**2025 IEEE World Haptics Conference (WHC)**
Suwon Convention Center, Suwon, Korea
July 8 ~ 11, 2025

# Classification of Vibrations Produced During Skin-on-Skin Versus Object-on-Object Tactile Gestures

Alexandra de Lagarde[1] *, Thomas Daunizeau[1,2] *, and Malika Auvray[3]

## I. Introduction

The sonication of movement on different surfaces revealed how sounds can convey rich information about tactile gestures, textures, and intentions [1], opening new paths to explore tactile perception. To investigate the case of skin-on-skin touch sonification, a recent study recorded the auditory information produced by tactile gestures [2]. Participants presented with these sounds were able to recognize both the different tactile gestures and their underlying socio-emotional intentions through their auditory conversion. In addition, their perception varied depending on whether the gestures occurred on human skins or on inanimate objects, sometimes favoring one, sometimes the other. However, this promising result remains silent on the acoustic cues driving human perception. To fill this gap, the present study used a classification approach on the vibrations produced during tactile gestures on different surfaces to investigate the extent to which gestures on the skin versus on objects can be reliably distinguished based on vibrations, both cutaneous and aerial.

## II. Materials and Methods

### A. Experimental setup

Six categories of tactile gestures illustrated in Fig. 1.A - hitting, tapping, punching, stroking, rubbing, and static contact - were performed under two surface conditions: skin-on-skin and cardboard-on-cardboard. Four individuals were asked to perform the gestures and were given a brief definition of each gesture before the session (e.g. “rubbing means a back-and-forth sliding of the hand on the forearm”). This was done to avoid confusion between gesture category while leaving them enough freedom to add variability within each gesture category. In the skin-on-skin condition, participants used their dominant hand on a forearm, either their own or someone else’s. In the cardboard-on-cardboard condition, participants used two out of three pieces available, each of the same material but of different density. The acoustic properties of these interactions were recorded in an anechoic chamber, using a dual microphone setup consisting of a piezoelectric contact transducer and an omnidirectional aerial microphone, with recordings sampled at $44.1\,\mathrm{kHz}$ (see Fig. 1.B). This setup resulted in a total of 1194 auditory stimuli. Each stimulus corresponds to one occurrence of a tactile interaction and lasts between $0.15\,\mathrm{s}$ and $2.8\,\mathrm{s}$, as exemplified in Fig. 1.C.

### B. Data analysis

A metric multidimensional scaling (MDS) was applied to reduce the dimensionality of our large dataset. First, a dissimilarity matrix was applied for all pairwise comparisons using dynamic time warping (DTW), a method commonly used in audio recognition. Unlike Euclidean distance, DTW can accommodate features that occur over different durations and time scales, making it well suited to capture the variability of unconstrained human tactile gestures. To limit the computational load, signals were downsampled by a factor of ten and then Z-score were normalized to emphasize shape over magnitude. Based on the stress criterion, a two-dimensional projection proved sufficient. Then, the 5% of outliers previously identified were removed via the Mahalanobis distance and overlaid shaded convex hulls for each recording condition.

## III. Results

The MDS results are shown in Fig. 2. Tactile gestures with intermittent contact (tapping, hitting, punching) form a more compact dataset than those with continuous contact

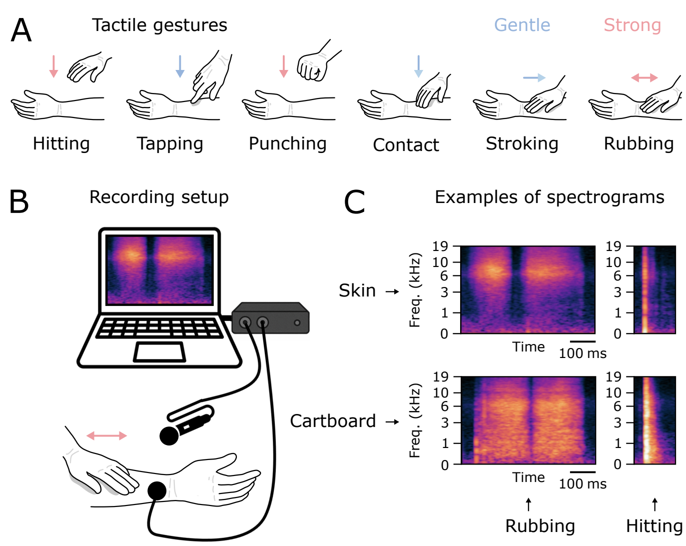


Fig. 1. Illustration of the setup used for touch sonification. A) Categories of tactile gestures. B) Combination of a piezoelectric contact transducer and an omnidirectional aerial microphone used to record vibrations produced by different tactile gestures. C) Examples of raw spectrograms which express frequency over time, with a $44.1\,\mathrm{kHz}$ sampling rate.

*Co-first authors.
[1]A. de Lagarde and T. Daunizeau are with the Institute of Intelligent Systems and Robotics (ISIR), Sorbonne University, F-75005 Paris, FRANCE.
[2]Current affiliation: École Polytechnique Fédérale de Lausanne (EPFL), CH-2000 Neuchâtel, SWITZERLAND.
[3]M. Auvray is with CNRS, the Institute of Intelligent Systems and Robotics (ISIR), Sorbonne University, F-75005 Paris, FRANCE.
e-mail: `malika.auvray@sorbonne-universite.fr`

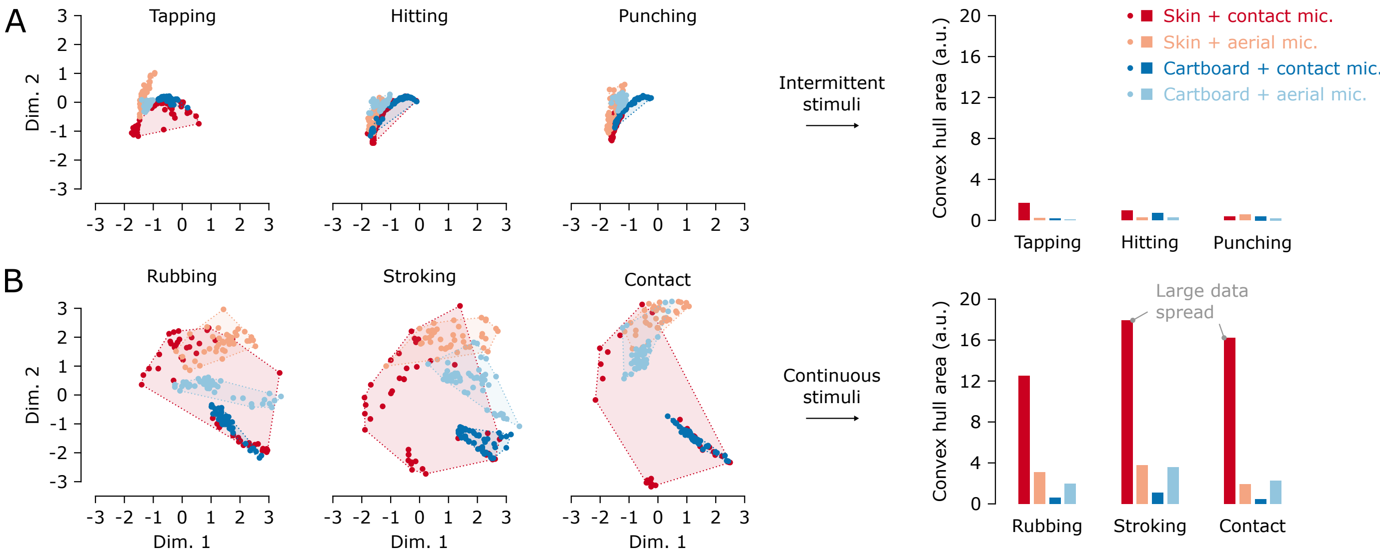


Fig. 2. Results obtained for A) three intermittent stimuli (tapping, hitting, punching) and B) three continuous stimuli (rubbing, stroking, contact), showing the differences as a function of the surfaces (skin in red and orange, cardboard in dark blue and light blue).

(rubbing, stroking, contact), as evidenced by their convex-hull areas. This pattern holds for both skin and cardboard contacts, indicating a shared underlying mechanism, possibly the $1/f$ frictional noise characteristic of Coulombic interactions [3]. Interestingly, skin-borne elastic vibrations recorded with the contact microphone exhibit the greatest scatter. This reflects the complex viscoelastic response of tissues, which depends heavily on subtle variations in the input stimulus. This variability is reflected in humans recognition responses [2], indicating underlying perceptual invariants. In addition, airborne vibrations for both skin and cardboard form a distinct group from contact signals, indicating that auditory feedback may enhance tactile information captured by mechanoreceptors.

## IV. Discussion

Tactile gestures are used to convey emotions and affective intentions to other humans. Each of these gestures has specific physical features such as applied pressure or velocity that can be recorded [4]. Our study allows the identification of structural differences in the acoustic signatures of gestures based on movement (intermittent contact or continuous friction) and surface dynamics (skin or cardboard). These characteristics allow highlighting why people in a previous study were able to identify different gestures without confusing intermittent and continuous contact. They also allow understanding why different results were obtained as a function of the surface involved (skin vs. inanimate object) [2]. The distinct grouping of airborne versus contact-based signals further reinforces the value of audition in conveying tactile information, especially in contexts where direct physical contact is not possible.

An abundant literature on sensory conversion revealed people's surprising abilities to make sense of information from one sensory modality when it is converted into another sensory modality [5], [6]. Our possibility to convert social touch into sounds sets the grounds for the opportunity to provide additional information on social touch at a distance. This would be of use in a broad diversity of remote experiences, for instance to convey richer socio-emotional intentions during human-to-human distant communications in the hope of compensating for the detrimental consequences of social isolation and lack of social touch. It may also enhance multisensory experiences to make gameplay more immersive or interactions with virtual agents more expressive, offering new possibilities in entertainment and therapeutic contexts.

## V. Acknowledgments

Project funded by the French ANR (project MATCH ANR-22-CE33-0010) and the French ANRT (CIFRE with Humans Matter, France).

## References

[1] A. Tajadura-Jiménez, N. Bianchi-Berthouze, E. Furfaro, and F. Bevilacqua, "Sonification of surface tapping changes behavior, surface perception, and emotion," *IEEE MultiMedia*, vol. 22, no. 1, pp. 48–57, Jan.-Mar. 2015.

[2] A. de Lagarde, C. Pelachaud, L. P. Kirsch, and M. Auvray, "Paving the way for social touch at a distance: Sonifying tactile interactions and their underlying emotions," *Proceedings of the National Academy of Sciences*, vol. 122, no. 19, p. e2407614122, May 2025.

[3] M. Wiertlewski, C. Hudin, and V. Hayward, "On the 1/f noise and non-integer harmonic decay of the interaction of a finger sliding on flat and sinusoidal surfaces," in *2011 IEEE World Haptics Conference*. Istanbul: IEEE, Jun. 2011, pp. 25–30.

[4] S. McIntyre, S. C. Hauser, A. Kusztor, R. Boehme, A. Moungou, P. M. Isager, L. Homman, G. Novembre, S. S. Nagi, A. Israr, E. A. Lumpkin, F. Abnousi, G. J. Gerling, and H. Olausson, "The language of social touch is intuitive and quantifiable," *Psychological science*, vol. 33, no. 9, pp. 1477–1494, Sep. 2022.

[5] L. P. Kirsch, X. Job, and M. Auvray, "Mixing up the senses: Sensory substitution is not a form of artificially induced synaesthesia," *Multisensory research*, vol. 34, no. 3, pp. 297–322, Jul. 2020.

[6] M. Auvray, "Sensory substitution," *Open Encyclopedia of Cognitive Science*, Feb. 2025.