9

Address-Event based Stereo Vision with Bio-inspired Silicon Retina Imagers

Jürgen Kogler[1], Christoph Sulzbachner[1], Martin Humenberger[1] and Florian Eibensteiner[2]
[1]*AIT Austrian Institute of Technology*
[2]*Upper Austria University of Applied Sciences*
Austria

1. Introduction

Several industry, home, or automotive applications need 3D or at least range data of the observed environment to operate. Such applications are, e.g., driver assistance systems, home care systems, or 3D sensing and measurement for industrial production. State-of-the-art range sensors are laser range finders or laser scanners (LIDAR, light detection and ranging), time-of-flight (TOF) cameras, and ultrasonic sound sensors. All of them are embedded, which means that the sensors operate independently and have an integrated processing unit. This is advantageous because the processing power in the mentioned applications is limited and they are computationally intensive anyway. Another benefits of embedded systems are a low power consumption and a small form factor. Furthermore, embedded systems are full customizable by the developer and can be adapted to the specific application in an optimal way.

A promising alternative to the mentioned sensors is stereo vision. Classic stereo vision uses a stereo camera setup, which is built up of two cameras (stereo camera head), mounted in parallel and separated by the baseline. It captures a synchronized stereo pair consisting of the left camera's image and the right camera's image. The main challenge of stereo vision is the reconstruction of 3D information of a scene captured from two different points of view. Each visible scene point is projected on the image planes of the cameras. Pixels which represent the same scene points on different image planes correspond to each other. These correspondences can then be used to determine the three dimensional position of the projected scene point in a defined coordinate system. In more detail, the horizontal displacement, called the disparity, is inverse proportional to the scene point's depth. With this information and the camera's intrinsic parameters (principal point and focal length), the 3D position can be reconstructed. Fig. 1 shows a typical stereo camera setup. The projections of scene point P are p_l and p_r. Once the correspondences are found, the disparity is calculated with

$$d = u_2 - u_1. \tag{1}$$

Furthermore, the depth of P is determined with

$$z = \frac{b \cdot f}{d}, \tag{2}$$

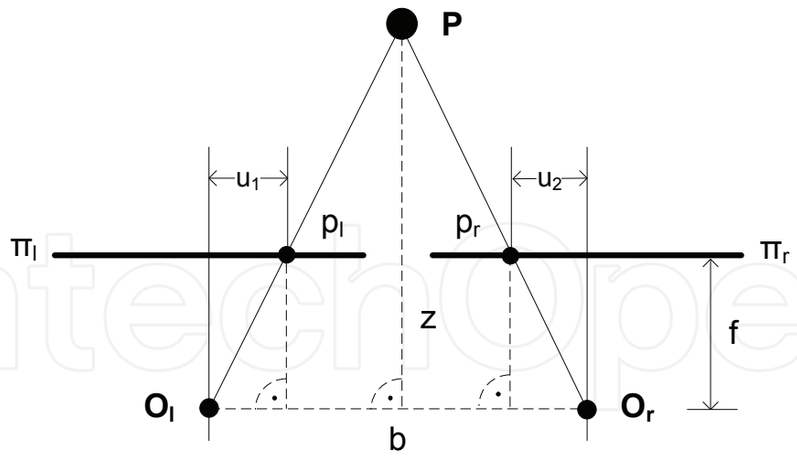


Fig. 1. Stereo vision setup; two cameras capture a scene point

where z is the distance between the camera's optical centers and the projected scene point P, b is the length of the baseline, d the disparity, and f is the focal length of the camera.
All stereo matching algorithms available for the mentioned 3D reconstruction are expecting images as captured from conventional camera sensors (Belbachir, 2010). The output of conventional cameras is organized as a matrix and copies slightly the function of the human eye. Thus, all pixels are addressed by coordinates, and the images are sent to an interface as a whole, e.g., over Cameralink. Monochrome cameras deliver grayscale images where each pixel value represents the intensity within a defined range. Color sensors additionally deliver the information of the red, green, and blue spectral range for each pixel of a camera sensor matrix.
A different approach to conventional digital cameras and stereo vision is to use bio-inspired transient senors. These sensors, called *Silicon Retina*, are developed to benefit from certain characteristics of the human eye such as reaction on movement and high dynamic range. Instead of digital images, these sensors deliver on and off events which represent the brightness changes of the captured scene. Due to that, new approaches of stereo matching are needed to exploit these sensor data because no conventional images can be used.

2. Silicon retina sensor

The silicon retina sensor differs from monochrome/color sensors in the case of chip construction and functionality. These differences of the retina imager can be compared with the principle operation of the human eye.

2.1 Sensor design

In contrast to conventional *Charge-coupled-Device* (CCD) or *Complementary Metal Oxide Semiconductor* (CMOS) imagers, which that encode irradiance of the image and produce constant amount of data at a fixed frame rate, irrespective of scene activity, the silicon retina sensor contains a pixel array of autonomous, self-signaling pixels which individually respond in real-time to relative changes in light intensity (temporal contrast) by placing their address on an asynchronously arbitrated bus. Pixels which are not stimulated by a change in illumination are not triggered; hence static scenes produce no output. In Fig. 2 an enhanced detail of the silicon retina chip is shown. The chip is equipped with the photo cells and the analogue circuits which emulate the function of the human eye.

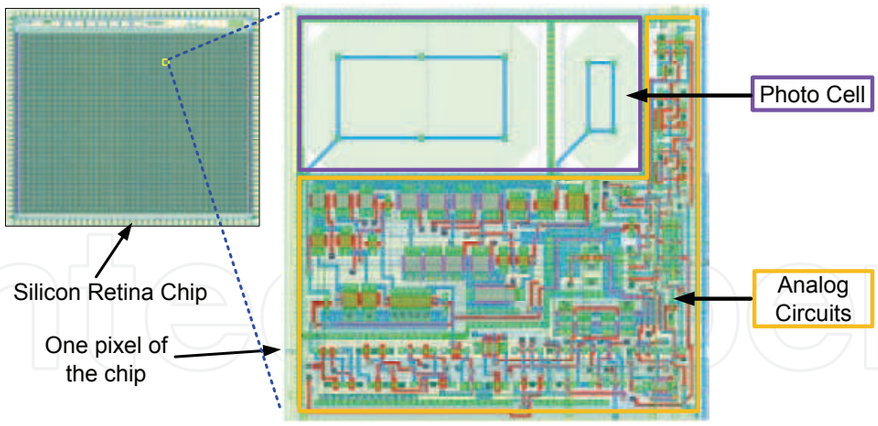


Fig. 2. Enhanced photo cell with analogue circuits of the silicon retina chip

Each pixel is connected via analog circuits with its neighbors. Due to these additional circuits on the sensor area, the density of the pixels is not as high as on conventional monochrome/color sensors, which results in a lower fill factor.
The research of this sensor type goes back to Fukushima et al. (Fukushima et al., 1970) who made a first implementation of an artificial retina in 1970. In this first realization, electronic standard components, which emulate the photo receptors and ganglion cells of the eyes, were used. A lamp array provided the visualization of the transmitted picture of the artificial retina. In 1988 Mead and Mahowald (Mead & Mahowald, 1988) developed a silicon model of the early steps in human visual processing. One year later, Mahowald and Mead (Mahowald & Mead, 1989) implemented the first retina sensor based on silicon and established the name *Silicon Retina*. The optical transient sensor (Häflinger & Bergh, 2002), (Lichtsteiner et al., 2004) used for the stereo matching algorithms described in this work, is a sensor developed at the AIT[1] and ETH[2] and is described in the work of Lichtsteiner et al. (Lichtsteiner et al., 2006).
The silicon retina sensor operates quite independently of scene illumination and greatly reduces redundancy while preserving precise timing information. Because output bandwidth is automatically determined by the dynamic parts of the scene, a robust detection of fast moving objects at variable lighting conditions is achieved. The scene information is transmitted event-by-event via an asynchronous bus. The pixel location in the pixel array is encoded in the event data using the *Address-Event-Representation* (AER) (see section 2.2) protocol.
The silicon retina sensor has three main advantages in comparison to conventional CCD/CMOS camera sensors. First, the high temporal resolution allows quick reactions on fast motion in the visual field. Due to the low resolution (128×128 with 40μm pixel pitch) and the asynchronous transmission of address-events (AEs) from pixels where an intensity change has been occurred, a temporal resolution of up to $1ms$ is achieved. In Fig. 3 (1) the speed of a silicon retina imager compared to a monochrome camera (Basler A601f@60fps) is shown.
The top image in column (1) of Fig. 3 shows a running LED pattern with a frequency of $450Hz$. The silicon retina can capture the LED changing sequence, but the monochrome camera can not capture the fast moving pattern and therefore, more than one LED column is visible in a single image.

[1] AIT Austrian Institute of Technology GmbH (`http://www.ait.ac.at`)
[2] Eidgenössische Technische Hochschule Zürich (`http://www.ethz.ch`)

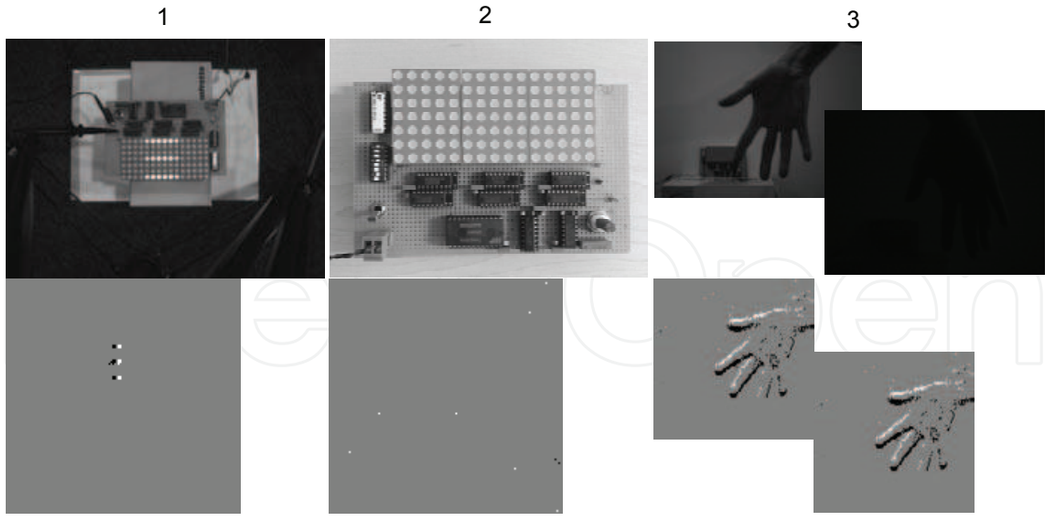


Fig. 3. Advantages of the silicon retina sensor technology, (1) high temporal resolution, (2) data transmission efficiency, (3) wide dynamic range

In Fig. 3 (2) the efficiency of the transmission is illustrated. The monochrome camera at the top of in the column (2) has no new information over time, nevertheless the unchanged image has to be transferred in any case. In case of silicon retina imagers, shown underneath, no information has to be transferred with exception of a few noise events which are visible in the field of view. Therefore, the second advantage is the on-sensor pre-processing because it reduces significantly both, memory requirements and processing power.
The third benefit of the silicon retina is the wide dynamic range of up to $120dB$, which helps to handle difficult lighting situations, encountered in real-world traffic and is demonstrated in Fig. 3 (3). The left image of the top pair shows a moving hand in an average illuminated room with an illumination of $\sim$1000 lm/m^2 and captured with a conventional monochrome camera. The second image of this pair on the right shows also a moved hand captured with a monochrome camera at an illumination of $\sim$5 lm/m^2. In case of the monochrome sensors only the hand in the well illuminated environment is visible, but the silicon retina sensor covers both situations, what is depicted in the image pair below in Fig. 3 (3).
The next generation of silicon retina sensors is a custom 304×240 pixel (near QVGA) vision sensor *Application-Specific Integrated Circuit* (ASIC) also based on a bio-inspired analog pixel circuit. The sensor encodes, as well as the described 128×128 sensor, relative changes of light intensity with low latentency, wide dynamic range, and communicates the information with a sparse, event based communication concept. The new sensor has not only a higher spatial resolution, the sensor has also a higher temporal resolution of up to $10ns$ and a decreased pixel pitch of 30μm. This kind of sensor is used for further research, but for the considerations in this work the 128×128 pixel sensor is used.

2.2 Address-event data representation

The silicon retina uses the so-called *Address-Event-Representation* (AER) as output format which was proposed by Sivilotti (Sivilotti, 1991) and Mahowald (Mahowald, 1992) in order to model the transmission of neural information within biological systems. It is a digital asynchronous multiplexing protocol and the idea is that the bandwidth is only used if it is necessary. The protocol is event-driven what means that only active pixels transmit their

output, and in contrast, the bus is unused if the pixels of the sensor cannot detect any changes. Different AER implementations have been presented in the work of Mortara (Mortara, 1998) and the work of Boahen (Boahen, 2000). In the work of Häflinger and Bergh (Häflinger & Bergh, 2002) an one-dimensional correspondence search takes place and the underlying data protocol is AER.
The protocol consists of the timestamp TS which describes the time when an event has occurred, the coordinates (x,y) define where the event has occurred, and the polarity p of the contrast change (event) which is encoded as an extra bit and can be ON or OFF, representing a fractional change from dark to bright or vice-versa. In the current version the timestamp is transmitted in absolut time which means it increases continuously from the start of the camera. The new protocol version sends a relative timestamp which saves transmission bandwidth.

3. Stereo processing with silicon retina cameras

The stereo matching is the elementary algorithm of each stereo vision application. Two cameras are placed in a certain distance (baseline) to observe the same scene from two different point views. Existing stereo matching algorithms deal with data from conventional monochrome/color cameras and cannot be applied directly to silicon retina data.
Existing methods for adjustment of the cameras, as well as calibration and rectification methods have to be extended and changed for the event-based stereo processing.
Also, for algorithm verification, existing data-sets could not be used, as these are based on frame-based representation of a scene. Thus, an event-based stereo verification method was implemented that describes a scene using geometric primitives. For verification purpose, ground truth information is essential, which could also be generated based on this scene description.

3.1 Stereo sensor setup

The goal of the stereo vision sensor described in this chapter is to detect fast approaching objects to forecast side impacts. For this reason, two silicon retina sensors are placed on a baseline to build up a stereo system. This stereo system is designed for pre-crash warning and consists of the stereo head and an embedded system for data acquisition and processing. The stereo vision sensor must fulfill requirements given by the traffic environment. In Fig. 4 a sketch of the side impact scenario including some key parameters is shown.
In the mentioned application, the stereo vision system has to detect closer coming objects and activates pre-safe mechanisms of the car. The speed of the approaching vehicle is defined with $60km/h$ and a minimal width of an object of $0.5m$. For activating the corresponding safety mechanisms of the car, we assume that the vehicle needs about $300ms$ which defines the detection duration of the camera system. A vehicle with a speed of $60km/h$ passes a distance of $5m$ in $300ms$, therefore the decision if an impact will occur or not has to be made $5m$ before the vehicle will impact. In Fig. 4 the detection distance and the critical distance, where a decision has to be made, are shown. These requirements define the key parameters of the optical system and the following embedded processing units.

3.2 Adjustment of the stereo sensor

Before the silicon retina stereo vision system can be used, a configuration has to be made. The focus of the lenses has to be set, the calculation of the calibration parameters has to be computed and for stereo matching the rectification parameters has to be extracted. In contrast

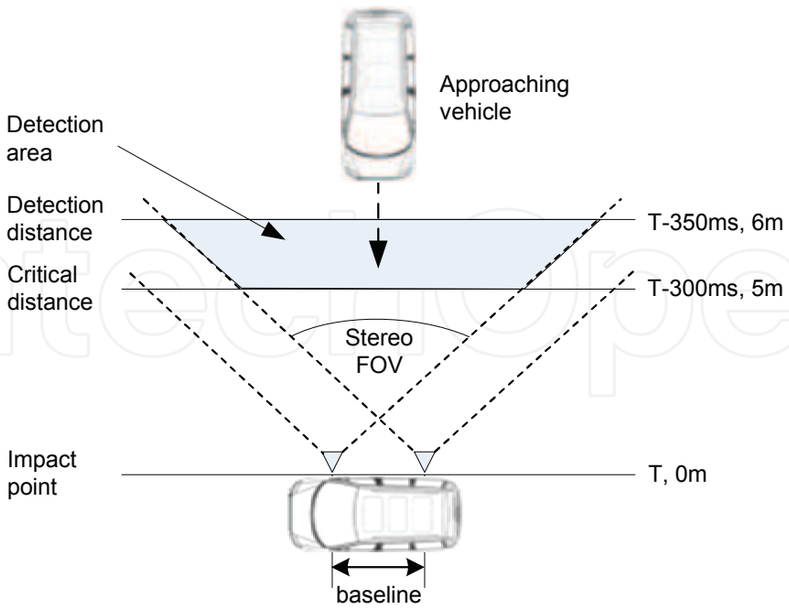


Fig. 4. Stereo vision setup for the use in a pre-crash warning side impact detection application

to conventional camera sensors, the silicon retina has no stable image which can be used for configuration and calibration purposes. Therefore, new methods for lens adjustment, calibration and rectification were implemented.

3.2.1 Lens configuration

Before the camera system can be used, the lenses must be brought in-focus with respect to the desired sensing range. The silicon retina sensor delivers image information only when changes in intensity happen. Therefore, an object must be moved in front of the camera, so that address-events will be generated. For this reason, a hardware which helps to adjust the lenses of silicon retina cameras was built. It allows the definition of thick or thin lines, which are moving in front of the camera to generate a stimulus. The hardware was built on a breadboard shown in Fig. 5. The board consists of a 15×5 LED matrix which is used

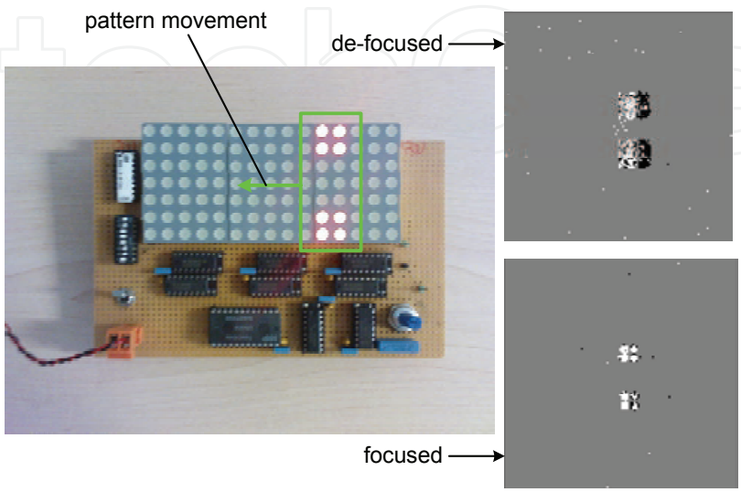


Fig. 5. Hardware for the adjustment of silicon retina camera lenses

to generate the stimuli for the silicon retina cameras. With the potentiometer the frequency (speed) of LED changes can be configured. After each cycle the pattern of highlighted LEDs is moved leftwards by one column. There is a software tool available, which allows a live view of the silicon retina output. This software transforms the address-events to an image frame and displays this stream on the screen. Using this tool, the impact of the lens adjustment can be directly observed on the screen. The images on the right side show the de-focused silicon retina image on the top and the correctly focused lens on the bottom. After the adjustment of the lenses, the data of the stereo vision system can be used.

3.2.2 Calibration and rectification

The acquired data from the cameras are not prepared for line-by-line matching, respectively event-by-event matching, because the epipolar lines (Schreer, 2005) are not in parallel. Therefore, a rectification of the camera data is carried out. Before this rectification can be done, the cameras have to be calibrated. With conventional cameras, the calibration pattern (Fig. 6 on the top) is captured in different views from the right and left camera. Then the

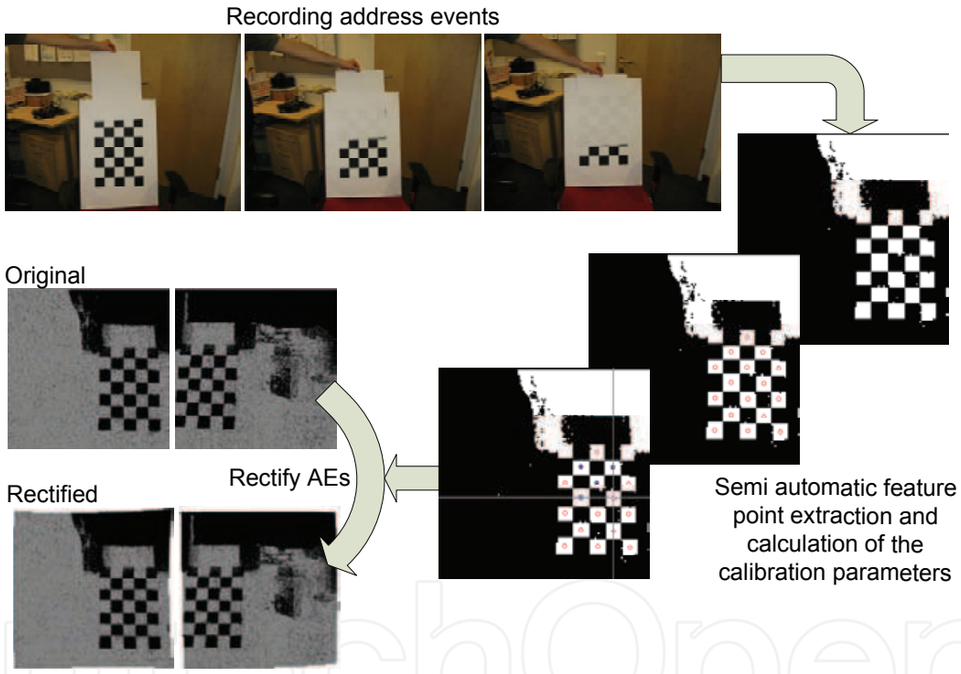


Fig. 6. Calibration and rectification of silicon retina cameras

corners of the pattern are used for calculation of the camera parameters. For silicon retina imagers, it is not possible to capture a static calibration pattern if there is no movement, more precisely no change in the intensity. Thus, an alternative approach is necessary. In Fig. 6 on the top, the calibration pattern is shown in a stable position and a white paper is moved up and down in front of the calibration pattern. During a period of time all address-events are collected and stored in an output file. The collected address-event data are converted into a binary image, which is used for the extraction of feature points. Instead of the corners from the calibration pattern, the center of gravity of the squares for extraction of corresponding features are used. The right side in Fig. 6 shows the semi-automatic extraction of the feature points, because not all centers are feasible for the calibration step. That means, the algorithm extracts more points but not all of them are supporting the calibration process and therefore,

the user has to choose manually which points should be used for the calibration step. For the calibration itself the method from (Zhang, 2002) in combination with the calibration toolbox from Caltech for Matlab (Bouguet, 2008) is used. All data extracted from the binary images are loaded via the external interface into the calibration engine and the results are applied on silicon retina data for the calibration and rectification step.
The left side of Fig. 6 shows an example of rectified silicon retina data from the left and right camera. In a next generation of calibration and rectification of silicon retina cameras LCD screens will be used where a pattern changes in a defined way in order to excite events.

3.3 Frame-based stereo matching

In the field of stereo matching exists different approaches for solving the stereo correspondence problem, but these approaches are developed for frame-based data from stereo vision systems based on conventional monochchrome/color cameras. If existing frame-based stereo matching algorithms shall be used with silicon retina cameras, the data of the silicon retina stereo vision system has to be converted into framed image/data streams before the frame-based stereo matching approaches can be applied.

3.3.1 Address-event to frame converter

Before the AE data can be used with full frame image processing algorithms, the data structure is changed into a frame format. For this reason an address-event-to-frame converter has been implemented.
The silicon retina sensor delivers permanently ON- and OFF-events, which are marked with a timestamp t_{ev}. The frame converter collects the address-events over a defined time period $\Delta t = [t_{start} : t_{end}]$ and inserts these events into a frame. After the time period the frame is closed and the generation of the next frame begins. The definition of an event frame is

$$AE_{frame} = \int_{t_{start}}^{t_{end}} AE_{xy}(t_{ev}) dt_{ev}. \tag{3}$$

Different algorithm approaches need a different frame format. The silicon retina stereo camera system used within this work is evaluated with two algorithms derived from two different categories. The first algorithm is an area-based approach, which works with the comparison of frame windows. The second algorithm is a feature-based variant which matches identified features. Both categories need differently constructed frames from the converter. Due to this reason, the converter offers configurations to fulfil these requirements. Fig. 7 shows on the left side the output frame of the converter with the collected ON- and OFF-events. The resolution

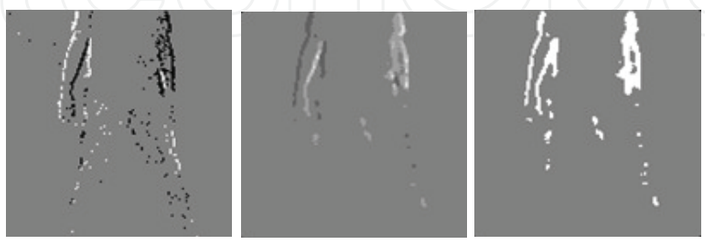


Fig. 7. Different results of AE to frame converter

of the timestamp mechanism of the silicon retina is $1ms$, but for the algorithm evaluated in this work a Δt of $10ms$ and $20ms$ is used. The Δt is changed for different conditions, which produce a different number of events.

The image in the middle of Fig. 7 shows a frame built for an area-based matching algorithm. For this reason each event received in the defined time period is interpreted as a gray value, with

$$AE_{frame} = \int_{t_{start}}^{t_{end}} graystep(AE_{xy}(t_{ev}))dt_{ev}. \tag{4}$$

The background of the frame is initialized with 128 (based on a 8 bit grayscale model) and each ON-event adds a gray value, and an OFF-event subtracts one. In Equation 5, the function for generating a gray value frame is shown. The 8 bit grayscale model limits the additions and subtractions of the $\Delta_{grayvalue}$ and saturates if an overflow occurs.

$$graystep(AE_{xy}(t_{ev})) = \begin{cases} AE_{frame_{xy}} + \Delta_{grayvalue} & AE_{xy}(t_{ev}) = ON_{event} \\ AE_{frame_{xy}} - \Delta_{grayvalue} & AE_{xy}(t_{ev}) = OFF_{event} \end{cases} \tag{5}$$

The right image in Fig. 7 shows a frame built for a feature-based image processing algorithm. In this case, multiple events received within the defined period of time will be overwritten instead of accumulated. Equation 6 shows the frame building and the used simplify function is illustrated in Equation 7.

$$AE_{frame} = \int_{t_{start}}^{t_{end}} simplify(AE_{xy}(t_{ev}), conv_{on})dt_{ev} \tag{6}$$

The simplify function gets a second parameter ($conv_{on}$) to decide the event variant (only ON or OFF). This frame is prepared for different kind of feature-based algorithms and also for algorithms based on segmentation.

$$simplify(AE_{xy}(t_{ev}), conv_{on}) = \begin{cases} ON_{ev} & AE_{xy}(t_{ev}) = ON_{ev} \wedge conv_{on} = 1 \\ OFF_{ev} & AE_{xy}(t_{ev}) = ON_{ev} \wedge conv_{on} = 0 \\ ON_{ev} & AE_{xy}(t_{ev}) = OFF_{ev} \wedge conv_{on} = 1 \\ OFF_{ev} & AE_{xy}(t_{ev}) = OFF_{ev} \wedge conv_{on} = 0 \end{cases} \tag{7}$$

Both specialized generated frames (middle and right in Fig. 7) can optionally be filtered with a median filter to reduce noise and small artifacts. With these settings every Δt, a new frame from the left and right address-event-stream is generated. These frames are now handled as images for the stereo matching algorithms described in the next section.

3.3.2 Area-based frame stereo matching

The area-based approach uses the neighborhood (block) of the considered pixel for the matching of each pixel and tries to match this with a corresponding block from the other camera image. These kind of algorithms are used if rectified stereo image pairs are available and if the output shall be a dense disparity map. Some algorithms using block-based techniques are shown in ((Banks et al., 1997), (Banks et al., 1999), (Zabih & Woodfill, 1994)).
For the demonstration of area-based processing with silicon retina data a *Sum of Absolute Differences* (SAD) correlation algorithm (Schreer, 2005) was chosen. A block matching, based on ON and OFF events, produces a lot of similar blocks and a lot of mismatches will appear. The grayscale images have more then two values and therefore, the statistical significance of a block is larger. Also in the work from Milosevic et al. (Milosevic et al., 2007) a silicon retina stereo matching algorithm based on SAD is shown. This algorithm uses an address-event-to-frame conversion to get grayscale images, which can be matched with the SAD technique. Milosevic et al. use in their work correlation windows of up to

15×15 without a proper rectification step to find the matches, perhaps this window size leads to a low processing performance. Therefore, the approach in our work uses different conversion outputs (see 3.3.1) and an adequate rectification step to enable smaller window sizes for the SAD and increase the performance of the computation. For the processing of the SAD algorithm the grayscale frames, as shown in Fig. 8 on the left side, are used. These

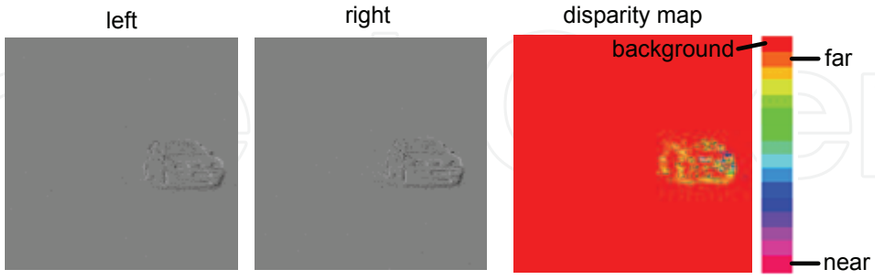


Fig. 8. Input stereo images for the SAD algorithm (left two images) and result of the matching process (right image)

input images consist of pixels with different grayscale values (accumulated address-events), therefore the matching results of the correlation are more confident. The found matches are used for the calculation of the disparity value, whereby the absolute value of the difference of both x-coordinates is taken (x-coordinate of the left and right image). In Fig. 8 on the right side the disparity image of the stereo pair on the left side is shown.
The disparity values are visualized in a color-coded manner according to the legend at the bottom of Fig. 8 on the right side. Due to the large amount of background pixels, the result is not a dense disparity map. The disparity map shall have the same or equal outlines as the original silicon retina input image.

3.3.3 Feature-based frame stereo matching

For feature-based stereo matching with silicon retina data, the address-event-data must be converted again, as described in section 3.3.1, before the features can be extracted from the image. Shi and Tomasi (Shi & Tomasi, 1994) give more details about features in their work and describe which features are good for tracking. Within their work they discuss e.g. the texturedness, dissimilarity and convergence of features. For the evaluation of the feature-based stereo matching with silicon retina cameras, a segment center matching approach is chosen. Tang et al. (Tang et al., 2006) describe in their work an approach for matching feature points. In Fig. 9, the left stereo pair shows the address-event-data converted into images feasible for the feature-based algorithm approach. If the image was not filtered

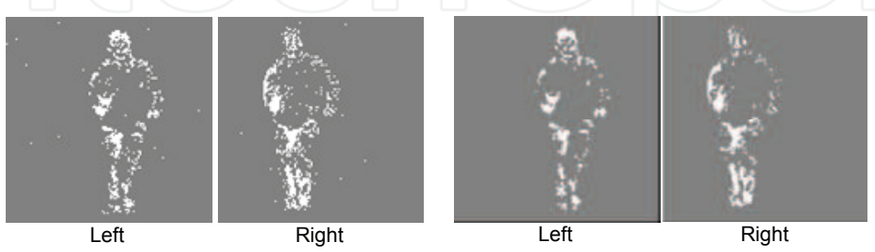


Fig. 9. Left: Input stereo images for the feature matching algorithm, Right: Input images filtered with a 3×3 median filter

during the conversion step, the image must be filtered now in order to remove noise in the

image. The right image pair in Fig. 9 shows the data after a 3×3 median filter has been applied.
In the next step some morphological operators are used to get linked regions of pixels which can be labeled as one connected area. The images are treated with some morphological operations (Gonzales & Woods, 2002) for the enhancement of features, which are required by the next step of the center matching algorithm. In the algorithm for the silicon retina images a square shape structuring element was used. The structuring element for the erosion has a size of 4×4 and the square for the dilation a size of 7×7. In the first row of Fig. 10, the silicon retina images after the dilation (left image pair) operation are shown and the results after the erosion (right image pair) are depicted. The images are now prepared for the segmentation

Fig. 10. Stereo images after the morphological operation dilation and erosion

and labeling step. For the region labeling a flood fill (Burger & Burge, 2005) algorithm is used. This algorithm labels all linked areas with a number in a way that the regions can be identified. The result of region labeling is shown in Fig. 11 in the left image pair. After region labeling, a

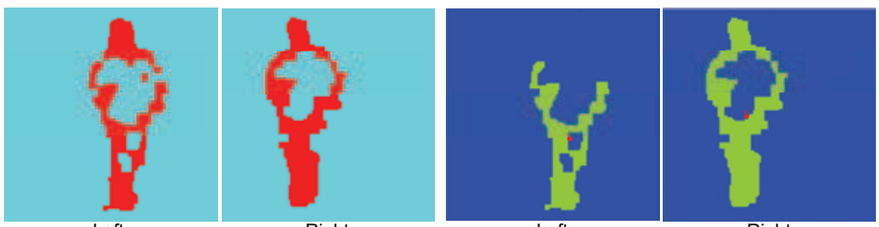


Fig. 11. Left: all found regions, Right: all found segments and the corresponding center of each segment (red dot)

few segments should be available which are used for matching. Before the matching can start, all segments with less than a defined amount of pixels, are removed. A region is a collection of more than one pixel and has a defined shape. A pixel-by-pixel matching is not possible and therefore, it must be defined how the whole feature (region) shall be matched. In a first step, the features are ordered downwards according to their area pixel count. This method is only useful if the found regions in the left and right image are nearly the same (also the same area pixel count). As representative point of the feature the center of the feature was chosen. The so called *center-of-gravity* must be searched for each feature. All found centers are marked with a red dot as shown in the right image pair in Fig. 11.
The center of the corresponding segment in the left and right frame can differ. Due to this reason the confidence of the found centers are checked. This mechanism checks the differences of center points, if they are too large, the center points are ignored for the matching. If the center points lie within the predefined tolerances, the disparity is calculated which represents the disparity of the whole object.

3.4 Event-based stereo matching

The usage of conventional block-based and feature-based stereo algorithms has shown a reduction of the advantage of the asynchronous data interface and throttle the performance of the silicon retina cameras. Due to this fact, a frame-less and therefore event-based stereo matching approach, which exploits the characteristics of the silicon retina technology, has to be developed. For this reason, a time-correlation algorithm for the correspondence search is used. This algorithm uses the time difference between events as the primary matching costs. In Fig. 12, the whole workflow of the event-based algorithm approach is shown.

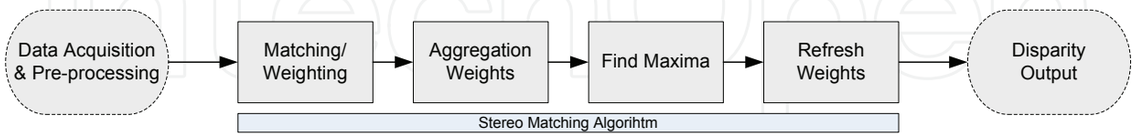


Fig. 12. Workflow of the event-based time correlation stereo matching algorithm

Data Acquisition and Pre-processing

Before the stereo matching workflow starts, the data from the silicon retina sensors are acquired, which means the data are read from the adapter board buffers and given to the rectification unit. For the rectification step, all needed parameters were calculated in a previous calibration step. The calibration determines the intrinsic camera parameters plus the rectification matrices for both sensors. For the calibration of silicon retina cameras the method described in section 3.2.2 is used and it is part of the pre-processing step.

Matching and Weighting

After the pre-processing, the stereo algorithm starts and uses the rectified data stream from silicon retina cameras. In the first step, the matching of the events is carried out where for each oncoming event a corresponding event on the opposite side is searched. For this search, all events of the current timestamp, as well as events from the past are used. This means also previous events are considered during the correspondence search. Due to the previous rectification, the search is carried out in a horizontal line within the disparity range.

In Fig. 13, on the top left side the event buffers are shown which store the current and the historical events. If there are possible matching candidates within the disparity range of actual and historical events, which have the same polarity, the timestamps are used for calculating the matching costs. In Fig. 13, the matching of an event at the x-coordinate 40 is illustrated. The left event is the reference event and the search takes place on the right side where three candidates with the same polarity and within the considered history are found. Now, the time difference between the timestamp of the left camera and the three found events of the right camera is calculated.

For determination of the costs of a found matched event pair, different weighting functions were used. In Fig. 14, all used weighting functions are depicted. On the abscissa of the diagrams the weighting costs which may achieve a maximum of 10 (derived from the maximal considered historical events) are plotted and on the ordinate of the diagrams the time difference is shown. In the example, the considered history is 10, which means from the current timestamp of an event 10 timestamps of the past are regarded for the current calculations. The left function shows a simple inverse linear relation between the time difference Δt and the weight. In the middle chart an inverse quadratic function is depicted which is faster declining and matches more current events and does not consider older events in the same amount. The Gaussian function shown in the right diagram of Fig. 14 increases,

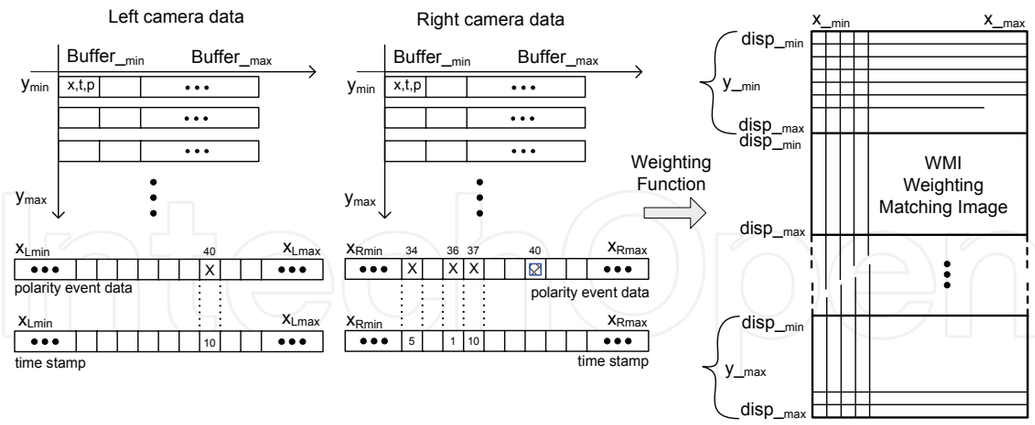


Fig. 13. Matching and weighting of corresponding address-events and writing of calculated costs into the WMI

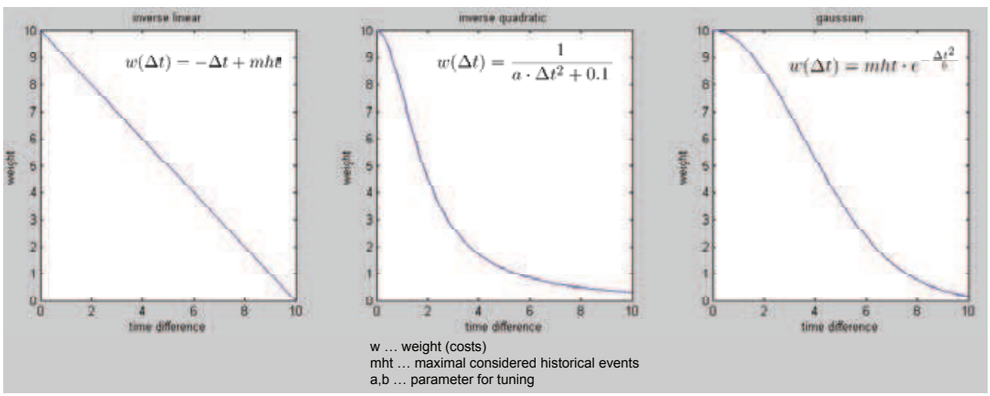


Fig. 14. Weighting function for calculating the costs of matched address -events

in comparison to the inverse linear function, the weights of current timestamps and decreases the older timestamps. Both functions on the right side, the inverse quadratic and the Gaussian can be tuned with a parameter for the adaption to different weighting needs. All the matched and weighted events are written into the *Weighted Matching Image* (WMI) shown in Fig. 13. This data storage is a two dimensional representation of a three dimensional space, where a place for each pixel coordinate and disparity level is reserved. The WMI is a dynamic data storage which is updated each processing cycle and only deleted if a reset takes place. That means all costs entered, stay in the WMI for a defined time till they are removed and so the matched costs from previous steps contribute to the results of the current calculations.

Aggregation Weights

The next step of the algorithm is the aggregation of the weights in the WMI. Therefore, the WMI structure is transformed logically and a filter kernel works on the weights with the same disparity. In the current algorithm, an average filter kernel with a variable window size is used.

Find Maxima
After the aggregation step, the maximum costs for each coordinate which represents the best matching disparity are searched.

Refresh Weights
In consideration that the WMI is a refreshing data structure, after the maximum search all weights are checked if they have to be deleted from the WMI, and therefore the weight itself is a lifetime counter. In each round, the weight is reduced with a defined value till the weight is zero and then deleted from the WMI or refreshed by a new match as well as a new weight.

Write Disparity Output
The results are written into the disparity map, which can be used from the application for further processing.

3.5 Verification of event-based stereo matching algorithms

Existing performance and quality metrics cannot be used within event-based data processing, thus a new approach has been implemented that redefines existing metrics for the event-based approach and describes the performance.
Early verification and validation approaches used in real-world environments were justified with a measuring tape. This was sufficient for some estimations whether an algorithm approach was generally computing or not. Predictions and declarations of the achieved quality of an algorithm were not possible.
A method for visualizing the performance of classifiers are receiver operating characteristics (ROC). Fawcett (Fawcett, 2004), (Fawcett, 2006) and Provost and Fawcett (Provost & Fawcett, 2001) give an introduction and practical considerations for ROC analysis in their work. Within verification of silicon retina stereo matching algorithms, we also address two-class problems. Each instance of the *ground truth* (**GT**) is mapped to one element of the set $\{p,n\}$, where p is an existing event and n is a missing ground truth event. The classification model represents the mapping from **GT** to a predictable class, the *disparity map* (**DM**) is mapped to the set $\{y,n\}$, where y is an existing and n is a missing disparity event. Based on these combinations, a two-by-two confusion matrix can be build. Metrics evaluation is based on comparing the disparity to the ground truth data set. Equation 8 defines both, the disparity and the ground truth set for metrics evaluation, where t_h defines the propagation delay of a set.

$$\begin{aligned} dm(x,y,t) &:= \int_{t-t_h}^{t} DM(x,y,t)dt \\ gt(x,y,t) &:= \int_{t-t_h}^{t} GT(x,y,t)dt \end{aligned} \tag{8}$$

– A *true positive* is defined by Equation 9 for both existing disparity and ground truth data with an error tolerance δ_d.

$$tp(t) := \sum_{x\in X, y\in Y} [|dm(x,y,t) - gt(x,y,t)| < \delta_d] \tag{9}$$

– A *false positive* is defined in Equation 10 for the same restrictions as *true positives*, though the error tolerance δ_d is exceed.

$$fp(t) := \sum_{x\in X, y\in Y} [|dm(x,y,t) - gt(x,y,t)| \geq \delta_d] \tag{10}$$

– A *false negative* $fn(t)$ is defined by an existing ground truth and a missing disparity value.
– A *true negative* $tn(t)$ is defined by both, a missing ground truth and a missing disparity value.

Based on these performance primitives, further performance metrics such as true positive rate or false positive rate of a time-slot can be computed.

4. Implementation

Due to the special characteristics of this novel sensor technology and the asynchronously data processing algorithm approach, existing data acquisition and processing techniques are not adequate. The high temporal resolution of the sensor results in data peaks up to 10*Meps* (Mega events per second).
Within the project, two system demonstrators were implemented. The PC based demonstrator is used for low-speed data processing and algorithm verification. The DSP demonstrator is intended to be used for high-speed real-time data processing. A third FPGA based demonstrator is outlined and represents the next step after the DSP implementation and gives and overview how the performance of event-based stereo matching can be increased.

4.1 PC demonstrator

The PC demonstrator is used for low-speed data processing, coupling the Ethernet interface of the imagers which offers AE without time-information. The timing information, which is essential for event based stereo matching, is assigned when acquiring the data.
The implemented tool is shown in Fig. 15. The tool consists of an viewer optimized for AE data and an embedded interpreter language, including the data objects for scene generation and verification. Scene generation contains geometric primitives that can be used for scene description. Also, recorded scenes can be used as objects for scene description. All geometric objects are inhered from a base object, afford generating ground truth information that is essential for verification. Using these base objects, also complex objects e.g. vehicles, can be compound.

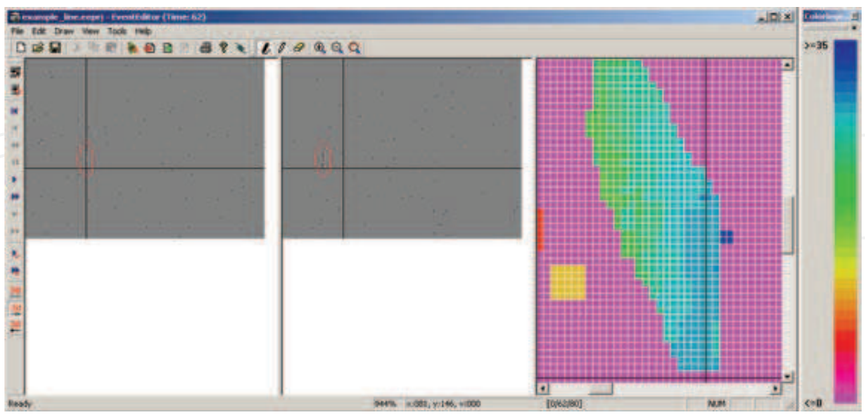


Fig. 15. Verification Tool; scenario shows a moving line from the upper left to the bottom right corner of the visualized time-slot. left window: left imager; middle window: right imager; right window: processed disparity map by the stereo matching algorithm

The tool handles and processes the address-event represented data asynchronously similar to the real environment. For data visualization, visualized time-slots are used, as shown

in section 3.3.1. The internal data management is realized using data containers enclosing timestamp-sorted AEs, the identifier, and the coordinates. For advanced visualizing of the data, the virtual reality modeling language (VMRL) (World Wide Web Consortium (W3C), 1995) is used. Fig. 16 shows a processed disparity map of a recorded scene of a pedestrian.

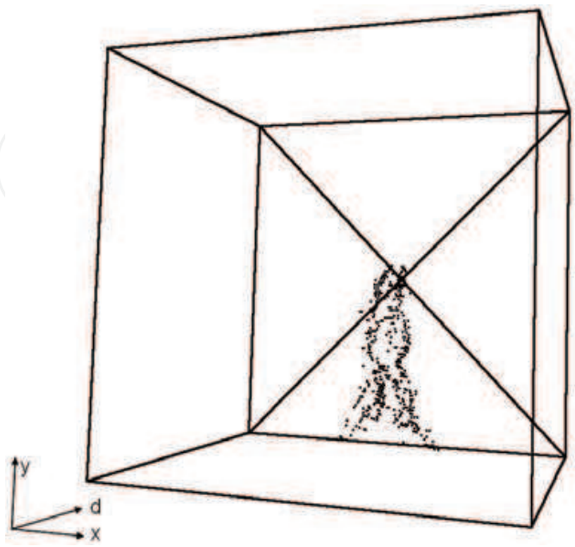


Fig. 16. Disparity map of scene with a pedestrian processed with an area-based stereo matching algorithm and visualized in VRML.

4.2 DSP demonstrator

The embedded system used for data acquisition and data processing is based on a distributed digital signal processing solution. Fig. 17 shows a schematic diagram of this demonstrator consisting of two silicon retina imagers connected to an adapter-board, that implements a memory mapped interface to the TMS320C6455. Both imagers stream data to the first in first out (FIFO) devices on this board. Once enough data are acquired, an interrupt is triggered and the direct memory access (DMA) controller flushes the FIFOs and transfers the data to the DSP, where it is available for further processing.

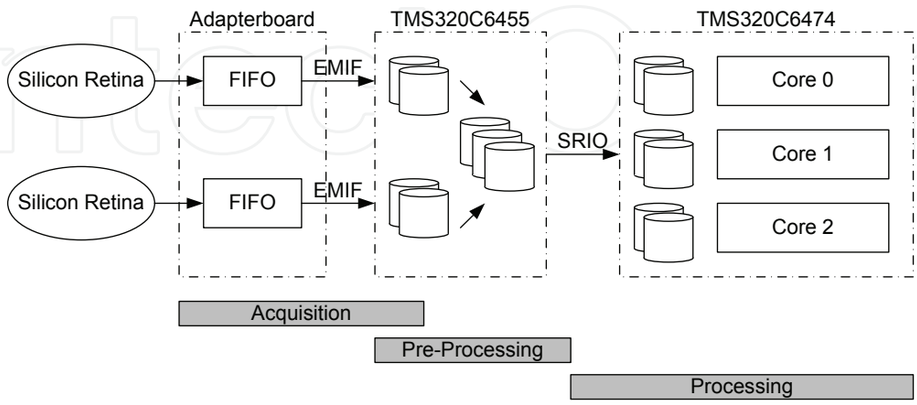


Fig. 17. Schematic diagram of the DSP-based demonstrator.

Further processing on the data acquisition processor includes noise filtering of events and a load balancer for partitioning the acquired amount of data on the multi-core processor that is

responsible for processing the stereo matching algorithm. Equation 11 shows the balancing criteria of the load balancer, where $\mathbf{E}(\mathcal{Y})$ is the y-coordinate of the event, N is the number of parallel units, H is the height of the sensor, and n is the processor identifier.

$$n = \frac{\mathbf{E}(\mathcal{Y})N}{H} \tag{11}$$

For data exchange between the single-core to the multi-core processor, a serial high performance interface is used which is intended for interconnecting distributed processing systems on chip-to-chip and board-to-board level. The data transfer is completely handled in hardware using a processor peripheral module. After transferring a burst of data, an interrupt is triggered on the specific core to initiate the stereo matching process.

4.3 FPGA demonstrator

The introduced DSP-based platform generally enables parallel computation of the proposed stereo vision algorithm by using multi-core processors. Additionally, the behavior of the used silicon retina sensors leads, in contrast to frame-based imagers, not only to less redundancy but also to a reduced amount of data because the retina only delivers data on intensity changes of the ambient light. Therefore, the underlying vision system usually have not to cope with a huge amount of data and so have to provide less memory bandwidth, which usually is the bottleneck in embedded vision systems. Unfortunately, the asynchronous characteristics of the silicon retina yields to a non-constant data rate, but any data peaks can be caught with a simple FIFO at the intput-stage.

Due to the computationally sophisticated and expensive nature of the presented event-based stereo matching algorithm, a more parallelized data processing would be obvious, and effectively necessary to fulfill the timing constraints of real-time applications and fully exploiting the high temporal resolution of the silicon retina sensors. This algorithms however, can significantly benefit from application depended customizations of the underlying system architecture: hence optimized memory access patterns, fast on-chip buffers or line-caches, and special computation units for data correlation are preferred. ASICs or even more FPGAs, can be used to put such customized architectures into practice and thus to exploit the immanent parallelism of stereo-vision algorithms (Porter & Bergmann, 1997).

However, an FPGA-based implementation will decrease the overall complexity of the embedded system because, e.g., in our case, as it is shown in Fig. 18, the data acquisition unit, the rectification module, the computation units, and finally the transmission of the disparity map can be integrated into one single chip, which obviously leads to a smart stereo vision systems (Belbachir, 2010). Nevertheless, by adapting the memory interfaces and access patterns, and the data path of the processing units to the asynchronous behavior of the silicon retina and the address-event format, the latency of the system would be reduced.

Furthermore, by using massively parallel hardware architectures the throughput of the system would be increased. The integration of all functional blocks into one customized FPGA or ASIC yields not only to a simplification but also to an improvement of the scalability and the overall energy efficiency of the system.

Fig. 18 shows a recommendation of a hardware-based implementation of the event-based stereo matching algorithm. First of all, data from the sensor must be gathered within the acquisition unit by a FIFO, which could be either on-chip if there is sufficient memory and the succeeding units are fast enough or even off-chip. In this case, the FIFO itself could be bigger and therefore the following processing units can operate at lower speed. After this, the events must be rectified, done by a pipelined online rectification unit. Unfortunately, this step is

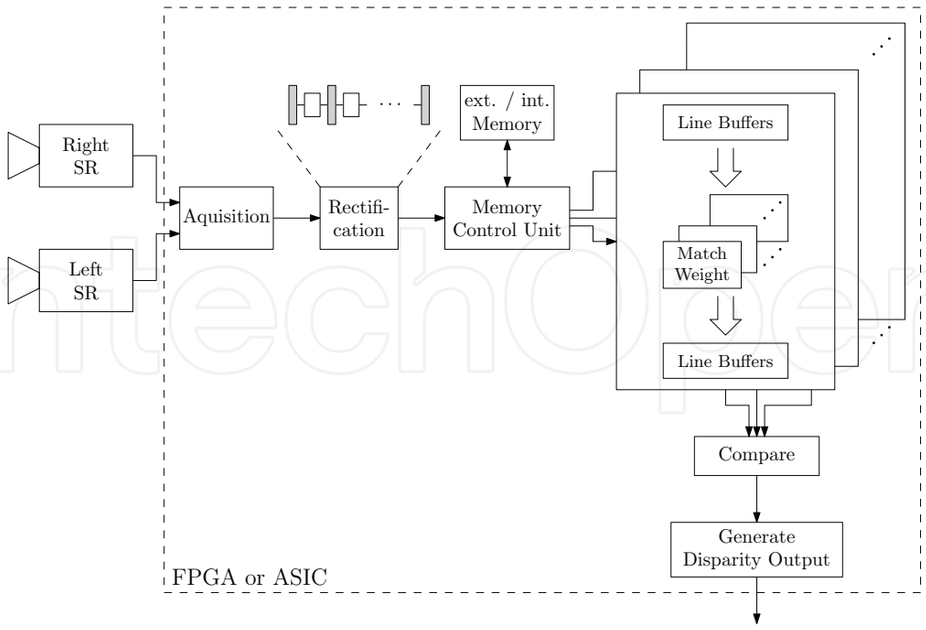


Fig. 18. Possible architecture of a hardware-based implementation

computationally very intensive even if tight tolerances should be attained, but this approach is very memory-efficient and, compared to the current used look-up table-based approach, can be parallelized as well. This processing step is one of the main challenges of the whole algorithm because it must entirely be accomplished before proceeding with the subsequent calculation.

An other key point of the system will be the memory control unit, because here all data accesses must be merged and a high bandwidth will be required. Here are also on- or off-chip memories possible, since the resources which are provided by up-to-date high-end FPGAs and the reduced amount of data supplied by silicon retina sensors, on-chip memory would be preferred. Additionally, the on-chip variant allows the usage of several dual-ported memories enabling the parallelization of the accesses and therefore a very high bandwidth. On the other side, using an off-chip memory leads to a simpler memory architecture and facilitates the deployment of a mid-end FPGA which yields also to a simplification of the overall system. In order to overcome this bottleneck, a further reduction of the address-event data should be done in any cases, e.g., using not the whole timestamp for an event, but only differences of timestamps corresponding to the depth of the considered history. Thus, memory accesses and space can be optimized even if off-chip memory with limited bandwidth will be used.

The matching and weighting of the address-events can be done in a parallel manner by selective pre-fetching of the events, although one match and weight unit processes one single line of pixels from the left and right sensor. If a block-based approach is used, the line buffers on the input side must be interconnected according to the block size. The criterion for matching and the weighting function are not important for the architecture as far as they can be mapped onto hardware. In the end, the results will be brought together and the disparity output will be generated.

5. Results

The different stereo matching approaches for address-event-data have been tested with a variety of parameters. This section compares the results of frame-based stereo matching divided into area-based approaches and feature-based approaches, as well as an event-based stereo matching approach. Each of the algorithm has defined parameters which can be used for the tuning of the algorithm results.

5.1 Results of frame-based address-event stereo matching

This section shows the results of frame-based stereo matching with silicon retina data. For this tests, the silicon retina data streams were converted into frames which can be used from stereo matching algorithms developed for conventional frame-based cameras.

5.1.1 Area-based address-event stereo matching

The algorithm parameter of the SAD is the size of the correlation window. We tested the algorithm with an object at three different distances ($2m$, $4m$, $6m$) and different settings of the address-event converter. In Fig. 19, the results of the SAD algorithm processing AE frames are given. On the x-axis the different converter settings at three different distances are shown. The first number represents the object distance in meters, the second value describes the time

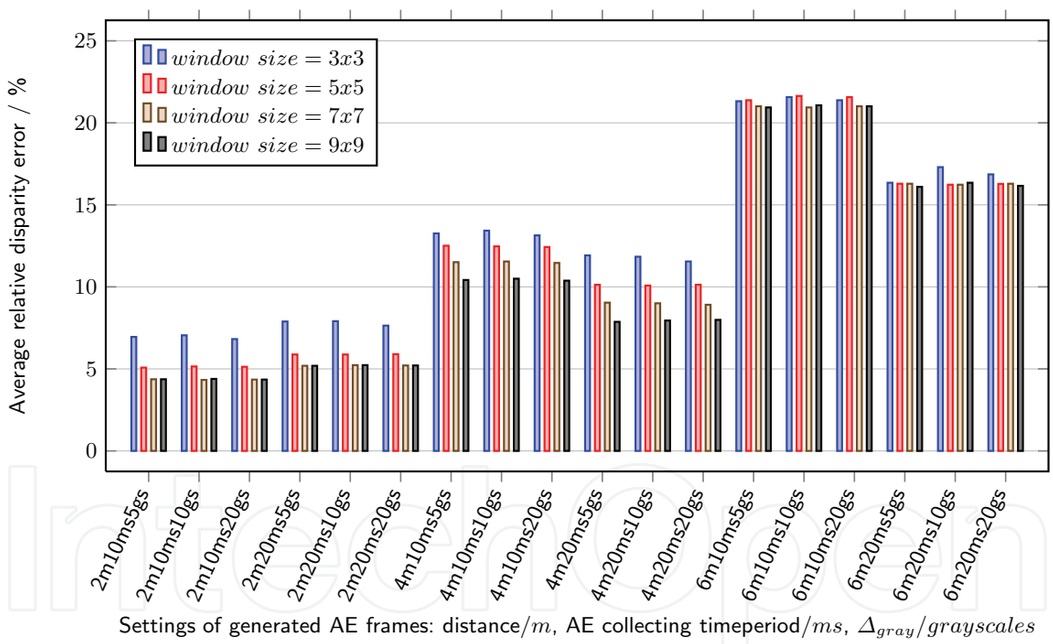


Fig. 19. Results of the area-based stereo matching algorithm on address-event frames

period for collecting address-events, and the last value represents the grayvalue stepsize for the accumulation function described in section 3.3.1. For each distance, all four converter settings with four different SAD correlation window sizes are evaluated. The output on the y-axis is the average relative error of the distance estimation based on 500 image pairs. The results in Fig. 19 show that the average relative disparity error increases with the distance of the object. In near distances, the results are influenced by the correlation window size, especially there is a significant difference between the usage of a 3×3 window and a 9×9 window. In the distance of $4m$ and $6m$, the results with a timestamp collection time Δt of 20ms

are better. The third parameter of the generated input AE frame is the grayscale step size which has no influence at any distance. Generally, we achieve with the SAD stereo matching approach used for AE frames in the main operating distance of 4m a minimal error of 8%. That is equivalent to an estimated distance range of $3.68m$-$4.32m$.

5.1.2 Feature-based address-event stereo matching

This section shows results of the feature-based stereo matching algorithm using AE frames. The parameters of the segment center matching are the morphological erosion and dilation function at the beginning of the algorithm. In Fig. 20, the results of the feature-based algorithm processing AE frames are given. For center matching only the collecting time period

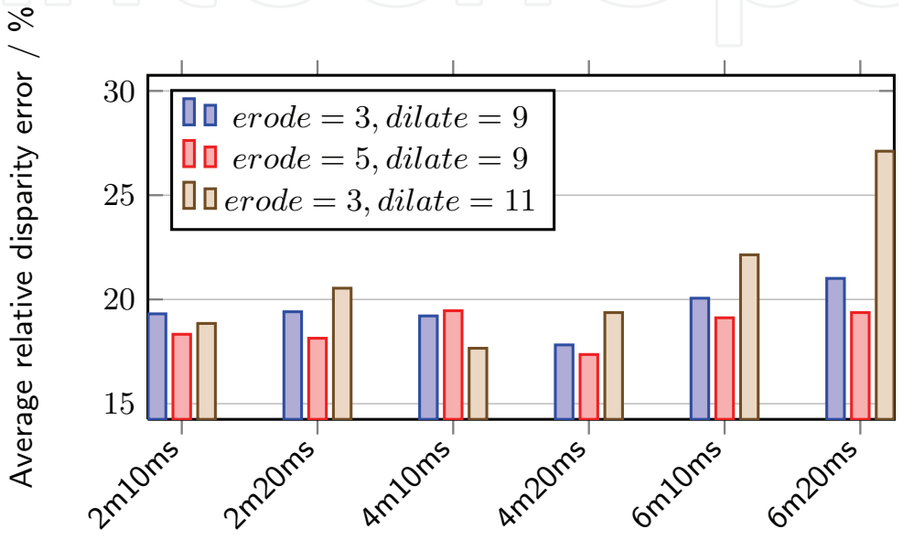


Fig. 20. Results of the feature-based stereo matching algorithm on address-event frames

Δt of the address-events is varied, which is shown with the second value from the descriptors on the x-axis. All converter settings with three different morphological erosion and dilation settings are evaluated. The structuring element is always a square. The results on the y-axis shows the average relative disparity error of the feature center matching at three different distances with two different address converter settings and with three different morphological function combinations. The results are based on 500 image pair samples. The achievements in Fig. 20 show that the average relative disparity error depends on the sizes of the structuring elements. At all distances, the morphological combination erosion=3 and dilation=5 produces the best results. The timestamp collection time Δt has only a significant influence at the distance of $6m$. In the main operating distance of 4m, the minimal error is 17%, which is equivalent to an estimated distance range of $3.32m$-$4.68m$.

5.2 Results of event-based address-event stereo matching

For the evaluation of the event-based stereo vision algorithm, a specific tool was used which is shown in Fig. 15. It is called Event-Editor and gives the opportunity to generate synthetic stereo data which allows a verification of the algorithm because of available ground truth information.

Before the evaluation, the parameters for the synthetic data have to be set. The detection range is between $6m$ and $5m$ what gives, considering the system configuration, a disparity range of 15 to 20. In the evaluation phase we considered a higher range of 35, which evaluates the capability of the algorithm to detect closer objects as well. The simplified model of the silicon retina used for the generation of synthetic silicon retina data is a suitable approximation of the real silicon retina sensor for the algorithm evaluation.
In Fig.21, the evaluation results of the algorithm are shown. Different aggregation window

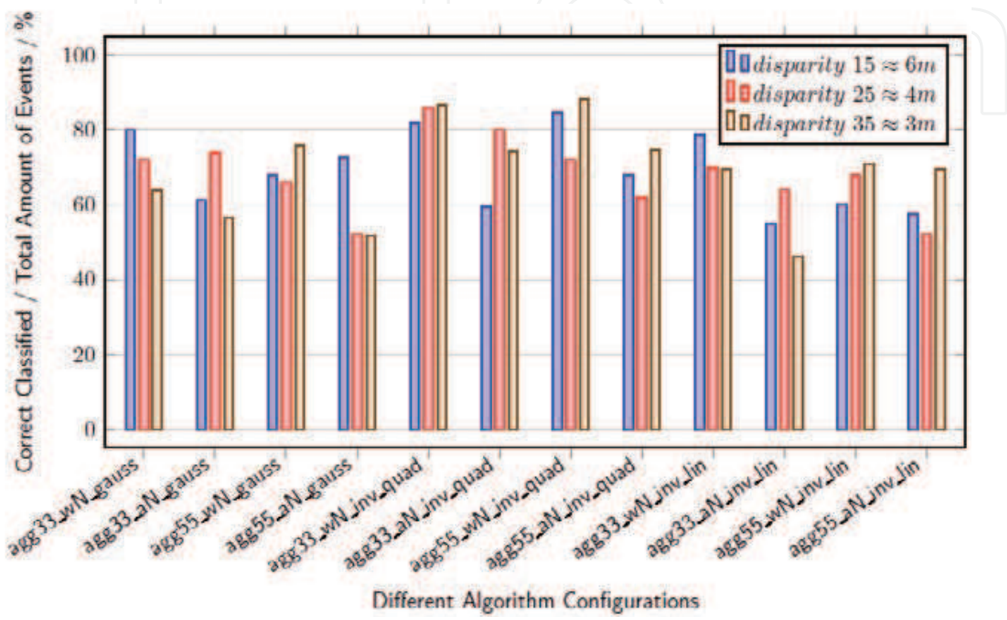


Fig. 21. Experimental results of the event-based stereo matching algorithm with different evaluation and algorithm settings

sizes, noise amounts, and weighting functions are compared. The value, in percent, plotted on the y-axis, gives the information of how many disparities were calculated correctly in relation to the total amount of events in the current matching step. The evaluation was carried out in three different disparity distances, depicted with the three colors in Fig. 21. The three parts of the values, plotted on the x-axis and divided by an underscore describe the algorithm evaluation settings. The first part shows the aggregation window size (3×3 or 5×5), the second part shows if there was a noise signal added (aN) or not (wN), and the last part describes the weighting function used. The results show that the quality of disparity calculation is independent of the evaluated distances. As expected, added noise effects the rate of correct classified disparity values in comparison to noise free data. The aggregation window size has only a small impact but especially when noise is added, the amount of correct classified disparities increase if the window size is enlarged. The highest influence has the weighting function (Gaussian function, an inverse quadratic function or an inverse linear function). In case of a linear quadratic function, the best results of more than 80% correct classified disparities could be achieved.

6. Conclusion and future work

The silicon retina is a new type of image sensor which opens up a new kind of sensor field next to conventional monochrome/color sensors and emulates, in terms of principle of operation,

the human eye better than the other sensors. This new type of sensor used in a stereo setup can be used for extracting depth information.
To do this, the correspondence problem has to be solved, but the silicon retina has a new data interface and therefore, novel approaches of stereo matching are needed. In a first step, the data of the silicon retina were adapted for stereo matching algorithms built for conventional image data. This method was not accurate enough and does not use the full potential of the silicon retina technology. Due to this fact, a new algorithm approach was implemented which uses the data of the silicon retina directly without conversion and exploits the novelty of the sensor.
In this approach, the time was used as the primary matching criterion to find corresponding events from the left and right camera. The results showed that the event-based stereo matching exploits the advantages of the silicon retina in comparison to the frame-based approaches. Even so, the results of the event-based approach needs an improvement in accuracy and confidence, which means that the event-based approach has to be enhanced. For this reason, new algorithm approaches which improve, in combination with the existing algorithms, the results of the stereo matching are implemented, or novel new algorithm approaches will be designed which achieve better results and significantly better accuracy.
The algorithm improvements may increase the accuracy and quality of the results, but for extensive algorithmic calculations also an adequate hardware performance is necessary. Therefore, new ways of hardware implementations are considered which can handle the amount of data and process the results in real-time. After the implementation of algorithms into a PC-based solution and migration to a optimized DSP-multicore solution, the event-based matching approach will be integrated into a FPGA, expecting not only a reduction of the latency and an improvement of the throughput of the system, but also an enhancement of the scalability and the overall energy efficiency of the system. A further advantage is that a FPGA-based platform facilitates fast prototyping and a high degree of flexibility because of reconfigurability. Hence the system can easily be adapted to changes in requirements, e.g. sensor size, timing constraints and communication interfaces.

7. Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement n° ICT-216049 (ADOSE).

8. References

Banks, J., Bennamoun, M. & Corke, P. (1997). Non-parametric techniques for fast and robust stereo matching, *Proceedings of the IEEE Region 10th Annual Conference on Speech and Image Technologies for Computing and Telecommunications*, Brisbane/Australia, pp. 365–368.

Banks, J., Bennamoun, M., Kubik, K. & Corke, P. (1999). A constraint to improve the reliability of stereo matching using the rank transform, *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Phoenix/USA, pp. 3321–3324.

Belbachir, A. (2010). *Smart Cameras*, Springer New York Dordrecht Heidelberg London.

Boahen, K. (2000). Point-to-point connectivity between neuromorphic chips using address events, *In IEEE Trans. on Circuits and Systems II* 47(5): 416–433.

Bouguet, J. (2008). Camera calibration toolbox for matlab, Publised in the Internet.

Computer Vision Research Group/Department of Electrical Engineering/California Institute of Technology - www.vision.caltech.edu/bouguetj/calib_doc/index.html.

Burger, W. & Burge, M. (2005). *Digital Image Processing An Algorithmic Introduction using JAVA - First Edition*, Springer-Science/Business Media LLC.

Fawcett, T. (2004). Roc graphs : Notes and practical considerations for researchers, *Technical Report HP Labratories* .

Fawcett, T. (2006). An introduction to roc analysis, *Pattern Recognition Letters* 27(8): 861–874.

Fukushima, K., Yamaguchi, Y., Yasuda, M. & Nagata, S. (1970). An electronic model of the retina, *Proceedings of the IEEE* 58(12): 1950–1951.

Gonzales, R. & Woods, R. (2002). *Digital Image Processing - Second Edition*, Prentice Hall/Pearson Education International.

Häflinger, P. & Bergh, F. (2002). An integrated circuit computing shift in stereo pictures using time domain spike signals, *Proceedings of the Conferenece NORCHIP*, Kopenhagen/Denmark.

Lichtsteiner, P., Kramer, J. & Delbruck, T. (2004). Improved on/off temporally differetiating address-event imager, *Proceedings of the 11th IEEE International on Electronics, Circuits and Systems*, TelAviv/Israel.

Lichtsteiner, P., Posch, C. & Delbruck, T. (2006). A 128×128 120db 30mw asynchronous vision sensor that responds to relative intensity change, *Proceedings of the IEEE International Solid-State Circuits Conference*, SanFrancisco/USA.

Mahowald, M. (1992). *VLSI analogs of neuronal visual processing: a synthesis of form and function*, Phd-thesis, California Institute of Technology.

Mahowald, M. & Mead, C. (1989). Silicon retina, *Analog VLSI and Neural Systems* pp. 257–278.

Mead, C. & Mahowald, M. (1988). A silicon model of early visual processing, *Neural Networks Journal* 1(1): 91–97.

Milosevic, N., Schraml, S. & Schön, P. (2007). Smartcam for real-time stereo vision - address-event based stereo vision, *Proceedings of Image Understanding / Motion, Tracking and Stereo Vision; INSTICC Inst. f. systems and technologies of information, control and communication; INSTICC Press*, Barcelona/Spain, pp. 466–471.

Mortara, A. (1998). A pulsed communication / computation framework for analog vlsi perceptive systems, *Neuromorphic Systems Engineering* 447(3): 201–215.

Porter, R. & Bergmann, N. (1997). A generic implementation framework for fpga based stereo matching, *Proceedings of the IEEE Region 10th Annual Conference on Speech and Image Technologies for Computer and Telecommunications*, Brisbane/Australia, pp. 461–464.

Provost, F. & Fawcett, T. (2001). Robust classification for imprecise environments, *Machine Learning* 42(3): 203–231.

Schreer, O. (2005). *Stereoanalyse und Bildsynthese*, Springer Verlag Berlin Heidelberg.

Shi, J. & Tomasi, C. (1994). Good features to track, *Proceedings of the IEEE Computer Vision and Pattern Recognition Conference*, Seattle/USA, pp. 593–600.

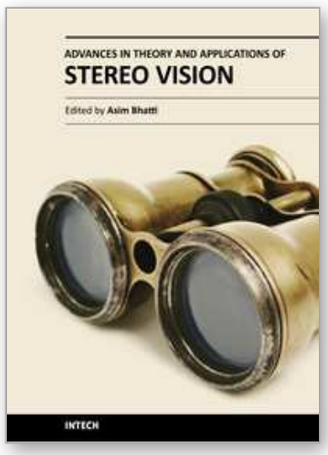
Sivilotti, M. (1991). *Wiring consideration in analog vlsi systems with application to field programmable networks*, Phd-thesis, California Institute of Technology.

Tang, B., AitBoudaoud, D., Matuszewski, B. & Shark, L. (2006). An efficient feature based matching algorithm for stereo images, *roceedings of the IEEE Geometric Modeling and Imaging Conference*, London/UK, pp. 195–202.

World Wide Web Consortium (W3C) (1995). Vrml virtual reality modeling language, http://www.w3.org/MarkUp/VRML.

Zabih, R. & Woodfill, J. (1994). Non-parametic local transforms for computing visual correspondence, *Proceedings of the third European conference on Computer Vision*, Stockholm/Sweden, pp. 151–158.

Zhang, Z. (2002). A flexible new technique for camera calibration, *Technical Report MSRTR9-71*, Microsoft Research.



Advances in Theory and Applications of Stereo Vision
Edited by Dr Asim Bhatti

ISBN 978-953-307-516-7
Hard cover, 352 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011

The book presents a wide range of innovative research ideas and current trends in stereo vision. The topics covered in this book encapsulate research trends from fundamental theoretical aspects of robust stereo correspondence estimation to the establishment of novel and robust algorithms as well as applications in a wide range of disciplines. Particularly interesting theoretical trends presented in this book involve the exploitation of the evolutionary approach, wavelets and multiwavelet theories, Markov random fields and fuzzy sets in addressing the correspondence estimation problem. Novel algorithms utilizing inspiration from biological systems (such as the silicon retina imager and fish eye) and nature (through the exploitation of the refractive index of liquids) make this book an interesting compilation of current research ideas.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jürgen Kogler, Christoph Sulzbachner, Martin Humenberger and Florian Eibensteiner (2011). Address-Event Based Stereo Vision with Bio-Inspired Silicon Retina Imagers, Advances in Theory and Applications of Stereo Vision, Dr Asim Bhatti (Ed.), ISBN: 978-953-307-516-7, InTech, Available from: http://www.intechopen.com/books/advances-in-theory-and-applications-of-stereo-vision/address-event-based-stereo-vision-with-bio-inspired-silicon-retina-imagers

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen