NEUROCOMPUTING

Neurocomputing 50 (2003) 341–364

www.elsevier.com/locate/neucom

Lateral interaction in accumulative computation: a model for motion detection

Antonio Fernández-Caballero[a,*], José Mira Mira[b], Ana E. Delgado[b], Miguel A. Fernández Graciani[a]

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

Received 22 March 2001; accepted 17 February 2002

Abstract

Some of the major computer vision techniques make use of neural nets. In this paper we present a novel model based on neural networks denominated lateral interaction in accumulative computation (LIAC). This model is based on a series of neuronal models in one layer, namely the local accumulative computation model, the double time scale model and the recurrent lateral interaction model. The LIAC model usefulness in the general task of motion detection may be appreciated by means of some significant examples of object detection in indefinite sequences of synthetic and real images. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Accumulative computation; Lateral interaction; Double time scale; Motion detection; Image sequences

1. Introduction

For the purpose of categorisation, computer vision techniques can be divided into two main groups [46]. (1) Model-based (or top-down) techniques use domain specific knowledge to construct template models of what is expected in an image and then try to fit the models to the image data. (2) Data-driven (or bottom-up) approaches that are appropriate when no a priori knowledge of the image contents is available, when the extent of variation of expected structures in the image can be high, or when construction of a model is very difficult. Such techniques exploit, as much as possible, the data in the image in three general steps. (a) The first is image pre-processing to

* Corresponding author.
E-mail address: caballer@info-ab.uclm.es (A. Fernández-Caballero).

PII: S0925-2312(02)00571-4

highlight useful data in the image and suppress noise or unwanted data. (b) The next step is segmentation, which results in pixels in the image being grouped together to form regions that can correspond to structures or structured parts. (c) The final step is image understanding, in which regions are described by an appropriate set of features, and these descriptions are used to determine correspondences between the segmented regions and expected structures.

Some of the major computer vision techniques make use of neural nets [3,24]. In this paper we present a novel model based on neural networks called lateral interaction in accumulative computation (LIAC) as well as its usefulness in the task of motion detection. This paper advances in the generalisation of the algorithmic lateral inhibition (ALI), introduced by Mira et al. [36] and Delgado et al. [4], by using a model of accumulative computation over temporal expansions of the input and output spaces and using a double time scale [11,14]. Furthermore, we explore the validity of our model for the detection of moving elements in real time and in noisy environments [12].

2. The LIAC model description

2.1. Inspiration

The usual paradigm in how to interpret the neuronal processes is the computational paradigm [27,31]. This paradigm begins differentiating between environment and system. The system's behaviour is described in terms of a set of inputs (stimuli), a set of outputs (answers) and a set of computing processes. A neuronal model in one layer is just a generalisation of the proposal of computation by layers [30,37]. Here, we have co-operative modules (local inference rules) that sample information in the input and output spaces according to the size and the shape of their functional receptive fields and discharge their results in the corresponding positions of the output space. Fig. 1 shows the general functional diagram of a layer of neurones [34].

The neuronal model presented in this paper is based on an accumulative computation function [13], followed by a set of co-operating lateral interaction processes performed on a functional receptive field organised as centre–periphery over temporal expansions of their input spaces [33,40]. The LIAC model is fully inspired in (a) the local accumulative computation model, (b) the double time scale model, and (c) the recurrent lateral interaction model. These three models are next introduced.

Notably different styles of local computation are present in the cortex. If computation has to be an appropriate paradigm to describe the nervous system, it should be distributed along the whole tissue with functional multiplicity in each neurone. The current state of the art of knowledge includes, among other, the following computational properties of a real neurone [34]. (1) There exist quick and slow synapses, allowing the coexistence of two intercommunication levels (answer and learning). (2) Synapses may be excitatory or inhibitory, and the whole analytic computation is based on this distinction (sum and subtract; accumulate and decrement, etc.). (3) The synapses of chemical nature are selective to different kinds of messengers, allowing the neuronal computation

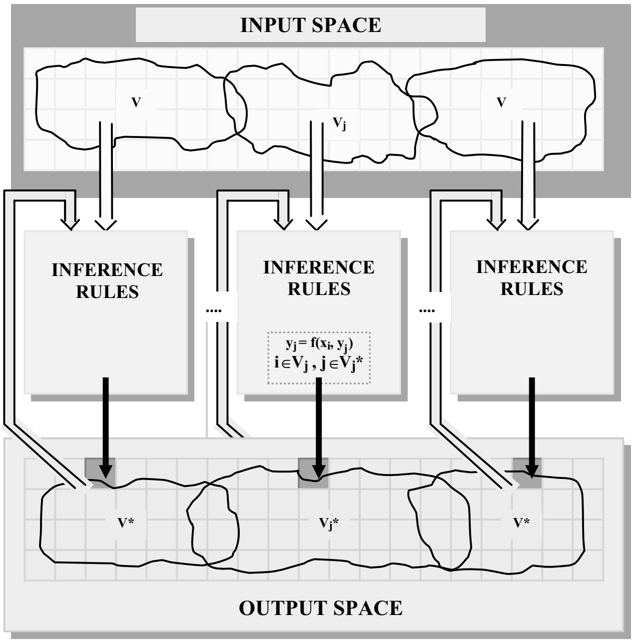


Fig. 1. Neural computation scheme in one layer.

to be interpreted as message passing and specific conditionals. (4) Synapses act in a co-operative and competitive way. (5) The most usual formulation of competitiveness is lateral inhibition that diminishes a unit's answer when its neighbours are active. (6) Neuronal contacts are plastic and they constitute one of the structural supports for self-programming (learning) and memory. Biology offers evidence of very complex local computation modes [33]. Some examples are: (a) a logarithmic translation in sensors, and a latter computation on that logarithmic representation; (b) a multiplicative pre- and post-synaptic inhibition and facilitation; (c) an adaptive threshold function with absolute and relative refractory periods; (d) a non-analytic computation that demands the use of conditionals and other control structures; and (e) a spontaneous activity and different response kinds. Some of the options presented so far have led to the local accumulative computation model [13]. Fig. 2 shows a diagram of the original local accumulative computation model.

Mature neurones are specialised cells in integrating and distributing global states of excitement in the receptive fields that respond as action potentials. However, the diversity of contacts (axon–soma, dendrite–dendrite, dendrite–soma, etc.), together with the computation specificity and the complexity made by these contacts, lead us to think of the existence of a sub-cellular micro-computation [33]. That is to say, there already exists enough experimental evidence to formulate a synaptic computational model before arriving to the neurone level. By only considering the existent local process in the dendrite field, we may find the following properties [14]: (a) a local

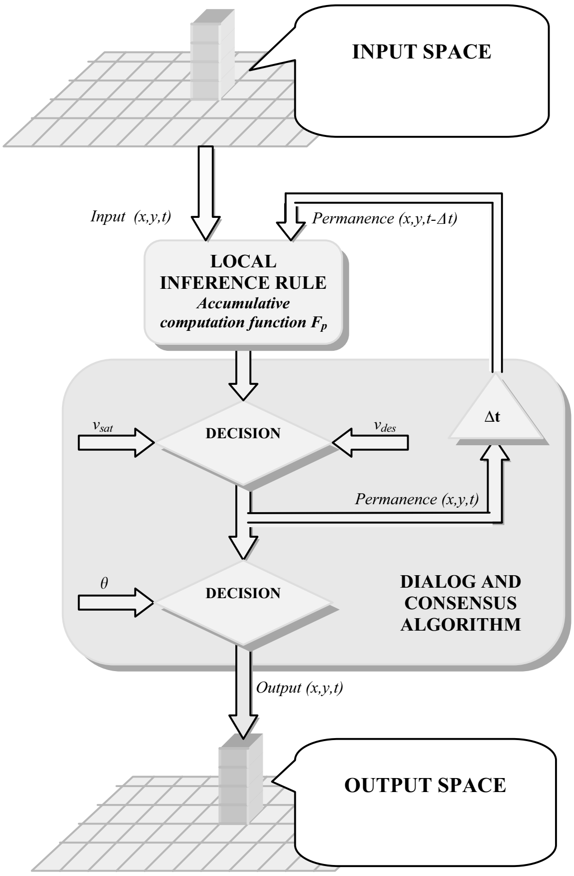


Fig. 2. Local accumulative computation model.

convergent process around each element, (b) a semiautonomous functioning, with each element capable of spatio-temporary accumulation of local inputs in time scale T, and conditional discharge, and (c) an attenuated transmission of these accumulations of persistent coincidences towards the periphery that integrates at global time scale t. Therefore, we face two different time scales: (a) the local time T, and (b) the global time t, $T \ll t$. Fig. 3, inspired by Mira et al. [35], shows an appropriate framework to model co-operative processes at symbolic level and to pick up the two time scales seen previously.

In lateral interaction models [35], you have a layer of modules of the same type with local connectivity, such that the response of a given element does not only depend on its own inputs, but also on the inputs and outputs of the element's neighbours (Fig. 4). From a computational point of view, lateral interaction nets divide the input space into three regions, centre, periphery, and excluded. The following steps have to be performed: (a) a processing over the central region, (b) a processing over the signals

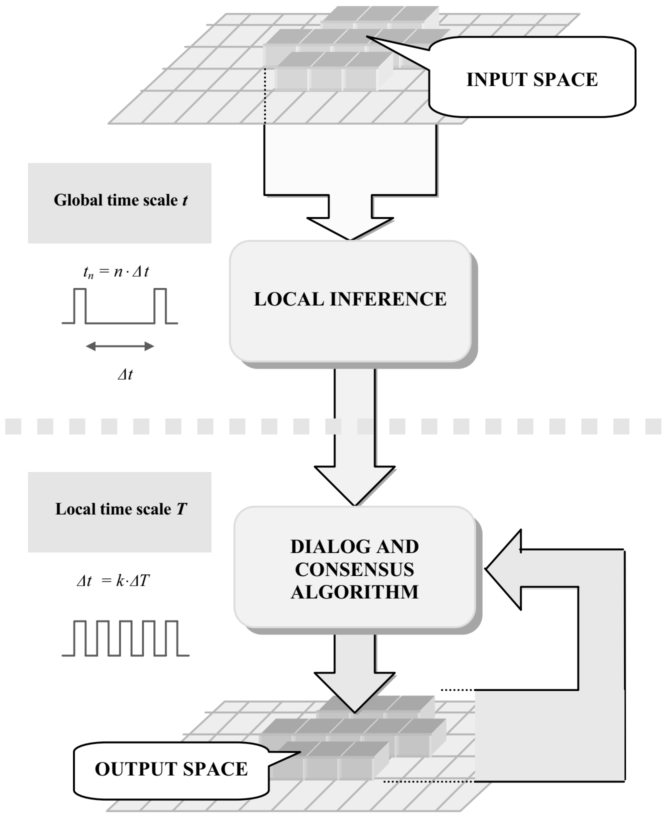


Fig. 3. Double time scale computation model.

coming from the periphery zone, (c) a comparison of the results of these operations and a local decision generation, and (d) a distribution over the output space.

Graphically, we have the recurrent lateral interaction model as depicted in Fig. 5.

The most interesting point in these kinds of recurrent inhibition connectivity schemes is that they exist in practically any level of vertebrates and spineless, from the first layers of the sensorial pathway (retina) up to the most central layers of the cerebral cortex. The other interesting aspect is that lateral interaction networks define computational families that can be projected according to their nature (analogic, logic or inferential) when selecting the synaptic “operator”.

2.2. Algorithmic lateral inhibition in perspective

There are some neuronal structures that appear from the neurogenesis up to the different integration levels (dendro-dendrite contacts, neurones, columns, etc.). Among these structures, lateral inhibition is probably the one that greater computational benefits has offered as an inspiration source for the design of neural artificial networks

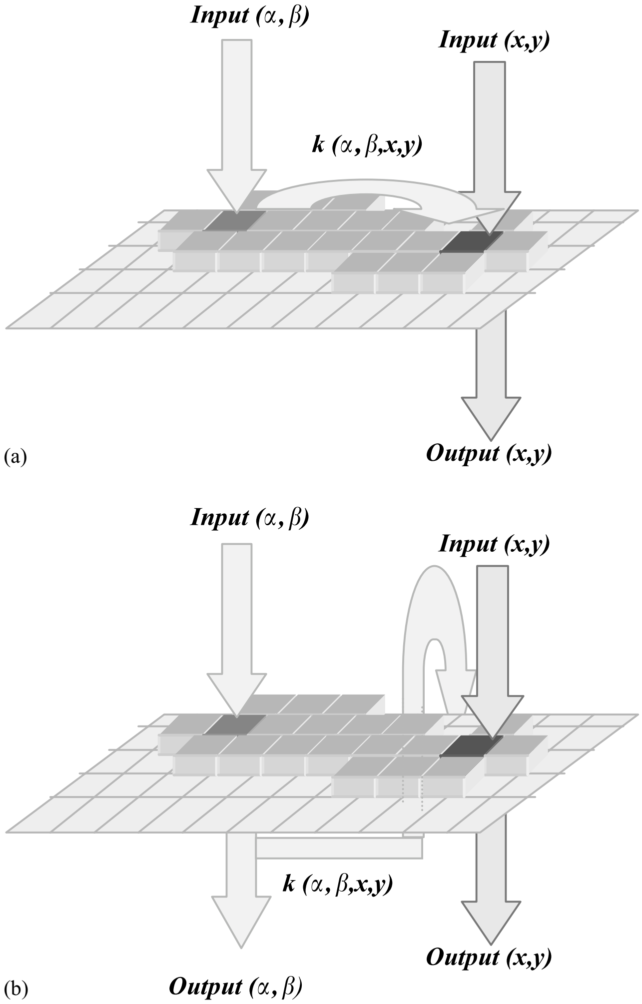


Fig. 4. Lateral interaction: (a) non-recurrent, and (b) recurrent.

able to detect spatio-temporal changes, to filter adaptively, or to pre-process data in self-organising networks, for example.

Lateral inhibition is present in neurocomputing since the first works of Ernst Mach on the interdependence of neighbouring elements in the retina and the subjective visual phenomena, which seemed to fully explain the apparent doubling of the single broad lines in the spectrogram [43]. The experimental research of Hartline and Ratliff on the composite eye of the limulus, and the precise mathematical formulation of the analytic lateral inhibition model, lead to explain and to predict the electro-physiological data observed [17].

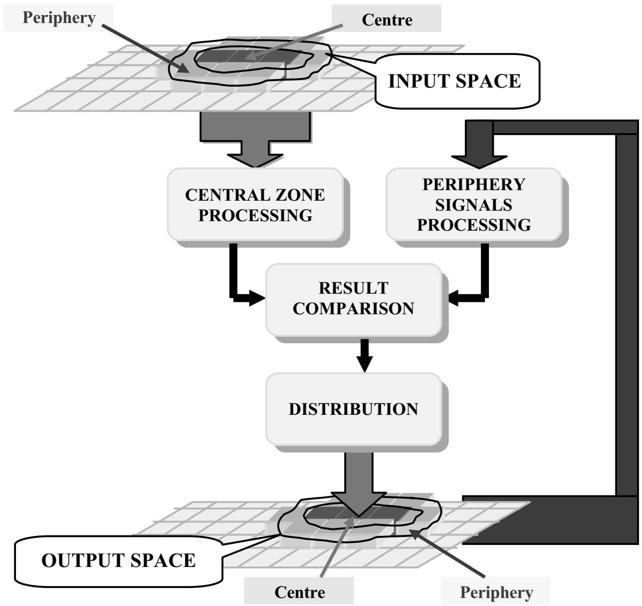


Fig. 5. Recurrent lateral interaction model.

Hubel and Wiesel [23], using the concept of receptive field, made an extensive use of the concept, and Delgado [8,9], Delgado et al. [4] and Mira et al. [38] extended the concept of lateral inhibition to embrace a wider group of operators than the lineal and non-lineal analytic one. We are computationally speaking of lateral interaction algorithms, with non-linealities of the type "if–then", local memory and sequential control. Finally, lateral interaction turns into a "problem-solving method" in the usual sense of the artificial intelligence [6]. It decomposes classification tasks where there is much more data than knowledge. This way, a lateral interaction network is a connectionist method of local calculation with learning capacity and appropriate to solve a wide family of spatio-temporal problems that need to reduce the dimensionality of the data to get an efficient treatment in real time [32].

The more usual algorithmic expression in neuronal calculation are the competitive nets, of the gradual or of the winner take all (WTA) type, previous, for example, to the mechanisms of self-organising learning [20,19]. Here, the maximum selection operator ("hard competition") and the corresponding "if–then" are the instantiation of the algorithmic character of lateral inhibition ("soft competition").

2.3. The lateral interaction in accumulative computation model

All general notions previously considered are incorporated in the model proposed. Common biological notions such as accumulative computation, lateral interaction and double delay are included.

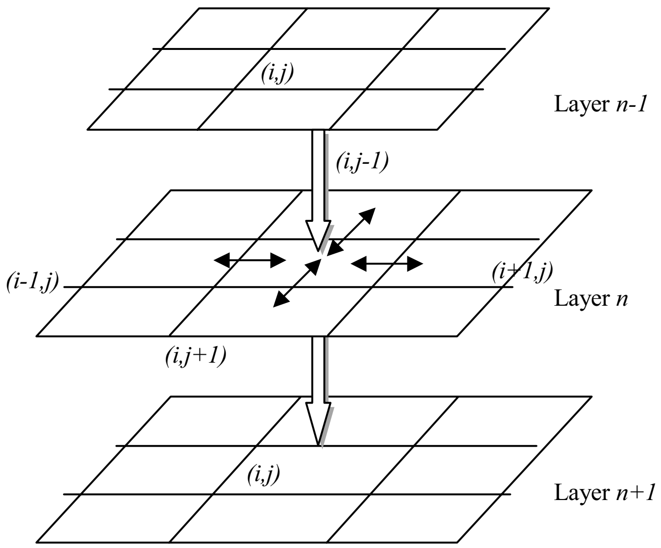


Fig. 6. Module (x, y) and its connections to the preceding layer, the following layer and its lateral connections to its four next neighbours.

Remember that at a symbolic level, we have representational input and output spaces and a set of co-operating processes that primarily perform local inference functions over data of a functional receptive field organised as centre–periphery and interchanging basic information given and received by each process in the dialogue area. After individual and local inference, a recurrent use of dialogue and consensus algorithms is performed comparing individual opinions with those of the periphery neighbours.

Our model is restricted to the following characteristics: (1) Application of accumulative computation to each central element starting from the charge value of the own element and the input coming from the previous layer. (2) Application of lateral interaction mechanisms (lateral inhibition) of recurrent type from periphery, where all neighbours have the same weight in lateral interaction, on the centre (x, y) by using specific channels incorporating the possibility of opening and closing themselves. (3) Global time scale t is used (a) for reading from the input of the previous layer, (b) during the process of accumulative computation, and (c) for writing of the output to the following layer. (4) Local time scale $T \ll t$ is used in all lateral co-operative interaction mechanisms.

This set of specific characteristics permit to unite the computational functions of the local accumulative computation model, the double time scale model and the recurrent lateral interaction model to produce the lateral interaction in accumulative computation model in one layer. These characteristics lead to a model where each module (x, y) in a layer may be represented as showed in Fig. 6. Graphically, we can represent the lateral interaction in accumulative computation model as shown in Fig. 7.

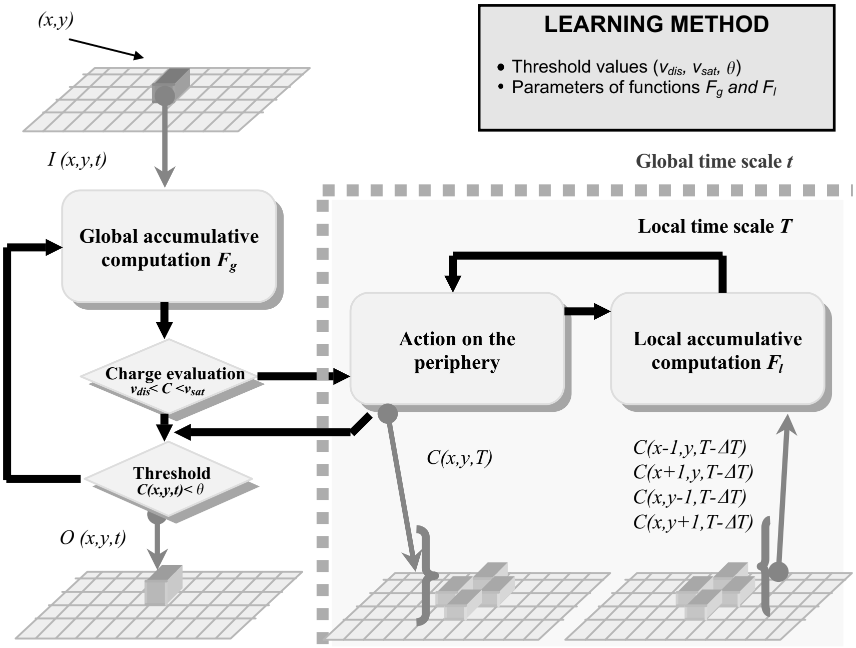


Fig. 7. Lateral interaction in accumulative computation neuronal model.

The co-operative process carried out by a functional group of neurones is characterised at symbolic level by the following parameters:

$$\langle G, I(x, y, t), O(x, y, t), C(x, y, t), F_{\mathrm{g}}, F_{\mathrm{l}}, v_{\mathrm{dis}}, v_{\mathrm{sat}}, \theta\rangle,$$

where G is the global function at knowledge level; to define, in each case, when using the lateral interaction in accumulative computation model applied to a specific problem, $I(x, y, t)$ the input from the previous layer into element (x, y) at instant t, $O(x, y, t)$ the output toward element (x, y) of the following layer at instant t; discharge value, $C(x, y, t)$ the charge value of element (x, y) at instant t, F_{g} the spatial accumulation function in global time scale t; this function is domain dependent of the global function G at knowledge level, F_{l} the spatio-temporal accumulation function in local time space T; this function is also domain dependent of the global function G at knowledge level, v_{dis} the minimum charge value; discharge value, v_{sat} the maximum charge value; saturation value, and θ the minimum discharge threshold value.

Global function G may easily be understood if thinking in a generic neuronal network, with classification functions, organised in different functional groups that calculate in a concurrent and/or sequential manner, as at least so many types of neurones as layers. This way, the different layers would be associated to the decomposition of the problem in different tasks, as, for example, characteristics extraction, distance measures, maximum selectors, etc.

Particular mention deserves the learning method. In this co-operative model it allows redefining all parameters, depending on the global function G. In our lateral interaction in accumulative computation model, the parameters that modify their value during the learning phase are $v_{\text{dis}}, v_{\text{sat}}$ and θ. The influence of threshold variability has largely been studied [5,25]. Other parameters with learning capacities will appear as consequence of the use of concrete F_{g} and F_{l} functions.

The concept of adaptive neighbourhood, widely used in artificial vision (e.g. [41,47]) to select a working scale at each point of the image and for each computation level, is of great interest in the lateral interaction in accumulative computation model. The idea is that the size and the pattern of the periphery depend on the characteristics of the image data through parameters that define the homogeneity measure of each pixel. The dynamic adaptation (learning) of different peripheries allows estimating different properties. Evidently, any similarity concept should be injected as external knowledge in the design of the net.

At last, let us face the stability of the model proposed. Stability is related to dynamic systems that incorporate nodes with positive feedback. But this it is not the case of our model, where the local dynamics is under the control of finite state automata and the dynamic range of the variables is enclosed between two saturation values. There are no coupled oscillators, in the sense of Terman and Wang [48], like in LEGION (locally excitatory globally inhibitory oscillator network) system [7].

2.4. The model's formulation

There are three general steps associated to the lateral interaction in accumulative computation model, as shown in Fig. 8. These are a spatial accumulation at global time scale, a spatio-temporal accumulation at local time scale, and a distribution at global time scale.

2.4.1. Step 1: Formulation at global time scale t

In first place, the global accumulative computation function computes the charge value of any of the central elements (x,y) at global time instant t, $C(x,y,t)$, from the charge value at the preceding instant, $C(x,y,t-\Delta t)$, and the input value $I(x,y,t)$, into the interval $[v_{\text{dis}}, v_{\text{sat}}]$.

$$C(x,y,t) = F_{\text{g}}[C(x,y,t-\Delta t), I(x,y,t)]$$

$$\text{if } C(x,y,t) \geqslant v_{\text{sat}} \quad \text{then } C(x,y,t) = v_{\text{sat}}$$

$$\text{if } C(x,y,t) \leqslant v_{\text{dis}} \quad \text{then } C(x,y,t) = v_{\text{dis}}.$$

This way, we have obtained the value of the individual opinion of element (x,y) starting from the input value to the element and the history (FIFO memory) of the state of the own element.

2.4.2. Step 2: Formulation at local time scale T

At local time scale T we have an iterative phase of upgrading the charge value from the lateral interaction of each element (x,y) with its four closest neighbours. This step

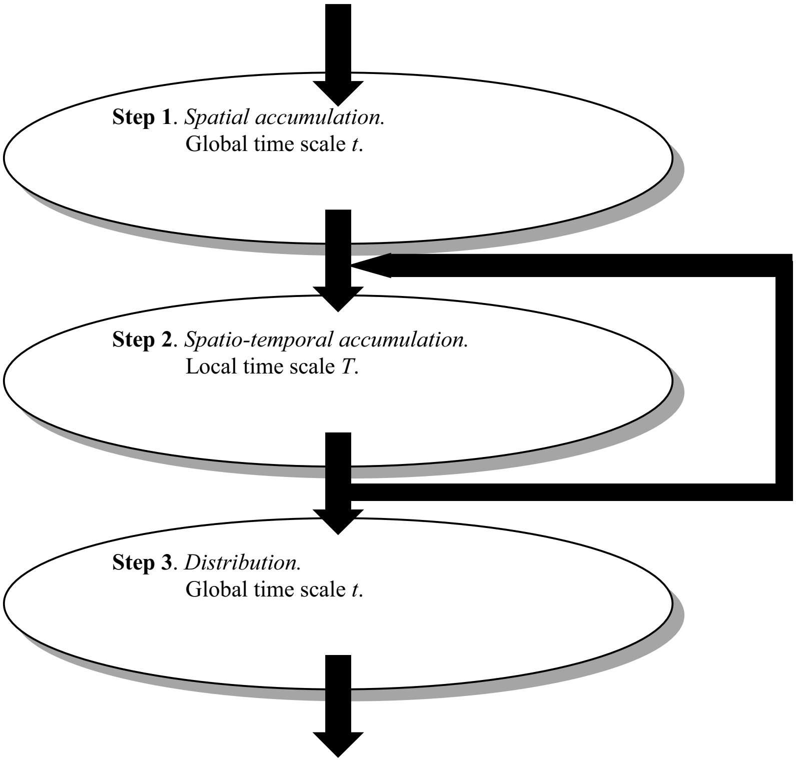


Fig. 8. General steps of the computational model.

takes place only if the central element (x, y) possesses a sufficient charge between limits v_{dis} and v_{sat}. Note that global time $t = n\Delta t$ is frozen during the dialog, but not local time $T = k\Delta T$.

A re-charge of module (x, y) is carried out, starting from the available information in the module's periphery.

$$\forall(\alpha, \beta) \in [x \pm 1, y \pm 1], \quad C(x, y, T) = F_1[C(x, y, T - \Delta T), C(\alpha, \beta, T - \Delta T)]$$

$$\text{if } C(x, y, T) \geqslant v_{\mathrm{sat}} \quad \text{then } C(x, y, T) = v_{\mathrm{sat}}$$

$$\text{if } C(x, y, T) \leqslant v_{\mathrm{dis}} \quad \text{then } C(x, y, T) = v_{\mathrm{dis}}.$$

This is the way each individual opinion is compared to that of its closest neighbours.

2.4.3. *Step 3: Formulation at global time scale t*

Lastly, back at global time scale t, the output $O(x, y, t)$ takes place. The output is a result of:

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

where θ is a threshold function. This output is the result of a consensus algorithm that has produced a new charge value starting from the combination of all the affected co-operative elements.

3. Application of LIAC model to motion detection

3.1. Problem statement

Detection of moving elements has been largely studied over the past decades. Sereno [44], and more recently Çesmeli and Wang [7] have performed an excellent classification of computational models in pattern motion. Models of local motion detection are gradient models [10,22,28,29], correlation models [1,4,18,50] and image difference models [15,45], whereas models of pattern motion measurement can be divided into models that incorporate multiple motion constraints [10], matching models [42,49], and models that use a smoothness constraint [21,22,26,51]. In particular, the studies on non-rigid object motion are among the most important cues in motion analysis [39]. The growing interest is motivated by their use in a great number of applications in areas as different as medicine or surveillance.

As stated by Çesmeli and Wang [7] the three major challenges in computational investigation of motion perception are the aperture problem [2], the blank wall problem [45] and the motion transparency problem [16]. These challenges are faced by our model applied to motion detection, as well as a fourth one, which arises when working with real moving non-rigid objects. This is deformation processing which is very expensive, just as motion may be highly complex and the object's shape is usually uncertain.

The problem we are putting forward is the discrimination of a set of objects capable of holding our attention in a scene. These objects are detected from the motion of any of their parts. Detected in an indefinite sequence of images, motion allows obtaining the silhouettes of the moving elements. If any element stops moving, it does not get any attention. This way, interest on that particular silhouette declines, so that it does not belong to the discriminated objects.

As it has already been mentioned, the way these elements are obtained will only provide information of existent motion in the scene. This is a characteristic problem of motion analysis. Up to some extent, the method can be generically classified into the models based on image difference. More concretely, it is linked to the generic behaviour of permanence memories [13]. Specifically, we will say that the observer is unable to discern any object unless it goes on moving. Therefore, an image where all the elements are static does not contribute to provide any information. In other words, the system will only pay attention to those pixels of an image of 512×512 pixels in which a variation in their grey level is detected.

In real scenes, all the components of the object do not have to move at the same time. For example, the human body (the object, in this case) is made up of a great number of members that do not move simultaneously. This way, a person sitting in front of the computer writing a letter will move his hands, part of his arms, his head, etc., but will be able to have his legs perfectly still. The system is able to detect those parts of the object in movement. That is to say, the object's silhouette is only composed of the hands, the arms and the head. Nevertheless, if other parts of the same object set in motion, it will be fundamental to be able to associate them to the parts previously detected, obtaining the object as one whole through the pertinent association mechanisms. This way, the proposed system is able to detect and even to associate all the parts of a moving object.

The problem is not limited to the observation, detection and pursuit of a single object in the scene, but rather it will be able to differentiate among all the objects that present some motion. Our particular solution for the detection of moving elements using the model presented is next illustrated.

3.2. Lateral interaction in accumulative computation model application

Firstly, the general accumulative computation equation present at global time scale has to be reformulated for the specific motion analysis problem we are facing.

We need k layers, one for each grey level present at the raw image to work with. Each element (x, y) at any of the sub-layers is capable of processing the motion of pixel $I(x, y, t)$, starting from its grey level and its charge value. The effect of function F_g on the algorithm for this first step is as follows:

$$
\begin{aligned}
C_k(x,y,t) &= F_g[C_k(x,y,t-\Delta t, I(x,y,t)] \\
&= \begin{cases} v_{dis} & \text{if } I(x,y,t) \neq k, \\ v_{sat} & \text{if } I(x,y,t) = k \text{ and } I(x,y,t-\Delta t) \neq k, \\ C_k(x,y,t-\Delta t) - v_{acc} & \text{if } I(x,y,t) = k \text{ and } I(x,y,t-\Delta t) \neq k, \end{cases} \\
&\text{if } C_k(x,y,t) \geqslant v_{sat} \text{ then } C_k(x,y,t) = v_{sat}, \\
&\text{if } C_k(x,y,t) \leqslant v_{dis} \text{ then } C_k(x,y,t) = v_{dis},
\end{aligned}
$$

where v_{acc} is the accumulative computation discharge value due to motion detection. This effect may be explained the following way:

(a) element (x, y) at grey level k is completely discharged if pixel $I(x, y, t)$ is not at grey level k,
(b) element (x, y) at grey level k gets the complete charge (saturation) if pixel $I(x, y, t)$ has changed to grey level k from instant $t - \Delta t$ to instant t,
(c) element (x, y) at grey level k is partially discharged by a value of v_{acc} if pixel $I(x, y, t)$ is at grey level k and has not changed since $t - \Delta t$.

Secondly, in local time scale T, and in an iterative way, the charge value is homogeneously distributed among all the elements that have the same grey level stripe value.

Four are the neighbours taken into account, namely those four surrounding element (x, y).

A homogeneous distribution may obviously be obtained by calculating the mean value of the five charge values. This is possible establishing the following formula for function F_l:

$$\forall(\alpha,\beta)\in[x\pm 1, y\pm 1],\quad \delta_{\alpha,\beta}=\begin{cases}1 & \text{if } C_k(\alpha,\beta,T-\Delta T)>v_{\mathrm{dis}},\\ 0 & \text{otherwise,}\end{cases}$$

$$\begin{aligned}
&C_k(x,y,T)\\
&\quad = F_l[C_k(x,y,T-\Delta T), C_k(\alpha,\beta,T-\Delta T)]\\
&\quad = \frac{1}{1+\delta_{x-1,y}+\delta_{x+1,y}+\delta_{x,y-1}+\delta_{x,y+1}}\\
&\qquad \times \left[\begin{array}{l} C_k(x,y,T-\Delta T)\\ \quad +\delta_{x-1,y}C_k(x-1,y,T-\Delta T)+\delta_{x+1,y}C_k(x+1,y,T-\Delta T)\\ \quad +\delta_{x,y-1}C_k(y-1,y,T-\Delta T)+\delta_{x,y+1}C_k(x,y+1,T-\Delta T)\end{array}\right].
\end{aligned}$$

In this way, attention can be maintained on the charged elements that are connected to saturate elements.

The last step consists now in distributing output values. We take the maximum value of all outputs of the k sublayers to show the silhouette of a moving object.

$$O_k(x,y,t)=\begin{cases}C_k(x,y,t) & \text{if } C_k(x,y,t)>\theta,\\ \theta & \text{otherwise,}\end{cases}$$

$$O(x,y,t)=\max_k[O_k(x,y,t)].$$

4. Data and results

The performance of the model applied to motion detection is demonstrated on two sets of image sequences. The first set includes synthetic scenes to test the model's behaviour in front of the aperture, the blank wall, and the motion transparency. The second set shows a real scene from a traffic control system, illustrating the deformation problem present in most scenes.

4.1. Synthetic images

The three synthetic image sequences are similar to those used by Çesmeli and Wang [7]. The main and important difference is that our image sequences are indefinite, composed of more than three frames. The parameter values used in all these sequences are explained next. The number of grey level bands used is 8. The parameters for accumulative computation are $v_{\mathrm{sat}}=255$, $v_{\mathrm{dis}}=0$, $v_{\mathrm{acc}}=95$.

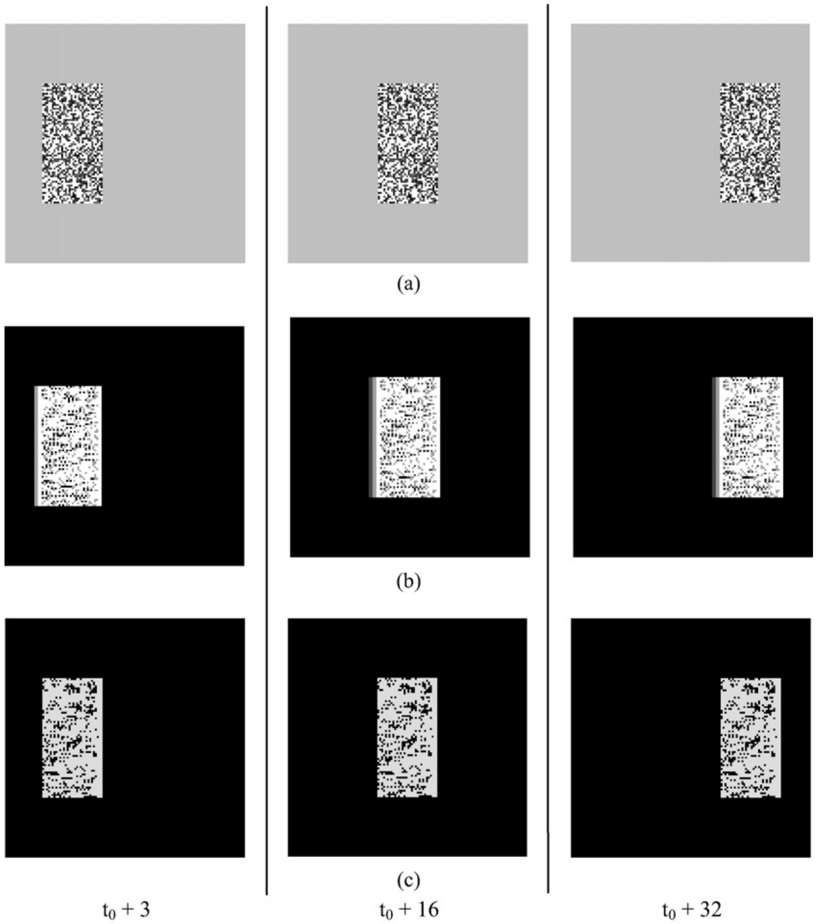
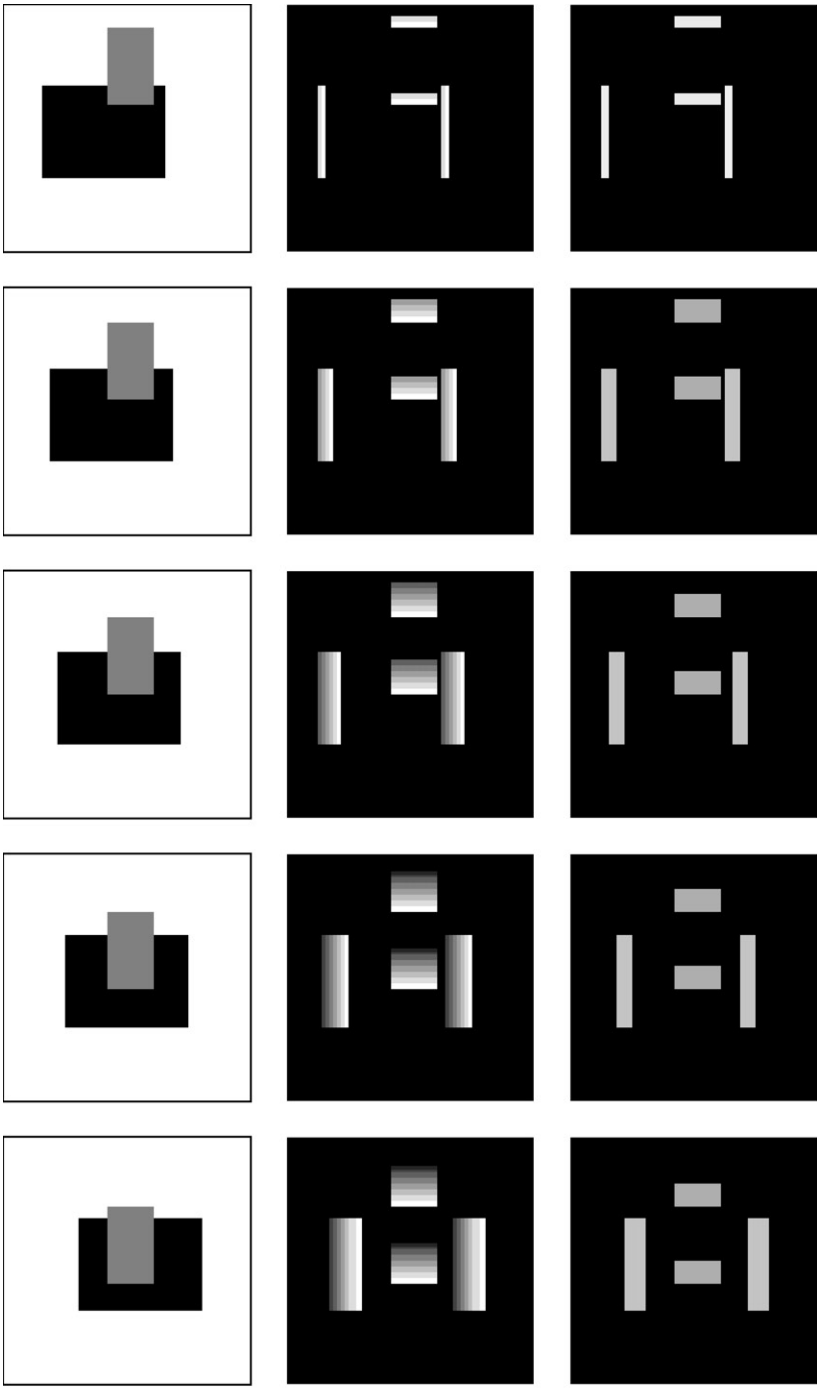


Fig. 9. Application of the model to a synthetic image sequence, illustrating the solution to the aperture problem: (a) test images series, (b) result after global accumulation, and (c) result after distribution.

The first sequence is depicted in Fig. 9. A textured rectangular region is moving two pixels per frame rightward on a homogeneous background as shown in Fig. 9a. Fig. 9b shows the method's output after accumulative computation, whereas Fig. 9c shows the output after lateral interaction. The moving element in Fig. 9b is composed of several charge values due to motion detection. Fig. 9c offers only two charge values for the whole image, namely black for those pixels considered not to have moved and a grey value for the pixels where motion has been detected. Now, it is this common charge value that faces the aperture problem successfully. The moving element is given a same charge value equivalent to a common "velocity" parameter to all pixels of the object.

In Fig. 10, two homogeneous rectangular regions are moving on a homogeneous background. The larger rectangle is moving two pixels per frame rightward and the



(a) (b) (c)

smaller one is moving three pixels per frame downward, occluding the former (Fig. 10a). Results of global accumulation and distribution are again shown in Fig. 10b and c, respectively. Notice that under these conditions the blank wall problem cannot be solved. Fig. 11 shows the same situation, but using textured rectangles. Now, the blank wall problem disappears. To our opinion, the real world is composed of elements where texture is widely present, for moving elements as well as for backgrounds. Thus, the blank wall problem is limited in our model for very special and unusual situations.

Lastly, in the third synthetic scene shown in Fig. 12, two random-dot rectangular regions (Fig. 12b1 and b2) are moving horizontally two pixels per frame in opposite directions (Fig. 12c) on a random-dot noise background (Fig. 12a). During this motion sequence, there is an overlapping area where both motions are simultaneously perceived. Our method perfectly segments moving regions. When both regions are separate, they appear with different charge values. But when they are perceived simultaneously, they get a single common charge value. There is no possibility to differentiate them when superimposed.

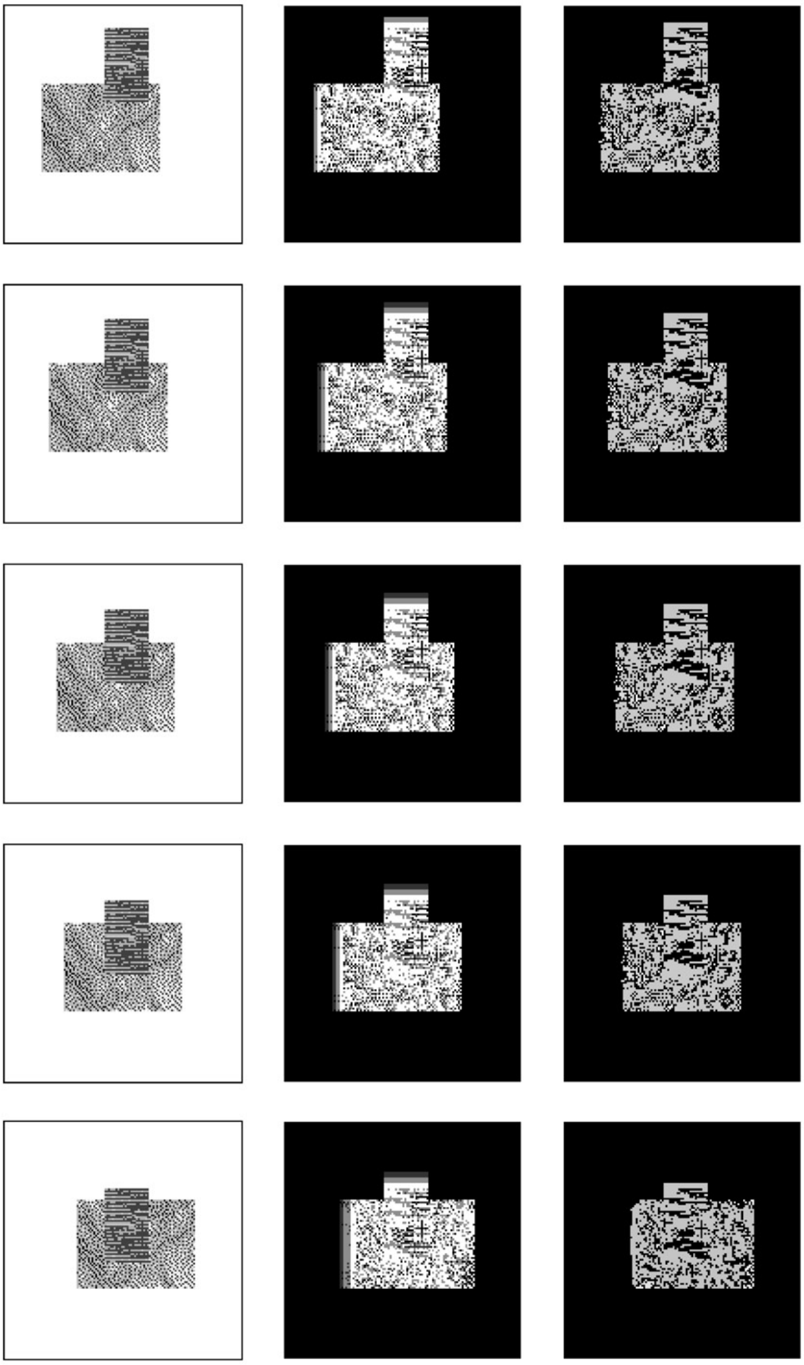
4.2. Real images

We have to highlight that our model applied to motion detection is really useful when used in real scenes. And this usefulness is even increased, as we are able to face deformation processes, and not only translational motion. Let us remember again that the number of images in a sequence is unlimited. In order to show all these advantages of the neuronal model for lateral interaction in accumulative computation in motion detection we have used a series of real scene test images as shown in Fig. 13, column (a). This sequence shows a road traffic scene. Column (b) of Fig. 13 shows the charge values as an output image after accumulative computation has taken place. Column (c), on the other hand, shows the output value after recursive tasks have been performed.

In the second column of that figure a black background can be appreciated where no motion is detected (elements set to value v_{dis}), as well as the conventional result of image difference in white colour. Image difference systems are capable of detecting pixels where motion occurs. See that precisely these pixels are the ones set to a permanence value of v_{sat} in our model. Any other colour appearing in Fig. 13 column (b) or (c) corresponds to one grey level value. So, the elements of Fig. 13, column (b) in any colour different from white or black are those taking an intermediate value between v_{dis} and v_{sat}. The third column of the figure offers more information on the object's shape. This column illustrates the result of applying the complete model on the images. It shows the final results of the use of accumulative computation followed by lateral interaction of maximally charged pixels on partially charged pixels.

So it can be appreciated how the silhouette rapidly appears after the model has been working with only three input images. Note that the model continues performing

Fig. 10. Application of the model to a synthetic image sequence, illustrating the blank wall problem: (a) test images series, (b) result after global accumulation, and (c) result after distribution.



(a) (b) (c)

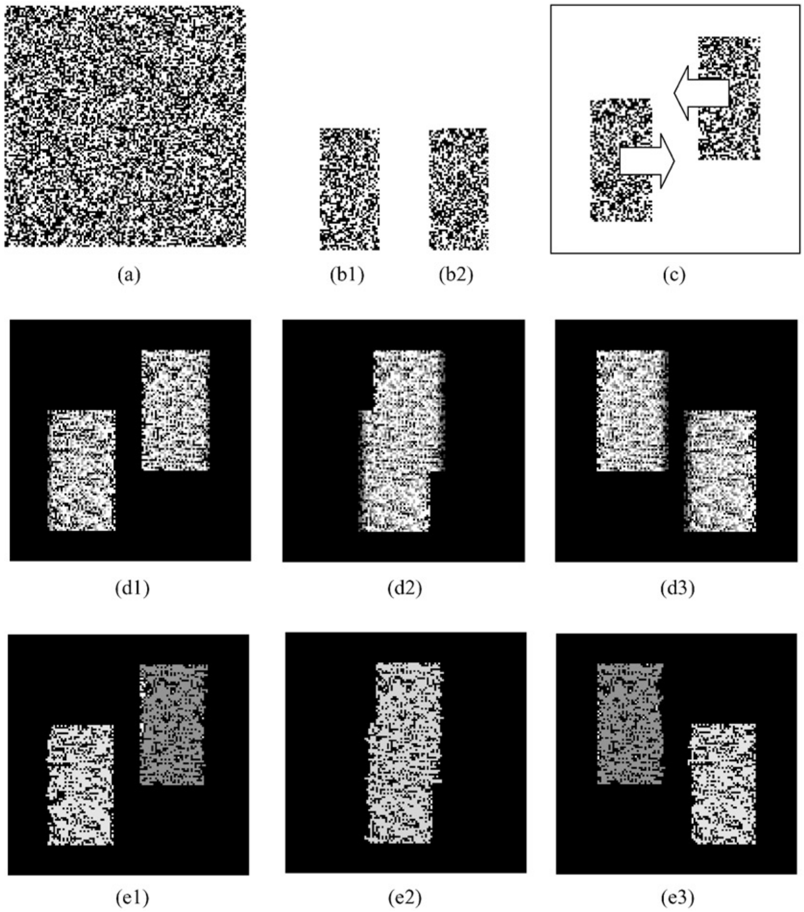


Fig. 12. Application of the model to a synthetic image sequence, illustrating the solution to the motion transparency problem: (a) test images series, (b) result after global accumulation, and (c) result after distribution.

excellent results through time. This example shows the functionality of the model, demonstrating that the presented mechanisms are adequate for moving object detection.

In comparison with other approaches, we have to highlight that the most significant contribution of our model is that it is capable of detecting all elements moving in an indefinite sequence of images with any kind of motion type. The most important limitation of the model applied to motion detection is the impossibility to differentiate among objects that are seen as a whole due to occlusions.

Fig. 11. Application of the model to a synthetic image sequence, illustrating the solution to the blank wall problem: (a) test images series, (b) result after global accumulation, and (c) result after distribution.

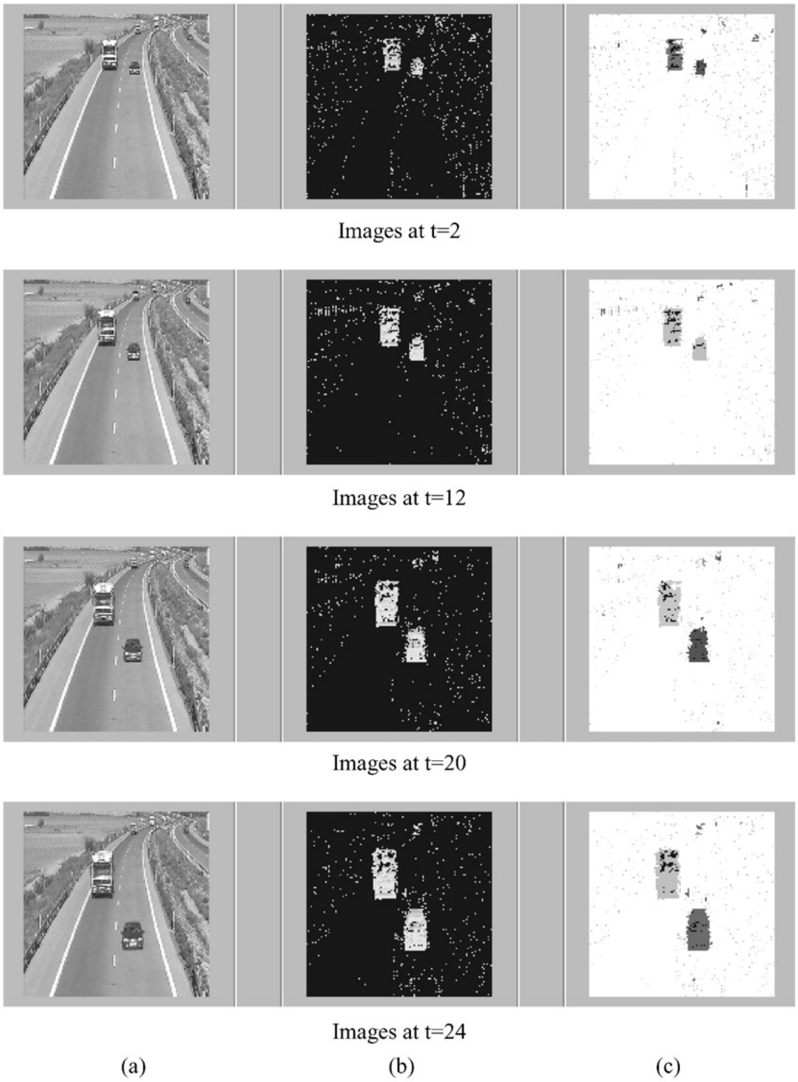


Fig. 13. (a) Test images series. (b) Result after global accumulation. (c) Result after distribution.

5. Conclusions

We have proposed in this paper a model for lateral interaction in accumulative computation and its application to motion detection in a neural architecture. This model, which can be considered as biologically plausible, is based on a series of neuronal models in one layer, namely the local accumulative computation model, the double time scale model and the recurrent lateral interaction model. The model applied to motion detection is capable of holding our attention on all moving objects in a scene composed of an indefinite sequence of images.

Our model is a 2-D approach to motion estimation. In these kinds of approaches, motion estimates are obtained from 2-D motion of intensity patterns. In these methods there is a general restriction: the intensity of the image along the motion trajectory must be constant, that is to say, any change through time in the intensity of a pixel is only due to motion. This restriction does not affect our model at all. This way, our algorithms are prepared to work with lots of situations of the real world, where changes in illumination are of a real importance.

The gradient-based estimates have become the main approach in the applications of computer vision. These methods are computationally efficient and satisfactory motion estimates of the motion field are obtained. Unfortunately, the gradient-based methods always present some restrictions, but our method does not. The disadvantages common to all methods based on the gradient also arise from the logical changes in illumination.

Obviously, a way of solving the former limitations of gradient-based methods is to consider image regions instead of pixels. In general, these methods are less sensitive to noise than gradient-based methods. Our particular approach takes advantage of this fact and uses all available neighbourhood state information as well as the proper motion information. On the other hand, our method is not affected by the greatest disadvantage of region-based methods. Our model does not depend on the pattern of translation motion. In effect, in region-based methods, regions have to remain quite small so that the translation pattern remains valid.

Thus, we can conclude that the proposed model seems promising for many different applications related to image processing. And, as it is a more generally applicable model, we are optimistic on future applications in other neurocomputing fields.

Acknowledgements

The authors would like to thank the anonymous referees for their very helpful comments that significantly improved the manuscript.

References

[1] E.H. Adelson, J. Bergen, Spatiotemporal energy models for the perception of motion, J. Opt. Soc. Amer. A 2 (2) (1985) 284–299.

[2] E.H. Adelson, J. Movshon, Phenomenal coherence of moving visual patterns, Nature 300 (1982) 523–525.

[3] S.M. Bandharkar, J. Koh, M. Suk, Multisale image segmentation using a hierarchical self-organizing map, Neurocomputing 14 (3) (1997) 241–272.

[4] H. Barlow, W. Levick, The mechanism of directionally selective units in rabbit's retina, J. Physiol. 178 (1965) 477–504.

[5] H. Bartsch, M. Stetter, K. Obermayer, The influence of threshold variability on the response of visual cortical neurons, Neurocomputing 32–33 (1–4) (2000) 37–43.

[6] J. Breuker, W. van der Velde, Common KADS Library for Expertise Modelling, IOS Press, Amsterdam, 1994.

[7] E. Çesmeli, D.L Wang, Motion segmentation based on motion/brightness integration and oscillatory correlation, IEEE Trans. Neural Networks 11 (4) (2000) 935–947.
[8] A.E. Delgado, Modelos Neurocibernéticos de Dinámica Cerebral, Ph.D. Dissertation, E.T.S. Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, 1978.
[9] A.E. Delgado, J. Mira, R. Moreno-Diaz, A neurocybernetic model of modal co-operative decision in the Kilmer-McCulloch space, Kybernetes 18 (1989) 48–57.
[10] C. Fennema, W. Thompson, Velocity determination in scenes containing several moving objects, Comput. Graph. Image Process. 9 (1979) 301–315.
[11] M.A. Fernandez, Una Arquitectura Neuronal para la Detección de Blancos Móviles, Ph.D. Dissertation, Facultad de Ciencias, UNED, 1995.
[12] A. Fernandez Caballero, Modelos de Interacción Lateral en Computación Acumulativa para la Obtención de Siluetas, Ph.D. Dissertation, Facultad de Ciencias, UNED, 2001.
[13] M.A. Fernandez, J. Mira, Permanence memory: a system for real time motion analysis in image sequences, in: Proceedings of the IAPR Workshop on Machine Vision Applications, 1992, pp. 249–252.
[14] 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.
[15] K.S. Fu, R.C. Gonzalez, C.S.G. Lee, Robotica. Control, detección, visión e inteligencia, McGraw-Hill, New York, 1988.
[16] E.J. Gibson, J.J. Gibson, O.W. Smith, H. Flock, Motion parallax as a determinant of perceived depth, J. Exp. Psychol. 58 (1) (1959) 40–51.
[17] K.K. Hartline, F. Ratliff, Spatial summation of inhibitory influences in the eye of limulus, Science 120 (1954) 781.
[18] B. Hassenstein, W.E. Reichardt, Functional structure of a mechanism of perception of optical movement, in: Proceedings of the First International Congress of Cybernetics, Namar 1956, pp. 797–801.
[19] M.H. Hassoun, Fundamentals of Artificial Neural Networks, MIT Press, Cambridge, MA, 1995.
[20] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice-Hall, Englewood Cliffs, NJ, 1999.
[21] E.C. Hildreth, The Measurement of Visual Motion, MIT Press, Cambridge, MA, 1984.
[22] B.K.P. Horn, B.G. Schunck, Determining optical flow, Artif. Intell. 17 (1981) 185–203.
[23] D.H. Hubel, T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, J. Physiol. 160 (1962) 106–154.
[24] G. Indiveri, L. Raffo, S.P. Sabatini, G.M. Bisio, A recurrent neural architecture mimicking cortical preattentive vision systems, Neurocomputing 11 (2–4) (1996) 155–170.
[25] L. Itti, C. Koch, J. Braun, A quantitative model relating visual neuronal activity to psychophysical thresholds, Neurocomputing 26–27 (1–3) (1999) 743–747.
[26] C. Koch, J. Marroquin, A. Yuille, Analog "neuronal" networks in early vision, Proc. Natl. Acad. Sci. USA 83 (1986) 4263–4267.
[27] T.S. Kuhn, The Structure of Scientific Revolutions, The University of Chicago Press, Chicago, 1970.
[28] T.B. Lawton, Outputs of paired Gabor filters summed across the background frame of reference predict the direction of movement, IEEE Trans. Biomed. Engng. 36 (1989) 130–139.
[29] D. Marr, S. Ullman, Directional selectivity and its use in early visual processing, Proc. Roy. Soc. London B 211 (1981) 151–180.
[30] W.S. McCulloch, Embodiments of Mind, MIT Press, Cambridge, MA, 1965.
[31] J. Mira, A.E. Delgado, Some reflections on the relationships between neuroscience and computation, in: J. Mira, R. Moreno-Diaz (Eds.), Biological and Artificial Computation: From Neuroscience to Technology, IWANN'97, Lecture Notes in Computer Science, Vol. 1240, Springer, Berlin, 1997, pp. 15–26.
[32] J. Mira, A.E. Delgado, What can we compute with lateral inhibition circuits? in: J. Mira, A. Prieto (Eds.), Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence, Lecture Notes in Computer Science, Vol. 2084, Springer, Berlin, 2001, pp. 38–46.

[33] 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.
[34] J. Mira, A.E. Delgado, J.G. Boticario, F.J. Diez, Aspectos Básicos de la Inteligencia Artificial, Editorial Sanz y Torres, SL, Madrid, 1995.
[35] 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, MIT Press, Cambridge, MA, 1996, pp. 244–255.
[36] 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, Vol. 3, D. Reidel, Dordrecht, 1979, pp. 273–280.
[37] J. Mira, R. Moreno-Diaz, A.E. Delgado, A theoretical proposal to embody cooperative decision in the nervous system, International Conference on World Problems and Systems Learning, Vol. II, 1983, pp. 687–690.
[38] J. Mira, et al., Cooperative organization connectivity patterns and receptive fields in the visual pathway: Application to adaptive thresholding, in: J. Mira, F. Sandoval (Eds.), From Natural to Artificial Neural Computation, Lecture Notes in Computer Science, Vol. 930, Springer, Berlin, 1995, pp. 15–23.
[39] A. Mitiche, P. Bouthemy, Computation and analysis of image motion: A synopsis of current problems and methods Int. J. Comput. Vision 19 (1) (1996) 29–55.
[40] R. Moreno-Diaz, F. Rubio, J. Mira, Aplicación de las transformaciones integrales al proceso de datos en la retina, Revista de Automática 5 (1969) 7–17.
[41] R.B. Paranjape, R.M. Rangayyuan, W.M. Morrow, H.N. Nguyen, Adaptive neighborhood image processing, Proc. SPIE Visual Commun. Image Process. 1818 (1992) 198–207.
[42] J.M. Prager, M.A. Arbib, Computing the optic flow: The MATCH algorithm and prediction Comput. Vision Graph. Image Proces. 24 (1983) 271–304.
[43] F. Ratliff, Mach Bands: Quantitative Studies on Neural Networks in the Retina, Holden-Day, Oakland, CA, 1965.
[44] M.E. Sereno, Neural Computation of Pattern Motion, MIT Press, Cambrige, MA, 1993.
[45] E.P. Simoncelli, Distributed Representation and Analysis of Visual Motion, Ph.D. Dissertation, MIT Press, Cambridge, MA, 1993.
[46] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, International Thomson Computer Press, 1993.
[47] F.R. Tamer, R.M. Rangayyuan, R.B. Paranjape, Adaptive neighborhood image deblurring, J. Electron. Imaging 3 (4) (1994) 368–378.
[48] D. Terman, D.L Wang, Global competition and local cooperation in a network of neural oscillators, Physica D 81 (1995) 148–176.
[49] W.B. Thompson, S.T. Barnard, Lower-level estimation and interpretation of visual motion, IEEE Computer 14 (1981) 20–28.
[50] A. Watson, A. Ahumada, Model of human visual-motion sensing, J. Opt. Soc. Amer. A 11 (2) (1985) 322–341.
[51] A.L. Yuille, N. Grzywacz, A computational theory for the perception of coherent visual motion, Nature 333 (1988) 71–74.

Antonio Fernandez 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, computer vision, neural networks, and agent technology.

Jose Mira 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-modeling discipline similar to electronics engineering, neural modeling 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.

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 modeling of biological structures, algorithmic lateral inhibition, cooperative processes of biological inspiration and fault-tolerant computation.

Miguel A. Fernandez received his degree in Computer Science from the University of Granada, Spain, in 1987, and received his Ph.D. from the Department of Artificial Intelligence of the National University for Distance Education, Spain, in 1995. Since 1991, he is an Assistant Professor with the Department of Computer Science at the University of Castilla-La Mancha, Spain. His research interests are in image processing, computer vision, and neural networks.