DogBOT

Team 6 –
Kuan-Ting Yu
Wei How Mo
Ping-Che Shiao
Issue Lee

Outline

- Gesture recognition - Kinect
- Human following - laser range finder
- Chase ball - color camera
- Facial expression - emotion

Gesture

- Stereo Vision
  - Compare two images to extract 3D info.
  - But, depth image become broken for area with less texture

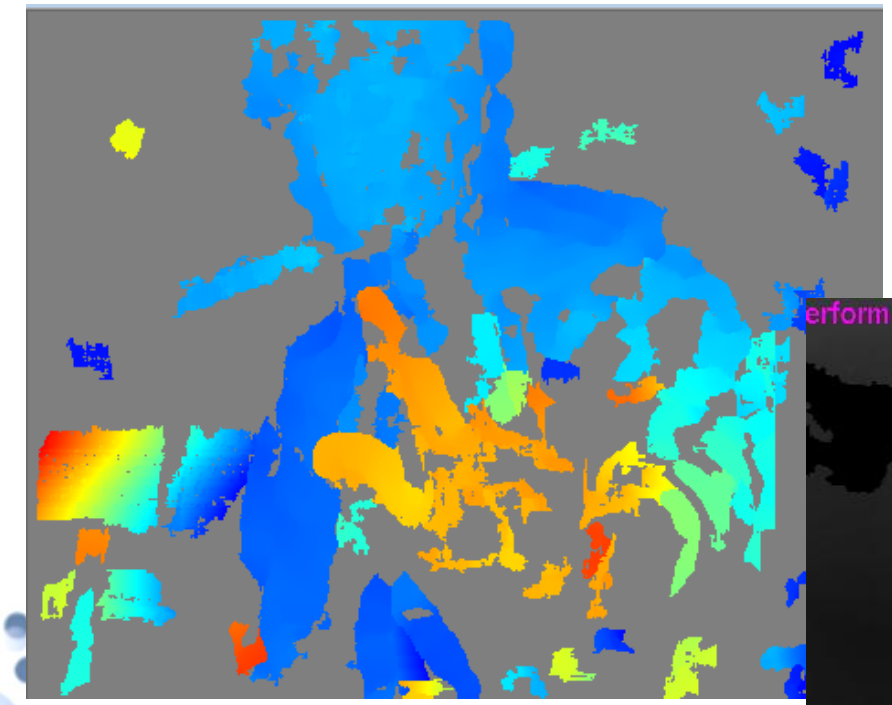
- Alternative: Kinect Sensor

Kinect

- Kinect is an active sensor which projects structured infrared light on objects.
  - use the dot pattern to derive 3D information
  - with built-in body parts tracker

http://boingboing.net/2010/11/23/dancing-with-invisib.html

Stereo v.s. Kinect



Gesture Recognition

- We have designed the finite state machine to recognize the 5 gestures for interaction.
- But, later we utilize Kinect's built-in gesture recognizer.

Gestures	State 1	State 2	State 3	State 4	Complete
HELLO					
INTRODUCTION					
COME					✓
CHASE					
STOP					

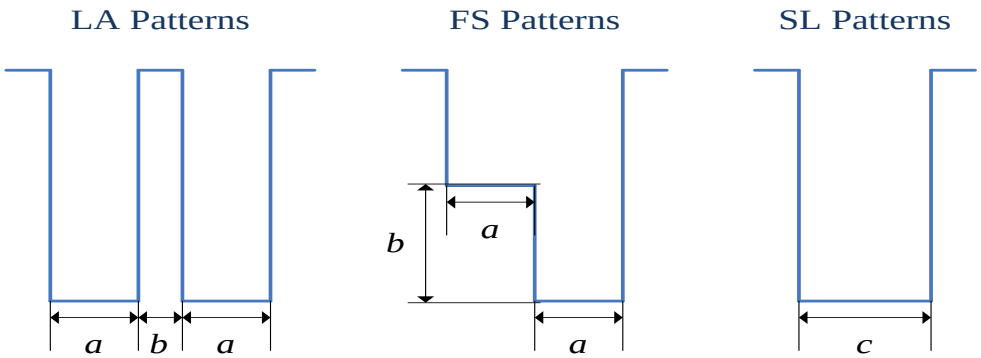
Human Following

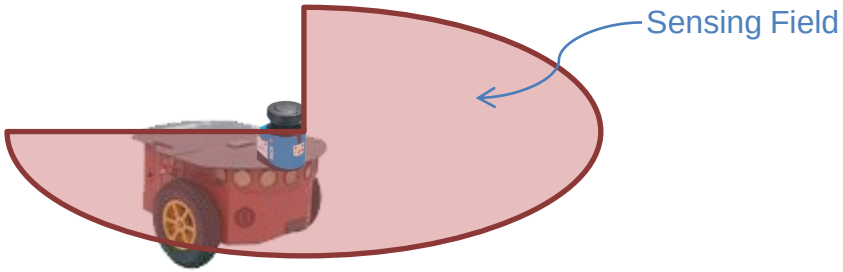
- Laser Scanner
  - ➢ High precision and scan rate
  - ➢ Invariant to different lighting conditions, perspective change, etc

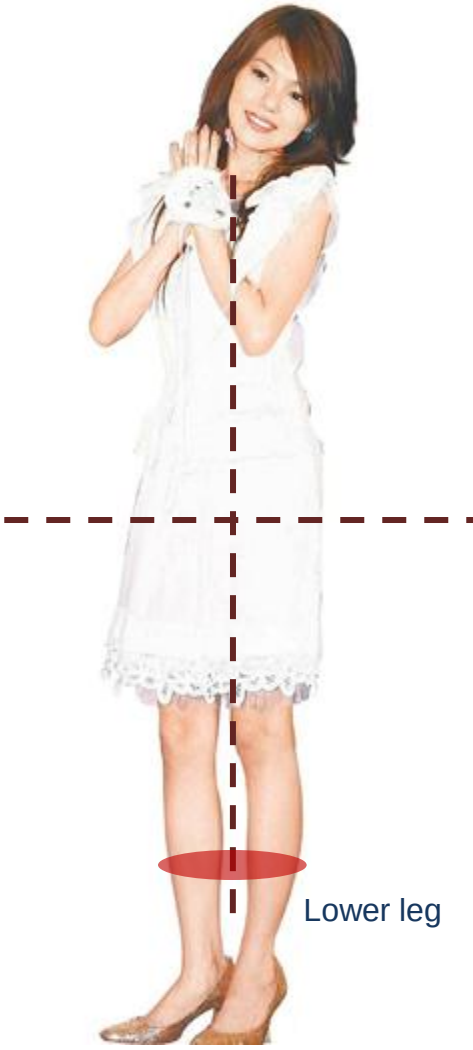
➢ Flow Chart:

Human Following

- Human detection using laser scanner
  - Detection position : lower legs





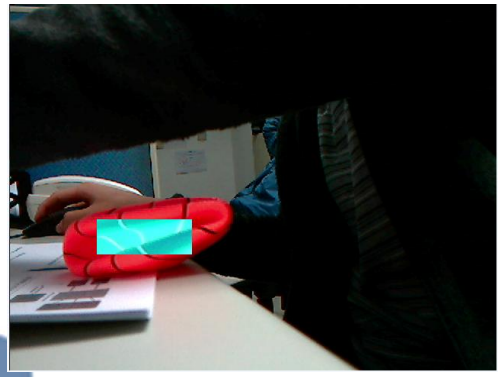


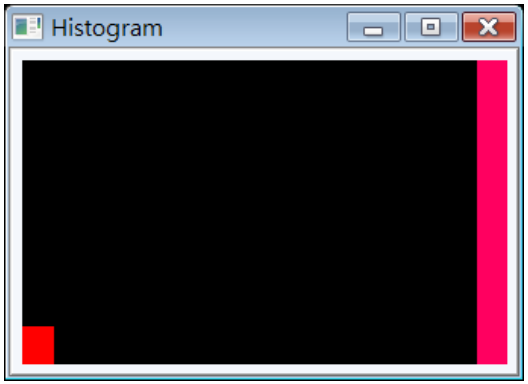
Human Following

- Obstacle Avoiding:
  - Laser based
  - Front of pioneer, degree 45 to degree 135

Ball chasing

1) calculate hue histogram of the ball
2) back projection on current image
3) thresholding and finding contour
4) coordinate transform from image to base

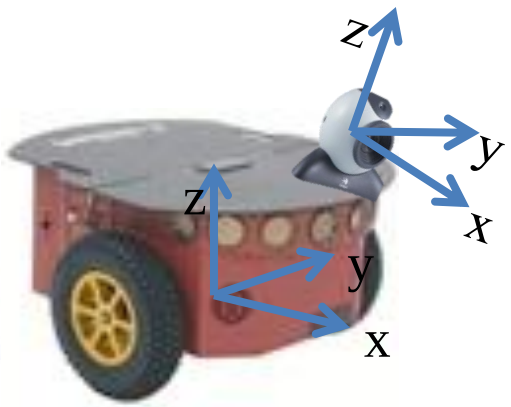




Coordinate transform

- To achieve higher level planning, precise target localization is a plus.

$$^{base}p = {}^{base}T_{cam}\,{}^{cam}T_{img}\,{}^{img}p$$



${}^{cam}T_{img}$ Recall

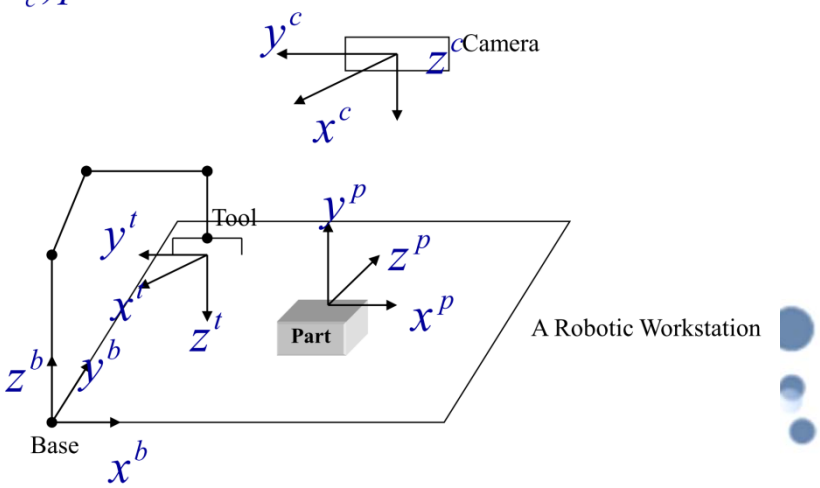
- Difficulty
  - The perspective transformation ${}^{image}T_{camera}$ is singular!

$${}^{camera}T_{image} \neq {}^{image}T_{camera}{}^{-1}$$

  - $p^{aug}(\lambda)$:augmented image coordinate of p^i
    - λ : is the depth of the corresponding p^c

$$p^{aug}(\lambda) = p^i - \left(\frac{\lambda}{\lambda - f}\right) z_c = \begin{bmatrix} p_1^i \\ p_2^i \\ -\frac{\lambda}{\lambda - f} \end{bmatrix} = Tran\left(\frac{\lambda}{f - \lambda} z_c\right) p^i$$

Depth is not fixed in our case!

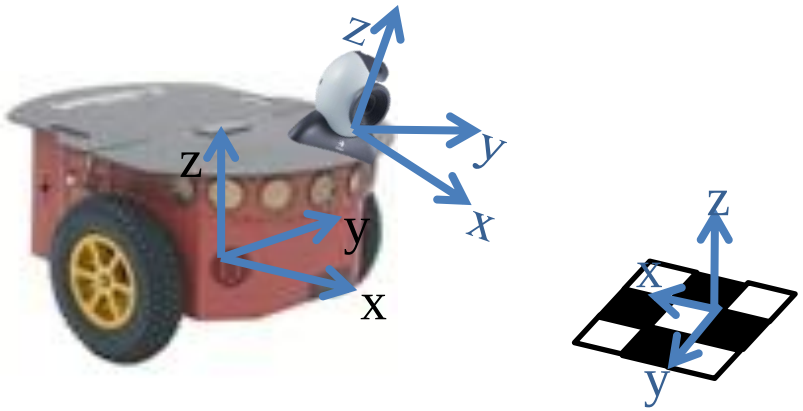


Our solution (1/2)

$$^{base}T_{cam}\ ^{cam}T_{board} = {}^{base}T_{board}$$

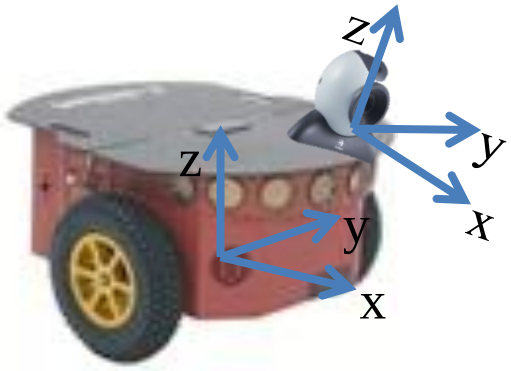
$^{cam}T_{board}$ is measured by calibration tool as extrinsic parameter.

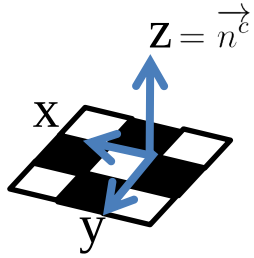
$^{base}T_{board}$ is measured by hand.



Our solution (2/2)

${}^{cam}T_{img}$, the inverse perspective transform, can be solved by extending the constraint of fixed depth, $\lambda = p_3^c$, to a general plane equation of the ground: $(\mathbf{p}^c - \mathbf{o}^c_{board}) \cdot \overrightarrow{n^c} = 0$





Experimental result

- The system can locate ball within 5m from camera with precision around 10cm.

Summery

- Gesture recognition - Kinect
- Human following - laser range finder
- Chase ball - color camera
- Facial expression - emotion