TapStrapGest: Elicitation and Recognition of Ring-based Multi-Finger Gestures

GUILLEM CORNELLA and EUDALD SANGENIS, University of California at Irvine, USA
MEHDI OUSMER and JEAN VANDERDONCKT, Université catholique de Louvain, Belgium
SANTIAGO VILLARREAL-NARVAEZ and BRUNO DUMAS, University of Namur, Belgium
ADRIEN CHAFFANGEON CAILLET, Berlin University of Applied Sciences and Technology (BHT), Germany and Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, France

Fig. 1. Usages of ring-based multi-finger gestures: (a) tapping on a hard surface, (b) tapping on a soft surface, (c) tapping in virtual reality, and (d) collaborative mid-air gestures.

We introduce TapStrapGest, a novel solution for customizable ring-based multi-finger gestures, encompassing the process from gesture elicitation to gesture recognition. Recognizing the growing demand for intuitive and customizable gesture interaction with fingers, TapStrapGest uses Tap Strap to enable users to perform simple and complex multi-finger gestures using smart rings. We conducted a gesture elicitation study, detailing the systematic process of soliciting and refining a custom set of user-defined ring-based finger gestures through participatory design and ergonomic considerations, including thinking time, goodness of fit, and memorization. Subsequently, we delve into the technical underpinnings of gesture recognition. We reduce the dimensionality of a dataset of 27 gesture classes from 21 to 15 by filtering, then from 15 to 5 by a Principal Component Analysis. We implement and compare four machine learning algorithms to show that a Quadratic Discriminant Analysis (precision=99.33%, recall=99.26%, and F1-score=99.26%) outperforms three other machine learning classifiers, *i.e.,* a Linear Discriminant Analysis, a Support Vector Machines, and a Random Forest, as well as existing recognizers from the literature, to accurately recognize such gestures without the need to call for Deep Learning. Through a performance analysis, we demonstrate that TapStrapGest is a versatile and admissible solution for ring-based multi-finger gesture interaction, opening avenues for "eyes-free" or "screen-free" human-computer interaction in various domains.

Authors' Contact Information: Guillem Cornella, cornellg@uci.edu; Eudald Sangenis, esangeni@uci.edu, University of California at Irvine, Irvine, USA; Mehdi Ousmer, mehdi.ousmer@uclouvain.be; Jean Vanderdonckt, jean.vanderdonckt@uclouvain.be, Université catholique de Louvain, Louvain Research Institute in Management and Organizations (LouRIM), Louvain-la-Neuve, Belgium; Santiago Villarreal-Narvaez, santiago.villarreal@unamur.be; Bruno Dumas, bruno.dumas@unamur.be, University of Namur, Namur Digital Institute (NADI), Namur, Belgium; Adrien Chaffangeon Caillet, Berlin University of Applied Sciences and Technology (BHT), Berlin, Germany and Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, Grenoble, France, adrien.chaffangeon@univ-grenoble-alpes.fr.

ACM 2573-0142/2025/6-ARTEICS001
https://doi.org/10.1145/3733047

CCS Concepts: • **Hardware** → **Sensor devices and platforms**; **Tactile and hand-based interfaces**; • **Human-centered computing** → *Interaction devices*; Laboratory experiments; *Graphical user interfaces*; Virtual reality; **Gestural input**; *Mobile devices*; Smartphones; **Participatory design**; • **Theory of computation** → *Support vector machines*; • **Computing methodologies** → *Machine learning approaches*; *Machine learning algorithms*.

Additional Key Words and Phrases: Finger tapping, Gesture elicitation study, Gesture input, Gesture recognition, Multi-fingers gestures. Quadratic Discriminant Analysis, Random forest, Support Vector Machines, Tap gestures

ACM Reference Format:
Guillem Cornella, Eudald Sangenis, Mehdi Ousmer, Jean Vanderdonckt, Santiago Villarreal-Narvaez, Bruno Dumas, and Adrien Chaffangeon Caillet. 2025. TapStrapGest: Elicitation and Recognition of Ring-based Multi-Finger Gestures. *Proc. ACM Hum.-Comput. Interact.* 9, 4, Article EICS001 (June 2025), 27 pages. https://doi.org/10.1145/3733047

1 INTRODUCTION

In recent years, the landscape of gesture interaction [18] has seen a remarkable evolution in terms of interaction capabilities in various contexts of use [52, 66], driven by the introduction of new devices, such as wearables [22]. Among these innovations, ring-based finger gesture interaction [24, 41, 56] stands out as a promising avenue, offering end users a seamless and intuitive means of interacting with interactive applications. To make the most of their benefits, we need to delve into the realm of customizable ring-based finger gesture interaction to present an innovative approach with significant potential to improve user experience across various contexts of use. In this work, we consider finger gestures, a type of microgesture [7], as well as all gestures that require moving the fingers using the wrist and/or arm.

We do not consider here vision-based finger recognition solutions (*e.g.,* with a Kinect [11]) as they are usually constrained by environmental conditions, such as lighting, occlusion, and visibility, which can be detrimental to the recognition process. Instead, wearable devices [54], particularly smart rings [19–21] equipped with sensors and communication capabilities, emerge as a compelling solution. Using the dexterity and expressiveness of finger movements [43], these devices allow users to interact with interfaces in a way that reproduces or imitates real-world gestures.

Central to the effectiveness and efficiency of the interaction of ring-based finger gestures is its customizability [13, 40, 50]. Unlike designer- or system-predefined gestures [46], unmodifiable gestures that are usually shipped with the device, user-defined [17, 65], customizable gestures empower users to define their gestures or to replace existing ones by their own [13], tailored to their preferences and usage scenarios. This level of customization not only enhances user satisfaction, but also fosters a deeper sense of engagement and control over the interface. Furthermore, the integration of Machine Learning (ML) algorithms [36] and sensor fusion techniques should further increase the capabilities of ring-based finger gesture interaction [62].

For example, the Tap Strap 2 ring-based device for finger gestures benefits from an extensive gesture vocabulary that covers letters, digits, symbols, and some commands (*e.g.,* back, undo) that are very well operated and recognized, but this gesture vocabulary is not entirely customizable and therefore not recognizable. Although Tap Strap 2 offers a range of predefined gestures and mappings, its vocabulary may still be limited compared to other input methods, gesture-based or not [31, 43]. Tap Strap 2 supports 8 generic gesture-based commands, while the average number of referents used for commands in gesture vocabularies can be higher (20 in [59] or 18 in [58]). To incorporate gestures for new commands, this device enables end-users to define a configuration where gestures can be combined, but this overwrites the existing gesture vocabulary and does not take into account non-tapping gestures in the customization. Therefore, the recognition of non-standard gestures should in principle rely on ML algorithms [36, 44].

To address the above limitations, we aim to explore the design principles, implementation challenges, and potential applications of the ring-based multi-finger gesture interaction that can be customized. We believe that the customizable ring-based finger gesture interaction holds real promise in improving the way we interact with

Table 1. Goal-Question-Metrics formulation of our research on ring-based multi-finger gestures.

Goal	Purpose	Allow
	Issue	effective and efficient interaction through
	Object	ring-based multi-finger gestures
	Viewpoint	from the end user's viewpoint and the interactive application's viewpoint
Research question	RQ_1	What are the most preferred user-defined gestures?
Metrics	M_1	Agreement rate
	M_2	Preference rank
Research question	RQ_2	What are the best recognizers?
Metrics	M_3	Precision
	M_4	Recall
	M_5	F1-score

gesture-based interfaces. Following the Goal-Questions-Metrics (GQM) [4] framework, a hierarchical model that follows a top-down approach where a goal is first specified and associated with research questions collected and measured through metrics, this paper (Table 1) makes the following contributions:

- After reviewing related work in Section 2, Section 3 reports the results of a gesture elicitation study executed with the Tap Strap 2, considered as a representative instance of a ring-based multi-finger system, for 18 commands for file operations and window management, thus resulting in a new consensus set.
- Section 4 consolidates the gesture proposals in Section 3 into an overall classification that can be used widely.
- Section 5 performs a Principal Component Analysis to reduce the dimensionality of the gesture set from 15 to 5. A Quadratic Discriminant Analysis outperforms three machine learning classifiers, and existing recognizers,
- Section 6 analyzes the performance analysis and suggests some implications for ring-based multi-finger gestures.
- Section 7 shows how the gestures considered in this paper can be represented by a formal notation.

2 RELATED WORK

2.1 Ring-based Finger Gesture Interaction

2.1.1 Single Finger. A range of studies have explored the potential of ring-based gesture interaction with a single finger. For example, Gummeson et al. [24] developed a ring-based device to be worn on the index that recognizes gestures on any available surface with an accuracy of 73% (for 23 gesture classes that include tapping, swiping, scrolling, and stroking for hand-written text entry) and wirelessly transmits gestures to the remote device. [22] introduced the concept of "bi-digit interaction" using a ring worn on a thumb, which takes advantage of the unique biomechanical properties against the index. While the thumb and the index are involved, only the thumb gestures are subject to recognition. The potential of smart rings is also demonstrated in various contexts of use such as ambient intelligence [19], assistive technology for people with motor impairments [20], and smart rooms [21], suggesting various design directions for their use. Again, this smart ring is worn on the index, therefore supporting single-finger gestures.

2.1.2 Multiple Fingers. We hereby distinguish *one-to-many approaches* (where one ring detects multiple fingers) from *many-to-many approaches* (where many rings detect multiple fingers). In the first sub-category, Wilhelm et al. [63] presented the eRing, a prototype that uses electric field sensing to detect finger gestures based on a single ring worn on one finger. In the second sub-category, finger-worn accelerometers, such as the Tap Strap 2, enhance touch interaction by identifying different parts of the finger and passive tools [25].

These accelerometers are also cost-effective, allow for unrestricted movement, and provide accurate 3D acceleration measurements. Furthermore, wrist-worn accelerometers have been found to be valid and reliable for measuring physical activity in adults, with low participant burden [39]. Finally, the use of accelerometers in gloves has been explored for identifying static gestures and as a mouse pointing device.

The location where a finger is tapped is not needed by TapGazer [27], thereby enabling end users to tap without looking at their hands ("hands-free" interaction [45]) or at any screen ("screen-free" interaction [42]) or anywhere ("location-free" [69]). Tap detection is ensured with any input device capable of discerning which finger is currently being tapped, such as finger-worn accelerometers, touch-sensitive surfaces, such as smart clothes, or visual finger tracking like UltraLeap's Leap Motion Controller (LMC). However, a LMC recognizes a two-handed model very accurately for much richer gestures than tapping and swipe gestures.

PeriSense [62] is a ring-based multi-finger device for which the sensor resolution was assessed at different distances to recognize finger gestures and unistroke gestures based on some gesture sets. PeriSense was able to reliably sense the change in conductive objects up to 2.5 cm/0.98 inch. TypeAnywhere [69] combines two Tap Strap devices to obtain a QWERTY-based text entry system for use in any location by decoding typing sequences based only on finger-tap sequences without relying on tap locations. The Pinch Sensor [60] is an elastic input device that senses the fine motion and pinch force of the index finger and thumb, the two digits most used in human hands for in-hand object manipulation skills. In addition to opening and closing, the device would allow a user to control a robotic or simulated two-finger hand to reorient an object in three different ways and their combinations. TextFlow [30] consists of a contextually aware system that uses mixed-initiative techniques to produce auditory message options tailored to the location, activity, and time of day. TextFlow allows end-users to browse and choose from suggested auditory messages using finger taps using a finger-worn device, thus supporting screen-free interaction.

Recent advances in recognition technology have made camera-based text entry using American Sign Language (ASL) finger spelling more viable. However, there are instances where this method may not be optimal or acceptable. In response, Martin et al. [41] introduced FingerSpeller, a Tap Strap-based solution that enables text entry without the need for a camera, using smart rings. Employing a Hidden Markov Model (HMM)-based backend with continuous Gaussian modeling, FingerSpeller achieves precise recognition, as evidenced in real-world testing. In isolated word recognition experiments off-line with a 1,164-word dictionary, FingerSpeller achieves an average character accuracy of 91% and word accuracy of 87% across three participants. An accuracy of 90% can be maintained for this specific dataset in this specific context involving these participants by focusing on 2 fingers instead of 5 fingers for recognition". This is, of course, specific to the dataset. If a gesture differentiates all finger movements, this is no longer applicable and its recognition is affected [55]. More recently, Jomyo et al. [29] demonstrated that multi-finger gestures tapped on the skin can be recognized using wrist-mounted piezoelectric film sensors. This approach is particularly useful for body-deictic tapping gestures, but not for other types of gestures than tapping.

In sum, ring-based gesture interaction provides an unobtrusive way to interact, provided that a wide palette of single/multiple finger gestures can be recognized. Although most existing work has focused on text input [25, 27, 30, 41], a highly complex task when performed without a screen, we note the predominance of the use of one [22] or two fingers [60], often the index finger and thumb, which are the most mobile fingers [60], and less of all fingers for gestural interaction, with the exception of TapType [51] and TypeAnywhere [69] that involve all ten fingers for text entry. When all fingers are mobilized, other devices come into play but require computer

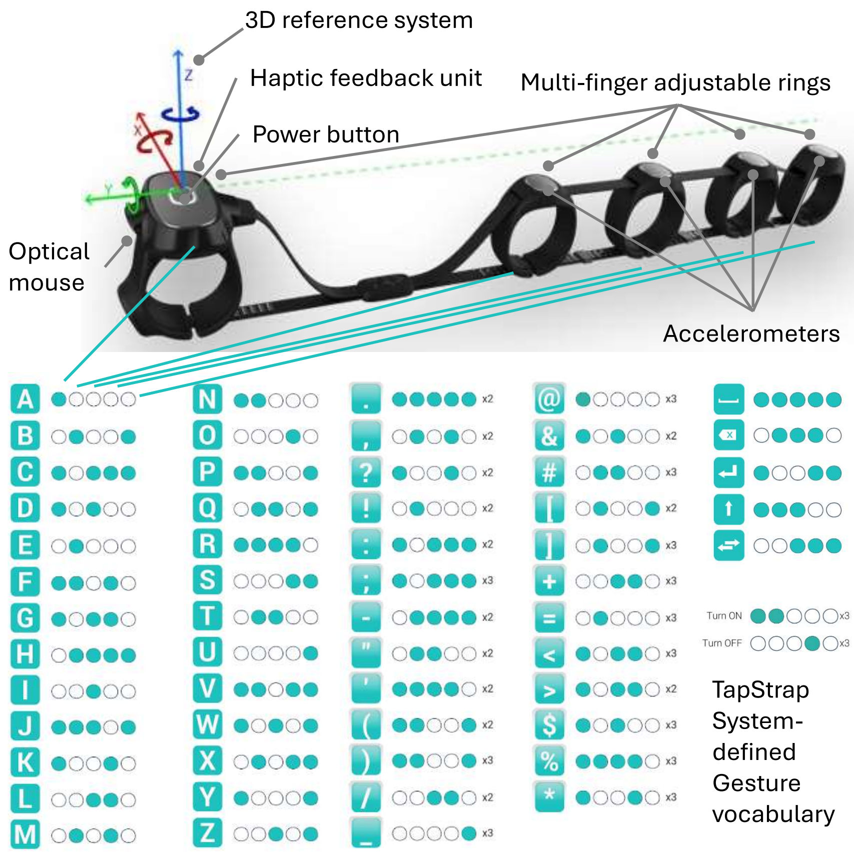

Fig. 2. How the Tap Strap 2 works to support ring-based multi-finger gestures and its vocabulary.

vision-based approaches. For example, TapXR recognizes finger movements well for gesture-based interaction but is again subject to good lighting conditions and reflection from the surface on which the fingers are placed. A light surface reflects finger position and movement less well than a dark surface, which negatively impacts gesture recognition. Based on these observations, we would like to investigate multiple rings for multi-finger gesture recognition, which is introduced in the next section.

2.2 Presentation and Evaluation of the Tap Strap 2

The Tap Strap 2 is a wearable device designed to transform any surface into a keyboard and mouse by featuring five finger rings equipped with accelerometers (Fig. 2) that detect finger movements and taps on a surface. This device enables end-users to input figures, letters, symbols, and control another device with finger tapping or simple gestures, eliminating the need for keyboards, mice, or other input devices. Tap Strap 2 utilizes Bluetooth connectivity to pair with any device such as a smartphone (Fig. 1-a), a tablet (Fig. 1-d), a computer (Fig. 1-c), and even virtual reality headsets (Fig. 1-c). Once paired, end-users tap their fingers on soft (*e.g.,* clothes in Fig. 1-b), irregular (*e.g.,* a sofa in Fig. 1-a) or hard (*e.g.,* a table in Fig. 1-c) surfaces or in mid-air (Fig. 1-d) to produce text or simple gestures that execute commands on the paired device. While the Tap Strap 2 consists of a one-handed wearable, a single user can wear one device on the dominant hand (Fig. 1-a), two devices on both hands to support bimanual input or two users (Fig. 1-d) can wear one or two devices to initiate a collaborative gesture. The Tap Strap 2 can be operated in five modes:

(1) *Keyboard Mode*: combine taps to type letters, numbers, symbols, and characters into the paired device.
(2) *AirMouse Mode*: input and control using close mid-air gestures into the paired device.
(3) *Optical Mouse Mode*: swipe the fingers to support navigation, selection, scrolling, dragging, and dropping in any software using any surface.
(4) *Controller Mode*: associate finger taps and air gesture swipes into a custom mapping that executes a command on the paired device.
(5) *Custom Mode*: use TapMapper web tool, the open source software development kit, its API or Unity plugin to engineer gesture input to the paired device.

The motivations to use a Tap Strap for gesture interaction stem from its portability, convenience, accessibility, productivity, and immersive capabilities. Whether the end user is on the go, in a crowded environment, or simply looking for a more intuitive way to interact, the Tap Strap offers unparalleled convenience and versatility. However, while ongoing development is already exploring the potential Tap Strap applications across various domains, ranging from general interaction and accessibility to gaming and immersive technologies, the research literature specifically dedicated to the Tap Strap still appears to be relatively limited regarding its capabilities for gesture interaction, beyond text entry and simple or dedicated commands. Dosdall et al. [16] used the chorded Tap Strap 2 keyboard as an input tool for Avar, a tool for the information visualization in augmented reality, for which a special mapping of the Tap Strap combinations was created for some commands programmed in Pharo. Tap4Light [3] controls Philips Hue smart lamps with 11 multi-finger (system-defined) tapping gestures captured with Tap Strap 2 in the context of ambient intelligence.

Using the accelerometers embedded in each ring, Tap Strap 2 can recognize which fingers are tapping at any given time, which makes it eligible for gesture recognition. Mrazek et al. [43] investigated what Tap Strap 2 is capable of by evaluating the usability of the device and by testing an automatic correct algorithm that reduces spelling errors. They found that performing 2 out of 10 drag drops and 5 out of 10 clicks. Simple combinations using one finger or consecutive fingers tend to be more accurate than keys such as 'q' and 'w', which require lifting the ring finger. Characters have varying recognition accuracy, ranging from the most accurate 'a' to the least accurate 'w'. There were instances where the right taps were produced but not properly recognized (true negatives [10]) or where almost right taps were produced and properly recognized (false positives [10]).

Videos demonstrate that the Tap Strap 2 works on both hard surfaces, such as a tabletop, and soft surfaces, such as a leg. But gesture recognition on a malleable surface seems less accurate than on a hard surface: the device does not register certain gestures or does misread them, unless surface calibration shoud be performed. In fact, gesture recognition performed best on hard surfaces (τ=84.7%), worst on soft surface (τ=76.07%) and better on irregular surfaces (τ=80.23%). Mrazek and Mohd [44] compared three LSTM models to interpret Tap Strap inputs from raw accelerometer data: a standard Long Short-Term Memory (LSTM) model, a model with CNN and LSTM layers, and a convolutional LSTM. The LSTM was found to have had the highest accuracy at an average of 97.47%, followed by the CNN-LSTM with an accuracy of 96.68% and the ConvLSTM with an accuracy of 96.65%. Although the model may suffer from model overfitting [26], it did not appear to have any impact on accuracy. To improve the accuracy of input readings from accelerometer data, various types of algorithms have been considered, such as Convolutional Neural Networks (CNNs) [47], LSTM [44]. Unfortunately, Tap Strap SDK allows only extracting raw input data, which therefore represents an opportunity, but also a challenge, to apply ML algorithms to reach the best accuracy possible. Before addressing the technical challenges related to ring-based multi-finger gesture recognition, we would like to explore user-defined gestures in this context of use, which is the goal of the next section.

3 GESTURE ELICITATION STUDY

Since the current gesture vocabulary of Tap Strap 2 (Fig. 2) consists of only multi-finger tapping gestures, among them are eight for commands, we would like to explore user-defined ring-based multi-finger gestures by conducting a Gesture Elicitation Study (GES), a participatory design method where participants are instructed to propose gestures to execute system commands [65]. Although many such studies have been reported (see [33] for a review of mid-air interaction, [59] for a macroscopic analysis of GESs, and [58] for a microscopic analysis of GESs), none of them seem to have reported gestures belonging to our field of investigation.

3.1 Positioning of Ring-based Multi-finger Gestures

Before proceeding, we provide our operational definition of a ring-based multi-finger gesture as follows:

> A "ring-based multi-finger gesture" is any movement of the device-wearing finger(s) that causes a detectable change in the rings' positions on a 2D surface or in a 3D spatial system of reference centered on the device.

By adopting this definition, we are compliant with the Vatavu and Bilius [56]' GestuRing space, which compiled a comprehensive survey of gestures with ring-based, ring-like, and ring-ready devices. In the first case of our definition, gestures are performed on any surface (*e.g.,* a flexible surface such as a cloth or a rigid surface such as a tabletop) and usually mimic interactions common on touchscreens, such as taps, pinch gestures, or directional swipes. In the second case of our definition, gestures are performed in space as the rings' accelerometers (Fig. 2) capture any movement. We adopt a comprehensive approach, reflected in our definition of ring-based multi-finger gestures, by examining both contact-based and mid-air input for such multi-finger rings, including combinations thereof. Thus, all the following examples are valid gestures for the purpose of our study:

> *Positive example*: Touch a surface with one or many fingers once, twice, or multiple times in a row. Tap a rhythmic pattern on the surface.
> *Positive example*: Push forward one or many fingers once, twice, or multiple times in a row.
> *Positive example*: Raise the hand by moving the fingers up.

Our definition assumes that any gesture is considered relevant as soon as the fingers are moving. Consequently, other movements where the fingers remain fixed while other limbs are moving, such as a moving wrist or a moving forearm with fixed fingers, do not belong to our definition since they are not detected.

> *Negative example*: Move the wrist or the forearm while keeping the fingers fixed in position.
> *Negative examples*: Touch the ring surface with one or many fingers, roll the rings.

Gheran et al. [19] define a "ring-based gesture" as any action performed *with* or *on* a smart ring or any movement of the wearing finger and/or hand that causes a detectable change in the ring's position and/or orientation in a system of reference centered on the user's finger or body. In contrast, our gestures involve five connected rings that do not move on the fingers and on which no gesture can be performed. Only the movements of the fingers are recorded. Therefore, gestures such as touching rings, tapping rings, or rolling them are out of the scope of our investigation. Fig. 3 shows a selection of gestures relevant to our field of investigation. Chaffangeon Caillet et al. [7] and Way and Paradiso [61] define a *microgesture* as "a quick and subtle movement of the fingers". Our definition encompasses these 3D microgestures [37] and goes beyond by assuming we have a gesture as soon as at least one finger moves in space, whatever the condition.

Multi-finger gestures, including microgestures, can be captured by devices other than rings, such as vision (*e.g.,* a Leap Motion Controller [49]), wearables [56], or radars [12]. Since the rings are worn on the five fingers of one hand (or two, since two can be used at once), the only physical limit to the gestures is their human movement in space. Other conditions, such as lighting, field of view, position, or material do not constrain the Tap Strap. The

Fig. 3. A selection of user-defined gesture proposals from our gesture elicitation study.

device can feel cumbersome to put on, to take off, to wear, even if we do not notice its weight (0.45 lbs, 204 g) after a while. Its price of $99, which may be a barrier to purchase or accessibility for different people, makes the Tap Strap still the most affordable and accurate off-the-shelf device for finger-tap detection [68]. Like all wearable devices, its operation life is limited by the life of the battery (10 Hours or 45 mAH of operation).

3.2 Participants

From internally maintained mailing lists of volunteers, we recruited thirty (*N*=30) no-handicap participants (14 female and 16 male, with a female/male ratio of 0.87) ranging from 13 to 63 years (with a mean age of 33.40, a standard deviation of 14.41, and a median of 25.50). Of the participants, 43% were employees in various areas of human activity, 7% were independent workers, 3% occupied executive positions, 3% were unemployed, and 44% were students. Three participants were left-handed and one of them was ambidextrous. The average score on the creativity tests was 60.92, slightly below the average of 63.45. None of the participants had prior knowledge of the device and all reported moderate to extensive usage of desktop computers, smartphones, and tablets, respectively.

3.3 Stimuli

We hereby refer to a "referent" [65] as any user interface feature that can be controlled independently with a command, which is equivalent to a "function" [59]. Table 2 presents the 18 referents of the experiment. We created a Microsoft PowerPoint presentation showing a *referent* [65] for each command via a before/after action represented on one slide at a time so that they can be viewed by participants. Each representation reproduced a simplified view of the screen with the cause and the effect.

Table 2. List of referents defined for the study.

Id.	Referent name	Description
1	Open window	Open a currently selected window which has the focus to its maximal viewing
2	Close window	Close a currently opened window
3	Switch windows	Switch between windows that are currently opened
4	Lock screen	Lock the current screen
5	Unlock screen	Unlock the current screen without requiring any password
6	Take a screen capture	Capture a screenshot of the current window in use
7	Go to previous item in a list	Go to the previous item in a list of items, such as in a list box our drop-down menu, or the previous window in a set of active windows
8	Go to next item in a list	Go to the next item in a list of items, such as in a list box our drop-down menu, or the next window in a set of active windows
9	Scroll up	Scroll up in a currently opened window or list
10	Scroll down	Scroll down in a currently opened window or list
11	Zoom in	Zoom in a currently active window
12	Zoom out	Zoom out in a currently active window
13	Play	Play a currently selected multimedia item, such as an MP4 video
14	Pause	Pause in the currently playing multimedia item
15	Turn up the volume	Increase the volume of the currently playing multimedia item
16	Turn down the volume	Decrease the volume of the currently playing multimedia item
17	Take a call	Answer a phone call
18	Hang up a call	Hand up a phone call

3.4 Procedure

Two members of the group welcomed the participants to the experimental site. The environment consisted only of a desk and a computer. Each participant must have no link with the members of the group questioning her/him to ensure the neutrality of the answers and avoid the Hawthorn effect. We divided the experiment into three phases.

Pre-test Phase. Participants were first asked to complete an approved consent form and a demographic questionnaire, including questions about their knowledge of connected devices, such as their frequency of device use. The participant is introduced to the Tap Strap device by watching a 1:30 min introductory video and a 1:15 min video for text entry. The participant then freely interacted with the device for five minutes based on TapManager.

Test Phase. The 18 referents were randomly presented to the participant [57]. Each referent was presented to the participant, one by one, to give them time to reflect on the gesture proposal that seemed the most preferred (until saying "I am ready"), the most appropriate for the order presented by the referent (Fig. 3). When desired, the participant could repeat the gesture proposal. Once the gesture proposal was produced and recorded, the participant was asked to provide the rank, goodness of fit, complexity, memorability, and fun character of the gesture proposal.

Post-test Phase. Once all referents have been presented, the participant leaves the experiment room and completes a questionnaire consisting of two parts: an IBM PSSUQ post-test questionnaire [35] and an additional questionnaire including closed questions with predefined answers and open-ended questions for participants to express their opinions.

3.5 Setup and Variables

Our setup follows a within-subjects design with the following dependent variables for each referent:

(1) Thinking time: a numerical variable defined as the average of times elapsed between the first showing of the referent and the moment when the participant knew which gesture would be preferred, was measured in seconds with a stopwatch.
(2) Goodness of fit: a numerical variable defined as the average rating from 1 to 10 expressing to what extent participants thought their gesture was appropriate.
(3) Agreement rate (AR): a numerical variable expressing the agreement among participants to propose a gesture class for a referent, as calculated by AGATe [57] along with its agreement magnitude (AR ≤0.1=low, 0.1< AR ≤0.3= moderate, 0.3< AR ≤0.5= high, 0.5< AR ≤1= very high).
(4) Complexity: a numerical variable defined as the average rating of 1=most complex to 7=least complex expressing to what extent participants thought their gesture was complex to produce or was inducing fatigue.
(5) Memorability: a numerical variable defined as the average rating of 1=least memorable to 7=most memorable expressing to what extent participants thought their gesture was easy to remember and reproduce.
(6) Fun: a numerical variable defined as the average rating of 1=least funny to 7=most funny expressing to what extent participants thought their gesture was funny.

Although other GESs have been conducted for one ring device [2, 19, 20], sometimes replicated [21], no existing GES covers a setup where a ring is on all five fingers simultaneously. However, this does not mean that the gestures resulting from our study are necessarily different from existing studies (*e.g.,* [70]), but that the setup

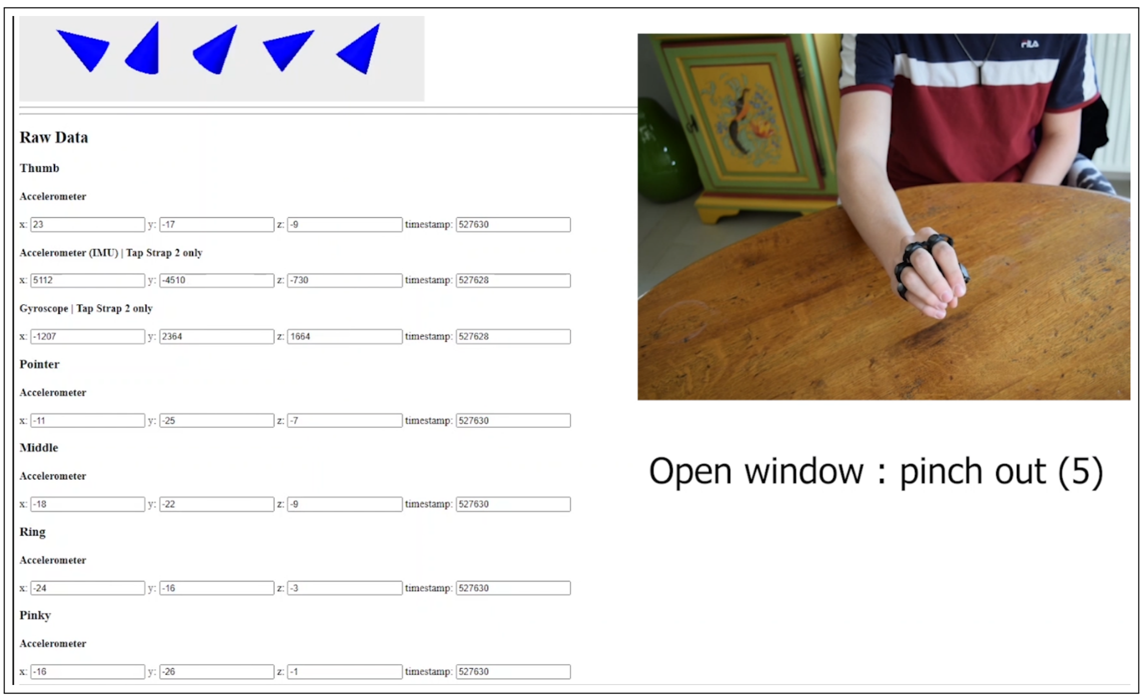

Fig. 4. Custom application for raw data capture of ring-based multi-finger gestures.

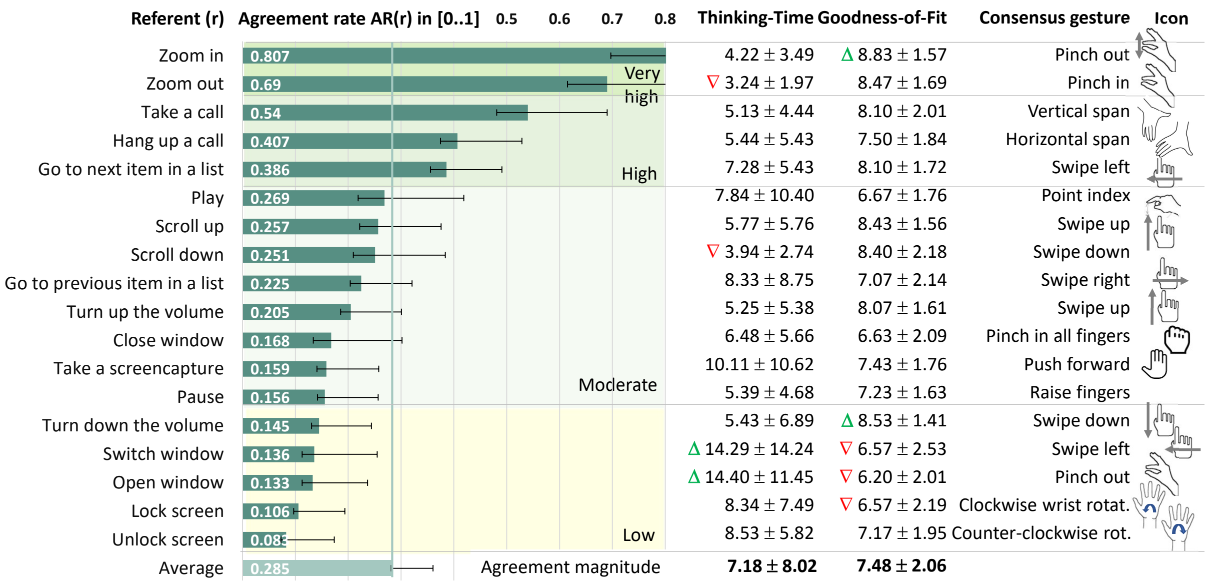

Fig. 5. Consensus set of gestures resulting from the gesture elicitation study: referent, Agreement rate (AR) and magnitude, Thinking time, Goodness of fit, consensus gesture, and icon. The two best and worst values of the two variables are depicted by a top-pointing arrow and a bottom-pointing arrow, respectively. The error bars show a confidence interval of 95%.

used to elicit them is unique. Our study is the only one specifically tailored for the Tap Strap 2 that goes beyond the text entry commands shipped with the device, as both classical referents (*e.g.,* for navigation) and original ones (*e.g.,* for window management) are involved with the six variables mentioned above, which goes beyond the mere agreement rate that is typically considered in other GESs. For example, Danyluk et al. [14] collected user-defined microgestures for augmented reality maps, but no ring device was involved. Same for [8, 37].

3.6 Data Collection

To collect raw data from the device for creating a dataset, we used Tap's Python SDK to extract raw sensor data intended for the training and evaluation of our models using a custom capture application (Fig. 4- see videos in the supplementary material). This SDK allows us to send and receive data and events by establishing a Bluetooth Low-Energy connection with Tap Strap 2 and to collect raw data from the sensors embedded within Tap Strap 2. As noted in Mrazek et al. [43], this SDK does not process or interpret raw data, which will be our responsibility to recognize them in Section 5. Log files were created with a timestamped format of 3D finger positions, such as:

```
Timestamp:850696 Thumb: 'X': -273.4375, 'Y': -664.0625, 'Z': 609.375 Index: 'X': -367.1875, 'Y': 179.6875, 'Z': 859.375 Middle: 'X': -281.25, 'Y': 85.9375, 'Z': 906.25 Ring: 'X': -226.5625, 'Y': 210.9375, 'Z': 867.1875 Pinky: 'X': -156.25, 'Y': 312.5, 'Z': 843.75 Timestamp:850697 Thumb: 'X': -281.25, 'Y': -679.6875, 'Z': 601.5625 Index: 'X': -367.1875, 'Y': 179.6875, 'Z': 859.375 Middle: 'X': -281.25, 'Y': 78.125, 'Z': 898.4375 Ring: 'X': -234.375, 'Y': 203.125, 'Z': 875.0 Pinky: 'X': -164.0625, 'Y': 312.5, 'Z': 859.375
```

3.7 Results and Discussion

We collected 30 participants × 18 referents = 540 gesture proposals. Each proposal is characterized by a log file and six measures, thus creating 540 × 6 measures = 3,240 data. Fig. 5 shows the consensus set of gestures resulting from the study, decomposed as the gesture name of the most agreed on gesture proposal (col. 5) by maximizing the Agreement rate (col. 2), its corresponding Thinking Time (col. 3), the Goodness-of-fit (col. 4), an icon representing the gesture class (col. 6), Agreement rate and its magnitude (col. 2). The references (col. 1) are sorted in decreasing order of their Agreement Rate. The Agreement Rates are overall moderate in magnitude (*M*=0.285,*SD*=0.20), ranging between 0.08 (low) for “Unlock screen” and 0.807 (very high) for “Zoom in”. In global sampling, $\frac{5}{18}$ = 28% of the rates belong to the low consensus category, $\frac{8}{18}$ = 44% of the rates belong to the moderate range, $\frac{3}{18}$ = 17% are high, and $\frac{2}{18}$ = 11% are very high. Apart from a few exceptions, most gestures received an agreement rate slightly higher than or close to those reported in the GES literature ([57] summarized agreement rates of 18 GESs).

Goodness of Fit generally received a high value that indicates subjective satisfaction of the participants in the gestures they proposed (*e.g.,* *M*=7.48, *SD*=2.06, *Mdn*=8). They mostly appreciated their gesture proposals for “Zoom in” (8.83 ±1.57) and “Turn up the volume” (8.53 ±1.41), and less for unfamiliar commands such as “Switch window” (6.57 ±2.53) *ex aequo* with “Lock screen” (6.57 ±2.19), and “Open window” (6.20 ±2.01). While “Lock screen” and “Unlock screen” were less satisfying in terms of Agreement rate and Goodness-of-Fit, they resulted in two symmetric consensus gestures: clockwise, respectively. counter-clockwise wrist rotations to mimic the opening and closing of a door using a key. The Thinking Time received quite reasonable values in general (*M*=7.18s, *SD*=8.02s), ranging from very short times (*e.g.,* “Zoom out” in 3.24 s and “Scroll down” in 3.94 s) to longer times, often associated with less familiar or frequent commands (*e.g.,* “Switch window” in 14.29 s). The Thinking time correlates negatively with the Goodness of Fit (Pearson’s $r_{n=19}=-.53$, R^2=.63-Fig. 6): referents with a higher thinking time have a lower goodness-of-fit, thereby suggesting that participants were less satisfied with their gesture proposal when they thought longer.

Fig. 7 shows the distribution of the last three dependent variables for the 18 referents. A Student test t shows that the means are significantly higher than the median *Mdn=4* with different levels of significance, mostly high.

The participants consistently rated Complexity (Cronbach’s α=.86, interpreted as ’good’) of 18 referents. A one-way ANOVA ($F_{17,30}$=2.87, p=.00011***, between groups) revealed that there was a statistically significant impact of Referent on Complexity. Tukey’s HSD Test for multiple comparisons found that the mean value of Complexity was significantly different between a few pairs of conditions, such as between “Scroll down”, “Scroll up”, “Turn up the volume”, “Play” and “Take a screen capture”, respectively. This is normal since the referent “Take a screen capture” was perceived as the most complex gesture. The referents perceived as the least complex were: “Turn down the volume”, “Turn up the volume”, “Scroll up”, “Scroll down”, and “Play”, which are probably among the most familiar and frequent actions that are meaningful for participants.

The participants also consistently rated Memorability (Cronbach’s α=.83, interpreted as ’good’) of 18 referents. A one-way ANOVA ($F_{17,30}$=3.44, $p\leq$.001***, between groups) revealed that there was a statistically significant impact of Referent on Memorability. The referents perceived as the easiest to remember were again “Scroll up” and “Scroll down”, “Zoom in”, “Turn down the volume” and “Turn up the volume”, and “Take a call”. The least memorable referents were “Open window” and “Take a screen capture” again, which seems consistent with previous results. Tukey’s HSD Test for multiple comparisons found that the mean value of Memorability was significantly different between a few pairs of conditions, such as between “Scroll up” and “Close window”, “Open window”, “Go to prvious item”, “Take a screen capture”, respectively.

The participants also consistently rated Fun (Cronbach’s α=.88, interpreted as ’good’) of 18 referents. A one-way ANOVA ($F_{17,30}$=1.57, p=.067, *n.s.*, between groups) revealed that there was no impact of Referent on

Fun. The participants did not express a different level of fun depending on the referent. The funniest referents were: "Play" and "Pause", "Scroll up" and "Scroll down", and "Hang up a call".

Following each study, participants were asked to complete *the IBM Post-Study System Usability Questionnaire (IBM PSSUQ questionnaire)* [35], a tool used for evaluating the usability of systems. The questionnaire consists of 16 items measuring user satisfaction, such as system usefulness, ease of use, and interface quality. The responses are collected on a 7-point Likert scale [38], ranging from strongly agree to disagree strongly, or the participant can opt to abstain from answering questions irrelevant to the study. We utilized a proposed questionnaire version by UIUXTrend consisting of 16 questions and maintained a scale of 1 to 7 (Fig. 8). The average score for most questions ranges from 4 to 6. However, participants deemed questions 7, 8, 9, and 10 off-topic, yet some still provided scores for these questions. Participants rated question 12 highly regarding clarity of information and question 13 for the acceptability of the proposed interface. The Information Quality was ranked the highest (*M*=5.40), followed by System usefulness (*M*=5.26), which is in turn slightly higher than the Overall satisfaction (*M*=5.20). The Information Quality was ranked the lowest (*M*=4.60), which is located slightly below the usual threshold of 5.

4 GESTURE CLASSIFICATION

The 540 gesture proposals were clustered into 21 gesture categories, which were further classified according to four dimensions as follows (Table 3) [19, 21, 57, 65]:

(1) *Range of motion*: expresses the distance from the participant's body to the locus of gesture decomposed into the following values: C=Close intimate, I=Intimate, P=Personal, S=Social, U=Public, and R=Remote. All gesture proposals were produced within the personal range, which is not too close of the body, but always at some distance that is typically the length of the forearm.

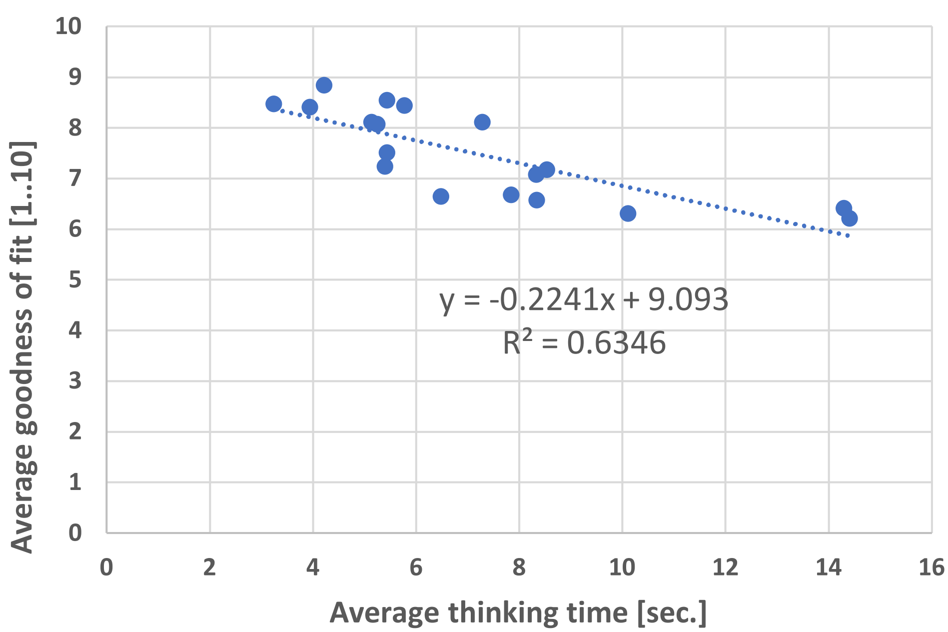

Fig. 6. Negative correlation between the Thinking time and the Goodness of fit.

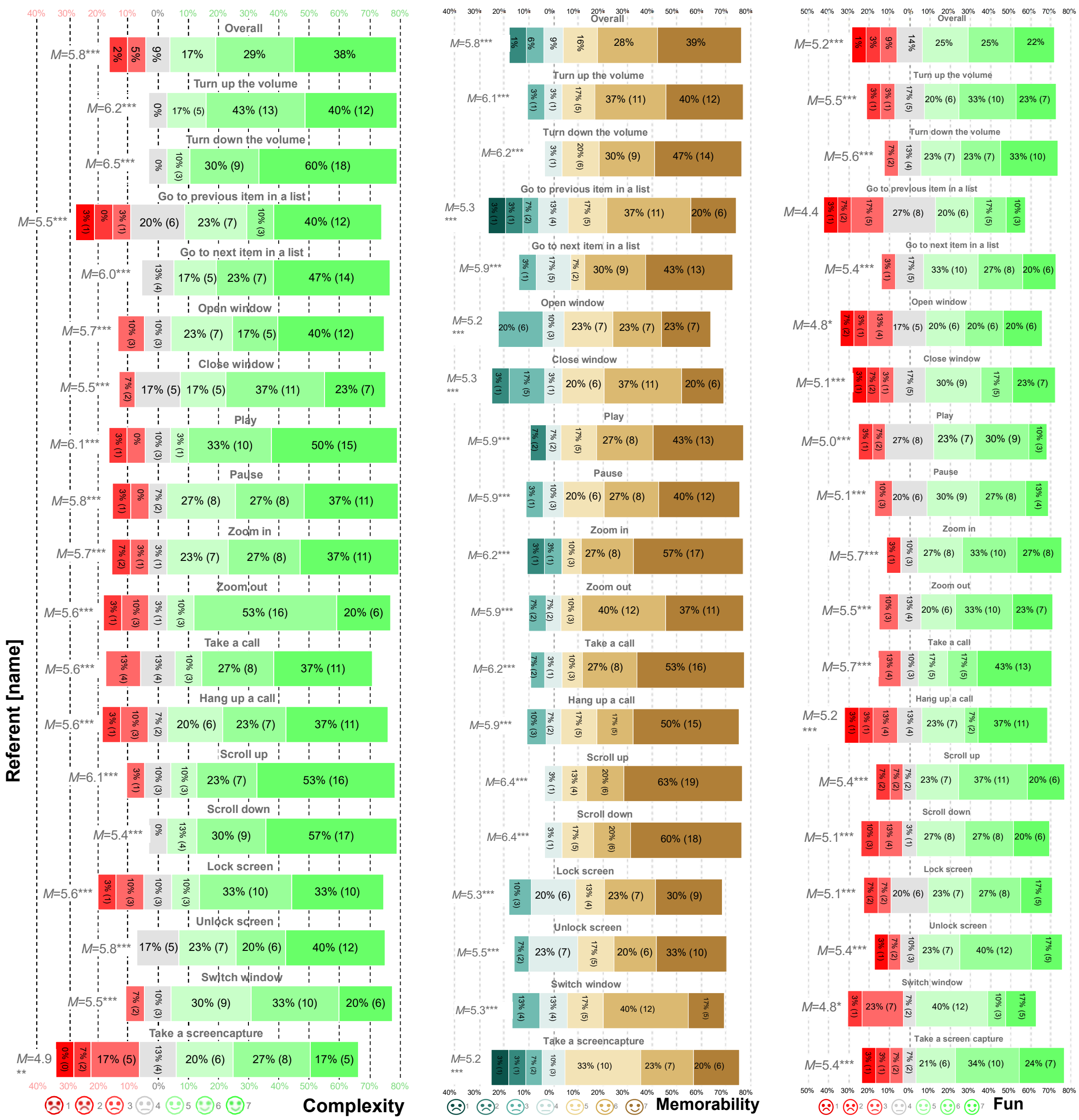

Fig. 7. Distribution of the last three dependent variables for the 18 referents: Complexity, Memorability, and Fun, each on 1-7 rating scale (M=mean, $p \leq .001^{***}$ or $p \leq .01^{**}$).

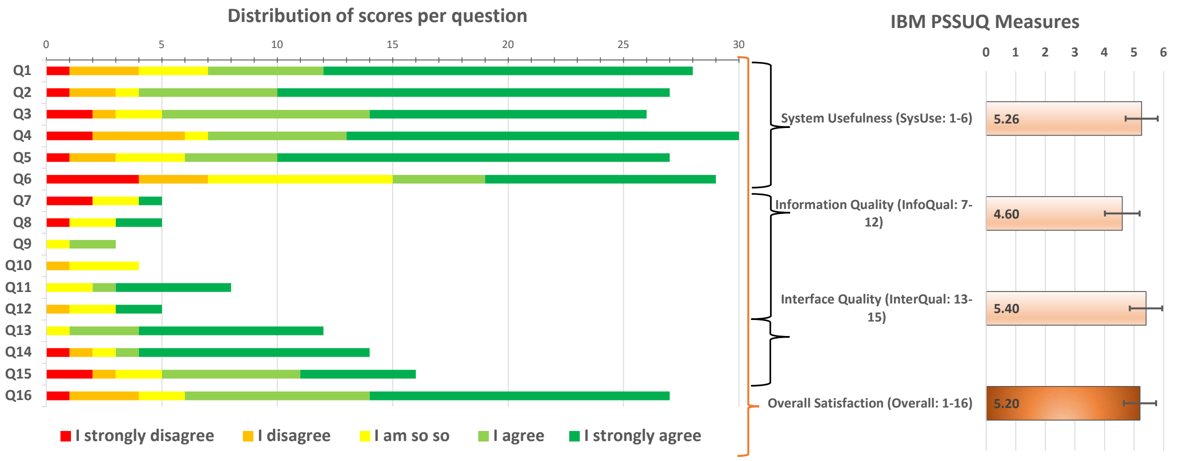

Fig. 8. Evaluation of the TapStrapGest according to IBM PSSUQ questionnaire [35].

(2) *Nature*: expresses the gesture nature decomposed into the following values: P=Pointing gesture, T=Static semaphoric, D=Dynamic semaphoric, S=Semaphoric stroke, A=Pantomimic, I=Static iconic, Y=Dynamic iconic, M=Manipulation gesture.
(3) *Laterality*: characterizes how hands were involved in the gesture proposal decomposed into D=Dominant unimanual, N=Non-dominant unimanual, S=Symmetric bimanual, A=Asymmetric bimanual [1]. Since only one Tap Strap 2 device was used, all participants agreed to put and calibrate the device on their dominant hand. While two devices could be used to combine gestures, either from the same person or from two different (therefore inducing bimanual gestures), this setup represents another type of experiment.
(4) *Form*: expresses the form of the gesture proposal, decomposed into the following values: S=Stroke, T=Static, M=Static with motion, D=Dynamic. Fig. 10 shows the distribution of the categories according to this dimension: 14% were strokes, only 1% was static, 66% were static with some form of motion, and 19% were dynamic.

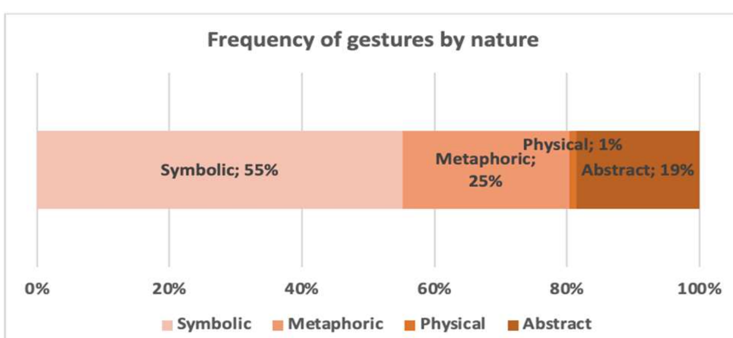

Fig. 9. Classification of gesture proposals by nature.

Table 3. Classification of gesture proposals. Legend: see text.

Category	Name	Range of motion	Nature	Laterality	Form
1	Single tap (any combinaison of fingers)	P	D	D	S
2	Double tap (any combination of fingers)	P	D	D	S
3	Swipe up (vertical axis)	P	S	D	M
4	Swipe down (vertical axis)	P	S	D	M
5	Swipe left	P	S	D	M
6	Swipe right	P	S	D	M
7	Swipe forward (horizontal axis)	P	S	D	M
8	Swipe backward (horizontal axis)	P	S	D	M
9	Wrist rotation counter-clockwise	P	A	D	D
10	Wrist rotation clockwise	P	A	D	D
11	Circular motion counter-clockwise	P	D	D	D
12	Circular motion clockwise	P	D	D	D
13	Pinch in	P	M	D	M
14	Pinch out	P	M	D	M
15	Point left	P	P	D	T
16	Point right	P	P	D	T
17	Point centrefront	P	P	D	T
18	Draw square	P	Y	D	D
19	Draw cross	P	D	D	D
20	Pick up (close fist and move up)	P	A	D	D
21	Hang up (close fist and move down)	P	A	D	D

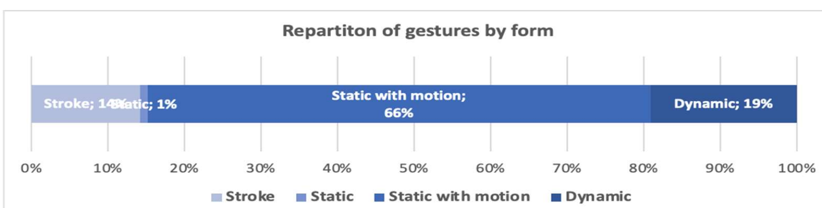

Fig. 10. Classification of gesture proposals by form.

5 GESTURE RECOGNITION

This section describes an implementation in Python and a test gesture recognition with TapStrapGest on ring-based multi-finger gestures issued by a Tap Strap by applying Machine Learning and Pattern Recognition techniques. First, the dimensionality of the dataset has been reduced by applying a Principal Component Analysis (PCA) [23]. Then, four gesture recognizers were implemented and trained using different techniques: (1) a Linear Discriminant Analysis (LDA) which was used to find whether any linear combination of features exists that characterizes or separates many gesture classes [67], (2) a Quadratic Discriminant Analysis (QDA), which is similar to LDA in which the measurements of each gesture class are expected to be normally distributed –an assumption that is challenging to prove in our case–, where there is no assumption that the covariance of each of

Table 4. The five first rows of the initial dataset (Acc.=accelerometer). Source: GitHub

Row	Acc. 1	Acc. 2	Acc. 3	Acc. 4	Acc. 5	Acc. 6	Acc. 7	Acc. 8	Acc. 9	Acc. 10	Acc. 11	Acc. 12	Acc. 13	Acc. 14	Acc. 15	Symbol
0	32	6	-8	11	-7	30	6	7	32	8	7	30	9	-7	31	A
1	33	6	-9	12	-8	29	6	7	32	8	7	30	8	-7	31	A
2	33	5	-8	11	-7	30	6	8	32	8	7	31	8	-6	31	A
3	33	5	-8	11	-7	30	5	7	32	8	7	30	8	-7	31	A
4	33	5	-8	11	-7	30	5	7	32	7	8	31	7	-6	32	A

Table 5. The five first rows of the initial dataset in Table 4, after normalization (Acc.=accelerometer).

Row	Acc. 1	Acc. 2	Acc. 3	Acc. 4	Acc. 5	Acc. 6	Acc. 7	Acc. 8	Acc. 9	Acc. 10	Acc. 11	Acc. 12	Acc. 13	Acc. 14	Acc. 15	Symbol
0	1.00	1.74	-0.61	-0.44	-1.51	1.23	-0.50	-1.02	1.66	-0.29	-2.08	1.75	-0.33	-2.66	1.52	A
1	1.09	1.74	-0.68	-0.37	-1.60	1.15	-0.50	-1.02	1.66	-0.29	-2.08	1.74	-0.40	-2.66	1.52	A
2	1.09	1.65	-0.61	-0.44	-1.51	1.23	-0.50	-0.90	1.66	-0.29	-2.08	1.82	-0.40	-2.54	1.52	A
3	1.09	1.65	-0.61	-0.44	-1.51	1.23	-0.56	-1.02	1.66	-0.29	-2.08	1.74	-0.40	-2.66	1.52	A
4	1.00	1.65	-0.61	-0.44	-1.51	1.23	-0.56	-1.02	1.66	-0.35	-1.91	1.82	-0.47	-2.54	1.62	A

Table 6. Some rows of the final dataset.

Row	Acc. 0	Acc. 1	Acc. 2	Acc. 3	Acc. 4	Symbol
4230	2.24	1.44	-1.26	-1.23	1.28	X
2925	5.04	2.58	2.2	-0.68	-0.75	Q
4894	-2.11	-1.04	-0.38	0.28	-1.18	*
19	-4.24	3.26	1.23	0.57	-1.29	A

the gesture classes is identical [53], (3) Support Vector Machines (SVM) [9], and (4) Random Forest (RF) [32]. These four classifiers were tested on a set of gestures in American Sign Language. An open source GitHub repository contains this implementation, as well as the dataset, and detailed results. Finally, a performance analysis of each classifier has been performed, considering precision indices and the time performance of the training and classification process. A video is accessible at https://vimeo.com/681846114.

5.1 Acquisition of Gesture Dataset

The Tap Strap SDK continuously streams data in raw sensors mode: five 3-axis accelerometers (one sensor per finger) and IMU data (3-axis accelerometer plus one gyroscope) located on the thumb. These raw data are measured with respect to the reference system located in the IMU on the thumb (Fig. 2). While the internal Tap Strap clock sends a sample of raw data with a frequency of 200 Hz (5 ms) in the format `timestamp; type (IMU, device); values`, we reduced the reading rate to 20 Hz (T=50 ms) to allow gesture recognizers to work properly with enough data without overloading them. A first analysis of the raw data showed that the 6 IMU values were not needed for this process. Therefore, a filter was applied to reduce the number of dimensions from 21 to 15 features (see Table 4 for some row examples). A custom application was used for raw data capture of the ring-based multi-finger gestures, further stored by class/label/letter. The full dataset is accessible at https://github.com/gcornella/TapStrap2-Hand-Gesture-Recognition-/blob/main/data_tap.txt.

5.2 Dimensionality Reduction

We normalized the dataset samples by removing the mean and scaling to unit variance using a Standard Scaler (see Table 5 for some row examples). A Principal Component Analysis (PCA) [23] was used to reduce the dataset

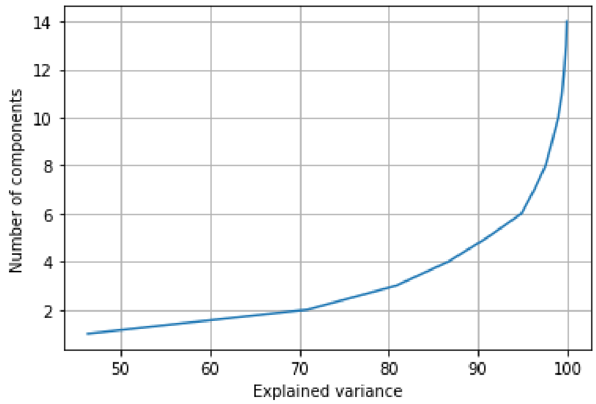

Fig. 11. Plot of the explained variance in the gesture dataset by the number of components.

dimensionality by projecting each point onto the first few principal components to obtain lower dimensional data while preserving as much of the data variability as possible. The plot represented in Fig. ?? was used for choosing the number of principal components: the fourth component explains 86.69% of the variability of the data, the fifth component explains 91.03% of the variability of the data, and the sixth component explains 94.86% of the variability of the data. We therefore chose to reduce the dimensionality from 15 dimensions to 5 for the next step (see Table 6 for some examples).

In sum, our publicly available dataset[1], recorded at a sample rate of 20 Hz (T=50 ms), contains 4,910 data samples and 27 American Sign Language alphabet characters (26 letters of the alphabet corresponding to the American Sign Language plus the "*" symbol), captured from a single subject wearing the device on the right (dominant) hand. Extensive tests with different subjects yielded similar results, thus suggesting minimal inter-subject variability. Even if this dataset was recorded user-dependent, the testing will be performed with many subjects in a user-independent scenario, which is far more demanding than user-dependent (see Fig. 13).

5.3 Testing Procedure

The gesture dataset was divided into four different subsets (data-train, label-train, data-test, label-test), where the training data contained 70% of the whole dataset and the testing data comprised 30% of the remaining part. For each recognizer, we repeated R=1, 000 times the recognition of a randomly selected sample from the subset to finally compute:

- The *precision* is the ratio $\frac{tp}{tp+fp}$ where tp is the number of true positives and fp the number of false positives. The precision is intuitively the ability of the recognizer not to label a negative sample as positive.
- The *recall* is the ratio $\frac{tp}{tp+fn}$ where tp is the number of true positives and fn the number of false negatives. The recall is intuitively the ability of the recognizer to find all the positive samples.
- The F-beta score can be interpreted as a weighted harmonic mean of the precision and recall, where an F-beta score reaches its best value at 1 and worst score at 0. The F-beta score weights recall more than precision by a factor of beta. In our case, we have β=1 to specify that recall and precision are equally important.
- The *support* is the number of occurrences of each class when true.

[1]See https://github.com/gcornella/TapStrap2-Hand-Gesture

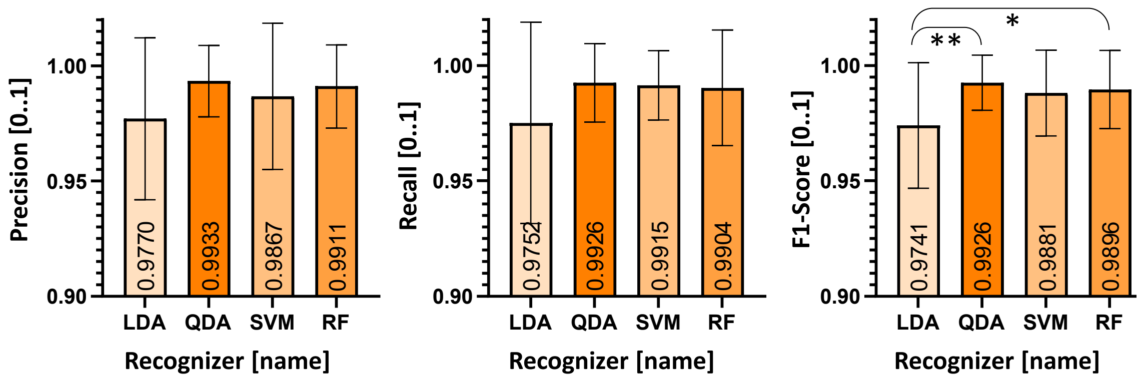

Fig. 12. Results of the performance analysis of the four recognizers: precision, recall, and F1-Score. LDA=Linear Discriminant Analysis, QDA=Quadratic Discriminant Analysis, SVM=Support Vector Machines, and RF=Random Forest. Error bars show a confidence interval of 95%.

5.4 Results and Discussion

Fig. 12 shows the average precision, recall, and F1 score for the four recognizers tested, with the most saturated color showing the best condition and the least saturated color showing the worst condition. A one-way ANOVA with a single factor reveals that the recognizer does not have any effect on precision ($F_{27,3}$ = 2.01, p=.12, *n.s.*) and recall ($F_{27,3}$ = 2.38, p=.07, *n.s.*), but well on F1-score ($F_{27,3}$ = 4.82, p=.0034**, with a small effect size of η^2=.12).

Regarding precision, QDA (M=0.9933, SD=0.0152) outperforms the three other conditions RF (M=0.9911, SD=0.0177), SVM (M=0.9867, SD=0.0312), and LDA (M=0.9770, SD=0.0344), respectively both in terms of the highest average and the smallest standard deviation, which confirms that this condition is the safest one for stability. Regarding recall, QDA (M=0.9926, SD=0.0167) outperforms SVM (M=0.9915, SD=0.0148), RF (M=0.9904, SD=0.0246), and LDA (M=0.9752, SD=0.0428), respectively. Finally, QDA displays the best F1-Score (M=0.9926, SD=0.0001) both in terms of highest average and smallest standard deviation, which makes QDA the best recognizer in our testing and LDA the worst one.

Then, we computed some normality tests to determine whether precision, recall, and F1-score follow some Gaussian distribution. Both Shapiro-Wilk [48] and d'Agostino-Pearson revealed that none of the conditions followed a normal distribution. With regard to the LDA, the precision was not normal (Shapiro-Wilk: *W-stat*=0.71, $p\leq$.001***, α=0.05; d'Agostino-Pearson: *DA-stat*=12.64, p=0.001795**), the recall was not normal (Shapiro-Wilk: *W-stat*=0.62, $p\leq$.001***, α=0.05; d'Agostino-Pearson: *DA-stat*=19.36, $p\leq$0.001***), and the F1-score was not normal neither (Shapiro-Wilk: *W-stat*=0.82, p=.000402***, α=0.05; d'Agostino-Pearson: *DA-stat*=6.61, p=0.036*).

Consequently, we computed a series of Kruskall-Wallis tests [34] for independent samples (α=0.05). No significant differences were found between the recognizers in terms of their precision (KW=4.507, p=0.21, *n.s.*) and their recall (KW=4.716, p=0.19, *n.s.*). However, a significant difference was found regarding their F1-Score (KW=12.71, p=0.0053**), which led us to perform a Dunn test for multiple comparisons: QDA is significantly better than LDA (M=−26.31, p=0.0061** with a medium effect size (Cohen's r=0.30 [9]) and RF is also significantly better than LDA (M= − 22.48, p=0.0299*) with a medium effect size (Cohen's r=0.37).

6 PERFORMANCE ANALYSIS AND DESIGN IMPLICATIONS

The performance of the testing is as follows for the R=1, 000 repetitions: LDA (5.95 s), QDA (8.79 s), SVM (182 s), and RF (864 s). For a single repetition, LDA and QDA benefit from a response time that is largely inferior to

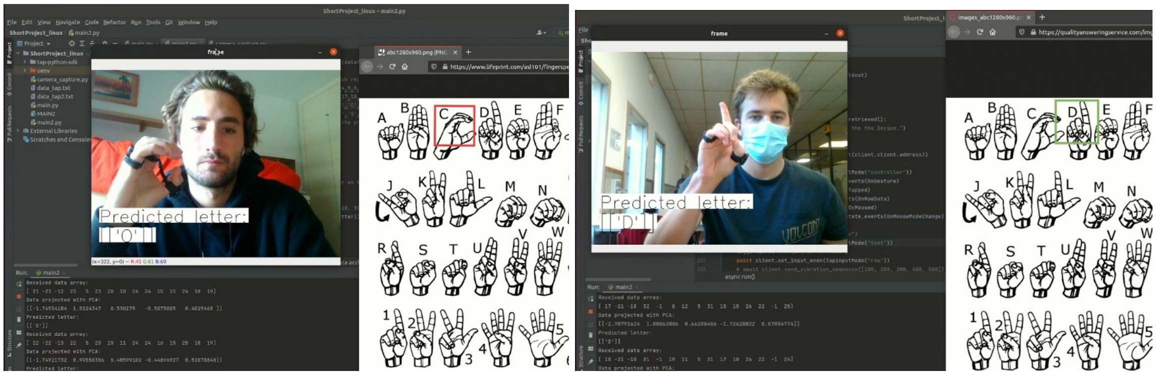

Fig. 13. Two examples of gesture recognition: an O letter as false positive (left), a D letter as true positive (right).

the minimal system response time of t=0.1 sec. Based on our insights and results of the gesture elicitation study (Section 3) and the gesture recognition (Section 5), we suggest the following list of implications for designing ring-based multi-finger gesture interfaces:

(1) *Support gesture customizability:* even if the system-defined gestures are relatively well defined and learned, the end user may still prefer one or more personal gestures, either to complement an existing command (which implies gestural redundancy), or to replace it (which induces inconsistency with the initial system). Customizability can be addressed in different ways: by changing the gesture set (we provide two new gesture sets that end users can choose from or customize by adding/deleting/replacing samples/classes without needing to entirely retrain the system), by adapting gesture mapping (since a gesture class is mapped to a system command, another class can be mapped also to the same command to address users' preferences and tailoring), and by tailored recognition (specific gesture classes can be activated/deactivated in TapStrapGest to tailor their recognition to a particular application or domain). However, TapStrapGest does not support gesture customizability by adjusting gesture sensitivity depending on the user (unless they are recorded in the templates) or sequences to combine multiple gestures into a new command (this is for future work).

(2) *Finger independence:* ensure that each finger's movement can be detected and distinguished from others. This allows users to perform intricate gestures that involve multiple fingers simultaneously without unintended actions. Calibration and sensitivity adjustment of the device might be necessary to accommodate different hand sizes and user preferences, as it may influence the recognition process.

(3) *Gesture consistency:* establishes a set of consistent gestures [15] that users can easily learn and remember. Consistency improves usability and reduces the cognitive load on users. For example, swiping inward with two fingers might always trigger a specific action, regardless of the application or context. If finger independence is ensured, the swiping could be recognized independently of the number of fingers, unless the number and position of the fingers are meaningful.

(4) *Feedback and confirmation:* provides visual or haptic feedback to confirm when a gesture is recognized and executed. This feedback reassures users that their input has been registered, enhancing the overall user experience. Feedback can include animations, sound cues, or slight vibrations depending on the device.

(5) *Customization:* enables end-users to customize their gestures to match their preferences and usage patterns [13, 50]. Offering options to adjust gesture sensitivity, assign custom gestures to specific actions, or even create user-defined gestures can improve user satisfaction and adoption of the interface.

(6) *Accessibility considerations:* ensures that the gestures are accessible enough to end-users with different levels of dexterity and mobility in the perspective of ability-based design [64]. Provide alternative input methods or adjustable settings for users with disabilities. Conduct usability testing with diverse user groups to identify and address any accessibility barriers.

Ring-based gesture recognition offers numerous advantages for the deaf and hard-of-hearing population, including increased accessibility and portability, as it eliminates the need for sôphisticated camera setups. These systems enable discreet, real-time communication through hand gestures, enhancing privacy and ensuring effective interaction across various environments. Without relying on visual data, they provide a secure and versatile solution that empowers users to communicate confidently and independently, significantly improving their overall accessibility and quality of life.

7 FORMALIZING RING-BASED MULTI-FINGER GESTURES

According to the definition of ring-based multi-finger gestures introduced in Section 3.1, we considered both microgestures and hand movements/rotations using the wrist and/or arm. Chaffangeon Caillet et al. [7] introduced μGlyph, a formal notation to describe a microgesture rigorously. Using μGlyph, a microgesture is described as a sequence of bio-mechanical movements, *e.g.,* flexion/extension, with the information of whether the finger is in contact with a surface. For example, a single-finger tap, *i.e.,* the entry and exit of a finger in contact with the opposite finger, is represented by ▼; ▲. Additional information such as the actuator and the surface can also be represented. For example, a single-index finger tap on the thumb is represented by •▼--- ; •▲--- . A multi-finger tap on the thumb is represented by •▼▼▼▼ ; •▲▲▲▲ . μGlyph can be used to represent two-handed interactions. For example, a touch of the index finger of the left hand on the thumb of the right hand is represented by _▼--- ----• . Finally, μGlyph can be used to describe contact on a surface other than a hand part by using the object glyph, customGlyph/object.pdf, regardless of whether the surface is another body part or an object. For instance, a touch of the index on the forearm is represented by: _▼--- forearm customGlyph/object.pdf while a touch of the index on a table is represented by _▼--- table customGlyph/object.pdf

As μGlyph only considers the movement of the fingers, gestures that use the wrist or the arm cannot be represented. Extending μGlyph to describe wrist movement seems straightforward because the wrist has the same biomechanical movements as the fingers, *i.e.,* extension, flexion, abduction, and adduction. However, as the information about whether the wrist is in contact is irrelevant, we only need to use the movement symbol under the hand representation in μGlyph. For example, wrist flexion is represented by:

_----
▼

Similarly, raising the hand, *i.e.,* extending the wrist with all fingers open, is represented as

■ ■ ■ ■
Op Op Op Op
■
Op
▲

. For the concise version of μGlyph, we propose to use the letter 'w' as the index of the movement glyph. Thus, wrist flexion is represented by : $_W$▼ while raising the hand is represented by $_W$▲//$_{t//i//m//t//p}$■.

However, extending μGlyph to gestures where the arm is used is not trivial because the forearm and shoulder introduce new biomechanical movements, such as pronating and supinating the forearm to rotate the wrist. Two

Table 7. Description in µGlyph of the user-defined gesture proposals from our gesture elicitation study

Gesture	µGlyph description
Index single tap in the air for "Open window"	_ ∘∘▼– – – ; _ ∘∘▲– – –
Pinch in all fingers for "Close window"	∘•▼ ∘•▼ ∘•▼ ∘•▼ ∘•▼ ●
Swipe right for "Next screen"	_ ∘•– – – – ; _ ••▶– – – ; _ •∘– – – –
Swipe left for "Previous screen"	_ ∘∘◀ ∘∘◀– –
Point index for "Play"	∘■ Op •■ Cl •■ Cl •■ Cl ∘■ Op
Pinch out two fingers for "Zoom in"	–∘▲ –∘▲– – –
Pinch in two fingers for "Zoom out"	∘–▼ ∘–▼– – –
Swipe up for "Turn volume up"	_ ∘∘▲– – –
Swipe backward "Turn volume down"	_ ••▼– – – table customGlyph/object.pdf
Raise fingers for "Pause"	∘■ Op ∘■ Op ∘■ Op ∘■ Op ∘■ Op ▲

options can be considered. The first is to introduce new symbols for all new biomechanical movements and modify the representation of the hand in µGlyph to include the forearm, elbow, and shoulder. The second is to keep µGlyph as it is and add additional information around the description of the movement and orientation of the hand in 3D, similar to GestureCard [28] or Charade [5]. However, finding the best notation for such movements is outside the scope of this paper. Nevertheless, Table 7 and Table 8 describe respectively the gestures of Figure 3 and Figure 5 with µGlyph, and extension to the gestures of the wrist, where appropriate.

8 CONCLUSION AND FUTURE WORK

In this paper, we first provided an operational definition of ring-based multi-finger gestures and compared it to other definitions, such as microgestures. Then, we conducted a gesture elicitation study with Tap Strap 2 to derive a consensus set of 18 user-defined ring-based finger gestures through participatory design and ergonomic considerations, all of them belonging to a well-defined category in a classification of these gestures. After performing a Principal Component Analysis to reduce the dimensionality of a gesture set from the 15 original dimensions of raw data captured to only 5, we implemented and compared four machine learning algorithms to show that a Quadratic Discriminant Analysis outperformed three other machine learning classifiers as well as existing recognizers without the need of Deep Learning. Through a performance analysis, we suggested some design implications. We demonstrated that TapStrapGest is an admissible solution for ring-based multi-finger gesture interaction, opening avenues for "eyes-free" or "screen-free" interaction in various domains. We therefore release two datasets acquired for ring-based multi-finger gestures, that can be used to train a recognizer for mid-air gestures. We finally show how the µGlyph notation can be effectively used to formalize most of the reported gestures, but not all, since wrist movement is not supported, opening an avenue to extend this notation.

Table 8. Description in µGlyph of the consensus set.

Gesture	µGlyph description
Pinch out	$\overset{-\circ}{\blacktriangle}\ \overset{-\circ}{\blacktriangle} - - -$
Pinch in	$\overset{\circ -}{\blacktriangledown}\ \overset{\circ -}{\blacktriangledown} - - -$
Point	$\_\ \overset{\circ\circ}{\blacktriangledown} - - -$
Swipe up	$\_\ \overset{\circ\circ}{\blacktriangle} - - -$ fast, with the palm facing down
Swipe down	$\_\ \overset{\circ\circ}{\blacktriangledown} - - -$ fast, with the palm facing down
Swipe right	$\_\ \overset{\circ\circ}{\blacktriangleright} - - -$ fast, with the palm facing down
Swipe left	$\_\ \overset{\circ\circ}{\blacktriangleleft} - - -$ fast, with the palm facing down
Pinch all fingers	$\overset{\circ\bullet}{\blacktriangledown}\ \overset{\circ\bullet}{\blacktriangledown}\ \overset{\circ\bullet}{\blacktriangledown}\ \overset{\circ\bullet}{\blacktriangledown}\ \overset{\circ\bullet}{\blacktriangledown}\ \bullet$
Raise fingers	$\overset{-\circ}{\blacktriangleleft}\ \overset{-\circ}{\blacktriangleleft}\ \overset{-}{\blacksquare}\ \overset{-\circ}{\blacktriangleright}\ \overset{-\circ}{\blacktriangleright}$

From an experimental point of view, we would like to test the same device in an extremely constrained context, to determine to what extent gesture recognition works acceptably, with precision. A comparison with similar devices that use finger gesture recognition, such as TapXR, is also worthwhile, to determine which device is best suited to which context and to obtain some benchmarking [6].

ACKNOWLEDGMENTS

Santiago Villarreal-Narvaez and Jean Vanderdonckt are supported by the EU EIC Pathfinder-Awareness Inside challenge "Symbiotik" project (1 Oct. 2022-30 Sep. 2026) under Grant no. 101071147. While the initial version of this paper was produced when Santiago Villarreal-Narvaez was in Université catholique de Louvain, the final revision took place when the author moved to University of Namur, for which we acknowledge funding of the OPTIMIS project by Pôle MecaTech (Convention nr. 8564) for Santiago Villarreal-Narvaez and Bruno Dumas. Adrien Chaffangeon Caillet is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Grant no. 466608952. He was also a scientific collaborator of LouRIM Research Institute, Université catholique de Louvain, during this period.

REFERENCES

[1] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, David Lindbauer, Alexandra Ion, Shengdong Zhao, and Jeffrey Tzu Kwan Valino Koh. 2012. *Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI.* Technical Report MSR-TR-2012-111. Microsoft. https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types-for-hci/

[2] Alexandru-Tudor Andrei, Laura-Bianca Bilius, and Radu-Daniel Vatavu. 2024. Take a Seat, Make a Gesture: Charting User Preferences for On-Chair and From-Chair Gesture Input. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 555, 17 pages. https://doi.org/10.1145/3613904.3642028

[3] Alexandru-Tudor Andrei, Alexandru-Ionut Siean, and Radu-Daniel Vatavu. 2022. Tap4Light: Smart Lighting Interactions by Tapping with a Five-Finger Augmentation Device. In *13th Augmented Human International Conference* (Winnipeg, MB, Canada) *(AH '22)*. Association for Computing Machinery, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3532525.3532535

[4] Victor R. Basili and David M. Weiss. 1984. A Methodology for Collecting Valid Software Engineering Data. *IEEE Transactions on Software Engineering* SE-10, 6 (1984), 728–738. https://doi.org/10.1109/TSE.1984.5010301

[5] Thomas Baudel and Michel Beaudouin-Lafon. 1993. Charade: Remote Control of Objects Using Free-Hand Gestures. *Commun. ACM* 36, 7 (jul 1993), 28–35. https://doi.org/10.1145/159544.159562

[6] F. M. Caputo, S. Burato, G. Pavan, T. Voillemin, H. Wannous, J. P. Vandeborre, M. Maghoumi, E. M. Taranta II, A. Razmjoo, J. J. LaViola Jr., F. Manganaro, S. Pini, G. Borghi, R. Vezzani, R. Cucchiara, H. Nguyen, M. T. Tran, and A. Giachetti. 2019. Online Gesture Recognition. In *Proceedings of Eurographics Workshop on 3D Object Retrieval*, Silvia Biasotti, Guillaume Lavoué, and Remco Veltkamp (Eds.). The Eurographics Association, 93–102. https://doi.org/10.2312/3dor.20191067

[7] Adrien Chaffangeon Caillet, Alix Goguey, and Laurence Nigay. 2023. μGlyph: a Microgesture Notation. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*, Albrecht Schmidt, Kaisa Väänänen, Tesh Goyal, Per Ola Kristensson, Anicia Peters, Stefanie Mueller, Julie R. Williamson, and Max L. Wilson (Eds.). Association for Computing Machinery, New York, NY, USA, 269:1–269:28. https://doi.org/10.1145/3544548.3580693

[8] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3403–3414. https://doi.org/10.1145/2858036.2858589

[9] Jacob Cohen. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, New York, NY, USA. https://doi.org/10.4324/9780203771587

[10] David Colquhoun. 2019. The False Positive Risk: A Proposal Concerning What to Do About p-Values. *The American Statistician* 73, sup1 (2019), 192–201. https://doi.org/10.1080/00031305.2018.1529622 arXiv:https://doi.org/10.1080/00031305.2018.1529622

[11] Harrison Cook, Quang Vinh Nguyen, and Simeon Simoff. 2015. Enabling Finger-Gesture Interaction with Kinect. In *Proceedings of the 8th International Symposium on Visual Information Communication and Interaction* (Tokyo, AA, Japan) *(VINCI '15)*. Association for Computing Machinery, New York, NY, USA, 152–153. https://doi.org/10.1145/2801040.2801060

[12] Klen Copič Pucihar, Nuwan T. Attygalle, Matjaž Kljun, Christian Sandor, and Luis A. Leiva. 2022. Solids on Soli: Millimetre-Wave Radar Sensing through Materials. *Proc. ACM Human-Computer Interaction* 6, EICS, Article 156 (jun 2022), 21 pages. https://doi.org/10.1145/3532212

[13] Adrien Coyette, Sascha Schimke, Jean Vanderdonckt, and Claus Vielhauer. 2007. Trainable Sketch Recognizer for Graphical User Interface Design. In *Proceedings of 11th IFIP TC 13 International Conference on Human-Computer Interaction, INTERACT '20'07, Rio de Janeiro, Brazil, September 10-14, 2007 (Lecture Notes in Computer Science, Vol. 4662)*, Maria Cecília Calani Baranauskas, Philippe A. Palanque, Julio Abascal, and Simone Diniz Junqueira Barbosa (Eds.). Springer, 124–135. https://doi.org/10.1007/978-3-540-74796-3_14

[14] Kurtis Danyluk, Simon Klueber, Aditya Shekhar Nittala, and Wesley Willett. 2024. Understanding Gesture and Microgesture Inputs for Augmented Reality Maps. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (IT University of Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 409–423. https://doi.org/10.1145/3643834.3661630

[15] Tilman Dingler, Rufat Rzayev, Alireza Sahami Shirazi, and Niels Henze. 2018. Designing Consistent Gestures Across Device Types: Eliciting RSVP Controls for Phone, Watch, and Glasses. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Montreal, QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173993

[16] Sarah Dosdall, Katrin Angerbauer, Leonel Merino, Michael Sedlmair, and Daniel Weiskopf. 2022. Toward In-Situ Authoring of Situated Visualization with Chorded Keyboards. In *Proceedings of the 15th International Symposium on Visual Information Communication and Interaction* (Chur, Switzerland) *(VINCI '22)*. Association for Computing Machinery, New York, NY, USA, Article 14, 5 pages. https://doi.org/10.1145/3554944.3554970

[17] Md Ehtesham-Ul-Haque and Syed Masum Billah. 2023. Abacus Gestures: A Large Set of Math-Based Usable Finger-Counting Gestures for Mid-Air Interactions. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 3, Article 93 (sep 2023), 30 pages. https://doi.org/10.1145/3610898

[18] Alois Ferscha and Stefan Resmerita. 2007. Gestural interaction in the pervasive computing landscape. *e & i Elektrotechnik und Informationstechnik* 124, 1-2 (2007), 17–25. https://doi.org/10.1007/s00502-006-0413-4

[19] Bogdan-Florin Gheran, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2018. Gestures for Smart Rings: Empirical Results, Insights, and Design Implications. In *Proceedings of the ACM International Conference on Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*, Ilpo Koskinen, Youn-Kyung Lim, Teresa Cerratto Pargman, Kenny K. N. Chow, and William Odom (Eds.). Association for Computing Machinery, New York, NY, USA, 623–635. https://doi.org/10.1145/3196709.3196741

[20] Bogdan-Florin Gheran, Radu-Daniel Vatavu, and Jean Vanderdonckt. 2018. Ring x2: Designing Gestures for Smart Rings using Temporal Calculus. In *Proceedings of the ACM Conference Companion on Designing Interactive Systems* (Hong Kong, China) *(DIS '18 Companion)*. Association for Computing Machinery, New York, NY, USA, 117–122. https://doi.org/10.1145/3197391.3205422

[21] Bogdan-Florin Gheran, Radu-Daniel Vatavu, and Jean Vanderdonckt. 2023. New Insights into User-Defined Smart Ring Gestures with Implications for Gesture Elicitation Studies. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 216, 8 pages. https:

//doi.org/10.1145/3544549.3585590

[22] Sarthak Ghosh, Hyeong Cheol Kim, Yang Cao, Arne Wessels, Simon T. Perrault, and Shengdong Zhao. 2016. Ringteraction: Coordinated Thumb-index Interaction Using a Ring. In *Proceedings of the ACM CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 2640–2647. https://doi.org/10.1145/2851581.2892371

[23] Michael Greenacre, Patrick J.F. Groenen, Trevor Hastie, Alfonso Iodice D'Enza, Angelos Markos, and Elena Tuzhilina. 2022. Principal component analysis. *Nature Reviews Methods Primers* 2, 100 (2022), 1–12. https://doi.org/10.1038/s43586-022-00184-w

[24] Jeremy Gummeson, Bodhi Priyantha, and Jie Liu. 2014. An energy harvesting wearable ring platform for gesture input on surfaces. In *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services* (Bretton Woods, New Hampshire, USA) *(MobiSys '14)*. Association for Computing Machinery, New York, NY, USA, 162–175. https://doi.org/10.1145/2594368.2594389

[25] Chris Harrison, Julia Schwarz, and Scott E. Hudson. 2011. TapSense: enhancing finger interaction on touch surfaces. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, Santa Barbara, CA, USA, October 16-19, 2011*, Jeffrey S. Pierce, Maneesh Agrawala, and Scott R. Klemmer (Eds.). Association for Computing Machinery, New York, NY, USA, 627–636. https://doi.org/10.1145/2047196.2047279

[26] Douglas M. Hawkins. 2004. The Problem of Overfitting. *J. Chem. Inf. Comput. Sci.* 4, 1 (2004), 1–12. https://doi.org/10.1021/ci0342472

[27] Zhenyi He, Christof Lutteroth, and Ken Perlin. 2022. TapGazer: Text Entry with Finger Tapping and Gaze-directed Word Selection. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 337, 16 pages. https://doi.org/10.1145/3491102.3501838

[28] Marc Hesenius and Volker Gruhn. 2019. GestureCards: A Hybrid Gesture Notation. *Proc. ACM Hum.-Comput. Interact.* 3, EICS, Article 22 (jun 2019), 35 pages. https://doi.org/10.1145/3331164

[29] Shumma Jomyo, Akira Furui, Tatsuhiko Matsumoto, Tomomi Tsunoda, and Toshio Tsuji. 2024. Finger-Tapping Motion Recognition Based on Skin Surface Deformation Using Wrist-Mounted Piezoelectric Film Sensors. *IEEE Sensors Journal* 24, 11 (2024), 17876–17884. https://doi.org/10.1109/JSEN.2024.3386333

[30] Pegah Karimi, Emanuele Plebani, Aqueasha Martin-Hammond, and Davide Bolchini. 2022. Textflow: Toward Supporting Screen-free Manipulation of Situation-Relevant Smart Messages. *ACM Trans. Interact. Intell. Syst.* 12, 4, Article 31 (nov 2022), 29 pages. https://doi.org/10.1145/3519263

[31] Tauheed Khan Mohd, Motti Kelbessa, Luke Heinrichs, and Ryan Freas. 2023. A Study of Supervised Clustering Methods for Optical Mouse Trajectory Data from Tap Strap 2. In *Proceedings of 2nd International Eastern Conference on Human-Computer Interaction* (Atatürk University, Erzurum, Turkey) *(IECHCI '23)*. 1–7. https://www.researchgate.net/publication/375913824_A_Study_of_Supervised_Clustering_Methods_for_Optical_Mouse_Trajectory_Data_from_Tap_Strap_2

[32] E.M. Kleinberg. 2000. On the algorithmic implementation of stochastic discrimination. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 5 (2000), 473–490. https://doi.org/10.1109/34.857004

[33] Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. 2019. Empirical Research in Mid-Air Interaction: A Systematic Review. *International Journal of Human–Computer Interaction* 35, 18 (2019), 1747–1768. https://doi.org/10.1080/10447318.2019.1572352

[34] William H. Kruskal and W. Allen Wallis. 1952. Use of Ranks in One-Criterion Variance Analysis. *J. Amer. Statist. Assoc.* 47, 260 (1952), 583–621. https://doi.org/10.1080/01621459.1952.10483441 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1952.10483441

[35] James R. Lewis. 2002. Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human–Computer Interaction* 14, 3-4 (2002), 463–488. https://doi.org/10.1080/10447318.2002.9669130

[36] Feifei Li, Yujun Li, Baozhen Du, Hongji Xu, Hailiang Xiong, and Min Chen. 2019. A Gesture Interaction System Based on Improved Finger Feature and WE-KNN. In *Proceedings of the 4th ACM International Conference on Mathematics and Artificial Intelligence* (Chegndu, China) *(ICMAI 2019)*. Association for Computing Machinery, New York, NY, USA, 39–43. https://doi.org/10.1145/3325730.3325759

[37] Guangchuan Li, David Rempel, Yue Liu, Weitao Song, and Carisa Harris Adamson. 2021. Design of 3D Microgestures for Commands in Virtual Reality or Augmented Reality. *Applied Sciences* 11, 14 (2021). https://doi.org/10.3390/app11146375

[38] Rensis Likert. 1932. A technique for the measurement of attitudes. *Archives of Psychology* 22, 140 (1932), 55–. http://psycnet.apa.org/record/1933-01885-001

[39] Fangyu Liu, Amal A Wanigatunga, and Jennifer A Schrack. 2021. Assessment of Physical Activity in Adults Using Wrist Accelerometers. *Epidemiologic Reviews* 43, 1 (07 2021), 65–93. https://doi.org/10.1093/epirev/mxab004 arXiv:https://academic.oup.com/epirev/article-pdf/43/1/65/48355086/mxab004.pdf

[40] Jani Mäntyjärvi, Juha Kela, Panu Korpipää, and Sanna Kallio. 2004. Enabling fast and effortless customisation in accelerometer based gesture interaction. In *Proceedings of the 3rd International Conference on Mobile and Ubiquitous Multimedia* (College Park, Maryland, USA) *(MUM '04)*. Association for Computing Machinery, New York, NY, USA, 25–31. https://doi.org/10.1145/1052380.1052385

[41] David Martin, Zikang Leng, Tan Gemicioglu, Jon Womack, Jocelyn Heath, William C Neubauer, Hyeokhyen Kwon, Thomas Ploetz, and Thad Starner. 2023. FingerSpeller: Camera-Free Text Entry Using Smart Rings for American Sign Language Fingerspelling Recognition. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 86, 5 pages. https://doi.org/10.1145/3597638.3614491

[42] Reeti Mathur, Aishwarya Sheth, Parimal Vyas, and Davide Bolchini. 2019. Typing Slowly but Screen-Free: Exploring Navigation over Entirely Auditory Keyboards. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 427–439. https://doi.org/10.1145/3308561.3353789

[43] Kristian Mrazek, Brian Holton, Tanner Klein, Izan Khan, Thomas Ayele, and Tauheed Khan Mohd. 2021. The TapStrap 2: Evaluating Performance of One-Handed Wearable Keyboard and Mouse. In *Proceedings of 23rd International Conference on Human-Computer Interaction - Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence* (Virtual event) *(HCI International 2021)*, Constantine Stephanidis, Masaaki Kurosu, Jessie Y. C. Chen, Gino Fragomeni, Norbert Streitz, Shin'ichi Konomi, Helmut Degen, and Stavroula Ntoa (Eds.). Springer International Publishing, Cham, 82–95. https://doi.org/10.1007/978-3-030-90963-5_7

[44] Kristian Mrazek and Tauheed Khan Mohd. 2021. Using LSTM Models on Accelerometer Data to Improve Accuracy of Tap Strap 2 Wearable Keyboard. In *Proceedings of 13th International Conference on Intelligent Human Computer Interaction* (Kent, OH, USA) *(IHCI '20'21)*. Springer-Verlag, Berlin, Heidelberg, 27–38. https://doi.org/10.1007/978-3-030-98404-5_3

[45] Cosmin Munteanu and Gerald Penn. 2017. Speech and Hands-free interaction: myths, challenges, and opportunities. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) *(MobileHCI '17)*. Association for Computing Machinery, New York, NY, USA, Article 66, 4 pages. https://doi.org/10.1145/3098279.3119919

[46] Dmitry Pyryeskin, Mark Hancock, and Jesse Hoey. 2012. Comparing elicited gestures to designer-created gestures for selection above a multitouch surface. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/2396636.2396638

[47] Md. Ashikur Rahman, Yousuf Mia, Mizanur Rahman Masum, Dm. Mehedi Hasan Abid, and Tariqul Islam. 2022. Real Time Human Activity Recognition from Accelerometer Data using Convolutional Neural Networks. In *Proceedings of 7th International Conference on Communication and Electronics Systems* (Coimbatore, India) *(ICCES '22)*. 1394–1397. https://doi.org/10.1109/ICCES54183.2022.9835797

[48] Samuel Sanford Shapiro and Martin Bradbury Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52, 3-4 (12 1965), 591–611. https://doi.org/10.1093/biomet/52.3-4.591 arXiv:https://academic.oup.com/biomet/article-pdf/52/3-4/591/962907/52-3-4-591.pdf

[49] Arthur Sluÿters, Mehdi Ousmer, Paolo Roselli, and Jean Vanderdonckt. 2022. QuantumLeap, a Framework for Engineering Gestural User Interfaces based on the Leap Motion Controller. *Proc. ACM Human-Computer Interaction* 6, EICS, Article 161 (jun 2022), 47 pages. https://doi.org/10.1145/3532211

[50] Arthur Sluÿters, Quentin Sellier, Jean Vanderdonckt, Vik Parthiban, and Pattie Maes. 2023. Consistent, Continuous, and Customizable Mid-Air Gesture Interaction for Browsing Multimedia Objects on Large Displays. *International Journal of Human–Computer Interaction* 39, 12 (2023), 2492–2523. https://doi.org/10.1080/10447318.2022.2078464 arXiv:https://doi.org/10.1080/10447318.2022.2078464

[51] Paul Streli, Jiaxi Jiang, Andreas Rene Fender, Manuel Meier, Hugo Romat, and Christian Holz. 2022. TapType: Ten-finger text entry on everyday surfaces via Bayesian inference. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 497, 16 pages. https://doi.org/10.1145/3491102.3501878

[52] Zhida Sun, Sitong Wang, Chengzhong Liu, and Xiaojuan Ma. 2022. Metaphoraction: Support Gesture-based Interaction Design with Metaphorical Meanings. *ACM Trans. Comput.-Hum. Interact.* 29, 5, Article 45 (oct 2022), 33 pages. https://doi.org/10.1145/3511892

[53] Alaa Tharwat. 2016. Linear vs. quadratic discriminant analysis classifier: a tutorial. *Int. J. Appl. Pattern Recognit.* 3, 2 (2016), 145–180. https://doi.org/10.1504/IJAPR.2016.079050

[54] Ayumu Tsuboi, Mamoru Hirota, Junki Sato, Masayuki Yokoyama, and Masao Yanagisawa. 2017. A proposal for wearable controller device and finger gesture recognition using surface electromyography. In *Proceedings of SIGGRAPH Asia 2017 Posters* (Bangkok, Thailand) *(SA '17)*. Association for Computing Machinery, New York, NY, USA, Article 9, 2 pages. https://doi.org/10.1145/3145690.3145731

[55] Jean Vanderdonckt, Paolo Roselli, and Jorge Luis Pérez-Medina. 2018. !FTL, an Articulation-Invariant Stroke Gesture Recognizer with Controllable Position, Scale, and Rotation Invariances. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (Boulder, CO, USA) *(ICMI '18)*. Association for Computing Machinery, New York, NY, USA, 125–134. https://doi.org/10.1145/3242969.3243032

[56] Radu-Daniel Vatavu and Laura-Bianca Bilius. 2021. GestuRING: A Web-based Tool for Designing Gesture Input with Rings, Ring-Like, and Ring-Ready Devices. In *Proceedings of The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21:)*, Jeffrey Nichols, Ranjitha Kumar, and Michael Nebeling (Eds.). Association for Computing Machinery, New York, NY, USA, 710–723. https://doi.org/10.1145/3472749.3474780

[57] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1325–1334. https://doi.org/10.1145/2702123.2702223

[58] Santiago Villarreal-Narvaez, Arthur Sluÿters, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2024. Brave New GES World: A Systematic Literature Review of Gestures and Referents in Gesture Elicitation Studies. *Comput. Surveys* 56, 5 (2024), 128:1–128:55. https://doi.org/10.1145/3636458

[59] Santiago Villarreal-Narvaez, Jean Vanderdonckt, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2020. A Systematic Review of Gesture Elicitation Studies: What Can We Learn from 216 Studies?. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 855–872. https://doi.org/10.1145/3357236.3395511

[60] Cong Wang, Deepak Vungarala, Kevin Navarro, Neel Adwani, and Tao Han. 2023. The Pinch Sensor: An Input Device for In-Hand Manipulation with the Index Finger and Thumb. In *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. 822–827. https://doi.org/10.1109/AIM46323.2023.10196268

[61] David Way and Joseph Paradiso. 2014. A Usability User Study Concerning Free-Hand Microgesture and Wrist-Worn Sensors. In *Proceedings of the 11th International Conference on Wearable and Implantable Body Sensor Networks* (Zurich, Switzerland) *(BSN '14)*. IEEE Press, 138–142. https://doi.org/10.1109/BSN.2014.32

[62] Mathias Wilhelm, Daniel Krakowczyk, and Sahin Albayrak. 2020. PeriSense: Ring-Based Multi-Finger Gesture Interaction Utilizing Capacitive Proximity Sensing. *Sensors* 20, 14 (2020). https://doi.org/10.3390/s20143990

[63] Mathias Wilhelm, Daniel Krakowczyk, Frank Trollmann, and Sahin Albayrak. 2015. eRing: multiple finger gesture recognition with one ring using an electric field. In *Proceedings of the 2nd International Workshop on Sensor-Based Activity Recognition and Interaction* (Rostock, Germany) *(iWOAR '15)*. Association for Computing Machinery, New York, NY, USA, Article 7, 6 pages. https://doi.org/10.1145/2790044.2790047

[64] Jacob O. Wobbrock, Krzysztof Z. Gajos, Shaun K. Kane, and Gregg C. Vanderheiden. 2018. Ability-based design. *Commun. ACM* 61, 6 (may 2018), 62–71. https://doi.org/10.1145/3148051

[65] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[66] Haijun Xia, Michael Glueck, Michelle Annett, Michael Wang, and Daniel Wigdor. 2022. Iteratively Designing Gesture Vocabularies: A Survey and Analysis of Best Practices in the HCI Literature. *ACM Trans. Comput. Hum. Interact.* 29, 4 (2022), 37:1–37:54. https://doi.org/10.1145/3503537

[67] Hua Yu and Jie Yang. 2001. A direct LDA algorithm for high-dimensional data — with application to face recognition. *Pattern Recognition* 34, 10 (2001), 2067–2070. https://doi.org/10.1016/S0031-3203(00)00162-X

[68] Bo Zhang, Lei Zhang, Mojun Wu, and Yan Wang. 2021. Using Auto-Encoder Neural Networks for Memory Fault Tolerance in Gesture Recognition System. In *Proceedings of 6th International Conference on Mathematics and Artificial Intelligence* (Chengdu, China). Association for Computing Machinery, New York, NY, USA, 25–33. https://doi.org/10.1145/3460569.3460571

[69] Mingrui Ray Zhang, Shumin Zhai, and Jacob O. Wobbrock. 2022. TypeAnywhere: A QWERTY-Based Text Entry Solution for Ubiquitous Computing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 339, 16 pages. https://doi.org/10.1145/3491102.3517686

[70] Hao Zhou, Taiting Lu, Yilin Liu, Shijia Zhang, Runze Liu, and Mahanth Gowda. 2023. One Ring to Rule Them All: An Open Source Smartring Platform for Finger Motion Analytics and Healthcare Applications. In *Proceedings of the 8th ACM/IEEE Conference on Internet of Things Design and Implementation* (San Antonio, TX, USA) *(IoTDI '23)*. Association for Computing Machinery, New York, NY, USA, 27–38. https://doi.org/10.1145/3576842.3582382

Received 16 February 2024; revised 23 September 2024; revised 30 September 2024; accepted 3 October 2024