Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 33 (2007) 955–967

Expert Systems with Applications

www.elsevier.com/locate/eswa

Modelling the Stereovision-Correspondence-Analysis task by Lateral Inhibition in Accumulative Computation problem-solving method

Antonio Fernández-Caballero [a,*], María T. López [a], José Mira [b], Ana E. Delgado [b], José M. López-Valles [c], Miguel A. Fernández [a]

[a] Departamento de Sistemas Informáticos, Escuela Politécnica Superior de Albacete, and Instituto de Investigación en Informática de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

[b] Departamento de Inteligencia Artificial, E.T.S. Ingeniería Informática, Universidad Nacional de Educación a Distancia, 28040 Madrid, Spain

[c] Departamento de Telecomunicación, Escuela Universitaria Politécnica de Cuenca, Universidad de Castilla-La Mancha, 13071 Cuenca, Spain

Abstract

Recently, the Algorithmic Lateral Inhibition (ALI) method and the Accumulative Computation (AC) method have proven to be efficient in modelling at the knowledge level for general-motion-detection tasks in video sequences. More precisely, the task of persistent motion detection has been widely expressed by means of the AC method, whereas the ALI method has been used with the objective of moving objects detection, labelling and further tracking. This paper exploits the current knowledge of our research team on the mentioned problem-solving methods to model the Stereovision-Correspondence-Analysis (SCA) task. For this purpose, ALI and AC methods are combined into the Lateral Inhibition in Accumulative Computation (LIAC) method. The four basic subtasks, namely "LIAC 2D Charge-Memory Calculation", "LIAC 2D Charge-Disparity Analysis" and "LIAC 3D Charge-Memory Calculation" in our proposal of SCA are described in detail by inferential CommonKADS schemes. It is shown that the LIAC method may perfectly be used to solve a complex task based on motion information inherent to binocular video sequences.

Keywords: Algorithmic Lateral Inhibition; Accumulative Computation; Lateral Inhibition in Accumulative Computation; Stereovision; Correspondence analysis

1. Modelling the Stereovision-Correspondence-Analysis task

1.1. The Lateral Inhibition in Accumulative Computation method

Recently, the Algorithmic Lateral Inhibition (ALI) method, as well as the Accumulative Computation (AC) method, has proven to be greatly efficient in modelling at the knowledge level for general-motion-detection tasks in video sequences. More precisely, the task of persistent-motion detection has been widely expressed by means of the AC method (Mira, Fernández, López, Delgado, & Fernández-Caballero, 2003), whereas the ALI method has been used with the objective of moving-objects detection (Mira, Delgado, Fernández-Caballero, & Fernández, 2004), labelling and further tracking (López, Fernández-Caballero, Mira, Delgado, & Fernández, 2006). This paper exploits our research team current knowledge on the problem-solving methods mentioned to model the Stereovision-Correspondence-Analysis (SCA) task (López-Valles, Fernández, Fernández-Caballero, & Gómez, 2005). For this purpose, ALI and AC methods are combined into the Lateral Inhibition in Accumulative Computation (LIAC) method (Fernández-Caballero, Fernández, Mira, & Delgado, 2003) as a powerful problem-solving method

* Corresponding author. Tel.: +34 967 599200; fax: +34 967 599224.
E-mail address: caballer@info-ab.uclm.es (A. Fernández-Caballero).

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.07.009

(PSM) in generic computer vision motion-motivated tasks.

A complete description of the ALI method is available in Mira et al. (2004). In the non-recurrent ALI case, each calculation element samples its data in the central (C) and periphery (P) part of the volume that its RF (receptive field) specified in the input space V. On these two data fields, the calculation element carries out evaluation inferences and results comparison. This comparison inference is made according to a set of criteria to generate a set of discrepancy classes as input to the final selection, where the output is obtained from the set of outputs associated with the different discrepancy classes, according to the specific discrepancy classes generated by the previous comparison inference. In an analogous manner there is the inferential scheme for the recurrent ALI circuits. Now each element of calculus starts to infer from data sampled in the central (C^*) and periphery (P^*) parts of its feedback receptive fields in the output space. The values in C^* (individual opinion before dialogue) are compared with the evaluation of the "opinions" of all the elements in the periphery. This comparison is made according to a set of rules for consensus to produce a discrepancy class. Finally, as in the non-recurrent case, this discrepancy is the input to a selection to provide the consensus output.

The AC method is based on the permanency effect. Accumulative computation has now been largely applied to moving objects detection, classification and tracking in indefinite sequences of images (e.g. (Mira et al., 2003)). The more general modality of AC is the charge/discharge mode, which may be described by means of the following generic formula:

$$Ch(x,y,t) = \begin{cases} \min(Ch(x,y,t-1) + C, Ch_{\max}) & \text{if "}\textit{property } P(x,y,t)\text{"} \\ \max(Ch(x,y,t-1) - D, Ch_{\min}) & \text{otherwise} \end{cases} \tag{1}$$

This way, the temporal accumulation of the persistency of the binary property $P(x,y,t)$ measured at each time instant t at each pixel (x,y) of the data field is calculated. Generally, if the property is fulfilled at pixel (x,y), the charge value at that pixel $Ch(x,y,t)$ goes incrementing by increment charge value C up to reaching $Ch_{\max}$, whilst, if property P is not fulfilled, the charge value $Ch(x,y,t)$ goes decrementing by decrement charge value D down to $Ch_{\min}$. All pixels of the data field have charge values between the minimum charge, $Ch_{\min}$, and the maximum charge, $Ch_{\max}$. Obviously, values C, D, $Ch_{\min}$ and $Ch_{\max}$ are configurable depending on the different kinds of applications, giving rise to all different operating modes of the accumulative computation. Values of parameters C, D, $Ch_{\max}$ and $Ch_{\min}$ have to be fixed according to the applications characteristics. The particular values $Ch_{\max}$ and $Ch_{\min}$ have to be chosen by taking into account that charge values will always be between them. The value of C defines the charge increment interval between time instants $t-1$ and t. Greater values of C allow arriving in a quicker way to saturation. On the other hand, D defines the charge decrement interval between time instants $t-1$ and t. Thus, notice that the charge stores motion information as a quantified value, which may be used for several classification purposes. In (Mira et al., 2003) the architecture of the accumulative computation module is shown. Some of the operating modes may be noticed there, demonstrating their versatility and their computational power.

Lastly, the LIAC method, understood as the combination or fusion of the ALI and AC methods, consists precisely in using the accumulative computation (or permanency computation) mechanisms as the algorithmic part of the LI (lateral inhibition).

1.2. The Stereovision-Correspondence-Analysis task

In a conventional stereoscopic approach, usually two cameras are assembled with a horizontal distance between them. As a consequence, objects displaced in depth from the fixation point are projected onto image regions which are shifted with respect to the image centre. Brown, Burschka, and Hager (2003) describe in their work a great variety of algorithms that have been developed to analyze the depth in a scene in a survey article. In many previous works, a series of restrictions are used to approach the correspondence problem. The most usual restriction is the disparity one, which considers that is not likely that there are objects very close to the camera. The scene is usually limited to a medium distance. This way, too high disparities are eliminated (Sumi, Kawai, Yoshimi, & Tomita, 2002). Koenderink and van Doorn (1976) expressed the necessary theory in the best initial works related to disparity restriction, and Wildes (1991) implemented some of their ideas (Wilson & Knutsson, 1989). More recently, disparity in stereoscopy continues showing a great interest (e.g. Muhlmann, Maier, Hesser, & Manner, 2002; Gutiérrez & Marroquín, 2004).

According to the correspondence techniques used, we may classify methods into correlation-based, relaxation-based, gradient-based, and feature-based. The main correlation-based technique is the area-correlation technique (e.g. (Zabih & Woodfill, 1994)). Area-based approaches have the advantage of generating dense disparity maps directly. Matching elements for area-based methods are the individual pixels over which the matching cost is evaluated; pixel-to-pixel correspondence is assessed on image intensity function and similarity statistics. For instance, the work by Binaghi, Gallo, Marino, and Raspanti (2004) investigates the potential of neural adaptive learning to solve the correspondence problem within a two-frame adaptive area matching approach. The method is based on the use of the zero mean normalized cross-correlation coefficient integrated within a neural network model which uses a least-mean-square delta rule for training. Another approach (Di Stefano, Marchionnia, & Mattoccia, 2004) proposes an area-based stereo algorithm suitable to real time applications, where the core of the algorithm relies on the uniqueness constraint and on a matching process that rejects previous matches as soon as more reliable ones are found. In (Goulermas & Liatsis, 2001) an algorithm for performing robust feature-based stereo-matching, without the ordering constraint, is proposed. The calculation of the disparity map is decomposed to a set of disjoint intra-row subproblems, each one having two objectives: the search for a high confidence intra-row matching and the enforcement of figural continuity at the inter-row level. The basic idea of relaxation techniques is that pixels to be set into correspondence perform "controlled estimations". In this kind of process, the correlation values of the neighbours of a pixel are of great importance for the evaluation of the correspondence (Grimson, 1985). A stereo-matching scheme using a genetic algorithm to improve the depth-reconstruction method of stereo-vision systems has also been presented (Han, Song, Chung, Cho, & Ha, 2001). The proposed approach considers the matching environment as an optimization problem and finds the optimal solution by using an evolutionary strategy. Methods based in the gradient or in the optical flow aim to determine local disparities between two images by formulating a differential equation that relates motion and luminance (Choi, Yoon, Lee, & Chien, 2003). Techniques based in features limit to reliable features, such as contours or curves (e.g. (Venkateswar & Chellappa, 1995)), at the analyzed regions. Feature-based approaches rely on the matching of explicit features extracted from the images, such as edges, which correspond to physical scene properties; these presume a degree of geometric invariance. Recently, an efficient and fast detection algorithm has been introduced for the maximally stable extremal regions (MSER) (Matas, Chuma, Urbana, & Pajdla, 2004). Some feature-based approaches use neural networks technology (e.g. Pajares & de la Cruz, 2001, & Pajares & de la Cruz, 2003).

The *Stereovision-Correspondence-Analysis Model* is presented in this article. This model is configured as a tool for obtaining continuous three-dimensional information about motion in a scene by means of the permanence effect. To make the comprehension of the model easier, we previously carry out a general description of it and later describe each subtask in detail.

For the rest of the *Stereovision-Correspondence Analysis* explanation, a sequence downloaded from the "http://www.labvisione.deis.unibo.it/smattoccia/stereo.htm" website is going to be used as a running example. It is the "IndoorZoom" sequence. Fig. 1 shows one couple of stereo frames from this sequence. These sequences were not filmed by us. Therefore, we have no control over the camera system's geometric parameter, nor over its features, such as focal distance or pixel size. Thus, the results will be affected by a scaling factor. Throughout this and subsequent sections, the figures with the results obtained at every stage of the "IndoorZoom" sequence will be shown.

The general description of the SCA model as well as its three basic subtasks, namely *LIAC 2D Charge-Memory Calculation*, *LIAC 2D Charge-Disparity Analysis*, and *LIAC 3D Charge-Memory Calculation* in our proposal of SCA are described in detail by inferential CommonKADS schemes (Schreiber et al., 2001; Breuker & van de Velde, 1994). This way, it is shown that the LIAC method may perfectly be used to solve a complex task based on motion information inherent to video sequences.

2. General description of the Stereovision-Correspondence-Analysis Model

The model which supports the *Stereovision-Correspondence Analysis* is the result of the analysis of the stereovision's geometric problem and the application of the pertinent restrictions, as well as the study of biological stereovision systems and the permanence and communication mechanisms on the local level, which our research group is very familiar with. Next, the proposed model is described in a general way, dedicating the rest of the paper to the detailed description of each subtask in which the solution to the problem is broken down in terms of elementary subtasks and inferences. The decomposition into subtasks is shown in Fig. 2.

The input is a three-dimensional scene where different moving elements will appear through time. Firstly, frame by frame, the permanence effect is applied for the purpose of *LIAC 2D Charge-Memories Calculation* to the stereo image pairs. The input at all times is the stereo image pair in grey levels, corresponding to a frame and the output to each frame is the state of a charge memory (calculated by means of accumulative computation), where the information associated with motion is stored. With the purpose of analyzing the periphery of the fixation point, we

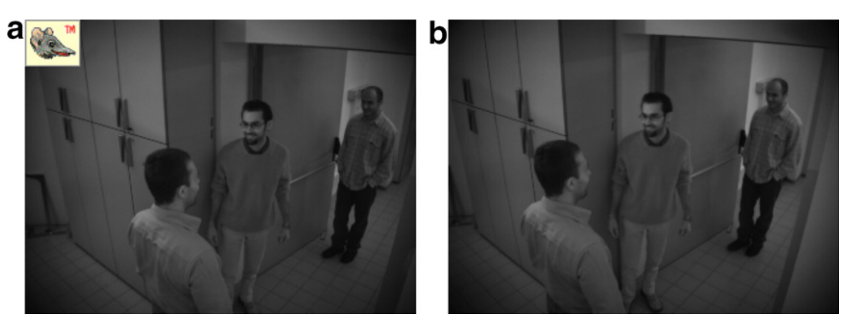

Fig. 1. One couple of frames from the running example. (a) Left input image and (b) right input image.

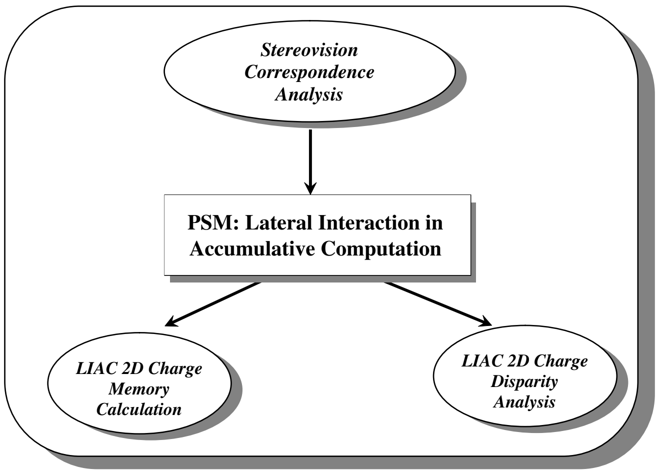

Fig. 2. "Stereovision-Correspondence-Analysis" task decomposition.

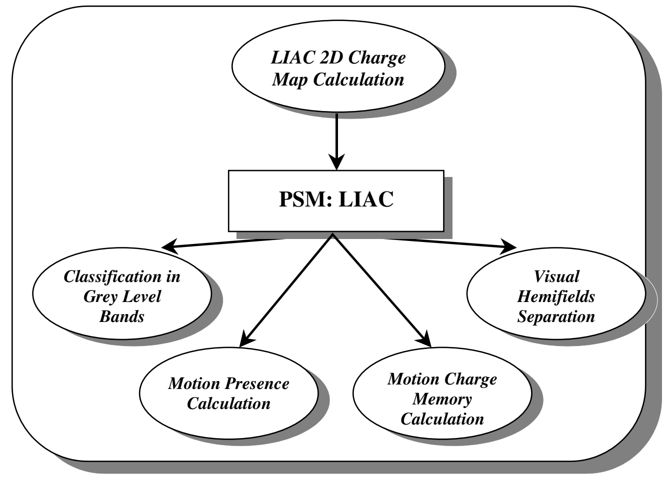

Fig. 3. "LIAC 2D Charge-Memory Calculation" subtask decomposition.

separated the right and left charge memories again in two areas, corresponding to the two associated visual hemifields. This option is justified by the intention of analyzing the stimuli which come from each side of the visual field separately, as found in the studied biological systems.

Then, the matching process between both right and left sequence charge memories for each frame is performed by means of the *LIAC 2D Charge Disparity Analysis* subtask. The input is the *2D Charge Memory* of each and every visual hemifield, and the output is a *Depth Memory* (obtained again through accumulative computation and lateral inhibition), where points (x, y, z) are activated in the place where motion in this frame is detected. This memory will also appear divided into its two corresponding hemifields. Every known concept in stereovision will be applied in this section: restrictions to the correspondences, complete primitives, etc., but applied to the charge memories obtained in the previous section instead of object-shape information, as usual. With this structure, we have tried to analyze motion in a scene, eliminating all static information and estimating the depth after performing a stereovision correspondence analysis, in all the three dimensions in space, for the objects in motion which appear in the scene.

3. LIAC 2D Charge-Memory Calculation

Its purpose is to represent two-dimensional motion for every input sequence in the permanence elements charge levels. Once these are obtained, the separation of the images into hemifields will be carried out to analyze separately the stimuli that come from each side of the visual field. The input is the stereo-sequence images in grey levels coming from camera signal digitalization. On the other hand, the output is the right and left permanence memories charge state, each one divided in two halves for the reasons stated in the previous paragraph.

Fig. 3 represents the division into subtasks of the *LIAC 2D Charge-Memory Calculation* task.

The *Classification in Grey-Level Bands* subsystem separates each of the frames into related regions of common grey-level band for the purpose of later analyzing its movements. The *Motion-Presence Calculation* subtask requires as its input the current image segmented into grey-level bands, as well as the previous image. The aim of this is to analyze which memory elements have skipped between the bands, detecting movement in the corresponding pixels. *Motion-Charge-Memory Calculation* obtains by means of accumulative computation on the negative of property *Motion Presence* a measure in the persistency of the motion present in the scene. All these subsystems have already been explained as such in López et al. (2006). Thus, we will only show the final results of these subtasks on the running example.

Fig. 4 represents the results of *Grey-Level-Bands Segmentation*. In order to do this, we used frame number 102 of the stereo sequence, segmenting it into eight grey-level bands. Here we only show the classification of the left image.

Fig. 5 shows the result of motion detection in three consecutive moments in time (frame numbers 100–102) for the "IndoorZoom" sequence. Out of all the segmented regions, the system only takes note of those pixels where there has been a jump in grey-level bands and movement has been detected.

In Fig. 6, it is possible to see the frame's evolution in the "IndoorZoom" sequence when applying the permanence effect on motion detection. The last movements that have taken place in the sequence are stored in the permanence memory. We assume that this process is continuous in this representation and so the 2D charge memories contain information from previous moments in time.

Finally, each of the visual-hemifields-separation subsystems uses the motion information coming from its permanence as its input and the two separate visual hemifields as its output. As output from this section, the two-dimensional charge memories are divided in two halves. The purpose of this is to separate the motion that takes place on

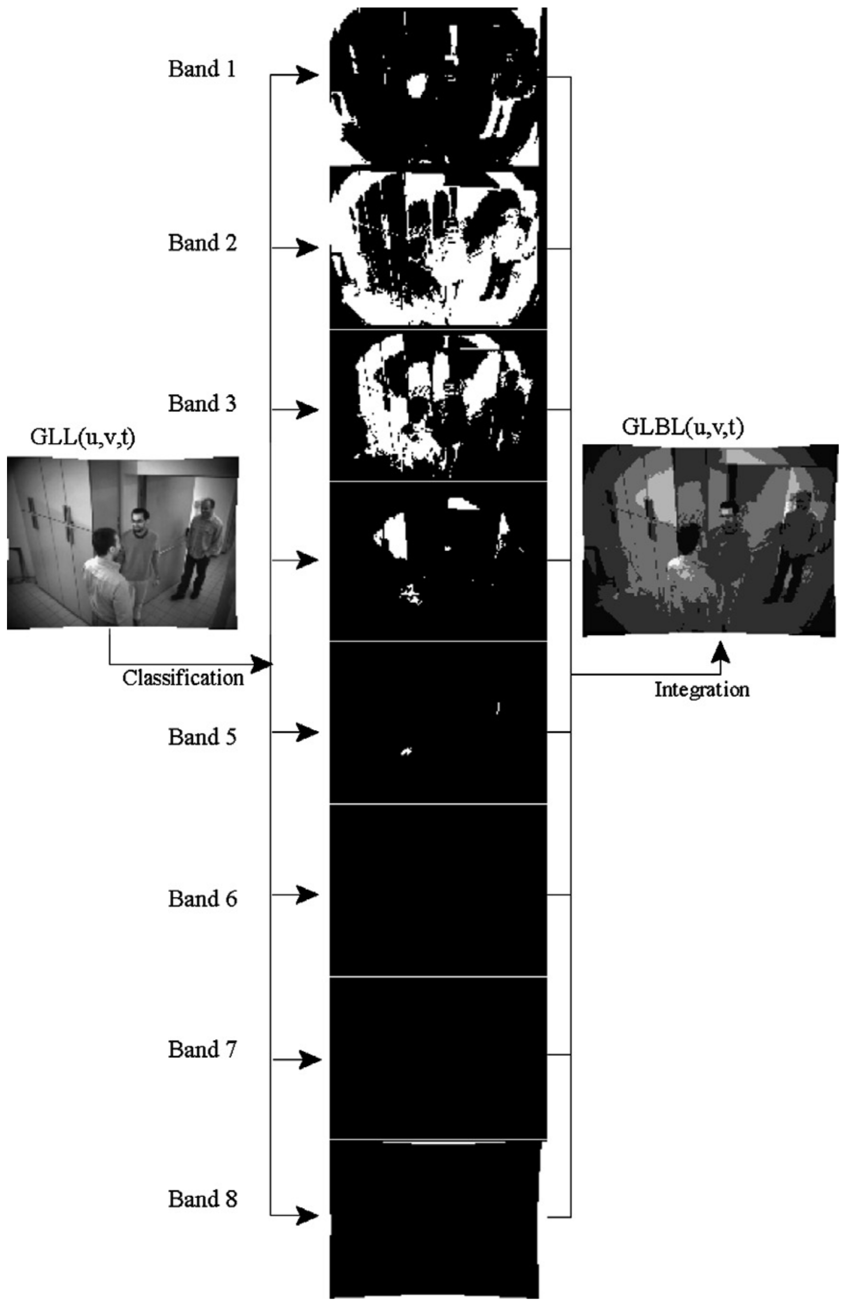

Fig. 4. ‘‘Classification in Grey-Level Bands’’ applied to the running example.

the right side of the visual field from the one which takes place on the left side. The static and dynamic roles associated with this subtask are represented in Fig. 7.

The *2D Left Visual Hemifield* Charge Memory will be called $H2_{L/R,L}(u,v,t)$ and the *2D Right Visual Hemifield* will be called $H2_{L/R,R}(u,v,t)$. The division of both hemifields is done horizontally down the middle of the charge memories. This is the reason for using this size ($H_{\max}$) as a static role. The first *select* inference extracts the left side of this complete memory, leaving the right side as an empty set. The right*select* inference has the same role, but this time it is the right visual hemifield which is extracted from the complete *2D Motion-Charge Memory*.

Thus, the visual hemifields' contents will be:

$$H2_{L/R,L}(u + H_{\max}/2, v, t) = C2_{L/R}(u - H_{\max}/2, v, t) \tag{2}$$

$$H2_{L/R,R}(u, v, t) = C2_{L/R}(u, v, t) \tag{3}$$

considering

$$0 < u \leqslant H_{\max}/2 \tag{4}$$

$$-V_{\max}/2 < v < V_{\max}/2 \tag{5}$$

$$0 < t < \infty \tag{6}$$

Fig. 8 represents a two-dimensional charge memory's *Visual Hemifields Separation*.

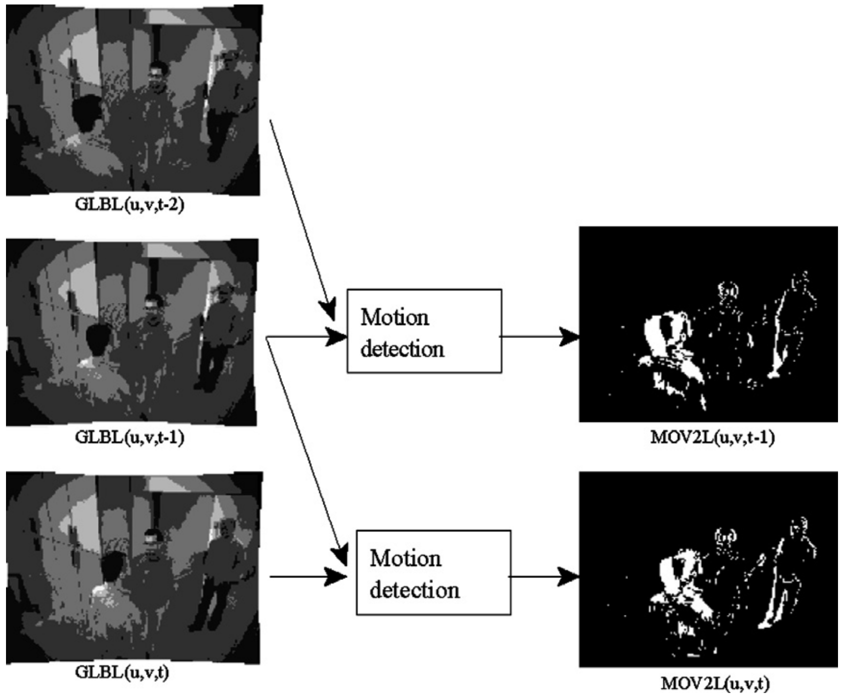

Fig. 5. "Motion-Presence Calculation" applied to the running example.

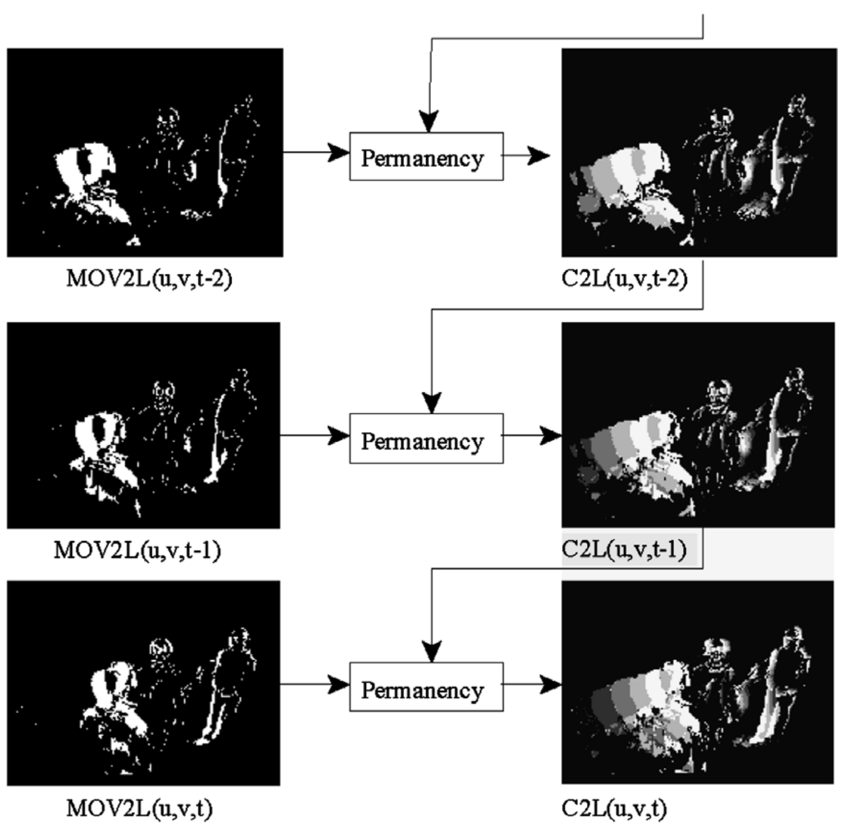

Fig. 6. "Motion-Charge-Memory Calculation" applied to the running example.

4. LIAC 2D Charge-Disparity Analysis

Now, the two left halves of the charge memories are arranged separately in relation to the two right halves, with the purpose of separating the stimuli that come from each visual hemifield. This way, two parallel data processing systems will be needed; each one of them will process its corresponding half of the visual field. The

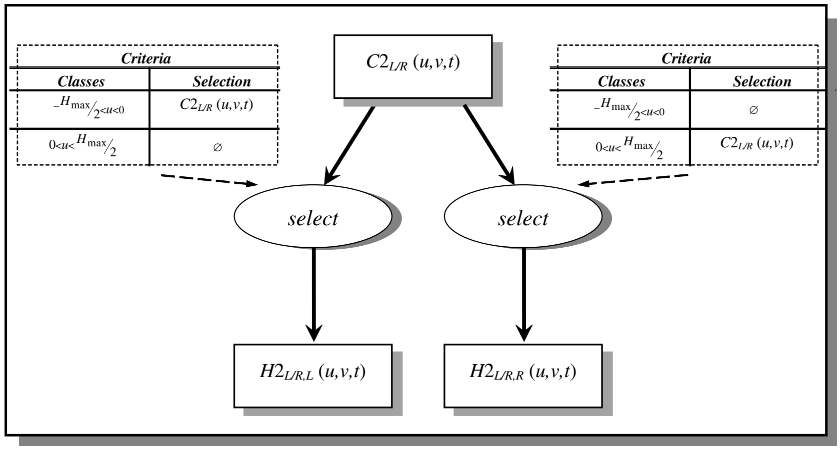

Fig. 7. ‘‘Visual-Hemifields-Separation’’ inferential scheme.

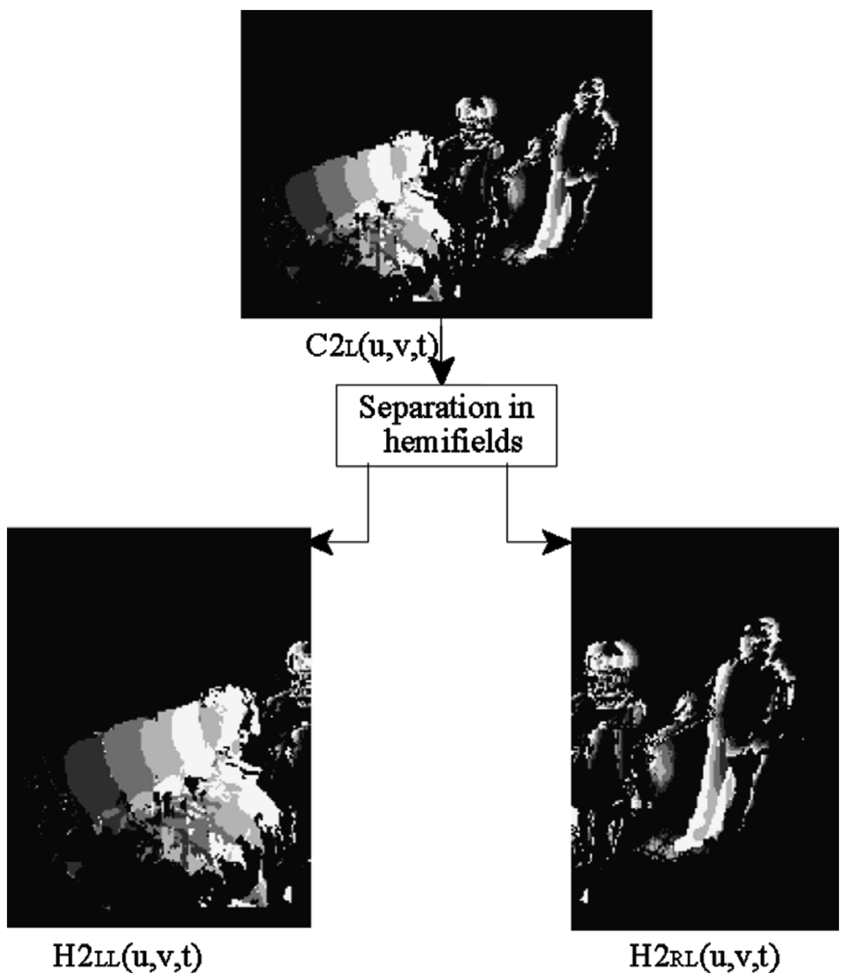

Fig. 8. ‘‘Visual-Hemifields Separation’’ applied to the running example.

output is a three-dimensional depth memory which shows the depth of the points in the scene where there has been movement.

Using the charge memories (*2D Left Visual Hemifield* and *2D Right Visual Hemifield*) as input has two important features:

- Only information about motion is used, filtering out all static information from the scene, whether it is 2D or 3D. Since our objective is to obtain a three-dimensional Memory of the scene's motion, it is more an advantage than a disadvantage to have a filtered Memory, as static elements produce noise to this project.

- On the other hand, object motion leaves trails in the permanence-memory-charge memories, which will have different characteristics, depending on its direction, instantaneous velocity and its motion history. However, the instantaneous motion of a single object will be represented in both charge memories as similar trails. Thus, trail matching of the moving objects in the sequence will be simple and robust.

The division of this subtask into elementary subtasks is represented in Fig. 9. We can see both elementary subtasks into which the subtask is divided. They are: on the one hand, *2D Charge Correspondence Analysis* and on the other hand, *3D Depth Memory Calculation*.

Both 2D visual-hemifield-charge memories are reorganized to process the stimuli that come from the *2D Left Visual Hemifield* separately from those in the *2D Right Visual Hemifield*. Therefore, the analysis process for the left-charge correspondences uses the left-visual hemifield from the left-charge memory and the left-visual hemifield from the right-charge memory as input. On the other hand, the analysis process for the right correspondences uses both right-visual hemifields from both two-dimensional charge memories as input.

The correspondence analysis output will be called *3D Charge Correspondence Memory* ($S3(u, v, d, t)$), where d stands for disparity. Here, the existing correspondences between the pixels in the right and left image hemifields for different disparities will be represented. From these charge correspondence memories, the maximum reliability depth for each of the visual hemifields coordinates (u, v) will be decided. This will be done by means of each *3D Depth Memory Calculation* process. The output will be called *3D Depth Memory*.

4.1. Charge-Correspondence Analysis

The purpose of this subtask is to prepare the necessary information to decide on the disparity with the greatest reliability for each of the processing elements in the input

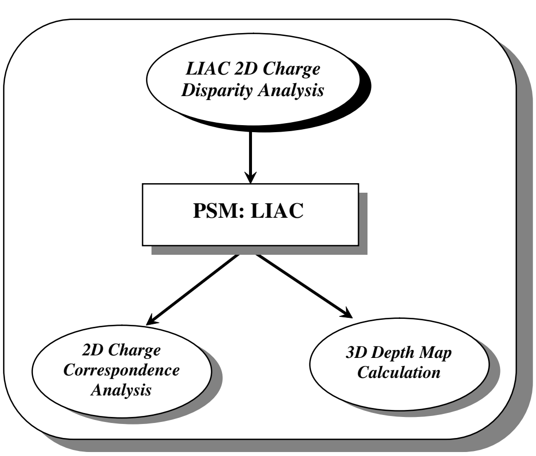

Fig. 9. "2D Charge-Disparity Analysis" subtask decomposition.

charge memories for *3D Depth Memory Calculation*. This task will mainly take into account the epipolar, arrangement and disparity restrictions. The application procedure will be seen in each of the steps taken.

With this, we can carry on with the analysis of the static and dynamic roles, which coincide in this subtask. Here, we see the equivalent hemifields for the right and left images, at the input, as dynamic roles and the *3D Charge Correspondence Memory* at the output. On the other hand, the size of the image, vertically and horizontally, and the maximum disparity value given by the restriction with the same name, appear as static roles.

It is common knowledge that the most robust correspondence primitives are those with the highest level, such as contours or regions. In our case, we intend to carry out a correspondence analysis per region. Therefore, we must group together those neighbouring charge elements whose corresponding elements have the same disparity. But before grouping the neighbouring elements together, we must define what we call corresponding charge elements. It is, basically, a question of finding which pixel has a similar history of movement in the opposite epipolar line and, consequently which processing element for the corresponding charge Memory has an instant charge level stored similarly.

Once the regions called "*constant disparity*" are established, it is convenient to establish a characteristic, which will allow us to set up a reliability criterion to decide on the correct disparity for each charge element (u, v). The characteristic chosen will depend on the position of each processing element and it has to do with the size of each "*constant disparity*" region. This size is calculated in two phases:

- First of all, we carry out a horizontal counting of all adjoining neighbours which belong to this region.
- Afterwards, the horizontal values found for all the adjoining vertical processing elements, which also belong to this region, are accumulated.

In the following subsections, these phases will be explained in the order in which they occur, that is to say, *Pixel-wise Charge Correspondence Analysis*, *Horizontal Charge Counting and Homogenizing*, and *Vertical Charge Accumulation and Homogenizing*, using the inferential diagrams.

4.1.1. Pixel-wise Correspondence Analysis

The *Pixel-wise Charge Correspondence Analysis* on the charge elements of two corresponding hemifields is carried out. By applying the epipolar restriction, each charge element from a hemifield is compared to those from another hemifield on the same row, although displaced horizontally, up to the maximum limit set by the disparity restriction.

The inferential diagram can be seen in Fig. 10. Here we notice that there are two *evaluate* inferences for each visual

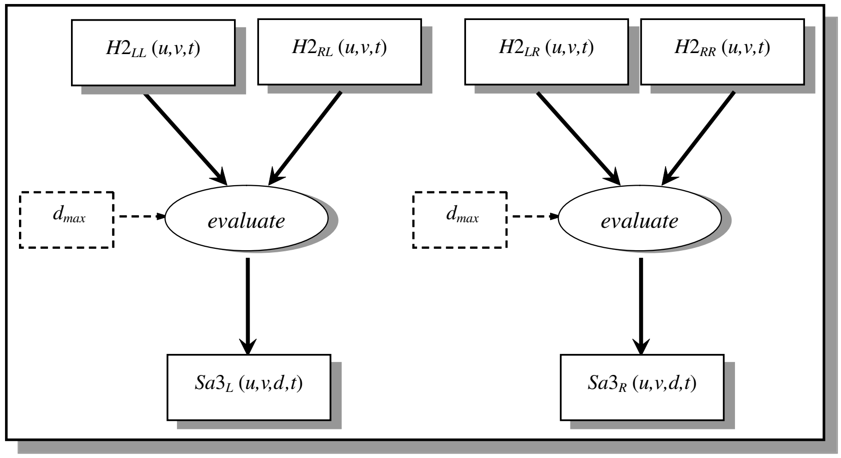

Fig. 10. "Pixel-wise Charge Correspondence Analysis" inferential scheme.

hemifield. Each one of these uses the corresponding halves $H2_{L/R,L}(u,v,t)$ and $H2_{L/R,R}(u,v,t)$ of the two-dimensional visual-hemifield-charge memories as input dynamic roles and the maximum disparity $d_{\max}$ value set up by the restriction, as static role. The output dynamic role for each inference is a three-dimensional matrix ($Sa3_{L/R}(u,v,d,t)$), which indicates whether or not there is a specific correspondence for each coordinate (u,v) and for each disparity value d.

The calculation expressions for the elements of this three-dimensional output matrix in each visual hemifield are as follows:

$$Sa3_L(u,v,d,t) = \begin{cases} 1 & \text{if } |H2_{LL}(u,v,t) - H2_{RL}(u+d,v,t)| \leqslant 1 \\ 0 & \text{otherwise} \end{cases}, \quad \forall d | 0 \leqslant d \leqslant d_{\max} \tag{7}$$

$$Sa3_R(u,v,d,t) = \begin{cases} 1 & \text{if } |H2_{LR}(u-d,v,t) - H2_{RR}(u,v,t)| \leqslant 1 \\ 0 & \text{otherwise} \end{cases}, \quad \forall d | 0 \leqslant d \leqslant d_{\max} \tag{8}$$

In Fig. 11, we can see a graph of the correspondence analysis. In the central matrix, white pixels indicate the detection of potential correspondences. The size of the areas is an indication of the correspondence confidence. Since the output role $Sa3_{L/R}(u,v,d,t)$ is a variable which depends on three spatial dimensions, as well as time, the variation in disparity has been represented in the figure in a mosaic-like way. For this reason, the central element of the figure, where the correspondences turn up, is made-up as a five-column, four-row matrix, where each column is also divided into two in order to house each charge Memory's right and left hemifield. In this figure, we also see that the character which appears in the left hemifield is the closest one, since its constant disparity region has a maximum size for $d = 18$. However, the right hemifield's character is further away and has a lower disparity ($d = 6$). The central character has a disparity of ($d = 15$) and it is quite static.

4.1.2. *Horizontal-Charge Counting and Homogenizing*

The purpose in this second step is to establish a charge elements' matrix the same size as the three-dimensional matrices $Sa3_{L/R}$ from the previous step, where each element would have the amount of horizontally corresponding adjacent input elements as the final result (those where $Sa3_{L/R}(u,v,d,t) = 1$). This process is carried out in two periods: the first (*Horizontal-Charge Counting*) running to the left counts the input elements set on 1 and stores the value in the corresponding output charge element ($Sb3_{L/R}(u,v,d,t)$). The inferential diagram of this first step can be seen in Fig. 12.

The process can be interpreted as three nested loops in the following way:

```
for u=1 to Hmax/2
  for v=1 to Vmax{
```

$$s1 = Sb3_{L/R}(u-1,v,d,t) + 1$$

$$s2 = 0$$

$$Sb3_{L/R}(u,v,d,t) = \begin{cases} s1 & \text{if } Sa3_{L/R}(u,v,d,t) = 1 \\ s2 & \text{if } Sa3_{L/R}(u,v,d,t) = 0 \end{cases} \tag{9}$$

```
}
```

Once the horizontal counting to the right is done, a charge homogenizing is carried out (*Horizontal-Charge Homogenizing*) so that all charge elements belonging to a horizontal-constant-disparity region acquire the same charge value. This acquired value will be maximal and it will correspond with the horizontal size of the constant disparity region, formed by all of them. This process' inferential diagram can be seen in Fig. 13 and its explanation is very similar to the previous one.

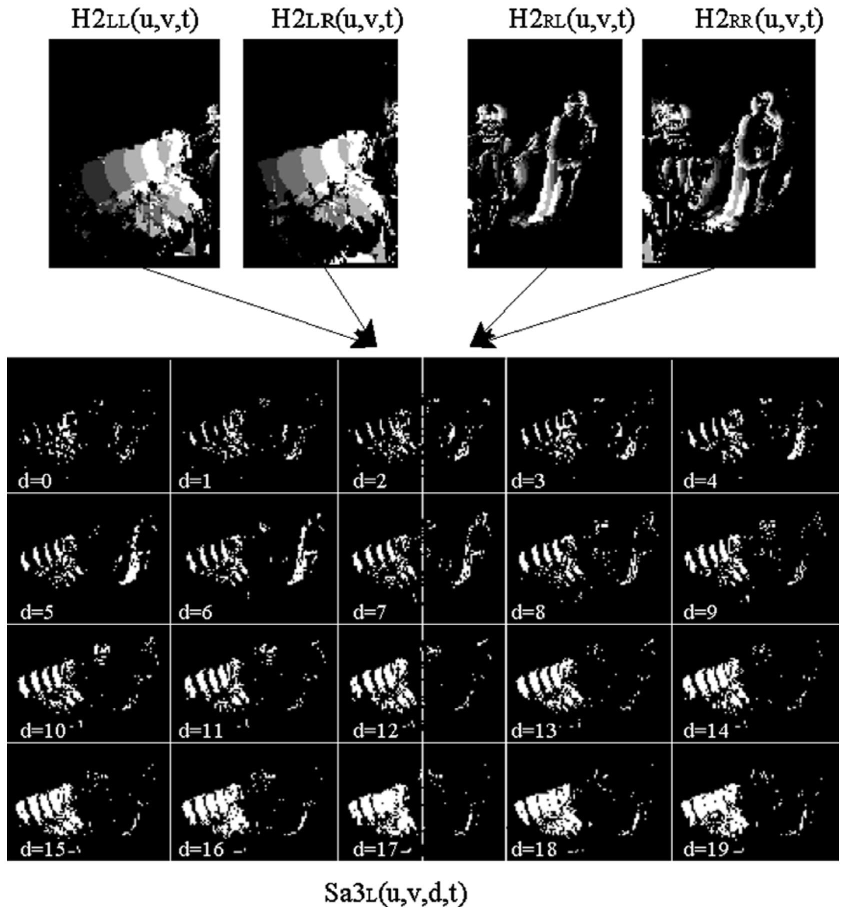

Fig. 11. "Pixel-wise Charge Correspondence Analysis" applied to the running example.

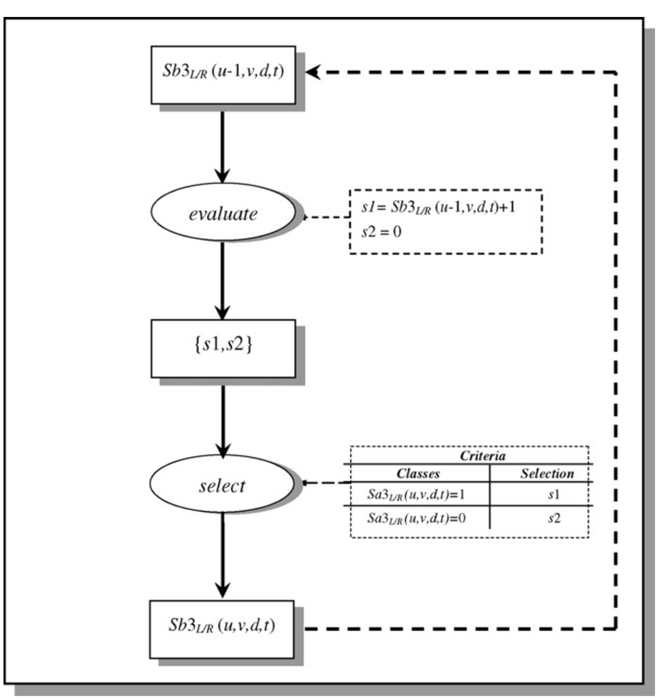

Fig. 12. "Horizontal-Charge Counting" inferential scheme: Recurrent spatial ALI.

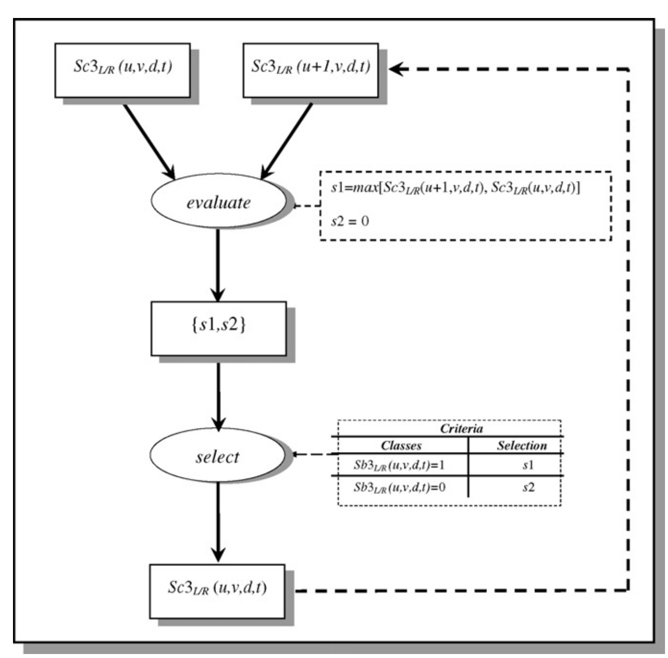

Fig. 13. "Horizontal-Charge Homogenizing" inferential scheme: Recurrent spatial ALI.

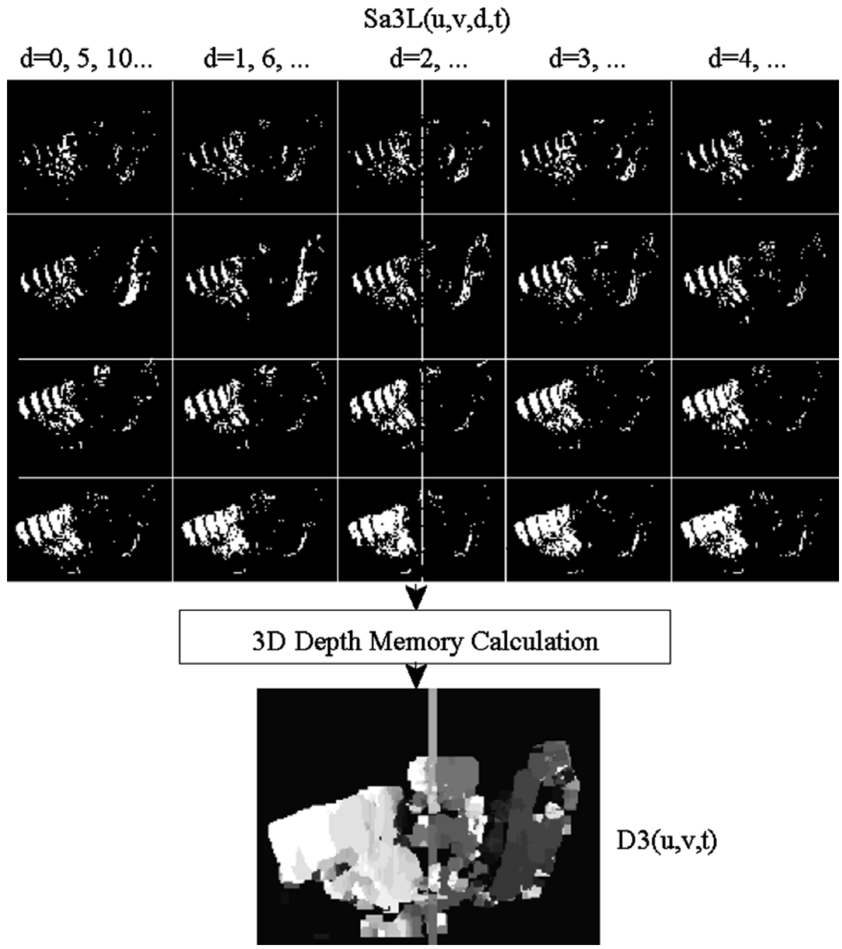

Fig. 14. "3D Depth-Memory Calculation" applied to the running example.

This diagram's processing can also be written as nested loops in this way:

for u=$(H_{max}/2-1)$ downto 1
  for v=1 to V_{max}
    for d=1 to d_{max}{

$$s1 = \max[Sc3_{L/R}(u+1,v,d,t), Sc3_{L/R}(u,v,d,t)]$$

$$s2 = 0$$

$$Sc3_{L/R}(u,v,d,t) = \begin{cases} s1 & \text{if } Sb3_{L/R}(u,v,d,t) = 1 \\ s2 & \text{if } Sb3_{L/R}(u,v,d,t) = 0 \end{cases} \tag{10}$$

}

4.1.3. Vertical-Charge Accumulation and Homogenizing

On this third step, we aim to establish a new charge element matrix of the same size as the three-dimensional matrices $Sc3_{L/R}$ from the previous step, where each element has, as final result, the charge accumulation from the neighbouring input elements, which are considered vertically corresponding. This process is done, as in the previous case, in two periods: a first running downwards (rising v values) counts the input elements other than 0 and stores the accumulated value in the corresponding output charge element ($Sd3_{L/R}(u,v,d,t)$). Thus, the process can be interpreted as three nested loops in the following way:

for u=1 to $(H_{max}/2)$
  for v=2 to V_{max}
    for d=1 to d_{max} {

$$s1 = Sd3_{L/R}(u,v,d,t) + Sd3_{L/R}(u,v-1,d,t)$$

$$s2 = 0$$

$$Sd3_{L/R}(u,v,d,t) = \begin{cases} s1 & \text{if } Sc3_{L/R}(u,v,d,t) = 1 \\ s2 & \text{if } Sc3_{L/R}(u,v,d,t) = 0 \end{cases} \tag{11}$$

}

Once the counting towards positive v values is done, charge homogenizing is carried out in such a way that all charge elements belonging to a constant disparity region vertically have the same charge value. This value will be maximal and it will correspond to that region's total size. Now, the processing can also be written as nested loops in this way:

```
for u=1 to (H_max/2)
  for v=(V_max-1) to 1
    for d=1 to d_max{
```

$$s1 = \max[Se3_{L/R}(u,v,d,t), Se3_{L/R}(u,v+1,d,t)]$$
$$s2 = 0$$
$$Se3_{L/R}(u,v,d,t) = \begin{cases} s1 & \text{if } Sd3_{L/R}(u,v,d,t) = 1 \\ s2 & \text{if } Sd3_{L/R}(u,v,d,t) = 0 \end{cases} \quad (12)$$

```
}
```

4.2. 3D Depth-Memory Calculation

Once we have calculated the region's sizes which each charge Memory correspondence belongs to, we need to associate, as maximum reliability disparity for each pixel, those values whose charge $S3_{L/R}$ (u,v,d,t) is maximal in d. With this, we are imposing the uniqueness restriction, since each processing element will only have a single disparity value as a final value.

This subtask has the *3D Charge-Correspondence Memory* $S3_{L/R}$ (u,v,d,t) as input and the maximum disparity imposed by the disparity restriction as static role. The processing carried out to obtain the disparity associated to each charge element is also shown in the following expression:

$$D3_{L/R}(u,v,t) = i | S3_{L/R}(u,v,i,t) \geqslant S3_{L/R}(u,v,j,t), \quad \forall (i,j), \ 0 \leqslant i,j \leqslant d_{\max} \quad (13)$$

This operation tries, basically, to find the value for i whose $S3_{L/R}(u,v,i,t)$ is maximum in the third dimension. The arrangement restriction is included in the method proposed for the charge-disparity-analysis subtask, since the specific correspondence verification and subsequent region configuration means maintaining the order of the found correspondences.

Based on the charge disparity calculation done and the camera system's geometric analysis, we can estimate the moving elements' depth. With this, we have obtained a stereoscopic motion Memory in which each moving element from the scene appears associated to its depth. Fig. 14 shows a graph of the input and output roles to the *3D Depth-Memory Calculation* subtask.

5. Conclusions

In this paper a combination of the ALI and AC methods into the Lateral Inhibition in Accumulative Computation (LIAC) method is presented as a powerful problem-solving method for the task of correspondence analysis in binocular stereovision. The current work has shown the convenience of modelling knowledge of tasks and methods in terms of a library of reusable components (inferential verbs "evaluate", "compare" and "select") and a set of input and output roles played by the entities of the application domain. For each one of the subtasks the results of the inferential scheme are illustrated.

Up to now, the common stereovision techniques are based on shape, analyzing disparity, and thus obtaining depth based on the system's geometry. However, they are basically static. This article proposes a new alternative which allows the continuous obtaining of three-dimensional information about motion in a scene. The motion trails of several moving objects in each frame will be different to each other due to the different nature of their movements. However, a single moving object will create very similar trails in both permanence memories that make up a stereo pair. This makes the trail-based correspondence analysis simple and robust at the same time.

The solution proposed involves a type of process which tries to take advantage of the use of high-order primitives and pixels only. On the one hand, the elements placed in correspondences are regions obtained from moving objects' motion trails by means of permanence memory interpretation. Again, this enables to obtain simple and robust correspondences. On the other hand, the fact that each pixel can decide, through a local analysis and based on motion trail overlapping, which disparity is more reliable creates a dense disparity memory, which is considered the biggest advantage in pixel-based correspondence systems.

Acknowledgements

This work is supported in part by the Spanish CICYT TIN2004-07661-C02-01 and TIN2004-07661-C02-02 grants, and the Junta de Comunidades de Castilla-La Mancha PBI06-099 grant.

References

Binaghi, E., Gallo, I., Marino, G., & Raspanti, M. (2004). Neural adaptive stereo matching. *Pattern Recognition Letters, 25*, 1743–1758.

Breuker, J., & van de Velde, W. (1994). *CommonKADS library for expertise modelling*. IOS Press.

Brown, M. Z., Burschka, D., & Hager, G. D. (2003). Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(8), 993–1008.

Choi, I., Yoon, J.-G., Lee, Y.-B., & Chien, S. I. (2003). Stereo system for tracking moving object using log-polar transformation and zero disparity filtering. In *Proceedings of the 10th international conference on computer analysis of images and patterns* (pp. 182–189).

Di Stefano, L., Marchionnia, M., & Mattoccia, S. (2004). A fast area-based stereo matching algorithm. *Image and Vision Computing, 22*, 983–1005.

Fernández-Caballero, A., Fernández, M. A., Mira, J., & Delgado, A. E. (2003). Spatio-temporal shape building from image sequences using lateral interaction in accumulative computation. *Pattern Recognition, 36*(5), 1131–1142.

Goulermas, J. Y., & Liatsis, P. (2001). Hybrid symbiotic genetic optimisation for robust edge-based stereo correspondence. *Pattern Recognition, 34*, 2477–2496.

Grimson, W. E. L. (1985). Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 7*, 17–34.

Gutiérrez, S., & Marroquín, J. L. (2004). Robust approach for disparity estimation in stereo vision. *Image and Vision Computing, 22*(3), 183–195.

Han, K.-P., Song, K.-W., Chung, E.-Y., Cho, S.-J-., & Ha, Y.-H. (2001). Stereo matching using genetic algorithm with adaptive chromosomes. *Pattern Recognition, 34*, 1729–1740.

Koenderink, J., & van Doorn, A. (1976). Geometry of binocular vision and a model for stereovision. *Biological Cybernetics, 21*, 29–35.

López, M. T., Fernández-Caballero, A., Mira, J., Delgado, A. E., & Fernández, M. A. (2006). Algorithmic lateral inhibition method in dynamic and selective visual attention task: Application to moving objects detection and labeling. *Expert Systems with Applications, 31*(3), 570–594.

López-Valles, J. M., Fernández, M. A., Fernández-Caballero, A., & Gómez, F. J. (2005). Stereovision disparity analysis by two-dimensional motion charge map inspired in neurobiology. In *1st International symposium on brain, vision and artificial intelligence* (pp. 459–468). *Lecture notes in computer science*, *Vol. 3704*.

Matas, J., Chuma, O., Urbana, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing, 22*, 761–767.

Mira, J., Fernández, M. A., López, M. T., Delgado, A. E., & Fernández-Caballero, A. (2003). A model of neural inspiration for local accumulative computation. In *9th International conference on computer aided systems theory* (pp. 427–435). *Lecture notes in computer science*, *Vol. 2809*.

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

Muhlmann, K., Maier, D., Hesser, J., & Manner, R. (2002). Calculating dense disparity maps from color stereo images, an efficient implementation. *International Journal of Computer Vision, 47*(1–3), 79–88.

Pajares, G., & de la Cruz, J. M. (2001). Local stereovision matching through the ADALINE neural network. *Pattern Recognition Letters, 22*, 1457–1473.

Pajares, G., & de la Cruz, J. M. (2003). Stereovision matching through support vector machines. *Pattern Recognition Letters, 24*, 2575–2583.

Schreiber, A., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., van de Velde, W., et al. (2001). *Knowledge Engineering and Management: The CommonKADS Methodology*. The MIT Press.

Sumi, Y., Kawai, Y., Yoshimi, T., & Tomita, F. (2002). 3D object recognition in cluttered environments by segment-based stereo vision. *International Journal of Computer Vision, 46*(1), 5–23.

Venkateswar, V., & Chellappa, R. (1995). Hierarchical stereo and motion correspondence using feature groupings. *International Journal of Computer Vision, 15*, 245–269.

Wildes, R. (1991). Direct recovery of three-dimensional scene geometry from binocular stereo disparity. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(8), 761–774.

Wilson, R., & Knutsson, H. (1989). A multiresolution stereovision algorithm based on Gabor representation. In *Proceedings of the IEE international conference on image processing and applications* (pp. 19–22).

Zabih, R., & Woodfill, J. (1994). Non parametric local transforms for computing visual correspondence. In *Proceedings of the third European conference on computer vision* (pp. 150–158).