Uncalibrated Visual Servoing

Ezio MALIS

Institut National de Recherche en Informatique et Automatique
Sophia-Antipolis, France
http://www-sop.inria.fr/icare/personnel/malis
Topics

- Hybrid Visual Servoing;
- Invariant Visual Servoing;
Topics

- Hybrid Visual Servoing;
- Invariant Visual Servoing;
Introduction

- Position-based visual servoing:
  - 3D model of the object
  - Possible getting out of the image
    (3D Cartesian control)
  - Coarse calibration sufficient
    (but unknown robustness domain)

- Image-based visual servoing:
  - Depth estimation or approximation
  - Potential problems of convergence
  - Coarse calibration sufficient
    (but unknown robustness domain)
2\textsuperscript{1/2}D visual servoing

- Control in the 3D space: rotation axis \( u \) and angle \( \theta \)
- Control in the 2D space: coordinate of an image point \((u, v)\)
- Control of the relative depth: \( \log(Z/Z) \)
- Task function \( e = (e_\nu, e_\omega) \):
  \[
  e_\nu = (u - u, v - v, \log(Z/Z)) \quad \text{and} \quad e_\omega = \theta u
  \]
Epipolar Geometry

Reference points \( p \) with camera \( K \)

Fundamental equation:

\[
Z_p = G_1(p + 1Z_c) = K_1R > K_c = K_1R > K_t
\]

Why is epipolar geometry useful?

If \( R = I \) and \( K = K_0 \):

\[
Z_p = p + 1Z_c
\]

If \( K = K_0 \) then one can estimate \( R \).
Reference points $p$ with camera $K$

Current points $p$ with camera $K$

Fundamental equation:

$$Z_p = G_1 (p + 1 Z_c)$$

$$G_1 = K^{-1} R > K_c = K^{-1} t$$

Why is epipolar geometry useful?

If $R = I$ and $K = K$:

$$Z_p = p + 1 Z_c$$

If $K = K$ then one can estimate $R$. 

EURON Summer School on Visual servoing
Epipolar Geometry

- Reference points \( p \) with camera \( K \)
- Current points \( p \) with camera \( K \)
- Rotation \( R \) and translation \( t \)

Fundamental equation:

\[
Z_p = G_1 (p + Z_c)
\]

\[
G_1 = K^{-1} R > K c = K^{-1} t
\]

Why is epipolar geometry useful?

If \( R = I \) and \( K = K' \):

\[
Z_p = p + Z_c
\]

If \( K = K' \) then one can estimate \( R \).
Reference points \( p \) with camera \( K \)

Current points \( p \) with camera \( K \)

Rotation \( R \) and translation \( t \)

Fundamental equation:
\[
\frac{Z}{Z} p = G_{\infty}(p + \frac{1}{Z}c)
\]
Epipolar Geometry

- Reference points \( p \) with camera \( K \)
- Current points \( p \) with camera \( K \)
- Rotation \( R \) and translation \( t \)
- Fundamental equation:
  \[
  \frac{Z}{Z} \mathbf{p} = G_\infty (\mathbf{p} + \frac{1}{Z} \mathbf{c})
  \]
  \[
  G_\infty = K^{-1} R^\top K
  \]

Why is epipolar geometry useful?

If \( R = I \) and \( K = K \), then one can estimate \( R \).
Epipolar Geometry

- Reference points $p$ with camera $K$
- Current points $p$ with camera $K$
- Rotation $R$ and translation $t$
- Fundamental equation:
  \[ \frac{Z}{Z} p = G_\infty (p + \frac{1}{Z} c) \]
  \[ G_\infty = K^{-1} R^\top K \]
  \[ c = -K^{-1} t \]

Why is epipolar geometry useful?
If $R = I$ and $K = K$:
\[ Z p = p + Z c \]

If $K = K$ then one can estimate $R$.
Epipolar Geometry

- Reference points \( p \) with camera \( K \)
- Current points \( p \) with camera \( K \)
- Rotation \( R \) and translation \( t \)
- Fundamental equation:
  \[
  \frac{Z}{Z} p = G_{\infty} (p + \frac{1}{Z} c)
  \]
  \[
  G_{\infty} = K^{-1} R^T K
  \]
  \[
  c = -K^{-1} t
  \]
- Why is epipolar geometry useful?
Epipolar Geometry

- Reference points $p$ with camera $K$
- Current points $p$ with camera $K$
- Rotation $R$ and translation $t$
- Fundamental equation:
  \[ \frac{Z}{Z} p = G_\infty (p + \frac{1}{Z} c) \]
  \[ G_\infty = K^{-1} R^\top K \]
  \[ c = -K^{-1} t \]

Why is epipolar geometry useful?
- If $R = I$ and $K = K$:
  \[ \frac{Z}{Z} p = p + \frac{1}{Z} c \]
Epipolar Geometry

- Reference points $p$ with camera $K$
- Current points $p$ with camera $K$
- Rotation $R$ and translation $t$
- Fundamental equation:
  \[ \frac{Z}{Z} p = G_{\infty}(p + \frac{1}{Z} c) \]
  \[ G_{\infty} = K^{-1} R^T K \]
  \[ c = -K^{-1} t \]
- Why is epipolar geometry useful?
  - If $R = I$ and $K = K$:
    \[ \frac{Z}{Z} p = p + \frac{1}{Z} c \]
  - If $K = K$ then one can estimate $R$. 

EURON Summer School on Visual servoing
**Interaction matrix**

- Non-singular block triangular matrix:

\[
\frac{de}{dt} = \begin{bmatrix} e_v \\ e_\omega \end{bmatrix} = \begin{bmatrix} L_v & L_{(v, \omega)} \\ 0 & L_\omega \end{bmatrix} \begin{bmatrix} M_v & M_{(v, \omega)} \\ 0 & M_\omega \end{bmatrix} v
\]

- \( L \) depends on \( Z \) and \( K \) (\( \det(L) = 1/(Z^3 \sin^2(\theta/2)) \)):

\[
L_v = L_v(Z, K), \quad L_{(v, \omega)} = L_{(v, \omega)}(K)
\]

\[
L_\omega = I - \frac{\theta}{2} [u]_x + \left( 1 - \frac{\sin(\theta)}{\sin^2(\theta/2)} \right) [u]^2
\]

- \( M \) depends on \( ^cR_e \) and \( ^cT_e \) (\( \det(M) = 1 \)):

\[
M_v = ^cR_e, \quad M_{(v, \omega)} = [^cT_e]_x ^cR_e, \quad M_\omega = ^cR_e
\]
Control law

- The task function should decrease exponentially:

\[ \dot{e} = -\lambda e \]

- Control law:

\[
v = -\lambda \begin{bmatrix}
\widehat{M}_v^{-1} & -\widehat{M}_v^{-1}\widehat{M}_{(v,\omega)}\widehat{M}_\omega^{-1} \\
0 & \widehat{M}_\omega^{-1}
\end{bmatrix}
\begin{bmatrix}
\widehat{L}_v^{-1} & -\widehat{L}_v^{-1}\widehat{L}_{(v,\omega)}\widehat{L}_\omega^{-1} \\
0 & \widehat{L}_\omega^{-1}
\end{bmatrix} \widehat{e}
\]

- In practice, only estimations or approximations;
- Robustness with respect to calibration errors.
Stability Analysis

Closed-loop differential system:

\[ \dot{e} = -\lambda Q(e) e \]

Necessary and sufficient condition for the local stability:

\[ \text{real}(\text{eig}(Q(0))) > 0 \]

Sufficient condition for the global stability:

\[ Q(e) > 0 \]

Known robustness domain with respect to:
- errors on \( Z \) and \( K \) (\( \hat{M} = M \))
- errors on \( Z, K, cR_e, c_t_e \)
Experimental Results (I)

Coarsely Calibrated Camera

Reference image

Initial image

Error $\|p - p\|$
Experimental Results (II)

Uncalibrated Camera

Reference image

Initial image

Trajectory

Error $\|p - p\|$

Control law $\nu$

Translation $e_\nu$

Control law $\omega$

Rotation $e_\omega$
Topics

- Hybrid Visual Servoing;
- Invariant Visual Servoing;
Introduction

- Uncalibrated eye-in-hand visual servoing:
  - unknown target model
  - unknown camera parameters
- Proof of the local stability:
  - robustness to camera calibration errors
  - larger stability domain with path planning
- Extension of teaching-by-showing approach:
  - different cameras for learning and servoing
  - the camera zoom during servoing
Standard Visual Servoing

- **Model-based visual servoing:**

- **Model-free visual servoing:**
Teaching-by-showing

- Reference image with $K$
- Extract $n$ reference points $p$
Teaching-by-showing

- Reference image with $K$
- Extract $n$ reference points $p$
- Current image with $K$
- Camera positioned if $p = p$

EURON Summer School on Visual servoing
Teaching-by-showing

- Reference image with $K$
- Extract $n$ reference points $p$
- Current image with $K$
- Camera positioned if $p = p$
- Not positioned if $K \neq K$
- Standard approaches cannot be used
Teaching-by-showing

- Reference image with $K$
- Extract $n$ reference points $p$
- Current image with $K$
- Camera positioned if $p = p$
- Not positioned if $K \neq K$
- Standard approaches cannot be used
- How to build an error invariant to $K$?
Camera Model

\[ p = K(t) m, \quad K = \begin{bmatrix} f(t) & f(t)s & u_0(t) \\ 0 & f(t)r & v_0(t) \\ 0 & 0 & 1 \end{bmatrix} \]
Invariance: points
Invariance: lines
Invariance with respect to $K$ (I)

Compute the following symmetric matrices:

$$S_p = \frac{1}{n} \sum_{i=1}^{n} p_i p_i^\top \quad \text{and} \quad S_m = \frac{1}{n} \sum_{i=1}^{n} m_i m_i^\top$$

Since $p = K m$ the two matrices are related by:

$$S_p = K \left( \frac{1}{n} \sum_{i=1}^{n} m_i m_i^\top \right) K^\top = K S_m K^\top$$

Cholesky decomposition of the positive matrices:

$$S_p = T_p T_p^\top \quad \text{and} \quad S_m = T_m T_m^\top$$
The two upper triangular matrices are related by:

\[ T_p = K \, T_m \]

Compute a point in the invariant space:

\[ q = T_p^{-1} p = T_m^{-1} K^{-1} K \, m = T_m^{-1} m \]
The two upper triangular matrices are related by:

\[ T_p = K \ T_m \]

Compute a point in the invariant space:

\[ q = T_p^{-1} p = T_m^{-1} \ K^{-1} K m = T_m^{-1} m \]
Invariance with respect to $K$ (II)

- The two upper triangular matrices are related by:

$$T_p = K T_m$$

- Compute a point in the invariant space:

$$q = T_p^{-1} p = T_m^{-1} K^{-1} K m = T_m^{-1} m$$
Compute the following symmetric matrices:

\[ S_p = \frac{1}{n} \sum_{i=1}^{n} p_i p_i^\top \quad \text{and} \quad S_m = \frac{1}{n} \sum_{i=1}^{n} m_i m_i^\top \]

Since \( p = K m \) the two matrices are related by:

\[ S_p = K \left( \frac{1}{n} \sum_{i=1}^{n} m_i m_i^\top \right) K^\top = K S_m K^\top \]

Cholesky decomposition of the positive matrices:

\[ S_p = T_p \ T_p^\top \quad \text{and} \quad S_m = T_m \ T_m^\top \]
Invariance with respect to $K$ (II)

The two upper triangular matrices are related by:

$$T_p = K T_m$$

Compute a point in the invariant space:

$$q = T_p^{-1} p = T_m^{-1} K^{-1} K m = T_m^{-1} m$$
The two upper triangular matrices are related by:

\[ T_p = K \cdot T_m \]

Compute a point in the invariant space:

\[ q = T_p^{-1} p = T_m^{-1} K^{-1} K m = T_m^{-1} m \]
The two upper triangular matrices are related by:

\[ T_p = K T_m \]

Compute a point in the invariant space:

\[ q = T_p^{-1} p = T_m^{-1} K^{-1} K m = T_m^{-1} m \]
Invariant Visual Servoing

Unified visual servoing approach: \( q = T_p^{-1} p \rightarrow q \)

- Model-based: given \( m = \begin{bmatrix} R & t \end{bmatrix} X \) compute
  \[ q = T_m^{-1} m \]

- Model-free: given \( p \) compute
  \[ q = T_p^{-1} p \]
The interaction matrix

Starting from $q_i = T_p^{-1}p_i = T_m^{-1}m_i$ the derivative of $q_i$ is:

$$\dot{q}_i = \dot{T}_m^{-1}m_i + T_m^{-1}\dot{m}_i$$

Knowing that $\dot{T}_m^{-1} = -T_m^{-1}\dot{T}_m T_m^{-1}$, the derivative becomes:

$$\dot{q}_i = T_m^{-1}(m_i - \dot{T}_m q_i)$$

Finally, since $m_i = L_m v$ and $\dot{T}_m q_i = M_i v$:

$$\dot{q} = T_m^{-1}(L_{im} - M_i)v$$

and the interaction matrix is:

$$L_{iq} = T_m^{-1}(L_{im} - M_i)$$
Control Law

- Define the vector $s = (q_1, q_2, ..., q_n)$
- Compute the interaction matrix ($s = L_q v$):

$$L_q = \begin{bmatrix} \Phi(Z) & \Psi \end{bmatrix} \begin{bmatrix} F_1(K) & 0 \\ 0 & F_2(K) \end{bmatrix}$$

- Using the epipolar geometry: $\hat{Z} = \kappa(t)Z$, with $\kappa(t) > 0$
- Define the task function: $e = \hat{L}_q^+(s - s(t))$

$$\hat{L}_q = \begin{bmatrix} \Phi(\hat{Z}) & \Psi \end{bmatrix} \begin{bmatrix} F_1(\hat{K}) & 0 \\ 0 & F_2(\hat{K}) \end{bmatrix}$$

- The control law for imposing $\dot{e} = -\lambda e$ is: $v = -\lambda e + \hat{L}_q^+ \frac{\partial s(t)}{\partial t}$
Stability Analysis

Closed-loop equation during the path tracking:

\[ \dot{e} = -\lambda A(t) \varepsilon + b(t) \]

where:

\[ A(t) = \left. \hat{L}_q^+ L_q \right|_{s=s(t)} = \begin{bmatrix} \kappa(t) F_1^{-1}(\hat{K}) F_1(K) & 0 \\ 0 & F_2^{-1}(\hat{K}) F_2(K) \end{bmatrix} \]

\[ b(t) = \left. (\hat{L}_q^+ L_q - I) \hat{L}_q^+ \right|_{s=s(t)} \frac{\partial s(t)}{\partial t} \]

- The system is locally stable if and only if \( \kappa(t) > 0, \hat{f} > 0 \).
- The tracking error is bounded if \( \hat{K}^{-1} K(t) > 0 \) and \( b(t) \) is bounded.
Experimental Results (I)

Learning with $f = 12\text{mm}$, servoing with $f = 6\text{mm}$

Reference image

Trajectory

Control law $\nu$

Translation

Initial image

Error $\|q - q\|$ 

Control law $\omega$

Rotation

EURON Summer School on Visual servoing
Experimental Results (II)

Learning with $f = 12mm$, servoing with $f = 6mm$

Reference image  
Trajectory  
Control law $\nu$  
Translation

Initial image  
Error $\|q - q\|$  
Control law $\omega$  
Rotation

EURON Summer School on Visual servoing
Experimental Results (III)

Zooming camera and non-planar object

Reference image

Focal length

Control law $\nu$

Translation

Initial image

Error $\|q - q\|$}

Control law $\omega$

Rotation

EURON Summer School on Visual servoing
Experimental Results (III)

Zooming camera and non-planar object

Reference image

Final image

Control law $\nu$

Translation

Initial image

Error $\|q - q\|$  

Control law $\omega$

Rotation
Problems with planar objects

The same invariants are obtained from different positions:

$$\exists m_i \neq m_i : q_i = q_i \quad \forall i$$

Indeed, since the points are on a plane $$\exists G$$:

$$p_i = G p_i$$

Thus, if $$G$$ is an upper triangular matrix:

$$p_i = K m_i \Rightarrow q_i = f_i(m_1, m_2, \ldots, m_n)$$

$$p_i = G K m_i \Rightarrow q_i = f_i(m_1, m_2, \ldots, m_n)$$
Solution

At the convergence we need $p = p$: 

- principal task: $q \rightarrow q$
- secondary task: $T_p \rightarrow T_p$

The algorithm works in both cases:

- if the object is not planar:
  - the principal task solve the problem
  - the secondary task can be any
- if the object is not planar:
  - the principal task is not enough
  - the secondary task helps solving the problem
Experimental Results (IV)

Zooming camera and planar object

Reference image

Initial image

Focal length

Error $\|q - q^*\|$  
Control law $\nu$

Control law $\omega$

Translation

Rotation

EURON Summer School on Visual servoing
Experimental Results (IV)

Zooming camera and planar object

Reference image  
Final image  
Control law $\nu$  
Translation

Initial image  
Error $\|q - \hat{q}\|$  
Control law $\omega$  
Rotation

EURON Summer School on Visual servoing
Experimental Results (V)

Zooming camera and non-planar object

Reference image

Initial image

Error $||q - q||$

Control law $\nu$

Control law $\omega$

Rotation

Focal length

Translation

EURON Summer School on Visual servoing
Experimental Results (V)

Zooming camera and non-planar object

Reference image

Final image

Control law $\nu$

Translation

Initial image

Error $\|q - q\|$

Control law $\omega$

Rotation

EURON Summer School on Visual servoing
Conclusion

- New uncalibrated visual servoing approach;
- Extension of teaching-by-showing approach;
- The visual servoing is:
  - locally stable
  - robust to calibration errors
- Path tracking despite camera is uncalibrated:
  - larger stability domain
  - bounded tracking error