

Apprentissage sur des données multivariées pour la modélisation de comportement

Florence DUCHÊNE

Projet ORION – INRIA Sophia-Antipolis

Introduction

Systeme d'apprentissage et de décision pour l'évaluation de l'état de santé et de la situation d'une personne

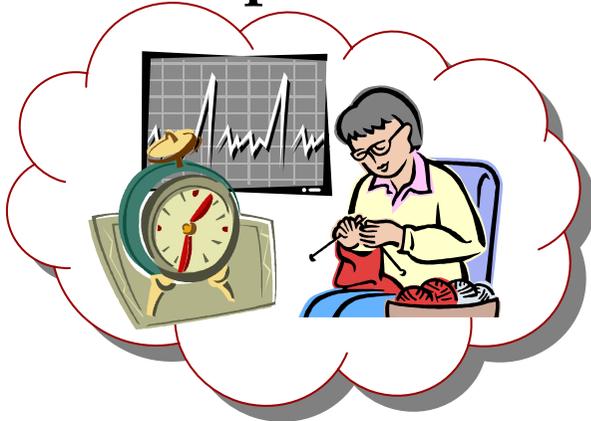
Dégradation de l'état de santé

→ Troubles du comportement à domicile

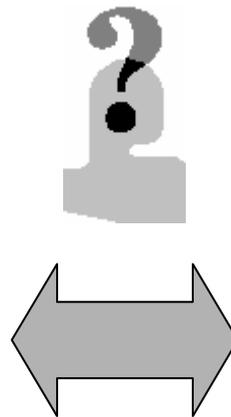


Évolution de la situation à long terme → **Étude des habitudes de vie**

Apprentissage d'un profil comportemental



Décision par comparaison au profil



Difficultés

Données temporelles, multicateurs, hétérogènes, complexes, imprécises, etc.

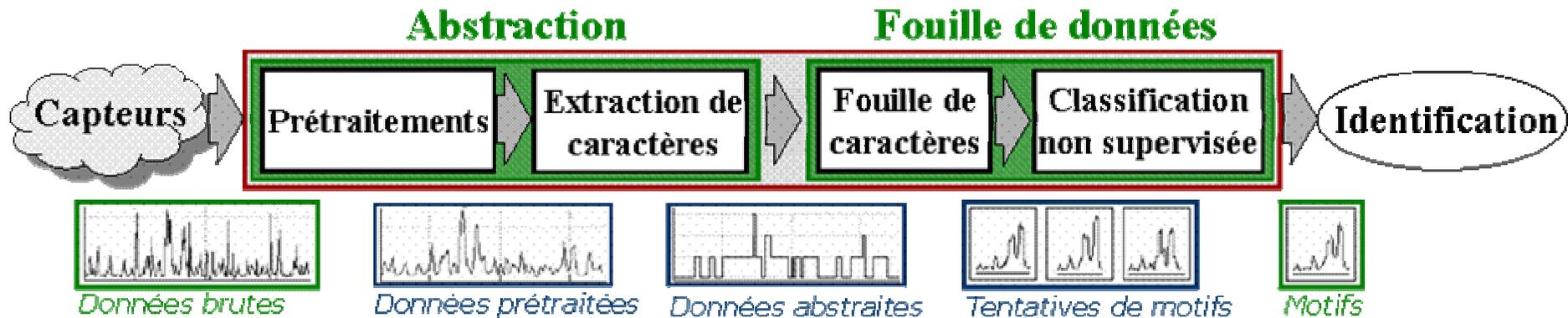
Apprentissage des habitudes de vie

Quoi?

➔ Modélisation du comportement dans les activités de la vie quotidienne

Comment?

➔ Extraction non supervisée de motifs temporels, multidimensionnels, hétérogènes



● Abstraction

- Filtrage moyen pondéré
- Discrétisation sur l'échelle des valeurs
- Agrégation dans le temps

● Fouille de données

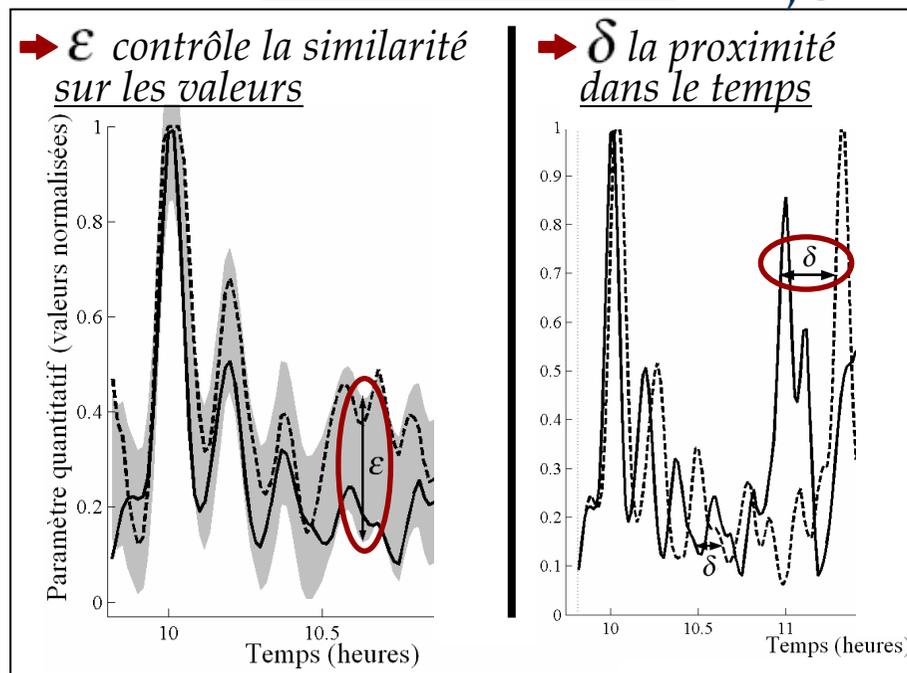
- Projections aléatoires (sous-séquences discrètes fréquentes)
- Synthèse des sous-séquences récurrentes (significatives, non redondantes)
- Classification sur la base d'une mesure de distance

Mesure de distance entre séquences multidimensionnelles et hétérogènes

Séquences bruitées : Variabilité des valeurs, interruptions, déformations temporelles.

Distance non métrique basée sur la plus longue sous-séquence commune – LCSS¹

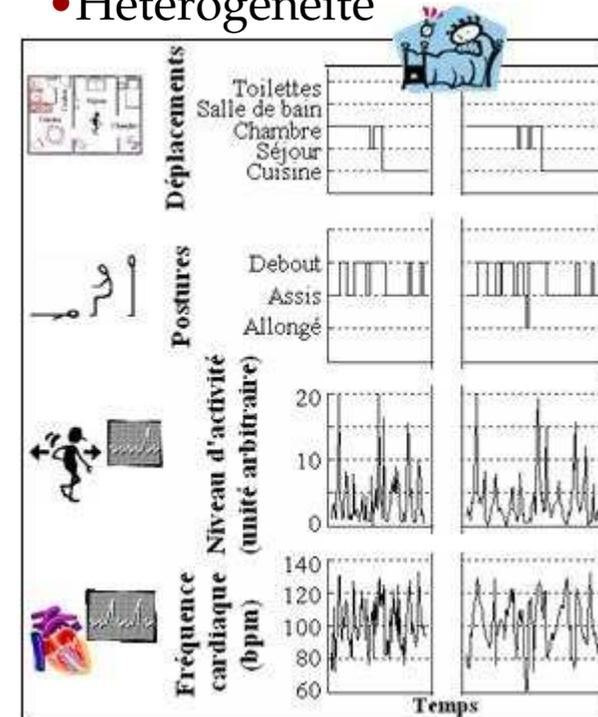
- Nombres de « points » similaires selon deux seuils de similarité ϵ, δ $LCSS_{\delta,\epsilon}(A,B)$



- Distance dans $[0,1]$: $D_{\delta,\epsilon}(A,B) = 1 - \frac{LCSS_{\delta,\epsilon}(A,B)}{\min(n,m)}$

Extension de la notion de similarité

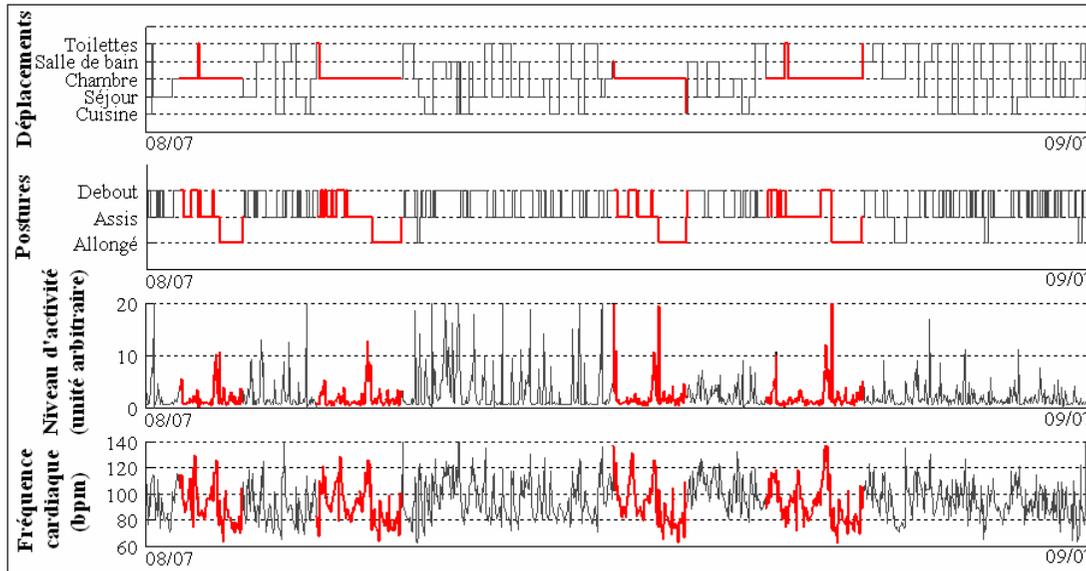
- Multidimensionnalité
- Hétérogénéité



(Données simulées)

¹M. Vlachos, G. Kollios, and G. Gunopulos, "Discovering Similar Multidimensional Trajectories," in *Proc. of the 18th ICDE*, San Jose, CA, 2002, pp. 673–684.

Expérimentation sur des données simulées

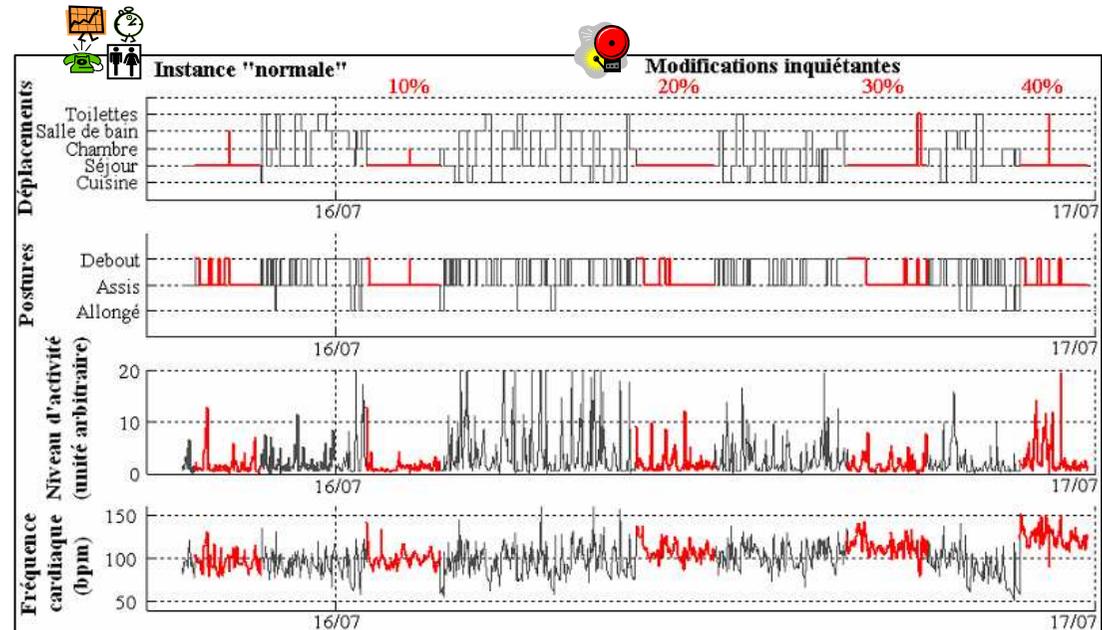


← **Sensibilité**

Indices	Identification		Classification	
	S_e	S_p	S_e	S_p
Moyenne	0.71	0.92	0.66	0.79
Écart-type	0.18	0.07	0.34	0.26
Indices parfaits	–	–	35%	60%
Classification parfaite			20%	

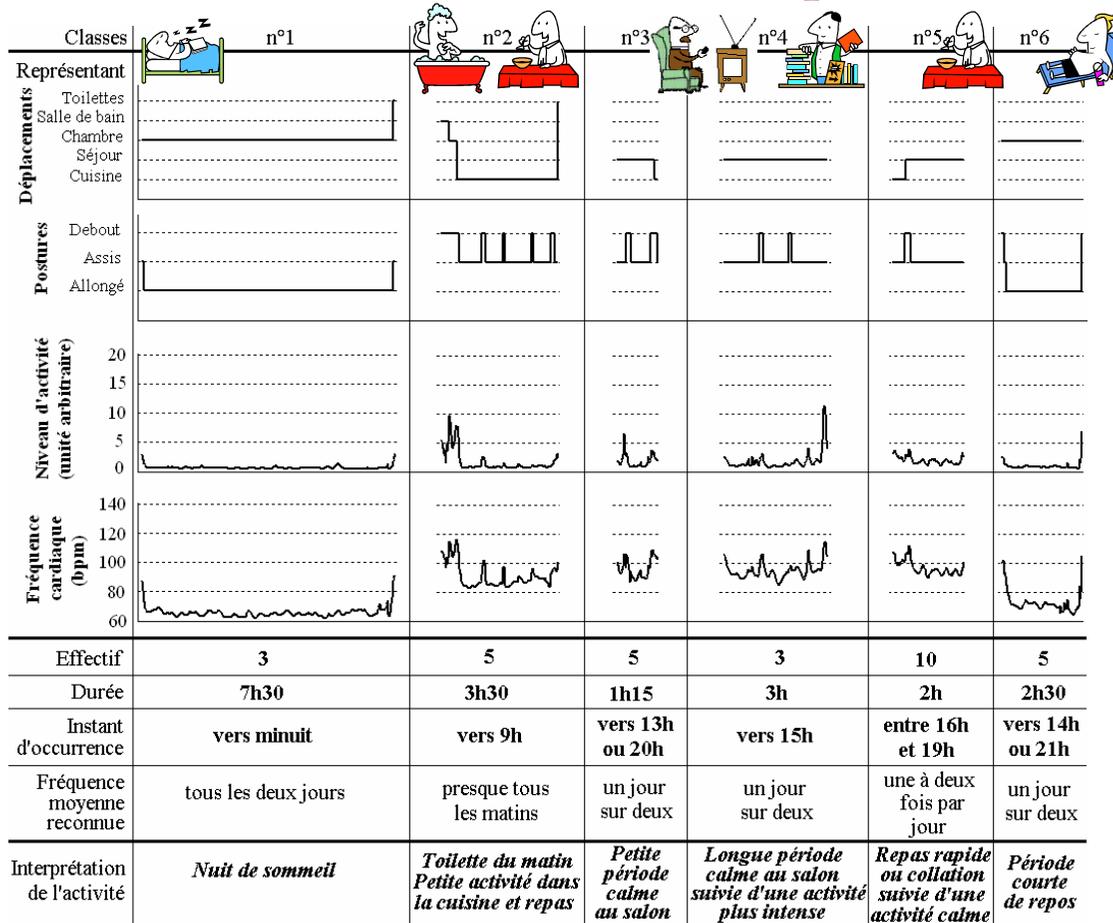
Spécificité →

Modifications	Normales	Inquiétantes			
		10%	20%	30%	40%
Taux de modification	0%	10%	20%	30%	40%
Taux de reconnaissance	77.5%	25%	10%	10%	5%

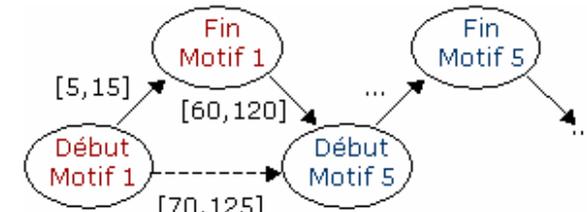


Modélisation du comportement

Extraction de motifs temporels



Construction d'un modèle de comportement



Décision par
comparaison
au profil

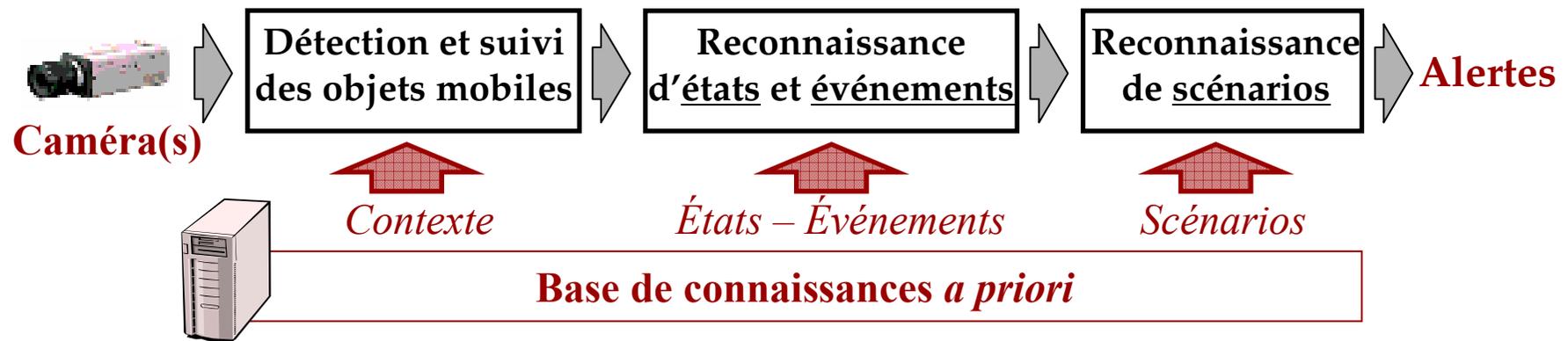
Système de décision



Apprentissage sur des données vidéos

Plateforme d'interprétation vidéos VSIP²

➔ **DÉCISION** - Reconnaissance de scénarios (fréquents/ indésirables)



➔ **APPRENTISSAGE** - Identification et caractérisation de :

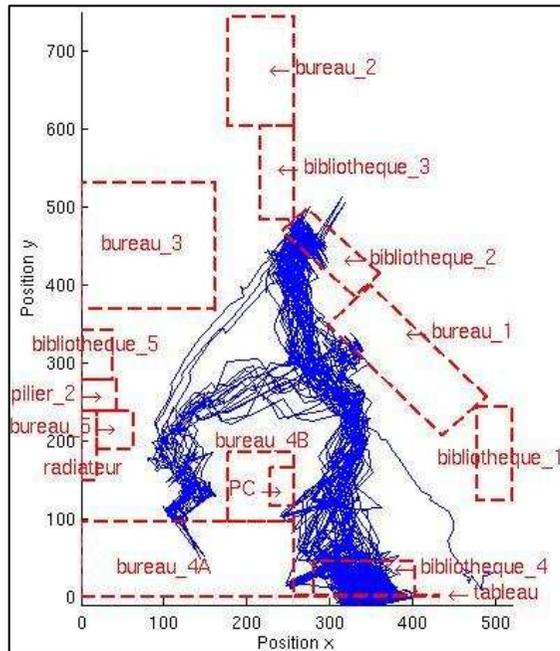
- Contexte : zones d'intérêt, objets mobiles
- États et événements
- Scénarios : **activités fréquentes**, scénarios indésirables

Apprentissage pour la surveillance vidéo

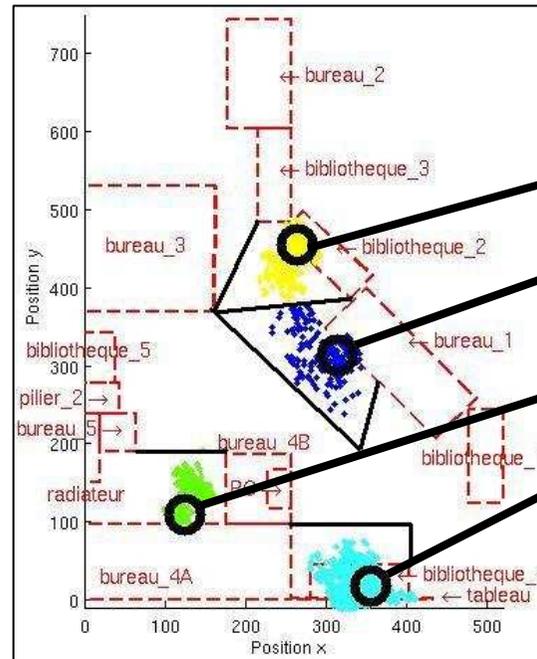
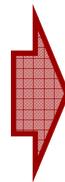
Zones géographiques d'intérêt (x,y)

➔ Partition du plan de la scène

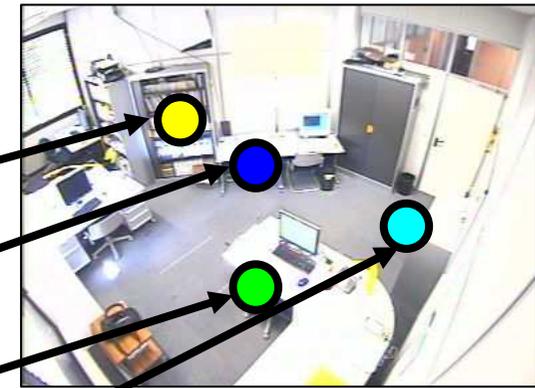
- *Clustering* sur une sélection de points : $vitesse < seuil$
- Suppression des *outliers* : $distance\ Mahalanobis < seuil$



Trajectoire dans le plan



Sélection de points occupés



Scène observée

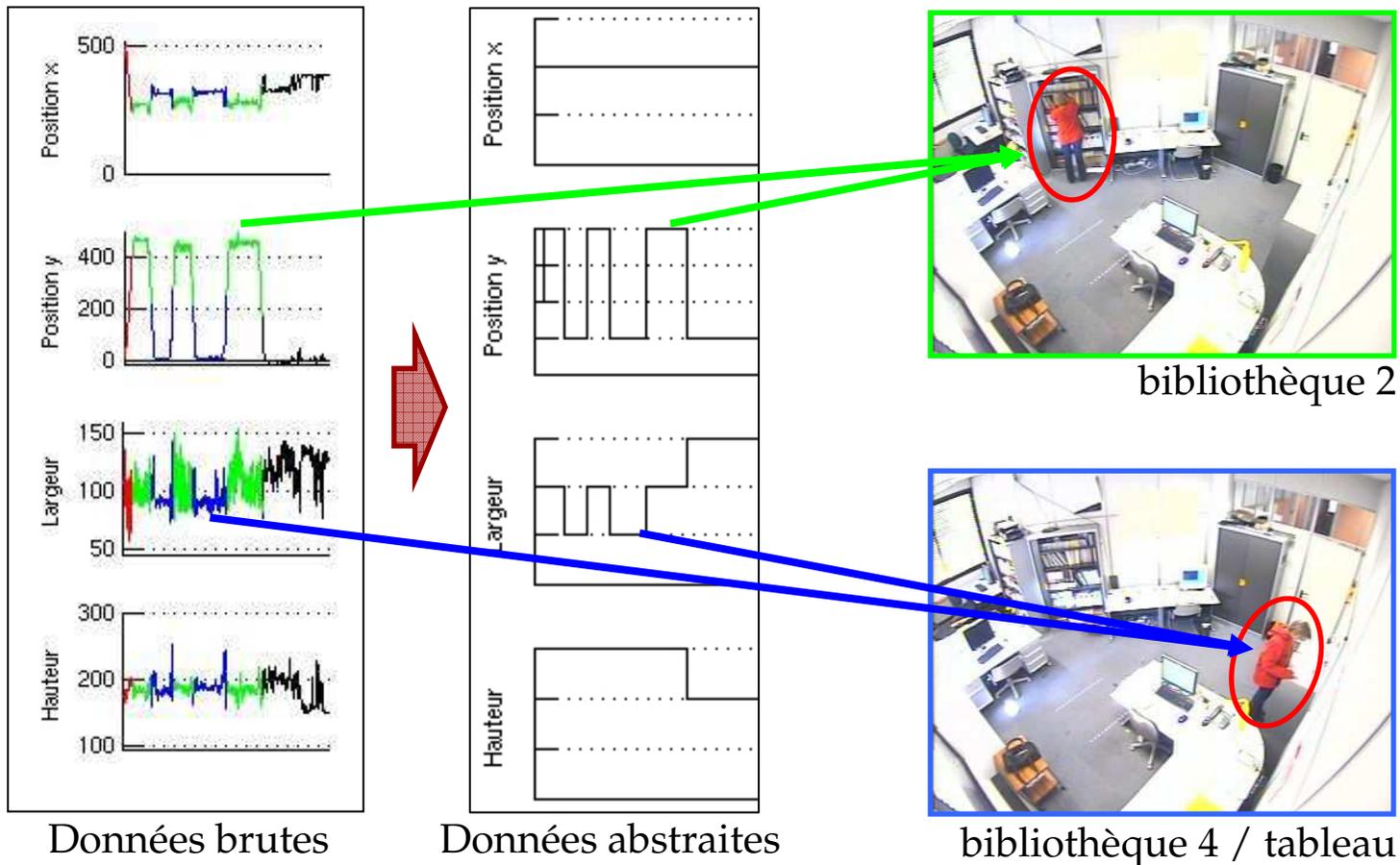
- Bibliothèque 2
- Bureau 1
- Bureau 4B
- Bibliothèque 4

Apprentissage pour la surveillance vidéo

États

➔ Symboles issus de l'abstraction

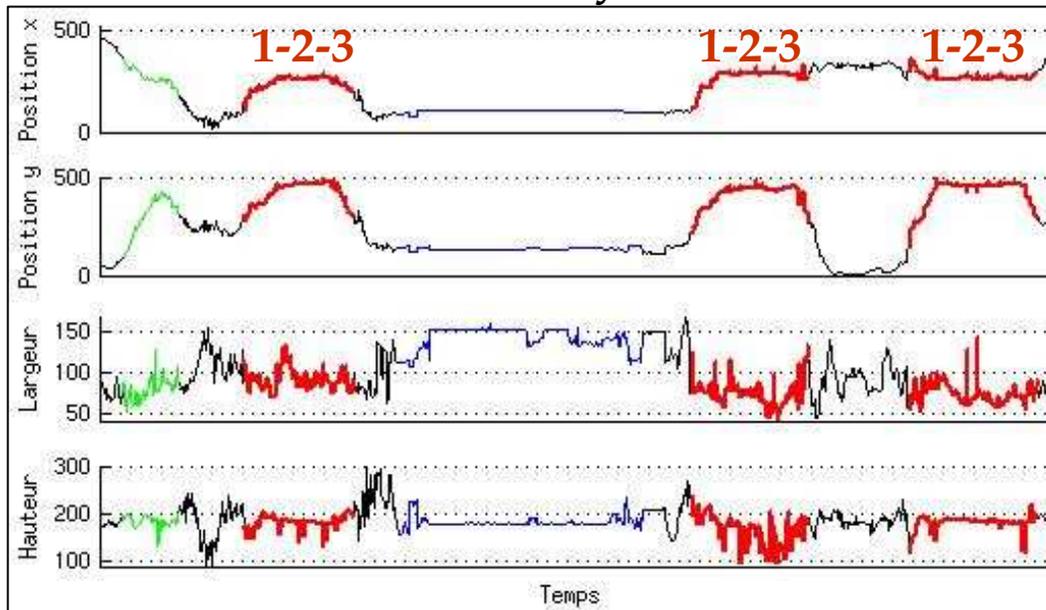
- vecteurs discrets
- **1 symbole** = continuité d'une même action réalisée



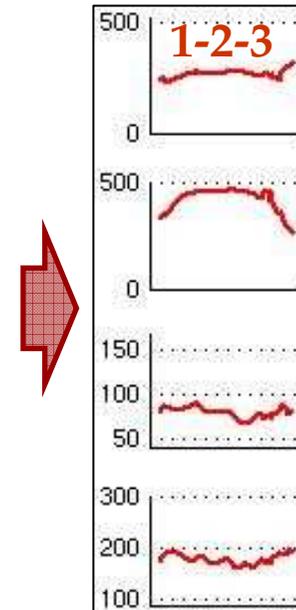
Apprentissage pour la surveillance vidéo

Événement - plus ou moins complexe

➔ Motif d'au moins deux symboles



Données brutes



Motif fréquent



« Aller à bibliothèque 2 »



« Prendre livre dans bibliothèque 2 »

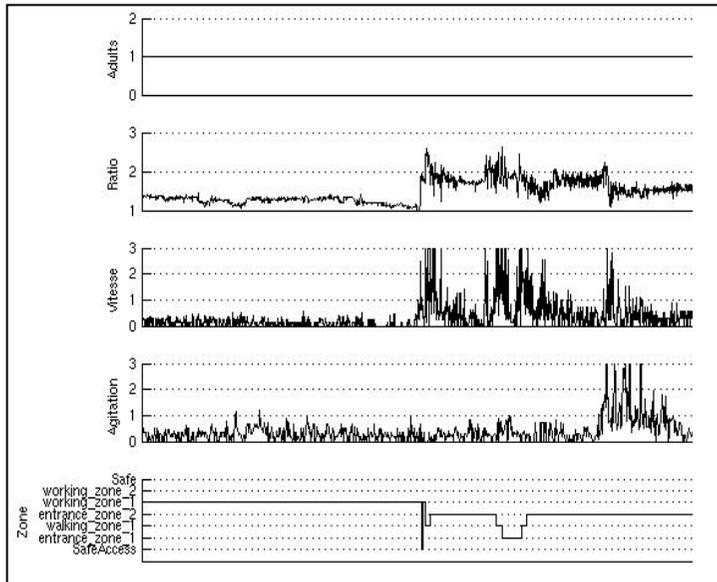


« Partir de bibliothèque 2 »¹⁰

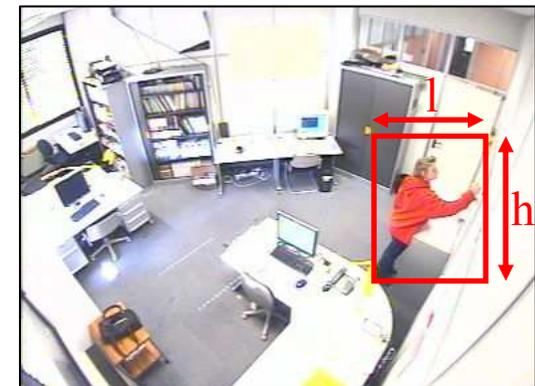
Apprentissage pour la surveillance vidéo

Scénario

- ➔ **Motif** : plusieurs symboles ou motifs « bas niveau »
 - Relations temporelles entre états et événements (motifs « bas niveau »)
 - Extraction de caractères plus « haut niveau », hétérogènes.



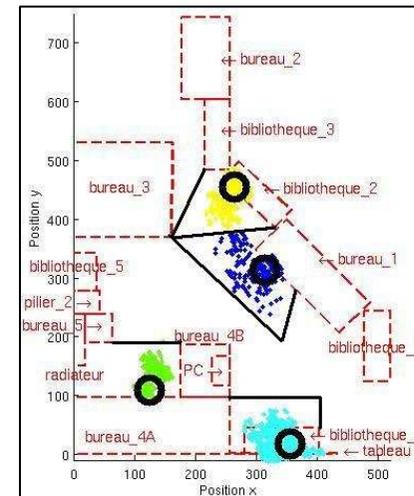
Agitation



Ratio



Vitesse



Zone

Applications à la télésurveillance médicale

Projet GER 'HOM (CSTB, INRIA, CHU Nice)

- Construction d'un laboratoire expérimental à Sophia-Antipolis.
- Intégration de plusieurs capteurs dont des **caméras vidéos**.
- Expérimentation prévue avec des personnes âgées.
(scénario de vie sur 1/2 journée)

Objectifs

- Reconnaissance de **situations inquiétantes** et **assistance** a la personne
- **Modélisation du comportement** habituel
- Intégration de **connaissances *a priori*** issues de l'EGS (Evaluation Gérontologique Standardisée)
- Analyse croisée des **incapacités** (*observations à domicile*) et **déficiences** (*évaluation par le gérontologue*)

