

ADOSE DELIVERABLE D6.7; PUBLIC SUMMARY
‘DEMONSTRATOR OF SILICON RETINA STEREO SENSOR’

*Issued by: AIT Austrian Institute of Technology
April 2011*

1. TARGET OF DEVELOPMENT

Workpackage 6 of ADOSE develops a bio-inspired stereo vision sensor for time-critical decision applications using the Silicon Retina sensor technology. These types of sensors are based on bio-inspired analogue circuits and have special characteristics. They don't deliver conventional image frames at a fixed frame-rate, but utilize an efficient, asynchronous, event-based data protocol that only delivers information on variations of intensity in a scene. Using conventional stereo vision approaches are of limited benefit as required data pre-processing throttles the performance significantly. Our approach is based on time-space correlation of the events without reconstructing images. As an application, the stereo sensor is used for pre-crash warning/preparation side impact detection. The deliverable discusses the development environment, analysis of recorded data regarding automotive requirements, the embedded laboratory prototype and the final DSP prototype as well as the concept for the final platform based on a FPGA or an ASIC.

2. ANALYSIS OF TEST DATA

The recording of test data took place in urban and highway environments. The recordings of scenarios with the vehicle speed 0 were taken in a parking area. The asynchronous address-event (AE) data from the ATIS cameras were time stamped with a resolution of 1 μ s during recording. To allow the visualization of the data and comparison with video image data in the following figures, AER data were collected into

image frames. To optimally represent the features of the scene we have used two reconstruction methods for converting (asynchronous) AER data into (synchronous) image frames using a fixed and dynamic frame length. The recorded data set includes: cluttered scenes, strong changes in illumination, partially occluded objects, moving objects and observer and typically driving situations. Figure 1 shows a sample scene with moving objects and observer.

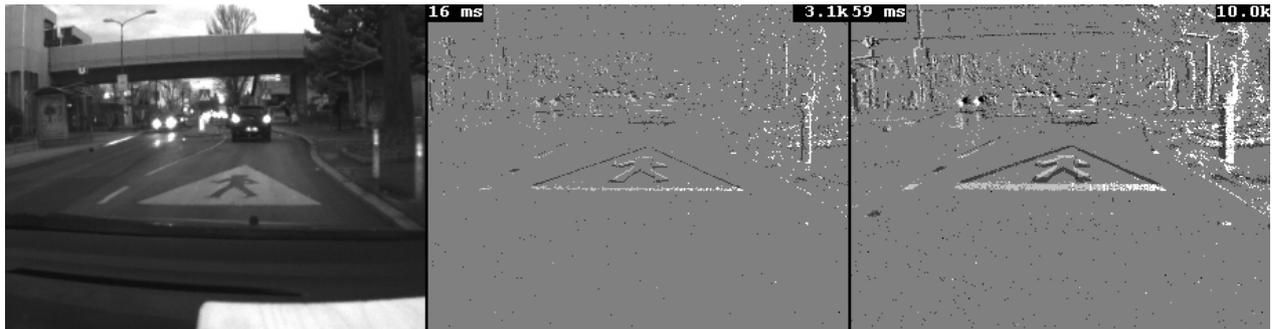


Figure 1: Scene with moving objects and observer; left: conventional imager, middle: fixed frame length of 16ms; right: dynamic frame length with sustained 10kAE.

Different scenarios have been recorded to show the capability of the Silicon Retina in different urban traffic and highway situations. The data rate of the recorded scenarios is analyzed to determine if the embedded system used has enough processing resources for handling the address-events data. Peak data rates occurred in a tunnel driving scene with $\sim 1.4\text{MAE/s}$. The cause for this data rate are the artificial AC fluorescent light sources in the tunnel with exhibit a strong 100 Hz flicker reflected on tunnel walls and the street surface. This flicker is received by the change sensitive pixels of the ATIS sensor. The data analysis was essential for algorithm development and as performance requirements.

3. DEVELOPMENT OF SILICON RETINA 3D STEREO ENGINE AND DETECTION ALGORITHM

Stereo Matching Algorithm: The stereo matching algorithm is based on the correlation of the events in time and space. Both paths are calculated separately and finally merged to a resulting disparity map. Figure 2 shows the overview of the stereo matching algorithm with all steps. The time-correlation path is based on correlating events due to their concurrency. The area-based path is based on a census transform.

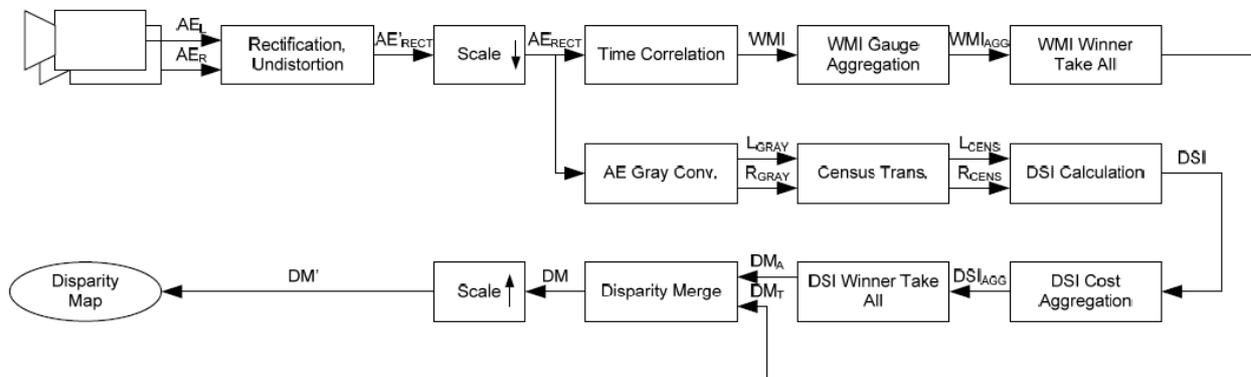


Figure 2: Overview of stereo matching approach

For verifying the detection-rate and precision, generated scenes including ground truth data were used. The time-correlation approach showed an average detection-rate of greater than 0.9 and an average precision of 0.55 - 0.75 without additional noise depending on the correlation history. The area-based approach showed a detection-rate of 0.65 – 0.85 and a precision of 0.95 without additional noise depending on the image history. Both approaches showed higher detection-rates with higher histories. The combined time-space correlation algorithm showed a detection-rate greater than 0.95 without and with noise and an average precision of 0.95 without and 0.15 with additional noise.

Detection algorithm: For tracking objects, a mean shift tracker is used for tracking point clouds in the sparse disparity map computed by the stereo matching algorithm. For detecting approaching objects, the trajectory of the detected object is computed and observed.

4. PROTOTYPES

Due to the special requirements of the Silicon Retina sensor technology, a development and verification tool-chain was established. This **Windows based tool** consists of an integrated scripting language for integration of custom add-ins e.g. for scene generation and verification. This tool is also used for data analysis and offline data processing.

The **embedded laboratory prototype** is a distributed embedded system based on a single-core and a multi-core fixed-point digital signal processor (DSP). An adapter-board is used for memory-mapping both Silicon Retina sensors to the single core DSP, that is required for data acquisition and pre-processing such as noise filtering and rectification. The main task of the data processing including stereo matching is realized on the multi-core DSP. The interface is able to required sustained 9MAE/s. For inter-chip-communication the serial high performance interface Serial Rapid IO is applied, using direct I/O transfers for data exchange and doorbells for interrupt triggering. For allowing multi-core signal processing, a static load balancing approach is realized on the single-core distributing the amount of data on the specific cores. Due to the large dimensions of the embedded laboratory prototype, an integration of the demonstrator in the vehicle is rather complicated. The prototype shows that both imagers are working and the DSP is able to receive and process data.

Figure 3 shows the hardware architecture of the **final prototype** that is an optimized and mechanically more robust version of the embedded laboratory prototype. The adapter-board and the pre-processing steps like noise filtering and event rectification and lens undistortion are realized in an FPGA design. This design is more robust against timing considerations caused by the cabling of the embedded laboratory prototype. The hardware consists of five stacked circuit boards: 2 ATIS board, motherboard, FPGA board and DSP board. The resolution and the lenses will be the same, but the baseline is smaller. Thus, the required depth resolution is lower meaning that the algorithm has to work much more precisely. The case concept is currently a first draft of a purely "functional" concept not dealing with any issues of a "fancy looking design". The major issues are condensation prevention on the windows/transparent cover and a VESA mounting interface. The BOM for a final industrialized DSP-based system is estimated at about 366 EUR for a 1000+ pieces quantity. Costs for the prototype board are 5356 EUR.

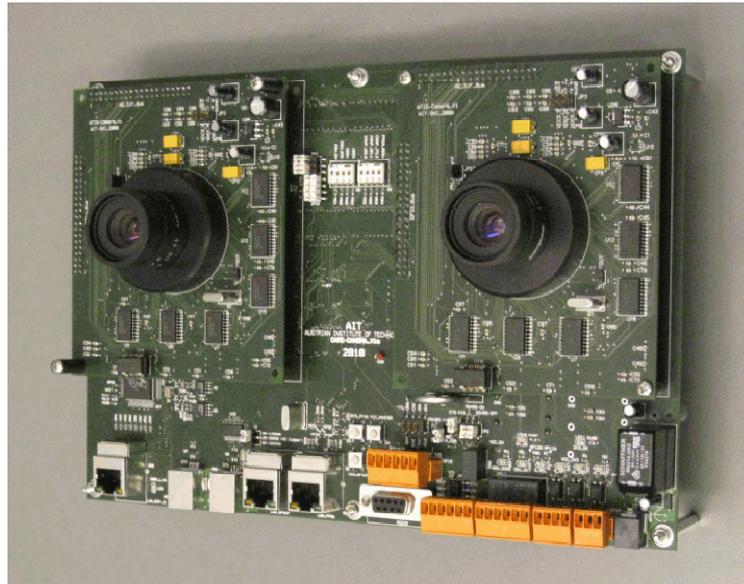


Figure 3: Hardware of final prototype consisting of five stacked circuit boards; the image shows two ATIS boards and motherboard; FPGA and DSP board are on the rear side.

Performance Intel: The algorithm is optimized for an i7-620M mobile architecture offering SSE4.2 using Microsoft Visual Studio 2005 and Intel C++ Studio XE 2011 as development environments. For the i7 architecture, we evaluated a performance boost of 3 for the compiler optimized version and 20.86 and 20.93 for the hand-optimized version.

Performance DSP: A performance boost of 45.92 was achieved from the functional behaviour to the platform optimized implementation. The key performance boost was achieved by an efficient memory management using caching and exhaustive usage of intrinsics by optimizing core functions. Further optimizations are possible by disabling caching and efficiently using DMA transfers for exchanging data between external and internal memory for data processing.

A live-demonstration of the tool-chain was demonstrated at the ITF Leipzig 2010 and the embedded laboratory prototype was demonstrated at the ADOSE progress meeting in February 2011 using a rotated wheel as moving object without ego-motion of the sensor. The actual status of the final DSP prototype is that the interfaces are being implemented and the existing optimized stereo matching and detection algorithm are being integrated.

5. FINAL SYSTEM CONCEPT

In terms of overall energy efficiency and cost-efficiency at higher quantities an ASIC solution would be preferred for a final system. Therefore, the time-correlation part of the stereo matching algorithm was realized in hardware. Verifications showed that with higher correlation history, sufficient detection-rates up to 0.9 can be realized. Synthesis of the design showed that approx. 122kLE and approx. 688kB memory are required. Due to the unknown chip consumption of the design of a specific ASIC technology, costs of a final ASIC are difficult to estimate.

Depending on the actual requirements of a final system two system size concepts are proposed. One is using Silicon Retina sensors with a high resolution and a shorter baseline of 23cm and a second with an approx. QVGA like resolution with a baseline of 45cm. The required baseline is strongly dependent on the actual application, depth resolution etc.