Real-time implementation of adaptive correlation filter tracking for 4K video stream – a demo

Marcin Kowalczyk  
AGH University of Science and Technology Krakow, Poland  
E-mail: kowalczyk@agh.edu.pl

Dominika Przewlocka  
AGH University of Science and Technology Krakow, Poland  
E-mail: dprze@agh.edu.pl

Tomasz Kryjak, Senior Member IEEE  
AGH University of Science and Technology Krakow, Poland  
E-mail: tomasz.kryjak@agh.edu.pl

I. INTRODUCTION

Object tracking is an important component of many vision systems. Example applications are tracking of: people in advanced video surveillance systems (AVSS), football, basketball, hockey players or objects like aircraft, drones, tanks in military systems. Information about the object location in consecutive frames allows also for a more complete analysis of its behaviour, e.g. human action recognition.

In this demo the implementation of a correlation filter tracker based on the MOSSE algorithm [1] in the Zynq UltraScale+ MPSoC system (Multi-Processor System on Chip), which works in real time for a video stream of $3840 \times 2160$ @ 60 fps, is presented.

II. SYSTEM OVERVIEW

The MOSSE algorithm can be divided into filter initialization, object tracking and filter update stages. It was decided to implement the initialization in the processor system due to many affine transformations. Other operations were implemented in programmable logic.

The system is designed for 4K resolution and the ZCU 104 board with Zynq UltraScale+ MPSoC XCZU7EV-2FFVC1156 device from Xilinx. The input is a 4K signal from a camera or computer’s graphics card and the output a 4K LCD monitor. The current version of the system supports a $64 \times 64$ tracking window. However, adapting the solution to power of 2 window sizes is rather straightforward.

The hardware module can be divided into two main parts. The first one is used to calculate the new location of the object based on correlation and the second to update the filter coefficients.

**Tracking:** In the beginning, it is checked whether the currently processed pixels are in the window around the object’s position in the previous frame. If so, the corresponding pixels are used to calculate the correlation.

Next, the image patch is pre-processed. Firstly, a LUT operation is applied. In this way the log function and scaling of input data is performed. Based on the number of pixels that have been pre-processed, the address for ROM containing the Hanning window coefficients is calculated. Then the pixels after the LUT correction are multiplied by the appropriate window coefficient.

After pre-processing, the data is sent to modules responsible for calculating a two dimensional Fourier transform and then multiplied element-wise with the filter coefficients. This way, the correlation of the window with the filter in the Fourier domain is obtained. The multiplication result is then sent to modules calculating the inverse Fourier transform. Next, the location of the correlation maximum is determined. The computed coordinates are sent to the part responsible for updating the filter coefficients.

**Update:** In the beginning, a window around new location of the object is read from the memory. Then its content is preprocessed and the Fourier transform is calculated. Based on the data stream from the described part, an address is generated for the ROM memory containing Fourier transform coefficients of a Gaussian distribution. It is used for calculation of the new coefficients of the filter, which are then saved to the BRAM memory. Figure 1 presents sample tracking results of the described system.

III. SUMMARY

The demo presents a hardware-software implementation of the MOSSE algorithm, which is an example of correlation filter based tracker. The presented solution is the basis for further research on the hardware implementation of these very effective tracking algorithms, like adding additional features (HOG or calculated by a deep convolutional neural network) or adding multi-scale tracking.

REFERENCES


The work presented in this paper was supported by the National Science Center project no. 2016/23/D/ST6/01389