Demo: An Embedded Vision System for High Frame Rate Visual Servoing

Zhenyu Ye, Yifan He, Roel Pieters, Bart Mesman, Henk Corporaal, Pieter Jonker
Eindhoven University of Technology, The Netherlands
Email: {z.ye, y.he, r.s.pieters, b.mesman, h.corporaal, p.p.jonker}@tue.nl

***Abstract*—The frame rate of commercial off-the-shelf industrial cameras is breaking the threshold of 1000 frames-per-second, the sample rate required in high performance motion control systems. On the one hand, it enables computer vision as a cost-effective feedback source; On the other hand, it imposes multiple challenges on the vision processing system. The authors have designed and implemented an FPGA-based embedded vision system in support of high frame rate visual servoing applications. The vision system will be demonstrated together with a mechanical system for vision based inkjet printing. This demonstration shows that, with off-the-shelf components, a robust, hard real-time, low delay embedded vision system is feasible for industrial applications. The research aspect of the experiment has been published in previous papers of the authors. This demonstration paper emphasises on the practical issues for the implementation of such system, and the lessons learned from this practice.**

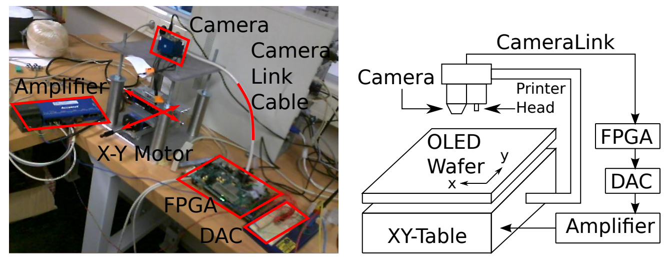
I. Introduction

Recent advances in image sensors and their interfaces enable high frame rate direct visual feedback as an attractive option for high performance motion control. However, it imposes multiple challenges on the vision processing system used in industrial applications. First, the vision processing system, including its interface to the camera, should be time-predictable. Second, the vision processing, besides being high frame rate, should have little delay. Third, the system should be cost-effective, e.g., not using costly components if they cannot improve the control performance at a large scale. While previous works have reported the implementations of 1000 frames-per-second (fps) visual servoing systems, the challenges mentioned above are not fully addressed. An earlier paper of the authors reported a case study to address these challenges [1]. This paper will briefly explain the experiment to be performed in the demonstration, followed by discussions on practical issues on the implementation of the system.

II. Visual Servoing for Inkjet Printing

The demonstration to be performed is vision-based inkjet printing for OLED display manufacturing. The experimental setup and its system diagram are shown in Fig 1. As a replication of an industrial inkjet printing machine for OLED display manufacturing, the setup has a static camera fixed at the top. The OLED wafer is moving with the X-Y table. During the manufacturing process, a static printer head, next to the camera, shoots chemical droplets at the pixel centers of the moving OLED wafer.

The vision processing pipeline, as shown Fig 2, detects the pixel centers of the OLED wafer while it is moving.



(a) (b)

Fig. 1. The (a) experimental setup and (b) system diagram.

The pipeline takes as an input a region of interest (ROI) of the OLED wafer, and outputs the coordinates of the pixel centers. The vision pipeline contains a few major algorithms. The Otsu optimum threshold algorithm [2] is used to handle variations in the lighting conditions. The multiple stage erosion is used to reduce the noise after binarization. The segmented OLED structures are reduced to two vectors, by horizontal and vertical projections, to enable a quick search of pixel centers.

III. FPGA-Based Embedded Vision System

To enable low delay vision processing, a dedicated vision architecture is designed and implemented on an FPGA, as shown in Fig 3. The timing breakdown of this vision system, processing one frame, is shown in Fig 4. The vision architecture, while it is optimized for resource utilization, consumes less than 15% of the resources on a Virtex II XC2VP30 FPGA, as shown in Fig 5. There is a large amount of data level parallelism in this pipeline, as analysed in another paper of the authors [3]. Therefore, if a higher performance of the vision processing is required, more resources can be utilized. Despite using a vision architecture not optimized for performance, the time of vision processing is not the limiting factor to obtain a higher frame rate and a lower delay, as shown in Fig 6. This is unexpected by the authors, and only discovered during the implementation of the system.

IV. Practical Issues

Multiple issues arose during the implementation of the embedded vision system. These practical issues are rarely reported in academic publications, but the authors think they are worth mentioning in this demonstration paper.

The first issue is the time-predictability of the vision processing system, which is crucial in high frame rate (above 1000

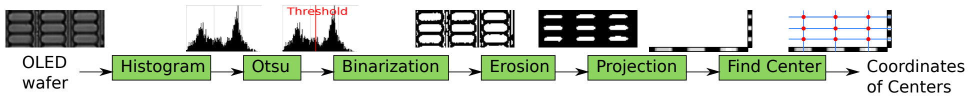


Fig. 2. Vision pipeline detecting the pixel ceters of the OLED wafer.

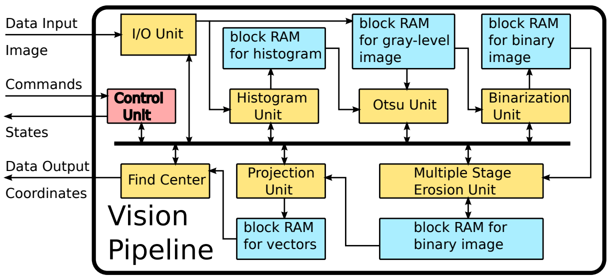


Fig. 3. Hardware architecture implemented on an FPGA.

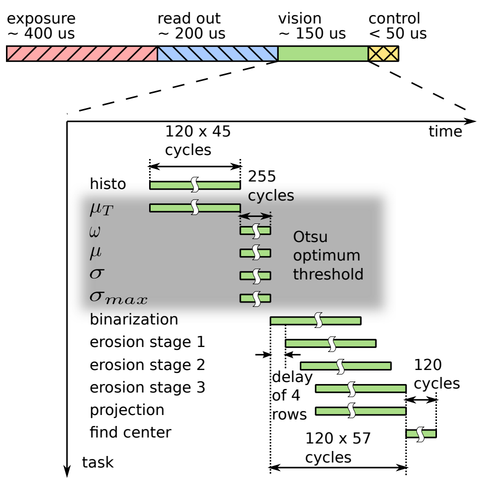


Fig. 4. The timing breakdown of the processing of one frame.

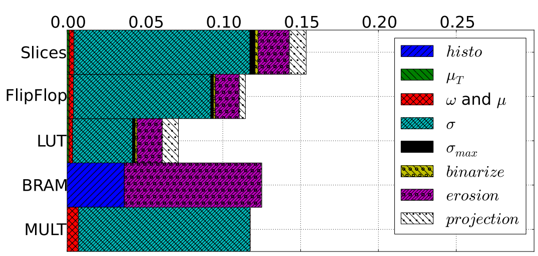


Fig. 5. Resource utilization correponding to the implementation in Fig 4.

fps) visual servoing. On a general-purposed processor, the jitter of the system can easily be larger than what is tolerable by the control law. The worst case execution time of general-purpose processors, if possible to be estimated, could easily be the bottleneck of the system.

The second issue is the choice of cameras. There are image sensors capable of functioning at one billion frames-per-second, but most of them are in-situ image sensors that buffer the frames on-chip. There are image sensors capable of streaming images off-chip at one million frames-per-second, but they require customized interfaces. Most industrial applications prefer the usage of cost-effective technologies and standard interfaces. The highest frame rate that can be supported by standard interfaces, e.g., CameraLink, is at the scale of 10000 fps. The camera used in this work functions at a frame rate at such scale.

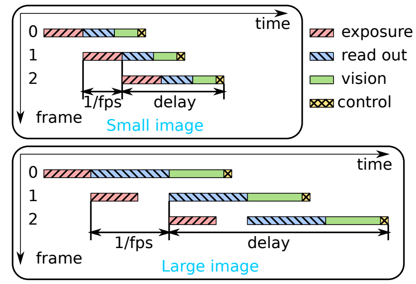


Fig. 6. Bottlenecks of frame rate (1/fps) and delay at variant image sizes. Frame rate is limited either by exposure time or image read out time.

The third issue is the camera interface. As discussed in our previous paper [1], the image read out time dominates the delay of the system. Most standard camera interfaces, e.g., GigE Vision, have high bandwidth, but introduce large delays, because the image needs to be copied multiple times through peripheral devices, e.g., a network controller. The added delay exacerbates what is already a bottleneck of the system. Moreover, the delay of these peripherals are mostly not time-predictable, either due to the protocols or to the black-box implementations of the protocols. Therefore, CameraLink turns out to be the only cost-effective industrial standard that satisfies our requirements.

The fourth issue is the lighting. The authors did not expect that the exposure time would dominate the delay of the system. The exposure time depends on the type of surface and the lighting condition. The datasheets of the image sensors cannot provide an estimation of the exposure time regarding to such variations. In our case study, even with strong lighting, the reflective glass surface of the OLED wafer requires a much longer exposure time than the standard exposure time described on the datasheet.

REFERENCES

[1] Z. Ye, Y. He, R. Pieters, B. Mesman, H. Corporaal, and P. Jonker, "Bottlenecks and tradeoffs in high frame rate visual servoing: A case study," in *IAPR Conference on Machine Vision Applications*, 2011.

[2] N. Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, no. 1, pp. 62 –66, 1979.

[3] Y. He, Z. Ye, D. She, B. Mesman, and H. Corporaal, "Feasibility analysis of ultra high frame rate visual servoing on fpga and simd processor," in *Advanced Concepts for Intelligent Vision Systems*, 2011.