Low-Middle level vision-on-chip

Javier Díaz., Department of Computer Architecture and Technology. University of Granada. Spain.

Bio-inspired Vision. *Engineering processing architectures designed for tasks that biological systems solve with impressive ease can benefit considerably by mimicking computing strategies developed by nature. However, the adaptation of such techniques is not straightforward, since the physical principles upon which biological tissues are based are very different from those characteristically used in electronic technology. Nevertheless, an "opportunistic attitude" which takes the key-functional principles that contribute to the outstanding performance of biological systems and also uses technology-motivated computing techniques to adapt those computing primitives must be of considerable interest.*

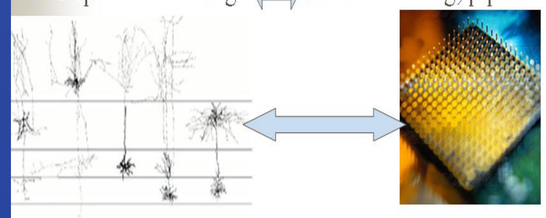
We use these techniques to develop powerful vision sensors for a multiscale, phase based, local features (phase, energy and orientation), motion and stereo-vision processing in real-time. Moreover, we address a simple condensation scheme capable to build sparse multimodal image descriptors with relevant information for middle-high level vision layers. The final system is based on FPGA technology which is suitable to be used on embedded systems, necessary for a large number of robotics and industrial applications, allowing exploration and interaction tasks in real-time attempting to close the perception-action loop.

From biological models to hardware systems

- Biological systems
  - High parallelism (3-D)
  - Slow processing
- Current Hardware devices
  - Low parallelism (2-D)
  - Fast processing

↓ ↓

- Population Coding ⟺ Time slicing, pipelining



Multiscale optical flow architecture

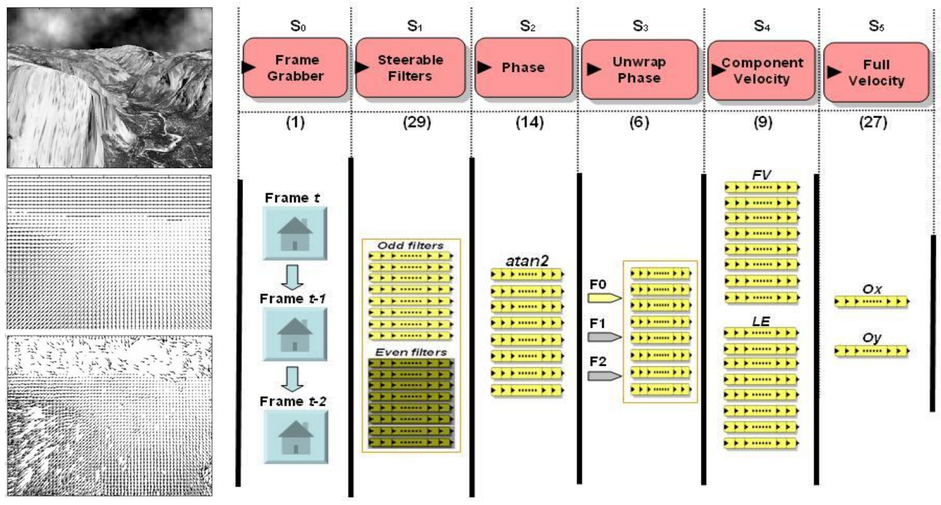
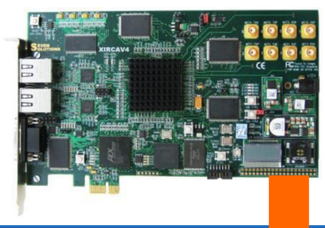
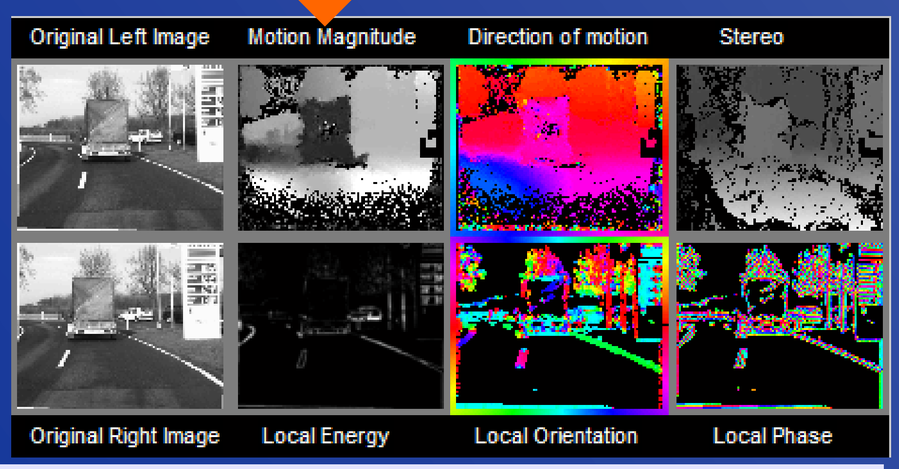


Image Resolution	Frame Rate (f/s)	
	sharing	no sharing
512 x 512	35,85	40,97
640 x 480	30,59	34,96
800 x 600	19,58	22,37
1024 x 1024	8,96	10,24

Processing platform

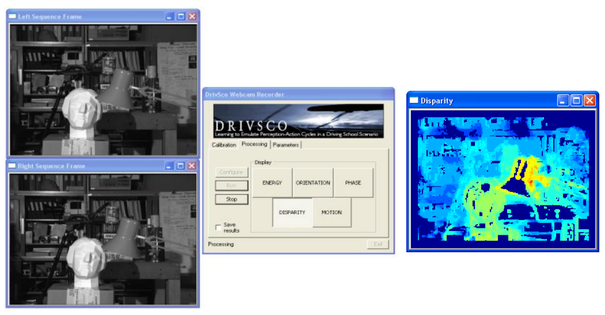


XircaV4 platform:
- Xilinx Virtex 4 FPGA xc4vfx100
- 4 ZBT memory banks and DDR memory
- PCI express
- Gigabit Ethernet



Multiscale stereo architecture

Image Resolution	Frame Rate (f/s)		
	3 banks/sharing	3 banks/ no sharing	4 banks
512 x 512	34,74	45,84	67,39
640 x 480	29,64	39,12	57,51
800 x 600	18,97	25,04	36,80
1024 x 1024	8,68	11,46	16,84



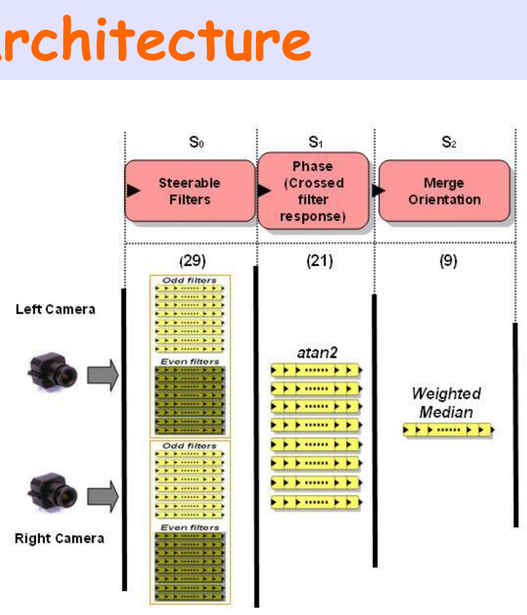


Image features condensation

On chip processing
Pixelwise
Local
Regular

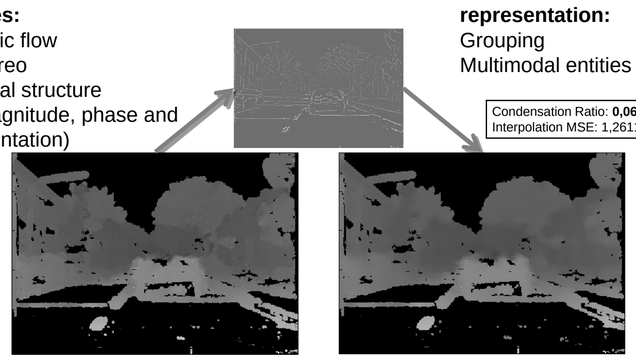
Condensation ⟶

⟵ **De-condensation**

Higher level modules
Based on multimodal entities
Corrective correlations
Multimodal structures

Low level vision cues:
Optic flow
Stereo
Local structure (magnitude, phase and orientation)

Middle level vision representation:
Grouping
Multimodal entities



- Compression motivated by bandwidth constraints
- Compatible with higher level stages
- Capable of embedding enhancing schemes (for instance from IMOs)

Local phase, energy and orientation architecture

Image Resolution	Frame Rate (f/s)
512 x 512	129 (PCIe limited)
640 x 480	129 (PCIe limited)
800 x 600	104,16
1024 x 1024	47,68