
VISIÓN EN ROBOTS MÓVILES

José María Cañas Plaza

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



INAOE, Puebla (México), agosto 2005

Contenidos

- Introducción
- Percepción en robots
- Experimentos
 - Hipótesis suelo
 - Taxias con filtros de color
 - Localización desde visión
 - Atención visual
 - Bordes, detección de puertas
 - 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

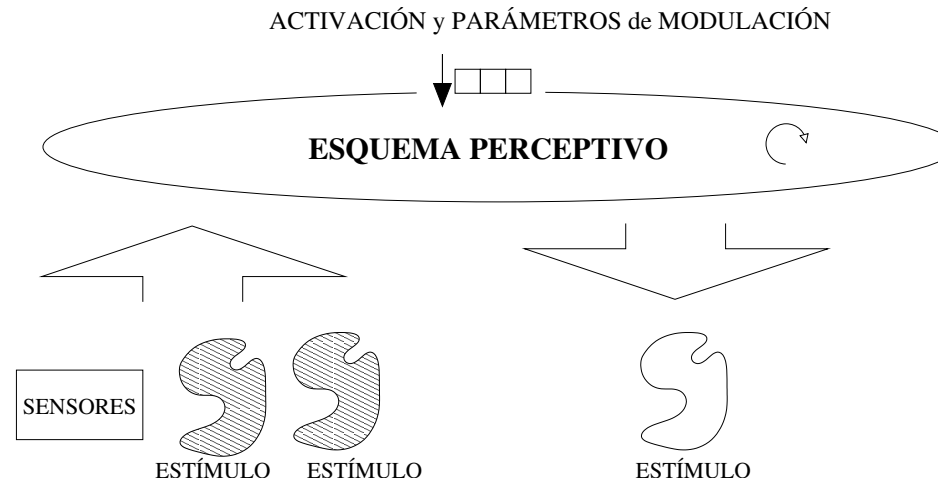
La *arquitectura* de un robot es la *organización* de sus capacidades sensoriales, de procesamiento y de acción para conseguir un repertorio de comportamientos inteligentes interactuando con cierto entorno

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

Arquitectura JDE

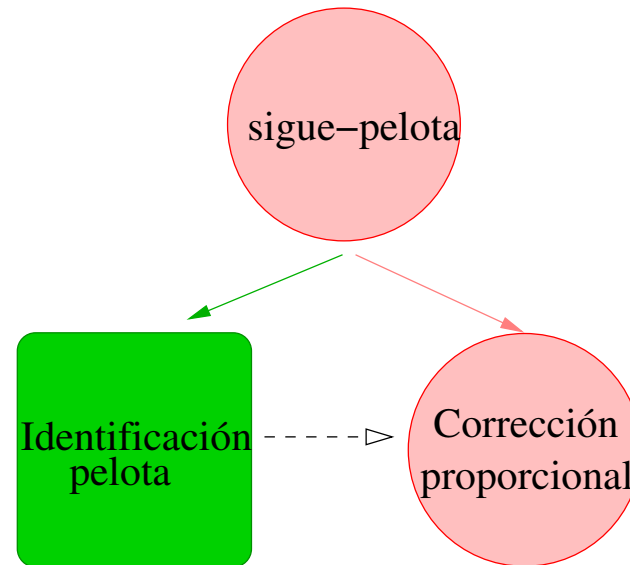
- Comportamiento = **percepción** y control
- Fragmentación en unidades (**esquemas**) asíncronas concurrentes
 - de percepción elaboran estímulos
 - de actuación toman decisiones
- La colección de esquemas se organiza en **jerarquía**
- Jerarquía Dinámica de Esquemas

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**)

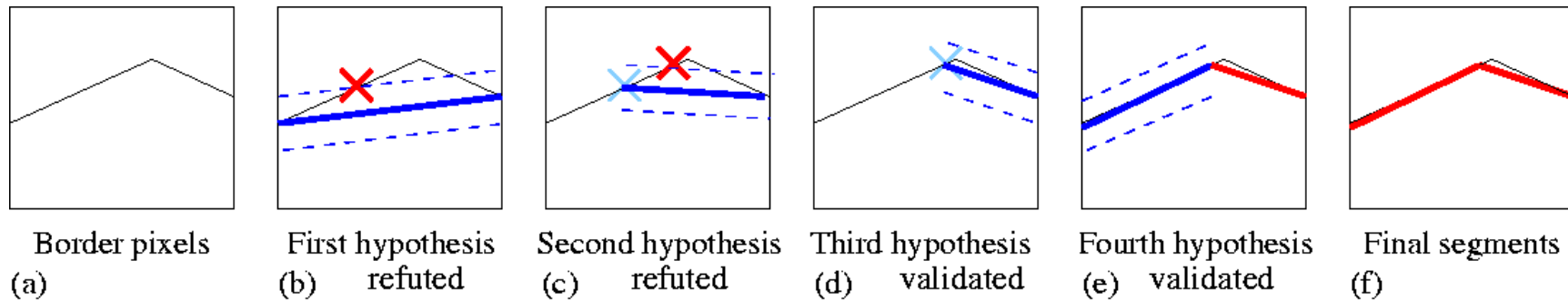
Hipótesis suelo

Comportamiento sigue-pared con visión



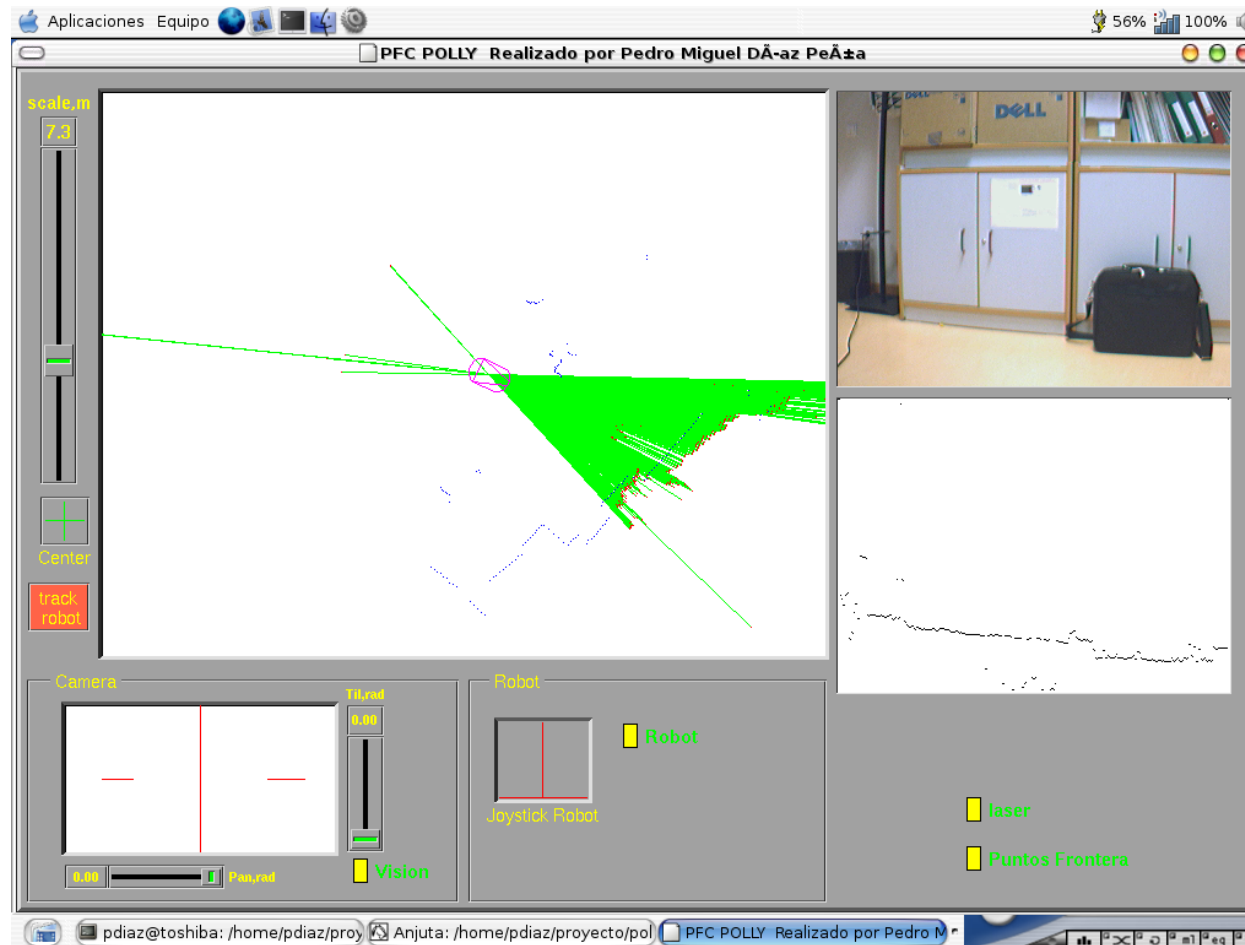
- identificar la frontera entre suelo y pared
- un esquema perceptivo y otro de actuación basado en casos
- no hay estimación explícita de distancia

SEGMENTACIÓN EN LA IMAGEN



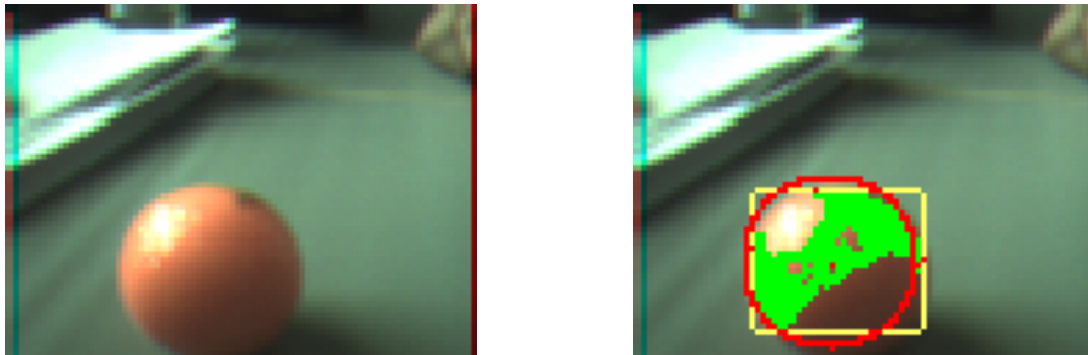
- abducir es hipotetizar: $a \rightarrow b$
- segmentación abductiva es muy rápida

Navegación local con visión monocular



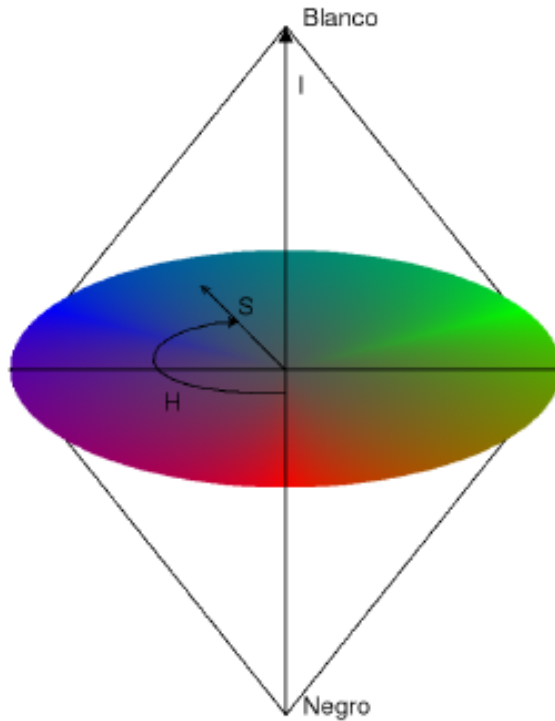
Taxias visuales con filtros de color

Comportamiento sigue-pelota con visión



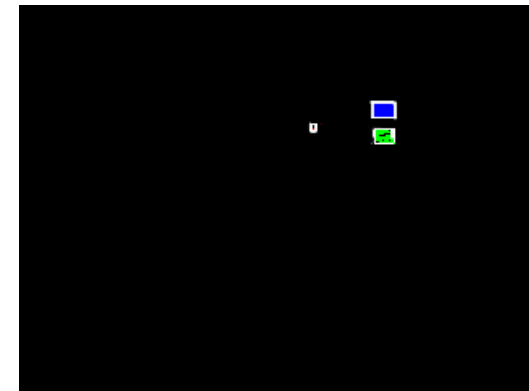
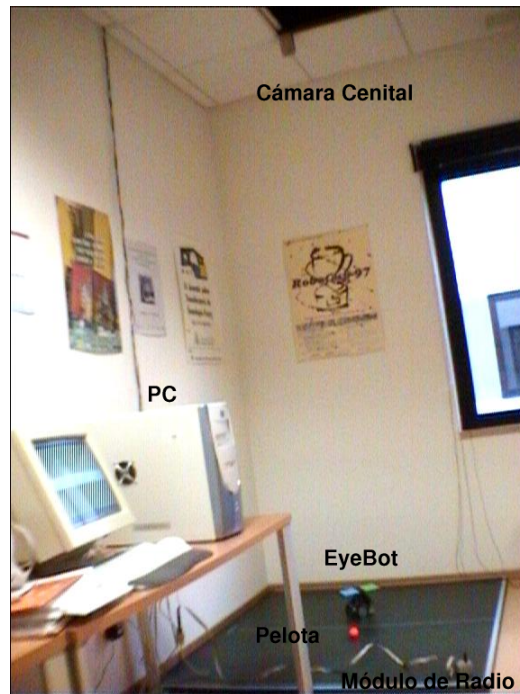
- esquema perceptivo: identificar posición de pelota
- esquema actuación: control corrección proporcional
- no hay estimación explícita de distancia

FILTRO DE COLOR



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

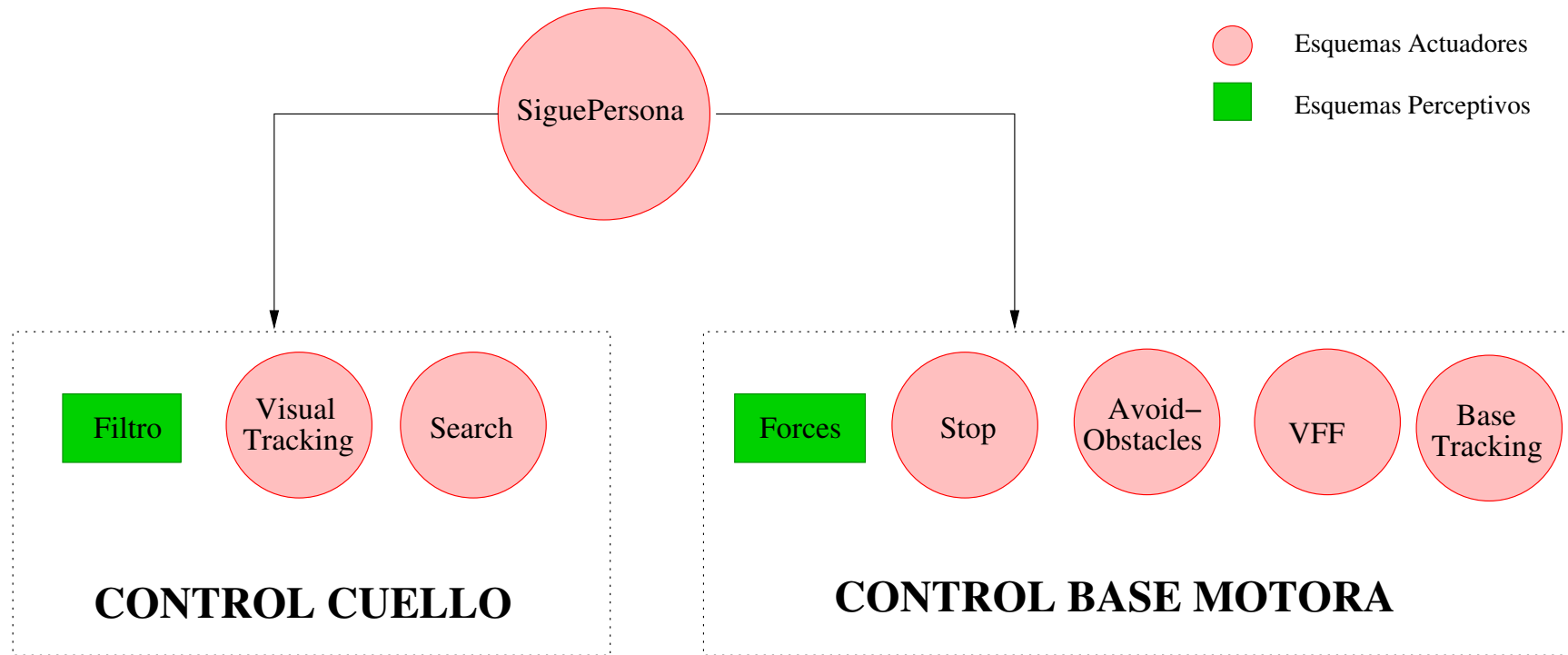
Comportamiento sigue-pelota con visión cenital



Comportamiento sigue-persona con visión



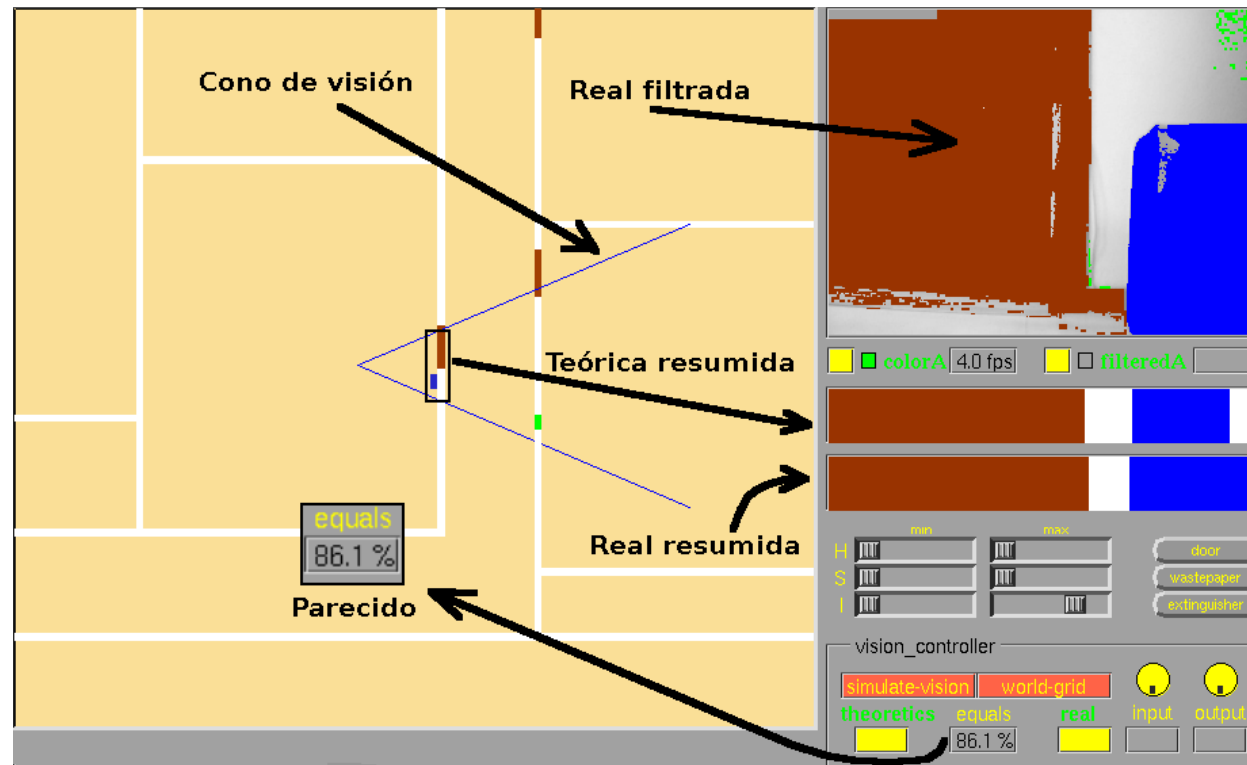
¿CÓMO?



TAMBIÉN EN EL PERRITO



Localización visual



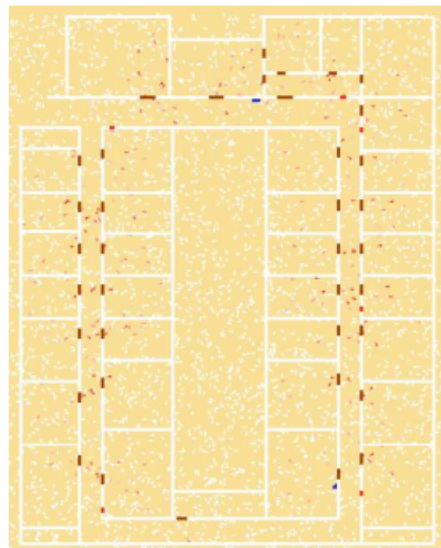
- Filtro de partículas: modelo observación, movimiento y remuestreo



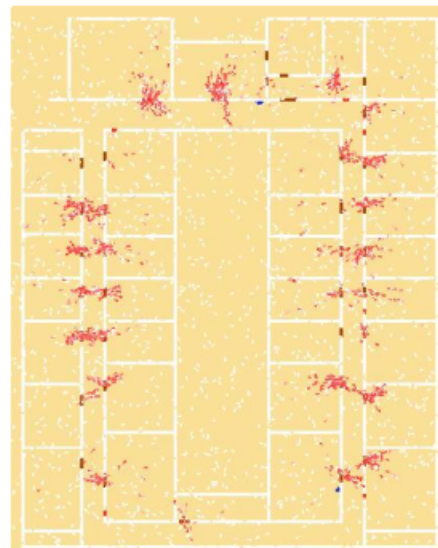
Observación 1



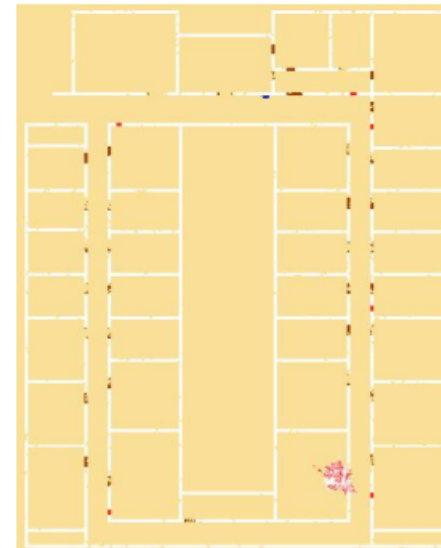
Observación 2



Distribución inicial
(uniforme)

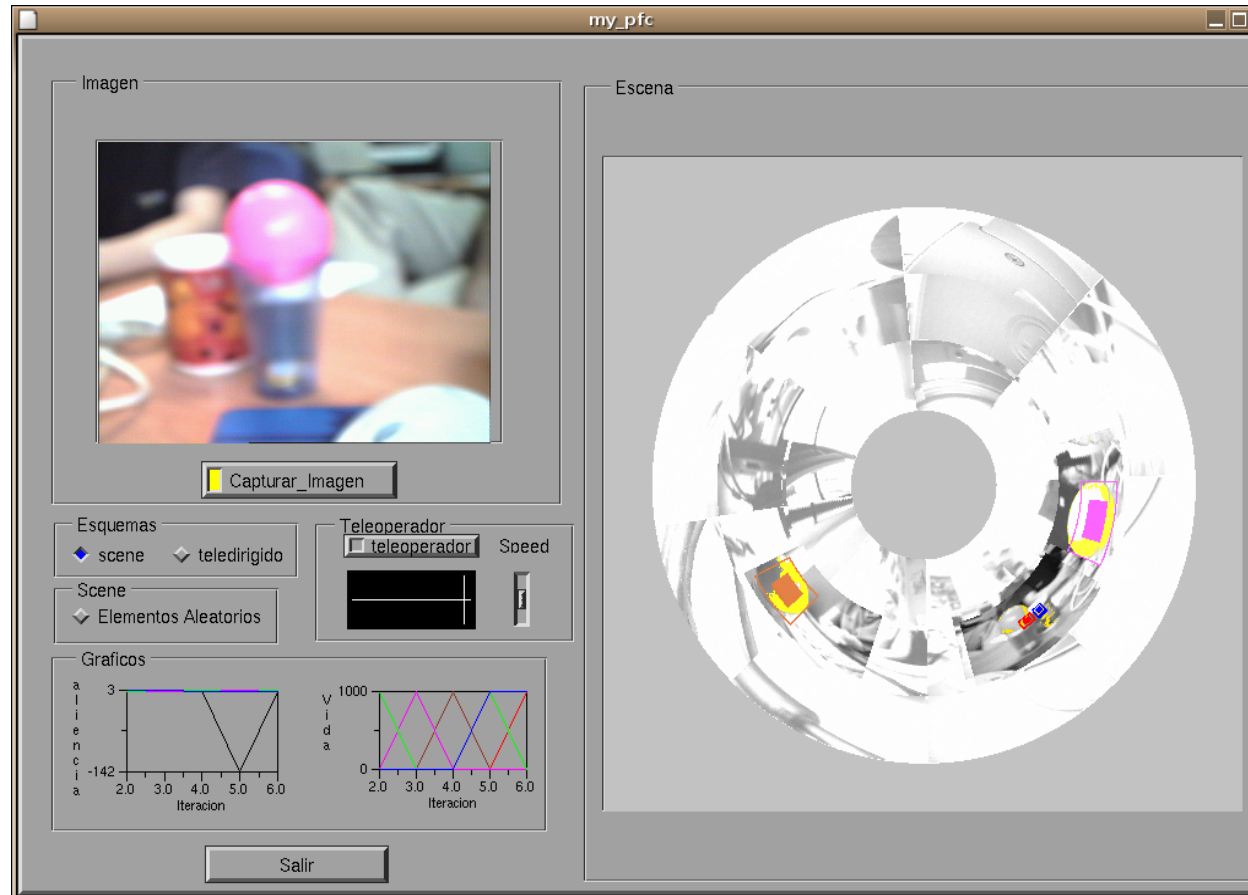


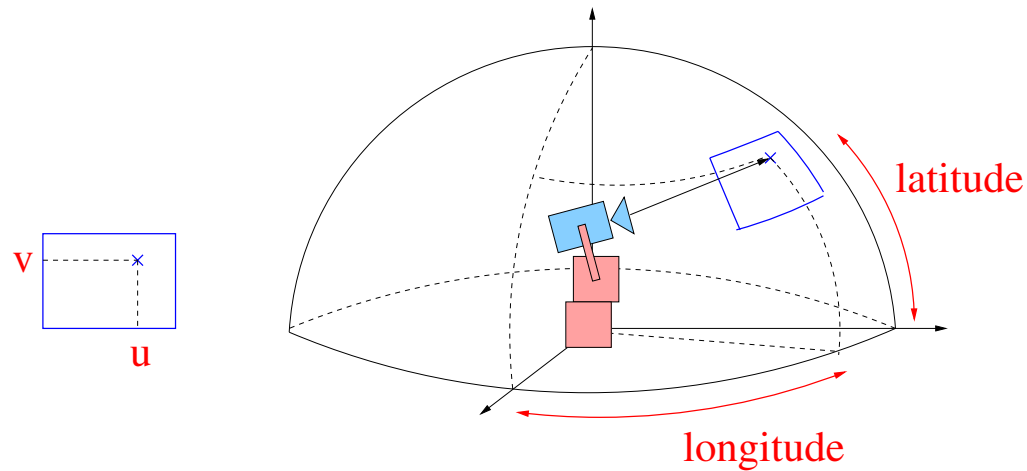
Acumulación en
zonas probables



Localización final

Atención visual

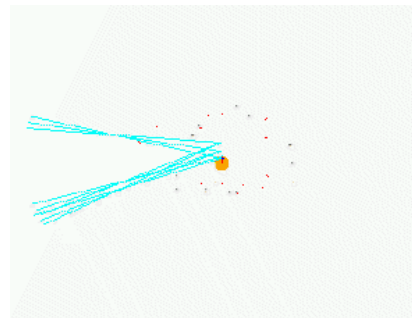
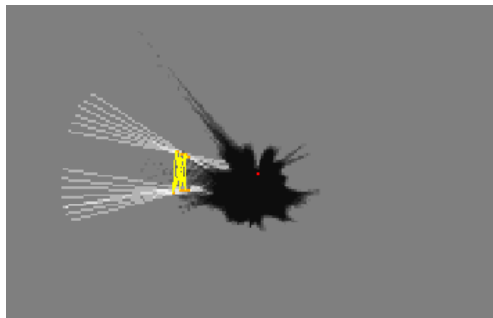




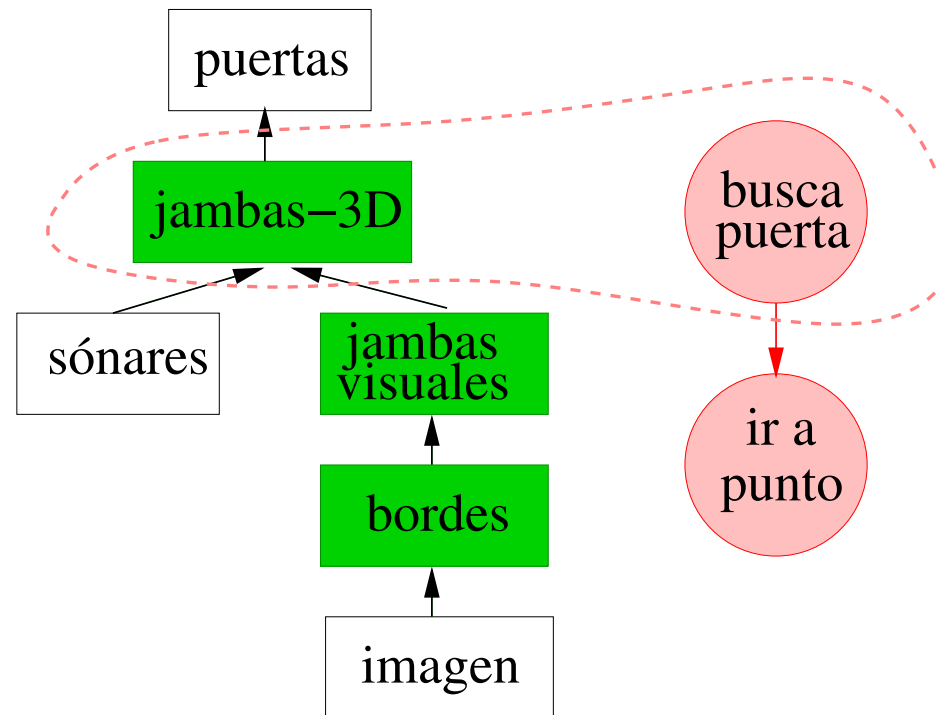
- La **imagen de escena** es más rica que la monocular
- Percepción activa: hay que mover el cuello mecánico
- Filtro para los colores relevantes
- Dinámicas temporales de saliencia y de vida
 - $liv(object, t) = liv(object, t - 1) - \Delta L_{time}$
 - $liv(object, t) = liv(object, t - 1) + \Delta L_{observation}$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

Detección de puertas

Profundidad fusionando evidencias



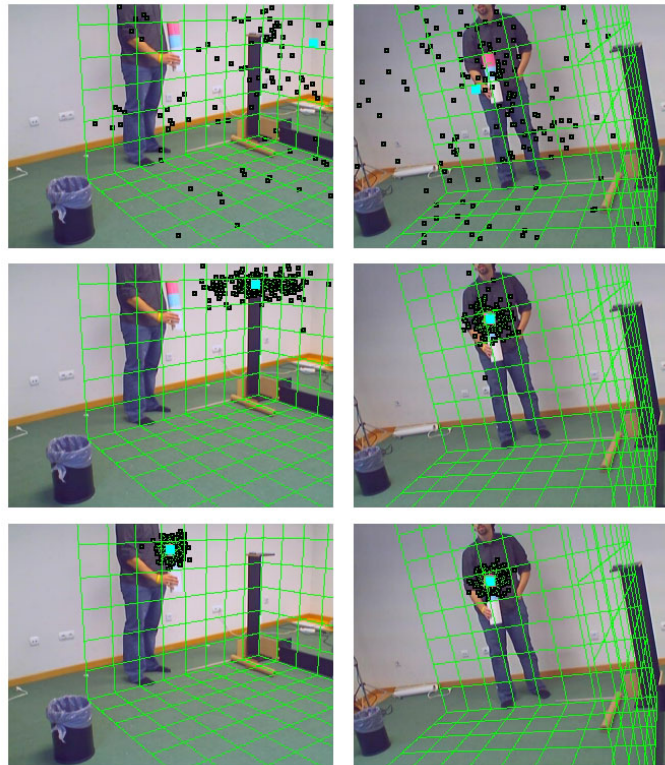
Percepción estructurada

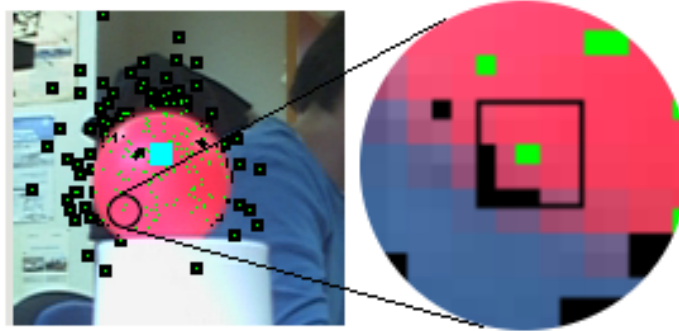


- Jerarquía de estímulos
- Interpretación y contexto de percepción
- Percepción activa: el robot se mueve para tomar más imágenes

Visión tridimensional, geometría proyectiva

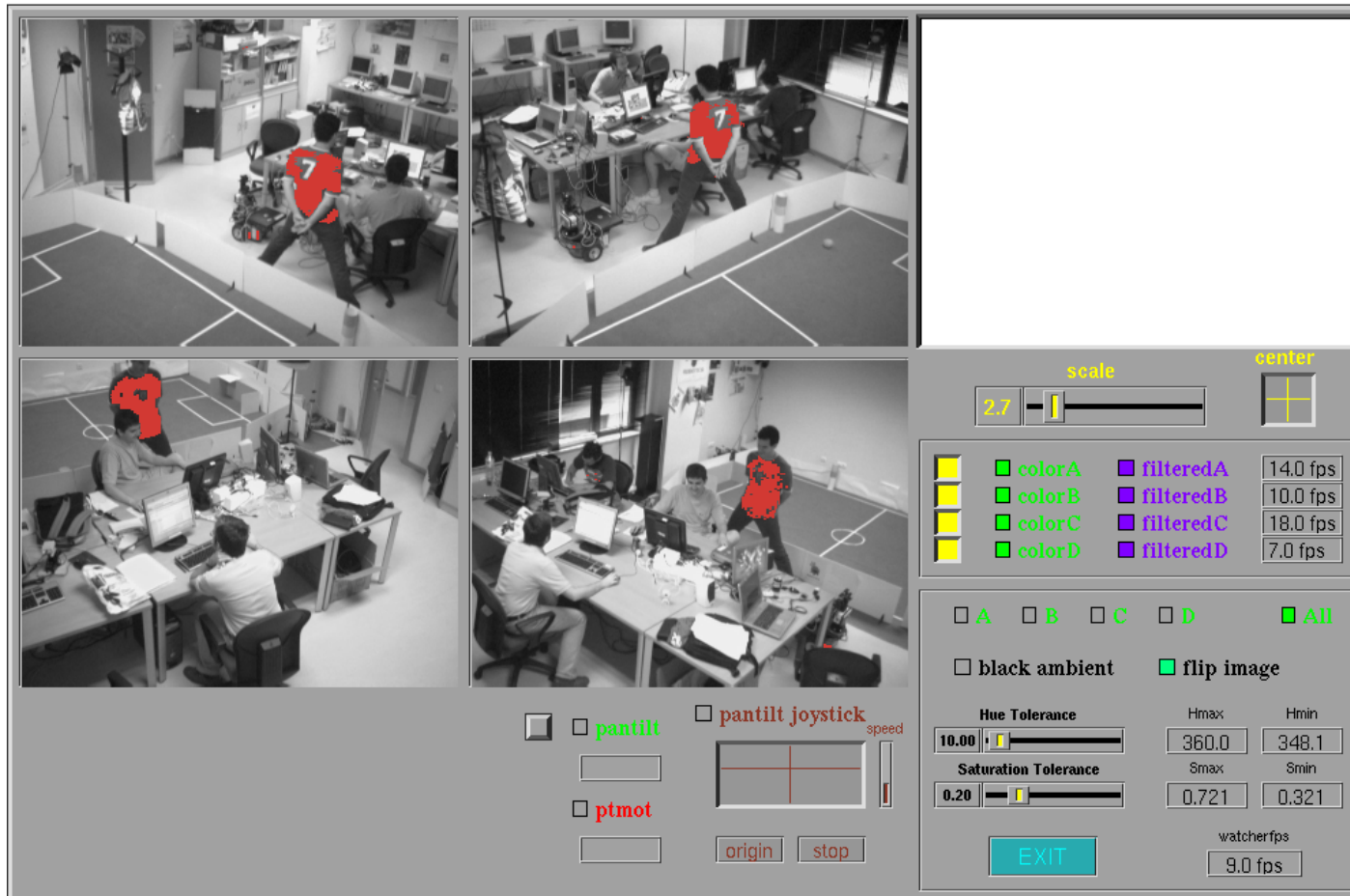
Seguimiento 3D con filtro de partículas





- 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

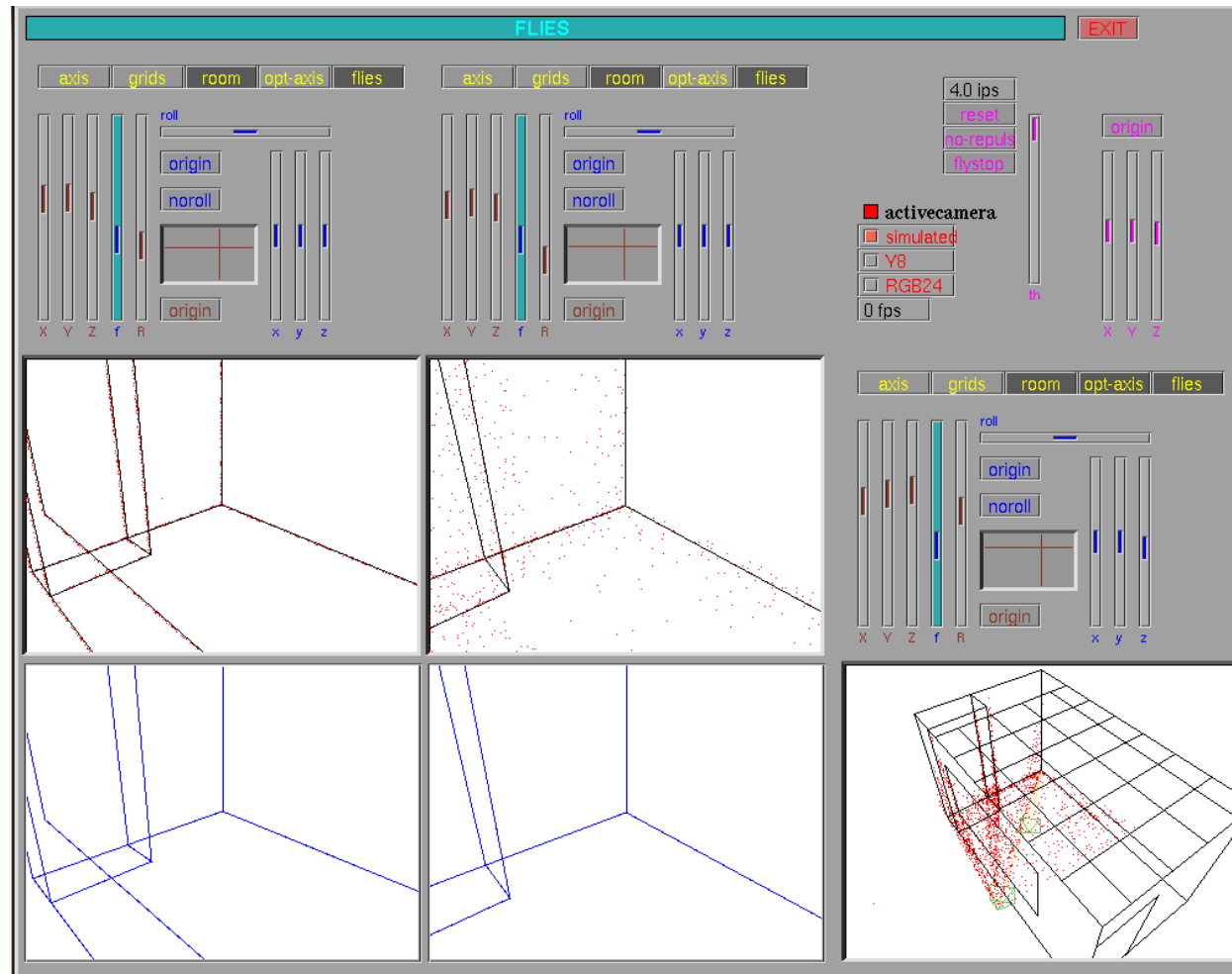


The interface displays four camera views of a room where a robot is being used. The robot is highlighted in red in the views. The control panel includes the following elements:

- Scale:** A slider set to 2.7.
- Center:** A crosshair icon.
- Color Selection:**

| | | | | |
|--------------------------|--------|--------------------------|-----------|----------|
| <input type="checkbox"/> | colorA | <input type="checkbox"/> | filteredA | 14.0 fps |
| <input type="checkbox"/> | colorB | <input type="checkbox"/> | filteredB | 10.0 fps |
| <input type="checkbox"/> | colorC | <input type="checkbox"/> | filteredC | 18.0 fps |
| <input type="checkbox"/> | colorD | <input type="checkbox"/> | filteredD | 7.0 fps |
- Tracking:**
 - A B C D All
 - black ambient flip image
- Tracking Parameters:**
 - Hue Tolerance:** 10.00 (Hmax: 360.0, Hmin: 348.1)
 - Saturation Tolerance:** 0.20 (Smax: 0.721, Smin: 0.321)
 - watcherfps:** 9.0 fps
- Navigation:**
 - pantilt
 - ptmot
 - pantilt joystick (with speed slider)
 - origin stop
- EXIT** button

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...
- Línea futura: Percibir a otro robot, combinación aditiva de estímulos