Wavelet and Shearlet-Based Image Representations for Visual Servoing

Rafael Reisenhofer\textsuperscript{1}, Lesley-Ann Duflot\textsuperscript{2,3}, Brahim Tamadazte\textsuperscript{3}, Nicolas Andreff\textsuperscript{3} & Alexandre Krupa\textsuperscript{2}

\textsuperscript{(1)} AG Computational Data Analysis, Universität Bremen
\textsuperscript{(2)} INRIA Rennes-Bretagne Atlantique
\textsuperscript{(3)} FEMTO-ST, AS2M, Univ. de Franche-Comté

NoKo 2017, Technische Universität Hamburg

5 May 2017
Outline

Visual Servoing
  Photometric Visual Servoing

Wavelet- and Shearlet-Based Visual Servoing

Experimental Validation
Some History

Determining Optical Flow

Berthold K.P. Horn and Brian G. Schunck
Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

ABSTRACT
Optical flow cannot be computed locally, since only one independent measurement is available from the image sequence at a point, while the flow velocity has two components. A second constraint is

ACKNOWLEDGMENT
This research was conducted at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract number N00014-75-C0643. One of the authors (Horn) would like to thank Professor H.-H. Nagel for his hospitality. The basic equations were conceived during a visit to the University of Hamburg, stimulated by Professor Nagel's long-standing interest in motion vision. The other author (Schunck) would like to thank W.E.L.
Visual Servoing

Video from Duflot et al. 2016
Visual Servoing

Basic setup:

▶ Use visual features extracted from visual sensors to control the motion of a robot in a closed-loop scheme.
▶ Minimize the error with respect to a set of desired visual features.

Two main approaches:

▶ Use high-level features such as detected objects and structures.
▶ Direct visual servoing: Only consider image intensities.
Photometric Visual Servoing - Basic Notation

- Image intensity in the image plane with respect to the robot pose $r(t) \in SE(3)$ at a time $t$:
  \[ I(x, y, t) \in \mathbb{R}. \]

- Feature vector for $N$ sampling points $(x_n, y_n)_{n \leq N} \subset \mathbb{R}^2$:
  \[ s_{ph}(t) = \left( I(x_1, y_1, t), I(x_2, y_2, t), \ldots, I(x_N, y_N, t) \right)^\top \in \mathbb{R}^N. \]

- Error regarding desired features $s_{ph}^* \in \mathbb{R}^N$:
  \[ e_{ph}(t) = s_{ph}(t) - s_{ph}^*. \]

- Six-dimensional velocity vector of the camera frame:
  \[ v(t) = (v_x(t), v_y(t), v_z(t), \omega_x(t), \omega_y(t), \omega_z(t))^\top. \]
1. Linearize the time-derivative of $s_{ph}$ in terms of $v$ and a so-called interaction matrix $L_{s_{ph}}(t)$:

$$\frac{ds_{ph}(t)}{dt} = L_{s_{ph}}(t)v(t).$$

2. Apply the Levenberg-Marquardt method to recover $v(t)$ at a time $t$ with $H(t) = L_{s_{ph}}(t)^T L_{s_{ph}}(t)$ and a positive gain parameter $\lambda$:

$$v(t) = -\lambda \left( (H(t) + \mu \text{diag}(H(t)))^{-1} L_{s_{ph}}(t)^T e_{ph}(t) \right),$$

where $\text{diag}(H(t))$ denotes the diagonal matrix given by the diagonal entries of $H(t)$ and $\mu$ is a damping parameter.
1. Linearize the time-derivative of $s_{ph}$ in terms of $v$ and a so-called interaction matrix $L_{s_{ph}}(t)$:

$$\frac{ds_{ph}(t)}{dt} = L_{s_{ph}}(t)v(t).$$

2. Apply the Levenberg-Marquardt method to recover $v(t)$ at a time $t$ with $H(t) = L_{s_{ph}}(t)^{T}L_{s_{ph}}(t)$ and a positive gain parameter $\lambda$:

$$v(t) = -\lambda \left((H(t) + \mu \text{diag}(H(t)))^{-1}L_{s_{ph}}(t)^{T}e_{ph}(t)\right),$$

where $\text{diag}(H(t))$ denotes the diagonal matrix given by the diagonal entries of $H(t)$ and $\mu$ is a damping parameter.
Photometric Visual Servoing - Computing $L_{sph}$

\[ \frac{ds_{ph}(t)}{dt} = L_{sph}(t)v(t) \]

1. Assuming temporal luminance constancy, we can write (Horn & Schunck, 1981):

\[ \frac{\partial I(x, y, t)}{\partial t} = -(\nabla_{xy}I(x, y, t))^T \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix}^T \]

2. Approximate velocities $\frac{dx}{dt}$ and $\frac{dy}{dt}$ using $v(t)$:

\[ \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix}^T \approx \begin{pmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -(1 + x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1 + y^2 & -xy & -x \end{pmatrix}^T v(t) \]

3. The interaction matrix is given by:

\[ L_{sph}(t) = - \begin{pmatrix} (\nabla_{xy}I(x_1, y_1, t))^T P(x_1, y_1) \\ \vdots \\ (\nabla_{xy}I(x_N, y_N, t))^T P(x_N, y_N) \end{pmatrix} \in \mathbb{R}^{N \times 6} \]

C. Collewet & E. Marchand; 2011
Wavelets

Wavelet systems in 2D can be constructed by considering generators $\psi \in L^2(\mathbb{R}^2)$ and the set of shifts and dilates

$$\Psi_\Gamma = \{ \psi_{a,m} = a^{1/2} \psi(a(\cdot - m)) : (a,m) \in \Gamma \},$$

where $\Gamma \subset \mathbb{R}^+ \times \mathbb{R}^2$ is a set of scaling and translation parameters.

The corresponding transform is defined for $f \in L^2(\mathbb{R}^2)$ as

$$\mathcal{W}_\psi(a,m) = \langle f, \psi_{a,m} \rangle_{L^2}.$$
Shearlets

Parabolic shearlet systems in 2D can be constructed by considering generators $\psi \in L^2(\mathbb{R}^2)$ and the set of shifts, dilates and shears

$$\Psi_\Gamma = \{ \psi_{a,s,m} = a^{3/2} \psi (S_s A_a (\cdot - t)) : (a, s, m) \in \Gamma \},$$

where $\Gamma \subset \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$ is a set of scaling and translation parameters, $S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ and $A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$.

The corresponding transform is defined for $f \in L^2(\mathbb{R}^2)$ as

$$S\mathcal{H}_\psi(a, s, m) = \langle f, \psi_{a,s,m} \rangle_{L^2}.$$
Wavelet-Based Feature Vector

- Image intensity in the image plane with respect to the robot pose $\mathbf{r}(t) \in SE(3)$ at a time $t$:

$$I(x, y, t) \in \mathbb{R}.$$

- Let $K \in \mathbb{N}$ the number of generators.

- Let $I_t := I(\cdot, \cdot, t)$.

- Feature vector for $N$ triples $$(k_n, a_n, m_n)_{n \leq N} \subset \{1, \ldots, K\} \times \mathbb{R}^+ \times \mathbb{Z}^2$$:

$$\mathbf{s_w}(t) = \left( \langle I_t, \psi^{(k_1)}_{a_1, m_1} \rangle_{L^2}, \langle I_t, \psi^{(k_2)}_{a_2, m_2} \rangle_{L^2}, \ldots, \langle I_t, \psi^{(k_N)}_{a_N, m_N} \rangle_{L^2} \right)^\top$$

$$\mathbf{s_{ph}}(t) = \left( I(x_1, y_1, t), I(x_2, y_2, t), \ldots, I(x_N, y_N, t) \right)^\top \in \mathbb{R}^N$$
Wavelet and Shearlet-Based Feature Vectors

photometric:

Subsampled wavelet transform

Non-subsampled shearlet transform
Wavelet-based Visual Servoing - Computing $\mathbf{L}_{sw}$

$$\frac{ds_{sw}(t)}{dt} = \mathbf{L}_{sw}(t)\mathbf{v}(t)$$

1. For continuously differentiable $\mathbf{I}$ with temporal luminance constancy, we can write:

$$\frac{d\langle I_t, \psi^{(k)}_{a,m} \rangle_{L^2}}{dt} = \int\int_{\mathbb{R}^2} \frac{\partial I(x,y,t)}{\partial t} \psi^{(k)}_{a,m}(x,y) \, dx \, dy$$

$$= - \int\int_{\mathbb{R}^2} (\nabla I_t(x,y))^\top \mathbf{P}(x,y) \mathbf{v}(t) \psi^{(k)}_{a,m}(x,y) \, dx \, dy$$

2. Denote: $I_t^{(i)}(x,y) = (\nabla I_t(x,y))^\top \mathbf{P}(x,y) \mathbf{e}_i$ for $i \leq 6$.

3. Wavelet-based interaction matrix:

$$\mathbf{L}_{sw}(t) = - \begin{pmatrix} \langle I_t^{(1)}, \psi^{(k_1)}_{a_1,m_1} \rangle_{L^2} & \cdots & \langle I_t^{(6)}, \psi^{(k_1)}_{a_1,m_1} \rangle_{L^2} \\ \vdots & \ddots & \vdots \\ \langle I_t^{(1)}, \psi^{(k_N)}_{a_N,m_N} \rangle_{L^2} & \cdots & \langle I_t^{(6)}, \psi^{(k_N)}_{a_N,m_N} \rangle_{L^2} \end{pmatrix} \in \mathbb{R}^{N \times 6}$$
Wavelet-based Visual Servoing

- Coarse approximations are noise-robust and associated to photometric visual servoing.

- Detail coefficients highlight features and are invariant to changes in illumination.

- Discrete wavelet and shearlet-based transforms can easily be computed with existing software libraries
  - MATLAB Wavelet Toolbox
  - FFST (Häuser & Steidl)
  - ShearLab (Kutyniok, Lim & R)
Experimental Validation

- Six degrees of freedom robotic system with CCD camera.
- Three conditions:
  - nominal
  - illumination
  - occlusion

- Two tests with different initial displacement for each condition:

<table>
<thead>
<tr>
<th></th>
<th>$\Delta T_x$ (mm)</th>
<th>$\Delta T_y$</th>
<th>$\Delta T_z$</th>
<th>$\Delta R_x$ (°)</th>
<th>$\Delta R_y$</th>
<th>$\Delta R_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>test 1</td>
<td>5</td>
<td>50</td>
<td>100</td>
<td>5</td>
<td>-5</td>
<td>-4</td>
</tr>
<tr>
<td>test 2</td>
<td>-20</td>
<td>-30</td>
<td>-20</td>
<td>10</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Experimental Results - Nominal Conditions

desired image: 

initial difference: 

wavelet:

shearlet:
Shearlet-Based Visual Servoing (Nominal Conditions)
Experimental Results - Partial Occlusion

desired image:

initial difference:

wavelet:

shearlet:
Wavelet-Based Visual Servoing (Partial Occlusions)

Video from Duflot et al. 2016
Summary

- Multiscale representations of images obtained from wavelet and shearlet transforms can be used to guide visual servoing tasks. The corresponding interaction matrix can be explicitly computed.

- Multiscale representations provide a natural compromise between intensity-based and feature-based visual servoing.

- First experiments suggest that emphasizing detail coefficients improves the performance under unfavorable illumination conditions.

- In contrast, focusing on the coarse approximation increases the stability regarding other types of noise and occlusions.

- Complex-valued wavelets and shearlets can provide smoother cost functions and maybe further improve the stability of the proposed scheme.
References


Shameless Advertising

Mathematical Signal Processing and Data Analysis

September 18-20, 2017
University of Bremen

Determining the characteristics of data is a crucial part of all data-driven science. Detecting the boundaries of a region of interest in an image, ascertaining trends in time-series data, and uncovering latent variables and the non-linear but dependent relationships between them are just some examples of problems in signal processing and data analysis. Often low complexity modelling, like sparsity/cosparsity assumptions, (non-)linear dimensionality reduction, manifold learning, etc. underlies the analysis. Other approaches are powered by variational methods, neural networks, Bayesian analysis, computational topology, and more.

The goal of this workshop is to bring together researchers of different aspects of mathematical signal processing and data analysis. A balance of theory, applications, and numerics will also be represented. To encourage younger researchers, there will be a poster session and, pending final funding, travel support.

The workshop is an official event of and supported by the DFG Research Training Group 2224 ´Parameter Identification - Analysis, Algorithms, Applications and also serves as the annual meeting of the GAMM Activity Group Mathematical Signal and Image Processing.
Thank you!