



# A real-time high dynamic range intensified complementary metal oxide semiconductor camera based on FPGA

Yi-Zheng Lang<sup>1</sup> · Yun-Sheng Qian<sup>1</sup> · Hong-Gang Wang<sup>2</sup> · Xiang-Yu Kong<sup>1</sup> · Sheng Wu<sup>1</sup>

Received: 14 May 2021 / Accepted: 20 March 2022 / Published online: 17 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

To capture more details and enhance image's contrast in a low-light-level (LLL) environment and improve the performance of the LLL night vision system, a method of extending the dynamic range for the LLL night vision system based on guided filter and adaptive detail enhancement is proposed. Moreover, a complete high dynamic range (HDR) intensified complementary metal oxide semiconductor (ICMOS) camera based on this method is implemented. Unlike conventional technologies to create HDR images, the proposed camera based on an HDR CMOS can better process images' edge information while avoiding artifacts' appearance and solve the storage problem in a multi-images fusion so that real-time image is realized processing. Based on this, the ICMOS camera can capture details in a high contrast environment with an extended dynamic range of up to 80 dB. As the core processing unit of this camera, a Xilinx Spartan 6 series FPGA has completed the work of CMOS driving, image detail enhancement, image storage control, video compression output. Finally, this camera can realize an output video with a resolution of 1280×1024 at the rate of 60fps, and experimental results show that this camera has a good performance under a  $10^{-4}$  lx environment.

**Keywords** Contrast · Low-light-level · ICMOS camera · HDR · Guided filter · FPGA

## 1 Introduction

The dynamic range that ordinary complementary metal oxide semiconductor (CMOS) cameras can capture is generally low, which often several orders of magnitude lower than natural environment'. This defect will inevitably lead to overexposure or underexposure in a high or low illumination environment, which inevitably brings about detail-miss in the captured images (Mendis et al. 1997). The emergence of high dynamic range (HDR) technology has solved the above problem in both software and hardware. The

---

✉ Yun-Sheng Qian  
yshqian@mail.njust.edu.cn

<sup>1</sup> School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, China

<sup>2</sup> School of Information and Electrical Engineering, Ludong University, Yantai 264025, Shandong, China

HDR technology can broaden the dynamic range of images, meanwhile, retain details that interest people (Lapray et al. 2016). There are already some HDR cameras used widely to solve problems in automotive imaging, tunnel inspection and metal welding real-time monitoring, etc. (Lee et al. 2010; Mann et al. 2012). Unfortunately, the number of HDR cameras used in a LLL environment is very minimal, so it's urgent to propose a LLL HDR camera with a good performance.

With the continuous development of LLL night vision technology, the vacuum night vision technology has evolved from a direct-view night vision device to a digital LLL device combing an image intensifier and a CMOS/CCD, which is called IC MOS/ICCD. The photosensitive surface of the CMOS/CCD is coupled with a relay element to receive the optical images from the image intensifier's fluorescent screen. To improve cameras' performance under a LLL environment, it's urgent to design an algorithm for image enhancement.

For the generation of HDR images, there are two main methods. Method one is to redesign the internal structure of the CMOS image sensor, which includes logarithmic response adjustments, multiple slope integration, time-to-saturation adjustments, and so on (Zhao et al. 2015). Method two is to capture some low dynamic range (LDR) images in different exposure times and then merge them into one HDR image, so images contain more information (Mann et al. 2012). In this paper, an FPGA-based HDR IC MOS camera based on Method One is designed to capture more details and improve the image's contrast in a LLL environment. It is furthermore improving the LLL night vision's performance.

This system uses HDMI display format to display the video on an OLED device. But the dynamic range of the HDR images processed by the FPGA is higher than the dynamic range of the OLED's, so it's necessary to use a tone mapping operator (TMO) to properly compress the HDR images to fit the display device (Karaduzovic-Hadziabdic et al. 2017; Abolbashari et al. 2012). TMOs can be divided into two main groups named global and local operators (Popovic et al. 2016). Global operators are spatially invariant because they apply the same transformation to all pixels. These TMOs usually have low complexity and high computational speed. Opposite to global operators, local operators are more flexible and adaptive, drastically improving local contrast in regions of interest. Since the local TMOs process pixels differently, they are computationally more expensive and resource-demanding.

According to various references, there are numerous HDR algorithms. However, only a few can be practically used because of the mismatch between algorithms and hardware. There are mainly hardware implementation platforms based on CPU, GPU, FPGA, and a mixture. Comprehensively considering the needs of image acquisition rate and computing speed, this paper chooses Xilinx Spartans 6 FPGA as the hardware implementation platform.

The HDR IC MOS camera proposed in this paper enhanced the details of images collected by HDR CMOS to expand further the dynamic range, which reaches up to 80 dB. The processed data is compressed by TMO and finally displayed on an OLED, with the resolution of  $1280 \times 1024$  and a frame rate of 60 fps. Experimental results show that this HDR IC MOS camera performs well and ensures the image quality under the illumination of  $10^{-4}$  lx.

The rest of this paper is organized as follows. The existing HDR imaging technologies are reviewed briefly in Sect. 2. The enhancement algorithm of the system is illustrated in Sect. 3. Section 4 elaborates on the hardware platform of this camera. Experimental results are presented in Sect. 5. Finally, the conclusion of this paper is shown in Sect. 6.

## 2 Related work

The acquisition and display of HDR images have earned lots of attention, and it has become a popular topic. However, most proposed algorithms have not considered the possibility of hardware implementation. In other words, due to the complexity of computation and poor adaptability, many algorithms only stay in the theoretical stage.

After more than ten years of development, some HDR CMOS cameras have gradually developed to capture images in the natural environment better. Compared with ordinary cameras, due to the particularity of the HDR algorithm, the possibility of software and hardware implementation should be considered while designing HDR cameras to implement the algorithm in hardware successfully. As mentioned in Sect. 1, most HDR cameras are based on CPU, GPU, FPGA, and combination (Tang et al. 2020). Hassan et al. proposed an FPGA-based architecture for a local TMO, which used Altera's FPGA to output a video at the rate of 60 fps with a resolution of  $1024 \times 768$ . A peak signal-to-noise ratio study shows that the architecture has a good performance (Hassan and Carletta 2007). Mann et al. presented a real-time video HDR prototype with applications such as mediated reality, augmented reality which consists of an EyeTap (electric glasses) welding helmet, with a wearable computer upon which are implemented a set of image processing (Mann et al. 2012). This system used lookup tables (LUT) to compute the HDR video in real-time with 120 fps on a Xilinx FPGA. Lapray et al. proposed a real-time hardware implementation of the HDR technique, and they built a dedicated smart camera that performs both capturing and processing HDR video through three exposures (Lapray et al. 2014). They also described a complete FPGA-based smart camera architecture, HDR-ARtiSt, which produces a real-time HDR live video stream through multiple captures (Lapray et al. 2016). They built this system on a standard B&W CMOS image sensor and a Xilinx FPGA to provide full sensor resolution (1.3 Mega pixels) at 60 frames per second in the form of a real-time HDR video stream. Nosko et al. described a novel FPGA architecture of an HDR video processing pipeline to capture a series of differently exposed images (Nosko et al. 2020). Also, they proposed an acquisition process enabling multi-exposure HDR and fast implementation of local TMO involving bilateral filtering. This system is based on Xilinx's FPGA, which can capture 30 fps of full high definition (HD) HDR video. Ahmad et al. introduced a novel feature vectors construction approach for face recognition using discrete wavelet transform (DWT). They achieved Xilinx's FPGA-based IP core implementation of transform block in face recognition systems. Through experimental comparison, the accuracy of facial recognition in this system is high (Ahmad et al. 2013). Jacquot et al. proposed a prototype CMOS camera system based on Xilinx's FPGA to create HDR images that approach CCDs' dynamic range (Jacquot et al. 2014). The camera was built on a commercial  $1280 \times 1024$  CMOS image sensor with 10-bits per pixel and up to 500 Hz full frame rate. What's more, this camera can reach higher frame rates by adjusting the available window. Chi-Yi Tsai et al. described an algorithm for the real-time enhancement of color images of poor visual quality based on GPU (Tsai et al. 2019). Their experimental results demonstrated the ability of the proposed GPU-accelerated color image enhancement method, which can render satisfactory enhancement results for both low- and high-intensity images. As can be seen, many HDR image processing algorithms can be successfully applied in hardware platforms. Because of its strong computing capabilities and excellent software adaptability, FPGA has been regarded as the first choice in the HDR image algorithm hardware platform.

However, it is easily found that most of the existing HDR systems still use Method Two to take multiple LDR images for the same scene to obtain HDR images. Consulting the literature, this method has been developed rapidly. With the deepening of research, the algorithm is continuously optimized, the calculation process is continuously simplified, and the effect of HDR images obtained by fusion has also been greatly improved. No doubt, there are still many shortcomings, such as ghosting and images' storage. Compared with Method Two, there are few researches on HDR cameras based on Method One. This is due to HDR sensors' high price and technical requirements, which is not convenient for people to carry out research.

For image processing, bilateral filter, as the simplest and most intuitive one among explicit weighted filters, is widely used in many applications (Tomasi et al. 1998; Liu et al. 2006; Durand et al. 2002; Winnemöller et al. 2006). Durand et al. applied bilateral filter to display HDR images successfully (Durand et al. 2002). Kuang et al. evaluated various HDR image TMOs, including Durand's (Kuang et al. 2006). The result showed that the images processed by Durand's algorithm are better than the images processed by others. Cadik et al. also conducted comparative analysis on several methods, and the results showed that the images processed by Durand's algorithm have the best performance, which is attributed to the algorithm using bilateral filter (Cadik et al. 2008). Branchitta et al. used a BF & DRP framework to visualize HDR infrared images (Branchitta et al. 2009). Zuo et al. proposed an HDR infrared image display and detail enhancement algorithm based on bilateral filter and achieved a good display effect (Zuo et al. 2011).

Although bilateral filter is top-rated in image processing, it is undeniable that it still has limitations. It has been noted in the literature, images processed by bilateral filter may lead to unwanted gradient reversal artifacts near edges, which affects the quality of the images (Zuo et al. 2011; Bae et al. 2006; Farbman et al. 2008; Zhou et al. 2016).

To solve the above shortcomings of existing FPGA-based systems and bilateral filter, this paper uses the NSC1105 HDR CMOS sensor developed by NIT company as the image acquisition device and then proposes an adaptive detail enhancement algorithm based on guided image to output video with the resolution of  $1280 \times 1024$ , 60 fps. Given the few research results, the contribution of this paper is to propose an image detail enhancement algorithm based on guided filtering, which performs well in the hardware platform. What's more, this paper develops an HDR camera based on an HDR CMOS sensor which has good performance under  $7 \times 10^{-4}$  lx. In view of there is rarely a report about ICMOS cameras, all these can provide a reference for related fields.

### 3 The algorithm of HDR ICMOS camera

ICMOS cameras are used in LLL imaging. However, the contrast of the LLL images is poor due to the environment. This paper proposes an adaptive detail enhancement algorithm to enhance the contrast of images based on guided filter to solve this problem. Firstly, using guided filter to decompose images into base layer images and detail layer images. Secondly, the base layer images are processed by gray stretching. After adaptive enhancement of detail layer images, the two images are merged into new images. It is worth noting that the sensor outputs analog signals, so this algorithm must be processed after analog-to-digital conversion. The block diagram of the algorithm is shown in Fig. 1. The details of the algorithm will be explained in the following sections.

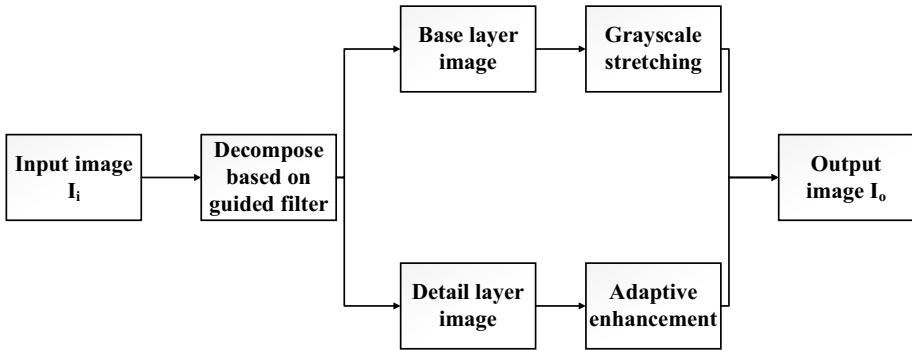


Fig. 1 The block diagram of algorithm

### 3.1 Decomposition of input image

As mentioned in Sect. 2, images processed by bilateral filter may lead to unwanted gradient reversal artifacts near edges. The reason is that the bilateral filter’s mechanism is closely related to a robust iterative procedure (i.e., the mean shift) which achieves edge-preserving filtering by searching for local modes in the joint spatial-range domain. One iteration of bilateral filter amounts to converge to the local mode (Zuo et al. 2011). However, when a pixel around an edge has a few similar pixels around it, the Gaussian weighted average is quite unstable, which will cause extra contours around the edges of the image. In this paper, we use another explicit filtering method called guided filter to solve this problem. Compared with bilateral filter, guided filter has a similar effect on smoothing the image, but it has better advantages in edge processing, which can avoid the appearance of artifacts effectively. Bilateral filter with a larger kernel window often has better performance but significantly increases computational complexity (Pham et al. 2011). The calculation time of guided filter has nothing to do with the size of the kernel window, so it’s flexible in selecting the appropriate filter window according to needs.

When guided filter is performed to decompose images, a guided image must be used. To retain the edge information of the input image, this paper selects the input image  $I$  itself as the guided image. The base layer image obtained by the procession of guided filter is:

$$I_b = \sum_{(x',y') \in W_{ij}} W_G(x',y') \times I \tag{1}$$

where  $I_b$  represents the base layer image. The notation  $(x',y') \in W_{x,y}$  indicates that  $(x',y')$  is a pixel in a filter window on which the pixel  $(x,y)$  centers.  $W_G(x',y')$  is the kernel weights function which can be used as the weighting coefficient to enhance the detail layer image. The kernel weight can be explicitly expressed as:

$$W_G(x',y') = \frac{1}{|w|^2} \sum_{\substack{(x',y') \in W_{x,y} \\ (x'',y'') \in W_{x',y'}}} \left( 1 + \frac{(I(x'',y'') - \mu_{x',y'})(I(x',y') - \mu_{x',y'})}{\sigma_{x',y'} + \varepsilon} \right) \tag{2}$$

In Eq. 2,  $w$  is the number of the pixels in the window  $W_{x,y}$ ,  $(x'',y'') \in W_{x',y'}$  indicates that  $(x'',y'')$  is a pixel in a filter window on which  $(x',y')$  centers.  $\mu_{x',y'}$  and  $\sigma_{x',y'}$  represent the mean and variance of  $I$  in  $W_{x',y'}$ , respectively. And  $\varepsilon$  is a regularization parameter which

describes the smoothing level of the filter. When  $\varepsilon$  is small (such as  $\varepsilon = 100$ ), the obtained detail layer image will consist of small details, such as background noise and tiny structural information. When  $\varepsilon$  is big, some textures will be ignored when images being processed, and some strong edge information will be focused on.

### 3.2 Base layer image processing

In order to improve the visibility of LLL images, it is necessary to enhance base layer images appropriately. Firstly, grayscale compression maps the base layer images into an acceptable range. Secondly, improving the brightness of the overall images through a linear transformation. The specific process is as follows:

$$I'_b = \log(I_b + o) \tag{3}$$

In Eq. 3,  $\log(\cdot)$  denotes the natural logarithm operation, and the parameter  $o = 1$  is set to prevent the log value from being negative after logarithmic calculation. At this time, the detail layer image in the logarithmic domain can be calculated:

$$I'_d = \log(I + o) - I'_b \tag{4}$$

Next, processing the base layer image by using a scale factor  $\beta$ :

$$U = \beta I'_b + \gamma \tag{5}$$

The parameter  $\gamma$  in Eq. 5 is to recover the contrast of the whole image. The dynamic range of image will be compressed if  $\beta < 1$ . On the contrary, the dynamic range will be broadened while  $\beta > 1$ . Supposing the basic contrast of the target image is  $T$ , the value of  $\beta$  can be calculated as follow:

$$\beta = \frac{\log(T)}{\max(I'_b) - \min(I'_b)} \tag{6}$$

In Eq. 6,  $\max(I'_b)$  and  $\min(I'_b)$  represent the maximum value and minimum value of  $I'_b$ , respectively. The whole contrast of image will change when image is processed by Eq. 5, so it's necessary to use factor  $\gamma$  to restore the contrast.

$$\gamma = (1 - \beta) \max(I'_b) \tag{7}$$

Finally, the image is indexed to obtain the enhanced base layer image:

$$I''_b = \exp(U) \tag{8}$$

### 3.3 Detail layer image processing

As mentioned above, because the detail layer image contains lots of tiny details that always determine the quality of the processed image, so the detail layer image must be enhanced. In the meantime, the detail layer image contains lots of noise coming from the input image. If using a fixed gain factor to enhance the whole image directly, the noise will also be expanded, affecting the quality of the final image. Psychophysical

experiments confirmed that noise in flat regions of the image will give rise to the illusion the observer may suffer from. While at sharp transitions in image intensity, the contrast and sensitivity of the human visual system will decrease with the sharpness of the change (Zuo et al. 2011). Therefore, adaptive enhancement based on image area information is essential to improve images' quality effectively. Katsaggelos et al. adopted the local variance as noise masking function  $M(x, y)$  to measure the spatial detail, and  $(x, y)$  is used to represent the specific coordinates of the pixel (Katsaggelos et al. 1991). They defined the noise visibility function as:

$$f(x, y) = \frac{1}{M(x, y) \times \theta + 1} \quad (9)$$

In Eq. 9,  $\theta$  is the adjustment parameter, the noise visibility function is a standardized function with a value range of  $[0, 1]$ . This paper replaces the noise visibility function in Eq. 9 with the kernel weights function described in Sect. 3.1 as follows:

$$f(x, y) = \frac{1}{W_G(x', y') \times \theta + 1} \quad (10)$$

It can be seen from Eq. 10 that the visibility function is close to 0 when the value of the kernel weights function is big, that is, where the image changes significantly, which means the noise has a little impact on the image, a larger gain factor can be used. On the contrary, the visibility function is close to 1 while the value of the kernel weights function is small, which means the noise has a great impact on the image. For the convenience of calculation, setting  $\theta = 1$ . In order to prevent the image quality from being seriously affected by excessive noise amplification, a smaller gain factor should be used to enhance the image quality. To make enhanced images neither detailedly missed or excessively enhanced, setting the gain factor  $g_{\min} = 1$  when  $f(x, y) = 1$ , and  $g_{\max} = 2.5$  when  $f(x, y) = 0$ . Now, the expression of the gain factor can be described as below:

$$g(x, y) = g_{\min} + [1 - f(x, y)](g_{\max} - g_{\min}) \quad (11)$$

The enhanced detail layer image  $I_d''$  can be obtained by Eq. 12.

$$I_d'' = \exp(I_d'(x, y)) \times g(x, y) \quad (12)$$

After the base layer image and the detail layer image are processed, we can get the enhanced image:

$$I_{\text{enhanced}} = I_b'' + \eta \times I_d'' \quad (13)$$

where  $\eta$  is the weight of the detail layer image. Here, we set  $\eta = 1.5$  in this paper.

#### 4 The hardware of HDR ICMOS camera

This paper designs an ICMOS camera based on a Xilinx Spartan 6 series FPGA. And FPGA has mainly completed CMOS driving, image adaptive detail enhancement, images' storage, output video, etc. The hardware block diagram of this system is shown in Fig. 2.

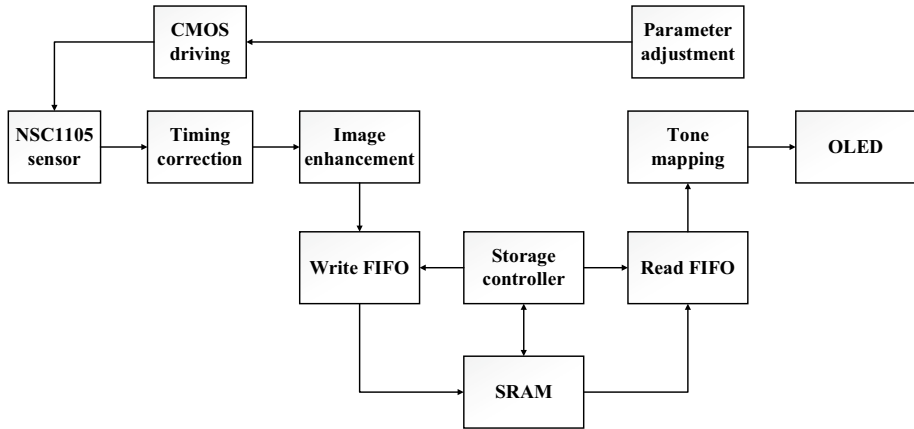


Fig. 2 System hardware block diagram

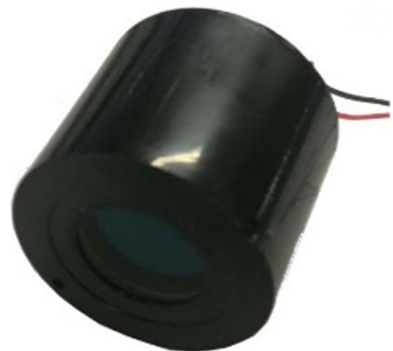
### 4.1 ICMOS composition and image acquisition

The size of the third-generation image intensifier’s phosphor screen is 1 inch, while the size of the photosensitive surface of the traditional CMOS sensor is mostly 1/2 inch. Therefore, it is necessary to couple the CMOS and the image intensifier with a light cone, but this will seriously reduce the imaging quality, and the image transfer function (MTF) will drop sharply. Traditional CMOS sensors are primarily linear in response. Due to the high gain characteristics of the image intensifier, even weak incident light received by the conventional CMOS sensor after being enhanced by the image intensifier can easily cause pixel saturation. In this paper, the NSCI105 CMOS has logarithmic response is selected. Even without special control, saturation will not occur under strong illumination. Therefore, the image’s contrast will not be significantly reduced, which is very suitable for HDR design.

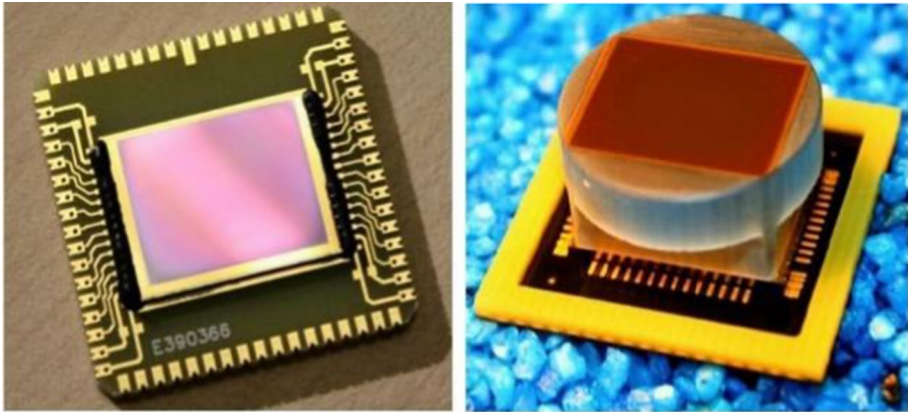
The NSCI105 used in this paper has been coupled with an optical fiber panel to match the size of the image intensifier’s phosphor screen and be directly coupled with the image intensifier to form an ICMOS. The physical image of the third-generation image intensifier and the CMOS coupled through the fiber plane are shown in Fig. 3, 4.

As mentioned above, FPGA is responsible for the driving of CMOS. To drive the CMOS sensor to work correctly. The NSCI105 sensor is driven with a 100 MHz clock

Fig. 3 Three-generation image intensifier







**Fig. 4** NSC1105 HDR sensor

cycle in this paper. At the same time, FPGA also needs to send three signals to choose the operating mode of the NSC1105 CMOS sensor: ZOOM, BIN, and HD. Each signal represents an operating mode. When selecting the operating mode, just need to control one of the signals to be a high level and the other signals to be a low level. The target resolution of this system is  $1280 \times 1024$ , which corresponds to the HD mode. Therefore, the HD signal is pulled high, ZOOM and BIN remain low.

After being driven correctly, the NSC1105 will output two differential analog signals. As all known, FPGA cannot process analog signals. So, we must convert the differential analog signal into a digital signal before processing. Comprehensively considering the AD device's power consumption and conversion rate, this paper selects the AD9649-20 device to complete the work of analog-to-digital conversion. FPGA needs to configure three signals for the AD chip through the serial peripheral interface (SPI) and provide two differential clock signals, thus making it work in the correct mode. The AD device converts the analog signal into a 14-bit digital signal before the digital signal is processed.

## 4.2 Image enhancement processing

Using guided filter to process the original image  $I$  to get the base layer image  $I_b$ . By subtracting the original image and the base layer image in the logarithmic domain, the detail layer image  $I'_d$  can be obtained. Enhancement process is divided into two parts: base layer image enhancement and detail layer image enhancement.

**Base layer image:** Performing logarithmic processing on  $I_b$  and then compressing the grayscale to reduce the grayscale ratio of the image, which can effectively increase the image's brightness and facilitate human eye observation.

**Detail layer image:** It can be seen from Sect. 3.2 that the decomposed detail layer image is in logarithmic form, so it must be exponentially calculated before the enhancement of the detail layer. After calculating the detail layer image, the corresponding gain coefficient may be determined and adaptively enhanced according to the intensity characteristics of each region.

After the base layer image and the detail layer image are enhanced, the two images are added together to obtain the final image, as shown in Eq. 13.

Images before and after image enhancement implemented in hardware are shown in Fig. 5(a-b)

### 4.3 Tone mapping

The image processed by the AD device has a 14-bit width that exceeds the display range of the OLED, which only has a display range of 8-bit. Therefore, the image must be compressed appropriately to be displayed on the OLED device. A hardware implementation of a TMO called a novel histogram adjustment method was proposed by Duan et al. (Duan et al. 2010), which is feasible and effective, and this method was verified in the HDR-ART-iSt designed by Lapray et al. (Lapray et al. 2016). This paper also uses this method to compress the processed 14-bit HDR image. This TMO uses the following process to compress the HDR image  $E_{x,y}$  to a displayable range  $D_{x,y}$ :

$$D_{x,y} = H \times (D_{\max} - D_{\min}) + D_{\min} \quad (14)$$

$$H = \frac{\ln(E_{x,y} + \tau) - \ln(E_{x,y(\min)} + \tau)}{\ln(E_{x,y(\max)} + \tau) - \ln(E_{x,y(\min)} + \tau)} \quad (15)$$

In Eq. 15,  $E_{x,y(\max)}$  and  $E_{x,y(\min)}$  represent the maximum and minimum of the HDR image, respectively. And  $\tau$  is a parameter that is inversely proportional to the display brightness, so the smaller the value of  $\tau$ , the brighter the display image. Considering that the OLED is an 8-bit device, so the  $D_{\max}$  and  $D_{\min}$  are set as 255 and 0, respectively. And then, needing to count that the maximum  $E_{x,y(\max)}$  and minimum  $E_{x,y(\min)}$  of the HDR image in Eq. 14, which are used to calculate the value of  $H$ . To ensure the stability and real-time performance of the TMO, the information calculation parameter  $H$  using the previous frame data is used for the tone mapping of the current frame by utilizing the characteristics of the correlation between successive two frames of images. In this way, HDR images can be compressed into displayable images. In addition, logarithmic operations are implemented based on LUT. Since the IC MOS camera designed in this paper is used in a LLL environment, higher image brightness is helpful to distinguish image details, so the parameter  $\tau$  is set as 1 in this paper.

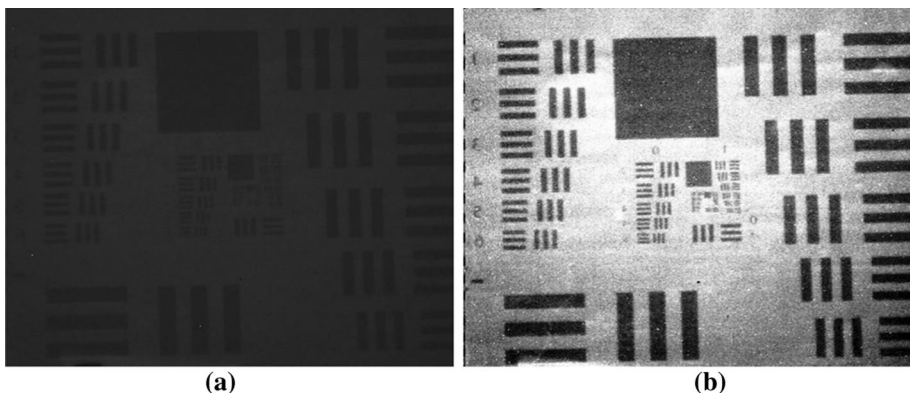


Fig. 5 a Image without enhancement, b Image enhancement in hardware

#### 4.4 OLED display

This system selects MDP02BBPFC OLED as the video display device, and its resolution is  $1280 \times 1024$ . To control the display effect of OLED, we use the method of controlling cathode voltage and OLED temperature to control the performance of OLED. The voltage divider resistor of MAX1763ESA adjusts the cathode voltage. Because the chip is a programmable logic device, its voltage divider resistance can be changed by changing its control signal. The operating temperature of OLED is an important parameter that affects operating characteristics, so a temperature sensor is equipped to monitor the operating temperature in real-time. When the temperature is higher than the optimum temperature, the cathode voltage is adjusted to reduce the temperature, vice versa.

The OLED used in this paper is RGB full-color 8-bit data output. Because this system uses black-and-white image output, three output signals must be given the same value simultaneously. After the mapping process in Sect. 4.3, although the image information is compressed, there is less data loss. Compared with the original image, the mapped image still has better-detailed information, and its dynamic range is more comprehensive than the original image's.

#### 4.5 System hardware analysis

The system proposed in this paper is described in Verilog language and integrated with Spartan 6 series FPGA using Xilinx's ISE Design Suite 14.7 toolset. This paper also uses Modelsim to test the system's image processing performance to verify the system's function.

Table 1 lists the comprehensive report of this system. It can be seen that the camera system designed in this paper consumes fewer FPGA resources, leaving enough space for future improvements.

### 5 Experiment and result

In this section, the experiments of the FPGA-based high dynamic ICMOS camera will be implemented. The physical picture of the FPGA-based HDR ICMOS camera is shown in Fig. 6.

Experiment 1: Test the high dynamic characteristics of the camera. To test the dynamic characteristics of the camera, we designed a scene with high contrast. In this case, the third-generation image intensifier should be removed to avoid being damaged due to oversaturation.

**Table 1** Summary of hardware synthesis report

DEVICE	SPARTAN-6-XC6SLX150
LUT	11,948/92152 (12.97%)
Registers	32,486/184304 (17.62%)
Memory	5123/21680 (23.63%)
PLL	3/6 (50%)

Fig. 6 HDR ICMOS camera



In this scene, the lamp's light on the desk contrasts sharply with the dim environment. Using the designed ICMOS camera and an ordinary camera to photograph the designed scene and then compare and analyze the images obtained from the two cameras. The images taken by the two cameras are as follows:

We can see from Fig. 7a that the text on the carton below the lamp and the edge information of the stacking boxes on the upper left of the picture are difficult to recognize, which shows that the ordinary camera's ability to work in a high-contrast environment is poor, and the images tend to lose a lot of details. In Fig. 7b, the text on the carton is recognizable. The details of the objects on the left side of the carton have been enhanced to see the texture. And we can distinguish the stacked boxes clearly.

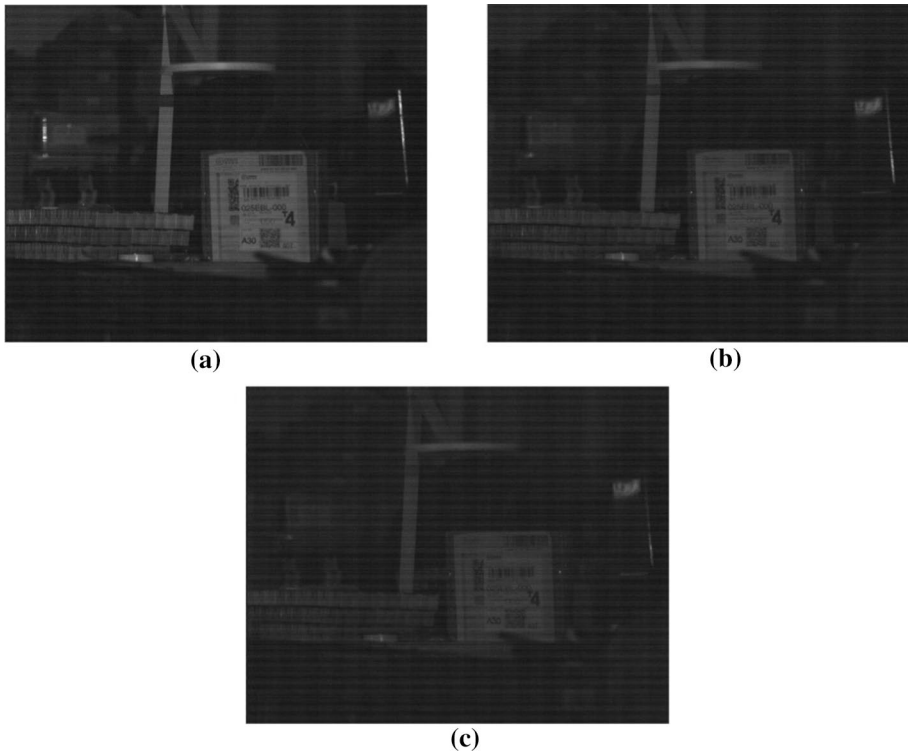
By comparison, the HDR camera designed in this paper has an excellent performance in a high-contrast environment that can capture more detailed information, better reflecting the original appearance of things.

Experiment 2: Test the performance of the HDR ICMOS camera under a LLL environment. The purpose of this experiment is to test the ultimate performance of HDR ICMOS cameras, that is, the minimum illumination that can distinguish things clearly. The scenes in this experiment are the same as shown in Fig. 7a. Images were captured under the illumination of  $5 \times 10^{-3}$  lx,  $1 \times 10^{-3}$  lx,  $7 \times 10^{-4}$  lx.

As shown in Fig. 8a, these boxes' text and outline information can be recognized clearly under this illuminance. Hence, the system has an excellent ability to distinguish things under this illuminance. In the  $1 \times 10^{-3}$  lx environment, the text on the box can still be distinguished. But the stacked boxes in the upper left corner of the image can only display the



Fig. 7 a Image captured by ordinary camera, b Image captured by ICMOS camera



**Fig. 8** **a** Image captured under  $5 \times 10^{-3}$  lx, **b** Image captured under  $1 \times 10^{-3}$  lx, **c** Image captured under  $7 \times 10^{-4}$  lx

blurred outline information, and it is no longer possible to distinguish image details. In the  $7 \times 10^{-4}$  lx environment, it is no longer possible to recognize the text on the box, and the information on the stacking boxes in the upper left corner has been completely lost.

## 6 Conclusion

In this paper, an HDR ICMOS camera based on adaptive detail enhancement and guided filter is designed and implemented to avoid artifacts and solve the storage problem in multi-images fusion. Furthermore, the proposed algorithm can effectively extend the dynamic range of the ICMOS camera, which reaches up to 80 dB, so that the performance of the LLL night vision system has been significantly improved. Unlike other HDR imaging methods, we focus on the HDR CMOS sensor, which saves the time in image processing and realizes real-time video display. Finally, this ICMOS camera provides an HDR live video with a resolution of  $1280 \times 1024 @ 60\text{fps}$  and has a good performance in a  $10^{-4}$  lx environment. What's more, the HDR ICMOS camera achieves a wider dynamic range than the 14-bit camera, so it can capture more detailed information in a LLL environment, which improves the night vision performance.

Since this paper uses RGB full-color 8-bit OLED as the display device of this system, the use of pseudo-color output video images is the following work.

**Funding** This work was supported by the Shandong Provincial Natural Science Foundation (Grant No. ZR2019MF010), the Science and Technology Program for Higher Education of Shandong Province (Grant No. J18KA320), and the National Defense Basic Research Program of China (Grant No. JCKY2018208B016).

## Declarations

**Conflict of interest** The authors declare no conflicts of interest.

## References

- Abolbashi, M., Magalhaes, F., Araujo, F.M.M., et al.: High dynamic range compressive imaging: a programmable imaging system. *Opt. Eng.* **51**(7), 071407 (2012). <https://doi.org/10.1117/1.OE.51.7.071407>
- Ahmad, A., Amira, A., Nichoil, P., et al.: FPGA-based IP cores implementation for face recognition using dynamic partial reconfiguration. *J. Real-Time Image Process.* **8**(3), 327–340 (2013). <https://doi.org/10.1007/s11554-011-0221-x>
- Bae, S.M., Paris, S., Durand, F.E.: Two-scale tone management for photographic look. *ACM Trans. Gr.* **25**(3), 637–645 (2006). <https://doi.org/10.1145/1141911.1141935>
- Branchitta, F., Diani, M., Corsini, G., et al.: New technique for the visualization of high dynamic range infrared images. *Opt. Eng.* **48**(9), 096401 (2009). <https://doi.org/10.1117/1.3216575>
- Cadik, M., Wimmer, M., Neumann, L., et al.: Evaluation of HDR tone mapping methods using essential perceptual attributes. *Computer Gr.-UK* **32**(3), 330–349 (2008). <https://doi.org/10.1016/j.cag.2008.07.007>
- Duan, J., Bressan, M., Dance, C., et al.: Tone-mapping high dynamic range images by novel histogram adjustment. *Pattern Recogn.* **43**(5), 1847–1862 (2010). <https://doi.org/10.1016/j.patcog.2009.12.006>
- Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Gr.* **21**(3), 257–266 (2002)
- Farbman, Z., Fattal, R., Lischinski, D., et al.: Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Gr.* **27**(3), 67 (2008). <https://doi.org/10.1145/1360612.1360666>
- Hassan, F., Carletta, J.E.: An FPGA-based architecture for a local tone-mapping operator. *J. Real-Time Image Process.* **2**(4), 293–308 (2007). <https://doi.org/10.1007/s11554-007-0056-7>
- Karadzovic-Hadziabdic, K., Hasi-Telalovi, J.H., Mantiuk, R.K.: Assessment of multi-exposure HDR image deghosting methods. *Computer Gr.-UK* **63**, 1–17 (2017). <https://doi.org/10.1016/j.cag.2017.01.002>
- Katsaggelos, A.K., Biemond, J., Schafer, R.W., et al.: A regularized iterative image restoration algorithm. *IEEE Trans. Signal Process.* **39**(4), 914–929 (1991). <https://doi.org/10.1109/78.80914>
- Lapray, P.J., Heyrman, B., Ginhac, D.: Hardware-based smart camera for recovering high dynamic range video from multiple exposures. *Opt. Eng.* **53**(10), 102110 (2014). <https://doi.org/10.1117/1.OE.53.10.102110>
- Lapray, P.J., Heyrman, B., Ginhac, D.: HDR-ARtiSt: an adaptive real-time smart camera for high dynamic range imaging. *J. Real-Time Image Process.* **12**(4), 747–762 (2016). <https://doi.org/10.1007/s11554-013-0393-7>
- Lee, S.H., Woo, H., Kang, M.G.: Global illumination invariant object detection with level set based bimodal segmentation. *IEEE Trans. Circuits Syst. Video Technol.* **20**(4), 616–620 (2010). <https://doi.org/10.1109/TCSVT.2010.2041824>
- Mendis, S.K., Kemeny, S.E., Gee, R.C., et al.: CMOS active pixel image sensors for highly integrated imaging systems. *IEEE J. Solid-State Circuits* **32**(2), 187–197 (1997). <https://doi.org/10.1109/4.551910>
- Nosko, S., Musil, M., Zemic, P., et al.: Color HDR video processing architecture for smart camera. *J. Real-Time Image Process.* **17**(3), 555–566 (2020). <https://doi.org/10.1007/s11554-018-0810-z>
- Popovic, V., Seyid, K., Pignat, E., et al.: Multi-camera platform for panoramic real-time HDR video construction and rendering. *J. Real-Time Image Process.* **12**(4), 697–708 (2016). <https://doi.org/10.1007/s11554-014-0444-8>
- Tang, X.D., Qian, Y.S., Kong, X.Y., et al.: A high-dynamic range CMOS camera based on dual-gain channels. *J. Real-Time Image Process.* **17**(3), 703–712 (2020). <https://doi.org/10.1007/s11554-019-00877-8>

- Tsai, C.Y., Huang, C.H.: Real-time implementation of an adaptive simultaneous dynamic range compression and local contrast enhancement algorithm on a GPU. *J. Real-Time Image Process.* **16**(2), 321–337 (2019). <https://doi.org/10.1007/s11554-015-0532-4>
- Winnemöller, H., Olsen, S.C., Gooch, B.: Real-Time video abstraction. *ACM Trans. Gr.* **25**(3), 1221–1226 (2006). <https://doi.org/10.1145/1141911.1142018>
- Zhou, Z.Q., Dong, M.J., Xie, X.Z., et al.: Fusion of infrared and visible images for night-vision context enhancement. *Appl. Opt.* **55**(23), 6480–6490 (2016). <https://doi.org/10.1364/AO.55.006480>
- Zuo, C., Chen, Q., Liu, N., et al.: Display and detail enhancement for high-dynamic-range infrared images. *Opt. Eng.* **50**(12), 127401 (2011). <https://doi.org/10.1117/1.3659698>
- Jacquot, B.C., Johnson-Williams, N.: Multiple-samples-method enabling high dynamic range imaging for high frame rate CMOS image sensor by FPGA and co-processor. 8th Conference of Optics and Photonics for Information Processing, 9216: 921609(2014). doi: <https://doi.org/10.1117/12.2062284>
- Kuang, J.T., Liu, C.M., Johnson, G.M.: Evaluation of HDR image rendering algorithms using real-world scenes. 30th International Congress of Imaging Science. 256- (2006).
- Liu, C., Freeman, W.T., Szeliski, R., et al.: Noise estimation from a single image. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. **1**(0), 901–908(2006).
- Mann, S., Lo, R.C.H., Ovtcharov, K., et al.: Realtime HDR (High Dynamic Range) video for eyetap wearable computers, FPGA-based seeing aids, and glasses (EyeTaps). 2012 25th IEEE Canadian Conference on Electrical and Computer Engineering. (2012). doi: <https://doi.org/10.1109/CCECE.2012.6335012>
- Pham, C.C., Ha, S.V.U., Jeon, J.W.: Adaptive guided image filtering for sharpness enhancement and noise reduction. 5th Pacific-Rim Symposium on Video and Image Technology. 7087 323–334(2011).
- Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. 6th International Conference on Computer Vision. 839–846 (1998). doi: <https://doi.org/10.1109/ICCV.1998.710815>
- Zhao, H., Shi, B.X., Fernandez-Cull, C., et al.: Unbounded high dynamic range photography using a modulo camera. 2015 IEEE International Conference on Computational Photography. 137: 46 (2015).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.