Vision Based hand tracking for Interaction

The 7th International Conference on Applications and Principles of Information Science (APIS2008)

Dept. of Visual Contents, Dongseo University

2008. 01. 29

Seung Il Han, Jeon Yeong Mi, Jung Hye Kwon, Hwang-Kyu Yang, Byung Gook Lee

hs12@dit.dongseo.ac.kr

Outline

1. Purpose
2. System overview
3. Flow chart
4. Segmentation of a hand region
5. Matching stereo data
6. Conclusion
7. Future work

1. Introduction to hand tracking for interaction

1.1 Purpose

- Any user can interact anytime without any condition of space.
- More comfortable to use digital-device on public place.

1.2 Optical vs. computational reconstruction

- Hardware
  - Bumblebee stereo camera
  - Display device
  - PC
- Software
  - Image processing
  - Stereo matching

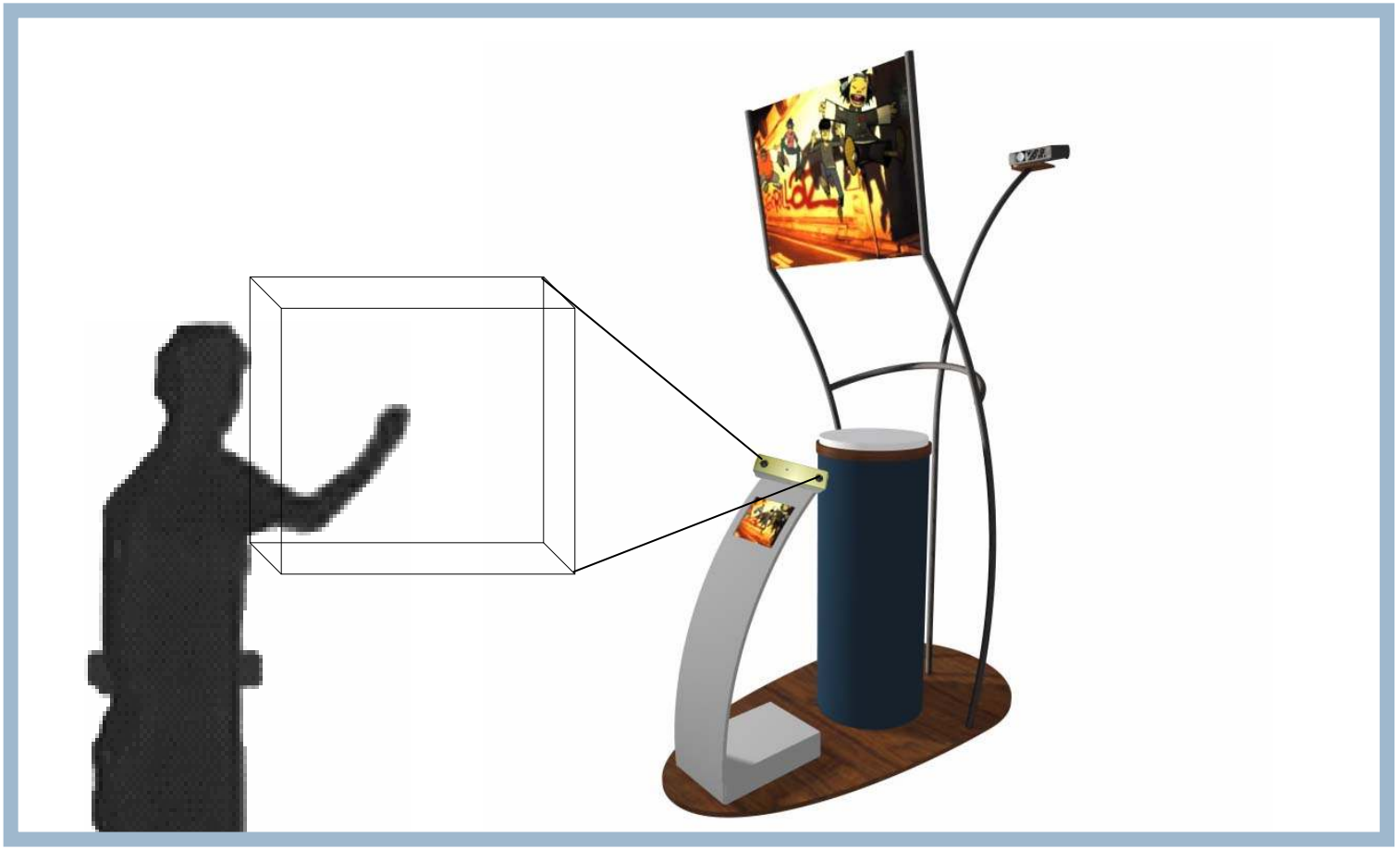
bumblebee™

Specifications:

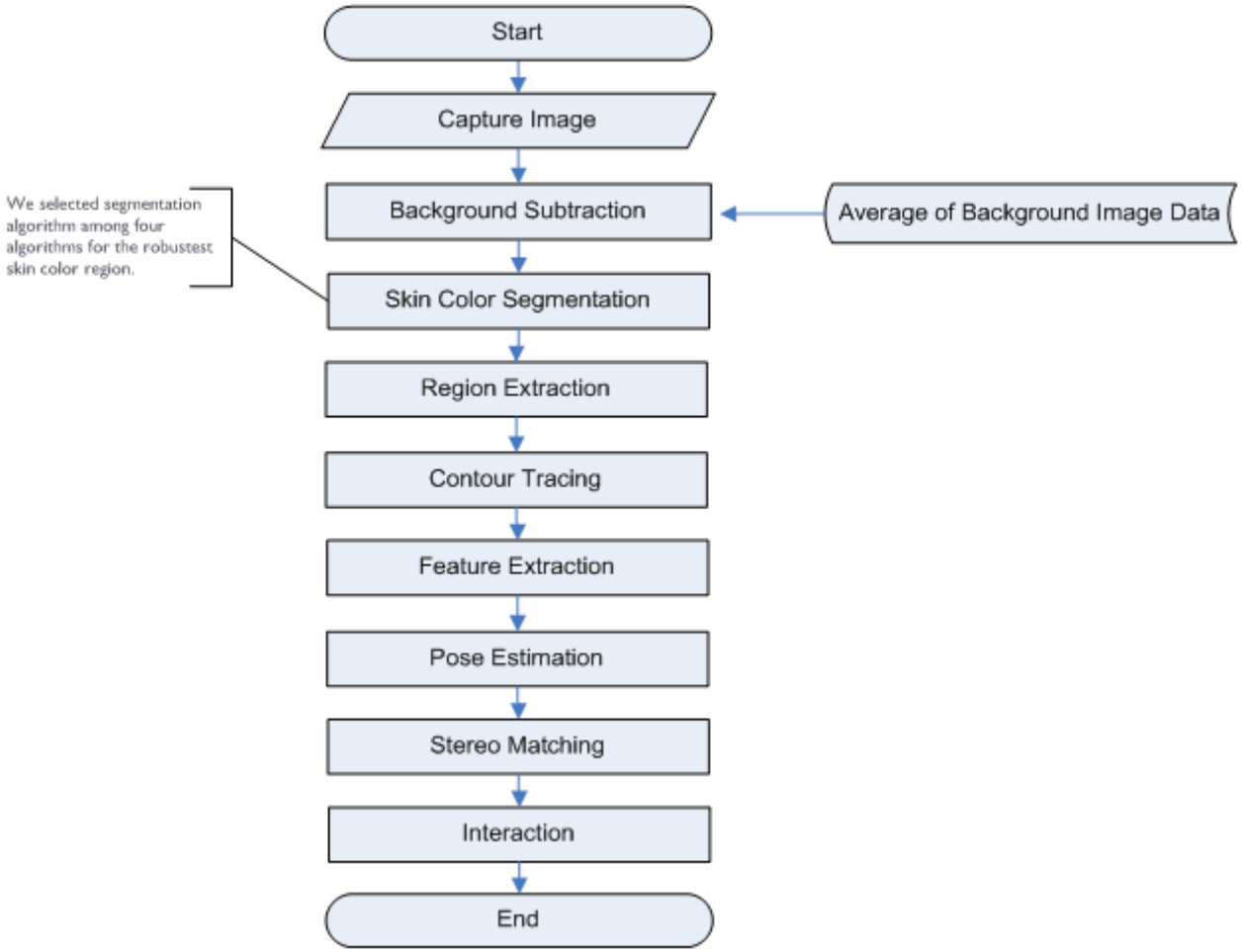
- 640x480 Option: two Sony ICX084 Color or BW CCDs
- 1024x768 Option: two Sony ICX204 Color or BW CCDs
- 1/3 inch, progressive scan CCDs
- 640x480 square pixels at 30 Hz frame rate
- 1024x768 square pixels at 15 Hz frame rate
- 10 bit A/D
- Shutter speed: 1/16000 to 1/30 second
- Baseline: 12cm
- Choice of 2mm, 4mm or 6mm focal length lenses (100°, 70°, and 50° HFOV respectively)

2. System overview

2.1 System architecture



2.2 Flow chart



2.3 The segmentation of a hand region

[Background]

[Captured]

[Foreground]

[Subtraction]

$$I_f = \begin{cases} 255, & |I_i - I_b| > Threshold \\ 0, & otherwise \end{cases}$$

2.3 The segmentation of a hand region

- Detection algorithm
  - Peer and Solina's 2003 [1]
  - Hsu and Mottaleb Jain's Central 2002 [2]
  - Hsu and Mottaleb Jain's Ellipse 2002 [2]
  - M. Jones and J. Rehg's Color Histogram [3, 4]

[1] Kovac, J. Peer, P. Solina, F. "Human Skin Color Clustering for Face Detection", *The IEEE Region*, vol.2, pp.144-148, 2003.

[2] R.L.Hsu, M.Abdel-Mottaleb, A.K.Jain, "Face Detection in Colour Images", *IEEE transactions, on Pattern Analysis and Machine Intelligence*, vol.24, no.5, pp 696-706, 2002 .

[3] M. Jones, J. Rehg. "Statistical color models with application to skin", *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol.1, pp.274-280, 2002.

[4] Shahzad Malik, "Real-time Hand Tracking and Finger Tracking for Interaction", CSC2503F Project Report, December 18, 2003.

2.3 The segmentation of a hand region

- Peer and Solina's 2003

* Uniform daylight illumination

$$R > 95, \quad G > 40, \quad B > 20 \quad and$$
$$max\{R, G, B\} - min\{R, G, B\} > 15 \quad and$$
$$|R - G| > 15 \quad and \quad R > G \quad and \quad R > B$$

$$OR$$

* Flashlight or light daylight

$$R > 220, \quad G > 210, \quad B > 170 \quad and$$
$$|R - G| \leq 15 \quad and \quad R > B \quad and \quad G > B$$

2.3 The segmentation of a hand region

- Hsu and Mottaleb Jain's Central 2002

$$W_{Ci}(Y) = \begin{cases} W_{LCi} + \frac{(Y - Y_{min})(W_{Ci} - W_{LCi})}{K_l = Y_{min}}; & Y < K_l \\ W_{HCi} + \frac{(Y_{max} - Y)(W_{Ci} - W_{HCi})}{Y_{max} - K_h}; & K_h < Y \\ W_{Ci}; & else \end{cases}$$

$$\bar{C}_b(Y) = \begin{cases} 108 + \frac{10(K_l - Y)}{K_l - Y_{min}}; & Y < K_l \\ 108 + \frac{10(Y - K_h)}{Y_{max} - K_h}; & K_h < Y \\ 108; & else \end{cases}$$

$$\bar{C}_r(Y) = \begin{cases} 154 - \frac{10(K_l - Y)}{K_l - Y_{min}}; & Y < K_l \\ 154 + \frac{22(Y - K_h)}{Y_{max} - K_h}; & K_h < Y \\ 108; & else \end{cases}$$

$RGB\ to\ YC_bC_r$

$$Y = 0.299R + 0.587G + 0.144B$$
$$C_b = -0.168R - 0.331G + 0.5B$$
$$C_r = 0.5R - 0.418G - 0.081B$$

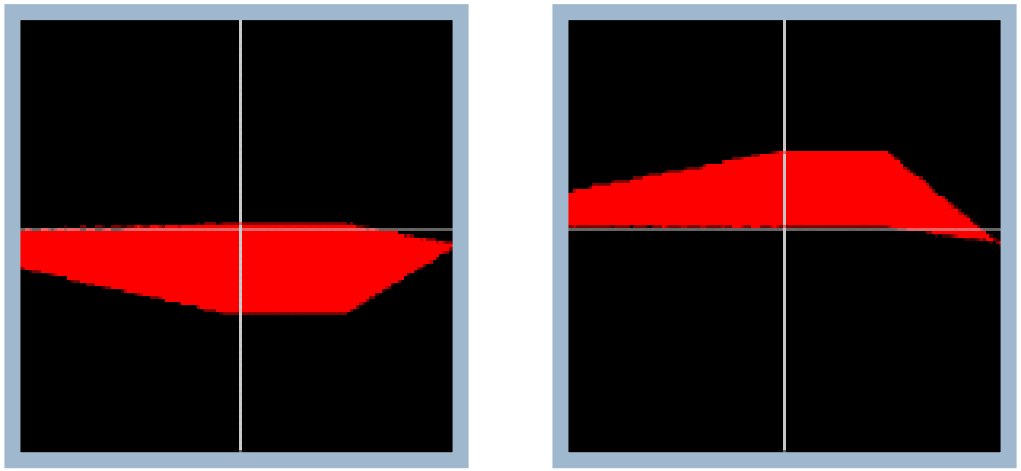
$where\ i\ in\ W_{Ci}(Y)\ is\ b\ or\ r, W_{Cb} = 46.97, W_{LCb} = 23, W_{HCb} = 14, W_{Cr} = 38.76,$
$W_{LCr} = 20, W_{HCr} = 10, K_l = 125, K_h = 188, Y_{min} = 16, Y_{max} = 235$

2.3 The segmentation of a hand region

- Hsu and Mottaleb Jain's Central 2002

$$\bar{C}_b(Y) - \alpha W_{Cb}(Y) \quad < C_b \quad < \bar{C}_b(Y) + \alpha W_{Cb}(Y), \quad \text{: a)}$$

$$\bar{C}_r(Y) - \alpha W_{Cr}(Y) \quad < C_r \quad < \bar{C}_r(Y) + \alpha W_{Cr}(Y) \quad \text{: b)}$$



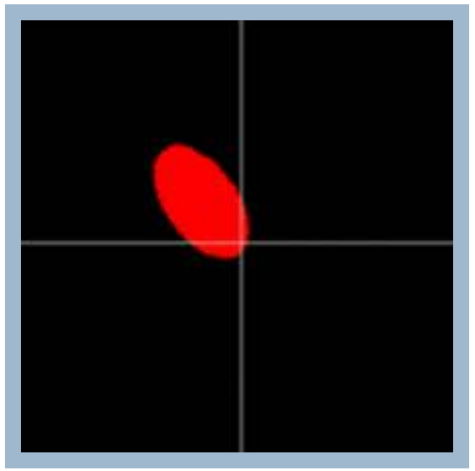
a) b)

2.3 The segmentation of a hand region

- Hsu and Mottaleb Jain's Ellipse 2002

$$\frac{(x-e_{cx})^2}{a^2}+\frac{(y-e_{cy})^2}{b^2}\leq 1$$

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}\begin{bmatrix} C_b-c_x \\ C_r-c_y \end{bmatrix}$$



$$c_x=109.38, c_y=152.02, \theta=2.53 radian$$

$$e_{cx}=1.6, e_{cy}=2.41, a=25.39, b=14.03$$

[2] R.L.Hsu, M.Abdel-Mottaleb, A.K.Jain, "Face Detection in Colour Images", *IEEE transactions, on Pattern Analysis and Machine Intelligence*, vol.24, no.5, pp.696-706, 2002 .

2.3 The segmentation of a hand region

- M. Jones and J. Rehg's color histogram

$$P(rgb|skin) = \frac{s[rgb]}{T_s}$$

$$P(rgb|-skin) = \frac{n[rgb]}{T_n}$$

$$P(skin|rgb) = \frac{P(rgb|skin)P(skin)}{P(rgb|skin)P(skin) + P(rgb|-skin)P(-skin)}$$

$$P(skin) + P(-skin) = 1$$

$$P(skin) = \frac{T_s}{T_s + T_n}$$

$$P(skin|rgb) \geq \sigma_s$$

H_s : The skin histogram

H_n : The non-skin histogram

$s[rgb]$: The pixel count in bin rgb of H_s

$n[rgb]$: The pixel count in bin rgb of H_n

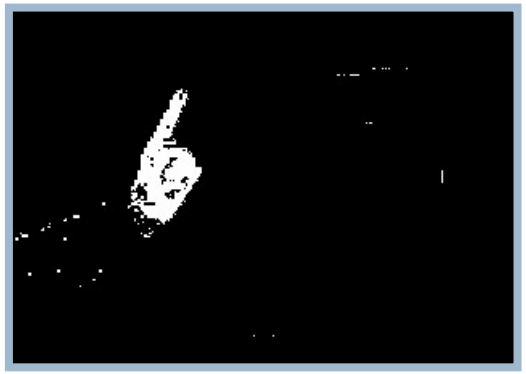
T_s : The total counts contained in H_s

T_n : The total counts contained in H_n

2.3 The segmentation of a hand region

- Technique of region extraction

[Color image of background subtraction]



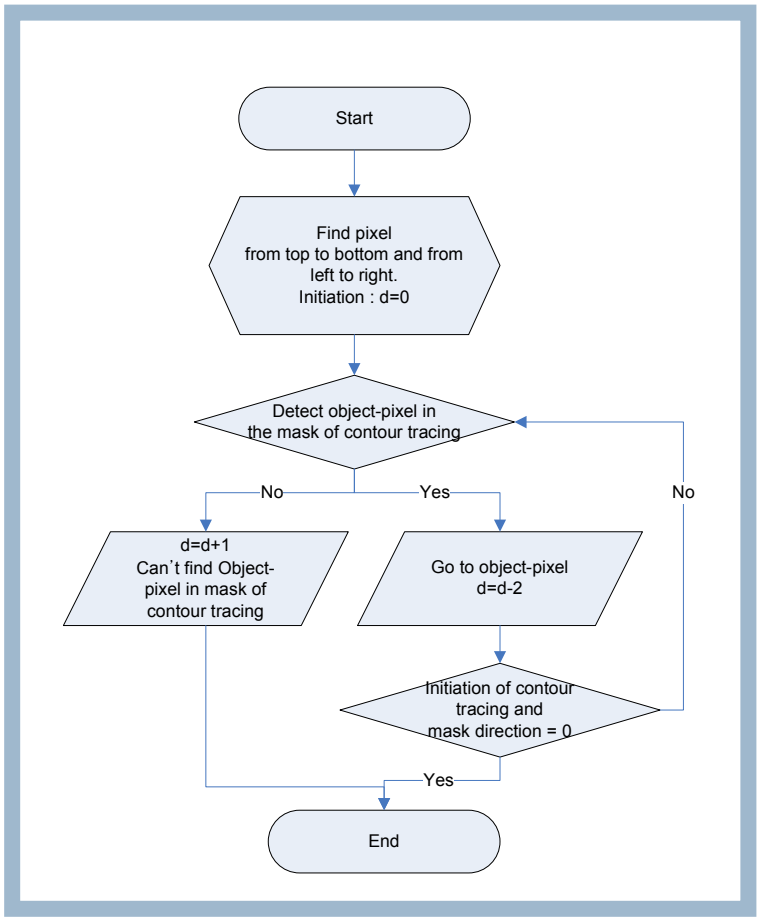
[Binary image of background subtraction]

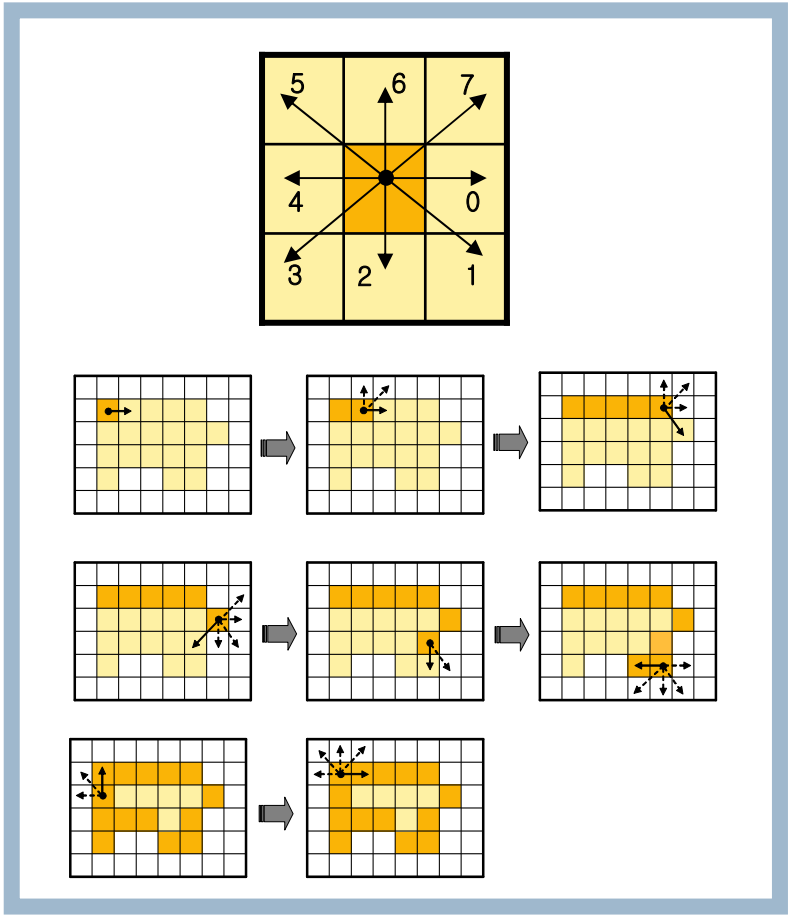
[Blob analysis of binary image]

[Binary morphological opening]

2.4 Detection of fingers

- Contour tracing with clockwise algorithm

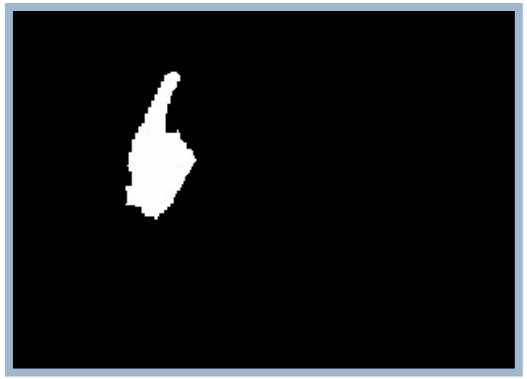




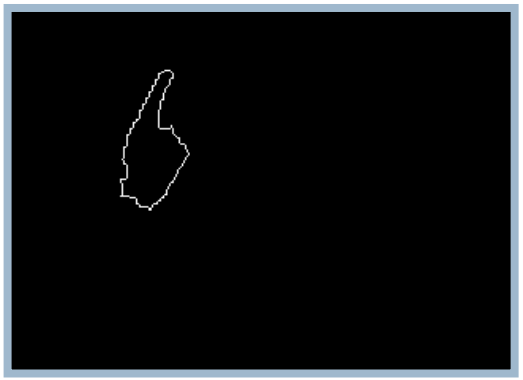
[5] Image Processing Program by Visual C++ , *Published by HANBIT Media*, pp.672-679, Vol. 10, No. 1779, 2007.

2.4 Detection of fingers

- Result of contour tracing



[Result of binary morphological opening]



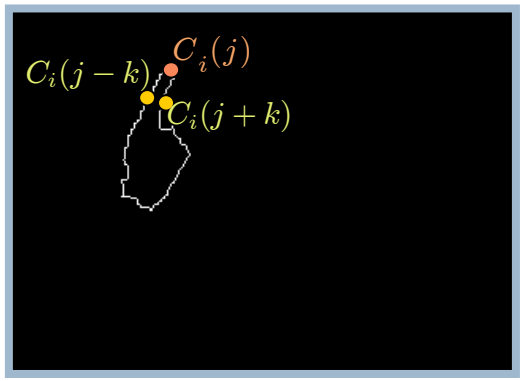
[Contour tracing result of feature extraction]

2.4 Detection of fingers

- Detection of peak point
  - At each pixel j in a hand contour i
  - Compute k-curvature θ_k
    - Angle between the two vectors

      $[C_i(j), C_i(j-k)]$ and $[C_i(j), C_i(j+k)]$ $where,\ k\ is\ constant$
  - Peak point and valley point
    - Peak point $\theta_k < \theta_{min}$
    - Valley point $\theta_k > \theta_{max}$

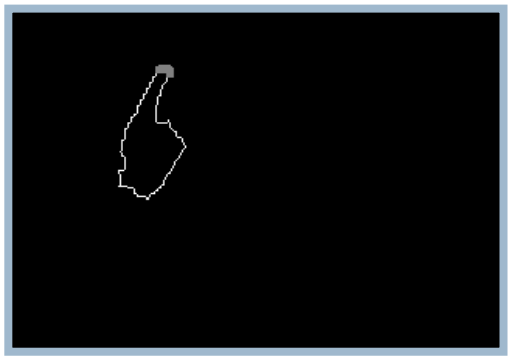
    $where,\ \theta_{min} = 90^\circ,\ \theta_{max} = 200^\circ$



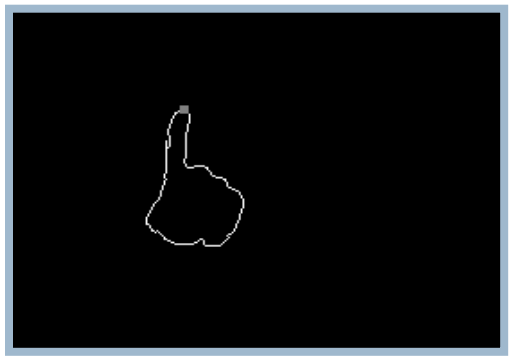
[4] Shahzad Malik, "Real-time Hand Tracking and Finger Tracking for Interaction", CSC2503F Project Report, December 18, 2003.

2.4 Detection of fingers

- Detection of peak point



[Peak point detection]



[Valley point detection]

2.4 Detection of fingers

- Detection of midpoint
  - Define $cnorm(x)$ as:

$$cnorm(x) = \begin{cases} x + N : x < 0 \\ x - N : (x + 1) > N \\ x : otherwise. \end{cases}$$

$where\ N\ represent\ the\ number\ of\ points\ along\ the\ contour\ perimeter.$

  - Midpoints represent the axis of the finger.
    - $P(cnorm(i + k))$ is a point that is k points to the left along a hand contour
    - $P(cnorm(i - k))$ is a point to the right
    - Let i_{ft} is the index of a finger tip

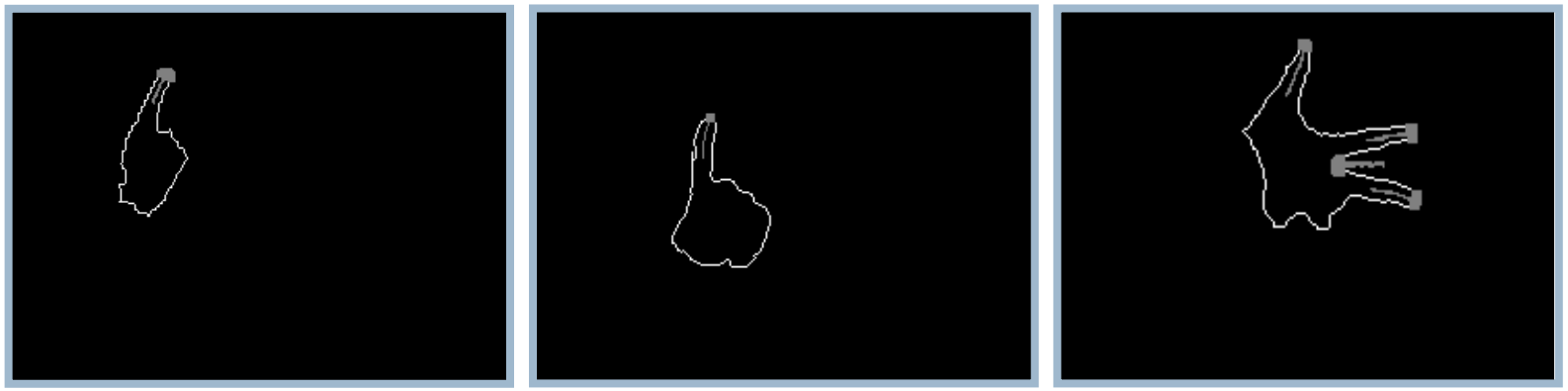
[4] Shahzad Malik, "Real-time Hand Tracking and Finger Tracking for Interaction", CSC2503F Project Report, December 18, 2003.

2.4 Detection of fingers

- Detection of midpoint
  - A midpoint $Q(k)$

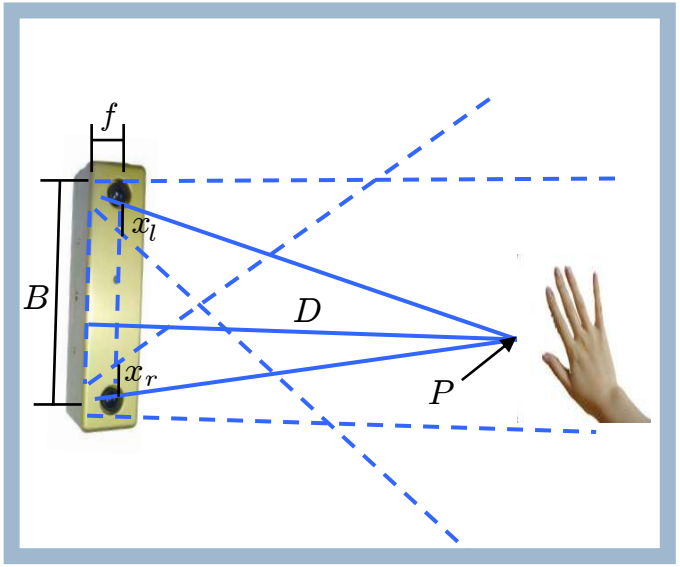
$$Q(k) = \frac{P(cnorm(i_{ft} + k)) + P(cnorm(i_{ft} - k))}{2}$$

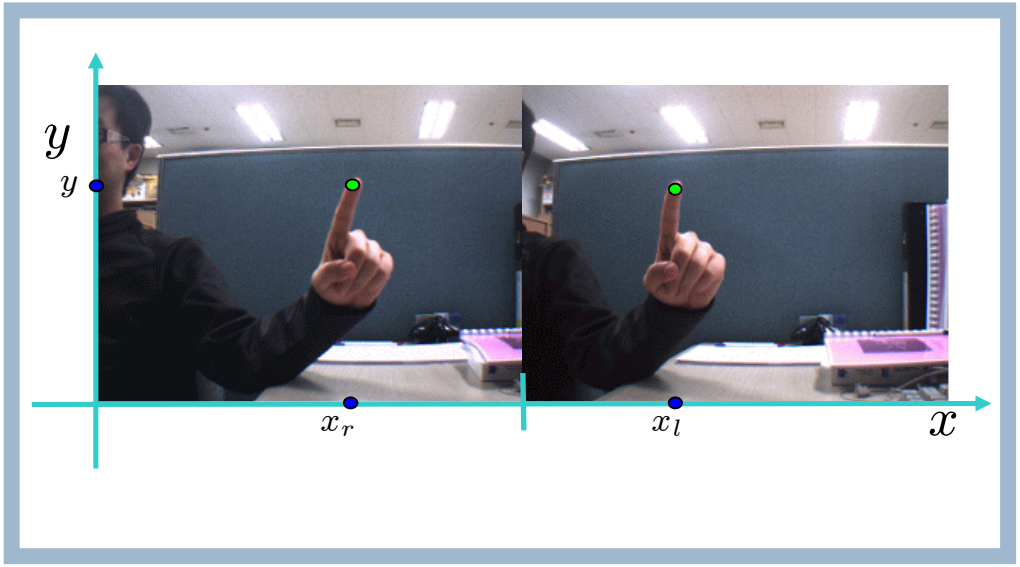
  - Compute $Q(k)$ for $k_{min} < k < k_{max}$



2.5 Matching stereo data

- Calculation of depth value





$$P_x = \frac{x_l + x_r}{2}, \quad P_y = \frac{y_l + y_r}{2} = y, \quad P_z = f\frac{B}{|x_l - x_r|}$$

$$D = \frac{P_z}{RC}$$

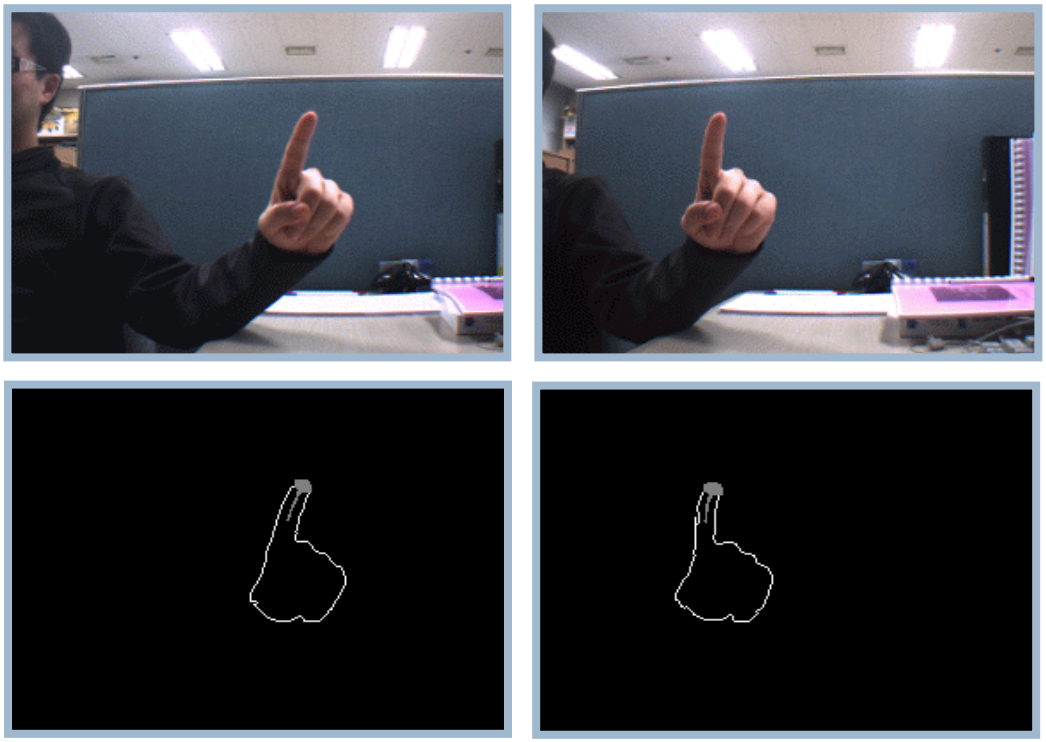
D : Distance using centimeter

R : Resolution ratio , C : CCD Sensor size

[6] Tomase Poggio, "Introductory Techniques for 3-D Computer Vision", *Massachussetts Institute of Technology*, pp. 139-145, 1998.

2.5 Matching stereo data

- Result of depth value
  - Real distance ≒ 66 Cm
  - Calculated depth value = 67.5 Cm



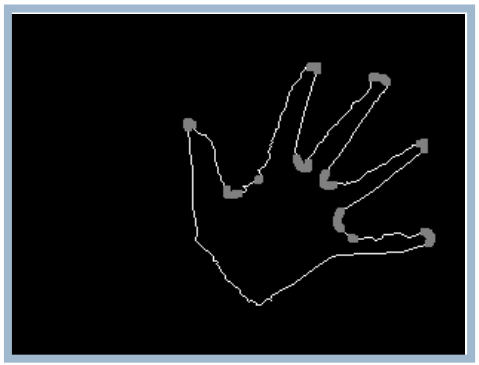
3. Conclusion

3.1 Conclusion

- The hardware costs very low.
- Hand tracking using an enhanced algorithm
- Interaction with system using user's hands
- Free of conditions of time & space

3.2 Future Work

- To perform a robust segmentation with ambient environment while detection the skin color.
- The determination of peak-point count and decision of location.
  - Gesture recognition performance
- To enhance the accuracy and speed of the system.
- Interaction using hands.



References

1. Kovac, J. Peer, P. Solina, F. "Human Skin Color Clustering for Face Detection", The IEEE Region, vol.2, pp.144-148, 2003.
2. R.L.Hsu, M.Abdel-Mottaleb, A.K.Jain, "Face Detection in Colour Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.5, pp.696-706, 2002 .
3. M. Jones, J. Rehg. "Statistical color models with application to skin", In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol.1, pp.274-280, 2002.
4. Shahzad Malik, "Real-time Hand Tracking and Finger Tracking for Interaction", CSC2503F Project Report, December 18, 2003.
5. "Image Processing Program by Visual C++", Published by HANBIT Media, pp.672-679, Vol. 10, No.1779, 2007.
6. Tomase Poggio, "Introductory Techniques for 3-D Computer Vision", Massachussetts Institute of Technology, pp.139-145, 1998.
7. George V. Paul, Glenn J. Beach, Charlesn J.Cohen, "A Real-time Object Tracking System using a Color Camera", pp.137-142, 2001.
8. J. Segen, S. Kumar. "3D hand pose estimation using a single camera", In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol.1, pp.479-485, 1999.

Thank You