



**HAL**  
open science

# Recognition of Gestures over Textiles with Acoustic Signatures

Pui Chung Wong, Christian Sandor, Alvaro Cassinelli

► **To cite this version:**

Pui Chung Wong, Christian Sandor, Alvaro Cassinelli. Recognition of Gestures over Textiles with Acoustic Signatures. SA '21: SIGGRAPH Asia 2021, Dec 2021, Tokyo Japan, France. pp.1-2, 10.1145/3476122.3484835 . hal-04358194

**HAL Id: hal-04358194**

**<https://hal.science/hal-04358194>**

Submitted on 21 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

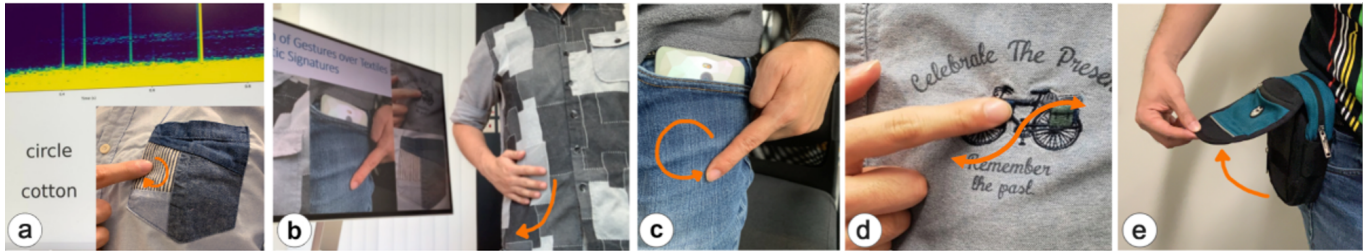


# Recognition of Gestures over Textiles with Acoustic Signatures

Pui Chung Wong  
tonywongpc@gmail.com  
City University of Hong Kong  
Hong Kong, China

Christian Sandor  
csandor@cityu.edu.hk  
City University of Hong Kong  
Hong Kong, China

Alvaro Cassinelli  
acassine@cityu.edu.hk  
City University of Hong Kong  
Hong Kong, China



**Figure 1:** a) An acoustic signature suffices for simultaneously recognizing the surface and the gesture performed over it. b) Presentation slides change by a hand swiping on a shirt; c) When the microphone is close to the fabric, recognition is possible even in a noisy environment. d) Textile patches can be exploited opportunistically. e) Interaction between certain types of fabric produce unmistakable acoustic signatures even if the signal travels over the air (here a hook-and-loop fastener).

## ABSTRACT

We demonstrate a method capable of turning textured surfaces (in particular textile patches) into opportunistic input interfaces thanks to a machine learning model pre-trained on acoustic signals generated by scratching different fabrics. A single short audio recording is then sufficient to characterize both a gesture and the textures substrate. The sensing method does not require intervention on the fabric (no special coating, additional sensors or wires). It is passive (no acoustic or RF signal injected) and works well using regular microphones, such as those embedded in smartphones. Our prototype yields 93.86% accuracy on simultaneous texture/gesture recognition on a test matrix formed by eight textures and eight gestures as long as the microphone is close enough (e.g. under the fabric), or when the patch is attached to a solid body transmitting sound waves. Preliminary results also show that the system recognizes the manipulation of Velcro straps, zippers, or the taping or scratching of plastic cloth buttons over the air when the microphone is in personal space. This research paves the way for a fruitful collaboration between wearables researchers and fashion designers that could lead to an interaction dictionary for common textile patterns or guidelines for the design of signature-robust stitched patches not compromising aesthetic elements.

## ACM Reference Format:

Pui Chung Wong, Christian Sandor, and Alvaro Cassinelli. 2021. Recognition of Gestures over Textiles with Acoustic Signatures. In *SIGGRAPH Asia 2021 Emerging Technologies (SA '21 Emerging Technologies)*, December 14-17, 2021. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3476122.3484835>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*SA '21 Emerging Technologies*, December 14-17, 2021, Tokyo, Japan

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8685-2/21/12.

<https://doi.org/10.1145/3476122.3484835>

## 1 INTRODUCTION

Since microphones are ubiquitous today, it make sense to consider passive acoustic sensing for activity recognition, including gestures while scratching textured surfaces. Leaving aside gesture recognition over everyday objects, we focus here on textiles, as these present an opportunity for design. A textile patch or a screen printed shirt generate a recognizable "acoustic signature" when scratched that can be mapped to simple commands for smartphones or home appliances (complementing voice-based smart home assistants). One can also envision an augmented wardrobe, giving a functional meaning to dressing codes: for instance recruiting a button-down shirt or a textured tie of formal business attires to generate workplace-specific commands.

## 2 PREVIOUS WORKS

There are various ways to make daily objects interactive but most are invasive (e.g., attaching widgets [Poupyrev et al. 2016], overlaying printed electronic, integrating inertial sensors [Laput et al. 2016], or using the speaker-microphone pair as an active sonar [Nandakumar et al. 2016]). Passive acoustic sensing does not requires any modification of the object as it relies on its naturally generated sounds. Recent advances in machine learning has opened wide this avenue. Among many others, Becker et al. [Xu et al. 2020] and Xu et al. [Becker et al. 2019] show 95% and 97% recognition of on-face gesture and sound-emitting gestures respectively using Convolutional Neural Net (CNN). Previous efforts relying on passive audio signatures mainly focus on input on planar surfaces (e.g. table and wall). We complement and extend this line of research by investigating the *simultaneous* recognition of gestures and substrate on patterned or non-patterned textiles.

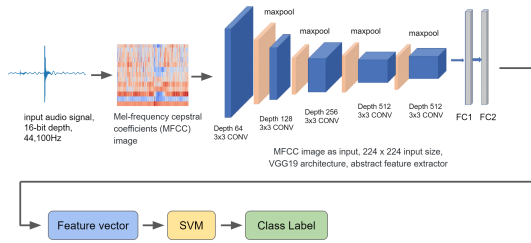


Figure 2: Classification pipeline

### 3 EXPERIMENTAL EVALUATION

#### 3.1 Algorithmic pipeline

Figure 2 shows our classification pipeline. First, the Mel-frequency cepstral coefficients (MFCC) image is extracted from the sample audio signal (16-bit depth, 48kHz, 1.4 seconds). To reduce training time and possible over-fitting for small data sets, we use a pre-trained VGG-19 model [Simonyan and Zisserman 2014] to then extract abstract features of MFCC images, that become the feature vector to classify. Our data samples are split 70/30 into a training and testing sets to train a support vector machine (SVM) for classification. We explore three conditions: 1) texture recognition, 2) texture-gesture recognition, 3) interaction recognition on an object.

#### 3.2 Data collection and augmentation.

With the microphone placed at close range (10cm) or under the surface, we collected 60 samples per condition. For texture recognition, we collected signals generated by sliding the index finger over cotton, linen, nylon and polyester, as well as some materials resembling parts of an attire (foam, tin, polyethylene bubble wrap, wood, cardboard, and paper). For texture/gesture recognition, we collected signals generated by performing eight gestures (swipe, circle, triangle, rectangle, tick, cross, rub, and star) on eight textures. We also trained the system on ten interactions with a waist bag (opening and closing the lid and zippers, tapping and swiping on the lid surface and interaction with the buckle and waist clip). To improve generalization, we perform time shifting and noise augmentation before converting samples to MFCC images. Time shifting creates new synthetic training samples and increases data size by 50%. To simulate real-world scenarios, street and office noise were injected into the samples with a 50% probability.

#### 3.3 Experiment Results.

We trained the models and tested accuracy with data collected under the three conditions described above, obtaining respectively: 97.22% for texture recognition alone; an overall test accuracy of 93.86% for texture/gesture recognition; and 93.89% for interaction recognition on a waist bag. Figure 3 shows the associated confusion matrices.

### 4 DISCUSSION AND FUTURE WORKS

Despite artificial noise injection, our system has been trained on a relatively quiet environment and may be too environment specific. To which extent one can leverage on classic noise cancellation algorithms for speech processing is an open question since the

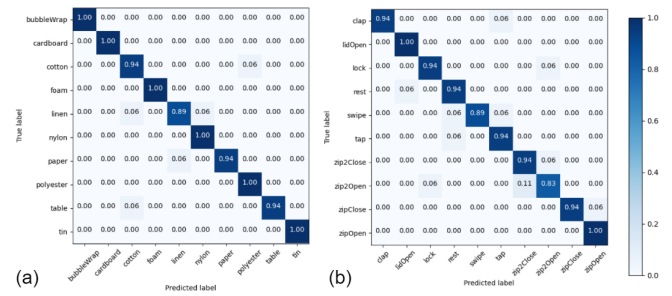


Figure 3: Confusion matrices for (a) textures and (b) waist bag interaction.

power spectrum of the scratching sound is much more similar to environmental noise than voice). Site-specific training (which in our preliminary study shows an accuracy surpassing 90%) would be too cumbersome. An interesting solution worth exploring is to pre-train the model in a noiseless environment, and then rely on transfer learning to reduce to a minimum the site-specific training phase [Carney et al. 2020]. Interestingly, this would also allow users to re-calibrate the system or even reconfigure their “textile functions” on the flight, recording site-specific models (workplace, home, street, etc) that they can invoke later, or be selected automatically by recognizing the environment from noise or by any other classic context aware computing strategy. Informed by textile design techniques, future work will explore ways to modify or create robust patterns for recognizing a set of standard gestures (directional swipes, taps, etc). These can be integral to the clothing, or presented as stitchable ornaments for T-shirts, cushions, tablecloth, etc. Such strategy for interactive cloth design will respect aesthetics while at the same time supporting ubiquitous computing.

### REFERENCES

- Vincent Becker, Linus Fessler, and Gábor Sörös. 2019. GestEar: Combining Audio and Motion Sensing for Gesture Recognition on Smartwatches. In *Proceedings of the 23rd International Symposium on Wearable Computers (London, United Kingdom) (ISWC '19)*. Association for Computing Machinery, New York, NY, USA, 10–19. <https://doi.org/10.1145/3341163.3347735>
- Michelle Carney, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen. 2020. Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3334480.3382839>
- Gierad Laput, Robert Xiao, and Chris Harrison. 2016. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (Tokyo, Japan) (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 321–333. <https://doi.org/10.1145/2984511.2984582>
- Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, and Shyamnath Gollakota. 2016. FingerIO: Using Active Sonar for Fine-Grained Finger Tracking. Association for Computing Machinery, New York, NY, USA, 1515–1525. <https://doi.org/10.1145/2858036.2858580>
- Ivan Poupyrev, Nan-Wei Gong, Shiho Fukuhara, Mustafa Emre Karagozler, Carsten Schwesig, and Karen E. Robinson. 2016. *Project Jacquard: Interactive Digital Textiles at Scale*. Association for Computing Machinery, New York, NY, USA, 4216–4227. <https://doi.org/10.1145/2858036.2858176>
- Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
- Xuhai Xu, Haitian Shi, Xin Yi, WenJia Liu, Yukang Yan, Yuanchun Shi, Alex Mariakakis, Jennifer Mankoff, and Anind K. Dey. 2020. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376836>