64 × 64 GM-APD array-based readout integrated circuit for 3D imaging applications

Jin WU1, Zhiming QIAN1, Yang ZHAO1, Xiangrong YU1, Lixia ZHENG1,2* & Weifeng SUN2

1Branch School of Southeast University, Wuxi 214135, China; 2National ASIC System Engineering Research Center, Southeast University, Nanjing 210096, China

Received 30 July 2018/Revised 24 October 2018/Accepted 29 November 2018/Published online 15 April 2019

Abstract Using the high sensitivity of the avalanche photodiode (APD) detector operated in the Geiger-mode (GM), an array readout integrated circuit (ROIC) comprising a two-segment time-to-digital converter (TDC) is employed for wide-dynamic time interval measurement, where a 1-bit low-segment TDC is implemented by discriminating a single-phase clock period. The proposed 64 × 64 GM-APD array ROIC fabricated using Taiwan semiconductor manufacturing company (TSMC) 0.18-µm complementary metal oxide semiconductor (CMOS) technology can operate at a maximum frequency of 500 MHz provided by an external phase-locked loop clock. The time resolution is reduced to <1 ns along with a maximum range of 4 µs; the differential non-linearity (DNL) and integral non-linearity (INL) are restricted to approximately −0.15 to 0.15 least significant bit (LSB) and −0.3 to 0.32 LSB, respectively; and the power consumption is 490 mW under a frame rate of 20 kHz. The developed ROIC is successfully used in imaging applications in two different ways.

Keywords single photon detection, avalanche photodiode, readout integrated circuit, time-to-digital converter, ranging imaging application


1 Introduction

Using the Geiger mode avalanche photodiode (GM-APD) detector and readout integrated circuit (ROIC), the single-photon detection technique can be employed for accurate time-of-flight (TOF) measurement and photon counting [1, 2]. It has been widely used in various applications, such as laser ranging and three-dimensional (3D) imaging [2], weak light detection, and accurate identification and tracking of camouflage targets [3].

The active quenching circuit (AQC) within the ROIC is capable of sensing the avalanche current pulse triggered by a single photon, and the generated STOP signal indicates the arrival of the specific photon. The TOF is defined as the time interval between two transient rising edges of EN and STOP signals, where the global synchronous EN signal provided by the delay controller is used to activate the detection system at the moment when the infrared photon laser pulse is emitted, and the STOP signal arrives randomly in the time domain. The ROIC can quantize the random TOF using the internal time-to-digital converter (TDC), and the measured TOF is directly related to the spatial distance L. The least significant bit (T_LSB) is defined as the resolution in time quantization, for a time resolution of 1 ns. The corresponding

*Corresponding author (email: zhenglx@seu.edu.cn)
space resolution is approximately 15 cm. For extending the dynamic range in a time interval or space distance, more digital conversion bits are needed.

Owing to the serious non-ideal effects in the array system, including noise interference, clock skew, and delay mismatch between critical signal paths, the performances of the array TDC are far inferior to those of the single-pixel TDC with the same circuit structure. With the continuous increase of the array scale, the performances improvement of the ROIC with regard to the accuracy, resolution, linearity, and dynamic range are difficult, and the system properties are strongly dependent on the ROIC array architecture.

For simultaneously satisfying the detection properties of accurate and range, the single TDC with a segment architecture is widely used. For array applications, a simple and effective architecture is multiple identical pixels in parallel with an embedded completed TDC [4]. In this structure, the clock phase number and the total digital bits of the low-segment TDC are difficult to extend owing to the limitation of the pixel area; thus, for a low resolution, a high clock frequency is needed, and the power consumption is thus increased. Because all the global logical control signals and counting clock signals must be transported to all the pixels, the interference noise between these critical signals is increased significantly.

In contrast, for the completely shared TDC architecture proposed by the Lincoln laboratory [5], all pixels are only used to latch the transient data of the TOF from one globally shared TDC (or several). Although the power consumption is significantly reduced, the reduction of the pixel area is not remarkable, owing to the necessary storage register array embedded in each pixel. The cross-talk between the high-frequency and high-density signal lines is considerable; thus, this completely shared architecture is only suitable for small-array applications. By connecting several small arrays of this kind in parallel, the array scale can probably be expanded.

The third array architecture is a trade-off between the previous two complementary structures [6, 7], where only a low-segment TDC is located outside pixels and shared by all (or some) pixels. In contrast, the high-segment TDC is located in the interior of each pixel. The global voltage controlled oscillator (VCO) clocks is directly utilized to drive the high-segment TDC, while its uniformly distributed multiphase signals are utilized to discriminate the clock period. If the transient phase states can be freely latched and stored by all array pixels when the corresponding STOP signal arrives, the shared low-segment TDC is implemented; thus, the dominant part of this TDC is constructed by the VCO located outside the pixel without area restriction. A fine resolution can be obtained with a relatively low clock frequency with a large number of multiphase clocks, and the low power consumption is beneficial for array extension.

In this study, the proposed 64 × 64 array infrared ROIC with the 1-bit low-segment TDC shared architecture is implemented using TSMC 0.18-μm CMOS technology. This unique structure is used to eliminate high frequency interference between multi-clock lines, and nonlinear errors are restricted by using optimal critical circuits and delay matching. The ROIC is designed to measure the photon TOF for range imaging and measure the photon quantity for intensity imaging.

2 Array ROIC implementation

2.1 System architecture

The ROIC system operated in the gated enable mode is illustrated in Figure 1, where all the global control signals are provided by the host. The RESET signal is used to reset each frame of the system to ensure accurate detection. The START, EN, and other clock signals are aligned at the rising edge; in this way, the initial phase mismatch between the photon emission and photon counting is completely eliminated in time measurement. For the array ROIC, the high-segment TDC is implemented by an 11-bit linear feedback shift register (LFSR) synchronous counter embedded in each pixel. For the low-segment TDC, if a VCO or voltage controlled delay line (VCDL) is used to provide the multi-phase clock, a corresponding decoding circuit is required. This incurs an area cost, and dynamic decoding may increase the bit error rate. To reduce the area and obtain a low error rate, it is preferential for the low segment to use a 1-bit
Figure 1  (Color online) Construction of the TOF detection system based on single photon detection.

Figure 2  (Color online) H- tree structure. (a) Pixel array; (b) the signal transmission path.

The TDC constructed by a two-phase HCK clock. The discriminated data with half-clock period resolution are stored in each pixel by a single D flip-flop (DFF). For this unique 1-bit subsection TDC, the VCO or VCDL is unnecessary. However, for extending the low-segment TDC from 1 bit to 3 bits or more, the multiphase clocks provided by the VCO or VCDL should be employed to replace the two-phase clock.

To ensure timing matching and reduce the clock jitter, the counting start signal EN, the counting termination signal STOP, the AQC reset signal REC, and the high-frequency counting clock signal HCK should be transmitted through the nested H tree networks to the individual pixel located within the array. The pixel array designed according to the aforementioned rules is shown in Figure 2(a). In this way, the clock skew concerning different pixels can be reduced to a minimum level through the exact time matching among different signal delay paths. The systematic offset in the time delay is nearly identical for all pixels. It is equivalent to a common-mode deviation and has little influence on the relative resolution between the pixels, that is, the quality of 3D imaging is not deteriorated.

By inserting the buffer, the transmission line can be divided into shorter segments to suppress the accumulation of the online capacitance. Meanwhile, the buffer can be used for signal level recovery and waveform shaping; thus, significantly reducing the rising/falling edge time and transmission delay. The rising/falling edge time of the signal within 10% of the signal cycle is necessary; thus, the size of the
driver should be reasonably designed according to the driving load. As shown in Figure 2(b), the clock line lengths of the 64 \( \times \) 64, 32 \( \times \) 32, 16 \( \times \) 16, 8 \( \times \) 8, 4 \( \times \) 4 and 2 \( \times \) 2 arrays are 2, 0.6, 0.3, 0.15, 0.05, and 0.05 mm, respectively. Therefore, the total clock line length from the root node to the leaf node is 3.15 mm, and a nine-level buffer driver must be inserted.

2.2 Timing diagram

As mentioned previously, the proposed ROIC is operated in a passive and configurable frame frequency mode, where a multi-channel high-precision delay controller is used as the buffer interface between the personal computer (PC) host and the ROIC to achieve an accurate delay matching between different modules. The key timing sequence is shown in Figure 3.

In each frame, the TDC starts counting from the rising edge of the EN signal, and the REC pulse at the start of the EN signal is used to reset the GM-APD for detection within a frame. Within the detectable window of 5 \( \mu \)s controlled by the EN gated signal, the dead time of approximately 12 ns is mainly consist of the REC pulse width, where the STOP signal cannot be generated within this region. The maximum range of the TOF is limited by the pulse width of the EN signal. During the current frame when EN is turned to a low level, the TOF data for the obtained pixel in each gated window are delivered outside in series for further processing.

2.3 Critical circuits

The completed single-pixel circuit is shown in Figure 4. To provide a sufficient overdrive voltage swing for the avalanche photodiode (APD), the AQC is operated under a supply voltage of 5 V (\( V_{DD} \)), while the power supply for the other logical circuits is 1.8 V. Thus, a voltage level shift circuit for interconversion is needed. Via current-to-voltage conversion, a voltage pulse is established in a sensed resistor or capacitor [8]. For reducing the area and achieving a fast transient response, the small anode-to-ground intrinsic capacitance \( C_C \) of the InGaAs APD is directly utilized to sense the avalanche current triggered by the photon. If the amplitude of the sensed voltage exceeds over the critical point of the comparator, the rising edge of the STOP signal is immediately generated; thus, the inverted \( STOPb \) signal can be used to turn on the M3 for quickly increasing the potential of node IN to \( V_{DD} \). In this way, the APD is immediately cut off to quench the avalanche current. For improving the signal-to-noise ratio, the preset critical point of the comparator is usually defined by a reference voltage, which should be high enough to suppress the possible spurious triggering or noise triggering. For reducing the pixel area, a single-ended inverter is used as a simple comparator to estimate the sensed voltage level, where the critical transition point is slightly larger than the threshold voltage of the n-metal-oxide-semiconductor field-effect transistor (NMOSFET) if its size is significantly larger.

In the two-segment TDC structure, the pixel-exclusive high-segment TDC is composed of an 11-bit LFSR, which includes a DFF chain and a feedback control loop. The loop closed by the XNOR gate according to the primitive polynomial, which can provide the total pseudo-random count up to \( 2^{11} - 1 \).
states (all “1” outputs are banned). The low-segment 1-bit TDC is implemented by a globally shared two-phase high-frequency clock signal HCK, and the DFF0 used to latch the data of the 1-bit TDC should be located in each individual pixel. The high-segment and low-segment TDCs are connected through multiplexer (MUX1). By simply breaking off the feedback loop, the LFSR can immediately shift from the counter mode to the data latching/series transmission mode. The data of the LFSR are transmitted to the outside of the chip, and static decoding is performed via the look-up table method; thus, the bit error rate is almost zero. Thus, as comparison with the binary synchronous counter, the dual-mode LFSR can significantly reduce the pixel area, the data-latching time, and the data error rate. To improve the frame clock frequency, the total array is divided into several uniform subsections for parallel data transportation, where the pixel data in each blocks are serially transported outside.

At the beginning of each frame, under the control of the RESET signal, the TDC circuit first performs a reset operation, clearing all the 12-bit data to 0 for setting the initial state. MUX1 to MUX4 are used to select different clocks and control signals in two operation modes. When EN = 1, the I ports of MUX1 and MUX4 are switched on. The high-frequency HCK is loaded into the LFSR for counting and simultaneously transmitted to the DFF0 for obtaining a fine phase resolution. When the STOP signal is triggered to high, it is selected as the control signal for freezing the data, and the LCK is selected for data transportation during EN = 0 instead of the previous HCK. The sensed data are latched in the register of the corresponding pixel. The LFSR counter in the EN = 1 period is switched to the shift register in the EN = 0 period by configuring MUX1 and MUX2. The data are collected by the data-acquisition card, transmitted to the PC, and then decoded by the PC for subsequent signal processing.

A pixel occupies an area of 50 \, \mu m \times 50 \, \mu m, which is exactly coincidence with the APD size. True single-phase clock DFFs with a relatively small area and short setup-hold time are employed for reducing the power consumption and restricting the bit error rate in data latching [9]. For random STOP triggering, if the rising edges of STOP and HCK are far closed as less than the setup-hold time of the DFF, a digital code error in latching the counting clock will probably occur. However, for the specific DFF structure, the minimal setup-hold time is difficult to further reduce. Under accurate delay matching control, a composite latching structure is employed for error code rejection. As shown in Figure 4, a buffer with a delay of \( t_{db} \), which is slightly larger than that of the transmission gate (TG), i.e., \( t_{dg} \), is inserted between MUX3 and DFF, so that the output of the TG is stable when the low-segment TDC latches the high-frequency clock phase. Furthermore, the setup-hold time of the TG is obviously smaller than that of the DFF; thus, the low-segment TDC can lock the counting clock phase with a low error rate. The composite sampling of TG+DFF is obviously superior to that of a single DFF. Additionally, the important delay stages or logical gates are all fully custom-designed to reduce the area and improve the matching property.

Owing to the simplicity of the circuit, the two-segment arrangement of the 11+1 structure can effectively restrain the non-linearity and error codes if the circuit timing sequence is optimized. However, the low-segment TDC has slightly less digital conversion bits than other work. A fine or accurate resolution corresponds to a high counting frequency and large power consumption.
3 Measurement and discussion

The proposed GM-APD 64 × 64 array ROIC is implemented via a TSMC 0.18-µm CMOS process, and the total chip area is 4.5 mm × 4.3 mm. For a single pixel, the area is 50 µm × 50 µm and the distance between pixels is 50 µm. The synchronous signal module (SYNC) and the clock controller are placed on the right side of the chip, covering area of 30 µm × 40 µm and 110 µm × 35 µm, respectively. A micrograph of the ROIC chip and the corresponding PCB-level test circuit are presented in Figure 5. In each pixel, the APD is interconnected with the AQC via the indium bump bonding technique.

The ROIC chip is first measured in the testing mode without cooperating with the APD detector array, where the EN, STOP and other related control signals are externally provided. The STOP signal under this condition is called STOP_TEST. The TOFs of different pixels both are synchronous under the same global control signals when the signals delay paths are consistent.

HCK is loaded into the array system during the EN detectable window by the field-programmable gate array (FPGA). Its rising edge is aligned with the EN signal for eliminating the initial offset. Under this condition, the time to be tested is first quantized by HCK with a value of $H$, which is the count value of the coarse measurement. The effective coarse time period is $(H-1)T_{clk}$ because the first HCK edge that starts to be counted is aligned with the EN signal, indicating that one counting clock period should be removed. The residual error time is measured as half of the counting period. When the value of the counting clock is latched high (1), the residual value is less than a half-cycle; thus, the 1-bit TDC data are 0. Conversely, if the value of the counting clock is latched low (0), the 1-bit TDC data are 1. In the end, the measured result is the sum of the two.

The typical waveform of random data within a frame is presented in Figure 6. All the global control signals provided outside follow the principles as shown in Figure 3. The counting frequency is 500 MHz (2 ns cycle). The EN signal is provided by the FPGA, and the STOP_TEST is provided by a digital delay to set the time interval freely. So, the transport delay mismatch between EN and STOP_TEST signal can hardly be avoided owing to the significantly different signal transmission paths. In the chip, there are 32 data output ports, and each the 128 (4 × 32) pixels share one port to serially output data. To increase the readability of the data, a word signal (WORD) is output every 12 clock cycles. The time interval is preset at 101 ns for eight adjacent pixels. By using WORD synchronous signals to recognize the logical level of the output waveform shown in the oscilloscope, the original 12 bits two-segment readout digital codes of these tested pixels are all 00100111111-0 as shown in Figure 6. According to the coding table and the calculation method described above, the decoded data is 53; thus, the measured result is
evaluated as \((53 - 1) \times 2 \text{ ns} + 1 \text{ ns} = 105 \text{ ns}\). An absolute offset error of 4 ns is formed owing to the signal transport delay mismatch between \(EN\) and \(STOP\). Fortunately, the offset error is almost fixed, and the impact on the relative accuracy is so small that it can be eliminated before the data processing. Additionally, in the practical operation mode, the \(STOP\) signal generated by the APD in each pixel only drives one pixel, whereas in the testing mode, the \(STOP\) signal must drive multiple pixels. Therefore, the offset error measured in the test mode is overestimated.

According to a more extensive test analysis, the measured full-scale linearity of the ROIC is presented in Figure 7, where the preset time interval linearly varies from 200 ns to 4 \(\mu\)s with a step of 200 ns. Obviously, the system has superior linearity over the full scale range.

For the nonlinear characteristic of the TDC, DNL mainly comes from the non-uniformity of the quantization steps caused by the transient variation of the clock frequency, and INL is primary arising from the long-term accumulation of errors of the clock period. By using a phase-locked loop clock and the matched timing control shown in Figure 4, a low bit error rate and good linearity are obtained in the proposed ROIC. The low-segment TDC uses a single minimum conversion bit with no decoding error, and the uniformity of the two phases can be well obtained. For the input TOF data between 1 and 1.1 \(\mu\)s, the measured DNL is \(-0.15\) to \(0.15\) LSB, and the INL is \(-0.3\) to \(0.32\) LSB. The detailed results are presented in Figure 8, respectively. As a theoretical prediction, with the increase of the counting clock frequency, the measured linearity of the TDC is deteriorated. Under different operation conditions, the DNL and INL are restricted within the range of \(-0.5\) to \(0.5\) LSB, indicating that the quantized noise is dominant compared with other nonlinear noises; thus, the effective number of bits (ENOB) keeps the initial setting of 12 bits unchanged.
Comparisons of the primary performances between the proposed array ROIC and similar ROICs from previous studies are presented in Table 1 [10–13]. The results show that our ROIC exhibits sufficient merits in a wide measurement range, as well as a small nonlinear error. Compared with the total power of the system given in [11], the total power consumption of our ROIC is acceptable. In the proposed array ROIC, the high segment adopts an 11-bit LFSR pseudo-random counter for range extension, and the low segment adopts 1-bit interpolation to realize half-periodic resolution. Under operation at a 500 MHz HCK counting frequency, the average time resolution is approximately 1 ns, and the measurement range of the TDC reaches to 4 µs; thus, the pulse width of the EN signal is set around 5 µs (adjustable). In imaging applications, the measured data of different results among the various pixels of the array are more critical than the common results of individual pixel. The proposed ROIC can provide a suitable resolution for long-distance object detection.

4 Application

The ROIC chip and the InGaAs detector array are interconnected via indium bump bonding and packaged into a camera box, which is connected to a commercial near-infrared-enhanced lens through a standard C-interface. A filter is placed in front of the lens, and a 3D laser camera is constructed. Compared with other imaging systems, the most outstanding advantage of this system is the simultaneous measurement of the 3D geometric information of the target and the intensity information of the laser echo signal.

A distance image is realized according to the principle of photon timing distance measurement. By quantifying the distance of the target object through the pixel TOF, contour imaging can be performed. If the object is approximately stationary, the test data between the frames are similar or even identical. Thus, the average processing of the counting data in multiple frames can be utilized to suppress the noises and errors. Furthermore, within a frame, a pixel TOF time shorter than the EN window is sensed, indicating that the photon is turned back. Otherwise, if the pixel TOF is exactly equal to the EN pulse width, no photon has returned. Thus, by counting the photon return times of the specific pixel in multiple successive frames, and with the help of the background processing, the intensity information related to the detected objects can be obtained. Clearly, for improving photon counting property of the ROIC, the frame frequency should be sufficiently high. The single-frame photon timing detection and multi-frame photon counting detection can be performed by the ROIC without mutual restriction.

Single-photon and multiple-photon detection have different avalanche probabilities for triggering the detectors, and the surface of the specific target has reflectance. If more photons are returned by intensive light sources with a small distance, the probability of triggering the detector is improved. The reflectivity of clothes for the laser photon varied significantly owing to the different materials of the objects; thus, the photon numbers detected via multi-frame detection differ significantly at nearly the same distance.
Table 2 Parameters of the laser

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photon wavelength</td>
<td>1064 nm</td>
</tr>
<tr>
<td>Photon pulse width</td>
<td>10 ns</td>
</tr>
<tr>
<td>EN frame frequency</td>
<td>10 kHz</td>
</tr>
<tr>
<td>Laser average power</td>
<td>0–200 mW (adjustable)</td>
</tr>
<tr>
<td>Transmitter-receiver field angle</td>
<td>7°</td>
</tr>
</tbody>
</table>

Therefore, clothes can be well distinguished in the intensity image; however, they can hardly be recognized in the range image, as the two objects have almost the same space distance. Using the developed 3D imaging receiver, an indoor 3D imaging experiment is performed. The critical parameters of the 3D camera are presented in Table 2.

The 3D imaging experimental results for range and intensity images at different distances are presented in Figure 9. For a long-distance test, the photons that can be detected by the detector are significantly reduced owing to the serious attenuation of the photons; thus, both the intensity image and the distance image are poor. In an extremely weak light environment, the range image is very sensitive to distance variation, but the sensitivity of the intensity image is relatively low. The system essentially performs a single-frame detection for the range image, which is similar to differentiation signal processing; thus, it is sensitive to external disturbances and noise but can immediately respond to dynamic variations of the target. Photon counting imaging technology performs multi-frame detection, corresponding to integral signal processing; thus, it is superior with regard to noise suppression. These two complementary methods combined are effective for achieving high-precision 3D imaging.

In the field of practical astronomical observation, when the targets are far beyond the maximum single
detectable window, a sufficiently large delay should be inserted between the laser emitting and photon counting, where the inserted delay is related to the object distance and can be determined by searching the detectable window via the iteration method. The rough distance of the object can also be evaluated using this large preset delay. To detect debris approximately 800 km away, the laser is continuously emitted for 10 s at a frame rate of 1 kHz. In every frame, the counter starts to count after a specific delay from photon emission, as mentioned previously, and the measured $t_{\text{count}}$ is shorter than the duration of the EN pulse, the object related to the specific pixel is detected. The area occupied by the object is more likely to reflect detectable photons than the vacuous space. According to the distribution of the detected number for corresponding pixel elements within 10000 laser emissions, we can obtain a rough outline of the object as shown in Figure 10.

5 Conclusion

Utilizing a unique 1-bit low-segment TDC shared architecture for restricting the high-frequency clock interference, as well as the optimal critical circuit with exact delay matching for eliminating nonlinear errors, the proposed $64 \times 64$ array ROIC implemented via TSMC 0.18-µm CMOS technology is capable of detecting a random TOF with nanosecond resolution and 4 µs dynamic range. The distance range can be extended via the detectable-windows searching technique, and the maximum range corresponding to the inserted delay is only restricted by the laser power. The ROIC designed to measure the photon TOF in a single frame for range imaging can be extended to photon quantity measurement via multi-frame detection for intensity imaging.

Acknowledgements This work was supported by Natural Key R&D Program of China (Grant No. 2016YFB0400904), National Natural Science Foundation of China (Grant No. 61805036), Natural Science Foundation of Jiangsu Province (Grant No. BK20181139), and Fundamental Research for Funds for Central Universities. We also appreciate the supporting in system testing and applications from the 44th Research Institute of China Electronics Technology Group Corporation.

References

6 Liu S, Zheng Y. A low-power and highly linear 14-bit parallel sampling TDC with power gating and DEM in 65-nm CMOS. IEEE Trans Very Large Scale Integr Syst, 2016, 24: 1083–1091