Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 00 (2014) 000–000

Procedia
Computer Science

www.elsevier.com/locate/procedia

6th International conference on Intelligent Human Computer Interaction, IHCI 2014

Logarithmic Image Sensor For Wide Dynamic Range Stereo Vision System

Christian Bouvier, Yang Ni

New Imaging Technologies SA, 1 Impasse de la Noisette, BP 426, 91370 Verrieres le Buisson France
Tel: +33 (0)1 64 47 88 58 christian.bouvier@new-imaging-technologies.com, yang.ni@new-imaging-technologies.com

Abstract

Stereo vision is the most universal way to get 3D information passively. The fast progress of digital image processing hardware makes this computation approach realizable now. It is well known that stereo vision matches 2 or more images of a scene, taken by image sensors from different points of view. Any information loss, even partially in those images, will drastically reduce the precision and reliability of this approach. In this paper, we introduce a stereo vision system designed for depth sensing that relies on logarithmic image sensor technology.

The hardware is based on dual logarithmic sensors controlled and synchronized at pixel level. This dual sensor module provides high quality contrast indexed images of a scene. This contrast indexed sensing capability can be conserved over more than 140dB without any explicit sensor control and without delay. It can accommodate not only highly non-uniform illumination, specular reflections but also fast temporal illumination changes.

Peer-review under responsibility of the Scientific Committee of IHCI 2014.

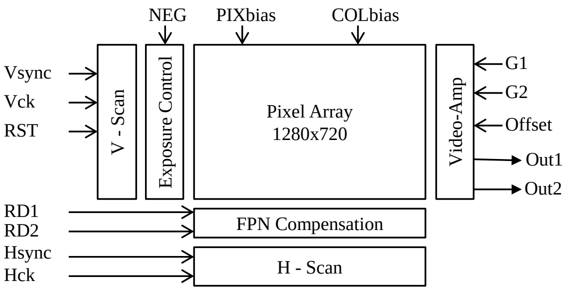
Keywords: HDR, WDR, CMOS sensor, Stereo Imaging.

1. Introduction

Stereo vision systems find application in a wide variety of fields such as robotics applications (Nasa Curiosity rover), automotive applications (collision avoidance, pedestrian detection and driving assistance) or security applications for people counting or industrial safety applications. Today, in computer vision, the geometric problems, such as correcting the projection distortions and calibrating a stereo pair are well known and can be dealt with through image processing and calibration steps [1,2,3,4]. When the stereo vision system is used to extract depth information, the sensor dynamic range and reactivity become critical because in saturated areas the details are lost and consequently depth extraction becomes impossible.

In this communication we present a stereo vision system that relies on two 1280x720 WDR (Wide Dynamic Range) logarithmic image sensors. In section 2 the stereo vision system architecture will be described and in section 3 we will give some indication of the stereo vision system performance.

Peer-review under responsibility of the Scientific Committee of IHCI 2014.



(a) Sensor general structure.

Optical format	1/2 inch
Pixel Size	5.6um x 5.6um
Pixel Array Size	1280x720
Fill Factor	30%
MicroLens	Available
Dynamic range	> 140dB
Output Format	Differential Analog Output
Maximum Frame rate	60Hz
Maximum H scanning rate	80MHz
CMOS process	Standard 0.18 1P3M
High light Log sensitivity	50mV/decade
Low light linear sensitivity	1V/lux*s @ t_{expo}=40ms
Random Noise	0.5mV over all DR
FPN	<1mV
Shutter Mode	Rolling Shutter
Power consumption	230mW

(b) Sensor Characteristics summary.

Fig. 1: Sensor Specifications

2. Stereo Vision Camera

When a stereo vision system is used to compute depth information, the sensor dynamic range and reactivity become critical because in saturated areas the contrast is lost and depth extraction is not possible. A stereo vision system has been developed to address those issues.

At the core of this camera we find 2 B&W WDR NSC1005 logarithmic sensors that are slaves to the same controller. The general specifications of the sensors are given by figure 1. The NSC1005 can deliver constant contrast sensitivity over 140 dB of dynamic range. Both sensors differential analog outputs are connected to 12 bits differential ADCs. In this design a FPGA generates all the sensors control signals (Vsync, Vck, Hsync, Hck, RST, RD1 and RD2) for both sensors resulting in a pixel level synchronization. The digitized left and right pixels levels are then sent to the FPGA. The FPGA will pack the 12 bits left and right channels to create a single 24 bits channel that is sent to a host PC (see fig. 2).

Thanks to the dynamic range provided by both logarithmic sensors, it is possible to set the exposure time and the frame rate at fixed values without risk of saturation. This means that gain and exposure time control becomes very simple for stereo vision systems. The wide dynamic range provided within a single frame for the 2 sensors makes the stereo vision system very robust to non-uniform illumination, specular reflections and fast temporal illumination changes. For example, in this stereo system design, the frame rate can be set to 25 or 30 Hz, the exposure time is set to the maximum possible value (the frame rate period) and the sensors gain are set from the host PC to the same value for both sensors. Another interesting property of logarithmic sensors is that contrast will be identical in areas seen by both left and right sensors regardless of the illumination intensity on both sensors. This property is important if we consider the stereo matching process that relies on contrast information.

3. System Performance

In order to operate the stereo vision system, a companion software was developed to perform the stereo calibration and the rectification processes. The stereo calibration module implements a camera calibration algorithm based on corner extraction from chessboard chart images[3,4] to correct both sensor modules distortions. Using the same sets of points the calibration module also computes the essential matrix, the fundamental matrix, the rotation matrix between the left and the right camera coordinate systems and the translation vector between the coordinate systems of the cameras. Finally the left and right images are rectified using Bouguet algorithm[4]. With this implementation

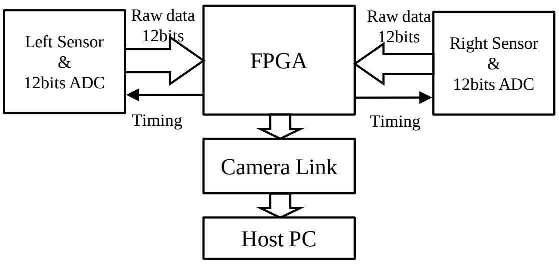


Fig. 2: Stereo vision system architecture.

of the calibration and rectification steps, we typically get a RMS re-projection error of the input calibration points of $E_{RMS} = 0.2$ pixel. After rectification of the left and right images, the relationship between the depth Z and the binocular disparity d is given by:

$$Z = \frac{f}{S_{pix}} \frac{T}{d} \tag{1}$$

where Z is given in meter, d is given in pixel, f is the lenses focal length, S_{pix} is the pixel size and T is the system baseline. The stereo vision system default configuration features low distortion lenses with a focal length of 5.5mm. Considering that the pixel size is 5.6um the maximum working Z_{max} range of the stereo system will be $Z_{max} \approx 50m$. The minimum working range will be limited by the disparity range allowed for the stereo matching. The closer the object will be to the camera, the higher the binocular disparity. Depending on the complexity of the stereo matching algorithm, a tradeoff should be made between the matching process complexity, the disparity range, the field of view and the processing speed. In our software implementation we typically set the maximum binocular disparity possible to 128 pixels. This gives a minimum working range $Z_{min} \approx 0.38m$. We can also compute the accuracy Z_{acc} of our calibrated and rectified system as a function of the RMS re-projection error and binocular disparity:

$$Z_{acc} = E_{RMS} \frac{f}{S_{pix}} \frac{T}{d^2} \tag{2}$$

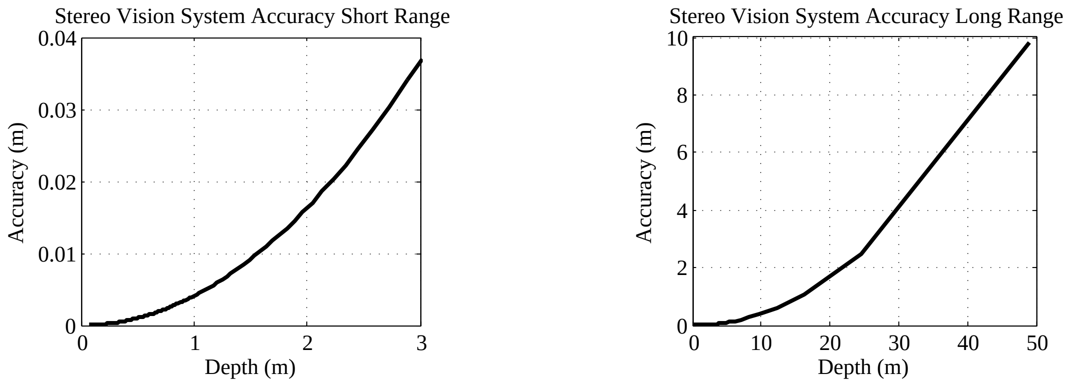
By combining eq. 1 and eq. eq. 2 we can compute the depth accuracy as a function of depth:

$$Z_{acc} = E_{RMS} \frac{S_{pix}}{fT} Z^2 \tag{3}$$

Figures 3a and 3b gives the short range and long range accuracy curves for $E_{RMS} = 0.2$. Those curves should be seen as the best accuracy that the stereo camera can achieved. Obviously, the absolute depth precision will dependent on the performance of the stereo matching algorithm. In the case of an active system, the presence of an illuminator generating a pattern will also help to achieve the accuracy given by fig. 3a and fig. 3b. Figure 4 gives some depth map sample images computed from the rectified left and right images using the stereo matching algorithm described in[5] without any kind of illuminator or pattern generator.

4. Conclusion

In this paper, we have presented a stereo vision system that relies on 2 logarithmic WDR 1280x720 sensors. This stereo vision system is able to provide high quality contrast indexed images of highly contrasted scene. This contrast indexed sensing capability can be conserved over more than 140dB without any explicit sensor control and without delay. The logarithmic photoresponse is also interesting if we consider the image robustness to illumination variabilities. This dynamic range and contrast conservation makes this stereo camera very effective for outdoor 3D stereo vision applications. Besides it also gives the unique possibility to combine passive 3D approaches with active illumination approaches without fear of losing information because of saturation.



(a) Stereo camera short range accuracy. (b) Stereo camera long range accuracy.

Fig. 3: Stereo Camera Accuracy

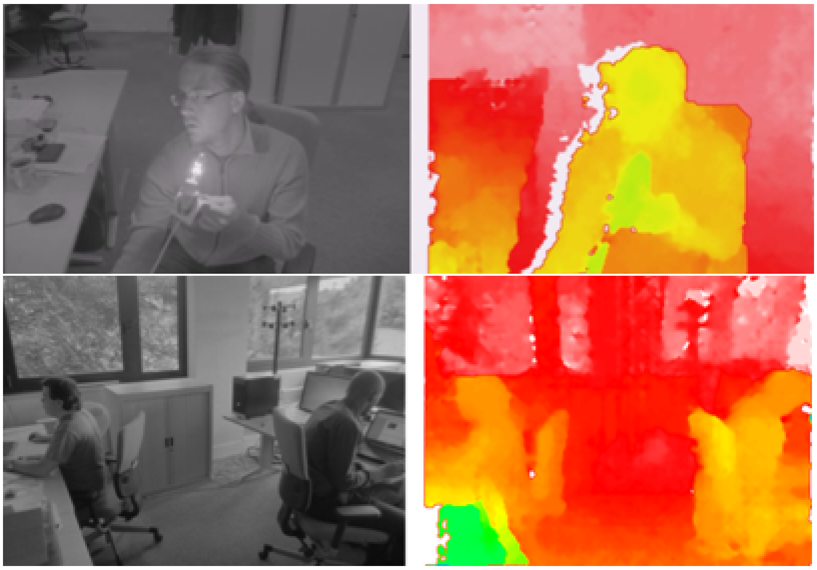


Fig. 4: Left sensor image and depth map sample image (frame rate 25Hz, exposure time 40ms)

References

1. Tsai, R.. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. 1987. URL: `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1087109`.
2. Kanade, T., Yoshida, A., Oda, K., Kano, H., Tanaka, M.. A stereo machine for video-rate dense depth mapping and its new applications. In: *Proceedings of the 15th Computer Vision and Pattern Recognition Conference (ICVPR '96)*. 1996, p. 196–202.
3. Zhang, Z.. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2000; **22**(11):1330–1334. doi:http://doi.ieeecomputersociety.org/10.1109/34.888718.
4. Bradski, G., Kaehler, A.. *Learning OpenCV*. O'Reilly Media Inc.; 2008. URL: `http://oreilly.com/catalog/9780596516130`.
5. Hirschmuller, H.. Accurate and efficient stereo processing by semi-global matching and mutual information. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*; CVPR '05. Washington, DC, USA: IEEE Computer Society. ISBN 0-7695-2372-2; 2005, p. 807–814. URL: `http://dx.doi.org/10.1109/CVPR.2005.56`. doi:10.1109/CVPR.2005.56.