# International Computer Vision Summer School 2010

ICVSS 2010

12 – 17 JULY

CONSORZIO CATANIA RICERCHE

SICILY, ITALY

# REAL-TIME ATTENTION SYSTEM ON-CHIP

**Barranco F., Díaz J., Tomasi M., Vanegas M., Ros E.**

{fbarranco, jdiaz, mtomasi, mvanegas, eduardo}@atc.ugr.es. Dep. of Computer Architecture and Technology University of Granada.

**Abstract.** This work proposes a novel architecture for the implementation of a real-time attention system on-chip. It integrates a bio-inspired attention model with the low-level visual primitives (optical flow and disparity). It computes a new saliency map biased with the weighted primitives, converting it in the most complex existing platform for scene understanding. The system is developed for hardware reconfigurable devices (FPGAs) as a hardware accelerator and could be used as a stand-alone platform.

## 1. Motivation

"*Vision is the process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information*" (David Marr).

**Attention**: A ***'stagelight'*** that successively focuses on different areas.

✓ *Selecting* an specific area.

✓ *Enhancing* the representation of the objects in this zone [1].

Two forms of attention (in parallel):

✓ ***Intrinsically*** salient stimuli, processed in a pre-attentive way (they are salient in a specific context) .

✓ ***Voluntary*** attention, with a changing criteria for the selection, it is more powerful than the first one (but about 10 times slower).

Most surprising fact: with a constrained computational capability, attention accomplishes near real-time performances, reducing intelligently the massive amount of data from the visual system (the optic nerve generates about $10^7 – 10^8$ b/s).

The Koch and Itti works [2][3]:

✓ Efficiently computational *bio-inspired model*.

✓ *Scene understanding* by rapidly selecting of the conspicuous locations.

Our goals:

✓ *Improve* and *extend* the previously defined attention model.

✓ Encapsulate the algorithm *on-chip*.

✓ Achieve a good trade-off between *accuracy* and *performances* (real-time).

## 2. Attention on-chip architecture

Our current system:

- Receives the input from a stereo pair of calibrated cameras
- Computes local contrast descriptors, optical flow, disparity (hierarchical computation)
- Obtains a good trade-off between accuracy and the real-time performances [4][5][6].
- *On-chip*, using an FPGA device (*low dimensions* and *power consumption*)

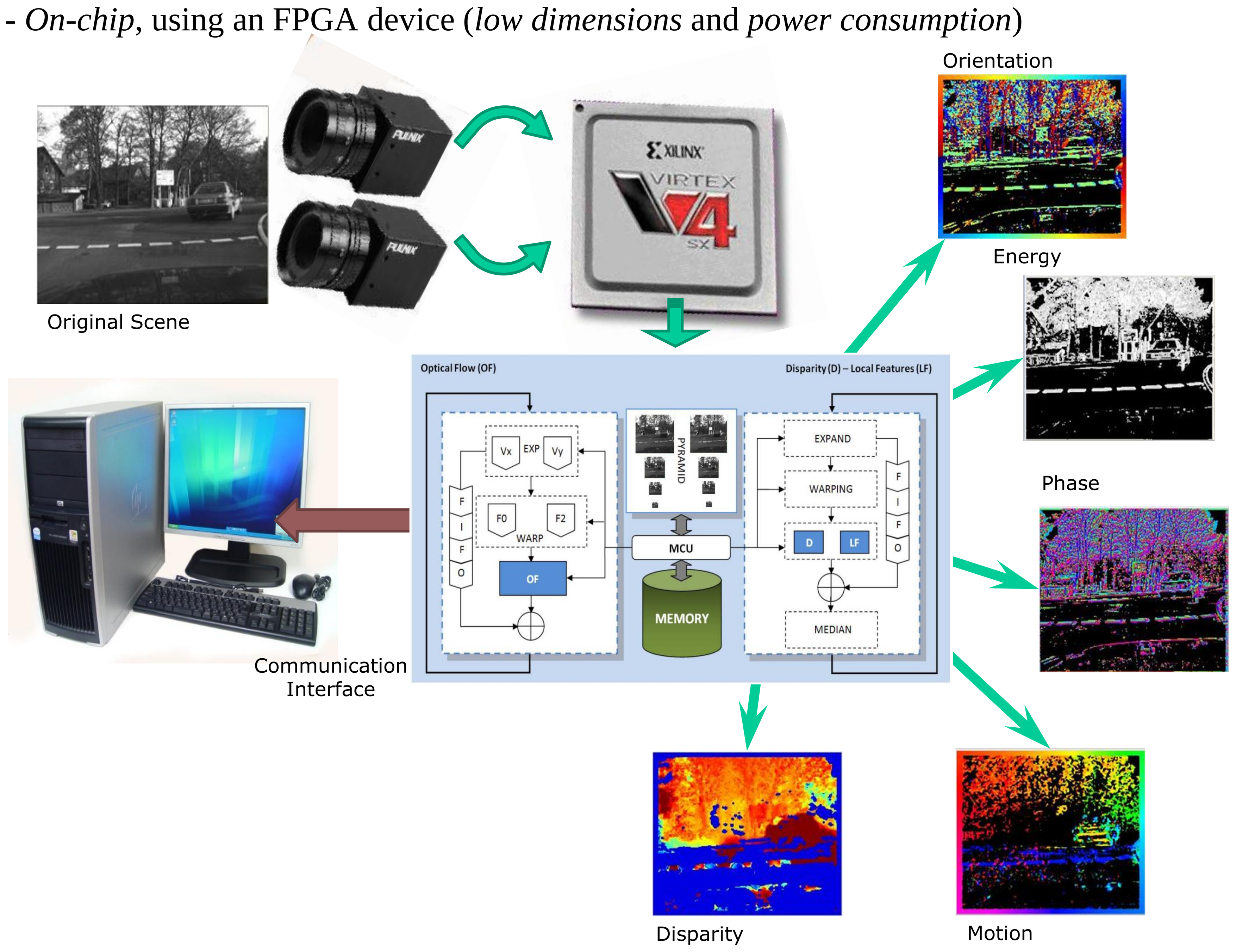



**Fig1.** Current architecture system. The system obtains in real-time the optical flow, disparity and local contrast descriptors processing the input from a stereo pair of calibrated cameras. More about the interface and details could be found in [7].

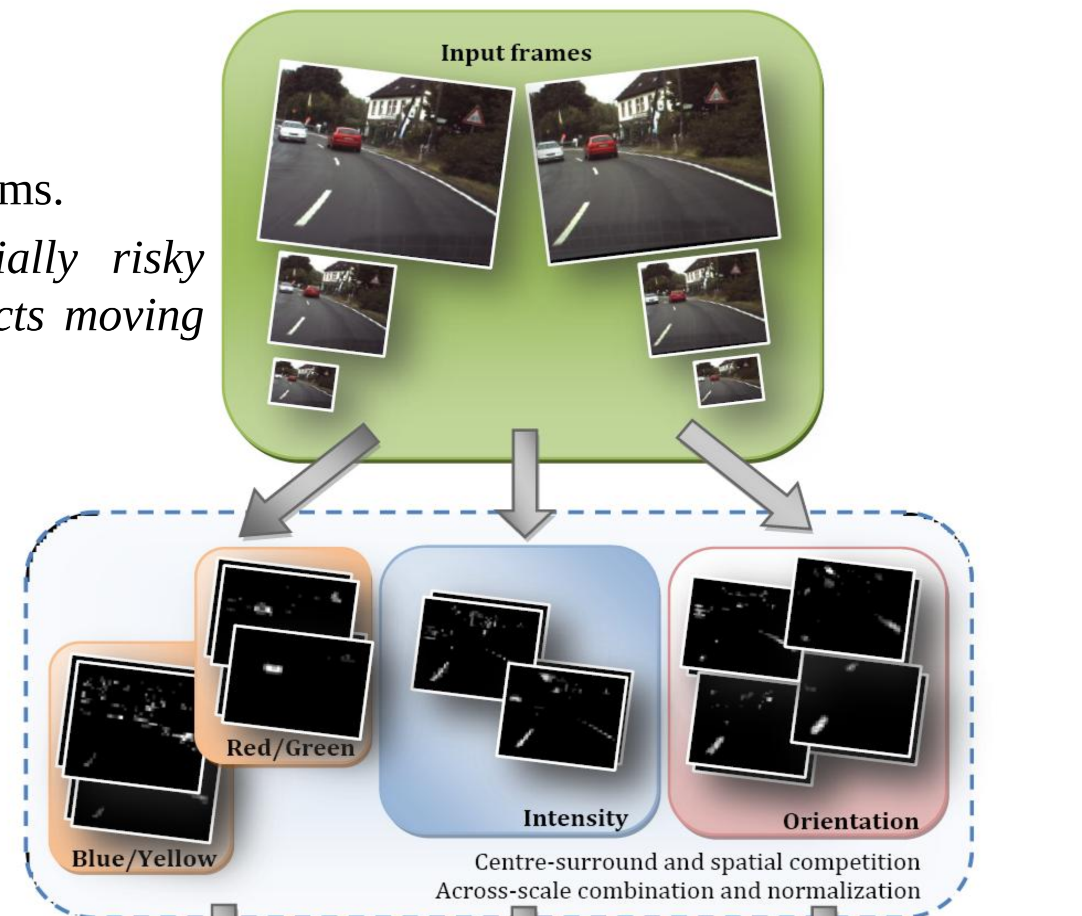
The proposed system:

- Obtains the *saliency maps biased* with the motion and the disparity.
- The *hierarchical motion* and *disparity* computation could be easily integrated to the scheme.
- *On-chip* and *real-time* performances
- Encapsulated in an FPGA (*co-processing* board or *stand-alone* platform)
- The *most complex* one developed until now (with these characteristics).
- Transforms our *low-level vision machine* system towards a *scene understanding engine*.

**Example**: Driver assistance systems.

*Focus the attention to potentially risky objects: weight the nearest objects moving towards ours*

The new system fixes the attention focus on the red car instead on the road, the signs or the other cars. The result can be checked comparing the resultant saliency map.



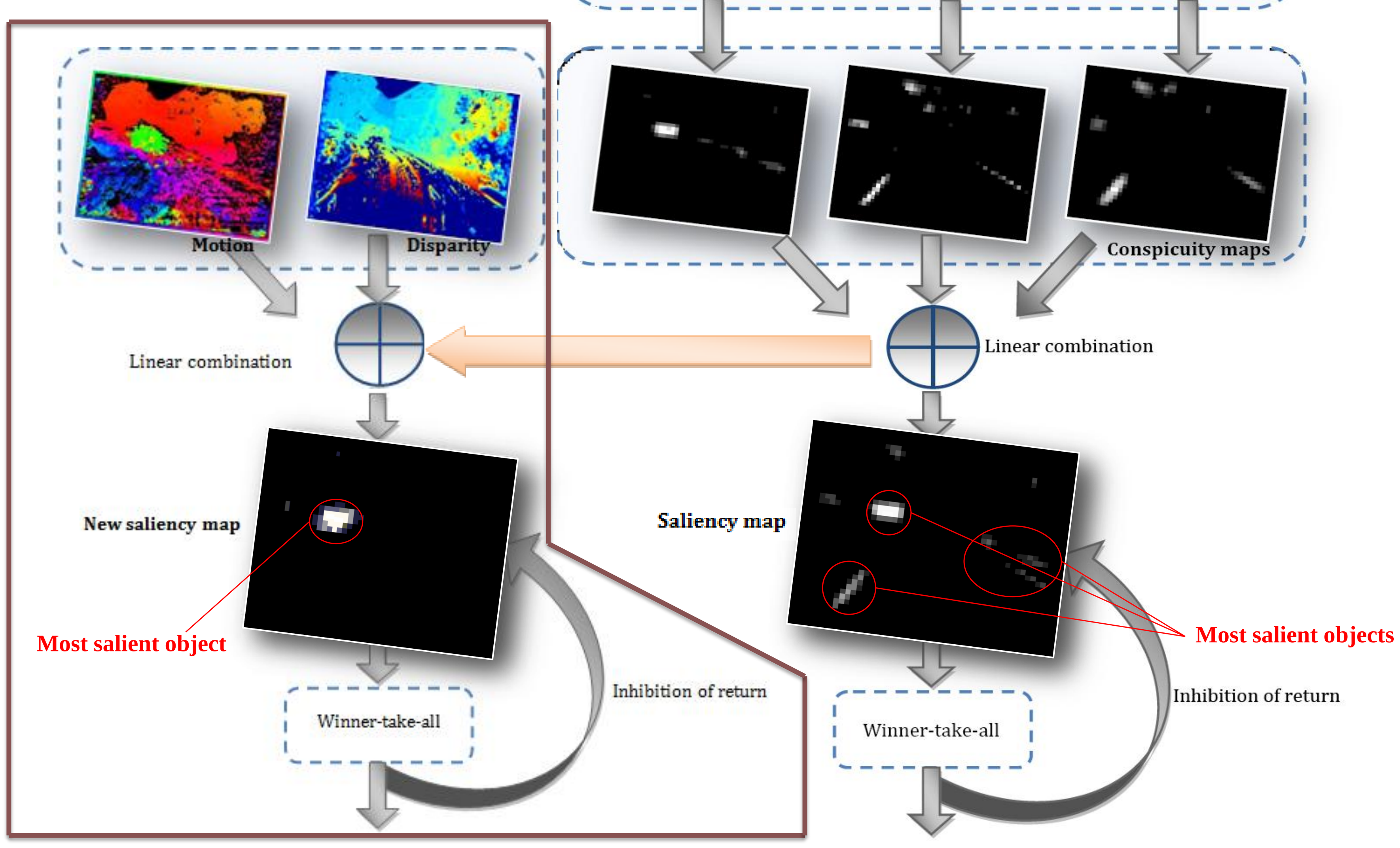




**Fig2.** New attention on-chip architecture. The right side shows the Itti & Koch model, the left one the new scheme. It is applied to a driving sequence example. The resultant saliency maps show the differences between the schemes. For the new one, the salient object is the red car (biasing the conspicuity maps with the movement of the car and the distance).

## References

[1] L., Itti, C., Koch, Computational Modelling of Visual Attention in *Nature Reviews Neuroscience*, 2001.

[2] D., Walther, C., Koch, Modeling attention to salient proto-objects in *Neural Networks*, 2006.

[3] L., Itti, C., Koch, E., Niebur, A Model of Saliency-Based Visual Attention for Rapid Scene Analysis in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998.

[4] J., Diaz, E., Ros, R., Agis, J. L., Bernier, Superpipelined High-Performance Optical-Flow Computation Architecture in *Comp. Vision and Image Understanding*, 2008.

[5] M., Tomasi, F., Barranco, M., Vanegas, J., Diaz, E., Ros, High performance optical flow architecture based on a multiscale and multi-orientation phase-based model *submitted to IEEE Trans. on Circuits and Systems for Video Technology*, 2009.

[6] M., Tomasi, M., Vanegas, F., Barranco, J., Diaz, E., Ros, A novel architecture for a massively parallel low level vision processing engine on chip in *IEEE Int. Symposium on Industrial Electronics*, 2010.

[7] Open rtVision project homepage, *http://code.google.com/p/open-rtvision/*

**Acknowledgment** This work has been supported by the Spanish project DINAM- VISION (DPI2007-61683), the regional projects (P06-TIC02007, TIC-3873) and the Programme Alβan, the EU Programme of High Level Scholarships for Latin America, No.(E06D101749CO).