Image Warping for Streaming Video Applications

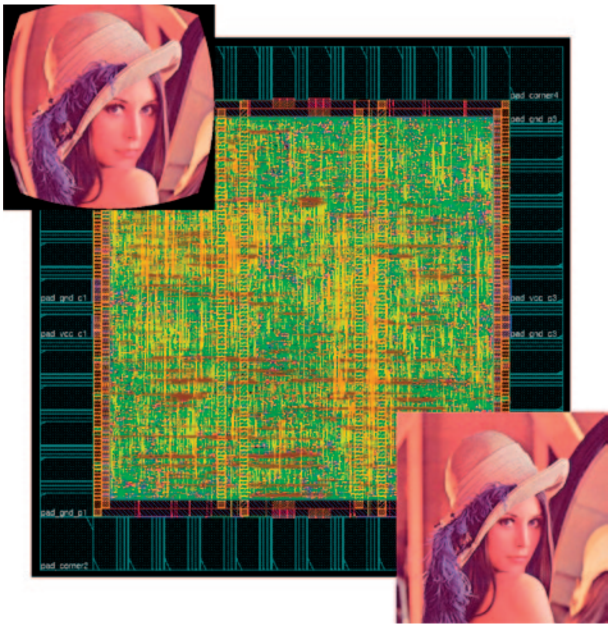
Personnel: Martin Polak, Dany Sünnen
Assistants: Pierre Greisen, Frank Gürkaynak, Simon Heinzle

Thesis: Semester Project

Partners: Disney

Real-world camera systems suffer from different non-idealities that corrupt the resulting video quality. In the current project we consider geometric errors, such as lens distortions, camera-internal parameter errors and relative positioning variations of several cameras. An image warping ASIC has been designed to digitally compensate geometric effects in real-time. The warping architecture uses configurable projective and non-linear transformations to determine the new location of each pixel. The main application is rectification of stereo video streams, but the architecture can be used for any kind of geometric image transformations.

The projective transformation requires multiplications with a 3-by-3 matrix, followed by two divisions. For non-linear lens correction, the Brown model is applied. The challenge is to reduce the very large fixed-point widths resulting from higher order terms in the correction formula (coordinate norm to the power of 6). The transformed coordinates serve as address to fetch the transformed pixel value. In general, the coordinates are not integer values which is why a bilinear interpolation of the surrounding pixels is applied. The transformation is a backward mapping: using the inverse mapping, we transform the coordinates in the target image to obtain the corresponding location in the source image. The final design runs at 125 MHz and covers a coordinate space of 2048 x 2048 pixels.



Layout of the warping VLSI design, sent for fabrication in 180 nm CMOS. The corner overlays show an example of a radially distorted and corrected image, respectively.

Real-Time Stereoscopic Depth Estimation

Personnel: Dinu Kaufmann, Simon Wright
Assistants: Pierre Greisen, Stefan Zwicky, Simon Heinzle

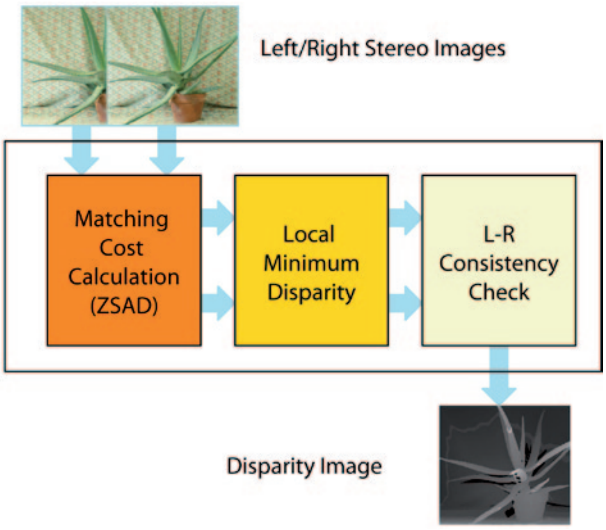
Thesis: Semester Project

Partners: Disney

Stereo vision systems work much like the human visual system: placing two cameras next to each other allows to obtain depth information if the two streams are properly combined. The depth value of a specific object is directly related to its disparity, which corresponds to the relative difference (in pixels) of the same object in the two viewpoints (cameras).

The goal of this project is an HDL implementation of a disparity estimation algorithm targeted for running on an FPGA. The employed algorithm, a local window matching algorithm, is currently default for real-time applications. A cost is calculated by matching a reference window in one image against a scan area in the other image. The minimal cost value provides a disparity value for the corresponding pixel. A subsequent consistency check, which combines left-right and right-left results, improves matching quality.

In this project, we opted for a cascaded window-based depth matching system, where we start by matching on a down-sampled image, and then refine the results in higher resolution. The results from the lower resolution thereby serve as offset to the higher resolution step. This allows for higher matching performance and down-sampling and additionally adds robustness against colorimetric and geometric distortions. The number of stages and the matching window sizes are configurable. The resulting design attains real-time performance for a stereo full-HD stream and a disparity range of 10 % of the image width.



Depth estimation overview: a stereo sequence is fed into a matching cost calculation block; for each pixel, the minimal cost value gives the disparity value, which is refined by a consistency check.