Touchscreen Everywhere: On Transferring a Normal Planar Surface to a Touch-Sensitive Display

Jingwen Dai
Postdoctoral Research Associate
University of North Carolina at Chapel Hill
dai@cs.unc.edu
Motivation

Bigger Display VS. Portability
Motivation

DLP Pico Projector
Single Image → 3D Action

~500mm

~400mm
Previews Works

- **Additional Sensors**
  - Light Touch (*IR optical sensors*)
  - Diamondtouch (*capacitive sensor array*)
  - Smartskin (*mesh-shaped antenna*)
  - Skinput (*bio-acoustic sensing array*)
  - LightSpace, Omnitouch (*Kinect*)

- **Computer Vision**
  - [Letessier2004] -- *Fingertip tracking, not touching detection*
  - [Kjeldsen2002, Hardenberg2001] -- *Delay-based scheme*
  - [Marshall2008] -- *Color change of the fingernail*
  - [Song2007, PlayAnywhere2005] -- *Shadow casted by finger*
  - [Fitriani2007] -- *Deformation on soft surface*
Overview

Imperceptible Structured Light

- Texture Image
- Subtraction Image
- Embedded Pattern Image

Hand Segmentation → Fingertip Detection → Touch Detection

Radiometric

Geometric

Priors in Pro-Cams
Priors in Projector-Camera System

- Geometric (Homography)

All these geometric priors can be derived through **2 projection-capture cycles** in initialization stage.
Embedding Codes into Video Projection


\[ I_i(x, y) = O_i(x, y) + P(x, y), \]
\[ I'_i(x, y) = O_i(x, y) - P(x, y), \]
\[ P(x, y) = \begin{cases} \Delta, & \text{when } B(x, y) = 1; \\ 0, & \text{when } B(x, y) = 0. \end{cases} \]

\[ S(x, y) = \max_i[C_i(x, y) - C'_i(x, y)], \quad i = \{R, G, B\}. \]
Embedded Pattern Design Strategy

<table>
<thead>
<tr>
<th>Method</th>
<th>Array Size</th>
<th>Win. Size</th>
<th>Alph. Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Morita 1988]</td>
<td>24 * 24</td>
<td>3 * 4</td>
<td>2</td>
</tr>
<tr>
<td>[Kiyasu 1995]</td>
<td>18 * 18</td>
<td>4 * 2</td>
<td>2</td>
</tr>
<tr>
<td>[Salvi 1998]</td>
<td>29 * 29</td>
<td>3 * 3</td>
<td>3</td>
</tr>
<tr>
<td>[Spoelder 2000]</td>
<td>65 * 63</td>
<td>2 * 3</td>
<td>2</td>
</tr>
<tr>
<td>[Albitar 2007]</td>
<td>27 * 29</td>
<td>3 * 3</td>
<td>3</td>
</tr>
<tr>
<td>[Desjardins 2007]</td>
<td>53 * 38</td>
<td>3 * 3</td>
<td>3</td>
</tr>
<tr>
<td>[Chen 2008]</td>
<td>82 * 82</td>
<td>3 * 3</td>
<td>7</td>
</tr>
</tbody>
</table>

Summary of typical spatial coding methods

**Constraints of Pattern Generation**

- Code Uniqueness
- Large Hamming Distance
Hand Segmentation

Contrast Saliency + Region Discontinuity
Histogram-based Contrast Saliency

Mean-Shift Region Smoothing

Precise Segmentation by Fusing

\[ C_F > \Delta \]

\[ C_F(k) = \frac{1}{\epsilon(L-1)} [\alpha \bar{S}(k) + \beta \bar{S}_N(k) + \gamma A(k)] \]
Fingertip Detection

- Binary Hand Image
- Hand Contour
- Fingertip Candidates
- Detected Fingertips
Touch Detection Through Homography and Embedded Codes

Projector Projection Plane
Binary Code: 100...000

Camera Image Plane
Binary Code: 100...000

Subtraction Image ($I_s$)

2D Texture Image ($I_c$)

Third-Person Perspective
From Resistive Touch to Capacitive Touch or Floating Touch

Homology
Experiments – System Setup

(a) (b)

- HDMI Power
- IEEE1394
- Tri. Sig.
- Syn. Sig.
- Proj.
- Cam.
Experiments – *System Initialization*

- Camera-Projector Homography Estimation

Projected chessboard

Captured image
Experiments – *System Initialization*

- **Camera-Table Homography Estimation**

Four Correspondences: $C_1 \sim O$, $C_2 \sim Px$, $C_3 \sim Py$, $C_4 \sim Pxy$
Experiments – *Display Quality Evaluation*

![Graph showing scores for flickering, image deterioration, and user satisfaction across different embedded intensities.](image-url)
Experiments – Hand Segmentation Accuracy

(a) Original Image
(b) Ground-truth
(c) Our Method
(d) SCM
(e) BkSub
(f) GB
Experiments – *Hand Segmentation Accuracy Evaluation*

![Graph showing hand segmentation accuracy evaluation for different methods: Our Method, SCM, BkSub, GB. The graph compares precision, recall, and F-Beta scores. Our Method consistently demonstrates the highest scores across all metrics.](image-url)
Experiments – *Touch Accuracy Evaluation*

[Image of experiment setup showing camera and third person views of a grid with fingertip trajectories marked.]

*[More HD videos](http://cs.unc.edu/~dai/Research/TouchEverywhere/TouchEverywhere.htm)*
Experiments – *Touch Accuracy Evaluation*

<table>
<thead>
<tr>
<th>Surface</th>
<th>Dark $\epsilon$(px)</th>
<th>Dark FRR/FAR(%)</th>
<th>Normal $\epsilon$(px)</th>
<th>Normal FRR/FAR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray</td>
<td>2.98</td>
<td>1.12/0.45</td>
<td>3.05</td>
<td>1.32/0.48</td>
</tr>
<tr>
<td>Yellow</td>
<td>3.04</td>
<td>1.23/0.57</td>
<td>3.12</td>
<td>1.54/0.61</td>
</tr>
<tr>
<td>Artifact</td>
<td>3.12</td>
<td>1.77/0.67</td>
<td>3.20</td>
<td>1.76/0.63</td>
</tr>
</tbody>
</table>

*Comparison with recent depth-camera sensing based methods*

In [Wilson 2010], the informal observed spatial error of finger detection on planar surface was *between 3-6 pixels*.

In [Omni-Touch 2011], the FRR and FAR of finger click detection on four different surfaces were *0.8% and 3.3%*. 
Experiments – Trajectory Tracking Evaluation
Experiments – *Multi-Touch Evaluation*
## Experiments – Efficiency Evaluation

<table>
<thead>
<tr>
<th>Subroutine</th>
<th>Hand Seg.</th>
<th>FTip Loc.</th>
<th>Touch Det.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (ms/frame)</td>
<td>14.63</td>
<td>1.32</td>
<td>1.74</td>
<td>17.69</td>
</tr>
</tbody>
</table>
Summary

- Using only off-the-shelf devices
- Precise hand segmentation in PROCAMS
- Achieving 3D sensing without explicit 3D reconstruction
Related Publications

Conference paper


Journal paper


http://cs.unc.edu/~dai/Publications/Publications.htm
THANK YOU!!