Scene Understanding (3)

perception, multi-sensor fusion, spatio-temporal reasoning and activity recognition.

Francois BREMOND

PULSAR project-team,
INRIA Sophia Antipolis, FRANCE

Francois.Bremond@sophia.inria.fr
http://www-sop.inria.fr/pulsar/

Key words: Artificial intelligence, knowledge-based systems, cognitive vision, human behavior representation, scenario recognition

Understanding systems

- Building systems
  - GERHOME: homecare monitoring for elderly
  - CARETAKER: management of large multimedia collections
    - Trajectory clustering
    - Activity clustering
    - Frequent Event Discovery in Videos (Learning Scenario Models)
- 3 types of tools:
  - performance evaluation
  - system control, optimization, supervision
  - learning
    - algorithm parameters
    - frequent trajectories
    - frequent activities

HealthCare Monitoring: (N. Zouba)

GERHOME (CSTB, INRIA, CHU Nice) :
Ageing population

Approach :
Multi-sensor analysis based on sensors embedded in the home environment
- Detect in real-time any alarming situation
- Measure the frailty level of the elderly
- Identify a person profile – his/her usual behaviors - from the global trends of life parameters, and then to detect any deviation from this profile

Gerhome laboratory

- **GERHOME (Gerontology at Home) : homecare laboratory**

http://www-sop.inria.fr/orion/personnel/Francois.Bremond/topicsText/gerhomeProject.html

- Experimental site in CSTB (Centre Scientifique et Technique du Bâtiment) at Sophia Antipolis

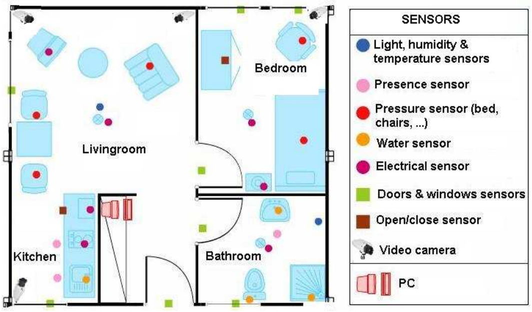
http://gerhome.cstb.fr

- **Partners:** INRIA, CSTB, CHU-Nice, CG06…



Gerhome laboratory

- Video cameras installed in the kitchen and in the living-room to detect and track the person in the apartment.
- Contact sensors mounted on many devices to determine the interactions with the person.
- Presence sensors installed in front of sink and cooking stove to detect the presence of people near sink and stove.



Position of the sensors in Gerhome laboratory

Sensors installed in Gerhome laboratory

Contact sensor in the cupboard door

Contact sensor in the window

Pressure sensor underneath the legs of armchair

Video camera in the living-room

Results: recognition of the "Fainting" activity

- The person is recognized with the postures "standing" "bending", and "sitting on the floor with outstretched legs".

Results: recognition of the "Falling down" activity

- The person is recognized with the postures "standing", " sitting on the floor with flexed legs ", and "lying with outstretched legs".

Event modelling

- We have modelled a set of activities by using an event recognition language.
- This is an example for "Meal preparation" event.

Composite Event (**Prepare_meal_1**, **"detected by a video camera combined with a contact sensors"**

Physical Objects ((p: Person), (Microwave: Equipment), (Fridge: Equipment), (Kitchen: Zone))

Components ((p_inz: PrimitiveState **Inside_zone** (p, Kitchen)) **"detected by video camera"**

(open_fg: PrimitiveEvent **Open_Fridge** (Fridge)) **"detected by contact sensor"**

(close_fg: PrimitiveEvent **Close_Fridge** (Fridge)) **"detected by contact sensor"**

(open_mw: PrimitiveEvent **Open_Microwave** (Microwave)) **"detected by contact sensor"**

(close_mw: PrimitiveEvent **Close_Microwave** (Microwave))) **"detected by contact sensor"**

Constraints ((open_fg **during** p_inz)

(open_mw **before_meet** open_fg)

(open_fg **Duration**>= 10)

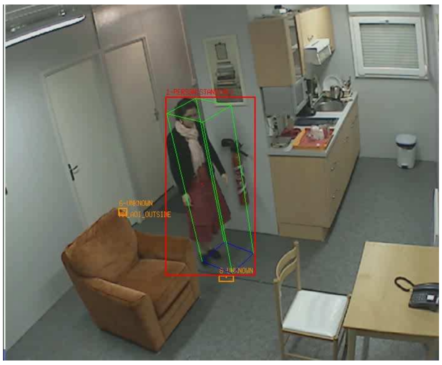
(open_mw **Duration**>=5))

Action (AText ("**Person prepares meal**")

AType ("**NOT URGENT**")))

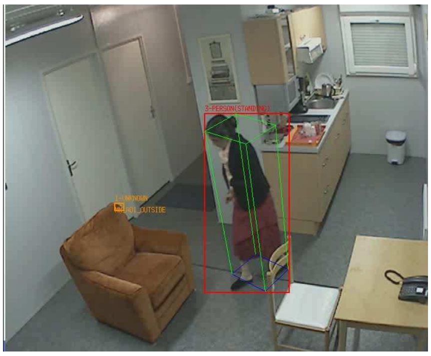
Recognition of the "Prepare meal" event

- The person is recognized with the posture "standing with one arm up", "located in the kitchen" and "using the microwave".



Recognition of the “Resting in living-room” event

- The person is recognized with the posture “sitting in the armchair” and “located in the living-room”.



Recognition Results

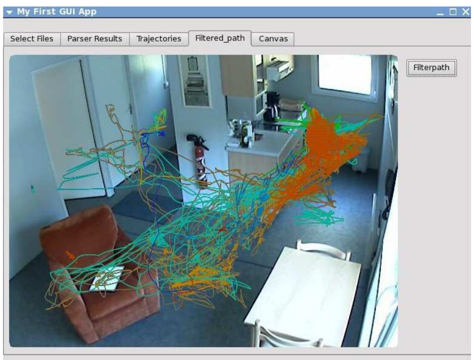
Recognition of the activity “**prepare a meal**” using only a video camera.

Recognition Results

- Recognition of the activity “sitting in the living-room” then “get up” and “go in the kitchen”.

Recognition Results

People localization during 3 hours



Different stationary areas during the 3 hours

Walked distance during the 3 hours = 3.71 km

Preliminary comparative results

Volontaires	Volontaire 1 (< 65 ans)			Volontaire 2 (> 80 ans)		
Activités / s.ous activités	**Temps moyen (min)**	**Durée totale (min)**	**Nb instances**	**Temps moyen (min)**	**Durée totale (min)**	**Nb instances**
-Utiliser le frigidaire	0 :12	2 :50	**14**	0 :13	1:09	**5**
-Utiliser les plaques chauffantes	0 :08	4 :52	**35**	0 :16	27:57	**102**
-Utiliser le micro-onde	0	0	**0**	0	0	**0**
-Utiliser l'eau chaude (cuisine)	0 :19	12:40	**40**	0 :09	3:31	**22**
-Utiliser l'eau froide (cuisine)	0 :28	9:36	**20**	0 :03	0:58	**19**
-Utiliser le tiroir haut	0 :51	21:34	**25**	4 :42	42:24	**9**
-Utiliser les toilettes	0	0	**0**	0 :56	0 :56	**1**
-TV allumée	42 :18	169:12	**4**	16 :19	65:18	**4**
-Assis sur la chaise1	0	0	**0**	92 :42	185:25	**2**
-Assis sur la chaise2	6 :07	73:27	**12**	0	0	**0**
-Assis sur le fauteuil	0	0	**0**	0 :01	0:06	**6**
-Etre debout immobile	0.09	30 :00	**200**	0.16	12 :00	**45**
-Se baisser	0.04	2 :00	**30**	0.20	5 :00	**15**
-Entrer dans le salon	1 :25	25 :00	**20**	2 :68	35 :00	**13**
-Entrer dans la cuisine	2 :45	12 :00	**5**	2 :66	8 :00	**3**
-Entrer dans la chambre	1 :25	5 :00	**4**	3 :28	23 :00	**7**
-Entrer dans la salle de bain	1 :00	2 :00	**2**	1 :00.	5 :00	**5**
-Etre dans l'entrée	2 :00	8 :00	**4**	1 :83	11 :00	**6**

Preliminary comparative results

Volontaires	Volontaire 1 (< 65 ans)			Volontaire 2 (> 80 ans)		
Activités / s.ous activités	**Temps moyen (min)**	**Durée totale (min)**	**Nb instances**	**Temps moyen (min)**	**Durée totale (min)**	**Nb instances**
-Utiliser le frigidaire	0 :12	2 :50	**14**	0 :13	1:09	**5**
-Utiliser les plaques chauffantes	0 :08	4 :52	**35**	0 :16	27:57	**102**
-Utiliser le micro-onde	0	0	**0**	0	0	**0**
-Utiliser l'eau chaude (cuisine)	0 :19	12:40	**40**	0 :09	3:31	**22**
-Utiliser l'eau froide (cuisine)	0 :28				0:58	**19**
-Utiliser le tiroir haut	0 :51				42:24	**9**
-Utiliser les toilettes	0				0 :56	**1**
-TV allumée	42 :18				65:18	**4**
-Assis sur la chaise1	0				185:25	**2**
-Assis sur la chaise2	6 :07	73:27	**12**	0	0	**0**
-Assis sur le fauteuil	0	0	**0**	0 :01	0:06	**6**
-Etre debout immobile	0.09	**30 :00**	**200**	0.16	12 :00	**45**
-Se baisser	0.04	2 :00	**30**	**0.20**	5 :00	**15**
-Entrer dans le salon	1 :25	25 :00	**20**	2 :68	35 :00	**13**
-Entrer dans la cuisine	2 :45	12 :00	**5**	2 :66	8 :00	**3**
-Entrer dans la chambre	1 :25	5 :00	**4**	3 :28	23 :00	**7**
-Entrer dans la salle de bain	1 :00	2 :00	**2**	1 :00.	5 :00	**5**
-Etre dans l'entrée	2 :00	8 :00	**4**	1 :83	11 :00	**6**

+ or - dynamic

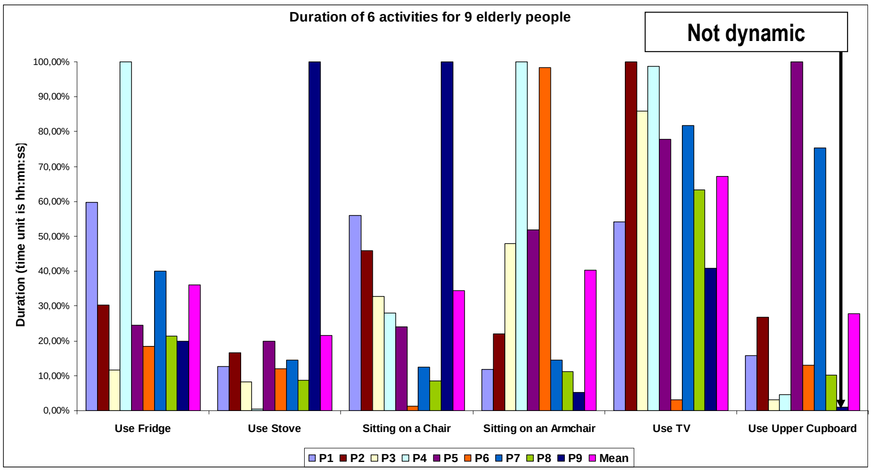
Preliminary comparative results

Volontaires	Volontaire 1 (< 65 ans)			Volontaire 2 (> 80 ans)		
Activités / s.ous activités	**Temps moyen (min)**	**Durée totale (min)**	**Nb instances**	**Temps moyen (min)**	**Durée totale (min)**	**Nb instances**
-Utiliser le frigidaire	0 :12	2 :50	**14**	0 :13	1:09	**5**
-Utiliser les plaques chauffantes	0 :08	4 :52	**35**	0 :16	27:57	**102**
-Utiliser le micro-onde	0	0	[illegible]	[illegible]	0	**0**
-Utiliser l'eau chaude (cuisine)	0 :19	12:40	**40**	0 :09	3:31	**22**
-Utiliser l'eau froide (cuisine)	0 :28	9:36	**20**	0 :03	0:58	**19**
-Utiliser le tiroir haut	0 :51	**21:34**	**25**	**4 :42**	42:24	**9**
-Utiliser les toilettes	0	0	**0**	0 :56	0 :56	**1**
-TV allumée	42 :18	169:12	**4**	16 :19	65:18	**4**
-Assis sur la chaise1	0	0	**0**	92 :42	185:25	**2**
-Assis sur la chaise2	6 :07	73:27	**12**	0	0	**0**
-Assis sur le fauteuil	0	0	**0**	0 :01	0:06	**6**
-Etre debout immobile	0.09	30 :00	**200**	0.16	12 :00	**45**
-Se baisser	0.04	2 :00	**30**	0.20	5 :00	**15**
-Entrer dans le salon	1 :25	25 :00	**20**	2 :68	35 :00	**13**
-Entrer dans la cuisine	2 :45	12 :00	**5**	2 :66	8 :00	**3**
-Entrer dans la chambre	1 :25	5 :00	**4**	3 :28	23 :00	**7**
-Entrer dans la salle de bain	1 :00	2 :00	**2**	1 :00.	5 :00	**5**
-Etre dans l'entrée	2 :00	8 :00	**4**	1 :83	11 :00	**6**

+ or – easy to use

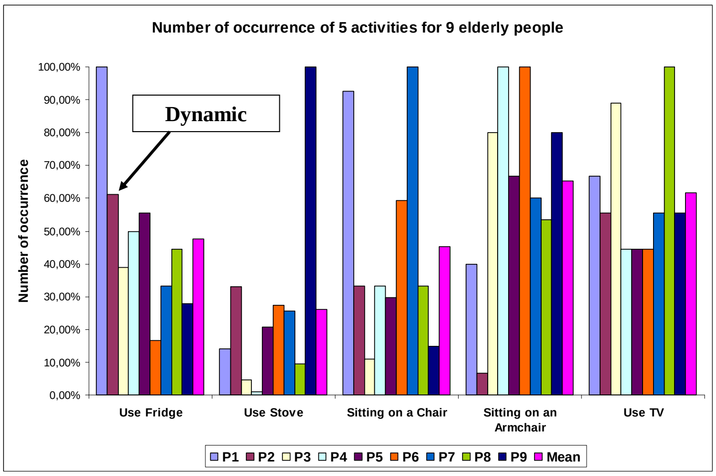
INRIA SOPHIA ANTIPOLIS

Preliminary comparative results



INRIA SOPHIA ANTIPOLIS

Preliminary comparative results



End-users

There are several end-users in homecare:

- Doctors (gerontologists):
  - Frailty measurement (depression, …)
  - Alarm detection (falls, gas, dementia, …).
- Caregivers and nursing home:
  - Cost reduction: no false alarm and reduction employee involvement.
  - Employee protection.
- Persons with special needs, including young children, disabled and elderly people:
  - Feeling safe at home.
  - Autonomy: at night, lighting up the way to bathroom.
  - Improving life: smart mirror, summary of user day, week, month, in terms of walking distance, TV, water consumption.
- Family members and relatives:
  - Elderly safety and protection.
  - Social connectivity.

Social problems and solutions

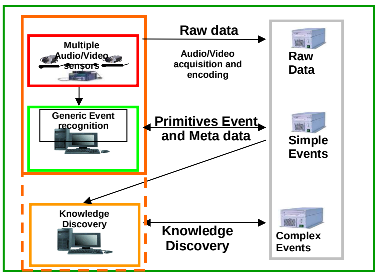
Problems	Solutions
Privacy confidentiality and ethics: video (and other data) recording, processing and transmission.	No video recording and transmission, only textual alarms.
Acceptability for elderly	User empowerment.
Usability	Easy ergonomic interface (no keyboard, large screen), friendly usage of the system.
Cost effectiveness	The right service for the right price, large variety of solutions.
Legal issues, no certification	Robustness, benchmarking, on site evaluation
Installation, maintenance, training, interoperability with other home devices	Adaptability, X-Box integration, wireless, standards (OSGI, …)
Research financing	? France (no money, lobbies), Europe (delay), US, Asia.

Video Understanding : Knowledge Discovery (E. Corvee, JL. Patino_Vilchis)

• **CARETAKER**: An FP6 IST European initiative to provide an efficient tool for the management of large multimedia collections.

• Applications to surveillance and safety issues, in urban/environment planning, resource optimization, disabled/elderly person monitoring.

• Validated on large underground video recordings (Torino, Roma).



Event detection examples

Data Flow

Object/Event Detection

Object Detection
- Id
- Type
- Info 2D
- Info 3D

Event Detection
- Id
- Type (inside_zone, stays_inside_zone)
- Involved Mobile Object
- Involved Contextual Object

Information Modelling

Mobile object table

Event table

Contextual object table

Mobile Objects

mob_obj_id	mob_obj_type	startframe	endframe	traj_type	shape_type	sig_evnt_id
6409	Person	81305	81365	2	small	inside_zone_Platform
6412	Person	81335	81335	1	tall	inside_zone_Platform
6392	PersonGroup	81130	81280	2	small	group_inside_zone_Platform
6400	PersonGroup	81210	81210	1	small	group_inside_zone_Platform
6399	Person	81200	81205	1	tall	inside_zone_Platform
6385	Person	81090	81150	2	tall	inside_zone_Platform
6381	PersonGroup	81030	81075	2	small	group_inside_zone_Platform
6374	Luggage	80950	81025	2	big	inside_zone_Platform
6370	Person	80900	80945	1	tall	inside_zone_Platform
6369	Luggage	80855	80935	2	large	inside_zone_Platform
6366	Person	80775	80775	1	tall	inside_zone_Platform
6353	Person	80205	80390	2	small	inside_zone_Platform
6361	Person	80310	80325	1	tall	inside_zone_Platform
5942	Person	76930	80265	2	tall	inside_zone_Platform
6342	Person	79780	80195	2	small	inside_zone_Platform
6322	PersonGroup	79565	79875	1	large	group_inside_zone_Platform

Knowledge Discovery: trajectory clustering

Objective: Clustering of trajectories into k groups to match people activities

- Feature set
  - Entry and exit points of an object
  - Direction, speed, duration, …
- Clustering techniques
  - Agglomerative Hierarchical Clustering.
  - K-means
  - Self-Organizing (Kohonen) Maps
- Evaluation of each cluster set based on Ground-Truth

Trajectory: Clustering Methods

System structure :

Input data : trajectories

- v_1: the entry point (entry)
- v_2: the exit point (exit)
- v_3: the direction (angle)
- v_4: the sum of the distance between all subsequent pairs of points (walkeddist)
- v_5: the distance from entry to exit
- v_6: the mean variation of the angle between all subsequent pairs of points

Clustering algorithm

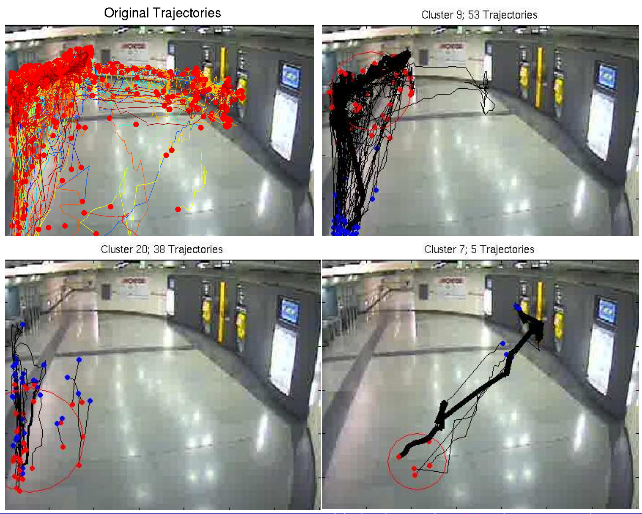
- •Classifies a trajectory into an existing cluster
- •Creates a new cluster if necessary

Cluster update

Trajectory clusters DB

End-user

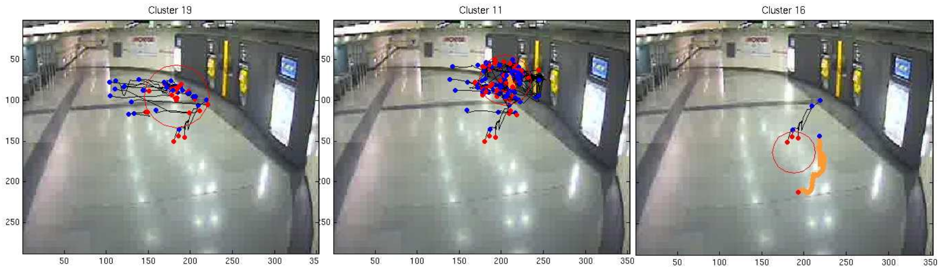
Results on Torino subway (45min), 2052 trajectories



- Entrance
- Exit

Trajectory: Analysis

SOM K-means Agglomerative



Groups with mixed overlap

Trajectory: Analysis

Intraclass & Interclass variance

	Agglomerative	K-means	SOM
Intraclass	1070.82	597.73	582.7
Interclass	10799.62	11272.71	11287.74

$$Intraclass = \frac{\sum_i \left[\sum_{j \in i} \left(v_j - \bar{v}_i \right)^2 \right]}{J}$$

$$Interclass = \frac{\sum_i \left(\bar{v}_i - \bar{v} \right)}{J}$$

$$\bar{v} = \frac{\sum_j v_j}{J}$$

- SOM algorithm has the lowest intraclass and higher interclass separation,
- Parameter tuning: which clustering techniques?

Trajectory clustering comparative evaluation

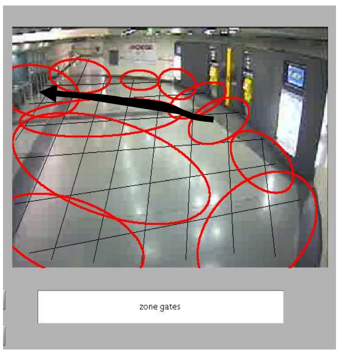
jlp1

		Clustring Technique			
		Agglomerative	K-means	SOM	
Validity index	Silhouette	0.32116	0.29501	0.3036	Higher is better
	Dunn	0.10098	0.064249	0.05031	Higher is better
	Davies Bouldin	0.58685	0.49352	0.37311	Smaller is better

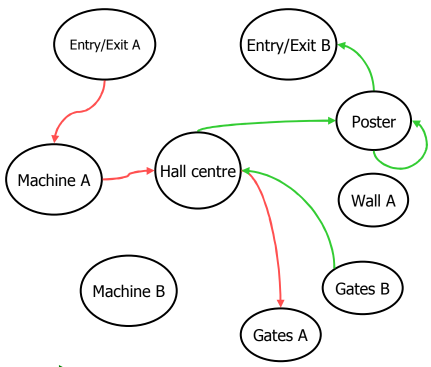
Validity indexes for 31 cluster configuration

Trajectory clustering evaluation

Trajectory : Zones of Interest **Definition for normal activities**



Trajectory label:
From Vending Machines to the gates



Go through the gates, stopping at a poster before exiting

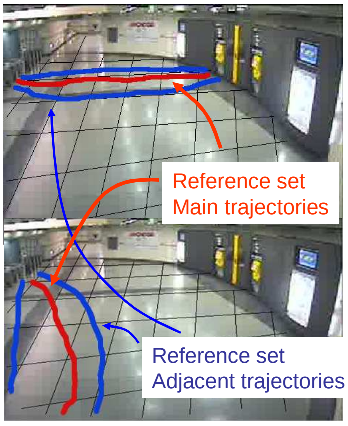
Enter, buy a ticket and go through the gates

Diapositive 31

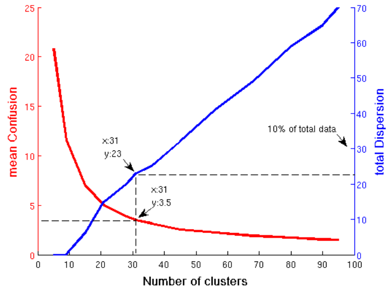
---

jlp1 Clustering
jlpatino; 08/04/2008

Trajectory clustering evaluation



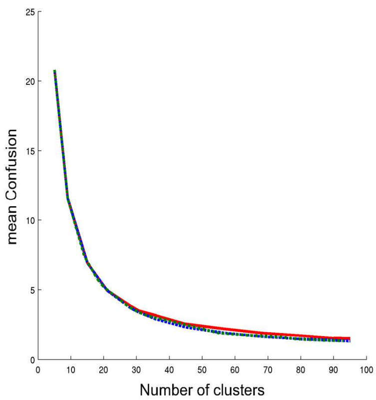
300 reference trajectories

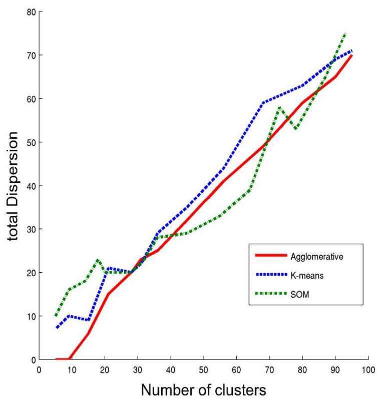


Confusion: nb. GT trajectory types set together in a cluster
Dispersion: percentage of GT trajectories of the same type per cluster

31 clusters represent a good compromise between confusion and dispersion

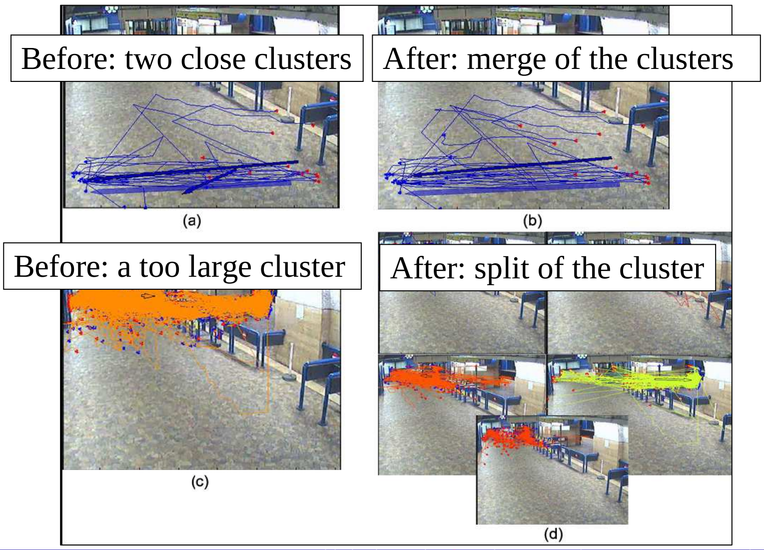
Trajectory clustering comparative evaluation





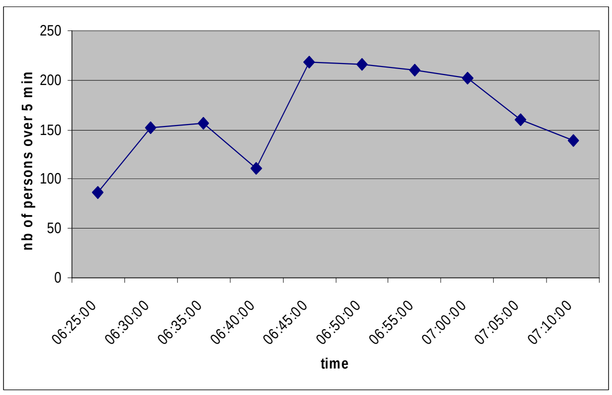
The agglomerative algorithm outperforms other techniques up to 31 clusters

Trajectory feature weight optimization : Results



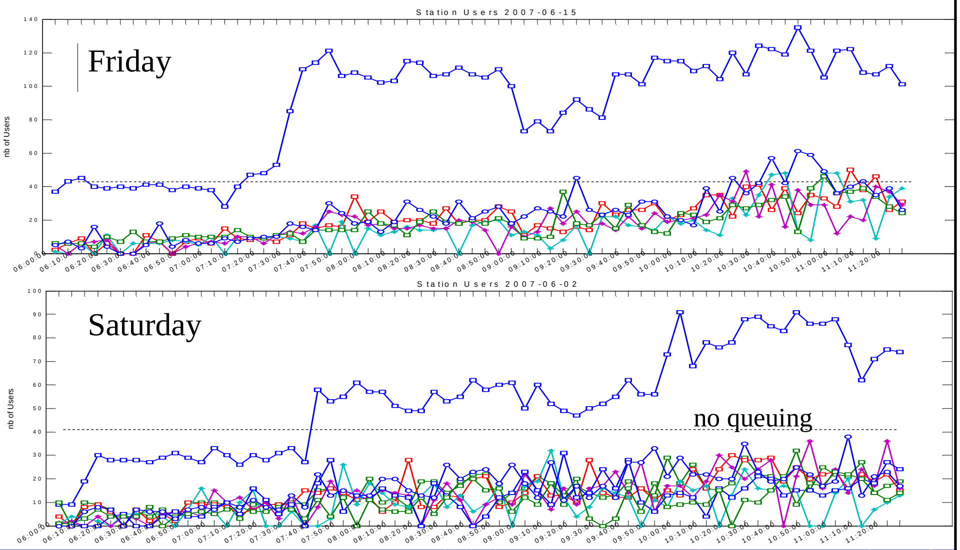
Mobile Object Analysis

Building statistics on Objects



There is an increase of people after 6:45

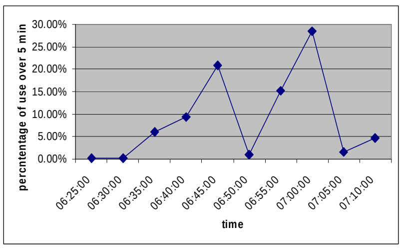
Two day analysis

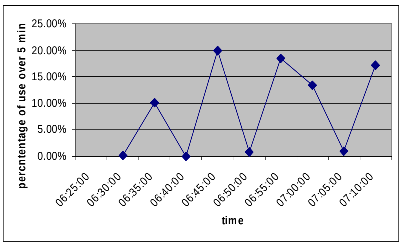


Contextual Object Analysis

Vending Machine 2

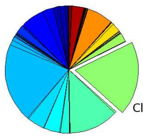
Vending Machine 1

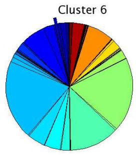


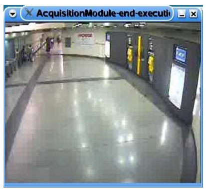


With an increase of people, there is an increase on the use of vending machines

Results : Trajectory Clustering



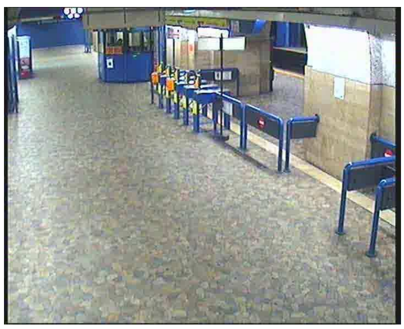




	Cluster 38	Cluster 6
Number of objects	385	15
Object types	types: {'Crowd'} freq: 385	types: {'Person'} freq: 15
Start time (min)	[0.1533, 48.4633]	[28.09, 46.79]
Duration (sec)	[0.04, 128.24]	[2.04, 75.24]
Trajectory types	types: {'4' '3' '7'} freq: [381 1 3]	types: {'13' '12' '19'} freq: [13 1 1]
Significant event	types: {'void '} freq: 385	types: {'inside_zone_Platform '} freq: 15

Knowledge Discovery: achievements

- Semantic knowledge extracted by the off-line long term analysis of on-line interactions between moving objects and contextual objects:
  - 70% of people are coming from north entrance
  - Most people spend 10 sec in the hall
  - 64% of people are going directly to the gates without stopping at the ticket machine
  - At rush hours people are 40% quicker to buy a ticket, …
- Issues:
  - At which level(s), clustering techniques should be designed : low level (image features)/ middle level (trajectories, shapes)/ high level (primitive events)?
  - uncertainty (noise/outliers/rare), what are the activities of interest?



Video Understanding:
Performance Evaluation (V. Valentin, R. Ma)

• **ETISEO**: French initiative for algorithm validation and knowledge acquisition: http://www-sop.inria.fr/pulsar/ETISEO/

• **Approach**: 3 critical evaluation concepts
- Selection of test video sequences
  - Follow a specified characterization of problems
  - Study one problem at a time, several levels of difficulty
  - Collect long sequences for significance
- Ground truth definition
  - Up to the event level
  - Give clear and precise instructions to the annotator
    - E.g., annotate both visible and occluded part of objects
- Metric definition
  - Set of metrics for each video processing task
  - Performance indicators: sensitivity and precision

Evaluation : current approach
(AT. NGHIEM)

- ETISEO limitations:
- Selection of video sequence according to difficulty levels is subjective
- Generalization of evaluation results is subjective.
- One video sequence may contain several video processing problems at many difficulty levels

- Approach: treat each video processing problem separately
- Define a measure to compute difficulty levels of input data (e.g. video sequences)
- Select video sequences containing only the current problem at various difficulty levels
- For each algorithm, determine the highest difficulty level for which this algorithm still has acceptable performance.

- Approach validation : applied to two problems
- Detect weakly contrasted objects
- Detect objects mixed with shadows

Evaluation :
Detection of weakly contrasted objects

- Problem definition:
  - the lower the object contrast, the worse the object detection performance

- For one algorithm, determine the lowest object contrast for which this algorithm has an acceptable performance

- Issue: one blob may contain many regions at several contrast levels

Evaluation :
Detection of weakly contrasted objects

- At each contrast level, the algorithm performance is x/m
  - m: number of all blobs containing current contrast level
  - x: number of blobs containing current contrast level and detected by given algorithm
- Algorithm capacity: the lowest contrast level for which algorithm performance is bigger than a given threshold

Number of blobs	Contrast level								
Sequence	0	1	2	3	4	5	6	7	8
ETI-VS1-BE-18-C1	0	19	46	29	13	5	5	2	0
ETI-VS1-BE-18-C4	0	15	16	6	11	12	14	10	9
ETI-VS1-MO-7-C1	0	3	30	30	30	30	1	0	0

Evaluation: Detection of weakly contrasted objects

Error rate on sequence ETI-VS1-BE-18-C4

Error rate	Contrast level							
Participant	0	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0
12	0	0.2	0.06	0.08	0.15	0.08	0	0.09
15	0	0.5	0.19	0.17	0.23	0.17	0.07	0.18
8	0	0.8	0.38	0.17	0.15	0.25	0.33	0.27
13	0	0.9	0.63	0.58	0.08	0.17	0.33	0.09

- Error rate threshold to determine algorithm capacity: 50%

Contrast level	Participants						
Sequence	1	8	9	11	13	15	22
ETI-VS1-BE-18-C4	0	2	-	0	4	2	0
ETI-VS1-MO-7-C1	0	2	0	0	-	-	-

INRIA SOPHIA ANTIPOLIS

Evaluation : conclusion

- A new evaluation approach to generalise evaluation results.
  - Implement this approach for 2 problems.
- Limitations: only detect the upper bound of algorithm capacity.
- The difference between the upper bound and the real performance may be significant if:
  - The test video sequence contains several video processing problems
  - The same set of parameters is tuned differently to adapt to several concurrent problems
- Ongoing evaluation campaigns:
  - VS, PETS at ECCV, CVPR, AVSS
  - TRECVid (NIST) with ILids video
- Benchmarking databases:
  - http://homepages.inf.ed.ac.uk/cgi/rbf/CVONLINE/entries.pl?TAG363
  - http://www.hitech-projects.com/euprojects/cantata/datasets_cantata/dataset.html

INRIA SOPHIA ANTIPOLIS

Video Understanding: Learning Parameters (B.Georis)

- **Objective**: a tool to automatically learn algorithm parameters based on experimental data

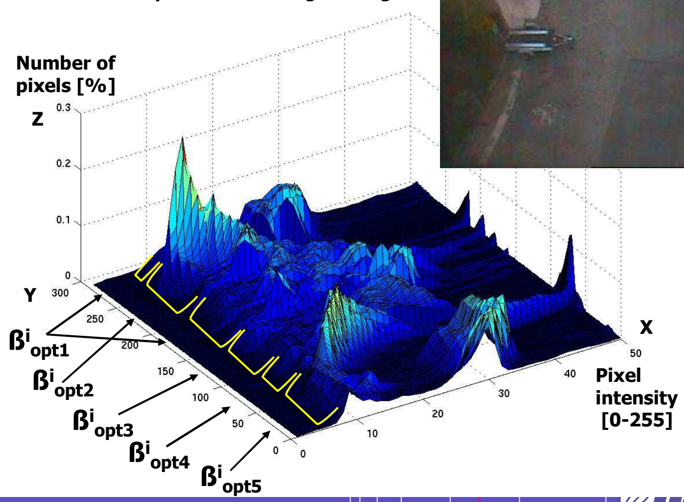
- Used for learning the segmentation parameters with respect to the illumination conditions

- **Method**
  - Identify a set of parameters of a task
    - 18 segmentation thresholds
  - depending on environment characteristics
    - Image intensity histogram
  - Study the variability of the characteristic
    - Histogram clustering -> 5 clusters
  - Determine optimal parameters for each cluster
    - Optimization of the 18 segmentation thresholds

Video Understanding: Learning Parameters

Camera View

Learning Parameters
Clustering the Image Histograms



Conclusion

A **global framework** for building video understanding systems:

- Hypotheses:
  - mainly fixed cameras
  - 3D model of the empty scene
  - predefined behavior models
- Results:
  - Video understanding real-time **systems** for Individuals, Groups of People, Vehicles, Crowd, or Animals …
  - **Knowledge** structured within the different abstraction levels
    - Formal description of the empty scene
    - Structures for algorithm parameters
    - Structures for object detection rules, tracking rules, fusion rules, …
    - Operational **language** for **event** recognition (more than 60 states and events), video event ontology
  - Tools for **knowledge** management
    - Metrics, tools for performance evaluation, learning
    - Parsers, Formats for data exchange
    - …

Conclusion: perspectives

Object and video event detection

- Finer human shape description: *gesture models*
- Video analysis robustness*: reliability computation*

Knowledge Acquisition

- Design of learning techniques to complement a priori knowledge:
  - visual concept learning
  - scenario model learning

System Reusability

- Use of program supervision techniques: *dynamic configuration of programs and parameters*
- Scaling issue: managing large network of heterogeneous sensors (cameras, microphones, optical cells, radars….)

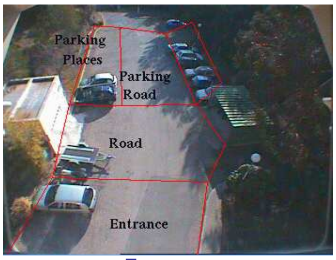
INRIA SOPHIA ANTIPOLIS

Video Understanding : Learning Scenario Models (A. Toshev)

or Frequent Composite Event Discovery in Videos event time series

INRIA SOPHIA ANTIPOLIS

Learning Scenarios: Problem Definition



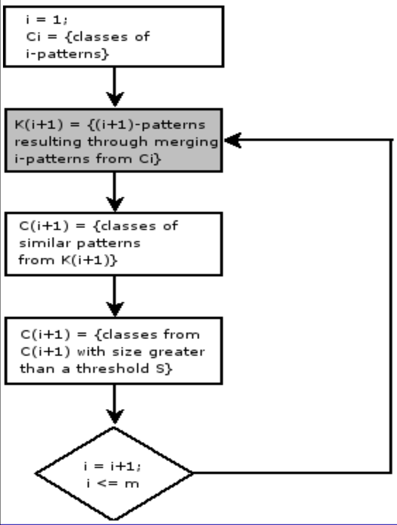
Zones

- Input: A set of **primitive events** from the vision module:
  object-inside-zone(***Vehicle***, ***Entrance***) [5,16]
- Output: **frequent event patterns.**
- A pattern is a set of events:
  object-inside-zone(*Vehicle*, *Road*) [0, 35]
  object-inside-zone(*Vehicle*, *Parking_Road*) [36, 47]
  object-inside-zone(*Vehicle*, *Parking_Places*) [62, 374]
  object-inside-zone(*Person*, *Road*) [314, 344]

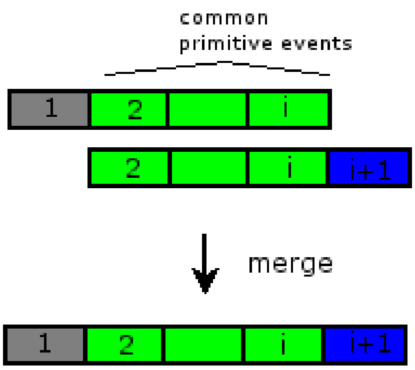
- **Goals**:
  - **Automatic data-driven modeling of composite events,**
  - **Reoccurring patterns of primitive events** ⇒ correspond to frequent activities,

  Find classes with large size & similar patterns.

Learning Scenarios: A PRIORI Method



Merge two *i*-patterns with *(i-1)* primitive events in common to form an (*i+1*)-pattern:



Learning Scenarios: Similarity Measure

2 types of Similarity Measure between event patterns :

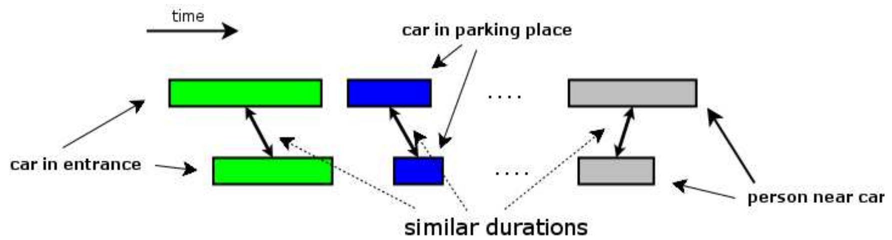
- similarities between event attributes
- similarities between pattern structures

Generic Similarity Measure :

- **Generic properties** when possible ⇒ easy usage in different domains,
- It should incorporate **domain-dependent properties** ⇒ relevance to the concrete application.

Learning Scenarios: Attribute Similarity

Attributes: the **corresponding events** in two patterns should have similar (same) attributes (duration, names, object types,...).



- Comparison between corresponding events (same type, same color).
- For numeric attributes: $G(x,y)= e^{-\frac{(x-y)^2}{\alpha xy}}$
- attr(p_i, p_j) = average of all event attribute similarities.

Learning Scenarios: Evaluation

Test data:

•Video surveillance at a **parking lot**,

•**4 hours** records from 2 days in **2 test sets**,

•Every test set contains appr. **100 primitive events**.

Results: In both test sets the following event pattern was recognized:

object-inside-zone(Vehicle, Road)
object-inside-zone(Vehicle, Parking_Road)
object-inside-zone(Vehicle, Parking_Places)
object-inside-zone(Person, Parking_Road)

Learning Scenarios: Evaluation

Test data:

•Video surveillance at a **parking lot**,

•**4 hours** records from 2 days in **2 test sets**,

•Every test set contains appr. **100 primitive events**.

Results: In both test sets the following event pattern was recognized:

object-inside-zone(Vehicle, Road)
object-inside-zone(Vehicle, Parking_Road)
object-inside-zone(Vehicle, Parking_Places)
object-inside-zone(Person, Parking_Road)

Learning Scenarios: Evaluation

Test data:

•Video surveillance at a **parking lot**,

•**4 hours** records from 2 days in **2 test sets**,

•Every test set contains appr. **100 primitive events**.

Results: In both test sets the following event pattern was recognized:

object-inside-zone(Vehicle, Road)
object-inside-zone(Vehicle, Parking_Road)
object-inside-zone(Vehicle, Parking_Places)
object-inside-zone(Person, Parking_Road)

Learning Scenarios: Evaluation

Test data:

•Video surveillance at a **parking lot**,

•**4 hours** records from 2 days in **2 test sets**,

•Every test set contains appr. **100 primitive events**.

Results: In both test sets the following event pattern was recognized:

object-inside-zone(Vehicle, Road)
object-inside-zone(Vehicle, Parking_Road)
object-inside-zone(Vehicle, Parking_Places)
object-inside-zone(Person, Parking_Road)

Maneuver Parking!

Learning Scenarios: Conclusion & Future Work

Conclusion:

- Application of a **data mining approach**,
- Handling of **uncertainty** without losing computational effectiveness,
- **General framework**: only a similarity measure and a primitive event library must be specified.

Future Work:

- Other similarities,
- Handling of different aspects of uncertainty,
- Qualification of the learned patterns,
  - Frequent equal interesting ?
- Different applications: different event libraries or features.