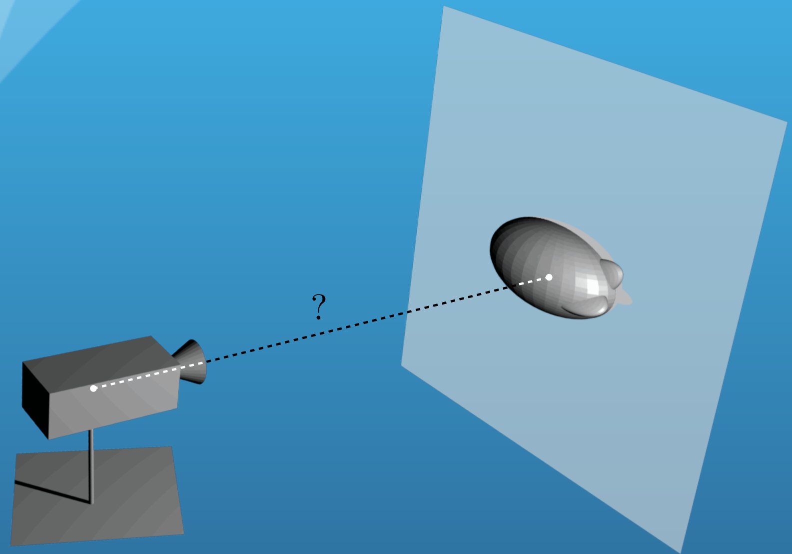
NEW DEPTH FROM FOCUS METHOD FOR 3D PTZ CAMERA TARGET TRACKING



Tiago Gaspar

(tgaspar@isr.ist.utl.pt)

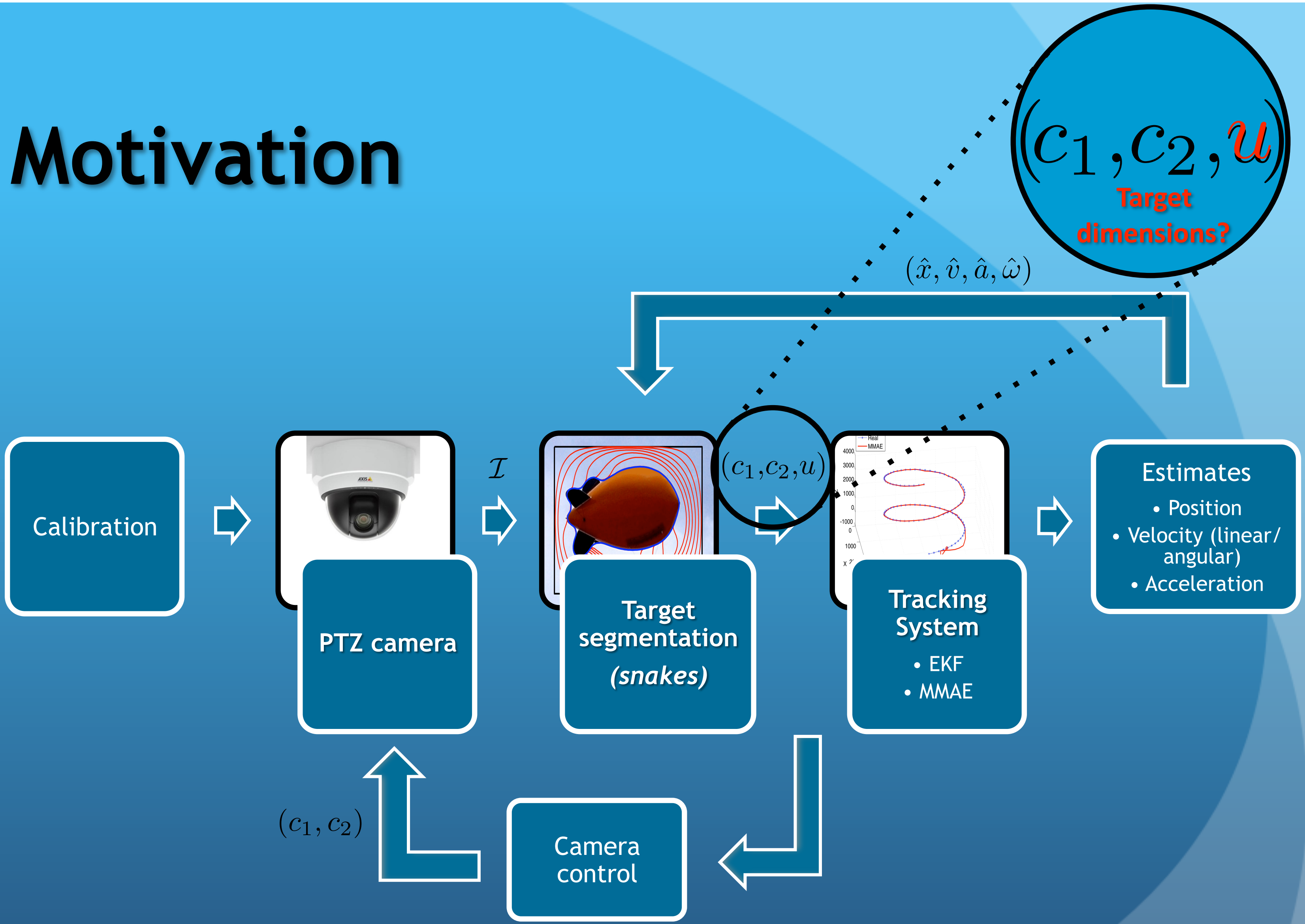
INSTITUTO SUPERIOR TÉCNICO

DSOR
dynamical systems and ocean robotics LAB

Presentation Outline

- Motivation / State of the art
- **Background on defocus theory**
- **Minimum blur focus value**
- **Depth estimation**
- **Experimental results**
- Conclusions

Motivation



(c_1, c_2, u)
Target dimensions?
$(\hat{x}, \hat{v}, \hat{a}, \hat{\omega})$
Calibration
$\mathcal{I}$
PTZ camera
Target segmentation (*snakes*)
(c_1, c_2, u)
Tracking System
• EKF
• MMAE
Estimates
• Position
• Velocity (linear/ angular)
• Acceleration
(c_1, c_2)
Camera control

State of the art

Depth estimate from blur by focusing

Depth from defocus

- ***Relation between depth and the amount of blur present in the image***
  - Model the blurring process of the camera, i.e. characterize the Point Spread Function (PSF)
  - Measure the amount of blur in the image: ill-posed inverse problem
  - Only needs 1 or 2 images

Depth from focus

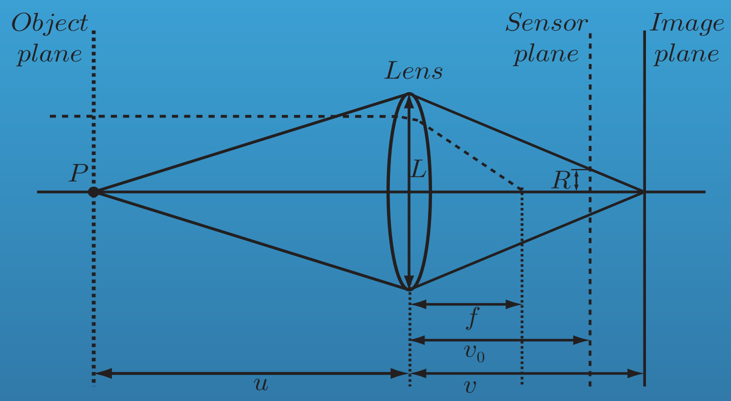
- ***Relation between depth and the focus value that leads to a sharp image***
  - Find the camera focus value that minimizes the amount of blur in the image according to a given criterion
  - Does not require a mathematical model for the camera blurring process (PSF does not have to be modeled)
  - Needs more images

Background on defocus theory (1)

Depth of field (of a camera with a given focus value):

• distance between the farthest and nearest planes, in relation to the camera, whose points appear with a satisfactory definition in the image, according to a given criterion

• cameras are unable to simultaneously focus planes at different depths



➢ Geometrical optics (image formation process)

➢ Thin lens model

➢ $v > v_0$

Gaussian Lens Formula: $$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

$$u = \frac{f v_0}{2Rf/L + v_0 - f}$$

An expression similar to this would be easily derived for the case $v < v_0$

Background on defocus theory (2)

$$u = \frac{f v_0}{2Rf/L + v_0 - f}$$

?

Depth from focus (DFF)

Find the sensor plane position (v_0) that leads to a sharply focused image, i.e. R = 0

- Optimize a cost function that depends on the amount of blur present in image points of interest

$$u = \frac{fv}{v - f}$$

Depth from defocus (DFD)

Estimate R by measuring the amount of blur present in image points of interest

- Express the defocused image (I_d) as the convolution of the focused image (I_f) with the PSF (h)

$$I_d(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_f(\alpha,\beta)h(x-\alpha,y-\beta)d\alpha d\beta$$

- Retrieve the PSF from the defocused image: solve an ill-posed inverse problem ⇨ **R**

Minimum blur focus value (1)

How to find the camera focus value that sharply focus the target?

- define a cost function that depends on the amount of blur that "corrupts" the target
- find the camera focus value that optimizes the cost function

Cost function

System goal: estimate the depth of a target

↓

maximize the image gradient magnitude across lines orthogonal to the target boundary (active contours)

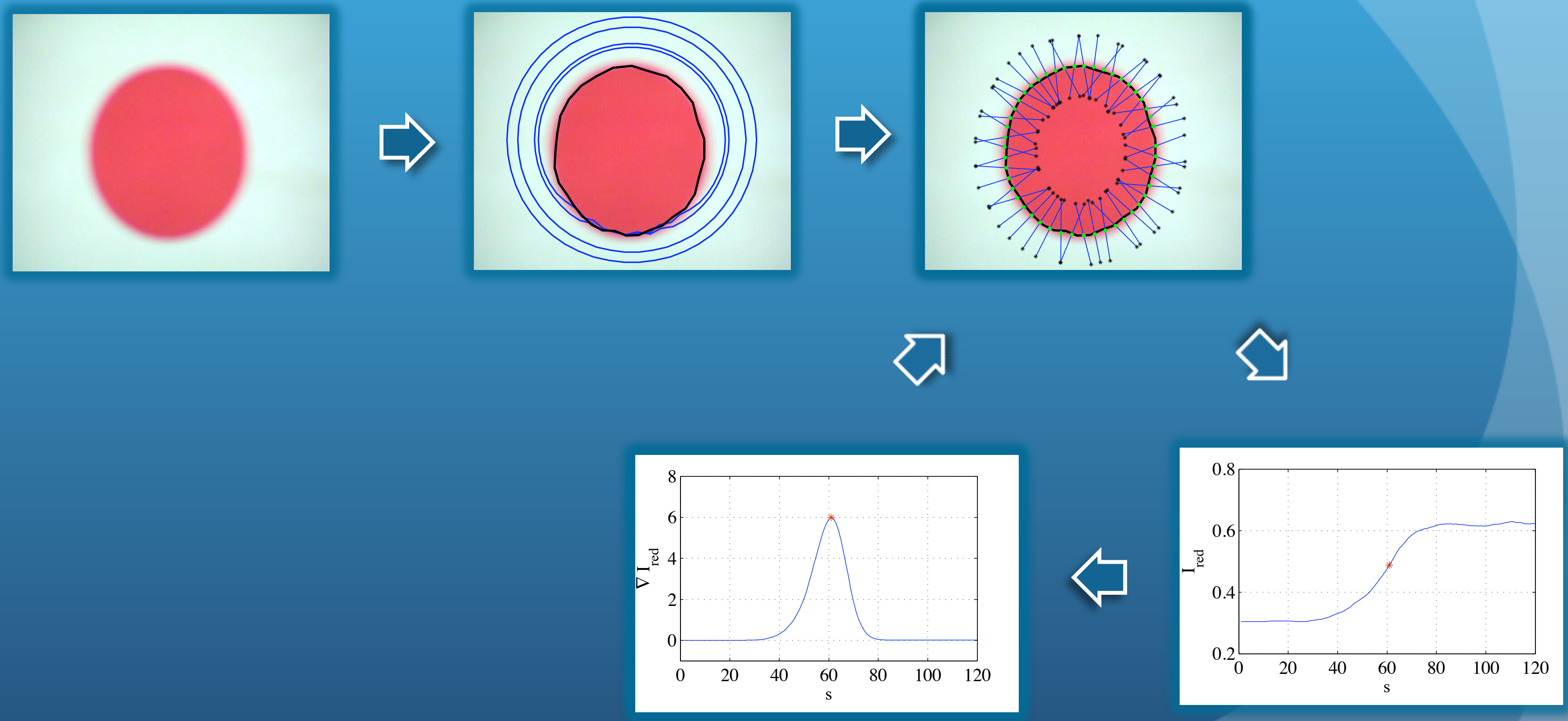
$$\min_{f_s} g(f_s) \Rightarrow g(f_s) = \frac{1}{\frac{1}{N_l}\sum_{i=1}^{N_l} \max_{(x,y)\in l_i} |\nabla I_{f_s}(x,y)|}$$

- f_s - focus setting
- I_{f_s} - image acquired with focus f_s
- (x, y) - image point coordinates
- l_i - i-th line
- N_l - number of lines

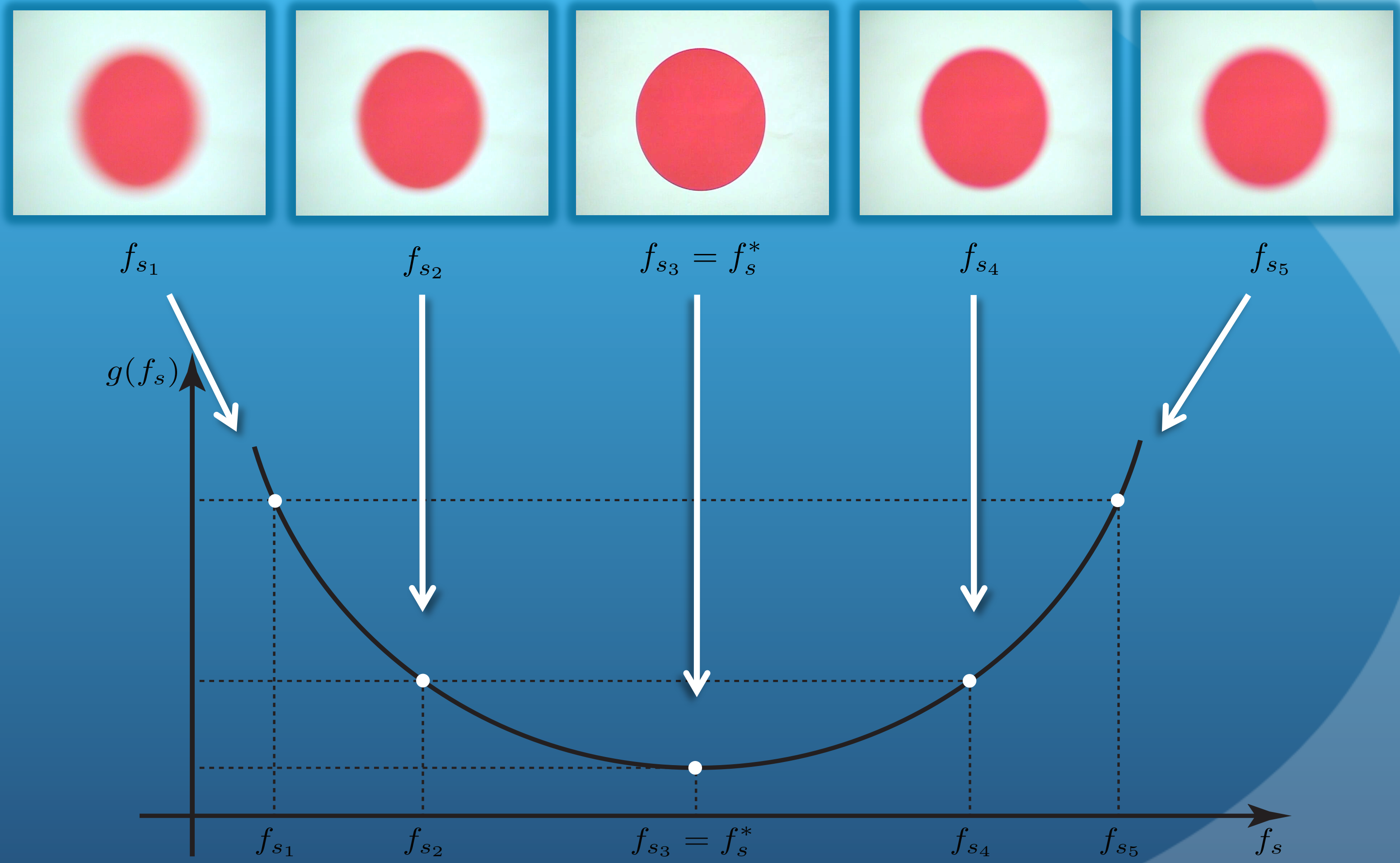
The formulation of this problem as the minimization of $g(f_s)$, instead of the maximization of its denominator, is related to the model that will be proposed for this function

Minimum blur focus value (2)

$$g(f_s) = \frac{1}{\frac{1}{N_l}\sum_{i=1}^{N_l} \max_{(x,y)\in l_i} |\nabla I_{f_s}(x,y)|}$$



Minimum blur focus value (3)



Minimum blur focus value (4)

Minimize the cost function

Data available is scarce (the acquisition of new images is time-consuming)

Moving target

↓

It is difficult to find the focus that minimizes the cost function in real time

Our approach: find a model for the cost function and estimate the position of its minimum using a small number of images

AXIS 215 PTZ

$$g(f_s) = a(f_s - f_s^*)^2 + b$$

⇩

$$g(f_s) = a' f_s^2 + b' f_s + c'$$

- estimate: a, b, f_s^* (3 unknowns)
- f_s^* - focus value that minimizes the cost function

$$a' = a$$
$$b' = -2af_s^*$$
$$c' = af_s^{*2} + b$$

Minimum blur focus value (5)

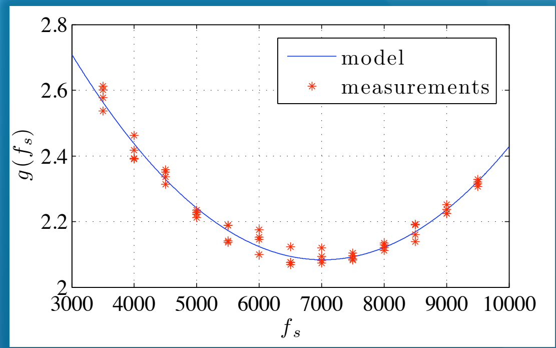
$$\min_{\mathbf{x}} ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2$$

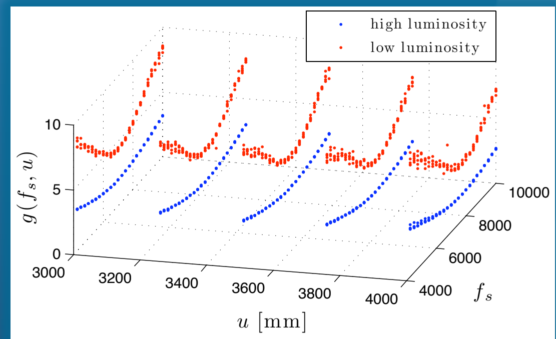
$$\begin{matrix} I_1 \rightarrow \\ I_2 \rightarrow \\ \\ I_N \rightarrow \end{matrix} \begin{bmatrix} f_{s_1}^2 & f_{s_1} & 1 \\ f_{s_2}^2 & f_{s_2} & 1 \\ & \vdots & \\ f_{s_N}^2 & f_{s_N} & 1 \end{bmatrix} \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} \quad \begin{bmatrix} g(f_{s_1}) \\ g(f_{s_2}) \\ \vdots \\ g(f_{s_N}) \end{bmatrix}$$

$$f_s^* = -\frac{b'}{2a'}$$

solved using: least squares method (optimal when the measurements are corrupted by white Gaussian noise)

(u = const.)





Considerations about the method:

- this algorithm must be repeated successively since the depth of the moving target (and therefore f_s^*) varies with time
- 3 unknowns: use at least 3 images acquired with different focus values
- this method is robust to variations in parameters such as zoom, aperture, and illumination, which change the shape of the cost function but not the position of its minimum

Depth estimation (1)

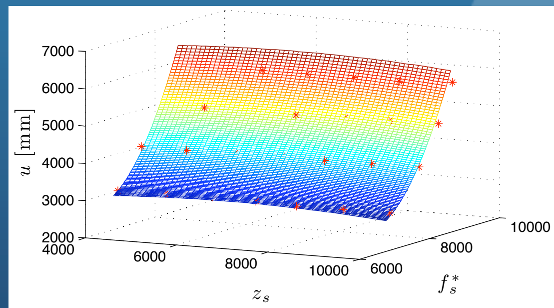
$$f_s^* \xrightarrow{?} u$$

1. $(f_s^*, z_s) \xrightarrow[v = h_2(f_s^*, z_s)]{f = h_1(f_s^*, z_s)} (f, v) \xrightarrow{u = \frac{fv}{v-f}} u$

? h_1 and h_2 difficult to calibrate

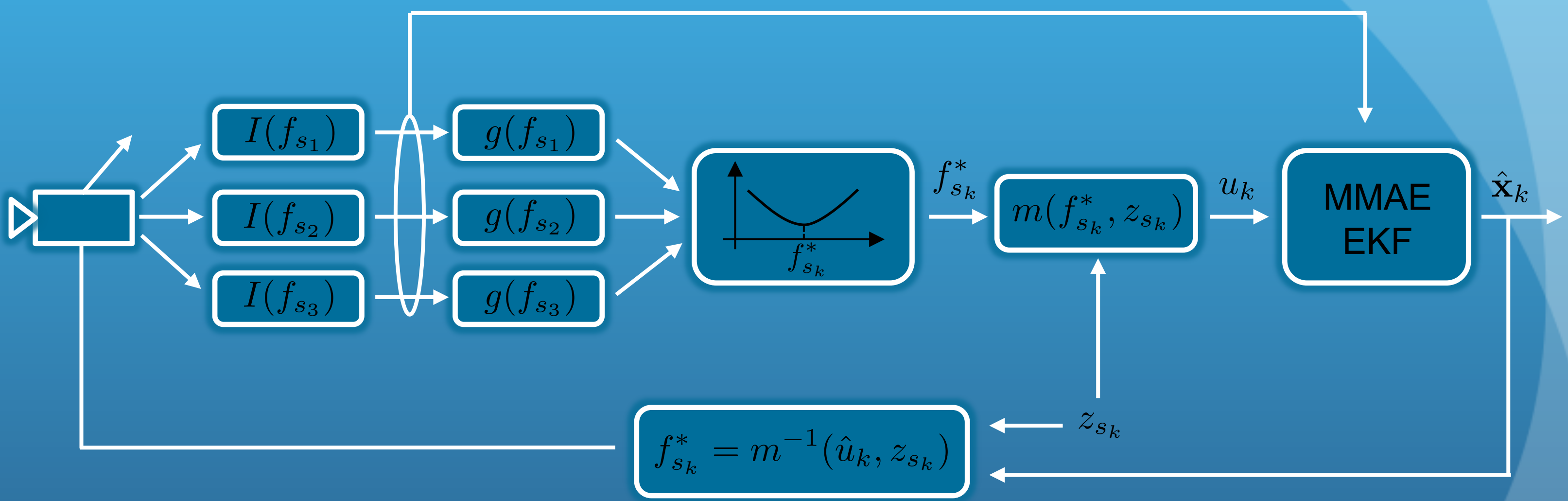
2. $(f_s^*, z_s) \xrightarrow{u = m(f_s^*, z_s)} u$

$$u = m(f_s^*, z_s) = a_1 z_s^2 + a_2 z_s + a_3 {f_s^*}^2 + a_4 f_s^* + a_5$$



Depth estimation (2)

System architecture:



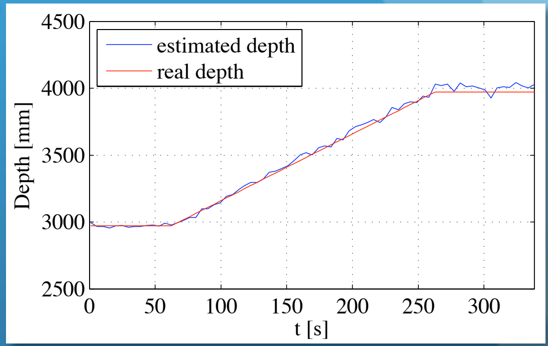
$$\begin{cases} f_{s_1} = f^*_{s_{k-1}} - \Delta \\ f_{s_2} = f^*_{s_{k-1}} \\ f_{s_3} = f^*_{s_{k-1}} + \Delta \end{cases}$$

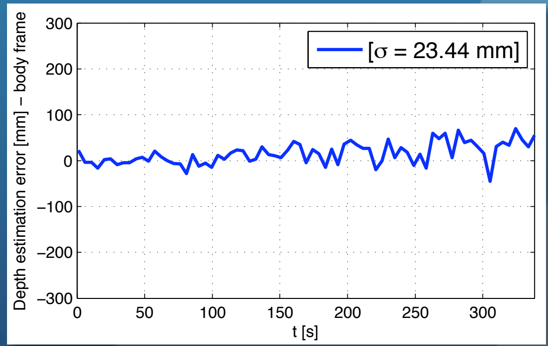
Experimental results (1)

Application description:

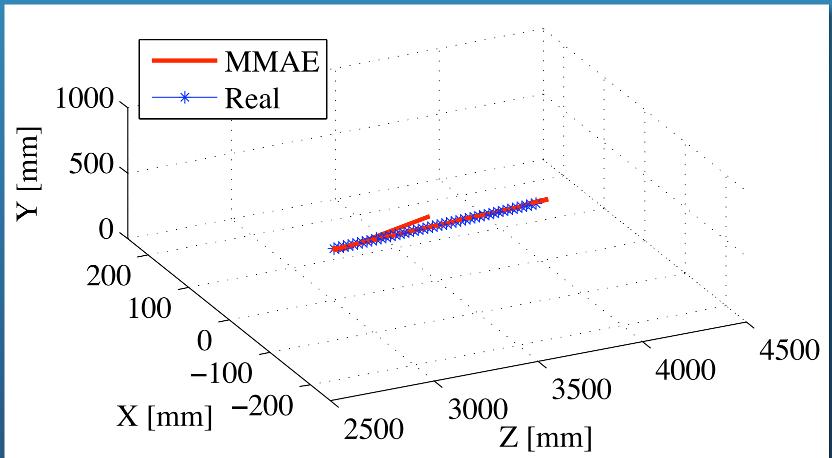
- camera: AXIS 215 PTZ
  - maximum focal length: 45.6 mm (aperture: F2.7)
  - focus setting: [1; 9999]
  - image size: 704x576
- target installed on a mobile platform
- MMAE: 4 models (4 different angular velocities)
- sampling interval: ~ 4.5s

Depth estimation





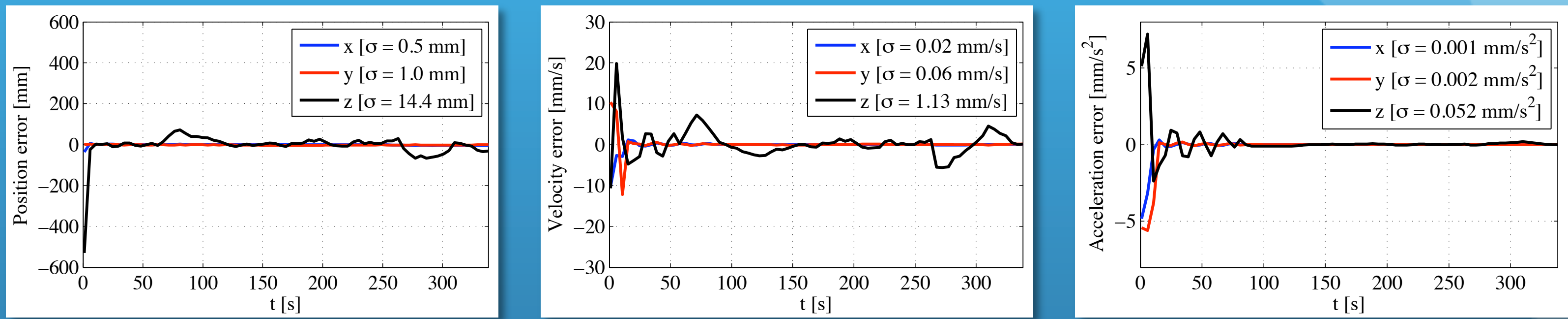
Application performance:



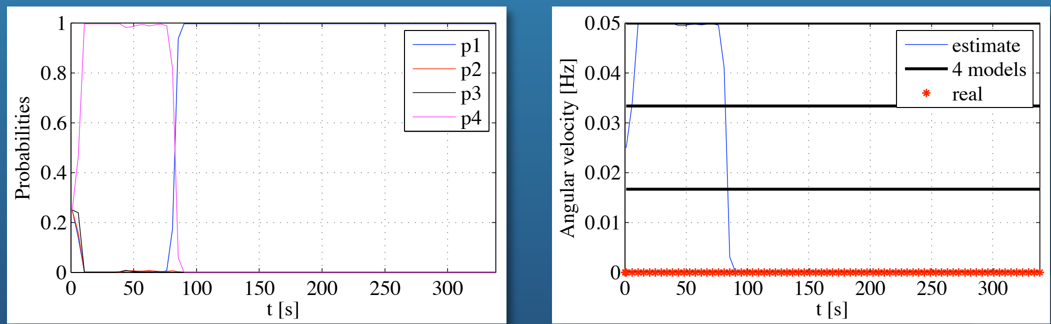
Experimental results (2)

MMAE - EKF

- ***Estimation errors***



- ***Angular velocity identification***



Conclusions

- **A new active depth from focus method** that estimates the depth of a moving target in real time, using a single PTZ camera, was presented, implemented, and tested in a real scenario
  - information present on the target boundary is used to infer depth, which is combined with a (sud-)optimal MMAE-EKF resulting in a full 3D nonlinear tracking system
  - depth and position estimates with accuracies on the order of few centimeters were obtained

- **Main limitation:** slow rate at which the target can change its depth

  (a consequence of)
  - slow velocity at which the camera acquires images with different focus values
  - 3 images per iteration of the algorithm

- **Future work:**
  - implement the algorithm in C
  - study new strategies to reduce the number of images required per iteration of the algorithm

Thank you.