Design of a Configurable Spike-Encoding Circuit Based on Focal Plane Array

Di Lu 1,2, Wenchang Li 1,2,*, Jian Liu 3,4, Gang Chen 5 and Zhigang Li 5

1 Laboratory of Solid-State Optoelectronic Information Technology, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China; ludi@semi.ac.cn
2 School of Integrated Circuits, University of Chinese Academy of Sciences, Beijing 100049, China
3 State Key Laboratory of Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China; liujian@semi.ac.cn
4 Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
5 Key Laboratory of Semiconductor Neural Network Intelligent Perception and Computing Technology, Chinese Academy of Sciences, Beijing 100803, China; chengang08@semi.ac.cn (G.C.); lizhg@semi.ac.cn (Z.L.)

* Correspondence: liwc@semi.ac.cn

Featured Application: A configurable readout circuit for the focal plane array is designed that transmits information in the form of spike-encoding and supports different encoding methods. Spiking neural networks with different structures can be matched according to the task.

Abstract: Spiking neural networks inspired by biological models are gaining popularity in artificial intelligence due to their ability to solve diverse problems while reducing energy consumption. As a result of the trade-off between the need to transmit large amounts of data and the power consumption of hardware deployment, artificial vision systems are particularly well-suited to construction using spiking neural networks (SNNs). How to communicate with the neuromorphic network effectively is one of the challenges associated with building systems that utilize SNN systems. It is necessary to convert the data to spike form before they can be processed by an SNN as input, unless neuromorphic or event-triggered sensing systems are employed. We present a configurable circuit based on a focal plane array (FPA) capable of providing spike-encoded readout data at the pixel level. With this type of circuit, the current signal of the photoelectric sensor can be encoded into two spike encodings with different precision, which are sent for processing to SNNs. This provides image information at two different scales for the artificial vision system based on SNNs. With this feature, we can use this circuit and different SNN structures to build an artificial target recognition system that is closer to the biological visual system.

Keywords: spiking neural networks; spike encoding; focal plane array; artificial vision system

1. Introduction

For decades, researchers have been investigating how the brain interacts with the eyeball, as this is one of the most sophisticated and powerful visual processing systems in existence. The semiconductor industry has made significant progress by utilizing bionics technology as a foundation for semiconductor imaging devices and image processing. Silicon-based focal plane array (FPA) imaging chips extend infrared and ultraviolet bands beyond visible light. By mimicking the brain's neurobiological system, a crucial structure in artificial intelligence was created: the spiking neural network (SNN).

Information in SNNs is encoded and transmitted through short pulses. These networks are composed of neurons whose action potential generation is described by mathematically non-differentiable equations that approximate the observed behavior of biological systems. SNNs may consume less energy than earlier versions of neural networks running on
dedicated hardware because pulses are only exchanged when input is processed [1]. There are still many issues with encoding methods, network structures, training methods, and hardware implementations that must be resolved before these principles can be utilized. This study focuses on the spike-encoding method and hardware implementation of an FPA.

Rate encoding and temporal encoding are the two types of spike-encoding methodologies utilized by artificial SNNs [2]. By determining whether the precise timing and order of pulses are essential to the presentation of the information, it is possible to divide all specialized coding schemes into these two categories. Information is encoded into the instantaneous or average spike generation rate of an individual or group of neurons, resulting in a value describing neuronal activity identical to the activation value of an artificial neuron. In temporal encoding methods, spike timing is used to encode information such as absolute timing relative to a global reference, relative timing among spikes from different neurons, or the sequence of spikes from different neurons.

Early studies in biology demonstrated that rate coding is a primary method for transmitting information within the nervous system [3]. However, subsequent research has shown that all sensory organs encode their perceptions into the timing of action potentials. Thorp spent years conducting experiments to validate the concept of precise pulse timing as an encoding scheme [4]. It is estimated that the human visual system requires 150 ms to process an object recognition task, which supports the theory that spike encodings in biological nervous systems have temporal properties, as rate encodings are insufficient to explain rapid biological responses [5]. These findings were supported by additional investigations of the visual system [6–9], auditory system [10], tactile system [11], and olfactory system [12,13]. In addition, experiments revealed a trade-off between rapid response time after stimulus onset and precision. In biomimicry studies of the human tactile system, for instance, the TTFS code has been shown to reliably transmit information about the direction of fingertip force and the shape of the surface in contact with the fingertip faster than possible rate codes and quickly enough to explain how tactile signals are used in natural manipulations.

On the other hand, although it has been demonstrated that temporal codes have greater information capacity, faster reaction times, and quicker transmission speeds, their architectures are more complex, and they lack practical training methods. The rate-based encoding scheme does not rely on the precise timing of each individual peak event but rather takes an average over the duration of the program. Their robustness against fluctuations and noise, as well as their simplicity due to their equivalence to current ANN activation values, persuade us.

A single spike-encoding method is commonly used for image processing systems based on SNNs. While this may be advantageous in terms of hardware implementation, it restricts the neural network’s ability to receive image data on a single scale. Due to their self-evolutionary properties, neural networks trained on prior information and backpropagation are able to comprehend data at multiple scales. As can be seen, a singular encoding method limits the adaptive scaling advantage of neural networks.

Research has indicated that bioelectrical signals are not transmitted in a fixed manner in the eye–brain biological visual system [6–9]. Attention in biological vision is constituted by the “scan mode,” a large-scale, low-loading mode of information, and the “gaze mode,” a small-scale, high-loading mode of information. As a means of imitating this mechanism, we conceived an FPA artificial vision system based on spike coding. This system can determine the position of the target through an accelerated “scan mode,” and then identify the type of target through a high-definition “gaze mode.” This paper presents two pulse-encoding methods used on this FPA system and their implementation.

2. System Requirement and Coding Theory

In this section, we analyze the limitations of existing FPA vision systems and propose a new FPA vision system based on spike coding. Next, we analyzed the main requirements
and technical difficulties of the design and, combined with theories, put forward the realization method.

2.1. Standard Digital FPA System

The digital FPA system is an important direction for its development. According to the different ways in which the analog-to-digital converter is integrated into the readout circuit, the digital readout scheme is divided into three architectures: chip-level, row/column-level, and pixel-level ADC. Among them, the pixel-level ADC digital readout circuit is the most advanced architecture at present.

Figure 1 depicts a block diagram of a standard digital FPA system. The sensor component is analogous to a photodiode (PD), which converts light intensity into an output current signal. The readout circuit utilizes a pixel-level ADC design, which converts the analog current signal to a digital signal before outputting it via the peripheral circuit. The digital signal is transmitted to the image processing module, processed by the digital morphological neural network, and then the recognition and classification tasks for the target are completed.

![Figure 1. Block diagram of a standard digital FPA system.](image)

Digitizing the signal within the FPA pixels is an efficient method of processing it; however, such systems still have limitations for artificial vision tasks.

First, based on the reusability of the pixel circuit, the digital encoding method of every pixel in the complete FPA is standardized. This implies that the image accuracy acquired in any area of the FPA cannot be configured and is instead determined by the number of conversion bits of the ADC at the outset of the system design. This goes against the pattern of “scan-gaze” in biological systems. Since the FPA has already digitalized the signal, it can only process the image encoding using a digital morphological neural network, such as ANNs or CNNs. The actual brain and nervous system utilize more efficient bioelectric impulse signals to transmit and process information, which are mechanically distinct from digital morphological neural networks. The final method for digitizing the signal within the FPA is to use multiple series-connected D flip-flops and SRAM to form an N-bit counter, which occupies up to 80% of the pixel area in existing product designs. This is an impediment to the further miniaturization of FPA pixels.
2.2. Spike FPA System

To address the aforementioned issues with the digital FPA system, we conceived an FPA artificial vision system based on spike encoding, which more precisely simulates the biological “scan-gaze” visual system. Figure 2 depicts a block diagram of this FPA system.

![Block diagram of a spike-encoding FPA system](image)

**Figure 2.** Block diagram of a spike-encoding FPA system.

The system’s working mode is divided into an accelerated mode (correspondingly, scan mode) and an accurate mode (correspondingly, gaze mode), which are switched via an externally input control signal.

2.2.1. Accelerated Mode

In accelerated mode, the luminance of each pixel is encoded using a coarse single spike, and a shallow SNN is used to construct the image processing block, imitating the saccade principle in biology to locate the area where the target is situated. This design serves to increase the transmission rate and processing efficiency of image data; however, at the expense of image resolution. In the neurons of the subsequent neural network, such pulses are received using leaky integrated–firing neurons (LIF). Each LIF neuron corresponds to a randomly divided area of the array. The output layer uses a classification voting mechanism to determine the area whose brightness is higher than the design threshold. The output area’s coordinates and the mode-switching signal are fed back to the array reader, as depicted in Figure 3. In their research, Yang Li and Xiang He have developed an SNN model derived from ANN that can accomplish nearly the same performance as ANN in target detection and segmentation tasks [14].

2.2.2. Accurate Mode

The FPA will enter accurate working mode once it receives the area coordinates and mode switching signal. In accurate mode, pixels that are not confined by region coordinates are disabled. The pixel-coding circuit in the framed portion will switch to a multi-pulse precise coding mode, and the 4-layer SNN that follows will conduct target recognition and classification, as depicted in Figure 4. This precise encoding necessitates longer encoding periods, increases pixel power consumption, and slows down image processing. However, given that only a portion of FPA pixels are active and the precision of target recognition, this trade-off is worthwhile. Precision mode replicates the gaze and object resolution mechanisms of biological visual systems. In their paper, Tengxiao Wang
and Cong Shi proposed a four-layer structure SNN that can be used for target recognition and classification [15] and discussed its hardware implementation method [16].

2.3. Technical Difficulties

In the early stages of research, the photoelectric sensor in the FPA system had a developed solution. In addition, the SNN used for target recognition and classification, including its model architecture, deployment method, and training method, had a reference design scheme [14–16]. Therefore, the technical difficulty of the system resides in the FPA FPA-integrated encoding circuit. Depending on the requirements of the system, the encoding circuit must support two distinct spike-encoding modes: a coarse single-spike encoding and a precise spike-sequence encoding. This requires judicious selection from existing coding theories. While designing, consideration must also be given to the circuit implementation’s dependability, reconfigurability, and reusability.

2.4. Spike-Encoding Methods

There are two commonly used spike-encoding methods: rate encoding and temporal encoding. In image processing, these methods involve SNNs and generate spike sequences that contain varying numbers, intervals, and levels of information. They involve SNNs that can be applied to a wide variety of problems.

2.4.1. Rate Encoding

Rate encoding can be implemented in three ways: count frequency encoding, density frequency encoding, and population frequency encoding [2,17].
• Count Frequency Encoding

Count rate (average over time) is the most common rate-encoding scheme. It is defined by the average emissivity

\[ v = \frac{N_{\text{spike}}}{T}, \]  

(1)

where \( N_{\text{spike}} \) is the spike count and \( T \) is the time window. Adrian and Zotterman found that the bending of frog muscles with varying weights alters the bioelectric signal’s firing rate [3]. Emissivity can be used to characterize any slowly shifting analog value in artificial contexts, such as pixel intensity or gas concentration.

• Density Frequency Encoding

Neural activity is measured through different simulations, and the results of neural responses are displayed in stimulus time histograms to visualize spiking activity. The spike density is defined as

\[ p(t) = \frac{1}{\Delta t} \frac{N_{\text{spike}}(t; t + \Delta t)}{K}, \]  

(2)

where \( N_{\text{spike}} \) is the average number of spikes throughout all iterations, expressed as a fraction of \( t \) and \( K \) iterations. This method of coding is not reasonable for biological behavior. In nature, predators do not hunt by calculating multiple averages of the perfectly consistent motion trajectories of the same prey. In an artificial SNN, however, it may be useful to use the same input to calculate the average of multiple simulated runs due to the uncertainty of the information transmission path caused by the network structure.

• Population Frequency Encoding

Based on similar properties of neurons in the population, its rate is defined by

\[ p(t) = \frac{1}{\Delta t} \frac{N_{\text{spike}}(t; t + \Delta t)}{N}, \]  

(3)

The number of spikes \( N_{\text{spike}} \) in the total population is determined in the time interval and divided by the duration \( t \) and the total number of neurons \( N \). This method of encoding is characterized by the fact that, for a given input value, the peak of the neuron can be uneven. If each neuron has a different activation curve for any input current value, then the overlapping spikes generated by a set of such neurons can be used to encode accurate information such as numbers, vectors, or matrices [18].

2.4.2. Temporal Encoding

Commonly used temporal coding can be divided into several subcategories: a global reference scheme, an ISI encoding scheme, and an associated synchronization scheme.

• Global Reference Scheme

The most basic temporal encoding scheme is **Time-to-First-Spike (TTFS)**, where information is encoded as the time delay between stimulus onset and the neuron’s first spike. In the most basic model, this time delay can be expressed as the reciprocal \( \Delta t = 1/a \) of the normalized signal amplitude \( a \), or as a linear relationship \( \Delta t = 1 - a \). In biology, studies by Johansson and Birznieks elucidate the phenomenon of TTFS encoding in the human tactile system. In a discrete fingertip event, the time difference between the onset of contact and the generation of the first spike contains information about the direction and magnitude of the force [11].

**Phase**-encoding theory encodes information as the time difference between a spike and a reference oscillation [19,20]. Gray, König, Engel, and Singer identified the existence of this coding theory in their studies of neurons in the cat visual cortex and identified the relationship between spike production and reference oscillations [21].

**Ranked-order coding (ROC)** theory encodes information as the order of firing among neuronal clusters relative to a global reference [22,23]. Compared with TTFS, the ROC
scheme avoids considering the precise time difference between the spike and the reference, and its working mode is closer to a discrete normalization filter. However, the absolute amplitude information of the stimulus is lost in ROC coding; therefore, the original signal with consistent amplitude cannot be reconstructed.

- ISI Encoding Scheme

In **inter-spike interval (ISI)** encoding (some papers call it “Latency Encoding”), information is defined by the time difference between spikes produced by different groups of neurons; this encoding scheme is observed in the neural activity of pyramidal cells [24]. Li and Tsien show that in this encoding theory, longer periods of silence can contain more information than periods of higher spike activity [25].

- Associated Synchronization Scheme

**Correlation and synchrony encoding** uses the spikes produced by groups of neurons to encode information, and the stimulation of a group of neurons will be converted into a set of spatiotemporal spike matrices for expression. Within this matrix, clusters of spikes with relatively short **ISIs** represent specific input patterns [17], information encoded as which neurons fire simultaneously. Sparse Distributed Representations (SDRs) [26,27], which are often used in compressed sensing theory, also belong to this encoding theory. Biological studies have identified this encoding scheme in the activity of cells in the somatosensory cortex of monkeys [28] and the visual cortex of cats [21,29].

2.4.3. Encoding Method Selection

In accelerated mode, every pixel of the sensor area is output. Therefore, the design seeks to reduce the number of pulses as much as possible and minimize the burden of information transmission. In accurate mode, only a small fraction of the sensor’s pixels are activated, allowing more pulses to deliver more precise information. It can be seen from the aforementioned encoding theory that the TTFS scheme in temporal encoding only generates one pulse, which is in line with the design pursuit in accelerated mode. A pulse interval is the reciprocal of the frequency for a fixed period of time (or counting period); therefore, the count frequency code can be obtained by repeating TTFS steps to generate a pulse sequence. To ensure circuit design reusability, this paper chose TTFS coding as the accelerated coding method and counting frequency coding as the accurate coding method.

3. Circuit Implementation and Simulation Results

We use a charge packet technique akin to digital coding to generate spikes in the spike-coded FPA system, and its working principle will be described in this section. In addition, we have designed this encoding circuit to be configurable, relying on the control of the feedback signal to select the encoding mode for the pixel luminance information.

3.1. Charge Packet Circuit

A charge packet circuit is hardware-friendly for pixel-level circuit implementation. The integrating capacitor is discharged by conducting the current through the photodiode. When the capacitor node potential falls below the threshold voltage, a comparator switches, producing a spike and activating a circuit reset to recalibrate the capacitor. Figure 5 depicts the circuit’s construction.

The circuit repeats this procedure during each counting cycle. In a comparator with a fixed threshold voltage $V_{ref}$, the pulse intervals vary with the magnitude of the PD input current, resulting in a frequency pulse code that is linearly proportional to the current.

3.2. Encoding Circuit A

Scheme A is based on the standard charge packet counting circuit’s front end, ignoring its asynchronous counterpart. Simultaneously, the frequency-encoding/temporal encoding switching control signal MODE is introduced: when $\text{MODE} = 1'b0$, the circuit operates in
frequency-encoding mode; when \( \text{MODE} = 1\text{’}b1 \), the circuit operates in temporal encoding mode, and the generated pulse width has the TTFS form, as depicted in Figure 6.

![Figure 5. Circuit diagram of charge packet pulse.](image1)

![Figure 6. Circuit diagram of encoding circuit A.](image2)

Scheme A is based on a frequency-encoding circuit and adds temporal encoding control via the NAND gate. The original frequency-encoding pulse retains only one pulse edge, and the following pulses are insulated (at the same time, the charge packet count is suspended for energy efficiency purposes).

In frequency encoding’s functioning mode, the integral capacitor must be periodically reset. Consequently, when designing the actual circuit, the latency of the comparator, the size of the reset switch, and the value of the integral capacitor are the most crucial factors that must be carefully considered. As depicted in Figure 7, a half-side Schmitt trigger has been implemented to reduce comparator latency. The input signal range and counting period must be utilized to optimize the size of the reset switch and the value of the integral capacitor. Figure 8 depicts the simulation results of the circuit in frequency-encoding mode and time-encoding mode, respectively.
The capacitance value is too small to encode temporal information. As a result, the TTFS determined by device size and cannot be adjusted externally. Meanwhile, the entire scheme (Figure 8) is added to the frequency-encoding circuit, especially the integrating capacitor.

As depicted in Figure 7, the circuit uses a Schmitt trigger (half side), a push–pull stage circuit structure, and local positive feedback, which accelerates the output inversion, to reduce the comparator delay. However, the threshold voltage for the Schmitt trigger is determined by device size and cannot be adjusted externally. Meanwhile, the entire scheme A is modified by the frequency-encoding circuit, especially the integral capacitance value. The capacitance value is too small to encode temporal information. As a result, the TTFS pulses of even the weakest input current signal account for only a small fraction of the total counting period.

Scheme B introduces an innovative comparator to solve the above two problems: a self-biased rail-to-rail comparator. In addition, for the temporal encoding mode, an additional integrating capacitor is added to the circuit.

3.3. Encoding Circuit B

This scheme is similar to the multi-color charge packet counting scheme in that it utilizes independent integration capacitors for frequency and temporal encoding modes, which optimize the capacitor capacitance accordingly. A newly developed comparator allows the reference signal to be adjusted freely [30], especially the ramp reference signal. Figure 9 depicts circuit construction.
The circuit in scheme B was primarily adjusted as follows:

- **Extra integral capacitor**

  Different integration capacitors are designed for frequency-encoding mode and temporal encoding mode, respectively (100 fF for frequency-encoding mode, 900 fF for temporal encoding mode), as shown in Figure 10. The integral capacitor working in frequency-encoding mode is small and can be quickly filled to generate pulses. The integration capacitor working in temporal encoding mode is relatively large. Under the minimum PD current, the capacitance integration time covers the entire technical cycle as much as possible. Only one pulse is generated to reduce circuit energy consumption.

- **Self-biased comparator**

  A specially designed rail-to-rail comparator with self-biased voltage is used. The comparator extends the comparison range between the power supply and ground to reduce delay. It is more sensitive to small PD currents, as depicted in Figure 11.
Appl. Sci. 2023, 13, x FOR PEER REVIEW...tistical effect on the activation of SNN neurons, thereby enhancing the recognition accuracy of the overall SNN network.

with a fixed DC level as the reference voltage. In temporal encoding mode, however, the wave starts at the beginning of the counting cycle, ends at the end of the technological cycle, and ends at the level of the power supply voltage. It is also possible to generate pulses with a fixed DC level as the reference voltage. In temporal encoding mode, however, the intersection position of the reference voltage and current is nonlinear in time due to different PD inputs. Consequently, we used the ramp wave as the comparator reference voltage in temporal encoding mode to reduce spike nonlinearity and improve interpretability.

The simulation of Scheme B is depicted in Figure 12, where YP is the temporal encoding mode output, $V_{\text{int1}}$ is the voltage waveform of the temporal encoding mode’s integrating capacitor, and $V_{\text{ref}}$ is the ramp reference signal. When the input current is low, the TTFS output pulse width is extremely broad and may extend the entire counting period.

Ramp wave reference voltage

Scheme B uses the ramp wave as a reference voltage for the comparator. The ramp wave starts at the beginning of the counting cycle, ends at the end of the technological cycle, and ends at the level of the power supply voltage. It is also possible to generate pulses with a fixed DC level as the reference voltage. In temporal encoding mode, however, the intersection position of the reference voltage and current is nonlinear in time due to different PD inputs. Consequently, we used the ramp wave as the comparator reference voltage in temporal encoding mode to reduce spike nonlinearity and improve interpretability.

The simulation of Scheme B is depicted in Figure 12, where YP is the temporal encoding mode output, $V_{\text{int1}}$ is the voltage waveform of the temporal encoding mode’s integrating capacitor, and $V_{\text{ref}}$ is the ramp reference signal. When the input current is low, the TTFS output pulse width is extremely broad and may extend the entire counting period.

3.4. Results and Comparison

We contrasted the operating currents of the two schemes as well as the widths of the generated spikes. The spike width is proportional to the integration time and the PD current when the value of the integration capacitance is held constant. The minor change in spike width in frequency coding mode can also be interpreted as a code with weak time information, which will have a statistical effect on the activation of SNN neurons, thereby enhancing the recognition accuracy of the overall SNN network.

Table 1 depicts the operating current of the complete pixel circuit in each mode.

---

Figure 11. Circuit diagram of the specially designed self-biased rail-to-rail comparator.

Figure 12. (a) Frequency encoding working mode, where YP is the pulse output, and $V_{\text{int0}}$ is the voltage waveform of the integrating capacitor. (b) Temporal encoding working mode, where YP is the TTFS pulse width output, and $V_{\text{int1}}$ is the voltage waveform of the integrating capacitor.
Table 1. Operating current of the pixel circuit. (CMOS 130 nm process; 1.2 V power supply voltage; 1 nA PD input current; 0 °C; typical corner).

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Frequency Encoding</th>
<th>Temporal Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme A</td>
<td>8.4 uA</td>
<td>0.23 uA</td>
</tr>
<tr>
<td>Scheme B</td>
<td>2.9 uA</td>
<td>2.5 uA</td>
</tr>
</tbody>
</table>

Table 2 depicts the widths of the generated spikes in each mode.

Table 2. Widths of the generated spikes in each mode (1.2 V power supply voltage; 25 °C; typical corner).

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Input Current</th>
<th>Spike Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme A</td>
<td>100 pA</td>
<td>327 ps</td>
</tr>
<tr>
<td></td>
<td>200 pA</td>
<td>327 ps</td>
</tr>
<tr>
<td></td>
<td>1 nA</td>
<td>327 ps</td>
</tr>
<tr>
<td></td>
<td>10 nA</td>
<td>328 ps</td>
</tr>
<tr>
<td></td>
<td>100 pA</td>
<td>496 ps</td>
</tr>
<tr>
<td></td>
<td>200 pA</td>
<td>522 ps</td>
</tr>
<tr>
<td>Scheme B</td>
<td>500 pA</td>
<td>553 ps</td>
</tr>
<tr>
<td></td>
<td>1 nA</td>
<td>583 ps</td>
</tr>
<tr>
<td></td>
<td>10 nA</td>
<td>739 ps</td>
</tr>
</tbody>
</table>

4. Discussion

During the actual simulation process, we discovered that the linearity problem between PD current and spike delay cannot be resolved using only a single slope reference signal (scheme B). Figure 13a depicts two signals connected to the comparator in temporal encoding mode, the PD signal and the \( V_{ref} \) signal, and the horizontal axis represents time (normalized by the integration period). The TTFS spike delay generated by this circuit exhibits strong nonlinearity with the input PD, as depicted in Figure 13b.

![TTFS Generation (PD Currents Clusters)](image)

**Figure 13.** (a) The PD signal versus the \( V_{ref} \) signal in the temporal encoding mode. (b) Linearity analysis of the PD signal and the \( V_{ref} \) signal.

In order to further improve the above nonlinear problem, a follow-up plan introduces a segmented-slope ramp wave in \( V_{ref} \), as depicted in Figure 14. In this circuit slice, a slope-enhanced current is introduced according to the time slice of a single integration cycle. A single-slope \( V_{ref} \) signal is substituted for the segmented-slope \( V_{ref} \) signal, and the relationship between the TTFS peak delay and input PD is depicted in Figure 15.
The non-linear factor has been improved from 0.099 to 0.0045, a 22-fold improvement. The part of the circuit structure that generates the segmented-slope $V_{ref}$ signal (Figure 12) is a shared circuit for the entire area array and does not occupy the space of a single-pixel readout circuit. In the circuit design, however, we implemented configurable segmented-slope $V_{ref}$ with memristors. This portion of the circuit will be implemented in future solutions as we are resolving the integration issue between the memristor and conventional CMOS technology.

Further, after the sensor and network part of the project is completed, we will conduct additional evaluations on the task accuracy and energy efficiency of the entire artificial vision system.

5. Conclusions
This work designs a readout circuit based on an FPA, which can encode the current signal of the sensor into a spike sequence and supports two different spike-encoding schemes in the same pixel. In contrast to conventional digital FPA readout circuits, which output a digital encoding, our scheme encodes image data as single spikes or spike sequences. Without the need for digital-to-spike conversion, this encoding can be received and processed directly as the input of spiking neurons in SNNs. Furthermore, compared with the single, non-configurable encoding scheme, our scheme can switch between temporal encoding and frequency encoding. The circuit is able to perform both full-array low-pulse encoding and regional pixel high-precision encoding, making it suitable for the various processing modes of SNNs. It is also able to bring the artificial vision system closer to the working mechanisms of the biological vision system when it is matched with different structural SNNs.
Author Contributions: Conceptualization, W.L.; Methodology, G.C.; Validation, Z.L.; Writing—original draft, D.L.; Writing—review & editing, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences, grant number [XDA18040400, XDB44000000].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to restrictions privacy. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the project containing this study is still in development.

Conflicts of Interest: The authors declare no conflict of interest.

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.