Two-Finger Gestures for 6DOF Manipulation of 3D Objects

Jingbo Liu
Oscar Kin-Chung Au
Hongbo Fu
Chiew-Lan Tai
6DOF Manipulation

• 3 independent translation along x, y, z axes
• 3 independent rotation about x, y, z axes
Multitouch 6DOF Manipulation

- Each touch point provides 2D input
- No straightforward mapping
- Existing methods require 3 touch points
Direct Manipulation

- Touch points need to be on the manipulating object
- Corresponding object points always under touching fingers
**Indirect Manipulation**

- Touch points need **NOT** be on the manipulating object
- Fingers can be placed **anywhere** on the screen
Our Design Rationale

Two-finger operations
• Hardware limitation
• Reduce occlusion
• Easy-to-use
• Not using one-finger operation

Unimanual interaction
• Suitable for mobile and small devices

Independent of fingers’ directness
• Suitable for small screens and small objects

Seamless operations
• No explicit switching between different DOFs
Existing Methods

Screen-Space Manipulation [Reisnam et al. 2009]

- Require 3 touch fingers
- **Direct** manipulation
- Hard to operate for small object / small screen
  - Hand occlusion & small operation area
Existing Methods

Screen-Space Manipulation [Reisnam et al. 2009]

- Require 3 touch fingers
- **Direct** manipulation
- Hard to operate for small object / small screen
  - Hand occlusion & small operation area
Existing Methods

Sticky Tool [Hancock et al. 2009]

- Require 3 touch fingers
- 2 direct & 1 indirect finger
- Depend on directness of fingers
- Bimanual interaction is demanded
Existing Methods

Sticky Tool [Hancock et al. 2009]

- Require 3 touch fingers
- 2 direct & 1 indirect finger
- Depend on directness of fingers
- Bimanual interaction is demanded
## Existing Methods

**DS3** [Martinet et al. 2012]

- Require 3 touch fingers
- Separate translation from rotation
- Depend on **directness** of fingers
- Bimanual interaction is demanded

<table>
<thead>
<tr>
<th>Method</th>
<th>DS3</th>
<th>1d</th>
<th>1d+1i</th>
<th>≥2d</th>
<th>≥2d+1i</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Translation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T_x</td>
<td></td>
<td><img src="image1" alt="Red dot" /></td>
<td><img src="image2" alt="Red dot" /></td>
<td><img src="image3" alt="Red dot" /></td>
<td><img src="image4" alt="Red dot" /></td>
</tr>
<tr>
<td>T_y</td>
<td></td>
<td><img src="image1" alt="Red dot" /></td>
<td><img src="image2" alt="Red dot" /></td>
<td><img src="image3" alt="Red dot" /></td>
<td><img src="image4" alt="Red dot" /></td>
</tr>
<tr>
<td>T_z</td>
<td></td>
<td><img src="image1" alt="Red dot" /></td>
<td><img src="image2" alt="Red dot" /></td>
<td>♀</td>
<td>♀</td>
</tr>
<tr>
<td><strong>Rotation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R_x</td>
<td></td>
<td><img src="image1" alt="Red dot" /></td>
<td><img src="image2" alt="Red dot" /></td>
<td><img src="image3" alt="Red dot" /></td>
<td><img src="image4" alt="Red dot" /></td>
</tr>
<tr>
<td>R_y</td>
<td></td>
<td><img src="image1" alt="Red dot" /></td>
<td><img src="image2" alt="Red dot" /></td>
<td><img src="image3" alt="Red dot" /></td>
<td><img src="image4" alt="Red dot" /></td>
</tr>
<tr>
<td>R_z</td>
<td></td>
<td><img src="image1" alt="Red dot" /></td>
<td><img src="image2" alt="Red dot" /></td>
<td><img src="image3" alt="Red dot" /></td>
<td><img src="image4" alt="Red dot" /></td>
</tr>
</tbody>
</table>

T_x, T_y OR R_x, R_y, R_z
Existing Methods

**DS3** [Martinet et al. 2012]

- Require **3** touch fingers
- Separate translation from rotation
- Depend on *directness* of fingers
- Bimanual interaction is demanded
Our Approach

Only two fingers are needed for all 6DOF operations
Based on moving characteristics of 2 fingers

Key Idea: Two operation modes

**Mode 2m**
- 2 moving fingers

**Mode 1m + 1f**
- 1 moving finger and 1 fixed finger
Mode 2m

- involves 2 moving fingers
- controls 4 DOF by an RST style gesture
Mode 2m

Panning for xy-translation

<table>
<thead>
<tr>
<th>Method</th>
<th>Our Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOF</td>
<td>2m</td>
</tr>
<tr>
<td>Translation</td>
<td></td>
</tr>
<tr>
<td>$T_x$</td>
<td>![Image]</td>
</tr>
<tr>
<td>$T_y$</td>
<td>![Image]</td>
</tr>
<tr>
<td>$T_z$</td>
<td>![Image]</td>
</tr>
<tr>
<td>Rotation</td>
<td></td>
</tr>
<tr>
<td>$R_x$</td>
<td>![Image]</td>
</tr>
<tr>
<td>$R_y$</td>
<td>![Image]</td>
</tr>
<tr>
<td>$R_z$</td>
<td>![Image]</td>
</tr>
</tbody>
</table>
Mode 2m

Pinching for z-translation
**Mode 2m**

**Swiveling for z-rotation**

<table>
<thead>
<tr>
<th>Method DOF</th>
<th>Our Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2m</td>
</tr>
<tr>
<td>Translation</td>
<td></td>
</tr>
<tr>
<td>T_x</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td>T_y</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td>T_z</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td>Rotation</td>
<td></td>
</tr>
<tr>
<td>R_x</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td>R_y</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td>R_z</td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Mode 2m

Integral RST-style gesture
## Mode 1m + 1f

- involve 1 moving and 1 fixed finger
- **Pin-panning gesture**
- control remaining 2 DOF

<table>
<thead>
<tr>
<th>Method</th>
<th>Our Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOF</td>
<td>2m</td>
</tr>
<tr>
<td>Translation</td>
<td></td>
</tr>
<tr>
<td>$T_X$</td>
<td></td>
</tr>
<tr>
<td>$T_Y$</td>
<td></td>
</tr>
<tr>
<td>$T_Z$</td>
<td></td>
</tr>
<tr>
<td>Rotation</td>
<td></td>
</tr>
<tr>
<td>$R_X$</td>
<td></td>
</tr>
<tr>
<td>$R_Y$</td>
<td></td>
</tr>
<tr>
<td>$R_Z$</td>
<td></td>
</tr>
</tbody>
</table>
Mode 1m + 1f

Pin-panning for xy-rotation
Main Features

**Seamless transition** between the two modes
Main Features

Independent of the directness of fingers
**Issue: Mode Classification**

Distinguish *mode 2m* and *mode 1m+1f*

Cannot rely on *finger immobility*
- Difficult to keep finger completely fixed
- Imprecise hardware input
Mode Classification

Our approach – Learning-based recognition

Define **feature vector as gesture descriptor**

- Speeds of touch points
- Magnitude of centripetal acceleration

Learning **boundaries of 2 modes** in descriptor space
Mode Classification

Our approach – Learning-based recognition

Adopt Support Vector Machines (SVM) classifier

Involve 12 users with 5000 training samples
Validation

High cross validation accuracy 96.03%
High classification accuracy 91.4%
User Study

3D docking task

Compare with the state-of-art interfaces
  • sticky tools
  • screen-space manipulation
  • DS3

Tested on
  • 2 screen sizes
  • 3 object sizes
  • 2 difficulty levels
Result – Completion Time

- Comparable to state-of-art 3 finger techniques
- Insensitive to screen size
- High performance for complex task
Result – Completion Time

- Comparable to state-of-art 3 finger techniques
-Insensitive to object size
Conclusion

• First time introduce a single-hand, two-finger multitouch technique for 6DOF manipulation of 3D objects

• Independent of fingers’ directness

• Seamless operations

• Support different sized multitouch screens

• Comparable to state-of-art 3 finger techniques