
ROBOTS QUE USAN SUS OJOS

visión en robots móviles

José María Cañas Plaza

<http://gsyc.esct.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

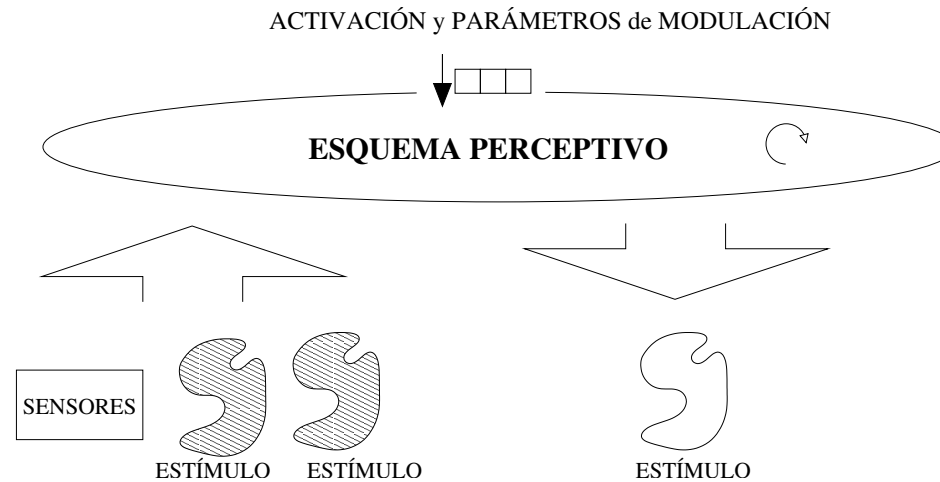
La *arquitectura* de un robot es la *organización* de sus capacidades sensoriales, de procesamiento y de acción para conseguir un repertorio de comportamientos inteligentes interactuando con cierto entorno

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

Arquitectura JDE

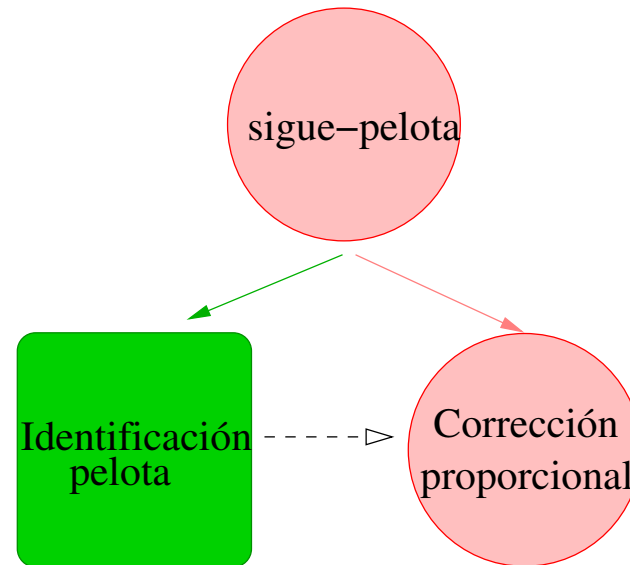
- Comportamiento = **percepción** y control
- Fragmentación en unidades (**esquemas**) asíncronas concurrentes
 - de percepción elaboran estímulos
 - de actuación toman decisiones
- La colección de esquemas se organiza en **jerarquía**
- Jerarquía Dinámica de Esquemas

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**)

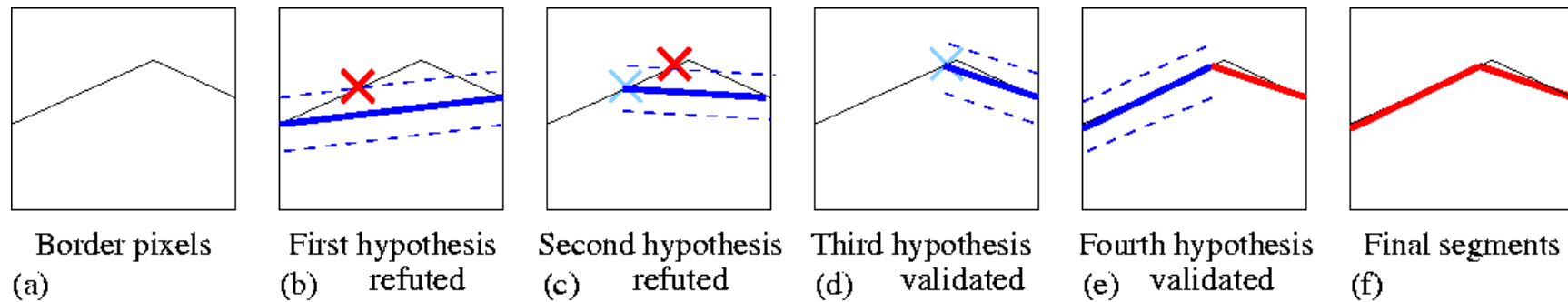
Hipótesis suelo

Comportamiento sigue-pared con visión



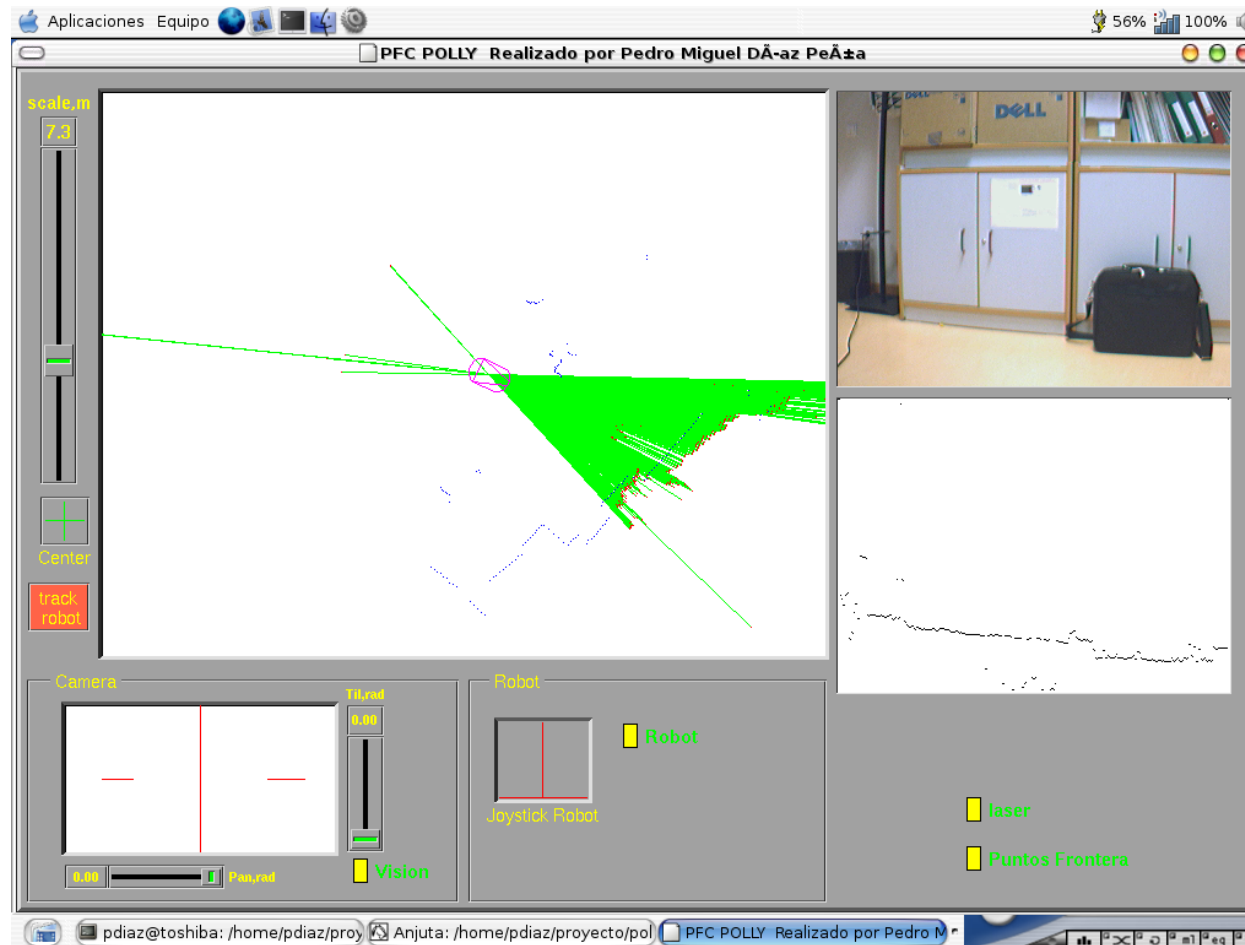
- identificar la frontera entre suelo y pared
- un esquema perceptivo y otro de actuación basado en casos
- no hay estimación explícita de distancia

SEGMENTACIÓN EN LA IMAGEN



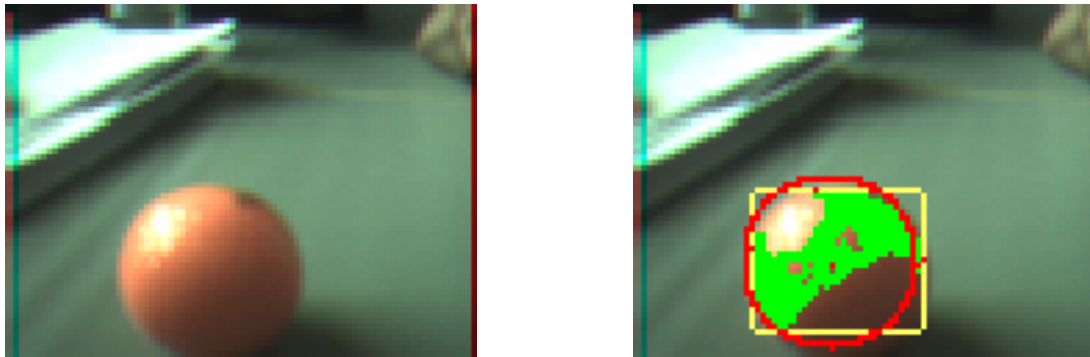
- abducir es hipotetizar: $a \rightarrow b$
- segmentación abductiva es muy rápida

Navegación local con visión monocular



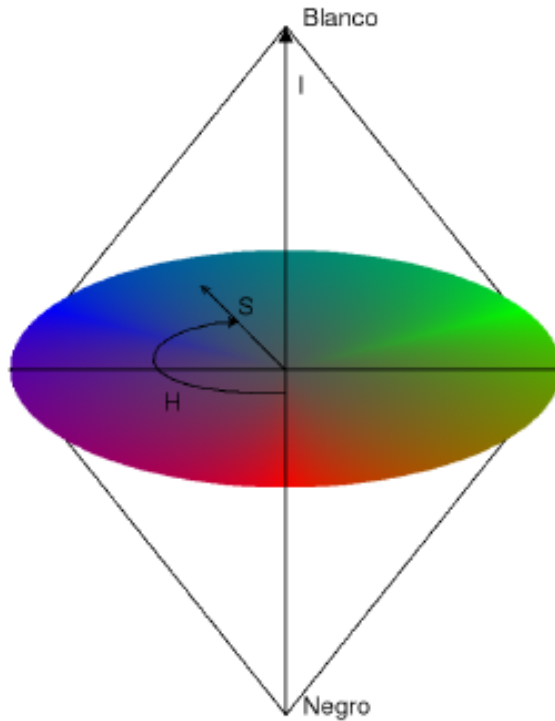
Taxias visuales con filtros de color

Comportamiento sigue-pelota con visión



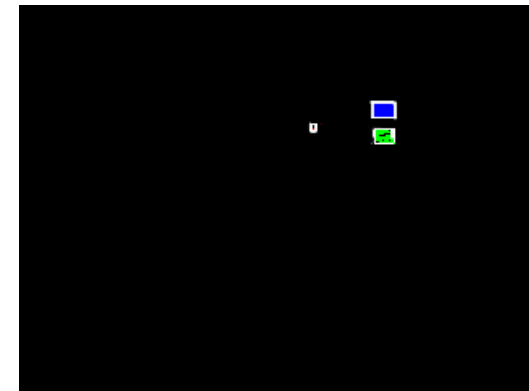
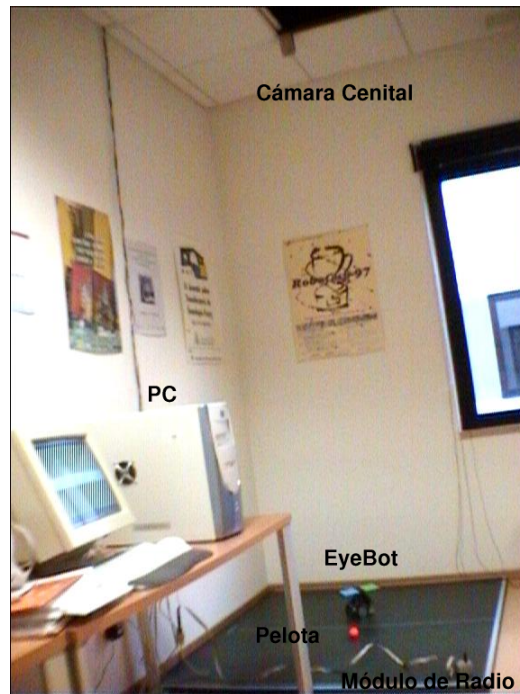
- esquema perceptivo: identificar posición de pelota
- esquema actuación: control corrección proporcional
- no hay estimación explícita de distancia

FILTRO DE COLOR



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

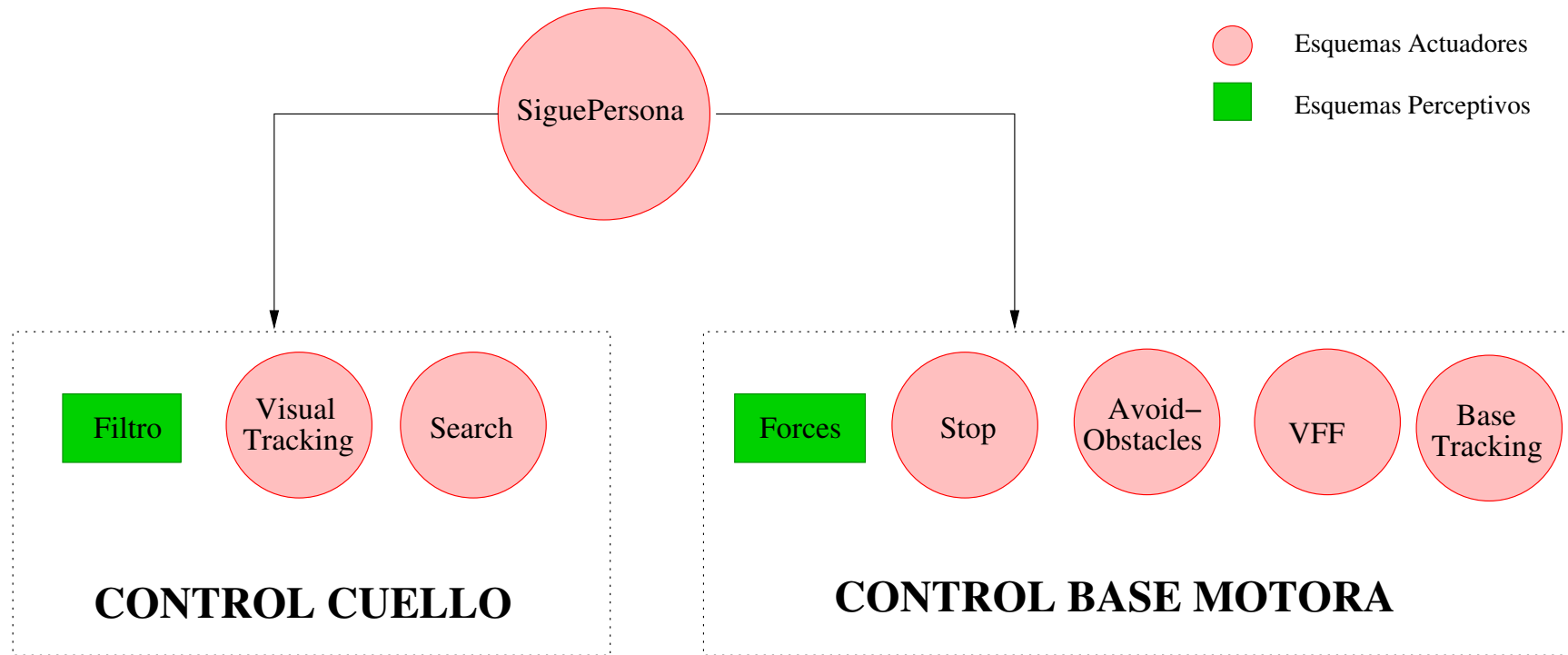
Comportamiento sigue-pelota con visión cenital



Comportamiento sigue-persona con visión



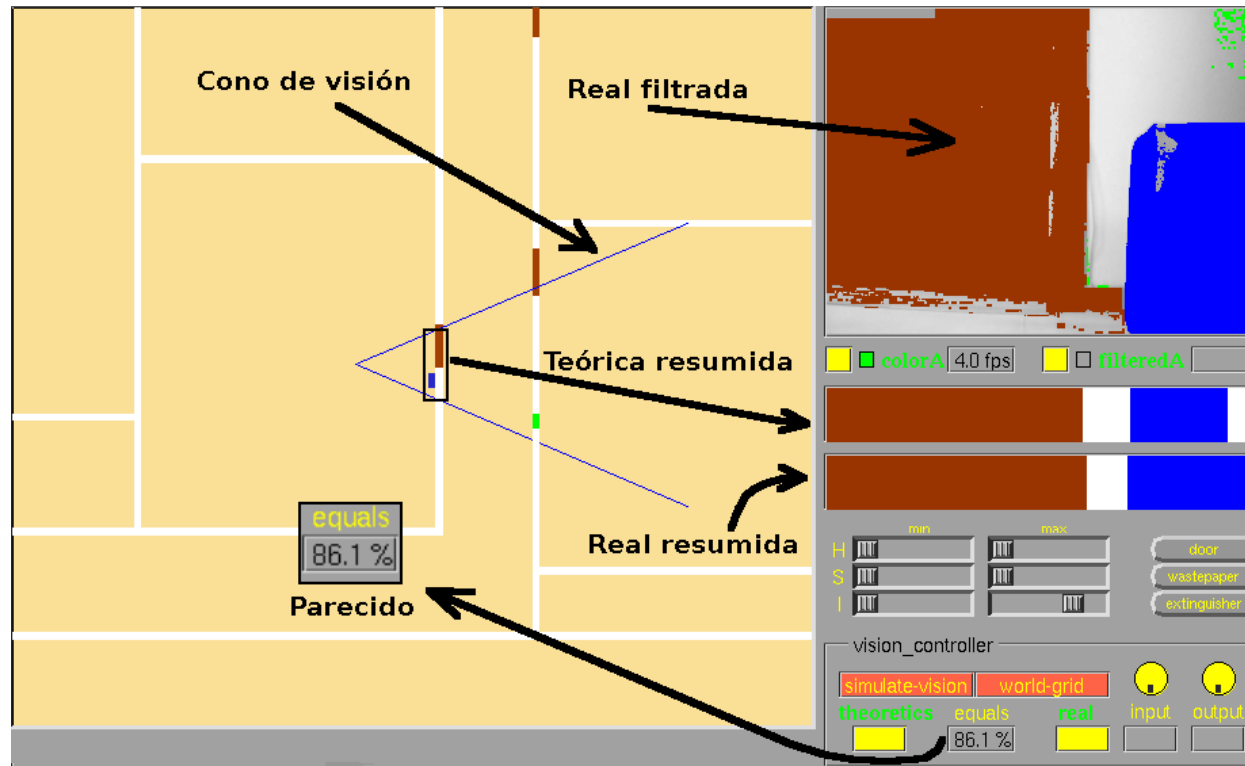
¿CÓMO?



TAMBIÉN EN EL PERRITO



Localización visual



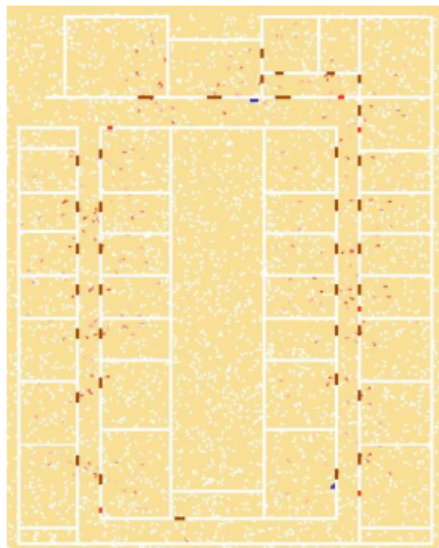
- Filtro de partículas: modelo observación, movimiento y remuestreo



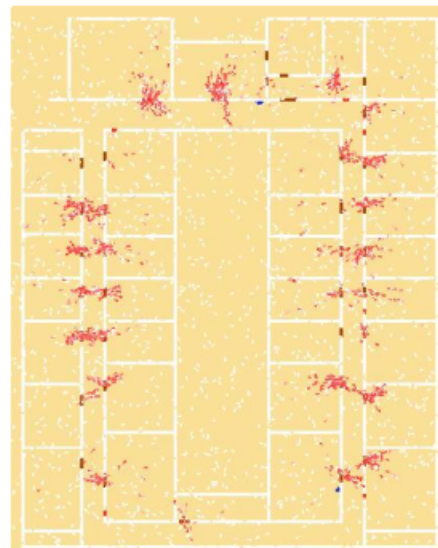
Observación 1



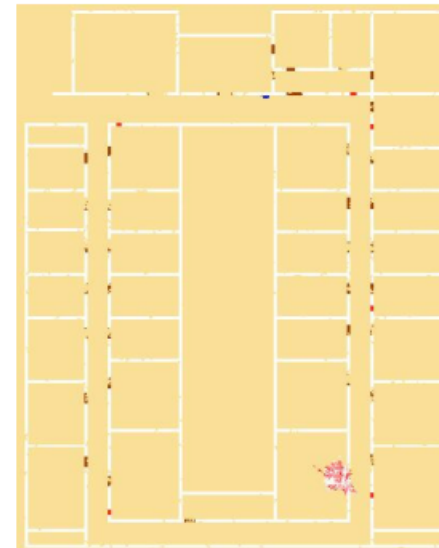
Observación 2



Distribución inicial
(uniforme)

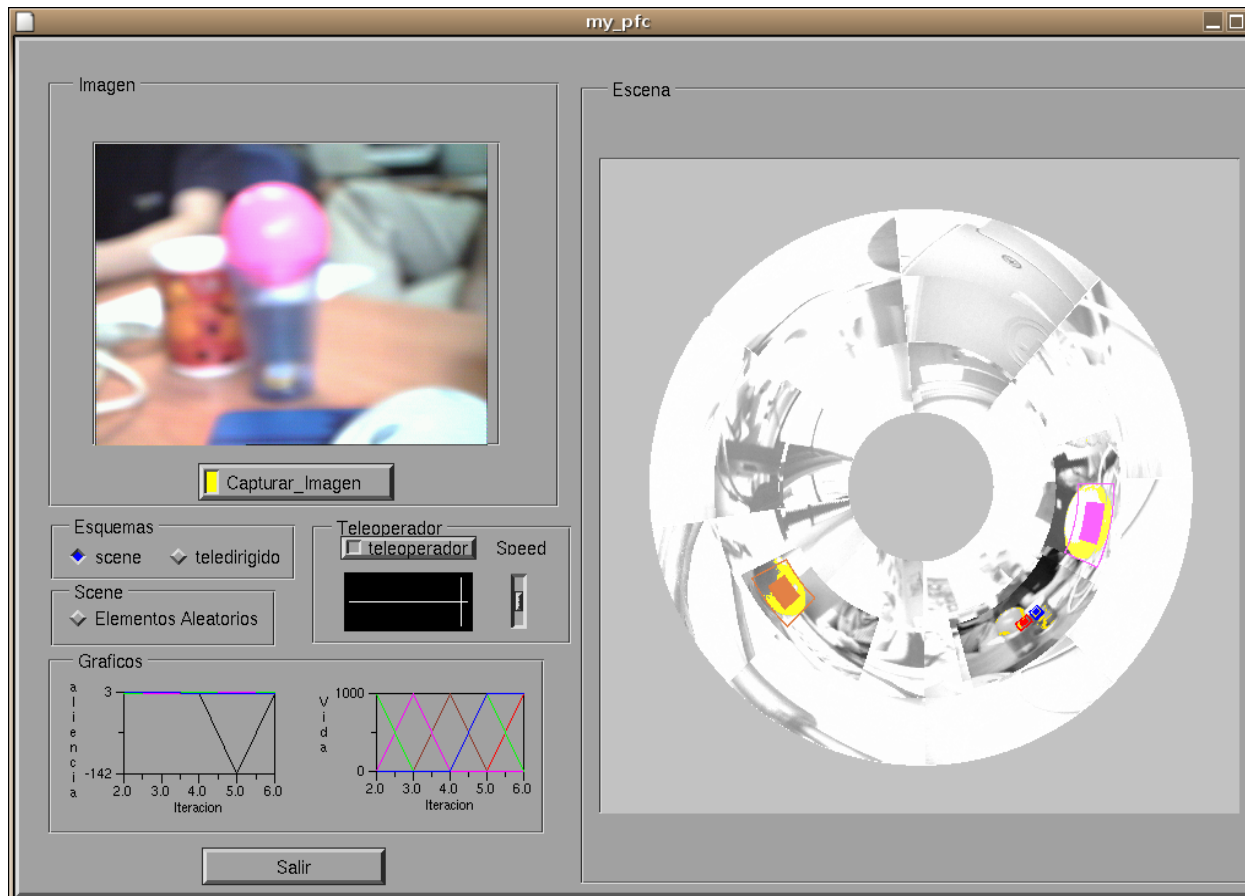


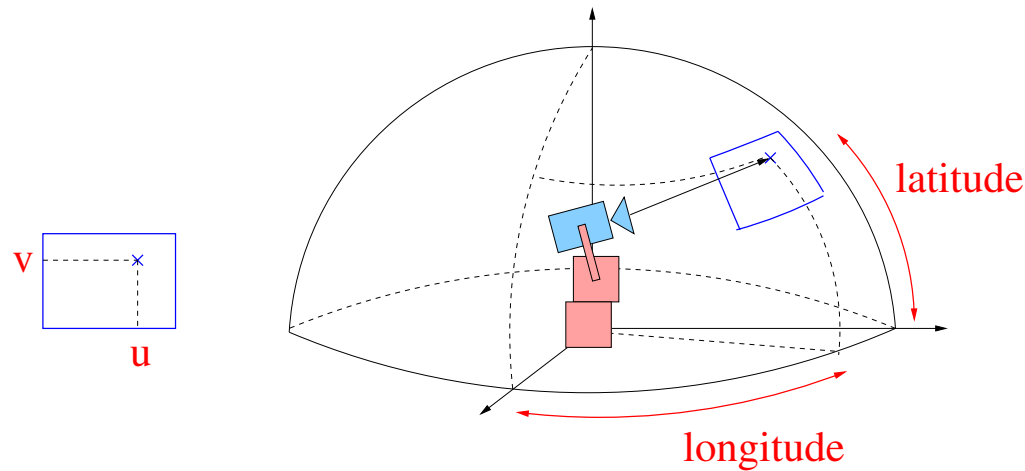
Acumulación en
zonas probables



Localización final

Atención visual



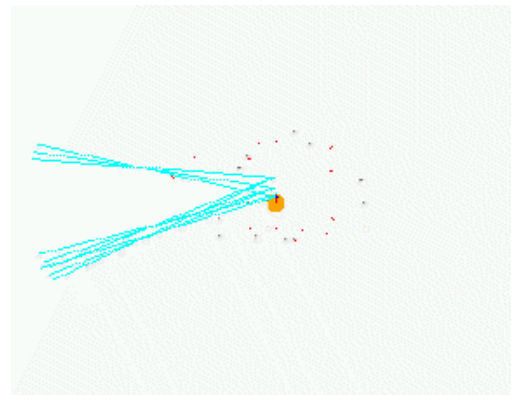
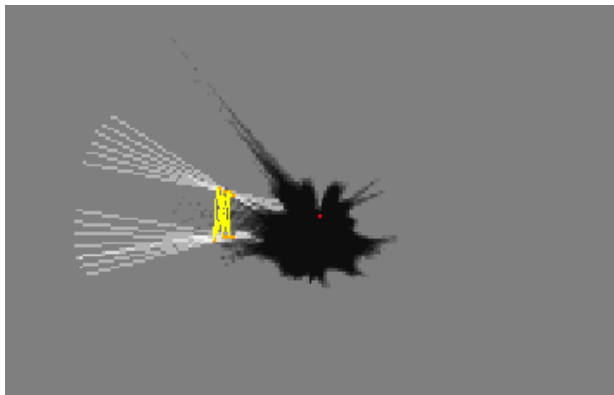


- La **imagen de escena** es más rica que la monocular
- Percepción activa: hay que mover el cuello mecánico
- Filtro para los colores relevantes
- Dinámicas temporales de saliencia y de vida
 - $liv(object, t) = liv(object, t - 1) - \Delta L_{time}$
 - $liv(object, t) = liv(object, t - 1) + \Delta L_{observation}$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

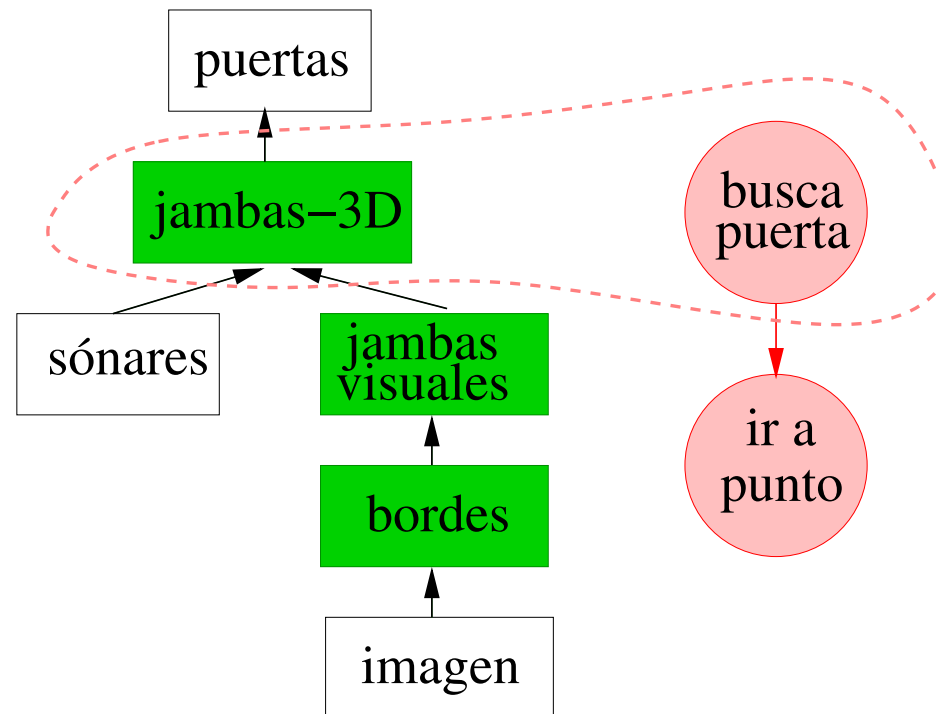
Combinación de información visual

Detección de puertas

PROFUNDIDAD FUSIONANDO EVIDENCIAS PROBABILÍSTICAMENTE



PERCEPCIÓN ESTRUCTURADA

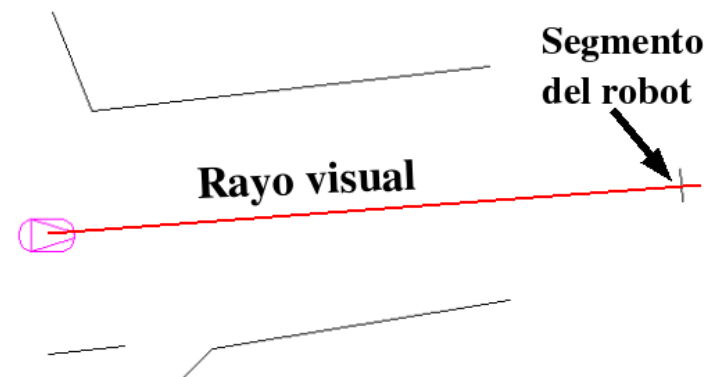
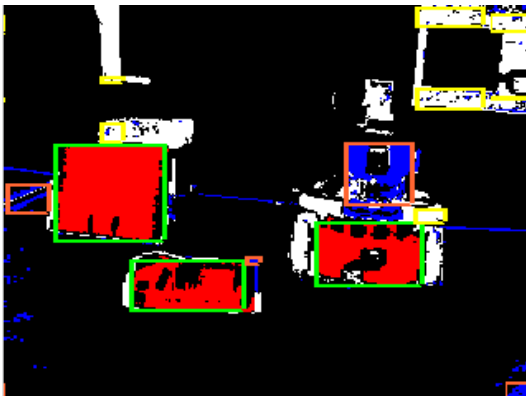
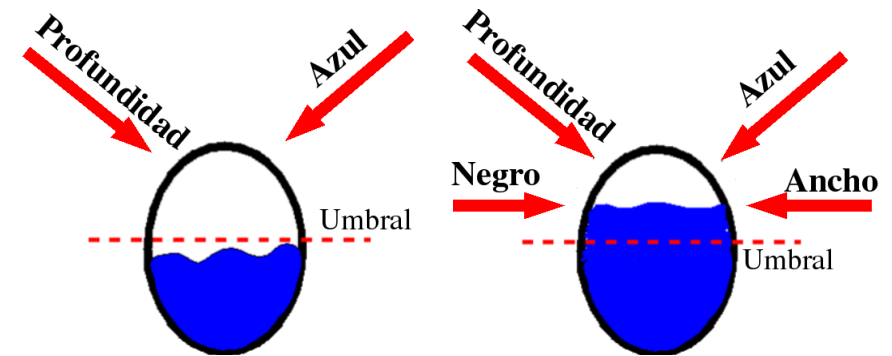


- Jerarquía de estímulos
- Interpretación y contexto de percepción
- Percepción activa: el robot se mueve para tomar más imágenes

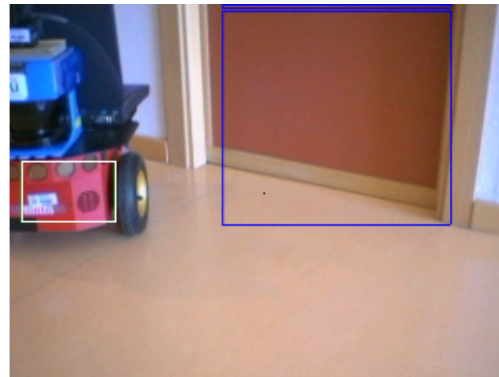
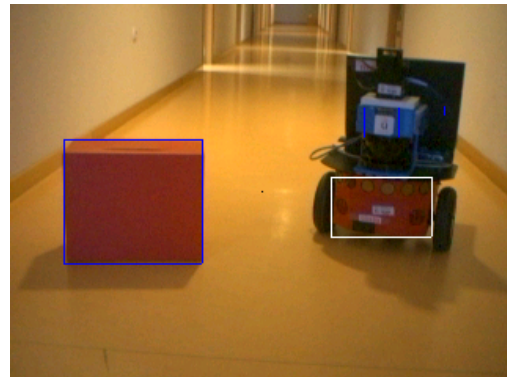
Reconocimiento de congénere

SUMA HETEROGÉNEA DE SUBESTÍMULOS

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

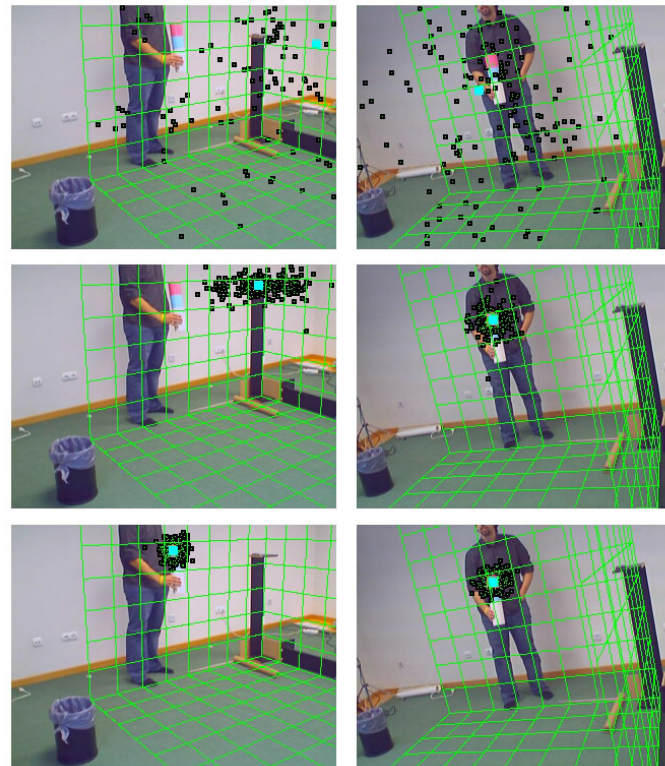


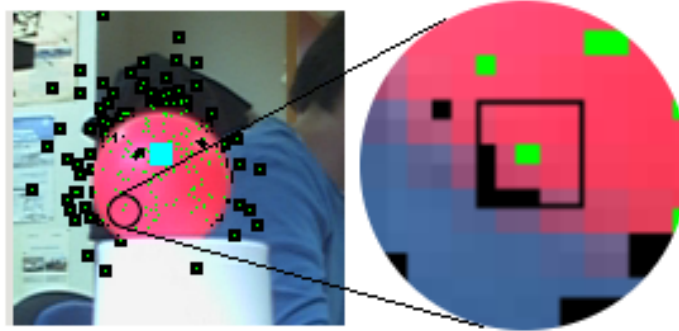
ES MUY ROBUSTO Y DISCRIMINANTE



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

Watcher JDE

Camera A Camera B Virtual Camera X:14777.3 Y:-2093.2 Z:4610.7 x:0.0 y:4425.0 z:-25.0 Camera Select

real cams simulated cams
 virtual camera on

Color Filter and Image Selectors

colorA 9.0 fps
 colorB 9.0 fps
 colorC 9.0 fps
 colorD 8.0 fps
 hsi filter
 motion filter

Camera C Camera D Cameras Control

A B C D All VC

Room Grid OptAxis Axis

black ambient flip image
Hue Tolerance 10.00 Saturation Tolerance 0.10
Hm Tolerance 30.00 Motion Tolerance -0.00
Hmax Hmin Smax Smin

Flies Filter Control Particles Filter Control Simulated Object Control Main Options

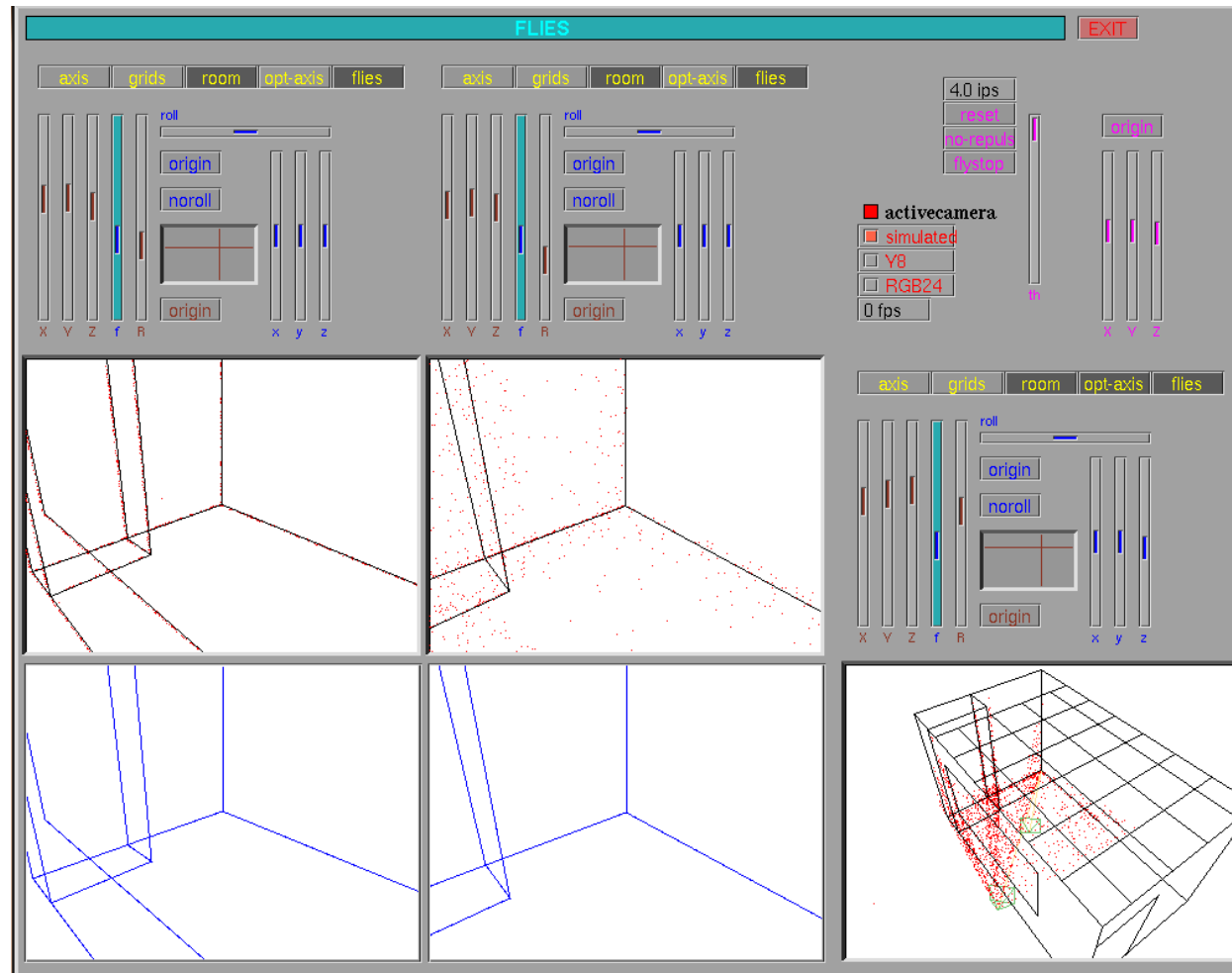
Flies Filter Control: Activate Reset color motion
View Elitism 30 Abductive 60 Thermalnoise 5 Random Mutation 5
NO OBJECT DETECTED NEAR TARGET
files fps Estimated: X:0.0 Y:0.0 Z:0.0 Diff: X:0.0 Y:0.0 Z:0.0

Particles Filter Control: Activate Reset color motion
View Std 100 Mean 0
WARNING! OBJECT DETECTED NEAR TARGET
particles fps Estimated: X:3707.5 Y:2363.6 Z:961.7 Diff: X:0.0 Y:0.0 Z:0.0

Simulated Object Control: 3d cube blue red pink
 target pink
Animation Center Lock
x y z Joystick Sp 0.05
3d cube - X:3701.2 Y:2363.6 Z:961.7
target - X:4788.0 Y:2674.0 Z:1000.0

Main Options: Simulated Real
watcher fps 26.0 fps
EXIT

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