Optical flow or image subtraction in human detection from infrared camera on mobile robot

Antonio Fernández-Caballero [a,b,*], José Carlos Castillo [a], Javier Martínez-Cantos [c], Rafael Martínez-Tomás [d]

[a] *Universidad de Castilla-La Mancha, Instituto de Investigación en Informática de Albacete (I3A), 02071-Albacete, Spain*
[b] *Universidad de Castilla-La Mancha, Escuela de Ingenieros Industriales de Albacete, Departamento de Sistemas Informáticos, 02071-Albacete, Spain*
[c] *MoviRobotics S.L., Parque Científico y Tecnológico de Albacete, Paseo de la Innovación 1, 02006-Albacete, Spain*
[d] *Universidad Nacional de Educación a Distancia, Departamento de Inteligencia Artificial, 28040-Madrid, Spain*

ARTICLE INFO

ABSTRACT

Perceiving the environment is crucial in any application related to mobile robotics research. In this paper, a new approach to real-time human detection through processing video captured by a thermal infrared camera mounted on the autonomous mobile platform mSecurit^{TM} is introduced. The approach starts with a phase of static analysis for the detection of human candidates through some classical image processing techniques such as image normalization and thresholding. Then, the proposal starts a dynamic image analysis phase based in optical flow or image difference. Optical flow is used when the robot is moving, whilst image difference is the preferred method when the mobile platform is still. The results of both phases are compared to enhance the human segmentation by infrared camera. Indeed, optical flow or image difference will emphasize the foreground hot spot areas obtained at the initial human candidates' detection.

1. Introduction

Perceiving the environment is crucial in any application related to mobile robotics research [1,2]. The information surrounding the robot can be used to navigate, avoid barriers and execute a given mission [3]. As an outstanding motion detection method, optical flow is being widely used in mobile robot navigation. Optical flow plays a central role in the ability of primates to recognize movement. Image flow divergence has been used to orient a robot within indoor hallways by estimating time to collision [4] and differences in optical flow magnitude have been used to classify objects moving at different speeds to simplify road navigation in traffic [5]. Optic flow has also been used in a dynamical model of visually guided steering, obstacle avoidance and route selection [6]. An approach uses optical flow information to track features on objects from the time they appear on screen until they interact with the local sensors of the robot [7].

Moreover, the use of autonomous robots or vehicles can provide significant benefits in the surveillance field [8,9]. And, many algorithms focusing specifically on the thermal domain have been explored. The thermal-based algorithms are inspired in the biological processes of many animals and insects, which are affected by the presence of thermal energy in their environment [10]. Indeed, diverse types of thermoreceptors are found in nature, which aid animals and insects in hunting, feeding and survival [11]. Now, in computer vision, the unifying assumption in most methods is the belief that the objects of interest are warmer than their surroundings [12]. Indeed, some animals can see in total darkness, or even see colors beyond the visual spectrum, that humans have never seen. Thermal infrared video cameras detect relative differences in the amount of thermal energy emitted/reflected from objects in the scene. As long as the thermal properties of a foreground object are slightly different (higher or lower) from the background radiation, the corresponding region in a thermal image appears at a contrast from the environment. In [13,14], a thresholded thermal image forms the first stage of processing after which methods for pose estimation and gait analysis are explored. In [15], a simple intensity threshold is employed and followed by a probabilistic template. A similar approach using Support Vector Machines is reported in [16]. Recently, a new background subtraction technique to robustly extract foreground objects in thermal video under different environmental conditions has been presented [17]. A recent paper [18] presents a real-time ego-motion estimation scheme that is specifically designed for measuring vehicle motion from a monocular infrared image sequence at night time.

* Corresponding author at: Universidad de Castilla-La Mancha, Instituto de Investigación en Informática de Albacete (I3A), 02071-Albacete, Spain. Tel.: +34 967 599200; fax: +34 967 599224.
E-mail address: caballer@dsi.uclm.es (A. Fernández-Caballero).

Fig. 1. The mSecurit™ mobile surveillance robot.

In the Robotics field, a new type of infrared sensor is described [19]. It is suitable for distance estimation and map building. Another application using low-cost infrared sensors for computing the distance to an unknown planar surface and, at the same time, estimating the material of the surface has been described [20]. Also, another paper [21] presents a vision-based approach for tracking people on a mobile robot using thermal images. The approach combines a particle filter with two alternative measurement models that are suitable for real-time tracking. Very recently, a people tracking system uses a combination of thermal and color information to robustly track persons has been introduced [22]. The thermal camera simplifies the detection problem, which, according to the authors of the paper, is especially difficult on a mobile platform. The system is based on a fast and efficient sample-based tracking method that enables tracking of people in real-time. The measurement model using gradient information from the thermal image is fast to calculate and allows detection of persons under different views.

In this paper, we introduce our approach to real-time human detection through processing video captured by a thermal infrared camera mounted on the indoor autonomous mobile platform mSecurit™ (see Fig. 1) developed by the Spanish private company MoviRobotics S.L. The approach starts with a phase of static analysis (on the current image frame) for the detection of human candidates. Then, a dynamic analysis [23,24] (taking the previous and the current images) by means of an optical flow algorithm based on the Lucas and Kanade without pyramids approach [25], or a subtraction-based approach [26–29], depending on whether the robot is moving or still, is run. The algorithm aligns two images to achieve the best matches and determines motion between both images. The approach assumes the images to be roughly aligned and uses Newton–Raphson iteration for the gradient of error. Lastly, the results of both phases are compared to efficiently segment the humans. The rest of the paper is as follows. Section 2 introduces the most relevant features of the mSecurit™ mobile surveillance robot. Then, the algorithms to detect humans from the infrared camera mounted on the mSecurit™ robot are explained in Section 3. Section 4 offers a complete set of data and results to assess the performance of the human detection algorithms described. Lastly, in Section 5 some conclusions are drawn.

2. Description of the mSecurit™ mobile platform

The mSecurit™ is a mobile platform over four skid steer wheels specially designed for surveillance tasks. The user can remotely control the robot in real-time and manage some platform parameters such as current state, battery level, Wi-Fi link quality and global positioning system (GPS) coverage, among others. Moreover, the user may run some autonomous tasks such as surveillance and navigation. The navigation of the mSecurit™ is mainly based on GPS. The route learning approach consists on remotely controlled "walks". During these, the robot stores points named "known points". Afterwards, the user may choose a path to be followed. The robot is able to patrol the selected path by following its known points. Several ultrasound sensors are installed on the chassis for dynamic obstacle avoidance. Whether a known point is unreachable, e.g. a closed fence or a fallen object in the path, the robot goes to the next known point, and continues patrolling. In addition, mSecurit™ carries a motorized pan-tilt-zoom dome camera, and a thermal infrared one. Furthermore, the robot is able to detect human intruders feeding the algorithms presented in this paper with thermal infrared frames. There is an intelligent switching between image subtraction and optical flow, depending on the platform movement and the processing load. Surveillance can be active on the whole route or only in certain known points.

Human detection methods suffer from false detections. This issue has been studied in order to obtain a robust surveillance system. On our system, the experimental studies show a low false detection rate. So, an alarm is triggered if a human is detected in at least three different close-in-time frames. The frame rate reaches 5–6 frames per second. When an alarm is triggered the robot stops other functions in order to reinforce the evidence. Then, the dome camera performs an exploration pattern while a local video recording of both cameras is performed. At the same time, an alarm is sent to the client application. Afterwards, the robot returns to the state previous to the alarm.

Finally, let us highlight that mSecurit™ can be included in any network surveillance system through a communication protocol. In order to efficiently manage several robots, the user is constantly shown the updated positions with external global positioning software (e.g. Google Earth). In standalone or collaborative modes, the robots cover their routes and are able to go to the known points on demand.

3. Human detection from infrared camera

Prior to describing in detail the proposed human detection algorithm, there are some working conditions that have to be clarified. The mobile platform can be in two states, namely static or mobile. When the images are taken in static state, it is well known that optical flow techniques are not optimal. This is because homogeneous regions in an image sequence provide no cues for optical flow computation. Moreover, the Lucas–Kanade optical flow method used in our approach does not provide optical flow estimation in uniform regions. Some other approaches (e.g. piecewise-smooth dense optical flow [30]) perform dense optical flow computation to solve this issue through extrapolating flow from informative segments to adjacent homogeneous regions. But, this leads to large errors in static homogeneous areas. The Lucas–Kanade approach is selected due to its low computational load and good performance.

On the other hand, segmentation approaches based on image subtraction have been widely studied so far (e.g. [31–33]). Subtraction-based approaches present an important drawback when processing images with a non-static background. This is the case for images captured from a moving robot. Thus, the proposed human detection system selects which operation mode is more suitable for each situation.

Now, the proposed human detection algorithm is explained in detail in the following subsections describing the different phases,

Fig. 2. (a) Input IR image frame. (b) Scaled frame. (c) Closed soft threshold.

Fig. 3. (a) Previous frame $I(r, c, t-1)$. (b) Current frame $I(r, c, t)$. (c) Difference image $I_d(r, c, t)$. (d) Moving objects image $I_\theta(r, c, t)$.

namely, human candidates' blob detection, image motion analysis, and human blob segmentation. In the image motion analysis stage there are two different processing approaches, namely optical flow and image subtraction.

3.1. Human candidates' blob detection

The algorithm starts with the analysis of a single image, $I(r, c, t)$, captured at time t by the infrared camera. This phase uses only the information of the current frame to perform the detection. Firstly, a change in scale, as shown in Eq. (1) is performed. The idea is to normalize all images to always work with a similar scale of values, transforming $I(r, c, t)$ to $I'(r, c, t)$. The normalization assumes a factor γ, empirically calculated as the value corresponding to the mean gray level value of an image, $\bar{I}(t)$, at a day time changing environment temperature (see Fig. 2(b)).

$$I'(r, c, t) = \frac{I(r, c, t) \times \gamma}{\bar{I}(t)} \tag{1}$$

where $I'(r, c, t)$ is the normalized image. Notice that $I'(r, c, t) = I(r, c, t)$ when $\bar{I}(t) = \gamma$.

The next step is the elimination of incandescent points. As the image has been scaled, the threshold θ_i calculated to eliminate these points is related to the normalization factor γ. Indeed,

$$\theta_i = 3 \times \frac{5}{4} \gamma \tag{2}$$

$\delta = \frac{5}{4}\gamma$ introduces a tolerance value of a 25% above the mean image value. And, $3 \times \delta$ provides a value high enough to be considered an incandescent image pixel. Thus, pixels with a higher gray value are discarded and filled up with the mean gray level of the image.

$$I'(r, c, t) = \begin{cases} I'(r, c, t), & \text{if } I'(r, c, t) \leq \theta_i \\ \overline{I'}(t), & \text{otherwise.} \end{cases} \tag{3}$$

The algorithm uses a threshold to perform a binarization for the aim of isolating the human candidates' spots. The threshold θ_h, obtains the image areas containing moderate heat blobs, and,
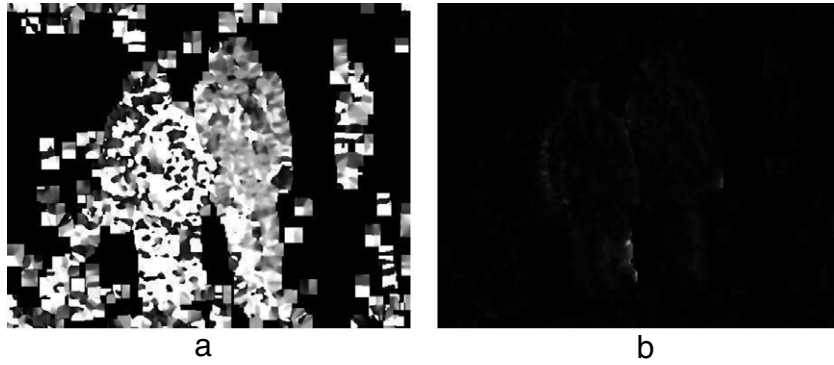
Fig. 4. Optical flow calculation. (a) Angles. (b) Moments.
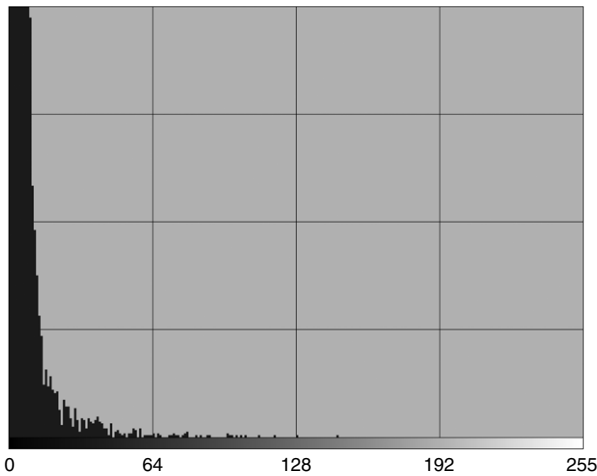

Fig. 5. Moments histogram.

therefore, belonging to human candidates. Thus, warmer zones of the image are isolated where humans could be present. The soft threshold is calculated as:

$$\theta_h = \frac{5}{4}(\gamma + \sigma_{I'}) \qquad (4)$$

where $\sigma_{I'}$ is the standard deviation of image $I'(r, c, t)$. Notice, again, that a tolerance value of a 25% above the sum of the mean image gray level value and the image gray level value standard deviation is offered.

Now, image $I'(r, c, t)$ is binarized using the threshold. Pixels above the threshold are set as maximum value max $= 255$ and pixels below are set as minimum value min $= 0$ (see Fig. 2(c)).

$$B_s(r, c, t) = \begin{cases} \min, & \text{if } I'(r, c, t) \leq \theta_h \\ \max, & \text{otherwise.} \end{cases} \qquad (5)$$

Afterwards, the blobs contained in the image are obtained. A minimum area, B_{min}, – function of the image size – is established for a blob to be considered to contain humans. As a result image $B_r(r, c, t)$ is obtained by eliminating non-human blobs from image $B_c(r, c, t)$.

3.2. Image motion analysis

As aforesaid, this phase is divided into two different processing techniques. If the robot is not moving, a subtraction-based technique is used in order to isolate moving regions that are candidates to contain humans. On the other hand, when the robot is moving, the previous technique is useless. Therefore an optical flow approach able to discard scene movement is taking into account.

3.2.1. Image subtraction

Now, we introduce a quite simple consecutive image difference method, which has to be applied when the mobile platform is completely stopped and the infrared camera is also still. The motion detection by image difference uses two consecutive normalized input images, $I^n(r, c, t - 1)$ and $I^n(r, c, t)$ (see Fig. 3(a) and (b)). A classical image subtraction is performed in order to obtain the movement in the scene (see Eq. (6)).

$$I_d(r, c, t) = |I^n(r, c, t - 1) - I^n(r, c, t)|. \qquad (6)$$

The obtained image, $I_d(r, c, t)$, contains two kinds of spots. On the one hand, there are moving objects at the current (and correct) position and, on the other hand, there are objects detected belonging to the previous position (see Fig. 3(c)). This last kind of spot is called the ghost in image difference. $I_d(r, c, t)$ is binarized in order to remove the residual movement corresponding to the previous position (see Fig. 3(d)). The threshold applied, $\theta = 20$, is empirically obtained from experience to obtain image $I_\theta(r, c, t)$.

Notice that in some cases (e.g. cold weather), "ties" have also to be added into image $I_\theta(r, c, t)$ in order to join possibly disconnected legs and heads.

3.2.2. Optical flow

The dynamic analysis by optical flow requires the calculation of the moments corresponding to each pixel movement on a couple of normalized input images, $I^n(r, c, t - 1)$ and $I^n(r, c, t)$. The optical flow calculation results into two gray level images, where each pixel reflects the angular moment detected, storing the movements in X and Y axes. Firstly, the algorithm performs the speed calculation of the optical flow. The selected optical flow approach is the Lucas–Kanade without pyramids algorithm. The calculated speeds, as a result of the optical flow, are turned into angles, $\alpha(r, c, t)$, and magnitudes, $m(r, c, t)$. Fig. 4(a) shows the direction of the movement (angles). Similarly, Fig. 4(b) shows the magnitudes (moments), that is to say, the amount of movement at each pixel (x, y) between $I^n(r, c, t - 1)$ and $I^n(r, c, t)$, in form of a moments image, $M(r, c, t)$. The results clearly indicate that angles are less important than moments for our purpose. Indeed, on the one hand, non-rigid objects' movements go into very different directions, and, on the other side, angles with low moments may be caused by image noise (See Figs. 4 and 5).

To efficiently use the moments image $M(r, c, t)$, its histogram has been studied for many cases. We came to the conclusion that most values fall into the [0, 64] interval, but very close to 0. Indeed, the average value is close to 1 in these moments images. Therefore, two thresholds, a moments soft threshold $\mu_s = 10$ and a moments hard threshold $\mu_h = 25$, are used to delimit the blobs of possible people. Firstly, to obtain the most representative values, the moments soft threshold μ_s is applied to the moments image $M(r, c, t)$ to obtain image $M_s(r, c, t)$, as shown in Fig. 6(c).
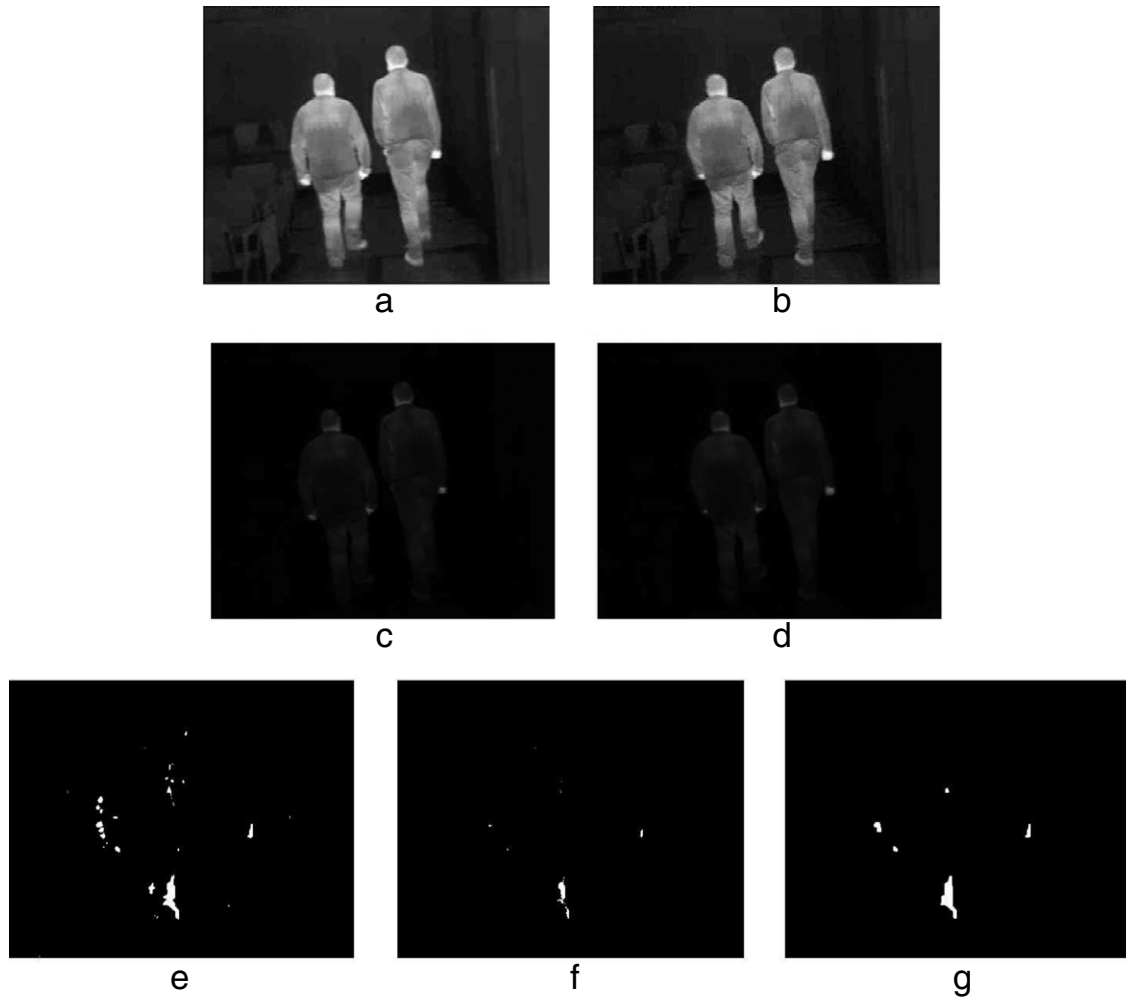
Fig. 6. (a) Previous frame. (b) Current frame. (c) Multiplied previous frame. (d) Multiplied current frame. (e) Soft thresholded moments $M_s(r, c, t)$. (f) Hard thresholded moments $M_h(r, c, t)$. (g) Matched thresholds $M_r(r, c, t)$.

Table 1

Image subtraction results for the mSecurit™ mobile platform images.

	Sequence 1	Sequence 2	Sequence 3
True positives	285	261	260
False positives	4	0	0
False negatives	58	26	35
Sensitivity	0.83090379	0.90940767	0.88135593
Precision	0.98615917	1	1
F-score	0.90189873	0.95255474	0.93693694
Area under the ROC (AUC)	0.8904519	0.95470383	0.94067797

Fig. 7. ROC curves for the image subtraction mode applied to our proper images.

Afterwards, an opening filter is applied to erase isolated pixels, getting $M_o(r, c, t)$. After this, the moments hard threshold, $\mu_h = 25$, is applied to $M(r, c, t)$ in order to obtain image $M_h(r, c, t)$ (see Fig. 6(d)) containing zones with elevated movement. Now, the blobs present in $M_o(r, c, t)$ are compared to the blobs of $M_h(r, c, t)$. The aim is to verify if each blob detected with the hard threshold is contained in a spot detected with the soft threshold. The spots that do not meet this condition are discarded. Finally, the resulting image, called refined moments image $M_r(r, c, t)$, and shown in Fig. 6(e), only contains the blobs (moving objects) that have met the previous condition.

3.3. Human blob segmentation

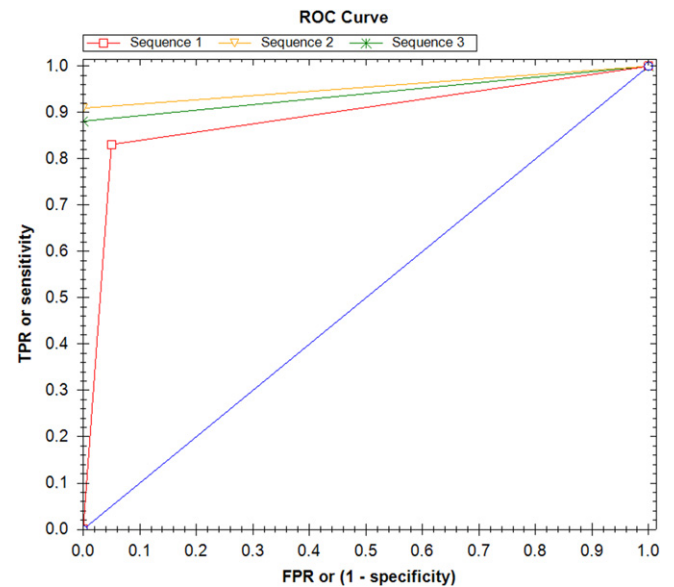This phase enhances the human detection by combining the results of the previous phases, that is, the human candidates' blob image either with the refined moments image in case optical flow has been computed, or with the image subtraction result.
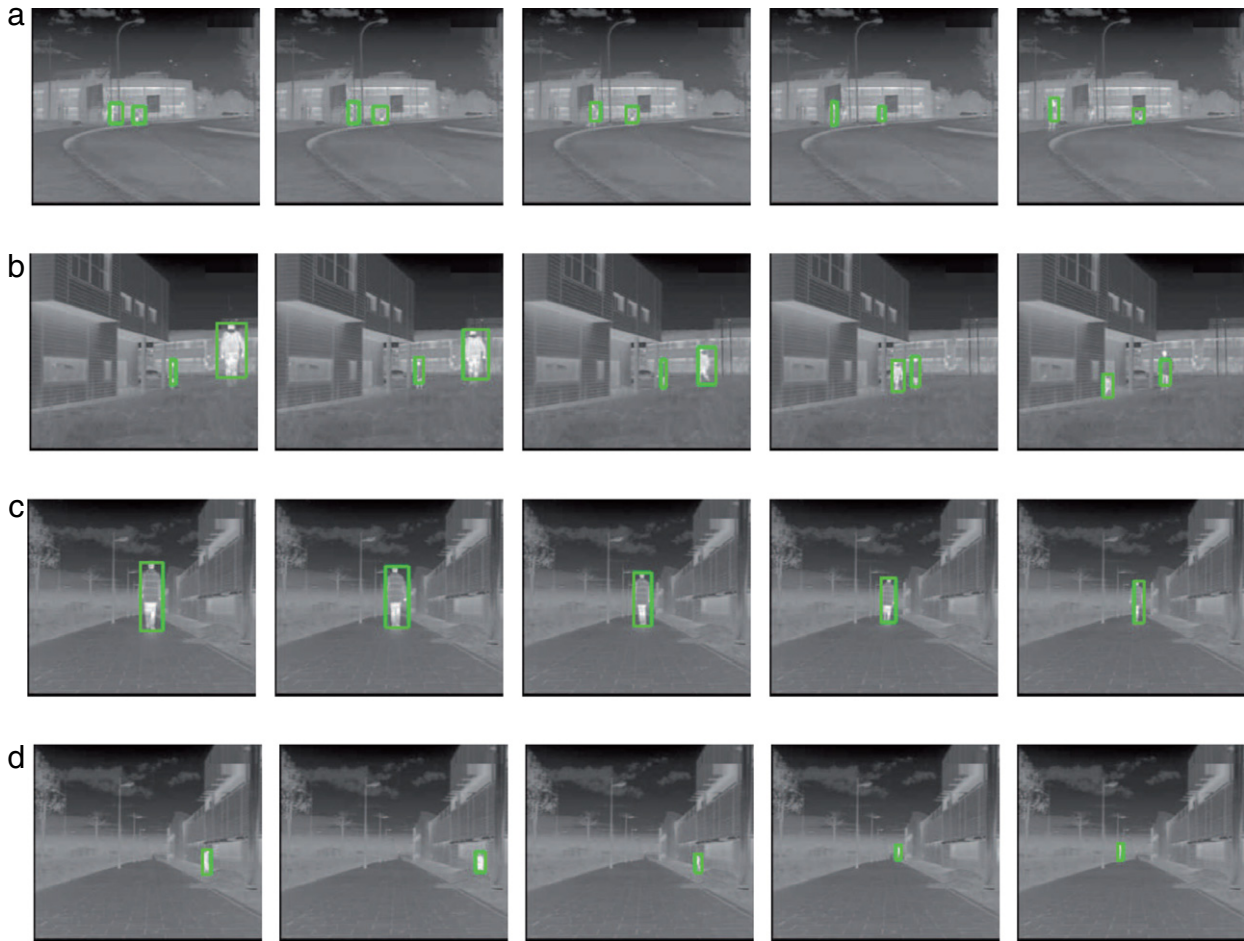
Fig. 8. Results for the image subtraction mode. Images captured from the mSecurit^{TM} mobile platform.

Table 2
Image subtraction results for the OTCBVS data sets.

	Sequence 1	Sequence 2	Sequence 3
True positives	285	261	260
False positives	4	0	0
False negatives	108	66	55
Sensitivity	0.72519084	0.79816514	0.82539683
Precision	0.98615917	1	1
F-Score	0.83577713	0.8877551	0.90434783
AUC	0.83759542	0.89908257	0.91269841

If optical flow technique has been applied in the previous step, this phase performs an *AND* operation between the output images of both previous phases, $B_r(r, c, t)$ and $M_r(r, c, t)$. The aim here is to take advantage of the optical flow information to improve the detection performed in the static analysis. This is, the optical flow emphasizes the moving areas obtained at the initial human candidates' detection. The possibilities that these moving shapes are humans are increased, as the resulting image

$$P_r(r, c, t) = B_r(r, c, t) \cap M_r(r, c, t) \qquad (7)$$

verifies if there exists a sufficient amount of pixels in movement within the zones delimited in $B_r(r, c, t)$.

On the other hand, if image subtraction was the technique used, this phase performs an *OR* operation between the images of the previous phases, $B_r(r, c, t)$ and $I_\theta(r, c, t)$. In case a spot is contained in both images, we keep the subtraction spot. This spot, experimentally, has shown a better fit with the object in the scene.

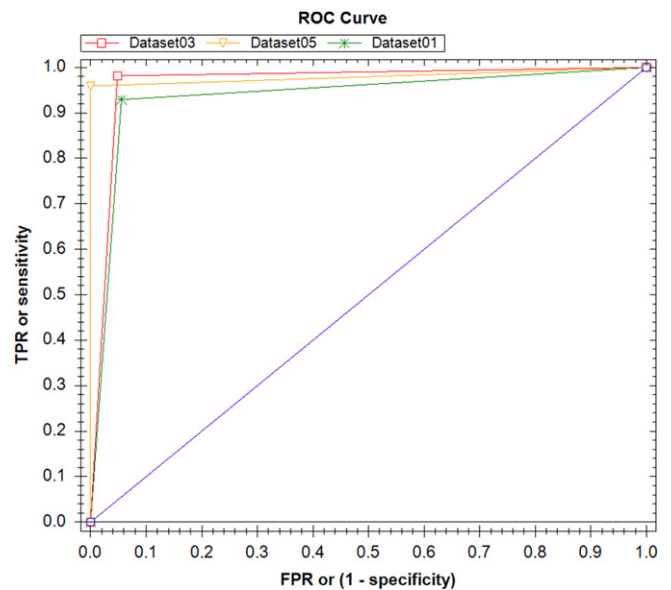$$P_r(r, c, t) = B_r(r, c, t) \cup I_\theta(r, c, t). \qquad (8)$$

Fig. 9. ROC curves for the image subtraction mode applied to the OTCBVS data sets.

4. Data and results

The complete human detection algorithm has been tested on an Intel® Celeron® *M* motherboard and processor at 600 MHz installed on the mSecurit^{TM} mobile robot. The RAM unit has a
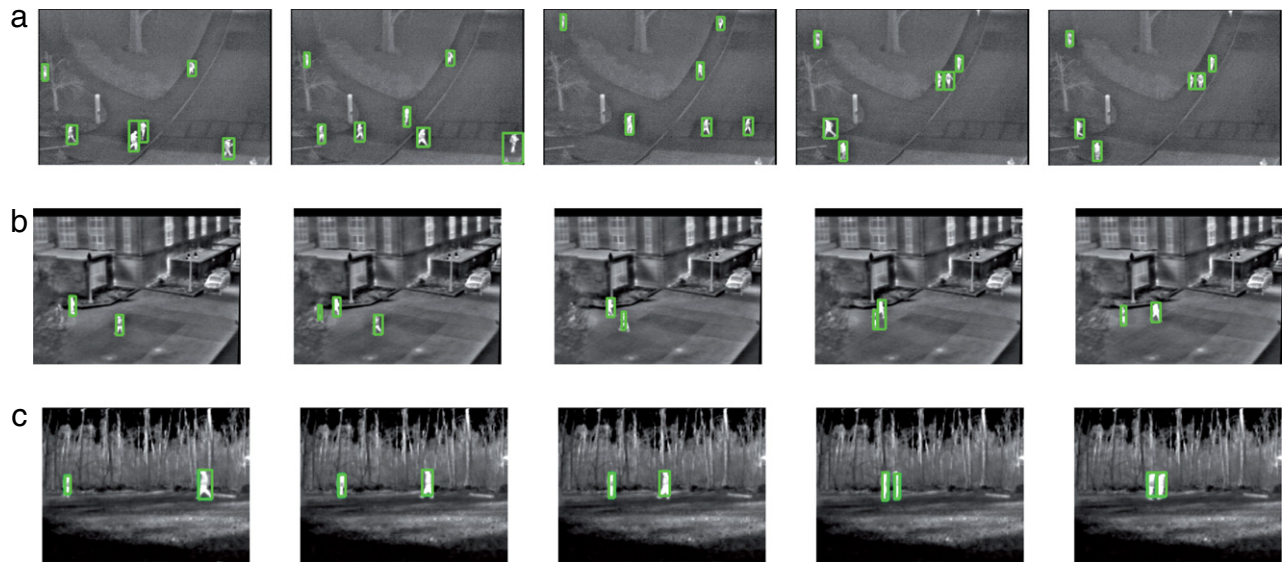
Fig. 10. Results for the image subtraction mode. Images obtained from the OTCBVS data sets. (a) Data set 05. (b) Data set 03. (c) Data set 01.

Table 3
Optical flow results for the mSecurit™ mobile platform images.

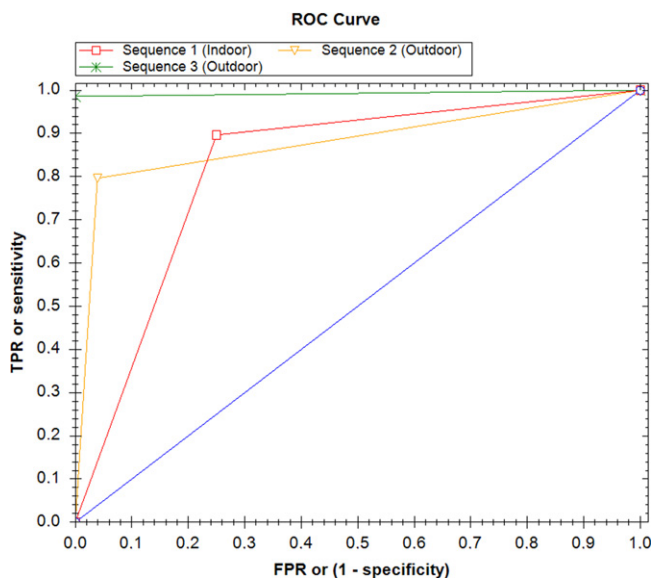|	Sequence 1 (indoor)	Sequence 2 (outdoor)	Sequence 3 (outdoor)
True positives	26	297	275
False positives	1	17	0
False negatives	3	76	4
Sensitivity	0.89655172	0.79624665	0.98566308
Precision	0.96296296	0.94585987	1
F-Score	0.92857143	0.86462882	0.99277978
AUC	0.82327586	0.87858309	0.99283154

Fig. 11. ROC curves for the optical flow mode applied to our proper images.

capacity of 1 GB. The performance results in terms of real-time capability of the algorithms described are excellent, as the method deals with the 6 frames per second provided by a FLIR camera (http://www.flir.com/) installed on the mSecurit™ mobile platform. The camera provides a frame rate of 5 frames per second and an image resolution of 768×576 pixels.

In the previous sections, the ability of our proposal to be adapted to two different operation modes has been described. For that reason two different kinds of tests were run on our system.

In first place, the performance of the subtraction-based approach was tested with several image sequences. As we wanted to assess the reliability of our proposal, not only sequences from the mobile platform were tested, but also images from three different data sets from the "OTCBVS Benchmark Data set Collection" (http://www.cse.ohio-state.edu/otcbvs-bench/). In second place, the approach including optical flow was tested; but, in this case only proper video sequences have been used.

4.1. Results for the subtraction-based mode

As already told before, the subtraction-based approach was tested with our proper videos and also with OTCBVS data sets. The results for our proper sequences are summarized in Table 1. As you may easily observe, the results show a high performance with a low amount of false positives. Please consider that our images are taken from the robot by means of the camera mounted in parallel to the ground at a height of 50 cm. Therefore, when humans are moving away from the camera along the z-axis, their silhouettes degrade and may confuse the algorithm. Precisely, false negatives are produced by this fact. In any case, the algorithm offers a good performance as depicted on the receiver operating characteristic (ROC) curves (see Fig. 7). A qualitative example of the algorithm's performance is also shown in Fig. 8.

We continued performing some subtraction-based tests by running our algorithms on three different data sets of the OTCBVS benchmark. These data sets contain images captured by an overhead infrared camera and an infrared camera mounted parallel to the ground (just like our proper images). As these images do not show any important displacement of humans in the z-axis, the performance obtained is highly increased in comparison to the previous results (see Fig. 8). This is shown in Table 2. Also the ROC
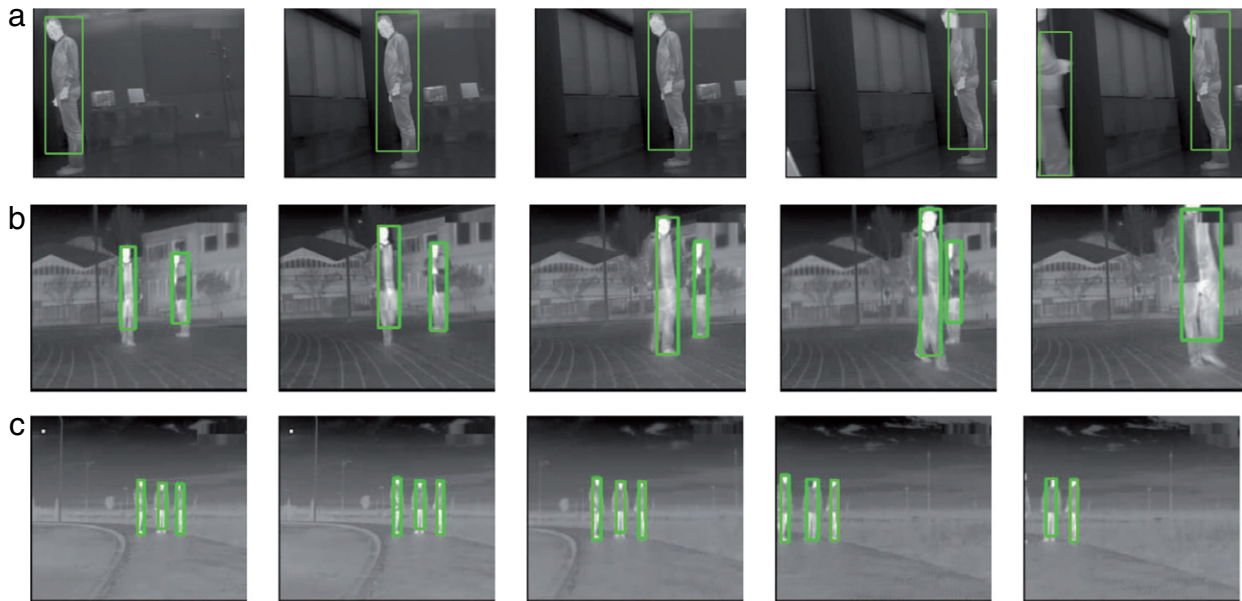
Fig. 12. Results for the optical mode. Images captured from the mSecurit™ mobile platform.

curves associated depict this enhanced behavior (see Fig. 9). Again, qualitative results are offered from the data sets execution. These are depicted in Fig. 10.

4.2. Optical flow-based results

To end this Data and results section, the second operation mode of the algorithm was also tested. That is, the optical flow approach thought to operate when the robot is moving was evaluated. As mentioned before, the images tested correspond to our proper captures from the camera mounted on the mobile platform. The performance is shown in Table 3 and depicted in Fig. 11. As it already happened with the subtraction-based approach, our implementation has a maximum detection range. The false negatives are produced when the humans are ahead of the camera. Some qualitative results are also depicted (see Fig. 12).

5. Conclusions

In the mobile robotics field, infrared sensors are being widely used for many applications, such as navigation and localization. In this paper, our approach to real-time human detection through processing video captured by a thermal infrared camera mounted on the Spanish private company MoviRobotics S.L. indoor autonomous mobile platform mSecurit™ has been presented. The mSecurit™ mobile platform has been specially designed for surveillance tasks. The robot is able to detect humans feeding the algorithms presented in this paper with thermal infrared frames. There is an intelligent switching between image subtraction and optical flow, depending on the platform movement and the processing load. Surveillance can be active on the whole route or only in certain known points.

The infrared-based approach starts with a phase of static analysis for the detection of human candidates. Then, a dynamic analysis by means of the optical flow algorithm based on Lucas and Kanade approach without pyramids or the standard image difference method is run. The results are promising and we are now engaged in performing tests in different real-world visual surveillance scenarios.

Acknowledgements

References

[1] J.M. Gascueña, A. Fernández-Caballero, Agent-based modeling of a mobile robot to detect and follow humans, Lecture Notes in Computer Science 5559 (2009) 80–89.

[2] A. Cherubini, G. Oriolo, F. Macrí, F. Aloise, F. Cincotti, D. Mattia, A multimode navigation system for an assistive robotics project, Autonomous Robots 25 (4) (2008) 383–404.

[3] L. Guo, M. Zhang, Y. Wang, G. Liu, Environmental perception of mobile robot, in: Proceedings of the 2006 IEEE International Conference on Information Acquisition, 2006, pp. 348–352.

[4] D. Coombs, M. Herman, T. Hong, M. Nashman, Real-time obstacle avoidance using central flow divergence, and peripheral flow, IEEE Transactions on Robotics and Automation 14 (1) (1998) 49–59.

[5] A. Giachetti, M. Campani, V. Torre, The use of optical flow for road navigation, IEEE Transactions on Robotics and Automation 14 (1) (1998) 34–48.

[6] B.R. Fajen, W.H. Warren, S. Temizer, L.P. Kaelbling, A dynamical model of visually-guided steering, obstacle avoidance, and route selection, International Journal of Computer Vision 54 (1–3) (2003) 13–34.

[7] A. Lookingbill, J. Rogers, D. Lieb, J. Curry, S. Thrun, Reverse optical flow for self-supervised adaptive autonomous robot navigation, International Journal of Computer Vision 74 (3) (2007) 287–30.

[8] J. Pavón, J.J. Gómez-Sanz, A. Fernández-Caballero, J.J. Valencia-Jiménez, Development of intelligent multi-sensor surveillance systems with agents, Robotics and Autonomous Systems 55 (12) (2007) 892–903.

[9] M.T. López, A. Fernández-Caballero, M.A. Fernández, J. Mira, A.E. Delgado, Visual surveillance by dynamic visual attention method, Pattern Recognition 39 (11) (2006) 2194–2211.

[10] R. Barrett, P.F.A. Maderson, R.M. Meszler, The pit organs of snakes, Biology of the Reptilia 2 (1970) 277–314.

[11] A.L. Campbell, R.R. Naik, L. Sowards, M.O. Stone, Biological infrared imaging and sensing, Micron 33 (2) (2002) 211–225.

[12] A. Yilmaz, K. Shafique, M. Shah, Target tracking in airborne forward looking infrared imagery, Image and Vision Computing 21 (7) (2003) 623–635.

[13] S. Iwasawa, K. Ebihara, J. Ohya, S. Morishima, Realtime estimation of human body posture from monocular thermal images, in: Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 15–20.

[14] B. Bhanu, J. Han, Kinematic-based human motion analysis in infrared sequences, in: Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, 2002, pp. 208–212.

[15] H. Nanda, L. Davis, Probabilistic template based pedestrian detection in infrared videos, in: Proceedings of the IEEE Intelligent Vehicle Symposium, vol. 1, 2002, pp. 15–20.

[16] F. Xu, X. Liu, K. Fujimura, Pedestrian detection and tracking with night vision, IEEE Transactions on Intelligent Transportation Systems 6 (1) (2005) 63–71.

[17] J.W. Davis, V. Sharma, Background-subtraction in thermal imagery using contour saliency, International Journal of Computer Vision 71 (2) (2007) 161–181.

[18] S.-H. Jung, J. Eledath, S. Johansson, V. Mathevon, Egomotion estimation in monocular infra-red image sequence for night vision applications, in: IEEE Workshop on Applications of Computer Vision, 2007, p. 8.

[19] G. Benet, F. Blanes, J.E. Simó, P. Pérez, Using infrared sensors for distance measurement in mobile robots, Robotics and Autonomous Systems 40 (4) (2002) 255–266.

[20] M.A. Garcia, A. Solanas, Estimation of distance to planar surfaces and type of material with infrared sensors, in: Proceedings of the 17th International Conference on Pattern Recognition, vol. 1, 2004, pp. 745–748.

[21] A. Treptow, G. Cielniak, T. Duckett, Real-time people tracking for mobile robots using thermal vision, Robotics and Autonomous Systems 54 (9) (2006) 729–739.

[22] G. Cielniak, T. Duckett, A.J. Lilienthal, Data association and occlusion handling for vision-based people tracking by mobile robots, Robotics and Autonomous Systems (2010) doi:10.1016/j.robot.2010.02.004.

[23] M.T. López, A. Fernández-Caballero, M.A. Fernández, J. Mira, A.E. Delgado, Motion features to enhance scene segmentation in active visual attention, Pattern Recognition Letters 27 (5) (2006) 469–478.

[24] J. Mira, A.E. Delgado, A. Fernández-Caballero, M.A. Fernández, Knowledge modelling for the motion detection task: the algorithmic lateral inhibition method, Expert Systems with Applications 27 (2) (2004) 169–185.

[25] Bruce D. Lucas, Takeo Kanade, An iterative image registration technique with an application to stereo vision, in: Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981.

[26] S. Yalamanchili, W.N. Martin, J.K. Aggarwal, Extraction of moving object descriptions via differencing, Computer Graphics and Image Processing 18 (2) (1982) 188–201.

[27] M. Piccardi, Background subtraction techniques: a review, in: IEEE International Conference on Systems, Man and Cybernetics, vol. 4, 2004, pp. 3099–3104.

[28] M.T. López, A. Fernández-Caballero, M.A. Fernández, J. Mira, A.E. Delgado, Dynamic visual attention model in image sequences, Image and Vision Computing 25 (5) (2007) 597–613.

[29] A.E. Delgado, M.T. López, A. Fernández-Caballero, Real-time motion detection by lateral inhibition in accumulative computation, Engineering Applications of Artificial Intelligence 23 (1) (2010) 129–139.

[30] T. Amiaz, N. Kiryati, Piecewise-smooth dense optical flow via level sets, International Journal of Computer Vision 68 (2) (2006) 111–124.

[31] P.L. Rosin, T. Ellis, Image difference threshold strategies and shadow detection, in: Proceedings of the 1995 British Conference on Machine Vision, vol. 1, 1995, pp. 347–356.

[32] S.M. Desa, Q.A. Salih, Image subtraction for real time moving object extraction, in: Proceedings of the International Conference on Computer Graphics, Imaging and Visualization, 2004, pp. 41–45.

[33] N. Ohta, A statistical approach to background subtraction for surveillance systems, in: Proceedings of the International Conference on Computer Vision, vol. II, 2001, pp. 481–486.

Antonio Fernández-Caballero received his degree in Computer Science from the Technical University of Madrid, Spain, in 1993, and received his Ph.D. from the Department of Artificial Intelligence of the National University for Distance Education, Spain, in 2001. Since 1995, he is an Associate Professor with the Department of Computer Science at the University of Castilla-La Mancha, Spain. His research interests are in Image Processing, Cognitive Vision, Neural Networks, and Intelligent Agents. A. Fernández-Caballero is an Associate Editor of the Pattern Recognition Letters journal and a member of the IAPR. He has authored more than 150 peer reviewed papers.
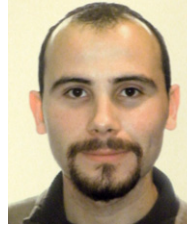
José Carlos Castillo received his degree in Computer Science from the University of Castilla-La Mancha, Spain, in 2006. For two years he was a researcher in the area of wireless sensor network routing protocols. Currently, he is working in moving objects segmentation through multiple camera types, and is pursuing his Ph.D. in multisensory data fusion techniques in the surveillance and safety domain.

Javier Martínez-Cantos received his degree in Computer Science from the University of Castilla-La Mancha, Spain, in 2004. Currently, he is a researcher in the Spanish company MoviRobotics S.L. dedicated to the development of mobile robots. He is pursuing his Ph.D. in mobile robotics applied to the surveillance and safety domain.

Rafael Martínez-Tomás received his degree in Physics from the University of Valencia, Spain, in 1983, and his Ph.D. from the Department of Artificial Intelligence of the National University for Distance Education, Spain, in 2000. Since 2001, he is an Assistant Professor with the Department of Artificial Intelligence of the National University for Distance Education. His research interests are in Knowledge Engineering, Knowledge-Based Systems, Spatial–Temporal Logics, Description Logics and Video-Sequence Semantic Interpretation.