We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Cognitive and Computational Neuroscience: Principles, Algorithms, and Applications in Surveillance Context

Lozada Torres Edwin Fabricio,
Martínez Campaña Carlos Eduardo and
Gómez Alvarado Héctor Fernando

Additional information is available at the end of the chapter

Abstract

Today, working with human behavior is vitally important, especially if we consider the impact neuroscience and security systems. The responsibility of monitoring in a conventional way is in charge of a human agent (vigilant). On the other hand, a vigilant cannot be aware at all times. He can only be aware for 20 minutes which is the time he can monitor four cameras simultaneously; after that, the task of surveillance ceases to make sense. This reveals one of the shortcomings of surveillance (SV) systems. Whether a surveillance system provides a warning of an activity or situation makes it as important as the selection of the technological elements that allowed it to be captured. Security systems based on intelligent technologies have had an accelerated development in recent times detection and identification of car registration numbers, detection of static objects in tracks, and detection of pedestrians circulating on not permitted routes. The reuse of methodologies, procedures, and ontologies is described in this chapter of the book.

Keywords: cognitive, surveillance, applications, algorithms, behavior

1. Introduction

The analyst of sequences of video is a very important topic at the time of using this strategy for surveillance of security places. The problem consists in understand any success of the real life saved in a video camera; this process for recognized persons, objects, vehicles, danger places, alarms, etc. is the principal goal of this work [1]. Use the term monitoring to conceptualize the process of collection and selection of activities according to the relevance of the situation that needs to be identified. This process is part from the monitoring of signals or images that allow

characterization of the situation of interest. Thus, the general objective of this monitoring is to identify suspicious situations in order to ensure that activities and situations are normal, so it informs about possible abnormalities that may occur [2].

Ubiquitous computing also supports the development of surveillance (SV) systems through the recognition of activities of the high semantic level based on multisensory monitoring which goes from the capture to the interpretation of the signal. Through the different stages, the abstracted information provided by the sensors is associated with the activities that happen on the stage—person watching television, person takes their medicine, and person calls by phone. In general, this type of multisensory recognition has been applied to infer activities of people's daily lives [3]. The analysis of multisensory information requires a high degree of abstraction from the low semantic level that does not produce the necessary details to understand the detail of what happened in a scenario. This semantic gap is clearly identified in SV systems that process multisensory signals as they pass directly from the sensory signal to interpret the situation. This interpretation always depends on the knowledge, the expression capacity, and the specific language of the scorer. Some researches propose solutions to eliminate the semantic gap. Most of them are based on structures that start from the low semantic level to obtain a high level that allows quality descriptions that help in search and recovery of activities and situations in the SV systems [4, 5]. In the next subchapter, we explain the solutions to SV based on cognitive neurosciences.

2. Symda project

In spite of all the research efforts, it has not been possible to integrate the SV systems in a single functional structure. This is an idea that would allow improvements in the interpretation of situations in a scenario. There are architectures that group multisensory systems in order to help the human operator to make decisions according to identifying a situation of interest, a theme must be developed from the technological combination. With this topic the SIMDA group has carried out projects that propose the integration of different technologies and the semantic conceptualization of situations [4]: AVANZA, CICYT 2004, CYCYT 2007, and INT3. The contribution of INT3 is fundamental to this work, as they obtained from Horus, a multisensory framework for monitoring and detecting activities, integrating multisensory systems into a single processing unit, as shown in **Figure 1**.

As shown in **Figure 1**, Horus is a modular architecture for management of multisensory inputs, incorporating a model of conceptualization that allows sharing information of interest among multiple scenarios. Multisensitive sources are mainly related to image sensors, since they are the most widespread for monitoring tasks although other technological sensors, such as wireless sensor networks (WSN), are also integrated into INT3-Horus generic objects. The framework is distributed and hybrid. The remote nodes perform not only the lower level processing but also data acquisition, while a central node is responsible for collecting the information and its coalescence. The proposal consists in the identification of things and the monitoring of human behavior [5]. This is eminently complex since there are multiple objects and types of behavior. The model has input and output interfaces, which allow reuse
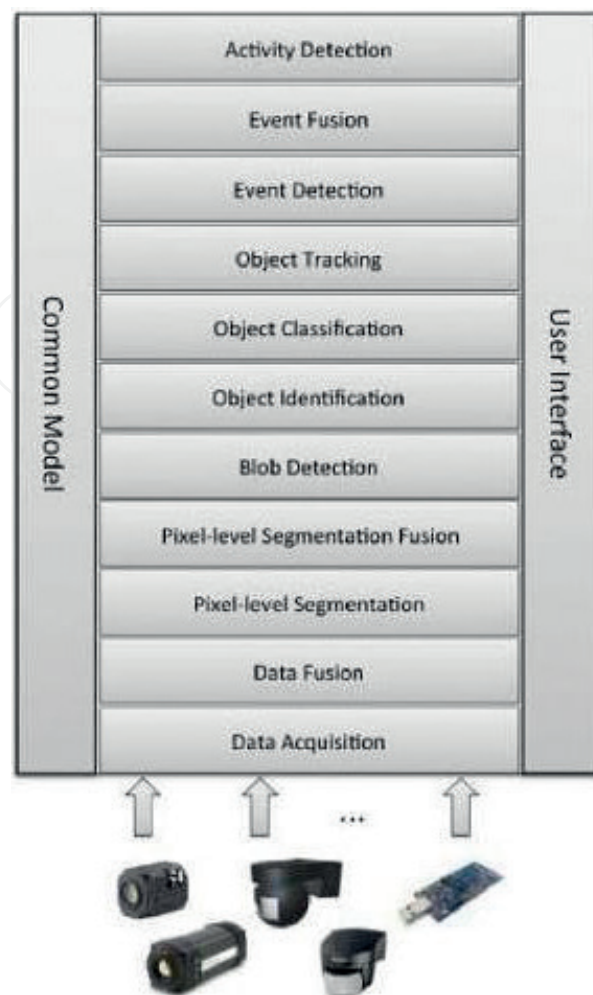
Figure 1. Horus modular system (Source: Castillo et al. (2012)).

and adaptability, all based on the security model in video surveillance systems. For this it is necessary to obtain elements such as entities, procedures, and the relationships between them. In event-based systems, the MVC provides information about changes in the application and provides a representation that adapts to the needs of the user. The model receives inputs to the application and interacts with it to update the objects and to represent the new information [5]. We propose to work with knowledge structures, which collect generalities and particularities of the situations of interest in order to automatically identify them in the monitored scenarios [6, 7]. During the last decade, the ontologies are used in applications for the areas of natural language processing, e-commerce, intelligent information integration, information consultation, database integration, bioinformatics, education, and semantic web, among others. These ontologies provide a vocabulary and organization of concepts that represent a conceptual framework for the analysis, discussion, or consultation of information from a scenario. But, there is a need to perform reasoning tasks which modules or tools must be integrated into a single conceptual, methodological, and technological framework. These modules must be coupled to Horus framework, in order to infer activities of the high semantic level (see **Figure 2**).
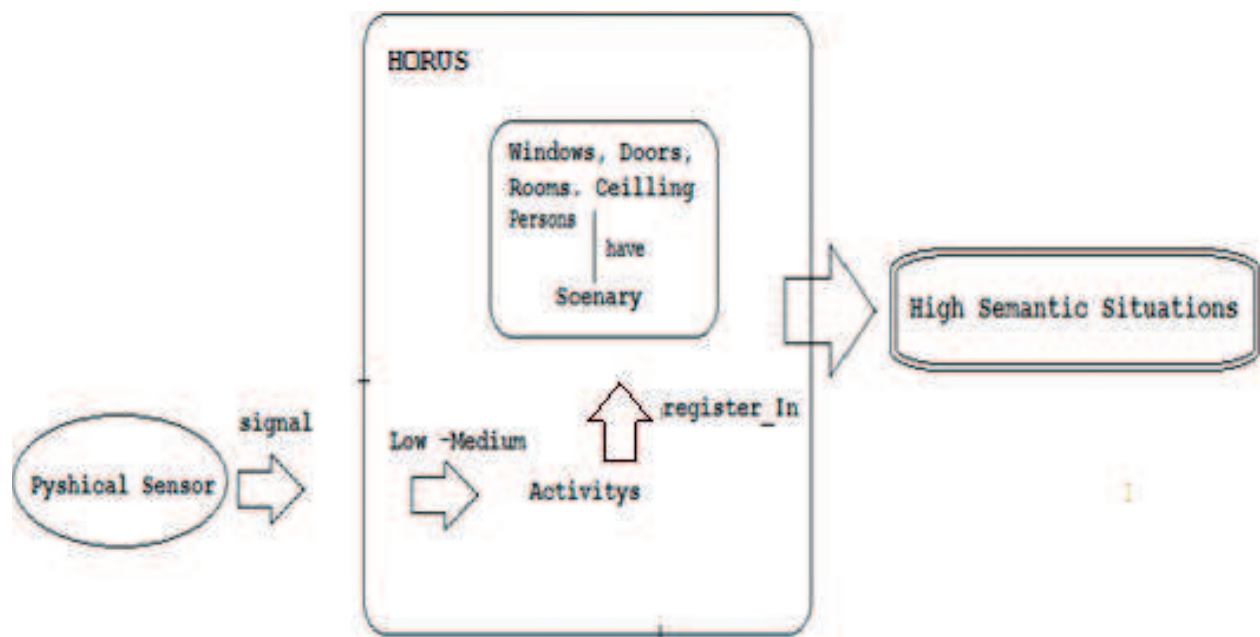
Figure 2. Semantic model: the knowledge structure is used to model scenarios, activities, and situations.

The hypothesis to be verified is that, through the design of ontologies and semantic technologies easy to use, reuse and modulate it can inferre situations of high semantic level and the rapid prototype of SV systems with a similar level of abstraction that a human agent has. To verify our hypothesis, ontology must fulfill the following aspects:

1. It is a semantic multisensory referential framework, which reduces the difficulty of working with different types of signals from different sensors that together with the semantic conceptualization of the signal allows to obtain the latest in the appropriate characteristics of the behaviors and activities developed in the context under study.

2. Systems based on ontologies conceptualize the information that comes from the case in the investigation. Applying this theory to video surveillance systems, it is possible through semantics to infer and have retrospective analysis, that is, perform the activities even after they have occurred.

3. Import ontologies. It adapts its structure to be combined with other ontologies developed in different domains. This is in order to reuse representations of knowledge in different areas of science.

4. Conceptualize and infer activities. It is the process in which the knowledge of the expert is used to conceptualize activities, or rules or axioms for inference are applied to the activities that are recorded in the scenario. We are tasked with inferring activities of the medium and high semantic level, while Horus is tasked with the low semantic level.

5. Conceptualize and infer situations. It is to allow the vigilant or expert to establish the relationships between the activities that happen in it and to design rules and semantic axioms

to infer a situation. In an alternate way, it is to have the semantic capacity needed to adapt its structure to the appearance of new conceptualizations of activities and situations as a product of learning acquired by computer algorithms. Based on the interest of knowledge situations, there are two types of tasks: (a) conceptualize and model the knowledge of the human expert, when it exists (depending on the basic activities recognizable from the sensors or the processing of video images), and (b) conceptualize and model situations where the knowledge does not exist although it is possible to find it in records (case bases) of the situations that are intended to be identified. In this case, the required process is particularly complex [8, 9], and it will require the use of intelligent algorithms for identification. Literals (a) and (b) are studied in this research, since we work with the knowledge of the expert when he can clearly describe the scenarios and situations, in addition to scenarios where there is some expert knowledge which is not precise and is intended to find automatically the situations of interest. Here, situations are composed for activities that individually are not clearly suspicious, but when analyzed in a certain sequence and repetition, they do reflect to be.

The use of SV systems based on closed-circuit television (CCTV) cameras has grown exponentially over the last decade. Especially, concern for security as a result of emerging international terrorism has led experts to anticipate a greater diffusion of these systems as well as their integration into a global remote monitoring network [10]. Analyzes the latest advances in the multisensory SV systems use by companies that produce this type of technology with an emphasis on their manufacturing, added value, other products differences, and its use. This analysis focuses on the SV systems based on cameras and sensors for surveillance. The result of the analysis allows answers to questions such as the following: Would it be useful to be able to track people in different areas and places? Is it possible to check for false alarms in establishments or simply monitor a trade from the comfort of home?

Applications of this type are already commercially available, allowing access from a single control center to images of CCTV systems in various geographically distributed environments. For example, the synchronized video acquisition system is developed to interoperate with SV systems in order to act as an object trajectory server. It consists of a series of navigational instruments that allow the direct geo-referencing of each of the images captured by the video camera in post-processing in a common reference time for all navigation instruments and for all sensors used in the capture system video. Remote sensing allows to have information about an object or surface through the analysis and processing of the data supplied by different sensors that are synchronized. In addition, it associates the time and Global Positioning System (GPS) with the image generated by the video [7]. As a result systems are capable of analyzing the video of different subsystems and interpreting what happens in the images. Applications of examples that can have video acquisition and synchronization systems are real-time monitoring of traffic conditions, forest fire control, natural disaster monitoring, and geo-referenced video projection in public virtual machines such as Google Earth and Virtual Visualizations [7].

3. Ontology and agents

Analyze surveillance system based on cameras which have been done by using three different methods: the first one has been developed using an expert knowledge, the second one was learned from recorded videos, and the third one has been developed as a refinement taking into account evaluation with ground truth. This project was deployed in Madrid-Barajas airport; this technology is used to support ground traffic management inside the Advanced Surface Movement. Model-Based Reasoning (MBR) modeling of semantic reasoning allows the resolution of problems in the identification of activities in space and time. This is a basic theme of the same, since the temporal management must be strictly linked to the occurrence of the facts. However, there are still two fundamental problems in this application: the degree of dependency between the model use and the domain and the reutilization on the systems when the domains change. The implementation aim is to help solving the problem of temporal diagnosis for environments of high conceptual complexity integrating MBR and ontologies for domain knowledge representation. A traditional system of security and vigilance is that the caretaker is alert toward what happens in the zone which needs security. In this kind of system, the quality of the vigilance has a direct relationship with the human capacities of the caretaker, which are incremented with the use of security cameras, motion sensors, etc. A minimum requirement for security systems is the ability to analyze multiple objects or groups of objects in real time [8, 11]. The main objective of this study is to calculate some parameters for the performance evaluation of the tracking system in order to identify an alarm human behavior. Here, it is necessary—the problem—to consider a sequence of previously recorded videos as well as subsequent processes in which a human operator takes notes of the images and places marked on each video frame. Using ontology with agents for running the system, we proposed the next system (**Figure 3**).

The first step only uses the two classes in the CARETAKER ONTOLOGY (**Figure 2**), person and objects, to determinate how the system function and learn from this. Without using the speak recognition and write recognition (see Crubèzy, Connor, Buckeridge, Pincus, Musen: Ontology-Centered Syndromic Surveillance for Bioterrorism.), the structure is similar at the bioterrorism ontology (**Figure 4**).

The CARETAKER ONTOLOGY has other classes and properties (**Figure 5**).

Now, in this job selection, the classes are person and object. To use these classes, it is necessary that the image catcher for the sensors at the office can be s processed to recognize people, objects, and activities. In CARETAKER it is possible to see these elements for the ontology (see A Real-Time Scene Understanding System for Airport Apron Monitoring: AVITRACK Project): the AVITRACK project has the same structure for the ontology.

Scene analysis terminology:

Scene: Place of development of activities, events, and activities.

Physical object and interest: Thing in tracking or tracking that is in the study area. These objects allow their relationship to obtain the contextual object, since this occurs after the relationship between physical objects. The movement of this type of objects can be random and unprovoked.

Contextual object: The object of a scene conditioned to the appearance of activities, events, and activities.

Objective tracked: The object of interest located in the area of interest, directly related to semantic tracking.

ROI: Region of interest. The context, area, or region of interest.

Oscillating movement: Constant movement related to the fact, it can be provoked or not.

After this work using the CAREGIVER:

We have proposed the syntax to describe states, events, and activities. These meta-concepts are described with a name and four parts:

Physical objects: Semantically related produce facts.

Components: What allows to obtain a description of a context.

Prohibited components: That does not correspond to the scene, activity, or context.

Restrictions: Relationships between the concepts that allow obtaining basic characteristics of the scene.
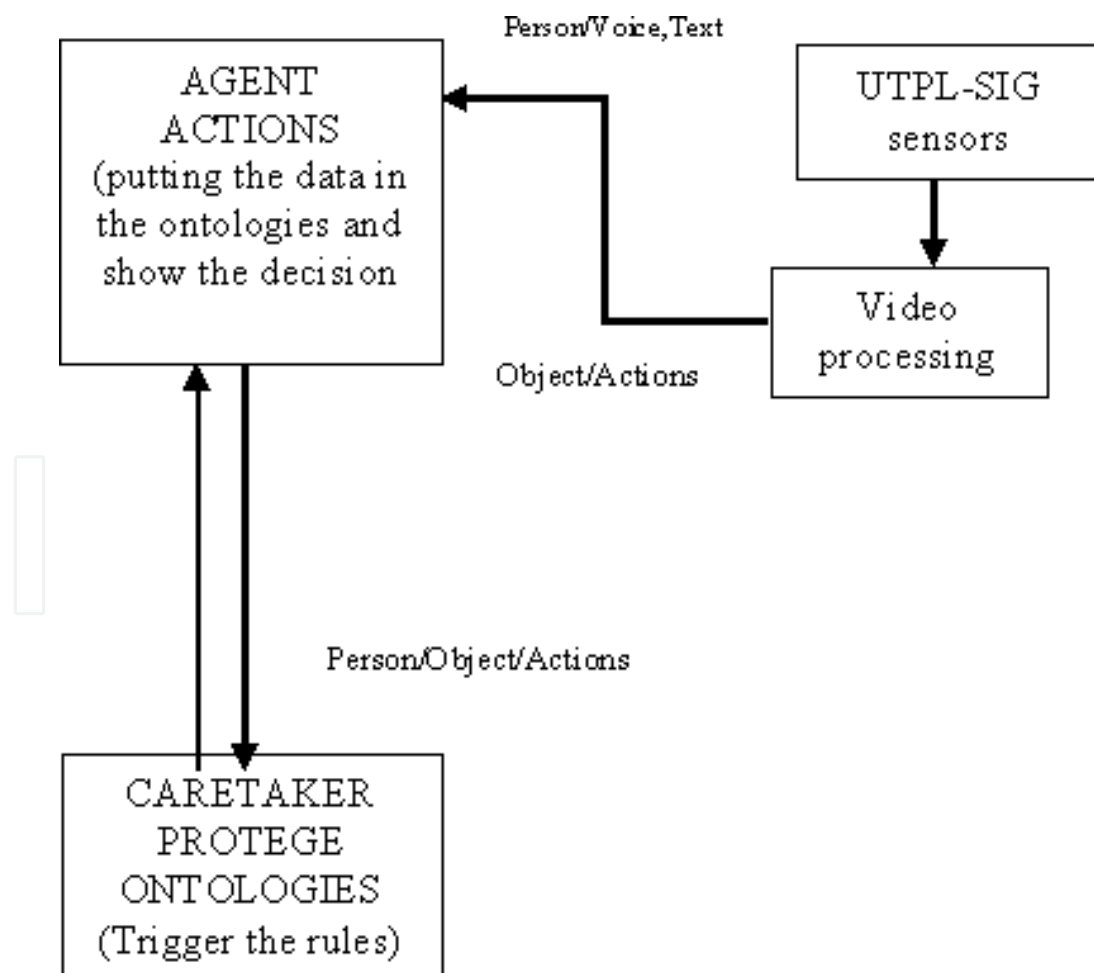
Figure 3. The sensors detecting an object or person, after the ontology using this information to determinate the actions and answer with the rules.
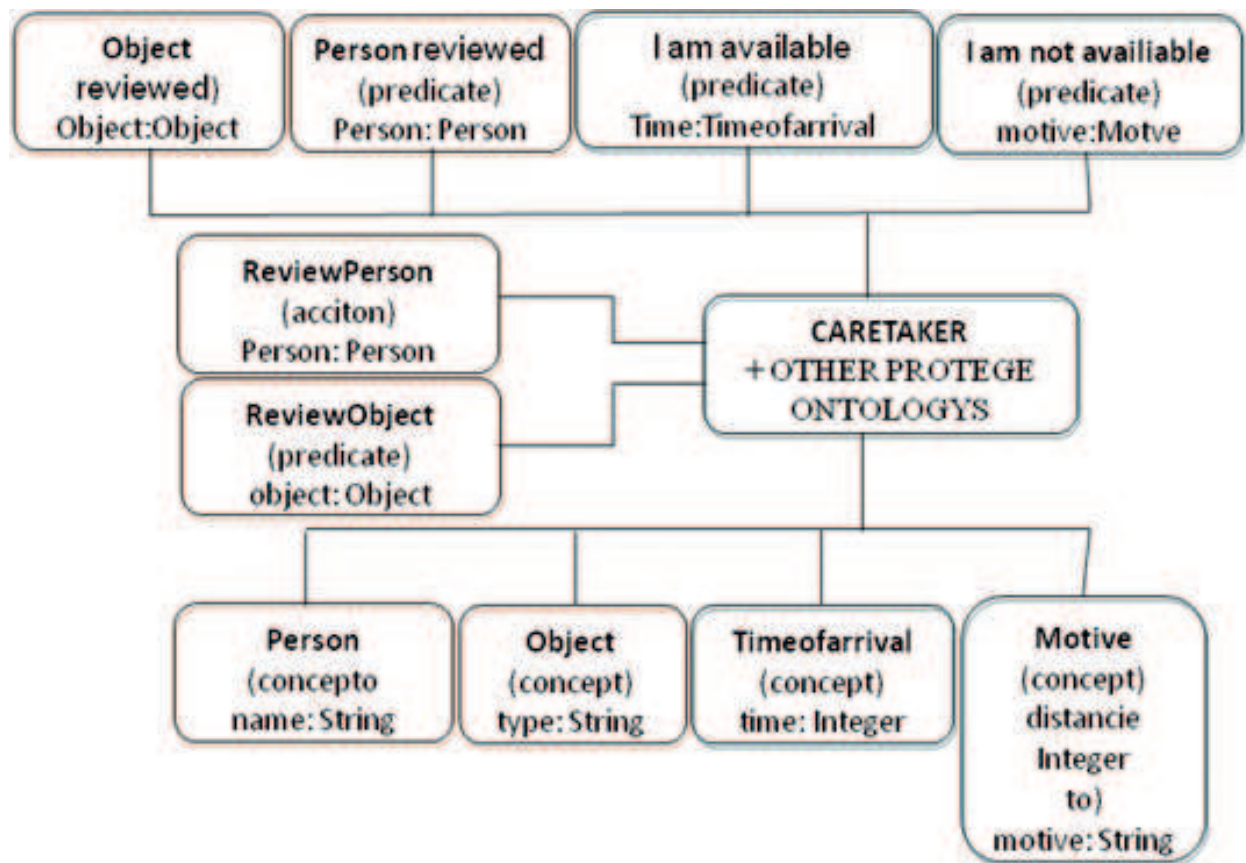
Figure 4. Our CARETAKER ONTOLOGY.

Description of a model and an associated instance, respectively, of a primitive state, a primitive event, and an event composed of multiple agents (see (CAREGIVER)). The code to conceptualization:

package surveillanceontology;

import jade.content.Concept;

public class Persona implements Concept {

  private String nombres;

  public String getNombres() {

    return nombres;

  }

  public void setNombres(String n) {

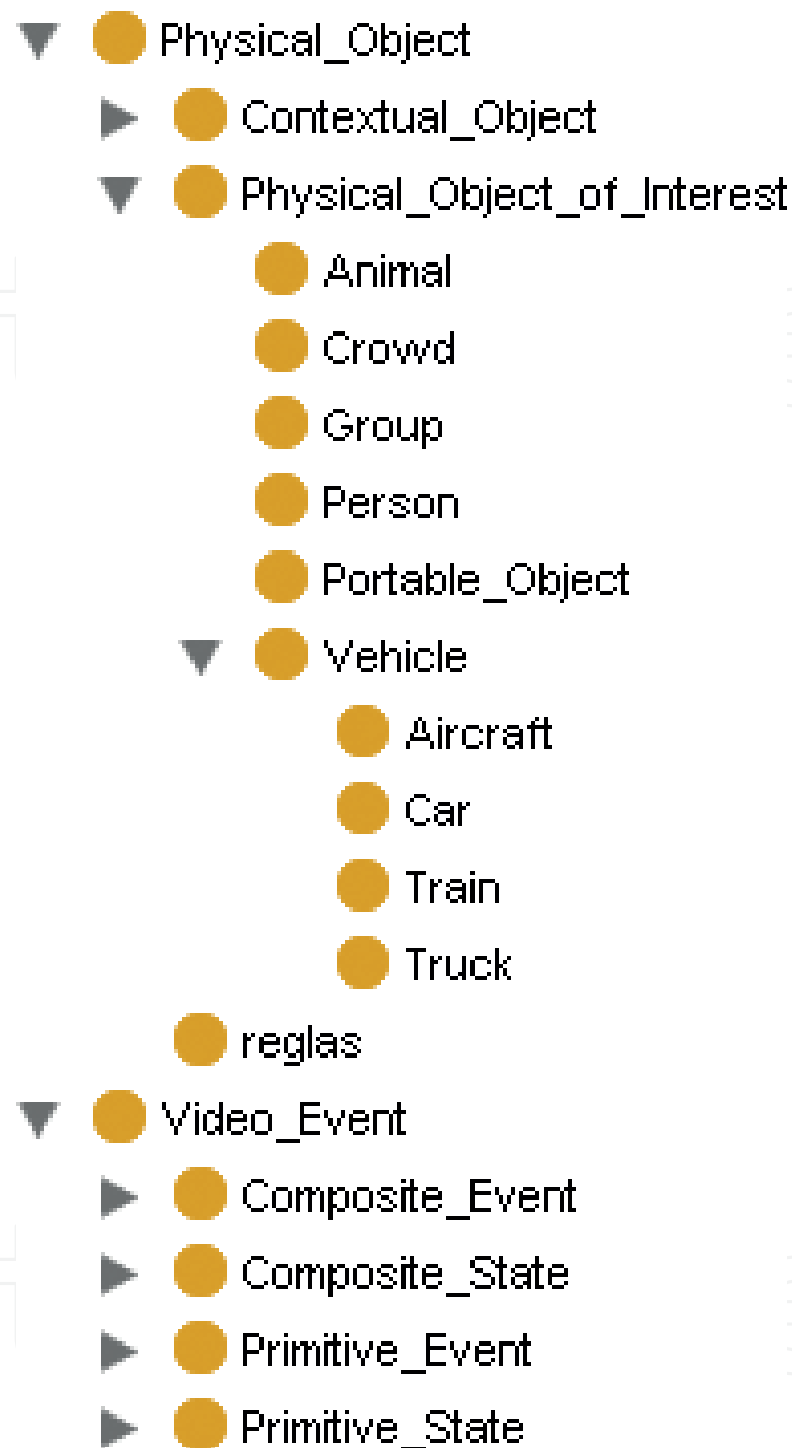   nombres = n;

  }

}

Figure 5. CARETAKER ONTOLOGY.

PrimitiveState_Inside_zone

PhysicalObjects_

(p: Person, z: Zone)

Constraints:

(p in z)

Instance:

S1: **PrimitiveState_Inside_zone_**(Hector, ZonaProhibida)

CompositeEvent SigOfficeEntrance

PhysicalObjects:

(e : Persona[worker], r: Persona**[Hector]**

Components:

((c1: PrimitiveState Inside_zone (e, "Back_Counter"))

 (c2: PrimitiveEvent Changes_zone(r, "Entrance_Zone", "Front_Count")

 (c3: PrimitiveState Inside_zone (e,"Safe"))

 (c4: PrimitiveState Inside_zone(r,"Safe")))

Constraints:

((duration-of(c3) >= 1 second)

 (c2 during c1)

 (c2 before c3)

 (c1 before c3)

 (c2 before c4)

 (c4 during c3))

Instance:

e2: CompositeEvent SigOfficeEntrance

Ontology has two uses: (a) semantics for the occurrence of events, through states and events with little granularity in which the occurrence of physical events is highlighted. Here, we describe the attributes of each thing and the relationships between them, in order to obtain a clear description of the scene under study. The levels of implementation allow the use of the restrictions based on time and space, which means that the states and events that occur in the video can have their location and semantic description. This basic corpus can be refined depending on the needs. Refined concepts are more difficult to extract from videos. For example, they may need posture analysis algorithms. The concept "holding an object" is perceived differently depending on the posture but also in the properties of the held object. Holding a gun is perceptually different from holding a luggage.

The proposed corpus should be seen as an extendable basis. The issue is now to define tools and protocols to allow a collaborative extension of the corpus.

4. Alarm-agent-caretaker

This job uses the Protégé software for using the CARETAKER ontologies and it configurates the ontology for our benefit. At the time, it is necessary to transform the classes in the CARETAKER ONTOLOGY in java classes using the BeanGenerator plugging for this process. This is important because the agents use the ontology in java code when the alarm generates the agent. Using this generation puts the data in the java CARETAKER ONTOLOGY, and this ontology returns as the decisions for the agent. For this reason, communication takes place in two ways: the first time from sensor-agent-ontology and the second after ontology agent.

5. The style

When the sensor catches a person/object/action, the system uses this steps to determinate if it is a person, object, or action. After these, data is sending at the CARETAKER ONTOLOGY for taking decisions. The equation for the image binary is:

$$img(i, j) = \begin{cases} imgG(i, j) > umbral\ entonces\ img(i, j) = 1 \\ \\ imgG(i, j) < umbral\ entonces\ img(i, j) = 0 \end{cases} \tag{1}$$
$$imgN(i, j) = 1 - img(i, j)$$

On the other side, detecting movement is necessary to establish the difference between the bottom image (principal image) and the image caught by sensors. Using this equation for knowing the difference

$$suma = \sum_m \sum_n (resta_{mn}) \tag{2}$$

This methodology was used in this job to recognize these activities:

1. Person in the office

2. Persons in the office

3. Objects in the desk

In this moment the system determinates if the person is in the office and if an object was subtracted from the office or another action in the scene. Good, the system has the words person and object in a file with the name saved; for example, in **Figure 3** the system recognizes the people, and this image is labeling with the word person recognition. The process is the same for object and actions.

Well, when the systems recognize a person in the class (CARETAKER ONTOLOGY), it is an instance because it is necessary to identify if the person has the permission for this site or not. In this moment the alarm is generated: the system recognizes a person and the rule. If the

person is in the office when the time is higher at 21:h00, then it calls the security in charge to review the person in the place. Security is subclass from person class.

The action reviews if the person is in the ontology due to the connection between person recognition and the decisions after the alarm. At the same time, when the security person reviews at people recognized in the site, the message will come back before reviewing the person.

A difficult remaining problem is the segmentation process. Indeed, in order to be classified, images have to be segmented to allow descriptor computation. The symbolic description made by the expert may help finding the image processing tasks required for extracting the pertinent information from the provided images.

As an example, an object described with the "granulated texture" concept may be segmented with a texture-based segmentation algorithm. The regions of interest selected by the expert (see 1) in this work use the img (i, j) to correct this problem.

The system uses the natural language for decisions: the agents are programmed using JADE (Java Agent Development Framework) and the natural language coded to communicate the agents:

Jade enterprise

Alarmado (generating the alarm)

Security central (receive the alarm and using the ontology for making a decisions)

On the other side, the system uses the file saved to communicate the sensors with the alarm agent. This is important because the alarm has code lines to review each 10 seconds if the file.txt has the word: person recognition, object recognition, or actions after producing the after process.

Author details

Lozada Torres Edwin Fabricio[1], Martínez Campaña Carlos Eduardo[1] and Gómez Alvarado Héctor Fernando[2]*

*Address all correspondence to: hf.gomez@uta.edu.ec

1 Universidad Regional Autónoma de los Andes, Ambato, Ecuador

2 Universidad Técnica de Ambato, Huachi Chico, Ambato, Ecuador

References

[1] Collins RT, Lipton AJ, Kanade T. A System for Video Surveillance and Monitoring. CiteSeeer; 2000. pp. 50-62

[2] Waters S. How to Identify Shoplifters [En línea]. 27 Octubre 2016. Available from: https://www.thebalance.com/how-to-identify-shoplifters-2890263 [Último acceso: Marzo 20, 2013]

[3] Chikhaoui B, Wang S, Pigot H. A new algorithm based on sequential pattern mining for person identification in ubiquitous environments. In: Proceedings of the 4th International Workshop on Knowledge Discovery form Sensor Data (ACM SensorKDD '10); 2010. pp. 20-28

[4] Simon F, Zhuang Y. A security model for detection suspicious patterns in physical environment. In: Third International Symposium of Information Assurance and Security; 2007. pp. 221-226

[5] Castillo JC, Fernandez-Caballero A, Loopez MT. A review on intelligent monitoring and activity interpretation. Revista Iberoamericana de Inteligencia Artificial. 2017;**20**(59):53-69

[6] Bredmond J, Corvee E, Patiño J, Thonnat M. CARETAKER Project [En línea]. 2008. Available from: http://www-sop.inria.fr/members/Francois.Bremond/topicsText/caretakerProject.html [Último acceso: Julio 8, 2011]

[7] Town C. Ontological inference for image and video analysis. Machine Vision and Applications. 2006;**17**:94-115

[8] Sunico V. Reconocimiento de Imágenes: Usuarios, Segmentos de Usuarios, Gestos, Emociones [En línea]. 2008. Available from: http://jm.sunico.org/4007/06/48/reconocimiento-de-imagenes-usuarios-segmentos-de-usuarios-gestos-emociones-y-empatia/ [Último acceso: Septiembre 21, 2010]

[9] Hu W, Wang L, Maybank S. Survey on visual surveillance of object motion and behaviors. IEEE Transactions on Systems, Man and Cybernetics, Part C. 2004;**34**(3):334-354

[10] Honovich J. IPSurveillance [En línea]. 2010. Available from: http://ipvideomarket.info/ [Último acceso: Noviembre 21, 2010]

[11] Albusac Jiménez JA. Vigilancia Inteligente: Modelado de Entornos Reales e Interpretación de Conductas para la Seguridad. España: Castilla-La Mancha; 2008