CNNA 2016, August 23-25, 2016, Dresden, Germany

A Real-time High Dynamic Range Vision System with Tone Mapping for Automotive Applications

Julien N. P. Martel *, Lorenz K. Müller *, Stephen J. Carey † and Piotr Dudek †
Email: jmartel@ini.ethz.ch lorenz@ini.ethz.ch stephen.carey@manchester.ac.uk p.dudek@manchester.ac.uk
* Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich 8057, Switzerland
† School of Electrical & Electronic Engineering, The University of Manchester, Manchester M13 9PL, United-Kingdom

***Abstract*—Automated and autonomously driving vehicles capture information from the environment using vision sensors (amongst others) and process it in order to detect lane markings, pedestrians, other vehicles, or to read traffic signs. A crucial issue arises when visual sensors are used in real-world situations: the observed scenes are subject to a wide range of lighting conditions. This difficulty adds to the computational complexity needed to reliably process visual information at low-latencies. We propose a system based on a programmable vision-chip that can process and output frames in real-time in a wide range of lighting conditions and intra-scene light variations using a novel parallel High Dynamic Range (HDR) tone mapping algorithm.**

I. INTRODUCTION

Control technologies developed for automated and autonomously driving vehicles rely on a variety of sensors such as lasers, radars, GPS or cameras. Among these, visual sensors hold a prominent place: they are relatively inexpensive, they can see at fairly long distances at daytime as compared to LIDAR and radar, they can see colour in the visible light spectrum, their resolution can be high, and they generally have no moving parts. Despite these many advantages, visual information sensing comes at the cost of a processing burden to reliably extract relevant features from uncertain environments such that they can be used in later processing stages. For practical purposes, this processing must be done at low-latency to achieve real-time performance allowing a safe and responsive control of the vehicle. Another major issue is to design vision systems that can operate in a wide range of lighting conditions. Considerable light variations can exist within a scene: caused by the shadows cast by trees, highly reflective surfaces used in road signs or even by radiation actively emitted in the case of traffic lights. In addition, substantially different light settings exist across scenes captured under bright-light at daytime on a sunny or cloudy day, when going through underpasses or tunnels with controlled illumination or at night with very limited lighting.

The first computational issue can be addressed using appropriate processing devices. Indeed we can equip automated vehicles with more processing power; the solution we propose in this work takes a rather different approach in that we use a vision-chip that couples sensing and processing using in-pixel processing circuitry. This has two advantages: first we move processing capabilities to the sensing device thus reducing the computational load on the processing core, secondly we can design new algorithms taking full advantage of this collocated sensing/computing architecture avoiding the information transfer bottleneck inherent to classical strategies where the information acquisition and its processing are separated.

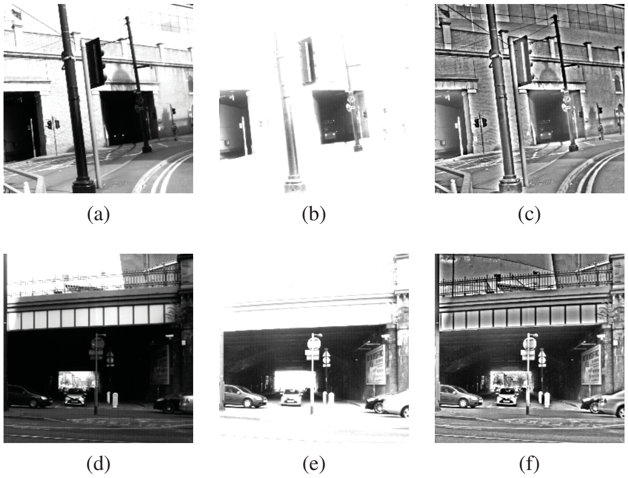


Fig. 1. Examples of images output by our HDR tone mapping vision system compared with under and over exposed images: (a) (d) are under exposed (b) (e) are over exposed (c) (f) are outputs of the system with the HDR tone mapping algorithm running

To address the second issue we suggest a new vision system to tackle the problem of lighting variations in automotive applications. Our system consists of a vision-chip equipped with an algorithm that produces tone mapped images containing High Dynamic Range (HDR) information, operating with virtually no overhead in terms of processing time (the computations are carried out during light integration in the image sensor). HDR imaging deals with the capture of light-intensity in scenes that can exhibit several orders of magnitude of illuminance. Tone mapping on the other hand, allows the efficient handling and representation of such data on a lower range (for e.g. a low number of bits) while perceptually preserving information from the many orders of magnitude that the illuminance spans. Our system addresses both problems jointly, hence the output frames are “conventional” in their format, i.e they do not require a higher number of bits for storage, and can be further processed by any classical artificial vision algorithm.

II. HDR IMAGING AND TONE MAPPING ON A VISION-CHIP

Principles

In order to capture a HDR image, a vision system must be able to deal with pixels subject to different radiances in different parts of the scene. The irradiance at a pixel is low in very dark regions of the scene and very high in bright parts

---

This work was funded by SNF grants 143947 and CRSII2_160756, and EPSRC grant EP/M019284/1

corresponding to the low and high values they would ordinarily take. Because the range of values the pixel can record is limited (typically to 8-bits), the principle of the system we design is to handle these intra-scene variations by adjusting the exposure time for each pixel individually. The range of values our system reports is not extended, (it is kept the same, i.e. 8-bits), but by varying the pixel integration times, the same recorded value (e.g. 128) is reached at different times on the array, and thus corresponds to different irradiance values at each pixel, producing a tone-mapped image.

A detailed description of the algorithm can be found in [1]. While the light integration is taking place, we check a criterion based on the intensities in pixel's neighbourhoods to decide whether the pixel should continue to integrate or not. If the pixel decides it should stop then it activates a flag that can be thought of as masking the pixel during the next operations so that it keeps its latest value.

During the whole duration of the integration we check this criterion as often as possible, that is every δt, which is the time to compute and evaluate the criterion. We update the pixels' activity flags accordingly. Because δt is small and the criterion is evaluated pixel-wise with overlapping neighbourhood it guarantees the resulting HDR image to be locally well-exposed with smoothly varying integration times.

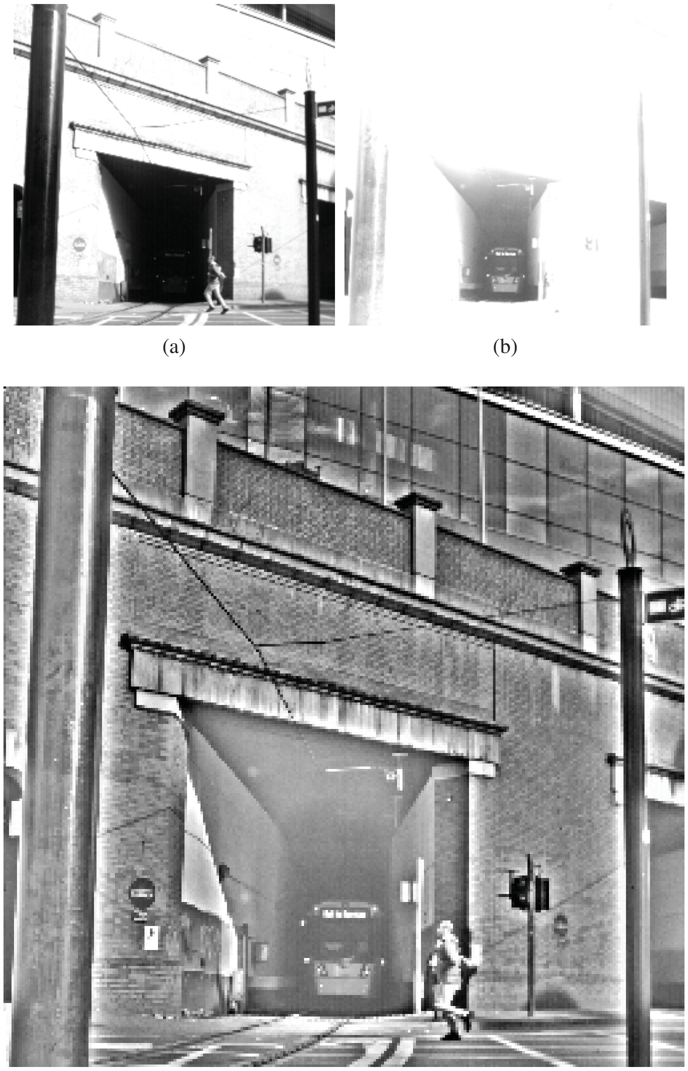
The pixel-wise integration on the vision chip can be as short as $\delta t = 2$µs and as long as several seconds. A longer integration time allows the non-irradiant pixels to get more light, in practice this integration time upper bound – specified by the maximal integration time a pixel can take – is chosen so that the system runs at the desired frame rate (e.g. the maximum integration time is about $40\mathrm{ms}$ to run at 25FPS).

To summarise, pixel-wise integration ranging from very short to long times are used to achieve HDR, while the resulting pixel values are tone-mapped, since they are based on the light intensity at the pixel with respect to its local neighbourhood only – pixels across the array with the same value have different irradiance, while preserving local contrast.

Implementation on the SCAMP-5 vision-chip

We implemented a system with our HDR imaging and tone mapping algorithm using the SCAMP-5 vision chip [2]. The device comprises an array of 256×256 pixels collocated with their processing elements (PEs) organised as a SIMD processor array. A PE on SCAMP-5 includes a small ALU (Arithmetic and Logic Unit) along with 7 analogue registers and 13 single bit digital registers. A diffusion network with a tunable resistance allows fast "blurring" operations and a flexible addressing architecture provides block summation as well as analogue and digital readouts.

Testing the system


(a) (b)

(c)

Fig. 2. An example where a pedestrian can be seen in the under-exposed setting in (a) but not in an over-exposed setting (b) which helps to see the coming tram. Only with the HDR frame (c) both the pedestrian and the tram can be seen.

The vision system we propose works in real-time ≥ 25FPS and produces HDR frames that are tone mapped so that they can be used in further vision processing stages if needed, for instance in recognition or detection applications. The further processing could be also done on the same vision chip, and/or using external devices. The whole system (USB interface, system controller and vision chip) runs at under 1.2W. It has been tested and operates across different light settings from 25lx to more than 25klx and with a scene contrast ratio of about $1 : 10^6$. Note that the SCAMP-5 vision-chip does not output colour images. Also the actual frame rate is limited in a low-light setting because of the long exposure time needed due to the relatively low sensitivity and small fill factor of the photodiode on this vision-chip [2]. This is not a restriction of the HDR tone mapping algorithm but rather a limitation in the design of the current vision-chip prototype.

III. CONCLUSION

We presented a vision system for simultaneous HDR imaging and tone-mapping. The proposed solution has near-zero latency, since the algorithm operates concurrently with light integration, producing a tone-mapped image for the current frame in time determined by the maximal desired exposure time.

REFERENCES

[1] J. N. P. Martel *et al.*, "Parallel hdr tone mapping and auto-focus on a cellular processor array vision chip," in *Proc. of the IEEE Internat. Symp. on Circuits and Systems, ISCAS'16*, 2016.

[2] S. J. Carey *et al.*, "A 100,000 fps vision sensor with embedded 535 GOPS/W 256×256 SIMD processor array," in *Proc. of the VLSI Circuits Symp.'13*, 2013, pp. C182–C183.