NANYANG TECHNOLOGICAL UNIVERSITY

# Human computer interface for quadriplegic people based on face position/gesture detection

Zhen-Peng BIAN *PhD student, IMI & EEE, NTU*
Supervisor *Assoc Prof Chau Lap-Pui, EEE, NTU*
Co-supervisor *Prof Nadia Magnenat THALMANN, IMI, NTU*

Accepted by ACM Multimedia 2014

institute for media innovation
The avant-garde place

# ABSTRACT


- This work proposes a human computer interface using a depth camera for quadriplegic people. The nose position is employed to control the cursor along with the commands provided by mouth's status. The detection of nose position and mouth's status is based on randomized decision tree algorithm. The experimental results show that the proposed interface is comfortable, easy to use, robust, and outperforms the existing assistive technology.

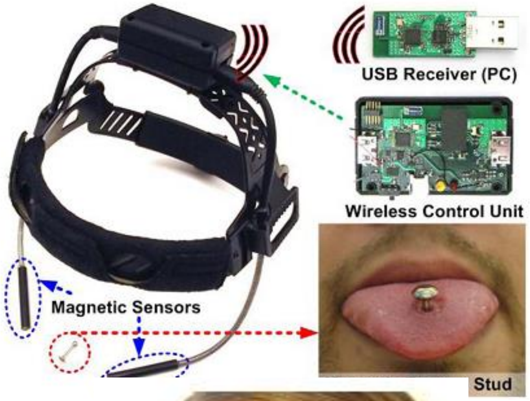
# Introduction

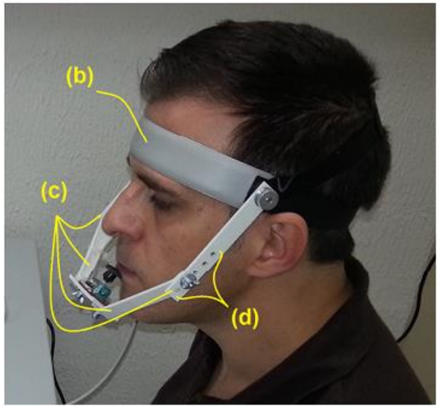
- Quadriplegic people have great difficulties to use standard input devices of personal computers, such as mouse. Some assistive technologies (ATs) have been developed to help them to control computers. According to [1][2], ATs can be roughly categorized as:

1. Physiological signals based, such as **E**lectro**E**ncephalo**G**ram, **E**lectro**M**yo**G**ram, and **E**lectro**O**culo**G**ram;
2. Voice commands based, such as speech recognition and non-verbal vocalization software;
3. Mechanical motion based such as sip-and-puff or mouth stick ;
4. Tracking movement of body parts and/or landmarks based, such as head and eye trackers.

[1]J. Music, et al.. Testing inertial sensor performance as hands-free human computer interface. WSEAS Trans. Comput., 2009.
[2]B. Youse, et al.. Quantitative and comparative assessment of learning in a tongue-operated computer input device. IEEE T INF TECHNOL B., 2011.

# Assistive Technologies

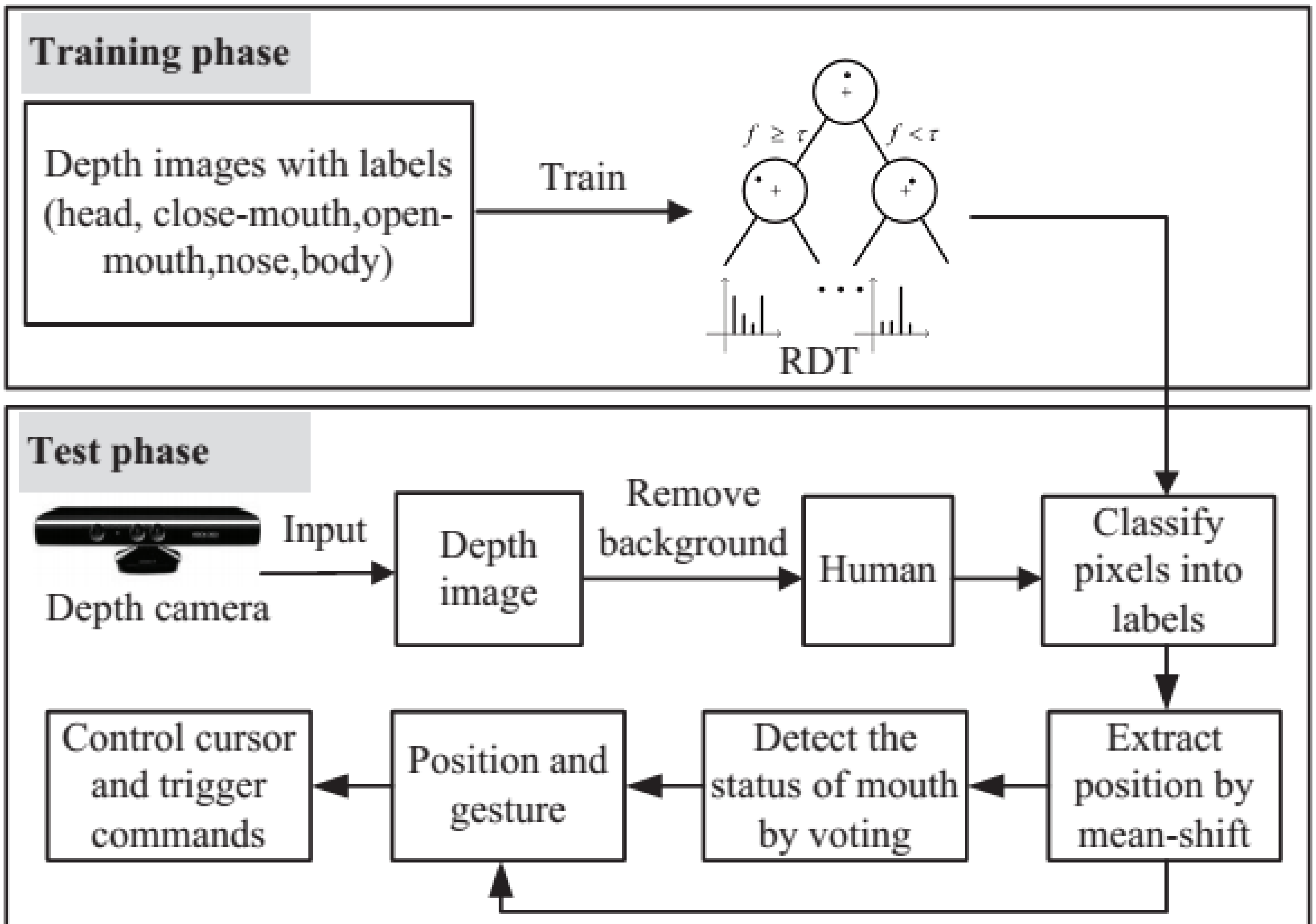

# Motivation

- Interface for quadriplegic people

1. comfortable (no body attachment),
2. easy to use (no initialization, no calibration, no guardian)
3. robust (insensitive to colour or illumination of lights).

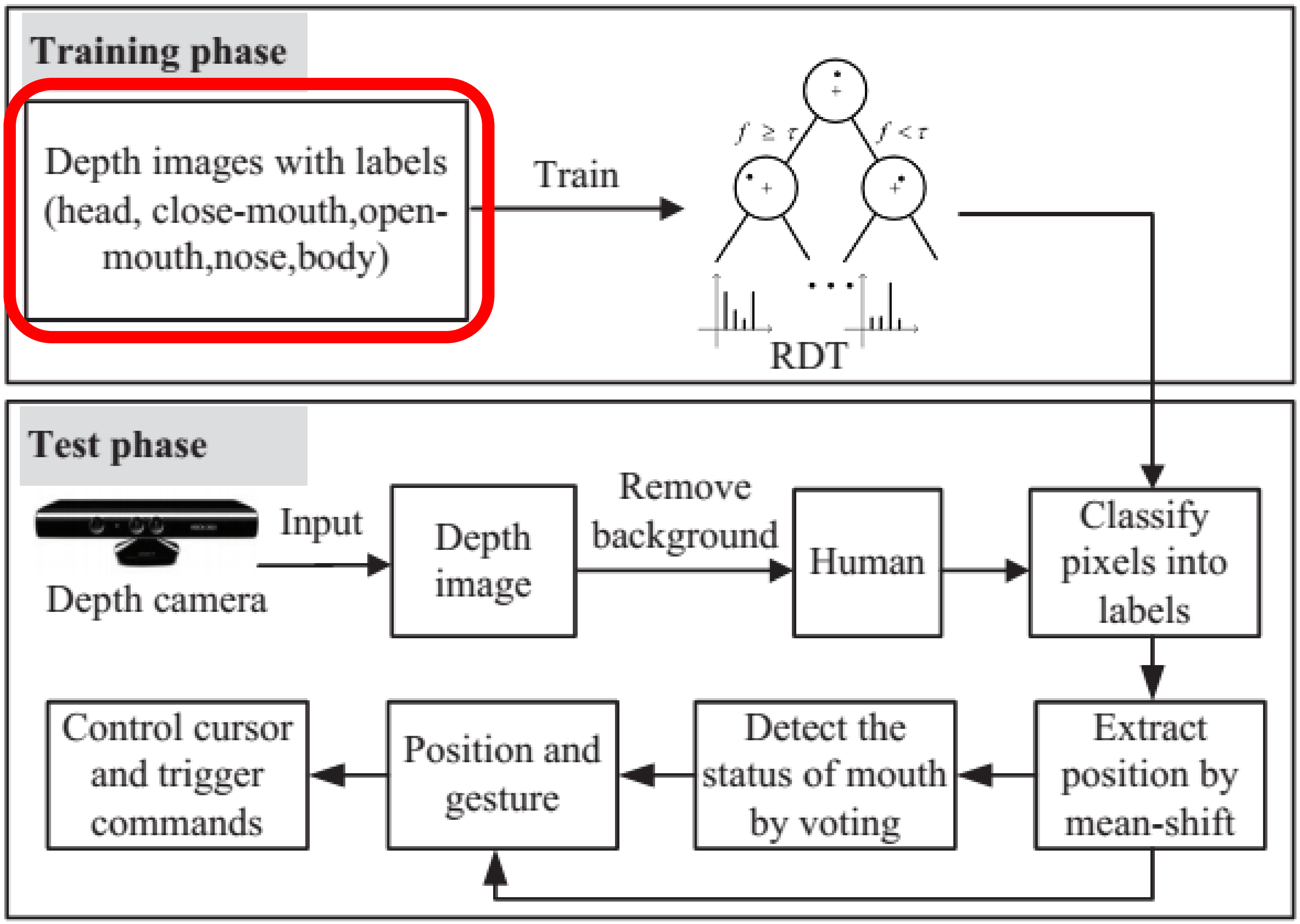
- It can enrich people with severe mobility impaired in their rehabilitation or daily life activities.

# Framework



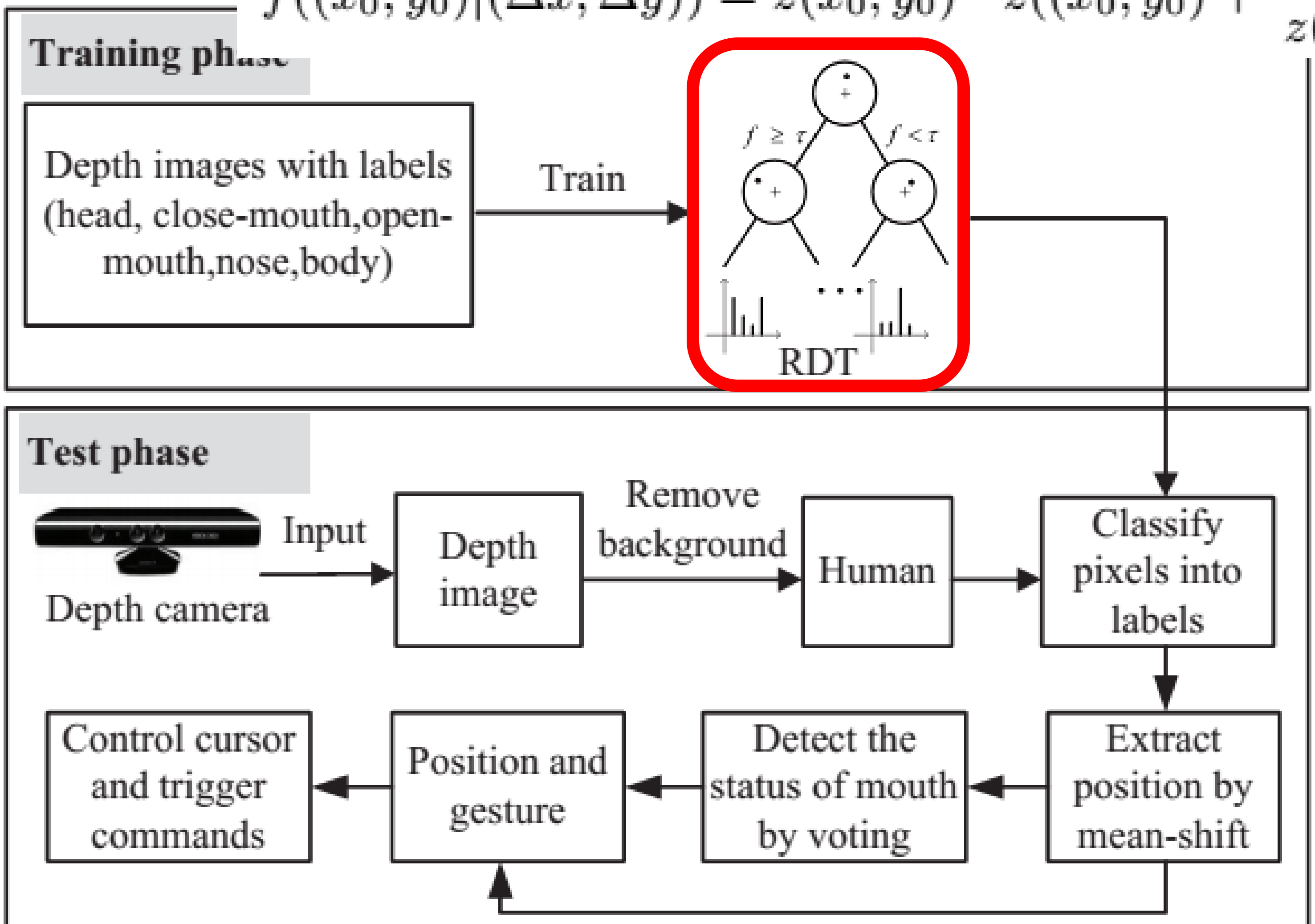
Training phase
Depth images with labels (head, close-mouth,open-mouth,nose,body)
Train
RDT
Test phase
Input
Depth camera
Depth image
Remove background
Human
Classify pixels into labels
Extract position by mean-shift
Detect the status of mouth by voting
Position and gesture
Control cursor and trigger commands

# Training data

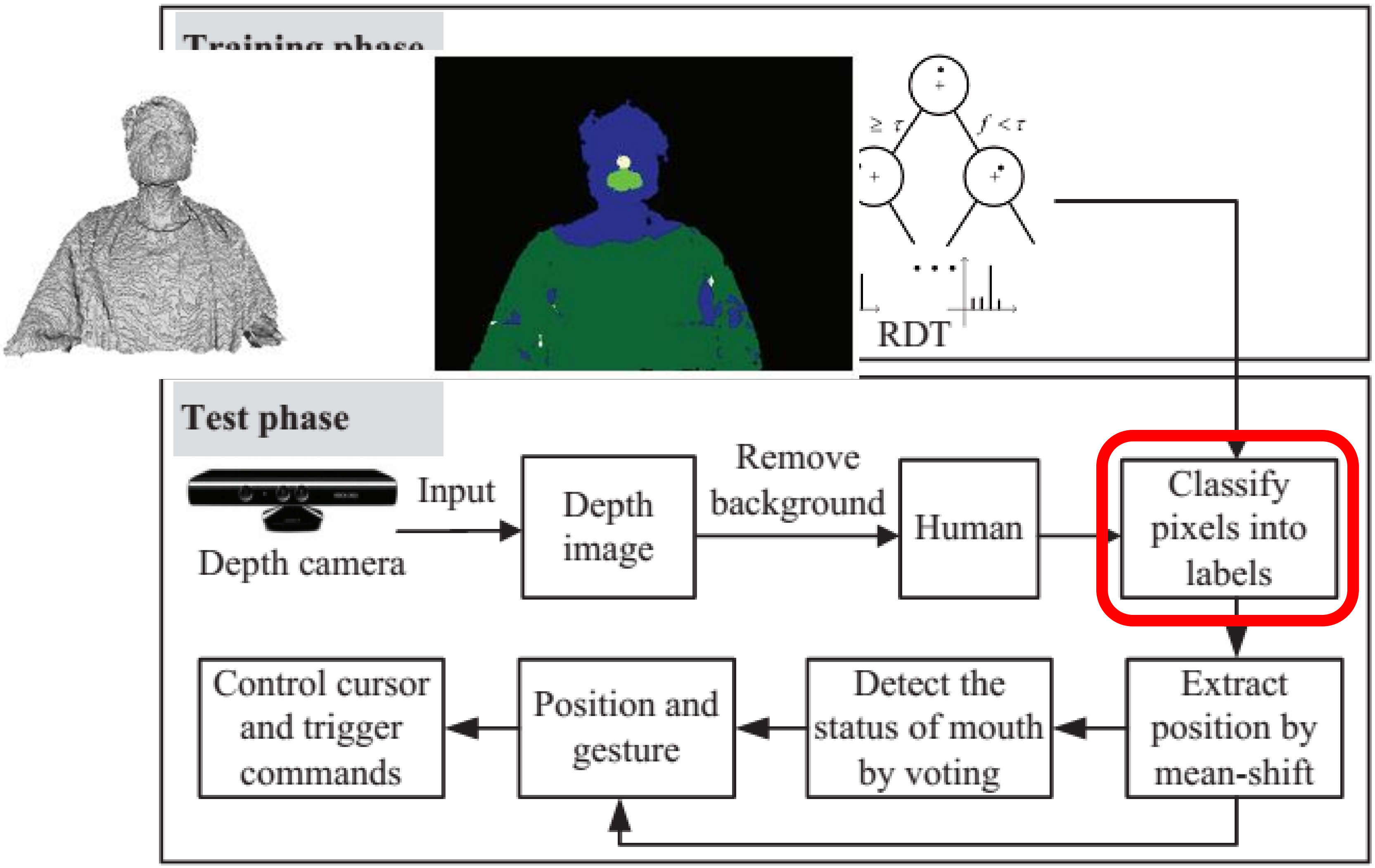


# Feature of randomized decision tree (RDT)

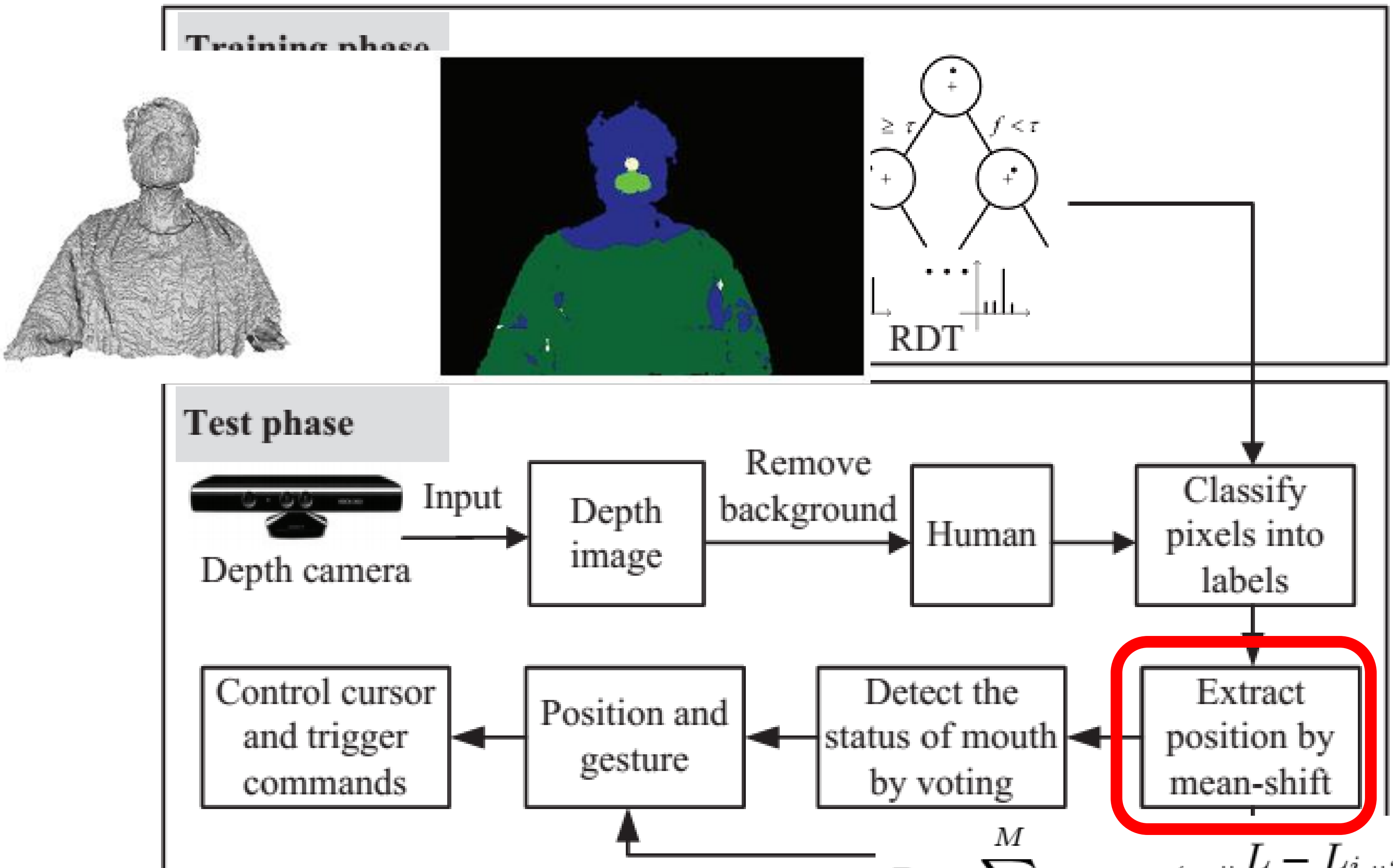
$$f((x_0, y_0)|(\Delta x, \Delta y)) = z(x_0, y_0) - z((x_0, y_0) + \frac{(\Delta x, \Delta y)}{z(x_0, y_0)})$$



# Classify pixels

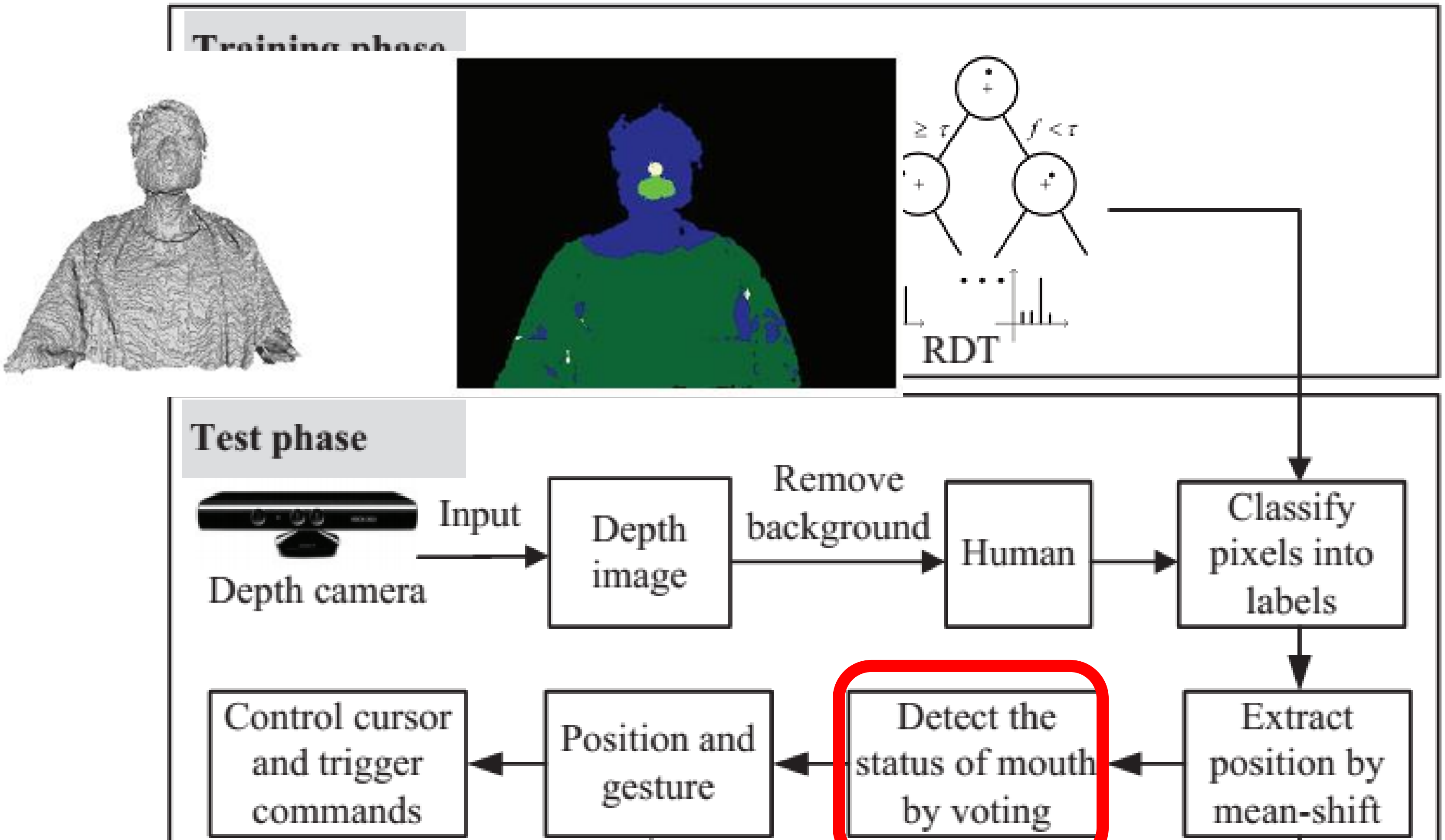


# Nose position detection




$$P_l \propto \sum_{i=1}^{M} w_{il} \exp(-\| \frac{L - L_i}{b_l} \|^2)$$

# Mouth status detection



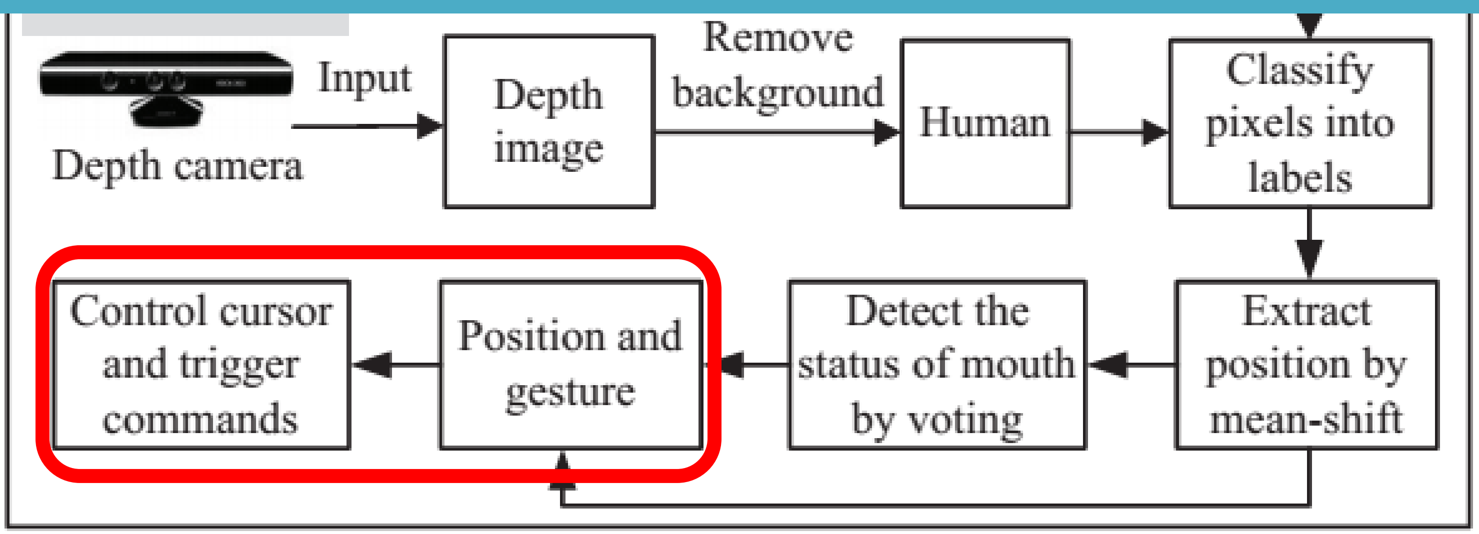

$$S_m^* = \arg\max_{S_m}(w_{S_m} \max_{L}(\sum_{i=1}^{M} w_{iS_m} \exp(-\|\frac{L - L_i}{b_{S_m}}\|^2)))$$

# Control function

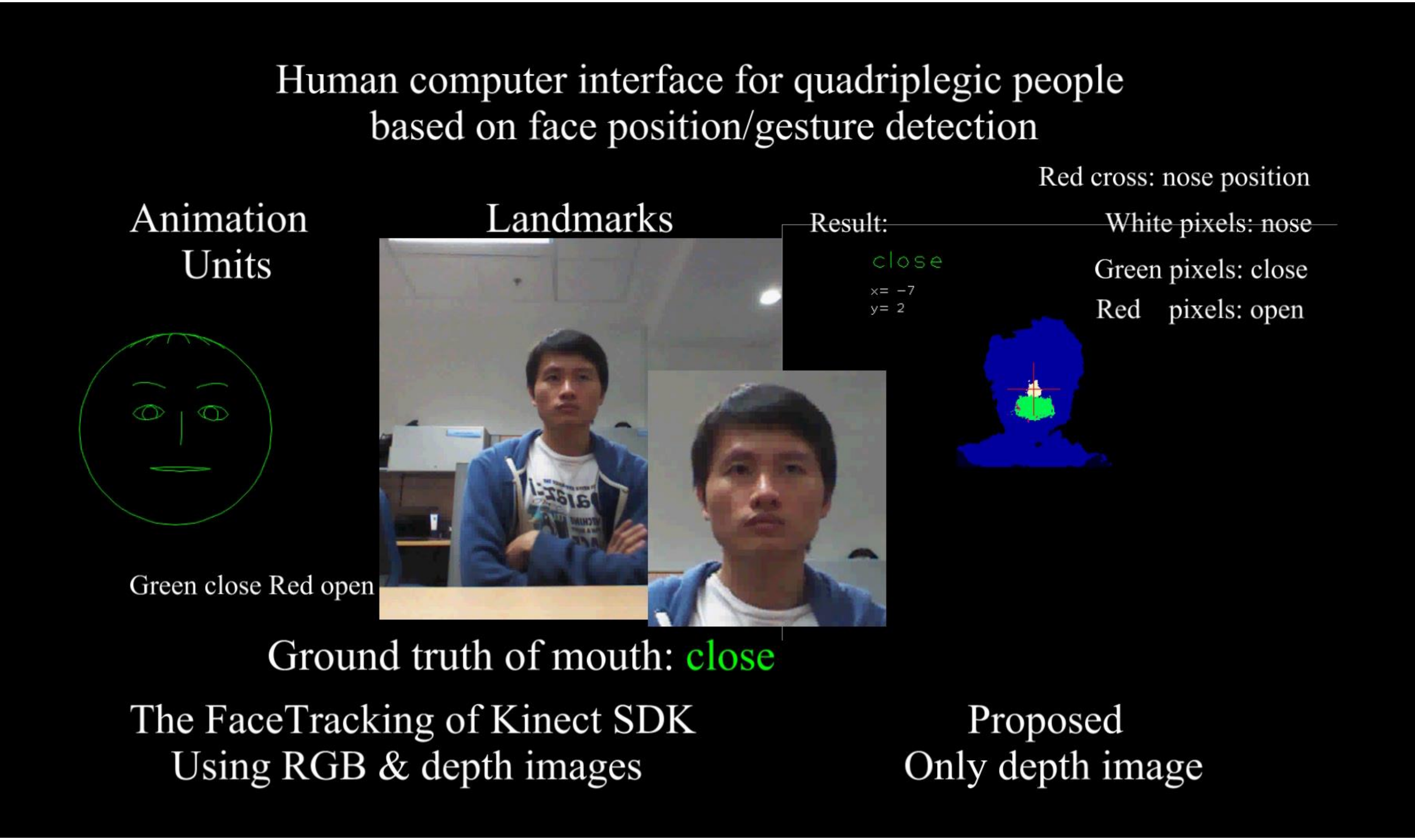
Close mouth– enable the movement of cursor
Open mouth– disable the movement of cursor

$$D_{x,cursor} = G_x * D_{x,nose}$$
$$D_{y,cursor} = G_y * D_{y,nose}$$

Single click-- open-close-open
Double click-- open-close-open-close



# Demo


Human computer interface for quadriplegic people
based on face position/gesture detection
Red cross: nose position
Animation
Units
Landmarks
Result:
close
x= -7
y= 2
White pixels: nose
Green pixels: close
Red pixels: open
Green close Red open
Ground truth of mouth: close
The FaceTracking of Kinect SDK
Using RGB & depth images
Proposed
Only depth image

# Evaluation

(a) (b) (c) (d) (e) (f) (g) (h)

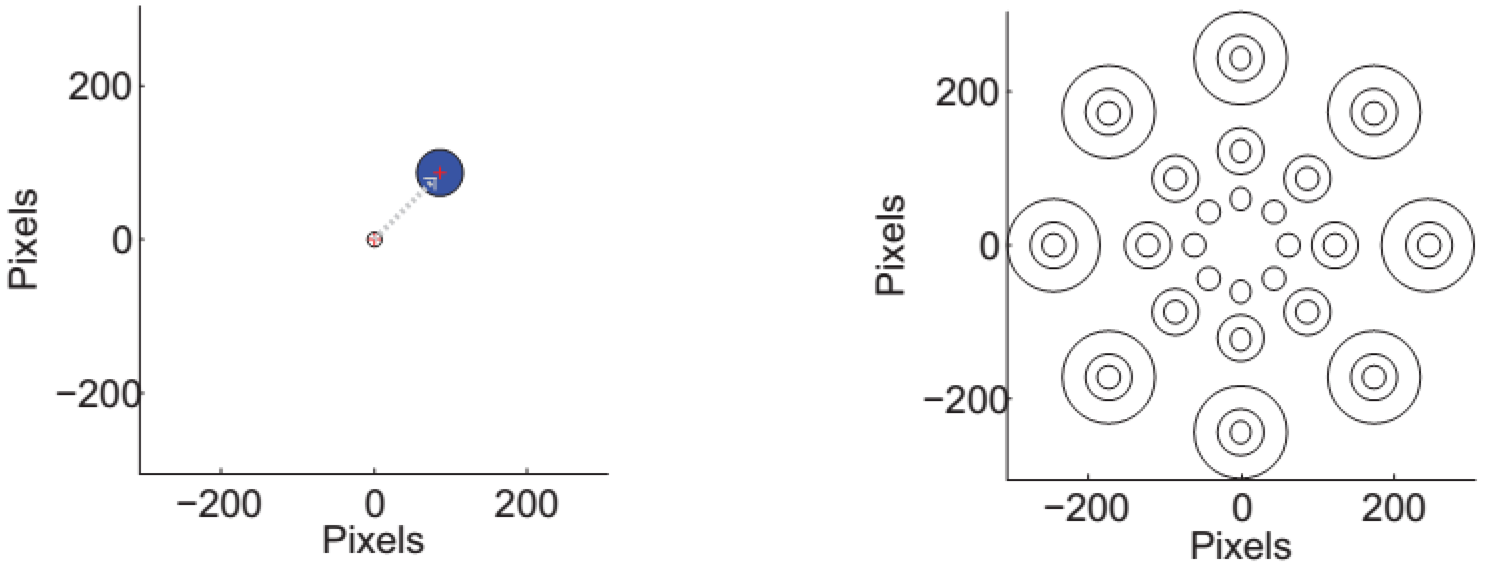
**Figure 3: Some examples of Bosphorus 3D face database. (a)-(d) Positive; (e)-(h)Negative.**

**Table 1: Errors of detection algorithms.**

| Error | Porposed | SVM1 | SVM2 | SVM3 | SVM4 |
|---|---|---|---|---|---|
| Positive | 2 | 6 | 4 | **1** | 35 |
| Negative | **5** | 6 | 12 | 17 | 9 |
| Total errors | **7** | 12 | 16 | 18 | 44 |
| Error rate(%) | **0.35** | 0.61 | 0.81 | 0.91 | 2.22 |

**The nose mean error of proposed method is 4.5 mm in 3D.**

# Evaluation




(a) (b)

GUI screen for center-out tapping (pointing and selecting). (a)One randomly appear target. (b)48 possible targets.

The tests main objectives are to obtain the throughput and task completion time.

# Evaluation

- **Throughput**: TP indicates that subjects deliver the amount of information to the computer though an interface, which is defined as

$$TP = \frac{ID}{MT},$$

- ID is the index of difficulty of the target
- MT is the movement time
- Higher TP is better.
- **Task Completion Time**: TCT is the total time to complete each round.
- Lower TCT is better.

# Evaluation

Results of the interface operation

| | FMf | FMo | FMc | CameraMouse | Mouse |
|---|---|---|---|---|---|
| TP(bit/s) | 1.8 | 2.3 | 2.3 | 1.5 | 4.4 |
| TPn(%) | 121 | 155 | 155 | 100 | 293 |
| TCT(min) | 5.1 | 4.6 | 4.2 | 5.6 | 2.3 |
| TCTn(%) | 91 | 82 | 75 | 100 | 41 |

Proposed interface FMf, FMo and FMc
FMf: freeze the cursor on the target as click
FMo: open the mouth to freeze the cursor
FMc: click by mouth action

CameraMouse: M. Betke, et al.. The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. IEEE TNEUR SYS REH, 2002.

# Conclusion

- We proposed a human computer interface for quadriplegic people based on face position/gesture detection
- It is

1. comfortable (no attachment on body),
2. easy to use (no initialization, no calibration, no guardian)
3. robust (insensitive to colour or illumination).

# Q & A