Activity Recognition and Uncertain Knowledge in Video Scenes

Rim Romdhane, Carlos Fernando Crispim-Junior, Francois Bremond, Monique Thonnat

Activity Recognition and Uncertain Knowledge in Video Scenes

Rim Romdhane, Carlos fernando Crispim junior, Francois Bremond and Monique Thonnat
INRIA
BP 93 06902 Sophia Antipolis Cedex FRANCE
[Rim.Romdhane, carlos-fernando.crispim junior, Francois.Bremond, Monique.Thonnat]@inria.fr

Abstract

Activity recognition has been a growing research topic in the last years and its application varies from automatic recognition of social interaction such as shaking hands, parking lot surveillance, traffic monitoring and the detection of abandoned luggage. This paper describes a probabilistic framework for uncertainty handling in a description-based event recognition approach. The proposed approach allows the flexible modeling of composite events with complex temporal constraints. It uses probability theory to provide a consistent framework for dealing with uncertain knowledge for the recognition of complex events. We validate the event recognition accuracy of the proposed algorithm on real-world videos. The experimental results show that our system can successfully recognize activities with a high recognition rate. We conclude by comparing our algorithm with the state of the art and showing how the definition of event models and the probabilistic reasoning can influence the results of real-time event recognition.

1. Introduction

In the literature, many video event recognition systems have been proposed. However, many challenging problems still remain to obtain a robust recognition because of noise, illumination changes, segmentation issues and occlusions. We propose a constraint-based approach for real-world video interpretation based on probabilistic reasoning for composite event recognition. The main goal is to improve the techniques of video data interpretation taking into account the imprecision and uncertainty of low level data.

The paper is organized as follows: In section 2, we review the related work. In section 3 and 4 we describe the proposed video interpretation framework for event recognition. The experiments realized to evaluate the proposed method are shown in section 5. Finally, we present the conclusion in section 6.

2. Related work

Automatic activity recognition is a very important and active area of research [7], [5], [19]. Activity recognition approaches can be divided into two main approaches: probabilistic approaches and description-based approaches. The probabilistic approaches include Bayesian Networks and Hidden Markov Models. The main characteristic of these techniques is to model explicitly the uncertainty.

Bayesian Network have been successfully applied to person interaction [17] such as 'shake hands', parking lot surveillance [13], traffic monitoring [8] and detection of left luggage [11]. However, Bayesian Network are not adapted to model temporal compositions of events. Temporal representation is often done using a static representation, where time points or time slices are represented as static processes. HMMs and their extensions [16], [6], [4] have been widely used. Their advantage compared to Bayesian Network is the ability to recognize sequence of events. However HMMs are limited in recognizing sequence of events which involve several mobile objects because the probability of being in a state for a mobile object has to be combined with the probability of being in another state for all other mobile objects. These combinations lead the recognition process to a combinatorial explosion.

Description-based approaches have also been largely used to recognize activities for few decades. Constraint Satisfaction Problem (CSP) has been applied to model activities as constraint networks [18]. Description-based approaches are suitable for recognizing high-level activities. They can more easily incorporate human knowledge into the systems and require less training data as pointed out by many researchers [16], [14], [21]. But the formalism of the event modeling and recognition is largely deterministic and convenient mechanism to handle uncertainty and compensate for the failures of low-level is generally unavailable.

The approaches combining logic and probabilistic reasoning have been designed to overcome the limitations of the previous approaches. Ryoo and Aggarwal [22], [23] have taken advantage of the concept of the hallucinated time intervals, similar to the one used in [12] to deal with un-

certainty. Tran and Davis [24] have adopted probabilistic graphical model, Markov logic networks (MLNs) to probabilistically infer events in a parking lot. In [2] authors have presented a probabilistic event logic (PEL) which uses weighted event-logic formulas to represent probabilistic constraints among events. However, they do not deal with the low-level uncertainty and consider only the recognition of primitive events of basketball game. Kwak et al. [9] have adopted constraint flows to summarize the combination of the primitive events composing a complex event.But the complexity of the recognition algorithm is not described. The logic-probabilistic combination is an promising field of research. However, it has not been fully explored and many efforts are still needed to provide a complete framework for fully handling the uncertainty of recognition. Thus, in this paper, we propose a new approach for complex event recognition which combine logic and probabilistic reasoning for a better performance of the recognition.

3. Event Representation

We propose a generic event representation formalism that is capable to represent all types of events used for the automatic video recognition and that is able to manage the uncertainty of recognition at the event modeling level. This formalism contains the Event description Language described in [25] which is declarative and intuitive, so that the experts of the application domain can easily define and modify it. The main limitation of this language is the lack of mechanism to handle the uncertainty of recognition. For this, we proposed 2 extensions, mainly, we propose (i) the notion of *'utility'* to deal with missed observations (see section 3.1) and we propose (ii) a specific relation in the representation of the event to manage the tracking identifier maintenance (see section 3.2).

There are four types of activities going from simple to more complex: **primitive states**, **composite states**, **primitive events** and **composite events**. An event model is composed of five elements:

- ***Physical objects***: objects of interest involved in the event. The type of the objects depends on the application domain. Physical objects includes mobile objects (e.g. person, vehicle), contextual objects (equipment, zones).

- ***Components***: the sub-events composing the event.

- ***Constraints***: conditions between the physical objects and/or the components including symbolic, logical, spatial and temporal constraints based on Allen predicates [1].

CompositeEvent (MatchingSheetsActivity,
PhysicalObjects ((p:Person), (z_1:Zone), (z_2:Zone), (z_3:Zone))
Components ((c1: **CompositeEvent** change_zones_coffee_Lib_TV (p, z_1, z_2, z_3)
(c2: **PrimitiveEvent** moveFrom_TV_coffeeCorner (p, z_3, z_1)))
Constraints (c1 *before* c2)
Alarm (Level: NOTURGENT))

CompositeEvent (change_zones_coffee_Lib_TV,
PhysicalObjects ((p:Person), (z_1:Zone), (z_2:Zone), (z_3:Zone))
Components ((c1: **PrimitiveEvent** moveFrom_coffeeCorner_Library (p, z_1, z_2))
(c2: **PrimitiveEvent** moveFrom_Library_TV (p, z_2, z_3)))
Constraints (c1 *before* c2)
Alarm (Level: NOTURGENT))

PrimitiveEvent (moveFrom_coffeeCorner_Library,
PhysicalObjects ((p: Person), (z_1: Zone), (z_2: Zone))
Components ((c1: **PrimitiveState** Inside_zone (p, z_1))
(c2: **PrimitiveState** Inside_zone (p, z_2)))
Constraints ((c1 *before* c2)
(z_1'Name = coffee Corner)
(z_1'Name = Library)))

Figure 1. Event Models.

- ***Alarm***: The alarm information describes the importance of the scenario model in terms of emergency (i.e. not urgent, urgent, very urgent). The alarm level can be used to filter the recognized events, for displaying only important events to the user.

The figure 1 illustrates an event model example 'MatchingSheetssActivity' and its sub-events. This activity is modeled with the help of clinicians to define the cognitive functioning status of patients and their motor skills. In this activity, the patient is asked to match a set of sheets (named A, B, C and D) in their specific placement dispersed over the room. The patient have to move from the coffee corner where there is the sheet named A, to the library (sheet B), to TV zone (sheet C, D) and go back to coffee corner.

3.1. Missed Observation

Occlusion and bad imaging conditions (e.g. dark, shadowed areas of the scene) are common conditions that prevent us from observing the occurrence of some events. When we miss the recognition of one of the sub-events the whole event is missed. To prevent from this, we propose a notion of *'utility'* in the definition of the event model by associating a coefficient to each sub-event. *Utility* which is defined by a human expert expresses the importance or priority of sub-events for the recognition of the whole event. Its range is in the interval]0,1], higher is the utility value higher is the importance of the sub-event in the recognition of the whole event. The value 1 means that the sub-event is required for the recognition. Figure 2 shows that the utility coefficient associated to *'PersonSlumping'* is choosen lower (i.e. 0.2) than the utility for *'PersonStanding'* and *'PersonSitting'* (i.e. 0.6). We make the choice to consider that the detection the primitive state *'PersonSlumping'* is not mandatory for the recognition of the event. This choice

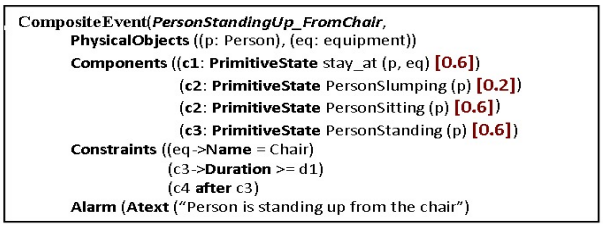

Figure 2. Utility coefficient associated to each sub-event of the event model 'PersonStandingUp-FromChair'.

can be explained by the fact that the posture algorithm is more performant to detect *'PersonStanding'* and *'PersonSitting'* than *'PersonSlumping'*.

3.2. Identity Maintenance

Identity maintenance is necessary when there exist multiple identities that actually refer to the same mobile object. It is caused by lack of visual information (appearance, shape, etc.) to make unique identity connections across observation gaps. Identity maintenance is a primary source of uncertainty for activity recognition. It affects more precisely the recognition of long-term events.

Our approach to solve this issue in the level of event modeling is to propose the use of specific relation ***'equal'*** in the representation of the event. More precisely, the identification whether the identifier of two objects A and B refer to the same object is represented by the relation **equal(A, B)**.

In this work, the evaluation of this relation is done using appearance matching (e.g. 3D height, 3D width, etc). This logic relation is very useful in the case where the vision algorithm (i.e. tracking algorithm) fails to match and maintain the tracking identifier of the detected mobile object. Many other identifying contextual cues about identities can be discussed in the litterature. These cues are based on the individuals belongings, closed place activity, knowledge and appearance as pointed in [24].

4. Event Recognition: a Generic Framework

For the recognition process, an event model tree is computed as described in[25]. The tree defines which sub-event triggers the recognition of which event: the sub-event which happens last in time triggers the recognition of the global event. The first step of the event recognition process is to recognize all the possible primitive statesby instantiating all the models with the detected objects. The second step consists in recognizing complex events according to the event model tree and the simple events previously recognized. The final step checks whether the recognized event at time t has been already recognized previously to update the event end-time or create a new event instance.

4.1. Probabilistic Elementary Event Recognition

The observations are inherently uncertain, hence a formal probabilistic approach is needed to reason under uncertainty. We propose to compute the conditional probability of the recognition of the event instance e belonging to an event model Ω given that the mobile physical objects in the model Ω have been observed and given that the constraints in the model Ω are satisfied by the observation O.

$$P(e \in \Omega | \zeta(\Omega, O), V_\Omega = po_e^O) = \frac{P(\zeta(\Omega, O) | e \in \Omega) \times P(V_\Omega = po_e^O | e \in \Omega) \times P(e \in \Omega)}{P(\zeta(\Omega, O), V_\Omega = po_e^O)} \quad (1)$$

- $e \in \Omega$, e is an instance of event model Ω.
- $\zeta(\Omega, O)$, the constraints of event model Ω are satisfied by observation O.
- $V_\Omega = po_e^O$, the tracked physical objects in the observation O correspond to physical object variables V in the model Ω of event instance e.

$P(e \in \Omega)$ is the prior probability that a certain scenario model Ω is detected. We can assume that all scenarios in a certain universe are equally probable, so as not to favor any scenario just because it happens more often. For example, the universe of the scenario models that describe a person posture is: (*PersonStanding*,*PersonSitting* and *PersonBending*).

$$P(e \in \Omega) = \frac{1}{Nbr.Scenario\Omega Universe} \quad (2)$$

$P(\zeta(\Omega, O) | e \in \Omega)$ is the probability that the constraints of the event model are verified given that the event e is true. This probability is computed as following:

$$P(\zeta(\Omega, O) | e \in \Omega) = \prod_{i=1}^{n} \frac{\sharp(\zeta(\Omega, O)_i \wedge e \in \Omega)}{\sharp(e \in \Omega \wedge V_\Omega^{\zeta_i} \in po_e^O)} \times P(\zeta_i(\Omega, O)) \quad (3)$$

The term $\sharp(\zeta(\Omega, O)_i \wedge e \in \Omega)$ implies that only frames where event e have been identified (i.e. annotated) as an instance of Ω are considered, and for each constraint of event model Ω, the number of frames where it is satisfied are counted. The term $\sharp(e \in \Omega \wedge V_\Omega^{\zeta_i} \in po_e^O)$ indicates that we only consider the frames of the training dataset where the event e is annotated and the physical objects are correctly tracked. The computation of $P(\zeta_i(\Omega, O))$ is detailed in section 4.3. $P(V_\Omega = po_e^O | e \in \Omega)$ is the probability that the physical object variables in the event model Ω have been detected given that e is an event instance of the event model

Ω. This probability is provided by the tracking algorithm as described in[3].

The probability $P(\zeta(\Omega, O), V_\Omega = po_e^O)$ is computed based on the following equation (4):

$$\begin{aligned} &P(\zeta(\Omega, O), V_\Omega = po_e^O) = \\ &P(\zeta(\Omega, O), V_\Omega = po_e^O | e \in \Omega) \times P(e \in \Omega) + \\ &P(\zeta(\Omega, O), V_\Omega = po_e^O | \ \rceil e \in \Omega^{'}) \times P(\rceil e \in \Omega^{'}) \end{aligned} \quad (4)$$

4.2. Probabilistic Complex Event Recognition

The probabilistic recognition of complex event is defined as a hierarchical Bayesian inference. The objective is to recognize the complex event e given an observation O. What we want to calculate here is :
'The probability to recognize a complex event instance e belonging to an event model Ω given that the components (sub-events) in the model Ω are observed and the constraints in the model Ω are satisfied in the observation'. The proposed way of calculating this is:

$$\begin{aligned} &P(e \in \Omega | SE(\Omega, O), \zeta(\Omega, O)) \\ &= \frac{P(SE(\Omega, O) | e \in \Omega) \times P(\zeta(\Omega, O) | e \in \Omega) \times P(e \in \Omega)}{P(SE(\Omega, O), \zeta(\Omega, O))} \end{aligned} \quad (5)$$

- $SE(\Omega, O)$, the components (sub-events) of the model Ω are observed in O.

The different probability terms are calculated in the same manner that for the primitive event case.

4.3. Probabilistic Constraint Verification

In this work, we consider the spatial constraints related to the mobile object speed, position and closeness to a given contextual objects (e.g. person near TV), the constraints related to the posture (e.g. person is sitting) and we use the Allen[1] temporal constraints. A main problem is the imprecision and uncertainty in the detection of the location of mobile objects due to low level detection errors (e.g. reflections, shadows or occlusions). Thus the verification of the constraint may fail. A solution to cope with this problem is to propose a probabilistic verification of the constraint. In the process of spatial constraint verification, we take into account *(i)* the geometrical uncertainty which is related to the verification of the constraint (e.g. verifying the spatial constraint 'person-inside-zone' consists in the geometrical computation whether a point representing the person is inside a polygon representing the zone). The first step consists in computing the distance $dist$ of the person to the contextual objects (i.e. zone), the second step is to find a probability distribution function (PDF) that maximize the value of probability when the distance $dist$ is small and a minimum value when the distance $dist$ is big. We validate experimentally that the distribution of the distances fit into the Gaussian distribution denoted by $\mathcal{N}(\mu, \sigma)$ (Eq. 6), the mean μ and the standard deviation σ are learned off-line.

$$\mathcal{N}(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(dist - \mu)^2}{2\sigma^2}) \quad (6)$$

We take also into account (ii) the uncertainty of the attribute values due to noisy data and low-level algorithm errors. This type of uncertainty handling is described in section 4.4 by proposing a new dynamic model for the re-estimation of the attribute values to deal with noisy data..

4.4. Dynamic Model for temporal attribute filtering

Observations with real video sequences can be corrupted by noise, thus our goal is to estimate more accurately an attribute value given its observed value. The proposed process of the temporal filtering of attributes works in two steps as described in [20]:
- *The first step (1)* consists in computing the expected value a_{exp} of an attribute a at the current instant t_c given the estimated value of a and its velocity at the previous time t_p.
- *The second step (2)* is to compute the estimated value a_{est} of the attribute based on the previous one.
- *The final value* $\bar{a}$ of the attribute is the mean between the expected and the estimated values of the attribute weighted by the expected and estimated reliability values

4.5. Dealing with the Tracking Identifier

The recognition of an event over time needs the maintaining of the same identifier for each mobile object when recognizing its sub-events, otherwise it will be considered as a different object. To deal with this tracking error at the event detection level, we propose the use of the recognition history of an event e, $\{e^1, ..., e^{t-2}, e^{t-1}\}$: the recognized events over time are stored in a buffer and for each time t, and for each detected event e, we propose to look at the change of its physical objects identifier. If the identifier of a physical object changes suddenly and/or for a short period of time, we do not consider the new identifier and we maintain the last identifier of the physical object.

5. Experimental Results

We have evaluated the event recognition accuracy of our algorithm on two real world health care applications (Tab.1and Tab.2) and have compared our results with the approach proposed in [25] (Tab.3). For the first dataset, Video recordings of *37* patients are used to assess the proposed framework performance. These patients are part of a clinical trial for Alzheimer's study and they are asked to perform a set of activities. The length of each video sequence is about $12(\pm 5)$ min, 8 fps. The

Events	*GT*	% R	FP	FN
Person sitting	21	76.2	6	5
Close phone	14	85.7	0	2
In coffee Corner	51	98	3	1
In reading zone	28	100	5	0
In zone library	19	95	2	1
In zone TV	14	100	2	0
Move from coffeeCorner to reading zone	17	100	0	0
Move from reading zone to coffeeCorner	15	100	2	0
Move from coffeeCorner to library zone	11	90	2	1
Move from library zone to coffeeCorner	10	100	0	0
Person reading	14	92	2	1
MatchingSheetssActivity	31	58	1	13

Table 1. Recognition Results of the proposed algorithm: the recognition rate (% R), the false positive (FP) and the false negative (FN). *37* patients, $12(\pm 5)$ min, 8 fps.

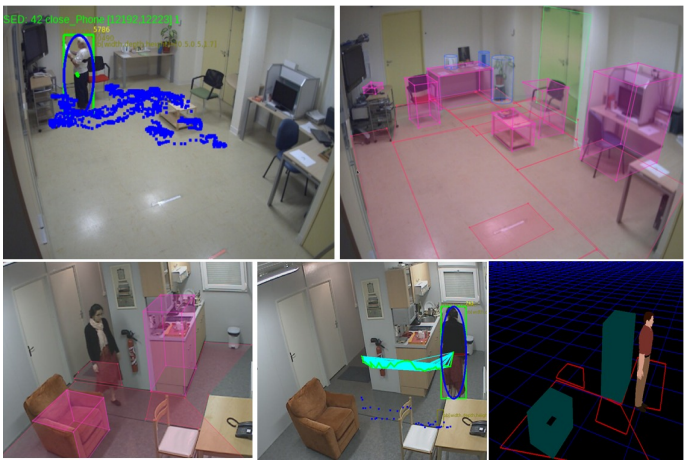

Figure 3. Activity detection evaluated on Health care videos.

Events	*GT*	% R	FP	FN
In Kitchen	12	100	7	0
In LivingRoom	12	91	6	1
Close Armchair	9	88	1	1
Person sitting	4	100	7	0
Sitting at armchair	4	100	7	0
Move-kitchen-LivingRoom	4	100	3	0
Move-zone-LivingRoom-kitchen	6	100	2	0

Table 2. Recognition Results of the proposed algorithm on Gerhome dataset: the recognition rate (% R), the false positive (FP) and the false negative (FN). The ground truth GT corresponds to 4 videos sequences, with a total of 9452 frames, 8 fps.

video dataset will be soon public. The second dataset consists in monitoring elderly observed in an experimental laboratory during 4 hours, the data is available on www-sop.inria.fr/members/Francois.Bremond/topicsText/gerhomeProject.html.

Table. 1 shows that we manage to successfully recognize primitive states (e.g. 'Close phone': 85.7%, 'In coffee Corner': 98%) with a low false detection rate. By avoiding miss detections of primitive states and using a flexible event description, the proposed system recognizes the complex events with a recognition rate about 58% for the 'MatchingSheetsActivity' event and 100% for the event 'Move from reading zone to coffeeCorner'(Tab. 1).

The comparison (Table 2) shows that the recognition rate of the complex event MatchingSheetssActivity in the case of the proposed algorithm (58%) is higher than the deterministic algorithm [25] (38%). This can be explained by the fact that the deterministic algorithm fails to recognize some primitive states because the person was not correctly detected. However, the proposed algorithm manages to recognize the primitive state and as a consequence the complex event. It can also be explained by the fact that the algorithm [25] does not manage the loss of tracking identifier which can deeply affect the recognition of long-term events even though its sub-events are detected.

We have evaluated also our approach on the 'Building Entrance' real world videos of ETISEO [15] dataset (fig.4). We have selected this public dataset because it has been used previously in several work [10].

6. Conclusions

We have proposed a description-based event recognition approach to describe and recognize video activities. The proposed approach allows flexible modeling of activities and manages the uncertainty of recognition. We have detailed the conditional probability estimation of activities as a Bayesian process. We have presented how we manage

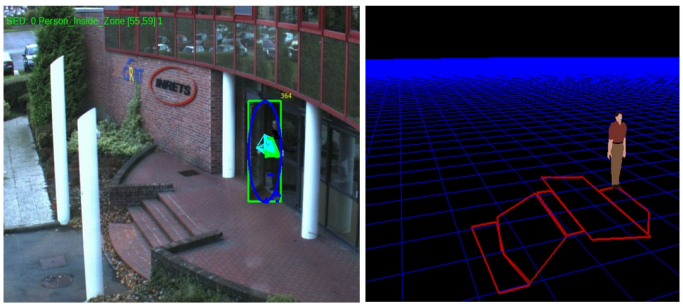

Figure 4. Illustration of the Etiseo activity detection.

Algorithm	*GT*	% R	FP	FN
Approach [25] MatchingSheetssActivity	31	38	0	19
Proposed algorithm MatchingSheetssActivity	31	58	1	13

Table 3. Comparison of recognition rate (% R), the false positive (FP) and the false negative (FN) of the proposed algorithm with the state of the art algorithm. The ground truth GT correspond to 31 videos sequences ($12(\pm 5)$ min, 8 fps)

low level uncertainty at the level of event modeling and at the level of event recognition. The proposed approach recognizes successfully activities of real world videos. We have finally compared our approach with the state of the art showing how the fexible modeling and the probabilistic reasoning can improve the results of real-time event recognition.

References

[1] J. Allen. Maintaining knowledge about temporal intervals. In *In Communications of the ACM*, 1983.

[2] W. Brendel, A. Fern, and S. Todorovic. Probabilistic Event Logic for Interval-Based Event Recognition. In *CVPR*, pages 3329–3336, 2011.

[3] D. P. chau, F. bremond, and M. thonnat. A multi-feature tracking algorithm enabling adaptation to context variations. In *The International Conference on Imaging for Crime Detection and Prevention (ICDP)*, November 2011.

[4] C. L. Chiang, C. C. Lien, and C. H. Lee. Scene-based event detection for baseball videos. In *Journal of Visual Communication and Image Representation*, pages 1–14, February 2007.

[5] V. Delaitre, D. Fouhey, I. Laptev, J. Sivic, A. Gupta, and A. A. Efros. Scene semantics from long-term observation of people. In *ECCV*, 2012.

[6] T. Duong, H. Bui, D. Phung, and S. Venkatesh. Activity recognition and abnormality detection with the switching hidden semi-markov model. In *CVPR*, 2005.

[7] D. F. Fouhey, V. Delaitre, A. Gupta, A. Efros, I. Laptev, and J. Sivic. People Watching: Human Actions as a Cue for Single-View Geometry. In *ECCV*, 2012.

[8] P. Kumar, S. Ranganath, H. Weimin, and K. Sengupta. Framework for real-time behavior interpretation from traffic video. In *IEEE Trans. on Intelligent Transportation Systems*, volume 6, pages 43–53, 2005.

[9] S. Kwak, B. Han, and J. H. Han. Scenario-Based Video Event Recognition by Constraint Flow. In *CVPR*, pages 3345–3352, 2011.

[10] G. Lavee, R. Michael, and R. Ehud. Propagating Uncertainty in Petri Nets for activity Recognition. In *ISVC*, 2010.

[11] F. Lv, X. Song, V. Wu, B. Kumar, and R. Nevatia. Left luggage detection using bayesian inference. In *Proc. of IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance*, pages 83–90, 2006.

[12] D. Minnen, I. Essa, and T. Starner. Expectation grammars: Leveraging high-level expectations for activity recognition. In *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 626–632, 2003.

[13] R. Nevatia, S. Hongeng, and F. Bremond. Video-based event recognition : activity representation and probabilistic recognition methods. In *Computer Vision and Image Understanding*, volume 2, pages 129–162, 2004.

[14] R. Nevatia, T. Zhao, and S. Hongeng. Hierarchical language-based representation of events in video streams. In *In IEEE Workshop on Event Mining*, 2003.

[15] F. Nghiem, A.T.and Bremond, M. Thonnat, and V. Valentin. ETISEO, performance evaluation for video surveillance systems. In *AVSS 2007*, 2007.

[16] N. Oliver, E. Horvitz, and Garg. Layered representations for human activity recognition. In *In IEEE International Conference on Multimodal Interfaces (ICMI)*, pages 3–8, 2002.

[17] S. Park and J. Aggarwal. A Hierarchical Bayesian Network for Event Recognition of Human Actions and Interactions. 2:164–179, 2004.

[18] S. Reddy, Y. Gal, and S. Shieber. Recognition of Users Activities Using Constraint Satisfaction. In *Springer Berlin / Heidelberg*, pages 415–421, 2009.

[19] R. Romdhane, F. Bremond, , and M. Thonnat. Uncertainty for Complex Event Recognition. In *AVSS*, 2010.

[20] R. Romdhane, F. Bremond, and M. Thonnat. Probabilistic Recognition of Complex Events. In *ICVS*, 2011.

[21] M. S. Ryoo and J. K. Aggarwal. Recognition of composite human activities through context-free grammar based representation. In *CVPR*, pages 1709–1718, 2006.

[22] M. S. Ryoo and J. K. Aggarwal. Semantic representation and recognition of continued and recursive human activities. In *International Journal of Computer Vision (IJCV)*, pages 1–24, 2009.

[23] M. S. Ryoo and J. K. Aggarwal. Stochastic Representation and Recognition of High-Level Group Activities. In *International journal of computer Vision*, 2010.

[24] S. Tran and L. S. Davis. Event modeling and recognition using Markov logic networks. In *European Conference on Computer Vision, ECCV 08*, octobre 2008.

[25] T. Vu, F. Bremond, and M. Thonnat. A Novel Algorithm for Temporal Scenario Recognition. In *The Eighteenth International Joint Conference on Artificial Intelligence*, 2003.