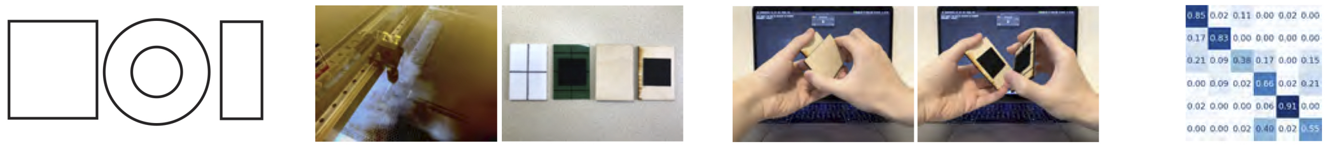
Let It Rip! Using Velcro for Acoustic Labeling

Tzu-Sheng Kuo, Eric Rawn
Stanford University
{tskuo, erawn}@stanford.edu



a) design labels b) cut and mount labels c) record Velcro rips d) classify labels

Figure 1. Our acoustic labeling process comprises a) designing label shapes using a SVG editor b) cutting the labels from the paired sides of Velcro strips and attaching them to surfaces c) recording the sound generated by separating the two sides d) classifying the labels based on their audio signal.

ABSTRACT
We present an early stage prototype of an acoustic labeling system using Velcro, a two-sided household adhesive product. We create labels by varying the shape of Velcro pieces to produce distinct sounds when the two sides are separated, and we use an automatic audio classification pipeline to detect and classify small sets of labels. We evaluate our classifier on four sets of three simple Velcro labels, present a demo highlighting potential use cases of these labels, and discuss future applications.

Author Keywords
Velcro; acoustic; labels; sound; activity sensing.

CCS Concepts
•Human-centered computing → Ubiquitous and mobile computing systems and tools;

INTRODUCTION
Sensing and making use of the signals present in everyday life creates potential for novel experiences. Prior work has explored the use of airborne electromagnetic noise [15], RF waves [13], physical vibrations [14], pressure variations [1], multi-modal signals [10], and even real-time crowdsourcing [9] for sensing human movements and appliances. Recently, researchers have found that sounds can also be used to recognize activities such as a human laugh or a vacuum in-use [8]. However, since not every action and object produces a sound on its own, alternative methods modify objects to produce audible cues, such as 3D-printed levers and etched lines [12, 5]. In practice, many of these audio labeling mechanisms require specialized tools to produce and create permanent changes to objects, which may not be desirable in certain applications.

UIST '20 Adjunct, October 20–23, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7515-3/20/10.
https://doi.org/10.1145/3379350.3416175

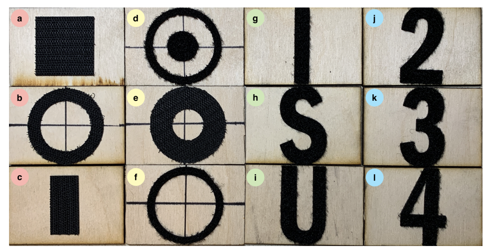


Figure 2. Our four sets of labels include the core set (a, b, c), the ring set (d, e, f), the letter set (g, h, i), and the number set (j, k, l).

We propose an acoustic labeling approach using compact, low-cost removable labels made from Velcro. We cut labels that vary in shape and size and show that these labels produce distinguishable sounds. We evaluate our automatic label classification pipeline on four sets of labels and show how these labels can be used to identify the removal of objects from a desk or wall in an example scenario. We then discuss further applications of this labeling method.

LET IT RIP: SYSTEM DESIGN
Our process begins by cutting commercially available Velcro into different shapes to create labels. Each label comes as a pair with mirrored geometry so that the two sides (the hook and loop sides) align when pressed together.

Cutting and Assembling Velcro
We design our labels in Adobe Illustrator and generate a SVG file from outline shapes. We then use a Brightstar Advantage-24 Laser Cutter to cut the shapes from large Velcro pieces into individual labels. While we use a laser cutter for precision, the labels we test are not particularly complex and could be cut carefully using scissors. For collecting sound data, we use stiff backing support slides to which we attach the labels. Each slide measures 2×3 inches, and we use two different materials: 1/8-inch thick plywood and acrylic. These two materials are representative of the different surfaces our Velcro may be adhered to for various applications.

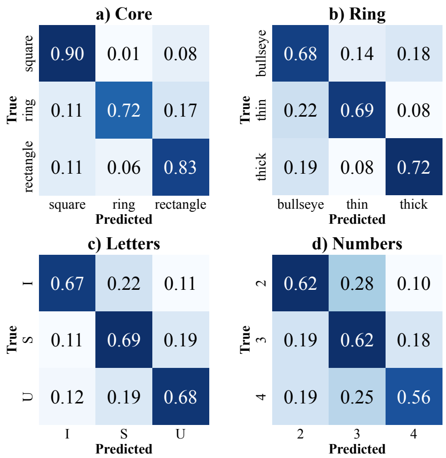


Figure 3. We evaluate our system on four sets of labels with an average accuracy of 82%, 70%, 68%, and 60% for the a) core, b) ring, c) letters, and d) numbers, respectively.

Recording and Processing Audio

To record our audio data, we use the built-in microphone of 2019 Macbook Pro Laptop with a sampling rate of 48 kHz. We record batches of 10-50 consecutive rips and use PyAudioAnalysis [3] to split them into individual rips by automatic silence detection. For each rip we compute five audio features – MFCC, chroma, mel spectrogram, spectral contrast, and tonnetz. Together these features describe the timbre, tone, and harmonic changes in the signal [11, 6, 4]. We then normalize these features across each dimension and feed them into a support vector machine (SVM) [2] for classification.

Classifier

To train the SVM, one author recorded 200 rips with 12 label shapes (Fig. 2) on two backing materials (24 slides total). For testing, four participants recorded 10 rips with each 24 slides. In total, our training and testing data consists of 4800 and 960 rips, respectively. Since the training and testing data were recorded in different places with different people, a covariate shift exists between the datasets. To account for this shift, we remove one testing rip for each label and add it into the training set. This process may be considered as a one-step calibration in real use cases.

EVALUATION

In our exploration of the feasibility of Velcro labels, we evaluate 12 labels, divided into four sets of three labels shown in Fig. 2. Each of these sets has different strengths and potential applications. Our evaluation results are shown in Fig. 3.

Our core set of labels contains three geometric shapes – square, rectangle, and ring – each with a maximum dimension of 2-inches. These labels vary noticeably in shape and area and produce consistent sounds when ripped. Our classifier achieves 90%, 83%, and 72% accuracy on three core shapes.

In some applications, limited space may be available for attaching labels. In this case, it is beneficial to have labels with identical outer boundaries but varying internal geometries that emit different sounds. We test a set of ring-shaped labels that all fit within a 2-inch diameter circular region. Our classifier is able to distinguish them at an overall accuracy of 70%.

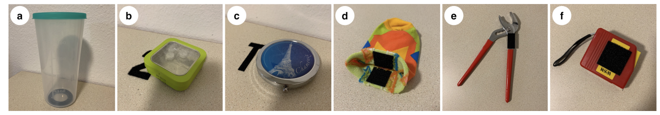


Figure 4. Our example application uses six labels and objects a) a cup with the ring label, b) mints with the 2 label, c) a mirror with a T label, d) a sock with a small square label, e) a wrench with a rectangle label, and f) a tape measure with a big square label.

For other applications, the appearance of labels themselves could be used as design features for the objects to which they are attached. We create labels with recognizable digits and letters to annotate objects both visibly and acoustically. We chose three letters and three digits for geometric diversity, and our classifier achieved 68% and 60% accuracy, respectively.

DEMO AND APPLICATIONS

As an application of our system, we take advantage of both the adhesive and audio properties of Velcro to show how our system could be used to track the removal of objects from their original location. We attach six labels to six objects on a desk or wall to evaluate the performance of the classifier applied to everyday objects (Fig.4). When any object is removed from its position, it emits an audible rip identified by our system.

For evaluation, we use the same training data recorded from our main study, and the same author record 50 rips with each of the six objects as testing data. Since the labels are attached to very different objects between the training and testing sets, we augment the training data with 3 new recorded rips on the actual objects to account for the covariate shift. We then calculate the accuracy of the classification using the remaining 47 rips for each of the six objects.

Our system achieve accuracy greater than 80% for three of our six labels, and greater than 50% accuracy on two additional labels. Although most of our original training data comes from the controlled lab study recorded with wood and acrylic slides, our classification pipeline is still able to classify labels when affixed to real objects with very different material properties, such as socks, and with very different shapes, like cups.

CONCLUSION AND FUTURE WORK

We presented an early stage prototype for using Velcro for acoustic labeling. We performed an initial analysis and demonstrated an application using household objects. Further development involves building a comprehensive vocabulary of Velcro patterns which are reliably distinguishable by our classifier, investigating the consistency of the classifier with respect to different speeds and directions of detachment, and enhancing the classifier with a superior machine learning model [7]. We are also interested in exploring other application scenarios, such as customizing the input devices of smart textiles or low-cost hardware by attaching Velcro to fabrics or cardboard and sensing its sound with mobile phones or gaming consoles. We look forward to leveraging off-the-shelf materials to make computers smarter and more aware of their surroundings.

ACKNOWLEDGEMENTS

We would like to thank Prof. Michael Bernstein and Evan Strasnick for their guidance. In addition, we thank Mackenzie Leake for her helpful discussion and assistance on the project.

REFERENCES

[1] Tim Campbell, Eric Larson, Gabe Cohn, Jon Froehlich, Ramses Alcaide, and Shwetak N. Patel. 2010. WATTR: A Method for Self-powered Wireless Sensing of Water Activity in the Home. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing (UbiComp '10)*. ACM, New York, NY, USA, 169–172. DOI:http://dx.doi.org/10.1145/1864349.1864378

[2] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* 2, 3, Article 27 (May 2011), 27 pages. DOI:http://dx.doi.org/10.1145/1961189.1961199

[3] Theodoros Giannakopoulos. 2015. pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PLOS ONE* 10, 12 (12 2015), 1–17. DOI: http://dx.doi.org/10.1371/journal.pone.0144610

[4] Emilia Gómez and Perfecto Herrera. 2004. Estimating The Tonality Of Polyphonic Audio Files Cognitive Versus Machine Learning Modelling Strategies. In *5th International Society for Music Information Retrieval (ISMIR) Conference*. 92–95. http://mtg.upf.edu/system/files/publications/G%C3%B3mezHerrera-ISMIR-2004-page-92.pdf

[5] Chris Harrison, Robert Xiao, and Scott Hudson. 2012. Acoustic Barcodes: Passive, Durable and Inexpensive Notched Identification Tags. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. Association for Computing Machinery, New York, NY, USA, 563–568. DOI: http://dx.doi.org/10.1145/2380116.2380187

[6] Christopher Harte, Mark Sandler, and Martin Gasser. 2006. Detecting Harmonic Change in Musical Audio. In *Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia (AMCMM '06)*. Association for Computing Machinery, New York, NY, USA, 21–26. DOI: http://dx.doi.org/10.1145/1178723.1178727

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778.

[8] Gierad Laput, Karan Ahuja, Mayank Goel, and Chris Harrison. 2018. Ubicoustics: Plug-and-Play Acoustic Activity Recognition. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 213–224. DOI: http://dx.doi.org/10.1145/3242587.3242609

[9] Gierad Laput, Walter S. Lasecki, Jason Wiese, Robert Xiao, Jeffrey P. Bigham, and Chris Harrison. 2015. Zensors: Adaptive, Rapidly Deployable, Human-Intelligent Sensor Feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1935–1944. DOI: http://dx.doi.org/10.1145/2702123.2702416

[10] Gierad Laput, Yang Zhang, and Chris Harrison. 2017. Synthetic Sensors: Towards General-Purpose Sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3986–3999. DOI: http://dx.doi.org/10.1145/3025453.3025773

[11] Tom LH. Li and Antoni B. Chan. 2011. Genre Classification and the Invariance of MFCC Features to Key and Tempo. In *Proceedings of the 17th International Conference on Advances in Multimedia Modeling - Volume Part I (MMM'11)*. Springer-Verlag, Berlin, Heidelberg, 317–327.

[12] Valkyrie Savage, Andrew Head, Björn Hartmann, Dan B. Goldman, Gautham Mysore, and Wilmot Li. 2015. Lamello: Passive Acoustic Sensing for Tangible Input Components. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1277–1280. DOI: http://dx.doi.org/10.1145/2702123.2702207

[13] Yang Zhang, Yasha Iravantchi, Haojian Jin, Swarun Kumar, and Chris Harrison. 2019. Sozu: Self-Powered Radio Tags for Building-Scale Activity Sensing. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 973–985. DOI: http://dx.doi.org/10.1145/3332165.3347952

[14] Yang Zhang, Gierad Laput, and Chris Harrison. 2018a. Vibrosight: Long-Range Vibrometry for Smart Environment Sensing. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 225–236. DOI: http://dx.doi.org/10.1145/3242587.3242608

[15] Yang Zhang, Chouchang (Jack) Yang, Scott E. Hudson, Chris Harrison, and Alanson Sample. 2018b. Wall++: Room-Scale Interactive and Context-Aware Sensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–15. DOI:http://dx.doi.org/10.1145/3173574.3173847