

Dynamic stereoscopic selective visual attention (DSSVA): Integrating motion and shape with depth in video segmentation

Antonio Fernández-Caballero^{a,b,*}, María T. López^{a,b}, Sergio Saiz-Valverde^a

^a Instituto de Investigación en Informática de Albacete (I3A), Universidad de Castilla-La Mancha, 02071 Albacete, Spain

^b Departamento de Sistemas Informáticos, Escuela Politécnica Superior de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

Abstract

Depth inclusion as an important parameter for dynamic selective visual attention is presented in this article. The model introduced in this paper is based on two previously developed models, dynamic selective visual attention and visual stereoscopy, giving rise to the so-called dynamic stereoscopic selective visual attention method. The three models are based on the accumulative computation problem-solving method. This paper shows how software reusability enables enhancing results in vision research (video segmentation) by integrating earlier works. In this article, the first results obtained for synthetic sequences are included to show the effectiveness of the integration of motion and shape features with depth parameter in video segmentation.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Visual attention; Video sequences; Stereovision; Motion detection; Software reusability

1. Introduction

Visual input is with no doubt the most powerful source of information used by humans to represent a dynamic scene (Buxton, 2003). Visual information consists of a series of redundant spatial data sets, which the brain processes fast and robustly. In order to advance in vision research it seems mandatory to provide robust reusable software (Costagliola, Francese, & Scanniello, 2003; Frakes & Kang, 2005). Indeed, according to Selby (2005): “Software reuse enables developers to leverage past accomplishments and facilitates significant improvements in software productivity and quality. Software reuse catalyzes improvements in productivity by avoiding redevelopment and improvements in quality by incorporating components whose reliability has already been established”.

The name dynamic selective visual attention comprises a set of image processing mechanisms to focus your vision on those regions where spatio-temporal relevant events take place (e.g., Ban & Lee, 2006; Le Meur, Le Callet, Barba, & Thoreau, 2006). A computational model capable of supporting the selective visual attention theory can be fundamental in dealing with the great amount of information in a video sequence. Our research group has recently developed a dynamic selective visual attention (DSVA) model (López, Fernández-Caballero, Fernández, Mira, & Delgado, 2007). Its application to moving objects detection and labeling has also been described (López, Fernández-Caballero, Mira, Delgado, & Fernández, 2006), as well as its approach to visual surveillance (López, Fernández-Caballero, Fernández, Mira, & Delgado, 2006a). We have approached the solution to the DSVA problem with a series of biologically inspired methods based on two fundamental mechanisms: (1) accumulative computation (Fernández et al., 1995) and (2) a generalized version of the calculation done by lateral inhibition networks called algorithmic lateral inhibition (Fernández-Caballero, Mira, Fernández, & López, 2001; Fernández-Caballero, Mira,

* Corresponding author. Address: Departamento de Sistemas Informáticos, Escuela Politécnica Superior de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain. Tel.: +34 967599200; fax: +34 967599224.

E-mail address: caballer@dsi.uclm.es (A. Fernández-Caballero).

Delgado, & Fernández, 2003). Our approach is similar to the attention model for dynamic vision presented by Backer and Mertsching (2003), where the characteristics extracted are symmetry, eccentricity, color, contrast and depth. In our case, the characteristics extracted and integrated to determine the attention focus are the difference in illumination level, size, height, width, density, height/width relationship, velocity, and acceleration (López, Fernández-Caballero, Fernández, Mira, & Delgado, 2006b), where velocity as well as acceleration is calculated using object motion in the visual scene as a base.

The greatest difference between the dynamic selective visual attention and the closest dynamic model (Backer & Mertsching, 2003) is that we did not use the depth characteristic, obtained thanks to the use of visual stereoscopy techniques. Currently, there are many algorithms developed to analyze depth in a scene. An excellent article on revision methods can be read in Brown, Burschka, and Hager (2003). Our group has also presented a comprehensive state of the art (López-Valles, Fernández, & Fernández-Caballero, 2005), in Spanish, as a previous step to the implementation of its own depth-securing model based on area correlation, which exploits the symbiosis between stereoscopy and movement (López-Valles, Fernández, & Fernández-Caballero, 2007). Recently, the stereovision-correspondence-analysis task has also been described in detail (Fernández-Caballero et al., 2007).

This article proposes the incorporation of depth as a fundamental characteristic of dynamic selective visual attention giving rise to the so-called dynamic stereoscopic selective visual attention (DSSVA). Accumulative computation problem-solving method is present in most steps throughout the DSSVA method introduced.

2. Proposal for dynamic stereoscopic selective visual attention

This section presents the proposed solution to the dynamic stereoscopic selective visual attention (DSSVA) problem. As Fig. 1 shows, DSSVA is broken into two sub-

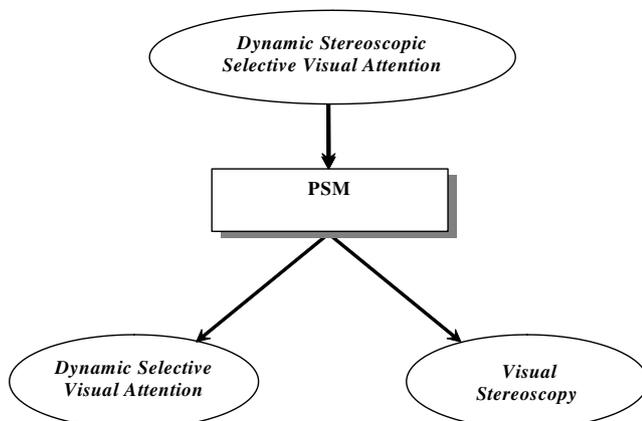


Fig. 1. Dynamic stereoscopic selective visual attention tasks.

tasks: the dynamic selective visual attention (DSVA) subtask and the visual stereoscopy (VS) subtask. This division does not mean that a selective attention technique is carried out first and later another one for stereoscopy. It is simply a name for the subtasks into which the system is broken down. In fact, DSVA is also carried out in the VS subtask choosing elements from a certain depth.

Fig. 2 shows the general diagram for the DSSVA system. The inputs to the system are the stereo image sequences and the observer's purpose, providing permitted values to the parameters size, height, width, density, height/width relationship, velocity, and acceleration. Thus, the parameters needed to accomplish the tasks are introduced into the system by means of the observer's purpose. The image sequences are grey-leveled (GL_L and GL_R) and correspond to the right and left image sequences, respectively.

The outputs supplied by this model are the right and left stereo attention foci (SAF_L and SAF_R), which are the set of figures of interest. If we observe the complete diagram in Fig. 2 we see that the DSVA is applied first to each input sequence in a parallel and independent way. We should point out that the DSVA feeds back on attention foci which come from VS.

The result of the DSVA for both input sequences is the monocular attention focus (MAF_L and MAF_R), and the segmented objects in focus (SOF_L and SOF_R) in current instant t and in the previous one $t - 1$. The MAFs are formed by the set of pixels in the objects and in the figures of interest in the input sequences with the observer's restrictions. In pixels where the MAFs include a figure, the SOF output contains the object selected from the input image, transformed into grey-level bands (GLB).

Finally, the observer's purpose restrictions related to stereovision, along with the aforementioned DSVA outputs, are the visual stereoscopy subtask inputs. Once the selection in the subtask is made, we go on to work with both sequences to obtain three-dimensional information by selecting the objects for a certain depth. This information is used to perfect DSVA. In the end, we get the attention foci (left stereo attention focus SAF_L and right stereo attention focus SAF_R) for the objects which have passed all the selection processes.

Next, we proceed to the specification of the two subtasks the system comprises.

2.1. Dynamic selective visual attention

This section briefly mentions the original DSVA solution (López, Fernández-Caballero, Mira, et al., 2006) which was developed to find a solution for selective visual attention in a single input sequence. Fig. 3 shows the DSVA breakdown into its four subtasks: (a) Feature Extraction: Obtaining dynamic characteristics from the pixels in the image, as well as the shape characteristics in the blobs and attention-grabbing figures. (b) Feature Integration: Application of the criteria established by the user to the characteristics extracted in the Feature Extraction

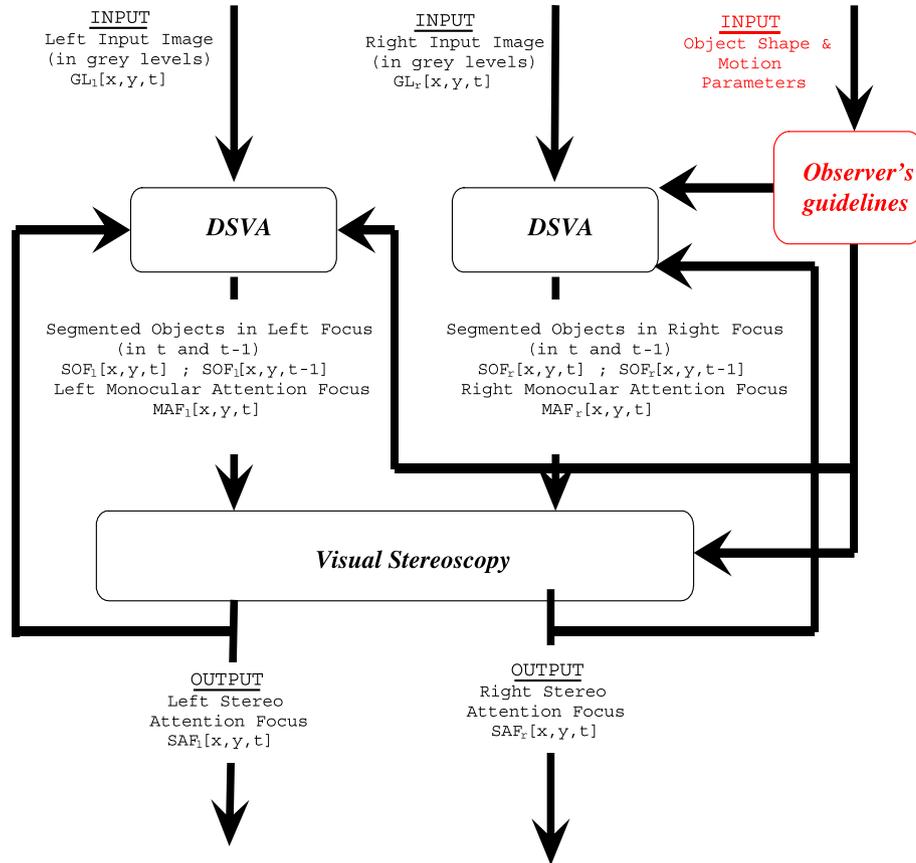


Fig. 2. Dynamic stereoscopic selective visual attention general diagram.

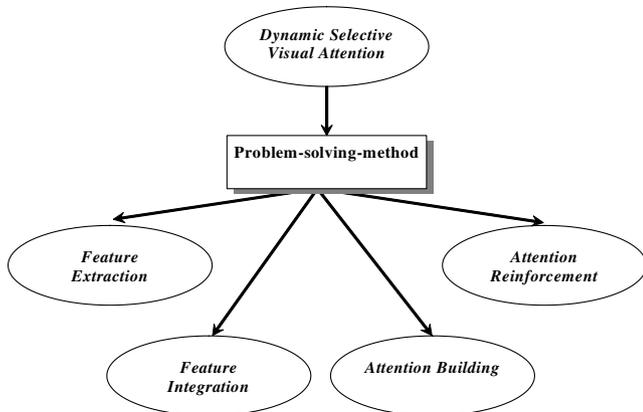


Fig. 3. Dynamic selective visual attention tasks.

subtask, filtering the pixels, blobs and the figures of interest to obtain pixels of interest. (c) Attention Building: Building blobs from the pixels of interest calculated in Feature Integration. (d) Attention Reinforcement: Building figures and holding the attention on certain figures (or objects) of interest in the sequence of images.

And, Fig. 4 shows the proposed new design for the DSVAs model, adapted for the integration with stereoscopy, as well as the subtasks' input and output. The input to the DSVAs subtask is a grey-leveled sequence of images

and the attention focus that comes from VS. In Fig. 4, this is shown as $GL_{l/r}[x, y, t]$ and $SAF_{l/r}[x, y, t]$, respectively. As output, we get the attention focus ($MAF_{l/r}[x, y, t]$) and the segmented objects present in the focus ($SOF_{l/r}[x, y, t]$ and $SOF_{l/r}[x, y, t - 1]$) in the current and previous instant. That is, the set of figures which comply with the observer's specifications. Notice that l/r means left or right in all cases.

According to Fig. 4, the more interesting associated formulas are explained next. Firstly, as part of subtask Attention Building, the input images captured in grey levels are segmented into a lower number ($nGLB$) of levels:

$$GLB_{l/r}[x, y, t] = \left\lfloor \frac{GL_{l/r}[x, y, t] \cdot nGLB}{GL_{max} - GL_{min} + 1} \right\rfloor + 1 \quad (1)$$

where GL_{max} and GL_{min} are the maximum and minimum grey-level values, respectively, for an input image. Typically, $GL_{max} = 255$ and $GL_{min} = 0$.

The subtasks Feature Extraction and Feature Integration obtain the Interest Map (IM), which filters image pixels according the parameters set up by the observer, reinforced by the current stereo attention focus (SAF). This map stores for each image pixel the result of the comparison with three discrepancy classes: "active", "inhibited" and "neutral", as a result of evaluating motion detection between two consecutive time instants and the observer's guidelines, as shown next:

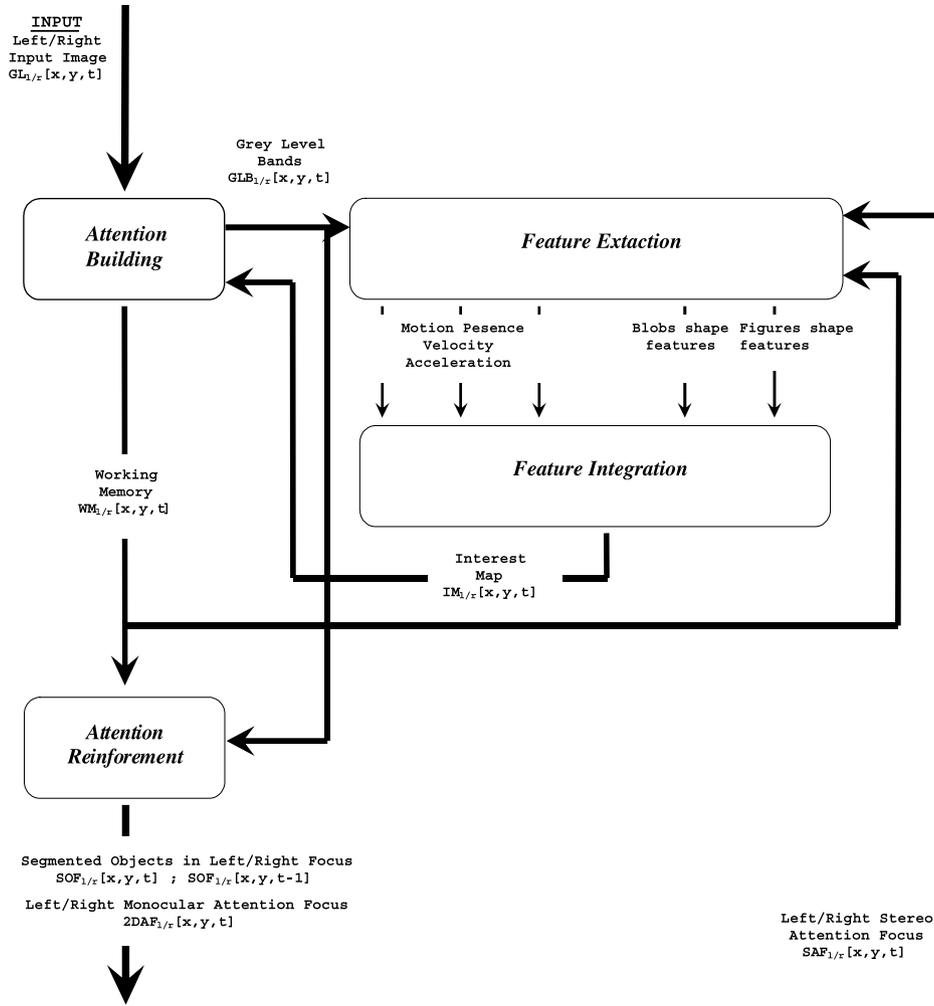


Fig. 4. Dynamic selective visual attention general diagram.

$$IM_{l/r}[x, y, t] = \begin{cases} v_{active}, & \text{if } (GLB_{l/r}[x, y, t] \neq GLB_{l/r}[x, y, t-1]) \\ & \wedge (guidelines - fulfilled) \\ v_{inhibited}, & \text{if } (GLB_{l/r}[x, y, t] \neq GLB_{l/r}[x, y, t-1]) \\ & \wedge \neq (guidelines - fulfilled) \\ v_{neutral}, & \text{if } (GLB_{l/r}[x, y, t] = GLB_{l/r}[x, y, t-1]) \end{cases} \quad (2)$$

Next the way in which the Working Memory (WM) is obtained for each grey-level band in subtask Attention Building is explained. The value for each pixel $[x, y]$ is v_{active} whenever the pixel is in “active” state in the IM. $v_{neutral}$ is assigned if the pixel is labeled as “neutral” and $v_{inhibited}$ if the pixel is “inhibited”:

$$WM_{i,l/r}[x, y, t] = \begin{cases} v_{active}, & \text{if } (GLB_{l/r}[x, y, t] = i) \\ & \wedge (IM_{l/r}[x, y, t] = v_{active}) \\ v_{neutral}, & \text{if } (GLB_{l/r}[x, y, t] = i) \\ & \wedge (IM_{l/r}[x, y, t] = v_{neutral}) \\ v_{inhibited}, & \text{otherwise} \end{cases} \quad (3)$$

The value at each pixel of the WM will be the maximum value of the WM_i calculated at each grey-level band:

$$WM_{l/r}[x, y, t] = \arg \max_i WM_{i,l/r}[x, y, t] \quad \forall i \in [1..nGLB] \quad (4)$$

Lastly, the monocular attention focus is obtained on a reinforcement basis in Attention Reinforcement subtask. Pixels that appear with a value equal to v_{active} in the WM reinforce attention in the MAF, whilst those that appear as a $v_{inhibited}$ decrement the attention value:

$$MAF_{l/r}[x, y, t] = \begin{cases} \max(MAF_{l/r}[x, y, t-1] - D_{MAF}, Ch_{min}), & \text{if } WM_{l/r}[x, y, t] = v_{inhibited} \\ \min(MAF_{l/r}[x, y, t-1] + C_{MAF}, Ch_{max}), & \text{if } WM_{l/r}[x, y, t] = v_{active} \end{cases} \quad (5)$$

where D_{MAF} and C_{MAF} are the discharge and charge constants, respectively, for the monocular attention foci. This is the first time when accumulative computation appears as a PSM (problem-solving method). Remember from Fernández-Caballero et al. (2007) that the general modality

of accumulative computation (AC) is the charge/discharge mode, which may be described by means of the following generic formula:

$$Ch[x,y,t] = \begin{cases} \min(Ch[x,y,t-1] + C, Ch_{\max}), & \text{if } (P[x,y,t] - \text{fulfilled}) \\ \max(Ch[x,y,t-1] - D, Ch_{\min}), & \text{otherwise} \end{cases} \quad (6)$$

The temporal accumulation of a persistency measure of the property $P[x,y,t]$ gotten at each time instant t at each pixel $[x,y]$ is calculated. If the property P is fulfilled, the charge value Ch goes incrementing by increment charge value C up to reaching Ch_{\max} , whilst, if property P is not fulfilled, the charge value goes decrementing by decrement value D down to Ch_{\min} . Values of C , D , Ch_{\max} and Ch_{\min} have to be fixed according to the application characteristics.

The segmented objects in focus are also obtained through a simple filtering:

$$SOF_{l/r}[x,y,t] = \begin{cases} GLB_{l/r}[x,y,t], & \text{if } MAF_{l/r}[x,y,t] > Ch_{\min} \\ -1, & \text{otherwise} \end{cases} \quad (7)$$

2.2. Visual stereoscopy

Fig. 5 shows the proposal for the obtaining of depth in stereoscopic image sequences (Fernández-Caballero et al., 2007; López-Valles et al., 2007), except adapted to the inclusion of selective visual attention, which was not accounted for in the original works. Said subtasks are: (a) Obtaining of the 2D Charge Map: This obtains the dynamic characteristics of the pixels in the input image sequences by means of accumulative computation mechanisms. (b) Charge Disparity Analysis: This subtask is in charge of the matching. That is, finding the correspondence between the pixels in the left images and those in the right images in such a way that they reference the same pixel in the real scene. (c) Obtaining of Depth: Its result is the potential objects which meet the requirements introduced by the observer and the information about the depth of the selected objects. (d) 3D Attention Reinforcement:

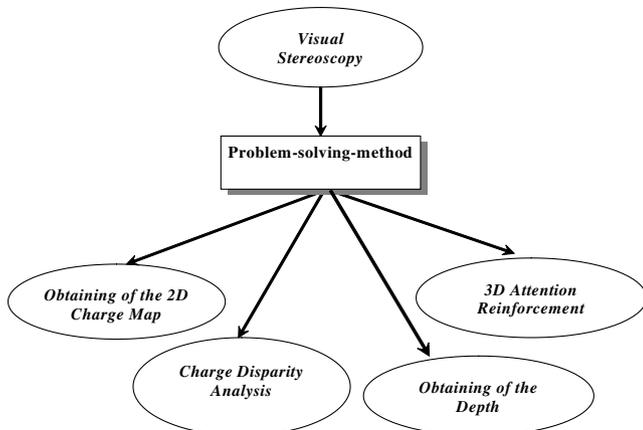


Fig. 5. Visual stereoscopy tasks.

Building figures and holding the attention on certain figures (or objects) of interest to the observer of the sequence of images. It returns the final stereo attention foci (SAF). Notice that this latter subtask was not present in the previous versions of visual stereoscopy.

In Fig. 6, we now see the complete design of the proposed model for VS. The inputs are the segmented objects in right and left focus (segmented objects in left focus $SOF_l[x,y,t]$ and $SOF_l[x,y,t-1]$, and segmented objects in right focus $SOF_r[x,y,t]$ and $SOF_r[x,y,t-1]$) and the left and right monocular attention focus ($MAF_l[x,y,t]$ and $MAF_r[x,y,t]$) that come from the DSVA subtask done in parallel on both initial (left l and right r) input sequences. This subtask's outputs are the same as those from the final system, that is, the final SAFs (left stereo attention focus $SAF_l[x,y,t]$ and right stereo attention focus $SAF_r[x,y,t]$). We also have the inputs and outputs for the four great subtasks which make up the VS subtask.

Fig. 6 is also explained in terms of the most relevant equations that obtain the parameters in the subtasks for the VS. Firstly, there is a calculation of the so-called motion charge memories through a couple of subtasks Obtaining of the 2D Charge Map, inspired in the biological visual hemifields. For this purpose, left and right motion charge memories (CM) are obtained by means of AC PSM to provide accurate information about movement:

$$CM_{l/r}[x,y,t] = \begin{cases} \max(CM_{l/r}[x,y,t-1] - D_{CM}, Ch_{\min}), & \text{if } SOF_{l/r}[x,y,t] = SOF_{l/r}[x,y,t-1] \\ \min(CM_{l/r}[x,y,t-1] + C_{CM}, Ch_{\max}), & \text{if } SOF_{l/r}[x,y,t] \neq SOF_{l/r}[x,y,t-1] \end{cases} \quad (8)$$

where D_{CM} and C_{CM} are the discharge and charge constants, respectively, for the motion charge memories.

For the left visual hemifield, we obtain the left–left and right–left motion charge memory ($MM_{ll}[x,y,t]$ and $MM_{rl}[x,y,t]$):

$$MM_{l/r,l} \left[x + \frac{NC}{2}, y, t \right] = CM_{l/r} \left[x - \frac{NC}{2}, y, t \right] \quad (9)$$

And, the operation is simpler for the left–right and right–right motion charge memory ($MM_{lr}[x,y,t]$ and $MM_{rr}[x,y,t]$):

$$MM_{l/r,r}[x,y,t] = CM_{l/r}[x,y,t] \quad (10)$$

The disparity map gotten in subtask Charge Disparity Analysis is a three-dimensional depth map, which shows the depth of points in the scene where there has been movement. This decision stands on the disparity with the greatest reliability, taking the epipolar, the ordering and the disparity restrictions into account:

$$DM[x,y,t] = d|CCM[x,y,d,t] \geq CCM[x,y,e,t] \quad \forall (d,e), 0 \leq d, e \leq d_{\max} \quad (11)$$

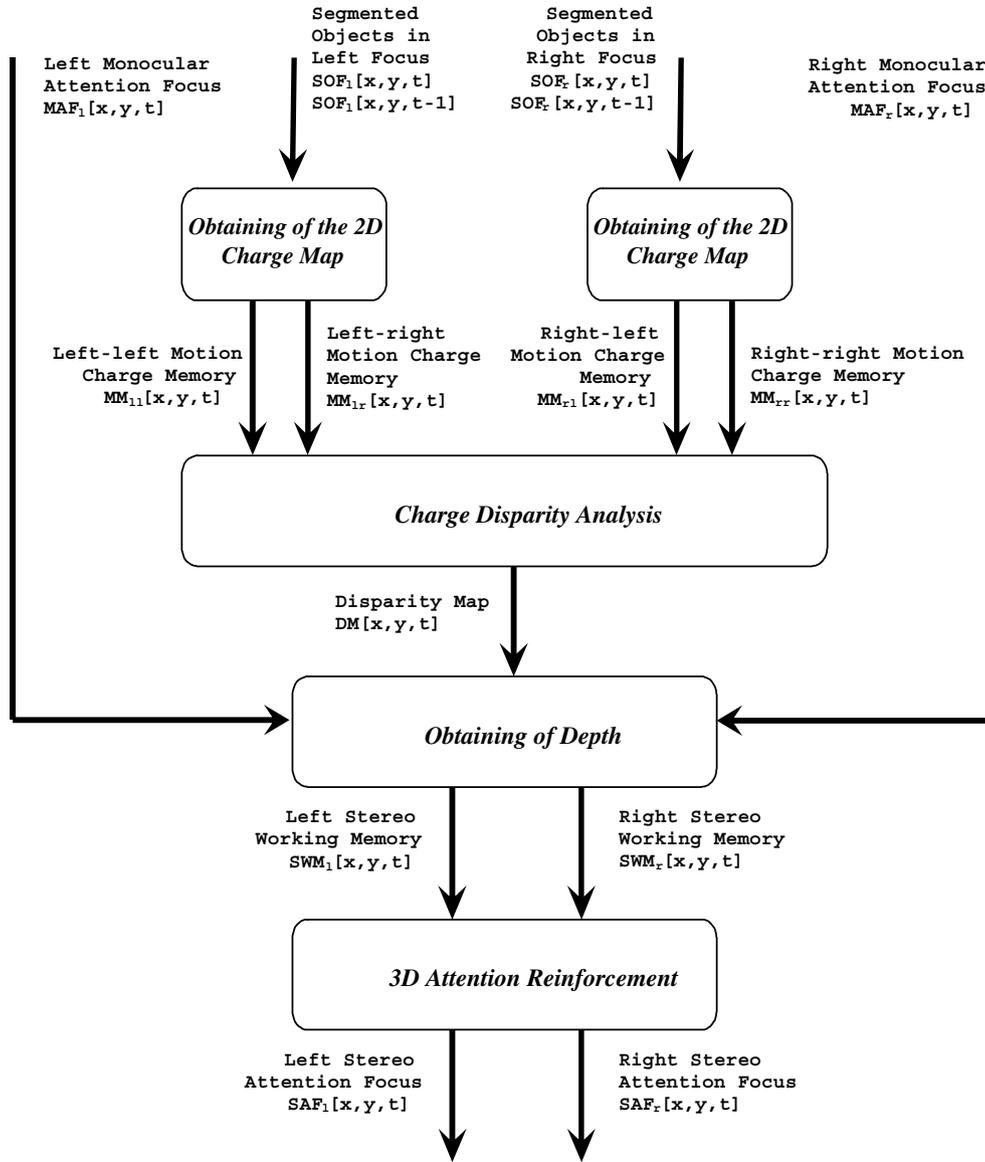


Fig. 6. Visual stereoscopia general diagram.

where CCM is the charge correspondence memory resulting from the comparison of the charges present in the motion charge memories obtained as hemifields, and d_{max} is the maximum permitted disparity between a couple of left and right images.

After Obtaining of Depth subtask, in a similar way as described in DSV, the left and right stereo working memories (SWM) are calculated:

$$SWM_{l/r}[x, y, t] = \begin{cases} v_{active}, & \text{if } (MAF_{l/r}[x, y, t] \neq Ch_{min}) \\ & \wedge (guidelines - fulfilled) \\ v_{inhibited}, & \text{otherwise} \end{cases} \quad (12)$$

Lastly, the couple of stereo attention foci (SAF) is obtained on a reinforcement basis in subtask 3D Attention Rein-

forcement – just as for the monocular attention foci. Pixels that appear with a value equal to v_{active} in the SWM reinforce attention in the SAF, whilst those that appear as a $v_{inhibited}$ decrement the attention value:

$$SAF_{l/r}[x, y, t] = \begin{cases} \max(SAF_{l/r}[x, y, t-1] - D_{SAF}, Ch_{min}), & \text{if } SWM_{l/r}[x, y, t] = v_{inhibited} \\ \min(SAF_{l/r}[x, y, t-1] + C_{SAF}, Ch_{max}), & \text{if } SWM_{l/r}[x, y, t] = v_{active} \end{cases} \quad (13)$$

where D_{SAF} and C_{SAF} are the discharge and charge constants, respectively. Again we are in front of the PSM accumulative computation.

3. Data and results

In this section, we show the results obtained in a series of input stereo image sequences after the proposed DSSVA method was applied.

3.1. Cubes scene

This synthetic scene, called Cubes, is made up of 240 frames which have a pixel size of 320×240 . The aim of this example is to hold the attention on an unbending object which does not change its perspective with respect to the cameras and which moves in space.

3.1.1. Example 1

In this first example, we hope to accomplish visual attention using only size and shape parameters as criteria. The parameters which stand out the most for the desired object are, in this case: height = 60 pixels, width = 60 pixels and maximum size = 3000 pixels. Notice that these are the observer's guidelines concerning the shape of the objects. The results obtained are presented in Fig. 7, where columns (c) and (d) show the attention focus obtained for the respective cameras.

We should point out that with the features displayed in the aforementioned table, we are able to focus the attention on the small cube throughout the sequence. The results can be rated as very good since we have selected and held the attention on the object throughout time, giving a specific size and shape criteria which excluded the rest of the image.

3.1.2. Example 2

In this second example, we use depth as the selection criterion. That is to say that we will give size and shape criteria

which will select both cubes (we increase shape = 500 pixels, width = 500 pixels and maximum size = 20,000 pixels), but we will only keep the one whose depth is close to the camera (minimum disparity = 6 pixels; maximum disparity = 8 pixels). The results obtained are presented in Fig. 8, where, again, columns (a) and (b) show the input images to the cameras and where column (c) now shows the depth map obtained and columns (d) and (e) show the attention focus obtained for the right and left cameras.

We also notice that in frame 60, with the exposed features in the table mentioned earlier, it is possible to focus the attention on both cubes. This is because, although we can see in the initial input images that the big cube is further away than in instant $t = 1$, there are areas situated in the indicated disparity and the small cube is close enough to appear too. In the next frame shown, $t = 120$, we see clearly that the cubes in the input images are further away. This has an effect on the output because none of the cubes appear there. In the last frame in this study, $t = 220$, the big cube has come closer. We know this because it is the only one seen in the final result.

3.2. Ball scene

This example uses depth as a selection criterion. Size and shape criteria, which select all the objects that appear in the scene, will be given. This is done this way because the objects' geometry varies a lot. They revolve around themselves and move close and move away constantly, making it very difficult to determine their shapes and sizes. For this reason, their depth will be used as a selection criterion. We have use the parameter maximum disparity $d_{\max} = 3$ pixels, which will enable us to film the figures at the back of the scene, located far away from the camera. The results are

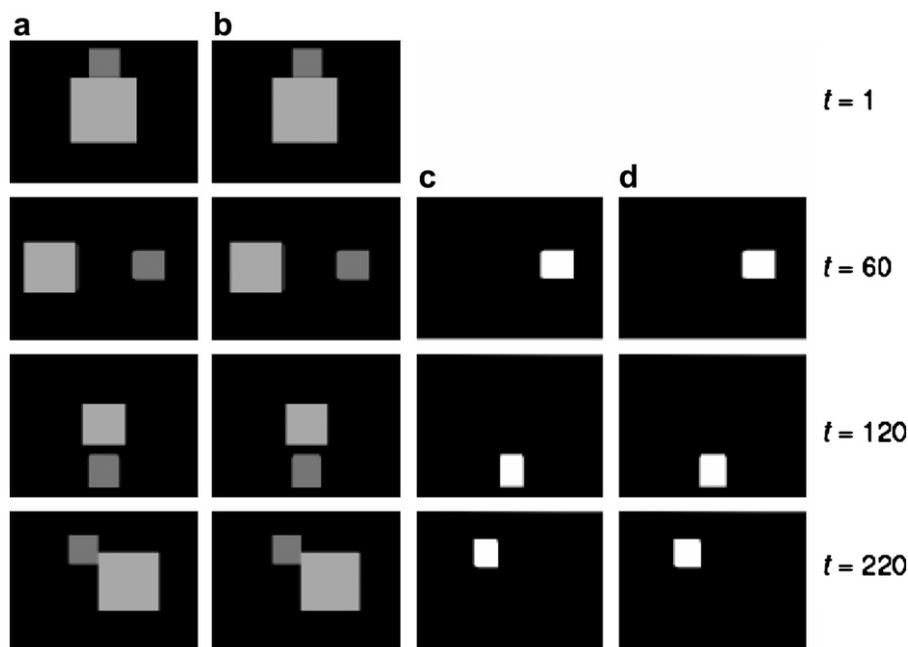


Fig. 7. Results for cubes scene, Example 1. (a) Left input image. (b) Right input image. (c) Left stereo attention focus. (d) Right stereo attention focus.

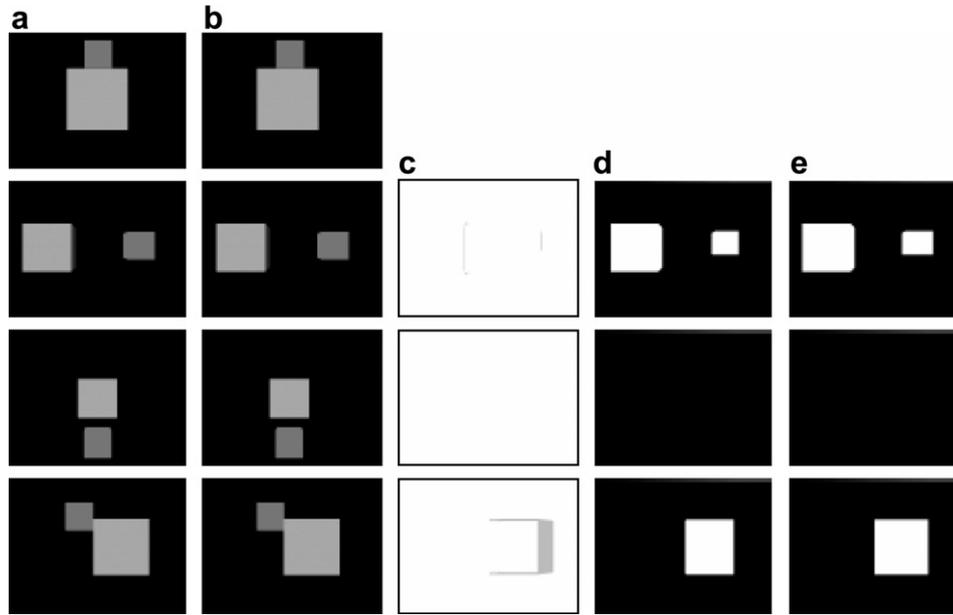


Fig. 8. Results for cubes scene, Example 2. (a) Left input image. (b) Right input image. (c) Depth map. (d) Left stereo attention focus. (e) Right stereo attention focus.

presented in Fig. 9, where columns (a)–(c) mean the same for Fig. 8. Columns (d) and (e) show the overlapping of the attention focus obtained with the corresponding input images for the respective cameras.

We see right away that the results are not as optimum as in the cubes scene. The reason for this is that if the figures rotate or do not have edges, the proposed system does not work properly and some pixels' disparity will be measured incorrectly. This small number of pixels almost always takes up as much disparity as possible, thus the Visual Stereoscopy subtask will allow the attention focus that comes from the dynamic selective visual attention subtask to go through.

We will now take a closer look at the results in Fig. 9. Since the cube and the ring body are at the back of the scene in frame 21, they display a dark gray level in the depth map. They cube does not show up in the depth map in frame 81. In the next frame, 111, the cube, which is moving closer to the camera and pivoting, should have disappeared from the Depth Map and therefore, from the attention focus. However, due to the turning motion, the system fails again. On the other hand, the ball blocks up the ring body, but not completely. Due to the closeness and to the restrictions imposed, there is a gap in the ring body's depth map, which belongs to the ball. In the last

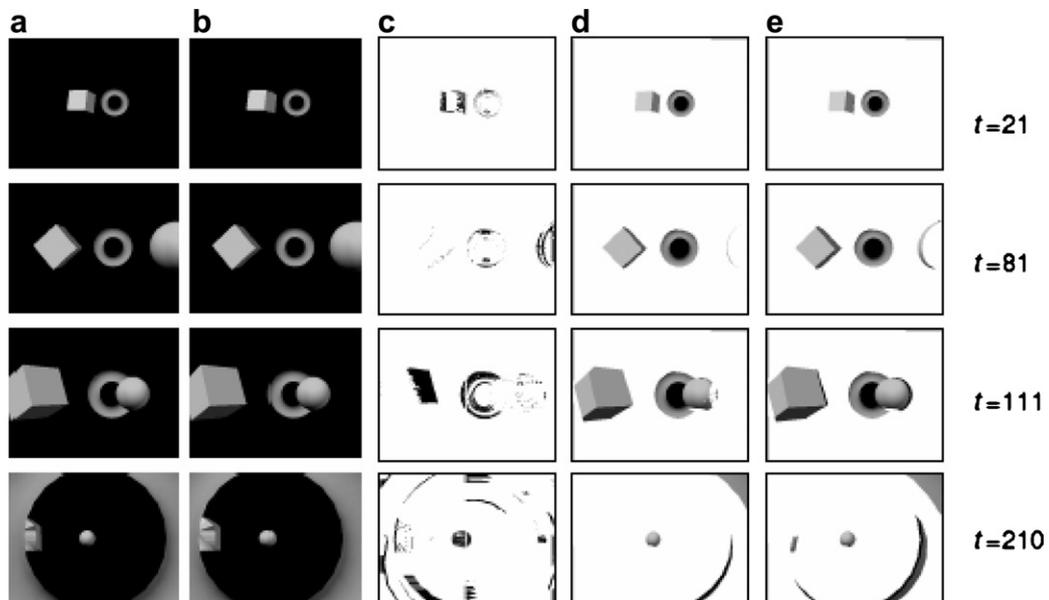


Fig. 9. Results for ball scene. (a) Left input image. (b) Right input image. (c) Depth map. (d) Left stereo attention focus on left input image. (e) Right stereo attention focus on right input image.

frame, the ball is situated in the darkest pixels, which are located at the far end.

4. Conclusions

This article has presented the inclusion of the depth parameter in selective visual attention, based on previous experience on both subjects – dynamic selective visual attention (DSVA) and visual stereoscopy (VS) – separately. We have shown that software reusability is perfect to enhance applications from previously acquired knowledge. As a result of the way we structured the task, by dividing the DSSVA task into two subtasks, DSVA and VS, we started out with a great advantage. It is possible to do only subtask DSVA and manage to focus and hold the attention with motion, shape and size parameters. By using subtask VS, we obtain depth information about the objects in the scene. Notice again, that our accumulative problem-solving method is used in many steps throughout the DSSVA model proposed. Thus, the model benefits from a well-proven and efficient base method.

With regard to the initial results on synthetic images, we can say that the success rate is encouraging. When the objects keep their perspective, their geometry is flat in relation to the camera and object motion is gentle and not too fast, the results are good, as has been proved in the quite complex cubes scene. Nonetheless, if the objects pivot, their geometry is not regular, they lack edges or move too fast, the results are not as good (Ball scene).

Acknowledgements

This work is supported in part by the Spanish CICYT TIN2004-07661-C02-02 grant and the Junta de Comunidades de Castilla-La Mancha PBI06-0099 grant.

References

- Backer, G., & Mertsching, B. (2003). Two selection stages provide efficient object-based attentional control for dynamic vision. In *Proceedings of the international workshop on attention and performance in computer vision*, 9–16.
- Ban, S.-W., & Lee, M. (2006). Selective attention-based novelty scene detection in dynamic environments. *Neurocomputing*, 69(13–15), 1723–1727.
- 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.
- Buxton, H. (2003). Learning and understanding dynamic scene activity: A review. *Image and Vision Computing*, 21(1), 125–136.
- Costagliola, G., Francese, R., & Scanniello, G. (2003). A visual system supporting software reuse in the banking legacy system context. *International Journal of Software Engineering and Knowledge Engineering*, 13(1), 83–101.
- Fernández-Caballero, A., López, M. T., Mira, J., Delgado, A. E., López-Valles, J. M., & Fernández, M. A. (2007). Modelling the stereovision-correspondence-analysis task by lateral inhibition in accumulative computation problem-solving method. *Expert Systems with Applications*, 33(4).
- Fernández-Caballero, A., Mira, J., Delgado, A. E., & Fernández, M. A. (2003). Lateral interaction in accumulative computation: A model for motion detection. *Neurocomputing*, 50C, 341–364.
- Fernández-Caballero, A., Mira, J., Fernández, M. A., & López, M. T. (2001). Segmentation from motion of non-rigid objects by neuronal lateral interaction. *Pattern Recognition Letters*, 22(14), 1517–1524.
- Fernández, M. A., Mira, J., López, M. T., Alvarez, J. R., Manjarrés, A., & Barro, S. (1995). Local accumulation of persistent activity at synaptic level: Application to motion analysis. *From Natural to Artificial Neural Computation*, 137–143.
- Frakes, W. B., & Kang, K. (2005). Software reuse research: Status and future. *IEEE Transactions on Software Engineering*, 31(7), 529–536.
- Le Meur, O., Le Callet, P., Barba, D., & Thoreau, D. (2006). A coherent computational approach to model bottom-up visual attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5), 802–817.
- López, M. T., Fernández-Caballero, A., Fernández, M. A., Mira, J., & Delgado, A. E. (2006a). Visual surveillance by dynamic visual attention method. *Pattern Recognition*, 39(11), 2194–2211.
- López, M. T., Fernández-Caballero, A., Fernández, M. A., Mira, J., & Delgado, A. E. (2006b). Motion features to enhance scene segmentation in active visual attention. *Pattern Recognition Letters*, 27(5), 469–478.
- López, M. T., Fernández-Caballero, A., Fernández, M. A., Mira, J., & Delgado, A. E. (2007). Dynamic visual attention model in image sequences. *Image and Vision Computing*, 25(5), 597–613.
- 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 labelling. *Expert Systems with Applications*, 31(3), 570–594.
- López-Valles, J. M., Fernández, M. A., & Fernández-Caballero, A. (2005). Concepts and techniques of stereo vision for computers. *Inteligencia Artificial – Revista Iberoamericana de Inteligencia Artificial*, 9(27), 35–62.
- López-Valles, J. M., Fernández, M. A., & Fernández-Caballero, A. (2007). Stereovision depth analysis by two-dimensional motion charge memories. *Pattern Recognition Letters*, 28(1), 20–30.
- Selby, R. W. (2005). Enabling reuse-based software development of large-scale systems. *IEEE Transactions on Software Engineering*, 31(6), 495–510.