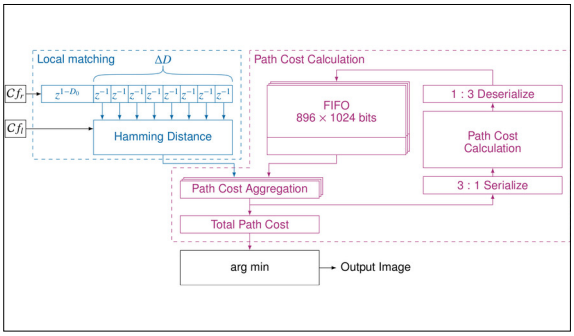
Roy
Seitz

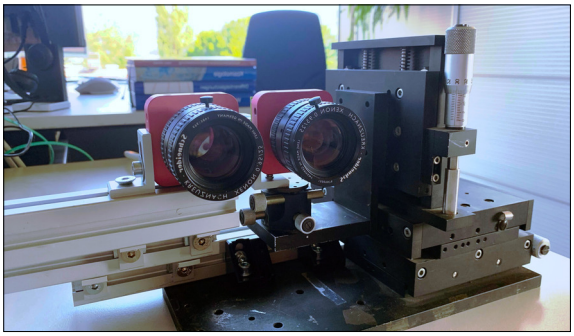
Graduate Candidate	Roy Seitz
Examiner	Prof. Dr. Paul Zbinden
Co-Examiner	Dr. Peter Schwider, Lachen SZ, SZ
Subject Area	Sensor, Actuator and Communication Systems
Project Partner	Photonfocus AG, Lachen, SZ

StereoVision

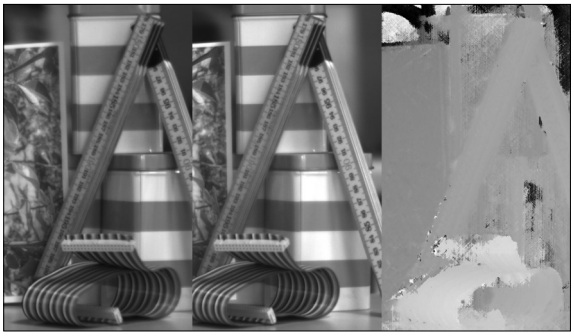
High-Speed Stream-Based Parallel Dense Disparity Map Calculation



StereoVision Core: Census transformed images are locally matched (blue) and smoothed along paths (red).
Own presentment



StereoVision Setup, featuring two cameras from Photonfocus. The FPGA is on a KC-705 evaluation board (not shown).
Own presentment



Test scene with left, right and calculated disparity image. Reflections and low texture regions prove difficult.
Own presentment

Introduction: Photonfocus is specialized in industrial camera applications and also offers tailor-made solutions. For this they rely on a constantly growing catalogue of IPs. Based on two images StereoVision allows recalculating the distance to objects due to their displacement. This is an idea on the roadmap of Photonfocus. StereoVision thus became the subject of this thesis, which consists of three main parts: a literature study for state-of-the-art systems, high-level modelling and evaluation of suitable algorithms, and implementation of the most suitable algorithm in an FPGA. Requirements are a stream-based calculation capable of processing multiple pixels per clock, preferably without external memory.

Approach / Technology: The used algorithms can be classified into three groups: local, global and semi-global. Local algorithms are lightweight and efficient but the achieved accuracy is rather poor. Global algorithms are precise and robust but require a full frame to be memorized, which is therefore unsuitable for an FPGA without external memory. Semi-global Matching (SGM) enhances local algorithms with a smoothness constraint along paths through the resulting disparity image, allowing both a stream-based implementation and high quality results.
The algorithms evaluated during the literature study were implemented in Matlab and their accuracy was compared. During this evaluation, SGM clearly outperformed all other choices. Standard SGM is, however, not capable of processing multiple pixels per clock. Thus the algorithm had to be slightly adapted prior to implementing it fully stream-based, with a modular approach for maximum flexibility.

Result: Photonfocus built a StereoVision platform around a Kintex-7 KC705 evaluation board, featuring two of their 2MP cameras to verify the StereoVision IP core. Configured for 128 displacement levels, 2 pixels per clock and 2048 pixel wide images, the core requires 166 block RAMs, 38.3 kLUTs and 68.5 kFlops. Clocked with only 100 MHz, this core is already capable of processing 200 MP/s. This equates to 651 fps at VGA (640 x 480 px), 254 fps at XGA (1024 x 768 px), or even 96 fps at full-HD (1920 x 1080 px) resolution.
More parallel pixels and higher clock speeds enable even higher throughputs, allowing fast and precise object recognition. Even fast moving objects can now be tracked visually, smoothing the way for a new range of applications.