FPGA-based Real-Time Disparity Computation and Object Location


Pedro Miguel Santos[1]

João Canas Ferreira[2]

University of Porto[1, 2]   INESC PORTO[2]

Brief Description

The system is capable of locating the objects, from a scene, that are close to the cameras, using the principles of stereoscopy.

General concept inspired by *FingerMouse* [1].

Principles of Operation:

1. Two side-by-side cameras capture the scene
2. Select a specific area in one of the images
3. Search for it in the other image
4. Measure the displacement of the area
5. Segment the image into objects, by clustering groups of pixels with similar displacement

Our implementation:

- Implemented on a Xilinx Virtex 2-Pro, using Verilog
- Works up to 40 fps, for images of 640x480 pixels
- Identifies disparities up to 135 pixels
- Uses two metrics for area matching – SAD and CENSUS –, and operates over 3x3-pixels windows
- Includes an innovative coarsening stage

System Stages

For this pair of input images, the goal of the system is to isolate and locate the hand.

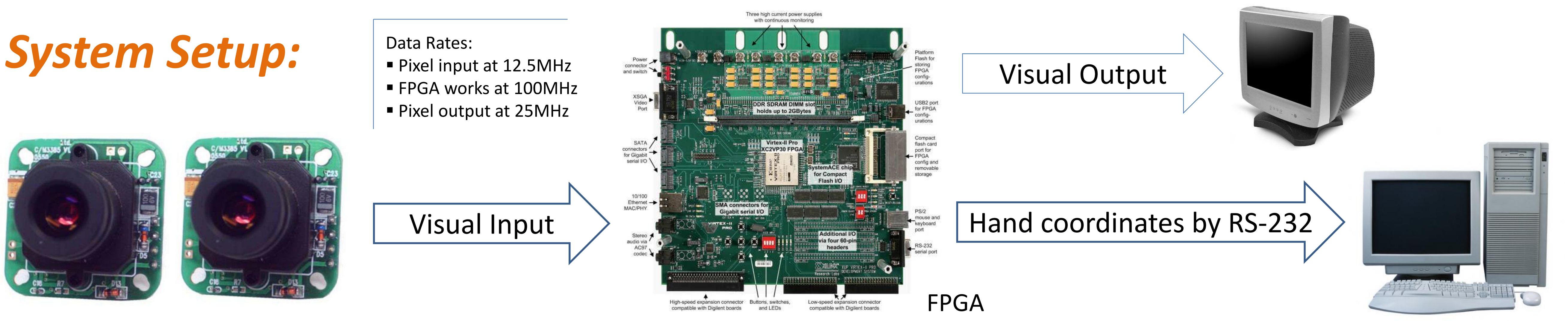
Disparity Computation

Binarization and Merging

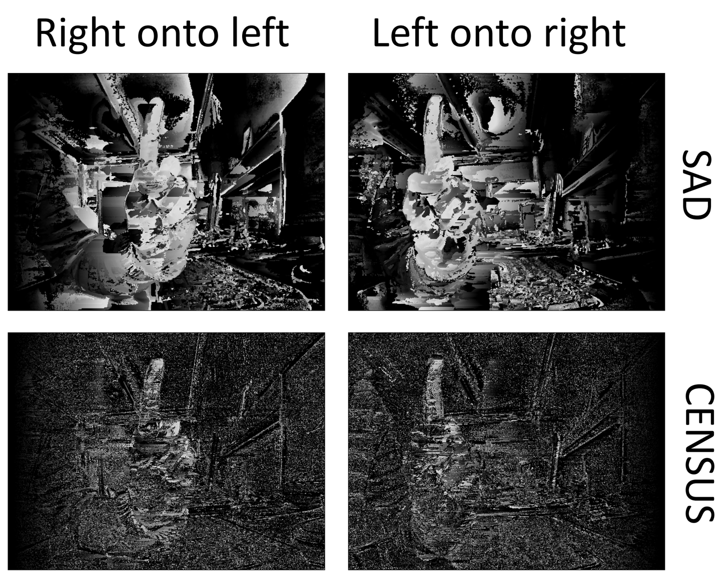
Consistency Check

Coarsening

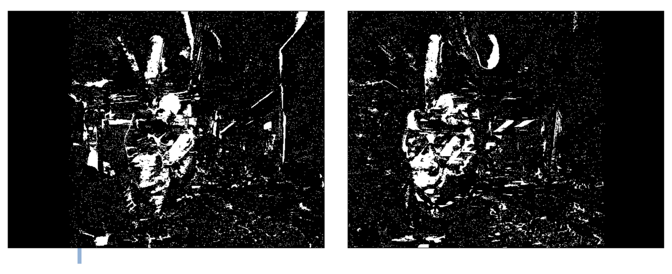
System Setup:



Stage Outcome: / *Brief Description:*



- For each pixel of an input image, a matrix of neighbouring pixels is selected, and searched for in the other image.
- The position of the best matching area yields the pixel's *disparity* (i.e., the displacement of that area w.r.t. its position in the reference image).
- Pixels with *high disparities* are *close* to the cameras. Using the disparity values as brightness values in a image results in a *disparity map* (shown in the left).
- Four disparity maps are computed: for both metrics (SAD and CENSUS), right image pixels are searched for in the left image and vice-versa.



- Pixels are classified as "foreground" or "background" pixels.
- Pixels whose disparity is higher than a threshold, by any of the metrics, are given a 1; all others are given a 0.
- The output are two bitmaps; each merges the results of SAD and CENSUS.

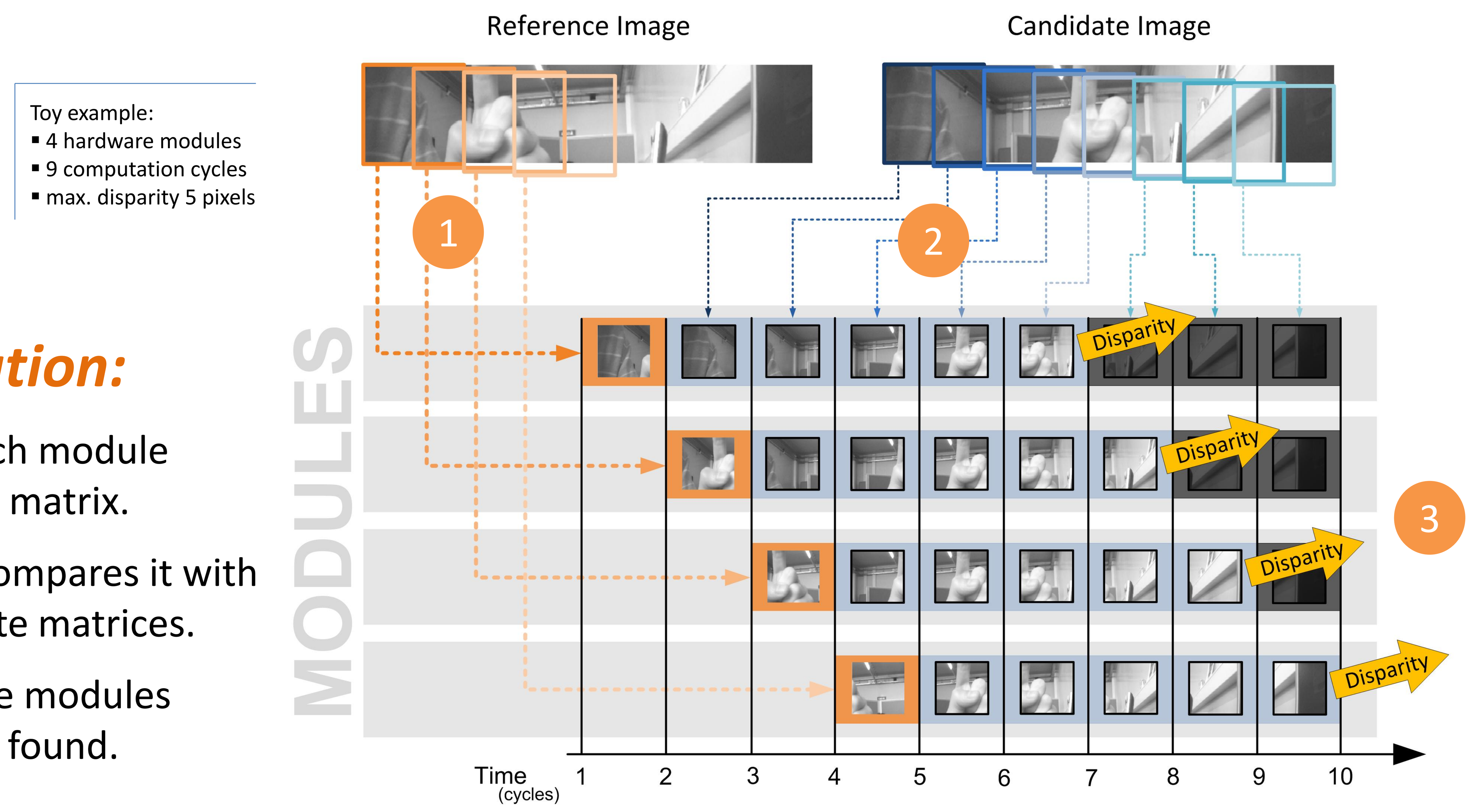
- The foreground pixels, of one of the bitmaps, are confirmed as such.
- For each foreground pixel, its disparity is used to check if, in the other bitmap, at the predicted location, a foreground pixel is also found. If so, the pixel is kept.
- Only one bitmap is checked against the other, resulting in a single output.

- To improve center of gravity computation, the previous output is coarsened.
- The bitmap is broken down into squares of 20x20-pixels. Squares with an high density of foreground pixels are assigned a 1, while low density areas get a 0.
- Using this, foreground areas are intensified while spurious areas are discarded.

[x=165, y= 548]   The center of gravity of the foreground pixels is computed, and the coordinates are sent to the computer.

Disparity Computation Stage

To match areas of one image in the other image in real-time is the most important and demanding stage of the system.



Hardware Architecture:

- An original architecture was developed
- 20 dedicated hardware modules are used
- Each handles one reference matrix
- All modules can be given the same stream of candidate matrices simultaneously

Hardware Operation:

1. In the first cycle, each module receives a reference matrix.
2. In the next cycles, compares it with a stream of candidate matrices.
3. In the last cycles, the modules output the disparity found.

Universidade do Porto FEUP Faculdade de Engenharia · INESCPORTO · FCT Fundação para a Ciência e a Tecnologia

Contacts:

1. pedro.miguel.santos@fe.up.pt
2. jcf@fe.up.pt

References:

[1] Hebb, J.; Koch, T.; Kuonen, S.: "*VLSI Implementation of the FingerMouse Algorithm*". Master Thesis, ETH Zürich, Winter Term 2004/2005.