Compression of dynamic tactile information in the human hand

Yitian Shao, Vincent Hayward, Yon Visell

Compression of dynamic tactile information in the human hand

Yitian Shao[1], Vincent Hayward[2,3,4], Yon Visell[1]*

A key problem in the study of the senses is to describe how sense organs extract perceptual information from the physics of the environment. We previously observed that dynamic touch elicits mechanical waves that propagate throughout the hand. Here, we show that these waves produce an efficient encoding of tactile information. The computation of an optimal encoding of thousands of naturally occurring tactile stimuli yielded a compact lexicon of primitive wave patterns that sparsely represented the entire dataset, enabling touch interactions to be classified with an accuracy exceeding 95%. The primitive tactile patterns reflected the interplay of hand anatomy with wave physics. Notably, similar patterns emerged when we applied efficient encoding criteria to spiking data from populations of simulated tactile afferents. This finding suggests that the biomechanics of the hand enables efficient perceptual processing by effecting a preneuronal compression of tactile information.

INTRODUCTION

The sense of touch, which is essential for skilled manipulation and object perception, relies on the encoding of mechanical signals collected by the skin and subcutaneous tissues into neural representations. While neural responses to tactile stimuli are often associated with mechanical inputs arising from small skin regions, we recently observed that dynamic touch elicits mechanical waves in the tactile frequency range that spread throughout the whole hand, with transient excitations decaying within 30 ms (1). Dynamic tactile inputs can thus drive widespread tactile afferent populations (2, 3). These touch-elicited waves have been found to facilitate fine perceptual discriminations (4, 5) and can be used to infer actions, the attributes of touched objects, and locations of contact with the hand (1, 6–8). Receptive fields of neurons in somatosensory cortical areas were observed to span large hand areas and multiple digits (9, 10). The large spatial scale integration at the early stages of processing (11, 12) induces cortical neurons to exhibit integrative responses to tactile inputs delivered to widespread limb regions (13–15). Thus, somatosensory processing could depend on information transported by mechanical waves that propagate in tissues to remote locations, distant from the loci of mechanical contact.

An analogy could be drawn to the cochlea, where the transport of dispersive mechanical waves via the basilar membrane imparts preneuronal filtering to auditory stimuli (16), supporting a frequency-place transformation (17–20). Similar processes have been observed for mechanical waves propagating in the hand (3). In the rodent vibrissal system, whisker mechanics also impart preneuronal processing to tactile stimuli (21, 22).

If the transport of mechanical waves in the hand facilitates efficient somatosensory information encoding, then it should be possible to describe tactile stimuli in terms of a smaller space of informative parameters. This would allow stimuli to be represented as combinations of a small number of primitive features, or tactile patterns. These representations are commonly observed in sensory systems. They correspond to an efficient sensory coding hypothesis that proposes neural circuitry to have evolved to capture relevant sensory information with the fewest physical and metabolic resources (23, 24). Studies of commonly encountered visual and auditory stimuli show that representations in the neural pathways for perceptual processing can emerge from the need to efficiently encode information in natural scenes (25–28).

Here, we show how mechanical waves in the hand produce an efficient encoding of tactile inputs. By optimally encoding a dataset of thousands of naturally occurring whole-hand tactile stimuli, we obtained a compact lexicon of primitive spatiotemporal patterns that sparsely represented information in the entire dataset, enabling it to be classified with an accuracy exceeding 95%. These primitive patterns reflected the interplay of the anatomy of the hand and the physics of tactile wave propagation and were evocative of hand sensory function, including the individuation of digits and the denser innervation of the distal ends of the fingers. We obtained notably similar patterns when we applied the efficient encoding criteria to spiking data from populations of simulated tactile afferents. These results reveal a possible important contribution of the hand biomechanics to early somatosensory processing, which may be compared to the role of cochlear mechanics in early auditory encoding. This new knowledge revises existing views of touch sensing and may aid the understanding of hand sensory function and deficits affecting the sense of touch. It also furnishes new principles that may guide the design of electronic tactile sensors that could leverage the ability of propagating waves to communicate touch information. These devices may yield important applications in robotics, prosthetics, and medicine.

RESULTS

We formulated the efficient encoding of tactile information as an optimal matrix factorization problem and evaluated its predictions using a database of whole-hand tactile stimuli, comprising spatiotemporal skin accelerations, $a(\mathbf{x}, t)$, that were captured at 30 different locations, $\mathbf{x}$, via a sensor array worn on the hand during performances of 13 manual gestures and 4600 interactions with objects (see Fig. 1 and Materials and Methods). Each of the 4600 captured stimuli was represented in the dataset by 18,000 samples.

[1]Department of Electrical and Computer Engineering, Media Arts and Technology Program, Department of Mechanical Engineering, and California NanoSystems Institute, University of California, Santa Barbara, Santa Barbara, CA, USA. [2]Sorbonne Université, Institut des Systèmes Intelligents et de Robotique, F-75005 Paris, France. [3]Centre for the Study of the Senses, School of Advanced Study, University of London, London, UK. [4]Actronika SAS, Paris, France.
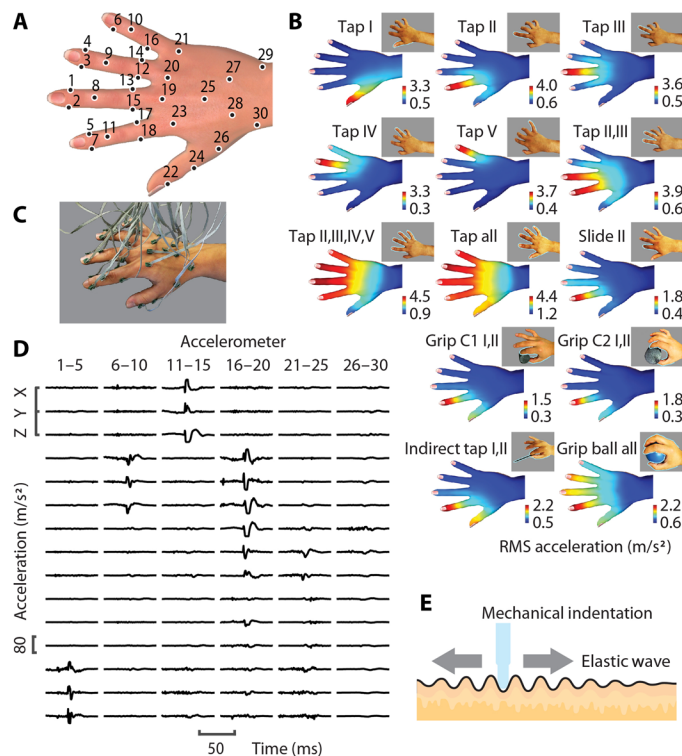*Corresponding author. Email: yonvisell@ucsb.edu

Fig. 1. Recording touch-elicited waves in the whole hand. (**A**) Sensor placements and (**C**) view of the 30 three-axis micro-electromechanical system accelerometers worn on the hand. (**B**) Representative patterns of propagating vibrations in the skin, drawn from the dataset. Each was elicited via hand-object contact during one of 13 manual gesture types (see Materials and Methods). (**D**) Example of 90 acceleration time series elicited by tapping Digit II. The signals rose quickly and faded within 30 ms after contact. (**E**) Skin-object contacts elicit vibrations that propagate throughout the hand as elastic waves.

These stimuli were encoded via a compact lexicon of M primitive spatiotemporal patterns, or "bases," $w_i(\mathbf{x}, t)$, weighted by time-dependent activations, $h_i(t)$, that were unique to each stimulus

$$\hat{a}(\mathbf{x},t) = \sum_{i=1}^{M} \sum_{\tau=0}^{T-1} h_i(t-\tau)\, w_i(\mathbf{x},\tau), \text{ where } w_i(\mathbf{x},t) \geq 0,\, h_i(t) \geq 0 \quad (1)$$

We computed an optimal encoding (Fig. 2A) by maximizing the information about every element of the dataset, $a(\mathbf{x}, t)$, that was gained by observing the estimate, $\hat{a}(\mathbf{x}, t)$, as determined by Eq. 1. The simultaneous optimization of the model with respect to $w_i(\mathbf{x}, t)$ and $h_i(t)$ (see Materials and Methods) yielded a set of "tactile basis patterns," $w_i(\mathbf{x}, t)$, that together produced an efficient encoding, revealing the latent structure hidden in the ensemble of stimuli. These basis patterns optimally represented the dataset in the sense of maximum likelihood (see Supplementary Text).

The bases may also be interpreted as an array of analysis filters that extracted information from the stimuli via different, complementary patterns of spatiotemporal integration of mechanical signals in the hand. These filters may be compared to spectrotemporal tuning functions in auditory processing (29) or spatiotemporal receptive field filters in retinal processing (30). In a minimal functional model of neural population coding, these filter outputs may be passed through nonlinear transfer functions to predict neural firing.

Model (1) included a non-negativity constraint to match the rectifying property of mechanotransduction (31). This encouraged a sparse encoding, (32, 33), as observed in mammalian visual (34) and auditory (35) cortices and in rodent barrel and somatosensory cortices (36, 37).

Although the analysis was blind to the conditions that gave rise to the signals, the tactile bases were evocative of hand sensory function (Fig. 2, A and B). Most were initially localized at the distal ends of single digits (the most densely innervated regions of the hand). They traveled proximally at rates of 1 to 10 m/s, while decaying over 10 to 30 ms, matching the causal physics of waves in the hand (see the Supplementary Materials). Other bases evolved from the distal region of individual digits to diffuse regions of the hand surface (Fig. 2A). In the frequency domain, pairs of bases exhibited similar spatial patterns but distinct frequency characteristics. For example, the encoding yielded pairs of bases that were both spatially localized within one digit but had different filtering properties: low pass, from about 20 to 80 Hz (Fig. 2B, basis 2), or high pass, from 80 to 160 Hz (Fig. 2B, basis 6). Similar patterns emerged when the encoding rank M, or number of bases, was adjusted (Fig. 3B) or when optimizing with different initial conditions, data subsets, or optimization objectives (figs. S1 to S6).

It could be hypothesized that the structure of our dataset favored such an encoding, in which several spatiotemporal basis patterns were associated with individual digits. For example, 45% of the 4600 analyzed tactile stimuli were elicited by gestures that produced contact at only one digit. To investigate this possibility, we applied the same analysis to a subset of the data that excluded tactile stimuli produced by single-digit gestures. We repeated the same analysis using an additional subset that only included stimuli produced via gestures involving contact with all five digits. In each case, the results were highly similar to those we obtained from encoding the entire dataset (fig. S9), including distinct basis patterns that were primarily localized in single digits.

The space of possible tactile stimuli is constrained by contact and continuum mechanics (Fig. 3A and fig. S8). To assess the number of bases, $w_i(\mathbf{x}, t)$, that were needed to capture information about the causal origin of the stimuli, we varied the encoding rank and trained support vector machine (SVM) classifiers to predict the gestures from the activation patterns. The encoding was not selected to optimize classification accuracy. Nonetheless, the classification accuracy increased with the number of bases and was greater than 90% if at least 7 bases were used, or greater than 95% if 12 bases were used. A high classification accuracy is not necessarily expected to be achieved when the number of input dimensions exceeds the number of classes. For example, there are binary prediction tasks, such as cancer prognosis prediction from images, where the consideration of multitudes of features is required to achieve moderately accurate classification (38).

The bases encoded the stimuli via a small number of time-dependent activation weights (Fig. 2, C and D). Stimuli elicited by multifinger gestures were encoded by several bases, while simpler gestures activated one or two. Tactile stimuli produced via similar gestures yielded similar activation patterns (fig. S5), while dissimilar gestures resulted in dissimilar activations, even when the same combinations of digits were involved.

The encoding residual decreased with the number of bases. Five bases were sufficient to maximize the accuracy (80%) with which stimuli from one participant could be classified using only data from the other participants (Fig. 3C). These five bases were highly conserved between individuals and were associated with individual
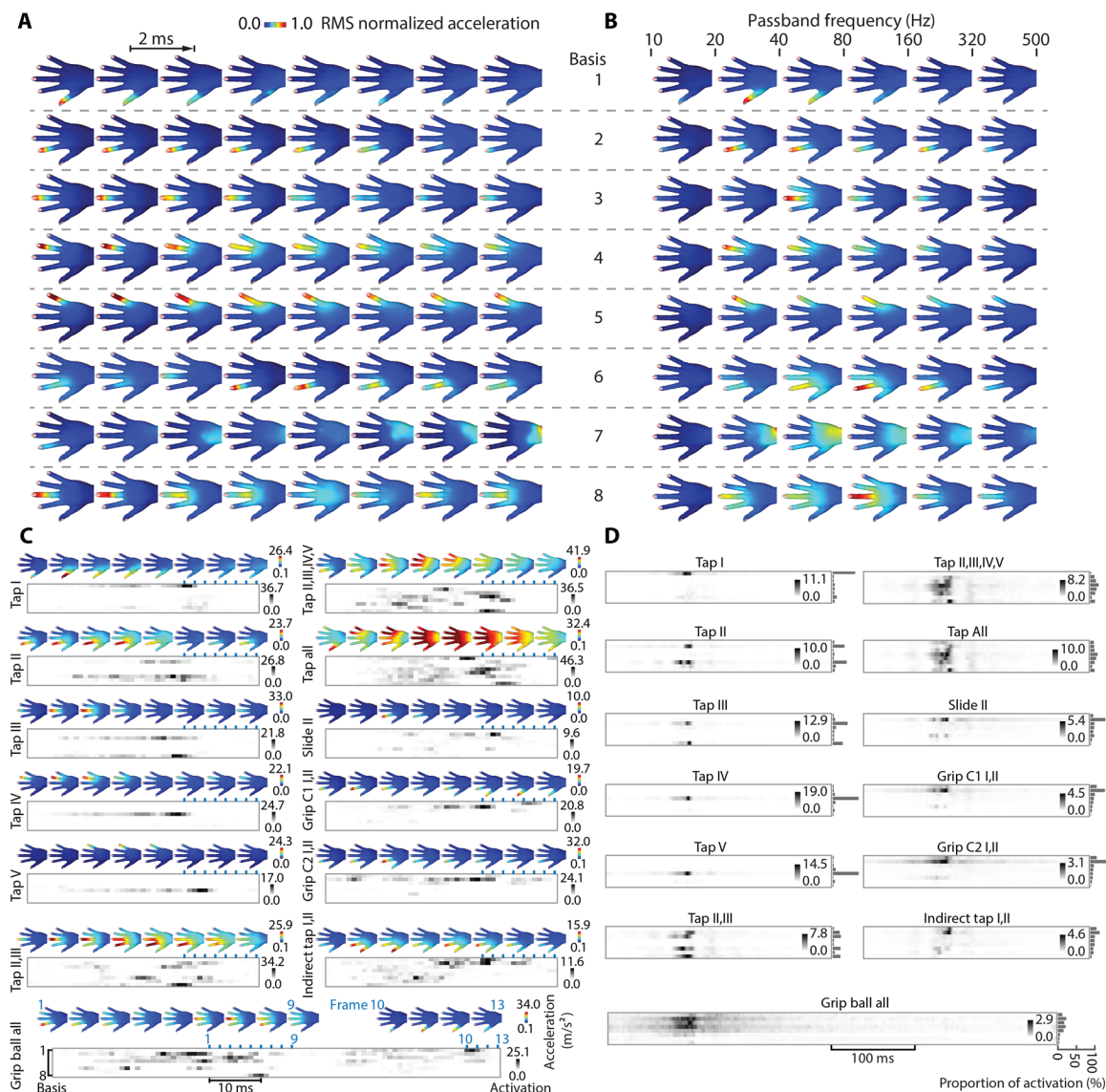
Fig. 2. Optimal spatiotemporal primitives and encodings. (**A**) Each row represents a spatiotemporal basis pattern. Motifs were similar when the rank, *M*, was varied (Fig. 3). The tactile bases, displayed at 2-ms intervals in descending order of activation, reflect the individuation of digits and larger representations of commonly used digits II and III. (**B**) Analysis within different frequency bands revealed that different basis patterns captured distinct frequency content. (**C**) Activations, $h_i(t)$ (shown in grayscale, temporal resolution: 1 ms), produced by encoding the displayed tactile stimulus (see Fig. 1). Blue ticks show the time instant for each displayed stimulus frame. (**D**) Mean activations for stimuli elicited by each gesture class, averaged across all encoded trials.

digits (Fig. 3B). The activations of the bases (Fig. 2, C and D) exhibited a high degree of sparsity preserved across many trials (see table S1). Those that were associated with multiple finger contact were less sparse and more diverse than those involving just one finger. Information independence among the basis activations decreased with the number of digits engaged.

The observed encoding efficiency of the mechanical signals was a consequence of spatiotemporal integration supplied by the tactile basis patterns. Prevailing physiological models leave little doubt that the spatial and temporal properties of touch-elicited mechanical signals are reflected in the volleys of afferent activations during natural hand interactions. However, extant methods preclude the simultaneous capture of neural signals from populations of peripheral afferents in the behaving hand. We instead computationally predicted the spiking

responses of a population of 773 vibration-sensitive afferents excited by the raw mechanical signals in the entire dataset (fig. S7). The neural simulation yielded 773 spike trains for each of the 4600 trials in the dataset.

We optimized the encoding of the predicted neural responses with the same method that was used for mechanical signals. The results were notably similar (Fig. 3, D and E). The spiking bases exhibited similar patterns of spatial integration to the bases that we obtained using the mechanical data, including individuation of digit representations and denser activation of the fingertips. The results of the classification tasks were qualitatively similar to those that we obtained from the mechanical data, despite the higher dimensionality of the input data (see Materials and Methods). This suggests that the encoding revealed organizational principles that would be preserved
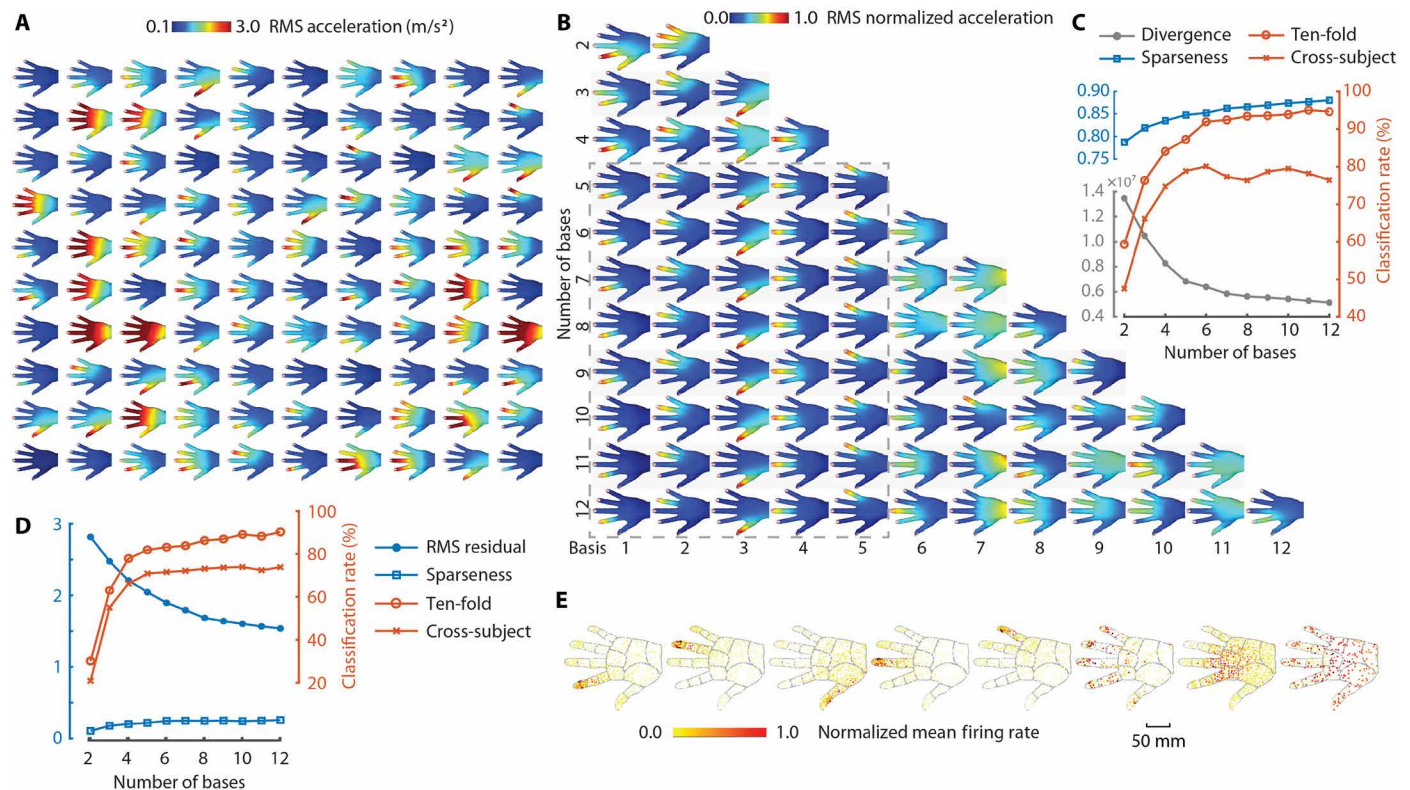
Fig. 3. Learning and evaluating efficient tactile encodings. (**A**) Random sampling of 100 tactile stimuli drawn from the dataset illustrates its diversity (time averages shown). (**B**) Maximizing the efficiency with which the stimuli were encoded yielded primitive bases, $w_i(\mathbf{x}, t)$ (time averages shown). Each row corresponds to an encoding of fixed rank, M, from $M = 2$ to 12 bases, arranged in order of increasing activation. Individuated digit representations were highly conserved. Higher-rank encodings included additional diffuse patterns. (**C**) The encoding residual decreased with the rank, while classification improved. (**D**) Results based on the neural simulations. As the number of spiking bases increased, the encoding residual decreased, and classification also improved. (**E**) Optimizing the encoding of spiking data from 773 simulated afferents yielded bases ($M = 8$ shown) that integrated activity in neural populations throughout the hand. They reflected the individuation of digits and denser activation of the fingertips, similar to results obtained with the mechanical data.

by neurotransduction and that went beyond the mere properties of skin vibrations.

DISCUSSION

The size and the diversity of our corpus of data were limited by experimental constraints, as in many other studies using corpora of motor or sensory data (*39*–*41*). Although we selected the gestures on the basis of the most reasonable assumptions available, a larger dataset could be captured during spontaneous manual activities outside of the laboratory or specified on the basis of on analyses of conditions in which our species evolved.

A useful comparison can be drawn to research on hand movements and grasping. Research in this area has shown how a relatively small number of coordination patterns ("synergies") can explain most of the variability in hand movement data. Similar coordination patterns have been observed in studies based on different laboratory datasets, or on spontaneous manual activities outside the laboratory, with some task dependency (*41*). The dimensionality of the analysis presented here is much larger than typically arises in hand kinematic studies. Nonetheless, analogous considerations may apply to our findings. Our analyses of subsets of the mechanical data yielded basis patterns that were very similar to those that we obtained from the combined dataset.

The tactile basis patterns were also invariably organized along a gradient from higher to lower finger individuation. This is opposite to the trend that is observed in grasping studies and may evince an important difference in organizational principles between the tactile and motor systems. The larger degree of individuation that our findings suggest appears to be a consequence of the physics of vibration transmission in the skin, which causes propagating vibrations to attenuate with increasing distance from their source. In contrast, hand movement studies reveal a higher degree of multidigit coordination, which is facilitated by the biomechanics of the limb.

While the behavioral relevance of propagating vibrations in the limb is not fully understood, previous research shows how these vibrations can mediate tactile perception (*4*, *5*). Further research is needed to clarify the relevance of the predictions from efficient tactile encoding to hand function and somatosensory processing.

Our findings suggest that the biomechanics of the hand can facilitate tactile perception by effecting the preneuronal compression of tactile information in the whole hand. This compression was produced by a compact lexicon of primitive tactile wave patterns. Spatiotemporal integration supplied by these basis patterns optimally encoded the tactile stimuli. Recent studies of neural correlates of somatosensory processing reveal that, at the earliest stages of cortical processing, individual neurons exhibit complex responses to tactile stimuli distributed throughout the extremities (*15*, *42*, *43*). These studies, together

with the new findings presented here, show how traditional depictions of receptive fields do not reflect the extent of early somatosensory integration (*44*), including effects of mechanical transmission in the body.

MATERIALS AND METHODS
Apparatus
In our previous study (*1*), we developed a custom array of 30 three-channel miniature accelerometers (model ADXL335; Analog Devices) attached to the skin to record the stimuli (Fig. 1C). The sensors were attached to the dorsal hand region, because collecting measurements from the volar region of the hand during natural activity remains technically prohibitive, due to the necessity to expose the glabrous skin to contact. However, we observed that patterns of mechanical wave propagation generated in the volar and dorsal regions are quite similar (fig. S8), indicating that similar results would be associated with the volar hand region.

Each accelerometer had a mass of 40 mg. We used noncontact laser vibrometry to verify that the small mass of the sensors did not significantly affect the measurements. They were soldered to a miniature two-sided printed circuit board (dimensions, 6 mm × 8 mm), had a wide bandwidth (0 to 1600 Hz in X and Y; 0 to 550 Hz in Z), and a dynamic range overlapping that of the vibrotactile system (±35.3 m/s^2). The accelerometers were affixed to the skin using an elastic, skin-compatible prosthetic adhesive (Pros-Aide, FX Warehouse, Philadelphia, PA). The 90 analog signals measured via this apparatus were sampled at a frequency of 2.0 kHz and quantized with a resolution of 12 bits by a data acquisition system (model PCIE-6321, National Instruments, Austin, TX).

Data
The data consisted of touch-elicited vibrations of the skin that were captured from four individuals (one female and three males, aged 19 to 23 years old). Experiments were conducted consistent with institutional ethics guidelines, and all participants gave their informed consent. No participant reported or exhibited abnormalities of the hands. All were right hand dominant and wore the accelerometer array as indicated in Fig. 1A. The accelerometers were positioned on the hand's dorsal surface so that they would not interfere with touch interactions. The signals were captured from the dorsal surface of the hand, which avoided introducing artifacts in the data. Because of the properties of vibration transmission in the hand, the captured signals are quite similar to those occurring in the volar surface (see fig. S8). The positions were anatomically standardized. Tactile signals in the hand were measured as individuals performed 13 different prescribed manual gestures during each of the 4600 trials. The gestures were selected to be similar to those used when interacting with the environment in everyday life. The majority involved coupled movement of multiple digits and contact between different parts of the hand and objects: tapping a steel plate with individual digits or combinations of digits, feeling the surface via sliding contact, two-finger grasping of a small or large plastic cylinder (diameter d = 40 or 56 mm, masses m = 31 g) with digits I and II, grasping a plastic ball (d = 63 mm, m = 26 g) with all fingers, and indirectly tapping a surface via a stylus (d = 6 mm, length L = 155 mm, m = 30 g) held in digits I and II. Participants were instructed to use forces of approximately 1 N. The gestures were otherwise unconstrained. Measurements were captured in successive blocks of identical activities. Each block of measurement

trials lasted 45 s and was composed of 20 trials (tapping gestures) or 10 trials (other gestures). Visual cueing helped participants to maintain a pace of 2 or 4 s per trial in respective cases. The tactile signals elicited by each gesture spanned 1 to 2 s of data. For analysis, we extracted the time-varying acceleration magnitude, $\|a_k(t)\|$, from each kth accelerometer (see below), truncated each trial to 600 ms, and downsampled the data to 1.0 kHz. The analyzed data from each trial thus consisted of 30 time-varying signals of 600-ms duration sampled at 1.0 kHz, yielding 18,000 data samples per trial. Thus, the nominal dimensionality of each of the 4600 spatiotemporal stimuli in the dataset was 18,000. The total storage of the dataset required 165 MB.

Spatiotemporal patterns of skin oscillations
Although the skin acceleration measurements were sampled at a discrete array of points, they provided a sufficient representation of information in the fields of tactile waves in the hand, because the wavelength, λ, was at least twice as large as the accelerometer spacing, thus satisfying a Nyquist criterion. From wave mechanics, summarized further on in Supplementary Text, λ satisfies a dispersion relation λ = $c(f)/f$, where $c(f)$ < 10 m/s was the frequency-dependent speed of surface wave propagation in the analyzed range of frequencies, $10 < f < 1000$ Hz. We determined that λ = $c(f)/f$ was larger than 10 mm for all frequencies. Further discussion is provided in Supplementary Text.

To match the rectifying properties of tactile afferents, we first computed signal magnitudes. The acceleration magnitude from the kth accelerometer at time frame t was computed as $a_k(t) = \|\mathbf{a}_k(t)\|_2$, where $\mathbf{a}_k(t)$ is the vector signal from the kth accelerometer. We translated the accelerometer signals into spatiotemporal skin motion by interpolating the acceleration magnitudes among nearby measurement locations, as shown in Fig. 1B, using an inverse-distance filter, informed by biomechanical measurements (*1*, *3*). The acceleration amplitude $a(\mathbf{x}, t)$ at each location $\mathbf{x} = (x_1, x_2, x_3)$ on the surface of the model hand was computed as a weighted sum of accelerations $a_k(t)$ at nearby sensors

$$a(\mathbf{x}, t) = \frac{\sum_{k=1}^{30} f(\phi_k(\mathbf{x})) \, a_k(t)}{\sum_{k=1}^{30} f(\phi_k(\mathbf{x}))},$$

$$\text{with } \phi_i(\mathbf{x}) = \frac{\beta}{d_k(\mathbf{x}) + \alpha} - C, \text{ and } f(\phi) = \begin{cases} \phi, \phi \geq 0 \\ 0, \phi < 0 \end{cases} \quad (2)$$

where $\phi_k(\mathbf{x})$ is a function of the distance $d_k(\mathbf{x})$ from position of the interpolated point $\mathbf{x}$ to that of the kth accelerometer and $f(\phi)$ is a half-wave rectifying operator replacing all negative values with zeros. We evaluated Eq. 2, using values α = 25.5 mm, β = 17.0 mm, and $C = 8.7 \times 10^{-2}$, which we selected on the basis of previously published data. (*3*) This yielded time-dependent spatiotemporal stimuli comprising trials with a duration of 100 ms (or 200 ms for sliding gestures). These were rendered on a scale-normalized three-dimensional hand to characterize patterns of skin vibrations during active touch (movie S1).

We analyzed frequency content in the stimuli by filtering them to extract content in separate frequency bands: 10 to 20 Hz, 20 to 40 Hz, 40 to 80 Hz, 80 to 160 Hz, 160 to 320 Hz, and 320 to 500 Hz (Fig. 2B). To avoid artifacts, filtering was performed using zero-phase finite impulse response filters.

Spatiotemporal encoding
The model of efficient spatiotemporal encoding is based on convolutive non-negative matrix factorization (*45*). This model may be

compared to that used to represent stimulus encoding in the auditory system (*40*). The model is mathematically simple; requires few arbitrary choices, the effects of which are readily analyzed; can accommodate physiologically motivated assumptions; and can be compared with models of sensory encoding in other modalities. It encoded the tactile stimuli, $a(\mathbf{x}, t)$, by determining the values of h_i and w_i that provided the best statistical estimate, $\hat{a}(\mathbf{x}, t) = a(\mathbf{x}, t) + \eta(\mathbf{x}, t)$, of $a(\mathbf{x}, t)$ as determined by model (1), where η is a residual error.

The same tactile basis patterns, $w_i(\mathbf{x}, t)$, encoded all 18,000 sample values of all 4600 stimuli in the dataset. The activation weights, $h_i(t)$, associated with each basis differed for each stimulus. Both factors were jointly estimated from the data. Each basis could assume arbitrary non-negative values for each position and time. No other statistical assumption was made about the data. The model was causal (Eq. 1); hence, the bases described responses that ensued with delays, τ. We set their duration, T, to the time required for mechanical waves in the hand to decay, about 30 ms, although our findings were robust to variations in duration (see fig. S4). This duration spanned 30 time samples at the sample rate of 1 kHz. Each basis pattern was therefore represented by 900 values. We computed optimal encodings with ranks $M = 2$ to 12, corresponding to 2 to 12 basis patterns. For each value of M, we determined the optimal basis set and per-stimulus activation weights via simultaneous iterative optimization over w_i and h_i, beginning from random initializations of each (*32*). This optimization maximized the statistical information about the stimulus, $a(\mathbf{x}, t)$, that is gained by observing the estimate, $\hat{a}(\mathbf{x}, t)$ (see Eq. 1), as measured by the Kullback-Leibler divergence

$$D_{KL}(a(\mathbf{x},t),\hat{a}(\mathbf{x},t)) = \sum_{\mathbf{x},t} \left(a(\mathbf{x},t) \ln \frac{a(\mathbf{x},t)}{\hat{a}(\mathbf{x},t)} - a(\mathbf{x},t) + \hat{a}(\mathbf{x},t) \right)$$

This measure quantified the dissimilarity between a and $\hat{a}$, regarding them as statistical distributions that encoded information. Under mild assumptions, minimizing the Kullback-Leibler divergence is equivalent to maximizing, with respect to $h_i(t)$ and $w_i(\mathbf{x}, t)$, the likelihood of model (1) to represent the data.

Solving this minimization problem involved the determination of parameters $h_i(t)$ and $w_i(\mathbf{x}, t)$ that best captured information in the ensemble of data.

Classifying manual interactions from tactile codes

The optimization yielding the tactile codes was performed in an unsupervised manner, without using knowledge about the manual interaction that produced the tactile signals, the touched objects, or any other factors.

To assess the number of bases, $w_i(\mathbf{x}, t)$, that were necessary to capture information about the causal origin of the stimuli, we varied the encoding rank (number of bases) and designed a classification task whose objective was to use the activation weight pattern to identify the gesture that elicited the stimulus. We integrated the weights over time, $\bar{h}_i = \frac{1}{T} \sum_{t=0}^{T-1} h_i(t)$, to eliminate the adverse effects arising from timing differences across trials. The task involved multiclass classification, which we implemented as thirteen 1-versus-12 classification tasks. We avoided classification methods, such as convolutional neural networks, that would require extensive model tuning. We instead opted for SVM classifiers, which require few choices and are theoretically sound, involving a convex optimization. All classifiers used a radial basis function SVM kernel (width, 5.0; selected using an independent validation set). We evaluated classification performance using a standard (10-fold) cross-validation method, with a 90% training and 10% testing data split. To assess the between-individuals generalizability of these inferences, we performed a cross-individual validation, in which we trained a classifier on data from three participants and tested it on data from the fourth and averaged the results across each left-out participant.

Evaluating the sparseness of the encoding

The model included a non-negativity constraint that encouraged sparseness. Sensory processing research has often connected neural representations with sparse codes (*34*–*37*).

The Hoyer sparseness measure, a normalized ratio of ℓ^1 and ℓ^2 norms, is often preferred, on the basis of criteria discussed in the literature (*46*)

$$\text{Sparseness}(h) = \frac{\sqrt{n} - (\sum |h_i|)/\sqrt{\sum h_i^2}}{\sqrt{n} - 1}$$

where n is the number of elements in h.

Encoding efficiency

We assessed the diversity of the encoding by computing the empirical Shannon information entropy of the activation values across the entire dataset, as follows. Discretize the activation values $h_i(t)$ and let p_k be the probability that a randomly drawn activation from any stimulus, channel, or time has a value lying in histogram bin k. The joint entropy of activations in all channels was

$$H_J = H(\{h_1, h_2, \ldots, h_M\}) = -\sum_{k=1}^{K} p_k \log_2 p_k$$

The distribution, p_k, was computed from all values of $h_i(t)$ in the entire dataset, for all i. The entropy H_J was maximized when all weight values were equally likely and decreased as the sparseness of the code increased. This measure revealed differences between encoded stimuli produced by different actions (table S1). The joint activation entropy H_J was highest for gestures involving multifinger contact and lowest for contacts of single fingers, suggesting that the model was most efficient at encoding gestures involving individual digits. For each basis, we computed the entropies, $H(h_i)$

$$H(h_i) = -\sum_k p(h_{i,k}) \log_2 p(h_{i,k}), \ i = 1, \ldots, M$$

where $h_{i,k}$ are the discrete values of the activation coefficients, h_i, from the histogram distribution, $p(h_{i,k})$, computed using the entire dataset. These entropies were then summed

$$H_S = \sum_{i=1}^{M} H(h_i)$$

giving a result that satisfied, $H_J \leq H_S$. Equality was reached if the activations for different bases were statistically independent; thus, $H_S - H_J$ measured the degree of dependence of the activations among the bases (table S1). When this quantity was minimized, the bases represented maximally independent components of the data.

Encoding neural population responses: Spiking simulations

Because no method is known to record neural activity in multiple peripheral afferents during natural manual interactions, we used a biologically justified neuron spiking simulation software [TouchSim (*47*)]. This software predicted, in silico, the firing patterns of 773

vibration-sensitive afferents distributed throughout a simulated hand in response to the touch-elicited vibrations of the skin that we captured in vivo.

We computed skin displacements from the skin acceleration data. The skin displacement signals were used to drive the TouchSim model. For each of the 4600 trials of the entire database, the model produced and output spike trains for each of the 773 simulated Pacinian corpuscle afferents [PC afferents form a class of afferent fibers terminating in Pacinian corpuscles thought to play a major role in the encoding of skin vibrations (48–50)]. We also computed mean firing rates for each. The output of the simulation was thus the spike train data and mean firing rate of the afferents for each trial. Representative trials, and mean firing rates for several gesture classes, are shown in fig. S7.

We analyzed the mean firing rate data for all 773 PCs using a non-negative matrix factorization procedure, similar to the one used in our analysis of the acceleration data. Informed by our analysis of the acceleration data, we performed the non-negative matrix factorization analysis of the simulated neural data for 2 to 12 bases, yielding 11 different encodings of increasing dimensionality. The eight-basis solution is shown in Fig. 3E. Each basis describes a distribution of mean firing rates used in the encoding. The bases bore a notable resemblance to those that we obtained by analyzing the acceleration data. We evaluated the quality of the encodings using a classification task, residual measure, sparseness measure, and a cross-participant classification task. The classification and evaluation procedures were exactly the same as those used for the accelerometer data and were not optimized for these data. Nonetheless, classification rates reached 90% with eight bases (Fig. 3D).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/16/eaaz1158/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. Y. Shao, V. Hayward, Y. Visell, Spatial patterns of cutaneous vibration during whole-hand haptic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4188–4193 (2016).
2. B. H. Pubols Jr., Effect of mechanical stimulus spread across glabrous skin of raccoon and squirrel monkey hand on tactile primary afferent fiber discharge. *Somatosens. Res.* **4**, 273–308 (1987).
3. L. R. Manfredi, A. T. Baker, D. O. Elias, J. F. Dammann III, M. C. Zielinski, V. S. Polashock, S. J. Bensmaia, The effect of surface wave propagation on neural responses to vibration in primate glabrous skin. *PLOS ONE* **7**, e31203 (2012).
4. J. W. Morley, M. J. Hawken, P. D. Burge, Vibratory detection thresholds following a digital nerve lesion. *Exp. Brain Res.* **72**, 215–218 (1988).
5. X. Libouton, O. Barbier, Y. Berger, L. Plaghki, J.-L. Thonnard, Tactile roughness discrimination of the finger pad relies primarily on vibration sensitive afferents not necessarily located in the hand. *Behav. Brain Res.* **229**, 273–279 (2012).
6. A. I. Weber, H. P. Saal, J. D. Lieber, J.-W. Cheng, L. R. Manfredi, J. F. Dammann, S. J. Bensmaia, Spatial and temporal codes mediate the tactile perception of natural textures. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17107–17112 (2013).
7. B. Delhaye, V. Hayward, P. Lefèvre, J.-L. Thonnard, Texture-induced vibrations in the forearm during tactile exploration. *Front. Behav. Neurosci.* **6**, 37 (2012).
8. L. R. Manfredi, H. P. Saal, K. J. Brown, M. C. Zielinski, J. F. Dammann III, V. S. Polashock, S. J. Bensmaia, Natural scenes in tactile texture. *J. Neurophysiol.* **111**, 1792–1802 (2014).
9. Y. Iwamura, M. Tanaka, M. Sakamoto, O. Hikosaka, Rostrocaudal gradients in the neuronal receptive field complexity in the finger region of the alert monkey's postcentral gyrus. *Exp. Brain Res.* **92**, 360–368 (1993).
10. M. Sur, Receptive fields of neurons in areas 3b and 1 of somatosensory cortex in monkeys. *Brain Res.* **198**, 465–471 (1980).
11. R. S. Johansson, I. Birznieks, First spikes in ensembles of human tactile afferents code complex spatial fingertip events. *Nat. Neurosci.* **7**, 170–177 (2004).
12. F. Bengtsson, R. Brasselet, R. Johansson, A. Arleo, H. Jörntell, Integration of sensory quanta in cuneate nucleus neurons in vivo. *PLOS ONE* **8**, e56630 (2013).
13. E. P. Gardner, R. M. Costanzo, Spatial integration of multiple-point stimuli in primary somatosensory cortical receptive fields of alert monkeys. *J. Neurophysiol.* **43**, 420–443 (1980).
14. M. Prsa, K. Morandell, G. Cuenu, D. Huber, Feature-selective encoding of substrate vibrations in the forelimb somatosensory cortex. *Nature* **567**, 384–388 (2019).
15. J. M. D. Enander, H. Jörntell, Somatosensory cortical neurons decode tactile input patterns and location from both dominant and non-dominant digits. *Cell Rep.* **26**, 3551–3560.e4 (2019).
16. R. Nobili, F. Mammano, J. Ashmore, How well do we understand the cochlea? *Trends Neurosci.* **21**, 159–167 (1998).
17. G. Von Békésy, *Experiments in Hearing* (McGraw-Hill, 1960).
18. E. de Boer, M. A. Viergever, Wave propagation and dispersion in the cochlea. *Hear. Res.* **13**, 101–112 (1984).
19. M. Chatterjee, J. J. Zwislocki, Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch. *Hear. Res.* **111**, 65–75 (1997).
20. A. Paredes Gallardo, B. Epp, T. Dau, Can place-specific cochlear dispersion be represented by auditory steady-state responses? *Hear. Res.* **335**, 76–82 (2016).
21. M. A. Neimark, M. L. Andermann, J. J. Hopfield, C. I. Moore, Vibrissa resonance as a transduction mechanism for tactile encoding. *J. Neurosci.* **23**, 6499–6509 (2003).
22. K. Bagdasarian, M. Szwed, P. M. Knutsen, D. Deutsch, D. Derdikman, M. Pietr, E. Simony, E. Ahissar, Pre-neuronal morphological processing of object location by individual whiskers. *Nat. Neurosci.* **16**, 622–631 (2013).
23. F. Attneave, Some informational aspects of visual perception. *Psychol. Rev.* **61**, 183–193 (1954).
24. H. B. Barlow, Possible principles underlying the transformation of sensory messages, in *Sensory Communication* (MIT Press, 1961), pp. 217–234.
25. B. A. Olshausen, D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).
26. Y. Dan, J. J. Atick, R. C. Reid, Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *J. Neurosci.* **16**, 3351–3362 (1996).
27. A. J. Bell, T. J. Sejnowski, The "independent components" of natural scenes are edge filters. *Vision Res.* **37**, 3327–3338 (1997).
28. M. S. Lewicki, Efficient coding of natural sounds. *Nat. Neurosci.* **5**, 356–363 (2002).
29. J. J. Eggermont, P. M. Johannesma, A. M. Aertsen, Reverse-correlation methods in auditory research. *Q. Rev. Biophys.* **16**, 341–414 (1983).
30. S. Hochstein, R. M. Shapley, Quantitative analysis of retinal ganglion cell classifications. *J. Physiol.* **262**, 237–264 (1976).
31. R. S. Johansson, J. R. Flanagan, Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* **10**, 345–359 (2009).
32. D. D. Lee, H. S. Seung, Algorithms for non-negative matrix factorization, in *Advances in Neural Information Processing Systems* (NIPS, 2001), pp. 556–562.
33. D. Donoho, V. Stodden, When does non-negative matrix factorization give a correct decomposition into parts?, in *Advances in Neural Information Processing Systems* (NIPS, 2004), pp. 1141–1148.
34. W. E. Vinje, J. L. Gallant, Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* **287**, 1273–1276 (2000).
35. H. Terashima, H. Hosoya, T. Tani, N. Ichinohe, M. Okada, Sparse coding of harmonic vocalization in monkey auditory cortex. *Neurocomputing* **103**, 14–21 (2013).
36. S. P. Jadhav, J. Wolfe, D. E. Feldman, Sparse temporal coding of elementary tactile features during active whisker sensation. *Nat. Neurosci.* **12**, 792–800 (2009).
37. S. Crochet, J. F. Poulet, Y. Kremer, C. C. Petersen, Synaptic mechanisms underlying sparse coding of active touch. *Neuron* **69**, 1160–1175 (2011).
38. A. H. Beck, A. R. Sangoi, S. Leung, R. J. Marinelli, T. O. Nielsen, M. J. van de Vijver, R. B. West, M. van de Rijn, D. Koller, Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci. Transl. Med.* **3**, 108ra113 (2011).
39. E. P. Simoncelli, B. A. Olshausen, Natural image statistics and neural representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).
40. M. S. Lewicki, T. J. Sejnowski, Coding time-varying signals using sparse, shift-invariant representations, in *Advances in Neural Information Processing Systems* (NIPS, 1999), pp. 730–736.
41. M. Santello, G. Baud-Bovy, H. Jörntell, Neural bases of hand synergies. *Front. Comput. Neurosci.* **7**, 23 (2013).
42. M. L. Lipton, M. C. Liszewski, M. N. O'Connell, A. Mills, J. F. Smiley, C. A. Branch, J. R. Isler, C. E. Schroeder, Interactions within the hand representation in primary somatosensory cortex of primates. *J. Neurosci.* **30**, 15895–15903 (2010).
43. G. Foffani, J. K. Chapin, K. A. Moxon, Computational role of large receptive fields in the primary somatosensory cortex. *J. Neurophysiol.* **100**, 268–280 (2008).
44. J. L. Reed, P. Pouget, H.-X. Qi, Z. Zhou, M. R. Bernard, M. J. Burish, J. Haitas, A. B. Bonds, J. H. Kaas, Widespread spatial integration in primary somatosensory cortex. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10233–10237 (2008).

45. P. Smaragdis, *International Conference on Independent Component Analysis and Signal Separation* (Springer, 2004), pp. 494–499.

46. N. Hurley, S. Rickard, Comparing measures of sparsity. *IEEE Trans. Inf. Theory* **55**, 4723–4741 (2009).

47. H. P. Saal, B. P. Delhaye, B. C. Rayhaun, S. J. Bensmaia, Simulating tactile signals from the whole hand with millisecond precision. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E5693–E5702 (2017).

48. N. Cauna, G. Mannan, The structure of human digital pacinian corpuscles (corpus cula lamellosa) and its functional significance. *J. Anat.* **92**, 1–20 (1958).

49. R. S. Johansson, U. Landström, R. Lundström, Responses of mechanoreceptive afferent units in the glabrous skin of the human hand to sinusoidal skin displacements. *Brain Res.* **244**, 17–25 (1982).

50. B. Stark, T. Carlstedt, R. G. Hallin, M. Risling, Distribution of human pacinian corpuscles in the hand: A cadaver study. *J. Hand Surg.* **23**, 370–372 (1998).

51. Y. Shao, H. Hu, Y. Visell, A wearable tactile sensor array for large area remote vibration sensing in the hand. arXiv:1908.08199 [cs.RO] (22 August 2019).

52. R. D. Blandford, K. S. Thorne, *Applications of Classical Physics* (California Institute of Technology, 2003).

53. T. J. Moore, A survey of the mechanical characteristics of skin and tissue in response to vibratory stimulation. *IEEE Trans. Man Mach. Syst.* **11**, 79–84 (1970).

54. H. Azhari, *Basics of Biomedical Ultrasound for Engineers* (John Wiley & Sons, 2010).

55. J. Achenbach, *Wave Propagation in Elastic Solids* (Elsevier, 2012), vol. 16.

56. J. M. Pereira, J. M. Mansour, B. R. Davis, Dynamic measurement of the viscoelastic properties of skin. *J. Biomech.* **24**, 157–162 (1991).

57. J.-X. Liu, D. Wang, Y.-L. Gao, C.-H. Zheng, Y. Xu, J. Yu, Regularized non-negative matrix factorization for identifying differentially expressed genes and clustering samples: A survey. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **15**, 974–987 (2018).

58. A. Sotiras, S. M. Resnick, C. Davatzikos, Finding imaging patterns of structural covariance via non-negative matrix factorization. *Neuroimage* **108**, 1–16 (2015).

59. T. T. Vu, B. Bigot, E. S. Chng, Speech enhancement using beamforming and non negative matrix factorization for robust speech recognition in the CHiME-3 challenge, in *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)* (IEEE, 2015), pp. 423–429.

60. D. D. Lee, H. S. Seung, Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).

61. D. J. Field, What is the goal of sensory coding? *Neural Comput.* **6**, 559–601 (1994).

62. Z. Qiu, T. Yao, T. Mei, Learning spatio-temporal representation with pseudo-3D residual networks, in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2017), pp. 5533–5541.

63. D. H. Kim, W. Baddar, J. Jang, Y. M. Ro, Multi-objective based spatio-temporal feature representation learning robust to expression intensity variations for facial expression recognition. *IEEE Trans. Affect. Comput.* **10**, 223–236 (2017).

64. J. Kim, K. P. Truong, G. Englebienne, V. Evers, Learning spectro-temporal features with 3D CNNs for speech emotion recognition, in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)* (IEEE, 2017), pp. 383–388.

65. P. Bashivan, I. Rish, M. Yeasin, N. Codella, Learning representations from eeg with deep recurrent-convolutional neural networks. arXiv:1511.06448 [cs.LG] (19 November 2015).

66. Y. Yang, E. Aminoff, M. Tarr, K. E. Robert, A state-space model of cross-region dynamic connectivity in MEG/EEG, in *Advances in Neural Information Processing Systems* (NIPS, 2016), pp. 1234–1242.

67. V. Ramanarayanan, L. Goldstein, S. S. Narayanan, Spatio-temporal articulatory movement primitives during speech production: Extraction, interpretation, and validation. *J. Acoust. Soc. Am.* **134**, 1378–1394 (2013).

Citation: Y. Shao, V. Hayward, Y. Visell, Compression of dynamic tactile information in the human hand. *Sci. Adv.* **6**, eaaz1158 (2020).

Science Advances

Compression of dynamic tactile information in the human hand

Yitian Shao, Vincent Hayward and Yon Visell