

# Human–Machine Interaction via Dual Modes of Voice and Gesture Enabled by Triboelectric Nanogenerator and Machine Learning

Hao Luo, Jingyi Du, Peng Yang, Yuxiang Shi, Zhaoqi Liu, Dehong Yang, Li Zheng,\* Xiangyu Chen,\* and Zhong Lin Wang\*



Cite This: <https://doi.org/10.1021/acsami.3c00566>



Read Online

ACCESS |

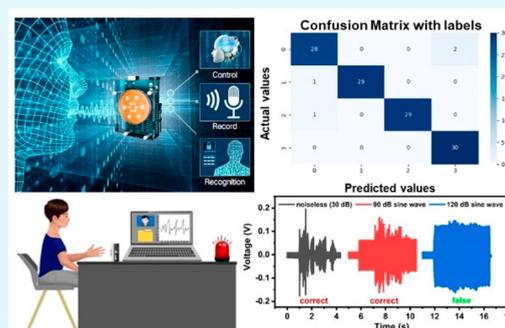
Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** With the development of science and technology, human–machine interaction has brought great benefits to the society. Here, we design a voice and gesture signal translator (VGST), which can translate natural actions into electrical signals and realize efficient communication in human–machine interface. By spraying silk protein on the copper of the device, the VGST can achieve improved output and a wide frequency response of 20–2000 Hz with a high sensitivity of 167 mV/dB, and the resolution of frequency detection can reach 0.1 Hz. By designing its internal structure, its resonant frequency and output voltage can be adjusted. The VGST can be used as a high-fidelity platform to effectively recover recorded music and can also be combined with machine learning algorithms to realize the function of speech recognition with a high accuracy rate of 97%. It also has good antinoise performance to recognize speech correctly even in noisy environments. Meanwhile, in gesture recognition, the triboelectric translator is able to recognize simple hand gestures and to judge the distance between hand and the VGST based on the principle of electrostatic induction. This work demonstrates that triboelectric nanogenerator (TENG) technology can have great application prospects and significant advantages in human–machine interaction and high-fidelity platforms.

**KEYWORDS:** triboelectric translator, silk protein, voice and gesture recognition, machine learning, human–machine interface



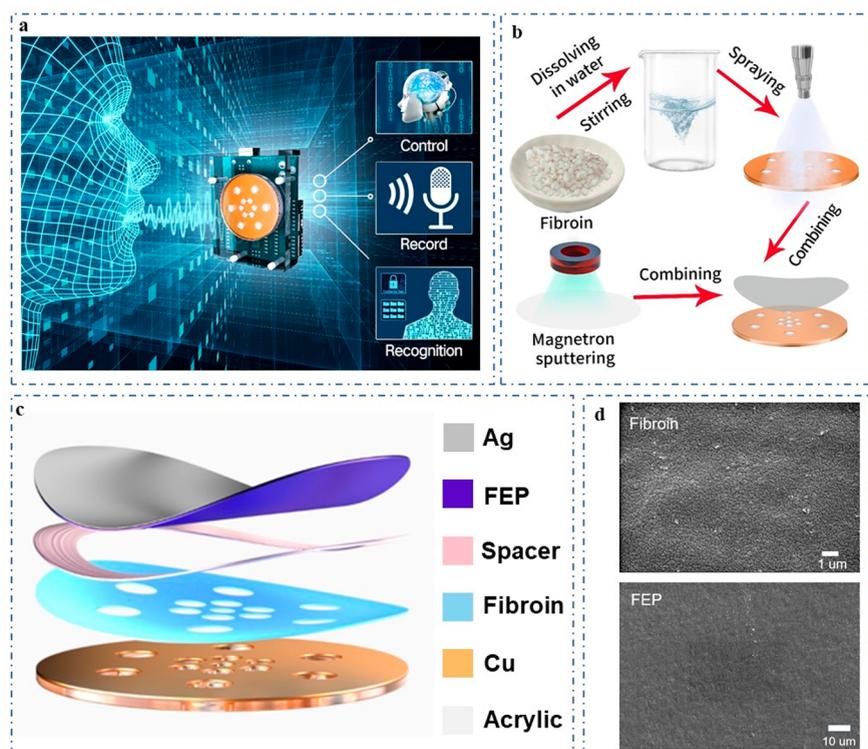
## 1. INTRODUCTION

Nowadays, the rapid development of artificial intelligence and the Internet of Things (IoT) has brought great changes to people's lives, and human–machine interaction sets application requirements and research ideas for artificial intelligence and the Internet of Things (IoT).<sup>1–7</sup> The technology of how to improve the efficiency of communication between human and machine is the most difficult problem for the researchers to overcome, in which the communication system<sup>8–15</sup> connecting human and machine is particularly important. Nowadays, there are two main methods of human–machine interaction: voice–machine interaction<sup>16–19</sup> and gesture–machine interaction.<sup>20–23</sup> Voice signals contain rich character information on humans and are easy to use, so people can use voice signals to make and realize accurate commands. At present, the main sensors currently used for speech recognition are capacitive sensors,<sup>24,25</sup> piezoresistive sensors,<sup>26,27</sup> piezoelectric sensors,<sup>28–30</sup> and triboelectric sensors.<sup>31–33</sup> However, some components in capacitive and piezoresistive sensors, such as conventional commercial microphones, rely on battery charging and are vulnerable to environmental noise. On the other hand, piezoelectric sensors have complex structures, troublesome manufacturing process and short working frequency ranges. Fortunately, triboelectric nanogenerators (TENGs) can effectively solve these problems. In 2012, Wang

et al. proposed the concept of triboelectric nanogenerator which is based on the coupling of contact electrification and electrostatic induction to collect various mechanical energy from the environment and convert it into electrical energy.<sup>34–41</sup> Triboelectric nanogenerators have the advantage of low production cost, simple structure, high sensitivity and no external power supply. TENGs are quite stable in speech recognition and have a good broadband effect. It can respond sensitively to human voice in the frequency range of 20–2000 Hz. Some TENG-based acoustic sensors have been reported to use electrospinning to prepare porous elastic nanofiber membranes<sup>42</sup> to collect sound signals using contact-separation mode. Although the membrane is thin and flexible and has good air permeability, the spinning process is cumbersome and many preparation materials are toxic. The final prepared membrane is not stable for long time use and a lot of information in frequency domain is missing. In 2018, Guo et al. invented a porous gold film TENG which has a broadband

**Received:** January 12, 2023

**Accepted:** March 12, 2023



**Figure 1.** Application of the VGST in human–machine interaction and its structure. (a) Applications of the VGST in human–machine interaction. (b) Fabrication process of the VGST device consisting of FEP, Ag, fibroin, and Cu. (c) 3D decomposition structure of the VGST. (d) Scanning electron microscope image of silk protein and FEP surface.

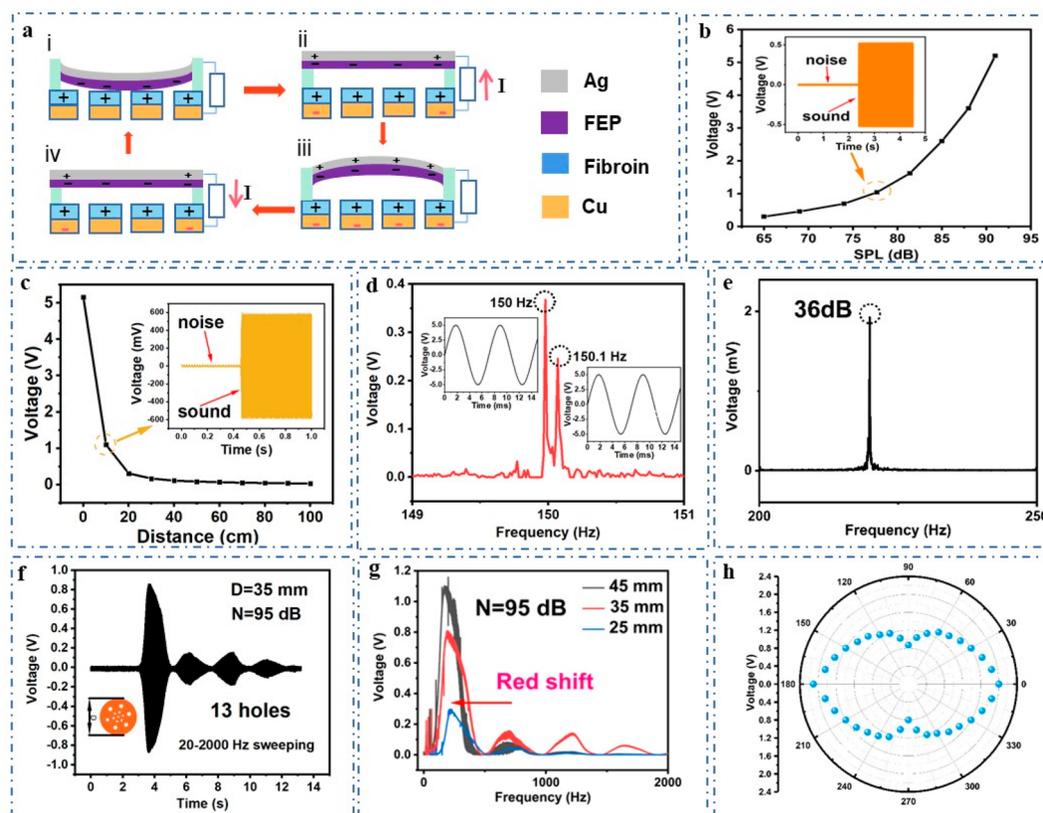
response,<sup>43</sup> but it has low output, not high enough sensitivity, and the gold film is expensive. Therefore, it is worth investigating to prepare an acoustic TENG with good performance, simple to fabricate and eco-friendly.<sup>44</sup>

For gesture recognition, now the general recognition methods are mainly based on visual image analysis, but they are easy to be affected by light and occlusion. TENG can use electrostatic signal detection technology to sense the change of electrostatic field around the object so as to detect and recognize the target.<sup>45,46</sup> It has the advantages of a small working blind area, simple system deployment, noncontact, and so on. However, a hybrid sensor for recognizing both voice and gesture sensor has never been realized by TENG technique.

In this paper, a voice and gesture signal translator (VGST) based on TENGs is designed and demonstrated for the first time. By spraying silk protein on the copper of the device, the VGST can be highly sensitive and has a wide range of frequency response (human ear hearing range 20–2000 Hz) with a higher sensitivity of 167 mV/dB. In comparison with previously reported acoustic sensors,<sup>43</sup> the sensitivity of our device is improved by 51.8%, due to the superior performance of the silk protein. It also has strong anti-interference ability and high accuracy in recovering music. In addition, the VGST can be used for voice and gesture recognition through machine learning algorithms,<sup>47</sup> and the recognition accuracy can reach 97%. In gesture recognition, four gestures can be clearly distinguished, and the distance from hand to device can be judged easily. This work demonstrates powerful functions of TENGs used in human–machine interfaces.

## 2. RESULTS AND DISCUSSIONS

TENG plays a very important role in many aspects of today's big data era. Figure 1a demonstrates the ability of a TENG-based translator to efficiently receive acoustic signals and accurately respond. It has a wide range of application prospects in machine control, sound recognition, and recovery. The fabricated VGST could complete information management and enhance information security for people to the greatest possible extent, and could also help people with hearing impairments to better recover their hearing. Figure 1b illustrates the process of VGST fabrication. The specific process is described in the Experimental Section. The front and side view of the entire device is shown in Figure S1, with dimensions of 52.5 mm × 67.5 mm × 20 mm. The detailed three-dimensional decomposition structure of the VGST is shown in Figure 1c. The entire membrane part of VGST consists of 80 nm of silver, 12.5 μm of FEP, 0.3 mm of spacer, 60 μm of copper and fibroin, and 2 mm of acrylic, showing that the device is very thin. Holes in the surface of the copper and fibroin are used to reduce air damping and to allow sound waves to better drive vibration of the membrane. To obtain the best sound detection, the size of the holes on the substrate is measured. If the holes are too small, then it is difficult to transmit the sound, thus the sound will be small and fine. If the holes are too large, then the high frequency sound will be sharp and harsh. Through lots of experiments, we found that it can obtain the best sound quality with radius of the outer holes of 3.2 mm and radius of the inner holes of 2 mm. Thus, weak sound of high frequency can also be accurately detected. We polarized the FEP film with a polarizer and injected charges to increase the surface charge density and then improve the electrification performance of FEP film. The picture of



**Figure 2.** Working mechanism and acoustic performance and characterization of the VGST. (a) Schematic diagram of the working principle of the VGST under sound driven. (b) Effect of sound pressure level on the output voltage of the VGST and (c) effect of distances on the output voltage of the VGST. Inset shows the comparison of output voltage of the VGST caused by sound and environmental noise, respectively. (d) Corresponding FFT spectrum under sine wave excitation of 150 and 150.1 Hz. Inset shows the imposed sine wave excitation on the VGST. (e) Frequency spectrum of the VGST derived through Fourier transform when the sound intensity is 36 dB. (f) Voltage signal measured from the VGST under impact of sound source with sweeping frequency range of 20 to 2000 Hz and sound intensity of 95 dB. (g) The frequency spectra of the VGST derived through Fourier transform when diameter of the device varies. (h) Shape-dependent directional patterns of the VGST under sound intensity of 100 dB ( $D = 35$  mm, where  $D$  is the diameter of the membrane).

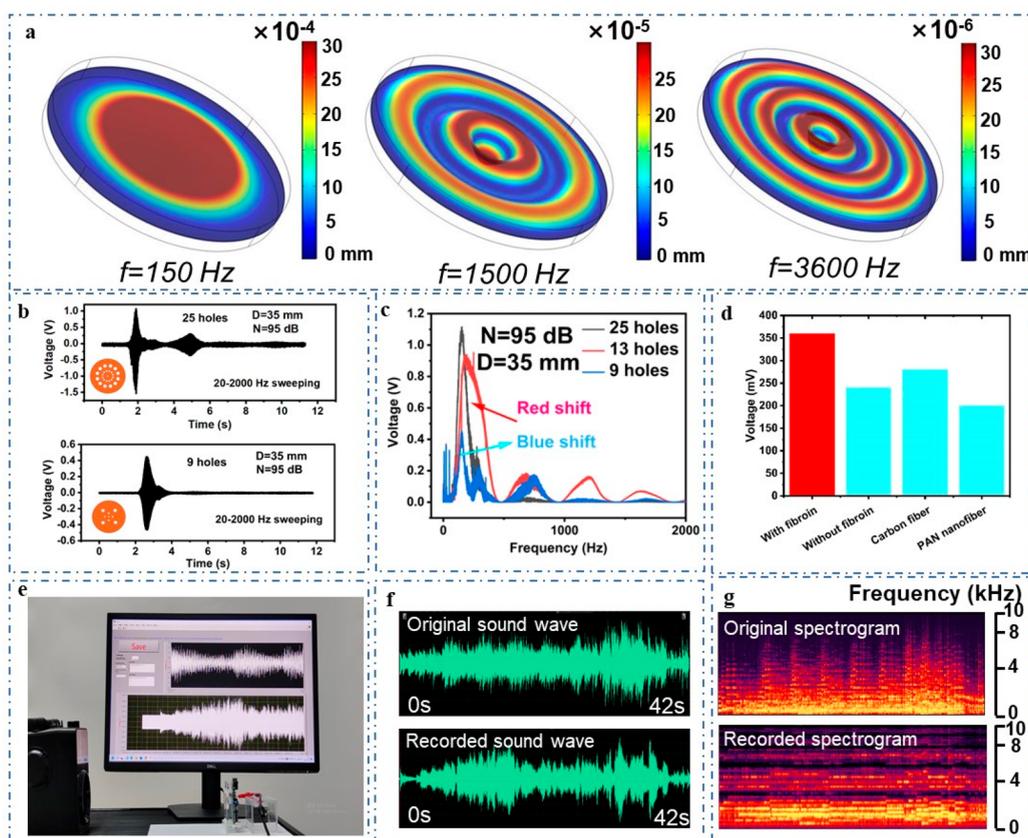
perforated copper sheet coated with silk protein is shown in Figure S2. The microstructure of FEP and fibroin is shown in Figure 1d. As can be seen from the SEM images, the surface of the FEP and fibroin is very smooth. As natural material with excellent physical and chemical properties, we sprayed silk protein on Cu film so as to fabricate sensitive VGST devices. Meanwhile, fibroin has high mechanical strength, lightweight and easy processing, and its strong ability to lose electrons can improve the output performance of the VGST. In addition, fibroin is nontoxic, harmless, and biocompatible, making it a good choice for triboelectric materials.

The working principle of the VGST can be attributed to the coupling effect of contact electrification and electrostatic induction, as shown in Figure 2a and Note S1, where the inset figures are cross sections of TENG components. To precisely control the amplitude and frequency, Adobe Audition is chosen to provide a stable sine wave. As shown in Figure 2b, under the excitation of sine-wave sound of 150 Hz and keeping the sound source 2 cm away from the VGST, the output voltage of the VGST increases with increasing sound pressure level, and it can be calculated that a higher sensitivity of 167 mV/dB can be achieved. The results show that under higher sound intensity, the vibration between membranes is more intense, and the output voltage is higher, which can be expressed as the following equations:

$$L_p = 20 \lg \left( \frac{P}{P_0} \right) \text{ (dB)} \quad (1)$$

$$V = S \cdot P = S \cdot P_0 \cdot 10^{L_p/20} \quad (2)$$

where  $V$  is the peak output voltage,  $S$  is the sensitivity,  $P$  is the sound pressure,  $P_0$  is the reference sound pressure as  $2 \times 10^{-5}$  Pa, and  $L_p$  is the sound pressure. As can be seen from the inset figure in Figure 2c, at 77 dB, the electrical output signal is well distinguished from system noise. As shown in Figure 2c, keeping sound intensity at 90 dB, the output voltage drops from 5.2 V to nearly 0 V when the distance between the sound source and the VGST increases from 0 to 100 cm, indicating that the sound wave attenuates during propagation as the distance increases. When a 150 Hz sine-wave and a 150.1 Hz sine-wave are separately imposed on the VGST (Figure 2d), the peaks of the two frequencies can be clearly identified in the frequency domain through fast Fourier transform, showing that the device has high accuracy to differentiate signals with a small frequency difference. When sine-wave sound source has a frequency of 150 Hz, a sound intensity of 36 dB is applied on the VGST. The output voltage is just larger than that of clutter signal, and its frequency peak can still be clearly seen through Fourier transform (Figure 2e), where 36 dB is almost close to the sound pressure level of quiet indoor environment, reflecting the good sensitivity of the fabricated



**Figure 3.** Other characteristics of the VGST and its application in recovering music. (a) Vibration patterns of FEP film under different frequencies (simulated using COMSOL under sound pressure of 1 Pa). (b) Output voltage measured by VGSTs with different number of holes (sweeping frequency range of 20–2000 Hz; sound intensity of 95 dB). (c) Frequency spectra of the VGST derived through Fourier transform when the number of holes of the VGST varies. (d) Comparison of output voltage of VGSTs with different electrification materials. (e) Application in recording the famous classical music “Blue Danube”. (f) Original music wave and recorded sound wave information and (g) their corresponding spectrograms.

VGST. Figure 2f shows the output voltage of the VGST (35 mm diameter, 95 dB sound intensity) under the impact of sound source with sweeping frequency range of 20–2000 Hz. It can be calculated from the setting time of sweeping frequency that when the sound frequency is 150 Hz, the voltage is maximum, which may be the resonant frequency of the VGST. Figure S3 shows that a 150 Hz sine-wave source is provided to VGST as an acoustic excitation. Since 150 Hz is the resonant frequency of the fabricated VGST device, maximum voltage output can be generated under the excitation of 150 Hz sound source. To further investigate the factors influencing the resonant frequency, we made VGSTs with different diameters of the membrane. Diameter is a key parameter affecting frequency response. Figure 2g shows that the VGST with a larger membrane diameter respond well to lower resonant frequency (red shift of frequency) with higher electric output. The results show that membranes with larger diameters can produce greater deformation and thus higher voltage output in response to the vibration of the acoustic wave, while the reduction in resonant frequency can be explained by the following equation:

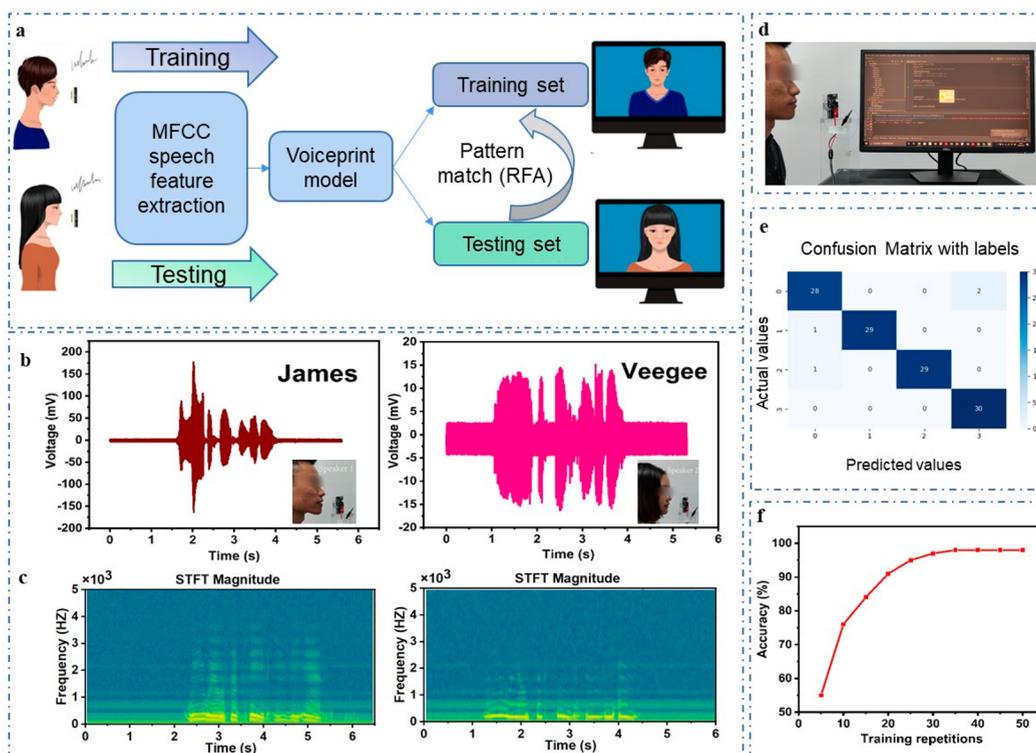
$$f_0 = \frac{c}{2\pi} \sqrt{\frac{S}{(l + 0.8d)V}} = \frac{c}{2\pi} \sqrt{\frac{\pi d^2}{4(l + 0.8d)V}} \quad (3)$$

where  $f_0$  is the resonant frequency of Helmholtz resonator,  $c$  is the speed of sound,  $S$  is the cross-sectional area of the neck or

opening,  $d$  is the diameter of the neck or opening,  $l$  is the length of the neck, and  $V$  is the volume of the vessel.

We also tested the voltage signal of the VGST (35 mm diameter) at different decibels under impact of sound source in frequency range of 20–2000 Hz (Figure S4a–c) and found that their resonance frequencies were almost the same (Figure S4d), which is a good validation of the above equation. The resonance frequency is independent with the sound intensity. To evaluate the directional sound response characteristics, we detected the output voltage signal of the VGST in every direction. The experimental results show a butterfly-shaped directional pattern with mirror symmetry, as can be observed from Figure 2h. The response signal keeps a stable output in every direction, and reaches the maximum value at 0 and 180 deg and the minimum value at 90 and 270°, where 0° means the direction where sound source is perpendicular to the front view of the VGST. It can be seen that VGST responds in every direction, providing a broad-angle range for sound detection.

In order to investigate the response of the VGST to sound source with different frequencies, finite element analysis method is adopted to simulate the deformation displacement of the circular film under three different acoustic frequencies. The modules and parameters of the finite element method are shown in Note S2. Usually, the first-order vibration mode provides the largest mechanical deformation of membrane, i.e., the largest deformation is under frequency of 150 Hz, while deformations under other frequencies are small, as shown in



**Figure 4.** Investigation on VGST voice recognition system. (a) Voice recognition process for TENG-based speaker identification. (b) Output voltage of two people saying “We will succeed in the end.” (c) Corresponding STFT spectrograms of the acquired sound waveforms for the two people. (d) Demonstration of the TENG-based voice recognition system. (e) Confusion matrix diagram for machine learning outcome. (f) Overall accuracy of voice recognition at different training repetitions using random forest algorithm.

**Figure 3a.** Figure S5 shows the deformation displacements of the membrane at other frequencies, and Figure S6 shows deformation displacements under different sound intensities. Therefore, we mainly focused on modulating these modes of the VGST to broaden the sound frequency range with good response performance.

The vibration mode of the membrane conforms to Helmholtz and time harmonic equations:

$$\nabla^2 U_m + k^2 U_m = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial U_m}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 U_m}{\partial \theta^2} + \frac{\omega^2}{C^2} U_m = 0$$

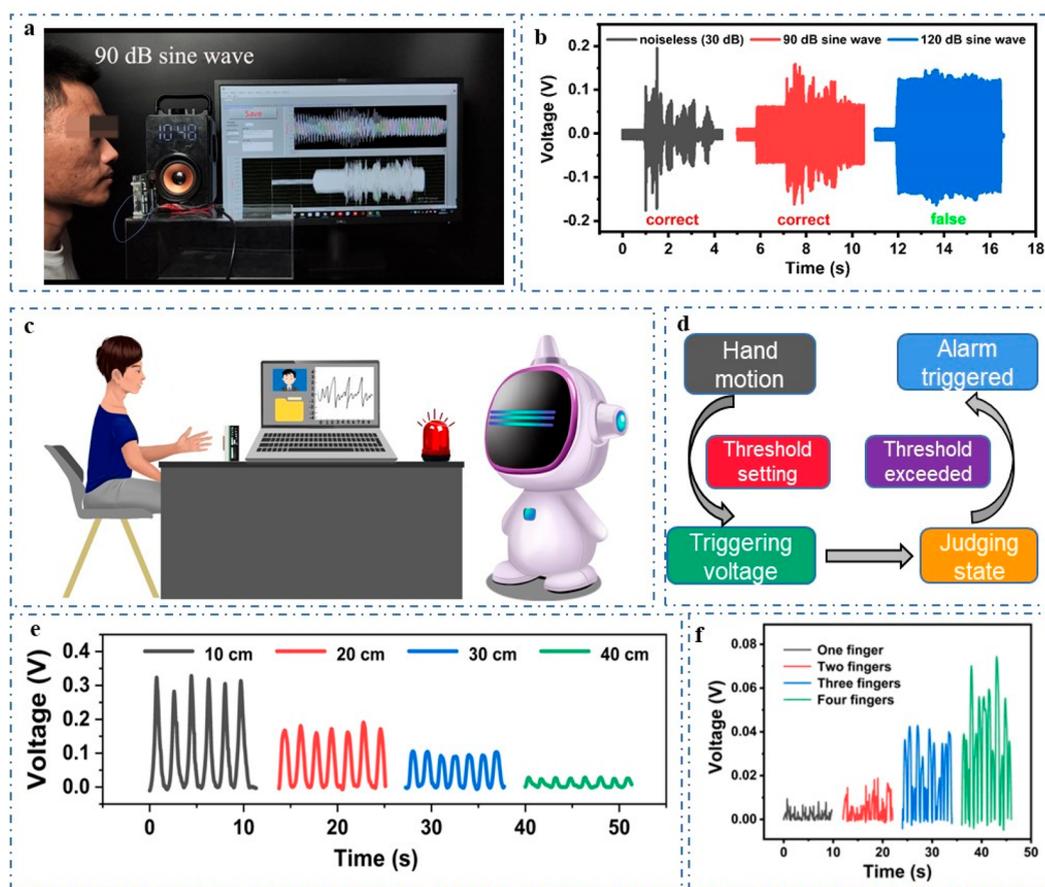
$$< \frac{D}{2} \quad (4)$$

$$U_m = 0 \rho = \frac{D}{2} \quad (5)$$

where  $U_m$  is the envelope of the film,  $\rho$  is the polar radius,  $k$  represents the wavenumber (equal to  $C = \sqrt{\frac{T}{\sigma}}$ ),  $\omega$  is the angular frequency,  $D$  is the diameter of the film,  $\frac{\omega}{c}$  is the wave velocity, and  $T$  and  $\sigma$  represent the stretching force and surface mass density of the film, respectively.

In order to investigate what kind of membrane has the best acoustic response performance, we tested the performance of the VGSTs with different numbers of holes in the membrane (Figure 3b). Through frequency sweeping and Fourier transform, it can be found that when the membrane has 13 holes, the resonant frequency is the largest, and the resonant frequencies with membranes of 25 and 9 holes are almost the same (Figure 3c), while the output voltage of the VGST increases as the number of holes increases. The results show

that the increase of hole numbers reduces the damping of the air to a greater extent, allowing the acoustic waves to drive membrane vibration better. Thus, the membrane displacement is increased, and the output voltage becomes larger. The resonant frequency is not determined by a certain factor, but it is determined by multiple factors. For example, the contact area between FEP and fibroin will decrease when the number of holes on the membrane is increased. Figure 3d depicts the performance comparison of the VGSTs with different triboelectric materials. It is easy to see that the output voltage is greater when copper is sprayed with silk protein than that without spraying. When the copper sheet is replaced by a carbon fiber mesh and PAN nanofiber films, the output voltage is smaller. By comparison, the effect of spraying silk protein on copper sheet is a better choice. This phenomenon is also due to the fact that silk protein is a very positive material and it generates a large amount of electron transfer upon contact with negative FEP film, resulting in a large output voltage. In addition, we have tested the stability of the VGST (Figure S7). A 90 dB, 220 Hz sine wave as an excitation source is supplied to the VGST every day, and it can be seen from Figure S7, the output of the VGST remained stable even after 30 days. To demonstrate the high-fidelity of VGST, we used VGST to listen to a short section of the famous international classical music “Blue Danube” (Figure 3e) to generate an electrical signal. Then we used the software Adobe Audition 2020 to compare and analyze the original and recovered music. As shown in Figure 3f,g, the sound waves of the recovered music are a little different from those of the original music because we used wavelet denoising by MATLAB to remove the burr signal from the waveform (Figure S8) and converted the signal into a waveform audio file through MATLAB. Calculated by



**Figure 5.** The VGST's antinoise performance and applications in gesture recognition. (a) Demonstration of the VGST's antinoise performance. (b) Speech waveforms at different noise intensity. (c) Application scenarios of gesture recognition. (d) Working flow of gesture recognition. (e) Output voltage corresponding to four different distances between hand and the device. (f) Output voltage corresponding to four different gestures.

MATLAB, the mean square error (MSE) between the original and processed music is  $1.0843 \times 10^{-4}$ , reflecting very high fidelity of the fabricated VGST. [Video S1](#) demonstrates the process of listening and recovering music by VGST, showing its excellent sound recovery performance.

Information security construction is an indispensable part of e-government construction and national security system. Password-based authentication is one of the most commonly used ways to consolidate information security in today's society,<sup>48</sup> including voice password. As an important kind of biometric feature, the human voice usually contains rich information, and the collection of human voice is very convenient, so the analysis of human voice signals has significant application values. [Figure 4a](#) depicts the overall process for voice recognition of VGST. First, two men and two women were selected as volunteers, each of whom speaks 100 sentences. We extract the eigenvalues of the sentences by calculating the Mel Frequency Cepstrum Coefficient (MFCC) to build a database. The specific process for obtaining MFCC features is described in [Note S3](#). Then 70 sentences out of 100 sentences are selected as the training set, and 30 sentences are selected as the testing set. As a new machine learning model, a random forest algorithm is used to establish a voice pattern model, which has high accuracy, good noise immunity, fast training speed, small calculation volume, and the ability to run effectively on large data sets. When the testing set is matched with the training set, it can judge which speaker each person's speech signal comes from. The principle of the algorithm is

simple which consists of a large amount of decision trees. When a new sample comes in, each decision tree in the forest makes a separate determination to see which category the sample should belong to. The phrase "We will succeed in the end" was chosen as the test statement in our experiment. Each person used this phrase to speak to the TENG, and the computer interface was able to display the speaker's name immediately. [Figure 4b](#) shows the output voltage waveforms of the voice signals of James and Veegee, and the corresponding information on their voice signals in frequency domain is shown in [Figure 4c](#). The output voltage waveform and frequency domain information on the other two volunteers are reflected in the [Figure S9](#). It has been observed that the frequency of a normal human voice is between 50 and 500 Hz, but each person's voice frequency corresponds to a different amplitude. By following the frequency information on each person, the VGST is able to distinguish their signals very well. In the voiceprint recognition system, the experimenter interacts with the VGST ([Figure 4d](#) and [Video S2](#)). The reliability of the VGST in voiceprint recognition was demonstrated in [Figure 4e](#), which is the confusion matrix plot, reflecting the recognition accuracy of 97%. [Figure 4f](#) displays that when training is repeated 5 times, the accuracy rate reaches 55%. As the number of training repetitions increases, the accuracy rate will increase and finally can reach 97%. We tested it using XGB algorithm compared with the random forest algorithm. [Figure S10](#) are the learning curves

calculated through both algorithms. It is not difficult to find that the accuracy of random forest algorithm is higher.

In order to test the antinoise performance of the VGST, we used an electronic speaker to create a 220 Hz sine wave as noise source. The noise source is 2 cm away from the device and it is vertical to the front side of the membrane. The electronic speaker released noise when the volunteer spoke to the VGST, as shown in Figure 5a and Video S3. In Figure 5b, human voice can be clearly distinguished in computer interface in a quiet environment and the human identity can be recognized accurately. In a noisy environment under 90 dB, the human voice can still be easily distinguished and the recognition is still correct. Human voice signals can be recognized correctly until the noise intensity continues to increase to 120 dB. The human voice waveform is almost drowned by the noise under sound intensity of 120 dB, where the recognition starts to be wrong, demonstrating the good antinoise performance of the VGST. Except for its application on voice recognition, the VGST can also be applied to gesture recognition. Human–machine interaction based on human hand movement is one of the most important methods and here we use the principle of noncontact electrostatic induction to judge the state of hand movement. The real-world demonstration of gesture recognition is shown in Figure 5c, and the circuit modules used are shown in Figure S11. The functions of the development board are described in Note S4. The whole process can be divided into four parts (Figure 5d). Once the threshold is set, the hand moves to generate an electrostatic induction signal and to trigger the alarm system. Then we can judge the hand state according to the output voltage generated from the device and displayed on the computer. If the voltage exceeds the set threshold, the alarm can be triggered to realize a warning effect. According to this working principle, the distance between the hand and the device can be judged according to the output voltage of the VGST (Figure 5e). In the initial state, the hand is kept directly in front of the VGST with a vertical distance of 50 cm, and there is no output signal. When the hand slowly moves to only 10 cm away from the VGST, the induced voltage is 0.32 V. With the decrease of the distance between hand and device, the induced output voltage becomes larger. A linear relationship between the output voltage and distance can be found, thus the distance between hand and the VGST can be judged by the output voltage. Next, we demonstrate four hand-gesture recognitions from 1 to 4. At the beginning, we keep the hand in the upper position before the VGST, and each gesture moves down slowly until the fingers completely pass by the VGST. According to Figure 5f, the more fingers in the gesture, the greater the generated output voltage signal. Therefore, the gesture can be judged according to the output voltage of the VGST. From this work, it can be seen that gesture recognition based on the principle of electrostatic induction has a good prospect, and it could be prospected that more complex gestures can be recognized by means of machine learning and three-dimensional spatial positioning.

### 3. CONCLUSIONS

In summary, we have designed and developed a thin triboelectric translator called a VGST that can translate the raw signals into electrical signals and realize both voice and gesture recognition with high sensitivity. By spraying silk protein on the copper of the device, it can detect voice in frequency range of 20–2000 Hz with resolution of 0.1 Hz and

sensitivity of 167 mV/dB. By changing the diameter and holes of the device, the frequency response of the VGST device can be changed and modified. This fabricated VGST can act as a high-fidelity listening platform, effectively recording and restoring sound signals. In addition, with the aid of machine learning, the VGST can recognize voice information with high accuracy of 97% under less than 120 dB noisy environment and has good antinoise performance. It can also be used for gesture recognition, judging the distance between human hand and the VGST device and recognizing different gestures. It can be seen that the demonstrated VGST with multifunctions can enrich the methods of human–machine interaction and has huge potential applications in many areas such as authentication, information security, high-fidelity platforms, machine control, and medical rehabilitation et al.

## 4. EXPERIMENTAL SECTION

**4.1. Fabrication of VGST.** First, an acrylic board with a thickness of 2 mm was cut into a circular substrate (diameter of 35 mm) by a laser cutting machine (LM-1390), and some holes were pierced on the circular substrate according to the preset sizes. The 60  $\mu\text{m}$  thick copper sheet was attached to the perforated acrylic substrate as the upper electrode, then a hobby knife was used to dig some holes as designed on the acrylic substrate to reduce air damping. 0.2 g of fibroin powder was dissolved in water to form a 15% solution (13.35 mL), then stirred for 5 min until uniform using a magnetic stirring apparatus. Later, the stirred solution was poured into a spray gun and sprayed evenly on the surface of the copper sheet until a white film is formed. The protein film was left for 20 min at room temperature of 20  $^{\circ}\text{C}$ , which can be used as an electrification layer. Then a layer of 80 nm silver was coated on the 12.5  $\mu\text{m}$  FEP film as the lower electrode using a magnetron sputtering machine. A polarizer was used for 10 min to provide a high voltage of 6 kV and some negative charges were injected on the surface of FEP. Three circular-ring paper sheets (each 0.1 mm thick) were used as spacers to separate the top and bottom membranes.

**4.2. Integration of VGST with Development Board.** An acrylic board was cut into a rectangle (6.75 cm long and 5.25 cm wide) and was cut with certain circular holes with a diameter of 3.3 cm in the center. Then glue was applied around the holes so that the VGST can be attached to the acrylic board. The four corners of the rectangular acrylic board were punched with holes, using nuts to integrate the VGST with development board.

**4.3. Characterization and Measurement.** SU8020 cold field scanning electron microscope was used to characterize the surface morphology of fibroin membrane as well as FEP membrane. Adobe Audition was used to provide stable sine wave with adjustable frequency. Norwest H3 Bluetooth speaker was used to drive the membrane vibration of the VGST. Voltage preamplifier (Stanford SR560) was used to measure the voltage signal of sound. Programmable electrostatic meter (Keithley 6514) was used to measure the gesture of voltage signal. Acquisition card NI USB-6356 was used for data acquisition. Software platform was built based on Labview for real-time data presentation. Analysis and voice recognition interface was built with Python. The *sklearn* library was used for machine learning training and testing.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

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

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsami.3c00566>.

Views of the VGST; output voltage and frequency spectra; COMSOL deformation of films; stability test of

the VGST; spectrograms of music and wavelet denoising by MATLAB; output voltage waveforms and the spectra of two other volunteers; illustration of the gesture recognition function of the VGST; electronic module; working mechanism of the VGST; simulation parameters using finite element method; specific process for obtaining MFCC features; functions of development board (PDF)

Video S1: Recording the famous classical music (MP4)

Video S2: Voice recognition process in demonstration of TENG-based speaker identification (MP4)

Video S3: Test of VGST antinoise performance (MP4)

## AUTHOR INFORMATION

### Corresponding Authors

**Li Zheng** – College of Mathematics and Physics, Shanghai Key Laboratory of Materials Protection and Advanced Materials in Electric Power, Shanghai University of Electric Power, Shanghai 200090, China; Email: zhengli@shiep.edu.cn

**Xiangyu Chen** – Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China; School of Nanoscience and Technology, University of Chinese Academy of Sciences, Beijing 100049, PR China; [orcid.org/0000-0002-0711-0275](https://orcid.org/0000-0002-0711-0275); Email: chenxiangyu@binn.cas.cn

**Zhong Lin Wang** – Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China; School of Nanoscience and Technology, University of Chinese Academy of Sciences, Beijing 100049, PR China; [orcid.org/0000-0002-5530-0380](https://orcid.org/0000-0002-5530-0380); Email: zhong.wang@mse.gatech.edu

### Authors

**Hao Luo** – College of Mathematics and Physics, Shanghai Key Laboratory of Materials Protection and Advanced Materials in Electric Power, Shanghai University of Electric Power, Shanghai 200090, China; Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China

**Jingyi Du** – College of Mathematics and Physics, Shanghai Key Laboratory of Materials Protection and Advanced Materials in Electric Power, Shanghai University of Electric Power, Shanghai 200090, China; Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China

**Peng Yang** – Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China; School of Nanoscience and Technology, University of Chinese Academy of Sciences, Beijing 100049, PR China

**Yuxiang Shi** – Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China; School of Nanoscience and Technology, University of Chinese Academy of Sciences, Beijing 100049, PR China; [orcid.org/0000-0001-5112-4159](https://orcid.org/0000-0001-5112-4159)

**Zhaoqi Liu** – Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China;

School of Nanoscience and Technology, University of Chinese Academy of Sciences, Beijing 100049, PR China

**Dehong Yang** – Beijing Key Laboratory of Micro-nano Energy and Sensor, Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, PR China; School of Nanoscience and Technology, University of Chinese Academy of Sciences, Beijing 100049, PR China

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acsami.3c00566>

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by the National Key R&D Project from Minister of Science and Technology (2021YFA1201601), the National Natural Science Foundation of China (Grant Nos. 62174014 and 22236005), the Beijing Municipal Science & Technology Commission (Z171100000317001), the Program for Professor of Special Appointment (Eastern Scholar) (TP2017076) at Shanghai Institutions of Higher Learning, Young Top-Notch Talents Program of Beijing Excellent Talents Funding (2017000021223ZK03), Beijing Nova program (Z201100006820063), and Youth Innovation Promotion Association CAS (2021165).

## REFERENCES

- (1) Pikul, J. H.; Ning, H. Powering the Internet of Things. *Joule* **2018**, *2* (6), 1036–1038.
- (2) Polyzois, C.; Rogdakis, K.; Kymakis, E. Indoor Perovskite Photovoltaics for the Internet of Things—Challenges and Opportunities toward Market Uptake. *Adv. Energy Mater.* **2021**, *11* (38), 2101854.
- (3) Ramalho, J.; Correia, S. F. H.; Fu, L.; Antonio, L. L. F.; Brites, C. D. S.; Andre, P. S.; Ferreira, R. A. S.; Carlos, L. D. Luminescence Thermometry on the Route of the Mobile-Based Internet of Things (IoT): How Smart QR Codes Make It Real. *Adv. Sci. (Weinh)* **2019**, *6* (19), 1900950.
- (4) Salvatore, G. A.; Sülzle, J.; Dalla Valle, F.; Cantarella, G.; Robotti, F.; Jokic, P.; Knobelspies, S.; Daus, A.; Büthe, L.; Petti, L.; Kirchgessner, N.; Hopf, R.; Magno, M.; Tröster, G. Biodegradable and Highly Deformable Temperature Sensors for the Internet of Things. *Adv. Funct. Mater.* **2017**, *27* (35), 1702390.
- (5) Shi, W.; Guo, Y.; Liu, Y. When Flexible Organic Field-Effect Transistors Meet Biomimetics: A Prospective View of the Internet of Things. *Adv. Mater.* **2020**, *32* (15), No. 1901493.
- (6) Wang, Y.; Liu, X.; Wang, Y.; Wang, H.; Wang, H.; Zhang, S. L.; Zhao, T.; Xu, M.; Wang, Z. L. Flexible Seaweed-Like Triboelectric Nanogenerator as a Wave Energy Harvester Powering Marine Internet of Things. *ACS Nano* **2021**, *15* (10), 15700–15709.
- (7) Wang, Z. L. Entropy Theory of Distributed Energy for Internet of Things. *Nano Energy* **2019**, *58*, 669–672.
- (8) Xu, Z.; Li, C.; Yang, Y. Fault Diagnosis of Rolling Bearings Using an Improved Multi-Scale Convolutional Neural Network with Feature Attention Mechanism. *ISA Trans* **2021**, *110*, 379–393.
- (9) Zhang, S.; Zhang, S.; Wang, B.; Habetler, T. G. Deep Learning Algorithms for Bearing Fault Diagnostics—A Comprehensive Review. *IEEE Access* **2020**, *8*, 29857–29881.
- (10) Zhang, W.; Li, X.; Jia, X.-D.; Ma, H.; Luo, Z.; Li, X. Machinery Fault Diagnosis with Imbalanced Data Using Deep Generative Adversarial Networks. *Measurement* **2020**, *152*, 107377.
- (11) Zhao, M.; Zhong, S.; Fu, X.; Tang, B.; Pecht, M. Deep Residual Shrinkage Networks for Fault Diagnosis. *IEEE Transactions on Industrial Informatics* **2020**, *16* (7), 4681–4690.
- (12) Wang, X.; Shi, Y.; Yang, P.; Tao, X.; Li, S.; Lei, R.; Liu, Z.; Wang, Z. L.; Chen, X. Fish-Wearable Data Snooping Platform for

Underwater Energy Harvesting and Fish Behavior Monitoring. *Small* **2022**, *18* (10), No. 2107232.

(13) Yang, P.; Shi, Y.; Li, S.; Tao, X.; Liu, Z.; Wang, X.; Wang, Z. L.; Chen, X. Monitoring the Degree of Comfort of Shoes In-Motion Using Triboelectric Pressure Sensors with an Ultrawide Detection Range. *ACS Nano* **2022**, *16* (3), 4654–4665.

(14) Jang, H.; Sel, K.; Kim, E.; Kim, S.; Yang, X.; Kang, S.; Ha, K. H.; Wang, R.; Rao, Y.; Jafari, R.; Lu, N. Graphene E-tattoos for Unobstructive Ambulatory Electrodermal Activity Sensing on the Palm Enabled by Heterogeneous Serpentine Ribbons. *Nat. Commun.* **2022**, *13* (1), 6604.

(15) Kim, M. K.; Parasuraman, R. N.; Wang, L.; Park, Y.; Kim, B.; Lee, S. J.; Lu, N.; Min, B.-C.; Lee, C. H. Soft-Packaged Sensory Glove System for Human-like Natural Interaction and Control of Prosthetic Hands. *NPG Asia Mater.* **2019**, *11* (1), 1230.

(16) Berry, M.; Lewin, S.; Brown, S. Correlated Expression of the Body, Face, and Voice during Character Portrayal in Actors. *Sci. Rep.* **2022**, *12* (1), 8253.

(17) Bruneel, L.; Danhieux, A.; Van Lierde, K. Training Speech Pathology Students in the Perceptual Evaluation of Speech in Patients with Cleft Palate: Reliability Results and the Students' Perspective. *Int. J. Pediatr. Otorhinolaryngol* **2022**, *157*, 111145.

(18) Eckstein, K. N.; Wildgruber, D.; Ethofer, T.; Bruck, C.; Jacob, H.; Erb, M.; Kreifelts, B. Correlates of Individual Voice and Face Preferential Responses during Resting State. *Sci. Rep.* **2022**, *12* (1), 7117.

(19) Lileikyte, R.; Irvin, D.; Hansen, J. H. L. Assessing Child Communication Engagement and Statistical Speech Patterns for American English via Speech Recognition in Naturalistic Active Learning Spaces. *Speech Communication* **2022**, *140*, 98–108.

(20) Ozcaliskan, S.; Lucero, C.; Goldin-Meadow, S. Blind Speakers Show Language-Specific Patterns in Co-Speech Gesture but Not Silent Gesture. *Cogn. Sci.* **2018**, *42* (3), 1001–1014.

(21) Pavlidou, A.; Chapellier, V.; Maderthaner, L.; von Kanel, S.; Walther, S. Using Dynamic Point Light Display Stimuli to Assess Gesture Deficits in Schizophrenia. *Schizophr. Res. Cogn* **2022**, *28*, 100240.

(22) Pyers, J. E.; Magid, R.; Gollan, T. H.; Emmorey, K. Gesture Helps, Only If You Need It: Inhibiting Gesture Reduces Tip-of-the-Tongue Resolution for Those with Weak Short-Term Memory. *Cogn. Sci.* **2021**, *45* (1), No. e12914.

(23) Ūnal, E.; Manhardt, F.; Özyürek, A. Speaking and Gesturing Guide Event Perception during Message Conceptualization: Evidence from Eye Movements. *Cognition* **2022**, *225*, 105127.

(24) Ren, Y.; Qing, L.; Li, L.; Hasicicholu; Zheng, H.; Ou, K. Facile Synthesis of Highly Conductive Polymer Fiber for Application in Flexible Fringing Field Capacitive Sensor. *Sensors and Actuators A: Physical* **2022**, *342*, 113616.

(25) Yang, C.-R.; Wang, L.-J.; Tseng, S.-F. Arrayed Porous Polydimethylsiloxane/Barium Titanate Microstructures for High-sensitivity Flexible Capacitive Pressure Sensors. *Ceram. Int.* **2022**, *48* (9), 13144–13153.

(26) Dong, W.; Li, W.; Guo, Y.; Wang, K.; Sheng, D. Mechanical Properties and Piezoresistive Performances of Intrinsic Graphene Nanoplate/Cement-based Sensors Subjected to Impact Load. *Construction and Building Materials* **2022**, *327*, 126978.

(27) Yong, S.; Chapman, J.; Aw, K. Soft and Flexible Large-strain Piezoresistive Sensors: On Implementing Proprioception, Object Classification and Curvature Estimation Systems in Adaptive, Human-like Robot Hands. *Sens. Actuators, A* **2022**, *341*, 113609.

(28) Huang, X.; Qin, Q.; Wang, X.; Xiang, H.; Zheng, J.; Lu, Y.; Lv, C.; Wu, K.; Yan, L.; Wang, N.; Xia, C.; Wang, Z. L. Piezoelectric Nanogenerator for Highly Sensitive and Synchronous Multi-Stimuli Sensing. *ACS Nano* **2021**, *15* (12), 19783–19792.

(29) Han, J. H.; Kwak, J.-H.; Joe, D. J.; Hong, S. K.; Wang, H. S.; Park, J. H.; Hur, S.; Lee, K. J. Basilar Membrane-Inspired Self-powered Acoustic Sensor Enabled by Highly Sensitive Multi Tunable Frequency Band. *Nano Energy* **2018**, *53*, 198–205.

(30) Evans, A. M.; Bradshaw, N. P.; Litchfield, B.; Strauss, M. J.; Seckman, B.; Ryder, M. R.; Castano, I.; Gilmore, C.; Gianneschi, N. C.; Mulzer, C. R.; et al. High-Sensitivity Acoustic Molecular Sensors Based on Large-Area, Spray-Coated 2D Covalent Organic Frameworks. *Adv. Mater.* **2020**, *32* (42), No. 2004205.

(31) Zhao, H.; Xiao, X.; Xu, P.; Zhao, T.; Song, L.; Pan, X.; Mi, J.; Xu, M.; Wang, Z. L. Dual-Tube Helmholtz Resonator-Based Triboelectric Nanogenerator for Highly Efficient Harvesting of Acoustic Energy. *Adv. Energy Mater.* **2019**, *9* (46), 1902824.

(32) Wang, Z.; Wu, Y.; Jiang, W.; Liu, Q.; Wang, X.; Zhang, J.; Zhou, Z.; Zheng, H.; Wang, Z.; Wang, Z. L. A Universal Power Management Strategy Based on Novel Sound-Driven Triboelectric Nanogenerator and Its Fully Self-Powered Wireless System Applications. *Adv. Funct. Mater.* **2021**, *31* (34), 2103081.

(33) Wang, F.; Wang, Z.; Zhou, Y.; Fu, C.; Chen, F.; Zhang, Y.; Lu, H.; Wu, Y.; Chen, L.; Zheng, H. Windmill-Inspired Hybridized Triboelectric Nanogenerators Integrated with Power Management Circuit for Harvesting Wind and Acoustic Energy. *Nano Energy* **2020**, *78*, 105244.

(34) Fan, F.-R.; Tian, Z.-Q.; Lin Wang, Z. Flexible Triboelectric Generator. *Nano Energy* **2012**, *1* (2), 328–334.

(35) Shi, Y.; Lei, R.; Li, F.; Li, S.; Tao, X.; Yang, P.; Chen, X.; Zhao, Q.; Wang, Z. L. Self-Powered Persistent Phosphorescence for Reliable Optical Display. *ACS Energy Letters* **2021**, *6* (9), 3132–3140.

(36) Tian, J.; Wang, F.; Ding, Y.; Lei, R.; Shi, Y.; Tao, X.; Li, S.; Yang, Y.; Chen, X. Self-Powered Room-Temperature Ethanol Sensor Based on Brush-Shaped Triboelectric Nanogenerator. *Research* **2021**, *2021*, 8564780.

(37) Lu, X.; Zheng, L.; Zhang, H.; Wang, W.; Wang, Z. L.; Sun, C. Stretchable, Transparent Triboelectric Nanogenerator as a Highly Sensitive Self-powered Sensor for Driver Fatigue and Distraction Monitoring. *Nano Energy* **2020**, *78*, 105359.

(38) Lei, R.; Shi, Y.; Ding, Y.; Nie, J.; Li, S.; Wang, F.; Zhai, H.; Chen, X.; Wang, Z. L. Sustainable High-voltage Source Based on Triboelectric Nanogenerator with a Charge Accumulation Strategy. *Energy Environ. Sci.* **2020**, *13* (7), 2178–2190.

(39) Lu, S.; Lei, W.; Gao, L.; Chen, X.; Tong, D.; Yuan, P.; Mu, X.; Yu, H. Regulating the High-voltage and High-Impedance Characteristics of Triboelectric Nanogenerator toward Practical Self-powered Sensors. *Nano Energy* **2021**, *87*, 106137.

(40) Zheng, Q.; Fang, L.; Tang, X.; Zheng, L.; Li, H. Indoor Air Dust Removal System Based on High-voltage Direct Current Triboelectric Nanogenerator. *Nano Energy* **2022**, *97*, 107183.

(41) Feng, Y.; Liang, X.; An, J.; Jiang, T.; Wang, Z. L. Soft-contact Cylindrical Triboelectric-electromagnetic Hybrid Nanogenerator Based on Swing Structure for Ultra-low Frequency Water Wave Energy Harvesting. *Nano Energy* **2021**, *81*, 105625.

(42) Jiang, Y.; Zhang, Y.; Ning, C.; Ji, Q.; Peng, X.; Dong, K.; Wang, Z. L. Ultrathin Eardrum-Inspired Self-Powered Acoustic Sensor for Vocal Synchronization Recognition with the Assistance of Machine Learning. *Small* **2022**, *18* (13), No. 2106960.

(43) Guo, H.; Pu, X.; Chen, J.; Meng, Y.; Yeh, M. H.; Liu, G.; Tang, Q.; Chen, B.; Liu, D.; Qi, S.; et al. A Highly Sensitive, Self-powered Triboelectric Auditory Sensor for Social Robotics and Hearing Aids. *Sci. Robot.* **2018**, *3* (20), No. eaat25516.

(44) Fan, X.; Chen, J.; Yang, J.; Bai, P.; Li, Z.; Wang, Z. L. Ultrathin, Rollable, Paper-Based Triboelectric Nanogenerator for Acoustic Energy Harvesting and Self-Powered Sound Recording. *ACS Nano* **2015**, *9* (4), 4236–4243.

(45) Tang, K.; Chen, X.; Zheng, W.; Han, Q.; Li, P. A Non-contact Technique Using Electrostatics to Sense Three-Dimensional Hand Motion for Human Computer Interaction. *J. Electrostat.* **2015**, *77*, 101–109.

(46) Xi, Y.; Hua, J.; Shi, Y. Noncontact Triboelectric Nanogenerator for Human Motion Monitoring and Energy Harvesting. *Nano Energy* **2020**, *69*, 104390.

(47) Lin, Z.; Zhang, G.; Xiao, X.; Au, C.; Zhou, Y.; Sun, C.; Zhou, Z.; Yan, R.; Fan, E.; Si, S.; Weng, L.; Mathur, S.; Yang, J.; Chen, J. A

Personalized Acoustic Interface for Wearable Human-Machine Interaction. *Adv. Funct. Mater.* **2022**, 32 (9), 2109430.

(48) Maharjan, P.; Shrestha, K.; Bhatta, T.; Cho, H.; Park, C.; Salauddin, M.; Rahman, M. T.; Rana, S. S.; Lee, S.; Park, J. Y. Keystroke Dynamics Based Hybrid Nanogenerators for Biometric Authentication and Identification Using Artificial Intelligence. *Adv. Sci.* **2021**, 8 (15), No. 2100711.