

# **Detecting Fingertip Method and Gesture Usability Research for Smart TV**

**Won-Jong Yoon and Jun-dong Cho**

# Outline

---

- **Previous work**
- **Scenario**
- **Finger detection**
- **Block diagram**
- **Result image**
- **Performance**
- **Usability test**
- **Conclusion**

# Previous work

---

- **Samsung Gesture TV (2014.01.07 CES)**
  - **Evolution kit(SEK-2000)**
    - Finger Gesture is developed from hand gesture.
    - Finger Gesture performs channel, volume, application, and scroll smart hub.
- **LG smart TV**
  - **One finger gesture recognition(2014.01.07 CES)**
    - <https://youtu.be/sNrtKQ3rAAs>
  - **Shaking open hand or close hand (2013.06)**
    - [https://youtu.be/900PY\\_JWSq0](https://youtu.be/900PY_JWSq0)

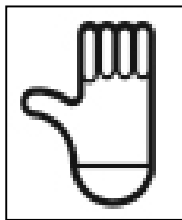
# Previous work(limitation)

---

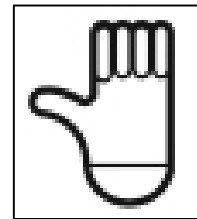
- Daeho Lee, and SeungGwan Lee. "Vision-based finger action recognition by angle detection and contour analysis." ETRI Journal, vol. 33, no. 3, pp. 415-422, June 2011.
  - When the wrist is displayed on the image, **this approach cannot detect the fingertips.**
- Qian, Chen, et al. "Realtime and robust hand tracking from depth", Computer Vision and Pattern Recognition (CVPR), pp. 1106-1113, 2014.
  - It could not separate the hand and the wrist. **User needs to wear black band.**

# User Scenario

- Initial Assumption
  - one hand is shown through the camera.
  - the palm heads to the camera.
  - the hand is ahead of your body.
  - each finger are detached.
  - hand is existed closely camera.



1



# User Scenario

---

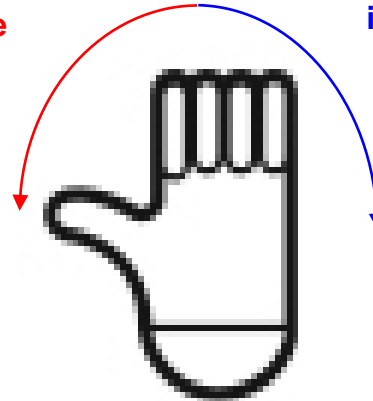
- Major functions
  - User can perform function of Smart TV through the gesture recognition.
  - During watching TV, the palm of user is ahead of body for adjusting TV using gesture recognition.
  - Gesture recognition is performed by number of fingertip.

# User Scenario

- **Finger gesture commands**
  - **5-finger**
    - Entertainment
  - **4-finger**
    - Volume control
  - **3-finger**
    - Channel control
  - **2-finger**
    - Double click
  - **1-finger**
    - Finger mouse

volume/channel  
decrease

volume/channel  
increase



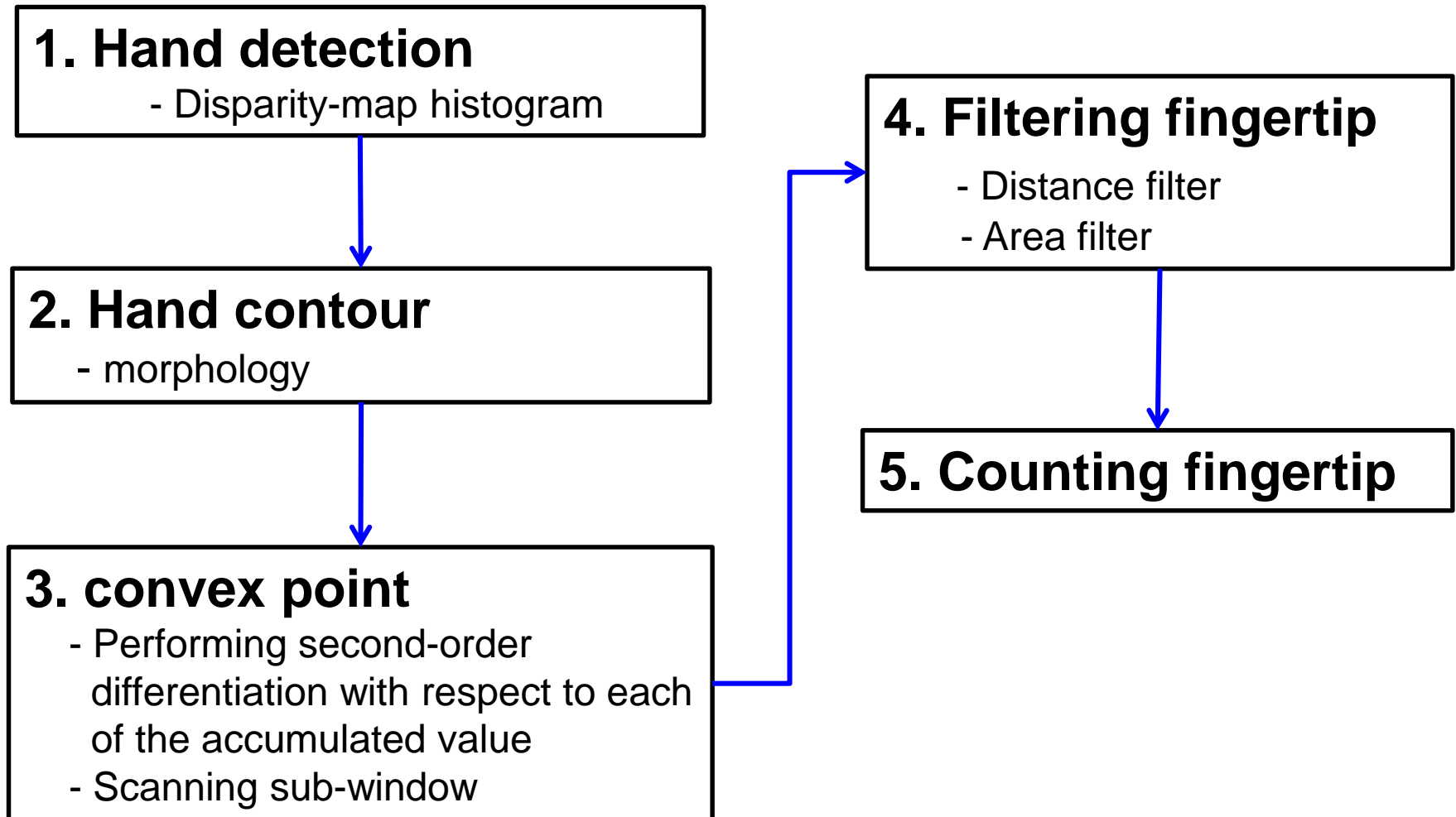
# Finger detection

---

- **Our proposed method detects convex feature-point in the object.**
- **Pros**
  - **Rotation-invariant and scale-invariant.**
    - **Even with freely moving the hands, we can detect the finger.**
  - **No need of wearing a band**
- **Cons**
  - **when user is apart far from camera, we cannot detect fingertip.**



# Block Diagram



# 1. Hand detection

- Skin color filter and histogram is used
- Threshold of hand is defined by the large value that is histogram of disparity-map.

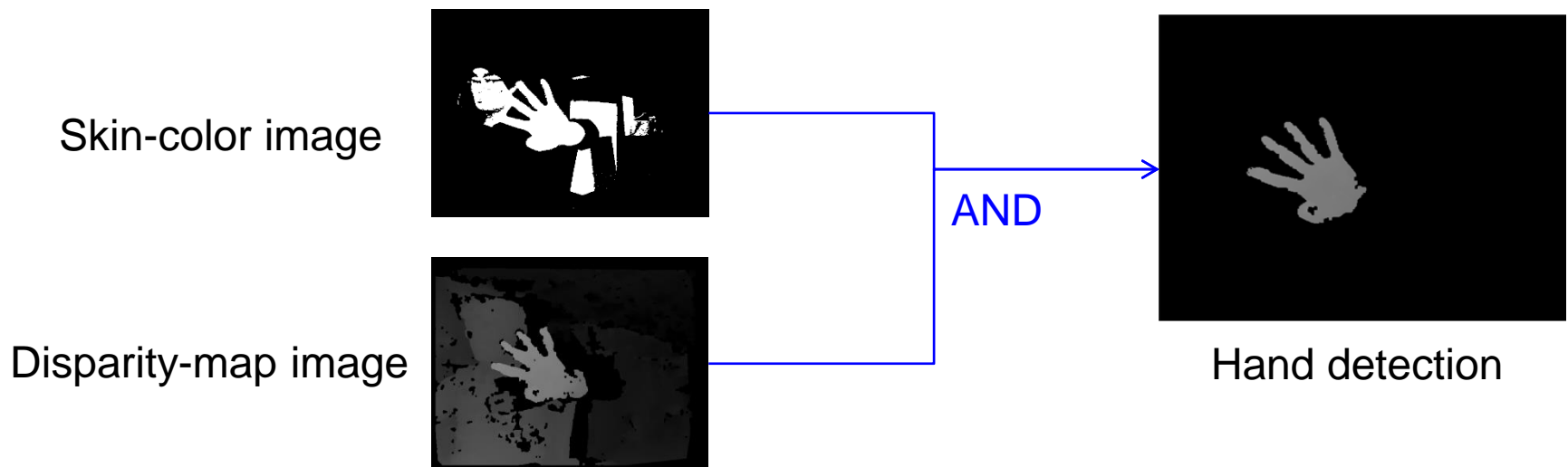
$$distance\ max = \frac{f \times T \times distance\ min}{f \times T + 100(mm) \times distance\ min}$$

$$HT = \frac{histogram(i + \Delta x) - histogram(i)}{\Delta x}, distance\ max \leq i \leq 255$$

- Distance min: max disparity value in the disparity-map image
- Distance max : the disparity value at 10cm from maximal disparity value
- HT : Hand Threshold value

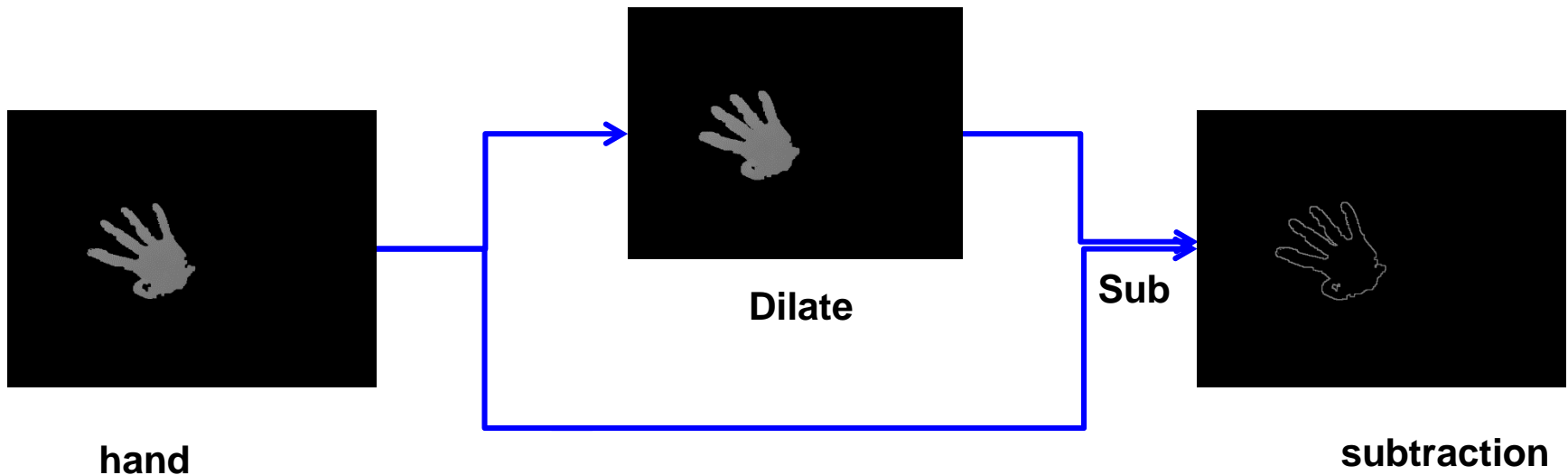
# 1. Hand detection (cont.)

- Hand detection is processed in the disparity-map by thresholding with Hand Threshold value.
- Hand detection images in the **disparity-map** and in the **skin-color image** were processed by **AND operation**.



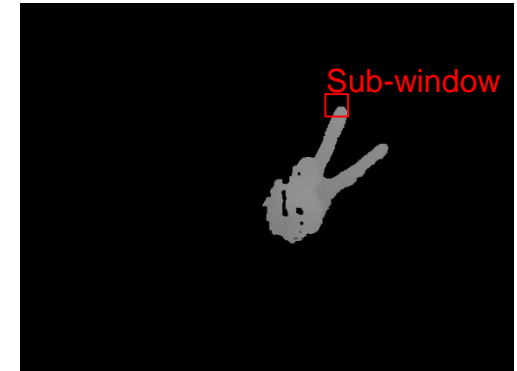
## 2. Hand contour

- hand contour image is extracted by morphology operation.
  - hand image process semiyearly dilate operation .
  - we process subtraction operation between hand image and dilate hand image.

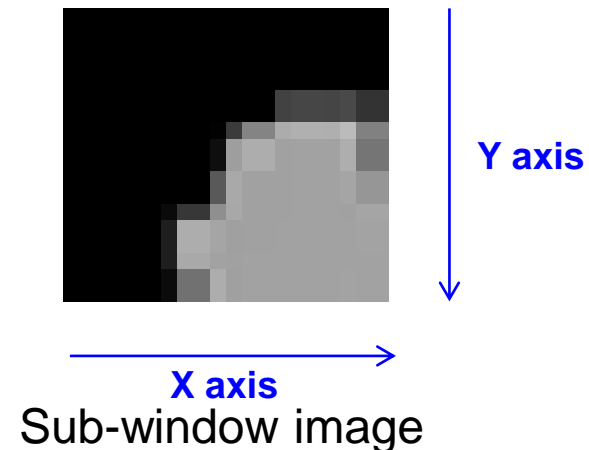


# 3. Convex point

- The convex point is the pixel which is the location of the Contour pixels in the region of Sub-window  $X$ , and we accumulates the pixel number relative to the  $Y$  axis.
- Since the fingertip is shaped as a curve, the value of the second derivative is present.
- we perform the second-order differential about the accumulated value.
- The total of two-order differential values of  $X$ -axis and  $Y$ -axis is equal to each of three or more, and we determine the candidate points.



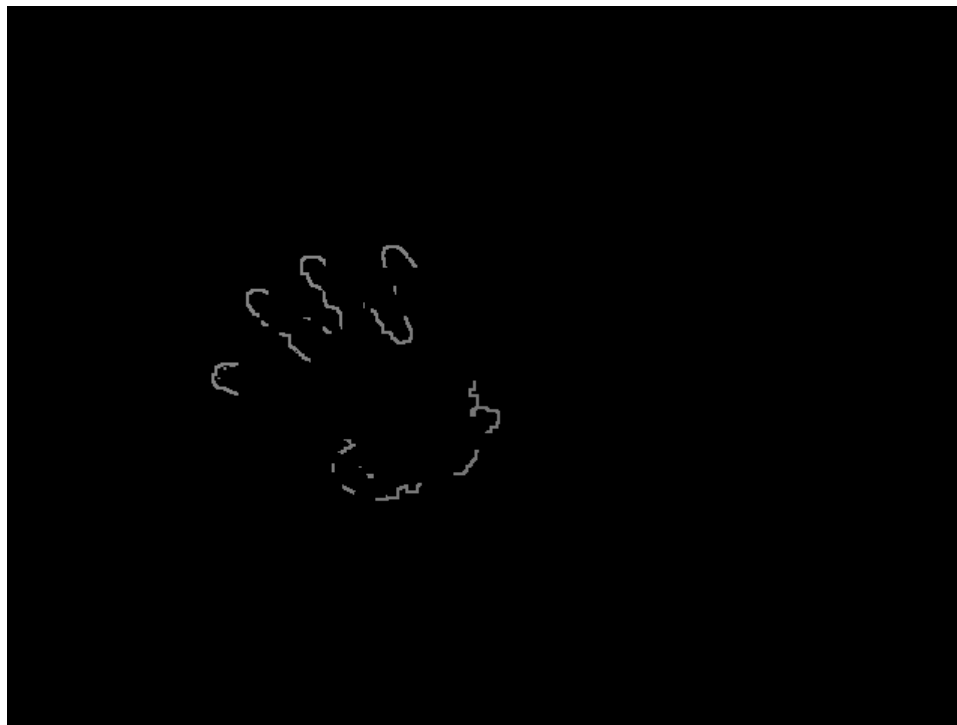
Hand detection



# 3. Convex point (cont.)

---

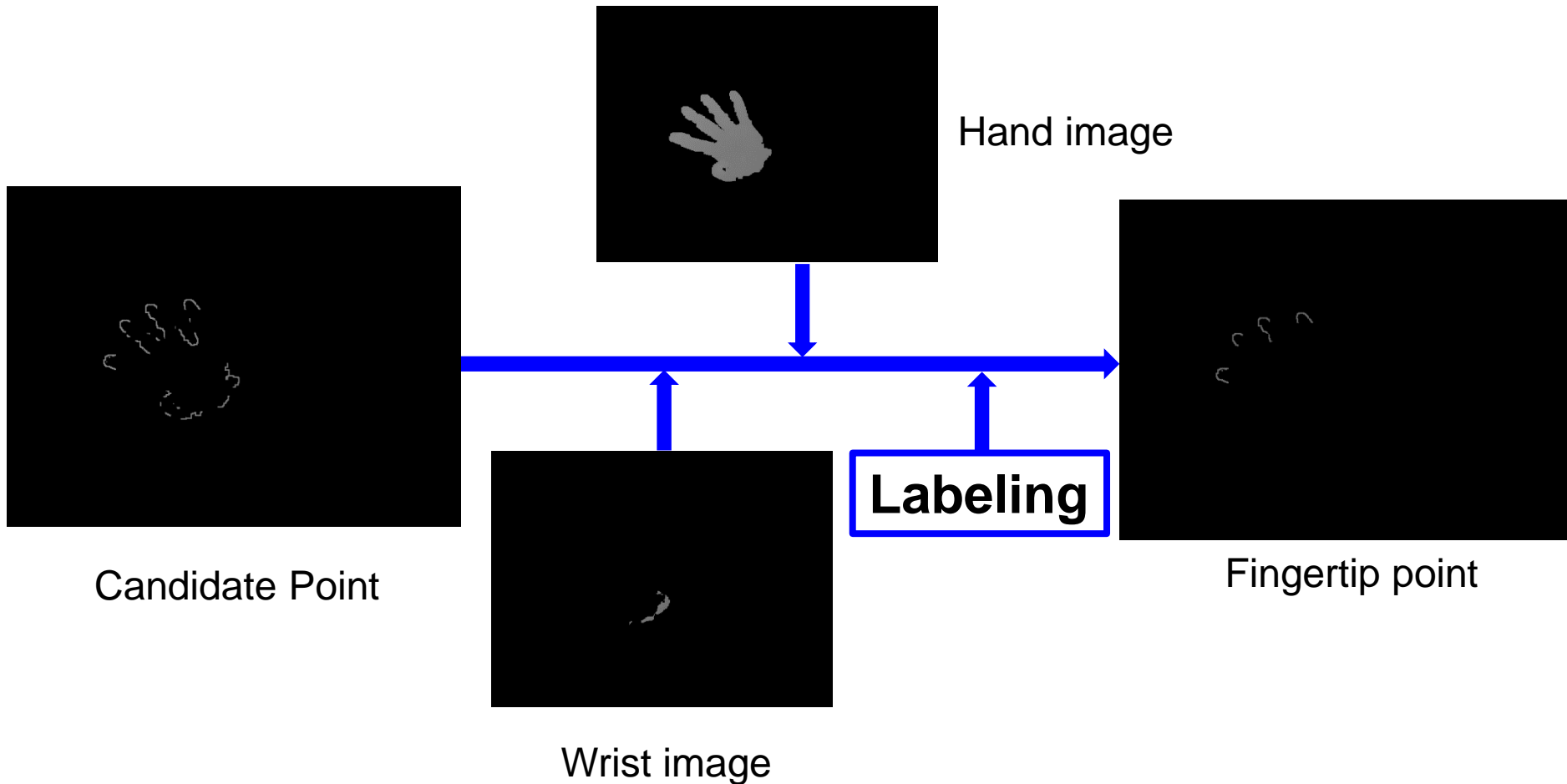
- Convex point image



# 4. Filtering fingertip

- We use distance filter and area filter for removing noise.
- Distance filter
  - Center of gradient of hand
    - distance : removing pixel within **hand region -> height 37%**
  - Center of gradient of wrist
    - distance : removing pixel within **hand region -> height 48%**
- Area filter
  - Using labeling algorithm
  - If area is less than 70, then we process to remove the noise.

# 4. Filtering fingertip (cont.)

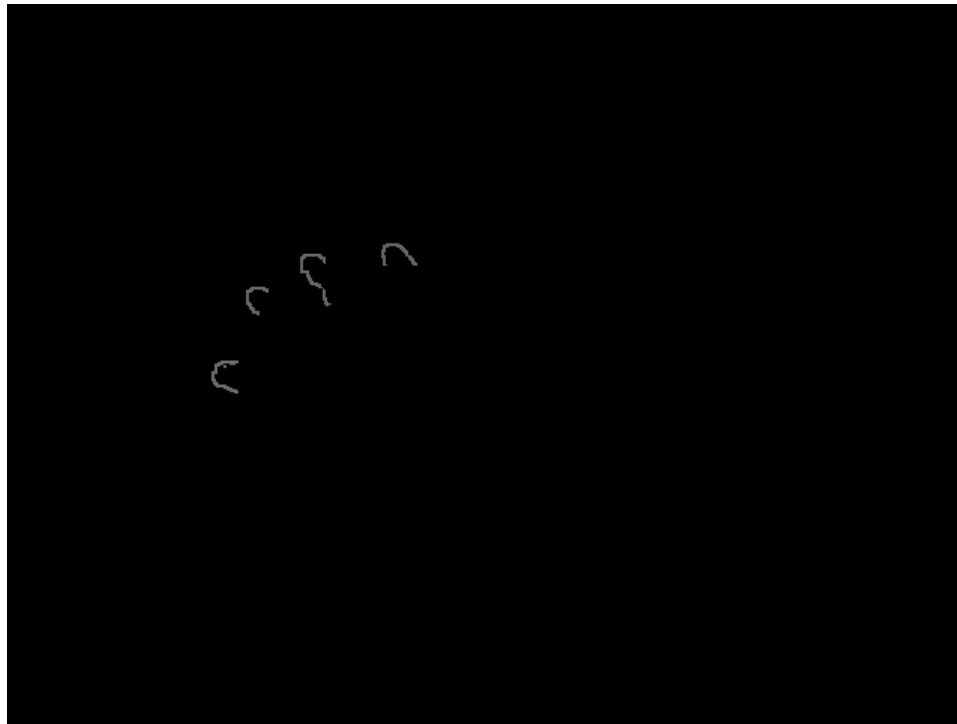




# 5. Counting fingertip

---

- Labeling algorithm



# Result image



Input image



output image

# Performance

Number of Fingertips	Correct / Total (Accuracy Rate : %)
1	110 / 111 (99.09%)
2	91 / 96 (97.79%)
3	78 / 80 (97.5%)
4	75 / 80 (93.75%)
5	98 / 101 (97.02%)
Total	<b>452 / 468 (96.58%)</b>

# Usability test

---

- **Tester : 29 person (The Twenties )**
- **We implement gesture recognition that is fingertip mouse, channel/volume function and entertainment**
- **Satisfaction, ease of handling, convenience and entertainment were evaluated by five-point scale.**

# Usability test(cont.)

- **Subject test result**

<b>Experiment item</b>	<b>Average / 5-point scale</b>	
	<b>Before experiment</b>	<b>After experiment</b>
Satisfaction	3.17 / 5	4.00 / 5
Ease of Handling	3.24 / 5	4.00 / 5
Convenience	3.48 / 5	4.13 / 5
Entertainment	2.37 / 5	4.44 / 5

# Usability test(cont.)

---

- **Question**
  - **Tester selects rank.(first/second/third)**
    - **Entertainment / volume / channel**
  - **Are you interested in using Smart TV with remote control and gesture recognition?**
    - **Positive Tester : 25 of 29 (person) : 86.2%**
  - **Are you interested in using a gesture recognition only for Smart TV?**
    - **Positive Tester : 22 of 29 (person) : 75%**

# Conclusion

---

- **Our proposed natural user interface method has advantage that is rotation-invariant and scale-invariant.**
- **If gesture recognition is used in conjunction with a remote control for Smart TV, user tests results in **convenience of 86.2%**.**
- **Our system resolved the limitation of the previous works [Daeho 2011, Qian 2014]**

---

**Thank you**

**Q & A**