

# Redes bayesianas temporales para reconocimiento de escenarios

Ahmed Ziani and Cina Motamed

Visión de Alto Nivel  
Dr. Enrique Sucar



Irvin Hussein López Nava  
Junio 2009

# Introducción (1)

Objetivo: aplicaciones de video-vigilancia

Los escenarios dinámicos constan de eventos:

- espaciales,
- temporales e
- interacción.

Reconocimiento automático:

- On-line: generar alarmas.
- Off-line: recuperación de comportamientos.



# Introducción (2)

Condiciones favorables de esta aplicación:

- los parámetros de la **cámara** y
- el **fondo** observado permanecen fijos.

Etapas de un sistema de video-vigilancia:

1. Detección de movimiento
2. Etapa de seguimiento
3. Interpretación de movimiento de alto nivel

# Introducción (3)

Los principales objetivos son:

- observar la escena,
- indexar las actividades y
- reconocer los escenarios modelados.

Los escenarios contienen

- trayectorias específicas espacio-temporal,
- interacción entre objetos estáticos y no estáticos, y
- una combinación de ambos.

# Detección de movimiento (1)

$$\begin{aligned} \text{if } & \left( \max_{c=R,G,B} |I^k(P) - R^{k-1}(P)| > \omega \right) \\ & \rightarrow D^k(P) = 1 \\ \text{else } & \rightarrow D^k(P) = 0 \end{aligned}$$

$D^k$  representa la decisión de detección:

- 1 = objeto en movimiento,
- 0 = fondo

Procedimientos de limpieza:

- Morfológicas: erosión y dilatación.
- Eliminación de regiones de poco interés.

# Detección de movimiento (2)



# Seguimiento

Parea las regiones detectadas de una **secuencia temporal** con otra teniendo en cuenta la *fusión y división* durante el movimiento de objetos.

Este algoritmo cuenta con un mecanismo de detección de confianza (creencia) que incrementa cuando:

- El objeto sigue un camino continuo.
- El tamaño del objeto permanece estable.
- No hay ambigüedad.

# Reconocimiento de Escenario

Un escenario está compuesto de un **conjunto de eventos** elementales relacionados.

Los “escenarios de actividad humana” presentan variabilidad.

Para la estructura lógica de estos eventos:

- Petri-net
- Redes bayesianas
- Modelos ocultos de Markov

# Redes bayesianas (1)

Es un grafo acíclico dirigido donde:

- Los nodos representan variables y
- Los arcos representan relaciones de dependencia.

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{parents}(X_i))$$

Principales ventajas de las RB:

- La capacidad de representar conocimiento
- Eficiente estructura de inferencia para aplicaciones en TR.

# Redes bayesianas (2)

La probabilidad conjunta está representada por:

$$P(X_j, H_i) = P(X_j | H_i) P(H_i)$$

El objetivo es estimar hipótesis  $H_i$  ( $i=1, \dots, n$ ) basados en la evidencia  $X_j$

$$P(H_i | X_j) = \frac{P(X_j, H_i)}{P(X_j)}$$

# RB y Reconocimiento visual de Escenario

- Descripciones textuales en una dinámica de vehículos y peatones.
- Reconocer actividades en un partido de futbol americano.
- RBN's han sido utilizadas para reconocimiento de comportamiento humano.

## Categorías de Modelado:

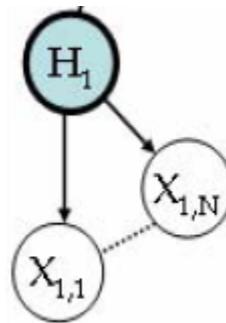
- Enfoque “*time slice*”. DBN.
- Enfoque “basado en eventos”. TNBN, NIEDT

# Modelo Propuesto (1)

- Capa de eventos atemporales

Cada **evento atemporal** está representado por una red bayesiana jerárquica.

1. Nivel bajo =  $X_j$ , características visuales: posiciones, velocidad, dirección y tamaño de los objetos en seguimiento.
2. Nivel alto =  $H_i$ , decisión de reconocimiento de eventos.



# Modelo Propuesto (2)

- Capa de razonamiento temporal

Conjunto de redes bayesianas que contienen nodos asociados con **información temporal**.

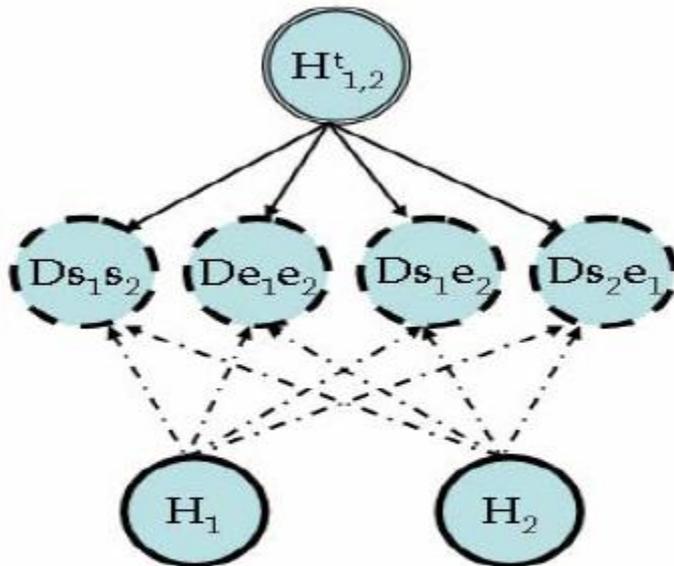
Evaluar cuantitativa y cualitativamente limitaciones temporales para cada evento y también estimar las relaciones entre eventos relacionados.

Nivel bajo = tiempo de “inicio”(s) y “fin”(e) de los eventos

Los datos son normalizados.

# Modelo Propuesto (3)

Se implementan las relaciones Allen para evaluar las relaciones temporales.



**Fig. 2.** Model of temporal relations

**Table 1.** Example of the relation 'Equal' of allen

$H_1$ equal $H_2$
$D_{s_1 s_2}$ true if $(s_2 - s_1 = 0)$
$D_{e_1 e_2}$ true if $(e_2 - e_1 = 0)$
$D_{s_1 e_2}$ true if $(e_2 - s_1 > 0)$
$D_{s_2 e_1}$ true if $(e_1 - s_2 > 0)$

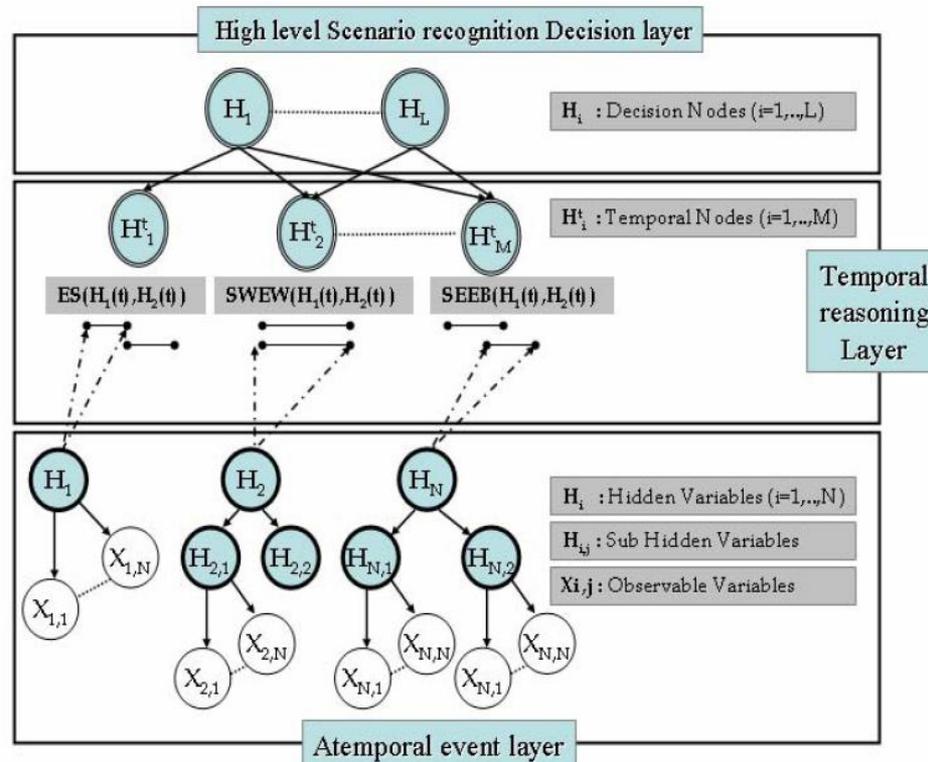
# Modelo Propuesto (4)

**Table 2.** Various situations of the relationships of allen

Relations	$Ds_1s_2$	$De_1e_2$	$Ds_1e_2$	$Ds_2e_1$
H <sub>1</sub> starts H <sub>2</sub>	$Ds_1s_2=0$	$De_1e_2<0$	$Ds_1e_2>0$	$Ds_2e_1>0$
H <sub>1</sub> finishes H <sub>2</sub>	$Ds_1s_2>0$	$De_1e_2=0$	$Ds_1e_2>0$	$Ds_2e_1>0$
H <sub>1</sub> during H <sub>2</sub>	$Ds_1s_2>0$	$De_1e_2<0$	$Ds_1e_2>0$	$Ds_2e_1>0$
H <sub>1</sub> equal H <sub>2</sub>	$Ds_1s_2=0$	$De_1e_2=0$	$Ds_1e_2>0$	$Ds_2e_1>0$
H <sub>1</sub> meets H <sub>2</sub> H <sub>1</sub> before H <sub>2</sub>	$Ds_1s_2>0$	$De_1e_2>0$	$Ds_1e_2>0$	$Ds_2e_1=0$
H <sub>1</sub> overlaps H <sub>2</sub>	$Ds_1s_2>0$	$De_1e_2>0$	$Ds_1e_2>0$	$Ds_2e_1>0$

# Modelo Propuesto (5)

Estructura general del modelo de escenario.



# Modelo Propuesto (6)

Modelo de escenario de una maleta abandonada.

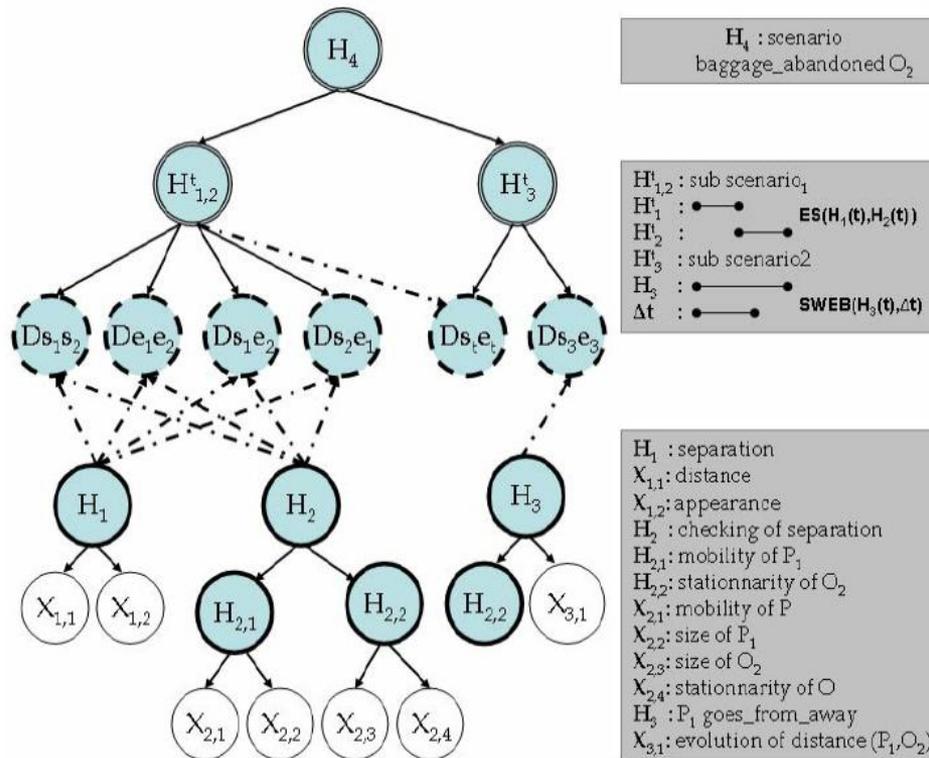


Fig. 4. Model of the scenario for an abandoned baggage

# Aprendizaje de Parámetros (1)

Los parámetros de la red se pueden aprender a partir de los datos experimentales.

$$P(X_i = x_k | \text{parent}(X_i) = c_j) = \theta_{i,j,k} = \frac{n_{i,j,k}}{\sum_K n_{i,j,k}}$$

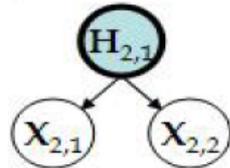
Problemas en la vigilancia visual:

- No se tienen suficientes ocurrencias para cada escenario.
- Se tienen entradas con incertidumbre.

Se puede utilizar el algoritmo *expectation-maximization*.

# Aprendizaje de Parámetros (2)

$H_{2,1}$  : mobility of  $P_1$



$X_{2,1}$  : mobility of  $P$

$X_{2,2}$  : size of  $P_1$

Table 3. Table of occurrence

$X_{2,1}$	$X_{2,2}$	$H_{2,1}$	Nb
1	1	1	18
1	1	0	1
0	1	0	3
1	0	0	2
?	1	1	1
1	?	1	2
0	0	0	14
0	?	0	3

Fig. 5. Learning of a naive bayesian network, an example

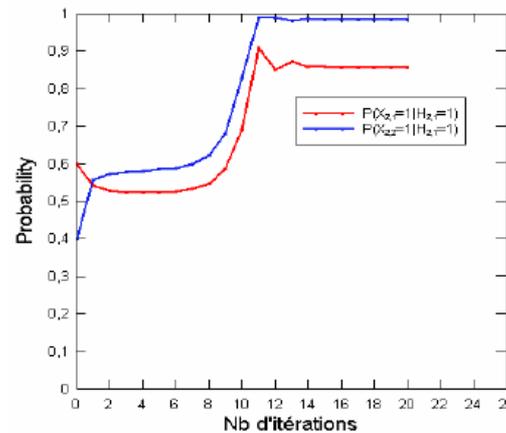
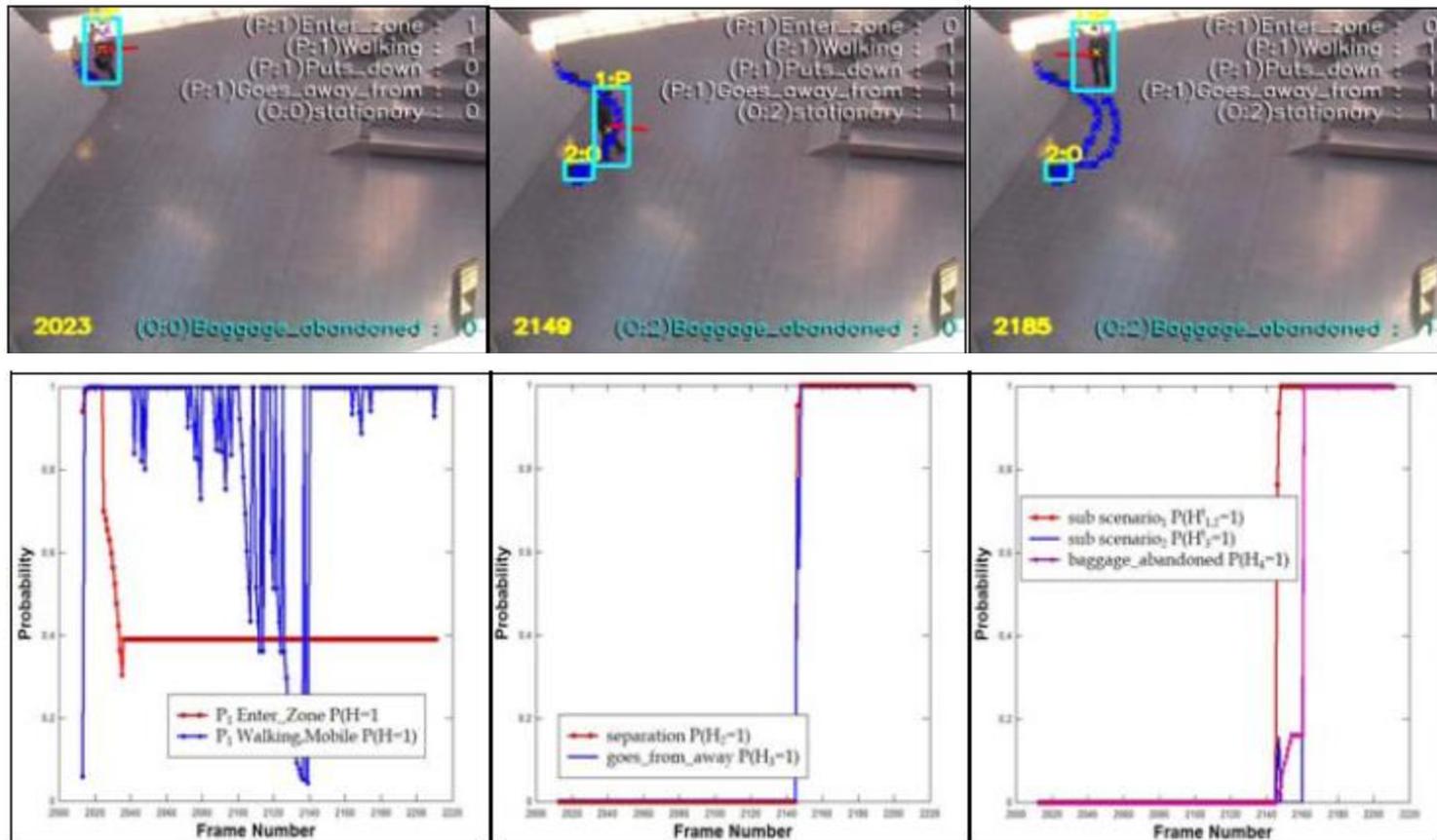


Fig. 6. Evolution of the estimated probabilities

# Resultados (1)



# Resultados (2)

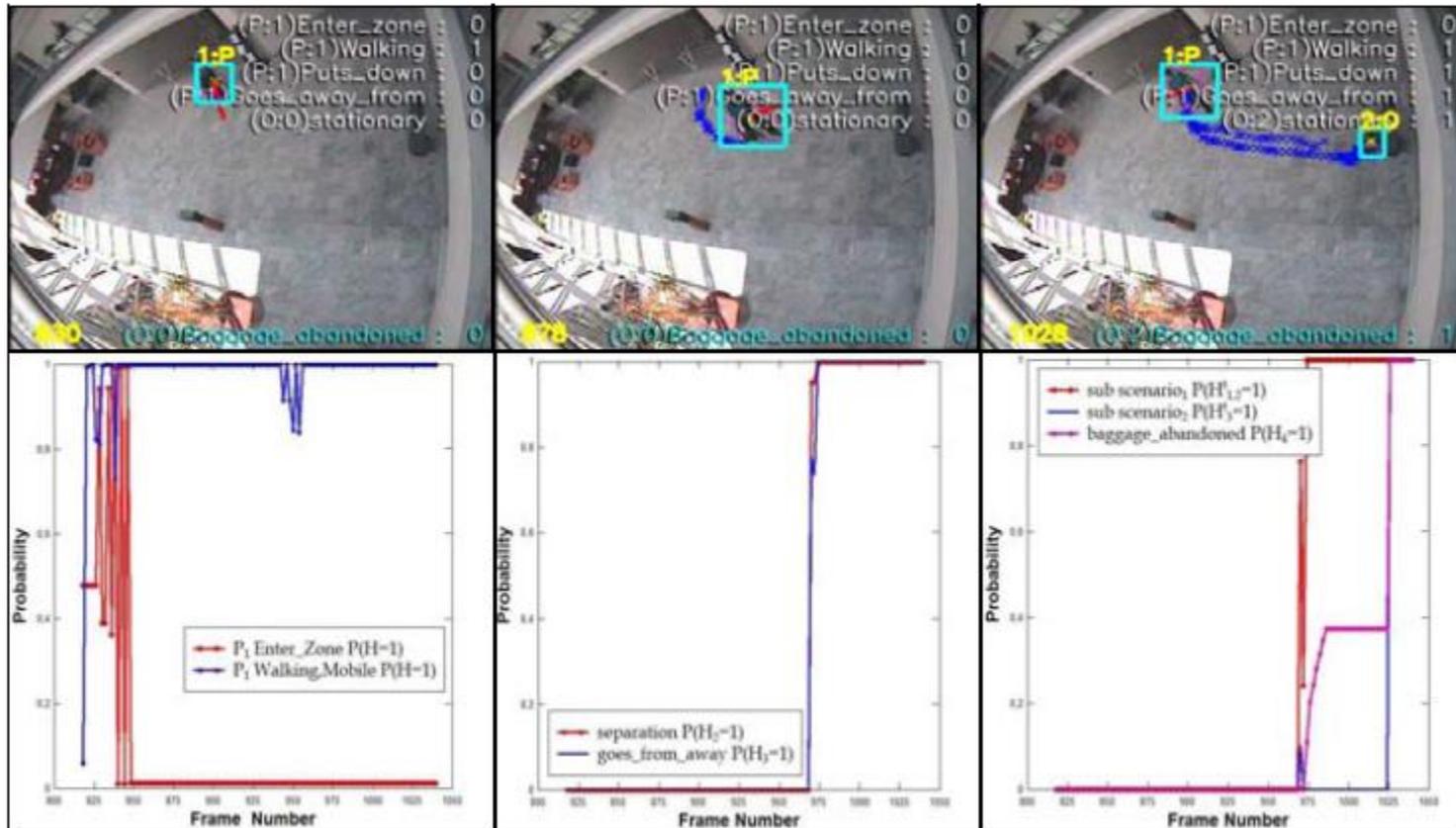


Fig. 8. Sequences from PETS'04 Workshop

# Conclusiones

## A FAVOR

- La aplicación propuesta es muy interesante, sobre todo la organización de la estructura

## EN CONTRA

- No detallan las pruebas de su aplicación.
- Los resultados están muy ligados al ambiente.