Assistive Robotics

IBVS
Image-Based Visual Servoing

Área de Ingeniería de Sistemas y Automática
Departamento de Informática e Ingeniería de Sistemas
Universidad de Zaragoza

Departamento de Informática e Ingeniería de Sistemas Universidad Zaragoza

Gonzalo López Nicolás

Index

- Introduction
- Robotic manipulation
- Vision systems
- Classification of visual control systems
- Position-Based Visual Servoing
- **Image-Based Visual Servoing**
  - Introduction
  - Interaction matrix
  - Control law
- Stability analysis
- Hybrid Visual Servoing
- Conclusion and bibliography

Image-Based Visual Servoing

- Image-Based Visual Servoing
  - The target is defined in image coordinates
  - The task is defined in the image
    - Can reduce computational delay, eliminating the need for image interpretation
    - Non-linear and coupled system
    - More robust to poor calibration

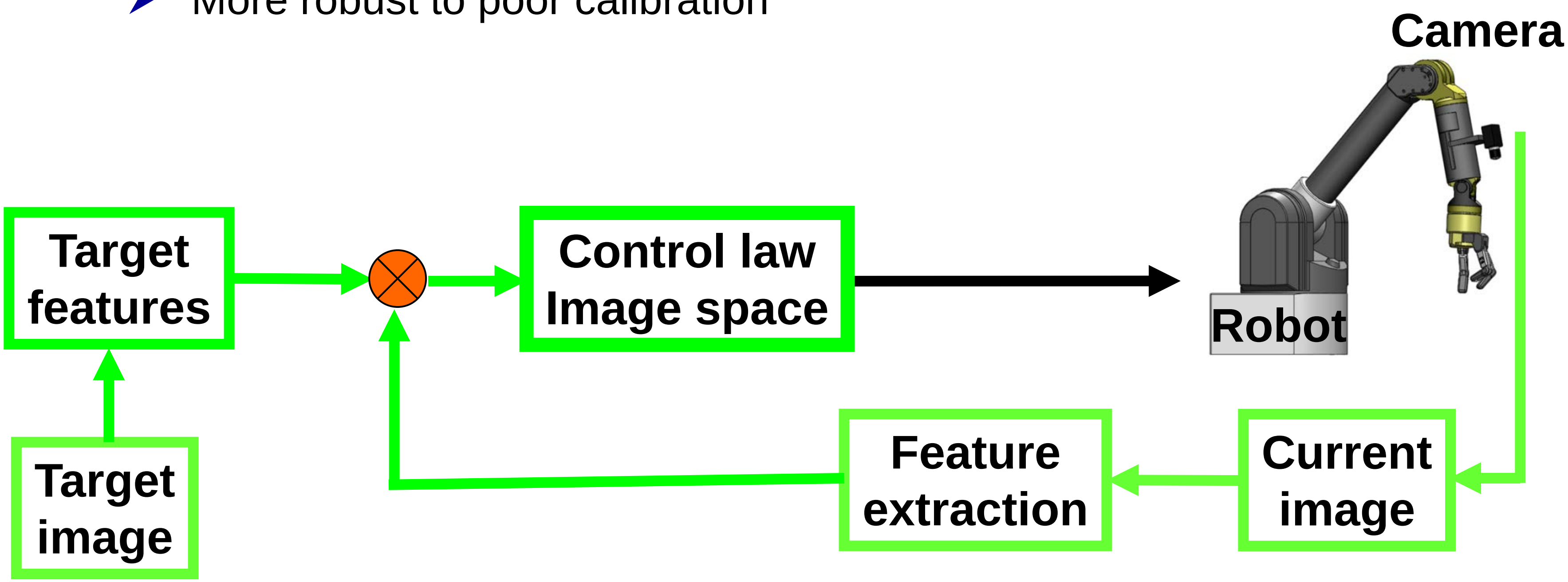


Image-Based Visual Servoing: Introduction

- Image-based visual servoing
  - The target is defined in image coordinates
  - The task is defined in the image space
  - Non-linear and coupled system
  - We seek for robustness to poor calibration

  - Related bibliography :
    - [Chaumette et al. 1991], [Espiau et al. 1992], [Corke & Good 1996], [Ma et al. 1996], [Matsumoto et al. 1996], [Hager 1997], [Christensen et al. 1999], [Drummond & Cipolla 1999], [Kragic et al. 2001], [Conticelli & Allotta 2001], [Graefe & Bischoff 2004], [Adachi & Sato 2004], [Argyros et al. 2005], [Remazeilles & Chaumette 2007]

Image-Based Visual Servoing

- Problem statement
  - Select k image features **s** suitable for controlling m degrees of freedom. $(\mathrm{k} \geq \mathrm{m}), (\mathrm{m} \leq 6)$
  - Determine the target $\mathbf{s}^*$
  - Regulate the error to zero $\mathbf{e}(\mathrm{t}) = (\mathbf{s} - \mathbf{s}^*) \rightarrow 0$

    $\text{if } \mathbf{s}^* = cte \Rightarrow \dot{\mathbf{e}} = \dot{\mathbf{s}}$
- Relation between temporal variation of **s** with camera velocity

$$\dot{\mathbf{s}} = \mathbf{L} \cdot \mathbf{v}, \text{ with } \mathbf{L} \in \Re^{k \times 6}$$

$$\mathbf{v} = \begin{pmatrix} v_x & v_y & v_z & \omega_x & \omega_y & \omega_z \end{pmatrix}^T$$

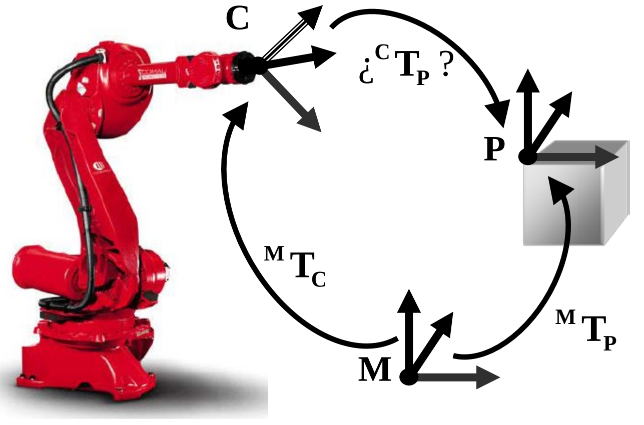
- $\mathbf{L}$: Interaction matrix (image Jacobian)
  - Relation between time variation of error with camera velocity

$$\dot{\mathbf{e}} = \mathbf{L} \cdot \mathbf{v}$$

Image-Based Visual Servoing: Interaction matrix

- Deduction of the interaction matrix
  - Location of a point with respect to the camera
    - World reference: **M**
    - Camera reference: **C(t)**
    - Point reference: **P**

$$¿\dot{\mathbf{s}} \leftrightarrow \mathbf{v}?$$

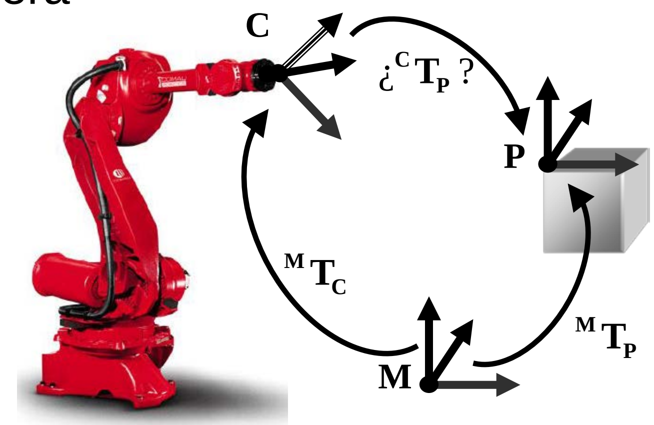


$$¿{}^{C}\mathbf{T}_{P}? \rightarrow {}^{C}\mathbf{T}_{P} = \left({}^{M}\mathbf{T}_{C}\right)^{-1} \cdot {}^{M}\mathbf{T}_{P}$$

$${}^{C}\mathbf{T}_{P} = \begin{pmatrix} {}^{C}\mathbf{R}_{P}(t) & {}^{C}\mathbf{t}_{P}(t) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} {}^{M}\mathbf{R}_{C}(t) & {}^{M}\mathbf{t}_{C}(t) \\ 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} {}^{M}\mathbf{R}_{P} & {}^{M}\mathbf{t}_{P} \\ 0 & 1 \end{pmatrix}$$

$$\begin{cases} {}^{C}\mathbf{R}_{P}(t) = {}^{C}\mathbf{R}_{M}(t) \cdot {}^{M}\mathbf{R}_{P} \\ {}^{C}\mathbf{t}_{P}(t) = {}^{C}\mathbf{R}_{M}(t) \cdot \left({}^{M}\mathbf{t}_{P} - {}^{M}\mathbf{t}_{C}(t)\right) \end{cases}$$

- Deduction of the interaction matrix
  - Derivative of point location with respect to camera



$${}^{C}\mathbf{t}_{P}(t)={}^{C}\mathbf{R}_{M}(t)\cdot\left({}^{M}\mathbf{t}_{P}-{}^{M}\mathbf{t}_{C}(t)\right)$$

$${}^{C}\dot{\mathbf{t}}_{P}={}^{C}\dot{\mathbf{R}}_{M}\cdot\left({}^{M}\mathbf{t}_{P}-{}^{M}\mathbf{t}_{C}\right)-{}^{C}\mathbf{R}_{M}\cdot{}^{M}\dot{\mathbf{t}}_{C}$$

$$\mathbf{R}\in SO(3)\rightarrow\dot{\mathbf{R}}=[\boldsymbol{\omega}]_{\times}\cdot\mathbf{R}$$

$${}^{C}\dot{\mathbf{R}}_{M}=\left({}^{M}\dot{\mathbf{R}}_{C}\right)^{T}=\left([\boldsymbol{\omega}^{M}]_{\times}\cdot{}^{M}\mathbf{R}_{C}\right)^{T}={}^{C}\mathbf{R}_{M}\cdot[\boldsymbol{\omega}^{M}]_{\times}^{T}={}^{C}\mathbf{R}_{M}\cdot[-\boldsymbol{\omega}^{M}]_{\times}=-{}^{C}\mathbf{R}_{M}\cdot[\boldsymbol{\omega}^{M}]_{\times}$$

$${}^{C}\dot{\mathbf{t}}_{P}=-{}^{C}\mathbf{R}_{M}\cdot[\boldsymbol{\omega}^{M}]_{\times}\cdot\left({}^{M}\mathbf{t}_{P}-{}^{M}\mathbf{t}_{C}\right)-{}^{C}\mathbf{R}_{M}\cdot{}^{M}\dot{\mathbf{t}}_{C}$$

$$\underbrace{{}^{C}\dot{\mathbf{t}}_{P}}_{\dot{\mathbf{p}}}=\underbrace{-{}^{C}\mathbf{R}_{M}\cdot\boldsymbol{\omega}^{M}}_{\boldsymbol{\omega}^{C}\rightarrow\boldsymbol{\omega}}\times\underbrace{{}^{C}\mathbf{R}_{M}\cdot\left({}^{M}\mathbf{t}_{P}-{}^{M}\mathbf{t}_{C}\right)}_{{}^{C}\mathbf{t}_{P}\rightarrow\mathbf{p}}-\underbrace{{}^{C}\mathbf{R}_{M}\cdot{}^{M}\dot{\mathbf{t}}_{C}}_{\mathbf{v}^{C}\rightarrow\mathbf{v}}$$

$$\boxed{\dot{\mathbf{p}}=-\boldsymbol{\omega}\times\mathbf{p}-\mathbf{v}}$$

v: camera velocity
ω: camera angular velocity

Image-Based Visual Servoing: Interaction matrix

- Deduction of the interaction matrix

$$\dot{\mathbf{p}} = -\boldsymbol{\omega} \times \mathbf{p} - \mathbf{v}$$

Jacobian

$$\dot{\mathbf{s}} = \begin{pmatrix} \dot{s}_1 \\ \vdots \\ \dot{s}_n \end{pmatrix} = \frac{\partial \mathbf{s}}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} = \mathbf{L} \cdot \mathbf{v} = \begin{pmatrix} \frac{\partial \mathbf{s}_1}{\partial r_1} & \cdots & \frac{\partial \mathbf{s}_1}{\partial r_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{s}_n}{\partial r_1} & \cdots & \frac{\partial \mathbf{s}_n}{\partial r_k} \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}$$

- Image features: $\mathbf{s}$
- Camera position: $\mathbf{r}$
- Camera velocity: $\mathbf{v} = \dot{\mathbf{r}}$
- Interaction matrix: $\mathbf{L}$

$$¿\dot{\mathbf{p}} \leftrightarrow \dot{\mathbf{s}} \leftrightarrow \mathbf{v}?$$

Image-Based Visual Servoing: Interaction matrix

- Deduction of the interaction matrix

$$\dot{\mathbf{p}} = -\boldsymbol{\omega} \times \mathbf{p} - \mathbf{v}$$

$$\mathbf{p} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \rightarrow \dot{\mathbf{p}} = \begin{pmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{pmatrix} = -\begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \times \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} Y \cdot \omega_z - Z \cdot \omega_y - v_x \\ Z \cdot \omega_x - X \cdot \omega_z - v_y \\ X \cdot \omega_y - Y \cdot \omega_x - v_z \end{pmatrix}$$

  - Points in the image (Projection of $\mathbf{p}$ in $\mathbf{x}$):

$$\mathbf{x} = (x, y)^T = (X/Z, Y/Z)^T$$

$$\dot{x} = \frac{\dot{X} \cdot Z - X \cdot \dot{Z}}{Z^2} = \frac{\dot{X} - x \cdot \dot{Z}}{Z}$$

$$\dot{y} = \frac{\dot{Y} \cdot Z - Y \cdot \dot{Z}}{Z^2} = \frac{\dot{Y} - y \cdot \dot{Z}}{Z}$$

$$¿\dot{\mathbf{s}} = \dot{\mathbf{x}} \leftrightarrow \mathbf{v}?$$

Image-Based Visual Servoing: Interaction matrix

- Deduction of the interaction matrix $\mathbf{L}$: $\dot{\mathbf{s}} = \mathbf{L} \cdot \mathbf{v}$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -1/Z & 0 & x/Z & x \cdot y & -(1+x^2) & y \\ 0 & -1/Z & y/Z & (1+y^2) & -x \cdot y & -x \end{pmatrix} \cdot \mathbf{v}$$

$$\mathbf{v} = (v_x \quad v_y \quad v_z \quad \omega_x \quad \omega_y \quad \omega_z)^T$$

  - For a point: $\mathbf{s} \in \Re^{\mathbf{2}}, \mathbf{L} \in \Re^{\mathbf{2\times 6}}$
  - For a set of n points: $\mathbf{s} \in \Re^{\mathbf{2n}}, \mathbf{L} \in \Re^{\mathbf{2n\times 6}}$

$$\mathbf{s} = (x_1 \quad y_1 \quad \ldots \quad x_n \quad y_n)^T$$

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_1 \\ \vdots \\ \mathbf{L}_n \end{pmatrix} = \begin{pmatrix} -1/Z_1 & 0 & x_1/Z_1 & x_1 \cdot y_1 & -(1+x_1^2) & y_1 \\ 0 & -1/Z_1 & y_1/Z_1 & (1+y_1^2) & -x_1 \cdot y_1 & -x_1 \\ & & & \vdots & & \\ -1/Z_n & 0 & x_n/Z_n & x_n \cdot y_n & -(1+x_n^2) & y_n \\ 0 & -1/Z_n & y_n/Z_n & (1+y_n^2) & -x_n \cdot y_n & -x_n \end{pmatrix}$$

Image-Based Visual Servoing: Interaction matrix

- Deduction of the interaction matrix in cylindrical coordinates
  - Considering the image points in cylindrical coordinates instead of Cartesian coordinates

$$\dot{\mathbf{s}} = \mathbf{L} \cdot \mathbf{v} \quad \begin{cases} \rho = \sqrt{\mathrm{x}^2 + \mathrm{y}^2} \\ \theta = \arctan \frac{\mathrm{y}}{\mathrm{x}} \to \mathrm{atan2}(\mathrm{y}, \mathrm{x}) \end{cases}$$

$$\dot{\rho} = \frac{\mathrm{x} \cdot \dot{\mathrm{x}} - \mathrm{y} \cdot \dot{\mathrm{y}}}{\rho} \qquad \dot{\theta} = \frac{\mathrm{x} \cdot \dot{\mathrm{y}} - \dot{\mathrm{x}} \cdot \mathrm{y}}{\rho^2}$$

$$\mathbf{s} = \begin{pmatrix} \rho_1 & \theta_1 & \ldots & \rho_n & \theta_n \end{pmatrix}^T$$

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_\rho \\ \mathbf{L}_\theta \end{pmatrix} = \begin{pmatrix} \frac{-\cos\theta}{\mathrm{Z}} & \frac{-\sin\theta}{\mathrm{Z}} & \frac{\rho}{\mathrm{Z}} & (1+\rho^2)\sin\theta & -(1+\rho^2)\cos\theta & 0 \\ \frac{\sin\theta}{\rho \cdot \mathrm{Z}} & \frac{-\cos\theta}{\rho \cdot \mathrm{Z}} & 0 & \frac{\cos\theta}{\rho} & \frac{\sin\theta}{\rho} & -1 \end{pmatrix}$$

$$\mathbf{v} = \begin{pmatrix} v_x & v_y & v_z & \omega_x & \omega_y & \omega_z \end{pmatrix}^T$$

  - Produces decoupling between columns 3 and 6 of the interaction matrix.

Image-Based Visual Servoing: Interaction matrix

- Deduction of the interaction matrix
  - Some possible features to use:
    - Points
      - Cartesian coordinates
      - Cylindrical coordinates
    - Segments
    - Straight lines
    - Circles
    - Cylinders
    - Contours of objects
    - Moments of shapes
      - Center of gravity
      - Object orientation
      - etc.

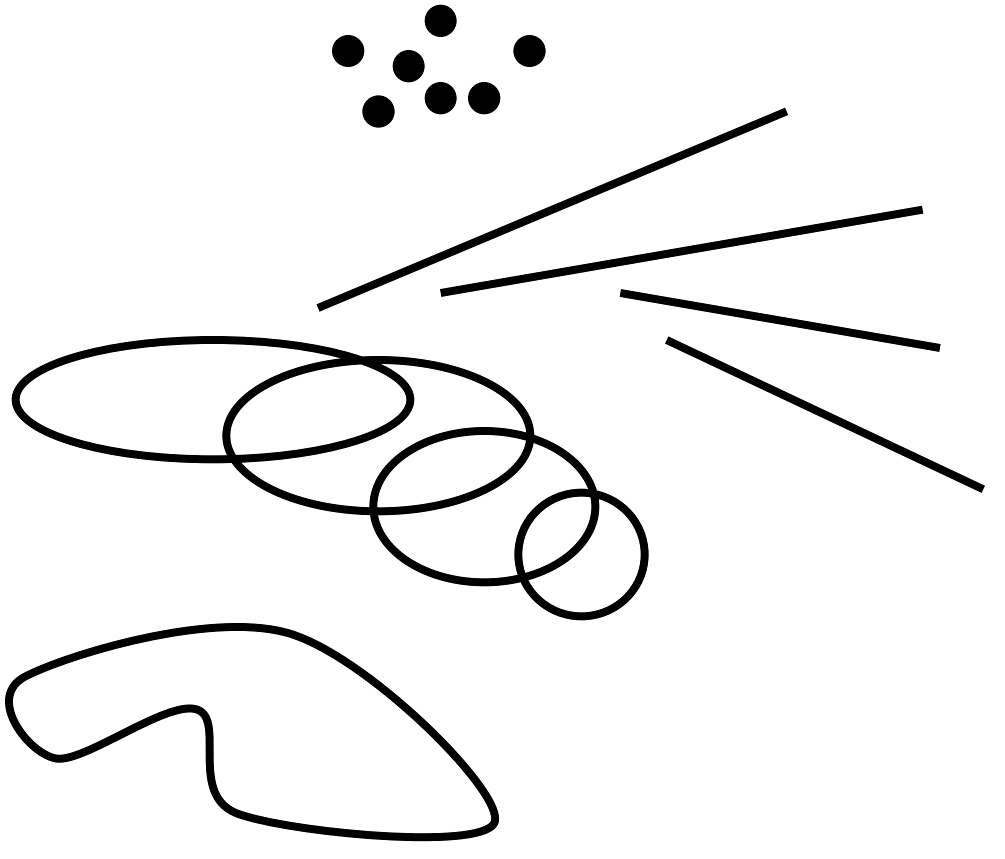


Image-Based Visual Servoing: Interaction matrix

- Computation of the interaction matrix $\dot{\mathbf{s}} = \mathbf{L} \cdot \mathbf{v} \rightarrow \mathbf{v} = -\lambda \cdot \hat{\mathbf{L}}^{+} \cdot \mathbf{e}$
  - Estimation in each control loop.
    - Computation using the analytical expression as a function of image features and depth estimation.

$$\hat{\mathbf{L}}(t) = \mathbf{L}(\mathbf{s}, \hat{Z})$$

      - Possibility of local minima and singularities
    - Computation of the interaction matrix at the target position. We use a constant Jacobian expression computed a priori using the features and depths at the target position.

$$\hat{\mathbf{L}} = \mathbf{L}(\mathbf{s}^{*}, \hat{Z}^{*}) = cte$$

      - Convergence is only ensured in the neighborhood of the target position.
  - Direct estimation
    - Estimation without taking into account its analytical expression.
    - Measure variations of **s** due to known camera motions

$$\Delta\mathbf{s} = \mathbf{L} \cdot \Delta\mathbf{v}$$

      - We have k equations and (kx6) unknowns. By performing N independent movements of the camera (N>6) it is possible to solve the system

Image-Based Visual Servoing: Control law

$$\dot{\mathbf{e}} = \mathbf{L} \cdot \mathbf{v}$$

- To achieve exponential error reduction we define:

$$\mathbf{v} = -\lambda \cdot \mathbf{L}^{+} \cdot \mathbf{e} \quad , \quad \lambda > 0$$

  - Where the Moore-Penrose pseudo inverse is used when $\mathbf{L}$ has rank 6.

$$\mathbf{L}^{+} = \left(\mathbf{L}^{T} \cdot \mathbf{L}\right)^{-1} \cdot \mathbf{L}^{T}$$

  - If k=6 and the determinant of $\mathbf{L}$ is nonzero

$$\mathbf{v} = -\lambda \cdot \mathbf{L}^{-1} \cdot \mathbf{e}$$

- In practice it is impossible to know perfectly the interaction matrix. It must be estimated: $\hat{\mathbf{L}}$
  - The control law remains in practice as: $\mathbf{v} = -\lambda \cdot \hat{\mathbf{L}}^{+} \cdot \mathbf{e}$
- Closed-loop behavior

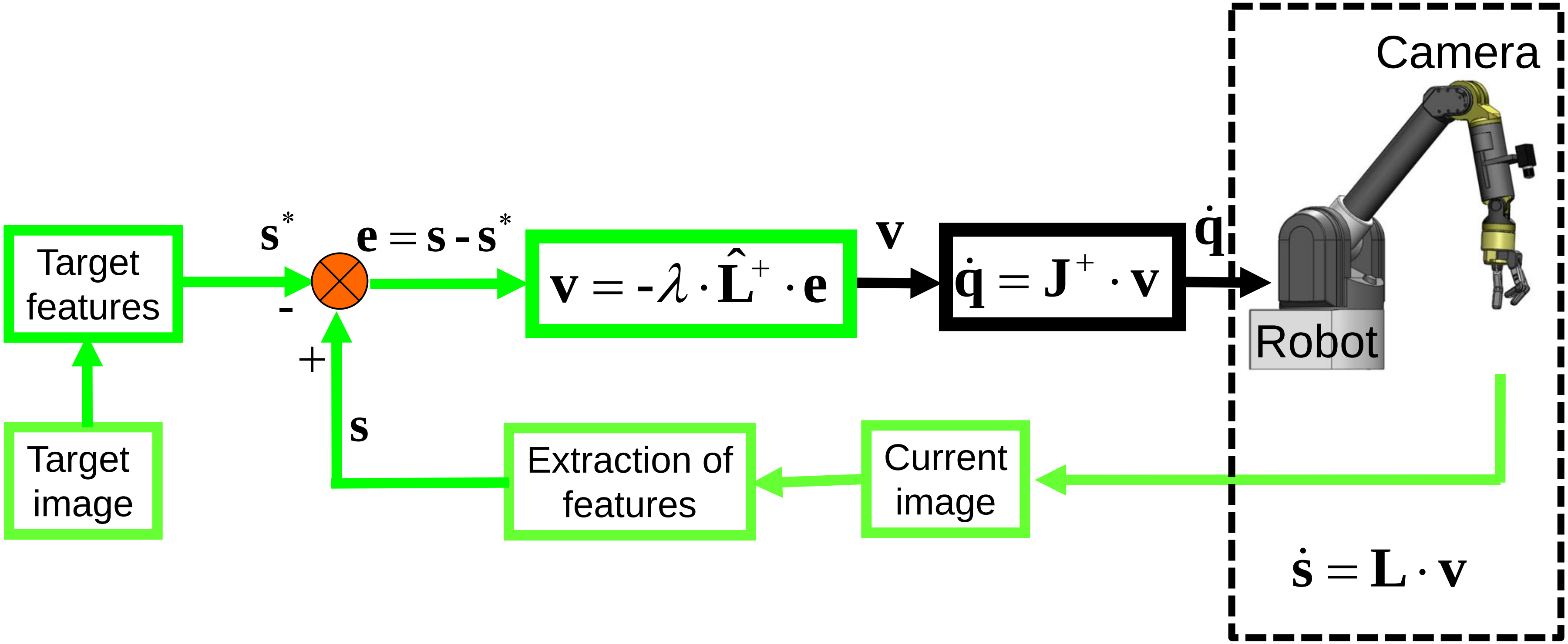
$$\dot{\mathbf{e}} = -\lambda \cdot \mathbf{L} \cdot \hat{\mathbf{L}}^{+} \cdot \mathbf{e}$$

  - Exponential error reduction if $\mathbf{L} \cdot \hat{\mathbf{L}}^{+} = \mathbf{I}_6 \Rightarrow \dot{\mathbf{e}} = -\lambda \cdot \mathbf{e}$
  - In any other case: Stability analysis...

$$\mathbf{L} \cdot \hat{\mathbf{L}}^{+} \neq \mathbf{I}_6 \Rightarrow ¿\,Stability\,?$$

Image-Based Visual Servoing: Control law

- **Image-based visual control loop**



Index

- Introduction
- Robotic manipulation
- Vision systems
- Classification of visual control systems
- Position-Based Visual Servoing
- Image-Based Visual Servoing

- **Stability analysis**
  - Lyapunov stability
  - Image-Based Visual Servoing stability

- Hybrid Visual Servoing
- Conclusion and bibliography