
ROBOTS QUE USAN SUS OJOS

visión en robots móviles

José María Cañas Plaza

<http://gsyc.escet.urjc.es/jmplaza>



SICFIMA, 13 marzo 2006

Contenidos

- Introducción
- Percepción en robots
- Experimentos
 - Hipótesis suelo
 - Taxias con filtros de color
 - Localización desde visión
 - Atención visual
 - Combinación de información visual
 - Geometría proyectiva
- Conclusiones

Introducción

Robótica ficción vs Robótica real



¿Qué es un robot?



Sistema informático con:

- Sensores
- Actuadores
- Computador

Hay que **programarlo** para que consiga sus objetivos y sea sensible a la situación

¿Qué es la visión computacional?

- Cámara, imágenes
- Procesamiento: bordes, transformadas, color, segmentación en objetos...
- Reconocimiento de formas, caras...
- Pares estéreo, línea epipolar...
- Reconstrucción 3D



Visión dentro de robótica

- Es un sensor más: proporciona información del entorno
- Potencialmente muy rico
- MUY barato (cmos, webcam)
- Extraer información útil es complejo
- Flujo desbordante de datos
- Los robots más modernos la usan



Percepción en robots

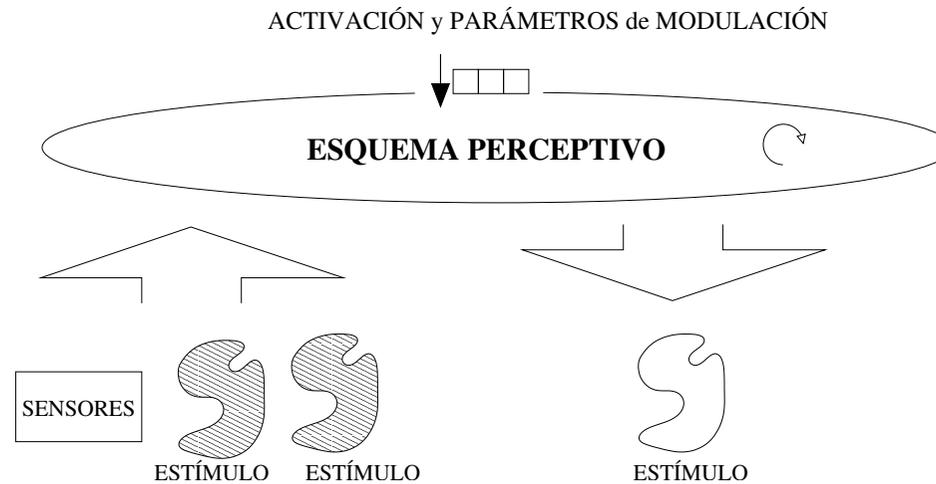
La *arquitectura* de un robot es la *organización* de sus capacidades sensoriales, de procesamiento y de acción para conseguir un repertorio de comportamientos inteligentes interactuando con cierto entorno

- La arquitectura determina el comportamiento observable
- Repertorio de comportamientos
- Información desbordante, incierta
- Selección de acción, atención

Arquitectura JDE

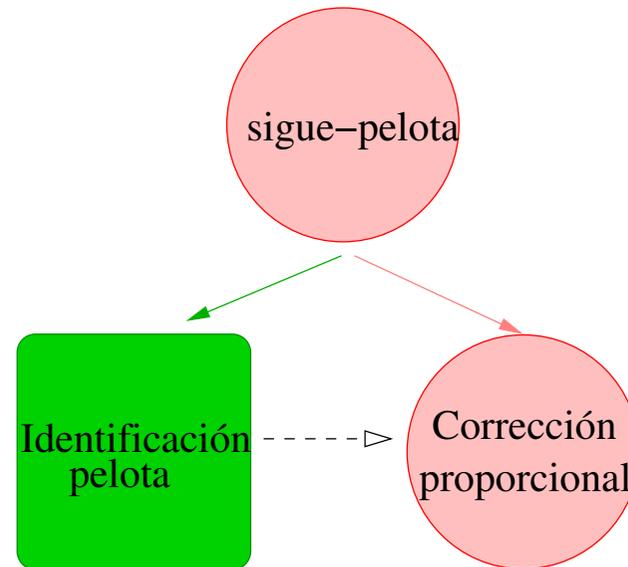
- Comportamiento = **percepción** y control
- Fragmentación en unidades (**esquemas**) asíncronas concurrentes
 - de percepción elaboran estímulos
 - de actuación toman decisiones
- La colección de esquemas se organiza en **jerarquía**
- Jerarquía Dinámica de Esquemas

Esquemas perceptivos



- Producen estímulos y los mantienen actualizados (**anclados**)
- un **estímulo** es *una pieza de información que al menos un esquema de actuación necesita para tomar sus decisiones*
- Simples lecturas sensoriales, transformaciones más elaboradas o depende de otros estímulos

Percepción en JDE



- Colección dinámica de estímulos (organizados en jerarquía)
- Carácter subsidiario de la percepción
- Coordinación con actuación: percepción situada, ¿cuándo?
- Sólo se buscan los estímulos que interesan (**atención**)

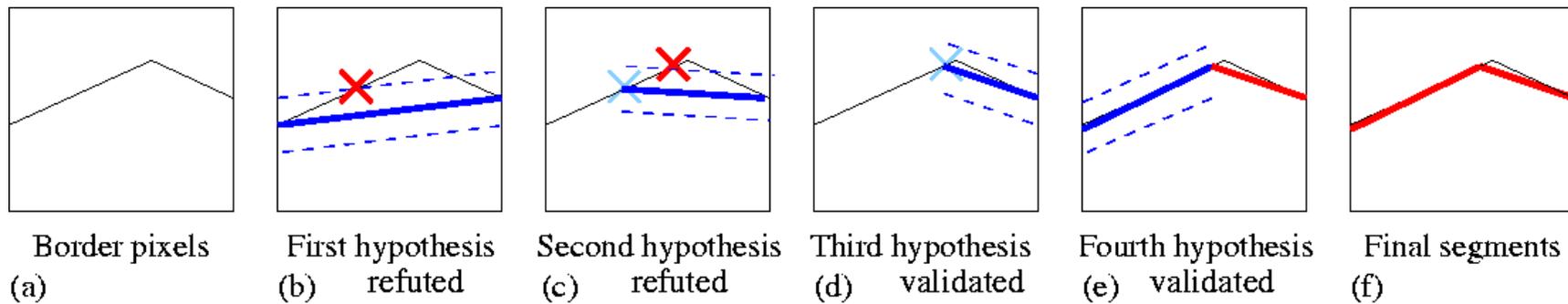
Hipótesis suelo

Comportamiento sigue-pared con visión



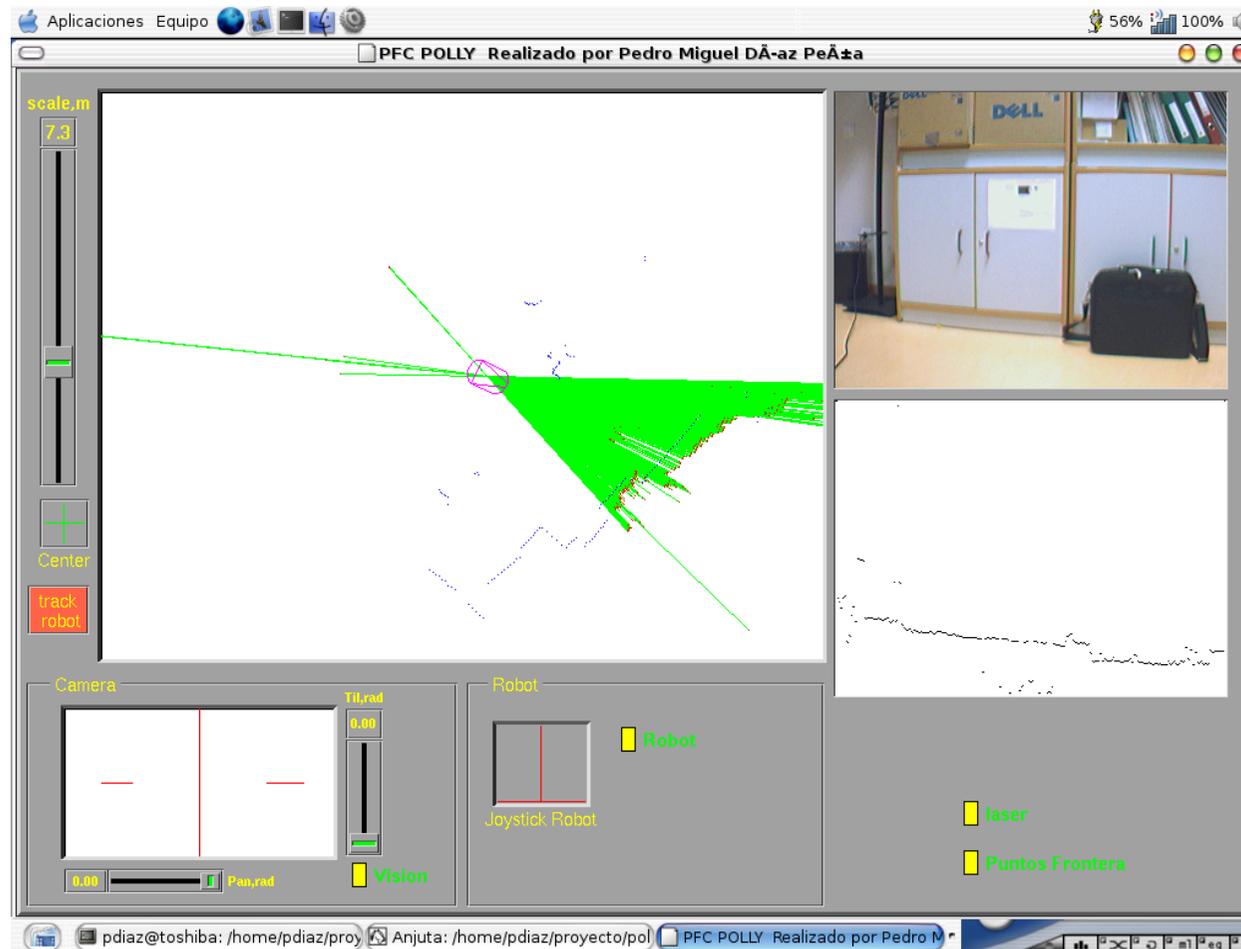
- identificar la frontera entre suelo y pared
- un esquema perceptivo y otro de actuación basado en casos
- no hay estimación explícita de distancia

SEGMENTACIÓN EN LA IMAGEN



- abducir es hipotetizar: $a \rightarrow b$
- segmentación abductiva es muy rápida

Navegación local con visión monocular



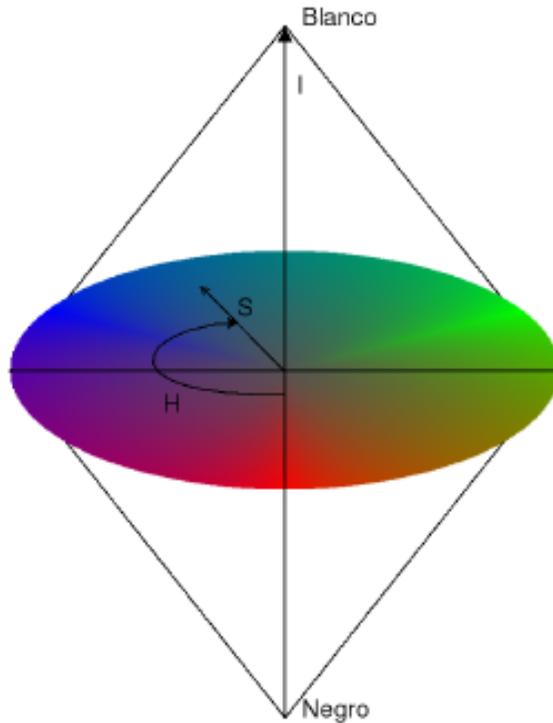
Taxias visuales con filtros de color

Comportamiento sigue-pelota con visión



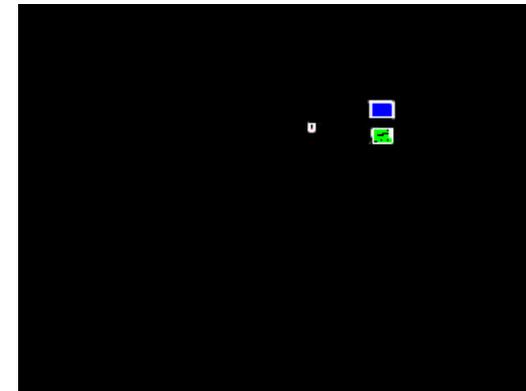
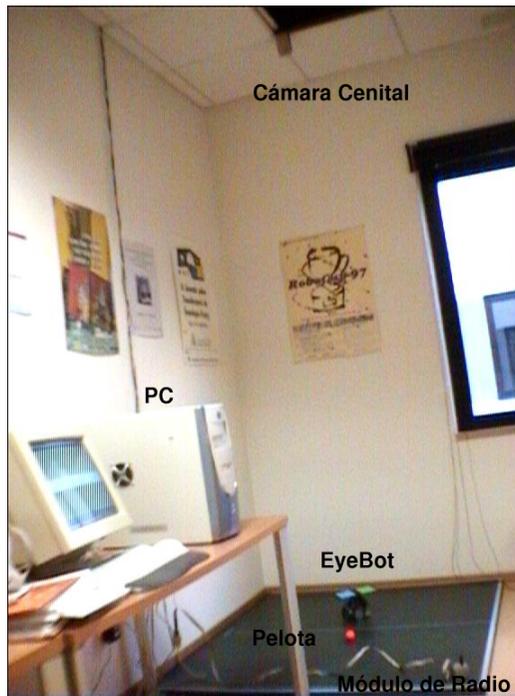
- esquema perceptivo: identificar posición de pelota
- esquema actuación: control corrección proporcional
- no hay estimación explícita de distancia

FILTRO DE COLOR



- espacios de color
- RGB:
Rmin,Rmax,Gmin,Gmax,Bmin,Bmax
- HSI:
Hmin,Hmax,Smin,Smax

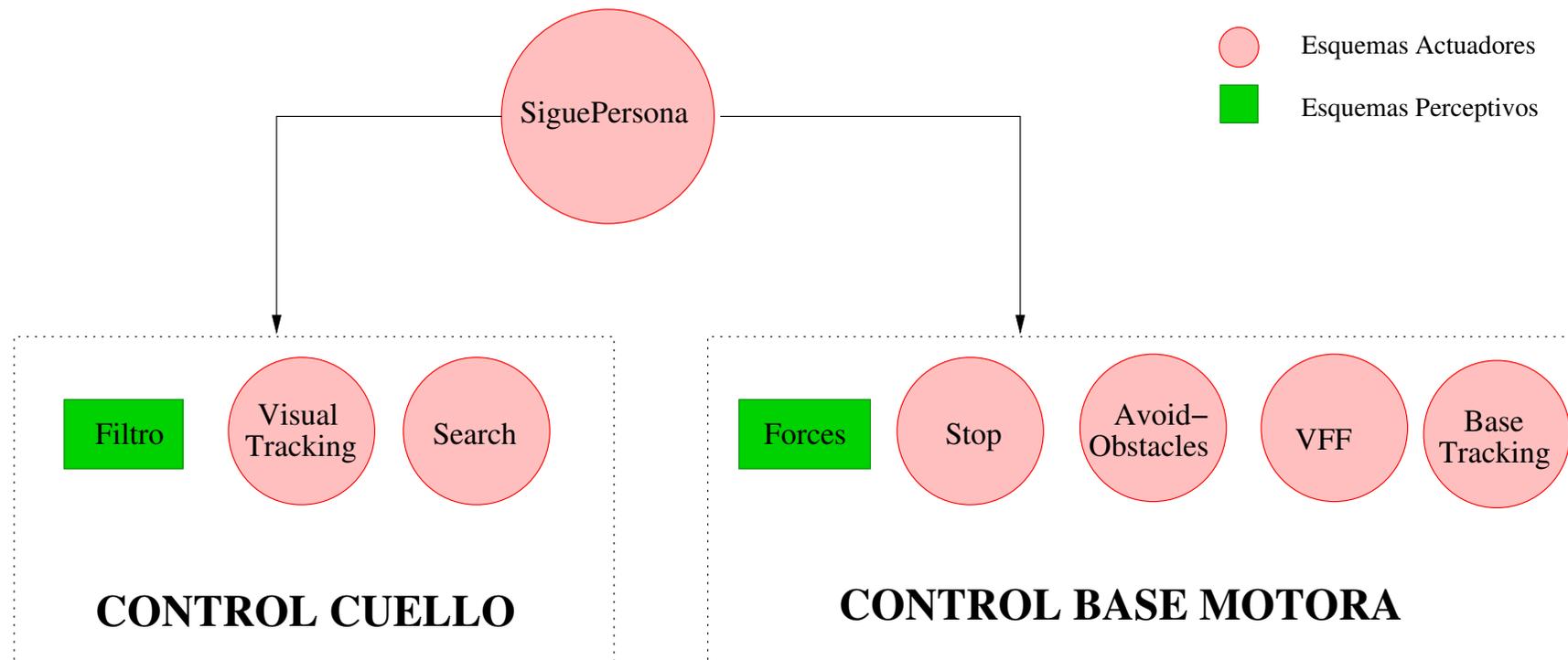
Comportamiento sigue-pelota con visión cenital



Comportamiento sigue-persona con visión



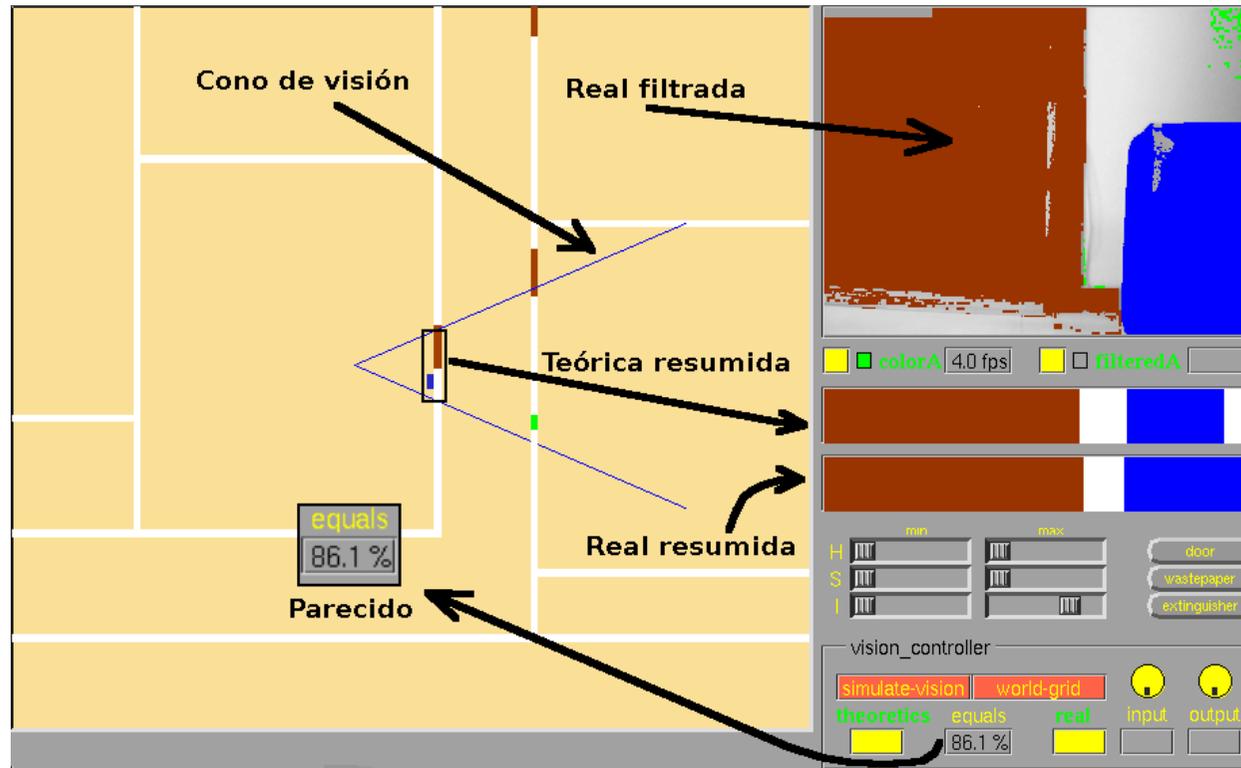
¿CÓMO?



TAMBIÉN EN EL PERRITO



Localización visual



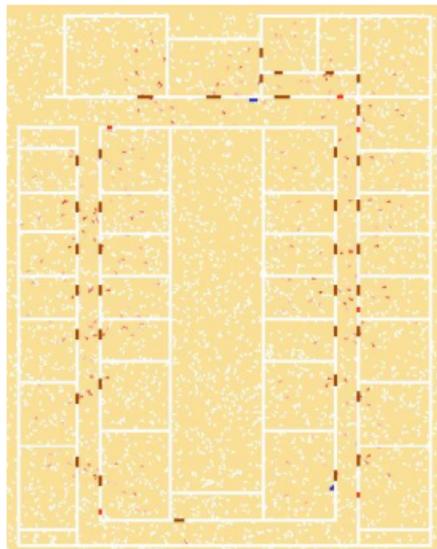
- Filtro de partículas: modelo observación, movimiento y remuestreo



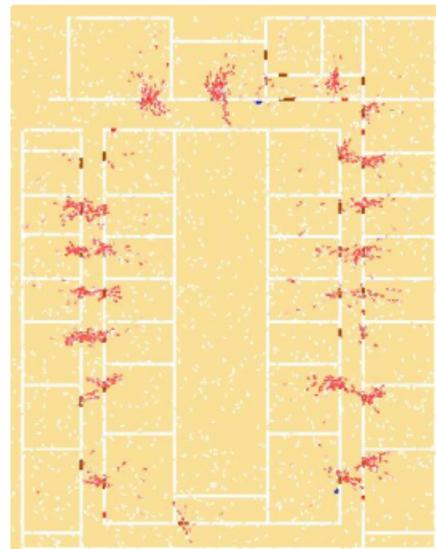
Observación 1



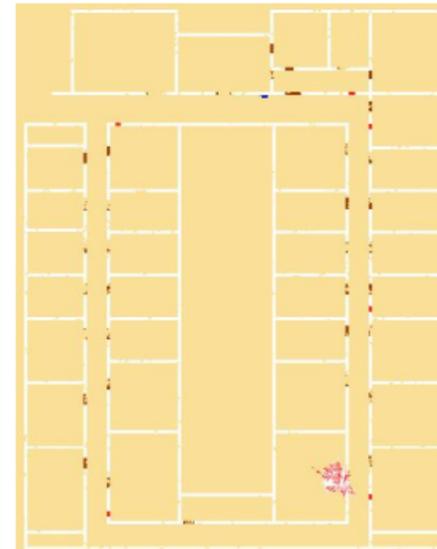
Observación 2



Distribución inicial
(uniforme)

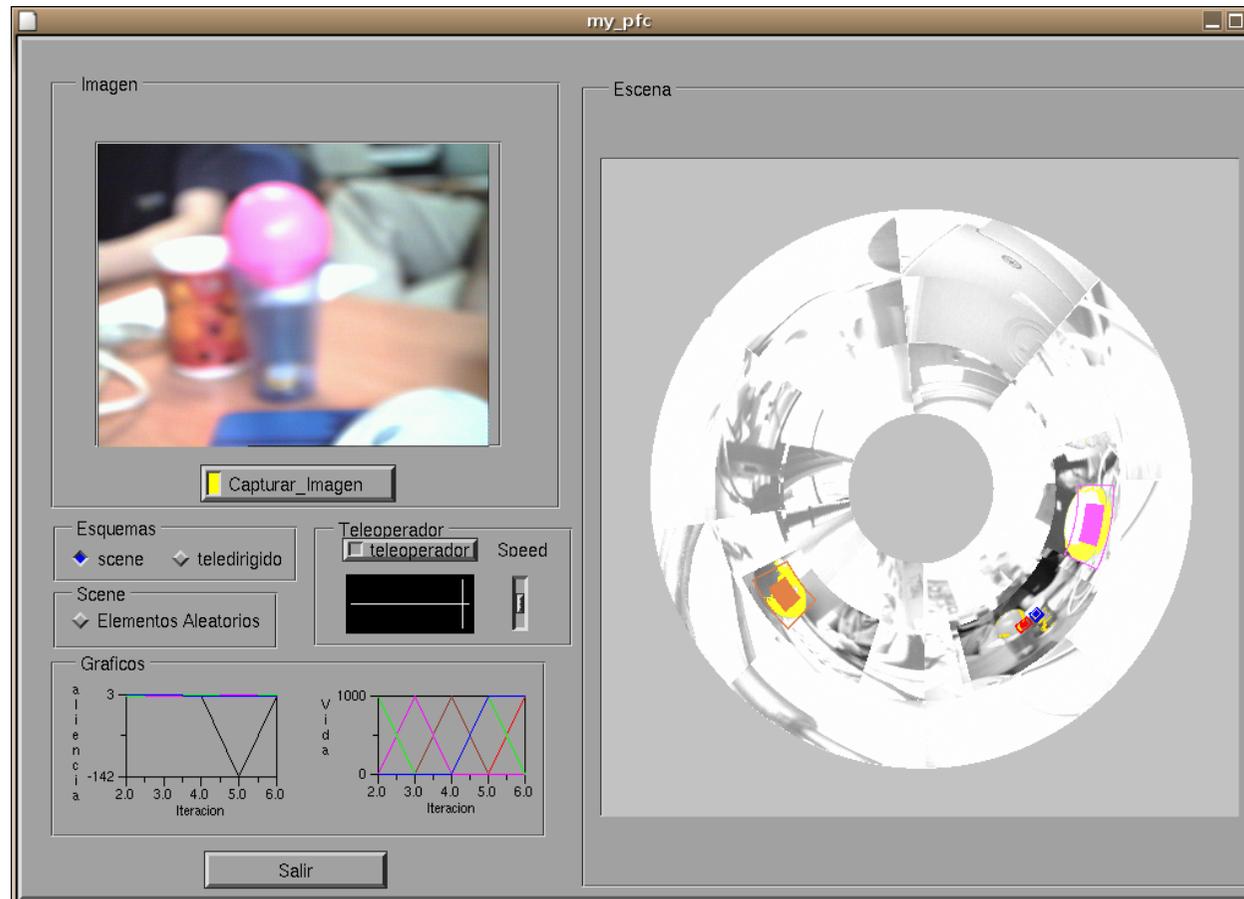


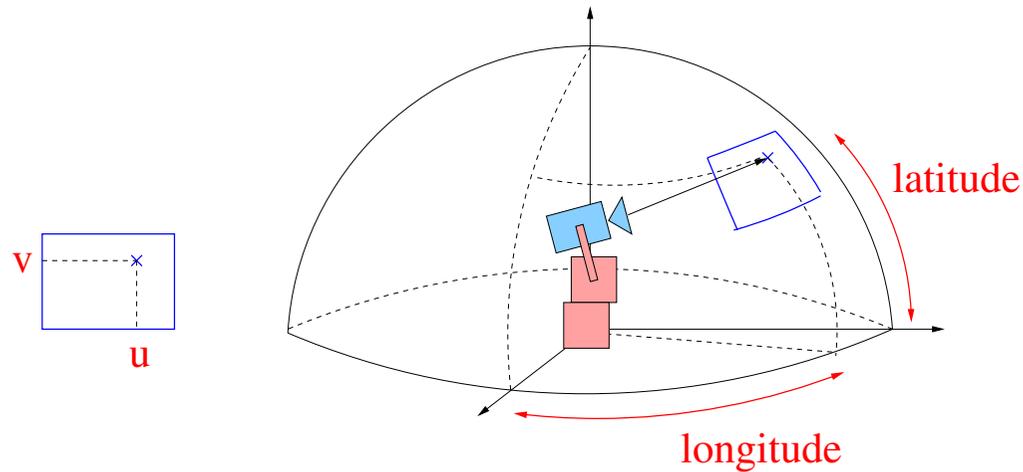
Acumulación en
zonas probables



Localización final

Atención visual



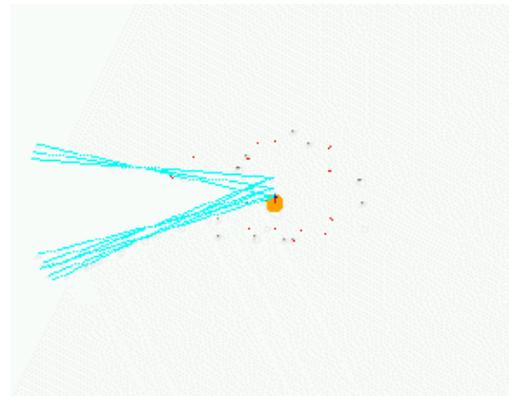
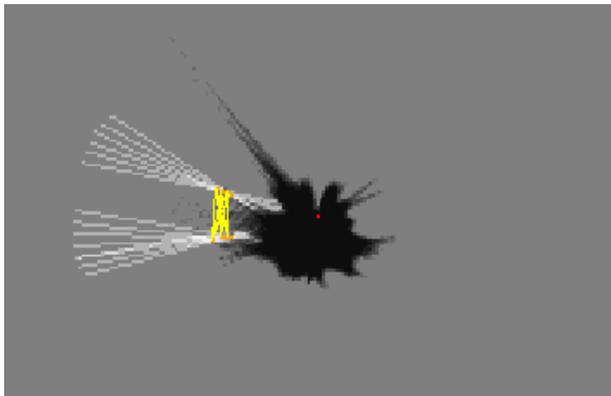


- La **imagen de escena** es más rica que la monocular
- Percepción activa: hay que mover el cuello mecánico
- Filtro para los colores relevantes
- Dinámicas temporales de saliencia y de vida
 - $liv(object, t) = liv(object, t - 1) - \Delta L_{time}$
 - $liv(object, t) = liv(object, t - 1) + \Delta L_{observation}$
 - $sal(fixp, t) = sal(fixp, t - 1) + \Delta S_{time}$
 - $sal(fixp, t) = 0$

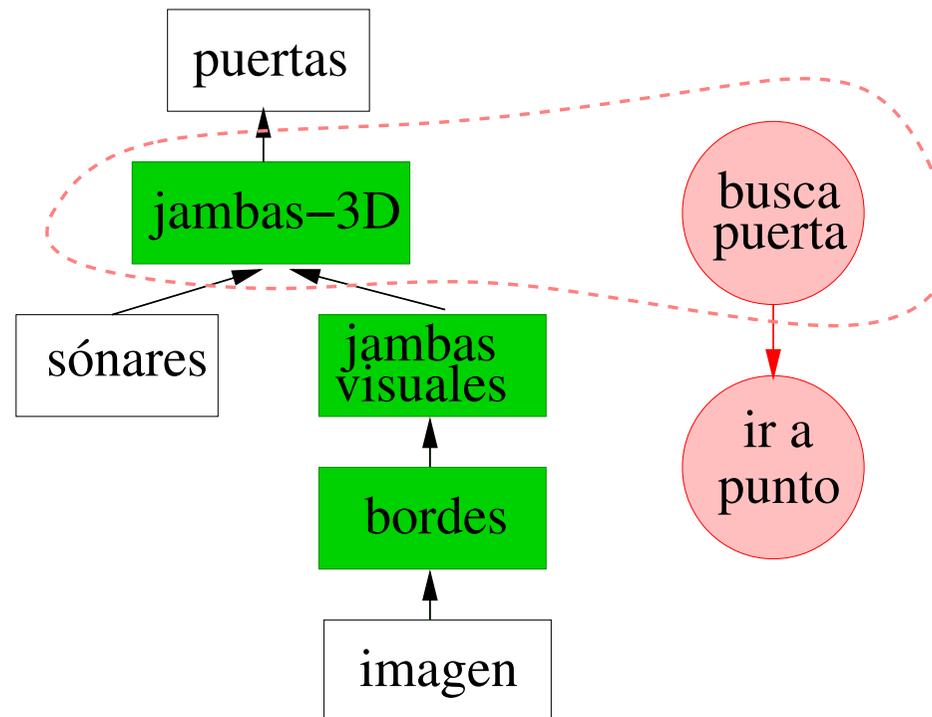
Combinación de información visual

Detección de puertas

PROFUNDIDAD FUSIONANDO EVIDENCIAS PROBABILÍSTICAMENTE



PERCEPCIÓN ESTRUCTURADA

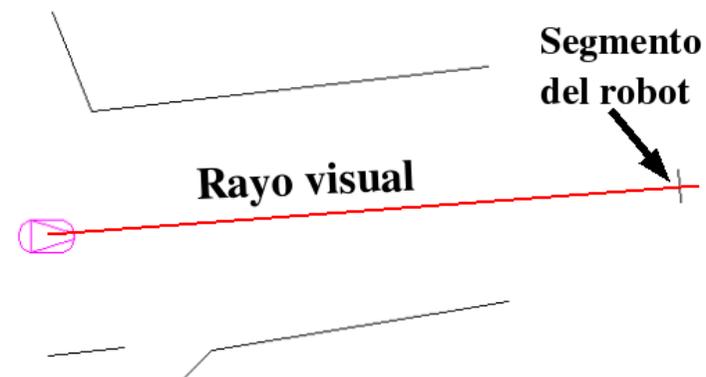
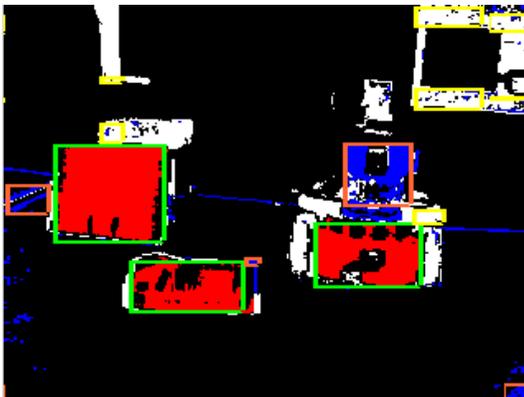
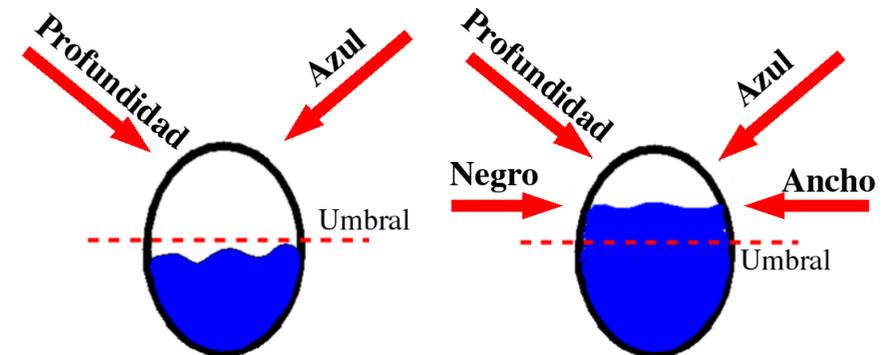


- Jerarquía de estímulos
- Interpretación y contexto de percepción
- Percepción activa: el robot se mueve para tomar más imágenes

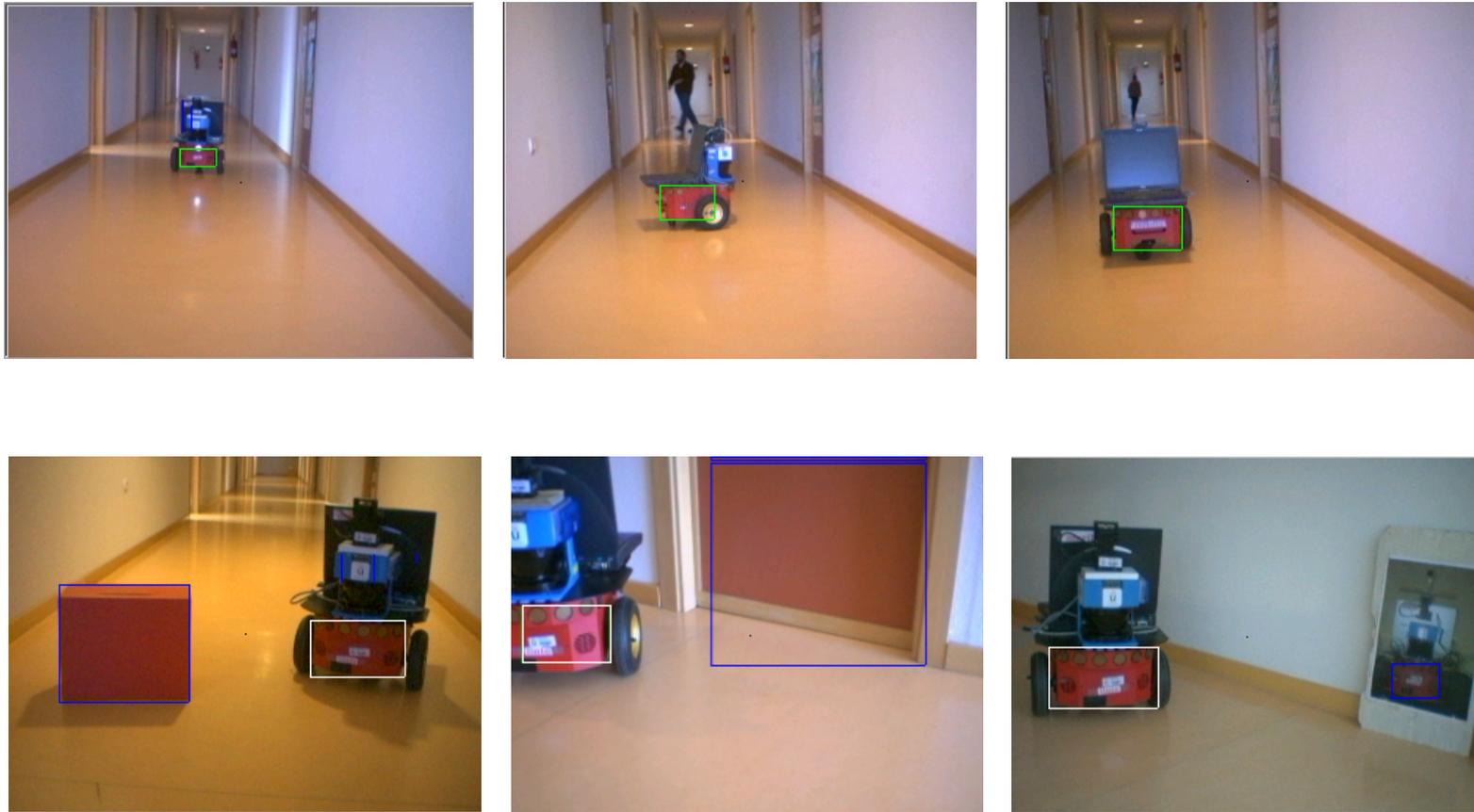
Reconocimiento de congénere

SUMA HETEROGÉNEA DE SUBESTÍMULOS

- mancha visual roja
- más ancho que alto
- mancha negra o azul encima
- sensación de profundidad

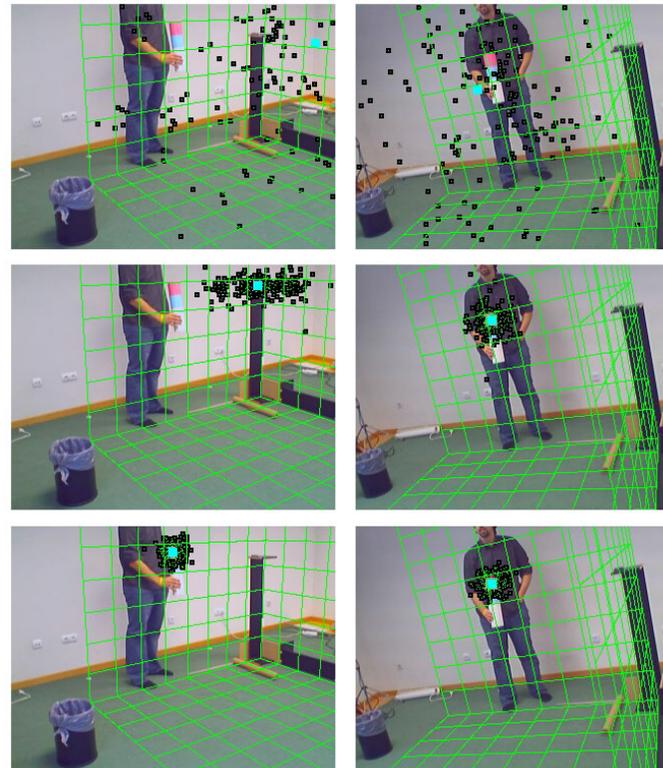


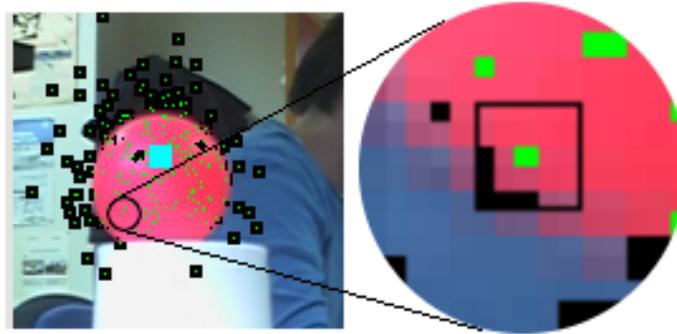
ES MUY ROBUSTO Y DISCRIMINANTE



Visión tridimensional, geometría proyectiva

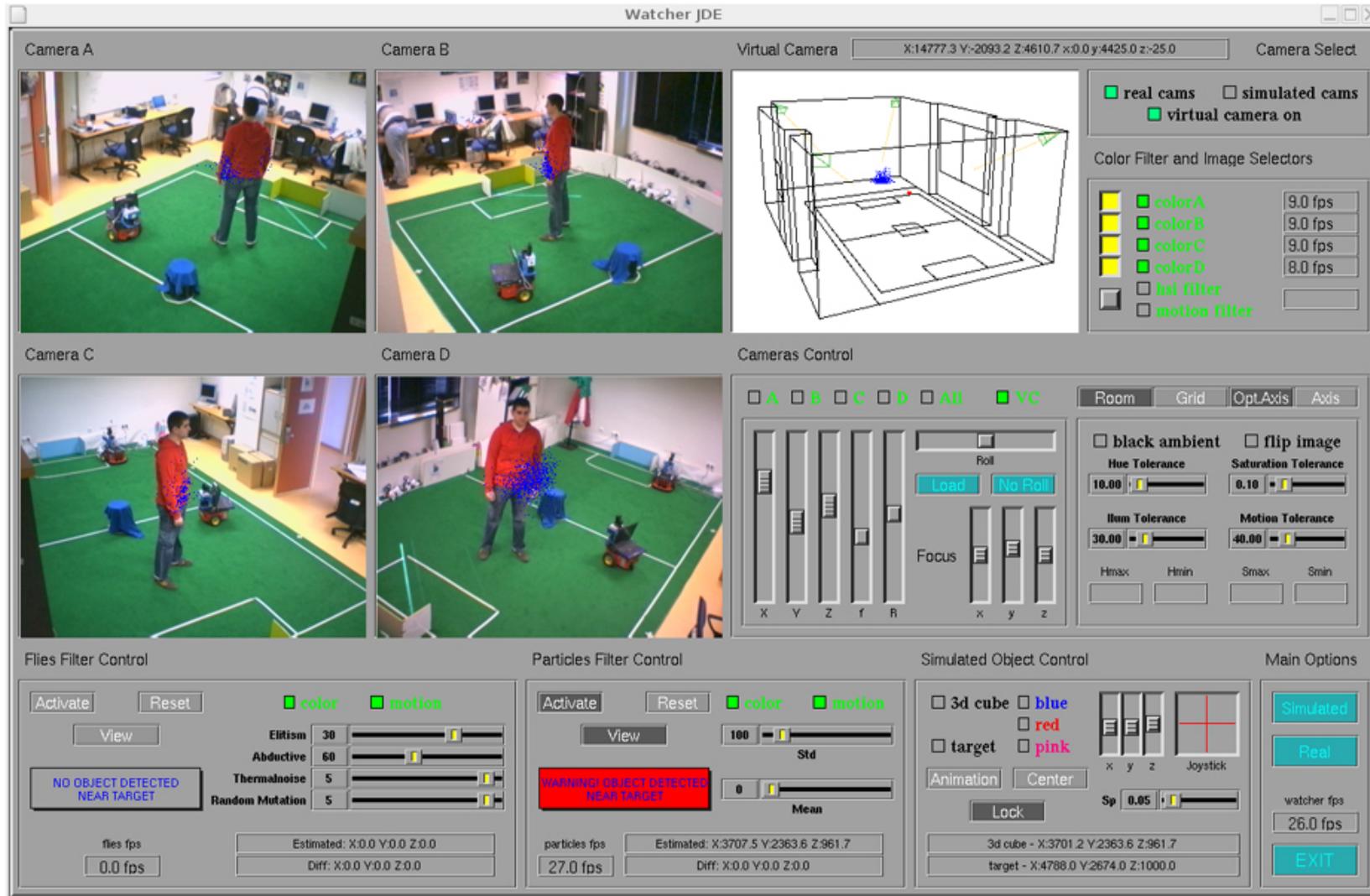
Seguimiento 3D con filtro de partículas





- Filtro de partículas: modelo observación, movimiento y remuestreo
- Población de partículas se mueve por el espacio de estados
- Cada partícula tiene un peso
- Iterativo, eficiente
- Sólo se explota el color

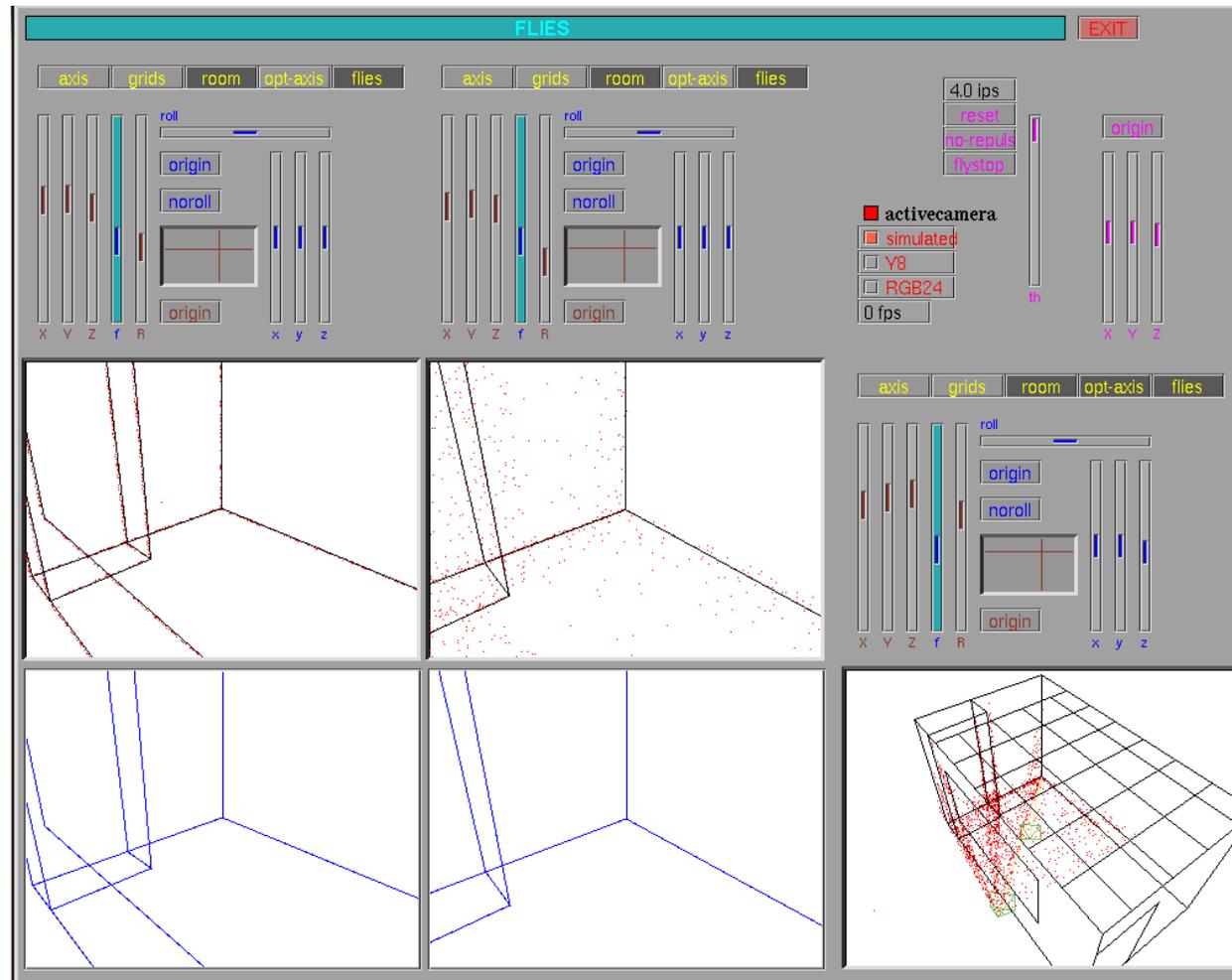
APLICACIÓN DE SEGURIDAD



The screenshot displays the 'Watcher JDE' application interface, which is used for monitoring a 3D environment. The interface is divided into several sections:

- Top Row:**
 - Camera A, Camera B, Camera C, Camera D:** Four real-time video feeds from different camera angles showing a person in a red shirt and a robot in a green field.
 - Virtual Camera:** A 3D wireframe view of the environment with a virtual camera's field of view indicated by a yellow cone. Coordinates are shown as X:14777.3 Y:-2093.2 Z:4610.7 x:0.0 y:4425.0 z:-25.0.
 - Camera Select:** A control panel with checkboxes for 'real cams', 'simulated cams', and 'virtual camera on'.
- Color Filter and Image Selectors:** A panel with color selection buttons (colorA, colorB, colorC, colorD) and checkboxes for 'hsi filter' and 'motion filter'. Each color has an associated FPS value (9.0 or 8.0).
- Cameras Control:** A panel with checkboxes for cameras A, B, C, D, All, and VC. It includes sliders for X, Y, Z, f, R and x, y, z, and buttons for 'Load' and 'No Roll'.
- Room, Grid, Opt.Axis, Axis:** A panel with checkboxes for 'black ambient' and 'flip image', and sliders for 'Hue Tolerance', 'Saturation Tolerance', 'Hm Tolerance', and 'Motion Tolerance'.
- Flies Filter Control:** A panel with 'Activate', 'Reset', and 'View' buttons. It includes sliders for 'Elitism', 'Abductive', 'Thermalnoise', and 'Random Mutation'. A status box shows 'NO OBJECT DETECTED NEAR TARGET' and 'files fps' as 0.0.
- Particles Filter Control:** A panel with 'Activate', 'Reset', and 'View' buttons. It includes sliders for 'Std' and 'Mean'. A status box shows 'WARNING! OBJECT DETECTED NEAR TARGET' and 'particles fps' as 27.0.
- Simulated Object Control:** A panel with checkboxes for '3d cube', 'blue', 'red', 'pink', and 'target'. It includes a joystick control and a 'Lock' button. Status boxes show coordinates for '3d cube' and 'target'.
- Main Options:** A panel with 'Simulated', 'Real', and 'EXIT' buttons, and a 'watcher fps' display showing 26.0.

Reconstrucción tridimensional con algoritmo evolutivo



Conclusiones

- La visión es un sensor más para los robots: barato, rico y complejo
- **El procesamiento visual ha de encajar en el planteamiento de percepción que ofrece la arquitectura.** JDE: colección dinámica y jerarquizada de estímulos
- **No hace falta procesamiento visual muy complicado para generar comportamiento complejo.** Ejemplos: hipótesis suelo, taxias de color, atención, seguimiento, reconstrucción, combinación subestímulos...
- Línea futura: atención 3D, reconstrucción 3D