An Optoneuronic Device with Realistic Retinal Expressions for Bioinspired Machine Vision

Dan Berco,* Diing Shenp Ang,* and Hai Zhong Zhang

Machine vision systems rely on communication between cameras and processor modules to capture and analyze visual information. This arrangement renders them as bulky and inefficient in terms of speed and power dissipation for futuristic big data applications that involve artificial intelligence algorithms. An apparatus able to imitate the operation of biologic eyes and function as a standalone platform would therefore present the next evolutionary step in machine visual perception. Neuromorphic computing is an alternative approach to the Von Neumann architecture that carries the potential for implementing such intelligent cameras. In this regard, artificial synaptic devices have been widely used in recent years to construct hardware-based neural networks mainly due to their adjustable electric parameters. Herein, a bioinspired, hybrid electrophotonic responsive neuronic device that mimics the combined functionality of retinal cones and bipolar cells is demonstrated. Under illumination, it features a hyperpolarization-like current response in an OFF state and a complementary depolarized reaction when toggled to an ON state. Furthermore, electrical pulsing done in conjunction with light stimulation can emulate the horizontal cell neurotransmitter release in center-surround biologic configurations. These devices may thus serve as building blocks for advanced visual systems, integrating self-healing sensory and neuromorphic computing into an artificial cognitive retina.

1. Introduction

Visual perception apparatuses (VPAs) are ubiquitous more than ever and play important roles in almost every aspect of modern civilization. They form the foundation for autonomous vehicle control and safety systems[5] and of artificial eyes in industrial fields such as pharmaceuticals, semiconductors, and even agriculture.[6] Another important milestone in this field was marked by the development of bionic eyes for the visually impaired.[7–11] However, these systems contain an inherent drawback that lies in their reliance on separate camera and processor modules to capture visual data and execute intelligent image processing algorithms. Current technologies do not present a viable workaround for this matter, and as a result, VPAs cannot surpass certain operation speed and power dissipation bounds. In addition, this construction renders them as somewhat bulky and inelegant, especially when intended to function as prosthetic eyes. Several recent studies suggested that hardware-based artificial neural networks could outperform pattern recognition software and serve as a foundation for futuristic VPAs.[12–13]

As for the implementations of such networks, solid-state devices that emulate biologic neurons and synaptic clefts were demonstrated as far back as the early 1990s over floating-gate metal-oxide-semiconductor (MOS) structures.[16,17] Alternatively, the study of memristors[18,19] as artificial synaptic devices has been gaining momentum since the discovery of the reversible resistive switching effect.[20] Within this realm, recent developments of hybrid electrophotonic responsive devices[21–24] may prove to be a leading indicator for the evolution of intelligent, artificial neuron-based VPAs.[25,26] Such systems were envisioned as being able to mimic the performance of biologic eyes without the need for external cameras or processors.[27] Bioinspired VPAs may thus present a quantum leap in robotic vision performance. Unfortunately, the majority of photoresponsive devices suffer from a multitude of complications that decimate their prospects to serve as a basis for an intelligent retina while integrated with current technologies. One major issue arises from their reliance on special materials that leads to high operating voltages, large current consumptions, complementary MOS (CMOS) incompatibility, and large (millimeter-sized) structures.[27]

This work presents an artificial neuronic device (AND) based on CMOS-compatible materials and a highly scalable structure, capable of a hybrid electrophotonic response that mimics the functionality of retinal rods and bipolar cells. Its combined reaction to external voltage and light stimuli is expressed as a modulated, extremely low output current. An “off-bipolar”
state, the device displays a hyperpolarization-like response under illumination that decays into a depolarized state in darkness. This state may then be toggled to an “on-bipolar” one, where it produces a depolarized output under illumination and hyperpolarizes in darkness. This functionality can be used to simulate self-healing sensory mechanisms found in living organisms’ retinas.\textsuperscript{28} Moreover, electrical pulsing applied with light stimulation can emulate horizontal cell neurotransmitter release in “center-surround” biologic configurations and regulate the output frequency of the device. A plurality of such devices may thus be used to form an array and operated in an “on-center off-surround” bitwise configuration for increased image acquisition accuracy. Considering that the mammalian retina is made of layered light-sensitive rods and cones that convert photonic stimulation into neural activity,\textsuperscript{29,30} visual data are conveyed by synaptic junctions of bipolar and horizontal cells via neurotransmitter signaling.\textsuperscript{31} The long-term potentiation and depression\textsuperscript{32–34} may be tuned by either electric stressing or light irradiation. In addition, different bias pulsing schemes could be used to emulate horizontal cell neurotransmitter release.\textsuperscript{35,36}

2. Results and Discussion

2.1. Implementation Considerations

The ANDs in this work were fabricated using a memristive crossbar structure, which is considered a promising candidate for achieving very-high-density networks similar to those of synaptic connections in the cerebral cortex.\textsuperscript{37} The structure features a thin-film large bandgap dielectric (hafnia) placed between intersecting top and bottom conducting electrodes (TE and BE, respectively). An important consideration for material choice was the CMOS-compatibility requirement. In this manner, the major drawbacks associated with light-responsive materials in synaptic devices, \textsuperscript{21–24} which simply cannot be integrated into the conventional CMOS process, were avoided. The TE and BE materials were chosen as indium tin oxide (ITO) and TiN. ITO is a well-known wide-bandgap semiconductor with good conduction and transparency properties. In this manner, ITO can facilitate photon–electron interaction at the TE–dielectric interface.

2.2. Hyperpolarized and Depolarized Responses

The basic approach for emulating the reaction of retinal cones and bipolar cells is through a photoinduced current response in the high and low resistive states (HRSs and LRSs, respectively) of the AND. The HRS was operated as an on-bipolar cell, whereas an LRS as an off-bipolar. Figure 1A,B shows a scanning electron microscope image along with a 3D depiction of the device and probing and illumination setup. A more detailed description of the measurement setup is given in the Supporting Information. Resistive states in the device are interchangeable as shown by the $I$–$V$ sweeps in Figure 1C. The figure highlights the stress-induced formation and annihilation of traps indicated by step-like transitions.\textsuperscript{38,39} A positive sweep (gray) over a slightly conductive LRS induced more trap formation and increased the conductivity even further, as verified by the successive black curve. The $I$–$V$ curves show a relatively symmetric behavior after the transition to an LRS as it effectively results in the formation of a conductive filament. A negative sweep (blue) resulted in a transition back to HRS due to trap annihilation and an evident current drop. The $I$–$V$ sweeps in Figure 1D verified that the conductivity alternation between an LRS (black) and an HRS (blue) was changed by more than four orders of magnitude.

Retinal cones release glutamate neurotransmitters when depolarized in darkness to interact with ON and OFF bipolar cells. $\alpha$-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors are excitatory to off-bipolar cells and cause them to depolarize or become excited. In contrast, mGluR receptors are inhibitory to on-bipolar cells and cause them to hyperpolarize (inhibited output). Under illumination, the opposite process takes place, as cones are now hyperpolarized, and their glutamate release is inhibited. The lack of neurotransmitters leads to the inhibition of off-cells and excitation of on-cells. In other words, off-bipolar cells follow the response of cones, whereas on-bipolar cells reverse it. This behavior is emulated by the AND where the output current takes the role of neurotransmitter release in the exited state.

Figure 1E shows the time-dependent current as the LRS device is exposed to 0.5 s illumination impulses with different power densities under a constant bias of 100 mV. The current drops in response to irradiation by up to four orders of magnitude and self-recovers once in darkness within a few seconds (self-healing sensory). Such a response corresponds to the hyperpolarization of off-bipolar cells under lighting and depolarization in darkness. The pharmacology of transmission at the synapse between a cone and an off-bipolar cell was studied by DeVries.\textsuperscript{28} Glutamate receptors on off-bipolar cells had over 1 nA drop after the application of 15 ms pulses of glutamate (2 mM). Desensitization was followed by gradual recoveries that ranged from a few seconds in a b7 cell to a minute for a b3 cell in a similar manner to the results shown in the figure.

As for Figure 1F, the time-dependent current behavior corresponds to depolarization of on-bipolar cells under lighting and hyperpolarization in darkness (note: the sampling rate is about 1/0.25 s). In this case, a saturated response was observed at a power density of 5×. The response patterns shown in Figure 1F are similar in nature to the depolarization response to a step of light in on-$\beta$ ganglion cells measured by Freed.\textsuperscript{40} In addition to the step-like action potential, ganglion cells produced a decaying response along with a high-frequency spiking while illuminated. This was attributed to a majority of type b1 bipolar synaptic connections that produced large quantal excitatory postsynaptic potential (EPSP) in the ganglion cell during initial transient depolarization. Type b2 and b3 cells caused small quantal EPSPs during subsequent depolarization and contributed to low-frequency responses. In the AND, hot-electron excitation is believed to produce the initial strong response to illumination, whereas slower detrapping results in a decaying depolarized current output. Device operation in the relatively higher-frequency regime will be addressed in the following section.

What possible mechanisms could account for the observed photoinduced response? The experimental results suggest that the main cause is a combination of surface plasmon-induced electron–photon excitations and trap-assisted hopping conduction.\textsuperscript{38,39} The schematic band diagrams shown in
Figure 1. Emulating the reaction of retinal cones and bipolar cells to illumination by a photoinduced current. A) Electron microscope image of the 200 × 300 nm² crossbar device (scale bar 1 μm). B) 3D visualization of the ITO/HfO₂/TiN device structure and measurement setup. C) I–V curves showing the stress-induced formation and annihilation of trap states in the hafnia as indicated by step-like transitions. D) I–V sweeps indicating over four orders of magnitude conductivity difference between an LRS (black) and an HRS (blue). E) Time-dependent current in an LRS with application of 0.5 s white light pulses (20 s in case “3”) and different irradiation power densities (the red numbers indicate multiples of 100 mW cm⁻² or a flux of 1.8 × 10⁶ photons per second on the device). F) After a transition to an HRS, the time-dependent current increases by over four orders of magnitude in response to the illumination pulses and decays back within a few seconds as well, once the light turned off. G) Energy band diagram of the HRS “on-bipolar” cell during a depolarization-like response. H) Conduction in an LRS without illumination where oxygen vacancies align to form a conductive filament. I) Energy band diagram illustrating the suggested mechanism of the LRS “off-bipolar” cell hyperpolarization-like response.
Figure 1G-I illustrate these mechanisms at work. As part of the setup, a gold-coated probe was incorporated to invoke surface plasmons (SPs) at the tip contact point where the ITO TE once illumination was applied. Plasmons have been demonstrated to effectively generate hot electrons at the interface of bandgap materials.\cite{41} In contrast, a tungsten probe was used with the BE to avoid symmetric excitations of the device and therefore maximize the electron–photon current.

Figure 1G shows the band structure and resulting current for an on-bipolar state (HRS-ON) under illumination. Hot electrons are transferred to the ITO conduction band at the Au probe–ITO contact point by plasmon-induced excitation. These energetic electrons follow the native band structure (work functions: ITO \( \approx 4.7 \) eV, \( \text{Au} \approx 5.1 \) eV, and TiN \( \approx 4.5 \) eV; hafnia electron affinity \( \approx 2 \) eV) and applied bias to drift diffuse toward the ITO–hafnia interface. There they get excited by 450 nm photons (the ITO film is highly transparent at this wavelength) and gain enough energy that allows them to overcome the ITO/hafnia 2.7 eV energy barrier.\cite{42} Excited electrons can traverse across the insulator to the either grounded or biased BE. The overall outcome is a positive photocurrent that decays relatively fast once the irradiation stops. At its HRS, the device is in fact a passive capacitor. As a result, the total amount of charge exited under illumination and the amount discharged in darkness are virtually the same and determined by time integration over the current.

The creation of abundant traps and defects in the dielectric is a key characteristic of the conductive LRS. In particular, oxygen vacancy traps have various charge states (marked as \( V_0^-, V_0^+, \)) and \( V_0^2 \) in Figure 1G) distributed between 0.7 and 2.9 eV below the hafnia conduction band.\cite{43,44} Additional traps are created due to the application of electric stress over the oxides in the HRS. This stress facilitates the formation of oxygen vacancies that align to form a conductive filament and act as stepping stones for electrons to traverse across the insulator. The device is thus being set to an LRS. When combined with rising temperatures within the dielectric, this stress drives both the creation and annihilation of traps. Reversing the bias polarity repels the negatively charged oxygen ions back toward the positively charged traps, where they may recombine, and the device is reset back to an HRS.\cite{38,39} This state is shown by the band diagram of Figure 1H and represents the off-bipolar state (LRS-OFF). The same state under illumination is shown in Figure 11. As the material is illuminated, electrons are excited from their trapped position causing a sudden drop in this current. They are then captured by other defects distributed at different energy levels throughout the dielectric. The charge build-up within the dielectric produces an opposing electric field that works to reduce the overall current in conjunction with the disruption. Once the light is turned off, the excitation is removed, trapped charged may be released, and the current recovers back to its original state. It is worth mentioning that previous works observed the light-induced, meta-stable Frenkel pair recombination that may take place and result in a HRS.\cite{35} Trapped charge lifetime of in high-k dielectrics was studied extensively in the context of CMOS bias temperature instability. It was shown that this lifetime has a huge span that varies from a few milliseconds to days on end.\cite{46} In this sense, the current behaviors observed are in agreement with this phenomenon as they reveal an initial sharp drop and a slower recovery associated with detrapping.

### 2.3. Stimulated Spiking Activity in HRS-ON

Most publications implementing time-dependent stimulations over synaptic devices rely on electrical modulation that in turn dictates the use of a FET construction. Such four-terminal devices are essential as each input signal requires a designated terminal. Needless to say, these anatomies decimate the integrability potential when large-scale, high-density networks are concerned. The presented approach takes advantage of double excitation processes to operate a two-terminal AND in the frequency domain, by inputting two modulating signals and producing a single correlated output. It was achieved using a charge pumping technique, as shown in Figure 2A. The light source (marked \( I \)) was pulsed, whereas the temporal current response \( I(t) \) recorded under bias \( B \).

Initially a low-frequency analysis was performed using 1 s illumination pulses having a cycle of 10 s with different power densities (10–50 mW cm\(^{-2}\)) and a bias of 100 mV without the application of an electric pulse (Figure 2B). The corresponding spectral decomposition is shown in Figure 2C (a.u. is used to denote arbitrary units). The figure shows the main spectral line associated with the excitation process located at \( \approx 0.1 \) Hz. The other peaks are believed to result from charge relaxation into deep trap states followed by discharging events to the BE. This modulation is clearly visible in the spatiotemporal current patterns of Figure 2B. As expected, the peaks show a direct correlation to the light intensity due to the generation of more hot electrons. The measurements were followed by lowering the light pulse duration to 0.5 s, effectively reducing the trap-charging phase to half, over the same time period allocated for relaxation. Figure 2D shows the modulation using these 0.5 s light pulses with a cycle of 10 s and a bias of 100 mV. The analysis in Figure 2E reveals a weaker main spectral line as detrapping processes were emphasized over excitation and trapping (0.5 s illumination in Figure 2D compared with 1 s in Figure 2B). The spatiotemporal pattern associated with relaxation and detrapping is more evident for all cases of \( I(t) \) in Figure 2E.

We believe that these results help to establish the theoretical explanation discussed earlier. Longer illumination pulses resulted in the generation of more hot electrons which led to stronger excitation currents. In this case, discharge currents were relatively weaker, as evident in the temporal current pattern. Once shorter pulses were used, the pattern changed due to weaker excitation and the discharge became more apparent. Next, the high-frequency response of the on-bipolar HRS device was characterized. Shifting the onset of electric modulation with respect to the photonic one determined the spike timing-dependent potentiation (STDP) and depression of the main spectral lines. A higher power output at the main spectral lines was termed “potentiation,” whereas a lower power “depression.” The results shown in Figure 3 reveal that electric versus light pulse timing can be used to channel spiking power to either higher- or lower-frequency modes. Such high low components may be separately filtered and used to simulate ganglion cells’ spiking activity. In this manner, differentiation between on-center and off-surround responses in a center-surround configuration could be done, as will be discussed later on.
Figure 3A shows a schematic depiction of the measurement setup along with the relative timing between electric and light stimulation. Voltage and light sources (marked $V$ and $L$) were operated, whereas the temporal current response $I(t)$ were recorded under bias $B$. The relative timing ($dt$) between electric and light pulse stimulations was used to modulate the said double excitation of hot electrons and produce a time-variant current. Voltage pulsing incorporated electric stressing to the band structure which in turn affected the concentration of charge at the ITO/hafnia interface. When combined with dielectric trap concentration variations, as defined by the state of the device (being either ON or OFF), the overall outcome was the modulation of the detrapping current.

This measurement was carried out using impulse stimulation (1 Hz, 50% duty cycle for both voltage and light), a sampling rate of 1/0.1 s, power density of 100 mW cm$^{-2}$, and 100 mV bias.

The voltage pulse $V$ was time shifted ($dt$) in relation to the illumination pulse and the resulting spectral output analyzed. Figure 3B,D shows the current as a function of time with excitation amplitudes for $-1$ and $+1$ V, respectively, whereas Figure 3C,E shows the corresponding spectral power decomposition of the main harmonics. As mentioned earlier, the approach used in the analysis aims to quantify the manner in which the devices respond in the frequency domain to temporal differences between the electric and photonic excitations. For this purpose, the power output for each excitation case was obtained by integration over the main spectral components and divided by the power output of the basic, nonelectrical stimulation output (black drop lines).

The analysis over normalized power summarized in Figure 3F shows that the output power was shifted by ±20% (error bars account for ±1%) as the signal’s energy was channeled to both
Figure 3. High-frequency and STDP response of an on-bipolar HRS device. A) Schematic depiction of the device’s cross section and measurement setup with the relative stimulation timing. B) Spatiotemporal current with an excitation amplitude ($V$) of $-1 \text{ V}$ (black: no electric pulse NP, red: $dt = 0$, blue: $dt = -0.25 \text{ s}$, and magenta: $dt = +0.25 \text{ s}$). C) Spectral power decomposition for the main harmonics (using matching colors). D) $I(t)$ with excitation amplitude of $+1 \text{ V}$. E) Associated spectral power decomposition. F) Normalized power outputs: $\tilde{P}_{-1V}$, $\tilde{P}_{+1V}$, and $\tilde{P}_0$ for each of the excitation cases (red: $-1 \text{ V}$, blue: $+1 \text{ V}$, black: no pulse). G) Main spectral power lines without the application of voltage pulses (NP), highlighting the main components used in the analysis (black drop lines) along with the high and low frequency bands (blue and red squares respectively). H) Channeling power to the high frequency band as a function of pulse polarity (total power by summation: $P_{HF}^{+1V}$ and $P_{HF}^{-1V}$) and offset time ($dt$). I) Channeling power to the low frequency band as a function of pulse polarity (total power by summation: $P_{LF}^{+1V}$ and $P_{LF}^{-1V}$) and offset time ($dt$).
higher- and lower-order harmonics. The normalized power was calculated according to

$$w_0 = \sum_{0.9 \leq f < 1.08} \delta f, NP \cdot P_{0, d_{ON}}(-0.25,0,0.25) = \frac{w_0}{w_0}$$

(1)

$$P_{-1V, d_{ON}}(-0.25,0,0.25) = \frac{1}{w_0} \sum_{0.9 \leq f < 1.08} \delta f, -1V$$

(2)

$$P_{+1V, d_{ON}}(-0.25,0,0.25) = \frac{1}{w_0} \sum_{0.9 \leq f < 1.08} \delta f, +1V$$

where $\delta f, V$ represents the spectral power for a given frequency $f$ and pulse bias $V$. The main spectral power lines that were used in the analysis are shown in Figure 3G with the low and high frequency bands (with respect to the main spectral components) being highlighted as well. These bands were defined as 0.8 Hz $\leq f_{LP} \leq$ 0.9 Hz and 1.08 Hz $\leq f_{HF} \leq$ 1.18 Hz. This definition was used to illustrate how energy is being shifted through the spectrum as the relative timing between stimuli was modified. Figure 3H,I shows channeling power to the high and low frequency bands as a function of pulse magnitude and as a function of pulse voltage polarity by summation over the spectral power in the respective bands ($P_{-1V, d_{ON}}$, $P_{+1V, d_{ON}}$, $P_{+1V}$, and $P_{+1V}$) and offset $dt$. They are obtained by summation over the spectral power in the bands as defined in the following equations

$$P_{LF, d_{ON}}(-0.25,0,0.25) = \sum_{0.8 \leq f < 0.9} \delta f, V$$

(3)

$$P_{HF, d_{ON}}(-0.25,0,0.25) = \sum_{1.08 \leq f < 1.18} \delta f, V$$

Higher-frequency harmonics may thus be emphasized by an AND configured to an on-center under illumination (e.g., by setting $dt < 0$, $+1V$), with respect to other ANDs without light stimulation, representing the off-surround state and producing a lower-frequency response (e.g., by setting $dt = 0$, $+1V$). Cone–bipolar pathways to on-beta ganglion cells were also shown to express high-frequency spiking in response to a step of illumination and low frequency in darkness. The current patterns in Figure 3 show both positive and negative parts. The reason is that at the HRS some of the excited electrons relaxed into deep state in the dielectric (a conductive filament is not yet present). The application of a voltage pulse shifted the electrode Fermi level which in turn affected charge occupation in these traps. Furthermore, it reduced the transport of hot electrons toward the ITO–hafnia interface during illumination. As an electric pulse precharged some of the trap states, once the stimuli were removed, this charged flowed back toward the TE.

2.4. Stimulated Spiking Activity in LRS-OFF

Figure 4 shows the STDP of spectral power in an off-bipolar LRS device. The measurement was carried out using voltage stimulation pulses of opposite $\pm 1V$ polarities (1 Hz, 50% duty cycle for both voltage and light) with a sampling rate of 1/0.1 s, irradiation power density of 100 mW cm$^2$, and bias of 100 mV. Figure 4A,C shows the current $I(t)$ with excitation amplitudes of $-1V$ and $+1V$ with the relative time shift between the pulses indicated next to the plots. The results show that the LRS-ON state outputs a distinctive spiking pattern depending on voltage polarity and pulse time shift. The spiking with inverse current polarity in the plots may be attributed to radio-telegraph noise events. Figure 4B,D shows the relevant spectral power decomposition for the main harmonics in each case (using corresponding colors). The plots reveal that the spatiotemporal current patterns have their low-frequency components emphasized for $dt = -0.25$ s, $+1V$ and $dt = +0.15$ s, $-1V$.

The normalized output power obtained by integration over the main spectral components (marked $P_{+1V}$ and $P_{-1V}$) and divided by the power output of nonelectrical stimulation ($P_0$) for each case (red: $-1V$, blue: $+1V$) is shown in Figure 4E. It is shown that this power may be adjusted by up to $\pm 60\%$ as a function of polarity and stimulation shift. Referring to the center-surround configuration discussed previously, in this case, devices may be operated to output either a high-power (e.g., $dt = 0$, $-1V$) or a low-power (e.g. $dt = 0$, $+1V$) low-frequency output and allow increased accuracy in differentiating between illuminated and dark off-surround cells. Figure 4F shows the spectrum for illumination without the application of voltage pulses (NP), highlighting the main spectral components used in the analysis.

2.5. Electrophotonic-Induced State Transitions

Resistive memory devices are known to allow reversible state transitions (i.e., HRS to LRS and vice versa) as well as gradual conductance adjustments. In the context of this work, it implies that a hybrid electrophotonic responsive AND may be operated to change its polarized response from off-bipolar to on-bipolar and back. Furthermore, the devices contain an additional degree of freedom for state toggling by displaying decremented conductance in response to illumination pulses. These reductions are attributed to light-induced, meta-stable Frenkel pair recombination events in the conductive filament trap states. In this manner, an AND could be operated to generate a direct current (DC) bias in direct proportion to its analog conductive state.

Figure 5 shows the manner in which the AND was operated to undergo a transition from a hyperpolarized response in LRS-OFF state to depolarized spiking of an HRS-ON without electric stimulation and larger irradiation power. The figure shows the different outputs associated with the hyperpolarized steady state and spiking under illumination and the manner in which the device may be toggled between the two modes. Gradual conductance changes after successive illumination sessions are shown in the $I$-$V$ sweeps of Figure 5A. The initial high conductive state is defined as the off-bipolar LRS, whereas the final low conductive state corresponds to an on-bipolar HRS. The time-dependent current in Figure 5B shows a significant hyperpolarization-like drop after being exposed to a flux of about $10^7$ photons per second for 5 s and a bias of 100 mV (the sampling rate was about 1/0.1 s). Self-recovery is evident after a relaxation period of roughly 10–15 s (marked by a green arrow), which corresponds in both shape and length to the trap-discharging pattern.
discussed previously. During recovery, two distinct phases were observed that could indicate the involvement of two types of defects in the process. These were manifested as an initial relatively fast transition (blue arrow), which lasted for few seconds, and a second much slower phase of about 20–30 s (red arrow).

During a second illumination session, using the same parameters, the current showed a similar behavior with a recovery to a lower conductive state that may be associated with Frenkel pair recombination events (Figure 5C). Additional third and forth illuminations with recoveries to even lower conductive states were recorded as well and are shown in Figure 5D,E. The recovery process in this case was temporarily fixated after the fast transition phase, which may be due to a lowered trap density.

Once the $I$–$V$ sweeps were performed (cyan and magenta curves in Figure 5A), the steady-state currents recovered back to their original levels. This helps to verify that indeed detrapping is the main mechanism involved in the recovery as soft electric stress helped to release these trapped electrons (Figure 1I). Once a transition to an HRS was made, the response was modulated using 0.5 s illumination pulses. Figure 5F shows this depolarized-like spiking activity in the HRS.

As mentioned earlier, conductance incremental setting through electric pulsing was accomplished as well. Figure 5G shows $I$–$V$ sweeps exhibiting these conductance changes that resulted from the successive application of 500 ms voltage pulses with different polarities and amplitudes to the TE. The amplitudes in this case were chosen to be relatively high to generate enough electric stress to facilitate trap generation–recombination

---

**Figure 4.** STDP of spectral power in an off-bipolar LRS device. A) Temporal current $i(t)$ with an excitation amplitude of $V = -1$ V for different relative pulse offset timings (the relative riming $dt$ between the light and electric pulses is labeled for each curve). B) Spectral power decomposition for each offset case with matching colors. C) $i(t)$ with an excitation amplitude of $V = +1$ V using the same relative pulse timings. D) Associated spectral power decomposition with corresponding colors. E) The normalized power output as a function of voltage polarity and offset ($dt$). F) Spectral power lines for the case without the application of voltage pulses (NP), highlighting the main components used in the analysis.
Directions of change are marked by colored arrows and corresponding labels indicating the bias and number of pulses. A positive polarity was applied to obtain an incremental increase in conductivity, starting from an initial state (black), whereas a negative polarity was used to gradually reduce it. Resistive memory devices, in particular in the submicrometer regime, were demonstrated to show nonlinear properties. For this reason, different pulse amplitudes were used to manipulate defect

Figure 5. Transition from a hyperpolarized response in an off-bipolar LRS to depolarized spiking of an on-bipolar HRS. A) $I-V$ curves showing the gradual conductance changes after successive illumination sessions. The initial conductive state is the off-bipolar LRS, whereas the final one corresponds to an on-bipolar HRS. B) Hyperpolarization-like current drop in response to a single 5 s light impulse (yellow) with a flux of about $10^7$ photons per second over the device showing a three-stage self-recovery. C) A second illumination session with a similar behavior and recovery to a lower conductive state. D) A third illumination session with a recovery to an even lower conductive state. E) Fourth illumination session. F) Depolarization-like spiking current modulation under the influence of 0.5 s illumination pulse train in an HRS. G) $I-V$ sweeps depicting conductance incremental setting through the application of successive electric stress. H) Long-term photonic depression measured at 100 mV. I) Long-term electric potentiation and depression measured at 100 mV.
creation and annihilation while in the same direction of change. The long-term photonic depression, measured after the exposure to the aforementioned 5 s light impulses, is shown by the orange curve of Figure 5H. This plot shows the rate at which each pulse reduced the steady-state conductance. In addition, the long-term electric potentiation and depression (measured at 100 mV) are shown by the colored curves of Figure 5I. The colors correspond to the pulse-induced conductance changes of Figure 5G.

2.6. Tonic and Phasic Receptor-Like Responses

In the mammalian retina, specialized bipolar cells are able to decompose the high- and low-frequency components of light signal amplitudes. A simplified schematic illustration for such cell arrangement is shown in Figure 6A. Transient bipolar cells produce a response proportional to the high-frequency component (high pass), whereas sustained cells produce a response proportional to the relatively slow changes (low pass). Their outputs are then combined and passed by ganglion cells to the optic nerve. Similarly, the self-recovery speed of ANDs in an HRS-ON state may be utilized, by manipulation of trap type and densities, to produce a similar behavior. This concept is shown by the band diagrams in Figure 6B. Bulk trap densities are directly correlated with the level of soft breakdown and can be controlled through a current compliance setting that limits the total amount of charge flow. Sparse and relatively isolated dielectric traps would capture less charge and result in a faster recovery time once the irradiation is removed. In contrast, abundant surface traps may also be used to generate a slow recovering signal, as demonstrated in this section.

The band diagrams in Figure 6B show oxide traps that are distributed not only spatially but also energetically throughout the dielectric. It should be emphasized that these densities are still far lower than the soft breakdown case which leads to nanoscale conductive filaments formation. The diagram shows a two-stage process. The first is the onset of illumination (marked “1”) where electrons are excited over the energy barrier to produce a photocurrent. In addition to this current, some electrons can relax and occupy empty traps within the dielectric. The second phase (marked “2”) takes place once the light is turned off. At this point, detrapping occurs and the released charge flows to the grounded electrode. A higher trap distribution thus leads to a slower decay of the photoinduced current as more electrons are involved in the process and their associated lifetime has a larger distribution as well.

Untreated (or unterminated) hafnia is known to be rich in surface traps. As part of the study, a photoinduced surface current was compared with the photoinduced bulk one. For this

Figure 6. Tonic and phasic receptor-like responses in an HRS. A) In the biologic retina, different types of bipolar cells respond to either fast or slow signals from photoreceptors. B) A band diagram depiction of low (left-hand side) and high (right-hand side) trap densities in a HRS. In case the dielectric medium has a high trap density, the detrapping process is faster, as depicted by the first diagram (high-pass, fast relaxation). The opposite situation is shown in the second diagram (low-pass, slow relaxation). C) Current mapping used to identify the Au edge and place the CAFM tip for I(t) measurement. The conductance between the tip and the grounded Au electrode occurs through surface states and defects distributed along, and in very close proximity to, the surface, within the area marked by a yellow dashed circle (blue scale bar 4 nm). D) Superimposed photoinduced, time-dependent current of the ITO/hafnia/TiN device and the test structure (illumination periods marked by yellow squares).
purpose, a planar structure was fabricated to allow conductance mainly through surface states. Initially, a current mapping of the surface was done using a conductive atomic force microscope (AFM) to approach a grounded Au electrode edge (Figure 6C). The time-dependent photocurrent was then recorded under irradiation with a power density of 20 mW cm\(^{-2}\). Both the surface and bulk (Figure 1E) photoinduced currents are superimposed in Figure 6D to highlight the difference in recovery times for a tonic and phasic receptor-like response. The associated bandwidth (or 3 dB frequency) is related to the overall decay time \( t \approx 3 \tau \) as \( f_{3dB} = 1/(2\pi\tau) \) with \( I(t) = I_{in}(1 - e^{-t/\tau}) \), indicating the response of the system to a Heaviside step function input. By this relation, the estimated bandwidths of the two structures are 240 and 12 mHz, respectively (note: the sampling rate is about 1/0.25 s).

### 2.7. Center-Surround Configuration

“Center-surround” is a term referring to photoreceptor and bipolar cell configurations that allow enhanced visual perception by differentiating illuminated and dark regions in the retina with better accuracy. In case where light hits a central cone while the perimeter cones are in darkness (commonly referred to as on-center off-surround), the dark-depolarized cones release glutamate and excite adjacent horizontal cells. These, in turn, release neurotransmitters to further inhibit the output of the central cone. A conjugated on-bipolar cell will then get excited and stimulate a ganglion cell. The output of each cone is connected to several bipolar cell types, resulting in parallel communication channels to the inner retina.[29,30,48] When the reverse situation occurs, only off-bipolar cells connected to the central cone will excite the ganglion cell. Ganglion cells fire at a higher frequency in response to an on-center off-surround illumination pattern due to the enhanced hyperpolarization of the central cone by a horizontal cell.

A method of operating the devices in a center-surround scheme, where pulsing can be set to emulate horizontal cell neurotransmitter release, is shown in Figure 7. An array of ANDs is configured to produce complementary signals in response to bitwise illumination patterns using the alternate placement of LRS-OFF and HRS-ON devices. Stimulating the ON and OFF devices with different biasing and nonidentical electro-optic pulse timings may be used to imitate the manner in which retinal ganglion cells spike in response to their conjugated bipolar cells’ polarization state. This functional concept was validated through a discrete-time simulation performed with behavioral models of LRS-OFF and HRS-ON devices over the MathWorks Simulink platform.

More detailed schematic block diagrams of the system are provided in the Supporting Information. Both electric and light modulations are accounted for by square-wave inputs. In this manner, the behavioral model outputs a sinusoidal waveform, having a modulated frequency depending on pulse timing between electric and light stimulations. Figure 7A,C,E,G,I shows various illumination and pulsing setups in a center-surround configuration, whereas Figure 7B,D,F,H,J shows the corresponding simulated proportional power output. It is shown that different output vectors are produced for each case, and the situations may be classified with better accuracy.

It is evident from the measurements that whenever a constant bias is supplied to the LRS-OFF devices they will produce a steady-state output in darkness and zero response under illumination (Figure 6). In addition, the transient output power in an HRS-ON device could be channeled to express a higher-frequency or lower-frequency spiking (relative to the light impulses) as a function of the electric pulse timing (Figure 3). In this manner, HRS may be used to differentiate between an on-center exclusively irradiated state (Figure 7C) and a fully illuminated state (Figure 7G). Negative pulse timing stimulation to the ON cells would result in a high-frequency output, whereas fully synchronized pulsing (\( dt = 0 \)) yields a low-frequency output (Figure 3). Complementary outputs could thus be produced to differentiate between the all-darkness (Figure 7A), only OFF-irradiated (Figure 7E), and fully illuminated (Figure 7G) states.

A simple image filter was implemented in MATLAB according to these principles. This filter was applied over three input images, and the results are shown in Figure 8. The figure shows the manner in which an image may be imprinted directly over an input layer of a cogni-retinal neural network as the center-surround configuration helps to correct blurring and increase contrast. As cells in this input layer would be configured in an ON-center OFF-surround manner, it would help improve the contrast and sharpness of imprinted images. This principal is shown in Figure 8. The filter was initially applied over a high-contrast image (Figure 8A). The output was virtually identical without contrast degradation, as shown in Figure 8B. However, when negative and positive blurred images were used (Figure 8C,E), the output images showed better contrast in comparison (Figure 8D,F). The filter contains a simplified definition of a threshold level, indicating whether each bit is illuminated or not. In the code, it is implemented as the 50% level of the color map. The result is that the horizontal line of the original figure is missing from the rendered images (Figure 8D,F) simply because it lies below this threshold. The filter is very rudimentary and given for illustrative purposes only. A threshold may thus be set to remove some details (horizontal line in the example) which would impede the performance efficiency (e.g., character recognition) of subsequent layers in the neural network for which the ANDs serve as an input.

### 3. Conclusions

In summary, this work demonstrates a bioinspired, CMOS-compatible neuronic device that mimics the combined functionality of retinal photoreceptor and bipolar cells. The presented experimental data give the reader an insight into the functionality under different pulsing and irradiation alignments while being compared with the biologic counterparts. In addition, it was shown that the long-term potentiation and depression parameters may also be incrementally tuned in either an electrical or a photonic manner. Such devices may thus serve as building blocks for intelligent artificial retinas, integrating self-healing sensory and neuromorphic computing into a single artificial neural network. An AND capable of hybrid photonic and electronic responses thus has direct implications for the feasibility of implementing an all-in-one image-sensing neural network. Visual information may be captured as an imprint of data directly
Figure 7. High-contrast bitwise operation in a center-surround configuration. A) In complete darkness, the on-center cell generates a low-power output signal as the AND is at an HRS. In contrast, DC biased off-surround LRS devices output a constant signal. B) The corresponding simulation configuration with an output RMS vector of (2, 0, 2). C) Once the on-center HRS cell is exposed to light, it generates a high-power low-frequency signal (spiking), whereas the unexposed off-surround cells maintain their constant output. D) A corresponding simulated output vector of (2, 1.11, 2). E) Application of DC bias to all cells in the complementary illumination state (on-center cell in darkness and off-surround illuminated) results in an overall zero output. F) A simulated zero output vector state. G) Pulse stimulation with Δt < 0 results in an overall spiking activity with a low-power output from the LRS-OFF devices. This scheme can be used to produce an output pattern that emphasizes the difference between the completely dark and completely illuminated cases. H) Corresponding output vector of (0, 1.61, 0). I) Stimulation with Δt = 0 channels the output power of the HRS-ON device to the low frequency band while maintaining a similar ratio with the outputs of the LRS-OFF devices. J) Corresponding output vector of (0, 1.11, 0).
onto the input layer by altering the bitwise frequency response of individual cells. Image processing can then be done by parallel feed forward propagation of the computation though successive layers of the same network based on the same devices. Weight patterns in each layer could be predetermined by a training phase onto the input layer by altering the bitwise frequency response of individual cells. Image processing can then be done by parallel feed forward propagation of the computation though successive layers of the same network based on the same devices. Weight patterns in each layer could be predetermined by a training phase based on the task at hand. Such an artificial cogni-retina (i.e., thinking eye) has the potential to perform image recognition tasks much faster than software algorithms executed over von Neumann machines.

4. Experimental Section

CAFM Sample Preparation: First, a blanket sample of ≈20 nm uncapped hafnia (HfO₂) thin film was fabricated on a p-Si substrate. The hafnia was formed by atomic layer deposition (ALD) using tetrakis(dimethylamino) hafnium as the metal precursor and H₂O vapor as the oxidizer. The growth temperature and pressure were 250 °C and 0.1 Torr, respectively, without postdeposition annealing. Second, a lithographic process was used to pattern 1 μm-wide rectangular structures over the oxide followed by a descum process and reactive ion etching of 5 nm into the hafnia to create a shallow trench. A gold thin layer at 110 °C for 10 min.

ITO/Hafnia/TiN Device Preparation: Electron-beam lithography (EBL) was first performed on a precleaned SiO₂/Si substrate, followed by DC reactive magnetron sputtering for TiN deposition (Ti:N ratio of 1:1:1) and lift-off to form the BE. ALD of the HfO₂ layer at 250 °C was then conducted with tetrakis(dimethylamino) hafnium as the precursor and H₂O as the oxidizing agent. After the second EBL step, ITO deposition was baked at 110 °C/C and 0.1 Torr, respectively, without postdeposition annealing. Second, a lithographic process was used to pattern 1 μm-wide rectangular structures over the oxide followed by a descum process and reactive ion etching of 5 nm into the hafnia to create a shallow trench. A gold thin film layer at 100 nm was then evaporated followed by a lift-off process and diluted isopropyl alcohol cleaning. Finally, the samples were baked at 110 °C for 10 min.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.
