Contributions to Active Visual Estimation and Control of Robotic Systems

Riccardo Spica

Lagadic group

Inria Rennes Bretagne Atlantique & IRISA

http://www.irisa.fr/lagadic

Perception and action are interconnected

- robotics is [B. Siciliano and O. Khatib, 2008]

  “the science that studies the intelligent connection between perception and action”

Perception and action are interconnected

- robotics is [B. Siciliano and O. Khatib, 2008]

  "the science that studies the intelligent connection between perception and action"

estimation accuracy affects control performance for tasks that depend on the estimation itself

system (controlled) trajectory affects estimation accuracy for nonlinear sensor-to-state mapping

Perception and action are interconnected

- robotics is [B. Siciliano and O. Khatib, 2008]

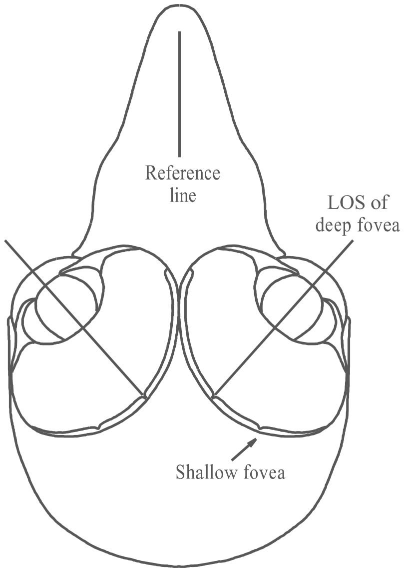
  "the science that studies the intelligent connection between perception and action"

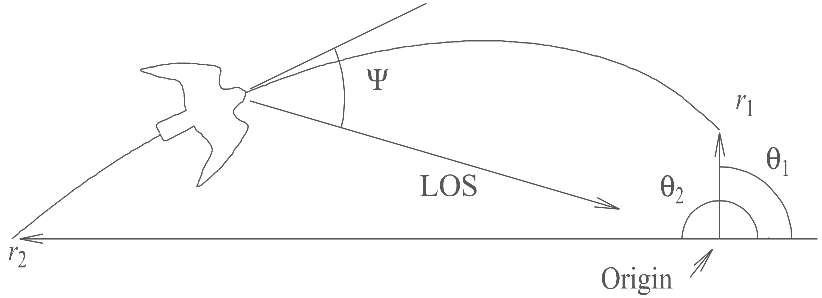

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$$

$$d = \sqrt{X^2 + Y^2 + Z^2}$$

An example from Nature

- raptors approach preys by following spiral trajectories
- falcons acute sight points approx. 45º to the side
- flying on a straight line, would force the falcon to turn its head to the side thus considerably increasing air drag
- joint maximization of both perception and action





[Tucker et Al. 2000]

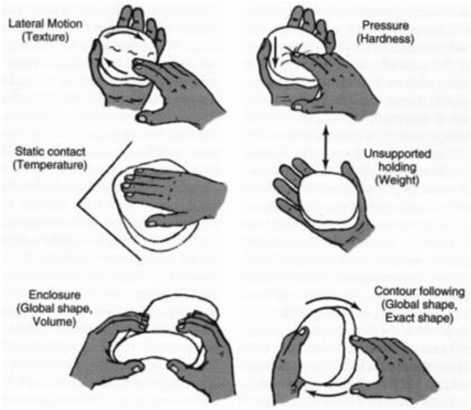
Outline

- introduction and motivation
- active structure estimation from controlled motion
- dense structure estimation from motion
- coupling visual servoing and active estimation
- conclusions and perspectives

Introduction and motivation

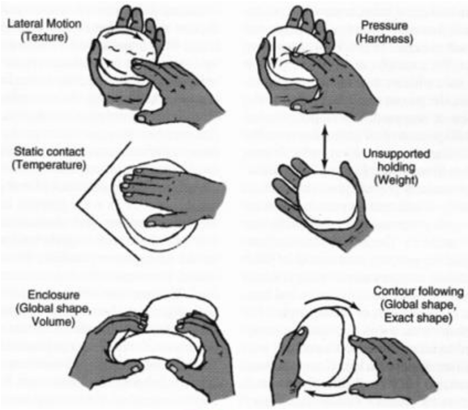
Active perception 1/2

- psychology – active perception
  - *"perceiving is active […].We don't simply see, we look".* [Gibson,1979].



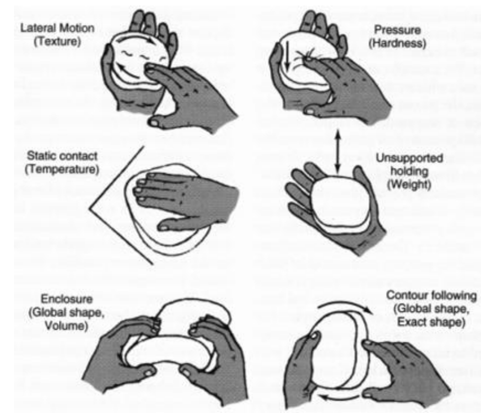
Active perception 1/2

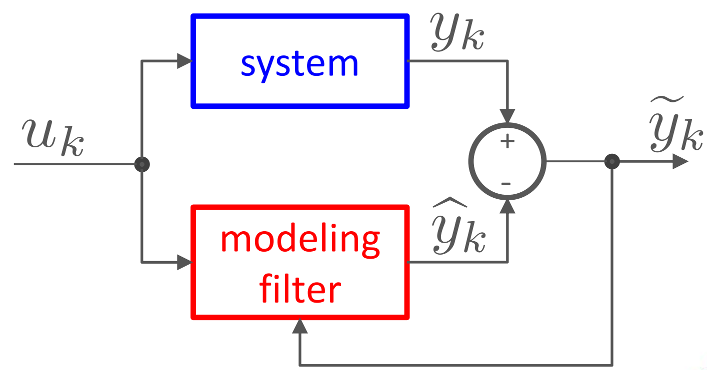
- psychology – active perception
  - *"perceiving is active […].We don't simply see, we look".* [Gibson,1979].
- statistics – experimental design
  - *"the estimation can never exceed the information supplied by the data"* [Fisher, 1947].
  - Cramer-Rao bound $\Sigma \succeq \mathcal{I}_F^{-1}$



Active perception 1/2

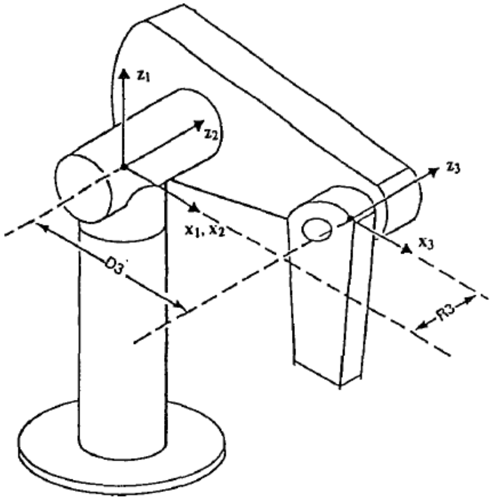
- psychology – active perception
  - *“perceiving is active […].We don’t simply see, we look”.* [Gibson,1979].
- statistics – experimental design
  - *“the estimation can never exceed the information supplied by the data”* [Fisher, 1947].
  - Cramer-Rao bound $\Sigma \succeq \mathcal{I}_F^{-1}$
- system identification and adaptive control – input design or optimal experiment design
  - persistence of excitation [Anderson, 1977]
  - observability Gramian [Kailath, 1980]





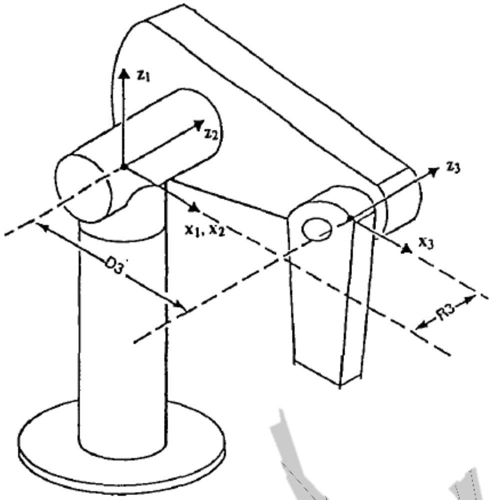
Active perception 2/2

- experimental robot kinematic/dynamic calibration
  - [Gautier, Khalil, 1992]

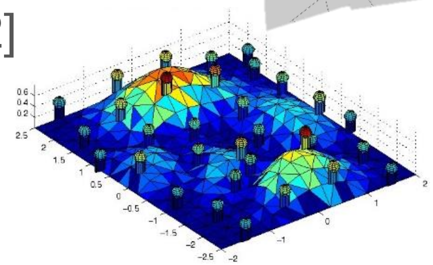


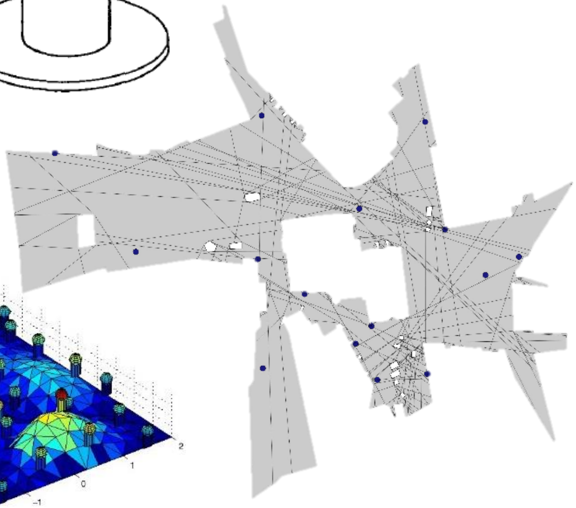
Active perception 2/2

- experimental robot kinematic/dynamic calibration
  - [Gautier, Khalil, 1992]

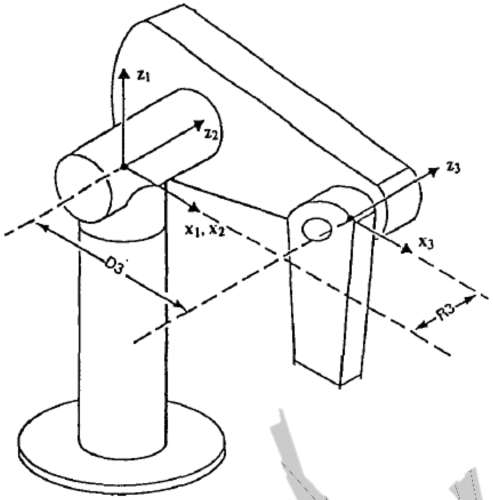


- optimal/dynamic sensor network placement
  - maximum coverage [Cortes et al. 2002]
  - art gallery problem [Borrmann et al. 2013]

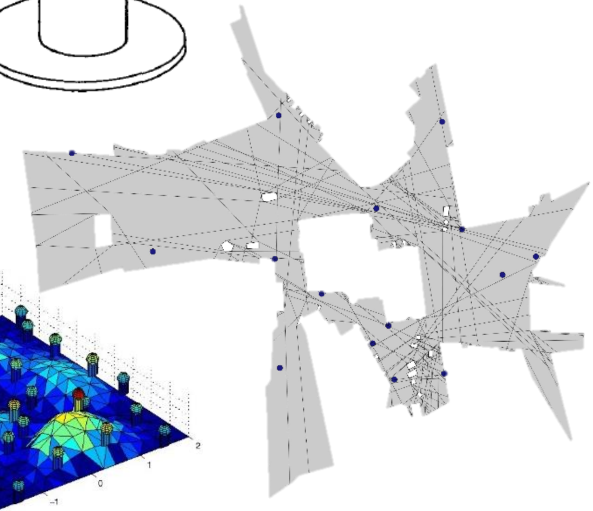


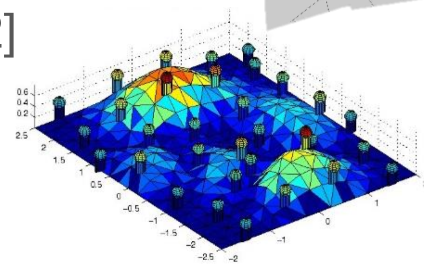


Active perception 2/2

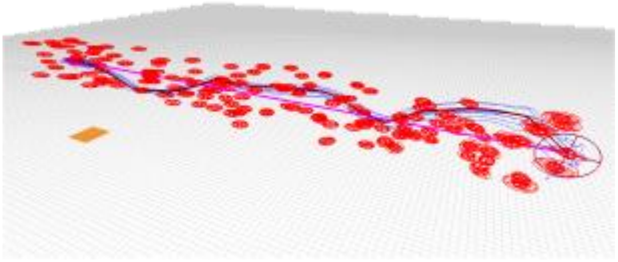


- experimental robot kinematic/dynamic calibration
  - [Gautier, Khalil, 1992]





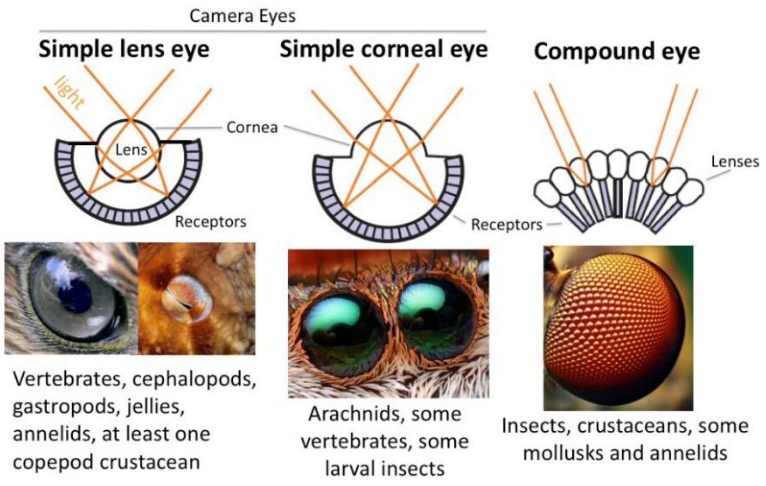
- optimal/dynamic sensor network placement
  - maximum coverage [Cortes et al. 2002]
  - art gallery problem [Borrmann et al. 2013]



- Active SLAM
  - exploration vs exploitation [Davison, Murray 2002; Achtelik et al. 2013]

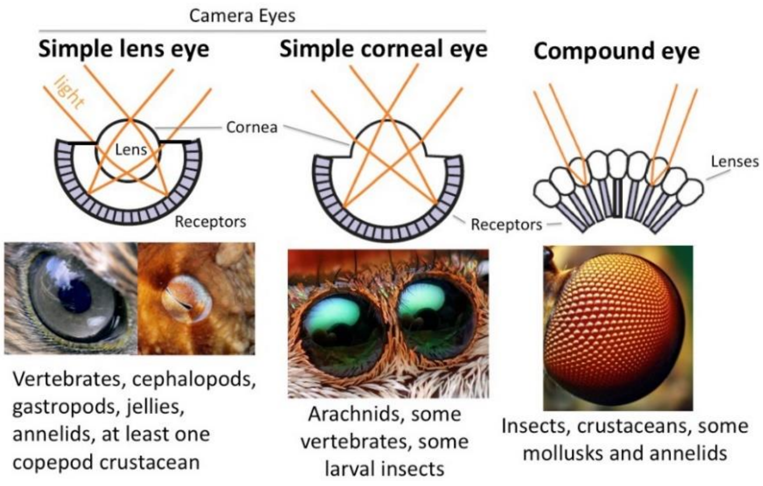
The sense of vision

- a powerful and complex sensor
  - vision takes up to 70% of the brain activity and 50% of neural tissue [Fixot, 1957]
  - almost all animals have eyes in a number of different forms [Land, 05]

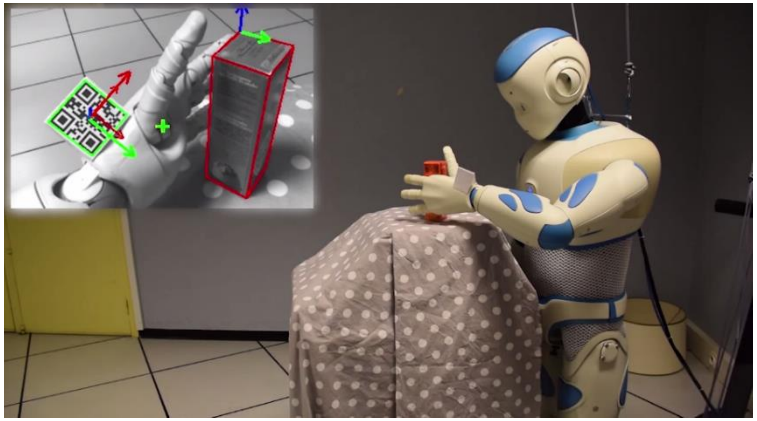


The sense of vision

- a powerful and complex sensor
  - vision takes up to 70% of the brain activity and 50% of neural tissue [Fixot, 1957]
  - almost all animals have eyes in a number of different forms [Land, 05]

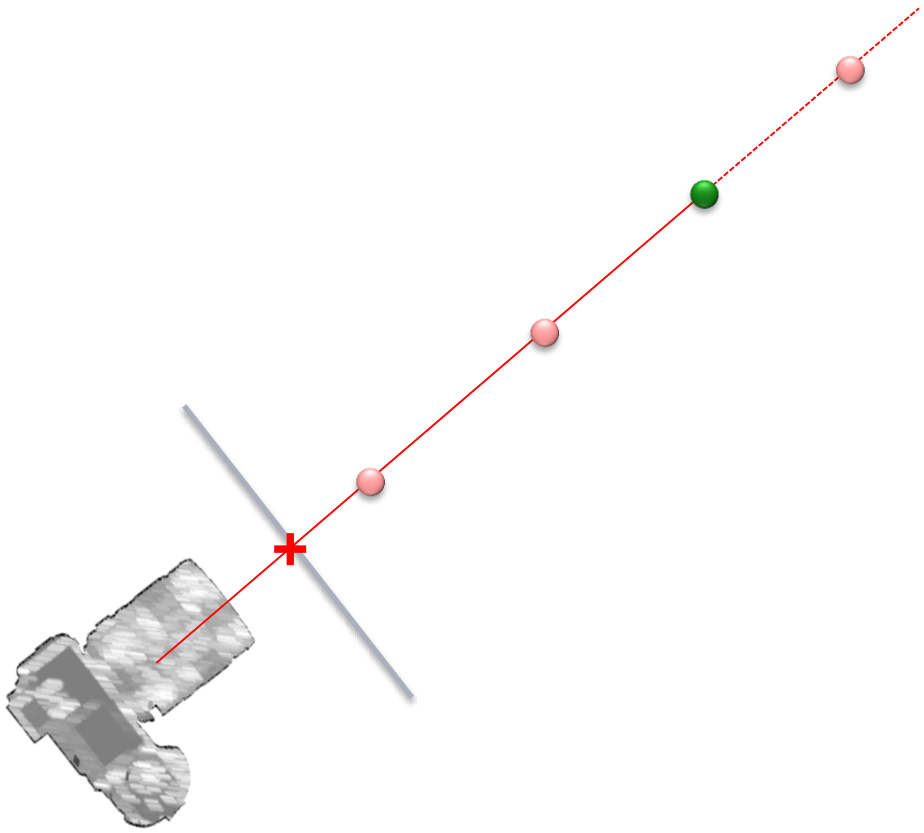


- a representative case study
  - 3D→2D (nonlinear) projection causes information loss [Ma et al., ‘03]
  - estimation performance depends on camera motion [Bajcsy, ‘88; Aloimonos et al. ‘87]

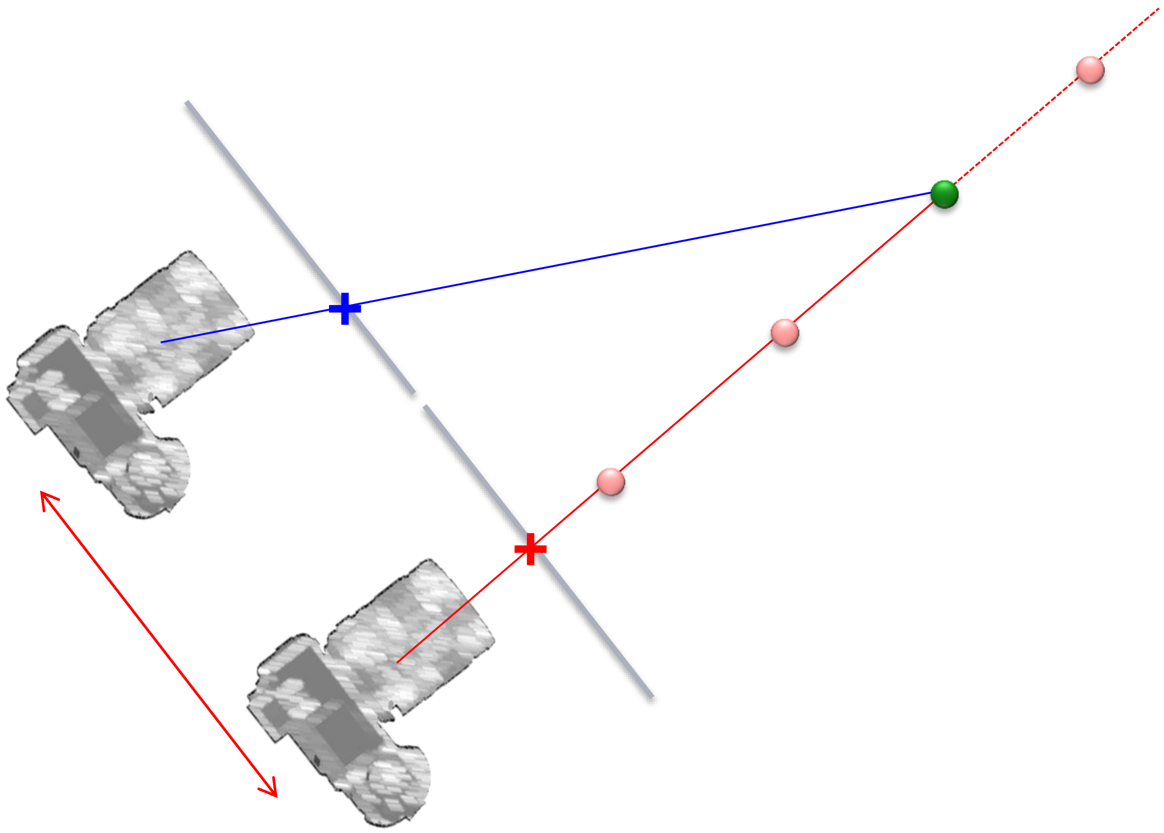


Vision based reconstruction

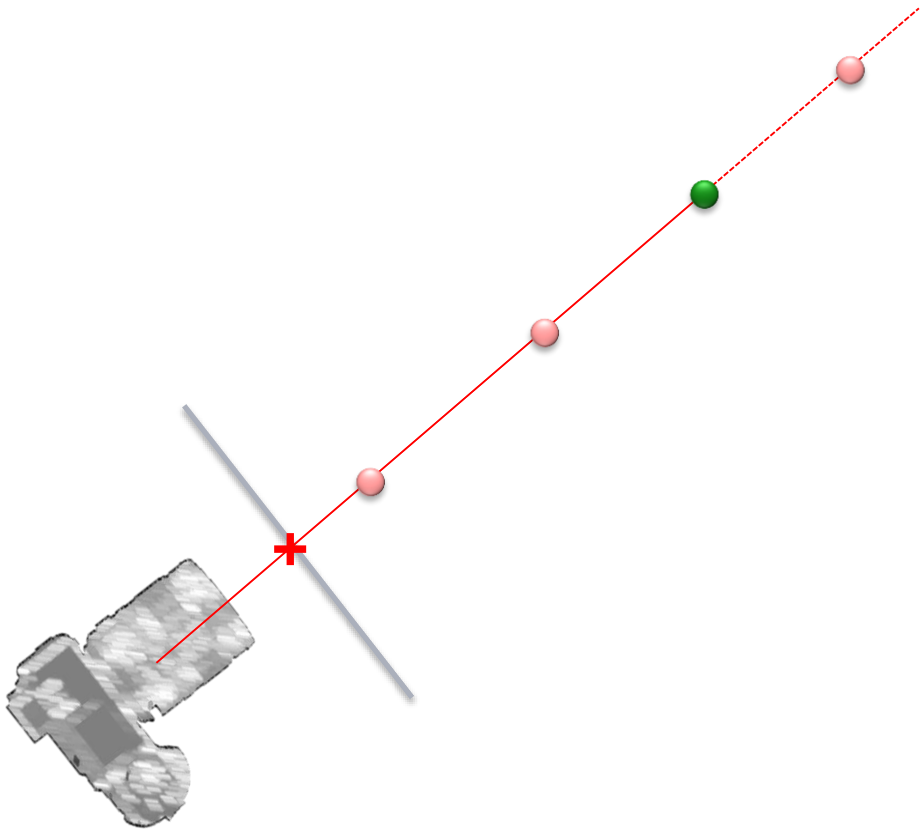
- an inverse problem



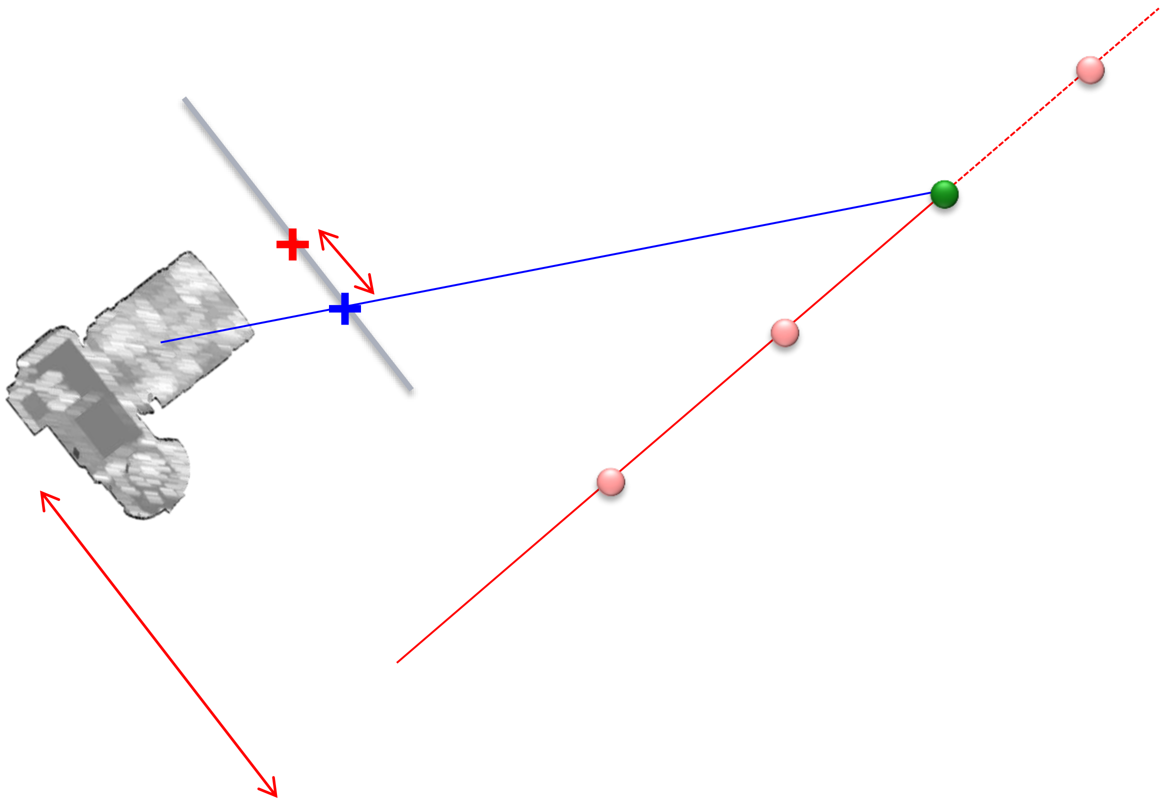
Stereo vision



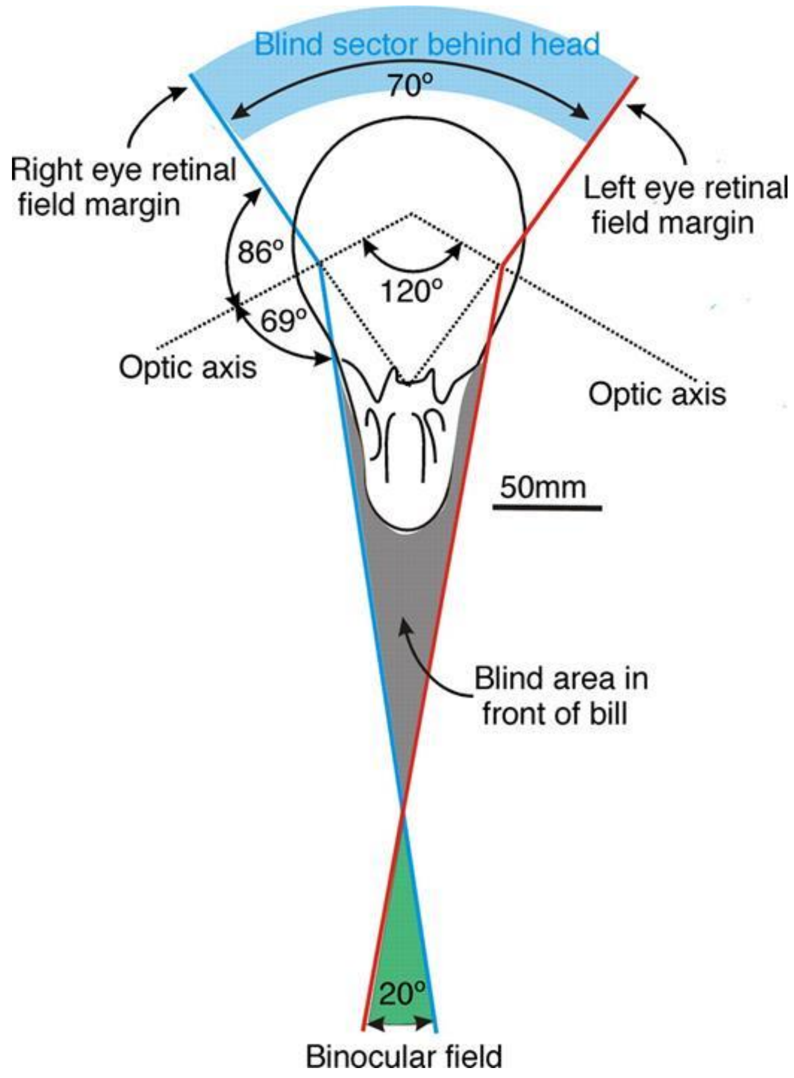
Structure from known motion



Structure from known motion

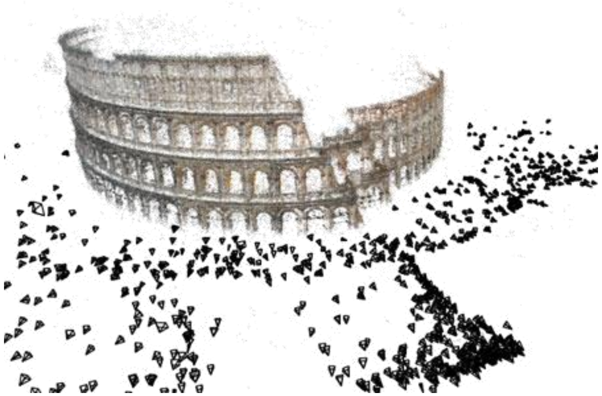


Structure from motion in Nature



Estimation from vision

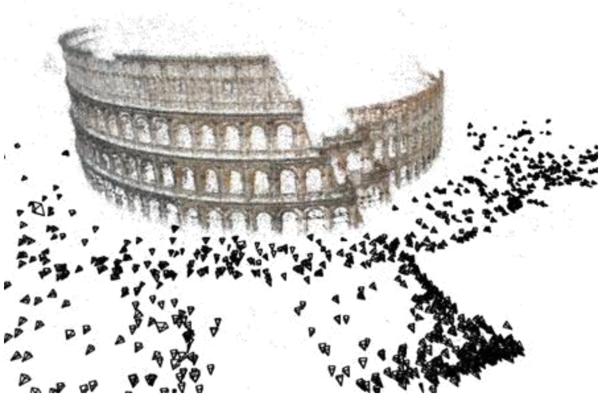
- Structure from Motion (SfM)
  - reconstruct complex scenes and camera poses (up to a scale factor)
  - batch off-line (bundle adjustment)



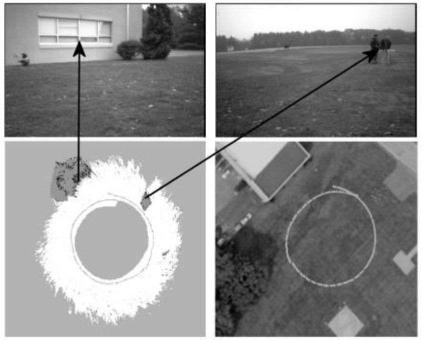
[Agarwal, et al. 2011]

Estimation from vision

- Structure from Motion (SfM)
  - reconstruct complex scenes and camera poses (up to a scale factor)
  - batch off-line (bundle adjustment)
- visual odometry
  - mainly reconstruct camera motion
  - ego-centric/local
  - sequential real-time processing



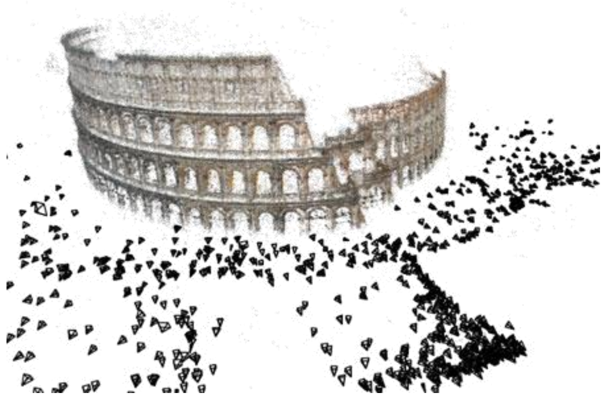
[Agarwal, et al. 2011]



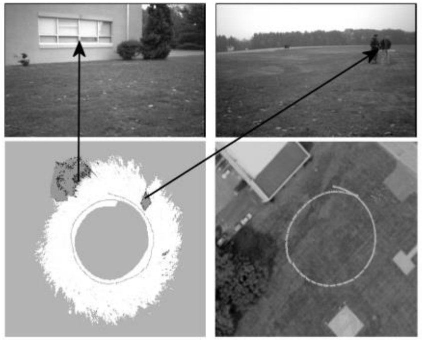
[Nister, et al. 2004]

Estimation from vision

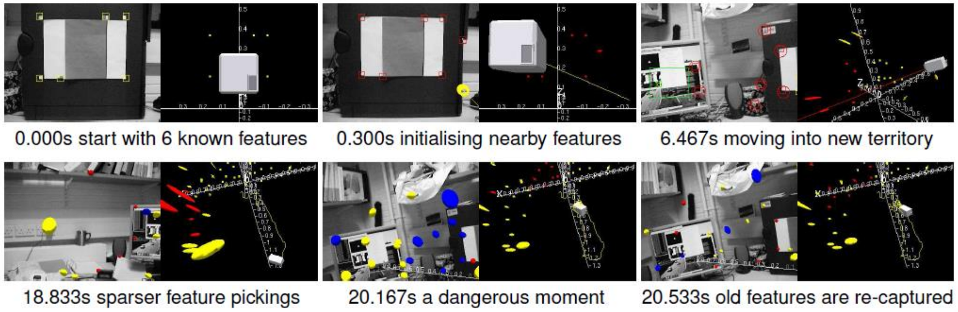
- Structure from Motion (SfM)
  - reconstruct complex scenes and camera poses (up to a scale factor)
  - batch off-line (bundle adjustment)
- visual odometry
  - mainly reconstruct camera motion
  - ego-centric/local
  - sequential real-time processing
- visual-SLAM
  - global map consistency
  - filtering + batch optimization (loop closure)



[Agarwal, et al. 2011]



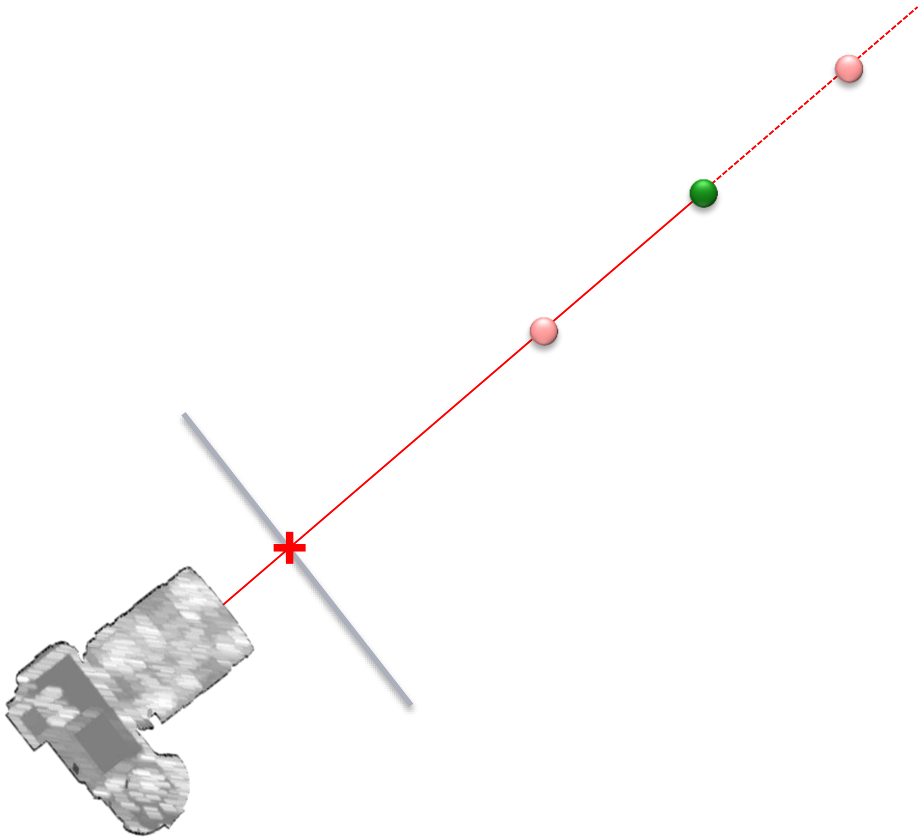
[Nister, et al. 2004]



[Davison, 2003]

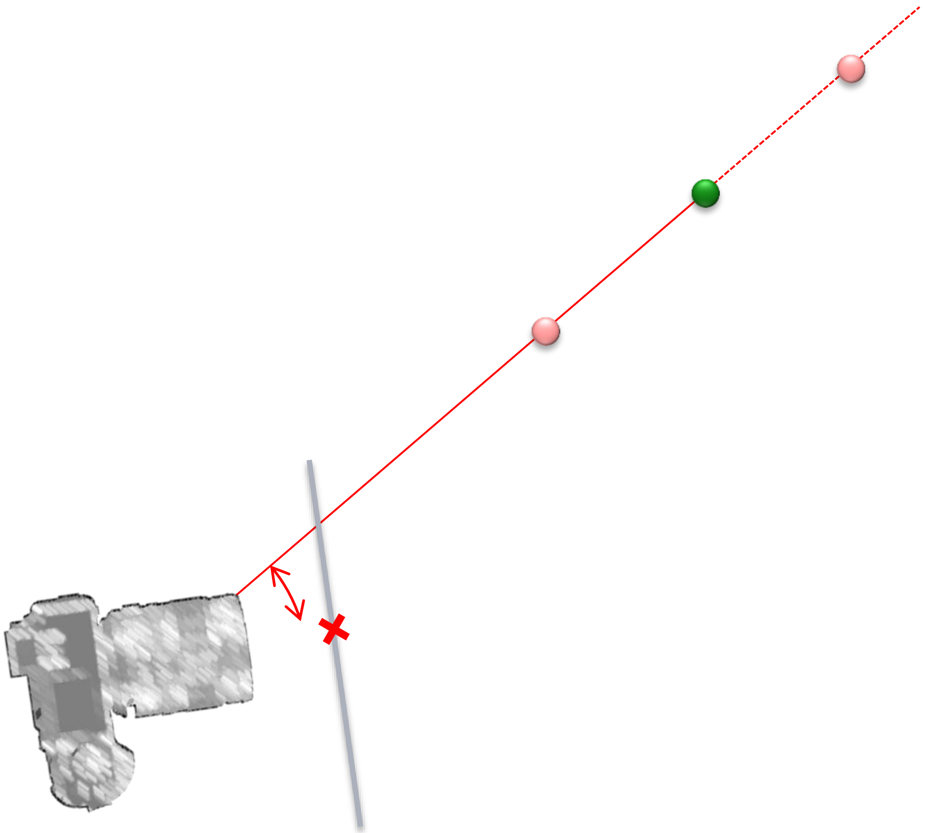
Effects of the camera motion 1/3

- pure rotations are not informative



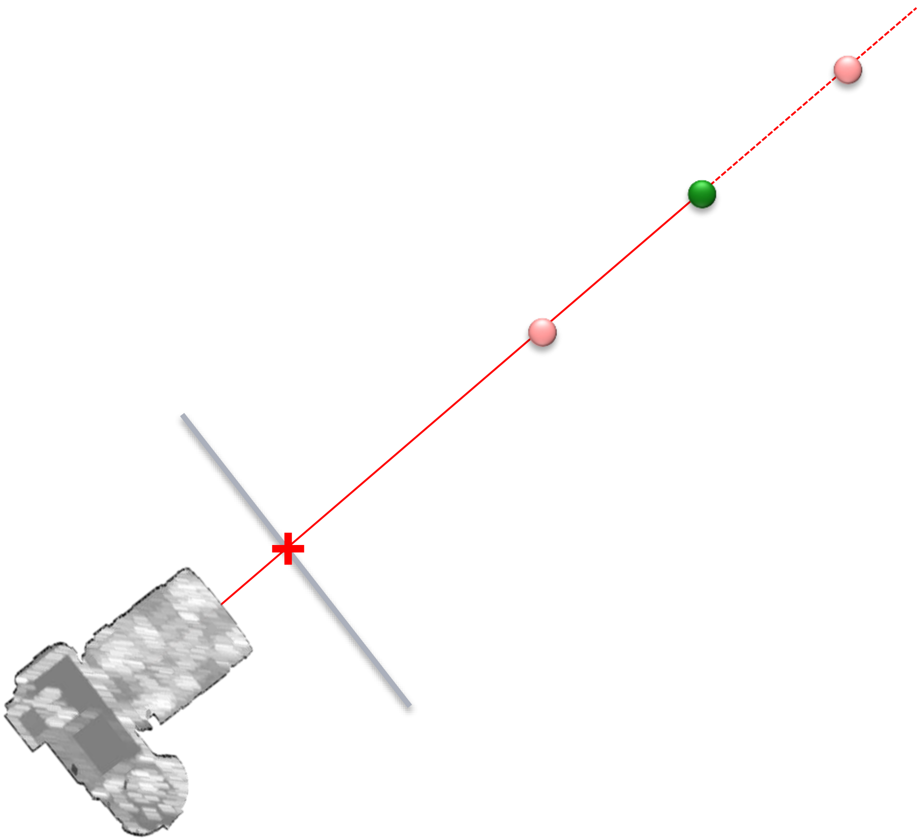
Effects of the camera motion 1/3

- pure rotations are not informative



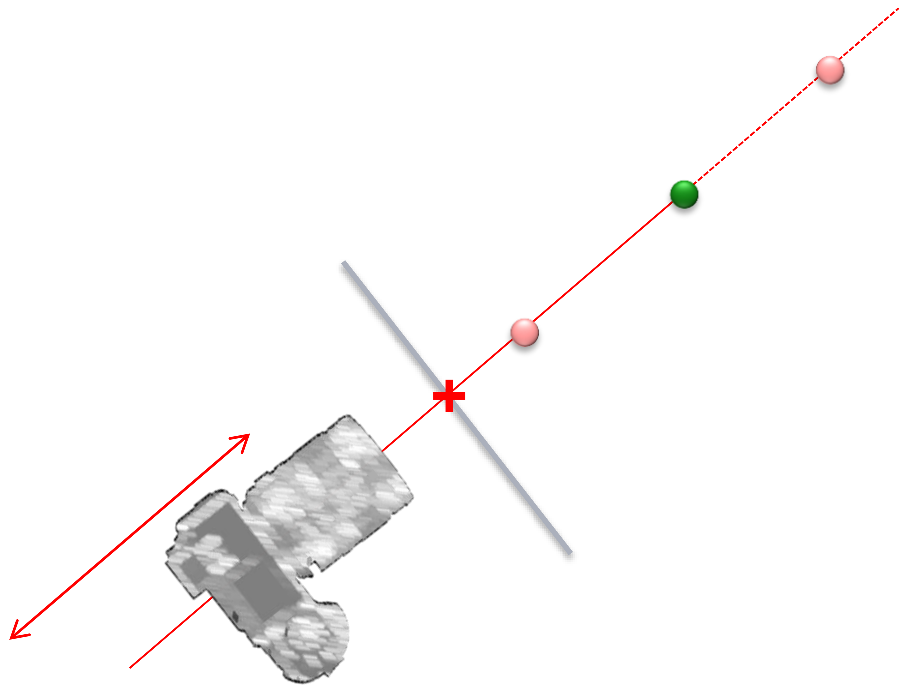
Effects of the camera motion 2/3

- translations along the projection ray of a point are not informative for that point



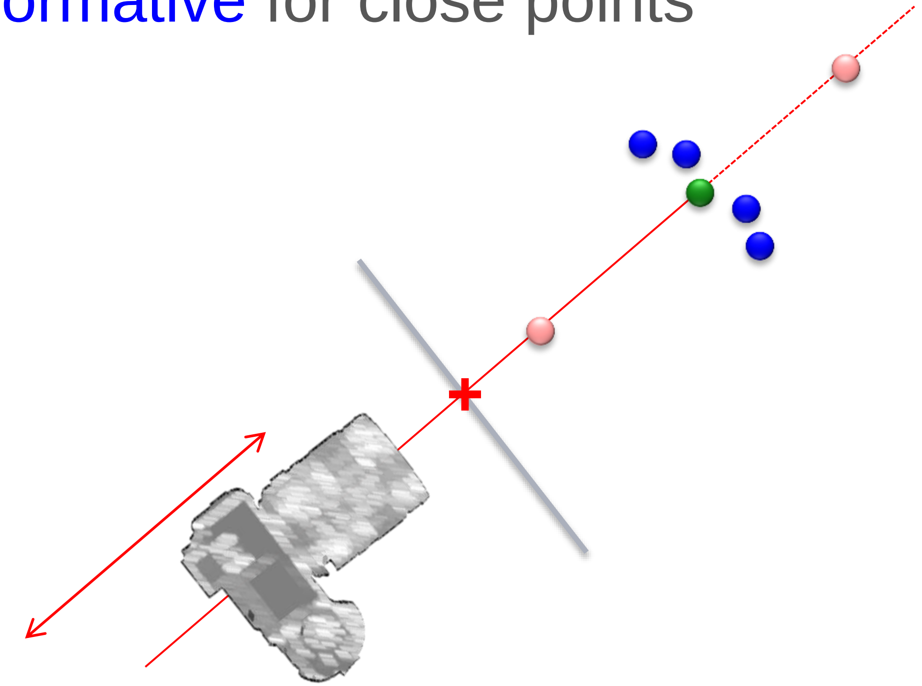
Effects of the camera motion 2/3

- translations along the projection ray of a point are not informative for that point



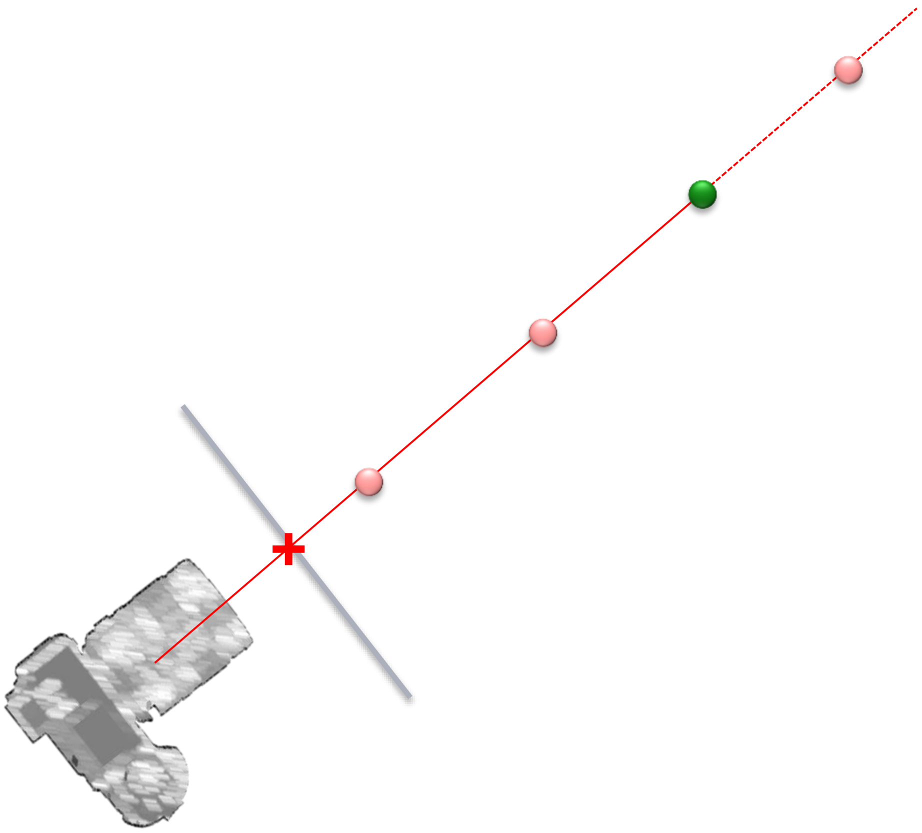
Effects of the camera motion 2/3

- translations along the projection ray of a point are not informative for that point
- and poorly informative for close points



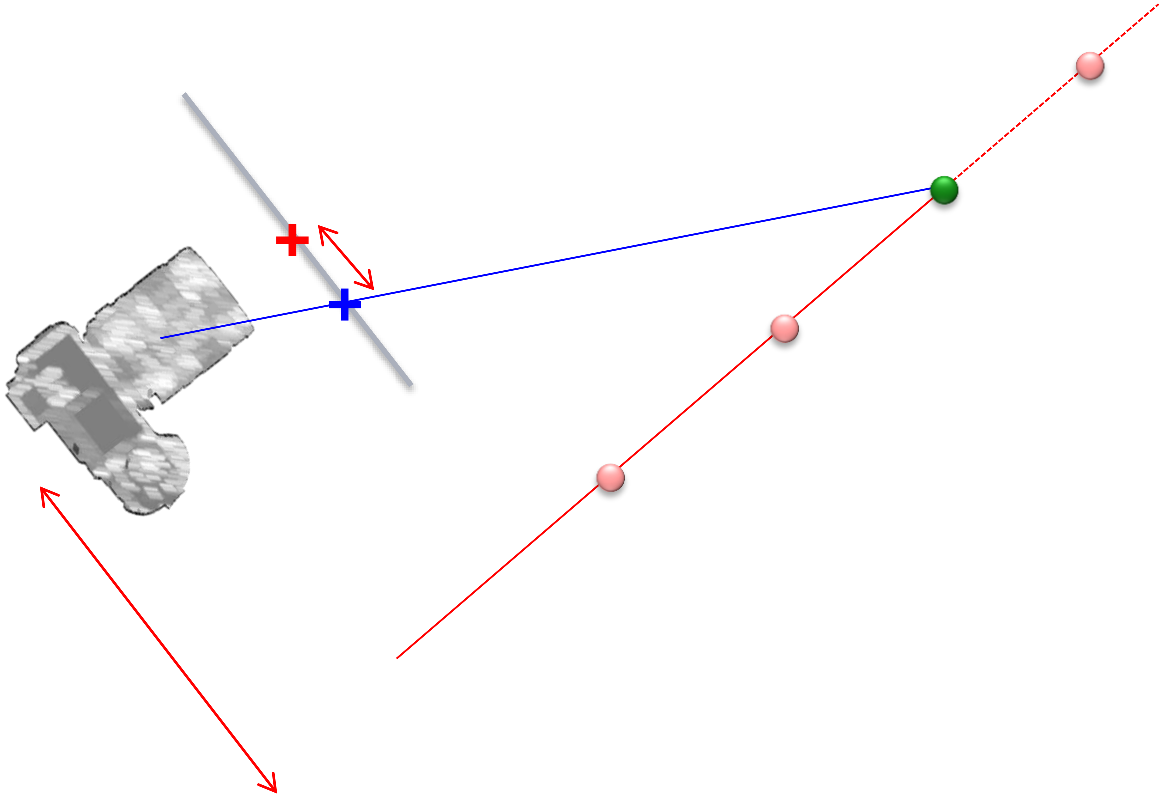
Effects of the camera motion 3/3

- optimal motion for a point – more complex in general



Effects of the camera motion 3/3

- optimal motion for a point – more complex in general



Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \boldsymbol{\chi})\boldsymbol{u}$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\boldsymbol{\chi} \in \mathbb{R}^p$: unmeasurable state component

Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

$$\dot{s} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \boldsymbol{u}$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\chi \in \mathbb{R}^p$: unmeasurable state component

Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

$$\dot{s} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \boldsymbol{u}$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\boldsymbol{\chi} \in \mathbb{R}^p$: unmeasurable state component

Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson ‘06]

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \chi)\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s}, \chi?)^{\dagger}(\boldsymbol{s} - \boldsymbol{s}^*)$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\chi \in \mathbb{R}^p$: unmeasurable state component
- poor approximation of χ can affect stability/performance [Malis et al., TRO 2010]. Possible solutions:

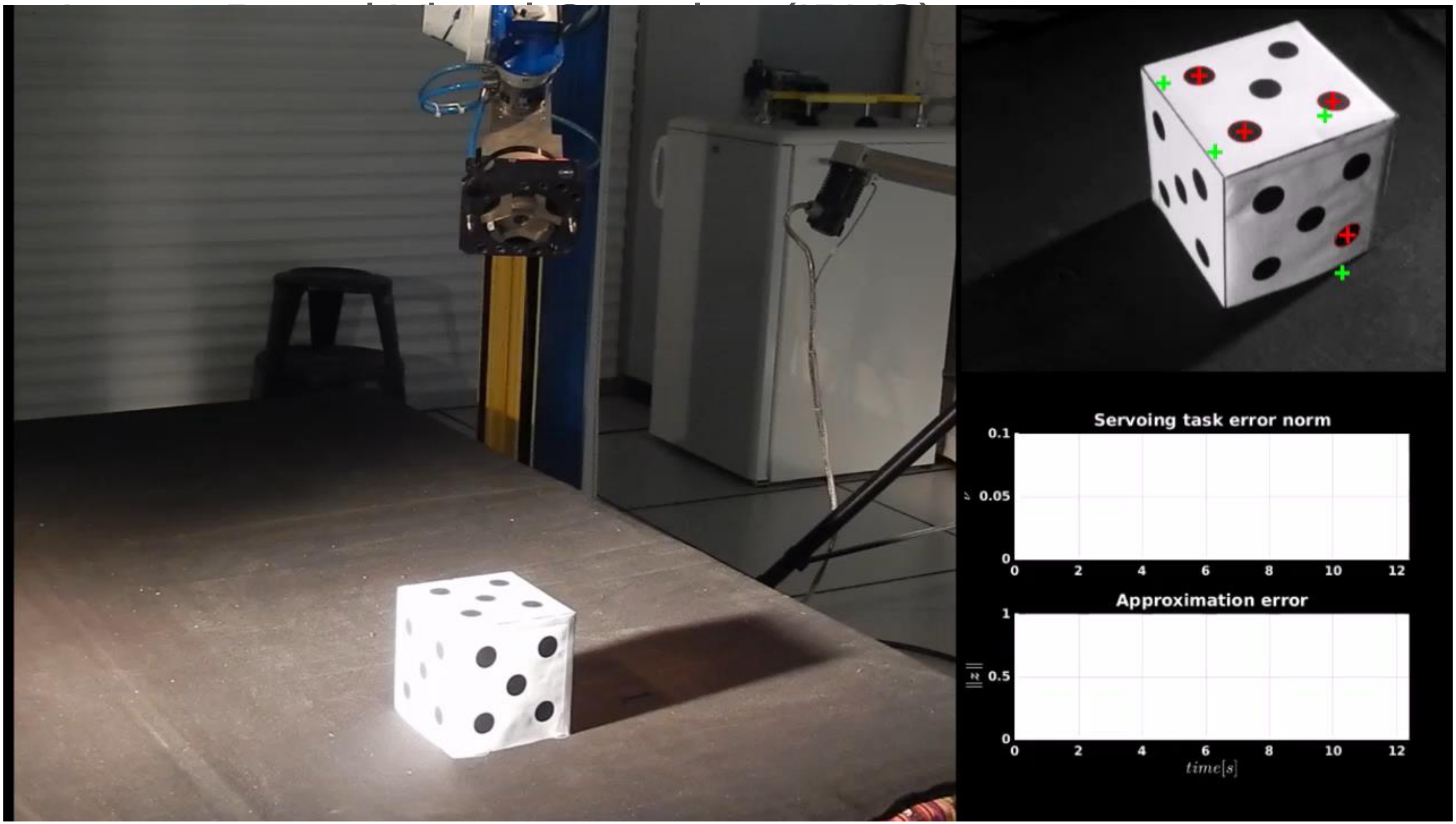
Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \boldsymbol{\chi})\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s}, \boldsymbol{\chi}^*)^\dagger(\boldsymbol{s} - \boldsymbol{s}^*)$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\boldsymbol{\chi} \in \mathbb{R}^p$: unmeasurable state component
- poor approximation of $\boldsymbol{\chi}$ can affect stability/performance [Malis et al., TRO 2010]. Possible solutions:
  - use the value $\boldsymbol{\chi}^*$ at desired pose (if known) [Espiau et al. 1992]

Visual control



Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \chi)\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s}, \chi^*)^\dagger(\boldsymbol{s} - \boldsymbol{s}^*)$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\chi \in \mathbb{R}^p$: unmeasurable state component
- poor approximation of χ can affect stability/performance [Malis et al., TRO 2010]. Possible solutions:
  - use the value χ^* at desired pose (if known) [Espiau et al. 1992]

Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \chi)\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s}, \hat{\chi})^{\dagger}(\boldsymbol{s} - \boldsymbol{s}^*)$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\chi \in \mathbb{R}^p$: unmeasurable state component
- poor approximation of χ can affect stability/performance [Malis et al., TRO 2010]. Possible solutions:
  - use the value χ^* at desired pose (if known) [Espiau et al. 1992]
  - estimate $\widehat{\chi}$: from known model or structure from motion (SfM) [Papanikolopoulos, Khosla '93; De Luca et al. '08; Mahony, Stramigioli '08]

Visual control

- Image Based Visual Servoing (IBVS) [Chaumette, Hutchinson '06]

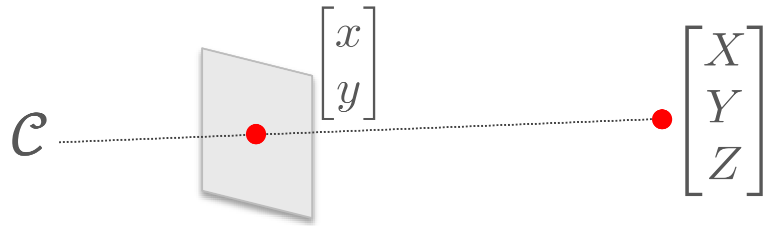
$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \chi)\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s}, \hat{\chi})^{\dagger}(\boldsymbol{s} - \boldsymbol{s}^*)$$

  - $\boldsymbol{u} = (\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^6$: camera (controllable) velocity in camera frame
  - $\boldsymbol{s} \in \mathbb{R}^m$: measurable feature vector
  - $\chi \in \mathbb{R}^p$: unmeasurable state component
- poor approximation of χ can affect stability/performance [Malis et al., TRO 2010]. Possible solutions:
  - use the value χ^* at desired pose (if known) [Espiau et al. 1992]
  - estimate $\widehat{\chi}$: from known model or structure from motion (SfM) [Papanikolopoulos, Khosla '93; De Luca et al. '08; Mahony, Stramigioli '08]
- SfM is non-linear $\longrightarrow$ observability depends on system inputs, i.e. on camera trajectory, and can be optimized

Active Structure from Controlled Motion

Problem statement

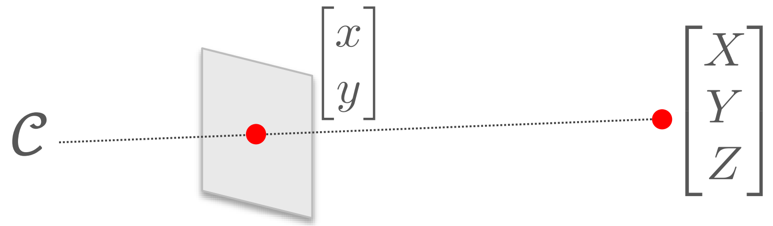
- use controlled monocular camera to reconstruct the structure of some basic 3D primitives in camera frame, e.g.



$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix} \rightarrow \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -\left(1+x^2\right) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \boldsymbol{u}$$

Problem statement

- use controlled monocular camera to reconstruct the structure of some basic 3D primitives in camera frame, e.g.

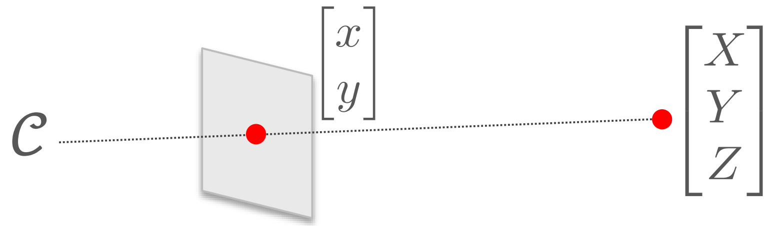


$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix} \rightarrow \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \boldsymbol{u}$$

  - calibrated camera (normalized coordinates)

Problem statement

- use controlled monocular camera to reconstruct the structure of some basic 3D primitives in camera frame, e.g.

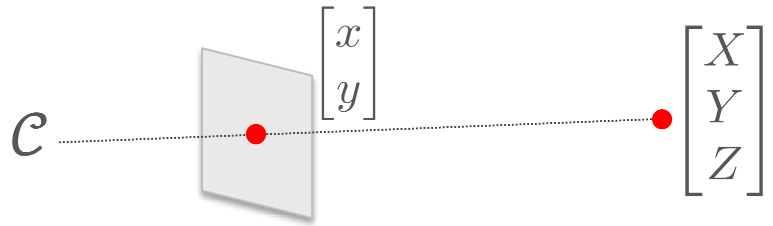


$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix} \rightarrow \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \boldsymbol{u}$$

  - calibrated camera (normalized coordinates)
  - known and controllable camera motion (in camera/body frame)

Problem statement

- use controlled monocular camera to reconstruct the structure of some basic 3D primitives in camera frame, e.g.

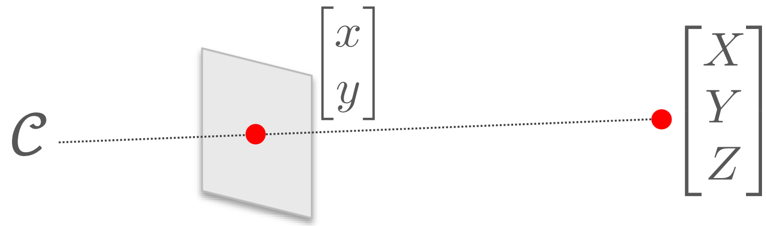


$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix} \rightarrow \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -\left(1+x^2\right) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \boldsymbol{u}$$

  - calibrated camera (normalized coordinates)
  - known and controllable camera motion (in camera/body frame)
  - best possible accuracy/convergence time for some $\max \|\boldsymbol{u}\|$

Basic idea

- dynamics of a point feature $[x\ y] = [X/Z\ Y/Z]$

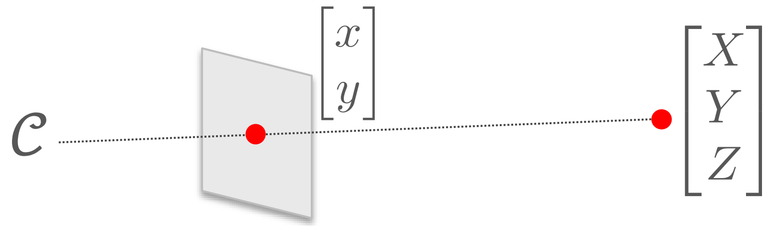


$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} xy & -\left(1 + x^2\right) & y \\ 1 + y^2 & -xy & -x \end{bmatrix} \boldsymbol{\omega} + \begin{bmatrix} xv_z - v_x \\ yv_z - v_y \end{bmatrix} \frac{1}{Z}$$

$\boldsymbol{\Omega}(x, y, \boldsymbol{v}) \approx$ sensitivity

Basic idea

- dynamics of a point feature $[x \ y] = [X/Z \ Y/Z]$



$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} xy & -\left(1 + x^2\right) & y \\ 1 + y^2 & -xy & -x \end{bmatrix} \boldsymbol{\omega} + \begin{bmatrix} x v_z - v_x \\ y v_z - v_y \end{bmatrix} \frac{1}{Z}$$

$\boldsymbol{\Omega}(x, y, \boldsymbol{v}) \approx$ sensitivity

- to maximize the convergence rate of an estimate of Z we must choose v to maximize (some norm of) $\boldsymbol{\Omega}$, e.g.

$$\boxed{\dot{\boldsymbol{v}} \propto \nabla_{\boldsymbol{v}} \|\boldsymbol{\Omega}\|}$$

A nonlinear observer for SfM

- $\boldsymbol{s} \in \mathbb{R}^m$, measurable state component (e.g. $\boldsymbol{s} = [x, y]^T$)
- $\boldsymbol{\chi} \in \mathbb{R}^p$, unmeasurable component (e.g. $\chi = 1/Z$)
- $\boldsymbol{u} = \left[\boldsymbol{v}^T \; \boldsymbol{\omega}^T\right]^T \in \mathbb{R}^6$ controllable input camera velocity
- $\boldsymbol{\Omega} \in \mathbb{R}^{p \times m}$, $\boldsymbol{f}_{\boldsymbol{s}} \in \mathbb{R}^m$, $\boldsymbol{f}_{\boldsymbol{\chi}} \in \mathbb{R}^p$ generic time-varying but known

$$\begin{cases} \dot{\boldsymbol{s}} = \boldsymbol{f}_{\boldsymbol{s}}(\boldsymbol{s}, \boldsymbol{\omega}) + \boldsymbol{\Omega}^T(\boldsymbol{s}, \boldsymbol{v})\boldsymbol{\chi} \\ \dot{\boldsymbol{\chi}} = \boldsymbol{f}_{\boldsymbol{\chi}}(\boldsymbol{s}, \boldsymbol{\chi}, \boldsymbol{u}) \end{cases}$$

A. De Luca, G. Oriolo, P. Robuffo Giordano. Feature Depth Observation for Image-based Visual Servoing: Theory and Experiments. The International Journal of Robotics Research, 27(10):1093-1116, October 2008.

A nonlinear observer for SfM

- $\boldsymbol{s} \in \mathbb{R}^m$, measurable state component (e.g. $\boldsymbol{s} = [x, y]^T$)
- $\chi \in \mathbb{R}^p$, unmeasurable component (e.g. $\chi = 1/Z$)
- $\boldsymbol{u} = \left[\boldsymbol{v}^T \; \boldsymbol{\omega}^T\right]^T \in \mathbb{R}^6$ controllable input camera velocity
- $\boldsymbol{\Omega} \in \mathbb{R}^{p \times m}$, $\boldsymbol{f}_{\boldsymbol{s}} \in \mathbb{R}^m$, $\boldsymbol{f}_{\boldsymbol{\chi}} \in \mathbb{R}^p$ generic time-varying but known

$$\begin{cases} \dot{\hat{\boldsymbol{s}}} = \boldsymbol{f}_{\boldsymbol{s}}(\boldsymbol{s}, \boldsymbol{\omega}) + \boldsymbol{\Omega}^T(\boldsymbol{s}, \boldsymbol{v})\hat{\boldsymbol{\chi}} \\ \dot{\hat{\boldsymbol{\chi}}} = \boldsymbol{f}_{\boldsymbol{\chi}}(\boldsymbol{s}, \hat{\boldsymbol{\chi}}, \boldsymbol{u}) \end{cases}$$

A. De Luca, G. Oriolo, P. Robuffo Giordano. Feature Depth Observation for Image-based Visual Servoing: Theory and Experiments. The International Journal of Robotics Research, 27(10):1093-1116, October 2008.

A nonlinear observer for SfM

- $\boldsymbol{s} \in \mathbb{R}^m$, measurable state component (e.g. $\boldsymbol{s} = [x, y]^T$)
- $\boldsymbol{\chi} \in \mathbb{R}^p$, unmeasurable component (e.g. $\chi = 1/Z$)
- $\boldsymbol{u} = \left[\boldsymbol{v}^T \ \boldsymbol{\omega}^T\right]^T \in \mathbb{R}^6$ controllable input camera velocity
- $\boldsymbol{\Omega} \in \mathbb{R}^{p \times m}$, $\boldsymbol{f}_{\boldsymbol{s}} \in \mathbb{R}^m$, $\boldsymbol{f}_{\boldsymbol{\chi}} \in \mathbb{R}^p$ generic time-varying but known

$$\begin{cases} \dot{\widehat{\boldsymbol{s}}} = \boldsymbol{f}_{\boldsymbol{s}}(\boldsymbol{s}, \boldsymbol{\omega}) + \boldsymbol{\Omega}^T(\boldsymbol{s}, \boldsymbol{v})\widehat{\boldsymbol{\chi}} - \boldsymbol{H}\widetilde{\boldsymbol{s}} \\ \dot{\widehat{\boldsymbol{\chi}}} = \boldsymbol{f}_{\boldsymbol{\chi}}(\boldsymbol{s}, \widehat{\boldsymbol{\chi}}, \boldsymbol{u}) - \alpha\boldsymbol{\Omega}(\boldsymbol{s}, \boldsymbol{v})\widetilde{\boldsymbol{s}} \end{cases}$$

$\widetilde{\boldsymbol{s}} = \widehat{\boldsymbol{s}} - \boldsymbol{s}$, $\widetilde{\boldsymbol{\chi}} = \widehat{\boldsymbol{\chi}} - \boldsymbol{\chi}$ estimation errors ($\widetilde{\mathbf{x}} = [\widetilde{\boldsymbol{s}}^T, \widetilde{\boldsymbol{\chi}}^T]^T \in \mathbb{R}^{m+p}$)

$\boldsymbol{H}$, α free gains

A. De Luca, G. Oriolo, P. Robuffo Giordano. Feature Depth Observation for Image-based Visual Servoing: Theory and Experiments. The International Journal of Robotics Research, 27(10):1093-1116, October 2008.

A nonlinear observer for SfM

- $\boldsymbol{s} \in \mathbb{R}^m$, measurable state component (e.g. $\boldsymbol{s} = [x, y]^T$)
- $\boldsymbol{\chi} \in \mathbb{R}^p$, unmeasurable component (e.g. $\chi = 1/Z$)
- $\boldsymbol{u} = \left[\boldsymbol{v}^T \ \boldsymbol{\omega}^T\right]^T \in \mathbb{R}^6$ controllable input camera velocity
- $\boldsymbol{\Omega} \in \mathbb{R}^{p \times m}$, $\boldsymbol{f}_{\boldsymbol{s}} \in \mathbb{R}^m$, $\boldsymbol{f}_{\boldsymbol{\chi}} \in \mathbb{R}^p$ generic time-varying but known

$$\begin{cases} \dot{\widehat{\boldsymbol{s}}} = \boldsymbol{f}_{\boldsymbol{s}}(\boldsymbol{s}, \boldsymbol{\omega}) + \boldsymbol{\Omega}^T(\boldsymbol{s}, \boldsymbol{v})\widehat{\boldsymbol{\chi}} - \boldsymbol{H}\widetilde{\boldsymbol{s}} \\ \dot{\widehat{\boldsymbol{\chi}}} = \boldsymbol{f}_{\boldsymbol{\chi}}(\boldsymbol{s}, \widehat{\boldsymbol{\chi}}, \boldsymbol{u}) - \alpha\boldsymbol{\Omega}(\boldsymbol{s}, \boldsymbol{v})\widetilde{\boldsymbol{s}} \end{cases} \Rightarrow \begin{cases} \dot{\widetilde{\boldsymbol{s}}} = -\boldsymbol{H}\widetilde{\boldsymbol{s}} + \boldsymbol{\Omega}^T(\boldsymbol{s}, \boldsymbol{v})\widetilde{\boldsymbol{\chi}} \\ \dot{\widetilde{\boldsymbol{\chi}}} = -\alpha\boldsymbol{\Omega}(\boldsymbol{s}, \boldsymbol{v})\widetilde{\boldsymbol{s}} + \boldsymbol{d}(\widetilde{\mathbf{x}}, \boldsymbol{u}) \end{cases}$$

$\widetilde{\boldsymbol{s}} = \widehat{\boldsymbol{s}} - \boldsymbol{s}$, $\widetilde{\boldsymbol{\chi}} = \widehat{\boldsymbol{\chi}} - \boldsymbol{\chi}$ estimation errors ($\widetilde{\mathbf{x}} = [\widetilde{\boldsymbol{s}}^T, \widetilde{\boldsymbol{\chi}}^T]^T \in \mathbb{R}^{m+p}$)

$\boldsymbol{H}$, α free gains

$\boldsymbol{d}(\widetilde{\mathbf{x}}, \boldsymbol{u}) = \boldsymbol{f}_{\boldsymbol{\chi}}\big|_{\boldsymbol{\chi}} - \boldsymbol{f}_{\boldsymbol{\chi}}\big|_{\widehat{\boldsymbol{\chi}}}$ vanishing disturbance ($\boldsymbol{d} \to 0$ as $\widetilde{\mathbf{x}} \to 0$)
with $\|\boldsymbol{d}\| \propto \|\boldsymbol{u}\|$

A. De Luca, G. Oriolo, P. Robuffo Giordano. Feature Depth Observation for Image-based Visual Servoing: Theory and Experiments. The International Journal of Robotics Research, 27(10):1093-1116, October 2008.

Persistence of excitation

- PE lemma (~ observability) [Marino, Tomei, 1995]:

$$\widetilde{\boldsymbol{\chi}} \rightarrow 0 \iff \int_{t}^{t+T} \boldsymbol{\Omega}(\tau)\, \boldsymbol{\Omega}^{T}(\tau)\, d\tau \geq \gamma \boldsymbol{I}_{p} > 0, \quad \forall t \geq t_0$$

Persistence of excitation

- PE lemma (~ observability) [Marino, Tomei, 1995]:

$$\widetilde{\chi} \to 0 \iff \int_t^{t+T} \boldsymbol{\Omega}(\tau)\,\boldsymbol{\Omega}^T(\tau)\,d\tau \geq \gamma \boldsymbol{I}_p > 0, \quad \forall t \geq t_0$$

  - convergence is exponential and semi-global w.r.t.

    $\dot{\chi} = 0 \Rightarrow \boldsymbol{d} = 0$ and convergence is global

Persistence of excitation

- PE lemma (~ observability) [Marino, Tomei, 1995]:

$$\widetilde{\chi} \to 0 \iff \int_{t}^{t+T} \mathbf{\Omega}(\tau)\,\mathbf{\Omega}^T(\tau)\,d\tau \geq \gamma \boldsymbol{I}_p > 0, \quad \forall t \geq t_0$$

  - convergence is exponential and semi-global w.r.t.

    $\dot{\chi} = 0 \Rightarrow \boldsymbol{d} = 0$ and convergence is global
  - if $m \geq p$ (more measures than unknowns) a sufficient condition is

$$\mathbf{\Omega}(t)\,\mathbf{\Omega}^T(t) \geq \frac{\gamma}{T}\boldsymbol{I}_p > 0, \quad \forall t \geq t_0$$

Persistence of excitation

- PE lemma (~ observability) [Marino, Tomei, 1995]:

$$\tilde{\chi} \to 0 \iff \int_t^{t+T} \boldsymbol{\Omega}(\tau)\,\boldsymbol{\Omega}^T(\tau)\,d\tau \geq \gamma \boldsymbol{I}_p > 0, \quad \forall t \geq t_0$$

  - convergence is exponential and semi-global w.r.t. $\|\boldsymbol{u}\|$

    $\dot{\chi} = 0 \Rightarrow \boldsymbol{d} = 0$ and convergence is global
  - if $m \geq p$ (more measures than unknowns) a sufficient condition is

$$\boldsymbol{\Omega}(t)\,\boldsymbol{\Omega}^T(t) \geq \frac{\gamma}{T}\boldsymbol{I}_p > 0, \quad \forall t \geq t_0$$

- since, $\boldsymbol{\Omega}(t) = \boldsymbol{\Omega}(\boldsymbol{s}, \boldsymbol{v})$ we can “optimize” the behavior by
  - controlling camera motion ($\boldsymbol{v}$) [Spica, Robuffo Giordano, CDC 2013].
  - selecting the set of measurements ($\boldsymbol{s}$) [Robuffo Giordano, Spica, Chaumette, ICRA 2015].
  - independently of the estimator (~ Fisher matrix and Gramian)

Estimation error dynamics assignment

$$\boldsymbol{\Omega} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$

$$\boldsymbol{H} = \boldsymbol{V}\begin{bmatrix}\boldsymbol{D}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{D}_2\end{bmatrix}\boldsymbol{V}^T$$

$$\boldsymbol{\epsilon} = \frac{1}{\alpha}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T\widetilde{\boldsymbol{\chi}}$$

$\sim$

$$\boxed{\ddot{\boldsymbol{\epsilon}}} = \boxed{(\check{\boldsymbol{\Pi}} - \boldsymbol{D}_1)\dot{\boldsymbol{\epsilon}}} - \boxed{\alpha\boldsymbol{\Sigma}^2\boldsymbol{\epsilon}}$$

~mass ~damper ~spring

- eigenvalues $\alpha\sigma_i^2$ of $\alpha\boldsymbol{\Omega}\boldsymbol{\Omega}^T$ determine the convergence rate

Estimation error dynamics assignment

$$\boldsymbol{\Omega} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$
$$\boldsymbol{H} = \boldsymbol{V}\begin{bmatrix}\boldsymbol{D}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_2\end{bmatrix}\boldsymbol{V}^T$$
$$\boldsymbol{\epsilon} = \frac{1}{\alpha}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T\tilde{\boldsymbol{\chi}}$$

$\sim$

$$\boxed{\ddot{\boldsymbol{\epsilon}}} = \boxed{(\check{\boldsymbol{\Pi}} - \boldsymbol{D}_1)\dot{\boldsymbol{\epsilon}}} - \boxed{\alpha\boldsymbol{\Sigma}^2\boldsymbol{\epsilon}}$$

~mass ~damper ~spring

- eigenvalues $\alpha\sigma_i^2$ of $\alpha\boldsymbol{\Omega}\boldsymbol{\Omega}^T$ determine the convergence rate
- in SfM $\boldsymbol{\Omega}(t) = \boldsymbol{\Omega}(\boldsymbol{s}, \boldsymbol{v}),\ \sigma_1^2(t) = \sigma_1^2(\boldsymbol{s}, \boldsymbol{v}) \leq \cdots \leq \sigma_p^2(t) = \sigma_p^2(\boldsymbol{s}, \boldsymbol{v})$

(e.g.) $$\dot{\boldsymbol{v}} = \boxed{\frac{k_1\boldsymbol{v}}{\|\boldsymbol{v}\|^2}(\|\boldsymbol{v}_0\| - \|\boldsymbol{v}\|)} + k_2\boxed{\left(\boldsymbol{I}_3 - \frac{\boldsymbol{v}\boldsymbol{v}^T}{\|\boldsymbol{v}\|^2}\right)}\boxed{\nabla_{\boldsymbol{v}}\sigma_1^2}$$

const. $\|\boldsymbol{v}\|$ null projector max. σ_1^2

Estimation error dynamics assignment

$$\boldsymbol{\Omega} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$
$$\boldsymbol{H} = \boldsymbol{V}\begin{bmatrix}\boldsymbol{D}_1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{D}_2\end{bmatrix}\boldsymbol{V}^T \quad \sim \quad \boxed{\ddot{\boldsymbol{\epsilon}}} = \boxed{(\check{\boldsymbol{\Pi}} - \boldsymbol{D}_1)\dot{\boldsymbol{\epsilon}}} - \boxed{\alpha\boldsymbol{\Sigma}^2\boldsymbol{\epsilon}}$$
$$\boldsymbol{\epsilon} = \frac{1}{\alpha}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T\tilde{\boldsymbol{\chi}}$$

~mass ~damper ~spring

- eigenvalues $\alpha\sigma_i^2$ of $\alpha\boldsymbol{\Omega}\boldsymbol{\Omega}^T$ determine the convergence rate
- in SfM $\boldsymbol{\Omega}(t) = \boldsymbol{\Omega}(\boldsymbol{s}, \boldsymbol{v}),\ \sigma_1^2(t) = \sigma_1^2(\boldsymbol{s}, \boldsymbol{v}) \leq \cdots \leq \sigma_p^2(t) = \sigma_p^2(\boldsymbol{s}, \boldsymbol{v})$

(e.g.) $$\dot{\boldsymbol{v}} = \underbrace{\boxed{\frac{k_1\boldsymbol{v}}{\|\boldsymbol{v}\|^2}(\|\boldsymbol{v}_0\| - \|\boldsymbol{v}\|)}}_{\text{const. } \|\boldsymbol{v}\|} + k_2\underbrace{\boxed{\left(\boldsymbol{I}_3 - \frac{\boldsymbol{v}\boldsymbol{v}^T}{\|\boldsymbol{v}\|^2}\right)}}_{\text{null projector}}\underbrace{\boxed{\nabla_{\boldsymbol{v}}\sigma_1^2}}_{\text{max. } \sigma_1^2}$$

$\nabla_{\boldsymbol{v}}\sigma_1^2$ is known in closed form

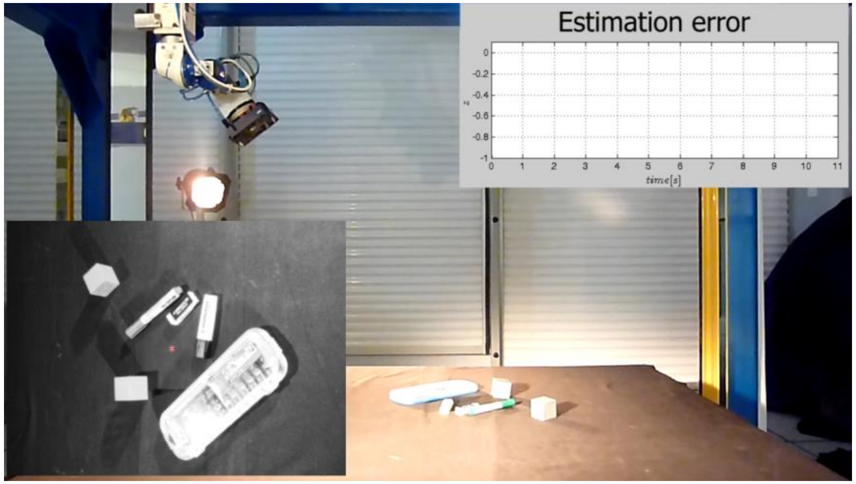
- one must act on $\dot{\boldsymbol{v}}$ (camera linear acceleration)

  $\boldsymbol{\omega}$ is free (can be used, e.g. to maintain visibility of $\boldsymbol{s}$)

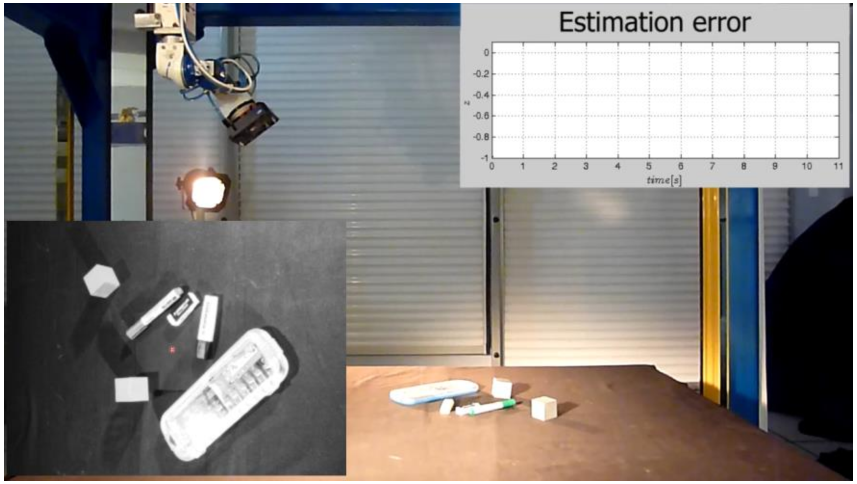
  α trades-off between convergence speed and noise

Experimental results for a point

Active estimation

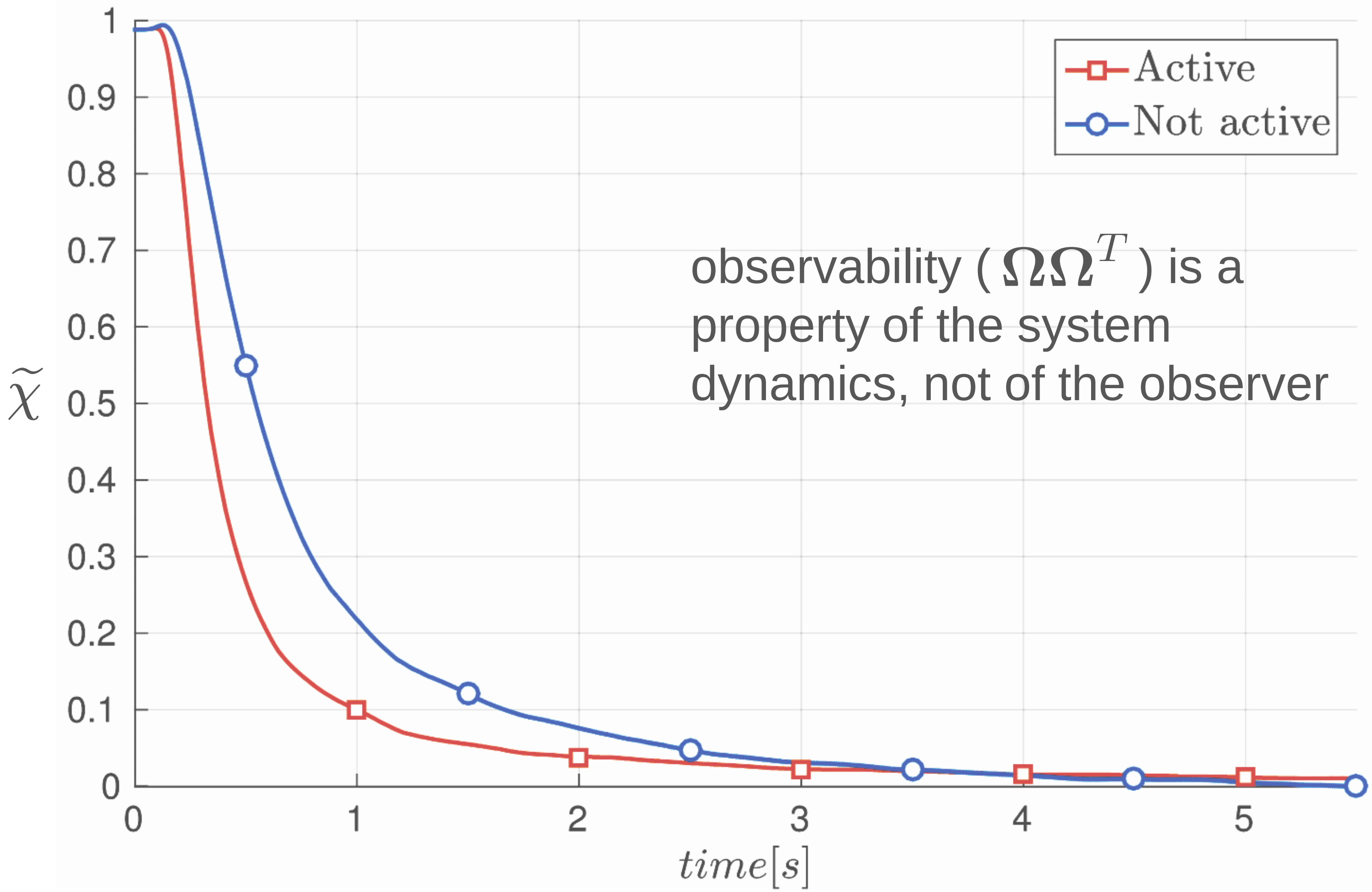


Constant linear velocity



R. Spica, P. Robuffo Giordano, and F. Chaumette, “Active Structure from Motion: Application to Point, Sphere and Cylinder,” IEEE Trans. on Robotics, vol. 30, no. 6, pp. 1499–1513, 2014.

Point depth estimation using a EKF filter

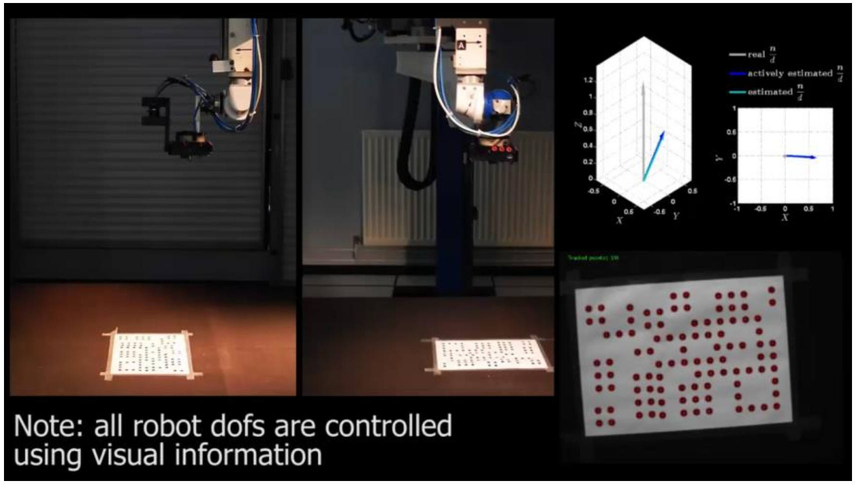


observability ($\Omega\Omega^T$) is a property of the system dynamics, not of the observer

Structure estimation for a plane

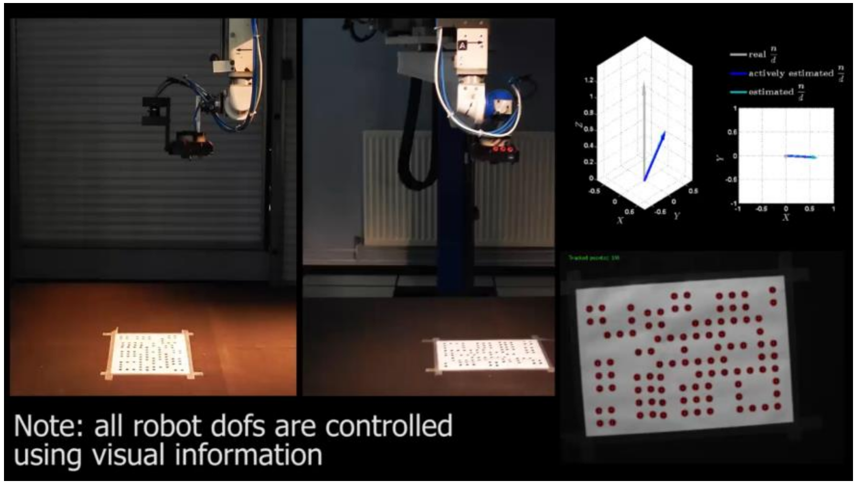
Plane fitting on estimated point cloud (active vs constant v)

$$s = (x_1, y_1, \dots, x_N, y_N)$$
$$\chi = \left(\frac{1}{Z_1}, \dots, \frac{1}{Z_N}\right)$$



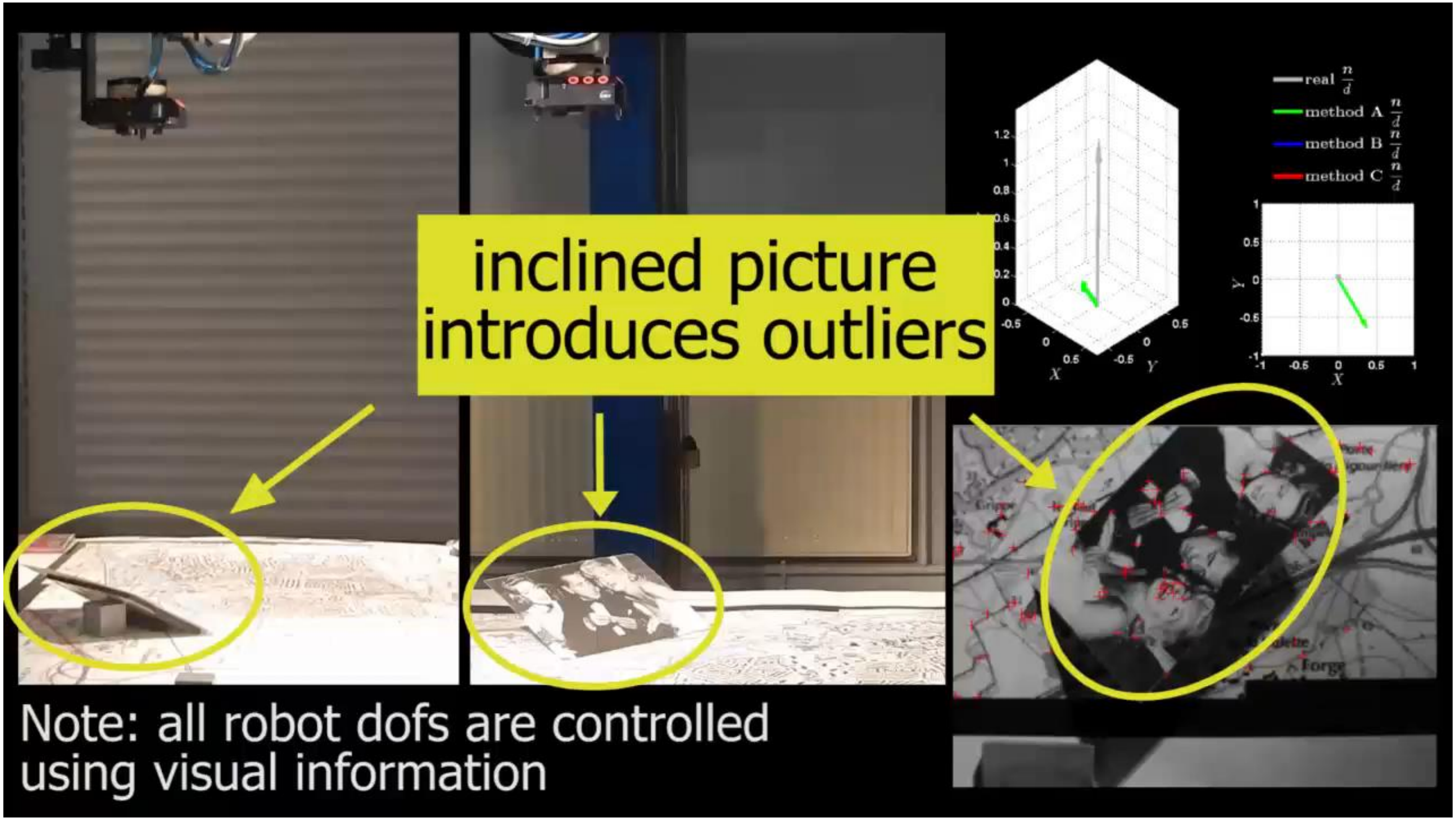
Direct estimation from image moments (active vs constant v)

$$s = (x_g,\ y_g,\ \mu_{20},\ \mu_{02},\ \mu_{11})$$
$$\chi = \frac{\boldsymbol{n}}{d}$$



R. Spica, P. Robuffo Giordano, F. Chaumette, "Plane Estimation by Active Vision from Point Features and Image Moments." in IEEE Int. Conf. on Robotics and Automation, ICRA'15, Seattle, Wa, May. 2015

Plane estimation in presence of outliers



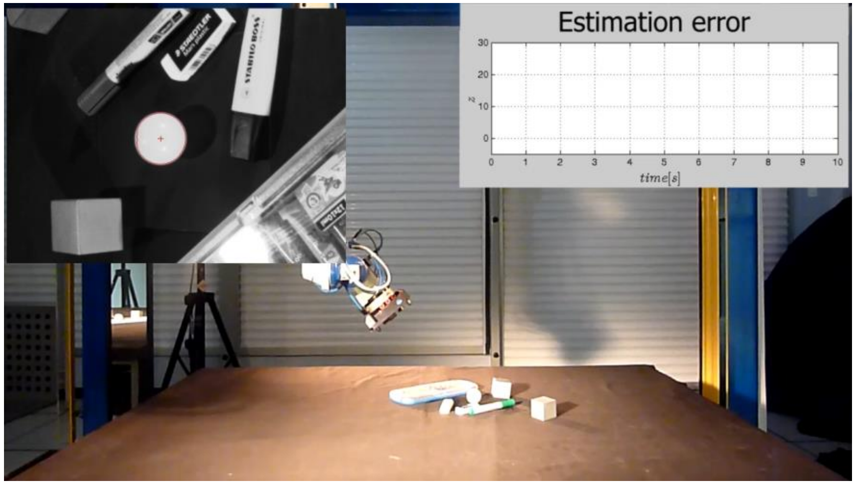
R. Spica, P. Robuffo Giordano, F. Chaumette, "Plane Estimation by Active Vision from Point Features and Image Moments." in IEEE Int. Conf. on Robotics and Automation, ICRA'15, Seattle, Wa, May. 2015

Experimental results for spheres and cylinders

Sphere

$$\boldsymbol{s} = \frac{\boldsymbol{p}_0}{R} \quad \chi = \frac{1}{R}$$

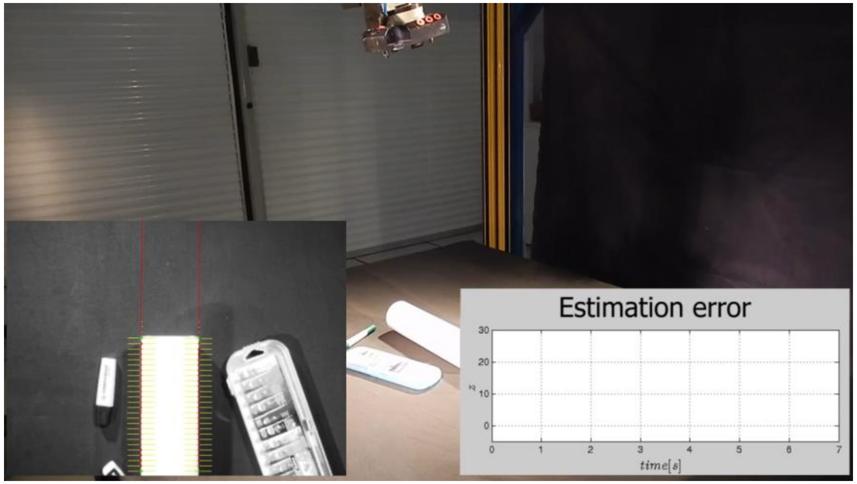
$$\sigma_1^2 = \|\boldsymbol{v}\|^2$$



Cylinder

$$\boldsymbol{s} = \frac{\boldsymbol{p}_0}{R} \quad \chi = \frac{1}{R}$$

$$\sigma_1^2 = \|\boldsymbol{v}\|^2 - (\boldsymbol{a}^T\boldsymbol{v})^2$$



R. Spica, P. Robuffo Giordano, and F. Chaumette, "Active Structure from Motion: Application to Point, Sphere and Cylinder," IEEE Trans. on Robotics, vol. 30, no. 6, pp. 1499–1513, 2014.

Adaptive image moments for plane estimation

- weighted discrete image moments

$$m_w(\boldsymbol{\theta}) = \sum_{k=1}^{N} w(x_k,\, y_k,\, \boldsymbol{\theta}) \overset{e.g.}{=} \sum_{j=1}^{\delta} \sum_{k=0}^{j} \boxed{\theta_{T_j+k}} \boxed{x^{(j-k)} y^k}$$

weights traditional moments

- using $\boldsymbol{s} = (m_w(\boldsymbol{\theta}_1),\, \ldots,\, m_w(\boldsymbol{\theta}_m)) \in \mathbb{R}^m$

$$\boldsymbol{\Omega}(\boldsymbol{\theta},\, \boldsymbol{v},\, t),\;\; \Phi_D(\boldsymbol{\theta},\, \boldsymbol{v},\, t) = \det\left(\boldsymbol{\Omega}\boldsymbol{\Omega}^T\right)$$

Adaptive image moments for plane estimation

- weighted discrete image moments

$$m_w(\boldsymbol{\theta}) = \sum_{k=1}^{N} w(x_k,\, y_k,\, \boldsymbol{\theta}) \overset{e.g.}{=} \sum_{j=1}^{\delta} \sum_{k=0}^{j} \underbrace{\boxed{\theta_{T_j+k}}}_{\text{weights}} \underbrace{\boxed{x^{(j-k)} y^{k}}}_{\text{traditional moments}}$$

- using $\boldsymbol{s} = (m_w(\boldsymbol{\theta}_1),\, \ldots,\, m_w(\boldsymbol{\theta}_m)) \in \mathbb{R}^m$

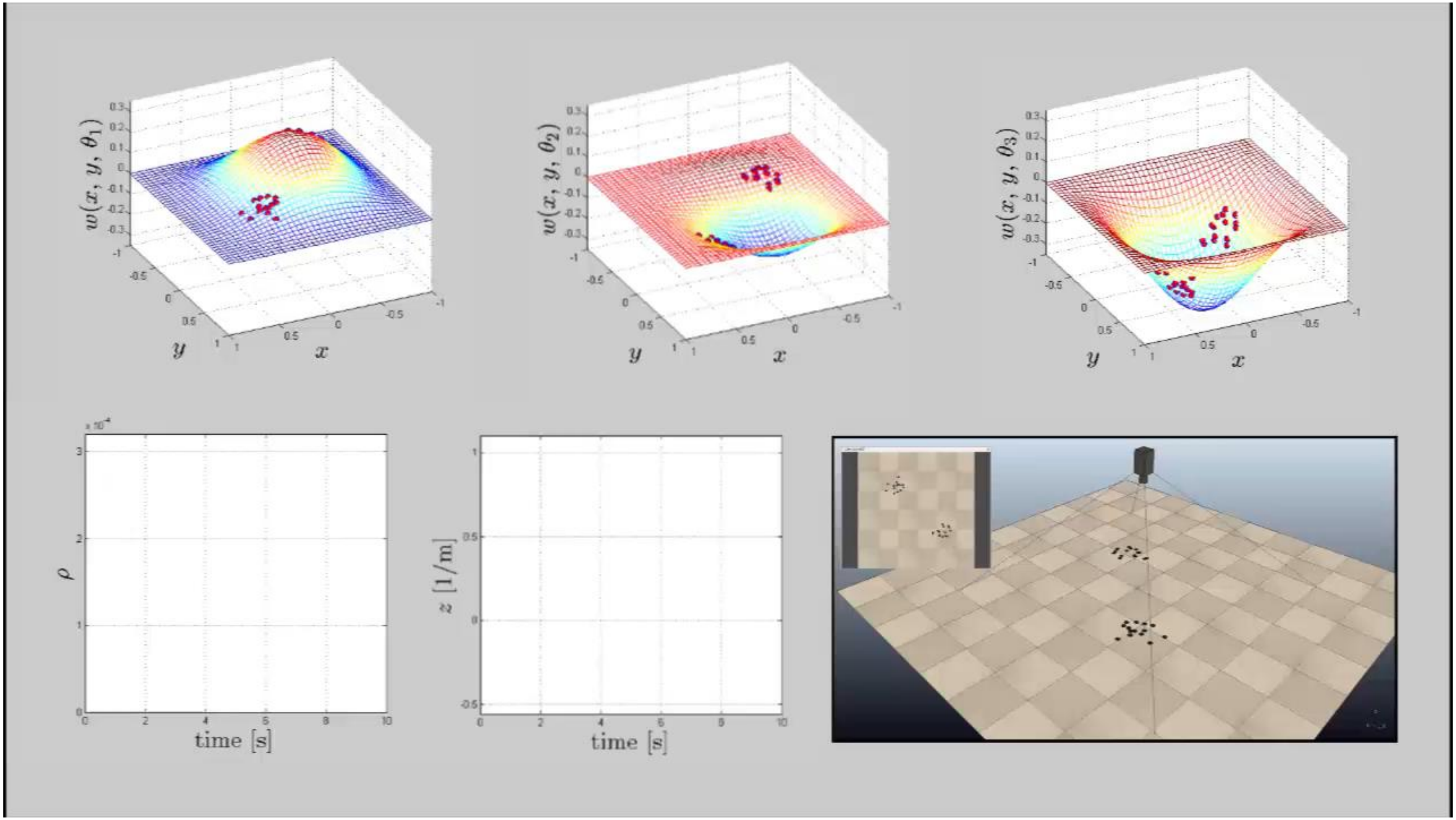
$$\boldsymbol{\Omega}(\boldsymbol{\theta},\, \boldsymbol{v},\, t),\;\; \Phi_D(\boldsymbol{\theta},\, \boldsymbol{v},\, t) = \det\left(\boldsymbol{\Omega}\boldsymbol{\Omega}^T\right)$$

- we can optimize w.r.t. $\boldsymbol{v}$ and/or $\boldsymbol{\theta}$ with

$$\dot{\boldsymbol{v}} = k_v \left(\boldsymbol{I}_3 - \frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}}\right) \nabla_{\boldsymbol{v}} \Phi_D \qquad \dot{\boldsymbol{\theta}} = k_\theta \left(\boldsymbol{I}_s - \frac{\boldsymbol{\theta}\boldsymbol{\theta}^T}{\boldsymbol{\theta}^T\boldsymbol{\theta}}\right) \nabla_{\boldsymbol{\theta}} \Phi_D$$

- additional constraints on $\boldsymbol{\theta}$ to impose $w(x_k,\, y_k,\, \boldsymbol{\theta}) = 0$ at the image borders (limited camera field of view)

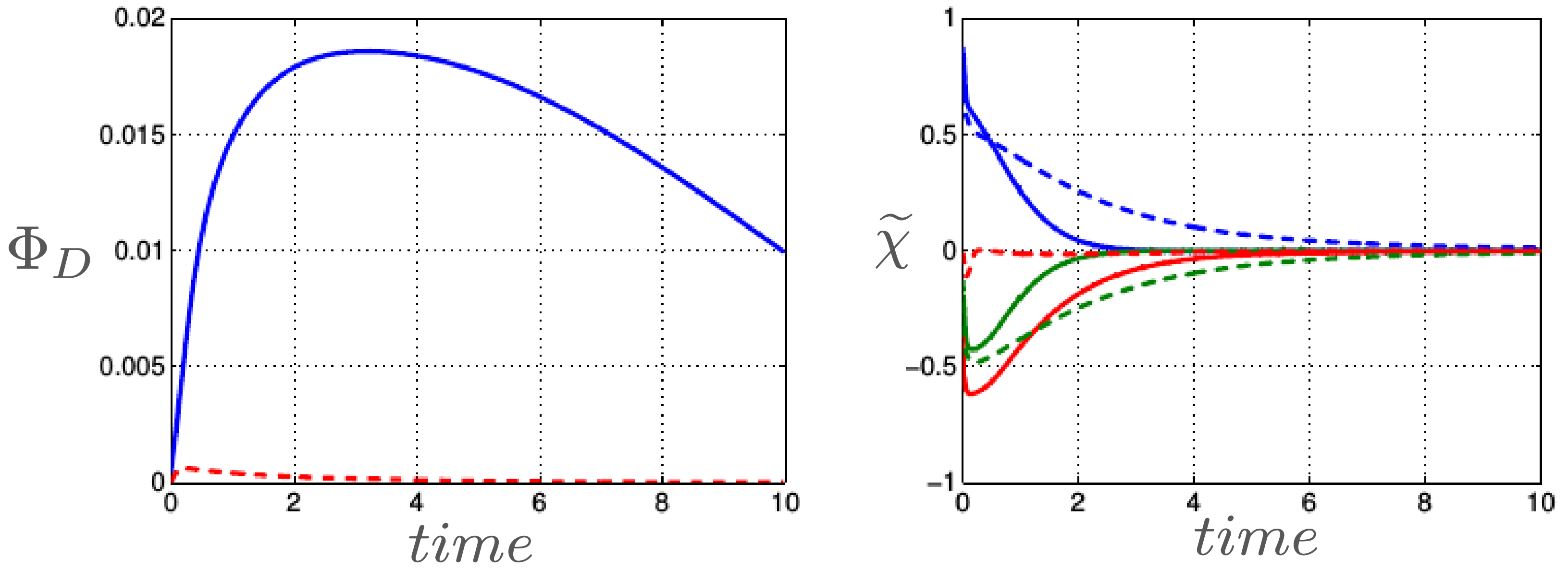
Adaptive image moments – simulation results



P. Robuffo Giordano, R. Spica, F. Chaumette, "Learning the Shape of Image Moments for Optimal 3D Structure Estimation." in IEEE Int. Conf. on Robotics and Automation, ICRA'15, Seattle, Wa, May 2015

Adaptive image moments – simulation results

- solid lines: $\boldsymbol{s} = (m_w(\boldsymbol{\theta}_1), \ldots, m_w(\boldsymbol{\theta}_3)) \in \mathbb{R}^3$
- dashed lines: $\boldsymbol{s} = (x_g, y_g, \mu_{20}, \mu_{02}, \mu_{11}) \in \mathbb{R}^5$



Dense Photometric Structure from Motion

In collaboration with Prof. Robert Mahony

Australian National University

NATURAM PRIMUM COGNOSCERE RERUM

ACRV

Problem statement

- assuming measurements of
  - luminance $Y(t, \boldsymbol{\pi})$ at each time and pixel $\boldsymbol{\pi} = [x, y, 1]^T$
  - camera velocity $[\boldsymbol{v}, \boldsymbol{\omega}]^T$
- estimate the dense depth map $Z(t, \boldsymbol{\pi})$

$Y(t, \boldsymbol{\pi})$

$Z(t, \boldsymbol{\pi})$

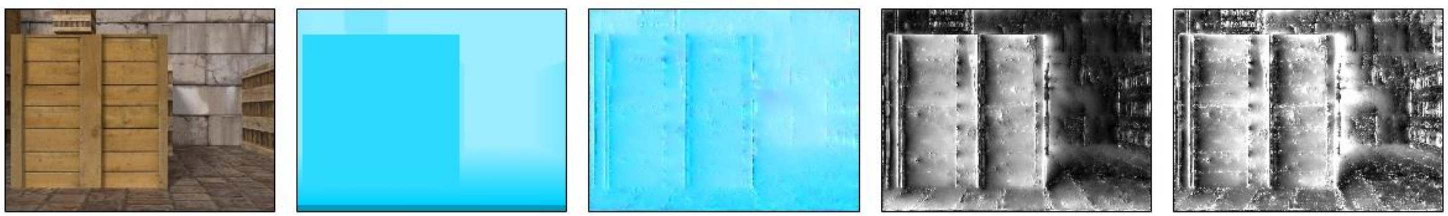
www.mip.informatik.uni-kiel.de

- avoid feature extraction, matching and tracking
- fast enough for typical robot dynamics (~100-200Hz)
- computationally affordable for robot onboard processing

Dense strategies for estimation/control

- visual servoing [Collewet, Marchand, 2011]

- optical flow estimation [Horn, Schunck, 1981; Adarve, Austin, Mahony, 2014]



- dense structure from motion [Matthies, Szelinski, Kanade, 1989; Zarrouati, Aldea, Rouchon, 2012]

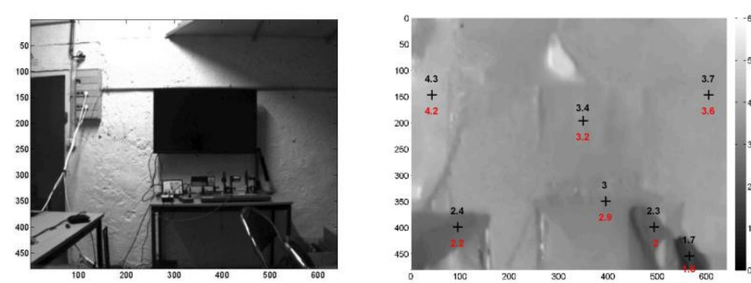


Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t, \boldsymbol{\pi}) = \frac{1}{Z(t, \boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t, \boldsymbol{\pi}) = \frac{1}{Z(t, \boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T (\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t, \boldsymbol{\pi}), \zeta(t, \boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t,\boldsymbol{\pi}) = \frac{1}{Z(t,\boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t,\boldsymbol{\pi}), \zeta(t,\boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

- optical flow $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v}(t),\ \Theta = -[\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times^2 \boldsymbol{\omega}(t)$

$$\dot{\boldsymbol{\pi}}(t,\boldsymbol{\pi}) = \zeta(t,\boldsymbol{\pi})\psi(\boldsymbol{v}(t),\boldsymbol{\pi}) + \Theta(\boldsymbol{\omega}(t),\boldsymbol{\pi})$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t,\boldsymbol{\pi}) = \frac{1}{Z(t,\boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t,\boldsymbol{\pi}), \zeta(t,\boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

- optical flow $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v}(t)$, $\Theta = -[\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times^2 \boldsymbol{\omega}(t)$

$$\dot{\boldsymbol{\pi}}(t,\boldsymbol{\pi}) = \zeta(t,\boldsymbol{\pi})\psi(\boldsymbol{v}(t),\boldsymbol{\pi}) + \Theta(\boldsymbol{\omega}(t),\boldsymbol{\pi})$$

- finally we obtain a system of PDEs

$$\frac{\partial Y}{\partial t} = -\nabla_{\boldsymbol{\pi}} Y^T \Theta - \zeta \nabla_{\boldsymbol{\pi}} Y^T \psi$$

$$\frac{\partial \zeta}{\partial t} = -\nabla_{\boldsymbol{\pi}} \zeta^T \Theta - \zeta \nabla_{\boldsymbol{\pi}} \zeta^T \psi + \zeta \boldsymbol{e}_3^T \left(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega}\right)$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t,\boldsymbol{\pi}) = \frac{1}{Z(t,\boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t,\boldsymbol{\pi}), \zeta(t,\boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

- optical flow $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v}(t)$, $\Theta = -[\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times^2 \boldsymbol{\omega}(t)$

$$\dot{\boldsymbol{\pi}}(t,\boldsymbol{\pi}) = \zeta(t,\boldsymbol{\pi})\psi(\boldsymbol{v}(t),\boldsymbol{\pi}) + \Theta(\boldsymbol{\omega}(t),\boldsymbol{\pi})$$

- finally we obtain a system of PDEs

$$\frac{\partial Y}{\partial t} = -\nabla_{\boldsymbol{\pi}} Y^T \Theta - \zeta \boxed{\nabla_{\boldsymbol{\pi}} Y^T \psi}$$

$\boldsymbol{\Omega}(Y, \boldsymbol{v})$ (sensitivity)

$$\begin{cases} \dot{\boldsymbol{s}} = \boldsymbol{f}_{\boldsymbol{s}}(\boldsymbol{s},\, \boldsymbol{\omega}) + \boldsymbol{\Omega}^T(\boldsymbol{s},\, \boldsymbol{v})\boldsymbol{\chi} \\ \dot{\boldsymbol{\chi}} = \boldsymbol{f}_{\boldsymbol{\chi}}(\boldsymbol{s},\, \boldsymbol{\chi},\, \boldsymbol{u}) \end{cases}$$

$$\frac{\partial \zeta}{\partial t} = -\nabla_{\boldsymbol{\pi}} \zeta^T \Theta - \zeta \nabla_{\boldsymbol{\pi}} \zeta^T \psi + \zeta \boldsymbol{e}_3^T (\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t,\boldsymbol{\pi}) = \frac{1}{Z(t,\boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t,\boldsymbol{\pi}), \zeta(t,\boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

- optical flow $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v}(t)$, $\Theta = -[\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times^2 \boldsymbol{\omega}(t)$

$$\dot{\boldsymbol{\pi}}(t,\boldsymbol{\pi}) = \zeta(t,\boldsymbol{\pi})\psi(\boldsymbol{v}(t),\boldsymbol{\pi}) + \Theta(\boldsymbol{\omega}(t),\boldsymbol{\pi})$$

- finally we obtain a system of PDEs and the observer

$$\frac{\partial \widehat{Y}}{\partial t} = -\nabla_{\boldsymbol{\pi}} Y^T \Theta - \widehat{\zeta} \nabla_{\boldsymbol{\pi}} Y^T \psi$$

$$\frac{\partial \widehat{\zeta}}{\partial t} = -\nabla_{\boldsymbol{\pi}} \widehat{\zeta}^T \Theta - \widehat{\zeta} \nabla_{\boldsymbol{\pi}} \widehat{\zeta}^T \psi + \widehat{\zeta} \boldsymbol{e}_3^T \left(\widehat{\zeta} \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega}\right)$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t,\boldsymbol{\pi}) = \frac{1}{Z(t,\boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T(\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t,\boldsymbol{\pi}), \zeta(t,\boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

- optical flow $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v}(t)$, $\Theta = -[\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times^2 \boldsymbol{\omega}(t)$

$$\dot{\boldsymbol{\pi}}(t,\boldsymbol{\pi}) = \zeta(t,\boldsymbol{\pi})\psi(\boldsymbol{v}(t),\boldsymbol{\pi}) + \Theta(\boldsymbol{\omega}(t),\boldsymbol{\pi})$$

- finally we obtain a system of PDEs and the observer

$$\frac{\partial \widehat{Y}}{\partial t} = -\nabla_{\boldsymbol{\pi}} Y^T \Theta - \widehat{\zeta} \nabla_{\boldsymbol{\pi}} Y^T \psi - h(\widehat{Y} - Y)$$

$$\frac{\partial \widehat{\zeta}}{\partial t} = -\nabla_{\boldsymbol{\pi}} \widehat{\zeta}^T \Theta - \widehat{\zeta} \nabla_{\boldsymbol{\pi}} \widehat{\zeta}^T \psi + \widehat{\zeta} \boldsymbol{e}_3^T \left(\widehat{\zeta} \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega}\right) + \alpha(\nabla_{\boldsymbol{\pi}} Y^T \psi)(\widehat{Y} - Y)$$

Image as a fluid

- assuming constant brightness [Horn, Schunck,1981] and static environment

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = 0 \qquad \zeta(t, \boldsymbol{\pi}) = \frac{1}{Z(t, \boldsymbol{\pi})} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta \boldsymbol{e}_3^T (\zeta \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega})$$

- assuming smooth $Y(t, \boldsymbol{\pi}), \zeta(t, \boldsymbol{\pi})$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} Y^T \dot{\boldsymbol{\pi}} + \frac{\partial Y}{\partial t} \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \nabla_{\boldsymbol{\pi}} \zeta^T \dot{\boldsymbol{\pi}} + \frac{\partial \zeta}{\partial t}$$

- optical flow $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v}(t)$, $\Theta = -[\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times^2 \boldsymbol{\omega}(t)$

$$\dot{\boldsymbol{\pi}}(t, \boldsymbol{\pi}) = \zeta(t, \boldsymbol{\pi}) \psi(\boldsymbol{v}(t), \boldsymbol{\pi}) + \Theta(\boldsymbol{\omega}(t), \boldsymbol{\pi})$$

- finally we obtain a system of PDEs and the observer

$$\frac{\partial \widehat{Y}}{\partial t} = -\nabla_{\boldsymbol{\pi}} Y^T \Theta - \widehat{\zeta} \nabla_{\boldsymbol{\pi}} Y^T \psi - h(\widehat{Y} - Y)$$

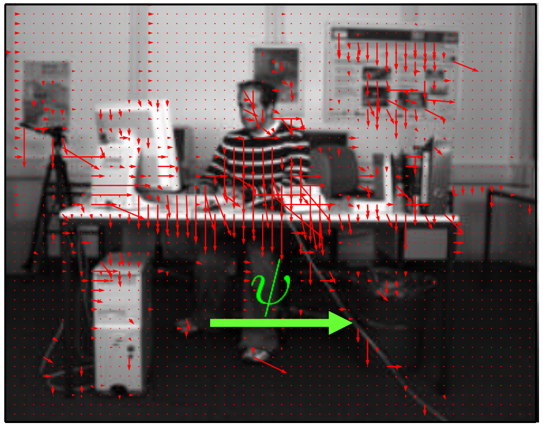
highly parallelizable

$$\frac{\partial \widehat{\zeta}}{\partial t} = -\nabla_{\boldsymbol{\pi}} \widehat{\zeta}^T \Theta - \widehat{\zeta} \nabla_{\boldsymbol{\pi}} \widehat{\zeta}^T \psi + \widehat{\zeta} \boldsymbol{e}_3^T \left(\widehat{\zeta} \boldsymbol{v} - [\boldsymbol{\pi}]_\times \boldsymbol{\omega} \right) + \alpha (\nabla_{\boldsymbol{\pi}} Y^T \psi)(\widehat{Y} - Y)$$

Dense disparity observability

- the estimation $\hat{\zeta}$ will not converge if $\sigma = \left|\nabla_{\boldsymbol{\pi}} Y^T \psi\right| \approx 0$:
  - if the camera does not translate $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v} = 0 \ \forall \boldsymbol{\pi}$
  - in non textured areas $\nabla_{\boldsymbol{\pi}} Y = 0 \ \forall \boldsymbol{v}$
  - in areas where the image moves along contours $\nabla_{\boldsymbol{\pi}} Y^T \psi = 0$

$\nabla_{\boldsymbol{\pi}} Y$



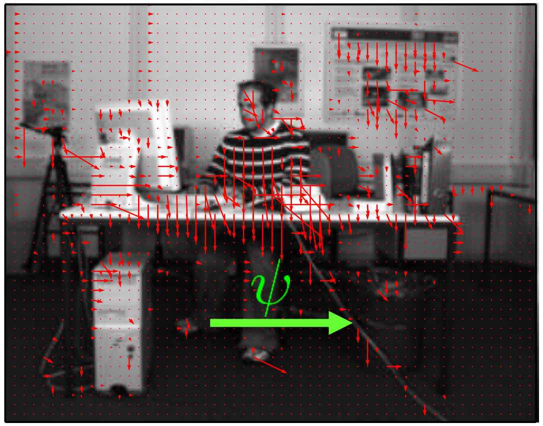
$\|\nabla_{\boldsymbol{\pi}} Y^T\|$

σ

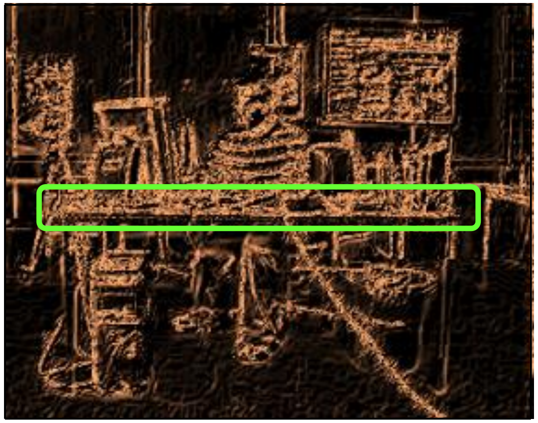
Dense disparity observability

- the estimation $\hat{\zeta}$ will not converge if $\sigma = \left|\nabla_{\boldsymbol{\pi}} Y^T \psi\right| \approx 0$:
  - if the camera does not translate $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v} = 0 \; \forall \boldsymbol{\pi}$
  - in non textured areas $\nabla_{\boldsymbol{\pi}} Y = 0 \; \forall \boldsymbol{v}$
  - in areas where the image moves along contours $\nabla_{\boldsymbol{\pi}} Y^T \psi = 0$

$\nabla_{\boldsymbol{\pi}} Y$



$\|\nabla_{\boldsymbol{\pi}} Y^T\|$

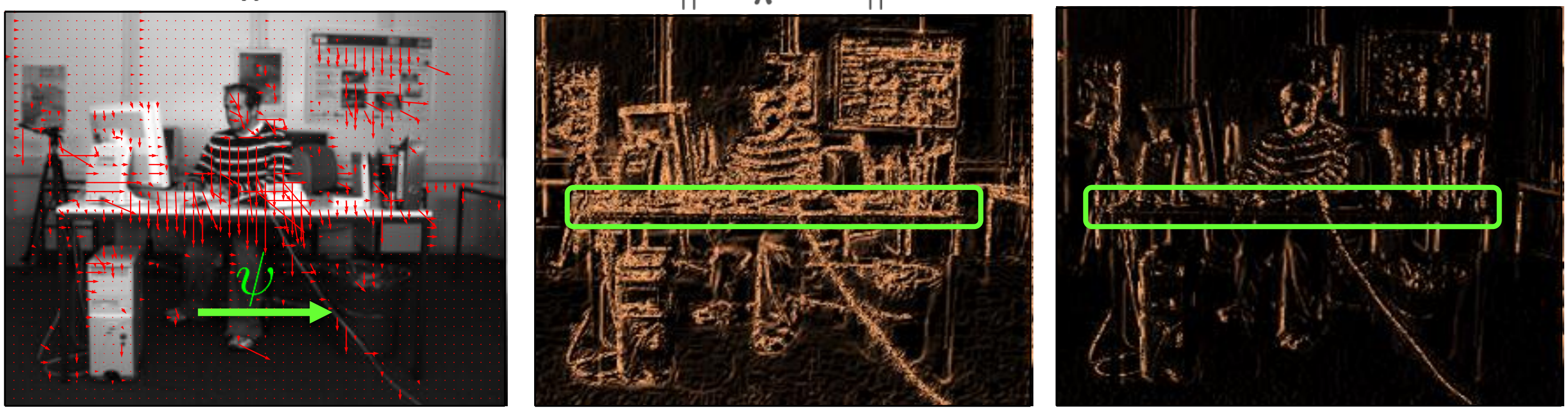


σ

Dense disparity observability

- the estimation $\hat{\zeta}$ will not converge if $\sigma = \left|\nabla_{\boldsymbol{\pi}} Y^T \psi\right| \approx 0$:
  - if the camera does not translate $\psi = [\boldsymbol{e}_3]_\times [\boldsymbol{\pi}]_\times \boldsymbol{v} = 0 \; \forall \boldsymbol{\pi}$
  - in non textured areas $\nabla_{\boldsymbol{\pi}} Y = 0 \; \forall \boldsymbol{v}$
  - in areas where the image moves along contours $\nabla_{\boldsymbol{\pi}} Y^T \psi = 0$

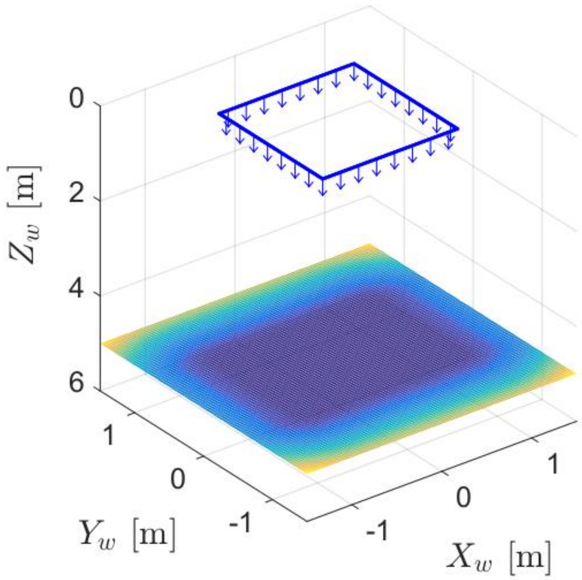
$\nabla_{\boldsymbol{\pi}} Y$ | $\|\nabla_{\boldsymbol{\pi}} Y^T\|$ | σ

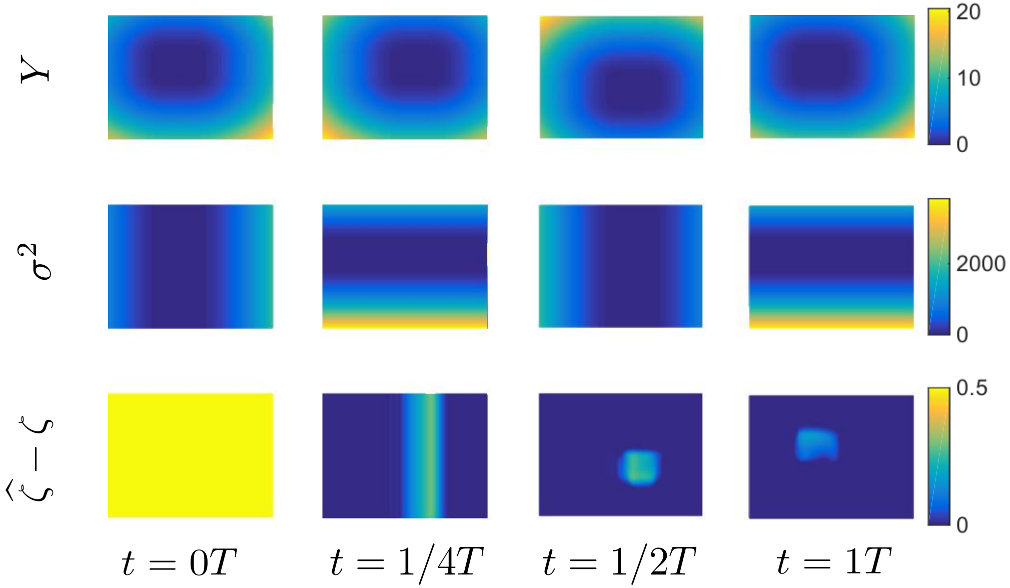


- regularization is necessary in unobservable areas, e.g.

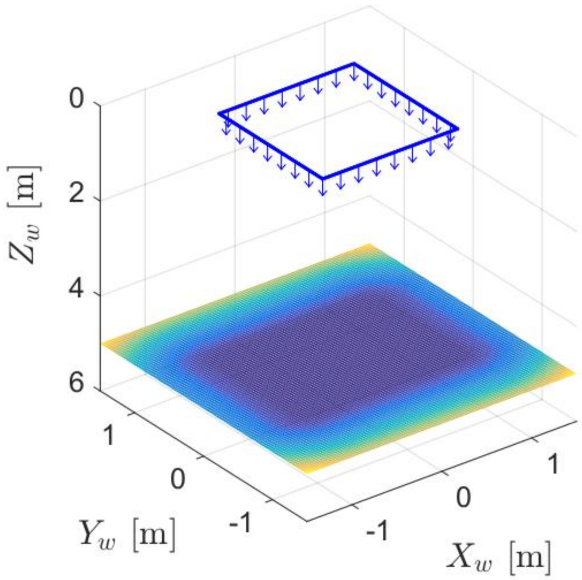
$$\frac{\partial \hat{\zeta}}{\partial t} = (\dots) \boxed{- q(t, \boldsymbol{\pi}) \left(\frac{\partial^2 \hat{\zeta}}{\partial x^2} + \frac{\partial^2 \hat{\zeta}}{\partial y^2} \right)} \qquad q(t, \boldsymbol{\pi}) = k \frac{\sigma_m}{\sigma(t, \boldsymbol{\pi}) + \sigma_m}$$

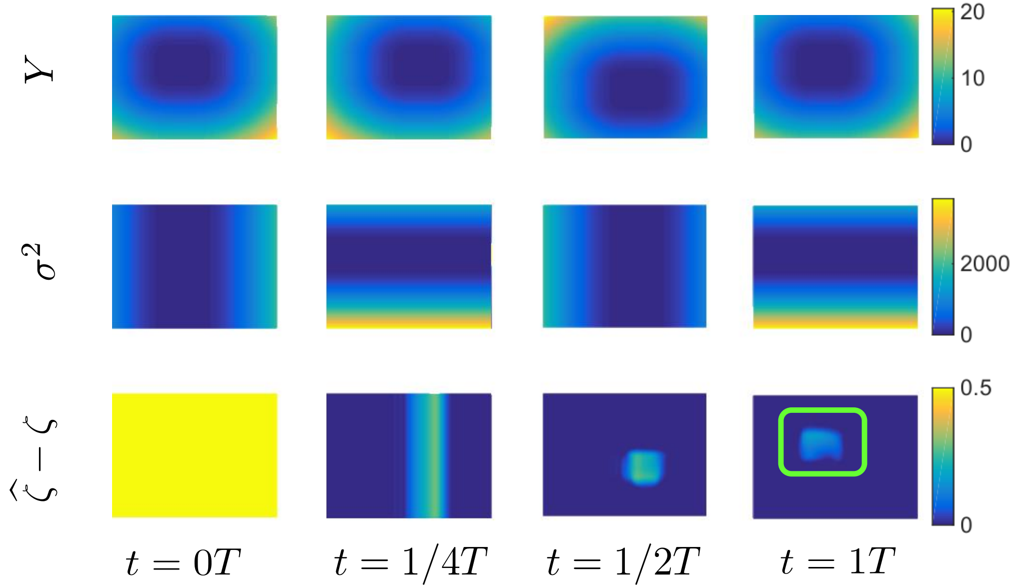
Simulation for planar surface w/o regularization



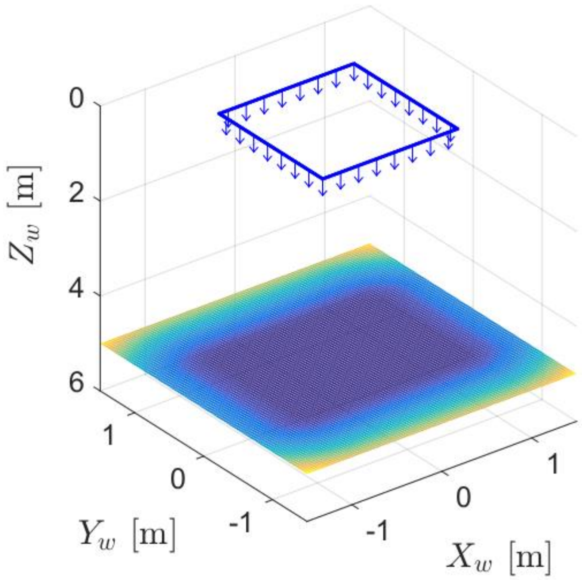


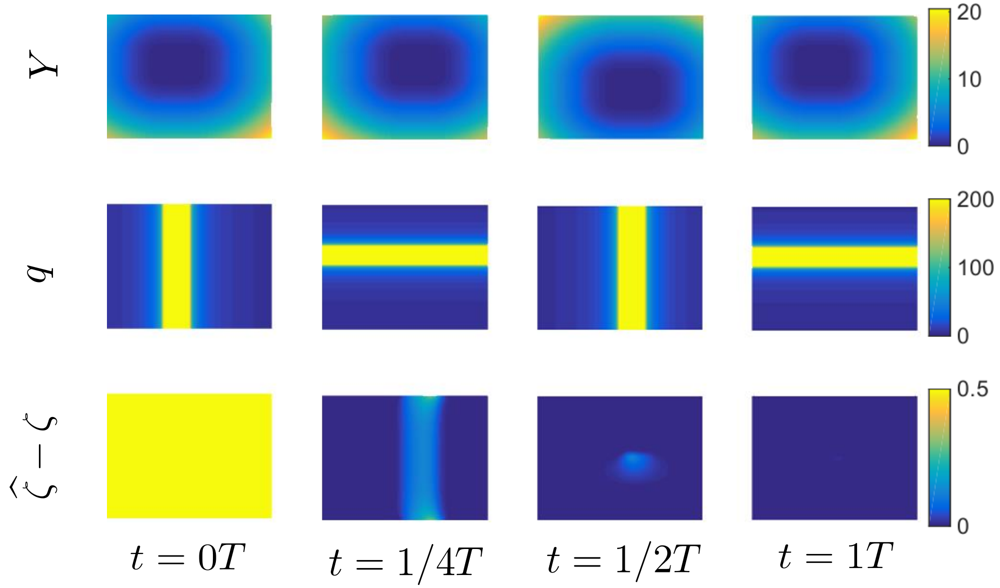
Simulation for planar surface w/o regularization





Simulation for planar surface with regularization





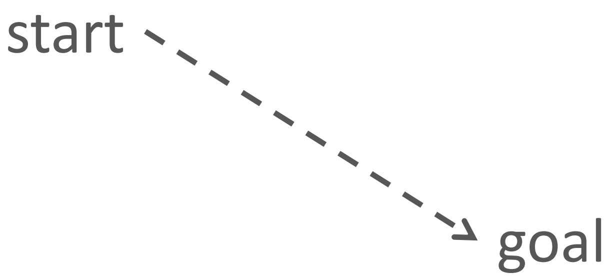
Coupling vision based control and active estimation

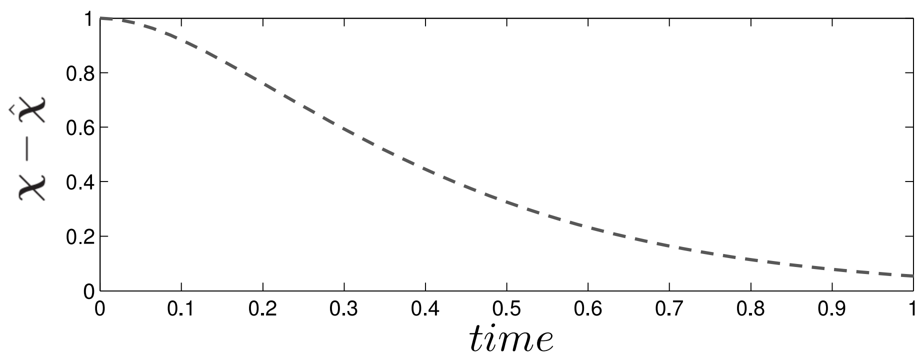
Problem statement

- standard visual servoing control law

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s},\chi)\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s},\hat{\chi})^{\dagger}(\boldsymbol{s}-\boldsymbol{s}^{*})$$

F. Chaumette, S. Hutchinson. Visual servo control, Part I: Basic approaches. RAM, 2006.



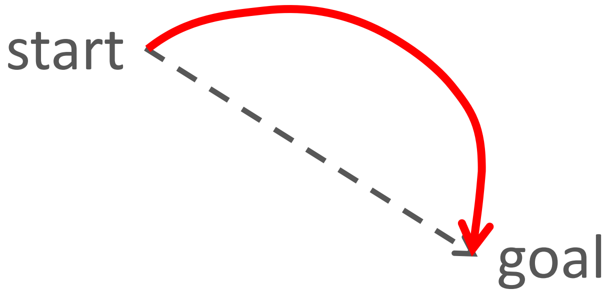


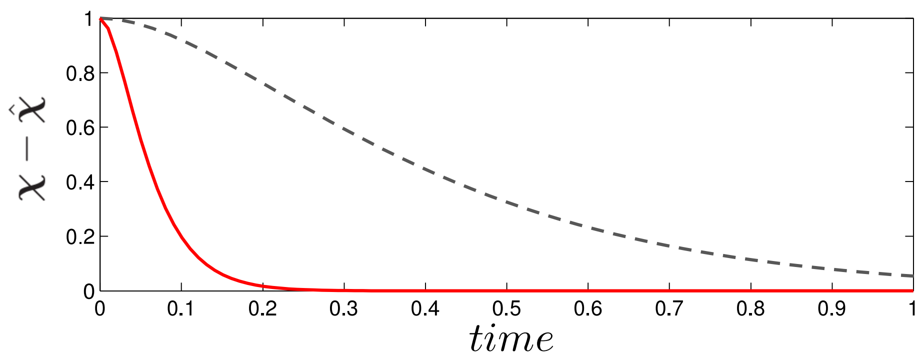
Problem statement

- standard visual servoing control law

$$\dot{\boldsymbol{s}} = \boldsymbol{L}(\boldsymbol{s}, \chi)\boldsymbol{u} \rightarrow \boldsymbol{u} = -k\boldsymbol{L}(\boldsymbol{s}, \hat{\chi})^{\dagger}(\boldsymbol{s} - \boldsymbol{s}^{*})$$

F. Chaumette, S. Hutchinson. Visual servo control, Part I: Basic approaches. RAM, 2006.





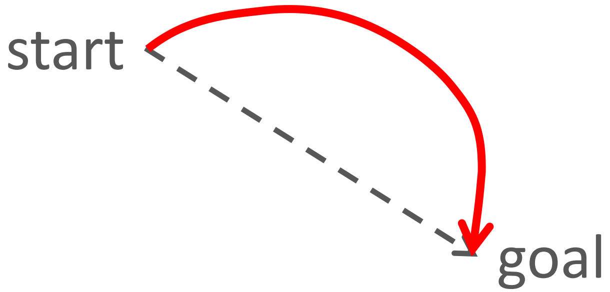
- act on the camera motion during the servoing transient so as to ‘optimally’ estimate the scene structure

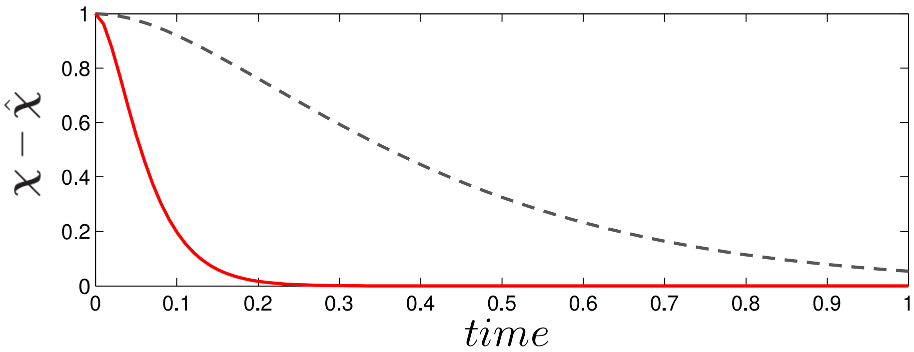
Problem statement

- standard visual servoing control law

$$\dot{s} = L(s, \chi)u \rightarrow u = -kL(s, \hat{\chi})^{\dagger}(s - s^*)$$

F. Chaumette, S. Hutchinson. Visual servo control, Part I: Basic approaches. RAM, 2006.





- act on the camera motion during the servoing transient so as to 'optimally' estimate the scene structure
- results:
  - better knowledge of the scene during task execution
    → task convergence closer to ideality
  - better knowledge of the scene at the end of the task
    → can be used for other purposes

2^{nd} order redundancy resolution

- excitation cost function $\mathcal{V} = \mathcal{V}(\sigma_i^2(\boldsymbol{s}, \boldsymbol{v})) = \mathcal{V}(\boldsymbol{s}, \dot{\boldsymbol{q}})$

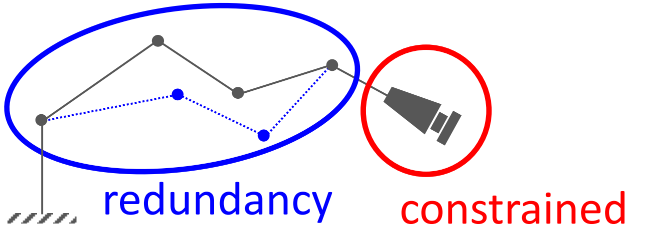
$\dot{\mathcal{V}} \approx \nabla_{\dot{\boldsymbol{q}}} \mathcal{V}^T \ddot{\boldsymbol{q}}$ requires acceleration action

- classical 2nd order projected gradient $\boldsymbol{e} = \boldsymbol{s} - \boldsymbol{s}^*$

$$\ddot{\boldsymbol{q}} = \boldsymbol{J}^\dagger(-k_v \dot{\boldsymbol{e}} - k_p \boldsymbol{e} - \dot{\boldsymbol{J}}\dot{\boldsymbol{q}}) + \underbrace{\boxed{(\boldsymbol{I} - \boldsymbol{J}\boldsymbol{J}^\dagger)}}_{\boldsymbol{P}} \nabla_{\dot{\boldsymbol{q}}} \mathcal{V} \qquad \widehat{\boldsymbol{J}} = \widehat{\boldsymbol{L}} \boldsymbol{J}_C$$

- if $\operatorname{rank} \boldsymbol{L} = 6$

$$\boldsymbol{P} \nabla_{\dot{\boldsymbol{q}}} \mathcal{V} = 0, \ \forall \mathcal{V}(\boldsymbol{v}(\dot{\boldsymbol{q}}))$$



  - $\boldsymbol{s}$ completely constraints camera motion
  - only "internal" redundancy → not useful for active SfM

R. Spica, P. Robuffo Giordano, and F. Chaumette, "Bridging Visual Control and Active Perception via a Large Projection Operator," IEEE Trans. on Robotics, **under review since April 2015**.

Redundancy maximization

- excitation cost function $\mathcal{V} = \mathcal{V}(\sigma_i^2(\boldsymbol{s}, \boldsymbol{v})) = \mathcal{V}(\boldsymbol{s}, \dot{\boldsymbol{q}})$

  $\dot{\mathcal{V}} \approx \nabla_{\dot{\boldsymbol{q}}} \mathcal{V}^T \ddot{\boldsymbol{q}}$ requires acceleration action

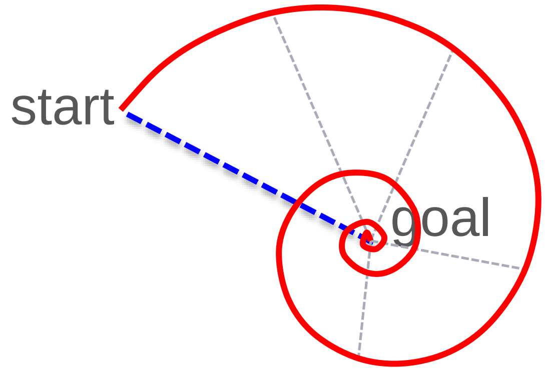
- redundant 2nd order projected gradient $\nu = \|\boldsymbol{s} - \boldsymbol{s}^*\|$ (extension of [Marey, Chaumette 2010])

$$\ddot{\boldsymbol{q}} = \boldsymbol{J}_\nu^\dagger(-k_v\dot{\nu} - k_p\nu - \dot{\boldsymbol{J}}_\nu\dot{\boldsymbol{q}}) + \underbrace{\boxed{(\boldsymbol{I} - \boldsymbol{J}_\nu\boldsymbol{J}_\nu^\dagger)}}_{\boldsymbol{P}_\nu}\nabla_{\dot{\boldsymbol{q}}}\mathcal{V} \qquad \widehat{\boldsymbol{J}}_\nu = \frac{1}{\nu}\boldsymbol{e}^T\widehat{\boldsymbol{J}}$$

- now (even if $\operatorname{rank} \boldsymbol{L} = 6$) in general

  $\boldsymbol{P}_\nu \nabla_{\dot{\boldsymbol{q}}} \mathcal{V} \neq 0,$

  - redundancy to maximize excitation



R. Spica, P. Robuffo Giordano, and F. Chaumette, "Bridging Visual Control and Active Perception via a Large Projection Operator," IEEE Trans. on Robotics, **under review since April 2015**.

Second order switching strategy

- alternative control laws:

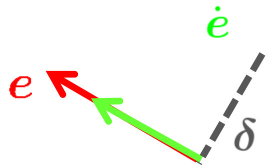
1) $\ddot{\boldsymbol{q}} = \boldsymbol{J}_\nu^\dagger(-k_v\dot{\nu} - k_p\nu - \dot{\boldsymbol{J}}_\nu\dot{\boldsymbol{q}}) + (\boldsymbol{I}_n - \boldsymbol{J}_\nu^\dagger\boldsymbol{J}_\nu)\ddot{\boldsymbol{q}}_w$

2) $\ddot{\boldsymbol{q}} = \boldsymbol{J}^\dagger(-k_v\dot{\boldsymbol{e}} - k_p\boldsymbol{e} - \dot{\boldsymbol{J}}\dot{\boldsymbol{q}})$

Second order switching strategy

- alternative control laws:
  1) $\ddot{\boldsymbol{q}} = \boldsymbol{J}_\nu^\dagger(-k_v\dot{\nu} - k_p\nu - \dot{\boldsymbol{J}}_\nu\dot{\boldsymbol{q}}) + (\boldsymbol{I}_n - \boldsymbol{J}_\nu^\dagger\boldsymbol{J}_\nu)\ddot{\boldsymbol{q}}_w$
  2) $\ddot{\boldsymbol{q}} = \boldsymbol{J}^\dagger(-k_v\dot{\boldsymbol{e}} - k_p\boldsymbol{e} - \dot{\boldsymbol{J}}\dot{\boldsymbol{q}})$
- (1) is singular when $\nu \approx 0 \longrightarrow$ we need to switch to (2)
- monotonic convergence only if $\boldsymbol{e}$, $\dot{\boldsymbol{e}}$ aligned before switch

$$\boldsymbol{\delta} = \left(\boldsymbol{I}_m - \frac{\boldsymbol{e}\boldsymbol{e}^T}{\boldsymbol{e}^T\boldsymbol{e}}\right)\dot{\boldsymbol{e}} = \boldsymbol{P}_{\boldsymbol{e}}\dot{\boldsymbol{e}} \approx 0$$



Second order switching strategy

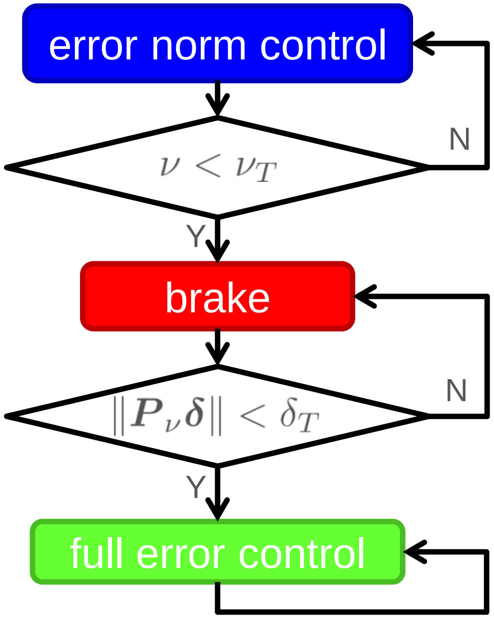
- alternative control laws:

1) $\ddot{\boldsymbol{q}} = \boldsymbol{J}_\nu^\dagger(-k_v\dot{\nu} - k_p\nu - \dot{\boldsymbol{J}}_\nu\dot{\boldsymbol{q}}) + (\boldsymbol{I}_n - \boldsymbol{J}_\nu^\dagger\boldsymbol{J}_\nu)\ddot{\boldsymbol{q}}_w$

2) $\ddot{\boldsymbol{q}} = \boldsymbol{J}^\dagger(-k_v\dot{\boldsymbol{e}} - k_p\boldsymbol{e} - \dot{\boldsymbol{J}}\dot{\boldsymbol{q}})$

- (1) is singular when $\nu \approx 0 \longrightarrow$ we need to switch to (2)
- monotonic convergence only if $\boldsymbol{e}$, $\dot{\boldsymbol{e}}$ aligned before switch

$$\boldsymbol{\delta} = \left(\boldsymbol{I}_m - \frac{\boldsymbol{e}\boldsymbol{e}^T}{\boldsymbol{e}^T\boldsymbol{e}}\right)\dot{\boldsymbol{e}} = \boldsymbol{P}_{\boldsymbol{e}}\dot{\boldsymbol{e}} \approx 0$$

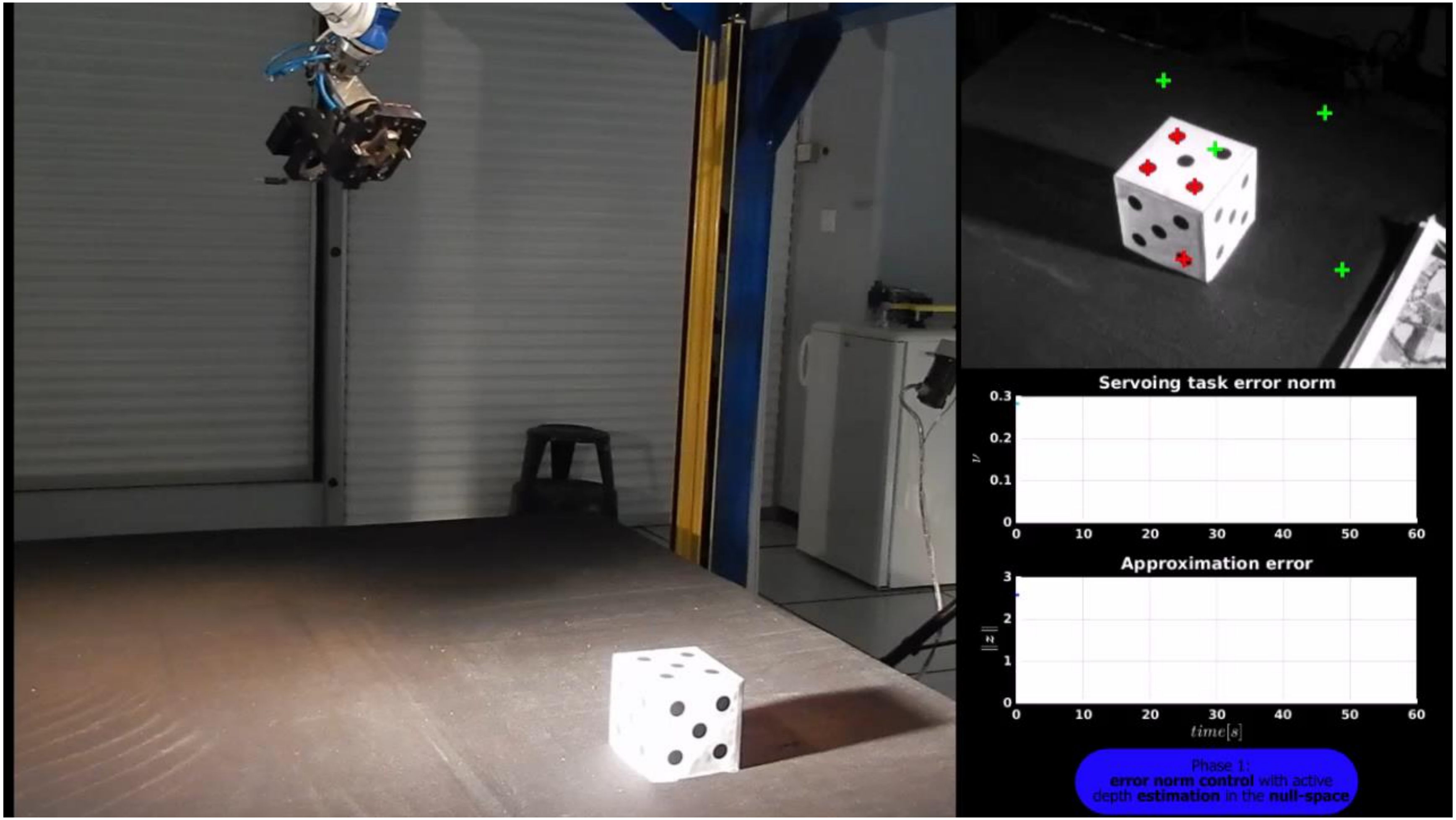


- finally the secondary task is

$\ddot{\boldsymbol{q}}_w = \nabla_{\dot{\boldsymbol{q}}}\mathcal{V}$ observability maximization

$\ddot{\boldsymbol{q}}_w = k_\delta\nabla_{\dot{\boldsymbol{q}}}\|\boldsymbol{\delta}\|$ alignment

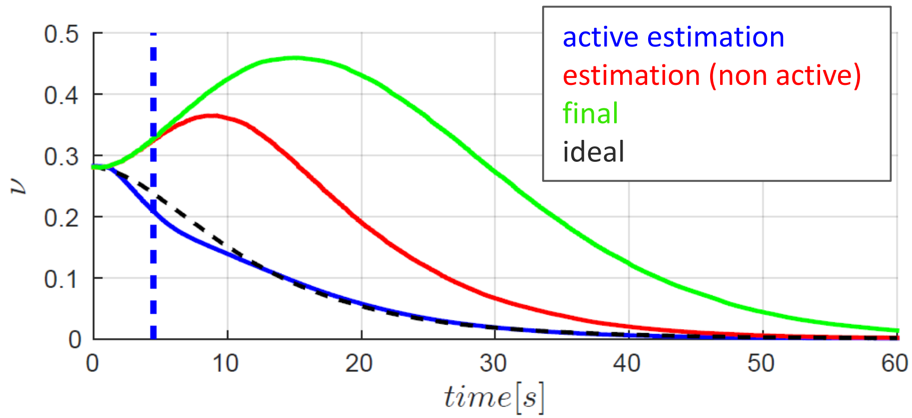
Experimental results



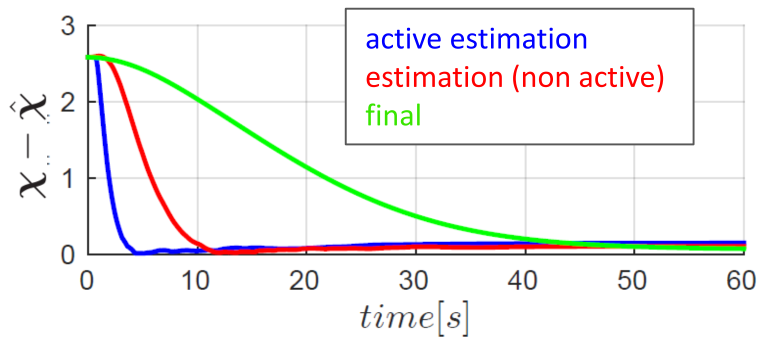
R. Spica, P. Robuffo Giordano, and F. Chaumette, “Bridging Visual Control and Active Perception via a Large Projection Operator,” IEEE Trans. on Robotics, **under review since April 2015**.

Experimental results for 4 point features

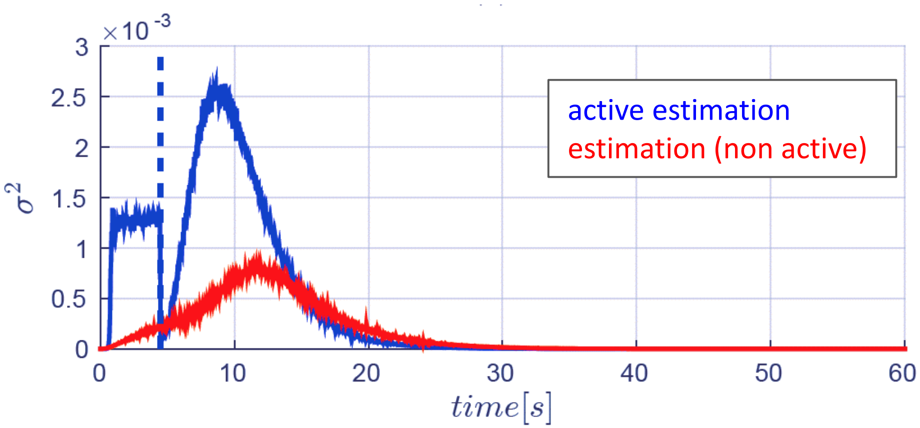
task error



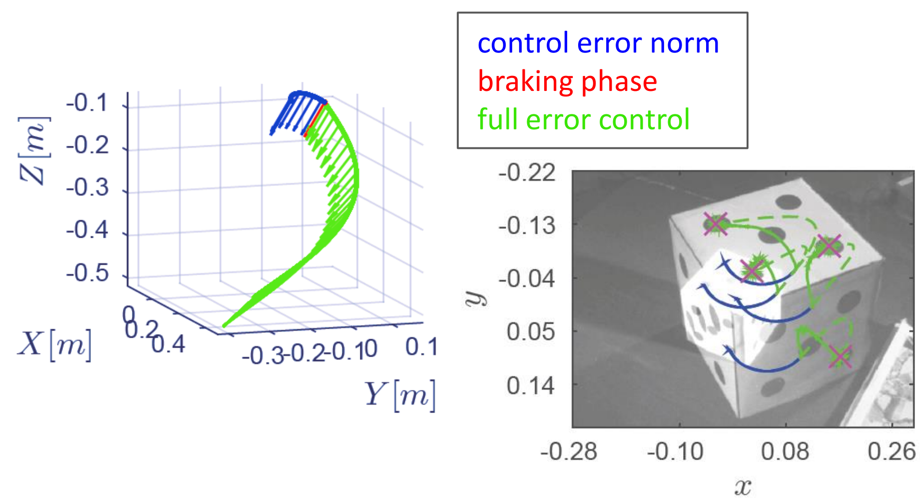
approximation error



observer eigenvalue

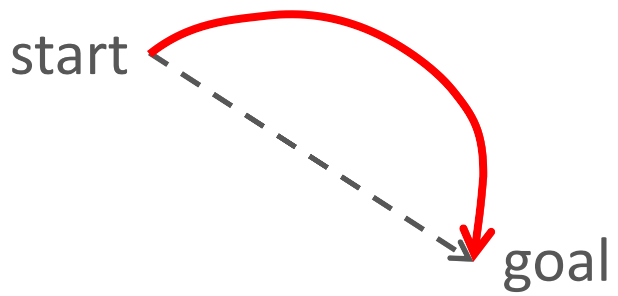


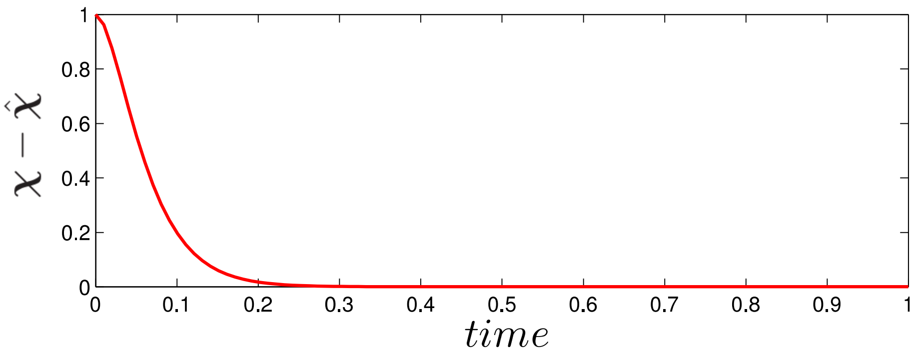
camera trajectory with active estimation



Active estimation triggering

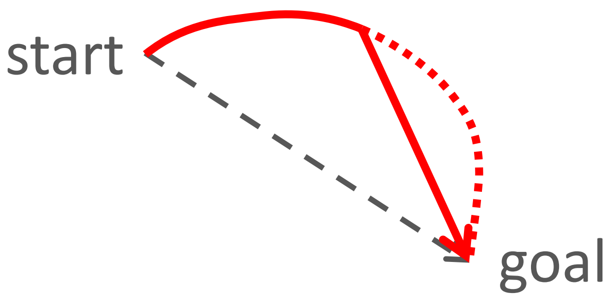
- weight active part depending on estimation status

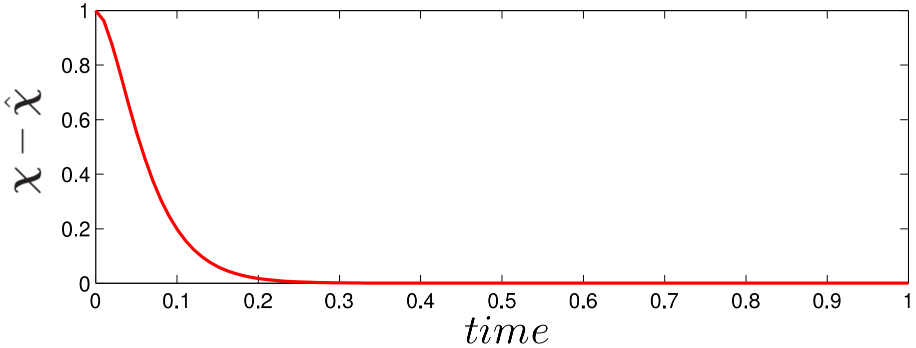




Active estimation triggering

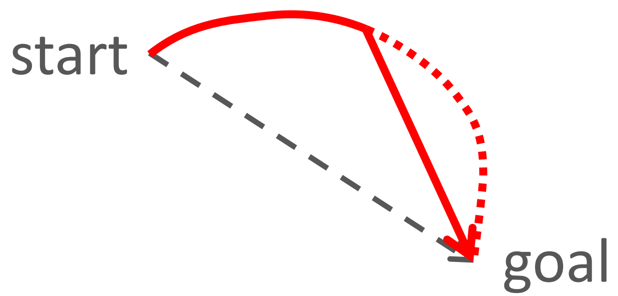
- weight active part depending on estimation status

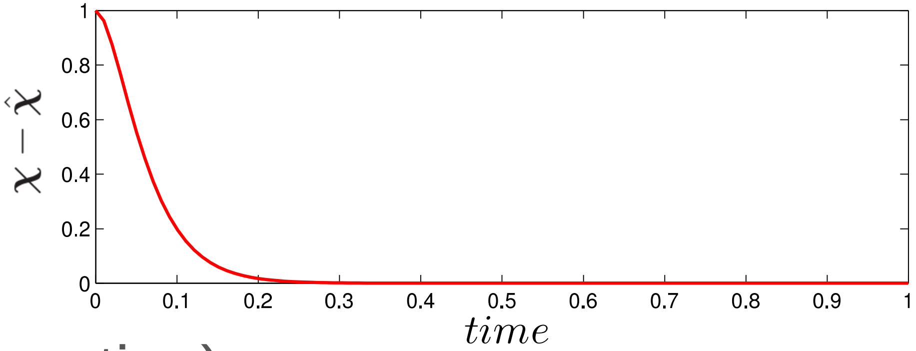




Active estimation triggering

- weight active part depending on estimation status



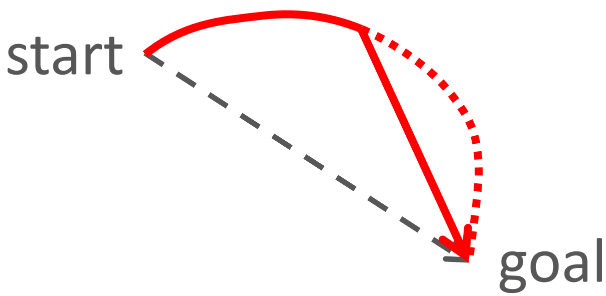


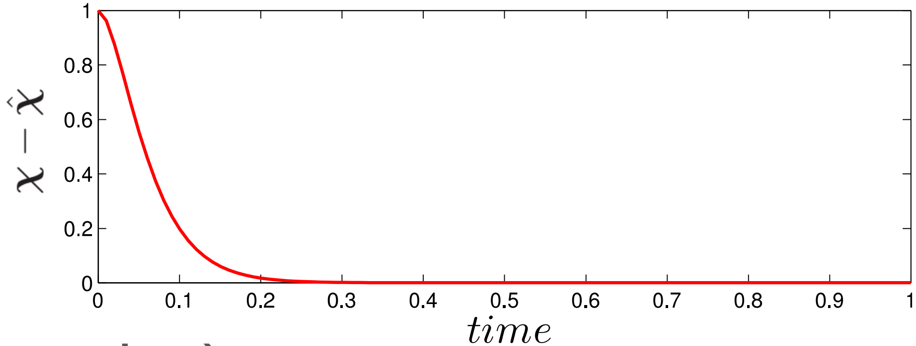
- if $\sigma_1^2 > 0$ (exciting camera motion)

$$E = \frac{1}{T}\int_{t-T}^{t} \|\boldsymbol{s}(\tau) - \hat{\boldsymbol{s}}(\tau)\|^2 d\tau \equiv 0 \iff \widetilde{\boldsymbol{\chi}} = 0$$

Active estimation triggering

- weight active part depending on estimation status

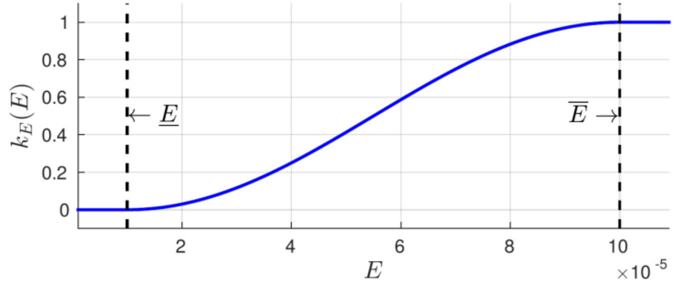


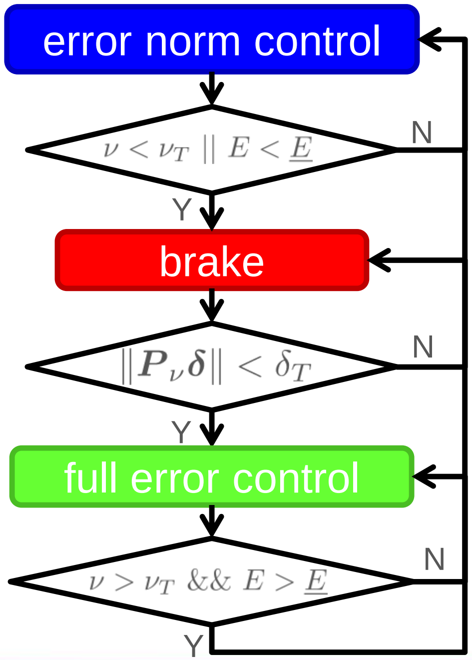


- if $\sigma_1^2 > 0$ (exciting camera motion)

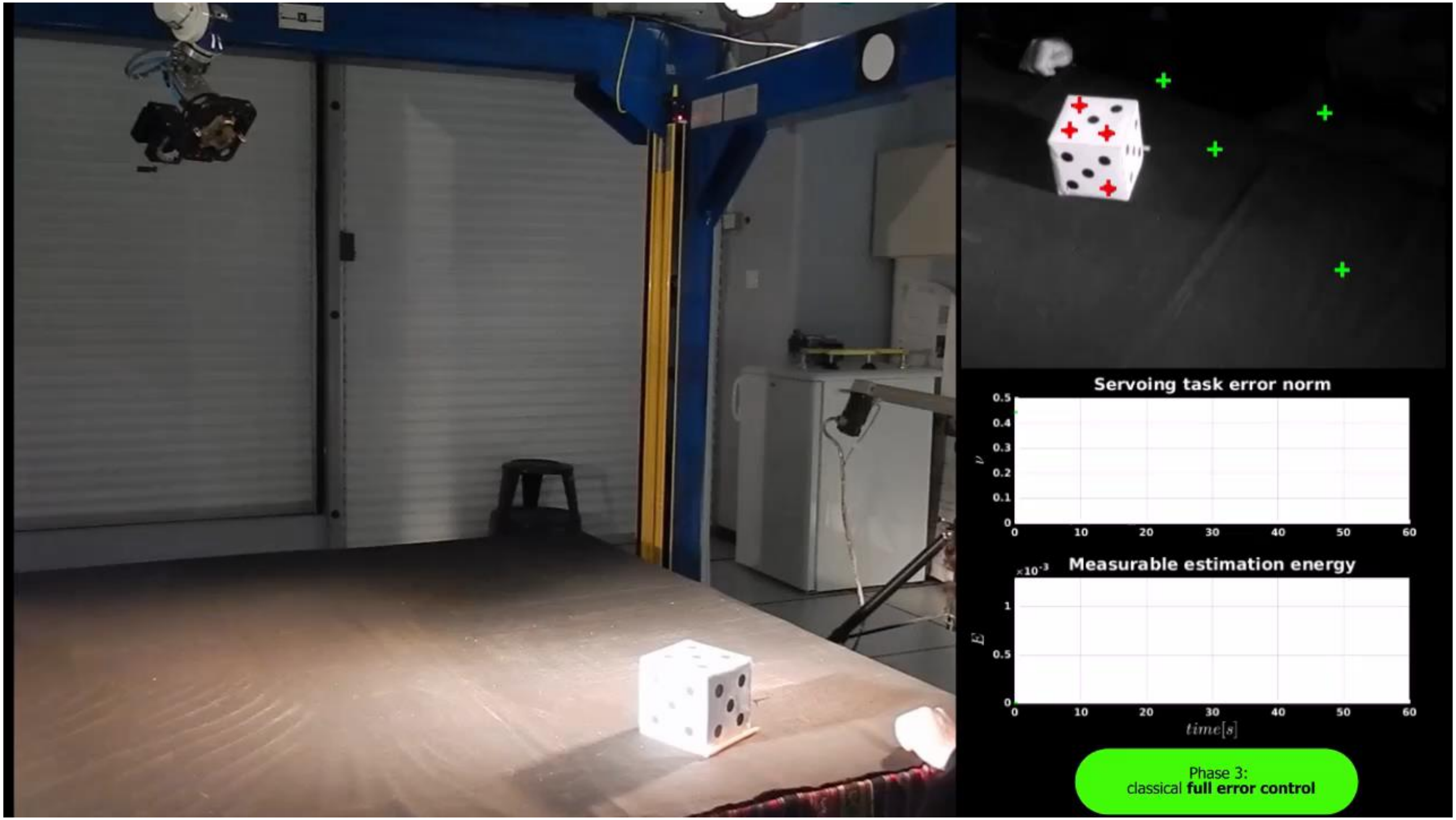
$$E = \frac{1}{T}\int_{t-T}^{t} \|\boldsymbol{s}(\tau) - \hat{\boldsymbol{s}}(\tau)\|^2 d\tau \equiv 0 \Longleftrightarrow \widetilde{\boldsymbol{\chi}} = 0$$

$$\mathcal{V}(\dot{\boldsymbol{q}},\, E) = k_E(E)\gamma \log\left(\frac{\gamma + \sigma_1^2}{\gamma}\right) - \frac{k_d}{2}\|\dot{\boldsymbol{q}}\|^2$$



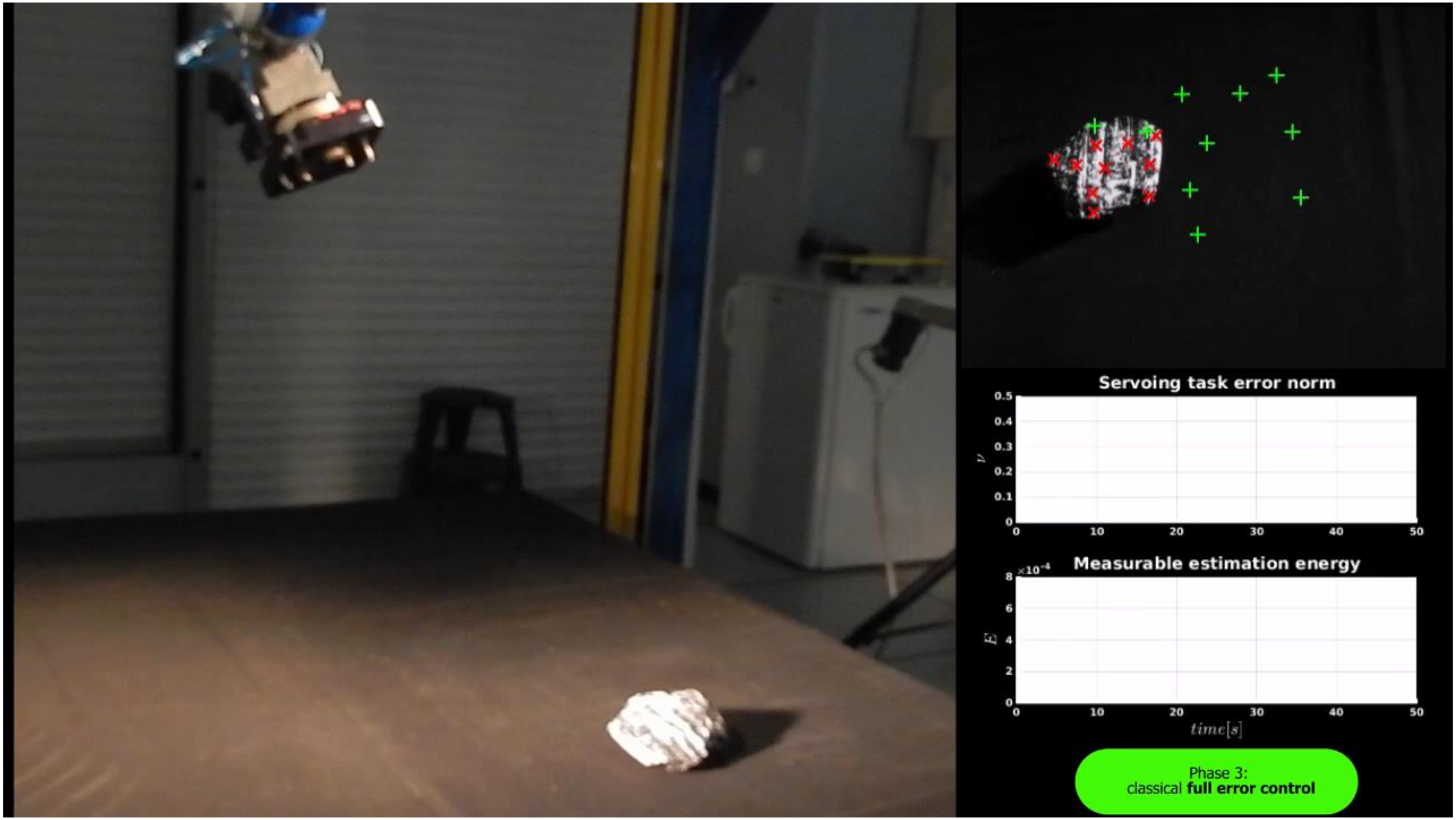


Experimental results with adaptive strategy 1/2



R. Spica, P. Robuffo Giordano, and F. Chaumette, “Bridging Visual Control and Active Perception via a Large Projection Operator,” IEEE Trans. on Robotics, **under review since April 2015**.

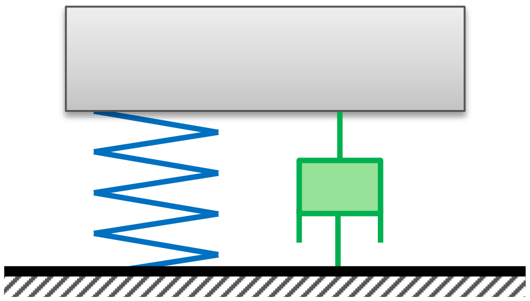
Experimental results with adaptive strategy 2/2



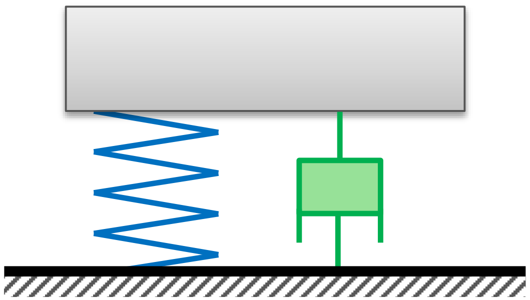
R. Spica, P. Robuffo Giordano, and F. Chaumette, “Bridging Visual Control and Active Perception via a Large Projection Operator,” IEEE Trans. on Robotics, **under review since April 2015**.

Conclusions 1/2

- generic strategy to: [CDC, 2013]
  - characterize the (nonlinear) dynamics of SfM
  - impose a transient ~linear 2nd-order reference system (i.e., assign the poles)

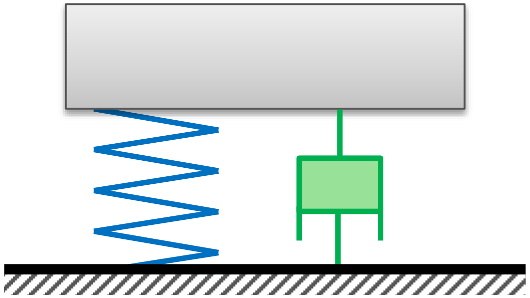


Conclusions 1/2



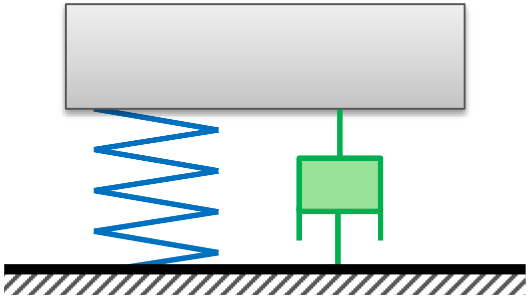
- generic strategy to: [CDC, 2013]
  - characterize the (nonlinear) dynamics of SfM
  - impose a transient ~linear 2nd-order reference system (i.e., assign the poles)
- by acting online on:
  - observer gains to fix damping factor
  - inputs $\boldsymbol{v}$ (“active” estimation) to maximize cut-off frequency

Conclusions 1/2



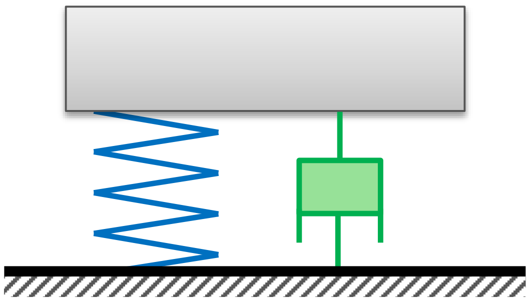
- generic strategy to: [CDC, 2013]
  - characterize the (nonlinear) dynamics of SfM
  - impose a transient ~linear 2nd-order reference system (i.e., assign the poles)
- by acting online on:
  - observer gains to fix damping factor
  - inputs $\boldsymbol{v}$ (“active” estimation) to maximize cut-off frequency
- byproduct:
  - optimized convergence for a given max $\|\boldsymbol{v}\|$
  - predictability (like linear system)

Conclusions 1/2



- generic strategy to: [CDC, 2013]
  - characterize the (nonlinear) dynamics of SfM
  - impose a transient ~linear 2nd-order reference system (i.e., assign the poles)
- by acting online on:
  - observer gains to fix damping factor
  - inputs $\boldsymbol{v}$ (“active” estimation) to maximize cut-off frequency
- byproduct:
  - optimized convergence for a given max $\|\boldsymbol{v}\|$
  - predictability (like linear system)
- experimental case studies: point [CDC‘13, TRO‘14], sphere [ICRA‘14, TRO‘14], cylinder [ICRA‘14, TRO‘14], plane [ICRA‘14/‘15]

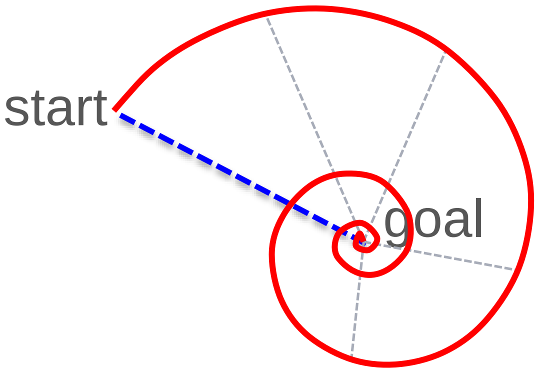
Conclusions 1/2



- generic strategy to: [CDC, 2013]
  - characterize the (nonlinear) dynamics of SfM
  - impose a transient ~linear 2nd-order reference system (i.e., assign the poles)
- by acting online on:
  - observer gains to fix damping factor
  - inputs $\boldsymbol{v}$ ("active" estimation) to maximize cut-off frequency
- byproduct:
  - optimized convergence for a given max $\|\boldsymbol{v}\|$
  - predictability (like linear system)
- experimental case studies: point [CDC'13, TRO'14], sphere [ICRA'14, TRO'14], cylinder [ICRA'14, TRO'14], plane [ICRA'14/'15]
- extension to dense state estimation (with Prof. Mahony)

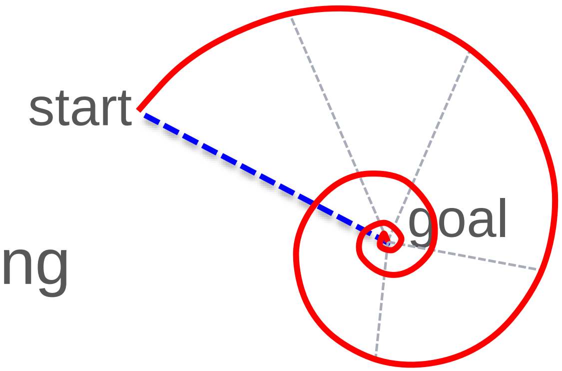
Conclusions 2/2

- general strategy for "deforming" the camera trajectory during IBVS transient to maximize observability [ICRA'14]
  - improved performance during task execution
  - better final estimation accuracy



Conclusions 2/2

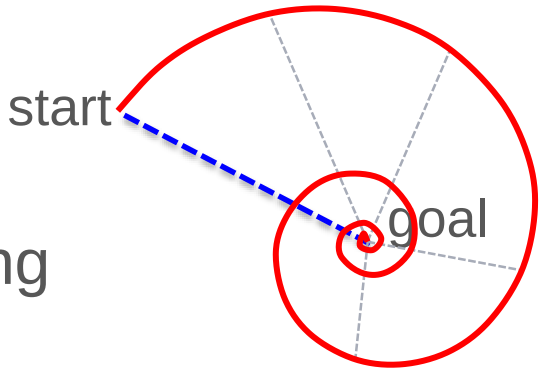
- general strategy for "deforming" the camera trajectory during IBVS transient to maximize observability [ICRA'14]
  - improved performance during task execution
  - better final estimation accuracy



- extension of [Marey, Chaumette 2010] to maximize redundancy while maintaining monotonic convergence [ICRA'14]

Conclusions 2/2

- general strategy for "deforming" the camera trajectory during IBVS transient to maximize observability [ICRA'14]
  - improved performance during task execution
  - better final estimation accuracy

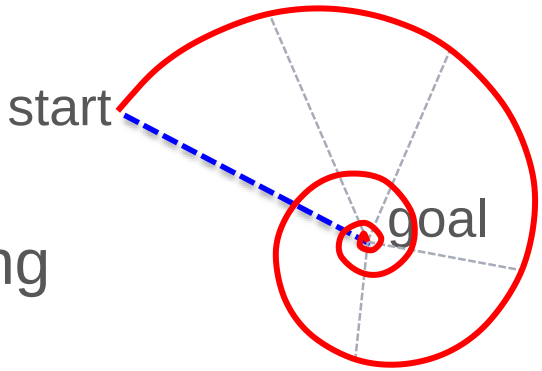


- extension of [Marey, Chaumette 2010] to maximize redundancy while maintaining monotonic convergence [ICRA'14]

- automatic activation/deactivation and tuning of active estimation depending on the current accuracy [TRO'15]

Conclusions 2/2

- general strategy for “deforming” the camera trajectory during IBVS transient to maximize observability [ICRA‘14]
  - improved performance during task execution
  - better final estimation accuracy



- extension of [Marey, Chaumette 2010] to maximize redundancy while maintaining monotonic convergence [ICRA‘14]
- automatic activation/deactivation and tuning of active estimation depending on the current accuracy [TRO‘15]
- experimental validation with simple target and realistic unstructured object [ICRA‘14, TRO‘15]

Open issues and perspectives – active SfM

- optimization of $\mathbf{\Omega}(t)\mathbf{\Omega}(t)^T$ might result in local minima
  - use extended horizon planning (and re-planning)

$$\int_{t}^{t+T} \mathbf{\Omega}(\tau)\,\mathbf{\Omega}^T(\tau)\,\mathrm{d}\tau$$

  - consider $m < p$ ($\mathbf{\Omega}(t)\mathbf{\Omega}(t)^T$ can never be full-rank)

Open issues and perspectives – active SfM

- optimization of $\boldsymbol{\Omega}(t)\boldsymbol{\Omega}(t)^T$ might result in local minima
  - use extended horizon planning (and re-planning)

$$\int_t^{t+T} \boldsymbol{\Omega}(\tau)\,\boldsymbol{\Omega}^T(\tau)\,\mathrm{d}\tau$$

  - consider $m < p$ ($\boldsymbol{\Omega}(t)\boldsymbol{\Omega}(t)^T$ can never be full-rank)
- deterministic framework → possibly not robust w.r.t. noise
  - probabilistic estimators
  - noise-aware metrics (Fisher information matrix)

$$\int_t^{t+T} \boldsymbol{\Omega}(\tau)\,\boldsymbol{R}^{-1}\boldsymbol{\Omega}^T(\tau)\,\mathrm{d}\tau$$

Open issues and perspectives – active SfM

- optimization of $\mathbf{\Omega}(t)\mathbf{\Omega}(t)^T$ might result in local minima
  - use extended horizon planning (and re-planning)

$$\int_t^{t+T} \mathbf{\Omega}(\tau)\, \mathbf{\Omega}^T(\tau)\, \mathrm{d}\tau$$

  - consider $m < p$ ($\mathbf{\Omega}(t)\mathbf{\Omega}(t)^T$ can never be full-rank)
- deterministic framework → possibly not robust w.r.t. noise
  - probabilistic estimators
  - noise-aware metrics (Fisher information matrix)

$$\int_t^{t+T} \mathbf{\Omega}(\tau)\, \boldsymbol{R}^{-1} \mathbf{\Omega}^T(\tau)\, \mathrm{d}\tau$$

- moving target ($m < p$)

Open issues and perspectives – dense SfM

- modeling/propagation of depth discontinuities
  - shock/rarefaction waves

Open issues and perspectives – dense SfM

- modeling/propagation of depth discontinuities
  - shock/rarefaction waves
- robustness w.r.t. numerical discretization
  - alternative (dis)similarity measurements to $\widehat{Y} - Y$ (e.g. mutual information)

Open issues and perspectives – dense SfM

- modeling/propagation of depth discontinuities
  - shock/rarefaction waves
- robustness w.r.t. numerical discretization
  - alternative (dis)similarity measurements to $\widehat{Y} - Y$ (e.g. mutual information)
- no stability proof
  - infinite dimensional port-Hamiltonian framework [van der Schaft, 2006]

Open issues and perspectives – dense SfM

- modeling/propagation of depth discontinuities
  - shock/rarefaction waves
- robustness w.r.t. numerical discretization
  - alternative (dis)similarity measurements to $\widehat{Y} - Y$ (e.g. mutual information)
- no stability proof
  - infinite dimensional port-Hamiltonian framework [van der Schaft, 2006]
- apply active estimation

Open issues and perspectives – control + SfM

- no formal stability proof for IBVS + active estimation
  - no separation principle
  - encouraging experimental results

Open issues and perspectives – control + SfM

- no formal stability proof for IBVS + active estimation
  - no separation principle
  - encouraging experimental results

- possible constraints violation (e.g. limited field of view, joints limits, …)
  - advanced task priority framework
  - model predictive control

Open issues and perspectives – control + SfM

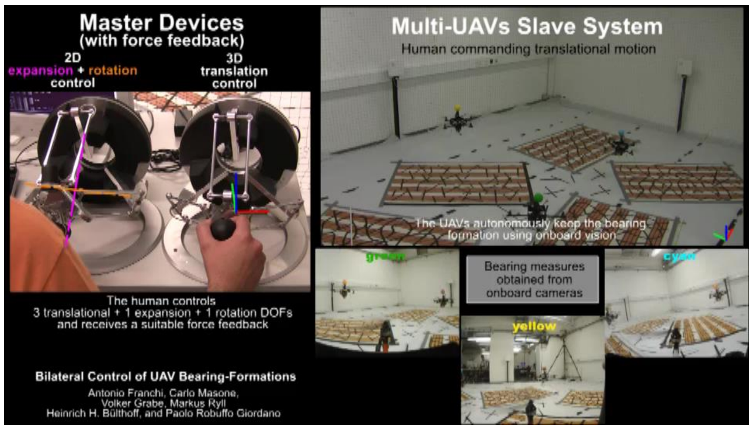
- no formal stability proof for IBVS + active estimation
  - no separation principle
  - encouraging experimental results

- possible constraints violation (e.g. limited field of view, joints limits, …)
  - advanced task priority framework
  - model predictive control

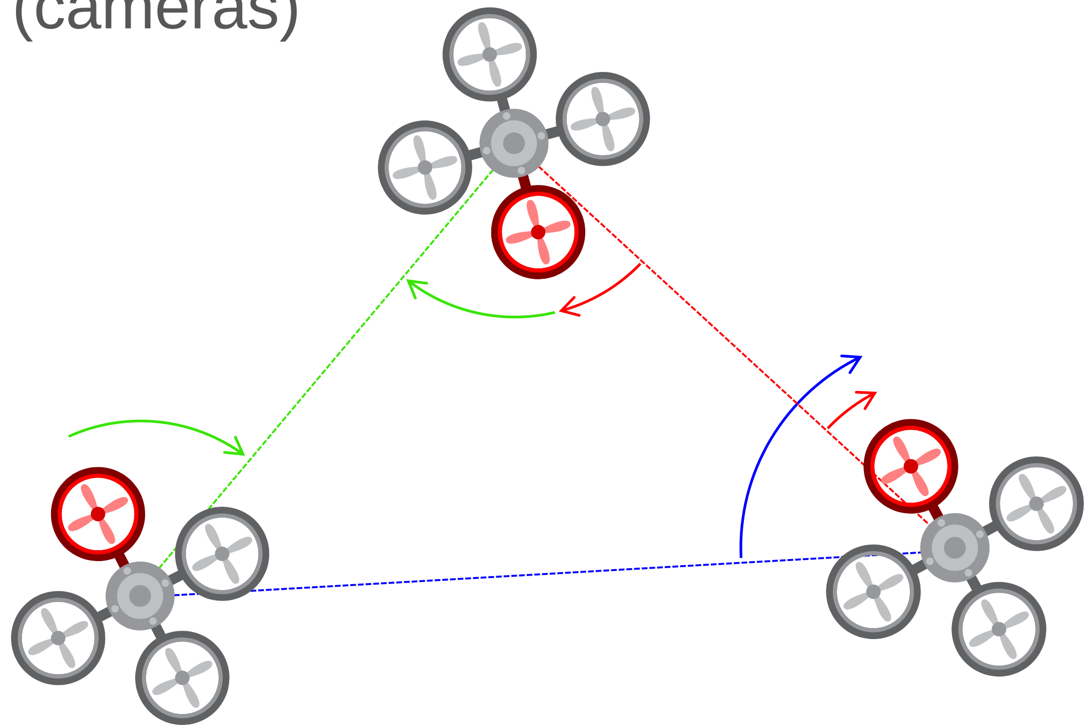
- need velocity measurement/control typically difficult on mobile robots and UAVs
  - acceleration measurements (IMU) and torque/force input
  - non-holonomic/under-actuated control

Possible application

- multi-robot systems, e.g. estimate/control formation from bearing measurements (cameras)



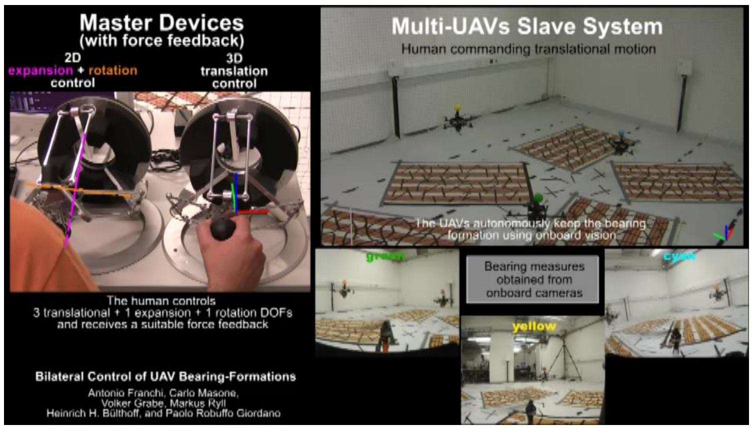
[Franchi et al. 2012; Bishop et al. 2011]



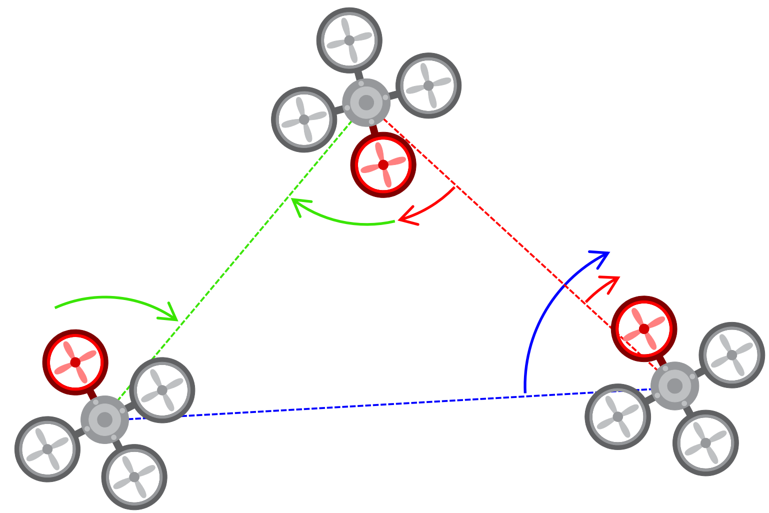
- many robots = many degrees of freedom to optimize
- decentralization

Possible application

- multi-robot systems, e.g. estimate/control formation from bearing measurements (cameras)



[Franchi et al. 2012; Bishop et al. 2011]



- many robots = many degrees of freedom to optimize
- decentralization

Thanks for your attention

Riccardo Spica

Lagadic group

Inria Rennes Bretagne Atlantique & IRISA

http://www.irisa.fr/lagadic