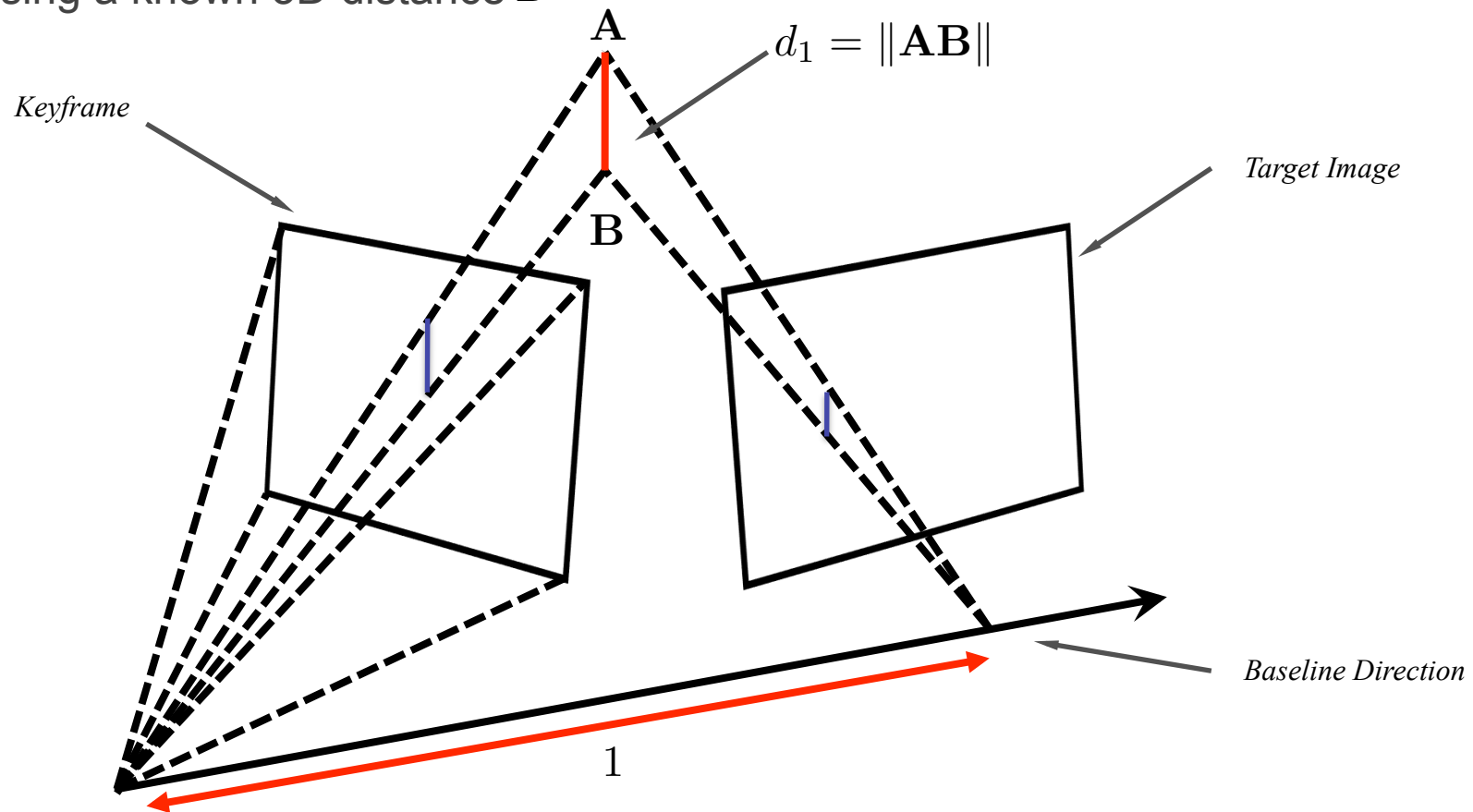


*Baseline Direction*



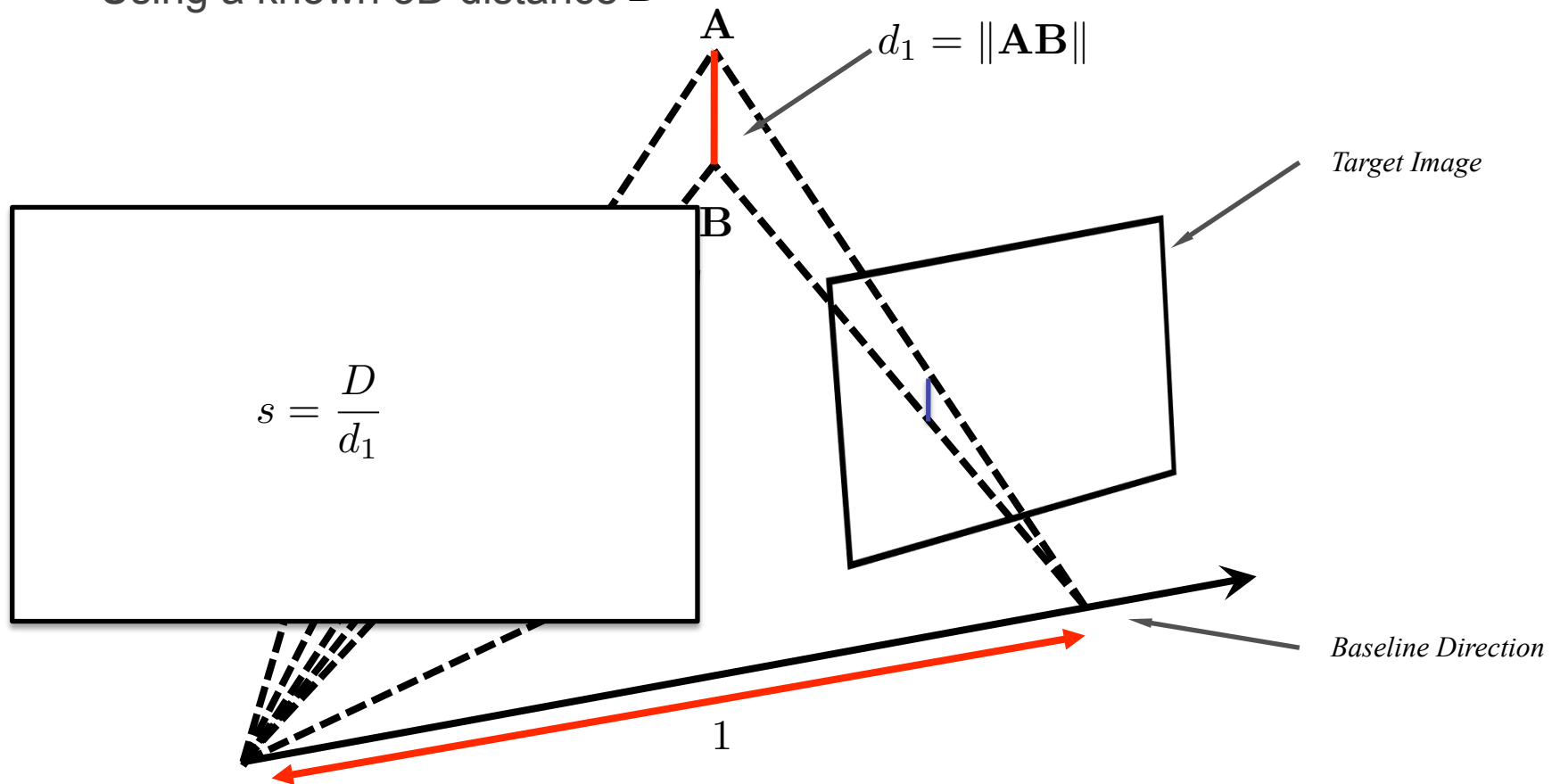
## Full Pose from Single Keyframe *Common Approach*

- Using a known 3D distance  $D$



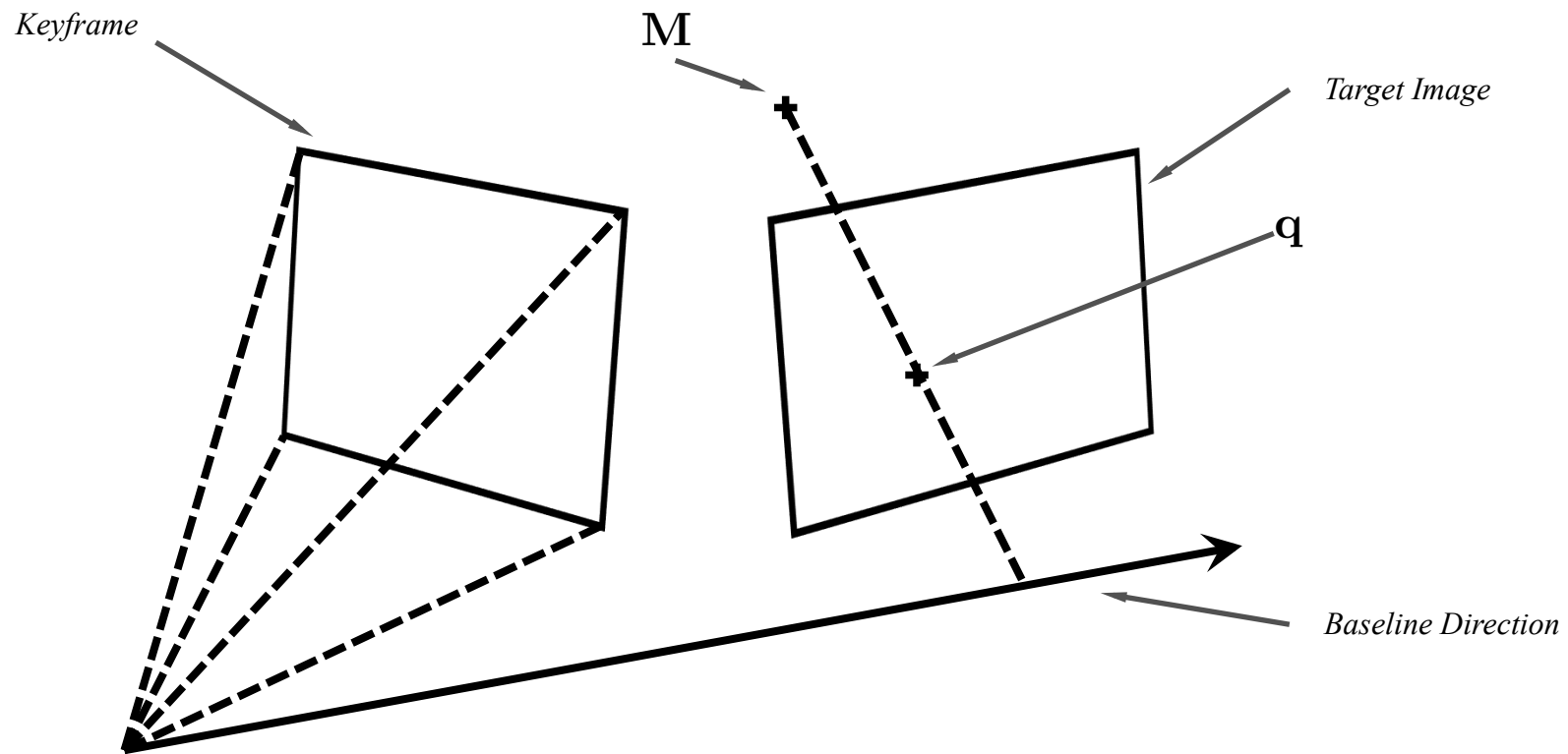
## Full Pose from Single Keyframe *Common Approach*

- Using a known 3D distance  $D$



## Full Pose from Single Keyframe *Common Approach*

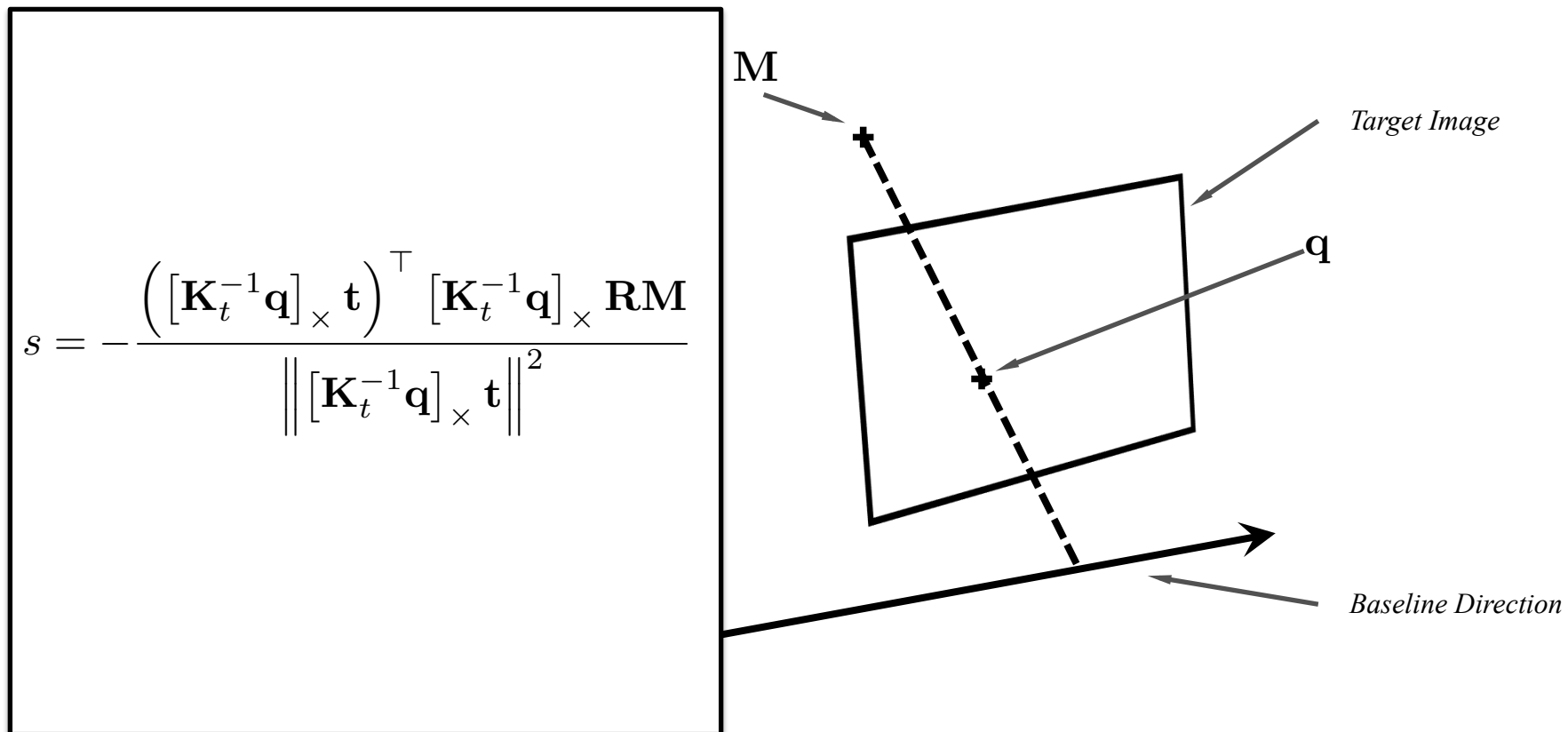
- Using the location of a known 3D point in the target image  $\langle \mathbf{M}, \mathbf{q} \rangle$





## Full Pose from Single Keyframe *Common Approach*

- Using the location of a known 3D point in the target image  $\langle \mathbf{M}, \mathbf{q} \rangle$



## Full Pose from Single Keyframe

### *Common Approach*

- Using a known 3D distance  $D$

$$s = \frac{D}{d_1}$$

- Using the location of a known 3D point in the target image  $\langle \mathbf{M}, \mathbf{q} \rangle$

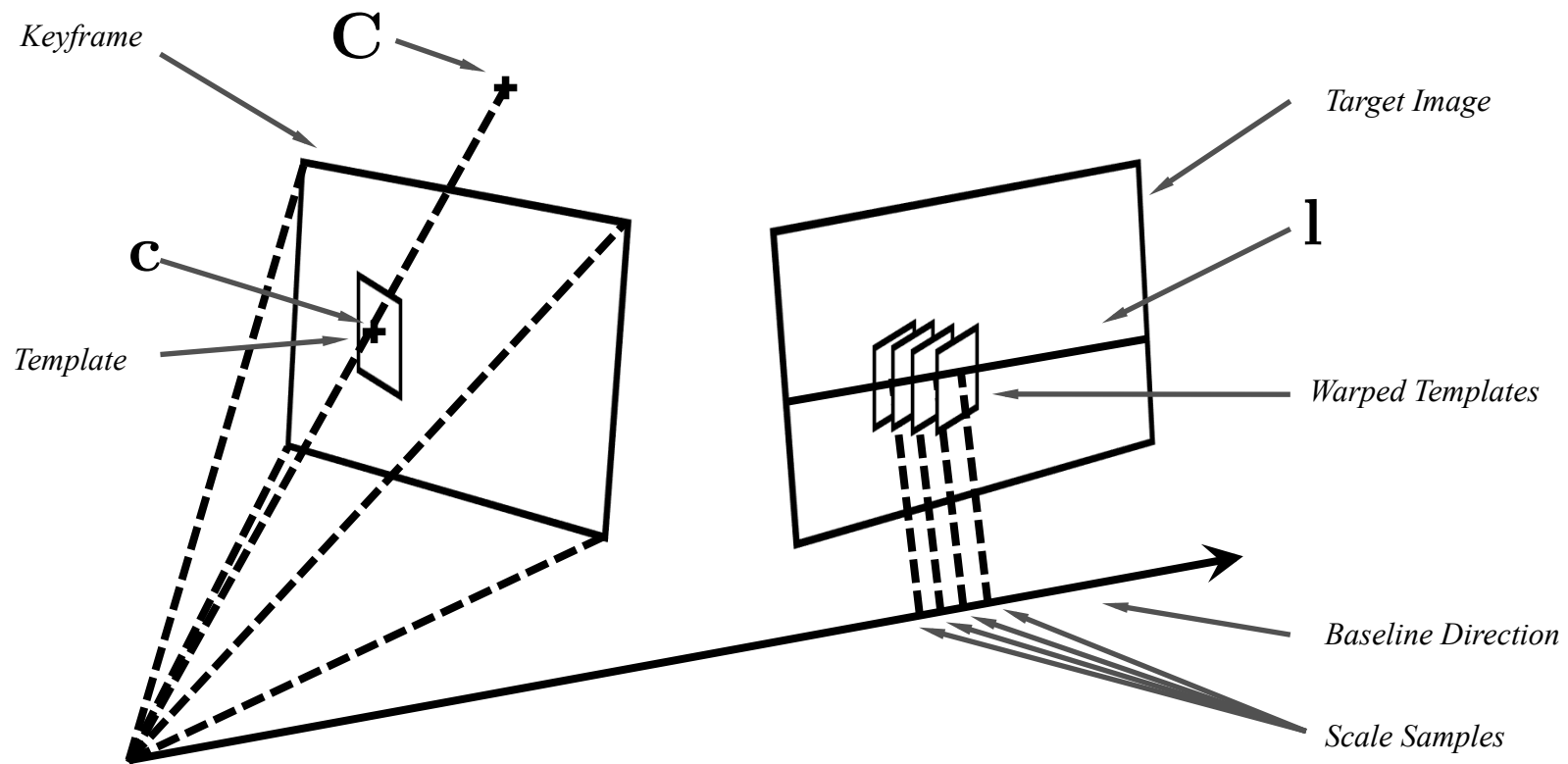
$$s = -\frac{\left( [\mathbf{K}_t^{-1} \mathbf{q}]_{\times} \mathbf{t} \right)^{\top} [\mathbf{K}_t^{-1} \mathbf{q}]_{\times} \mathbf{R} \mathbf{M}}{\left\| [\mathbf{K}_t^{-1} \mathbf{q}]_{\times} \mathbf{t} \right\|^2}$$

**Both methods requires interactions**

## Full Pose from Single Keyframe

### Method Overview

- We suppose that we know  $\mathbf{R}$  and  $\mathbf{t}$  ( $\|\mathbf{t}\| = 1$ )
- We search for  $S$

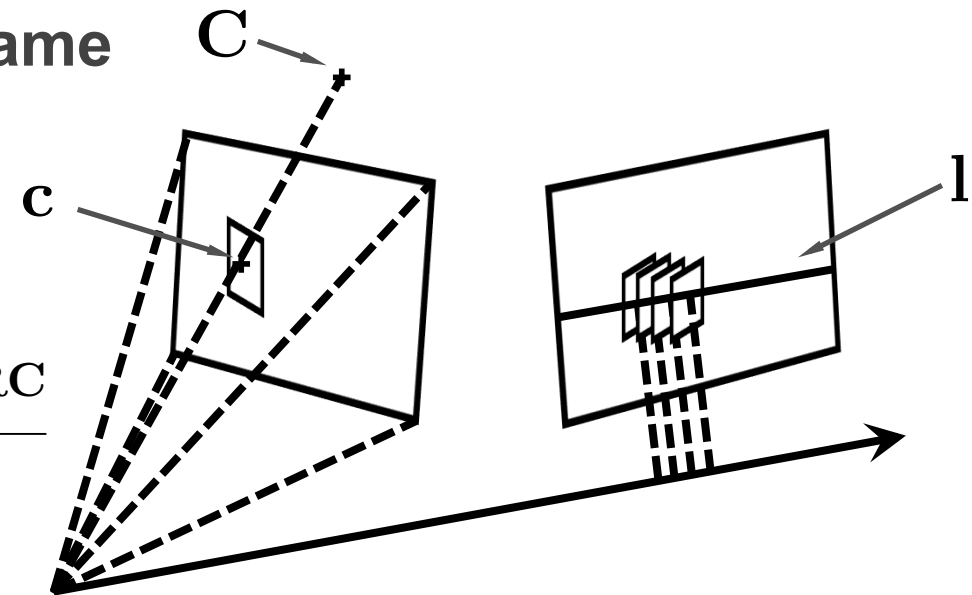


## Full Pose from Single Keyframe

### *Initial Estimates*

- Every point on a line gives a scale sample

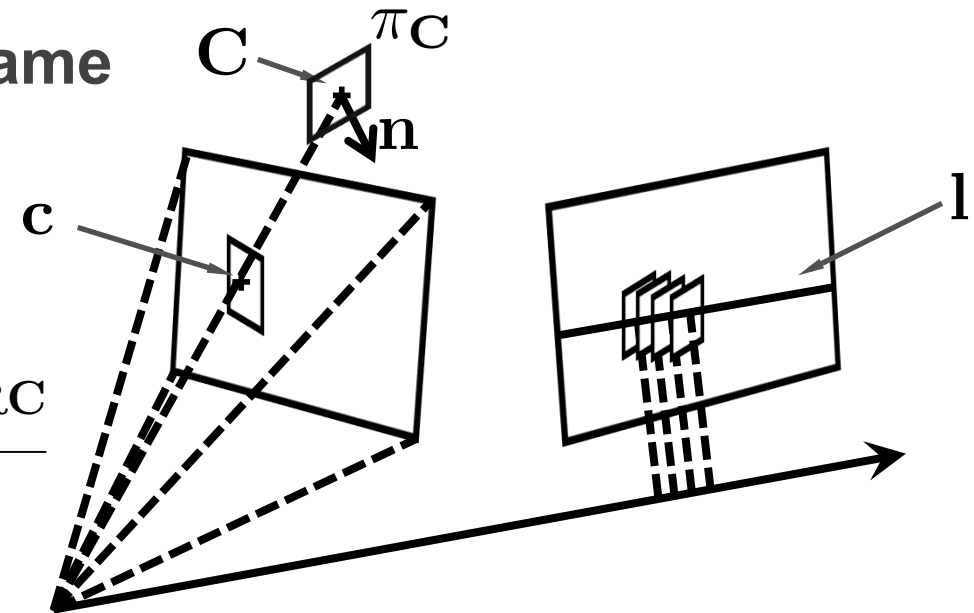
$$\forall \mathbf{c}' \in \mathbf{l}, s = -\frac{\left([\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t}\right)^{\top} [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{RC}}{\left\|[\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t}\right\|^2}$$



## Full Pose from Single Keyframe *Initial Estimates*

- Every point on a line gives a scale sample

$$\forall \mathbf{c}' \in \mathbf{l}, s = -\frac{\left( [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t} \right)^{\top} [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{RC}}{\left\| [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t} \right\|^2}$$



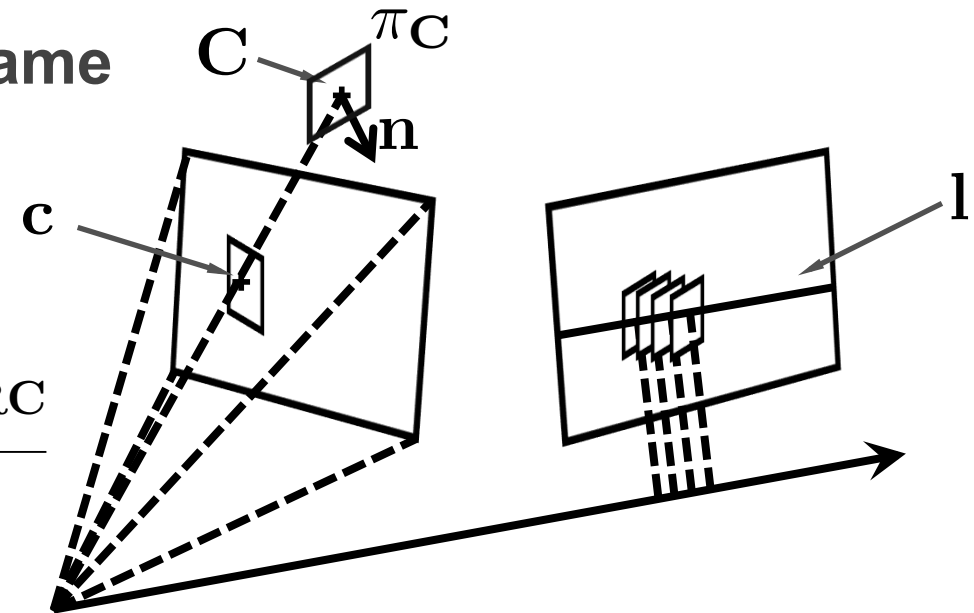
## Full Pose from Single Keyframe *Initial Estimates*

- Every point on a line gives a scale sample

$$\forall \mathbf{c}' \in \mathbf{l}, s = -\frac{\left([\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t}\right)^{\top} [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{R} \mathbf{C}}{\left\|[\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t}\right\|^2}$$

- We can define a local warping from the source to the target image

$$\mathbf{H}(s, \pi_{\mathbf{C}}) = \mathbf{R} - s \frac{\mathbf{t} \mathbf{n}^{\top}}{d}$$

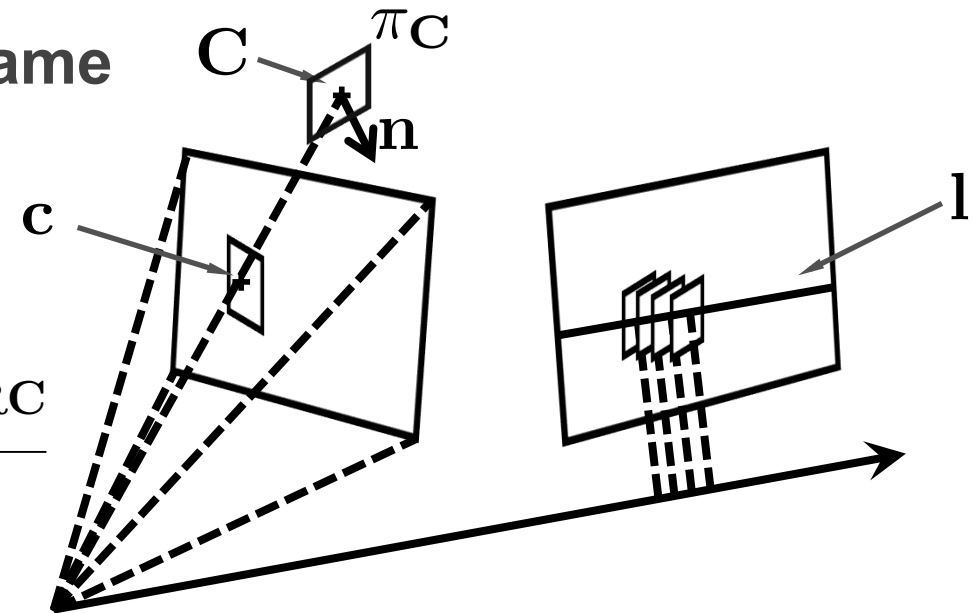


## Full Pose from Single Keyframe

### *Initial Estimates*

- Every point on a line gives a scale sample

$$\forall \mathbf{c}' \in \mathbf{l}, s = - \frac{\left( [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t} \right)^{\top} [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{RC}}{\left\| [\mathbf{K}_t^{-1} \mathbf{c}']_{\times} \mathbf{t} \right\|^2}$$



- We can define a local warping from the source to the target image

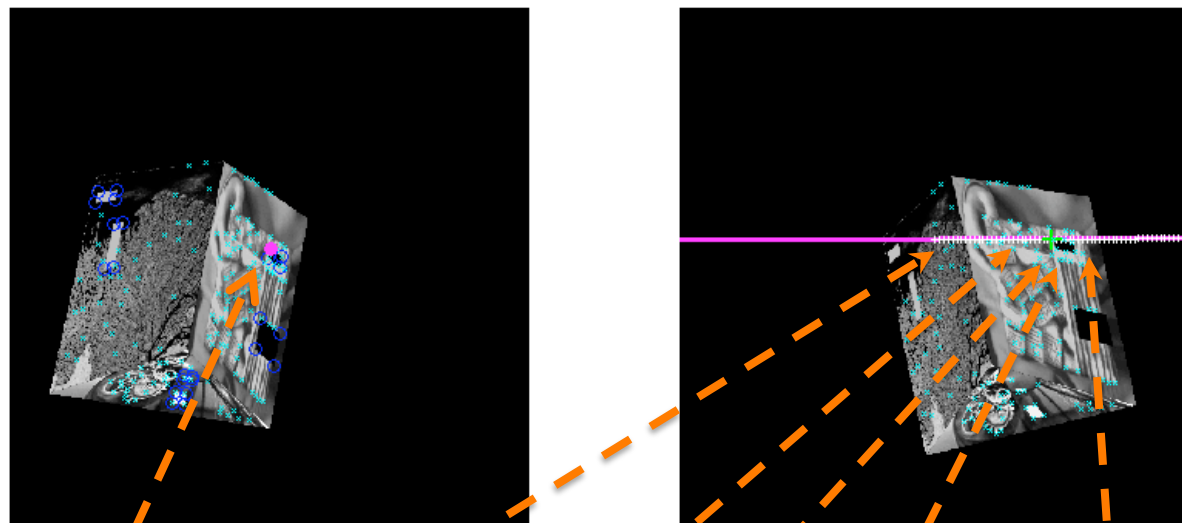
$$\mathbf{H}(s, \pi_{\mathbf{C}}) = \mathbf{R} - s \frac{\mathbf{t} \mathbf{n}^{\top}}{d}$$

- Template search

$$f(s) = \text{SM}(\mathcal{S}, \mathbf{H}^{-1}(s, \pi_{\mathbf{C}})(\mathcal{T}))$$

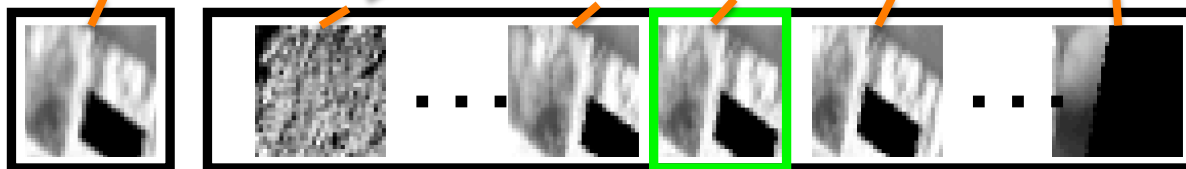
# Full Pose from Single Keyframe

## Initial Estimates - Overview



Keyframe

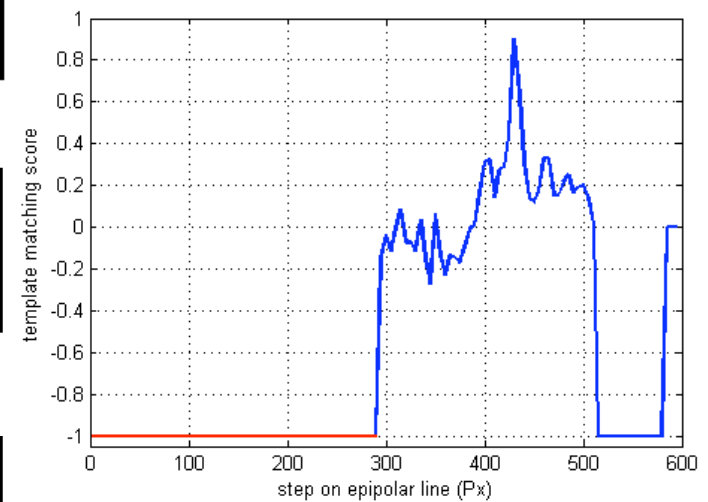
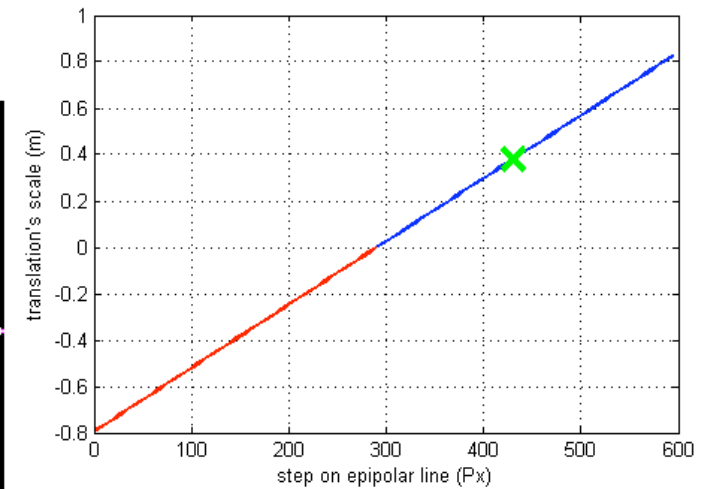
Target Image



Template

Warped Templates from Target

NCC	-0.146	0.437	0.906	0.631	0.164
-----	--------	-------	-------	-------	-------





## Full Pose from Single Keyframe *Nonlinear Refinement*

- Sub-optimal solution

$$\sum_i \text{SM} (\mathcal{S}, \mathbf{H}^{-1} (s, \pi_{c_i}) (\mathcal{T}))$$

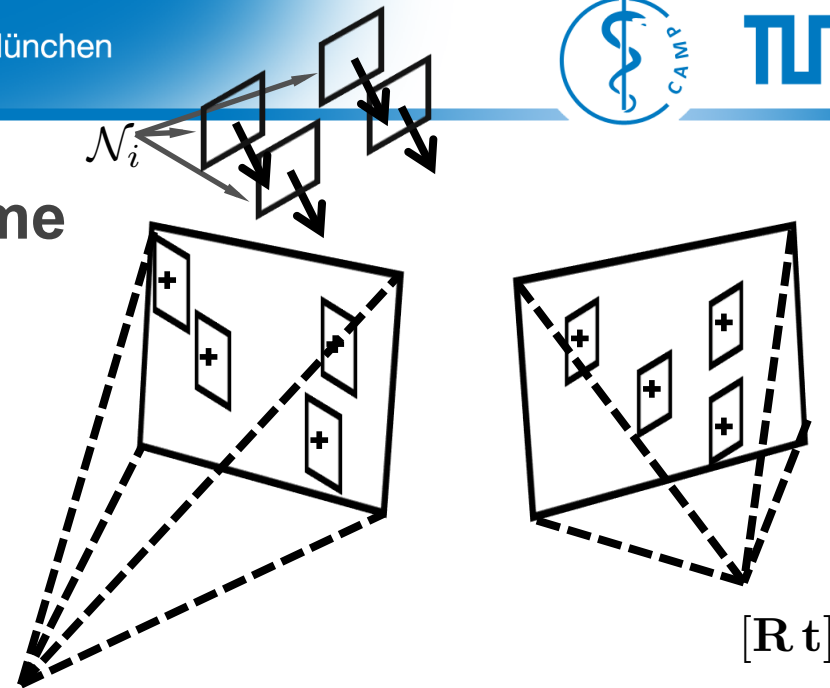
## Full Pose from Single Keyframe *Nonlinear Refinement*

- Sub-optimal solution

$$\sum_i \text{SM} (\mathcal{S}, \mathbf{H}^{-1} (s, \pi_{c_i}) (\mathcal{T}))$$

- We introduce a quadratic cost  $\mathbf{t} \leftarrow s\mathbf{t}$

$$\mathcal{C}_{\mathcal{P}} (\mathbf{R}, \mathbf{t}) = \sum_{j=1}^m \sum_{\mathbf{X}}^{\mathcal{N}_{\mathcal{C}_j}} \|\mathcal{S} (\mathbf{K}_s \mathbf{w} (\mathbf{X})) - \mathcal{T} (\mathbf{K}_t \mathbf{w} (\mathbf{R}\mathbf{X} + \mathbf{t}))\|^2$$



## Full Pose from Single Keyframe Nonlinear Refinement

- Sub-optimal solution

$$\sum_i \text{SM} (\mathcal{S}, \mathbf{H}^{-1} (s, \pi_{c_i}) (\mathcal{T}))$$

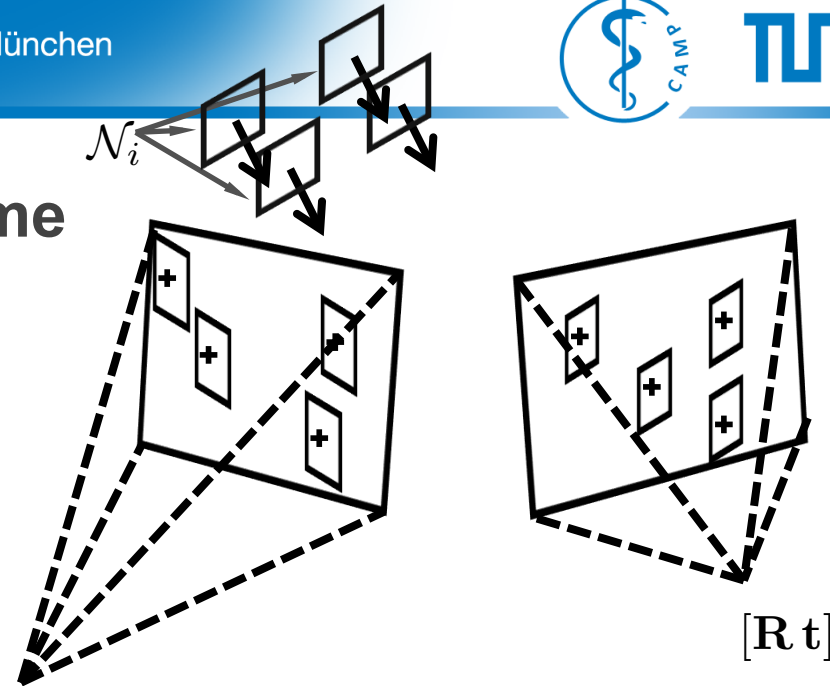
- We introduce a quadratic cost  $\mathbf{t} \leftarrow s\mathbf{t}$

$$\mathcal{C}_{\mathcal{P}} (\mathbf{R}, \mathbf{t}) = \sum_{j=1}^m \sum_{\mathbf{X}}^{\mathcal{N}_{\mathcal{C}_j}} \|\mathcal{S} (\mathbf{K}_s \mathbf{w} (\mathbf{X})) - \mathcal{T} (\mathbf{K}_t \mathbf{w} (\mathbf{R}\mathbf{X} + \mathbf{t}))\|^2$$

- Least square minimization

$$\arg \min_{\mathbf{M}_i, \mathbf{R}, \mathbf{t}} \mathcal{C}_{\mathcal{G}} (\mathbf{M}_i, \mathbf{R}, \mathbf{t}) + \mathcal{C}_{\mathcal{P}} (\mathbf{R}, \mathbf{t})$$

$$\text{with } \mathcal{C}_{\mathcal{G}} (\mathbf{M}_i, \mathbf{R}, \mathbf{t}) = \sum_{i=1}^n + \begin{matrix} \|\mathbf{K}_s \mathbf{w} (\mathbf{M}_i) - \mathbf{p}_i\|^2 \\ \|\mathbf{K}_t \mathbf{w} (\mathbf{R}\mathbf{M}_i + \mathbf{t}) - \mathbf{q}_i\|^2 \end{matrix}$$



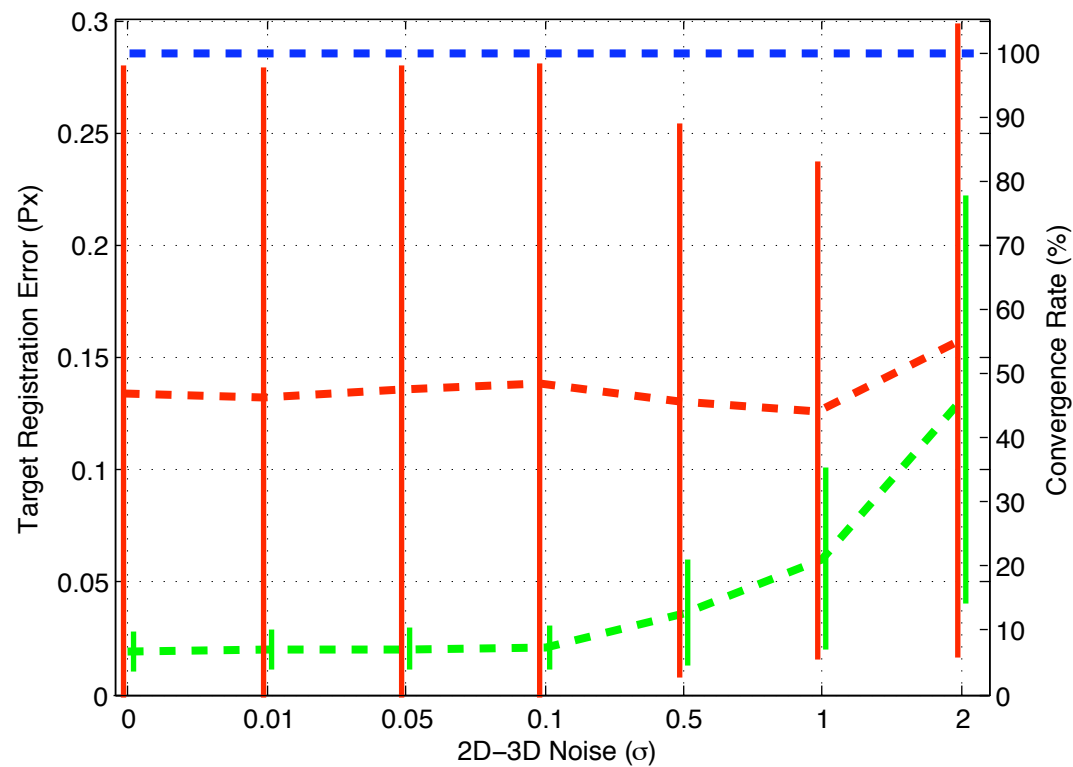
# Full Pose from Single Keyframe

## *Algorithm Overview*

1. Estimate  $E$
2. Decompose  $E$  in  $R$  and  $t$
3. for each 2D-3D correspondences
  - Find scale  $s$
4. end
5. Select the scale  $s$  with best consensus
6. Nonlinear estimation

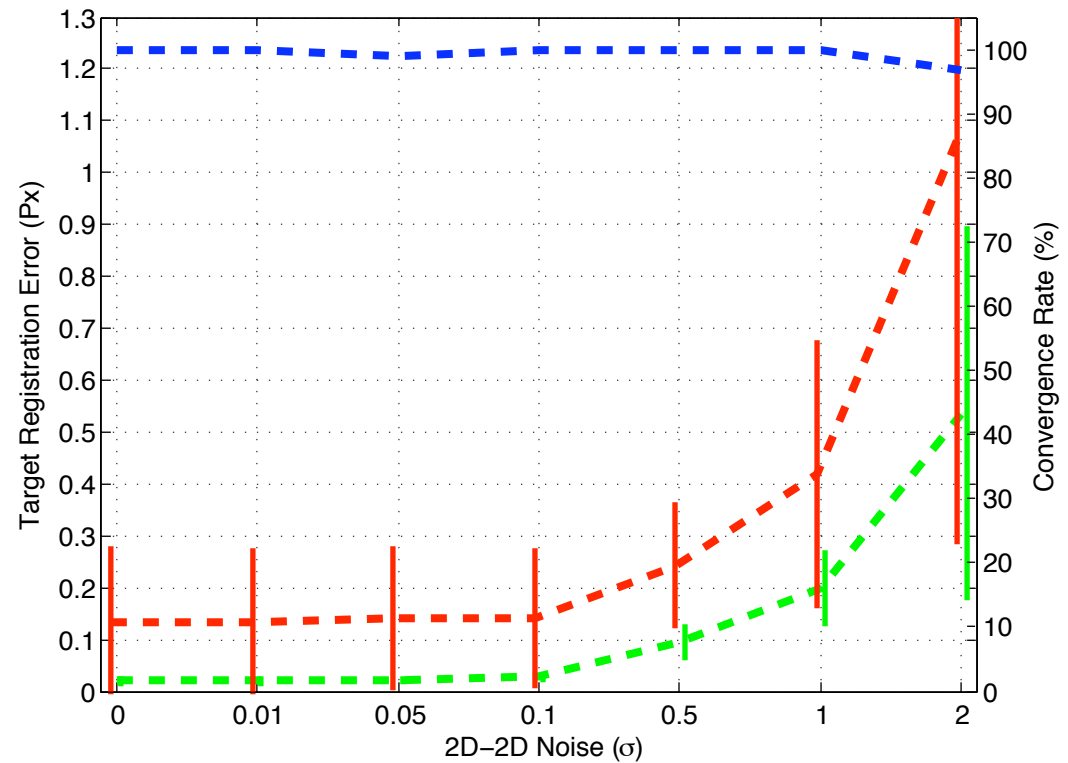
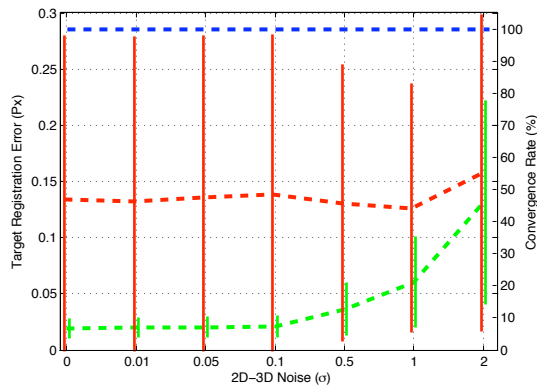
# Full Pose from Single Keyframe

## *Synthetic Experiments*



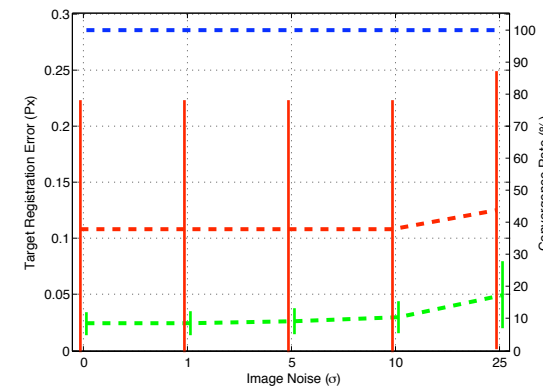
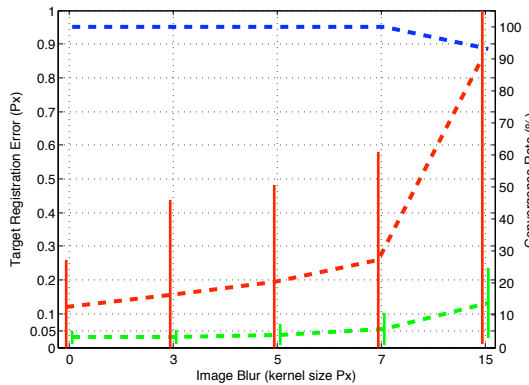
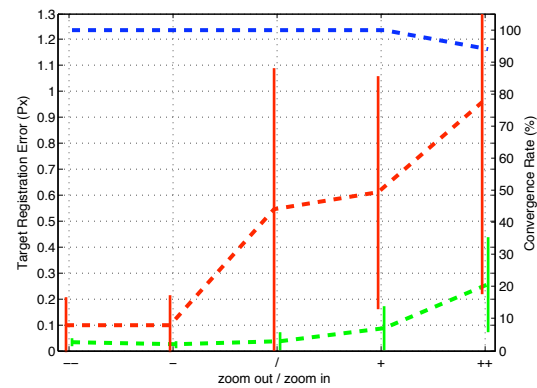
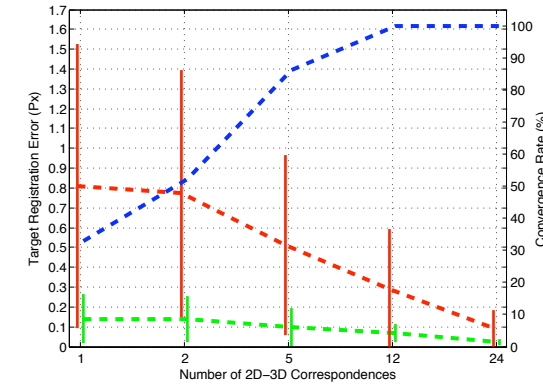
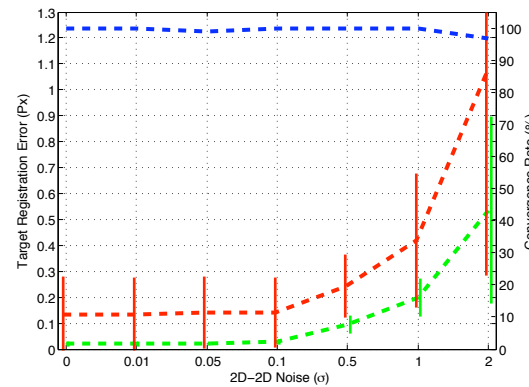
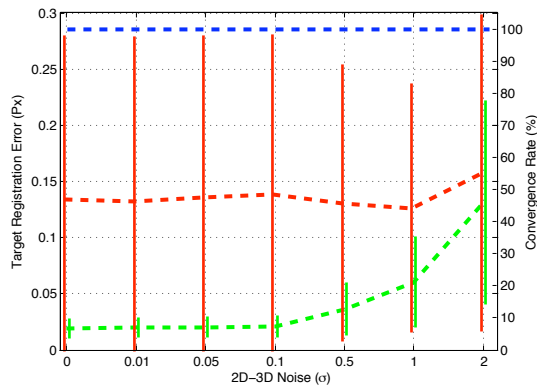
# Full Pose from Single Keyframe

## *Synthetic Experiments*

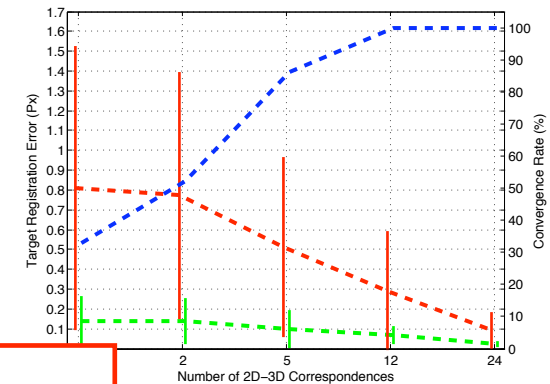
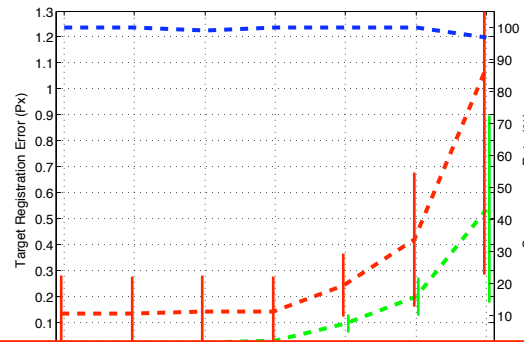
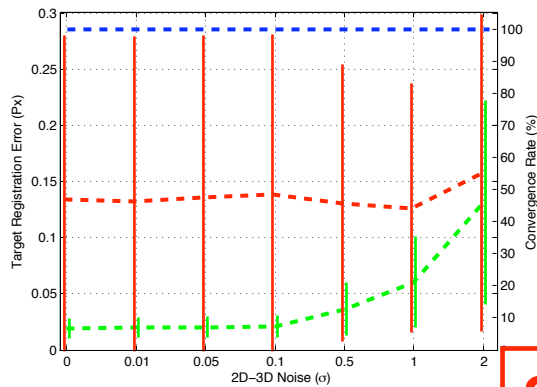


# Full Pose from Single Keyframe

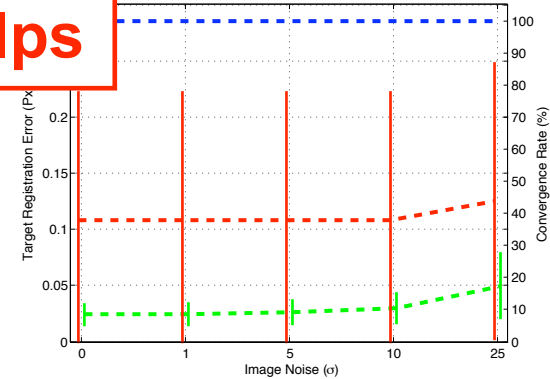
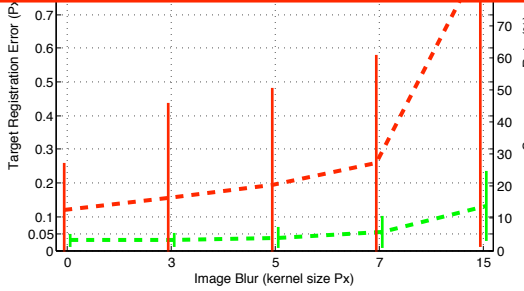
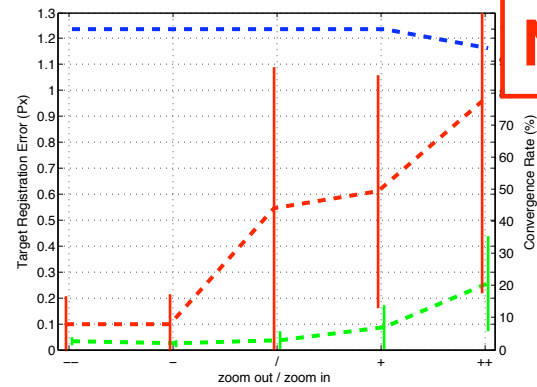
## Synthetic Experiments



# Full Pose from Single Keyframe Synthetic Experiments



**Stable and precise results  
Non-linear refinement helps**

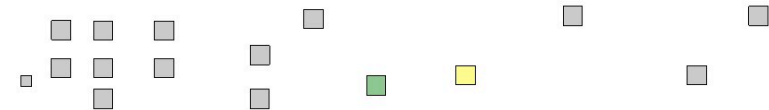




# Discrepancy Check using Augmented Reality

## *Keyframe Computation*

- Landmark-based registration
  - Register images to the CAD coordinate system
- Approach
  - Extract landmarks (Anchor-Plates) from images
  - Find corresponding 3D landmarks
  - Compute the pose (Rotation, Translation)
- Information stored
  - Image
  - Full pose
  - 3D points: *location and normal*.

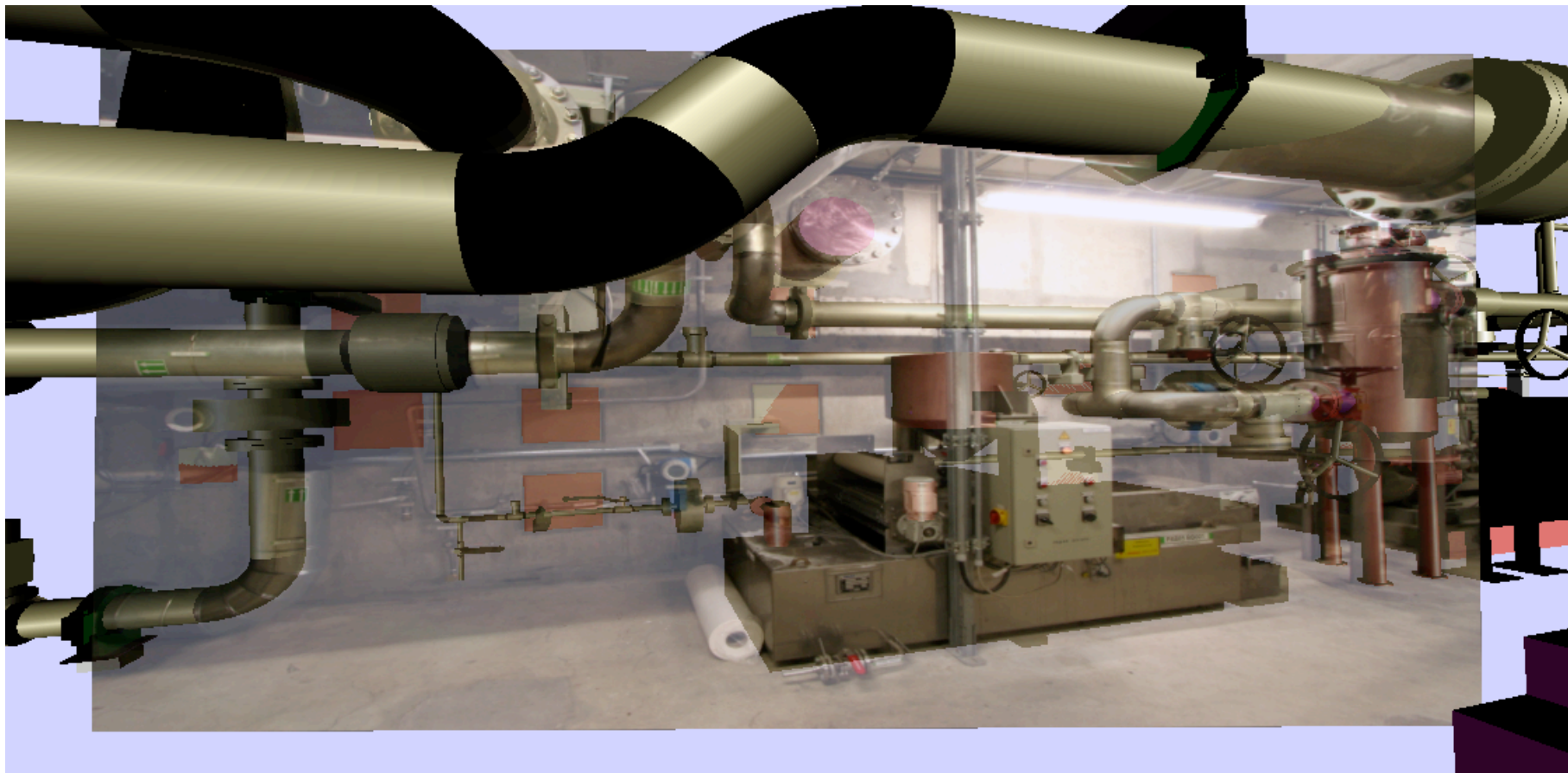


# Discrepancy Check using Augmented Reality *Results*



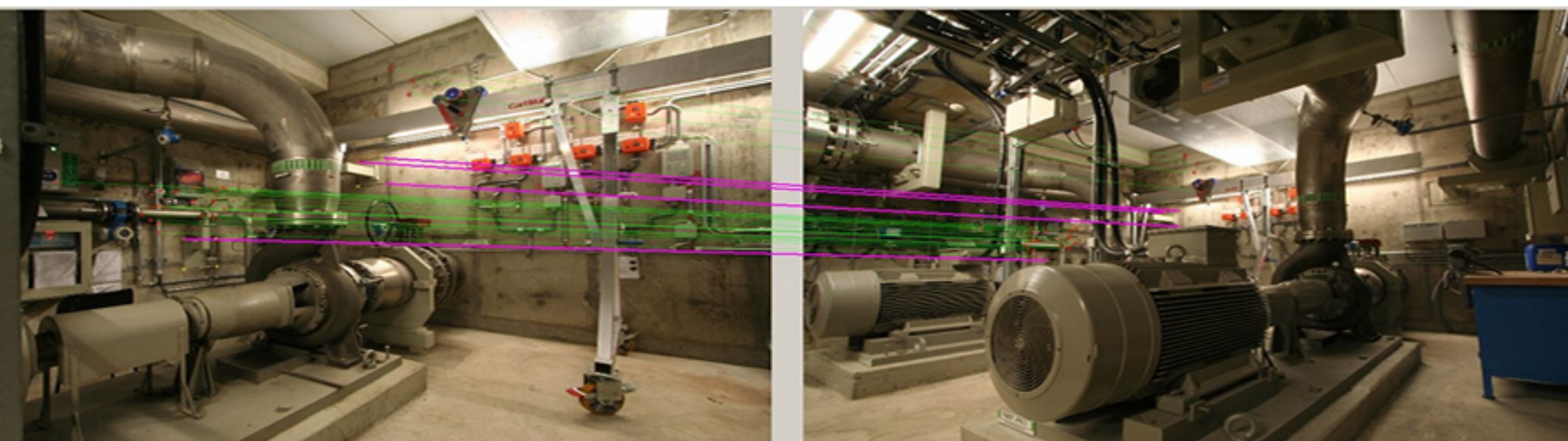
*Robust Features Matches and Propagated 2D-3D Correspondences*

## Discrepancy Check using Augmented Reality *Results*



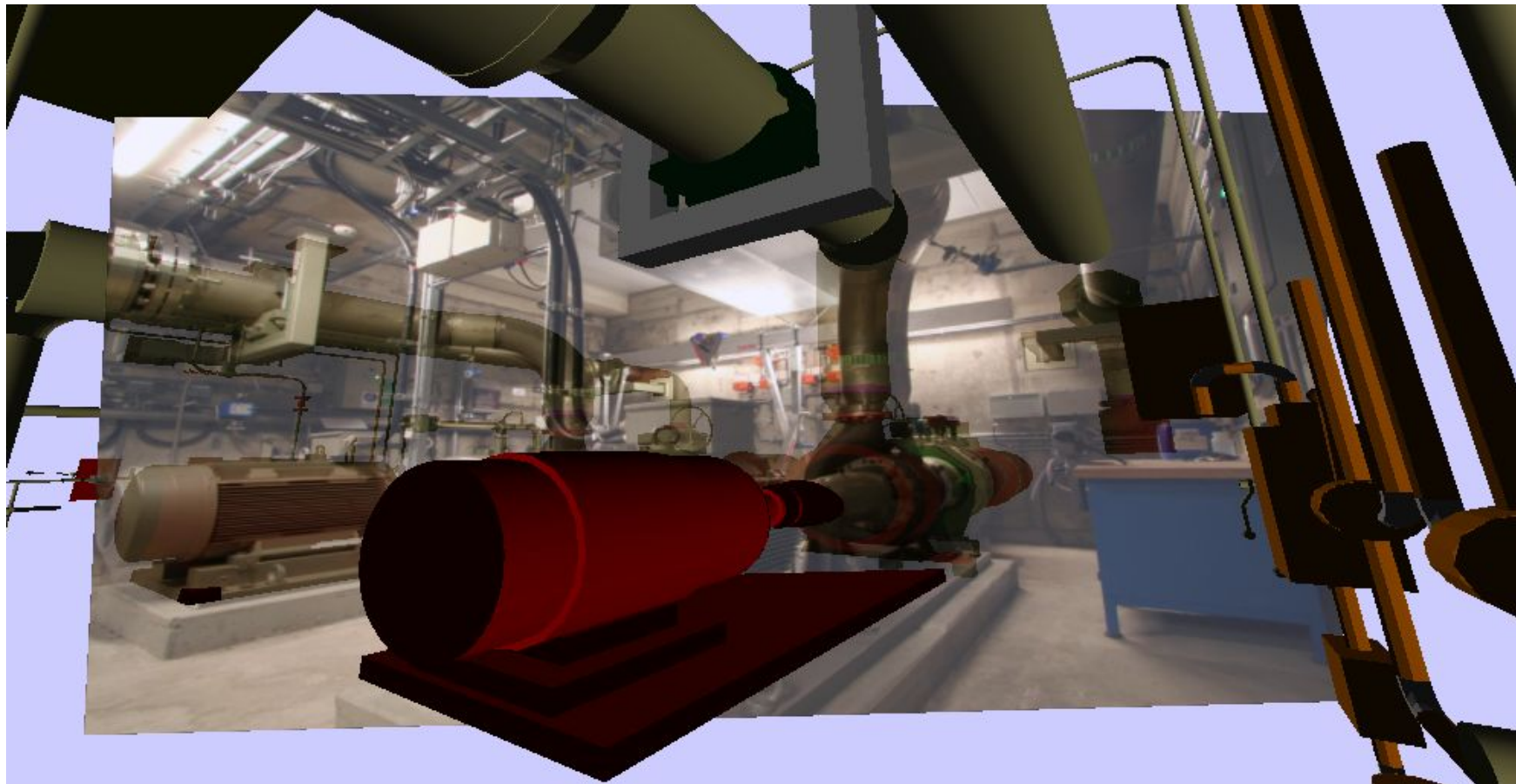
*Resulting Augmentation*

# Discrepancy Check using Augmented Reality *Results*



*Robust Features Matches and Propagated 2D-3D Correspondences*

## Discrepancy Check using Augmented Reality *Results*

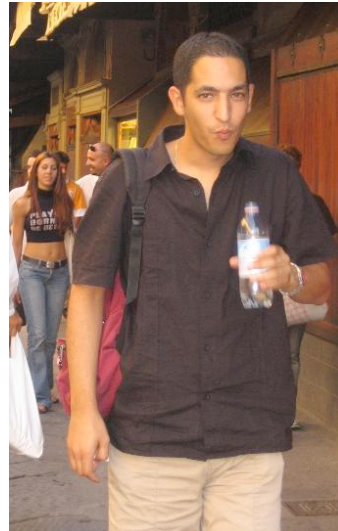


*Resulting Augmentation*

## I did not do this work alone



*Pierre Schroeder*



*Selim Benhimane*



*Juergen Sotke*



*Stefan Hinterstoiser*



*Stefan Holzer*

But also

Adrien Bartoli (universite d'Auvergne)

Bernard Ziesl

Mirko Appel (siemens CT)

Sumit Paranjape, Xinxing Feng, Xavier Fernandez, Hagen Kaiser

## Conclusion and Perspectives

### About this method

- ✓ Automatic full pose estimation
  - ✓ Perspectively corrected template matching
  - ✓ New non-linear cost function
- ✓ Solution is used on site for inspection

### About the application

- ✓ Landmark based registration
- ✓ Automatic 6 DoF stereo registration
- ✓ Bundle Adjustment with fix 3D points
- ✓ 3D Measurements
- ✓ Inspection / Documentation

## Future Work

- Extension to more than two images
- Non calibrated case
- Estimation of the normals
- Mobile AR using keyframes
- Guided 3D reconstruction
- 3D Annotations



**Thank for the attention**

(Any) **Questions?**



## Bibliography

- [1] Goose *et al.* Speech-enabled augmented reality supporting mobile industrial maintenance. Pervasive Computing 2003.
- [2] Pentenrieder *et al.* Augmented Reality-based factory planning - an application tailored to industrial needs. ISMAR, 2007.
- [3] Webster *et al.* Architectural Anatomy. Presence, 1995.
- [4] Schoenfelder & Schmalstieg. Augmented Reality for Industrial Building Acceptance. IEEE VR, 2008.
- [5] Schall *et al.* Virtual redlining for civil engineering in real environments. ISMAR, 2008.
- [6] Lowe. Fitting parameterized three-dimensional models to images. IEEE Trans. PAMI, 1991.
- [7] Drummond & Cipolla. Real-time tracking of complex structures with on-line camera calibration. BMVC, 1999.
- [8] Chia *et al.* Online 6 dof augmented reality registration from natural features. ISMAR, 2002.
- [9] Vacchetti *et al.* Stable real-time 3d tracking using online and offline information. IEEE Trans. PAMI, 2004.
- [10] Platonov *et al.* A mobile markerless AR system for maintenance and repair. ISMAR, 2006.
- [11] Hung and Faugeras. Some properties of the E matrix in two-view motion estimation. IEEE Trans. PAMI, 1989.
- [12] Georgel *et al.* A Unified Approach Combining Photometric and Geometric Information for Pose Estimation. BMVC, 2008.
- [13] Georgel *et al.* An Industrial Augmented Reality Solution For Discrepancy Check. ISMAR, 2007.