Real-time vision guided movement with reconfigurable Hardware

Christian Morillas, Samuel Morillas, Eduardo Ros, Antonio F. Díaz, Begoña del Pino and Francisco J. Pelayo

Departamento de Arquitectura y Tecnología de Computadores, Universidad de Granada, 18071, Granada, Spain
eduardo@atc.ugr.es, fpelayo@ugr.es

Abstract. This work summarizes the implementation and test of a vision guided mobile system. A direct sensor-motor interaction scheme leads the mobile towards the direction in which less optic flow is detected. The described system is inspired in the visuo-motor system of some insects, it uses a low-cost CMOS camera, whose digital output is captured and processed by a UP1x board of Altera. The designed digital modules that form the processing kernel of the system have been defined in VHDL. They implement real-time compression and change detection in both lateral sides of the visual field, with thresholds that are adapted depending on the global luminance of scene.

1 Introduction

The extraction and processing tasks of the visual information in real time need of high computational power. Furthermore, on one hand the visual information extraction usually requires such a computational complexity that makes difficult the use of low cost systems, but on the other hand visual information represents a very useful source for autonomous mobile systems. Clear examples of these systems are the micro-robots, in which is easy to incorporate vision front-ends through low cost micro cameras, but it is difficult to exploit this kind of sensorial information due to the low processing capabilities of these micro systems and the processing complexity required by the visual structure extraction task. For example, the visual systems of some insects such as the Dropsophila or domestic fly, extract information of the optic flow mostly driven by the ego-motion of the insect. It has been proved the existence of a very direct interaction between the sensor elements (composed eye specially sensitive movements in certain directions) and the motor elements that drive the wings. Such a direct feed-forward interaction (by means of short neuronal connection paths) provides these insects a high flying control efficiency despite their rudimentary neuronal systems [1, 2].

The hardware implementation of processing schemes based on these biological visual systems represent a valid option because simplified models [3] may be viable despite the computational resource constraints of the current implementation technologies. The current Field Programmable Logic Devices (FPLDs) are specially indicated for these kind of implementations because of their high parallelism

Dynamic Perception; R.P. Würtz, M. Lappe (Eds.); Aka and IOS Press
ISBN 3-89838-032-7
ISBN 1-58603-305-0

possibilities. This allows to allocate in the same chip the visual information extraction modules and other sensorial sources modules (multi-modal perception schemes). All this can work concurrently with the processing kernel that deal with these different information sources and makes decisions to evolve the global system in its environment.

A mobile platform (FrankeBot) [4] has been developed within the framework of a docent innovation project supported by the University of Granada (Docent Quality and Evaluation Department) that incorporates multiple sensors, digital and analog communication elements using microcontrollers and FPLDs as computational substrates. Although the system described in this work can be used in any mobile platform based on FPLDs, it was originally conceived to be integrated in the FrankeBot, in order to provide in real time measurements of the radial optic flow detected in both sides of the visual field. This optic flow is produced by the relative shift of the present features with respect to the mobile system depending on spatial and temporal differences. The optic flow provided by these features shifts in the visual field increases with higher spatial contrast patterns and when they are closer to the mobile system.

In the next section, the structure of the reference model is briefly introduced. In Section III is described the implementation of the model with diverse VHDL modules that have been synthesised with the environment Max+PlusII of Altera [5]. Finally in Section IV the final implementation is tested in a mobile platform.

2 Processing Module Structure

The flies have two composed eyes that are composed of multiple small eyes (elementary sensors) whose outputs are cooperatively collected to generate an activity pattern when a coherent movement is detected in a certain direction. A direct implementation of this movement information extraction scheme (reduced to one dimension and based on discrete analog optic sensors) is described in [2]. Furthermore, diverse VLSI approaches have been proposed that combine in the same chip, the sensors and the required analog processing circuits to extract the optic flow [3], that are called Focal-plane solutions.

In our case, the visual information is captured from a Back and White CMOS camera with digital output. In a first step the visual field is divided in two areas (left and right). Both areas will be processed separately producing different activity levels that will drive the mobile system. Each of these two areas is composed of set of elementary sensors, able to detect changes in the light intensity that reaches the receptive fields. In Fig. 1 is represented the interaction between these elementary sensors (only three in each side in this example), adding their contributions in order to produce a final estimation of the optimum direction in which the movement should evolve.

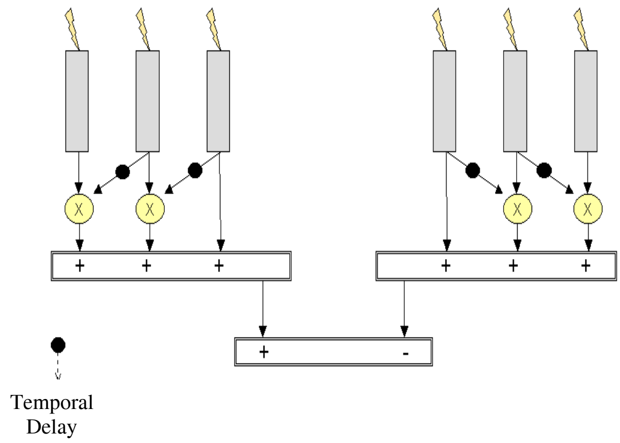

Fig. 1. Basic processing scheme to evaluate the radial optic flow. Each sensor element is a differentiator that contributes to the accumulated activity at each side of the visual field when a temporal change in the light intensity is detected.

The temporal delay and the product elements that link the neighbour sensors facilitate he contribution of those stimuli that move from the centre toward the sides and with a certain velocity. This is the movement pattern that produces an approaching object. Test results obtained from the software implementation of this model and other simplified versions have motivated the incorporation of the delay elements. The accumulated activity in one or other side of the visual field will be higher when more lateral radial flow is detected in these areas. The global accumulated activity will be calculated as the difference of these two levels (left and right sides) as illustrated in Fig. 1. This global estimation can be directly used to drive the mobile system.

In the fly the interaction between the sensors (composed eyes) and the actuators (wing motors) is almost direct; the accumulated activity is used to control the intensity that drives the wing motors. The relative movement of those efficiently controls the fly movement direction, leading to a natural tendency to get away of objects or to avoid any object with an approaching trajectory that would produce a "repelling" global optic flow signal. This natural tendency that facilitates the navigation avoiding objects is combined with an antagonist persecution and capture tendency that helps the male fly to track and reach the female fly. For this purpose the male fly eyes have a specific zone in the superior frontal eye called "love spot" [1, 2].

3 Hardware implementation of the model

As indicated in the previous section, the implementation here described uses a CMOS camera (model M4088, [6]) that uses a chip of OmniVision (OV5017, [7]).

This camera provides images of 384 columns and 288 rows (sampling the data in rows in ascending order to cover the different columns).

We make two grouping processes in order to compress the image in origin and reduce drastically the memory requirements. In a first step, we group all the pixels in the same column (adding the data). This grouping process is much easier if the scanning is done in a column order instead of a row order, and this motivates that the final orientation of the camera is rotated 90º degrees. After this reallocation, the camera provides images of 288 columns and 384 rows. In this way, the 384 data of each column are easily added (during the scanning process), obtaining what we call "macro-columns". The first consequence of this groping is that we lose any sensibility to movements in any vertical orientation, therefore we restrict our system to be able to compute only horizontal optic flow. In a second step, we group some adjacent macro-columns, computing the average, the resulting data are called "macro-pixels". In this way we gain robustness to noise although we loose resolution, we are not able to detect slight horizontal movements that could take place in this macro-pixels. Fig. 2 illustrates how each image is compressed spatially. Grouping 8 macro-columns to form a macro-pixel, we finally have 36 macro-pixels (numbered from 0 to 35) that will be used as elementary sensors for our system. The activity produced by each of this cells will be computed as the difference between absolute values calculated through the grouping procedures and the values corresponding to the previous image (temporal changes).

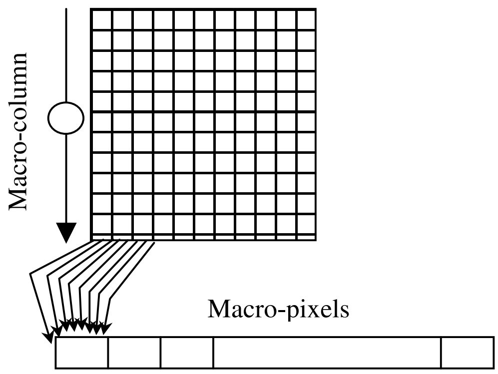

Fig. 2. Grouping of macro-columns and macro-pixels of the image captured with a rotated (90º) camera.

Now we distribute the elementary sensors to conform the left and right eye: the left eye is composed by the sensors 2 to 15 and the right eye is composed by the sensors 20 to 33. The central zone of the image (macro-pixels 16 to 19) has been eliminated as well as the lateral boundaries (macro-pixels 0,1 and 34,35).

For each frame is also obtained the average global activity that is related with the scene illumination conditions. This value is used to choose the range of significant bits in the accumulated activity, and provides the system with a certain robustness to changes in the illumination conditions (this changes are much more frequent in mobile systems than in static scenes).

The VHDL design of the system has been structured in the modules that perform the different tasks. The complete system, including the bits range selector, uses approximately 30 % of the logic cells of a CPLD-SRAM Flex-10K70 of Altera, and a 12 % of the 18Kbits of its memory blocks (EAB). The most complex module is the one that implements activity calculation stage, that uses 22 % of the logic cells of the CPLD (about 15400 logic gates). This module is structured as a three stage segmented processing pathway. This module compares the value of the captured macro-pixel with the previous value, and the result is multiplied by the delayed activity of the neighbour macro-pixel. This final magnitude is accumulated sequentially in both sides of the visual field.

4 Test of the system

For the test of the system we have used a mobile platform based on the PICBOT-2 of Microsystems Engineering [8], with an added UP-1X board of Altera with a CMOS camera. Figure 3. shows the complete system set up. An additional camera and a micro RF broadcast video to enable the remote recording of sequences from the point of view of the mobile system.

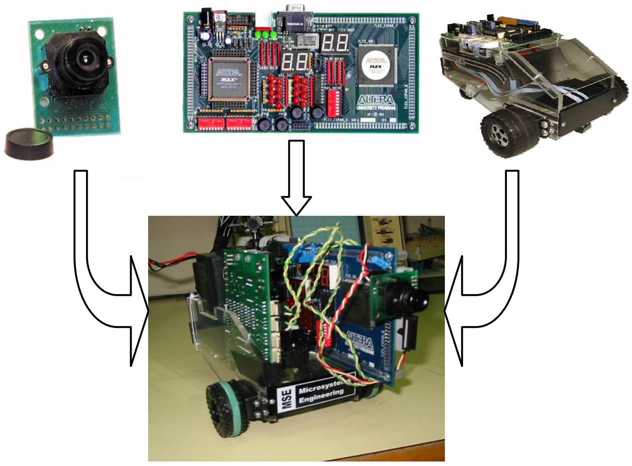

Fig. 3. Final system set up: CMOS camera, CPLD board and PICBOT-2 platform.

In Fig. 4.a can be seen the result of integration of the macro-pixels of the image (the VGA synchronism signals generation module described in [9] has been used for the visualization task). On the screen appears a dark band following the position of the black cylinder waved in front of the camera (in this case on the right side of the visual field). Finally, Fig. 4.b shows a photograph of one of the experiments of the mobile platform moving through the black cylinder wood. The response speed of the system is highly dependent of the number of processed frames per second. Further work will focus on adapting this number depending on the optic flow intensity detected in each instant.

Fig. 4. (a) Macro-pixels capturing a moving cylinder. (b) Experimental set up: Mobile robot in the cylinder wood.

Acknowledgement

This work has been carried out in the framework of the Docent Innovation Project called Hardware/Software Environment for experiments based on micro-robots, supported by the University of Granada. It has also received support from the EU research projects CORTIVIS (QLK6-CT-2001-00279, http://cortivis.umh.es) and ECOVISION (IST-2001-32114, http://www.pspc.dibe.unige.it/ecovision).

References

1. FlyBrain: *An Online Atlas and Database of the Drosophila Nervous System.* http://flybrain.neurobio.arizona.edu/
http://student.biology.arizona.edu/honors96/group11/URLS.htm
2. N. Franceschini, J.M. Pichon and C. Blanes: *From insect vision to robot vision,* Phil. Trans. Royal Society of London B 337, pp 283-294 (1992).
3. R.R. Harrison: *An analog VLSI motion sensor based on the Fly Visual System,* Ph.D. Thesis. California Institute of Technology. (1992). http://www.klab.caltech.edu/~harrison/abstracts/thesis.html
4. R. Agís, R. Carrillo, A. Cañas, B. del Pino, F.J. Pelayo: *Entorno Hardware-Software para experimentación basado en un micro-robot*. II Jornadas sobre Computación Reconfigurable y Aplicaciones, Granada, 18-20 Sept., 2002.
5. Altera. http://www.altera.com/
6. M4088 http://www.electronic-kits-and-projects.com/kit-files/cameras/d-m4088.pdf
7. OmniVision http://www.ovt.com/
8. Microsystems Engineering. http://www.microcontroladores.com/
9. Hamblen et al.: *Rapid Prototyping of Digital Systems.* Kluwer Academic Publishers. 2001.