EMBEDDED SMART CAMERA FOR HIGH SPEED VISION

M. Litzenberger, A.N. Belbachir, P.Schön and C.Posch

Austrian Research Centers GmbH – ARC
Donau City Strasse 1, 1220 Vienna, Austria
martin.litzenberger@arcs.ac.at

ABSTRACT

The architecture and prototype applications of an embedded vision system containing a neuromorphic temporal contrast vision sensor and a DSP are presented. The asynchronous vision sensor completely suppresses image data redundancy and encodes visual information in sparse Address-Event-Representation (AER) data. Due to the efficient data pre-processing on the focal plane, the sensor delivers high temporal resolution data at a low data rate. Hence, a compact embedded vision system using a low-cost, low-power digital signal processor can be realized. The one millisecond timestamp resolution of the AER data stream allows to acquire and process motion trajectories of fast moving objects in the visual scene. Various post processing algorithms, such as object tracking, vehicle speed measurement and object classification have been implemented on the presented embedded platform. The system's low data rate output, low power operation and Ethernet connectivity make it ideal for use in distributed sensor networks. Results from traffic-monitoring and object tracking applications are presented.

Index Terms— real time embedded vision system, temporal contrast imager, traffic monitoring, object tracking

1. INTRODUCTION

In the past, object tracking applications for different purposes [9] have emerged from an academically pattern recognition problem [2]. Such object tracking applications include pedestrian and vehicle tracking [14] and surveillance [10], bubbles tracking [5] and soccer players tracking [8], but are not exclusively limited to these applications.

Vehicle tracking is one of the most spread vision applications due to the increasing traffic and safety demand. The majority of the existing vehicle-tracking systems are based on the video image processing. Many video tracking systems identify vehicles by virtue of their motion. In cases where objects are moving continuously past the sensing camera, the motion segmentation techniques are fast and robust. Unfortunately, in cases where the sensing camera observes a largely stationary vehicle queue, e.g. in front of traffic lights, motion estimation based systems mostly fail to follow-up the objects.

Previous research in automotive tracking systems has not been completely successful. In [9] Kalman-Snakes technique is used to provide automobile contours after initial motion segmentation step. These contours are used for tracking purpose. Block matching techniques to find optical flow have been another approach [3], which is combined with a priori knowledge of the road geometry to handle stationary vehicles. Isolation of foreground objects “blobs” by a background estimation technique has been reported in [15]. Principal component analysis is used afterwards to classify the blobs and estimate their orientation.

In addition to the limitation of the developed tracking methods in terms of performance, video systems usually produce a huge amount of data that likely saturate any computational unit responsible for data processing. Thus, real-time object tracking based on video data processing requires large computational effort and is consequently done on high-performance computer platforms. As a consequence, the design of video-tracking systems with embedded real-time applications, where the algorithms are implemented in Digital Signal Processor (DSP) is a challenging task.

A further weakness of video detection is the limitation of conventional camera systems to operate under wide dynamic range lighting, which is typical for outdoor applications. Therefore, real-time video-based tracking applications are mostly constrained with limited resources at the price of the optimal performance.

There also exist other systems for vehicle tracking using the support of satellites for the vehicle follow-up [1]. Although those systems might be efficient, they are very costly as they are requiring a sensor, a communication link and a workstation mounted on each vehicle.

This paper presents a low-cost smart camera system capable of high speed vision for embedded applications. The system has been developed using an asynchronous temporal contrast vision sensor [12] and the algorithms for

1-4244-1354-0/07/$25.00 ©2007 IEEE

object tracking and traffic data acquisition have been implemented on a DSP. The system has already been demonstrated its performance for vehicle speed estimation and vehicle counting [13][4]. This work presents the embedded smart camera and its applications to traffic monitoring and object tracking.

The paper is structured as follows. In section 2, the smart camera architecture is described. The applications and experimental results are described in section 3. Section 4 gives a conclusion.

2. SMART CAMERA ACHITECTURE

2.1. Vision Chip

In contrast to traditional CCD or CMOS imagers that encode image irradiance and produce constant data volume at a fixed frame rate, irrespective of scene activity, the asynchronous vision sensor used in the presented smart vision system contains an 128×128 pixel array of autonomous, self-signaling pixels which individually respond in real-time to relative changes in light intensity by placing their address on an asynchronous arbitrated bus. Pixels that are not stimulated by a change in illumination are not triggered; hence static scenes produce no output. Because there is no pixel readout clock, no time quantization takes place at this point.

The sensor operates largely independent of scene illumination, directly encodes object reflectance, and greatly reduces redundancy while preserving precise timing information. Because output bandwidth is automatically dedicated to 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 to a DSP via an asynchronous bus. The pixel locations in the imager array are encoded in the event data that are reflected as 14 bits coordinates in the image space in the form of Address-Events (AE) [12]. Additional to the pixel address the polarity of the contrast change is encoded in a 15^{th} bit as "on" of "off" event, representing a change from dark to bright or vice-versa.

The high dynamic range of the photosensitive element (>120dB or 6 decades) makes the imager ideal for applications with uncontrolled light conditions. Figure 1 shows typical AE data of a highway monitoring scene acquired with the transient imager and a still video image for comparison. In order to visualize the AE data, events have been collected for a 50 millisecond interval and rendered like a video frame. The different gray shadings encode pixel activity per unit time, with black being the highest activity of "off" events and white indicating high "on" event activity. Note that, due the sensitivity of sensor to even small contrast changes all vehicles have almost the same representation in the data, independent of vehicle color. The most effective way of processing address-event data however does not make use of video framing but takes advantage of the efficient coding of the visual information by directly processing the asynchronous spatio-temporal information on moving objects contained in the AE data stream.

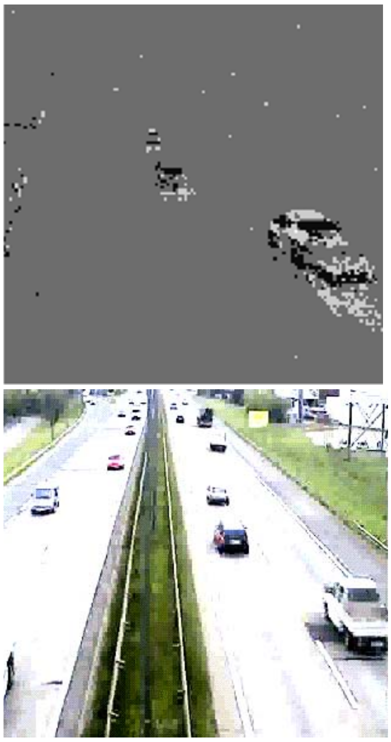

Fig.1 Example of asynchronous AE data represented as a video frame (top) and a video still image of the same scene (bottom).

2.2. Embedded System

Figure 2 depicts the general architecture of the smart vision system, which comprises the temporal contrast imager, a First-In, First-Out (FIFO) buffer memory and a simple low-cost and low-power fixed-point DSP (Analog Devices Blackfin® BF537). It has a maximum frequency of 600 MHz, 128 KB internal memory and 32 MB external SDRAM memory. This limited memory resource might be far-too low for the data processing of any high-resolution video system as it does not fit the traditional video processing needs. The imager and the DSP consume in total roughly 2.5 W of electrical power. The presented embedded system has been specifically designed to process asynchronous address-event data in an efficient way. The location (address) of the event generating pixel within the array is transmitted implementing a simple 4-phase handshake protocol. Following a 'data available' request from the sensor, the 15-bit event is taken from the asynchronous bit-parallel address-event bus, stored in a

FIFO and an acknowledge signal is sent back to the sensor. The FIFO is placed between the imager sensors and the DSP to cope with peaks of AE activity and is capable of handling up to 40 MHz memory access frequency. A process running on the DSP buffers the data for further processing as long as the FIFO EMPTY signal is not active. In the processing stage, every AE received by the DSP is labeled by attaching the processor clock ticks with 1 millisecond precision as a timestamp. These data are the basis for traffic data acquisition and object tracking algorithms.

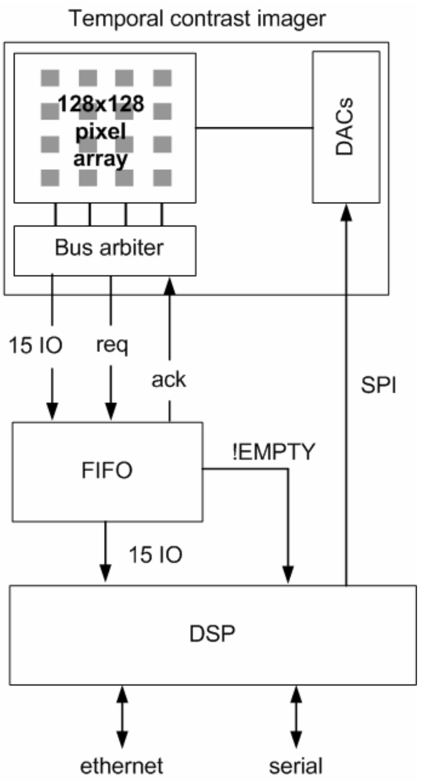

Fig. 2. Schematics of the smart camera architecture.

The DSP also controls on chip digital-analog-converters that generate the internal bias voltages for the imager via a serial interface. These control bias voltages allow for an on-the-fly adjustment of functional parameters like e.g. the contrast thresholds. Computation results such as vehicle counts or object tracks are sent via Ethernet to a host computer. Alternatively, the Ethernet connectivity allows connecting to other smart camera systems attached to a network of sensors and exchanging information about tracked objects together with geo-location information. The processing power remaining after completion of tracking subroutines would be sufficient to process this information to e.g. track objects from one sensors detection range to another sensors location. The full processing comprises the AE acquisition and time stamping, object clustering and tracking. Other AE processing algorithms for the vehicle speed estimation [7],[13] counting [4] and classification have also been implemented on this DSP.

Due to its low power consumption the system is suitable for autonomous solar or battery supply operation. Figure 3 shows a photograph of the compact smart camera system including the optics, the sensor and processor board with a total size of 7×7×7 cm. For a detailed specification of the smart camera system refer to the datasheet [16].

Fig. 3. Photograph of the smart camera system.

3. APPLICATIONS AND RESULTS

3.1. Traffic Data Acquisition

The smart camera has been installed at a highway test site for monitoring 4 lanes of traffic simultaneously. Single vehicle detection and data acquisition comprising time stamp, lane number, vehicle speed and length and classification into two classes (cars and trucks) have been implemented in the firmware of the system. The vehicle detection is based on monitoring the event activity in pre-defined regions-of-interest (ROI) corresponding to the highway lanes [4]. The speed and length estimations benefit from the high temporal resolution and the continuous information on the vehicle tracks found in AE data [13]. Length measurement is realized by measuring the presence of the vehicle within the ROI with millisecond precision and correcting it by the ROI’s length, optical/geometric considerations derived from the known viewing angle. A rough classification of vehicle type is based on this length estimation.

Figure 4 shows representative results of traffic data acquisition accumulated from two lanes at a test site with a 80km/h speed limit, over the course of two days.

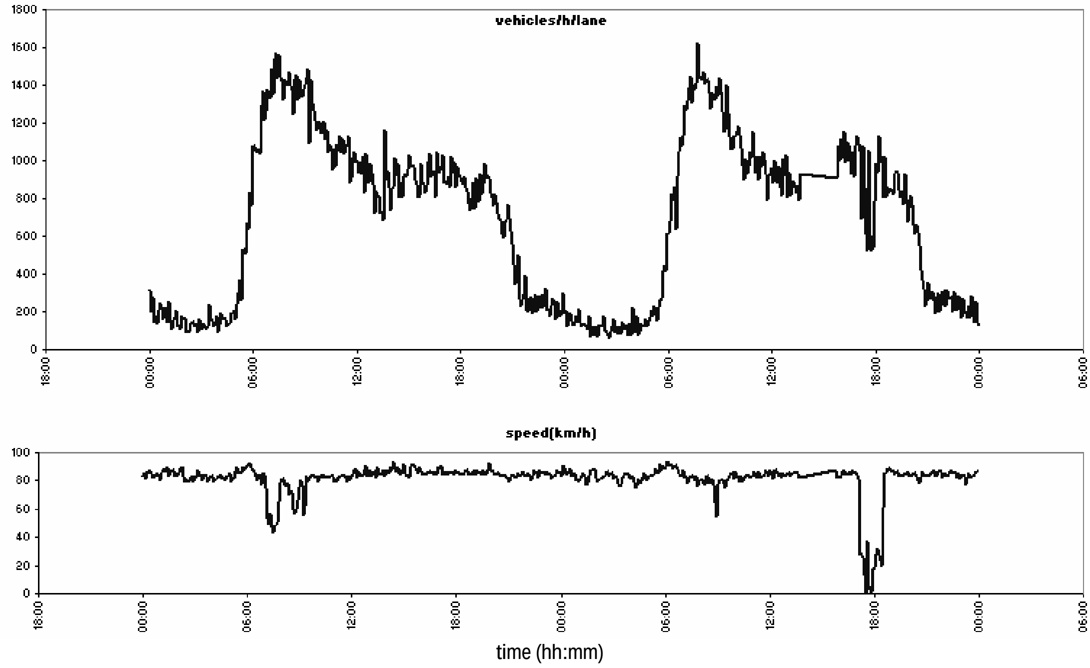

Fig. 4. Vehicle quantity and average speeds given in 5 minutes time intervals for two days of traffic data acquisition.

The traffic volume (top) and the average speed (bottom) are shown in 5 minutes resolution, scaled to vehicle count per hour per lane. As city inbound traffic flow is monitored, traffic peaks around 7am.Traffic congestions, recognizable by distinct drops in average speed, are observed in the morning of the first day and the evening of the second day.

The vehicle counting error has been evaluated against 6 hours of manually annotated video recordings containing approx. 7000 vehicles. Speed estimation errors in the speed range from 20 to 120 km/h been evaluated by measurement with calibrated light barriers as reference [13].

The connectivity of the smart camera system would allow to exchange information about traffic flow situation between different cameras installed at several locations along a route, thus enabling a network of smart cameras to assess the overall traffic situation along a certain route section. The differences of vehicles quantities measured at several successive camera locations e.g. allow to derive the vehicle density within a certain road section by the camera network itself, which is an important parameter in traffic telematics.

3.2. Object Tracking

The tracking algorithm implemented in the smart camera firmware uses a continuous clustering of AE's and tracking of clusters. The algorithm processes each event on-the-fly, as it is received without the need to buffer large amounts of AE data (little memory consumption) [11]. This is of special importance when using low-cost and low memory resource systems.

The simple tracking algorithm is outlined as follows: Each new event received by the DSP is assigned to an existing cluster if it fulfills a distance criterion, and the AE' pixel coordinate is used to update (shift) this clusters center position for tracking [6]. If an event received can not be assigned to an existing cluster a new cluster is seeded and kept hidden as long as the number of events attributed to it lies below a threshold. If an existing cluster did not receive enough events during a period of time it is deleted.

The algorithm fully exploits the one millisecond time resolution of the AE data and consumes only a small amount of memory as only the cluster list has to be kept in memory. Trials showed that a cluster list of ~20 clusters is sufficient to solve most of the test scenarios. As only a few variables are used to describe the features of a cluster, 2 kB of memory is sufficient for the cluster list. The computational complexity is moderate. However, event-address to cluster-center distances have to be computed several times for each new event. Using quadratic clusters instead of circular ones can further decrease the computational effort. The features of each cluster have to be updated once for every event.

Fig. 5 shows a demonstration of the algorithm on simulated AE data for a person tracking application. The AE data have been simulated from a video sequence with a resolution of 140×180 pixels that is close to the 128×128 imager. On the left part, images from a 4 second video sequence showing people leaving and elevator cabin have been extracted for illustration. On the right hand side, the simulated AE data of the scene including the tracking results is provided. It shows the different persons locations indicated by circles, a unique ID number identifying the object and an arrow indicating the direction and speed of movement. As an example, cluster with ID 198 has been correctly tracked throughout the sequence, while cluster with ID 227 was produced by a shadow effect and disappeared in the next sequence.

The interconnection of several smart cameras with overlapping observation areas allows to continuously track an identified object on its way e.g. through a building by sending the object and track information (motion vectors) to neighboring smart cameras.

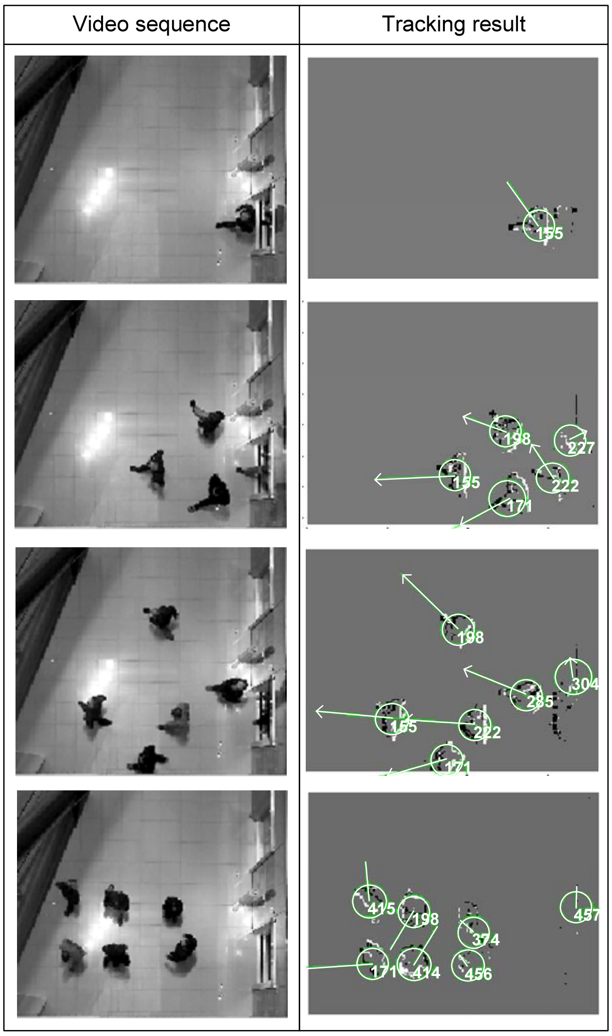

Fig. 5. People tracking example. Video frames and AE data frame representation with tracked object ID numbers and velocity vectors on top.

Table 1 gives an overview of key performance parameters for the presented traffic monitoring and the object tracking applications.

4. CONCLUSION

An embedded smart camera system capable of traffic monitoring and high speed object tracking has been presented. The signal processing in the system strongly benefits from the sparse data and high temporal resolution delivered by the temporal contrast vision sensor used. Traffic parameter acquisition and object tracking are performed in real time by a low-cost, low-power DSP. Representative traffic and object tracking data examples have been presented. The system is equipped with an Ethernet connection for data transfer which enables to connect multiple cameras via TCP/UDP protocol to form a network of smart cameras.

Table 1: Overview of traffic data acquisition and object tracking key performance values.

Traffic data acquisition	
Lanes monitored	up to 4
Speed estimation error	< 3%
Counting error	1h interval: < 3%
	3min interval: <10%

Object tracking	
Temporal resolution	1 ms
Objects tracked	up to 20
Data memory consumption	2 kB for cluster list

5. ACKNOWLEDGEMENTS

The authors acknowledge Patrik Lichtsteiner and Tobi Delbrück of INI-ETH/UNI Zürich for their work on the temporal contrast imager. We are thankful to the Austrian highway authority ASFiNAG-VTG and Bernhard Schalko for realization of the highway test site.

6. REFERENCES

[1] L. Alexander and M. Donath, “Differential GPS based control of a heavy vehicle,” *International Conference on Intelligent Transportation Systems*, 1999.

[2] Y. Anzai, “Pattern Recognition & Machine Learning,” Morgan Kaufmann, 1992.

[3] F. Bartolini, V. Capellini, and C. Giani, “Motion Estimation and Tracking for Urban Traffic Monitoring,” *International Conference on Image Processing*, volume 3, 1996

[4] D. Bauer, P. Bühler, N. Donath, H. Garn, B. Kohn, M. Korber, M. Litzenberger, J. Meser, C. Posch and P. Schön “Embedded Vehicle Counting System with ‘Silicon Retina’

Optical Sensor” *Workshop on information Optics* 2006, Toledo, Spain, June 5-7, 2006.
[5] D.C. Cheng and H. Burkhardt, “Bubble Tracking in Image Sequences,” *International Journal of Thermal Sciences*, Volume 42, Number 7, pp. 647-655(9), July 2003.
[6] D. Comaniciu and V. Ramesh, “Mean Shift and Optimal Prediction for Efficient Object Tracking,” *International Conference on Image processing*, vol.3, p.70-73, 2000.
[7] Bauer, A. N. Belbachir, N. Donath, et al., “Embedded Vehicle Speed Estimation System Using an Asynchronous Temporal Contrast Vision Sensor,” *EURASIP Journal on Embedded Systems*, vol. 2007, Article ID 82174, 12 pages, 2007. doi:10.1155/2007/82174
[8] P. Figueroa, N, Leite, R.M.L. Barros, I. Cohen and G. Medioni, “Tracking soccer players using the graph representation,” *Proceedings of the 17th International Conference on Pattern Recognition*, UK, 2004.
[9] D. Koller, J. Weber and J. Malik, “Towards Realtime Visual based Tracking in Cluttered Traffic Scenes”, *Proc. of the Intelligent Vehicles Symposium*, Paris, France, October 1994.
[10] D. Li, K.D. Wong, Y.H. Hu and A.M. Sayeed, “Detection, Classification and Tracking of Targets in Distributed Sensor Networks”, *IEEE Signal Processing Magazine*, 19(2):17--30, March 2002.
[11] M. Litzenberger, C. Posch, D. Bauer, A.N. Belbachir, P. Schön, B. Kohn, and H. Garn, "Embedded Vision System for Real-Time Object Tracking Using an Asynchronous Transient Vision Sensors". *12th IEEE Workshop on Digital Signal Processing and Signal Processing Education DSP/SPE 2006*, Wyoming, USA September 2006, in John Pierre and Cam Wright (Eds), pp. 173-178.
[12] P. Lichtsteiner, C. Posch and T. Delbruck, “A 128x128 120dB 30mW Asynchronous Vision Sensor that Responds to Relative Intensity Change,” in *IEEE International Solid-State Circuits Conference* (ISSCC2006), San Francisco, USA, Februay 2006.
[13] M. Litzenberger, A.N. Belbachir, N. Donath, G. Gritsch, H. Garn, B. Kohn, C. Posch and S. Schraml “Estimation of Vehicle Speed Based on Asynchronous Data from a Silicon Retina Optical Sensor,” *IEEE Conference on Intelligent Transportation Systems*, Canada, September 2006.
[14] N.T. Siebel and S.J. Maybank , “ Real-Time Tracking of Pedestrians and Vehicles,” *Proceedings 2nd IEEE Int. Workshop on PETS*, Kauai, Hawaii, USA, December 9 2001
[15] H. Veeraraghaven and O. Masoud. N. Papanikolopoulos, “Managing Suburban Intersections Through Sensing,” *Technical Report at Intelligent Transportation Systems Institute University of Minnesota* December 2002
[16] www.smart-systems.at/products/products_video_tds_en.html