BIOELECTRONIC DEVICES

Deciphering facial movements

Wearable piezoelectric thin films integrated with multiphysics modelling and three-dimensional digital image correlation can decode facial expressions in real time.

Xiaoxiang Gao, Xiangjun Chen and Sheng Xu

Human skin — especially on the face — provides psychological, physiological and pathological information[1]. It is thus unsurprising that the deformation dynamics of facial skin have been leveraged in human–machine interactions[2,3] and for health monitoring[4]. Facial movements are also an efficient mode of non-verbal communication[5]. However, accurately measuring and interpreting dynamic skin deformations is challenging. Stereophotogrammetry — the use of multiple cameras to generate three-dimensional (3D) images — can map strains in three dimensions at high spatial resolution when used with 3D digital image correlation (3D-DIC; a method for optical tracking and image registration). But the equipment needed is bulky, which hinders the use of this approach for measuring skin deformations on the human body in the real world[6]. Shape-sensing systems based on flexible fibre Bragg gratings (narrowband reflectors built into optical fibres) can also accurately map strain distributions, but the fibres are not stretchable and hence cannot be used to decode the motion of human skin[7]. Canan Dagdeviren and colleagues now report in *Nature Biomedical Engineering* an integrated system for the decoding of surface strains on the face and for predicting facial kinematics[8]. The system consists of a low-cost wearable piezoelectric device for strain sensing, 3D-DIC for calibration and for guiding device placement, multiphysics modelling for processing the data, and an algorithm for the real-time classification of the piezoelectric signals.

Dagdeviren and co-authors' strain sensor is a 50-μm-thick 2 × 2 array of piezoelectric elements connected by serpentine metal electrodes (Fig. 1a); this configuration provides connectivity while maintaining overall device softness and allowing for conformal contact with facial skin underneath a medical tape. Each element is composed of piezoelectric aluminium nitride thin films sandwiched by Mo electrodes and encapsulated with SiO_2 and polyimide, on an elastomeric polydimethylsiloxane substrate. Rather

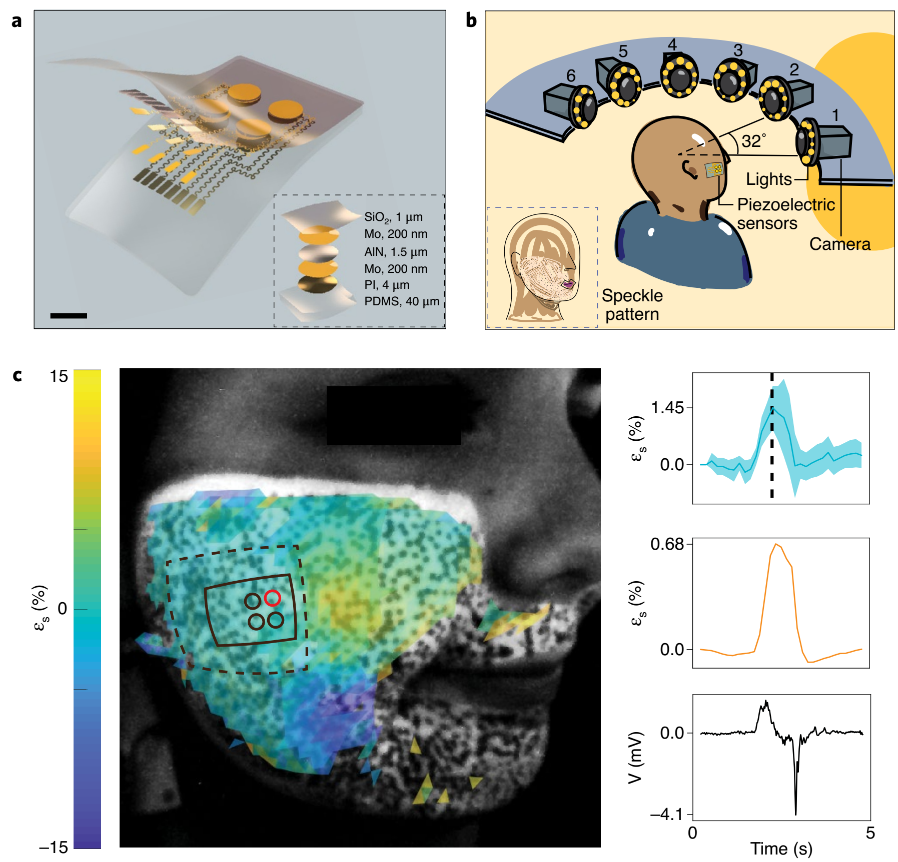

Fig. 1 | An integrated system for decoding facial strains and for the prediction of facial kinematics. a, Wearable conformable piezoelectric sensors. The inset describes the various layers. Scale bar, 5 mm. **b**, Stereophotogrammetry and 3D-DIC. An array of cameras in a circular arc acquires two-dimensional images of a face covered with a pattern of speckles (inset). Facial strain distributions are calculated via 3D reconstruction and the subsequent analysis of the speckle patterns. **c**, Left: map of the facial strain (ε_s) on a smiling face, measured by DIC. Right: simultaneously collected facial strain (blue) and device-voltage (black) waveforms for the sensor element marked by a red circle. The shaded band in the blue waveform indicates the standard deviation. The orange curve shows the theoretical strain waveform predicted from the device's voltage output. The mapped strain data correspond to the timing of the peak strain (dashed black line) in the waveform on the right. The solid rectangle and the circles indicate the location of the device and its four sensing elements, respectively. The outline of the covering medical tape is indicated by the dashed rectangle. AlN, aluminium nitride; PI, polyimide; PDMS, polydimethylsiloxane. Figure adapted with permission from ref. [8], Springer Nature Ltd.

than maximizing strain-induced peak voltages, as needed in energy harvesting[9], piezoelectric strain sensing requires establishing an accurate voltage–strain correlation. The authors characterized the sensor's output-voltage waveform under various mechanical deformations. The sensor differentiated compressive and tensile deformations according to the polarity of the voltage waveforms, which cannot be achieved via the resistive or capacitive strain-sensing mechanisms of non-piezoelectric materials[10]. Piezoelectric voltage waveforms measured in vitro under different deformation modes (concave and convex), rates and magnitudes were all quantitatively matched with those from analytical multiphysics modelling or from finite-element modelling (FEM).

To verify the system's decoding accuracy, Dagdeviren and co-authors combined stereophotogrammetry and 3D-DIC to generate ground-truth data for spatiotemporal strain mapping (Fig. 1b). To track dynamic facial deformations, they painted random black speckles on a white background on the human face. A set of cameras, with their relative positions calibrated, acquired a series of two-dimensional images of the speckle pattern with high spatiotemporal resolution. The images were then stitched together to form a full 3D view. 3D facial strain distributions were then mapped by comparing and calculating the position shifts of the speckles on the reconstructed 3D surface (Fig. 1c). Because FEM can consider all the material layers in the stack and thus has a higher accuracy than multiphysics modelling, the authors only used FEM to quantify strain values from the voltage waveform (by deriving the 'transfer function' between the strain of the sensor's top surface and the sensor's output voltage). For healthy individuals and for patients with amyotrophic lateral sclerosis, the decoded dynamic strains from the voltage waveforms were qualitatively close to those determined by 3D-DIC. The authors also observed that the piezoelectric sensor and 3D-DIC showed distinct biokinetic strain signatures during various facial deformations and device positions. By correlating signatures of different facial motions to the voltage waveforms, they generated a library of facial motions that could be inferred from the waveforms in real time via a *k*-nearest-neighbours classification algorithm. They used a training dataset containing many voltage waveforms for different known skin motions (with one motion corresponding to one or more waveforms) to train the algorithm, which was then used to identify, for each recorded motion, the waveforms in the library that were most similar. The authors also show that the four piezoelectric elements in the device, each generating a voltage waveform, can be laminated on facial areas with distinct deformation patterns, as identified by 3D-DIC. This increases the diversity of voltage waveforms fed to the algorithm, improving its classification accuracy.

Beyond capturing facial cues in patients with amyotrophic lateral sclerosis, Dagdeviren and co-authors' technology could be used to digitize a broad span of human non-verbal communication. This would be difficult with the authors' current devices, as they need to be calibrated to each subject, owing to natural anatomical variations in faces. A population-based method for calibrating the strain–voltage relationship using devices with identical material properties may mitigate this limitation. Still, recalibration may be required owing to material fatigue: the mechanical stability of the polycrystalline structure resulting from the deposition of aluminium nitride by reactive sputtering is affected by repeated deformations[11]; single-crystal thin films should better resist mechanical fatigue and hence help to reduce recalibration frequency. Moreover, as noted by the authors, larger and denser arrays of smaller sensing elements would improve the spatial resolution of strain mapping, and thus the algorithm's classification accuracy. Smaller sensing elements would also reduce the influence of non-uniform strains on each sensing element (FEM assumes a uniform strain in the area covered by each sensor), and facilitate the analysis of the sensors' behaviour under biaxial compression and stretching; after all, human skin is omnidirectionally stretchable. Deep learning may also improve sensor performance[12]. Ultimately, capturing and understanding facial expressions continuously in real time and at high fidelity may aid the study and monitoring of neuromuscular diseases in patients, and facilitate the communication needs of those with impaired facial muscles. ❐

Xiaoxiang Gao[1], Xiangjun Chen[2] and Sheng Xu[1,2,3,4] ✉

[1]*Department of Nanoengineering, University of California San Diego, La Jolla, CA, USA.* [2]*Material Science and Engineering Program, University of California San Diego, La Jolla, CA, USA.* [3]*Department of Bioengineering, University of California San Diego, La Jolla, CA, USA.* [4]*Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, USA.*

✉*e-mail: shengxu@ucsd.edu*

Published online: 22 October 2020
https://doi.org/10.1038/s41551-020-00629-1

References

1. Chung, H. U. et al. *Nat. Med.* **26**, 418–429 (2020).
2. Jeong, J. W. et al. *Adv. Mater.* **25**, 6839–6846 (2013).
3. Yamada, T. et al. *Nat. Nanotechnol.* **6**, 296–301 (2011).
4. Ershad, F. et al. *Nat. Commun.* **11**, 3823 (2020).
5. Hossain, M. S. & Yousuf, M. A. *Int. Arab J. Inf. Technol.* **15**, 278–288 (2018).
6. Chen, Z. et al. *J. Biomed. Opt.* **22**, 095001 (2017).
7. Shih, B. et al. *Sci. Robot.* **5**, eaaz9239 (2020).
8. Sun, T. et al. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-020-00612-w (2020).
9. Xu, S. et al. *Nat. Commun.* **1**, 93 (2010).
10. Amjadi, M. et al. *Adv. Funct. Mater.* **26**, 1678–1698 (2016).
11. Lei, Y. et al. *Nature* **583**, 790–795 (2020).
12. LeCun, Y. et al. *Nature* **521**, 436–444 (2015).

Competing interests

The authors declare no competing interests.