
Memoria visual atenta para navegación autónoma

Julio Vega, Jose M^a Cañas
julio.vega@urjc.es, jmplaza@gsyc.es



Grupo de Robótica, 6 Septiembre 2011

Contenidos

1. Introducción
2. Memoria visual
3. Sistema de atención visual
4. Experimentos
5. Conclusiones

1. Introducción

- **Sistemas de visión**, de los sensores más usado en robótica autónoma
- La dificultad está en **extraer información** útil de las imágenes
- En **robots autónomos** es imprescindible un mecanismo de control de atención

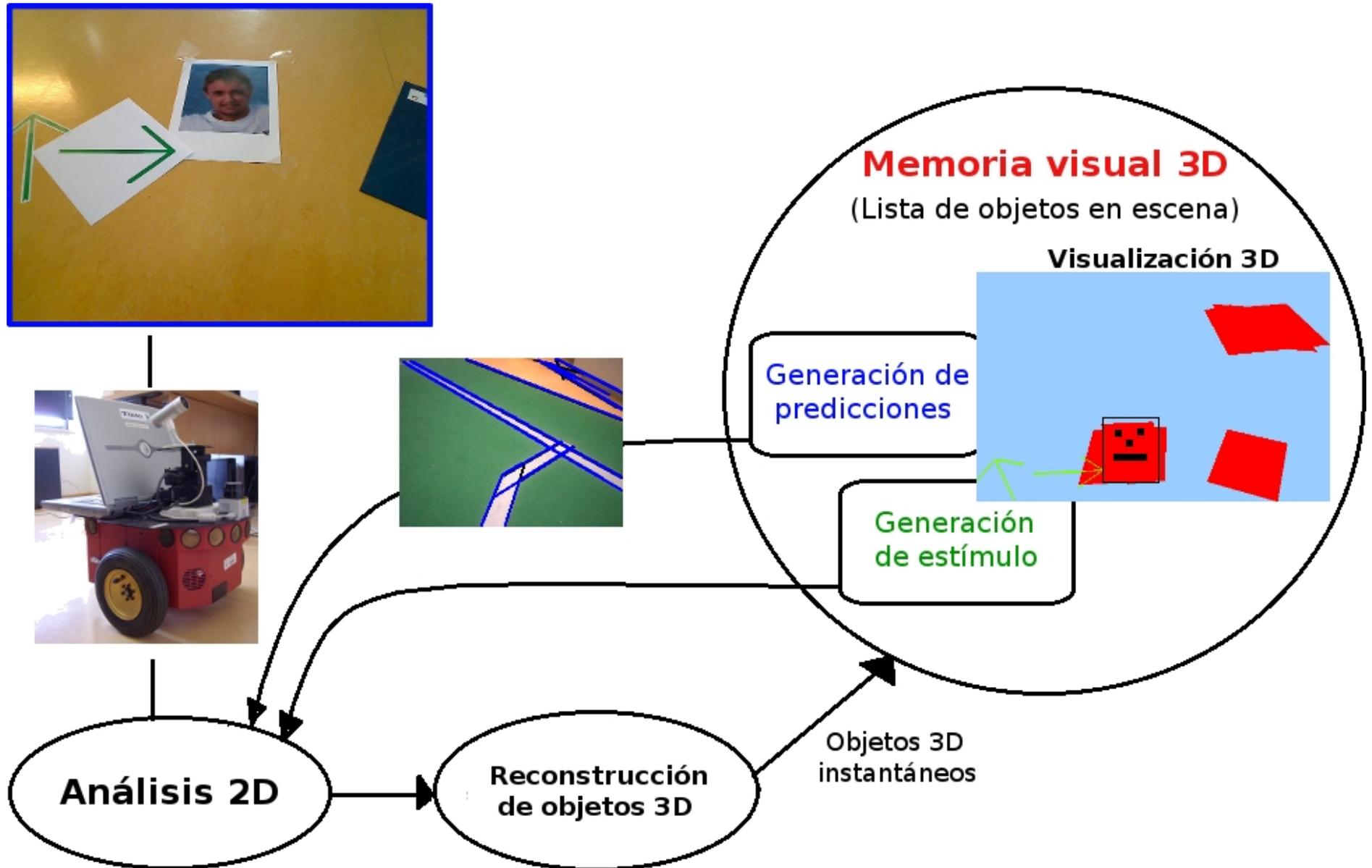
Atención visual en robots

- **Fijación** en uno o varios aspectos interesantes de la realidad y prescindir de los restantes
- Las cámaras de los robots proveen de un **amplio flujo de datos**
- Dos **vertientes** de atención:
 - Local** Seleccionar *dentro de una imagen* lo que interesa
 - Global** Seleccionar *del entorno* lo que interesa y dirigir la mirada hacia ellos

2. Memoria visual

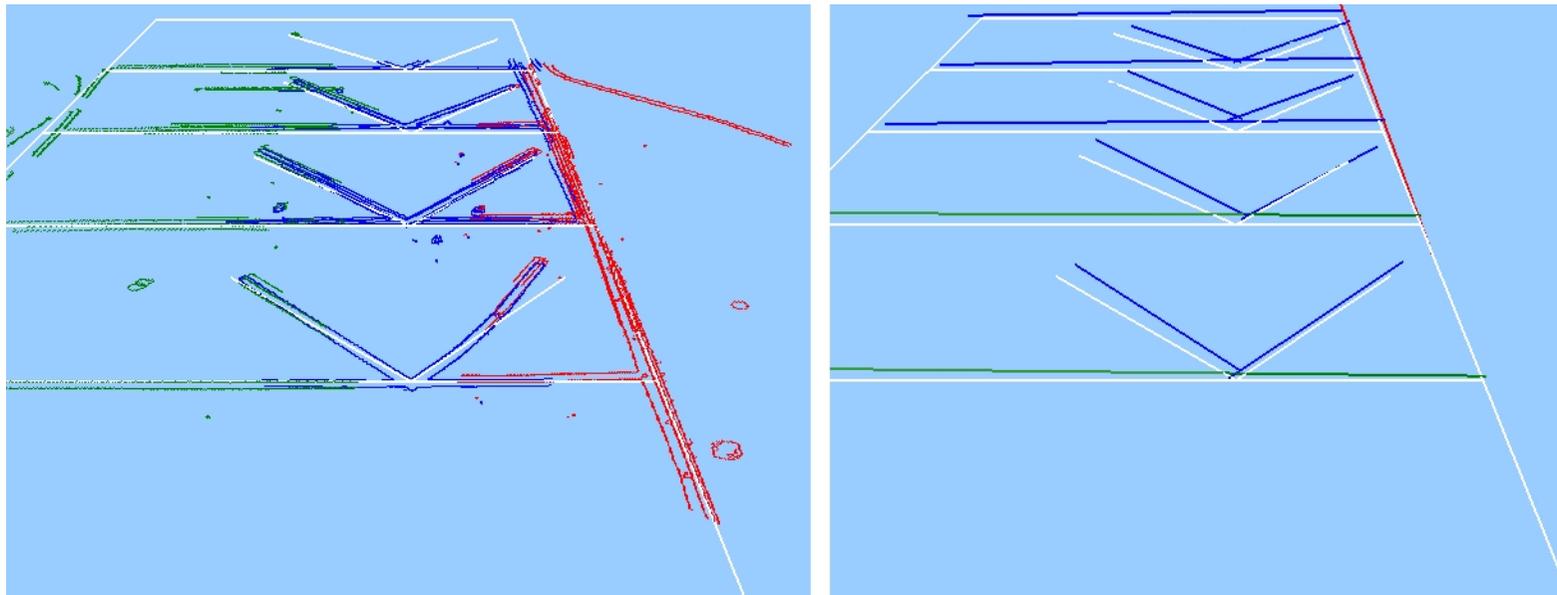
Objetivo: realizar un seguimiento visual de los objetos en escena

- *Análisis 2D:* detectar segmentos 2D en imagen actual
- *Reconstrucción 3D:* ubicar segmentos en espacio 3D (hip. suelo)
- *Memoria 3D:* almacenar posición 3D, generar hipótesis y predicción

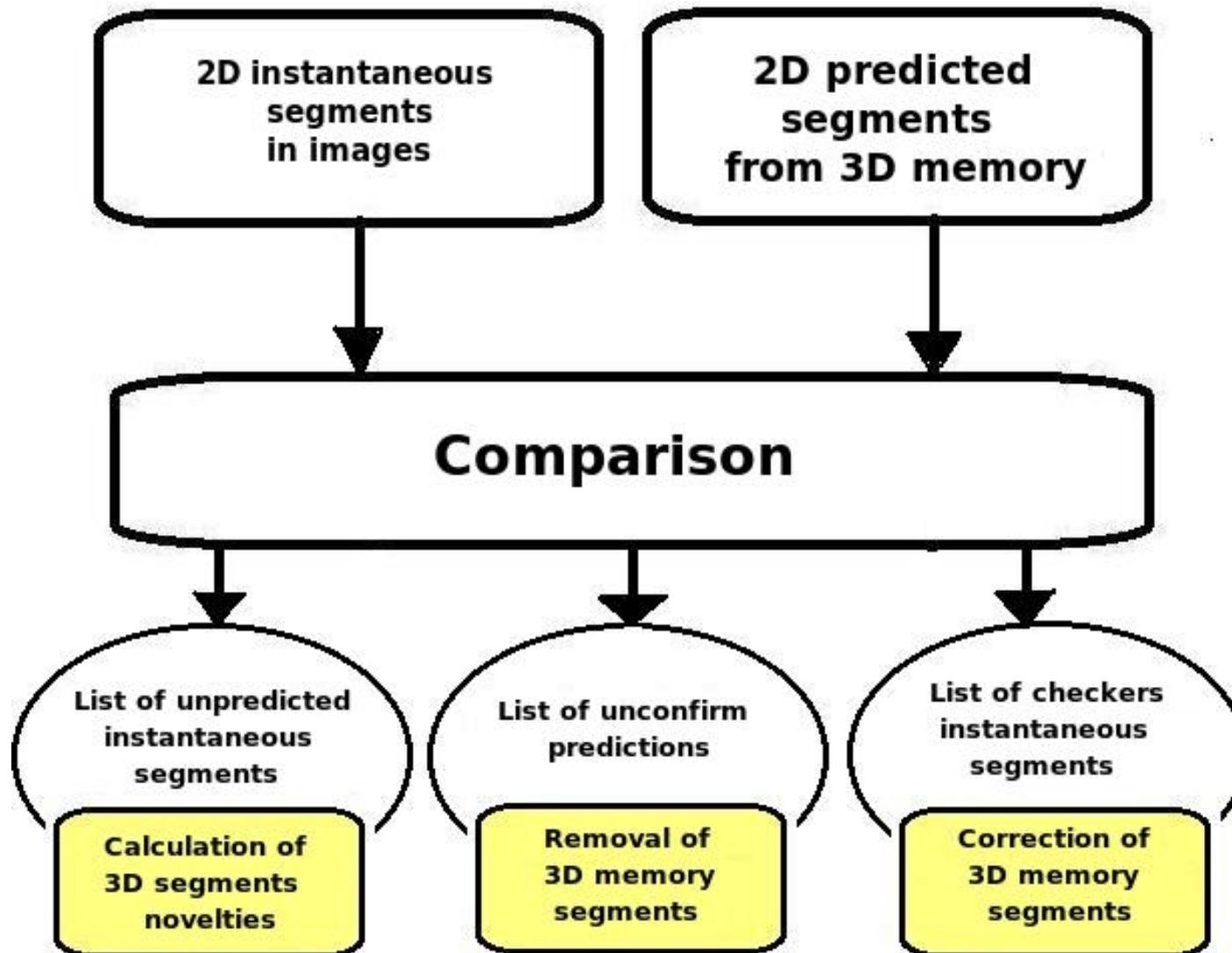


1. Procesamiento de imagen 2D

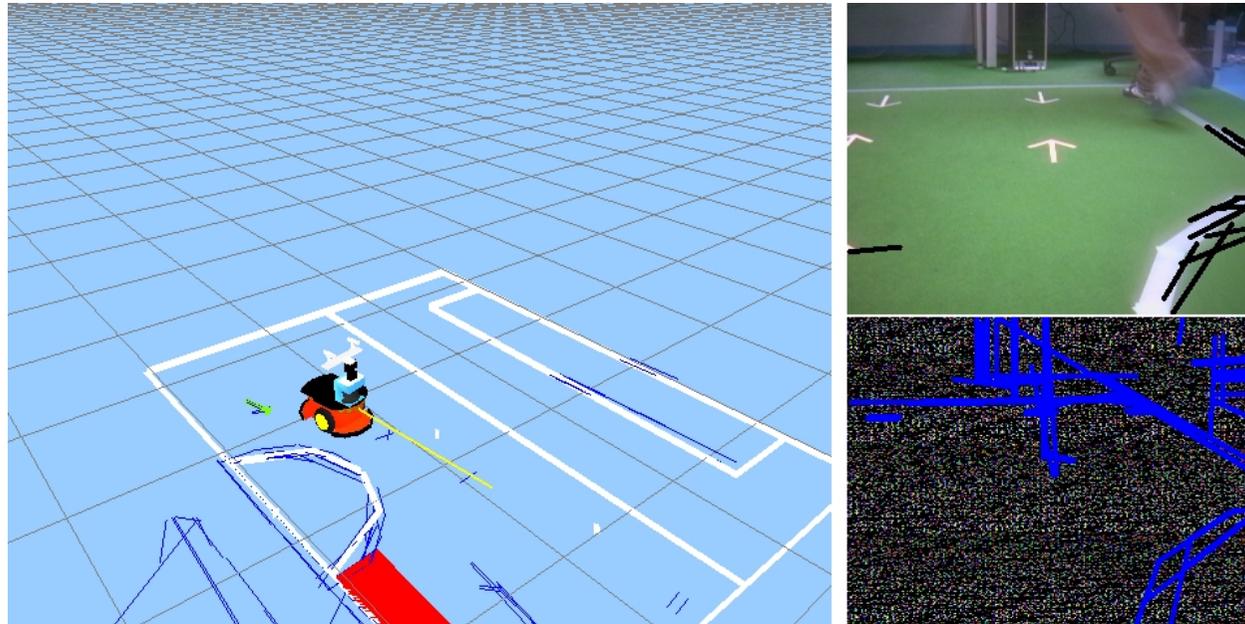
Objetivo: extraer segmentos rectos 2D como primitiva básica



- Confirmar/refutar segmentos almacenados previamente
- Fases: filtro de bordes y Transformada de Hough.
- Biblioteca *OpenCV*



2. Reconstrucción instantánea con segmentos 3D



- **Hipótesis suelo**: para situar los segmentos 2D en el espacio 3D
- Antes de inserción, post-procesamiento para eliminar duplicados
- **Modelo geométrico**: coordenadas respecto al robot

3. Inserción de segmentos 3D

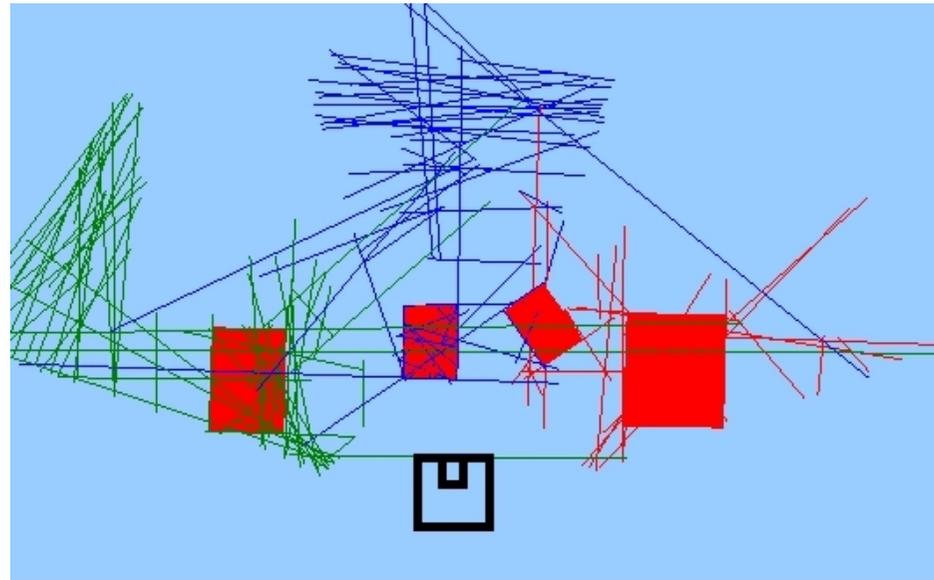
- Estructuras *Segment3D* posiblemente relacionadas entre ellas
- **Comparación** entre los segmentos actuales y los ya almacenados
- En caso de orientación similar tomamos:
 - La longitud del mayor de ellos (evitar miopía)
 - La orientación del más reciente (consistencia con la realidad)

4. Predicciones: borrado y corrección

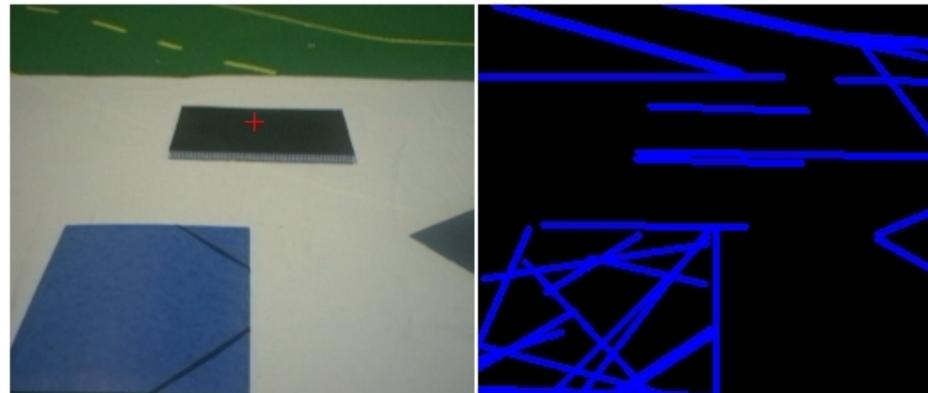
- **Nuevo segmento**: detectado actualmente pero no en predicción
- Puede reemplazar o corregir a uno existente (según condiciones)
- **Incertidumbre**: parámetro que aumenta con el paso del tiempo

5. Generación de hipótesis perceptivas

- Los segmentos pueden pertenecer a **estructuras más complejas**
- **Modelo de objeto**: para cuando no se ve la estructura completa
Ej.: con tres vértices podemos estimar un paralelogramo
- Comparar vértices y ángulos entre segmentos (robustez a oclusiones)
- Abstracción de **flechas**: útiles para la navegación



(a)

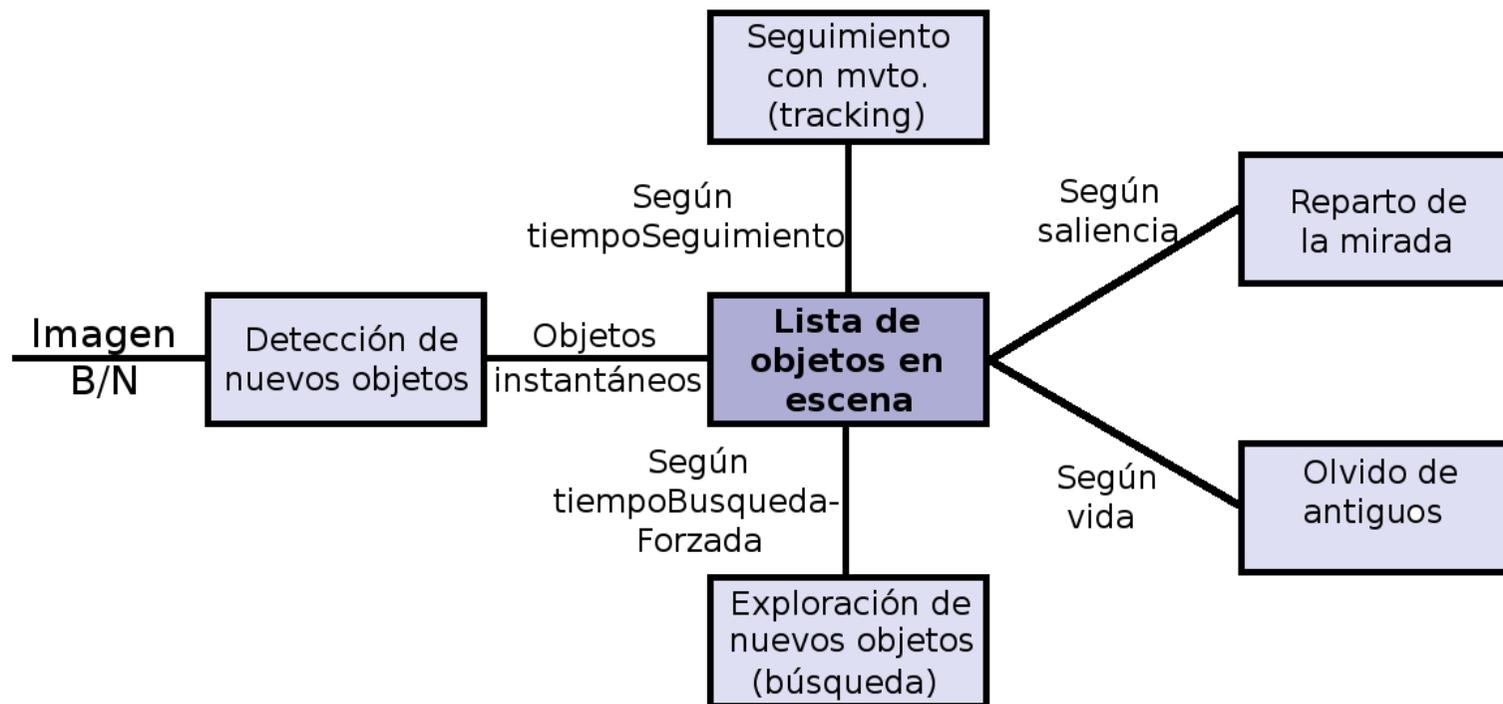


(b)

(c)

3. Sistema de atención visual

- Ampliado el **campo de visión** de la cámara
- El cuello mecánico sobre el que está permite moverla a voluntad
- **Transformación de coordenadas** de imagen a coordenadas en escena
- Y de coordenadas en la escena a **comandos** en el cuello mecánico



1. Reparto de la mirada. Saliencia

- Movimiento del cuello mecánico para **dirigir** la atención
- Con varios objetos hay que decidir a **dónde mirar** después
- Cada objeto tiene asociado un foco de atención y una saliencia
- El **punto de atención** es la posición del objeto en escena

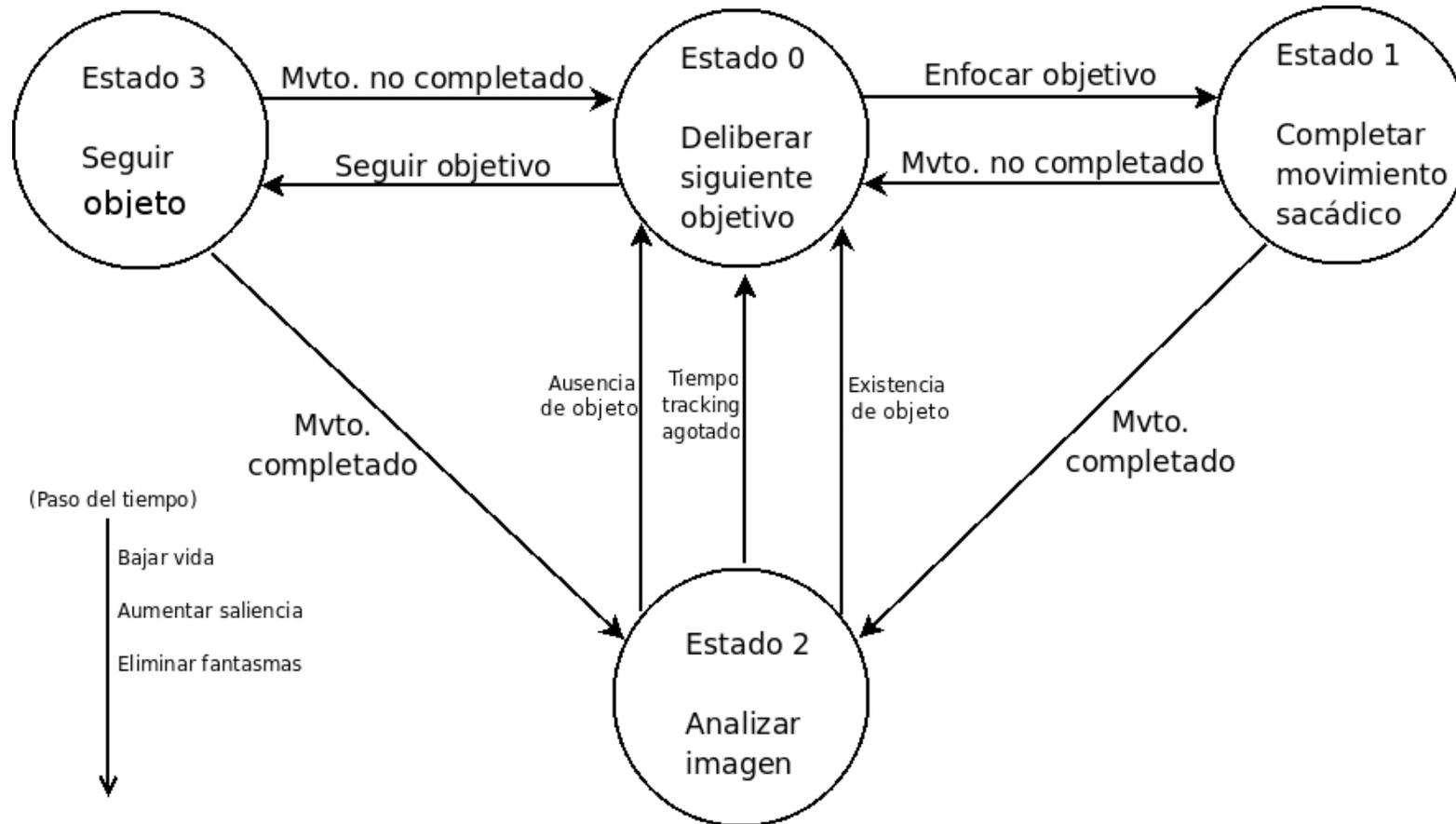
Saliencia

- Es todo aquello que *llama la atención* según el instante
- Valor que *aumenta* con el paso del tiempo

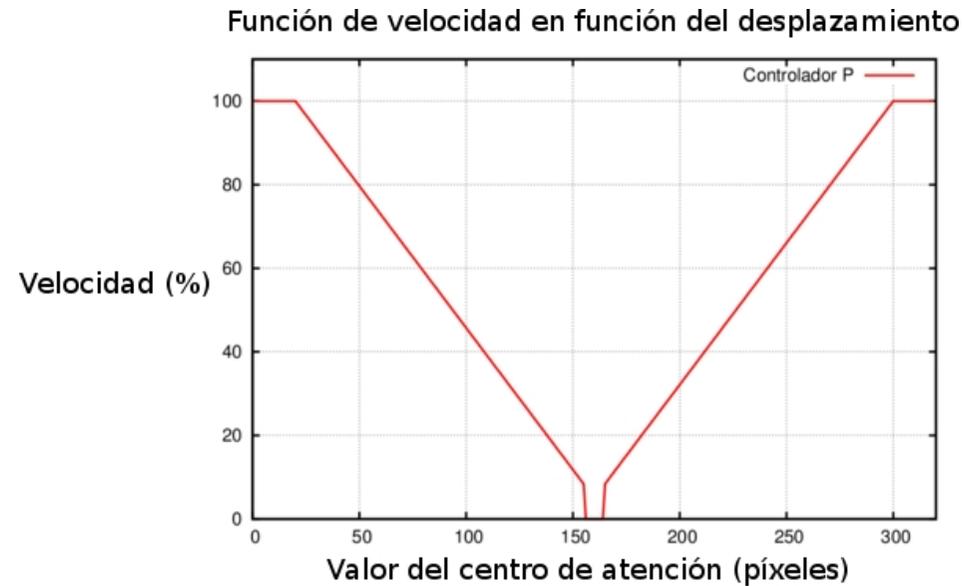
El sistema dirige la atención hacia el objeto con **máxima saliencia**

- Se anula cada vez que se visita tal objeto (**inhibición de retorno**)

2. Seguimiento con movimiento



- Cuando el sistema elige un objeto lo *sigue* espacialmente un tiempo



- Controlador P para conseguir un movimiento suave del cuello mecánico
- Mantiene el objetivo en el centro de la imagen

3. Exploración de nuevos objetos

- En cualquier momento puede interesar la **búsqueda** de nuevos objetos
- Inserta en memoria local **puntos de exploración** con alta saliencia
- Dos tipos: aleatorios y de recorrido
- Son visitados rápidamente para ver si existen objetos de interés
 - Si es así, entran en memoria para el reparto de mirada
- Tiempo entre búsquedas *proporcional* al nº de objetos en memoria

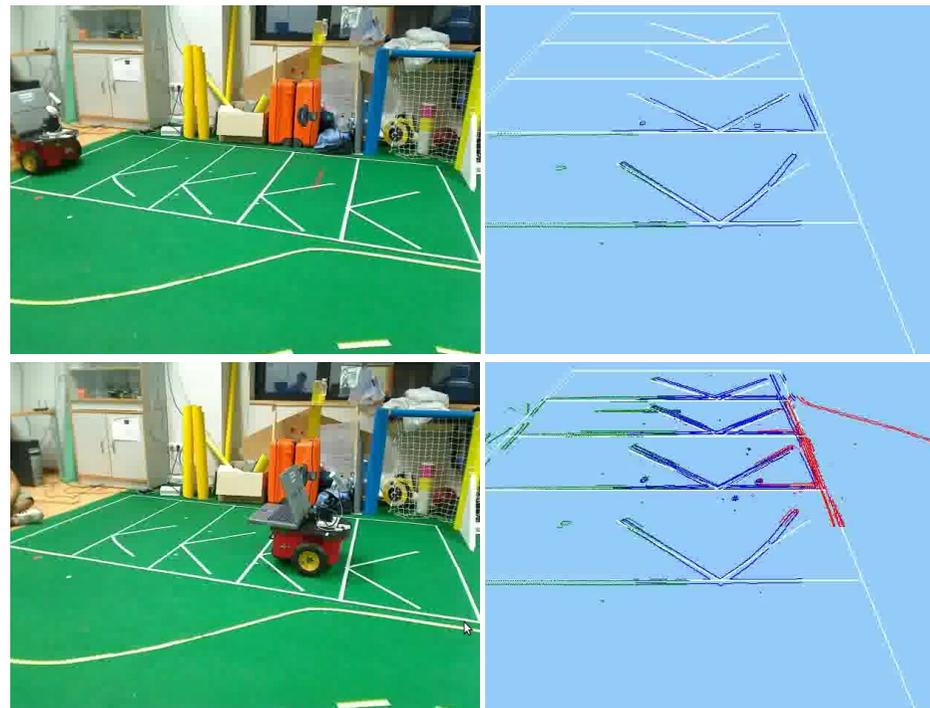
4. Representación interna del entorno. Vida

- Los objetos pueden **desaparecer** de la escena
- El sistema ha de *olvidar* antiguos objetos que ya no están
- **Vida**: dinámica inversa a la saliencia
- Cada vez que se visita un objeto su vida se incrementa un poco
 - Un objeto frecuentemente visitado tiene mayor vida
 - Límite máximo para evitar saturación
- Si la vida de un objeto es inferior a un umbral es que ha *desaparecido*

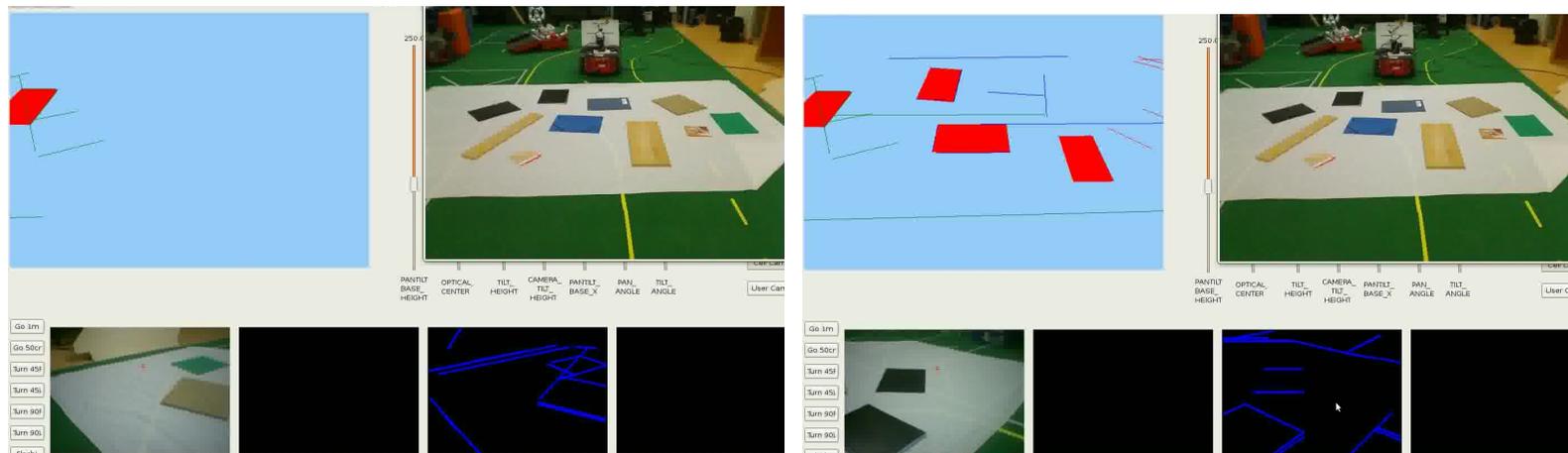
4. Experimentos

- Robot real Pioneer 2DX con unidad pan-tilt y cámara iSight

1. Reconstrucción del suelo



2. Abstracción de paralelogramos



3. Abstracción de flechas. Navegación



5. Conclusiones

- Sistema **atención visual global** busca y sigue **objetos**
 - Campo visual de la escena mayor que el instantáneo
 - Reparte la mirada si hay varios objetos de interés
 - Representación coherente **olvidando** los que ya no están
- Mecanismo de dinámica concurrente entre **vida** y **saliencia**
- **Memoria local** de corto plazo como campo de visión de toda la escena

Trabajos futuros

- Reconstrucción 3D de todo el entorno. Uso de dos cámaras
- **Obstáculos** de navegación también como focos de atención