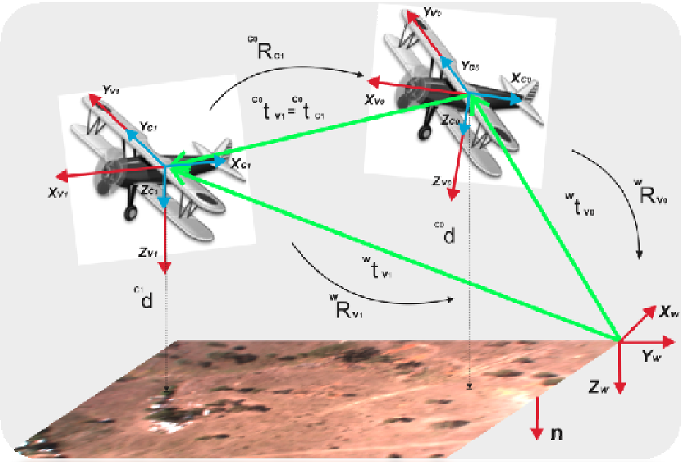
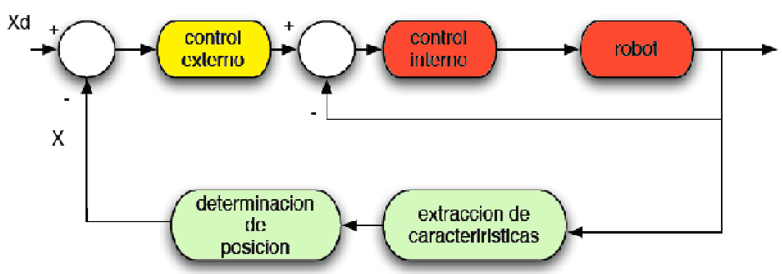
Curso: Técnicas Avanzadas de Visión por Computador
Máster Automática y Robótica

Pose Estimation and Position Based Visual Servoing





Carol Martínez
May 2011

OUTLINE

- Introduction
- Position Based Visual Servoing PBVS
- Pose estimation techniques
- Results
- Summary
- References

INTRODUCTION

Problem

To track and follow the car with one of the **ROBOTS** we have (ground or aerial)

Visual servo control,
Visual servoing:
The use of computer vision data to control the motion of a robot

INTRODUCTION

Problem

To track and follow the car with one of the **ROBOTS** we have (ground or aerial)

To achieve the task we need: detection, segmentation, tracking, recognition, alignment-visual servoing.

INTRODUCTION

Problem

To track and follow the car with one of the **ROBOTS** we have (ground or aerial)

To achieve the task we need**:**

- Robust perception

- Robust control

Perceptual Robustness

- Camera configuration static or moving
- Number of cameras
- Calibrations issues
- Image processing techniques

According to
K. Toyama and G. Hager, Incremental focus of attention for robust visual tracking, CVPR 1996

*Robustness is the **ability of a vision-based tracking** system to track accurately and precisely during or after visual **circumstances that are less than ideal**. … The robust vision-based tracking problem is therefore a vision-based tracking sub-problem – the problem of **coping with a complex environment***

Vision system requirements

1. **Handling** temporal inconsistencies in appearance and **occlusions** of the target object

2. **Handling** situations when the object is **outside of the FOV** (reinitialization)

3. **Adapt** to unpredictable object motion

4. Be insensitive to **lighting conditions** and specular reflections

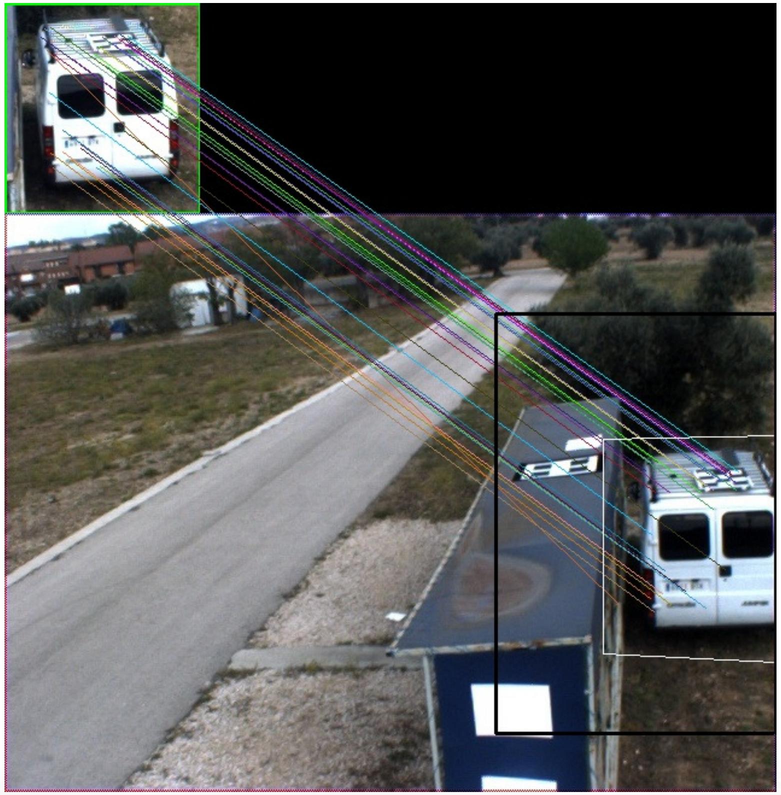
5. **Detect errors** (in tracking or detection) and **to recover** the tracking afterwards

6. Produce estimates in **Real-Time**

7. Use **minimum a-priori knowledge** about the object

Lack of robustness due to

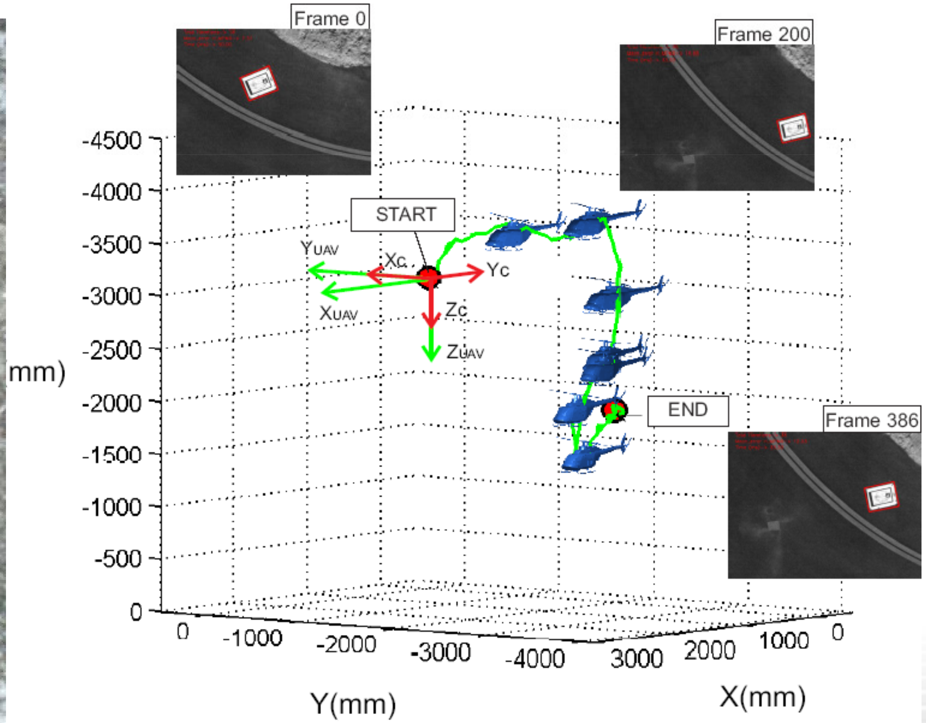
1- Figure-ground **segmentation** (detection of the target or **initialization** of tracking sequence)



2- **Matching** across images (in particular in the presence of large and varying inter-frame motions)

Lack of robustness due to

3- **Inadequate modeling** of motion (to enable prediction of the target in new images)



Lack of robustness due to

3- **Inadequate modeling** of motion (to enable prediction of the target in new images)

3 parameters (Tx, Ty, Rz)

4 parameters (Tx, Ty, Rz,scale)

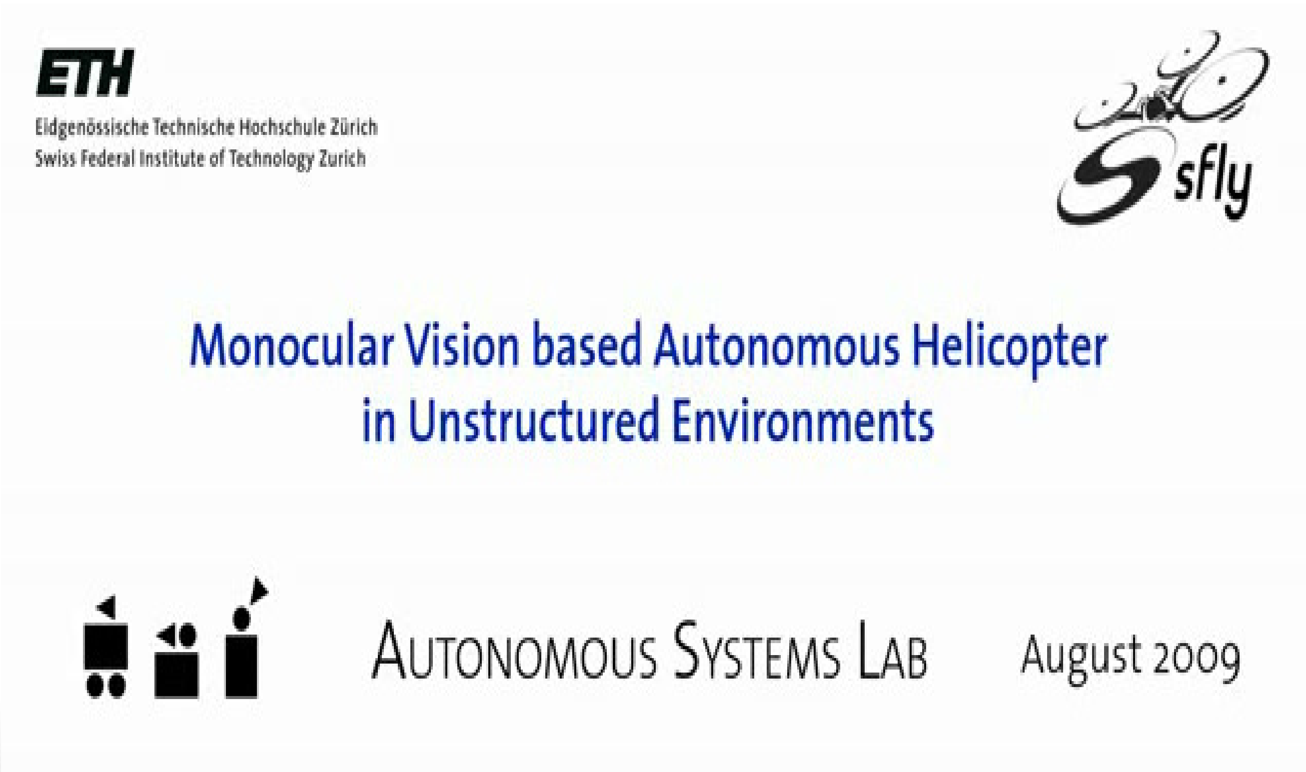
However …

There have been successful works using vision for controlling purpose

On-board cameras: monocular or stereo

However …

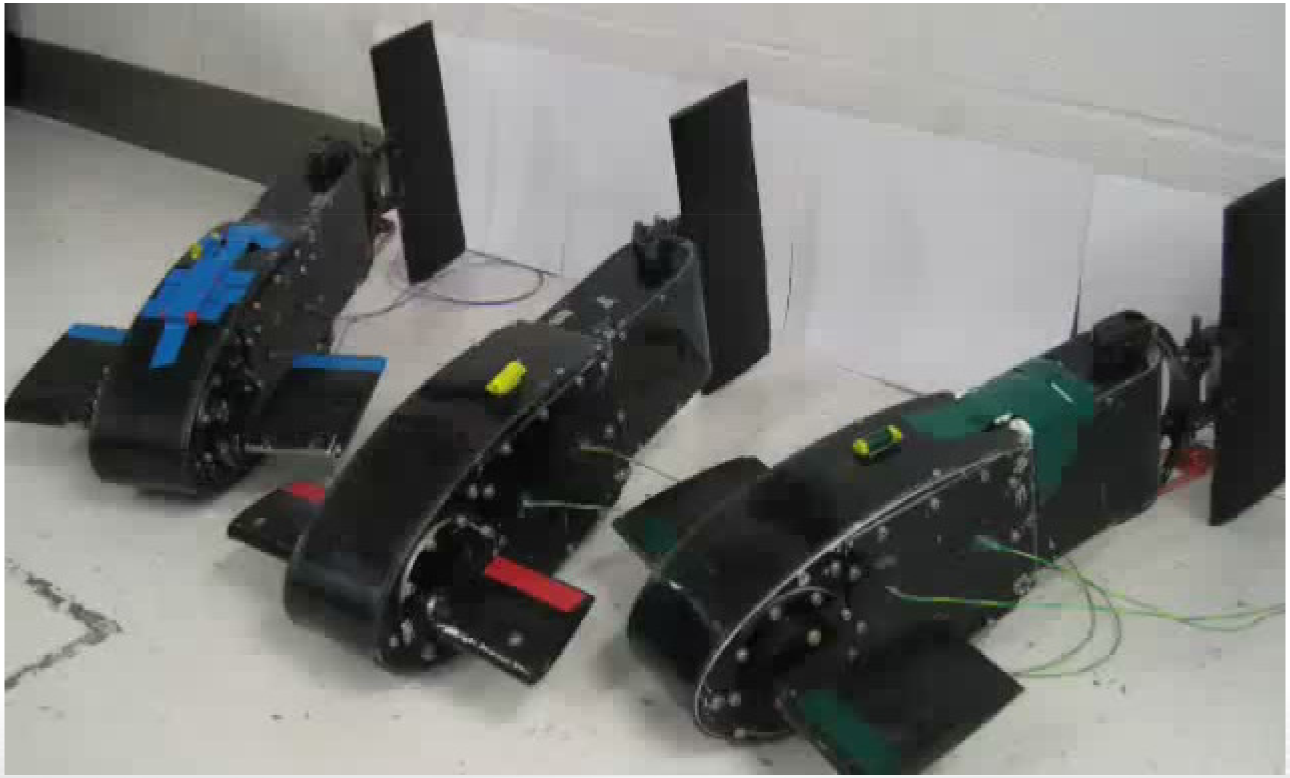
There have been successful works using vision for controlling purpose



On-board camera: monocular

However …

There have been successful works using vision for controlling purpose

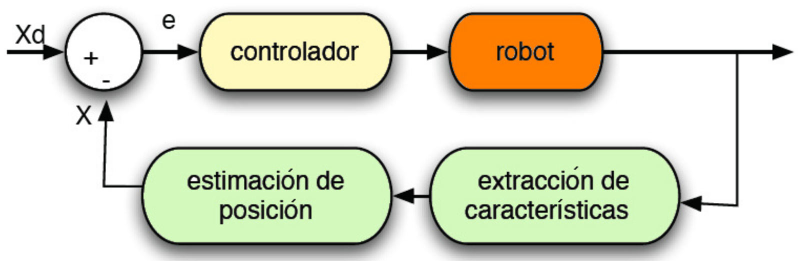


External camera systyem

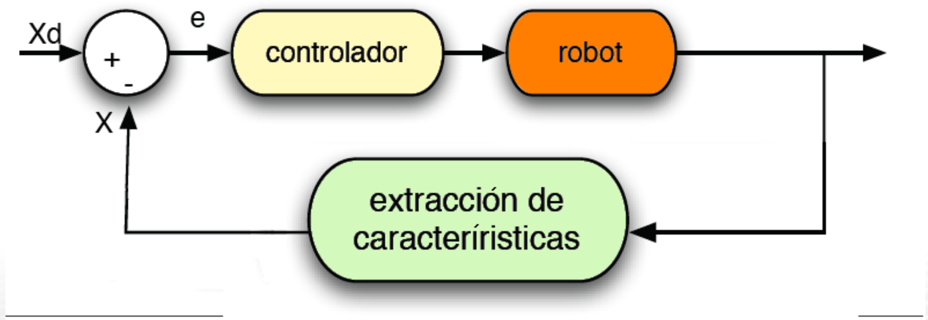
Position Based Visual Servoing PBVS

Classification of visual servoing techniques:

- Image-Based (IBVS)
- Position-Based (PBVS)
- Hybrid approach



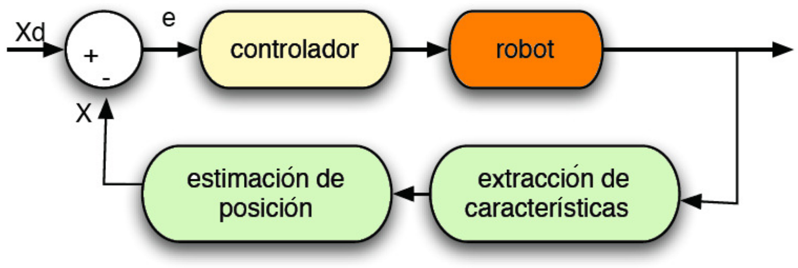
PBVS



IBVS

Position Based Visual Servoing PBVS

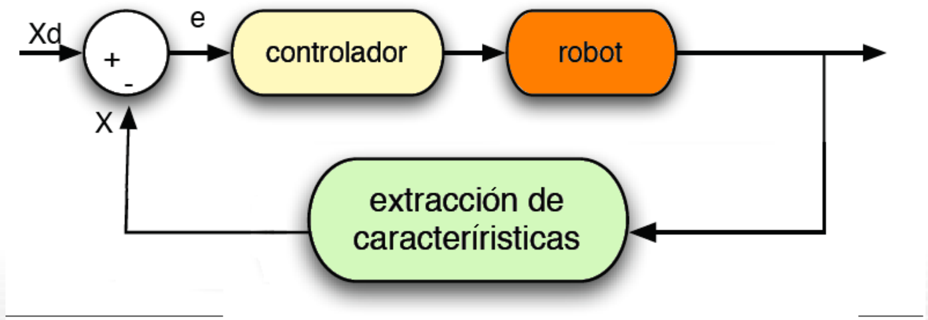
PBVS and **IBVS** are **different** in the nature of **the inputs used** in their respective control schemes



PBVS

Both approaches give **satisfactory results:**

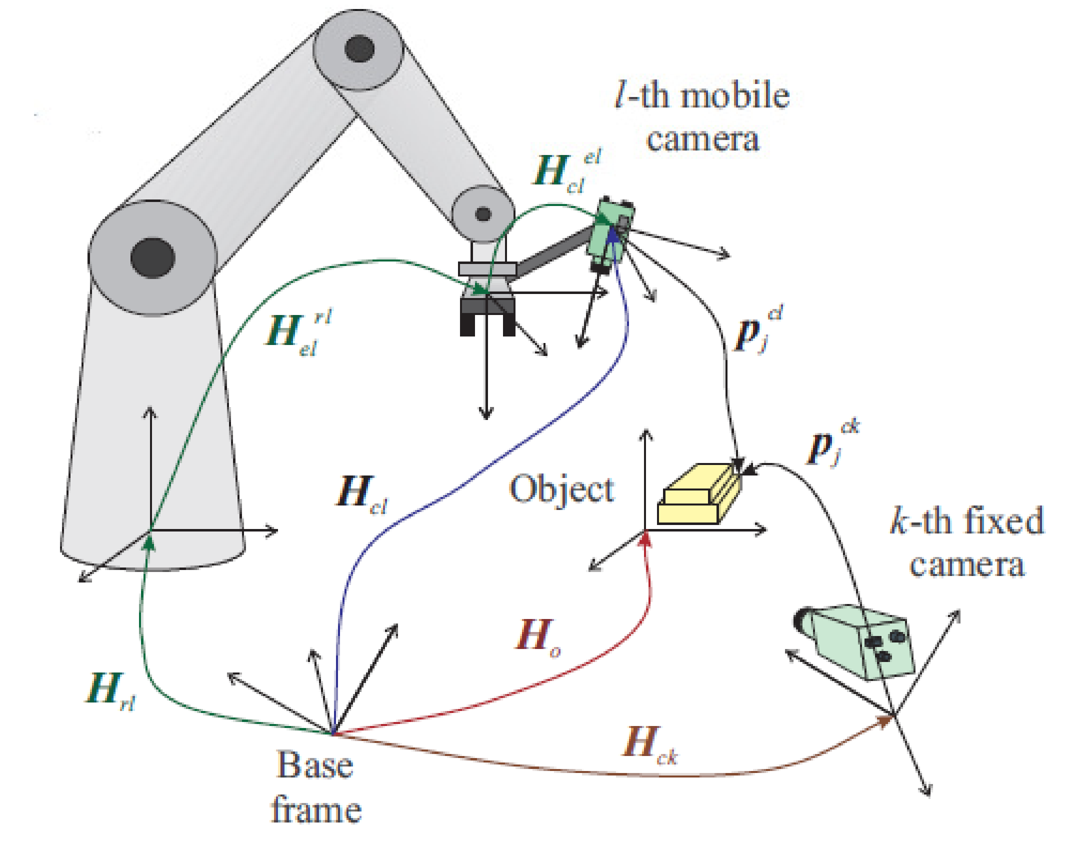
convergence, stability, robust to camera calibration errors, measurements errors.



IBVS

Position Based Visual Servoing PBVS

Depending on the camera-robot configuration



- Eye to hand
- Eye in hand
- Hybrid approach

Position Based Visual Servoing PBVS

Depending on the number of cameras

•**Monocular**:
There is a **lost of information** (depth), make the control more complicated.

Positioning tasks look for solving this problem:

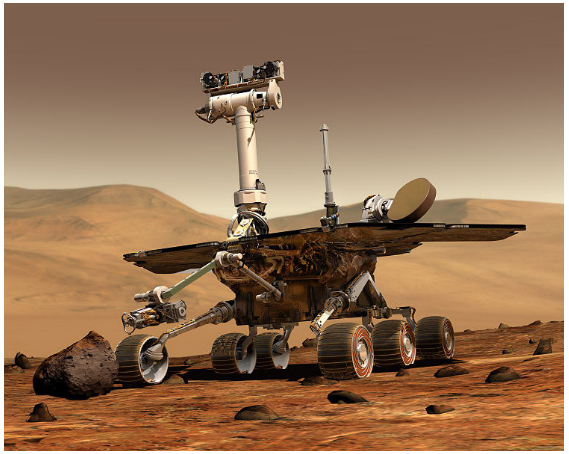
- Estimating depth before the tasks, or with metric information of the object.

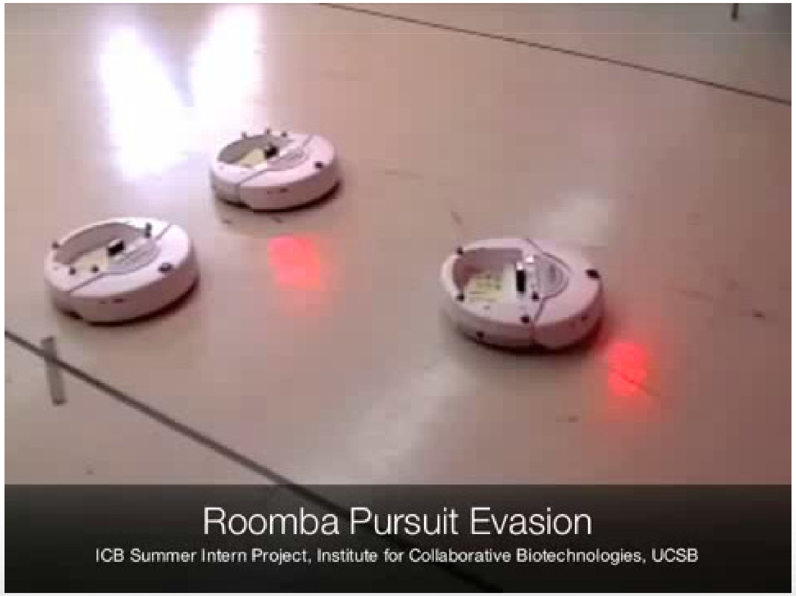
"Almost used with eye in hand configuration"

Position Based Visual Servoing PBVS

Depending on the number of cameras

Stereo:
3D information can be obtained



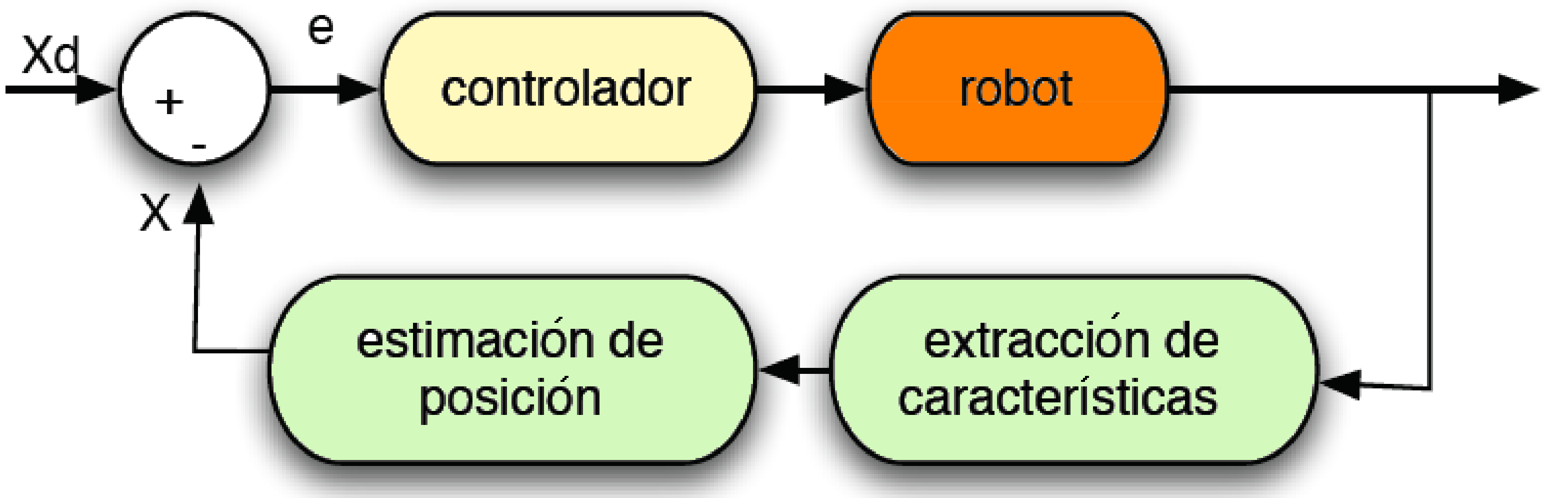


Two autonomous robots try to catch a remotely controlled evader

Redundant system:
3D information can be obtained.
Adding robustness.
Processing time increases

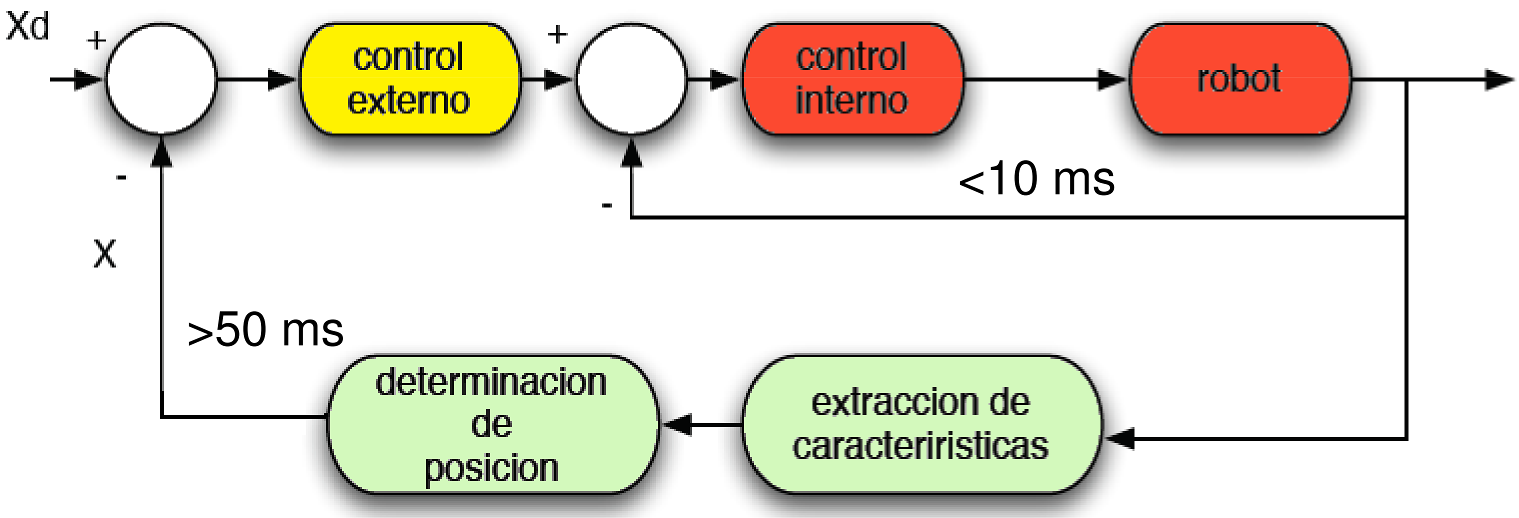
Position Based Visual Servoing PBVS

Control structure: direct visual control



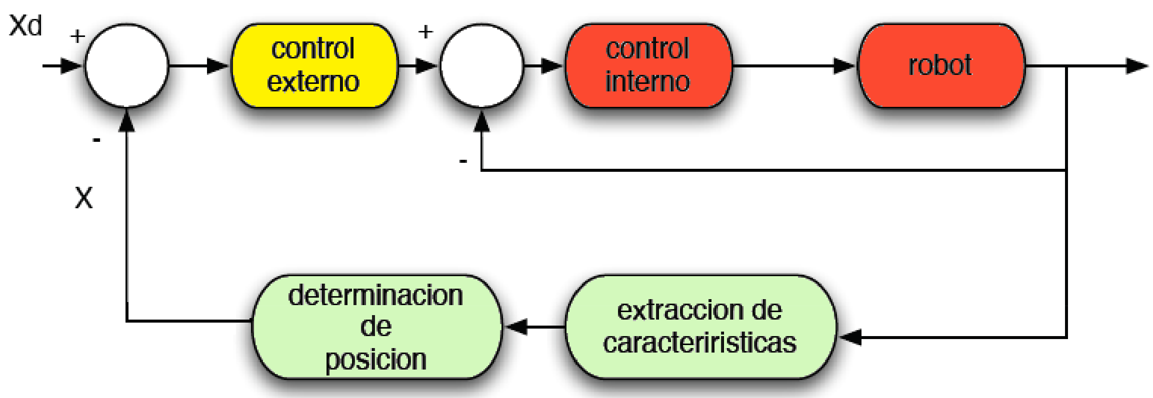
Position Based Visual Servoing PBVS

Control structure: indirect visual servoing, dynamic look and move



Position Based Visual Servoing PBVS

$$e(k) = p^*(k) - p(k)$$

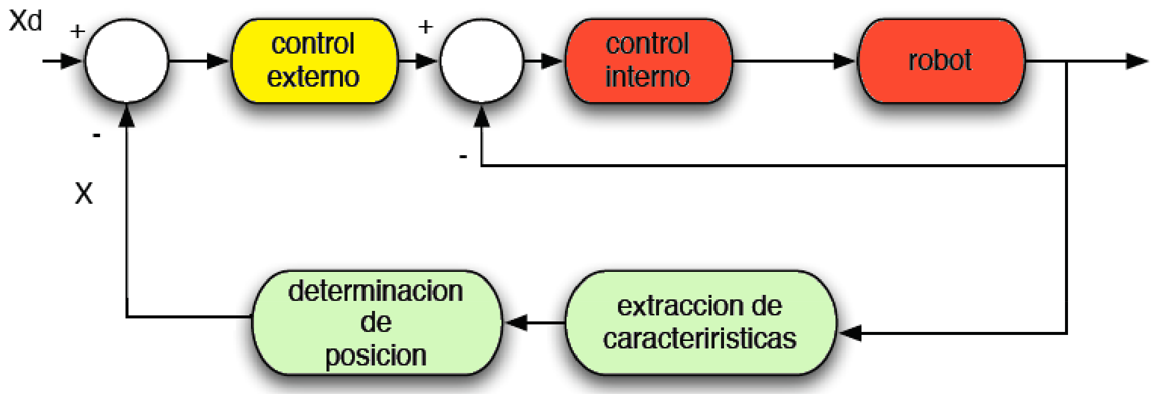


Error function based on the **3D** cartesian space, It is also called pose-based visual servoing.

Image features are extracted as well, but are additionally **used to estimated 3D** information (pose of the object in the cartesian space), hence it is servoing in 3D

Position Based Visual Servoing PBVS

$$e(k) = p^*(k) - p(k)$$

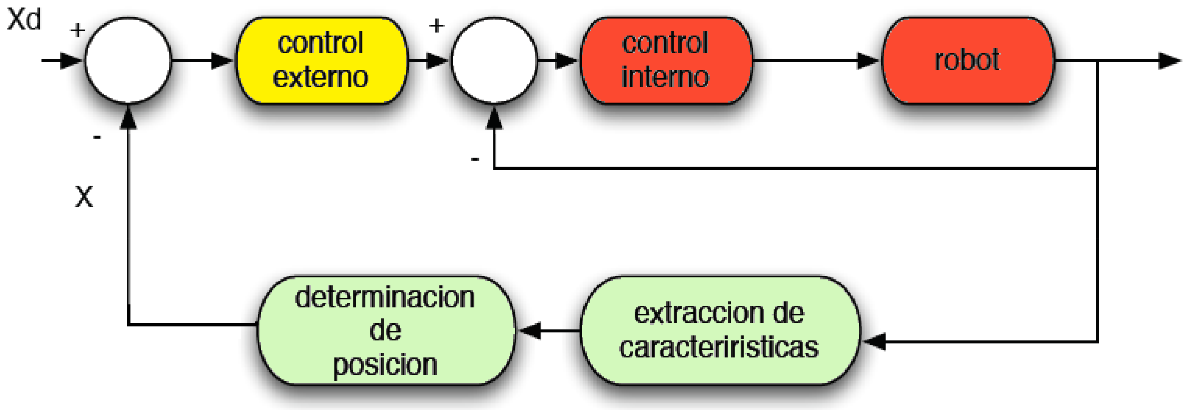


Error function based on the **3D** cartesian space, It is also called pose-based visual servoing.

Image features are extracted as well, but are additionally **used to estimated 3D** information (pose of the object in the cartesian space), hence it is servoing in 3D

Geometric models: **required**
Camera calibration: **required**
Camera robot transformation: **required**

Position Based Visual Servoing PBVS



Because **there is not direct control in the image plane**, the object can **go out** the **field of view** of the camera during the control task.

Solution: observing the object and the robot.

Solving the pose estimation problem

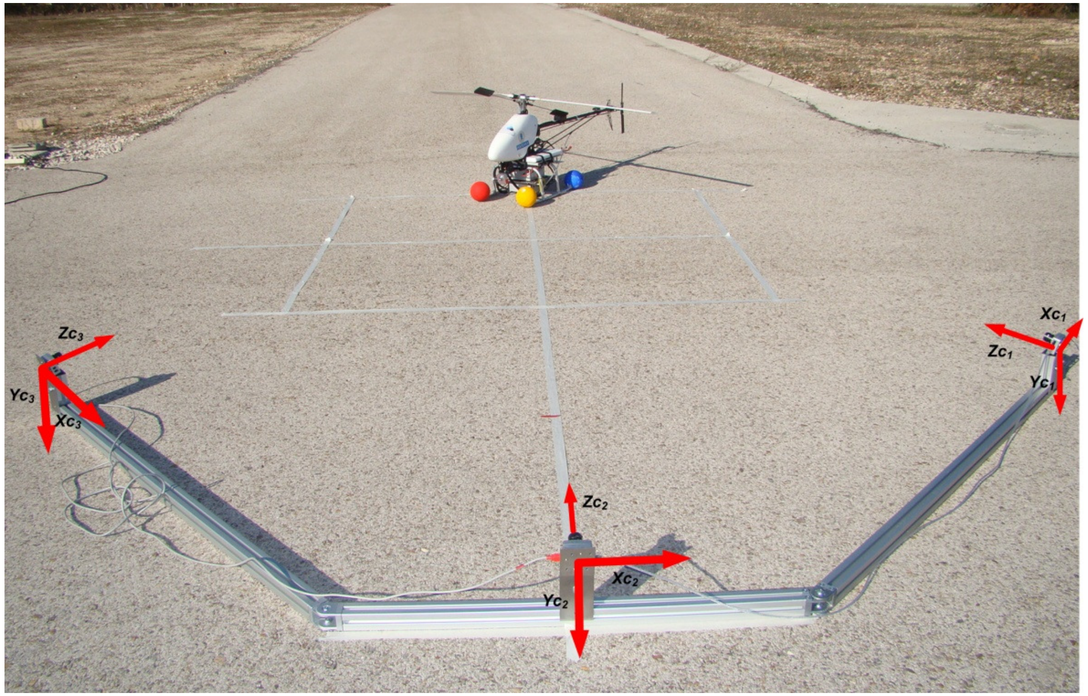
How to recover 6DOF?

Pose estimation using an on-board camera. There is not a specific object to follow

Solving the pose estimation problem

Pose estimation using an on-board camera. Following a specific object

Solving the pose estimation problem

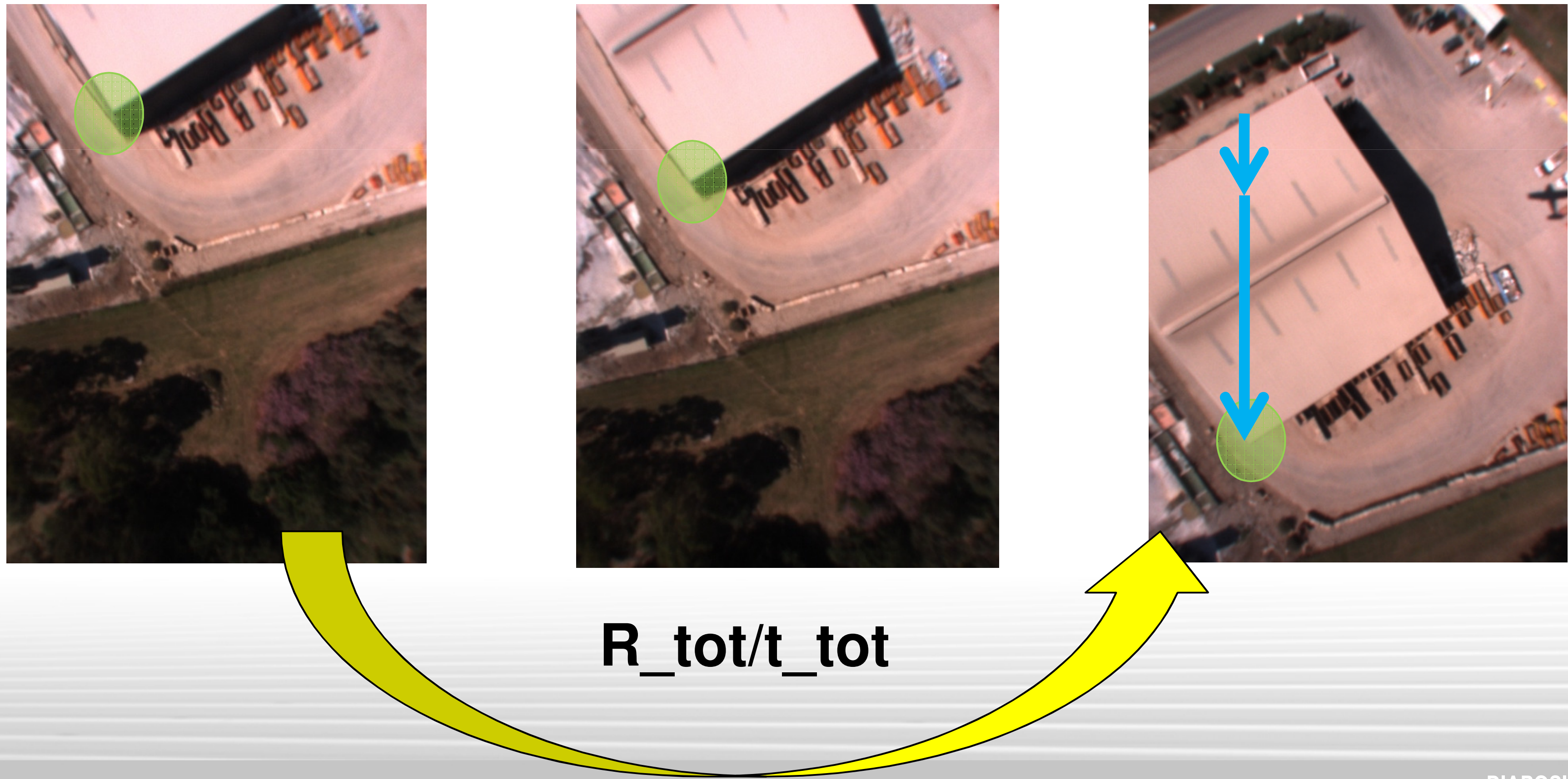


Pose estimation using an external camera system.

Pose estimation Problem

Pose estimation Problem

Assuming flat terrain, dominant movement is due to vehicle movement



Tracking of features

Frame to Frame Motion

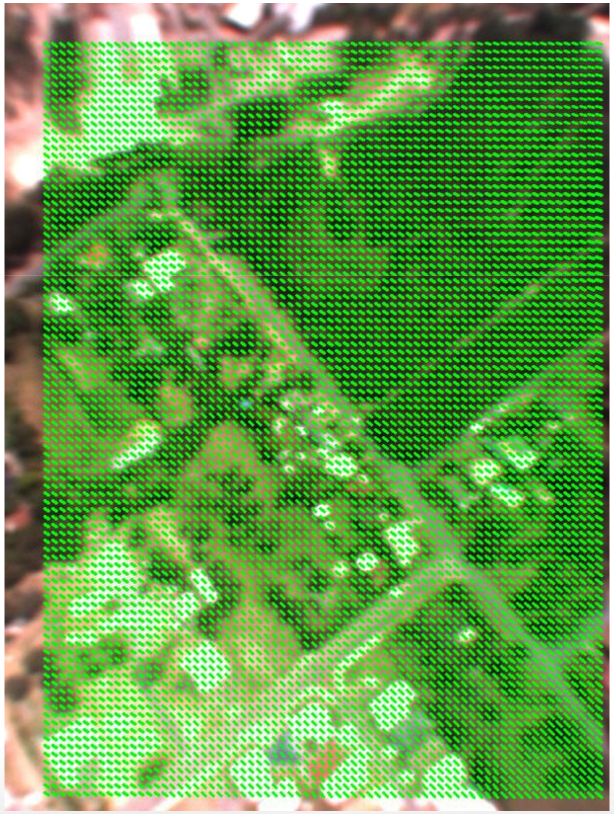
Feature-based

Direct methods

Recovering different motion models:

- Translation
- Rotation
- Scale
- Homography

Pose estimation problem



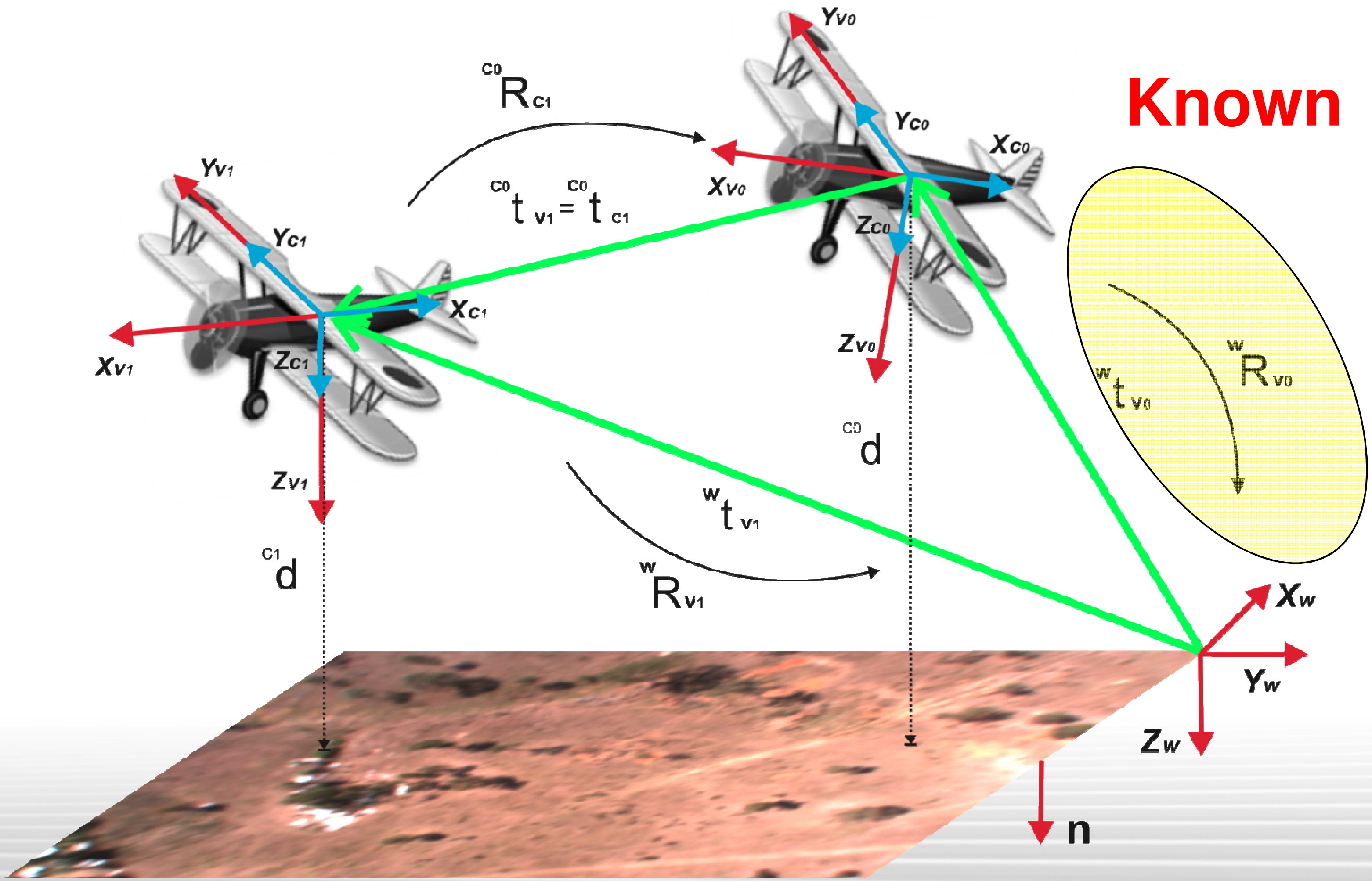
Frame to Frame Motion

Homography

$$\mathbf{x}' = \begin{bmatrix} 1+p_1 & p_2 & p_3 \\ p_4 & 1+p_5 & p_6 \\ p_7 & p_8 & 1 \end{bmatrix} \mathbf{x}$$

$$\mathbf{H_e} = {}^{\mathbf{c_2}}\mathbf{R}_{\mathbf{c_1}} + \frac{1}{d} {}^{\mathbf{c_2}}\mathbf{t}_{\mathbf{c_1}} \mathbf{n}^{\mathrm{T}}$$

Pose estimation problem



Pose estimation problem

$$\lambda\ {}^{i}\mathbf{x} = \mathbf{M}\ {}^{w}\mathbf{x}$$

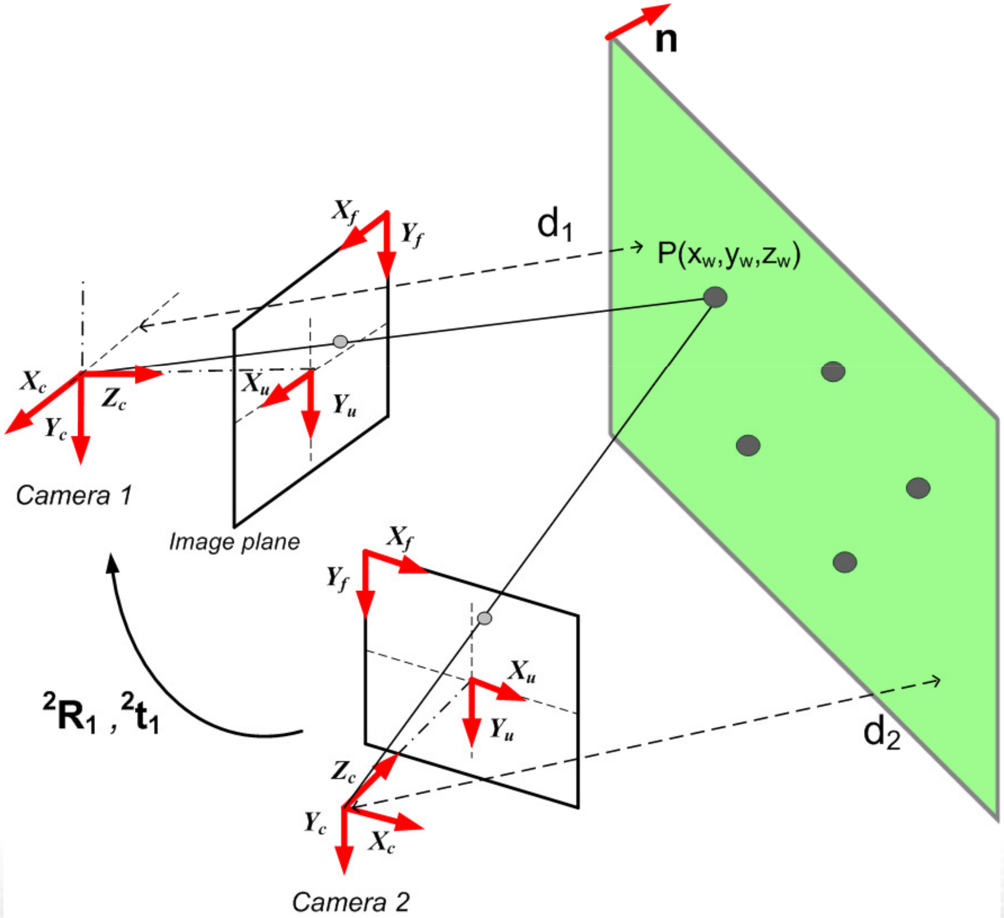
Scale factor $\Rightarrow$ $\lambda = \dfrac{Z}{f}$

Rigid transformation

$${}^{c_2}\mathbf{x} = {}^{c_2}\mathbf{R}_{c_1}\,{}^{c_1}\mathbf{x} + {}^{c_2}\mathbf{t}_{c_1}$$

$${}^{c_2}\mathbf{x} = \left({}^{c_2}\mathbf{R}_{c_1} + \frac{1}{d}\,{}^{c_2}\mathbf{t}_{c_1}\mathbf{n}^{T}\right){}^{c_1}\mathbf{x}$$

$$\mathbf{H}_{e} = {}^{c_2}\mathbf{R}_{c_1} + \frac{1}{d}\,{}^{c_2}\mathbf{t}_{c_1}\mathbf{n}^{T}$$



Frame to frame estimation

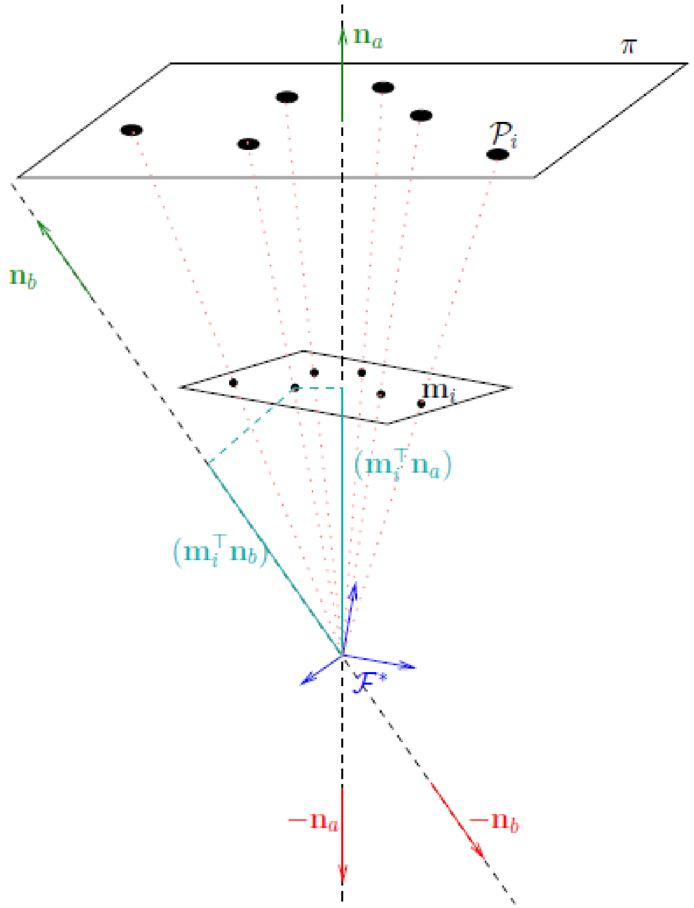
Pose estimation problem: H decomposition

1- H decomposition: Method in book→ "An invitation to 3D vision"

Solution 1	$R_1 = W_1 U_1^T$ $N_1 = \widehat{v}_2 u_1$ $\frac{1}{d} T_1 = (H - R_1) N_1$	Solution 3	$R_3 = R_1$ $N_3 = -N_1$ $\frac{1}{d} T_3 = -\frac{1}{d} T_1$
Solution 2	$R_2 = W_2 U_2^T$ $N_2 = \widehat{v}_2 u_2$ $\frac{1}{d} T_2 = (H - R_2) N_2$	Solution 4	$R_4 = R_2$ $N_4 = -N_2$ $\frac{1}{d} T_4 = -\frac{1}{d} T_2$

Pose estimation problem: H decomposition

2- From 4 solutions to 2: applying visibility constraint



All points seen by the camera must lie in front of it

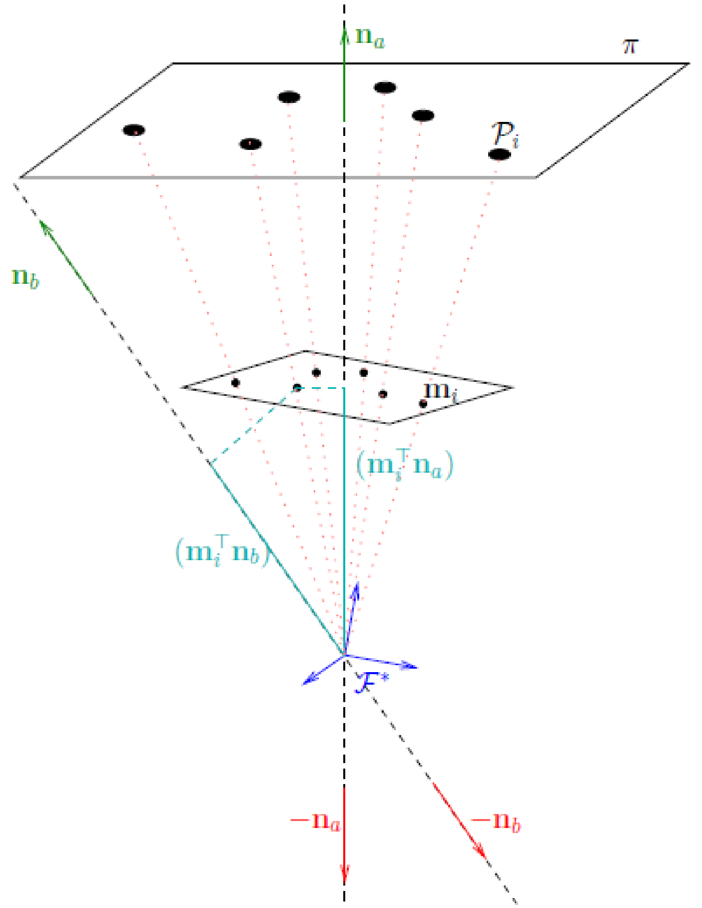
$$\mathbf{m}^* = \mathbf{K}^{-1}\mathbf{p}^*$$

$$\mathbf{m}^{*\top}\mathbf{n}^* > 0$$

TWO SOLUTIONS

Pose estimation problem: H decomposition

3- From 2 solutions to 1: assuming flat terrain



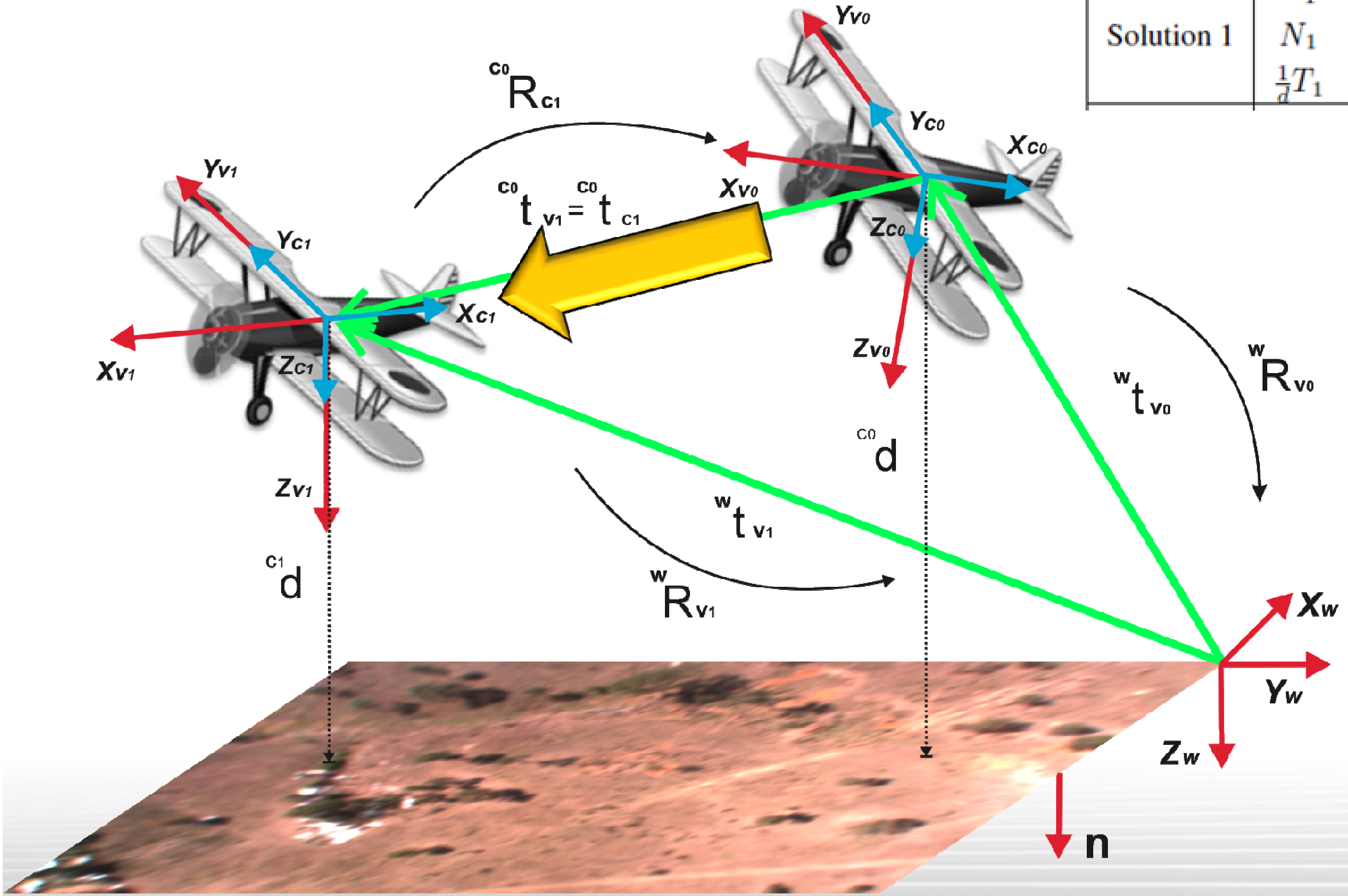
n=[0, 0, 1]

one SOLUTION

Solution 1	$R_1 = W_1U_1^T$ $N_1 = \widehat{v_2}u_1$ $\frac{1}{d}T_1 = (H - R_1)N_1$	Solution 3	$R_3 = R_1$ $N_3 = -N_1$ $\frac{1}{d}T_3 = -\frac{1}{d}T_1$
Solution 2	$R_2 = W_2U_2^T$ $N_2 = \widehat{v_2}u_2$ $\frac{1}{d}T_2 = (H - R_2)N_2$	Solution 4	$R_4 = R_2$ $N_4 = -N_2$ $\frac{1}{d}T_4 = -\frac{1}{d}T_2$

Pose Estimation

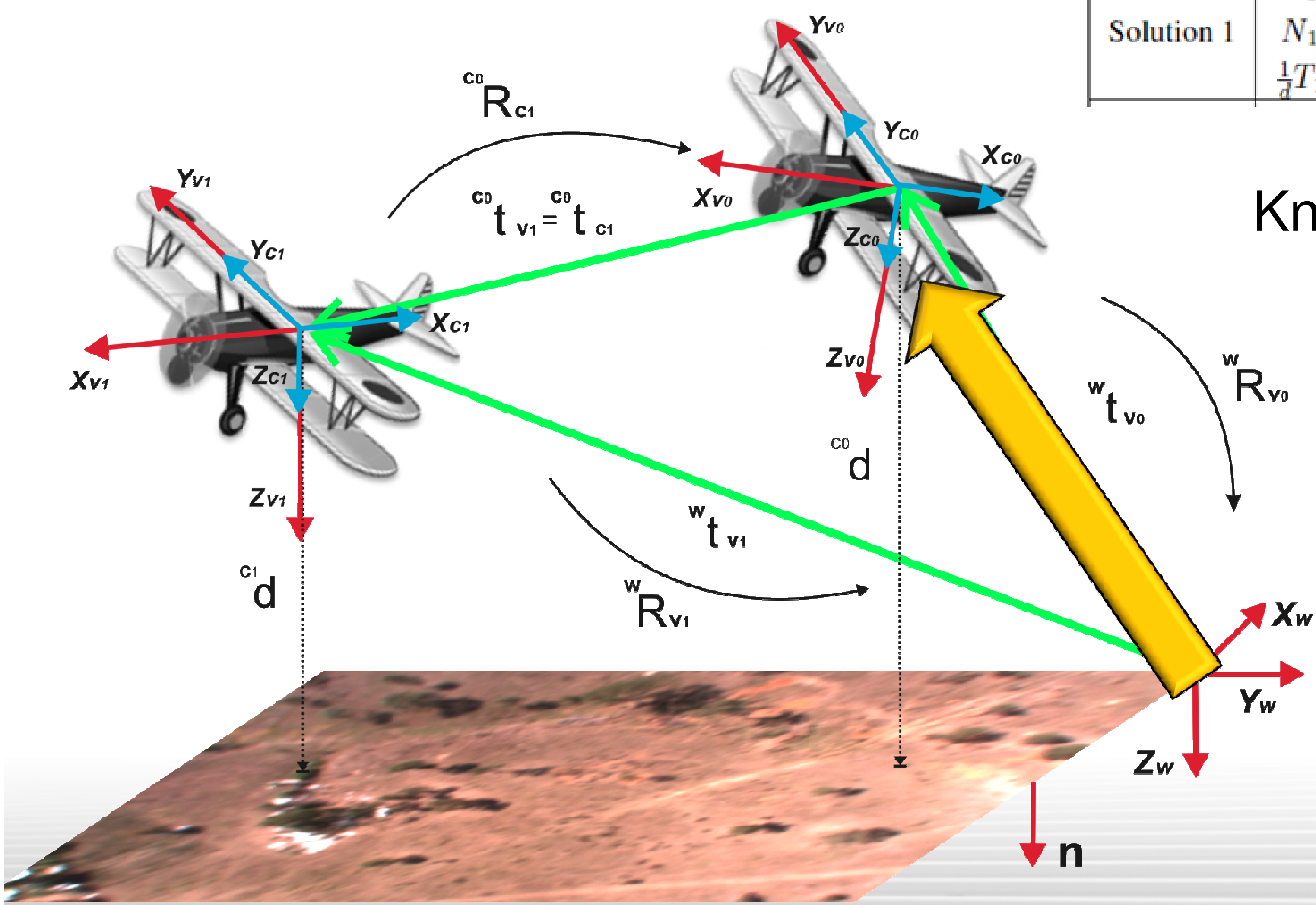
Solution 1	R_1	=	$W_1 U_1^T$
	N_1	=	$\widehat{v_2} u_1$
	$\frac{1}{d} T_1$	=	$(H - R_1) N_1$



Pose Estimation

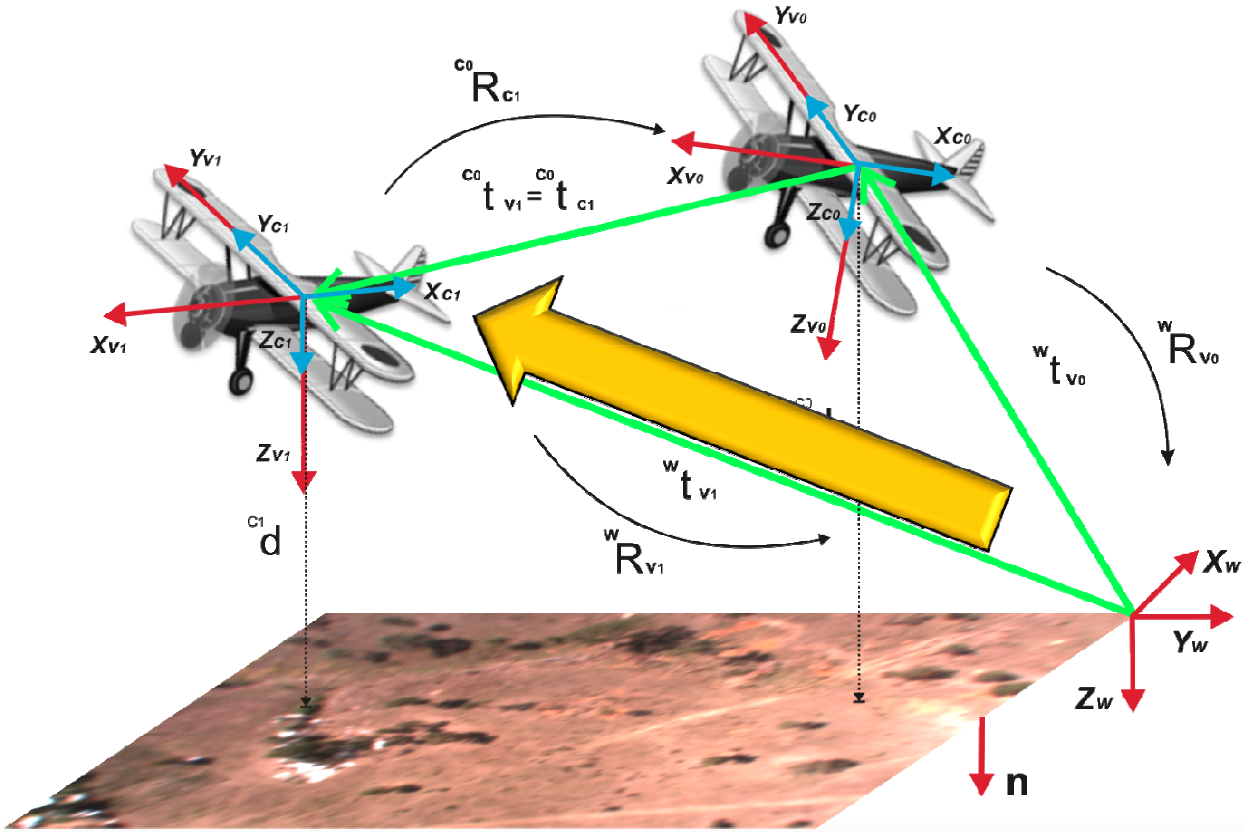
Solution 1	R_1	=	$W_1 U_1^T$
	N_1	=	$\widehat{v_2} u_1$
	$\frac{1}{d} T_1$	=	$(H - R_1) N_1$

Knowing Initial position



Pose Estimation

Solution 1	$R_1 = W_1 U_1^T$ $N_1 = \widehat{v_2} u_1$ $\frac{1}{d} T_1 = (H - R_1) N_1$



6 DOF are recovered

$$^{w}\mathbf{R}_{c_1} = {}^{w}\mathbf{R}_{c_0}\, {}^{c_0}\mathbf{R}_{c_1}$$

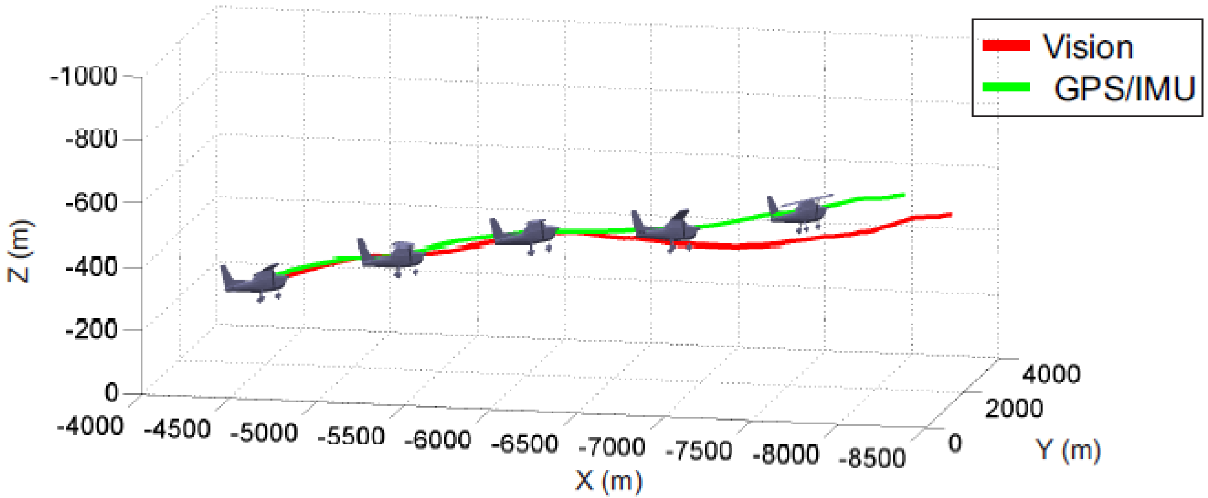
$$^{w}\mathbf{t}_{c_1} = {}^{w}\mathbf{R}_{c_0}\, {}^{c_0}\mathbf{t}_{c_1} + {}^{w}\mathbf{t}_{c_0}$$

Pose Estimation Results

- This strategy has been **used for pose estimation** of aerial vehicles using **frame to frame motion** estimation.

Pose Estimation Results

Results: take-off

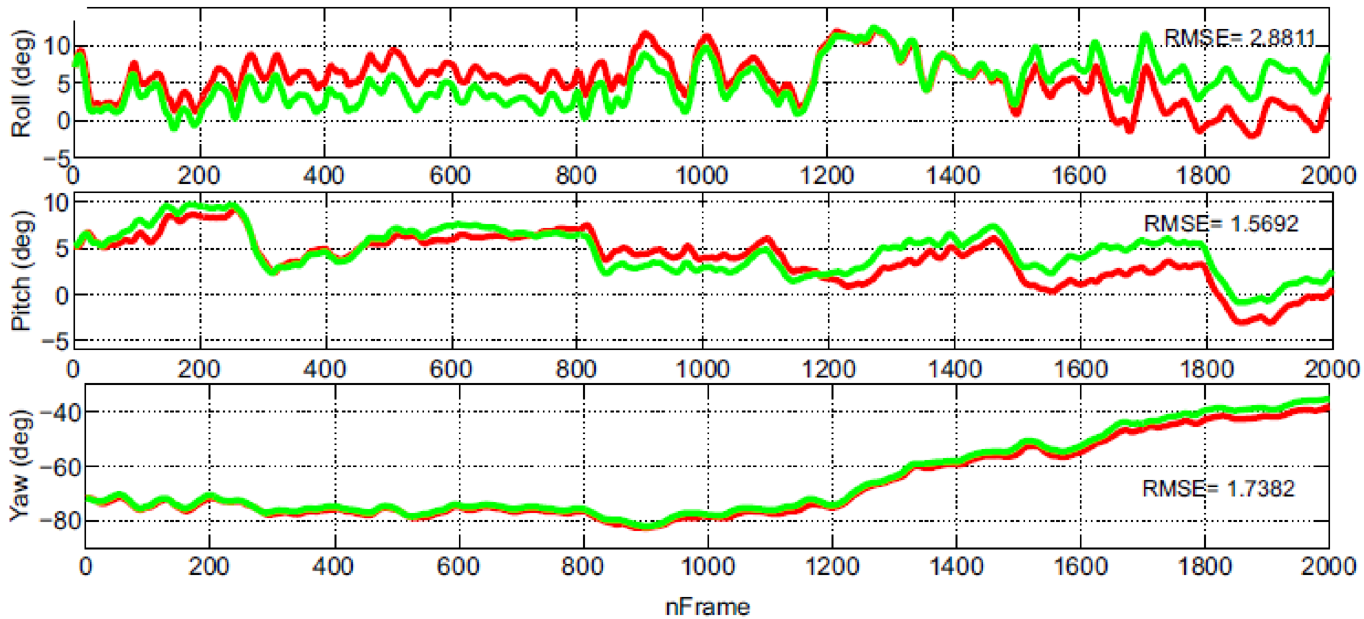


Results:
- Similar Behavior
- Low drift, only based on visual information

1 minute flight

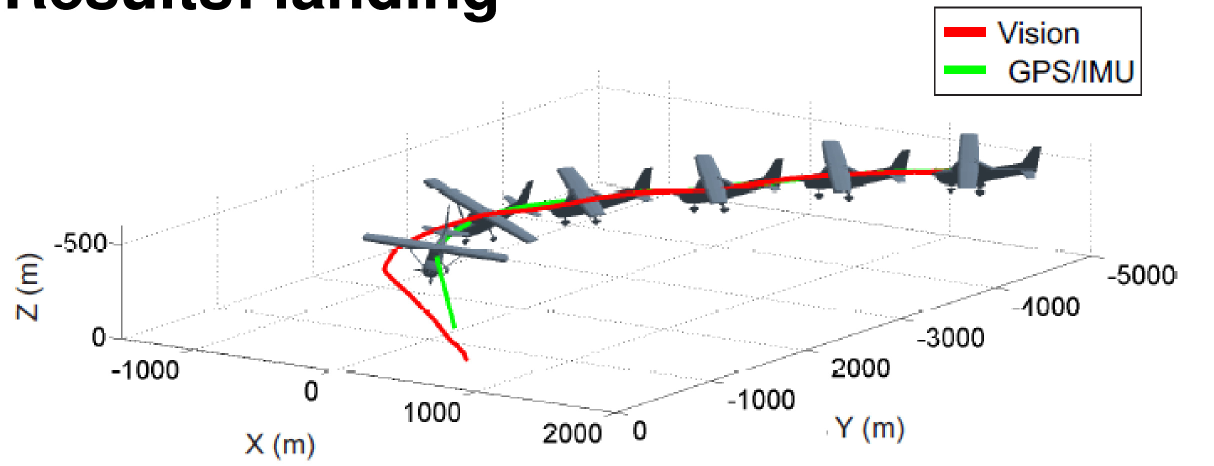
MAPE x,y,z
$[13.74\%, 3.32\%, 7.26\%]$

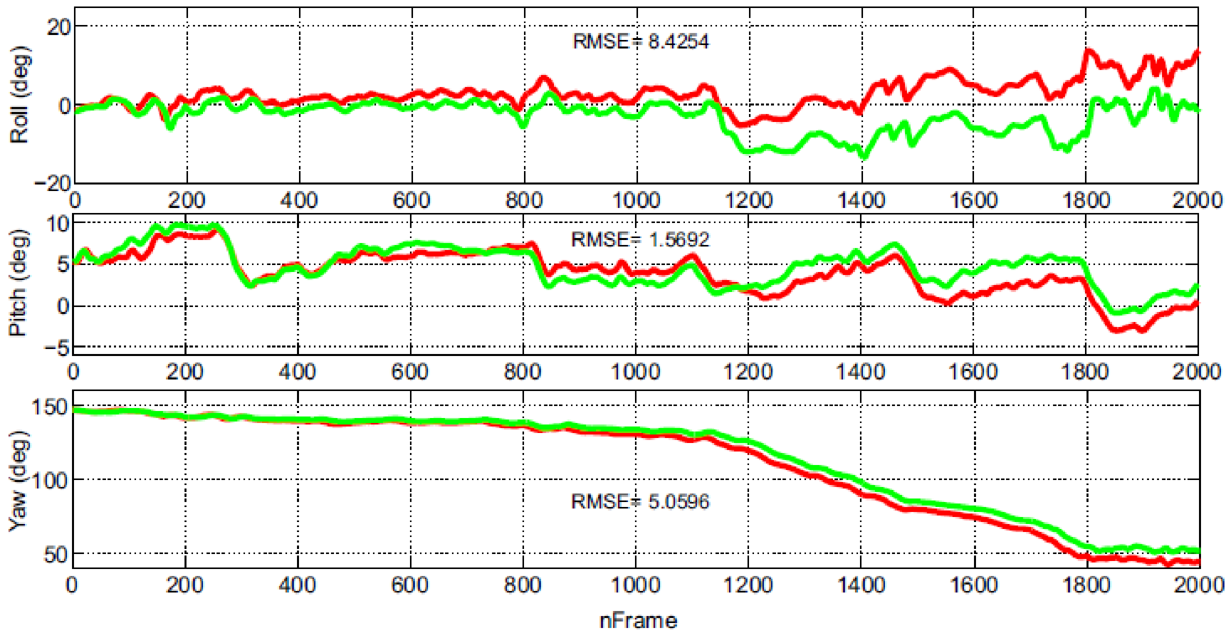
RMSE roll, pitch, yaw
$[2.8, 1.5, 1.7]$ deg



Results: cruise

Results: landing





Results:
- Similar Behavior
- Low drift, only based on visual information

MAPE x,y,z

$[8.12\%, 15.44\%, 3.70\%]$

RMSE roll, pitch, yaw

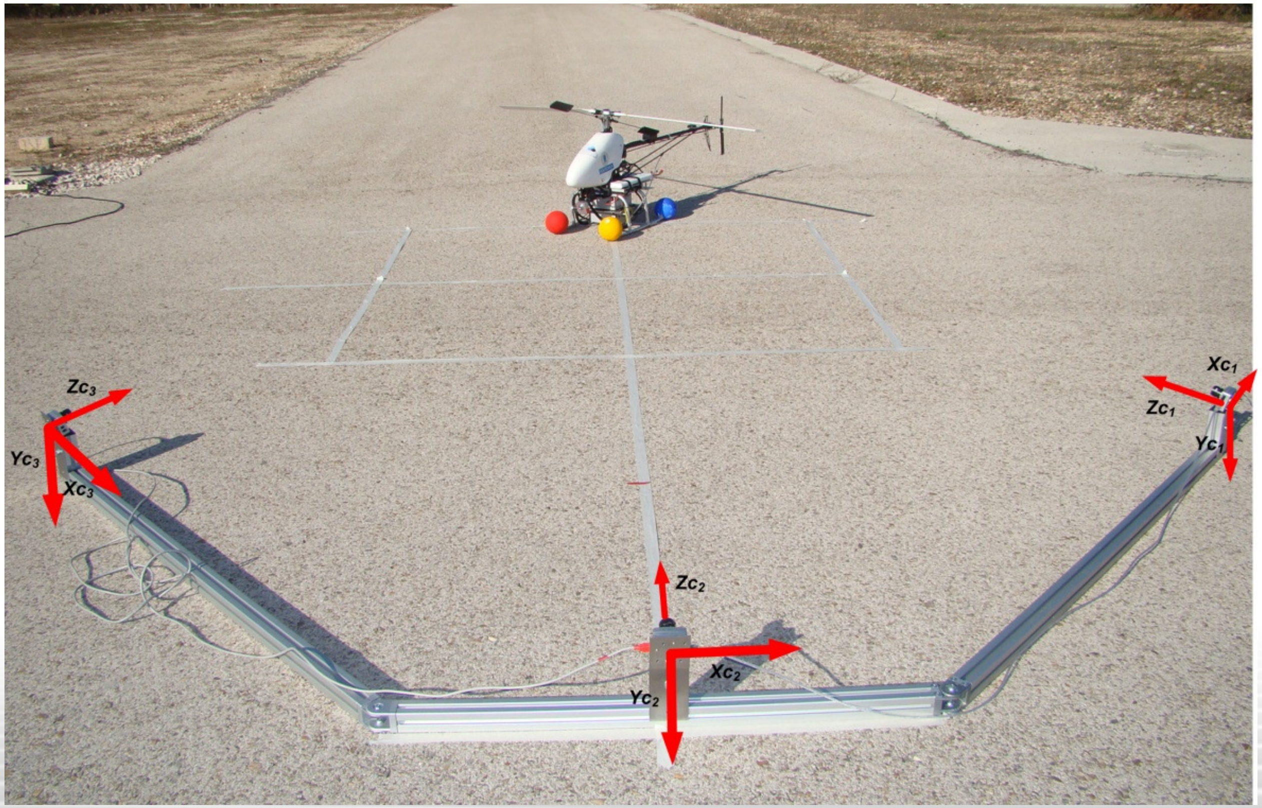
$[8.4, 1.5, 5]$ deg

Problems

- Planar assumption
- Drift due to integration of the data.
- What if there is a frame to frame error, it is integrated

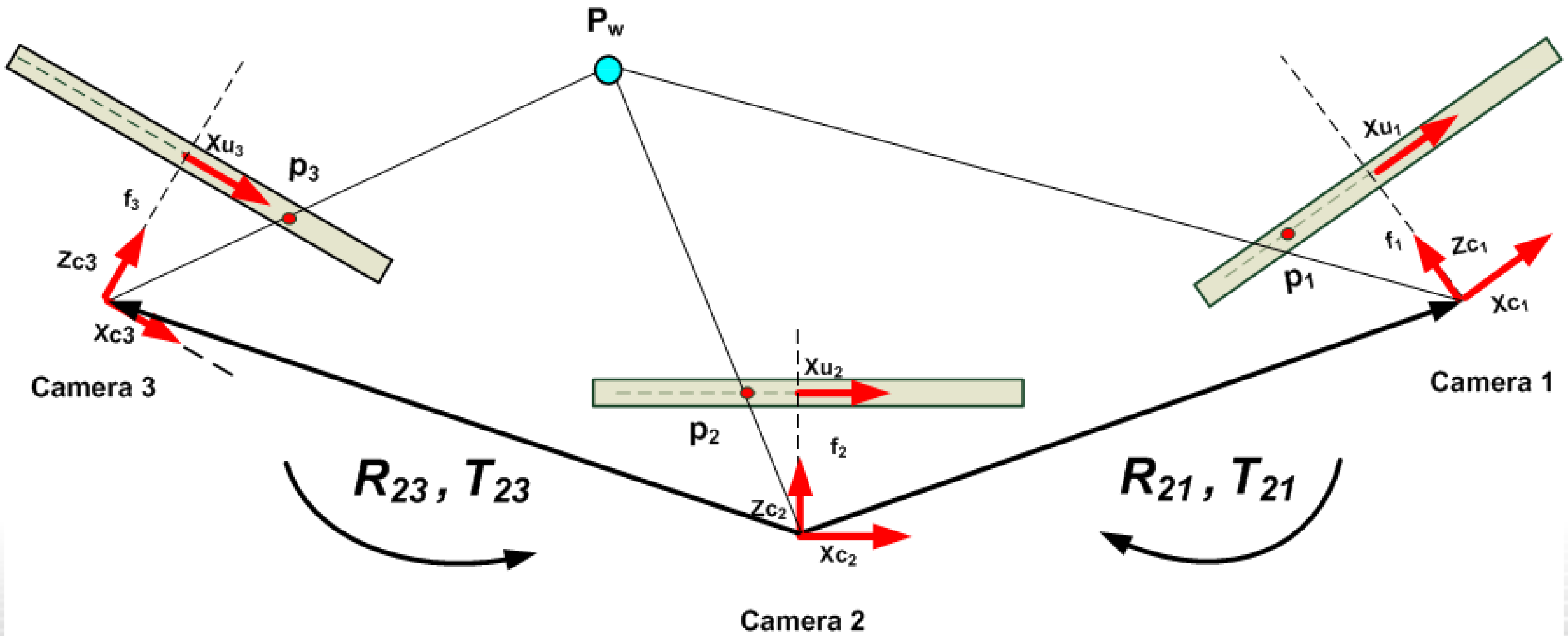
Using a external camera system

Define 3D position by detecting the coordinates of the object in each image

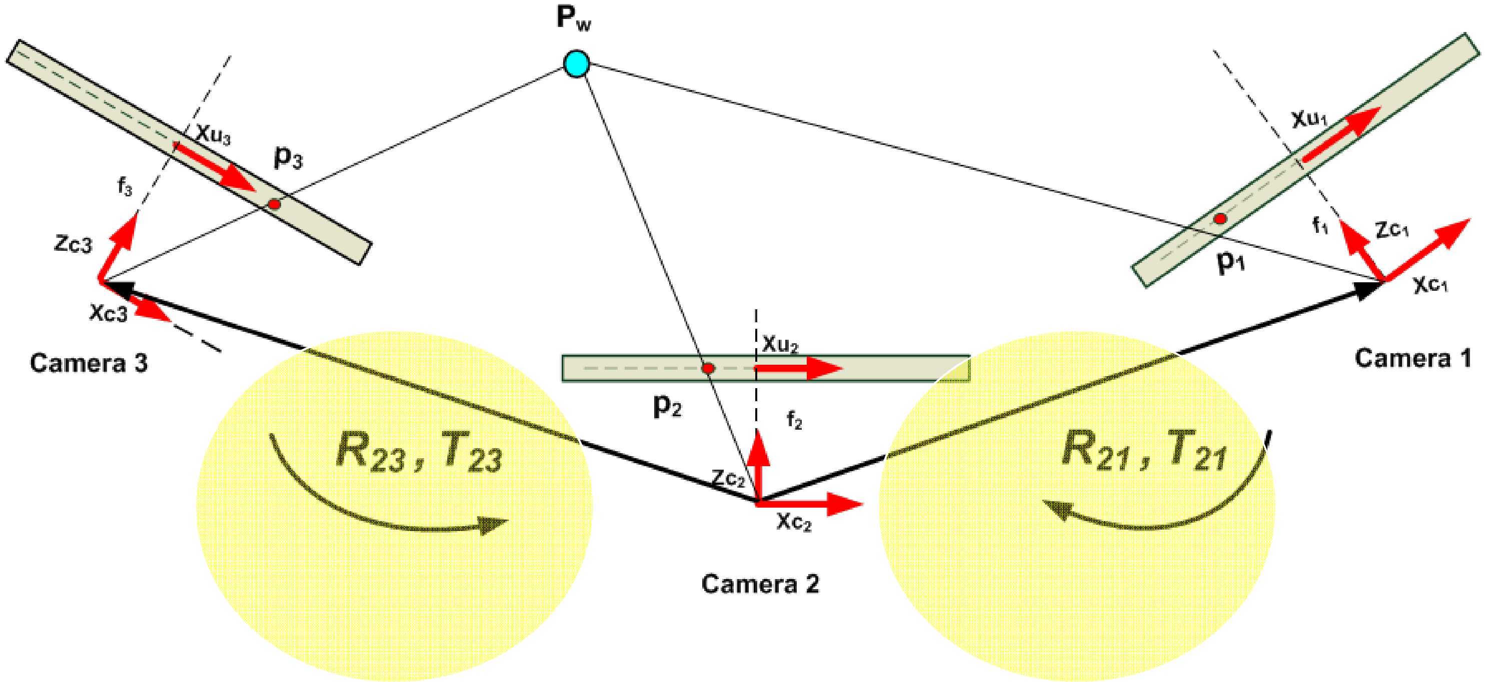


Using a external camera system

Define 3D position by detecting the coordinates of the object in each image

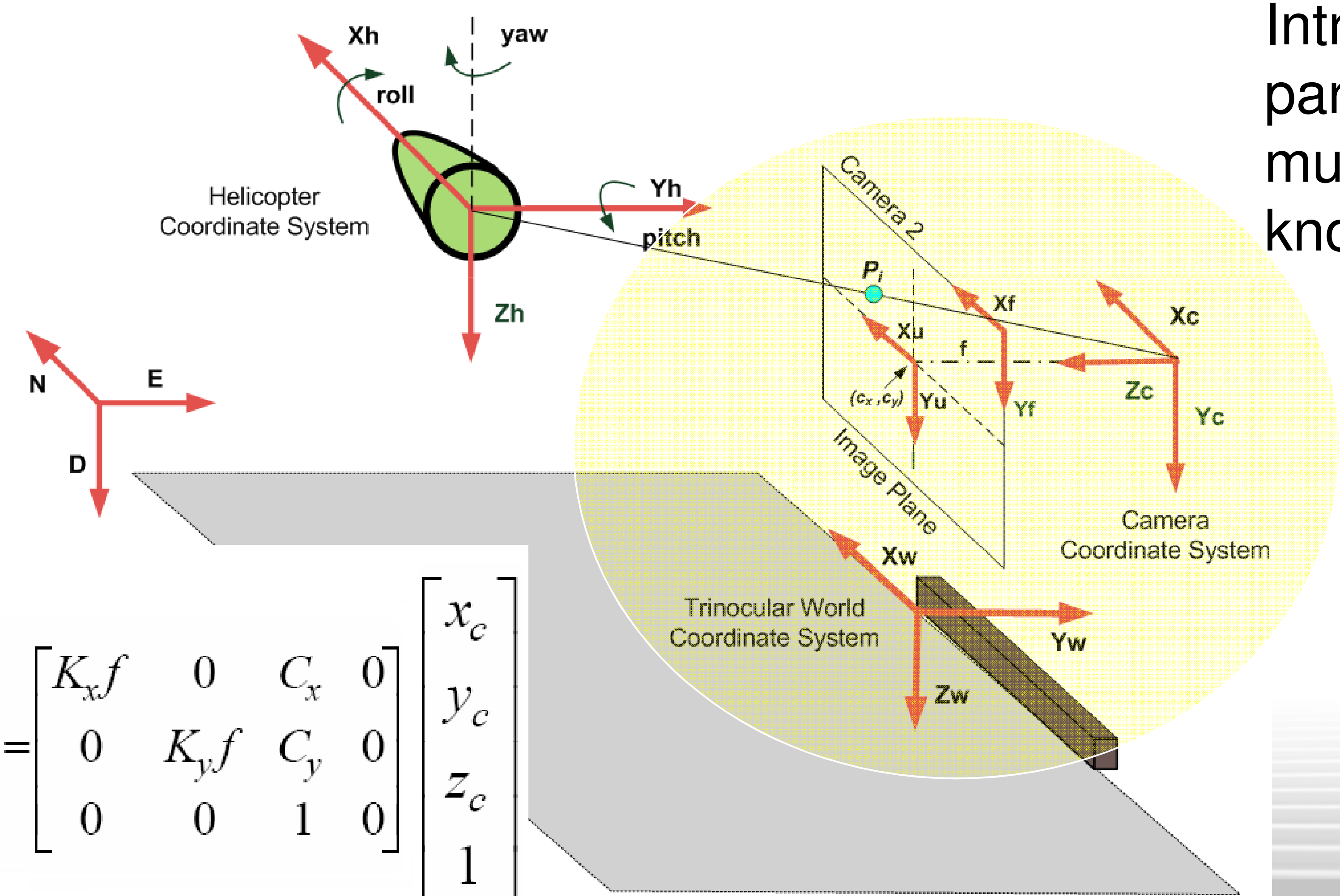


Using a external camera system



Extrinsic parameters must be known

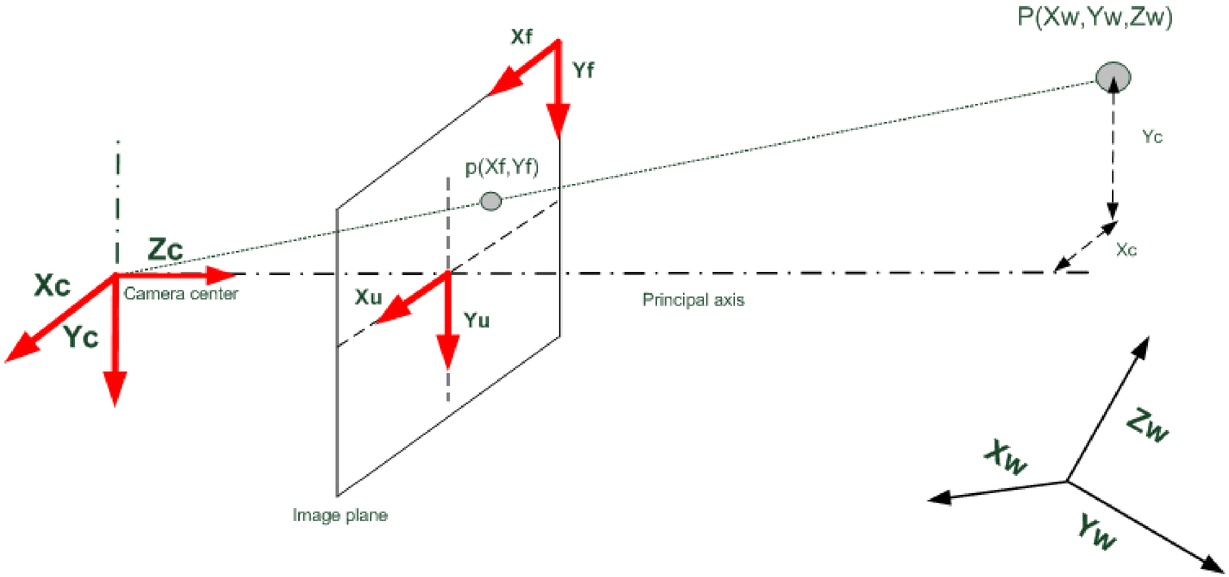
Using a external camera system

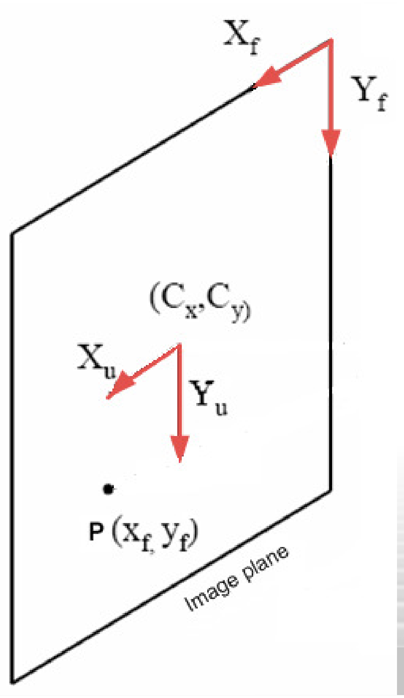


Intrinsic parameters must be known

$$\begin{bmatrix} n\,x_f \\ n\,y_f \\ n \end{bmatrix} = \begin{bmatrix} K_x f & 0 & C_x & 0 \\ 0 & K_y f & C_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}$$

Pose Estimation





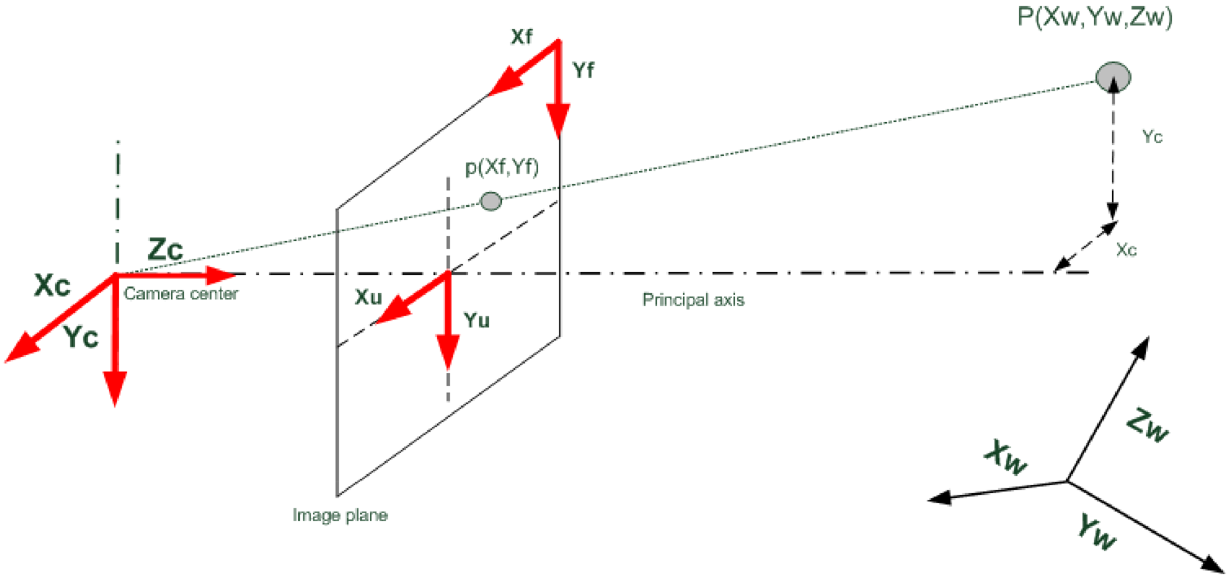
$$x_d = \frac{x_f - C_x}{K_x f} \qquad y_d = \frac{y_f - C_y}{K_y f}$$

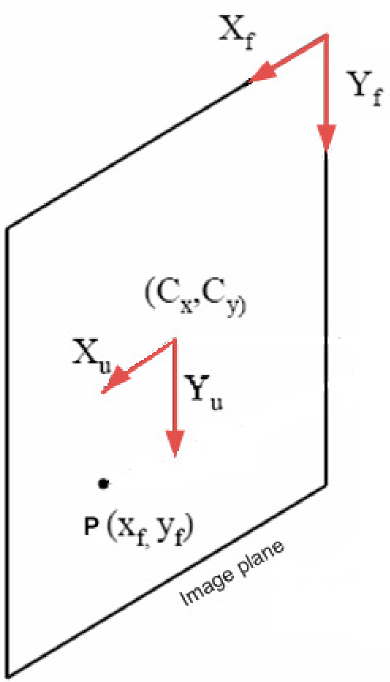
$(x_f, y_f) \Rightarrow (x_u, y_u)$ *Complicado. Iterativo con valor inicial*

$$(x_u, y_u) \equiv (x_d, y_d) \quad ; \quad r^2 \equiv x_d^2 + y_d^2 \quad \Rightarrow \quad (x_{fu}, y_{fu})$$

$$x_u \approx x_d - D_x(x_u, y_u) \qquad y_u \approx y_d - D_y(x_u, y_u)$$

Pose Estimation



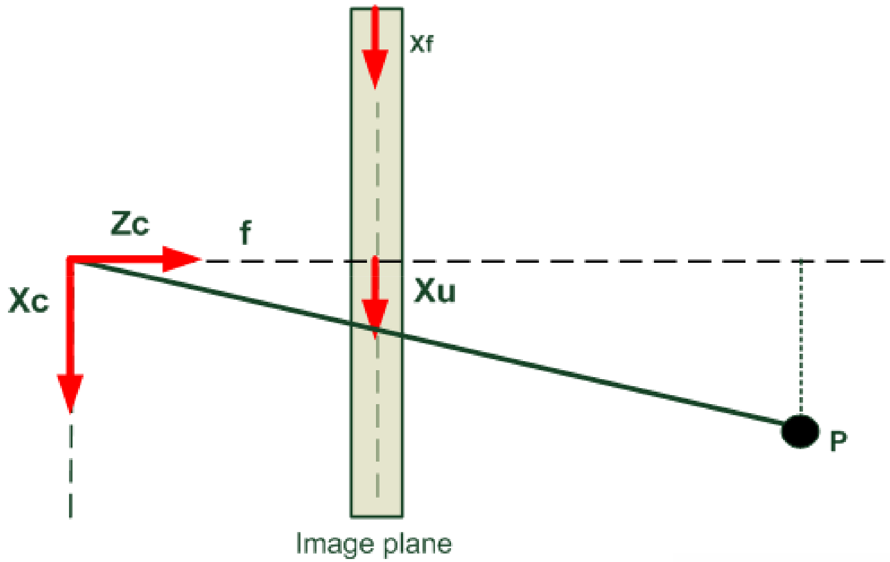


$$x_u \approx x_d - D_x(x_u, y_u) \qquad y_u \approx y_d - D_y(x_u, y_u)$$

$$d_{rad} = \left(1 + k1\, r^2 + k2\, r^4 + k5\, r^6\right) \qquad \mathbf{dx} = \begin{bmatrix} 2p_1 x_u y_u + p_2\left(r^2 + 2x_u^2\right) \\ 2p_2 x_u y_u + p_1\left(r^2 + 2y_u^2\right) \end{bmatrix}$$

$$x_u = \frac{x_u - \mathbf{dx}(1)}{d_{rad}} \qquad y_u = \frac{y_u - \mathbf{dx}(2)}{d_{rad}}$$

Pose Estimation



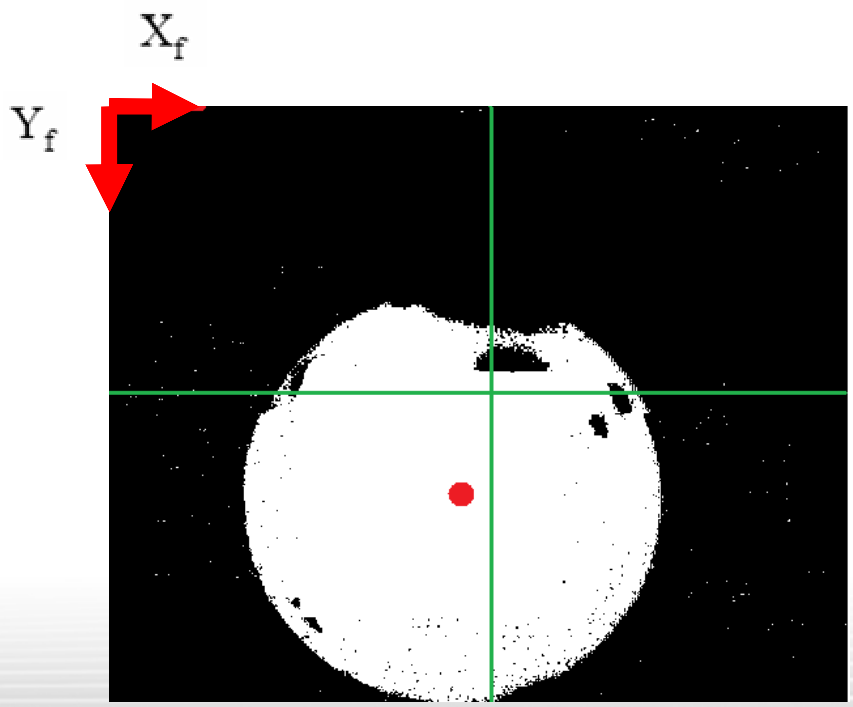
$$\frac{x_u}{f} = \frac{x_c}{z_c}$$

$$\frac{y_u}{f} = \frac{y_c}{z_c}$$

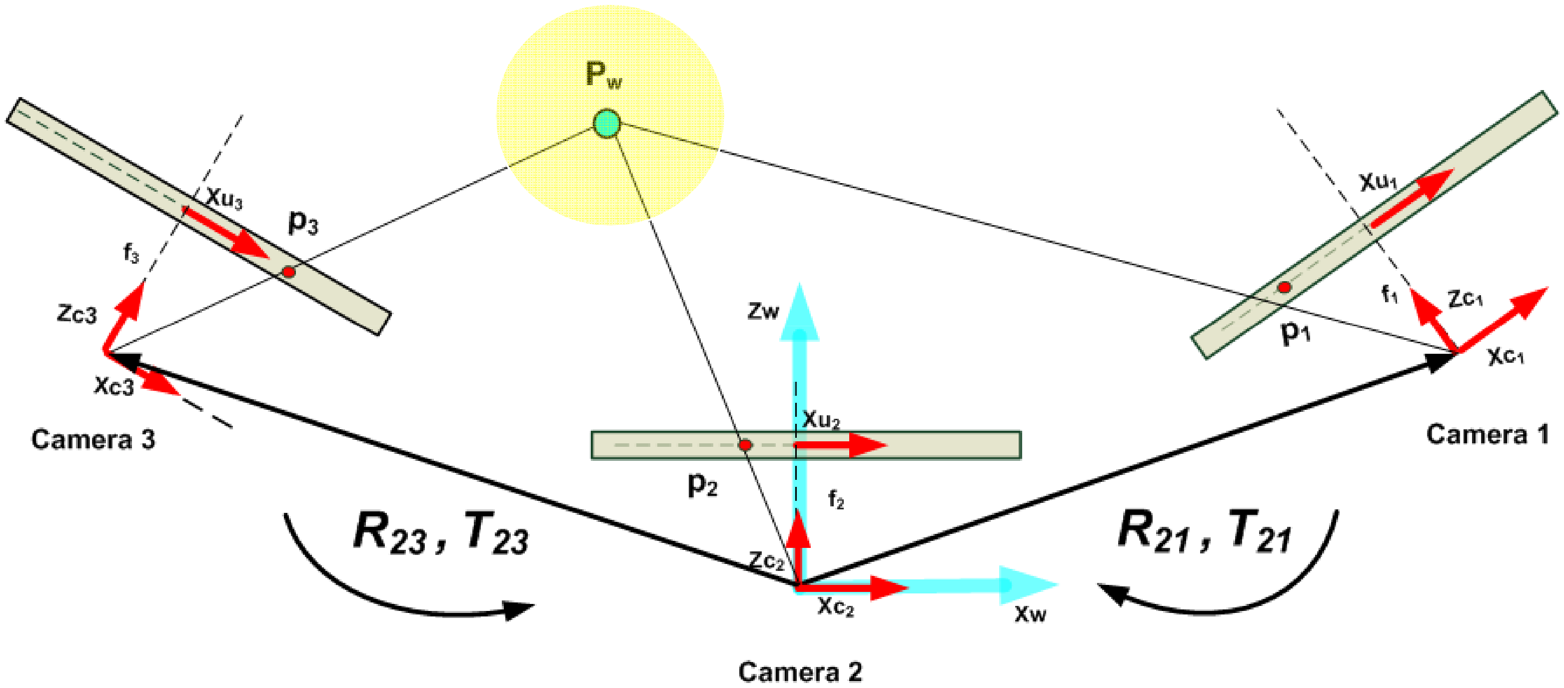
$$\begin{bmatrix} n x_u \\ n y_u \\ n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}$$

Feature extraction and tracking

- Color information
- Feature: center of gravity



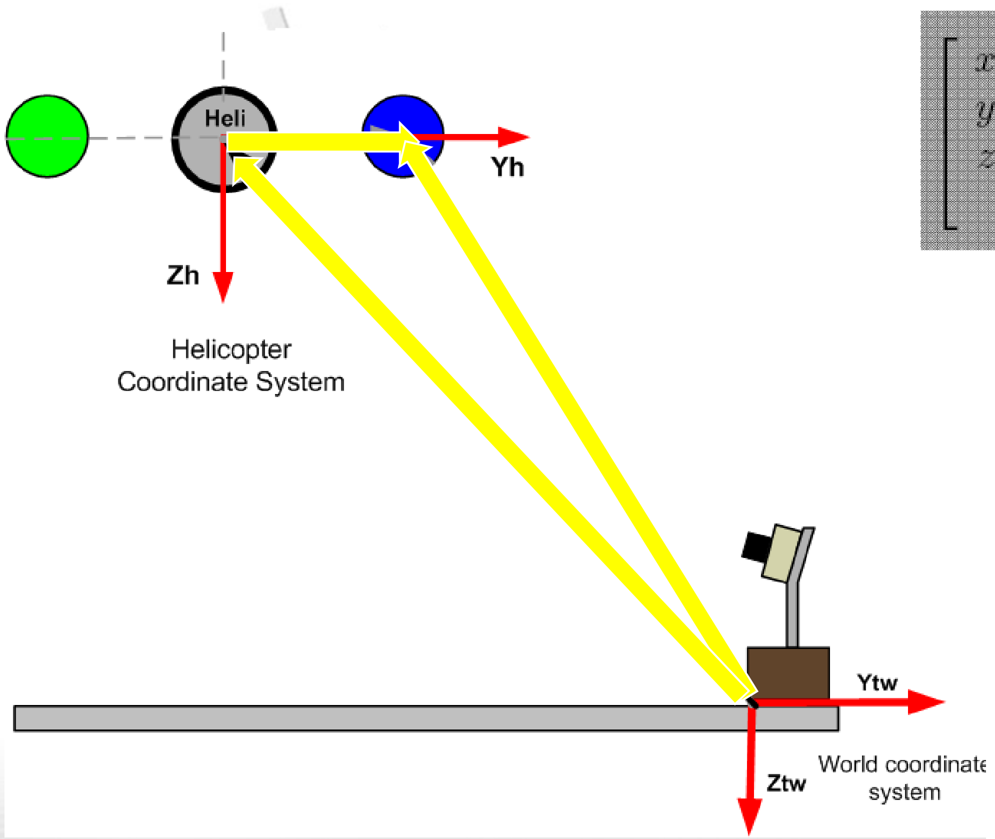
Using a external camera system



$$x_{u1_i} = f\frac{r^1_{11}X_w+r^1_{12}Y_w+r^1_{13}Z_w+t^1_x}{r^1_{31}X_w+r^1_{32}Y_w+r^1_{33}Z_w+t^1_z} \qquad y_{u1_i} = f\frac{r^1_{21}X_w+r^1_{22}Y_w+r^1_{23}Z_w+t^1_y}{r^1_{31}X_w+r^1_{32}Y_w+r^1_{33}Z_w+t^1_z}$$

$$A_i L_i = b_i,$$

Using a external camera system

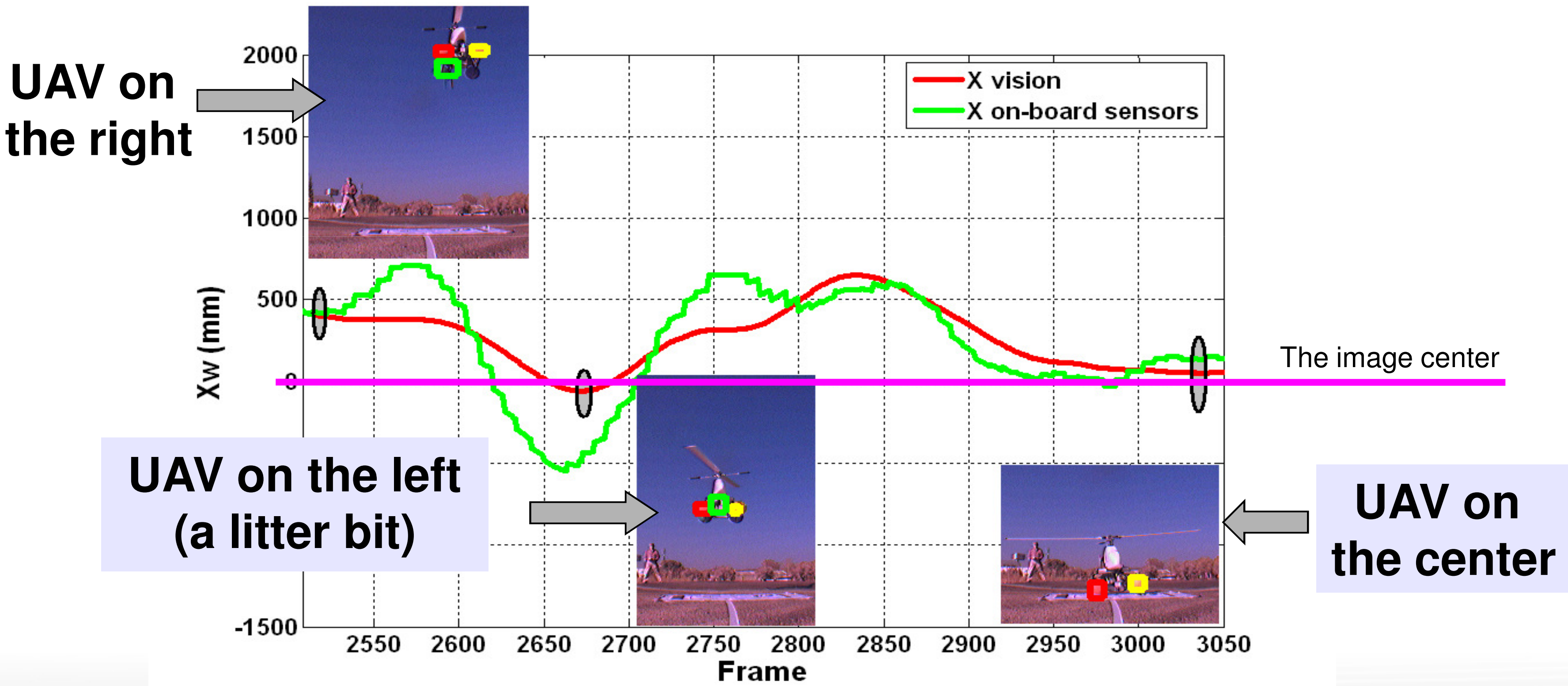


$$\begin{bmatrix} x_{tw} \\ y_{tw} \\ z_{tw} \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 & t_x \\ \sin(\theta) & \cos(\theta) & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_h \\ y_h \\ z_h \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x^1_{tw} \\ y^1_{tw} \\ z^1_{tw} - z^1_h \\ \vdots \\ z^4_{tw} - z^4_h \end{bmatrix} = \begin{bmatrix} x^1_h & -y^1_h & 1 & 0 & 0 \\ y^1_h & x^1_h & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y^4_h & x^4_h & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \\ E \end{bmatrix}$$

Pose Estimation --> TRINOCULAR SYSTEM

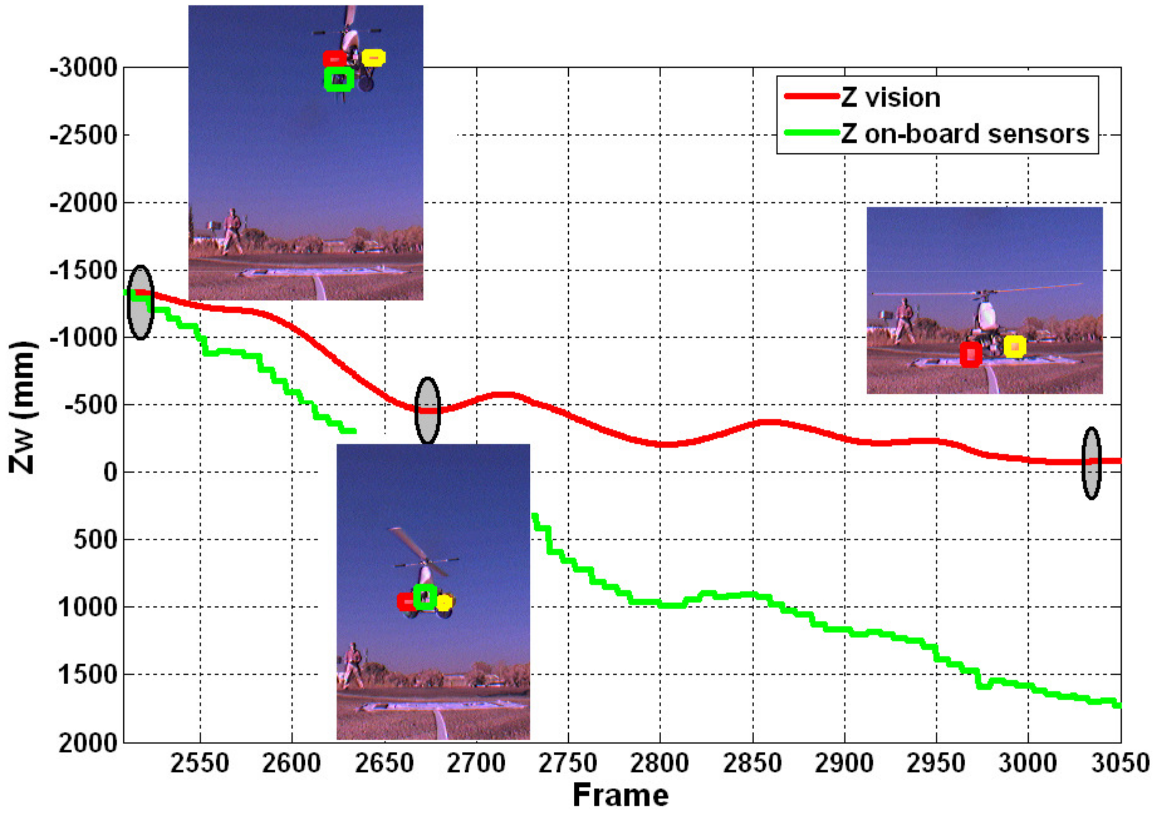
Position estimation during a **landing** task in **manual mode** (RC)



* Visual estimation **corresponds** with **real UAV position**

Pose Estimation --> TRINOCULAR SYSTEM

Position estimation during a **landing** task in **manual mode** (RC)

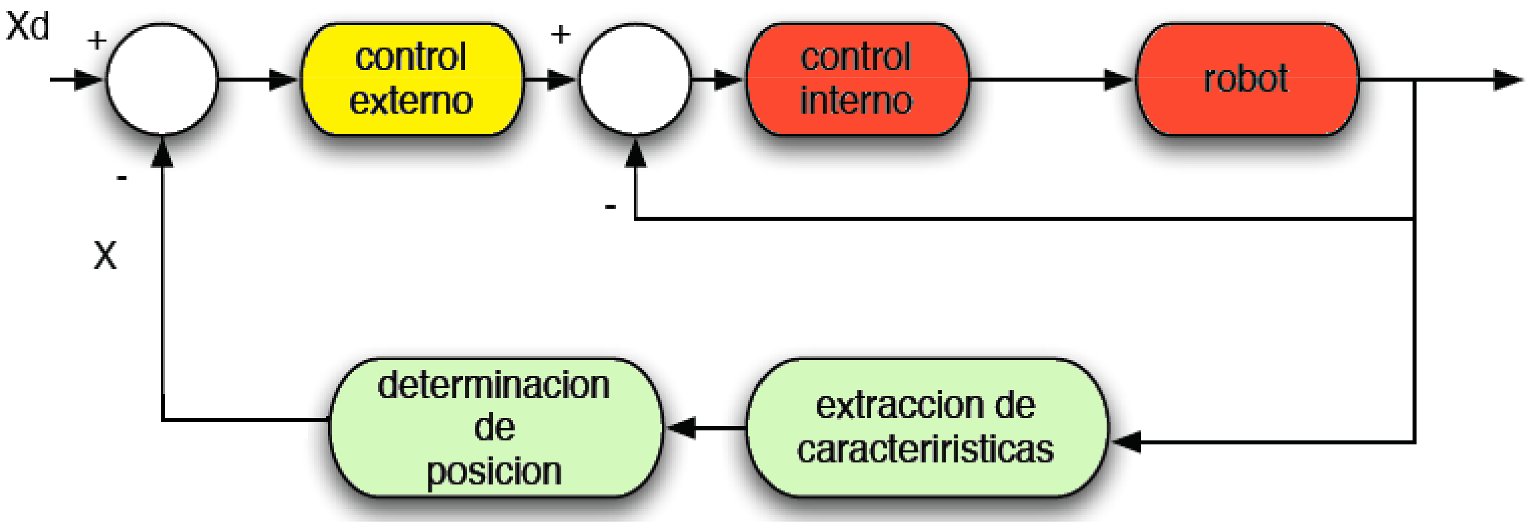


* Visual estimation **corresponds** with **real UAV position**
* **Improvement** of the UAV's position estimation

Results

UAV's YAW ANGLE ESTIMATION
USING AN EXTERNAL TRINOCULAR SYSTEM
Computer Vision Group
www.vision4uav.com
Universidad Politécnica de Madrid
POLITÉCNICA

Position Based Visual Servoing PBVS Results



Robo-Tenis

Monocular eye in hand, dynamic look and move strategy

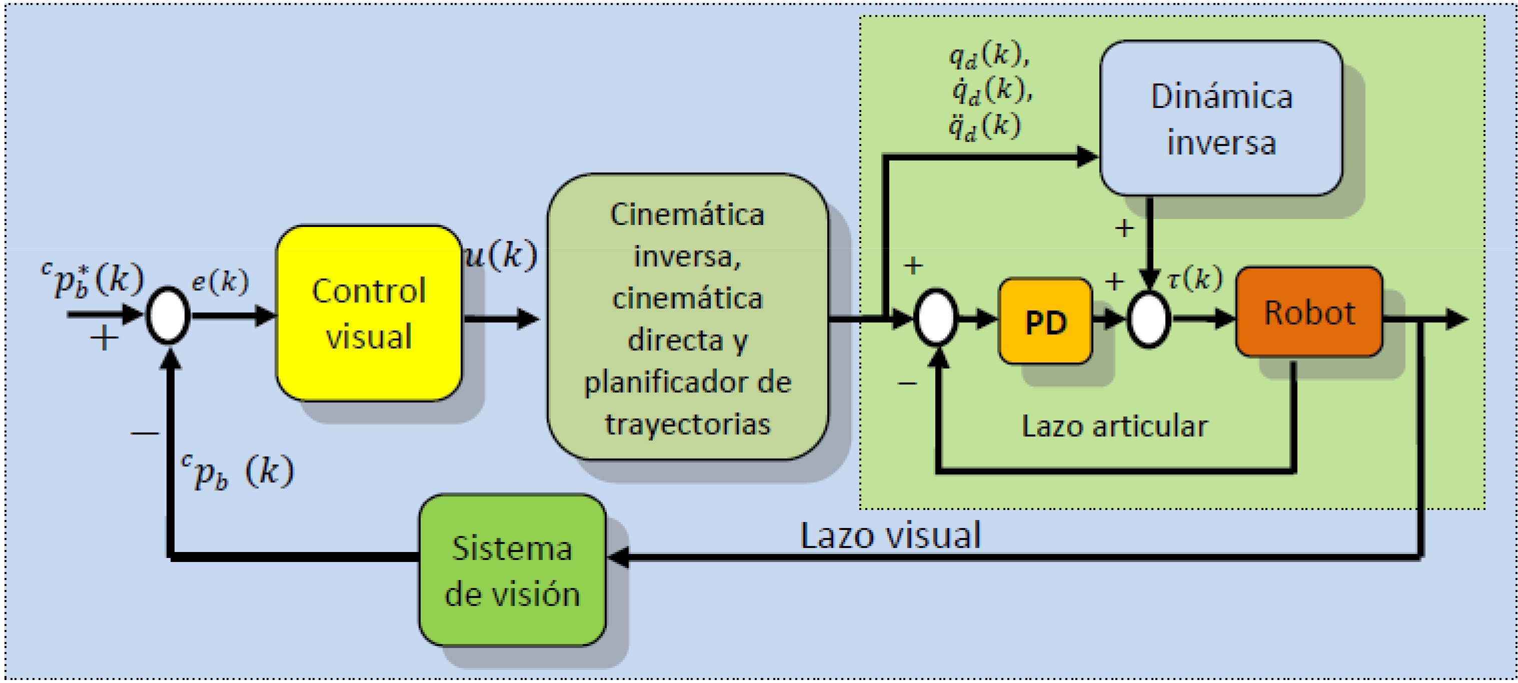


Fig. 6.2 Esquema básico de control del sistema Robotenis.

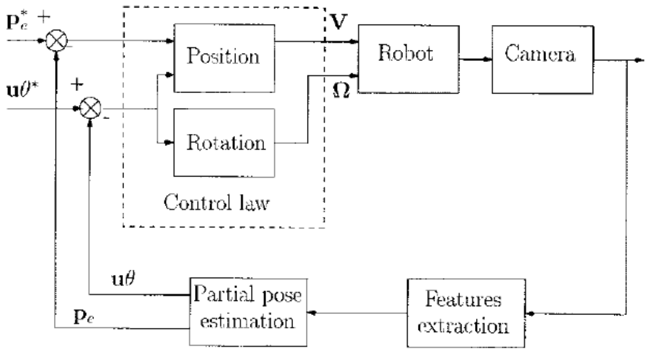
$$e(k) = {}^{c}p_b^*(k) - {}^{c}p_b(k)$$

$$e(k) = {}^{c}p_b^*(k) - {}^{c}R_w\left({}^{w}p_b(k) - {}^{w}p_c(k)\right)$$

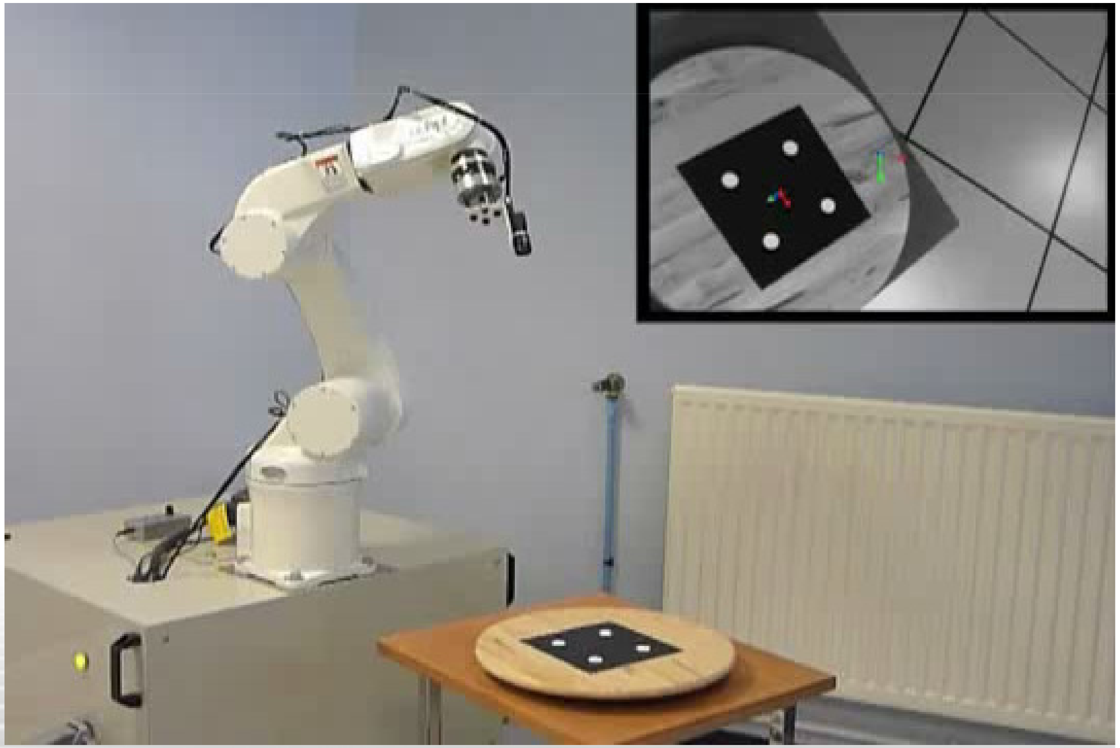
Robo-Tenis

Monocular eye in hand, dynamic look and move strategy

PBVS results

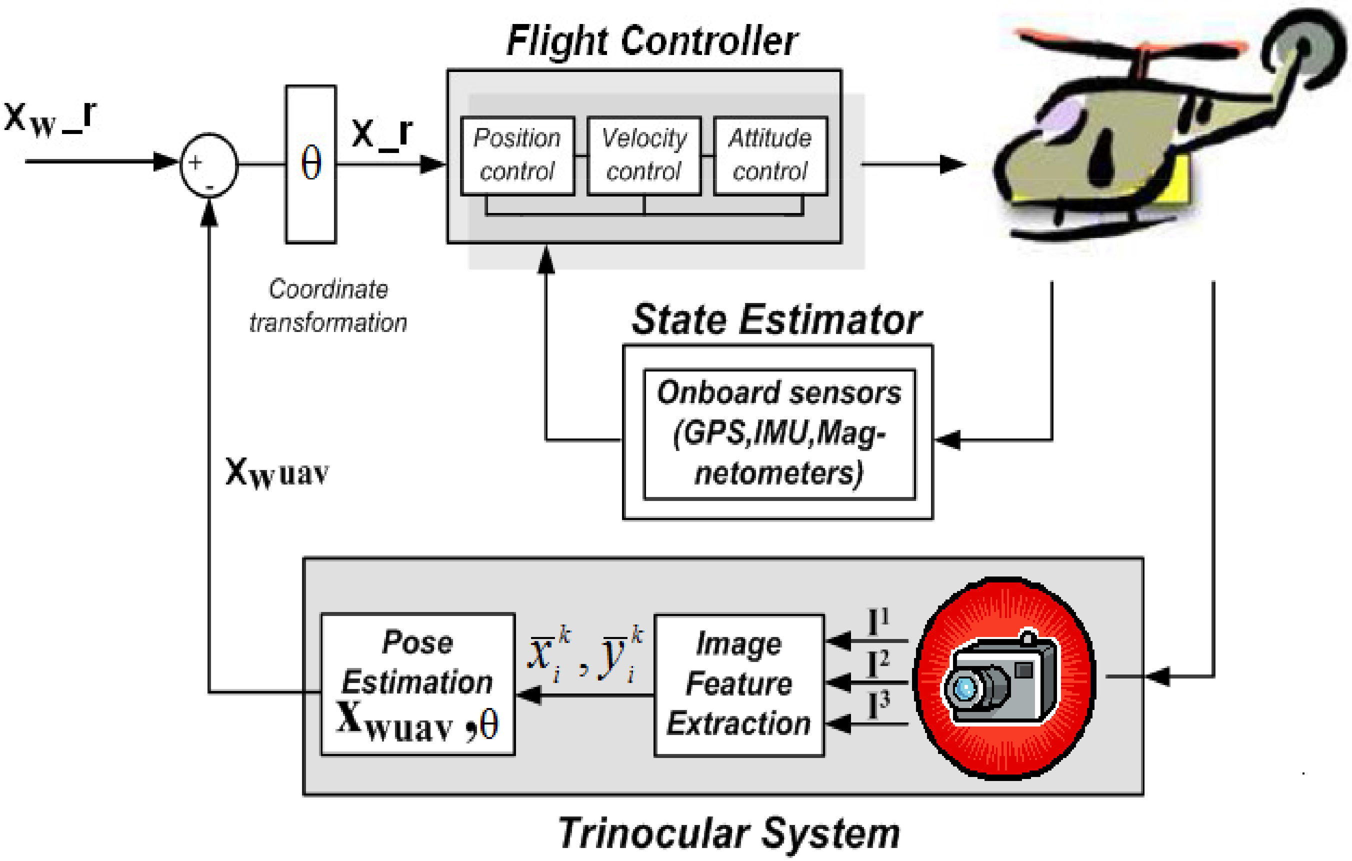


Hybrid approach

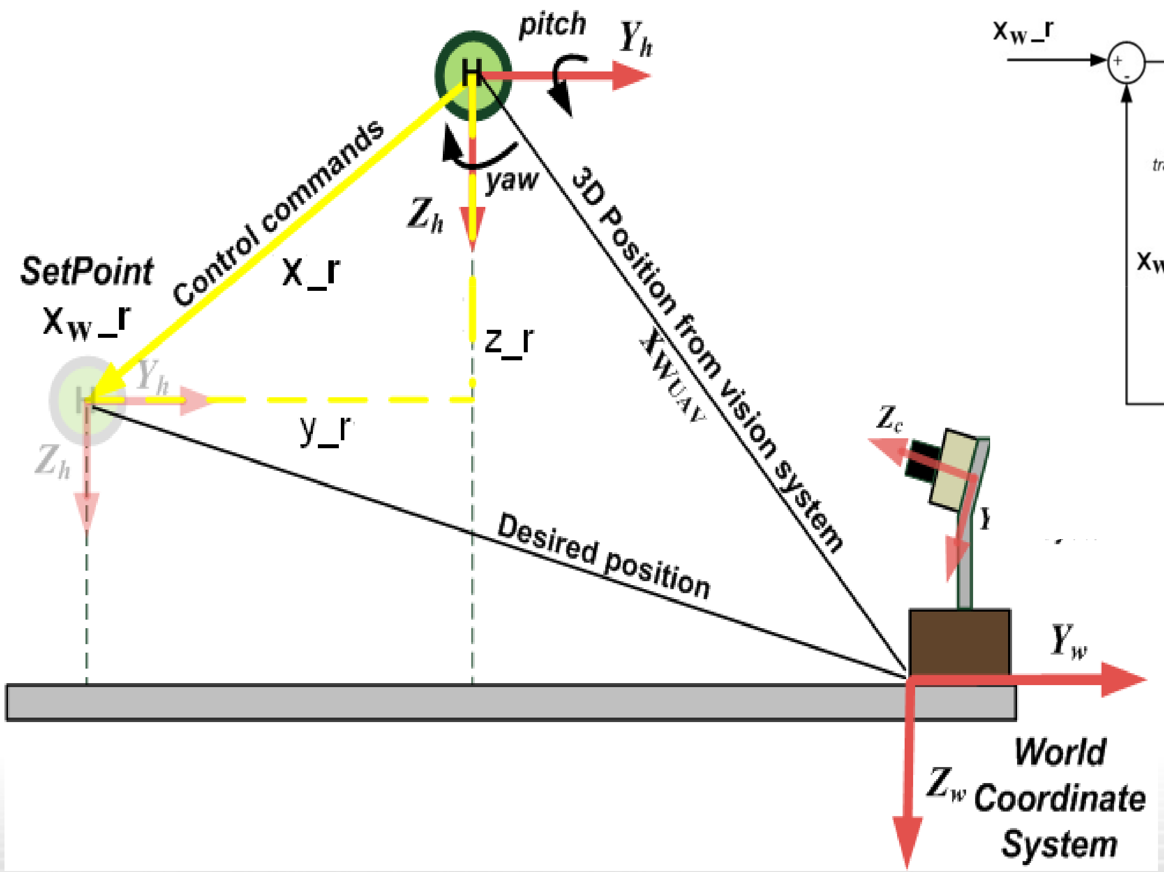


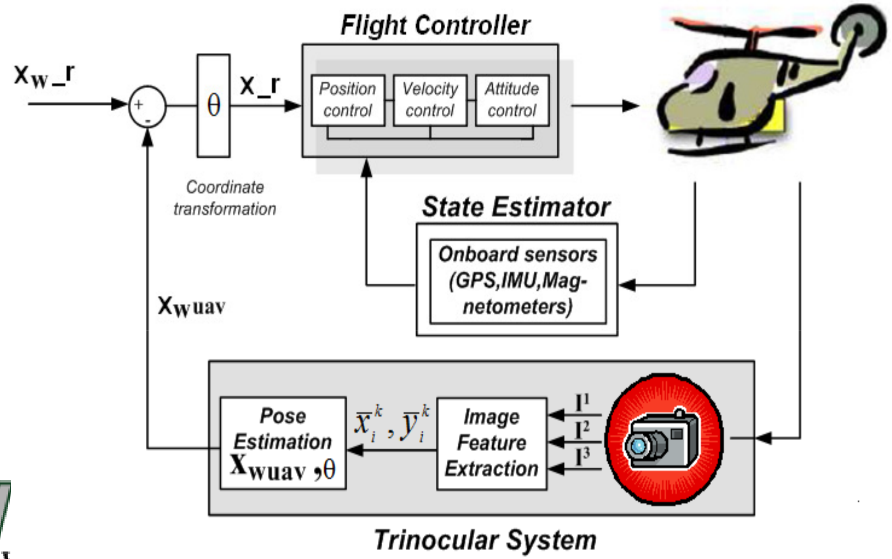
External camera system

Trinocular eye to hand, dynamic look and move strategy



External camera system





External camera system

External camera system

Vicon system: http://www.vicon.com/

Summarizing …

-Different visual control strategies depending on the error function:

- PBVS
- IBVS
- Hybrid

Summarizing …

-Different visual control strategies depending on the error function:

- PBVS
- IBVS
- Hybrid

-Position based visual servoing **depends on the pose estimation algorithm**

Summarizing …

-Different visual control strategies depending on the error function:

- PBVS
- IBVS
- Hybrid

-Position based visual servoing **depends on the pose estimation algorithm**

-Pose estimation algorithms (**depending on the number of cameras**):

- Monocular: require **additional information** to solve the depth
- Multi-camera systems: by **triangulation**, problem speed.

Summarizing …

-Different visual control strategies depending on the error function:

- PBVS
- IBVS
- Hybrid

-Position based visual servoing **depends on the pose estimation algorithm**

-Pose estimation algorithms (**depending on the number of cameras**):

- Monocular: require **additional information** to solve the depth
- Multi-camera systems: by **triangulation**, problem speed.

- Depending on the references: PBVS with **velocity commands or position commands.**

References

Slides based on:

Thesis:
Pablo Lizardo Pari
Control Visual Basado en Características de un Sistema Articulado. Estimación del Jacobiano de la Imagen Utilizando Múltiples Vistas

Thesis
Alberto Traslocheros Michel
Desarrollo, Implementación y Evaluación de Estrategias de Control Servo Visual para Robots Paralelos: Aplicación a la Plataforma Robotenis

Thesis
Luis Mejías Alvarez
Control Visual de un Vehículo Aéreo Autónomo Usando Detección y Seguimiento de Características en Espacioes Exteriores

References

Papers

Agin, G.J., "Real Time Control of a Robot with a Mobile Camera". Technical Note 179, SRI International, Feb. 1979.

S. A. Hutchinson, G. D. Hager, and P. I. Corke. A tutorial on visual servo control. IEEE Trans. Robot. Automat., 12(5):651--670, Oct. 1996.

F. Chaumette, S. Hutchinson. Visual Servo Control, Part I: Basic Approaches. IEEE Robotics and Automation Magazine, 13(4):82-90, December 2006

F. Chaumette, S. Hutchinson. Visual Servo Control, Part II: Advanced Approaches. IEEE Robotics and Automation Magazine, 14(1):109-118, March 2007

F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In D. Kriegman, G. Hager, and S. Morse, editors, The confluence of vision and control, volume 237 of Lecture Notes in Control and Information Sciences, pages 66–78. Springer-Verlag, 1998.

References

Papers

P. Corke and S. A. Hutchinson. A new partitioned approach to image-based visual servo control. IEEE Trans. Robot. Autom., 17(4):507–515, Aug. 2001.

E. Malis, F. Chaumette and S. Boudet, 2.5 D visual servoing, IEEE Transactions on Robotics and Automation, 15(2):238-250, 1999

E. Marchand, F. Spindler, F. Chaumette. ViSP for visual servoing: a generic software platform with a wide class of robot control skills. IEEE Robotics and Automation Magazine, Special Issue on "Software Packages for Vision-Based Control of Motion", P. Oh, D. Burschka (Eds.), 12(4):40-52, December 2005.

¡Thanks!


The Flying Machine Arena
Quadrocopter Ball Juggling
IDSC
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich