VSLAM using System on Chip Based Vision

Lars Asplund and Jörgen Lidholm

Mälardalen University

Abstract. A system utilizing reconfigurable hardware of 8 million gates and one to four CMOS cameras can to be used as a platform for a cooperation project for a video system for VSLAM.

1 Introduction

Many algorithms used in digital signal processing for image analysis require large computational resources. The most commonly approach is to use a DSP (Digital Signal Processor), such as Texas Instruments TMS320C6201 and C6701, Philips TriMedia TM1100 and Analog Devices Sharc ADSP 21160M. These processors are variations of SIMD architectures, and they contain several processing units. By the internal pipelining and by using the processing units in an optimal way quite high throughputs can be achieved. Standard PCs are naturally used for image analysis, but for real-time applications these systems has in the past not been powerful enough.

Reconfigurable hardware has most often been regarded as a means for speeding up standard processors or DSP's. Operating systems implemented in an FPGA as an accelerator can in a Real-Time system guarantee that the system is fully predictable [13].

There have also been attempts to use reconfigurable hardware for Image Processing, [12] and [1].

The current proposal aims at building a vision system (a vision system using image analysis in configurable hardware – FPGA, Field Programmable Gate Array) as a basis for a VSLAM (Video SLAM) for new robot designs. The system will analyze the output from one up to four digital cameras in order to find objects and by using two or more a 3D recognition is possible.

The design is based on an estimated throughput of 15-30Hz.

2 Overview of the current camera system

The FPGA used in this work is equivalent to one million gates and is mounted on a six layer PCB. On the board there is also 64 MB RAM and 32 MB Flash EPROM. The two CMOS cameras are from OmniVision Technologies Inc. type OV7610 and are directly coupled to the FPGA. The output from the camera is either RGB or YUV. In the current design only 8 wires are used for data transfer, and the format of data used is RGB. The cameras are used in non-interlaced format and the color filter pattern is the Bayer-pattern, this implies that the

green-value is new for each pixel, but the red and blue are alternating. We treat two pixels as one pixel, and thus giving full information of both Red, Green and Blue. The camera chip also has an I^2C bus. This serial bus uses two wires and it can be used to control the camera. Typical controls are gain and white balance.

The internal hardware design (further on referred to as the program) of the FPGA is stored in the Flash–memory in a so called bit–file. A CPLD (Complex Programmable Logic Device) handles the loading of the program into the FPGA. There is a selector–switch with which one of eight programs can be selected for loading. This selection can also be performed by the micro controller (AVR ATmega16) mounted on the PCB for the disc. One of the eight programs is a program for loading the Flash. Thus, each camera–disc can have seven different programs stored in the Flash, thus seven different algorithms can be selected during run-time. The time to load a new program is less than one second.

In this paper three different programs are described. The first program is using one camera for finding the ball and the other camera for finding white lines. The two cameras have different gain–settings for the individual colors. A stereo matching algorithm is the second program. Detected feature transitions from a row for both cameras are stored in a CAM for fast comparison and matching. The stereo matching program successfully detects and matches feature transitions and depth information, this can be obtained from two camera images in real-time. The third program implements Harris and Stephens combined corner and edge detection algorithm, and it uses both cameras. With a frame–rate of 13Hz corners are outputted from left and right cameras.

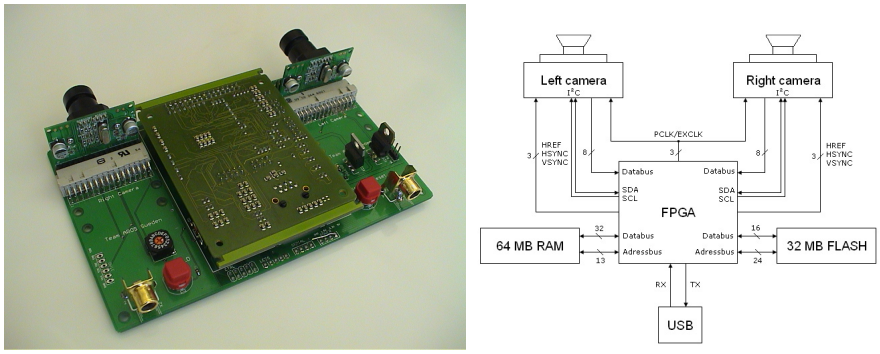

Fig. 1. PCB holding the FPGA-board and the two cameras and the corresponding block diagram

3 Future cooperation

For any kind of robot the sensor system is crucial for its observation of the environment. Of various sensors, vision is at least for humans but also for a

robot the most powerful sensor. The main way vision is implemented today is to use ordinary computers. Although a modern PC has very high performance there is always a trade–off between frame–rate and resolution, and given that the vision system should be part of a mobile robot, size is also an important factor.

The results from previous studys show that; by using an FPGA with more than 1 million gates, it is possible to achieve a frame–rate of close to real-time on a stereo–camera setup, where for instance all corners are detected in real–time.

The limitation of frame–rate is due to the number of hardware multipliers, and the clock–frequency of the FPGA. By increasing the number of multipliers the frame–rate can be turned up to the maximal 25Hz, which is then limited by the frame–rate of the cameras.

References

[1] K Benkrid, D Crookes, J Smith and A Benkrid, "High Level Programming for Real Time FPGA Based Video Processing", IEEE Symposium on Field-Programmable Custom Computing Machines, FCCM2001, April 2001.

[12] L Petit and J.D. Legat, "Hardware Techniques for the Real-Time Implementation of Image Processing Algorithms", Microelectronics Laboratory, University Catholique de Louvian, Belgium.

[13] L. Lindh, "FASTHARD - A Fast Time Deterministic Hardware Based Real-Time Kernel", IEEE press, Real-Time Workshop, Athens, January, 1992.