

On the Accelleration of Viola-Jones Face Detection Algorithm on FPGAs

Curtis Bader, Peter Irgens, Theresa Le, Devansh Saxena, Cristinel Ababei

Department of Electrical and Computer Engineering

Marquette University, Milwaukee WI, USA

Email: {curtis.bader,peter.irgens,theresa.le,devansh.saxena,cristinel.ababei}@marquette.edu

Abstract—We present an FPGA based hardware accelerator for the popular Viola-Jones face detection algorithm. We start by first presenting a detailed discussion of the algorithm. The scope of this discussion is to understand the internal structure of the algorithm and thus identify its primary computational bottlenecks. These computational bottlenecks represent the main candidate portions for speed-up via efficient implementation and parallelization on FPGAs. Selected portions of the algorithm are then coded in VHDL. The VHDL implementation is verified on a DE2-115 board, which uses a Cyclone IV FPGA chip. Experimental results are compared against simulations performed using in-house both pure C++ as well as CUDA programmed implementations of the Viola-Jones algorithm. The FPGA accelerated version is $4\times$ faster.

Index Terms—face detection; Viola-Jones algorithm; CUDA programming; FPGA accellerator

I. INTRODUCTION

Field programmable gate arrays (FPGAs) have been gaining a lot in popularity because the gap between their performance and power consumption, compared to specialized ASICs, has been closing in recent years. FPGAs are currently used in virtually any application domain.

II. RELATED WORK

Here, we describe previous attempts to speed-up the Viola-Jones algorithm via various hardware accellerators.

III. BACKGROUND ON VIOLA-JONES FACE DETECTION ALGORITHM

In this section, we present a high level description of the Viola-Jones face detection algorithm. Such a description is necessary as it will help us better understand how to port some of its key tasks into FPGA based accellerators later on. The pseudocode description of the Viola-Jones algorithm is presented in Fig.1. In the following paragraphs we discuss several key ingredients that made this algorithm the first successful realtime face detection algorithm.

In a face detection algorithm, we must use an accurate numerical description such that it sets human faces apart from other objects in a given image. Such characteristics can be extracted with a committee algorithm called Adaboost [1]. Such a committee can be created with weak classifiers to form a strong classifier by employing a voting mechanism. The Viola-Jones algorithm [2] uses Haar-like *rectangle features* to construct classifiers. A Haar-like rectangle feature is a scalar product between the image and some Haar-like *pattern*

Algorithm: Viola-Jones Face Detection Algorithm

```

1: Input: original test image
2: Output: image with face indicators as rectangles
3: for  $i \leftarrow 1$  to number of scales in image pyramid do
4:   Downsample image to create  $image_i$ 
5:   Compute integral image,  $image_{ii}$ 
6:   for  $j \leftarrow 1$  to number of shift steps of sub-window do
7:     for  $k \leftarrow 1$  to number of stages in cascade classifier do
8:       for  $l \leftarrow 1$  to number of filters of stage  $k$  do
9:         Filter detection sub-window
10:        Accumulate filter outputs
11:      end for
12:      if accumulation fails per-stage threshold then
13:        Reject sub-window as face
14:        Break this  $k$  for loop
15:      end if
16:    end for
17:    if sub-window passed all per-stage threshold checks then
18:      Accept this window as a face
19:    end if
20:  end for
21: end for

```

Fig. 1. Pseudocode of the Viola-Jones face detection algorithm.

or template. An example of a Haar-like pattern is shown in Fig.2.a.

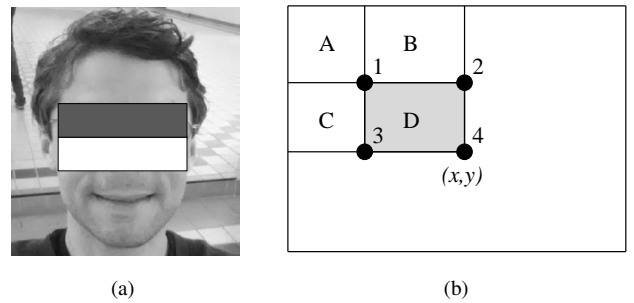


Fig. 2. (a) Haar-like rectangle feature is calculated only by using the pixels inside the Haar-like pattern (i.e., black and white rectangles). This example pattern illustrates that the area covering the eyes is usually darker than the area just above the cheeks. (b) The sum of the pixels inside rectangle D can be computed with four array references on values of the integral image: $4 + 1 - (2 + 3)$

A crucial element of the Viola-Jones algorithm is a technique to compute rectangle features very rapidly [2], [3]. This technique uses an intermediate representation for the image, the so called *integral image*. The integral image at

location (x, y) contains the sum of the pixels above and to the left of (x, y) . Instead of summing up all the pixels inside a rectangular window, this technique mirrors the use of cumulative distribution functions. Using the integral image any rectangular sum can be computed in four array references as shown in Fig.2.b.

The Viola-Jones algorithm uses so called *cascade classifiers*. A cascade classifier is constructed as a sequence of stages. At each stage a list of filters are applied to the area within the sliding sub-window. An example of such a multistage filter with 25 stages is shown in Fig.3. Each time the sliding sub-window shifts (typically pixel by pixel, but it can be more pixels at a time to further speed things up), the new region within the sliding sub-window is processed through the cascade classifier stage-by-stage. At each stage, the rectangle feature is evaluated and the weak classifier is computed. Then, a threshold check is used to see if the region is rejected as a face candidate or if it needs to continue to be processed in the next stage.

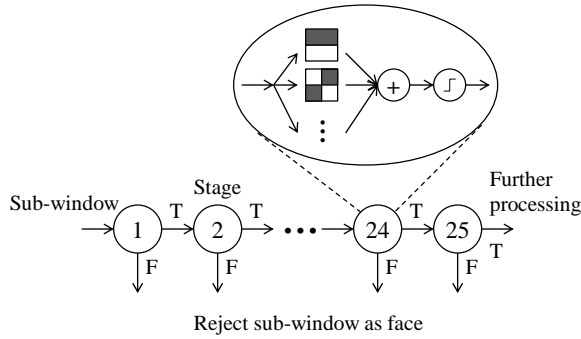


Fig. 3. Illustration of a cascade classifier with 25 stages as a decision tree, where at each node a threshold check is done to decide if the sub-window is rejected as no face (False, F) or if it is passed for further processing to the next stage (True, T), which means that the sub-windows still has chances to contain a face.

To be able to detect faces of different sizes, the algorithm works with a pyramid of scaled images (see Fig.4). This effectively allows sweeping using the same set of Haar-like patterns different scaled versions of the initial image. Thus, sliding sub-windows will sweep each of the images from the pyramid as illustrated in Fig.5.

When the outer-most for loop in the pseudocode description from Fig.1 finishes its execution, the Viola-Jones algorithm would have found and marked with rectangle indicators all faces present in the original test image as well as in the scaled versions of the image. For example, Fig.6 shows the result of running the Viola-Jones algorithm on the test image from Fig.5.

IV. COMPLEXITY ANALYSIS OF VIOLA-JONES ALGORITHM

Here, we perform a computational complexity of the Viola-Jones algorithm. We do that by using the GNU *gprof* tool on an in-house C++ implementation of the algorithm.

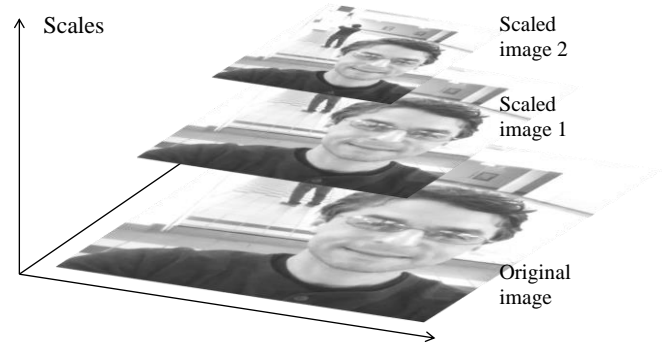


Fig. 4. Pyramid of scaled images. The bottom image is the original image. The others are scaled images obtained via downsampling of the original image using neighboring pixels.

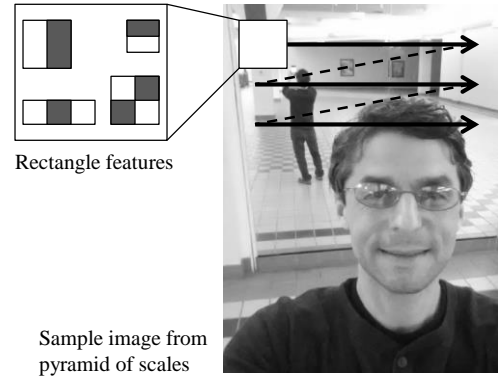


Fig. 5. Illustration of the process of sliding a sub-window. At each step during the sweeping process, the region within the sub-window has the Haar-like feature computed and processed through a cascade classifier.

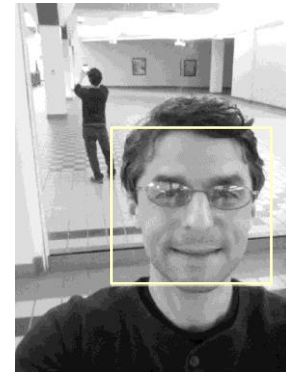


Fig. 6. The output of running the Viola-Jones algorithm shows detected faces via rectangle indicators superimposed on the original test image.

V. VHDL IMPLEMENTATION

Once portions of the VJ algorithm have been identified as the main candidates for implementation on FPGAs, we code their functionality in VHDL.

VI. EXPERIMENTAL AND SIMULATIONS RESULTS

Here, we perform report results of the comparison between the FPGA based accelerated VJ algorithm and the C++ im-

plementation running on a Linux machine.

VII. DISCUSSION

In our experiments, we found that.

VIII. CONCLUSION

We proposed an FPGA based implementation of the first face detection algorithm that was designed for real time applications. Our entire simulation framework is publicly available at [4].

REFERENCES

- [1] Y. Freund and R.E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. of Computer and System Sciences*, vol. 55, no. SS971504, pp. 119-139, 1997.
- [2] P.A. Viola and M.J. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [3] Y.-Q. Wang, "An analysis of the Viola-Jones face detection Algorithm," *Image Processing On Line (IPOL)*, ISSN 2105-1232, 2014.
- [4] Software downloads at MESS Lab, Marquette University, 2014. [Online]. Available: <http://dejazz.com/software.html>.