Galileo: Perceiving Physical Object Properties by Integrating a Physics Engine with Deep Learning

Jiajun Wu*  Ilker Yildirim*  Joseph Lim  Bill Freeman  Josh Tenenbaum

NIPS 2015

(* indicates equal contribution)
What can we learn from this video?
Understanding Dynamic 3D Scenes

- What can we learn from this video?
  - 3D object shapes
  - Physical events
  - Physical object properties
    - Masses
    - Coefficients of frictions

- Humans recover rich information from short videos.
- Generalization: Humans easily answer questions like
  - What will happen next?
A Simple Scenario in the Real World

• Galileo’s inclined surface experiment
Rich Physics in Simple Scenarios
Modeling the Physical World

Descriptive Physical Properties
- Velocity
- Acceleration
- Extended Distance
- Bounce Height

Intrinsic Physical Object Properties
- Coeff. Friction
- Material
- Coeff. Restitution
- Density
- Mass
- Volume

Videos
Our Model: Galileo

Physical object $i$
- Mass ($m$)
- Friction coefficient ($k$)
- 3D shape ($S$)
- Position offset ($x$)

Draw two physical objects

3D Physics engine

Simulated velocities ($v_{s1}, v_{s2}$)

Likelihood function

Observed velocities ($v_{o1}, v_{o2}$)

Tracking algorithm

...
Our Model: Galileo

Physical object $i$
- Mass ($m$)
- Friction coefficient ($k$)
- 3D shape ($S$)
- Position offset ($x$)

Draw two physical objects

3D Physics engine

Simulated velocities ($v_{s1}, v_{s2}$)

Likelihood function

Observed velocities ($v_{o1}, v_{o2}$)

Tracking algorithm
Our Model: Galileo

Physical object $i$
- Mass ($m$)
- Friction coefficient ($k$)
- 3D shape ($S$)
- Position offset ($x$)

Draw two physical objects

3D Physics engine

Simulated velocities $(v_{s1}, v_{s2})$

Likelihood function

Observed velocities $(v_{o1}, v_{o2})$

Tracking algorithm

...
Our Model: Galileo

Physical object $i$
- Mass ($m$)
- Friction coefficient ($k$)
- 3D shape ($S$)
- Position offset ($x$)

Draw two physical objects

3D Physics engine

Simulated velocities $(v_{s1}, v_{s2})$

Likelihood function

Observed velocities $(v_{o1}, v_{o2})$

Tracking algorithm
Our Model: Galileo

Physical object $i$
- Mass (m)
- Friction coefficient ($k$)
- 3D shape ($S$)
- Position offset ($x$)

Draw two physical objects

3D Physics engine

Simulated velocities ($v_{s1}, v_{s2}$)

Likelihood function

Observed velocities ($v_{o1}, v_{o2}$)

Tracking algorithm
Our Model: Galileo

Physical object $i$
- Mass ($m$)
- Friction coefficient ($k$)
- 3D shape ($S$)
- Position offset ($x$)

Sampling

Draw two physical objects

3D Physics engine

Simulated velocities $(v_{s1}, v_{s2})$

Likelihood function

Observed velocities $(v_{o1}, v_{o2})$

Tracking algorithm

\ldots
Results
Galileo as an Interpretation of Physical Scenes

Descriptive Physical Properties
- Velocity
- Acceleration
- Bounce Height
- Extended Distance

Intrinsic Physical Object Properties
- Coeff. Friction
- Coeff. Restitution
- Mass
- Density
- Volume

Physical object $i$
- Mass (m)
- Friction coefficient (k)
- 3D shape (S)
- Position offset (x)

Draw two physical objects

3D Physics engine

Simulated velocities $(v_{s1}, v_{s2})$

Likelihood function

Observed velocities $(v_{o1}, v_{o2})$

Tracking algorithm

Videos
Galileo as an Interpretation of Physical Scenes

**Descriptive Physical Properties**
- Velocity
- Acceleration
- Bounce Height
- Extended Distance

**Intrinsic Physical Object Properties**
- Coeff. Friction
- Coeff. Restitution
- Density
- Mass
- Volume

**Physical object i**
- Mass (m)
- Friction coefficient (k)
- 3D shape (S)
- Position offset (x)

**Videos**

**3D Physics engine**
- Simulated velocities \( (v_{s1}, v_{s2}) \)
- Observed velocities \( (v_{o1}, v_{o2}) \)

**Likelihood function**

**Tracking algorithm**
...
Galileo as an Interpretation of Physical Scenes

Descriptive Physical Properties
- Velocity
- Acceleration
  - Bounce Height
  - Extended Distance

Intrinsic Physical Object Properties
- Coeff. Friction
- Coeff. Restitution
- Density
- Mass
- Volume

Videos

Physical object $i$
- Mass (m)
- Friction coefficient (k)
- 3D shape (S)
- Position offset (x)

Draw two physical objects

3D Physics engine

Simulated velocities $(v_{s1}, v_{s2})$

Likelihood function

Observed velocities $(v_{o1}, v_{o2})$

Tracking algorithm

...
Galileo as an Interpretation of Physical Scenes

Descriptive Physical Properties
- Velocity
- Acceleration
  - Bounce
  - Height
  - Extended Distance

Intrinsic Physical Object Properties
- Coeff. Friction
- Coeff. Restitution
- Density
- Mass
- Volume

Physical object i
- Mass (m)
- Friction coefficient (k)
- 3D shape (S)
- Position offset (x)

Draw two physical objects

3D Physics engine

Simulated velocities \( (v_{s1}, v_{s2}) \)

Observed velocities \( (v_{o1}, v_{o2}) \)

Tracking algorithm

Videos
Generative + Recognition Model

If the model has prior knowledge like humans do...

Initialize MCMC

Cropped image

Log Likelihood

Number of MCMC sweeps
Behavior Tests

- Prediction task
  - How far will the object travel after collision?

```
<table>
<thead>
<tr>
<th>Material</th>
<th>Human Mean</th>
<th>Galileo Mean</th>
<th>Uniform Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>cardboard</td>
<td>75.3</td>
<td>71.9</td>
<td>106.4</td>
</tr>
<tr>
<td>dough</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hollow wood</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>metal coin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>metal pole</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>plastic block</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>plastic doll</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>plastic toy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>porcelain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wooden block</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wooden pole</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Behavior Tests

- Comprehension task
  - Which of the two objects is heavier?

- Transfer task
  - Will the object slide shown in a new scene setting?

<table>
<thead>
<tr>
<th></th>
<th>Comprehension</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human vs. Galileo</td>
<td>0.51</td>
<td>0.56</td>
</tr>
<tr>
<td>Human vs. Truth</td>
<td>0.68</td>
<td>0.42</td>
</tr>
<tr>
<td>Galileo vs. Truth</td>
<td>0.52</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Correlation coefficients between pairs of responses