



Self-powered high-sensitivity all-in-one vertical tribo-transistor device for multi-sensing-memory-computing

Received: 17 August 2022

Accepted: 13 December 2022

Published online: 23 December 2022

 Check for updates

Yaqian Liu^{1,2}, Di Liu^{1,3}, Changsong Gao^{1,3}, Xianghong Zhang ^{1,3}, Rengjian Yu^{1,3},
Xiumei Wang^{1,3}, Enlong Li^{1,3}, Yuanyuan Hu⁴, Tailiang Guo^{1,3} &
Huipeng Chen ✉^{1,3}

Devices with sensing-memory-computing capability for the detection, recognition and memorization of real time sensory information could simplify data conversion, transmission, storage, and operations between different blocks in conventional chips, which are invaluable and sought-after to offer critical benefits of accomplishing diverse functions, simple design, and efficient computing simultaneously in the internet of things (IOT) era. Here, we develop a self-powered vertical tribo-transistor (VTT) based on MXenes for multi-sensing-memory-computing function and multi-task emotion recognition, which integrates triboelectric nanogenerator (TENG) and transistor in a single device with the simple configuration of vertical organic field effect transistor (VOFET). The tribo-potential is found to be able to tune ionic migration in insulating layer and Schottky barrier height at the MXene/semiconductor interface, and thus modulate the conductive channel between MXene and drain electrode. Meanwhile, the sensing sensitivity can be significantly improved by 711 times over the single TENG device, and the VTT exhibits excellent multi-sensing-memory-computing function. Importantly, based on this function, the multi-sensing integration and multi-model emotion recognition are constructed, which improves the emotion recognition accuracy up to 94.05% with reliability. This simple structure and self-powered VTT device exhibits high sensitivity, high efficiency and high accuracy, which provides application prospects in future human-mechanical interaction, IOT and high-level intelligence.

With the upsurge of artificial intelligence applications and the rapid development of internet of things (IOT), the time-efficient sensing information acquisition, energy-efficient data memorizing and processing capabilities of terminal electronic systems are indispensable¹. However, in conventional chips, the separated sensing, storage, and

processing units always need to collect signals by external sensors, convert signals into digital format data subsequently, and then transfer the data to memory units and processors for subsequent processing tasks, which limits the data conversion and movement and results in low computing efficiency and huge power consumption². Thus,

¹Institute of Optoelectronic Display, National & Local United Engineering Lab of Flat Panel Display Technology, Fuzhou University, Fuzhou 350002, China.

²School of Physics and Electronic Engineering, Zhengzhou University of Light Industry, Henan 450002, China. ³Fujian Science & Technology Innovation Laboratory for Optoelectronic Information of China, Fuzhou 350100, China. ⁴State Key Laboratory for Chemo/Biosensing and Chemometrics, School of Physics and Electronics, Hunan University, Changsha 410082, China. ✉e-mail: hpchen@fzu.edu.cn

benefiting from in-memory-computing device³, which can both store data and compute simultaneously by device physics or other physical laws, an integrated sensing-memory-computing (SMC) system with high-efficiency perception, storage and calculation functions could greatly simplify transmission operations, decrease hardware bulk, and increase efficiency in comparison to the traditional chips^{1,4–8}.

Meanwhile, most creatures possess multiple sensory organs with which they can simultaneously perceive a wide variety of physical changes in the environment^{9–16}. Thus, SMC systems that can integrate and organize various sensory information (multi-sensing-memory-computing (MSMC)) are fundamental to effective perception and cognitive function^{1,4,11,17–20}. Noticeably, MSMC system desires various numbers of sensory receptors and processing nodes (tactile, auditory, and visual, etc) to keep producing multi-sensing raw data and processing different types of sensory information, respectively^{15,21–28}. However, the increased number of receptors slows down the working speed due to the subsequent data transmission and processing, and the separated sensory receptors and processing nodes also lead to transmission speed discrepancy, which would further limit the conversion speed and increase energy consumption^{24–26}. Therefore, developing a single device integrated with the function of MSMC is of great significance to address the above-mentioned challenges and improve sensory perceptual efficiency in current electronic devices.

Noticeably, triboelectric nanogenerator (TENG) is energy harvesting device that can convert various forms of energy such as human motion, sound vibration, and light energy into electric power^{29–36}. The native advantages of TENG render it a promising power supply device as well as a multisensory receptor^{33,37–42}. Meanwhile, vertical organic field effect transistors (VOFETs) composed of vertically stacked gate/source/drain electrodes and promising short channel possess small subthreshold swing (SS), high working frequency, and promising mechanical stability, ensuring their great practical applications in memory and computing devices^{43–46}. Thus, the inherent vertically stacked electrodes of VOFETs are suitable to integrate with TENG, which possibly provides an effective approach for achieving multifunctional all-in-one-device with sensing-memory-computing function.

Hence, in this work, we develop a vertical tribo-transistor (VTT) device based on TENG and VOFET to implement the multi-sensing-memory-computing function and the interaction of multisensory integration. This VTT is based on a simple configuration of VOFET without any redundant layers, and MXenes function as the top electrode of TENG, source electrode of transistor, and the light collection layer of multisensory, simultaneously. The VTT allows electrostatic induction and tribo-potential to tune ionic migration and Schottky barrier, and thus the triboelectrification sensory information can be amplified to the source-drain current in a self-energy transducing manner, which improves the sensitivity by 711 times over a single TENG. Meanwhile, processing functions of different sensory perception and multisensory interaction in brain superior colliculus are achieved based on the individual VTT device. Furthermore, artificial conscious response is generated by the integration of robot hand with VTT device, and the open angle of robot hand is successfully controlled with different sensory stimulation, demonstrating that this system can straightforward enhance the accuracy of relevant event. Finally, a multi-model emotion recognition system is constructed to detect and distinguish typical mood. The self-powered VTT device shows great potential in next-generation high-performance in-sensor-memory-computing artificial intelligent system and human-computer interaction interface applications.

Results

Self-powered vertical tribo-transistor with multi-sensing function

A schematic diagram of biological multisensory integration nervous system is illustrated in Fig. 1a. The system is composed of several

sensing units for perception, transmission pathways for transfer information, and brain for memory and computing. Here, a vertical tribo-transistor (VTT) device based on TENG (Fig. 1b) and VOFET (Fig. 1c) is fabricated to implement the multi-sensing-memory-computing function, as illustrated in Fig. 1d. Clearly, this VTT integrates the function of a TENG and a VOFET. It is based on a VOFET configuration without any redundant layer, which is attributed to the fact that both TENG device and VOFET device share the vertical structure. The VTT composes removable Kapton substrate, removable Au gate electrode, ion-gel insulating layer, MXenes network source electrode, PDVT-10 channel layer, and Au source/drain electrode. Meanwhile, the detailed fabrication process of VTT is shown in Supplementary Fig. 1, and the cross-section SEM is depicted in Supplementary Figure 2. Noticeably, MXenes, which possess regulated work function, outstanding mechanical and optical properties, have raises significant interest for optoelectronic, energy storage and Schottky-barrier based devices. Thus, MXenes are promising alternative electrode of TENG and network source electrode of VTT here. The sensing abilities, touch and hearing perception are emulated by moving of gate electrode and vibration of VTT, respectively, and the visual perception is achieved by the light-sensitive $\text{Ti}_3\text{C}_2\text{T}_x$ MXenes electrode.

The scanning electron microscope (SEM) image of MXenes network is presented in Fig. 1e, which clearly shows that the MXenes are interconnected with each other and form a network structure. Meanwhile, the transmission electron microscope (TEM) of MXenes is shown in Fig. 1f, where MXenes exhibit nanosheets structure, and the thickness of MXenes network source electrode is about 1 nm. Fig. 1g shows the high-resolution TEM image, and the corresponding lattice spacing in this layer is 0.3 nm⁴⁵. The electrostatic induction and triboelectrification between gate electrode and insulating layer would induce the electrode-double-layer effects, which manifests in a transient channel current, thus realizing the multi-sensing-memory-computing function.

We first investigated the sensing function of this VTT device, and the equivalent circuit of VTT is demonstrated in Fig. 2a. The gate electrode in Fig. 2a can be moved, and the distance between gate electrode and insulating layer is defined as *d*. Compared with traditional plane transistor integrated with TENG (Supplementary Fig. 3a, 3b), this VTT not only enhances the device integration level, but also simplifies the equivalent circuit (Supplementary Fig. 3). Clearly, VTT can achieve the function of TENG by only leading out the gate and source electrode, while if external voltage is applied to the source-drain electrodes, it can function as VOFET. The output performance of individual TENG is shown in Fig. 2b. With the increase of the distance between gate electrode and insulating layer from 0 to 1000 μm , the open-circuit voltage (V_{oc}) increases from 0 to 1.71 V. The working mechanism of TENG is illustrated in Supplementary Fig. 4. Here, ion-gel is employed as the triboelectric material in TENG. The Au electrode and MXenes electrode are employed as the bottom electrode and top electrode in TENG, respectively. When the distance between bottom electrode and triboelectric layer changes, and the confined charges in ion-gel attract the counterions. In the original state, the bottom electrode and ion-gel layer are not in contact with each other, and then when bottom electrode is fully contacted with ion-gel layer (IL), electrons flow to the IL because of the triboelectric effect. When bottom electrode is moved away, positive and negative ions in the IL accumulate at the air/ion-gel and ion-gel/top electrode interfaces, respectively, forming an electrical double layer (EDL). Accordingly, an output voltage is recorded with the electrons flowing from top electrode to bottom electrode. It is noteworthy that the high capacitance of EDL can remarkably improve the performance of TENG (as illustrated in Supplementary Note 1)^{37,45,46}.

Then, we investigated the property of the devices as a VOFET. The typical transfer curve of VOFET is illustrated in Fig. 2c, where gate voltage is modulated through external voltage, and the gate electrode is contacted with the insulating layer (Supplementary Fig. 5). From the transfer characteristics of VOFET⁴³, the subthreshold swing (SS) of

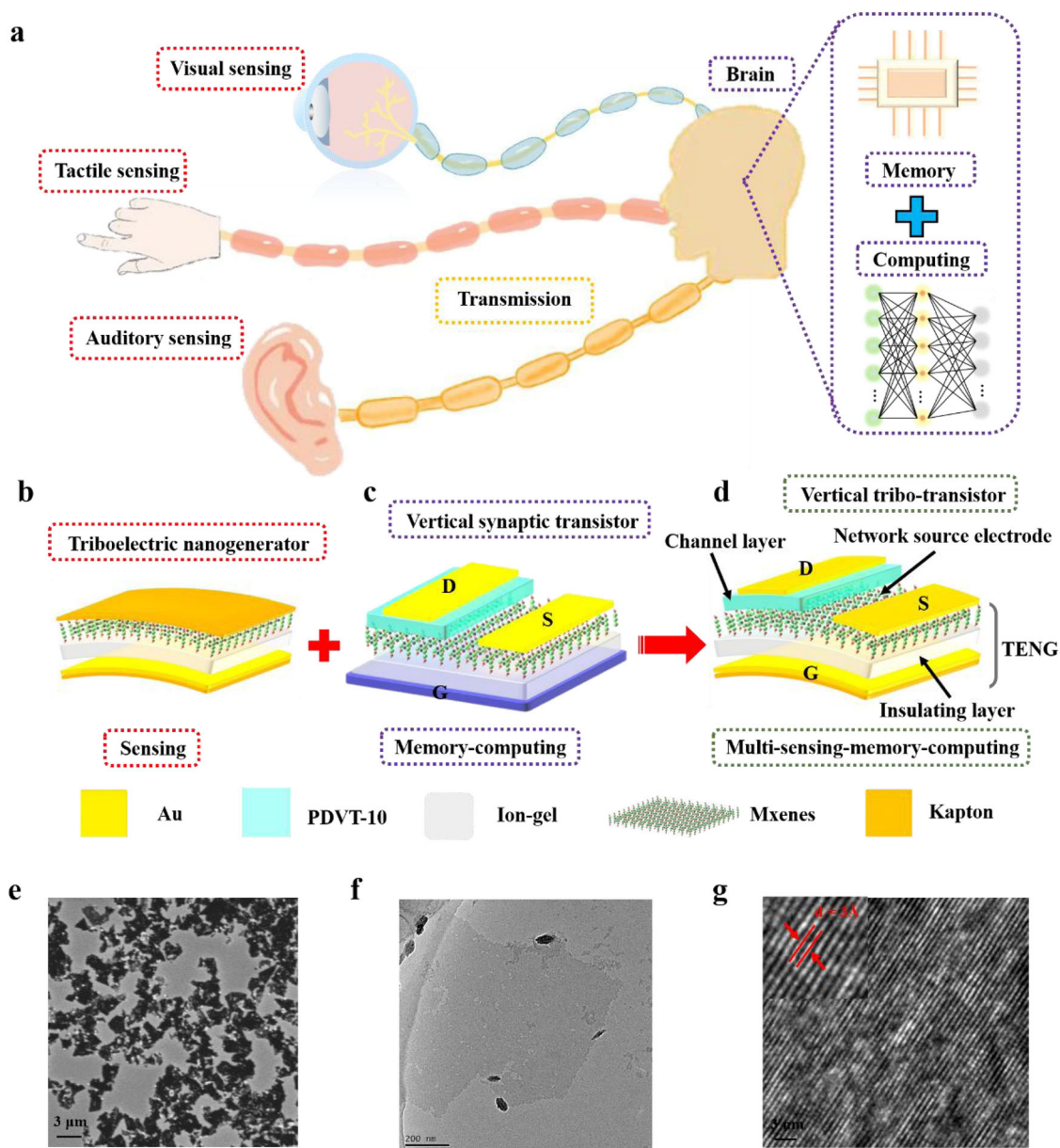


Fig. 1 | Schematic diagram of multi-sensing-memory-computing system, the device structure of VTT and its morphology characteristics. **a** The biological multi-sensing-memory-computing system. The sensory receptors collect sensory data and transmitted the data to human brain for further memory and computing.

b–d Schematic illustration the structure of TENG, vertical transistor, and VTT, respectively. **e** SEM images of MXenes network source electrode. **f** SEM images of MXenes network source electrode. **g** High-resolution TEM image of MXenes network source electrode.

91.8 mV/dec is extracted, while lower SS would result in higher ΔI ($I_{DS1}-I_{DS2}$) at same ΔV ($V_{GS1}-V_{GS2}$), which shows the great potential of the device in achieving high sensitivity sensory perception. Noticeably, the V_{OC} of TENG can be considered as a gate-source voltage and then convert into a transient channel current. Figure 2d and Fig. 2e show the transfer and output curve of VTT under different V_{OC} from TENG, respectively. When the distance between gate electrode and insulating layer increases from 0 to 1000 μm (Fig. 2d), the VTT exhibits p-type transistor behavior, and the drain-source current (I_{DS}) increases by nearly 4 orders. Moreover, the on-state current density is greater than 5 mA/cm² (Fig. 2e), providing sufficient current density to drive organic electronic devices. This high output current density of VTT is mainly ascribed to its ultrashort channel length. For comparison, the transfer curve of planar transistor is depicted in Fig. 2c left, which shows a higher SS compared with vertical transistor.

Furthermore, we investigated the sensitivity of these two devices individually when they are used for mimicking different sensory perception. Fig 2f, g show the tactile sensitivity of an individual TENG and VTT, respectively. The tactile sensitivity of TENG is defined as eq. 1:

$$S = (\Delta V/V_{\min})/\Delta d \quad (1)$$

where ΔV is the change of open-circuit V_{oc} ($V-V_0$), V_{\min} is the minimum V_{oc} , and Δd is the changed distance (d) between gate electrode and insulating layer. When d is increased from 0 μm to 450 μm , the sensitivity is calculated to be 0.0334 μm^{-1} , and then a lower sensitivity of 0.0038 μm^{-1} is estimated when the d increases from 450 μm to 1000 μm . In comparison, the tactile sensitivity of VTT is defined as eq. 2:

$$S = (\Delta I/I_{\min})/\Delta d \quad (2)$$

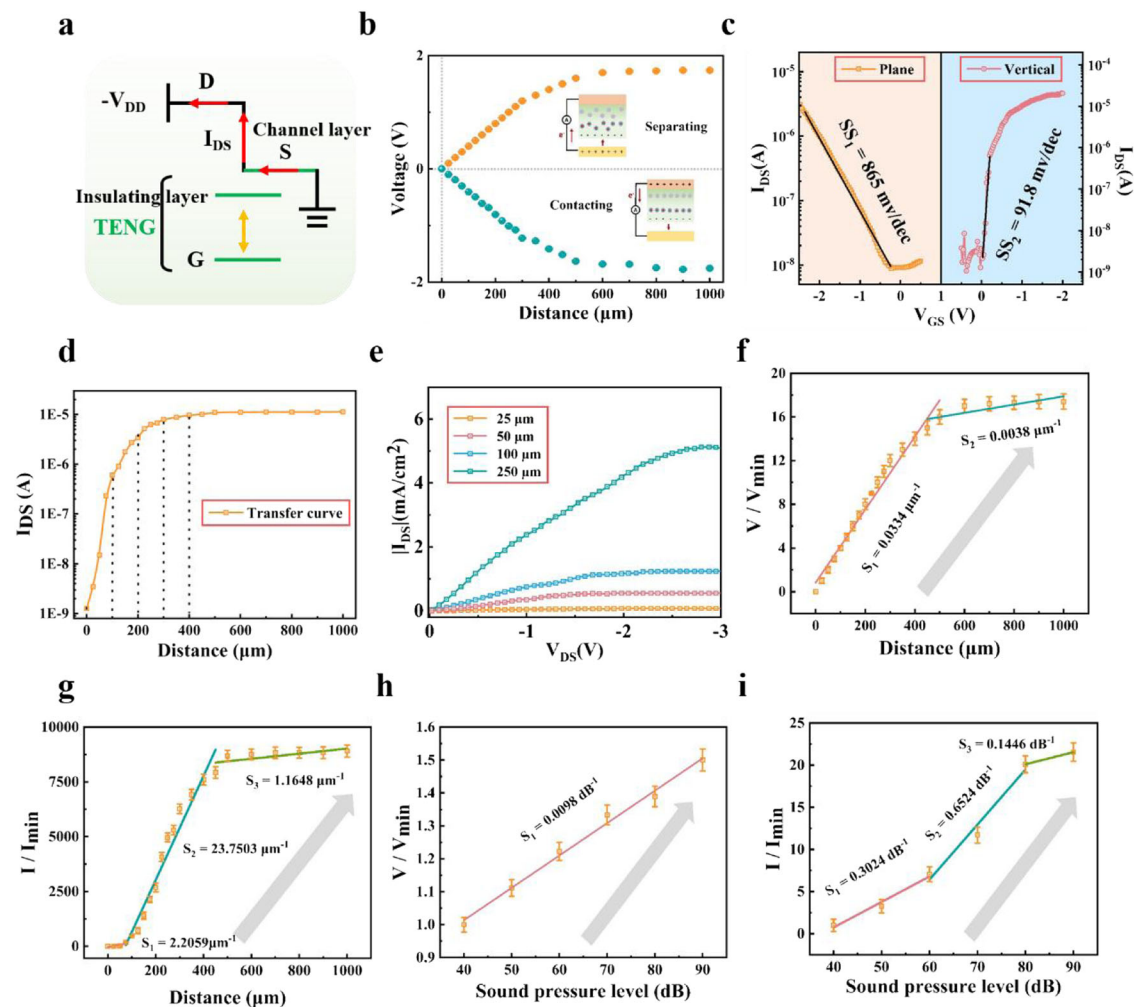


Fig. 2 | The electrical output and sensing properties of TENG and VTT.

a Equivalent circuit of the VTT device. **b** The TENG output voltage versus distance. **c** The SS of VTT with external gate voltage. **d, e** Transfer curve and output curve of VTT with TENG. **f, g** The tactile sensitivity of single TENG and VTT with different

distance. **h, i** The auditory sensitivity of single TENG and VTT with different sound pressure level. The error bars in f-i means the values of V/V_{min} or I/I_{min} within 10 cycles.

where ΔI is the change of drain-source current ΔI_{DS} ($I-I_0$), I_{\min} is the minimum I_{sc} , and Δd is the changed distance (d) between gate electrode and insulating layer. As depicted in Fig. 2g, the VTT initially exhibits a sensitivity of $2.2059 \mu\text{m}^{-1}$, and then, when the d is increased from $75 \mu\text{m}$ to $400 \mu\text{m}$, the sensitivity (S_2) is calculated to be $23.7503 \mu\text{m}^{-1}$, which is 711 times higher than that of the individual TENG. Finally, with the d exceeding $400 \mu\text{m}$, which corresponds with a saturated I_{DS} and a lower ΔI , a rapid decrease of the sensitivity to $1.1648 \mu\text{m}^{-1}$ is observed. Thus, different distance between gate electrode and insulating layer would result in different charge distribution, V_{oc} , I_{DS} , and sensitivity.

Besides, we characterized the auditory and visual stimuli transduction sensitivity of individual TENG and VTT, respectively. For the auditory stimuli transduction sensitivity, a loudspeaker with tunable sound pressure level (SPL) and frequency is used as the acoustic source to trigger the VTT. The auditory sensitivity is defined as eq. 3:

$$S = (\Delta V(\Delta I) / V_{\min}(I_{\min})) / \Delta \text{SPL} \quad (3)$$

where ΔSPL is the changed sound pressure level of the loudspeaker. As the sound pressure level increases from 40 to 90 dB, the individual TENG device exhibits a nearly linear response and a SPL sensitivity of 0.0098 dB^{-1} (Fig. 2h). Meanwhile, the dependence of sensitivity of VTT

on SPL is shown in Fig. 2i, and the highest sensitivity is 0.6524 dB^{-1} in the range from 60 to 80 dB. Furthermore, the absorbance of MXenes is presented in Supplementary Fig. 6, which exhibits particularly high absorption intensity in ultraviolet (UV) region, and the wavelength light source used is 325 nm next. The photosensitive MXenes function as the top electrode of TENG, source electrode of transistor, and the light collection layer of multisensory, simultaneously, and the visual sensitivity of individual TENG and VTT are illustrated in Supplementary Fig. 7. Clearly, VTT exhibits higher sensitivity than the individual TENG for each sensory perception, and the higher sensitivity of VTT is attributed to the amplification, low SS and high on/off ratio of VOFET. Moreover, the function of each material/layer in VTT with different sensing is shown in Supplementary Fig. 8, which further indicates that this VTT is beneficial to achieve multi-sensing function.

The working mechanism of self-powered vertical tribo-transistor (VTT)

Based on the above results, the working mechanism of VTT unit for tactile signal is illustrated in Fig. 3. In the initial state, the gate electrode is separated with the IL, and then the VTT is set to the original state (the process was depicted Supplementary Fig. 9). As the gate electrode is fully contacted with IL, positive charges and negative charges accumulated on the surface of gate electrode and IL interface because of

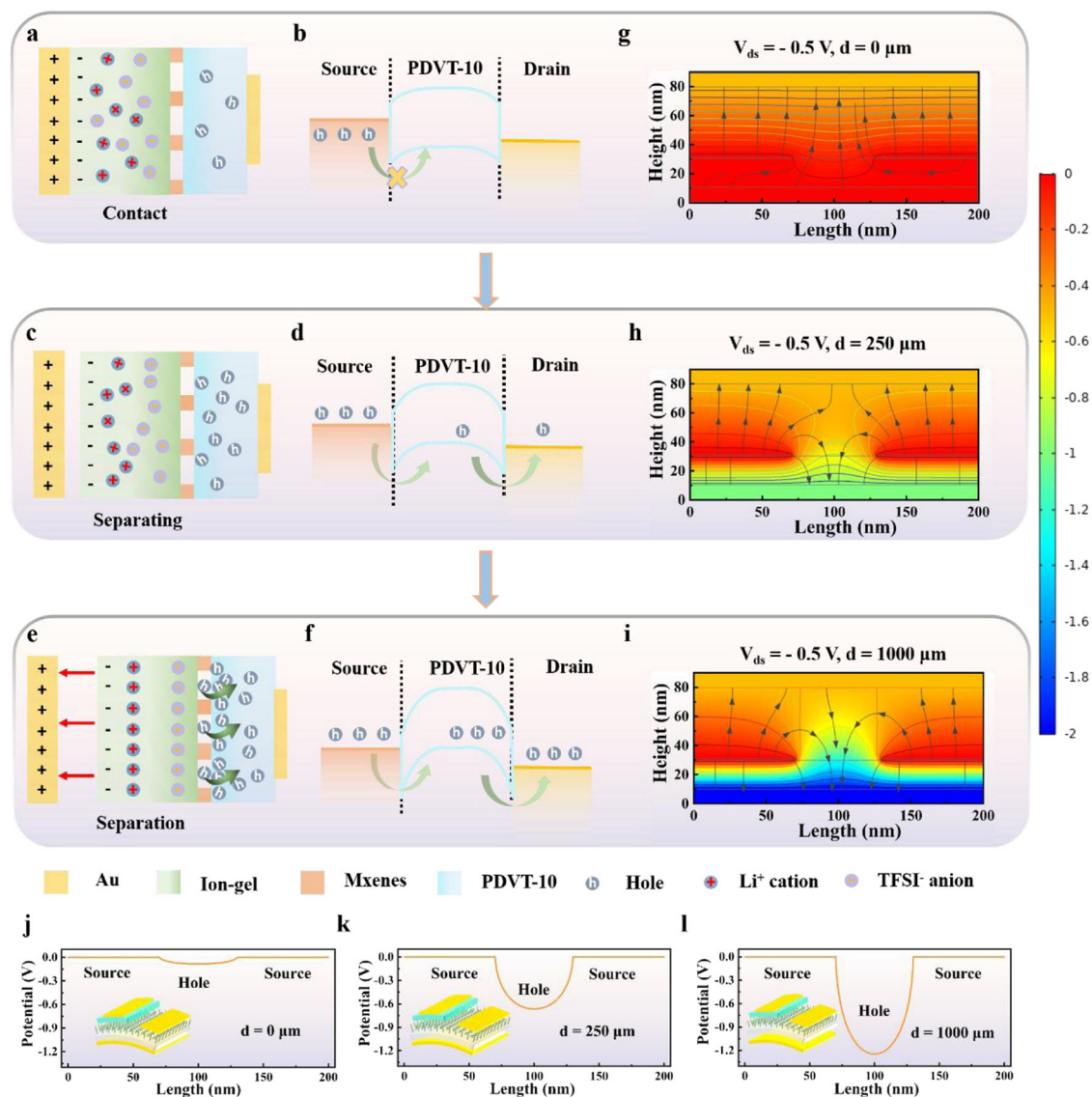


Fig. 3 | The working mechanism of multi-sensing VTT unit. a Charge distribution of the source/channel/drain layers when gate is in contact with the IL layer. **b** Corresponding band diagram of each layer at contact state. **c** Charge distribution of the source/channel/drain layers when gate is separated from the IL layer. **d** Corresponding band diagram of the source/channel/drain layers when gate is

separated from the IL layer. **e** Charge distribution of each layer at separation state. **f** Corresponding band diagram of each layer at separation state. **g–i** The electrical field distributions of VTT under three different states. **j–l** Potential distribution at MXenes electrode.

the electrostatic induction and triboelectrification, respectively, as depict in Fig. 3a. Noticeably, the identical positive charges and negative charges lead to the off state of VTT, and there is no current recorded when a source-drain voltage is applied. The corresponding band diagram of this state is illustrates in Fig. 3b. It is clear that energy band bending would occur at MXenes/PDVT-10 interface when MXenes come into contact with PDVT-10 semiconductor. The high Schottky barrier height between the MXenes and PDVT-10 causes small source-drain current. Next, when the gate electrode and IL are separated, the tribo-potential would regulate the ionic migration of ion-gel layer, and positive ions in the IL are induced based on the charge balance effect (Fig. 3c). Simultaneously, an EDL is gradually generated at the IL, where the number of negative ions balanced the positive charges. Accordingly, temporary holes are accumulated near IL/PDVT-10 interface, which would decrease Schottky barrier height between the IL and PDVT-10, and then a conductive channel from MXene to drain electrode is formed (Fig. 3d). Finally, when gate electrode and IL are further separated, the EDL is entirely generated at the IL because of

the tribo-potential, as shown in Fig. 3e. At this stage, massive holes accumulate at the IL/PDVT-10 interface, and thus the width of the Schottky barrier between the HOMO energy level of PDVT-10 and MXenes work function is narrowed, as shown in Fig. 3f, which enables the holes to easily flow into the semiconductor layer. Thus, holes are injected into PDVT-10 semiconductor from MXenes-network source electrode under the source/drain voltage and then are vertically transferred from PDVT-10 channel layer to drain electrode, resulting in the formation of drain-source current (on state).

To further understand the working mechanism and charge transport in VTT, a 2D cross-section simulation model is exploited to demonstrate the electrical field distributions and electrical potential during the aforementioned three states, as depicted in Fig. 3g–l. Notably, the drain-source current of VTT is dependent on the distance between gate electrode and IL, and different distances would result in different gate voltage. Thus, the structure of VTT in COMSOL Multiphysics is simply considered as a typical VOFET with varied gate voltage, as depicted in Supplementary Fig. 10. Figure 3g–i show the

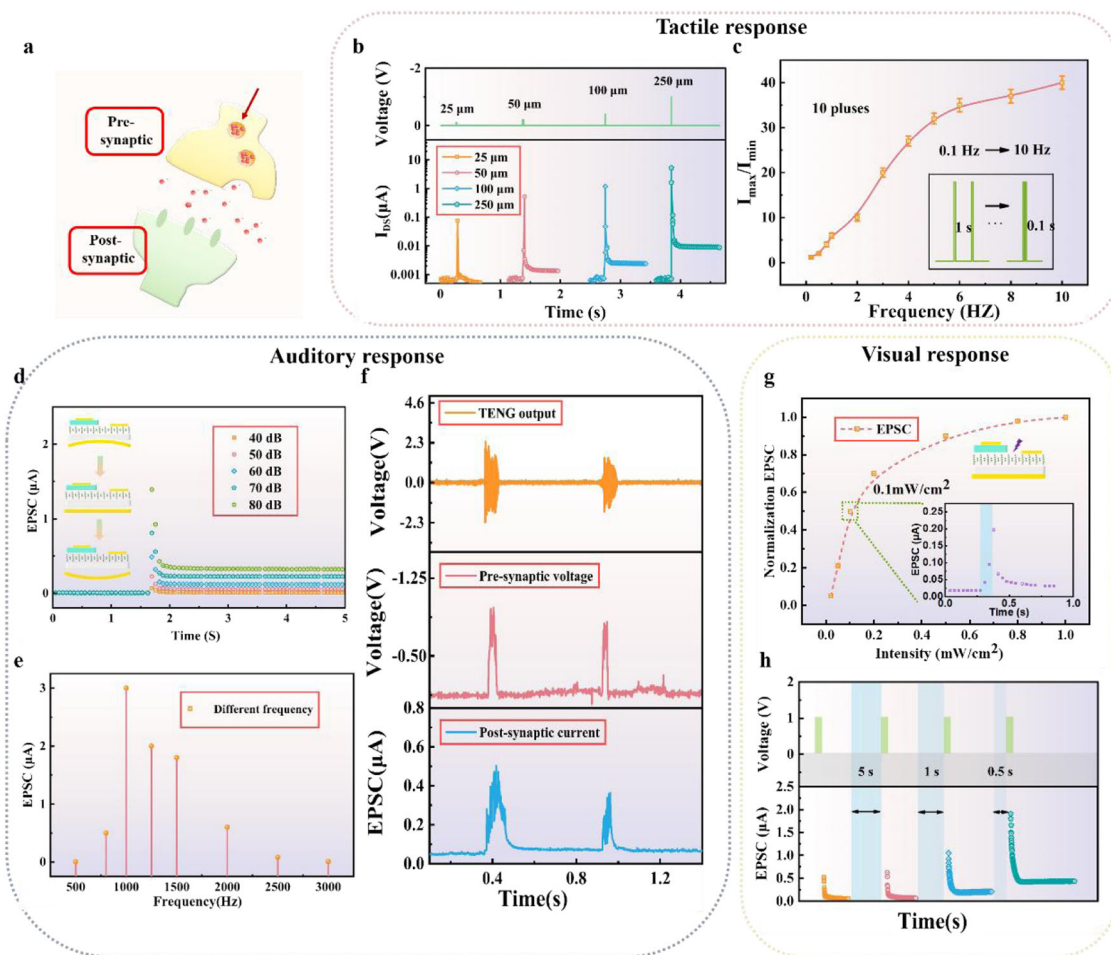


Fig. 4 | Synapse response of tactile, auditory, and visual. **a** The sensory transduction and memory-computing process. The tactile response of our VTT: **b** The EPSC with different distance and **c** I_{\max}/I_{\min} with different pulse frequencies. The error bar in **c** means the value of I_{\max}/I_{\min} within ten cycles. The auditory response of our VTT: **d** EPSCs of VTT with sound pressure level ranging from 40 dB to 80 dB.

e EPSCs of VTT with different sound frequencies. **f** Recorded TENG output, pre-synaptic voltage and corresponding post-synaptic current. **g** The visual response of our VTT Normalization EPSCs with different light intensity. **h** EPSC amplitudes with two successive light pluses with different interval time.

electrical field distributions of VTT under three different states, which indicates that the distance between gate electrode and IL could effectively control the charge behavior at the IL/PDVT-10 interface. Furthermore, Fig. 3j–l show the 1D potential distribution of MXenes electrode position (Supplementary Fig. 10). It is observed that the potential becomes more negative as the distance increased, indicating stronger band-bending at the Mxene/semiconductor interface, which is favorable for the injection of holes into the PDVT-10 channel. Moreover, the working mechanism of acoustic signal into electrical signal is illustrated in Supplementary Fig. S11.

Self-powered vertical tribo-transistor with multi-sensing-memory function

Furthermore, the multi-sensing-memory function of our VTT is demonstrated. Note that EDL effect is generated at ion-gel layer, which is accompanied with retentive effects and the ion transport in EDL, resulting in synaptic behaviors of VTT. In our sensing-memory-computing system, the sensory transduction process can be regarded as the presynaptic process, and the memory-computing process can be regarded as the postsynaptic process, as depicts in Fig. 4a. We first study the synaptic plasticity of the VTT with different tactile, auditory, and visual input signals. Fig 4b, c show the tactile response of this VTT, where the gate voltage is controlled by the distance between gate electrode and IL, and the drain-source voltage is fixed at a

constant value of -0.5 V. With the distance increased from $25 \mu\text{m}$ to $250 \mu\text{m}$, the equivalent gate voltage is increased from 0.5 V to 1.5 V, and the post-synaptic current amplitude of I_{DS} is augmented from $0.1 \mu\text{A}$ to $8 \mu\text{A}$, as shown in Fig. 4b. Meanwhile, the EPSC peak as a function of pulse time with a fixed distance of $25 \mu\text{m}$ is depicted in Supplementary Fig. 12. The EPSC firstly shows a linear enhancement when the pulse time is below 300 ms and then becomes saturated when the pulse time increases from 300 ms to 1000 ms. Supplementary Fig. 13 shows the responding time and decay time of the tactile EPSC, and the responding time is increased from 5 ms to 40 ms with the distance increased from $25 \mu\text{m}$ to $250 \mu\text{m}$. Clearly, larger distance will result in longer time of gate electrode and ion-gel insulating layer to contact with each other and then induce more ions and carries. Moreover, paired-pulse facilitation (PPF) is also presented in Supplementary Fig. 14. As the interval time of two continuous pluses is shorter than the ion transfer time, an obvious increase of the second EPSC amplitude can be observed. Thus, we further investigated the EPSC amplitude under different pulse frequencies. As illustrated in Fig. 4c, with the frequency changed from 1 Hz to 10 Hz, I_{\max}/I_{\min} (I_{\max} is the 10th EPSC amplitude, I_{\min} is the 1st EPSC amplitude) increases from 1.2 to 40 , which is critical for future neuromorphic computing such as dynamic high-pass filter.

Additionally, for the auditory response, the EPSCs of the VTT with a sound pressure level ranging from 40 dB to 80 dB are shown in

Fig. 4d. When the auditory pulse (100 ms) is applied to VTT, the EPSC rapidly reaches a peak value and then gradually decays to resting current. Meanwhile, the magnitude of EPSC increases from 0.5 μA to 3.2 μA when the sound pressure level increases from 40 dB to 80 dB, and the responding time and decay time of EPSC is also increased with the sound pressure level increased (Supplementary Fig. 12). Moreover, the effect of different sound frequencies (100 ms) on the EPSCs of the VTT is depicted in Fig. 4e. The EPSC amplitudes first increases and then decreases with the frequencies ranging from 500 Hz to 3000 Hz, and the EPSC reaches its peak at the frequency of 1000 Hz. Fig 4f shows the V_{oc} of TENG, transferred pre-synaptic voltage, and corresponding post-synaptic current is driven by two different voice signals. Clearly, different voice signals can be distinguished with different post-synaptic currents.

Furthermore, the visual response of VTT is illustrated in Fig. 4hg. The temporal response of normalized EPSCs as function of light intensity is demonstrated in Fig. 4h, and the inset shows a typical EPSC response with the intensity of 0.1 mW/cm². The amplitudes of EPSC increase almost linearly within a small light intensity range below 0.1 mW/cm², and then a saturated EPSC would be recorded with further increase of the light intensity. Meanwhile, the responding time and decay time are also shown in Supplementary Fig. 12. The working mechanism of VTT under light is shown in Supplementary Fig. 15.

Figure 4g illustrates the application of two successive light pulses with different interval time (5 s, 1 s, and 0.5 s), where the second EPSC peak increases with the decreases of the interval time. The increased EPSC value after the first light pulse and the second (consecutive) light pulse are denoted as EPSC₁ and EPSC₂, respectively. Supplementary Fig. 16 depicts the ratio of EPSC₂/EPSC₁ as a function of interval time, which is an indicator of enhancement of synaptic connections. The increased ratio suggests that the ion transport-mitigating layer contribute notably to maintain synaptic connections. Furthermore, VTT shows high endurance and stability with several cycles and days, respectively, as shows in Supplementary Fig. 17. These multi-sensing synaptic functions of our VTT lay foundation for realizing artificial intelligence system with sensing-memory-computing ability.

Self-powered vertical tribo-transistor with multi-sensing-memory-computing (MSMC) function

Importantly, the function of MSMC in the human brain can associate and learn crossmodal information. Inspired by this fact, we then investigated the MSMC function of our VTT device. Figure 5a depicts the traditional digit MSMC system, which contains several components, such as the signal acquisition part (receptors), the power supply part, the preliminary processing parts, and the processing unit. However, the separated analog-digital hybrid circuit, memory, computing,

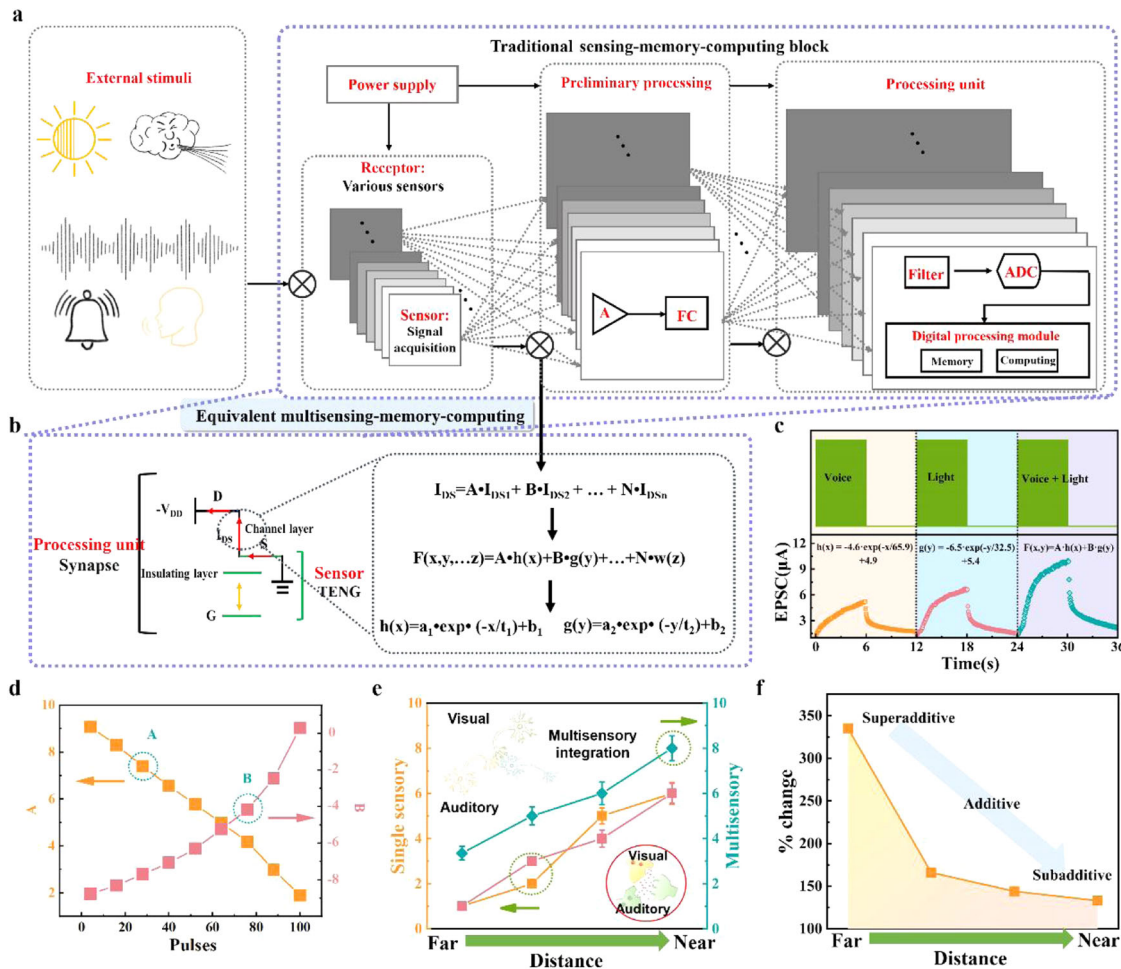


Fig. 5 | Multi-sensing-memory-computing model and mimicking superior colliculus. a External stimuli of our environment and traditional sensing-memory-computing block, which consists of power supply unit, receptor unit, preliminary processing unit, and processing unit. **b** Equivalent multi-sensing-memory-computing of our VTT and corresponding cross-correlation functions between single sensing-memory-computing and multi-sensing-memory-computing. **c** EPSC

with different stimuli, which can be well-fitted by cross-correlation functions. **d** The change of A and B with increased external sensing pulses (single voice pulses, and light pulses; or voice+light pulses). **e** Single sensory stimuli and multisensory stimuli with different sensory strength. The error bar in (e) means the stimuli of single sensory and multisensory within ten cycles. **f** The % change from superadditive, additive, to subadditive.

convolution computation and other hardware result in slow processing speed and high power consumption. Our self-powered VTT device is illustrated in Fig. 5b, the TENG component in VTT is regarded as the sensor of the MSMC system, and VTT is regarded as the preliminary and processing unit. Since the working mechanism of VTT under mechanical and optical stimuli is different, such device is attractive for processing multi-sensing information without an extra preliminary processing unit. Here, the output current (I_{DS}) of VTT is depicted as $I_{DS} = A \times I_{DS1} + B \times I_{DS2} + \dots + N \times I_{DSn}$, where I_{DS1} , I_{DS2} , and I_{DSn} are the EPSC with different external stimuli (visual, auditory, tactile, etc.). Meanwhile, the expression of I_{DS} is depicted as eq. (4):

$$F(x, y, \dots, z) = A \times h(x) + B \times g(y) + \dots + N \times w(z) \quad (4)$$

where $h(x)$, $g(x)$, and $w(z)$ are I_{DS1} , I_{DS2} , and I_{DSn} of VTT, respectively, A , B , and N are constants. For example, as shown in Fig. 5c, the EPSC of VTT under auditory stimuli is fitted as $h(x) = -6.5 \times \exp \times (-x/65.9) + 4.9$, and the EPSC of VTT under optical stimuli is fitted as $g(y) = -6.5 \times \exp \times (-y/32.5) + 5.4$. Based on the change of EPSC amplitude, the response under auditory and optical stimuli simultaneously can be fitted as $F(x, y) = A \times h(x) + B \times g(y)$. Noticeably, with the increases of stimuli pluses, the constant of A decreases, while B increases, as illustrated in Fig. 5d. We further investigate the MSMC function of VTT with tactile and visual response simultaneously, as illustrated in Supplementary Fig. 18. This phenomenon further demonstrates that our artificial multisensory device with the MSMC function could facilitate the parallel processing of large amounts of external stimuli information, which decreases the data exchange between storage and computation, increases working speed, and decreases power consumption compared with traditional CMOS architecture.

Furthermore, considering that multisensory integration is a critical and ongoing determinant of human behavior. Noticeably, inputs from different sensory neurons are converged on cells in the superior colliculus, and then result in multisensory integration. Here, we mimicked the response of superior colliculus in the human brain upon the approaching of a dog to demonstrate the multisensory integration (as illustrated in Supplementary Fig. 19). As the dog became closer, the individual information (visual and auditory) shows higher impulses, while VA responses became proportionately weaker (Fig. 5e). In addition, the change between single sensory and multi-sensory is defined as $VA/(V_{max}, A_{max}) \times 100\%$ (%change). Clearly, the %change is enhanced with the dog being closer, as illustrated in Fig. 5f. The neural computation involved in their integration is changed from super-additive to additive and then to sub-additive, and the detailed illustrations are provided in Supplementary Note 2. Moreover, as multisensory integration could enhance the single sensory signal and the behaviors that depend on them, we further constructed an artificial stimulus-response system to further prove the multisensory enhancement concept. The robot hand system supported by our VTT is also illustrated in Supplementary Note 3.

Self-powered vertical tribo-transistor with MSMC function for multi-mode emotion recognition

All of the above characterizations and analyses imply that our VTT exhibits excellent MSMC function in a single device. Thus, based on the MSMC function, we construct a multi-model emotion recognition system to further demonstrate the capability of VTT in extending artificial intelligent boundary. Noticeably, the emotion is tightly entangled with behavior of human, while the complete emotion information cannot be obtained only through visual or auditory perception. Thus, combining the characteristic information and extracted feature information from the two perceptions would improve the emotion recognition accuracy and reliability. As depicts in Fig. 6, through single VTT device, we achieve the multi-model emotion recognition. Noticeably, the effective features are extracted from

visual and auditory mode by our devices (the EPSC with different light intensity is been separated into 16 states, and each state corresponds to 16 numerical values of colorful images, as depicted in Supplementary Fig. 22), and then the information is transferred to input layer for further processing, as indicates in Fig. 6a and Supplementary Fig. 23. Here, the select fusion model is data-level fusion, which directly combines the most original data collect by our devices without special processing to construct a group of new data. Then the several basic discrete emotions, such as sadness, fear, disgust, surprise, excitement and anger can be recognized by operation methods to compute and process data from multiple data sources.

We first demonstrate the motion recognition accuracy with visual model, auditory model, and multi-model, respectively, as illustrates in Figs. 6b–f. Several images are selected from a video without sound information, sound wave, or visual information, and the recognition accuracy of six emotions is summarized at Figs. 6c, e, and 6f. Note that, the accuracy of multi-model emotion recognition is significantly higher than that of a single model, which further indicates that our multi-model emotion recognition system is suitable for recognizing the complex emotional category through a certain combination mode. Thus, the comparison of recognition accuracy between different models is further shown in Fig. 6g. Clearly, the visual + auditory model exhibits the highest recognition accuracy of 94.05% after 140 epochs, while the accuracy of other single models is lower than that of the multi-model system. Moreover, the accuracy is higher than the values reported in previous works as illustrates in Table 1. These results apparently demonstrate that MSMC VTT device can well retain the data information on each modal sensor to avoid the loss of information, maintain the integrity of information, and significantly enhance the certainty of emotion.

Finally, the superiority of our self-powered multi-sensing-memory-computing device is concluded. (i) simple structure and self-powered device: the receptors and synapses are integrated in a single device based on a simple configuration of VOFET without any redundant layers, which enhances packing density, simplifies fabrication procedures and reduces chip cost, and further decreases power consumption with the self-power function of TENG. (ii) high sensitivity: The amplification, low SS (91.8 mV/dec) and high current density (5 mA/cm²) of VTT ensure high sensing sensitivity (the tactile sensitivity improves 711 times over individual TENG). (iii) high efficiency and conversion speed: the multi-sensing-memory-computing function and the interaction of multifunction integration of our VTT increases the conversion speed, which is associated with the decreased physical separation between sensory receptors and processed nodes compared with the traditional CMOS architecture. (iv) high recognition accuracy: the multi-model emotion recognition is successfully achieved based on the multi-sensing-memory-computing ability of our VTT, which significantly improves the computing functionality, recognition variety and accuracy.

Discussion

In summary, we experimentally demonstrate a multi-sensing-memory-computing device that integrated a TENG and a transistor in a single device with the configuration of VOFET. Compared with a single TENG device, the sensing sensitivity is significantly enhanced due to the excellent SS and high current density of VTT. Meanwhile, in addition to tactile and auditory perception, the visual perception can be mimicked with the unique advantage and optical-sensitivity performance of MXene electrode. By calculating the cross-correlation functions, the memory-computing between multi-sensing is demonstrated, and the superior colliculus function and artificial stimulus-response system is also mimicked thanks to the multi-sensing integration property of our VTT. Finally, a multi-model emotion recognition system based on our VTT is achieved, which enables 94.05% recognition accuracy of emotion via data-level model fusion. This proof-of-concept work realized

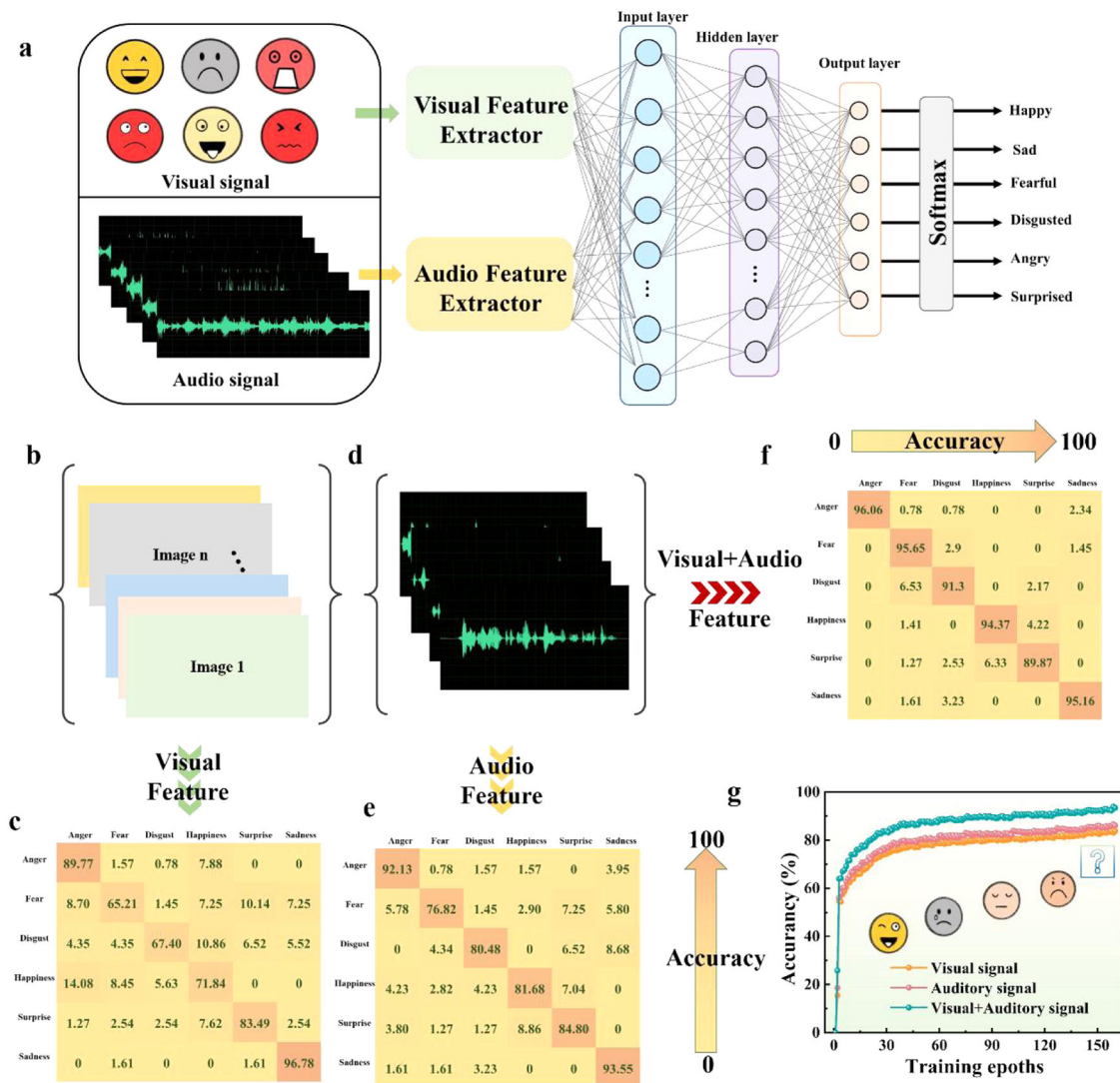


Fig. 6 | Multi-modal emotion recognition system. **a** The schematic diagram of multi-modal emotion recognition system to recognize the six emotions (anger, fear, disgust, happiness, surprise, and sadness), where the recognition process includes information fusion, feature extract, data classification, emotion recognition. **b** The pictures extract from database without sound information for emotion recognition. **c** The recognition accuracy of different emotion with single visual

model. **d** The sound extract from database without visual information for emotion recognition. **e** The recognition accuracy of different emotion with single auditory model. **f** The recognition accuracy of different emotion with multi-modal model. **g** The recognition accuracy of our multi-modal emotion recognition system at different model.

Table 1 | Accuracy of multi-modal recognition with different model

Multi-model	Model	Mode	Device number	Accuracy	Ref.
Visual+Auditory	Emotion	Software	/	90.8%	48
Visual +Auditory	Emotion	Software	/	77.5%	49
Visual + Auditory	Emotion	Software	/	85.9%	50
Visual+Auditory	Emotion	Software	/	82.9%	51
Visual+Auditory	Emotion	Software	/	80.3%	52
Visual+Tactile	Digit	Device	Two	90%	53
Visual+Tactile	Digit	Device	Two	86.8%	19
Visual+Tactile	Alphabet Letters	Device	/	92%	24
Visual+Auditory	Emotion	Device	Single	94.05%	This work

tactile/auditory/visual sensing-memory-computing within a single device, which has the potential to decrease the data movement between sensor, memory, and computing units, and paves the way for brain-inspired computing paradigms.

Methods

Device fabrication

The prepared ion-gel solution was spin-coated on Si substrate as the friction layer and insulation layer of TENG and VOFET, respectively.

The ion-gel solution was obtained by mixed PAN, $[\text{Li}^+\text{TFSI}^-]$, EC, and PC in a weight ratio of 7:1:1 and then the mixture was magnetically stirred at 500 rpm for 6 h. Then, the 2 mg/ml MXenes aqueous solution was selected as the top electrode and source network electrode of TENG and VOFET, respectively. Next, the prepared PDVT-10 solution was spin-coated on the electrode layer as the semiconductor layer, and then the patterned source and drain electrodes were thermally evaporated through a shadow mask. Finally, the fabricated device was transferred to the thermally evaporated gate electrode as VTT device.

Device characterization

The output performance of TENG was measured by an oscilloscope. The electrical performance of the synaptic transistor was characterized by Keithley 4200. The surface morphology of MXenes was examined by using scanning electron microscopy (SEM, Verios G4), transmission electron microscopy (TEM, Tecnai G2 F20). The UV-Vis absorption spectra of the films were measured by ultraviolet-visible near infrared spectrophotometer (Shimadzu UV-3600 Plus).

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

Code availability

The codes used for the simulations are available in [<https://github.com/YaqianLiuFZU/NCOMMS-22-33550>]⁴⁷.

References

- Wan, T. et al. Neuromorphic sensory computing. *Sci. China Inform. Sci.* **65**, 1–14 (2022).
- Zeng, X. & Hu, Y. Sensation and perception of a bioinspired flexible smart sensor system. *ACS Nano*. **15**, 9238–9243 (2021).
- Ielmini, D. & Wong, H.-S. P. In-memory computing with resistive switching devices. *Nat. Electron.* **1**, 333–343 (2018).
- Wang, T.-Y. et al. Reconfigurable optoelectronic memristor for in-sensor computing applications. *Nano Energy*. **89**, 106291 (2021).
- John, R. A. et al. Self healable neuromorphic memtransistor elements for decentralized sensory signal processing in robotics. *Nat. Commun.* **11**, 4030 (2020).
- Gao, J. et al. Intrinsic polarization coupling in 2D $\alpha\text{-In}_2\text{Se}_3$ toward artificial synapse with multimode operations. *SmartMat* **2**, 88–98 (2021).
- Tsodyks, M., Pawelzik, K. & Markram, H. Neural networks with dynamic synapses. *Neural Comput.* **10**, 821–835 (1998).
- Zeng, M., He, Y., Zhang, C. & Wan, Q. Neuromorphic devices for bionic sensing and perception. *Front. Neurosci.* **15**, 690950 (2021).
- Alvarado, J. C., Vaughan, J. W., Stanford, T. R. & Stein, B. E. Multisensory versus unisensory integration: contrasting modes in the superior colliculus. *J. Neurophysiol.* **97**, 3193–3205 (2007).
- Kang, H., Agus, T. R. & Pressnitzer, D. Auditory memory for random time patterns. *J. Acoust. Soc. Am.* **142**, 2219 (2017).
- Stein, B. E. & Stanford, T. R. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**, 255–266 (2008).
- Meredith, M. A. & Stein, B. E. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.* **56**, 640–662 (1986).
- Bu, X. et al. Ion-Gated Transistor: An Enabler for Sensing and Computing Integration. *Adv. Intell. Syst.* **2**, 2000156 (2020).
- Liu, Y. et al. Stretchable motion memory devices based on mechanical hybrid materials. *Adv. Mater.* **29**, 1701780 (2017).
- Wang, T. Y. et al. Ultralow power wearable heterosynapse with photoelectric synergistic modulation. *Adv. Sci.* **7**, 1903480 (2020).
- Shim, H. et al. Stretchable elastic synaptic transistors and neurologically integrated engineering systems. *Sci. Adv.* **5**, eaax4691 (2019).
- Kim, J. S. et al. Sensing and memorising liquids with polarity-interactive ferroelectric sound. *Nat. Commun.* **10**, 3575 (2019).
- Zhu, Y. et al. Light-Emitting Memristors for Optoelectronic Artificial Efferent Nerve. *Nano Lett.* **21**, 6087–6094 (2021).
- Wu, X. et al. Artificial multisensory integration nervous system with haptic and iconic perception behaviors. *Nano Energy*. **85**, 106000 (2021).
- Zhang, Z. H. et al. All-in-one two-dimensional retinomorphic hardware device for motion detection and recognition. *Nat. Nanotechnol.* **17**, 27–32 (2021).
- Cho, S. W., Kwon, S. M., Kim, Y.-H. & Park, S. K. Recent progress in transistor-based optoelectronic synapses: from neuromorphic computing to artificial sensory system. *Adv. Intell. Syst.* **3**, 2000162 (2021).
- Hou, Y. X. et al. Large-scale and flexible optical synapses for neuromorphic computing and integrated visible information sensing memory processing. *ACS Nano*. **15**, 1497–1508 (2021).
- Ji, X., Zhao, X., Tan, M. C. & Zhao, R. Artificial perception built on memristive system: visual, auditory, and tactile sensations. *Adv. Intell. Syst.* **2**, 1900118 (2020).
- Tan, H., Zhou, Y., Tao, Q., Rosen, J. & van Dijken, S. Bioinspired multisensory neural network with crossmodal integration and recognition. *Nat. Commun.* **12**, 1120 (2021).
- Wan, C. et al. An artificial sensory neuron with visual-haptic fusion. *Nat. Commun.* **11**, 4602 (2020).
- Wang, M. et al. Gesture recognition using a bioinspired learning architecture that integrates visual data with somatosensory data from stretchable sensors. *Nat. Electron.* **3**, 563–570 (2020).
- Gao, Z. et al. An artificial olfactory system with sensing, memory and self-protection capabilities-Elsevier Enhanced Reader. *Nano Energy*. **86**, 106078 (2021).
- Wan, C. et al. An artificial sensory neuron with visual-haptic fusion. *Nat. Commun.* **11**, 1–9 (2020).
- Cai, Y.-W. et al. A flexible ultra-sensitive triboelectric tactile sensor of wrinkled PDMS/MXene composite films for E-skin. *Nano Energy*. **81**, 105663 (2021).
- Chen, J. et al. Micro-cable structured textile for simultaneously harvesting solar and mechanical energy. *Nature Energy*. **1**, 16138 (2016).
- Cao, W. T. et al. A Stretchable Highoutput Triboelectric Nanogenerator Improved by MXene Liquid Electrode with High Electro-negativity. *Adv. Funct. Mater.* **30**, 2004181 (2020).
- Liu, Y. et al. A one-structure-layer PDMS/Mxenes based stretchable triboelectric nanogenerator for simultaneously harvesting mechanical and light energy. *Nano Energy*. **86**, 106118 (2021).
- Dudem, B., Ko, Y. H., Leem, J. W., Lim, J. H. & Yu, J. S. Hybrid energy cell with hierarchical nano/micro-architected polymer film to harvest mechanical, solar, and wind energies individually/simultaneously. *ACS Appl. Mater. Interfaces*. **8**, 30165–30175 (2016).
- Cho, Y. et al. Sustainable hybrid energy harvester based on air stable quantum dot solar cells and triboelectric nanogenerator. *J. Mater. Chem. A*. **6**, 12440–12446 (2018).
- Jiang, C. et al. All-electrospun flexible triboelectric nanogenerator based on metallic MXene nanosheets. *Nano Energy*. **59**, 268–276 (2019).
- Yang, Y. et al. Stretchable energy-harvesting tactile interactive interface with liquid-metal-nanoparticle-based electrodes. *Adv. Funct. Mater.* **30**, 1909652 (2020).
- Ye, B. U. et al. Electrospun ion gel nanofibers for flexible triboelectric nanogenerator: electrochemical effect on output power. *Nanoscale* **7**, 16189–16194 (2015).

38. Yu, J. et al. Contact-electrification-activated artificial afferents at femtojoule energy. *Nat. Commun.* **12**, 1581 (2021).
 39. Zhang, D. et al. Diversiform sensors and sensing systems driven by triboelectric and piezoelectric nanogenerators. *Coord. Chem. Rev.* **427**, 213597 (2021).
 40. Shao, Y. et al. Nanogenerator-based self-powered sensors for data collection. *Beilstein J. Nanotech.* **12**, 680–693 (2021).
 41. Guo, W. et al. Self-powered electrical stimulation for enhancing neural differentiation of mesenchymal stem cells on graphene-Poly(3,4-ethylenedioxythiophene) hybrid microfibers. *ACS Nano.* **10**, 5086–5095 (2016).
 42. Zhang, H. et al. Ion gel capacitively coupled tribotronic gating for multiparameter distance sensing. *ACS Nano.* **14**, 3461–3468 (2020).
 43. Kleemann, H., Krechan, K., Fischer, A. & Leo, K. A review of vertical organic transistors. *Adv. Funct. Mater.* **30**, 1907113 (2020).
 44. Hu, D. et al. High performance flexible nonvolatile memory based on vertical organic thin film transistor. *Adv. Funct. Mater.* **27**, 1703541 (2017).
 45. Li, E. et al. MXene based saturation organic vertical photoelectric transistors with low subthreshold swing. *Nat. Commun.* **13**, 2898 (2022).
 46. Zhang, X. et al. Programmable neuronal-synaptic transistors based on 2D MXene for a high-efficiency neuromorphic hardware network. *Matter* **5**, 1–18 (2022).
 47. Liu Y., et al. Self-powered high-sensitivity all-in-one vertical tribo-transistor device for multi-sensing-memory-computing. *GitHub*, <https://doi.org/10.5281/zenodo.7418909> (2022).
 48. Nguyen, D. et al. Deep spatio-temporal feature fusion with compact bilinear pooling for multimodal emotion recognition. *Comput. Vis. Image Und.* **174**, 33–42 (2018).
 49. Dobrisek, S. et al. Towards efficient multi-modal emotion recognition. *Int. J. Adv. Robot. Syst.* **10**, 53 (2013).
 50. Zhang, S. et al. Learning affective features with a hybrid deep model for audio–visual emotion recognition. *IEEE T. Circ. Syst. Vid.* **28**, 3030–3043 (2017).
 51. Kansizoglou I., Bampis L., Gasteratos A. An active learning paradigm for online audio-visual emotion recognition. *IEEE Trans. Affect. Comput.* **13**, 756–768 (2019).
 52. Raju J., Gaus Y. F. A., Breckon T. P. Continuous Multi-modal Emotion Prediction in Video based on Recurrent Neural Network Variants with Attention. 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 688–693 (2021).
 53. Yu, J. et al. Bioinspired mechano-phonic artificial synapse based on graphene/MoS₂ heterostructure. *Sci. Adv.* **7**, eabd9117 (2021).
- for Distinguished Young Scholars of Fujian Province (2020J06012), Fujian Science & Technology Innovation Laboratory for Optoelectronic Information of China (2021ZZ129).

Author contributions

H.P.C. T.L.G. and Y.Y.H. conceived the project, Y.Q.L. designed and performed the experiments and collected the data. Y.Q.L., D.L., C.S.G., X.H.Z., R.J.Y., X.M.W. and E.L.L. analyzed and discussed the data. H.P.C. supervised the project. Y.Q.L. and H. P. C. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-35628-0>.

Correspondence and requests for materials should be addressed to Huipeng Chen.

Peer review information *Nature Communications* thanks Sherajul Islam, and the other, anonymous, reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

Acknowledgements

The authors are grateful for financial support from National Natural Science Foundation of China (U21A20497) Natural Science Foundation