Estimating 3D Motion and Forces of Person-Object Interactions from Monocular Video

Zongmian Li $^{1,2}$, Jiri Sedlar $^{3}$, Justin Carpentier $^{1,2}$
Ivan Laptev $^{1,2}$, Nicolas Mansard $^{4}$, Josef Sivic $^{1,2,3}$

$^{1}$ Inria $^{2}$ DI ENS, PSL $^{3}$ CIIRC, CTU in Prague $^{4}$ LAAS-CNRS
Motivation: learn how to interact with objects

“Atlas” by Boston Dynamics
Goals

Input:
• A monocular RGB video
Method: a two-stage approach

Recognition stage

2D pose estimation

Input

[Cao et al., CVPR 2017]
Method: a two-stage approach
Method: a two-stage approach

Recognition stage

Contact recognizer

2D human pose

Input

[He et al., CVPR 2016]
Method: a two-stage approach
Method: a two-stage approach

Recognition stage

Object 2D endpoint estimation

2D human pose

Contact states

Input

[He et al., ICCV 2017]
Method: a two-stage approach
Method: a two-stage approach

Recognition stage
- 2D human pose
- Contact states

Estimation stage
Person-object trajectory optimization

[Tassa et al., IROS 2012]
[Carpentier et al., TRO, 2018]
Method: a two-stage approach

Recognition stage

Input

2D object endpoints

2D human pose

Contact states

Estimation stage

Person-object trajectory optimization

Estimated 3D motion and forces

[Tassa et al., IROS 2012]
[Carpentier et al., TRO, 2018]
Estimation stage

\[
\minimize_{x,u,c} \sum_{e \in \{h,o\}} \int_0^T l^e(x, u, c) \, dt
\]
Estimation stage

\[
\min_{\bar{x}, \bar{u}, \bar{c}} \sum_{e \in \{h, o\}} \int_{0}^{T} l^e(x, \bar{u}, c) \, dt
\]

Person-object

person-ground

contact forces
Estimation stage

\[
\text{minimize } \sum_{e \in \{h, o\}} \int_{0}^{T} l^e (x, u, c) \, dt
\]
Estimation stage

\[ \text{minimize} \sum_{e \in \{h,o\}} \int_{0}^{T} l^e (x, u, c) \, dt \]

Subject to

1. Contact motion model: \[ \left\| p_{\text{person joint}} - p_{\text{contact point}} \right\| = 0 \]
Estimation stage

\[
\begin{align*}
\text{minimize} & \quad \sum_{e \in \{h,o\}} \int_0^T l^e (x, u, c) \, dt \\
\text{Subject to:} & \\
2. \text{Contact force constraints:} & \\
\quad & \text{- Prevent the feet from sliding}
\end{align*}
\]
Estimation stage

\[
\min_{x,u,c} \sum_{e \in \{h,o\}} \int_{0}^{T} l^e(x, u, c) \, dt
\]

Subject to:

3. Lagrangian dynamics equation

\[
M(x) \ddot{x} + b(x, \dot{x}) = \tau(u, c)
\]

Experiments
A new dataset: Handtool videos

Barbell

Hammer

Scythe

Spade
At the poster (#202):

- Additional details
- Quantitative results
- Discussion

Zongmian Li $^{1,2}$, Jiri Sedlar $^3$, Justin Carpentier $^{1,2}$

Ivan Laptev $^{1,2}$, Nicolas Mansard $^4$, Josef Sivic $^{1,2,3}$

$^1$ Inria
$^2$ DI ENS, PSL
$^3$ CIIRC, CTU in Prague
$^4$ LAAS-CNRS