Making any planar surface into a touch-sensitive display by a mere projector and camera

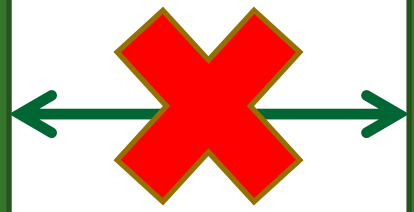
Jingwen Dai *Ronald Chung*
Computer Vision Laboratory
Dept. of Mech. and Automation Engineering
The Chinese University of Hong Kong

PROCAMS2012, Providence, RI, 17 June 2012

Introduction & Motivation

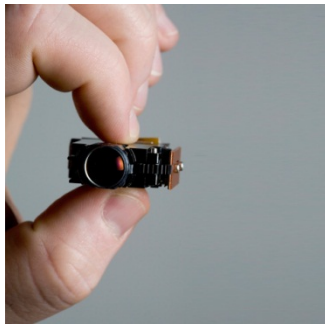
VS.

Bigger Display



Portability

Introduction & Motivation



DLP Pico Projector

Mobile Phone

DC

DV

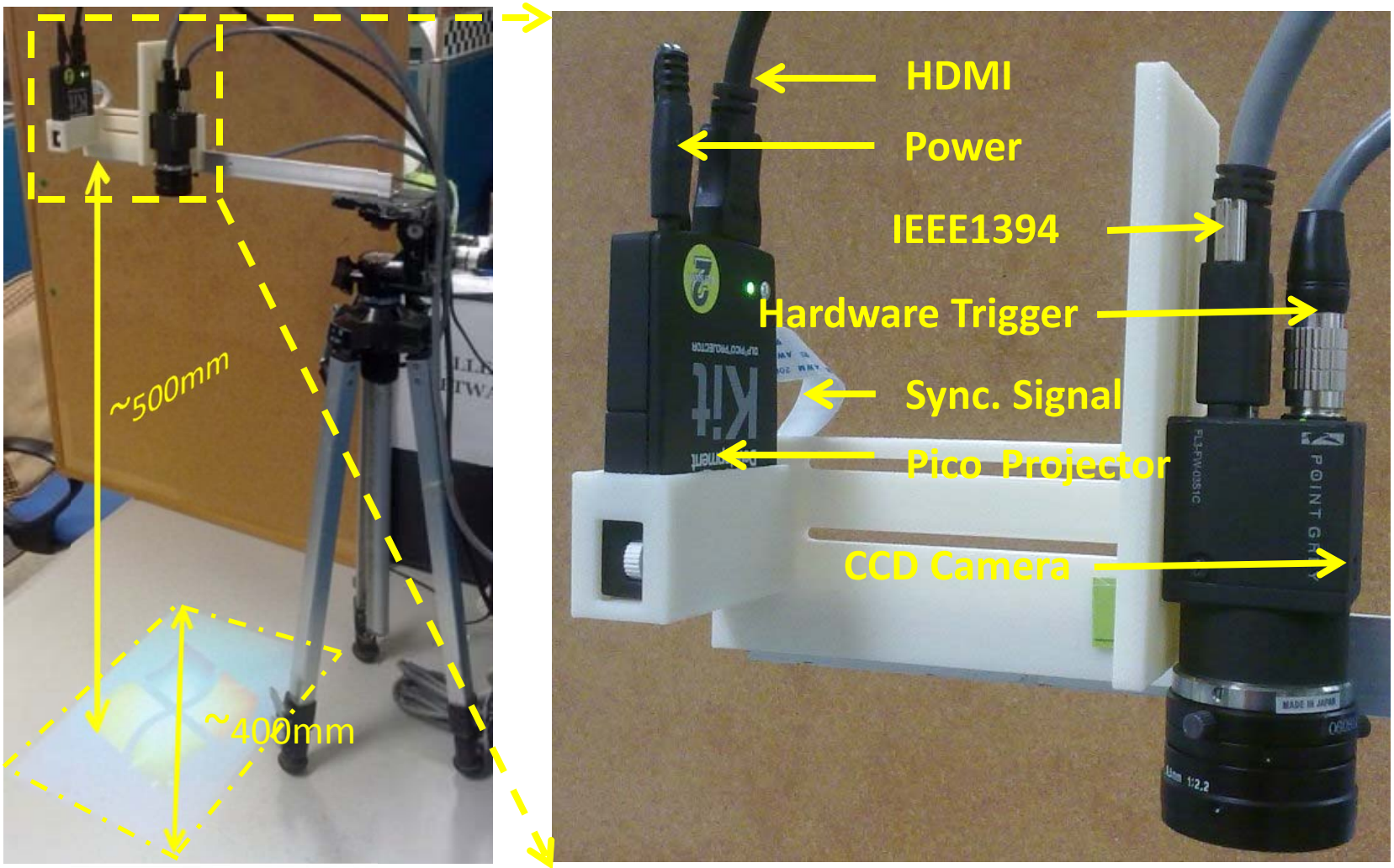
Previews Works

- **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

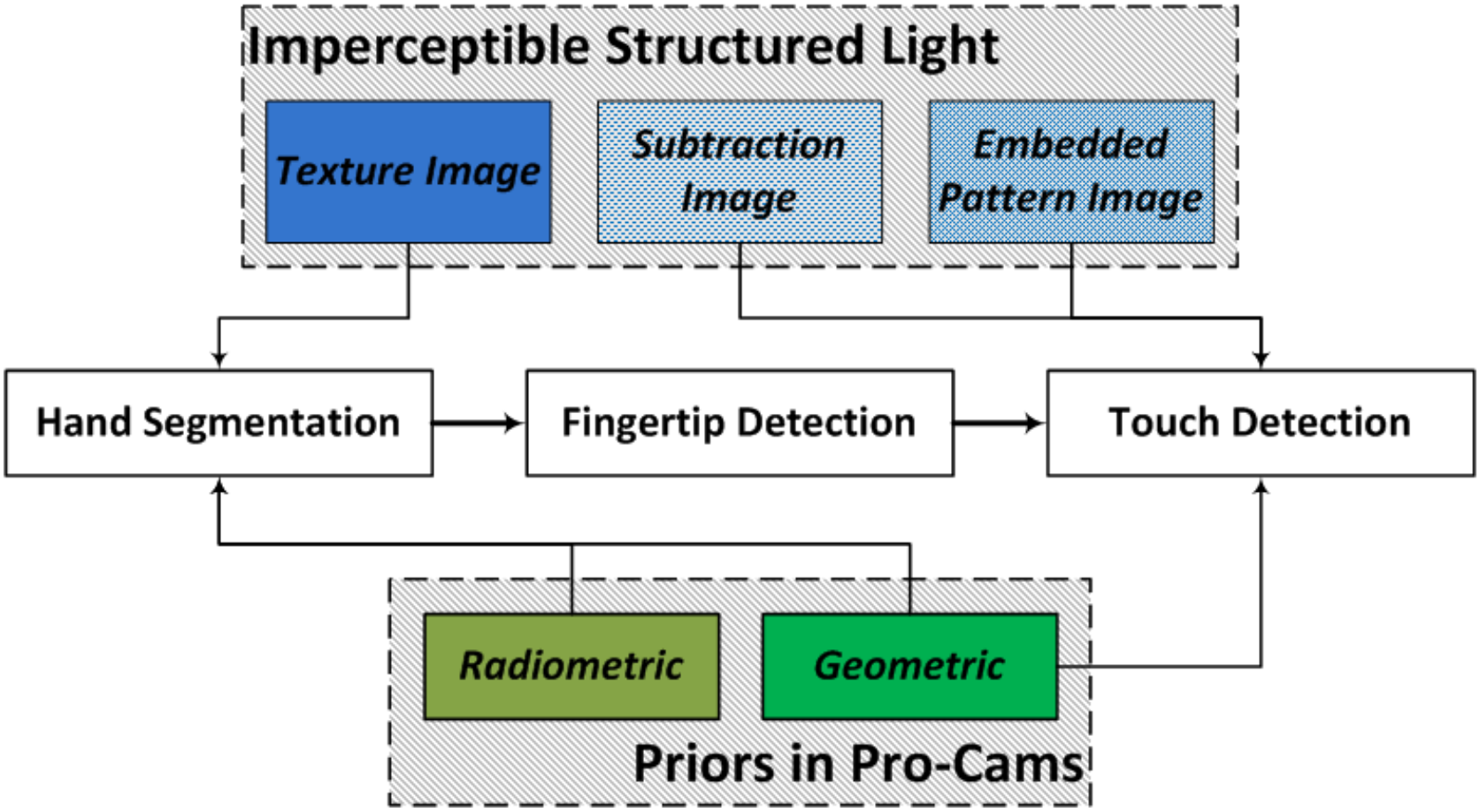
Main Contributions

- **Using only off-the-shelf devices**
- **Achieving 3D sensing without explicit 3D reconstruction**
- **Use of prior knowledge to enhance robustness**

System Prototype

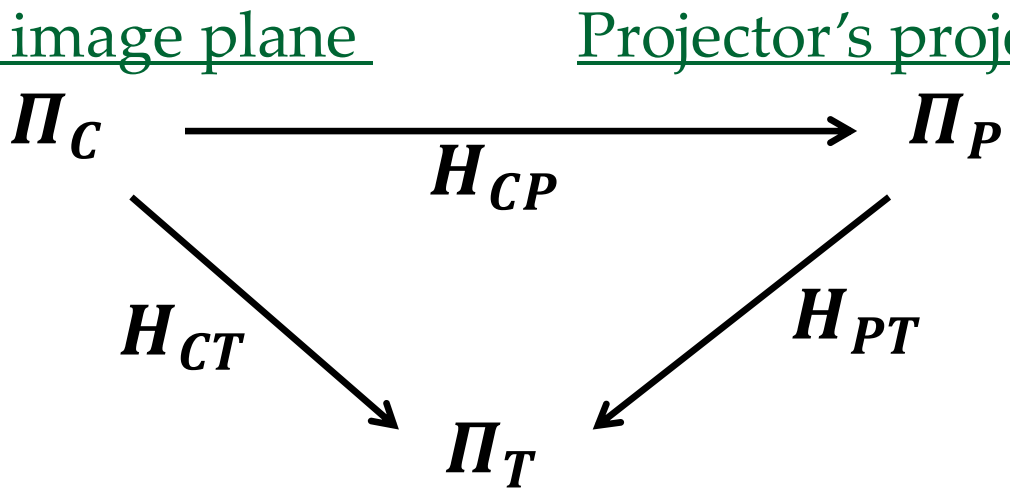


Overview


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

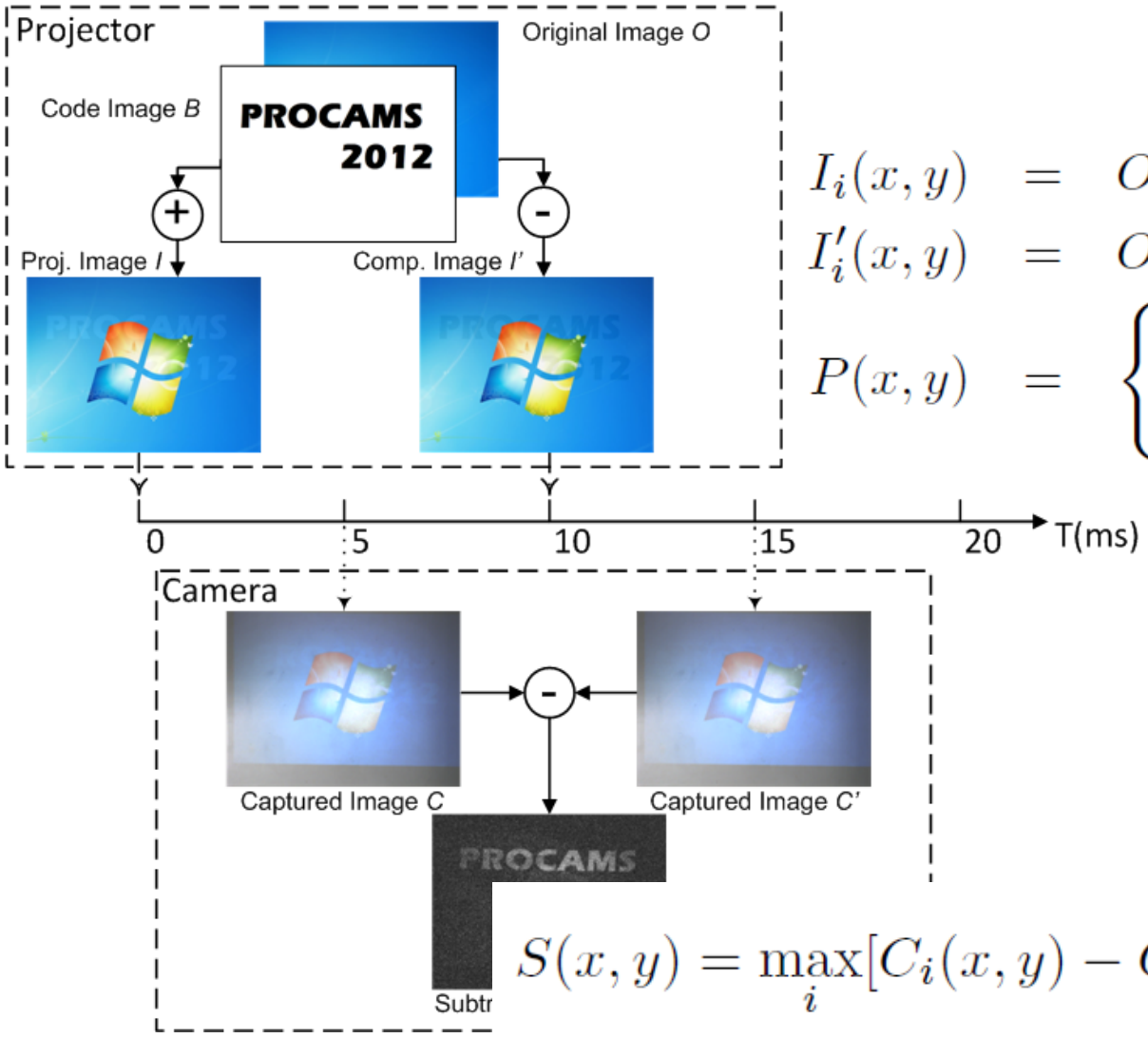
- **Geometric (Homography)**



- **Radiometric**

$$C_{pre} = VP + C$$

Embedding Codes into Video Projection



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

$$S(x,y) = \max_i [C_i(x,y) - C'_i(x,y)], i = \{R,G,B\}.$$

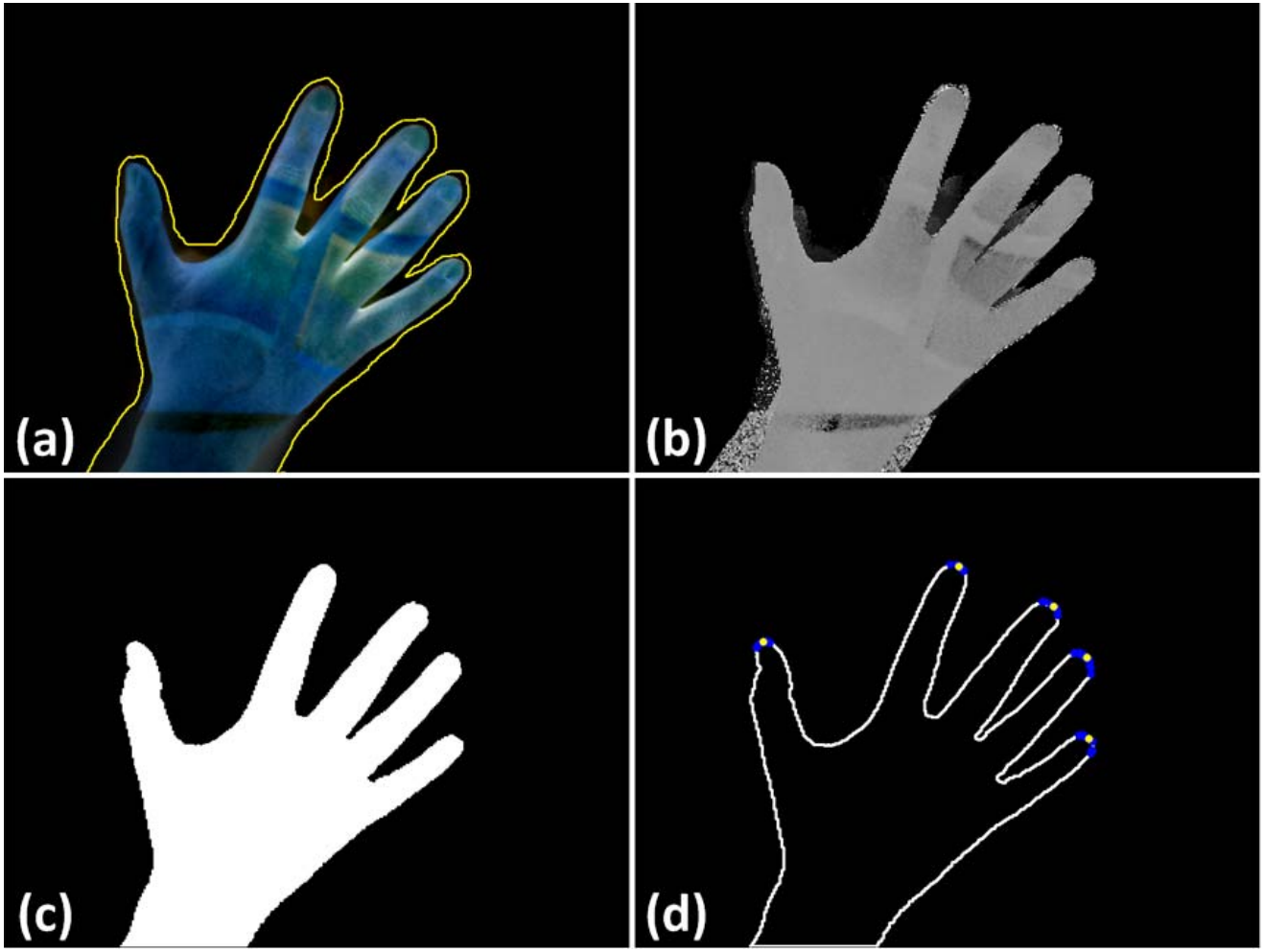
Embedded Pattern Design Strategy

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

Summary of typical spatial coding methods

- **Constraints of Pattern Generation**
  - Code Uniqueness
  - Large Hamming Distance

Hand Segmentation & Fingertip Detection



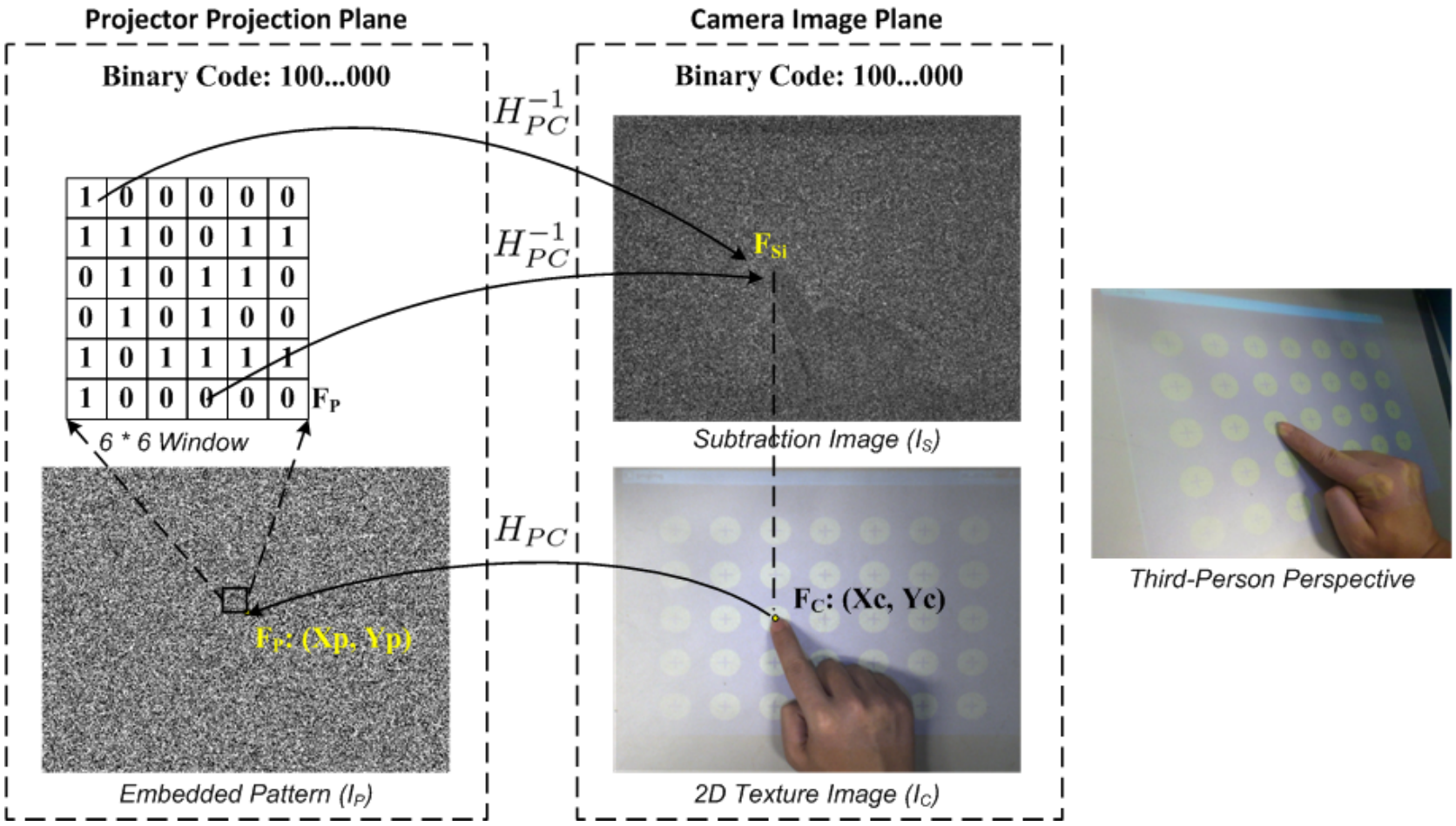
(a) Approximate segmentation

(b) H-channel

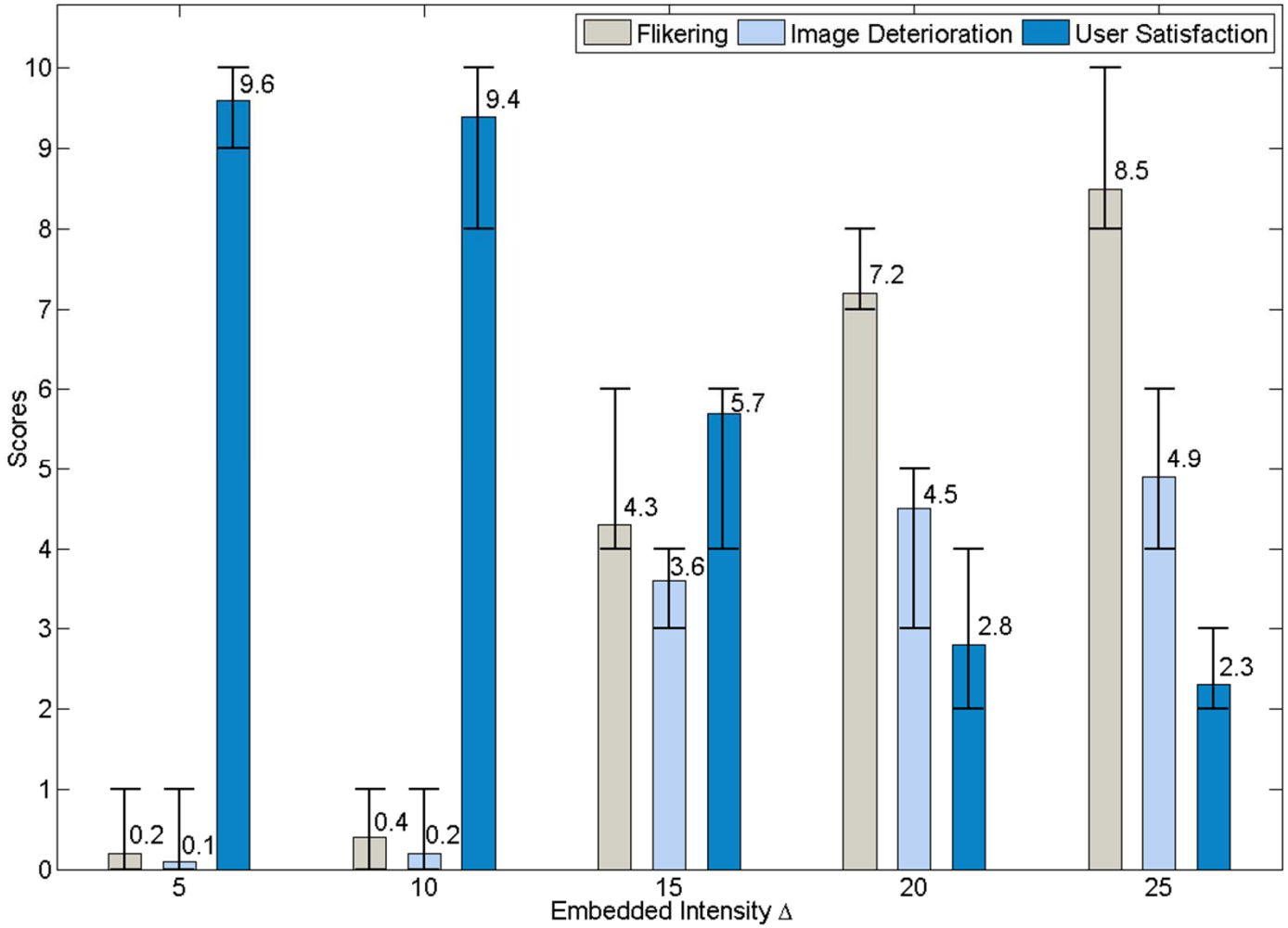
(c) Refined hand region

(d) Hand contour and detected fingertips

Touch Detection Through Homography



Experiments -- *Display Quality Evaluation*



Experiments -- *Touch Accuracy Evaluation*

Surface	Illumination			
	Dark		Normal	
	ϵ(px)	FRR/FAR(%)	ϵ(px)	FRR/FAR(%)
Gray	2.98	1.12/0.45	3.05	1.32/0.48
Yellow	3.04	1.23/0.57	3.12	1.54/0.61
Artifact	3.12	1.77/0.67	3.20	1.76/0.63

Comparison with recent depth-camera sensing based methods

In [2], the informal observed **spatial error** of finger detection on planar surface was between **3-6 pixels**,

In Omni-Touch [6], the **FRR** and **FAR** of finger click detection on four different surfaces were **0.8%** and **3.3%.**

Experiments -- *Efficiency Evaluation*

Subroutine	Hand Seg.	FTip Loc.	Touch Det.	Total
Time (ms)	14.63	1.32	1.74	17.69

Average processing time

Conclusion

- **This paper explores the possibility of replacing the display panel and the mouse-and-keyboard by a mere projector and camera.**

- **Limitations**
  - Hand segmentation depends on radiometric parameters
  - Too fast hand movement
  - Single hand operation