

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.

Index Terms—HDR, WDR, CMOS sensor, Stereo Imaging.

I. INTRODUCTION

We are all familiar with our ability, as human, to extract 3D information from our environment thanks to our stereoscopic vision. This level of perception is achieved through the sensing capability of the human eye and the processing capability of the human brain. In principle, extracting 3D information from an artificial stereoscopic vision system is simple. In the ideal case, if we are able to identify identical scene primitives with both sensor units, we can extract directly the depth information by triangulation. Stereo vision systems find application in a wide variety of fields such as robotics. For example Nasa Curiosity rover has several stereo camera systems to scan Mars surface for navigation, hazard avoidance or for terrain measurements. Stereo vision systems can also be found in automotive application. Car manufacturer Subaru proposes a stereo vision system that is used for collision avoidance, pedestrian detection and driving assistance. Stereo camera can also find application in security 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]. But even if we consider an ideal stereoscopic camera, the stereo correspondence process remains one of the most challenging task in order to extract depth information. When the stereo vision system is used to extract depth information, the sensor dynamic range and reactivity become critical because in saturated areas

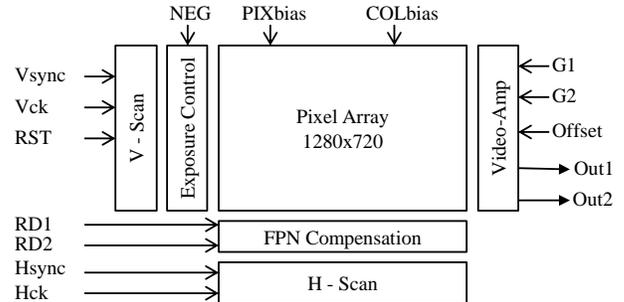


Fig. 1. Sensor general structure.

the details are lost and consequently depth extraction becomes impossible. The sensor reactivity will also be important to prevent saturation in changing environments.

In this communication we present a stereo vision system that relies on two 1280x720 WDR (Wide Dynamic Range) logarithmic image sensors. In section II we will detail the characteristics of the sensors embedded in the stereo vision system. In section III the stereo vision system architecture will be described. Finally, in section IV we will give some indication of the stereo vision system performance.

II. SENSOR CHARACTERISTICS

Similarly to biological photoreceptors, the sensors embedded in the stereo vision system presented in this paper follow a logarithmic law. This kind of sensor response is very interesting in imaging applications because it can deliver contrast indexed images over a very wide dynamic range with a single frame. Previous work on logarithmic sensors were based on current to voltage converting using MOS transistor in sub-threshold mode [5], [6], [7], [8], [9], [10]. The NSC1005 sensor presented in this paper features the pixel design that was described in a previous communication [11], [12]. In this pixel design the photodiode is used in solar cell mode in order to get a natural logarithmic law. The logarithmic pixel design proposed in [12] resolves the problems of the previous logarithmic image sensors. Internal FPN compensation is included, the sensitivity has been drastically improved and the lag problem have also been addressed.

The general structure of the sensor is given by figure 1. Pixel lines are selected sequentially. The FPN correction starts before each line readout. The selected line of the pixel array

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

TABLE I
SENSOR CHARACTERISTICS SUMMARY.

will be loaded to the first column memory buffers by RD1. Then the RST signal resets the current line. On RD2 we load the dark signals to the second column memory buffers. A differential amplifier is used to subtract the contents of these two memory buffers for the FPN compensation. The gain of the amplifier can be set to 4 different values (0dB, 4dB, 8dB and 12dB) using two digital control bits G0 and G1. An offset can be applied by adding an OFFSET voltage to match the sensor output with the input voltage range of an external differential ADC. Table I summarized the sensor characteristics.

It was demonstrated in [12] that the solar cell mode pixel photoelectric response V_D is easily derived from the open-circuit voltage of the illuminated photodiode after reset operation:

$$V_D = V_T \ln \frac{I_{ph} + I_S}{I_{ph} \exp\left(-\frac{(I_{ph} + I_S)t}{V_T C_D}\right) + I_S} \quad (1)$$

where $V_T = \frac{kT}{q}$ and k , T , q , I_S , I_{ph} , t and C_D are respectively the Boltzmann constant, the absolute temperature, the elementary charge, the junction current, the photocurrent, the exposure time and the junction capacitance. A quick analysis of eq. 1 denominator shows that for low illumination or short exposure time, the sensor response will be linear whereas the sensor response will follow a logarithmic law for high illumination [12]. Figure 2 shows the measured sensor photoelectric response for an exposure time $t = 40ms$. We can see on figure 2 that the proposed photoresponse model is valid. For low illuminations, the sensor is clearly operating in linear mode. From 0.8Lux, figure 2 shows that the sensor enters the logarithmic zone of the photoresponse. For those measures, the sensor has been illuminated using a green LED at 525nm. The dynamic range was measured to be over 140dB using laser illumination. Figure 3 shows a crop image sample taken from the sensor looking directly at an intense halogen light.

III. STEREO VISION CAMERA

When a stereo vision system is used to compute depth information, the sensor dynamic range and reactivity become

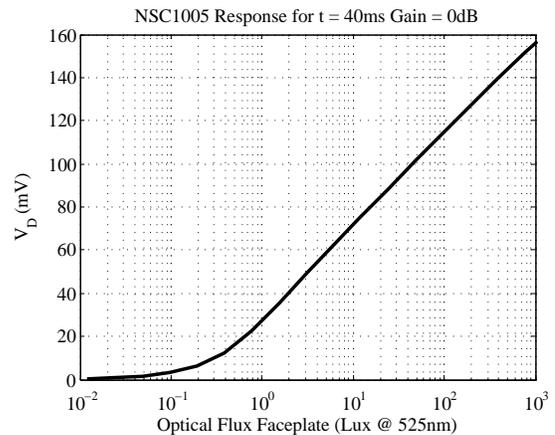


Fig. 2. Sensor Photoelectric Response.



Fig. 3. Sample image (frame rate 25Hz, exposure time 40ms)

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. Table II summarized the stereo camera main characteristics.

At the core of this camera we find 2 B&W WDR NSC1005 sensors that are slaves to the same controller. Both sensors differential analog outputs are connected to 12 bits differential ADCs. In this design a FPGA generates all the sensors control signals (V_{sync} , V_{ck} , H_{sync} , H_{ck} , 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. 4).

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 even if the sensors are directly illuminated with strong light (Fig. 3).

Sensors	2 x NSC1005
Data Output	24 bits CameraLink channel
A/D converter	12-bits (x 2)
Baseline	5cm

TABLE II
STEREO CAMERA MAIN CHARACTERISTICS.

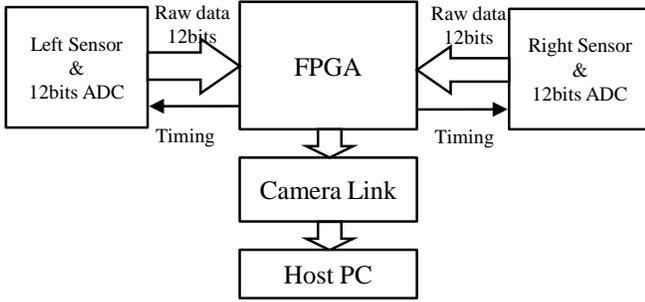


Fig. 4. Stereo vision system architecture.

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.

The logarithmic conversion function also offers interesting properties regarding both left and right images processing. If we consider the optical intensity $OS(\lambda, x, y)$ incident on a pixel at coordinates (x, y) , $OS(\lambda, x, y)$ can be written as the product of :

$$OS(\lambda, x, y) = L(\lambda, x, y) \times R(\lambda, x, y) \times T(\lambda, x, y) \quad (2)$$

where $L(\lambda, x, y)$, $R(\lambda, x, y)$ and $T(\lambda, x, y)$ are respectively the illuminance illuminating the scene, the reflectance of the elements in the scene and the overall transmittance [13]. This optical signal falling on a pixel will generate a corresponding photocurrent :

$$I_{ph}(x, y) = \alpha(\lambda, x, y) \times OS(\lambda, x, y) \quad (3)$$

where $\alpha(\lambda, x, y)$ is the spectral sensitivity of the photodiode. In section II, we saw that with enough light ($I_{ph} \gg I_S$) the sensor enters the logarithmic portion of its photoelectric response (0.8Lux for an exposure time of $t = 40ms$) and V_D can then be approximated by $V_D \approx V_T \ln(I_{ph}/I_S)$ and:

$$\begin{aligned} V_D(x, y) &\approx V_T \ln \left(\frac{\alpha(\lambda, x, y) \times OS(\lambda, x, y)}{I_S} \right) \\ &\approx V_T (\ln L + \ln R + \ln T + \ln \alpha - \ln I_S) \quad (4) \\ &\approx V_T (\ln L + \ln R + \ln T + \ln \alpha + FPN) \end{aligned}$$

We see that, when the sensor operates in logarithmic mode, the conversion function will lead to a pixel output that is the sum of the log-luminance, log-reflectance, log-transmittance and log-spectral sensitivity. In most of cases, the illumination on a scene is quite smooth and uniform. As a result, the illumination level, seen as an information vehicle, is transformed into an offset by the logarithmic response and so is the spectral sensitivity of the photo-detector. The suppression of

this offset will normalize the reference between both left and right sensors. The resulting signals will then index contrast instead of intensity. In our FPGA implementation we use the minimum level of each image to normalize the left and right images. Once this illumination information is removed, we see that the sensors signals now index contrast instead of light intensity. Furthermore, once the sensor is working in logarithmic mode, the contrast sensitivity will be constant providing that both sensor amplifier are set with the same gain. The direct consequence of this property 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 interesting if we consider the stereo matching process that relies on contrast information.

IV. 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 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:

$$Z = \frac{f}{S_{pix}} \frac{T}{d} \quad (5)$$

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:

$$Z_{acc} = E_{RMS} \frac{f}{S_{pix}} \frac{T}{d^2} \quad (6)$$

By combining eq. 5 and eq. eq. 6 we can compute the depth accuracy as a function of depth:

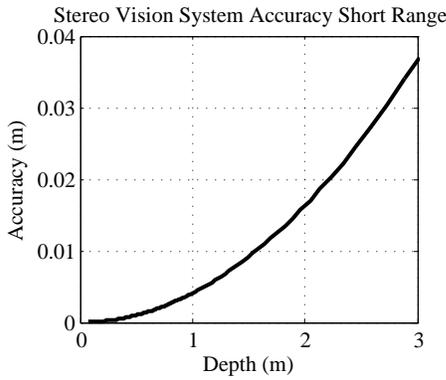


Fig. 5. Stereo camera short range accuracy.

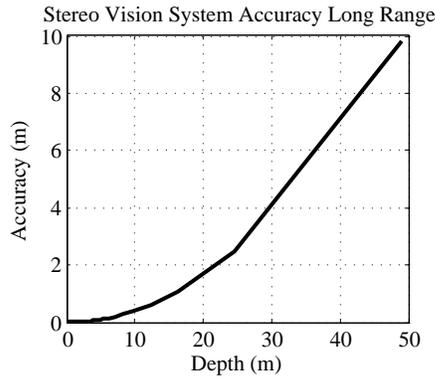


Fig. 6. Stereo camera long range accuracy.

$$Z_{acc} = E_{RMS} \frac{S_{pix}}{fT} Z^2 \quad (7)$$

Figures 5 and 6 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. 5 and fig. 6. Figure 7 gives some depth map sample images computed from the rectified left and right images using the stereo matching algorithm described in [14] without any kind of illuminator or pattern generator.

V. 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

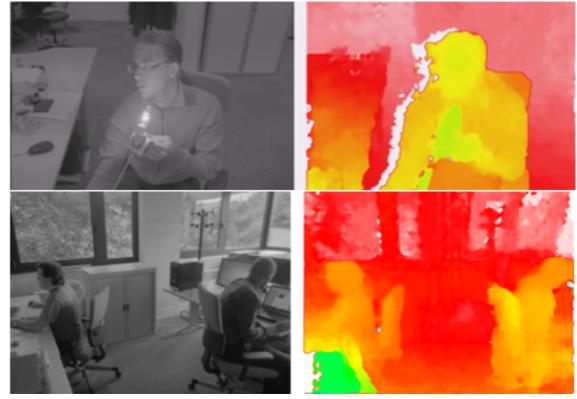


Fig. 7. Left sensor image and depth map sample image (frame rate 25Hz, exposure time 40ms)

approaches with active illumination approaches without fear of losing information because of saturation.

REFERENCES

- [1] R Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," 1987.
- [2] Takeo Kanade, A. Yoshida, K. Oda, H. Kano, and M. Tanaka, "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)*, June 1996, pp. 196–202.
- [3] Zhengyou Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [4] Gary Bradski and Adrian Kaehler, *Learning OpenCV*, O'Reilly Media Inc., 2008.
- [5] S.G. Chamberlain and J. Lee, "A novel wide dynamic range silicon photoreceptor and linear imaging array," *IEEE Journal of Solid-State Circuits*, vol. SC-19, no. 1, pp. 41–48, 1984.
- [6] N. Ricquier and B. Dierickx, "Pixel structure with logarithmic response for intelligent and flexible imager architectures," in *Solid State Device Research Conference, 1992. ESSDERC '92. 22nd European*, 1992, pp. 631–634.
- [7] T. Delbrck and C.A. Mead, "Analog vlsi phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits," California Institute of Technology, Computation and Neural Systems program, CNS Memorandum 30, 1994.
- [8] K. Takada and S. Miyatake, "Logarithmic-converting ccd line sensor and its noise characteristics," in *IISW 1997*, 1997.
- [9] S. Kavadias, B. Dierickx, and G. Meynants, "A self-calibrating logarithmic image sensor," in *IEEE Workshop on CCD&AIS*, 1999.
- [10] M. Loose, K. Meier, and J. Schemmel, "A self-calibrating single-chip cmos camera with logarithmic response," *IEEE Journal Solid-State Circuits*, vol. 36, no. 4, pp. 586–596, 2001.
- [11] Y. Ni and K. Matou, "A cmos log image sensor with on-chip fpn compensation," in *ESSCIRC'01*, Villach, Austria, September 2001, pp. 128–132.
- [12] Yang Ni, YiMing Zhu, and Bogdan Arion, "A 768x576 logarithmic image sensor with photodiode in solar cell mode," in *International Image Sensor Workshop*. International Image Sensor Society, June 2011.
- [13] B. Hoefflinger, *High-Dynamic-Range (HDR) Vision: Microelectronics, Image Processing, Computer Graphics (Springer Series in Advanced Microelectronics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.
- [14] Heiko Hirschmuller, "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*, Washington, DC, USA, 2005, CVPR '05, pp. 807–814, IEEE Computer Society.