PERGAMON

Pattern Recognition 36 (2003) 1131–1142

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Spatio-temporal shape building from image sequences using lateral interaction in accumulative computation

Antonio Fernandez-Caballero[a,*], Miguel A. Fernandez[a], Jose Mira[b], Ana E. Delgado[b]

[a] *Departamento de Informatica, Universidad de Castilla-La Mancha, Campus Universitaro, 02071 Albacete, Spain*
[b] *Departamento de Inteligencia Artificial, UNED, c/Senda del Rey, 9, 28040 Madrid, Spain*

Abstract

To be able to understand the motion of non-rigid objects, techniques in image processing and computer vision are essential for motion analysis. Lateral interaction in accumulative computation for extracting non-rigid shapes from an image sequence has recently been presented, as well as its application to segmentation from motion. In this paper, we introduce a modified version of the first multi-layer architecture. This version uses the basic parameters of the LIAC model to spatio-temporally build up to the desired extent the shapes of all moving objects present in a sequence of images. The influences of LIAC model parameters are explained in this paper, and we finally show some examples of the usefulness of the model proposed.

Keywords: Segmentation from motion; Image analysis; Shape representation; Silhouette recognition; Lateral interaction; Accumulative computation

1. Introduction

There has been a great deal of research interest in motion tracking [1–3] because of its great applicability in a wide variety of applications. Non-rigid motion is difficult by its proper nature. There has also been much work carried out on the extraction of non-rigid shapes from image sequences. In general, all papers take advantage of the fact that the image flow of a moving figure varies both spatially and temporally.

Little and Boyd [4] found it reasonable to suggest that variations in gaits are recoverable from variations in image sequences. There have been several attempts to recover characteristics of gait from image sequences by Polana and Nelson [5,6]. Polana and Nelson [7] characterise the temporal texture of a moving figure by summing the energy of the highest amplitude frequency and its multiples. They use Fourier analysis. Their more recent work [8] emphasises the spatial distribution of energies around the moving figure.

Bobick and Davis [9] introduced the motion energy image (MEI), a smoothed description of the cumulative spatial distribution of motion energy in a motion sequence. Bobick and Davis [10] enhanced the MEI to form a motion-history image (MHI), where pixel intensity is a function, over time, of the energy in the current binary motion energy and recent activity, which they extend in later work [11]. The MEI [9] is arrived at by binary threshold of motion displacements computed as the threshold of the pixelwise summed squared difference between each image and the first, over an entire sequence.

Yang and Ahuja [12] segment an image frame into regions with similar motion. The algorithm identifies regions in each frame comprising the multiscale intraframe structure [13]. Regions at all scales are then matched across frames. Affine transforms are computed for each matched region pair. The affine transform parameters for region at all scales are then used to derive a single motion field that is then segmented to identify the differently moving regions between two frames.

Olson and Brill [14] propose a general purpose system for moving object detection and event recognition where moving objects are detected using change detection and tracked using first-order prediction and nearest neighbour matching.

* Corresponding author. Tel.: +34-967-59-92-00; fax: +34-967-59-92-24.
E-mail address: caballer@info-ab.uclm.es (A. Fernandez-Caballero).

0031-3203/03/$30.00
PII: S0031-3203(02)00116-4

The goal of this paper is to present a novel method for spatio-temporally shape building taking advantage of the inherent motion present in image sequences. With this new model, we can parameterise the aspect of gradually emerging shapes due to motion.

2. Lateral interaction in accumulative computation

Lateral interaction in accumulative computation has recently been introduced [15–18], as well as its application to segmentation from motion [19]. For it, a generic model based on a neural architecture was presented. We shall now remind of the most important characteristics of this model.

The proposed model is based on accumulative computation function [16,17] followed by a set of co-operating lateral interaction processes [20,21]. These are performed on a functional receptive field organised as centre–periphery over non-linear and temporal expansions of their input spaces [5,21,22].

A lateral interaction model [6,23] consists of a layer of modules of the same type with local connectivity, such that the response of a given module does not only depend on its own inputs, but also on the inputs and outputs of the module's neighbours. From a computational point of view, the aim of the lateral interaction nets is to partition the input space into three regions: centre, periphery and excluded. The following steps have to be done: (a) processing over the central region, (b) processing over the feedback of the periphery zone, (c) comparison of the results of these operations and a local decision generation, and, (d) distribution over the output space.

We also incorporate the notion of double time scale present at sub-cellular micro-computation [17]. So, the following properties are applicable to the model: (a) local convergent process around each element; (b) semiautonomous functioning, with each element capable of spatio-temporal accumulation of local inputs in time scale T, and conditional discharge; and (c) attenuated transmission of these accumulations of persistent coincidences towards the periphery that integrates at global time scale t.

Therefore, we are in front of two different time scales: (1) the local time T and (2) the global time t ($t = nT$). Global time is applicable to steps (a) and (d) of our neuronal lateral interaction model, whereas steps (b) and (c) use local time scale T.

3. LIAC architecture for spatio-temporal shape building

In first place, and in the following figure, the complete structure chosen as the modular computational solution to apply the model to spatio-temporal shape building is presented. In Fig. 1, four layers can be appreciated that form the architecture of the lateral interaction in accumulative computation method.

Now we are going to explain the role of each of these four layers in the task of shape building.

3.1. Layer 0: segmentation by grey level bands

This layer covers the need to segment the image at a predefined group of n grey level bands. Each element (x, y) is capable of processing motion from input grey level value $IN(x, y, t)$ and its proper charge value. Let $GLS(k, x, y, t)$ be the presence or absence of grey level k at element (x, y) at time t

$$GLS(k,x,y,t) = \begin{cases} -1 & \text{if } IN(x,y,t) \neq k, \\ 1 & \text{otherwise,} \end{cases} \qquad \forall k \in [0,255],$$

where n is the number of grey level bands, and, k is a particular grey level band. In other words, we have to determine in what grey level band a certain pixel falls. So, we are not evaluating, at this level, if there is motion in a grey level band for a given pixel. This task is left to the following layer.

It must be clear that one, and only one, of the outputs of all the detecting modules of the grey level bands can be activated at a given instant. This fact, although obvious, is of a great interest at the higher layers of the architecture, since it will avoid possible conflicts among the values offered by the different grey level bands. Indeed, only one grey level band will contain valid values.

3.2. Layer 1: lateral interaction for accumulative computation

This layer has been designed to obtain the permanence value $PM(k, x, y, t)$ on a decomposition in grey level bands basis. We will have n sub-layers and each one of them will memorise the value of the accumulative computation present at global time scale t for each element. Lateral interaction in this layer is thought to reactivate the permanence charge of those elements partially loaded and that are directly or indirectly connected to maximally charged elements. The permanence charge of each element will be offered to the following layer as output.

Firstly, at global time scale t, permanence memory charge or discharge due to motion detection is performed. This information, given as input from layer 0, is associated to sub-layer k of layer 1 (grey level band k). The accumulative computation equation may be formulated as

$$PM(k,x,y,t) = \begin{cases} l_{dis} & \text{if } GLS(k,x,y,t) = -1, \\ l_{sat} & \text{if } GLS(k,x,y,t) = 1 \text{ and} \\ & \quad GLS(k,x,y,t-\Delta t) = -1, \\ \max(PM(k,x,y,t-\Delta t) - dv, l_{dis}) & \\ \quad \text{if } GLS(k,x,y,t) = 1 \text{ and } GLS(k,x,y,t-\Delta t) = 1, & \end{cases}$$

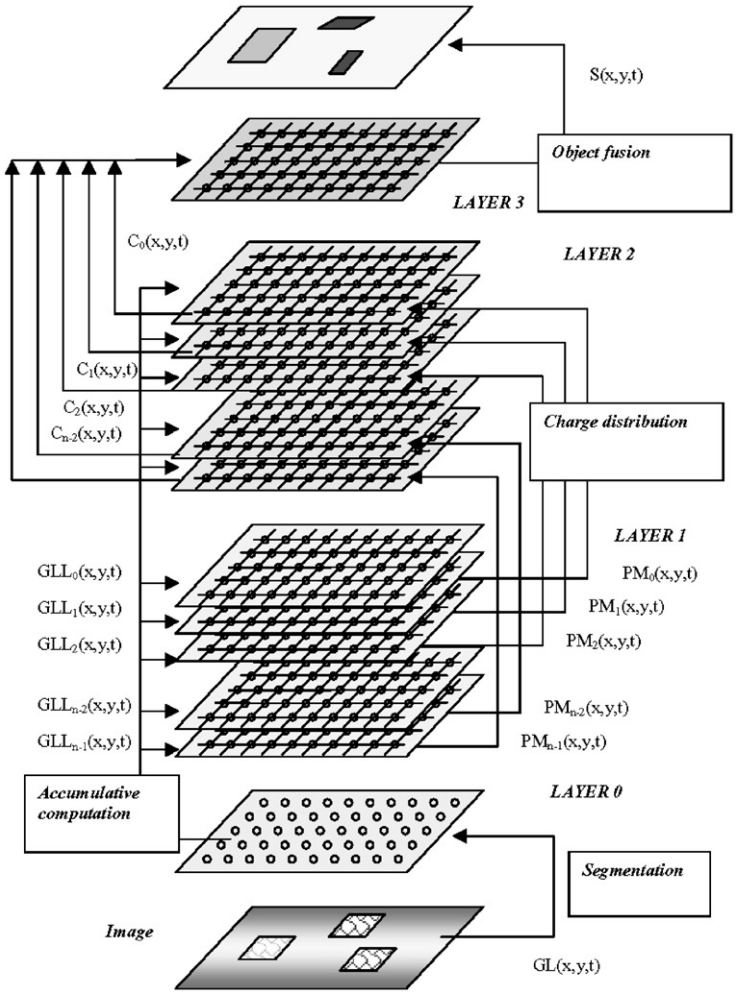

Fig. 1. Multi-layer configuration.

where l_{dis} is the discharge or minimum permanence value, l_{sat} is the saturation or maximum permanence value, and dv is the discharge value due to motion detection.

Note that Δt determines the sequence frame rate and is given by the capacity of the model's implementation to process one input image. At each element (x, y) we are in front of three possibilities: (1) The sub-layer does not correspond to the grey level band of the image pixel. The permanence value is discharged down to value l_{dis}. (2) The sub-layer corresponds to the grey level band of the image pixel at time instant t, and it did not correspond to the grey level band at the previous instant $t - \Delta t$. The permanence value is loaded to the maximum of saturation l_{sat}. (3) The sub-layer corresponds to the grey level band of the image pixel at time instant t, and it also corresponded to the grey level band at the instant $t - \Delta t$. The permanence value is discharged by a value dv (discharge value due to motion detection); of course, the permanence value cannot get off a minimum value l_{dis}.

The discharge of a pixel by a quantity of dv is the way to stop maintaining attention to a pixel of the image that had captured our interest in the past. As it will be seen later on, if a pixel is not directly or indirectly bound by means of lateral interaction mechanisms to a maximally charged pixel (l_{sat}), it goes down to the total discharge with time.

Secondly, an extra charge rv (recharge value due to neighbouring) is added to the permanence memory in those image pixels that receive a stimulus from a maximally charged element almost l_1 pixels far away in any of four directions. This recharge can only happen one time, and provided that none neighbour element up to the maximally charged element is discharged. l_1 is called number of neighbours in accumulative computation. This recharge mechanism allows maintaining attention on those pixels directly or indirectly

connected to maximally charge pixels. This mechanism is even able to reinforce the permanence memory value if the $rv > dv$.

$$PM(k,x,y,T) = \min(PM(k,x,y,T-\Delta T) + \varepsilon\, rv, l_{sat}),$$

where

$$\varepsilon = \begin{cases} 1 & \text{if } \exists(1 \leqslant i \leqslant l_1)!\, \forall(1 \leqslant j \leqslant i), \\ & (PM(k,x,y+i,T-\Delta T) \\ & \quad = l_{sat} \cap PM(k,x,y+j,T-\Delta T) > l_{dis}) \cup \\ & \quad (PM(k,x+i,y,T-\Delta T) \\ & \quad = l_{sat} \cap PM(k,x+i,y,T-\Delta T) > l_{dis}) \cup \\ & \quad (PM(k,x,y-i,T-\Delta T) \\ & \quad = l_{sat} \cap PM(k,x,y-j,T-\Delta T) > l_{dis}) \cup \\ & \quad (PM(k,x-i,y,T-\Delta T) \\ & \quad = l_{sat} \cap PM(k,x-1,y,T-\Delta T) > l_{dis}) \\ 0 & \text{otherwise} \end{cases}$$

Lastly, back at global time scale t, the permanence value at each pixel (x, y) is threshold and sent to the next layer.

$$PM(k,x,y,t) = \begin{cases} PM(k,x,y,t) & \text{if } PM(k,x,y,t) > \theta_1, \\ \theta_1 & \text{otherwise.} \end{cases}$$

In order to explain the central idea of this layer 1, we will say that the activation toward the lateral modular structures (up, down, right and left) is based on the following basic ideas: (1) all modular structures with maximum permanence value l_{sat} (saturated) output the charge toward the neighbours; (2) all modular structures with a not saturated charge value, and that have been activated from some neighbour, allow to pass this information through them (they behave as transparent structures to the charge passing); and (3) the modular structures with minimum permanence value l_{dis} (discharged) stop the passing of the charge information toward the neighbours (they behave as opaque structures). Therefore, we are in front of an explosion of lateral activation beginning at the structures with permanence memory set at l_{sat}, and that spreads lineally toward all the addresses, until a structure appears in the pathway with a discharged permanence memory.

3.3. Layer 2: lateral interaction for charge redistribution by grey level bands

Layer 2 is also formed of n sub-layers, where, by means of lateral interaction, charge redistribution among all connected neighbours in a surrounding window of $l_2 * l_2$ pixels that hold a minimum charge, is performed. Besides distributing the charge $C(k,x,y,t)$ in grey level bands, at this level, the charge due to the motion of the background is also diluted. The new charge obtained in this layer is offered as an output toward layer 3.

Starting from the values of the permanence memory in each pixel on a grey level band basis, we will see how it is possible to obtain all the parts of an object in movement. A part of an object concretely means the union of pixels that are together and in a same grey level band. The discrimination of each one of the parts that compose the objects is equally obtained by lateral co-operation mechanisms. In case of layer 2, the charge will be homogenised among all the pixels that pertain to the same grey level band and that are directly or indirectly united to each other. This way, a double objective will be obtained: (1) diluting the charge due to the false image background motion along the other pixels of the background. This way, there should be no presence of motion characteristic of the background, but we will rather keep motion of the objects present in the scene. (2) Obtaining a parameter common to all the pixels of the part of the object in a surrounding window of $l_2 * l_2$ pixels with a same grey level band.

Initially, at global time scale t, the charge value at each pixel (x, y) and at each sub-layer k is given the value of the permanence value from the previous layer

$$C(k,x,y,t) = PM(k,x,y,t).$$

Afterwards, at local time scale T, provided that the neighbour input charge values are high enough, the centre element (x, y) calculates the mean of its value and the partially charged neighbours in a surrounding window of $l_2 * l_2$ pixels. l_2 is denominated number of neighbours in charge redistribution

$$C(k,x,y,T) = \frac{C(k,x,y,T-\Delta T) + \sum_{i=-l_2}^{l_2} \sum_{j=-l_2}^{l_2} \delta_{x+i,y+j} C(k,x+i,y+j,T-\Delta T)}{1 + \sum_{i=-l_2}^{l_2} \sum_{j=-l_2}^{l_2} \delta_{x+i,y+j}},$$

$$\forall (i,j) \neq (0,0),$$

where

$$\delta_{\alpha,\beta} = \begin{cases} 1 & \text{if } C(k,\alpha,\beta,T-\Delta T) > l_{dis}, \\ 0 & \text{otherwise.} \end{cases}$$

Again at global time scale t, the charge value at each pixel (x, y) is threshold and sent to the next layer

$$C(k,x,y,t) = \begin{cases} C(k,x,y,t) & \text{if } C(k,x,y,t) > \theta_2, \\ \theta_2 & \text{otherwise.} \end{cases}$$

3.4. Layer 3: lateral interaction for spot fusion

In each element of layer 3, we have an input from each corresponding element of the n sub-layers of layer 2. This layer has as purpose the fusion into uniform spots of the

objects in a surrounding window of $l_3 * l_3$ pixels. That is why it takes the input charges of each one of the grey level bands and performs a fusion of these values, obtaining uniform parts of all the moving objects of the original image. Its output is a set of spots $S(x, y, t)$.

Up to now attention has been captured on any moving objects in the scene by means of co-operative calculation mechanisms in all grey level bands. Motion due to background has also been eliminated. It is now necessary to fix as a new objective to clearly distinguish the motion of the different objects. This discrimination is obtained equally by lateral co-operation mechanisms. Again we will connect the modular structures of this layer in a mesh form in layer 3. Nevertheless, now we will no longer work with sub-layers, but rather all information of the n sub-layers of layer 2 end up in a single layer. In layer 3, we will homogenise the charge values among all the pixels that contain some charge value superior to a minimum threshold and that are physically connected to each other in a radius of l_3 pixels.

Firstly, the spot charge value at each pixel (x, y) is given the charge value of the maximally charged sub-layer k from the previous layer.

$$S(x, y, t) = \max(C(k, x, y, t)), \quad \forall k \in [0, 255].$$

At local time scale, provided that the neighbour input charge values are high enough, the centre element (x, y) calculates the mean of its value and the partially charged neighbours in a surrounding window of $l_3 * l_3$ pixels. l_3 is denominated number of neighbours in object fusion.

$$S(x, y, T) = \frac{S(x, y, T - \Delta T) + \sum_{i=-l_3}^{l_3} \sum_{j=-l_3}^{l_3} \delta_{x+i, y+j} S(x + i, y + j, T - \Delta T)}{1 + \sum_{i=-l_3}^{l_3} \sum_{j=-l_3}^{l_3} \delta_{x+i, y+j}},$$

$$\forall (i, j) \neq (0, 0),$$

where

$$\delta_{\alpha,\beta} = \begin{cases} 1 & \text{if } S(\alpha, \beta, T - \Delta T) > l_{dis}, \\ 0 & \text{otherwise.} \end{cases}$$

Back to global time scale t, the spot charge value at each pixel (x, y) is threshold.

$$S(x, y, t) = \begin{cases} S(x, y, t) & \text{if } S(x, y, t) > \theta_3, \\ \theta_3 & \text{otherwise.} \end{cases}$$

4. Influence of the basic LIAC parameters in shape building

Now we will comment the influence of the most important parameters of the lateral interaction in accumulative computation model applied to spatio-temporal shape building. This will be carried out by using the results of applying a model's implementation to a sequence of images taken in our research laboratory. In this sequence three people are moving freely.

Table 1
Basic LIAC parameters

	Parameter description	Layer
n	Number of grey level bands	1
dv	Discharge value due to motion detection	1
rv	Recharge value due to neighbouring	1
l_1	Number of neighbours in accumulative computation	1
θ_1	Permanence value threshold	1
l_2	Number of neighbours in charge redistribution	2
θ_2	Charge value threshold	2
l_3	Number of neighbours in object fusion	3
θ_3	Shape value threshold	3

The basic parameters, as previously seen in the three layers of our architecture, are depicted in Table 1.

4.1. Influence of parameter "number of grey level bands"

To explain the influence of parameter number of grey level bands (n) we shall work with three possible values of this parameter and we will offer the obtained results after applying up to layer 3 of the model (see Fig. 2). Results appear as images, where (a) the first column offers the input images, (b) the second column offers the results of applying case study A ($n = 4$), (c) the third column offers the results of applying case study B ($n = 8$), and, (d) the fourth column offers the results of applying case study C ($n = 32$). As you may appreciate, a higher value of number of grey level bands usually enables to better discriminate the whole shapes of the moving non-rigid objects. Nevertheless, a too high value of this parameter may include some image background into the shapes. This may even lead to fuse more than one different shape into one single silhouette.

4.2. Influence of parameter "discharge value due to motion detection"

To explain the influence of parameter discharge value due to motion detection (dv) we will work with three possible values of this parameter ($dv = 255$, 127 and 15), and we will offer the results obtained as output of layer 1 of the model (Fig. 3). Once again, the first column of Fig. 3 offers the input images to the implementation of the lateral interaction in accumulative computation model. The following columns show the result of applying the model to the three case studies. On the column second you may appreciate in white colour the image pixels that have just changed their grey level band since the last image. Lowering the discharge value due to motion detection, we go obtaining the trail of the movement in the last images. This way, the third column offers the story of the last three images. That is, in white the difference among the images at t and $t - 1$, and in light

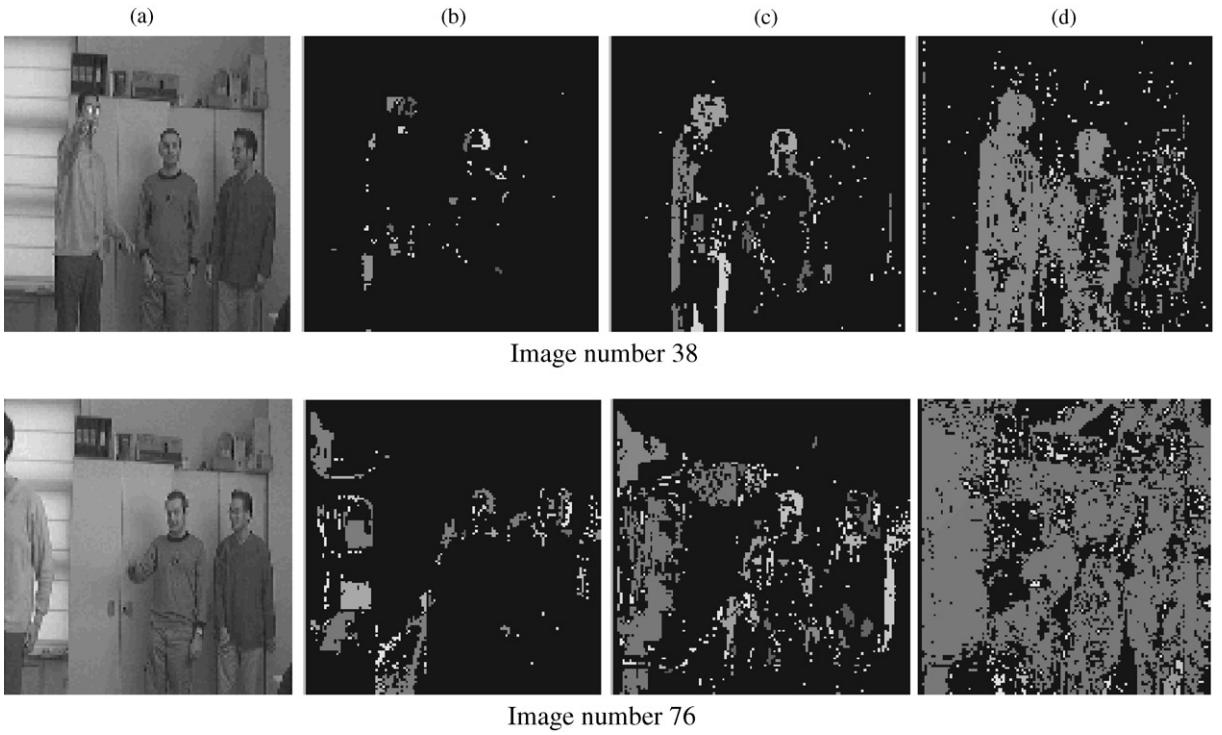

Fig. 2. Influence of parameter number of grey level bands: (a) input images; (b) result after layer 3 for case study A; (c) result after layer 3 for case study B; (d) result after layer 3 for case study C.

grey the difference between the images at $t-1$ and $t-2$. The fourth column shows more clearly the history of the last images. Indeed, by means of little value of discharge value due to motion detection, we obtain more information of the history of the movement through the offered trail.

4.3. Influence of parameter "recharge value due to neighbouring"

To explain the influence of the parameter recharge value due to neighbouring (rv) we will work equally with three possible values of this parameter ($rv = 0$, 63 and 127), and we will offer the results obtained as output of layer 1 (see Fig. 4). In the first column of Fig. 4, the input images to the implementation of the lateral interaction in accumulative computation model are shown. The successive columns show the result of applying the model in the three cases.

The second column of Fig. 4 shows the results for the first study case. Here we have decided not to recharge the memories, that is to say, we use a recharge value due to neighbouring equal to 0. You may appreciate three different charge levels. (a) In black, corresponding to grey level value 0—or minimum permanence value (l_{dis})—there is the background, where no motion has been detected. (b) In dark grey, that is to say, for a grey level around 128, we have motion detected between time instants $t-2$ and $t-1$. (c) In white, corresponding to grey level value 255—or maximum permanence value (l_{sat})—there is motion detected between time instants $t-1$ and t. The third column shows the results of applying the model to the second case study, using a recharge value due to neighbouring (dv) different to 0, but inferior to the discharge value due to motion detection (rv). In this case a charge increase is appreciated in those pixels that in the column second were in dark grey (a charge value sufficient to be recharged). These pixels are now in clear grey. Note that the recharge has as secondary effect, recovering part of the history of motion.

The last case study presents the highest possible value of recharge value due to motion detection (rv). The result is that of big pixel areas loaded to the maximum permanence value (l_{sat}).

4.4. Influence of parameters "number of neighbours in accumulative computation", "number of neighbours in charge redistribution" and "number of neighbours on object fusion"

As there is no appreciable difference in the explanation of the influence of parameters number of neighbours in accumulative computation (l_1), number of neighbours in object fusion (l_3) and number of neighbours in charge redistribution (l_2), we will only offer an example for l_2. To explain

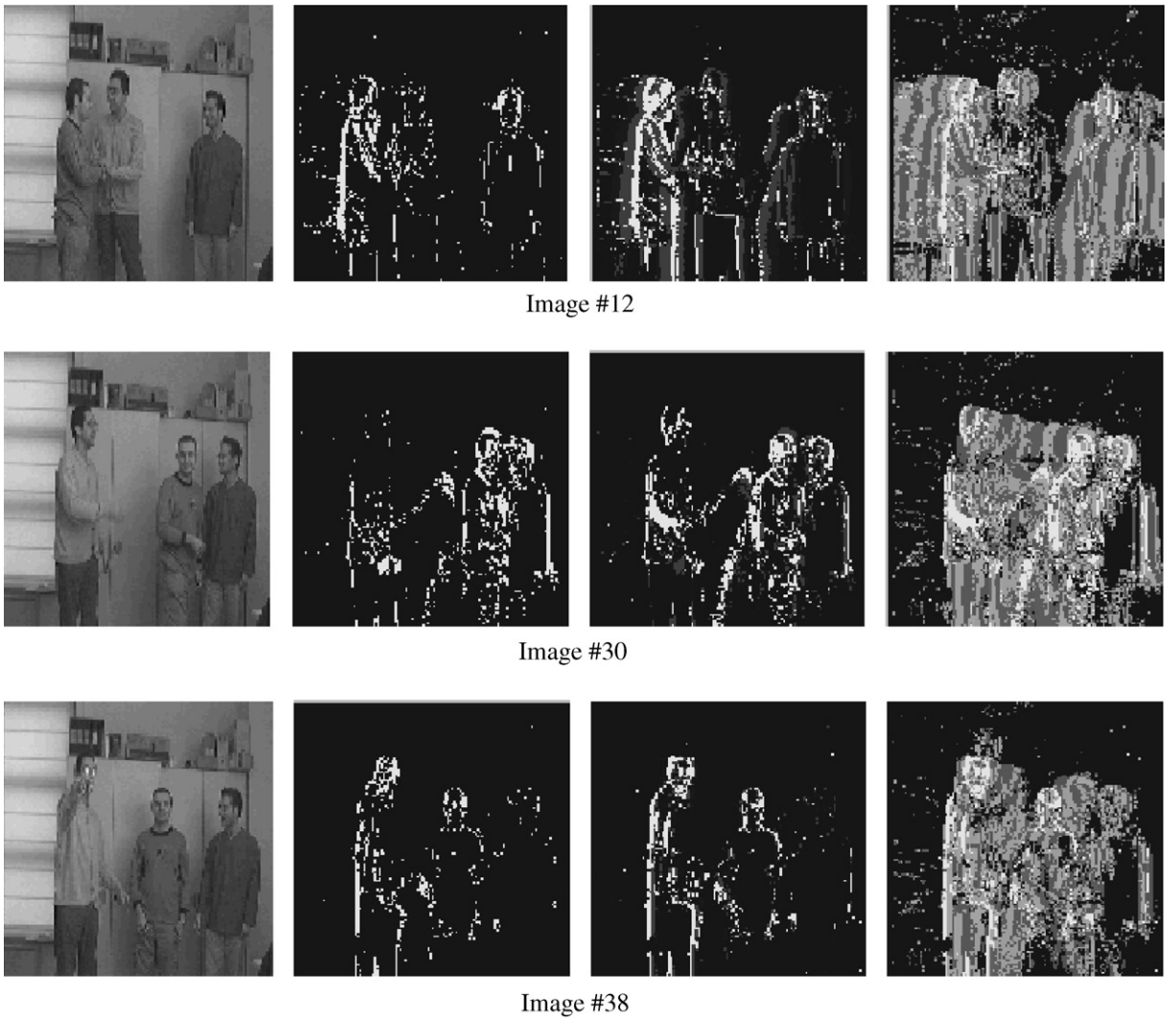

Fig. 3. Influence of parameter discharge value due to motion detection.

the influence of parameter number of neighbours in charge redistribution (l_2) we will offer on Fig. 5 two different cases, denominated case study A ($l_3=64$) and case study B ($l_3=8$), as reflected in Table 1. We offer the results equally in form of images of the output of layer 2 of the model.

In general, one can affirm that parameters number of neighbours in accumulative computation, number of neighbours in charge redistribution and number of neighbours in object fusion are conceived to define the outreach or influence area of lateral interaction of the model. Logically, the higher the action radius, the greater the area affected by the neighbours charges.

If you want to fuse small pixel areas, a relatively low value for number of neighbours in accumulative computation, number of neighbours in charge redistribution or number of neighbours in object fusion is required. This way it is possible to maintain the details of the non-rigid objects. On the contrary, if you want to fuse greater pixel areas a higher value has to be used. This other way allows fusing more extensive areas of pixels in a single shape. You may appreciate on Fig. 5 how a lower value of number of neighbours in charge redistribution guards more details of the shapes (compare column 3 with column 2).

4.5. Influence of parameters "permanence value threshold", "charge value threshold" and "shape value threshold"

Once again, we shall take advantage of the fact that there is no appreciable difference in the explanation of the

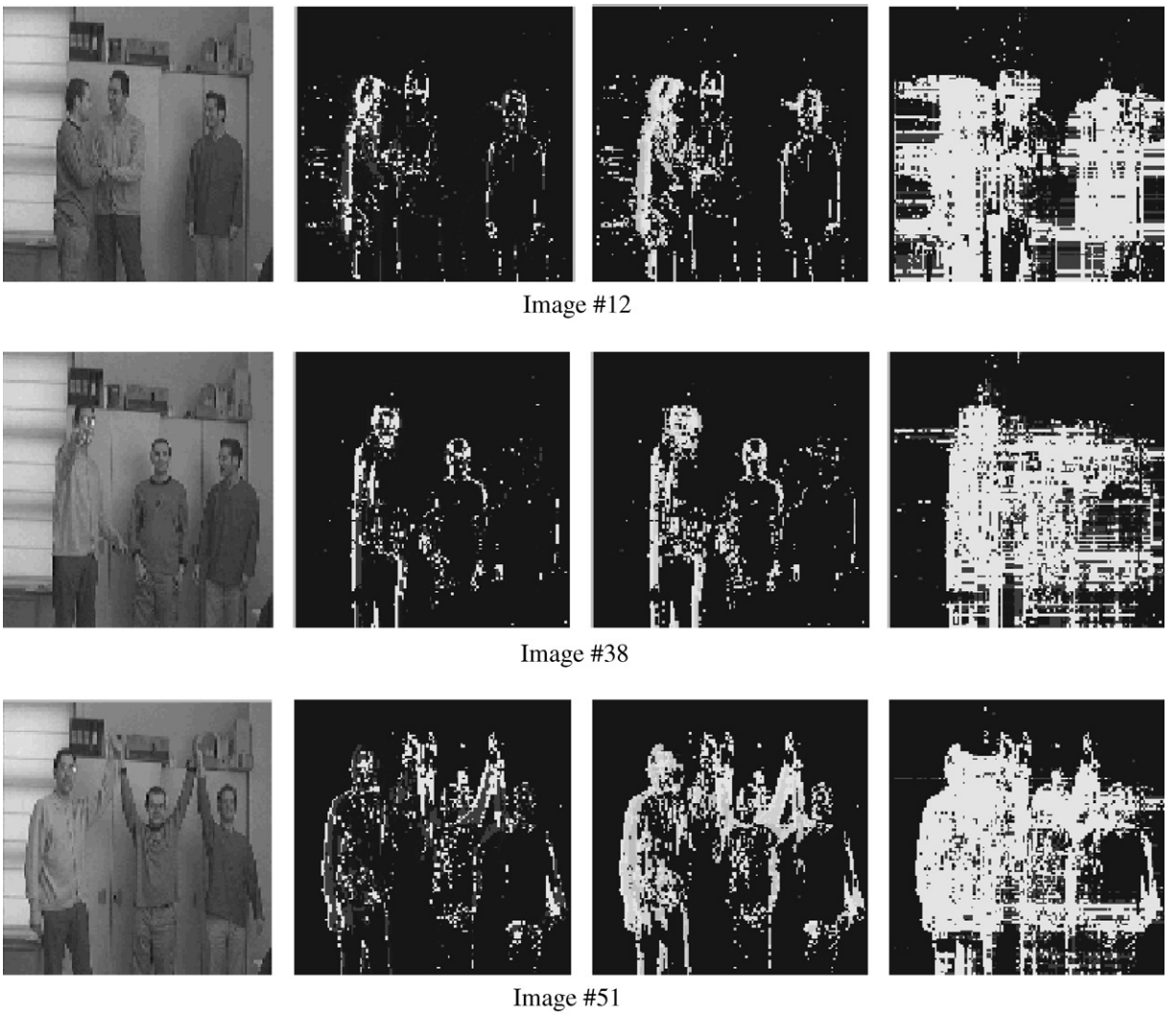

Fig. 4. Influence of parameter recharge value due to neighbouring.

influence of parameters permanence value threshold (θ_1), shape value threshold (θ_3) and charge value threshold (θ_2). Thus, to explain the influence of parameters permanence value threshold (θ_1), charge value threshold (θ_2) and shape value threshold (θ_3) we will work with two values of parameter θ_2, namely ($\theta_2 = 150$ and 200). Results are shown in Fig. 6.

As you can easily appreciate, parameter charge value threshold (θ_2)—or permanence value threshold (θ_1), shape value threshold (θ_3)—is a threshold that limits the image pixels charge value at layer 2—layer1, layer 3—of the lateral interaction in accumulative computation model. A higher value of permanence value threshold (θ_1), charge value threshold (θ_2) or shape value threshold (θ_3) is able to better discriminate the number of pixels to be used at this level.

5. Some significant examples

To show the usefulness and the multi-functionality of the parameterised model proposed, we show in this section three significant examples where the algorithms described may be applied, namely: (1) silhouette detection (SD), (2) motion detection (MD), and (3) direction detection (DD) and motion tracking. All three examples are the result of applying our model to a same image sequence, namely *TwoWalkNew*, downloaded from University of Maryland Institute for Advanced Computer Studies, copyright © 1998 University of Maryland, College Park. This sequence used to test the real time visual surveillance system W^4 [24] shows two people that walk trough a scene. The different results are obtained just by using different values for the same LIAC model (see Table 2).

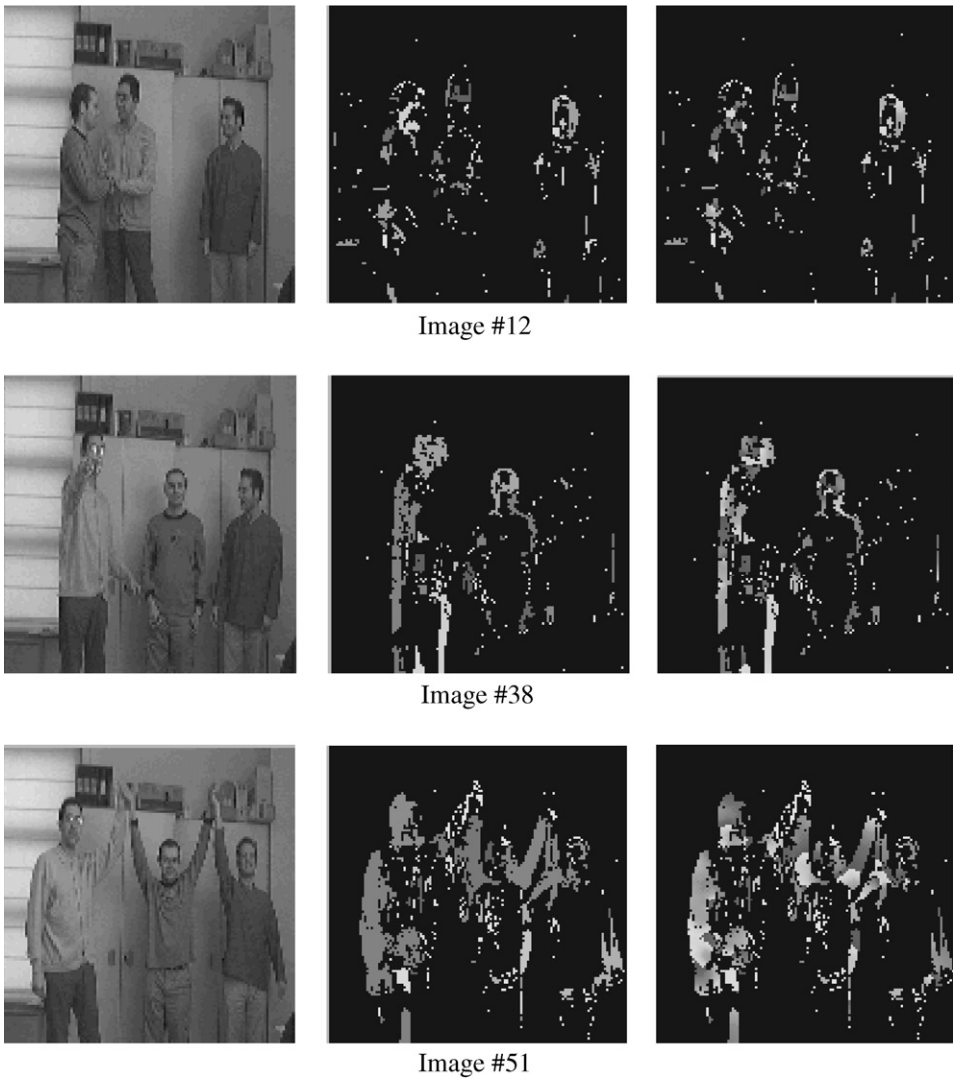

Fig. 5. Influence of the number of neighbours in charge redistribution.

Table 2
Parameter values for the three significant examples

	n	l_{dis}	l_{sat}	dv	rv	l_1	θ_1	l_2	θ_2	l_3	θ_3
SD	8	0	255	64	0	1	192	1	192	1	192
MD	16	0	255	192	4	8	63	16	63	32	63
DD	8	0	255	16	4	1	127	1	127	1	127

Fig. 7 shows the model output for the three examples. The pure output is shown on the first column for each example. The second column shows the result superimposed on the input image. The results of silhouette detection are drawn in blue colour (Fig. 7a), motion detection in red colour (Fig. 7b) and direction detection in green colour (Fig. 7c). In this last example notice that motion direction is shown by means of the intensity of green colour. Direction has to be interpreted going from clearer to darker green.

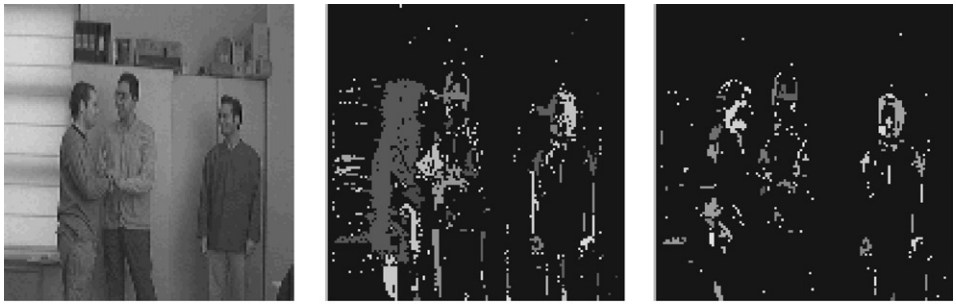

Image #12

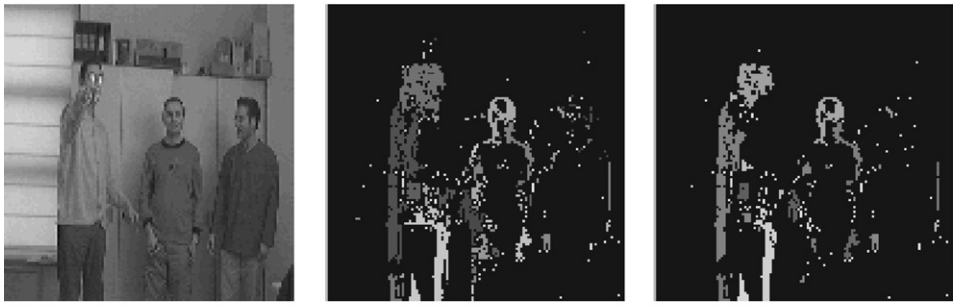

Image #38

Image #51

Fig. 6. Influence of parameter charge value threshold.

6. Conclusions

In this paper, we have introduced a multi-layer architecture that uses the basic parameters of the LIAC model to spatio-temporally build the shapes of all moving objects present in a sequence of images. This novel method for spatio-temporally shape building takes advantage of the inherent motion present in image sequences. With this new model, we can parameterise the aspect of gradually emerging shapes due to motion.

A high enough value of number of grey level bands enables to discriminate the whole shapes of the moving non-rigid objects. Nevertheless, a too high value of this parameter may include some image background into the shapes. This may even lead to fuse more than one different shape into one single silhouette. By tuning the value of discharge value due to motion detection, we obtain more or less information of the history of the movement through the offered trail. Recharge value due to neighbouring allows maintaining attention on those pixels directly or indirectly connected to maximally charge pixels. Parameters number of neighbours in accumulative computation, number of neighbours in charge redistribution and number of neighbours in object fusion are conceived to define the outreach or influence area of lateral interaction of the model. Logically, the higher the action radius, the greater the area affected by the neighbours charges. Lastly, notice that a higher value of permanence value threshold (θ_1), charge value threshold (θ_2) or shape value threshold (θ_3) is able to better discriminate the number of pixels to be used.

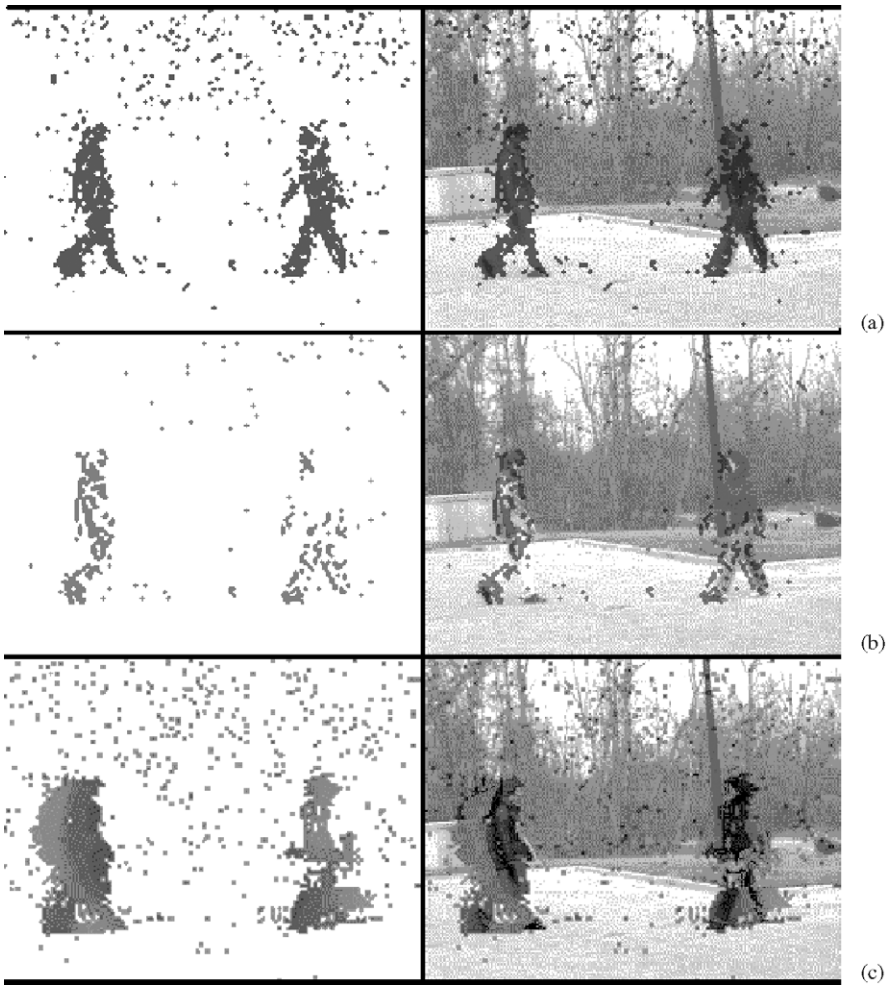

Fig. 7. Three significant examples: (a) silhouette detection, (b) motion detection, (c) direction detection.

Our model may be compared to background subtraction or frame difference algorithms in grey-scale images in the way motion is detected. Then, a region growing technique is performed to define moving objects up to the desired extent. In contrast to similar approaches no complex image pre-processing must be performed and no reference image must be offered to our model.

We also have to highlight that our method has no limitation in the number of non-rigid shapes to differentiate. As in other approaches [25] no object model is used, so our algorithms can segment any kind of moving elements. Our method facilitates any higher-level operation by taking advantage of the common charge value of parts of the moving objects.

The model seems to be promising in a lot of different applications related to image processing. We conclude affirming that the proposed neuronal lateral interaction in accumulative computation mechanisms offer a promising tool for image segmentation as a first approach to pattern recognition. We are currently testing the model in very different real world applications.

Acknowledgements

ToWalkNew sequence, courtesy of University of Maryland Institute for Advanced Computer Studies, College Park.

References

[1] T.S. Huang, Image Sequence Analysis, Springer, Berlin, 1983.
[2] J.K. Aggarwal, N. Nandhakumar, On the computation of motion from sequences of images—a review, Proc. IEEE 76 (8) (1988) 917–935.
[3] J. Wang, T.S. Huang, N. Ahuja, Motion and Structure from Image Sequences, Springer, Berlin, 1993.

[4] J.J. Little, J.E. Boyd, Recognizing people by their gait: the shape of motion, Videre, J. Comput. Vision Res. 1 (2) (1998) 2–32.
[5] J. Mira, A.E. Delgado, J.G. Boticario, F.J. Diez, Aspectos Basicos de la Inteligencia Artificial, Editorial Sanz y Torres, SL, Madrid, 1995.
[6] J. Mira, A.E. Delgado, A. Manjarres, S. Ros, J.R. Alvarez, Cooperative processes at the symbolic level in cerebral dynamics: reliability and fault tolerance, in: R. Moreno-Diaz, J. Mira (Eds.), Brain Processes, Theories and Models, The MIT Press, Cambridge, MA, 1996, pp. 244–255.
[7] R. Polana, R. Nelson, Detecting activities, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1993, pp. 2–7.
[8] R. Polana, R. Nelson, Recognition of nonrigid motion, Proceedings DARPA Image Understanding Workshop, 1994, pp. 1219–1224.
[9] A.F. Bobick, J.W. Davis, An appearance-based representation of action, Proceedings 13th International Conference on Pattern Recognition, 1996, pp. 307–312.
[10] A.F. Bobick, J.W. Davis, Real-time recognition of activity using temporal templates, IEEE Workshop on Applications of Computer Vision, 1996, pp. 1233–1251.
[11] J.W. Davis, A.F. Bobick, The representation and recognition of human movement using temporal templates, Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 928–934.
[12] M.-H. Yang, N. Ahuja, Extracting gestural motion trajectories, Proceedings Second International Conference on Automatic Face and Gesture Recognition, 1998, pp. 10–15.
[13] N. Ahuja, A transform for multiscale image segmentation by integrated edge and region detection, IEEE Trans. Pattern Anal. Machine Intell. 18 (12) (1996) 1211–1235.
[14] T. Olson, F. Brill, Moving object detection and event recognition algorithms for smart cameras, Proceedings DARPA Image Understanding Workshop, 1997, pp. 159–175.
[15] M.A. Fernandez, J. Mira, Permanence memory: a system for real time motion analysis in image sequences, IAPR Workshop on Machine Vision Applications MVA'92, 1992, pp. 249–252.
[16] M.A. Fernandez, J. Mira, M.T. Lopez, J.R. Alvarez, A. Manjarres, S. Barro, Local accumulation of persistent activity at synaptic level: application to motion analysis, in: J. Mira, F. Sandoval (Eds.), From Natural to Artificial Neural Computation IWANN'95, Lecture Notes in Computer Science, Vol. 930, Springer, Berlin, 1995, pp. 137–143.
[17] M.A. Fernandez, Una arquitectura neuronal para la detección de blancos móviles, Ph.D. Dissertation, UNED, 1996.
[18] A. Fernandez-Caballero, Modelos de interaccion lateral en computacion acumulativa para la obtencion de siluetas, Ph.D. Dissertation, UNED, 2001.
[19] A. Fernandez-Caballero, J. Mira, M.A. Fernandez, M.T. Lopez, Segmentation from motion of non-rigid objects by neuronal lateral interaction, Pattern Recognition Lett. 22 (14) (2001) 1517–1524.
[20] A.E. Delgado, J. Mira, R. Moreno-Diaz, A neurocibernetic model of modal co-operative decision in the Kilmer–Mcculloch space, Kybernetes 18 (3) (1989) 48–57.
[21] J. Mira, A.E. Delgado, R. Moreno-Diaz, Cooperative processes in cerebral dynamic, in: D.G. Lainiotis, N.S. Tzannes (Eds.), Applications of Information and Control Systems 3, D. Reidel Pub. Comp., Dordrecht, Holland, 1979, pp. 273–280.
[22] J. Mira, A.E. Delgado, J.R. Alvarez, A.P. de Madrid, M. Santos, Towards more realistic self contained models of neurons: high-order, recurrence and local learning, in: J. Mira, J. Cabestany, A. Prieto (Eds.), New Trends in Neural Computation IWANN'93, Lecture Notes in Computer Science, Vol. 686, Springer, Berlin, 1993, pp. 55–62.
[23] R. Moreno-Diaz, F. Rubio, J. Mira, Aplicacion de las transformaciones integrales al proceso de datos en la retina, Rev. Automat. 5 (1969) 7–17.
[24] I. Haritaoglu, D. Harwood, L.S. Davis, W^4: Who? When? Where? What? A real time system for detecting and tracking people, Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 222–227.
[25] L. Wiskott, Segmentation from motion: combining Gabor- and Mallat-wavelets to overcome the aperture and correspondence problems, Pattern Recognition 21 (10) (1999) 1751–1766.

About the Author—ANTONIO FERNANDEZ-CABALLERO received his M.Sc. in Computer Science from the Polytechnic University of Madrid (Spain) in 1993 and his Ph.D. from the Department of Artificial Intelligence of the National University for Distance Education (Spain) in 2001. His research interests are mainly in Image Processing, Neural Networks and Agents Technology. He is currently an Associate Professor in the Department of Computer Science of the University of Castilla-La Mancha (Spain).

About the Author—MIGUEL A. FERNANDEZ received his M.Sc. in Physics from the University of Granada (Spain) in 1987 and his Ph.D. from the Department of Artificial Intelligence of the National University for Distance Education (Spain) in 1996. His research interests are mainly in Image Processing and Neural Networks. He is currently an Assistant Professor in the Department of Computer Science of the University of Castilla-La Mancha (Spain).

About the Author—JOSE MIRA is a Professor of Computer Science and Artificial Intelligence and Head of the Department of Artificial Intelligence, National University for Distance Education (Spain). His current research interests include AI fundamentals from the perspective of a knowledge-modelling discipline similar to electronics engineering, neural modelling of biological structures (and application of these models to the design of more realistic artificial neural nets), and integration of symbolic and connectionist problem solving methods in the design of hybrid knowledge based systems. He is the General Chairman of the International Work Conferences on Natural and Artificial Neural Networks.

About the Author—ANA E. DELGADO is an Associate Professor of Computer Science and Artificial Intelligence of the Department of Artificial Intelligence, National University for Distance Education (Spain). Her current research interests include neural modelling of biological structures, algorithmic lateral inhibition, co-operative processes of biological inspiration and fault-tolerant computation.