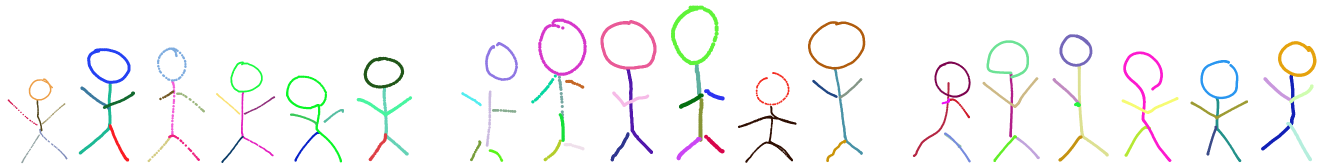
Synthesizing Stroke Gestures Across User Populations: A Case for Users with Visual Impairments

Luis A. Leiva
Sciling
Valencia, Spain
name@sciling.com

Daniel Martín-Albo
PRHLT Research Center
Universitat Politècnica de València, Spain
damarsi1@upv.es

Radu-Daniel Vatavu
MintViz Lab | MANSiD
University Ştefan cel Mare of Suceava, Romania
vatavu@eed.usv.ro



(a) (b) (c)

Figure 1. Examples of stroke gestures produced on touchscreens by people without visual impairments (a), people with visual impairments (b), and *synthetic gestures* (c) generated with our method, which automatically computes and employs the gesture articulation characteristics of a target population (b) to restyle the visual and kinematic appearance of gesture templates produced by another population (a). This way, we can generate synthetic gestures that have the same articulation characteristics as the originals (b), even if templates come from people outside that population (a).

ABSTRACT
We introduce a new principled method grounded in the Kinematic Theory of Rapid Human Movements to automatically generate synthetic stroke gestures *across user populations* in order to support ability-based design of gesture user interfaces. Our method is especially useful when the target user population is difficult to sample adequately and, consequently, when there is not enough data to train gesture recognizers to deliver high levels of accuracy. To showcase the relevance and usefulness of our method, we collected gestures from people *without* visual impairments and successfully synthesized gestures with the articulation characteristics of people *with* visual impairments. We also show that gesture recognition accuracy improves significantly when using our synthetic gesture samples for training. Our contributions will benefit researchers and practitioners that wish to design gesture user interfaces for people with various abilities by helping them prototype, evaluate, and predict gesture recognition performance without having to expressly recruit and involve people with disabilities in long, time-consuming gesture collection experiments.

Author Keywords
Gesture Synthesis; Bootstrapping; Kinematic Theory; Sigma-Lognormal Model; Rapid Prototyping; Touch Gestures

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces; I.5.2 Pattern Recognition: Design Methodology

CHI 2017, May 06–11, 2017, Denver, CO, USA

ISBN 978-1-4503-4655-9/17/05…$15.00

DOI: http://dx.doi.org/10.1145/3025453.3025906

INTRODUCTION
The popularity of stroke gesture input for graphical user interfaces has vastly increased with the prevalence of touchscreen devices. Stroke gestures represent fast movements produced by one or more fingers in contact with a touch-sensitive surface that reports a temporal sequence of $\{x, y\}$ coordinates mapped to a specific action in the user interface [26]. Compared to traditional interactions based on mouse and keyboard input, gestures have the potential to reduce users' cognitive load and visual attention [6,72] and to increase usability by replacing standard shortcuts with more accessible function triggers [27]. As touch interfaces become even more ubiquitous, it is crucial to provide equal access for people with all abilities, such as people with visual impairments, who face considerable challenges interacting with touchscreens that expose interfaces almost exclusively designed for visual input [7,22,23,44].

Previous research has highlighted many differences between the gesture preferences and articulations of people with and without visual impairments and, thus, valuable design guidelines are available today for practitioners to rely on [11,23]. Nevertheless, designing gesture-based user interfaces for people with visual impairments is still challenging because of the limited understanding in the community regarding users' gesture articulation and adoption in actual practice. This limited understanding is caused by little data available in the literature to inform how people with visual impairments actually produce touch gestures [11] and how accurately their gesture articulations can be recognized [23]. This state of things is exacerbated by the fact that access to participants with visual impairments to repeatedly collect gesture data and evaluate touch gesture interfaces as part of the iterative design cycle is a demanding process. Moreover, to our best knowledge, there are no public gesture datasets for the community to build on and to advance knowledge quickly in this direction, despite the importance of designing for accessible touch input.

We propose a method to *automatically generate synthetic gestures* that exhibit the articulation characteristics of a particular user population by using gesture templates collected from people outside that population. To achieve this goal, we use the Kinematic Theory [46] and its associated Sigma-Lognormal model ($\Sigma\Lambda$M). To showcase our method, we quantify the variation in gesture articulation produced by people *with* visual impairments, which we apply to gesture templates produced by people *without* visual impairments. Our greater goal is to automatically synthesize gesture training sets for any user population starting from a few gesture executions that are readily available, such as those produced by designers themselves.

This paper makes the following contributions:

1. We show that significant differences exist between gestures produced by people with and without visual impairments from the perspective of their velocity profiles.
2. We develop a generic, principled method to estimate distortions in the Sigma-Lognormal model of gesture velocity profiles in order to simulate human variability in gesture articulation for a target user population.
3. We use our method to synthesize stroke gestures for people with visual impairments by using gesture samples collected from people without visual impairments, showing that (i) synthetic gestures possess the same statistical characteristics as the originals produced by people with visual impairments, and (ii) synthetic gestures do increase recognition accuracy significantly.

Our method and the accompanying software are of special relevance to practitioners who wish to develop gesture-driven applications tailored to users with various gesture articulation abilities, without having to expressly recruit and involve users in preliminary, time-consuming collection experiments. Ultimately, our results can be used to inform theoretical and practical developments to synthesize stroke gestures for potentially any user population.

RELATED WORK

We discuss in this section previous work on gesture analysis, gesture interfaces for people with visual impairments, and techniques for generating synthetic gesture samples to boost the classification accuracy of gesture recognizers.

Techniques and tools for gesture articulation analysis

Users' stroke gesture articulations have been studied in the literature in terms of consistency between and within users [5], gesture preferences of various user populations [23,39,62,71], and the impact of gesture implementers, such as finger vs. pen input, or gesture articulation performance [5,60]. Fine-grained analyses of users' gesture articulations are also possible to understand how users vary their gestures relative to each other and relative to recognizers' canonical template forms [65,66]. Also, gesture recognition algorithms have been using many gesture features, such as path length, articulation time, or average speed [8,52], that can be repurposed as gesture performance measures to evaluate users' gesture articulations [51,68]. However, most of these measures, although useful for gesture classification, lack descriptive power for gesture analysis because they focus on the global characteristics of a gesture as a whole. In contrast, the Gesture RElative Accuracy Toolkit (GREAT) [65], which we employ in this work, enables access to fine-grained measurements on the gesture path that reveal and help understand the subtleties of users' gesture articulations. More specifically, the GREAT measures describe the many ways in which gestures unfold in time, space, stroke structure, and appearance, characterizing gesture articulations in terms of their closeness to a reference form, analogous to MacKenzie et al.'s accuracy measures for pointing tasks [31].

Gesture interfaces for people with visual impairments

The literature on designing accessible touch interfaces for people with visual impairments has focused significantly on applications and interaction techniques [7,10,44], while studies on how people with visual impairments use touch input or articulate gestures have been scarce. Nevertheless, the studies that exist have reported valuable and useful data. For example, Kane et al. [23] showed that blind people prefer gestures that use an edge or a corner of the device and Buzzi et al. [11,12] reported preferences for round-shaped gestures, one-finger input, one-stroke gestures, and short trajectories. Detailed examination of gesture articulation paths [23] showed that blind people produce touch gestures that are different in size, speed, number of strokes, and gesture shape than the gestures produced by sighted people. Our work looks in more depth at the differences between gestures articulated by people with and without visual impairments by considering the new perspective of the *velocity profiles* of the hand producing touch gestures in the context of the formalism of the Kinematic Theory [46].

Bootstrapping gestures by automatic synthesis

The amount and quality of training data are key factors for competitive gesture recognition. For example, the Freehand Formula Entry System [55] suggests 20–40 examples per symbol per user, and classifiers become more accurate when retrained with new samples [2,49]. Consequently, synthesizing new samples can improve recognition performance effectively.

Several techniques have been proposed in the literature to produce synthetic gestures with the goal to speed up development and to increase the accuracy of gesture recognizers. For example, Gesture Script [30] allows developers to describe the structure of a stroke gesture and, by using this information, the tool can synthesize new gesture samples by varying the relative scale and rotation of the gesture's components. Unfortunately, Gesture Script only works with unistroke gestures articulated in predefined ways. MAGIC Summoning [24] and Gesture Follower [13] are other tools that enable designers with means to generate synthetic gesture samples in 3D. MAGIC Summoning adds local perturbations to a gesture's resampled points, whereas Gesture Follower introduces variations into a gesture shape by using Viviani's 2/3 power law [69]. Both approaches are promising, although synthetic gestures might perform poorly for gesture recognition because of insufficient variation required for high-quality training [49]. However, this prior work has put forward the importance of increasing gesture recognition accuracy with large training datasets.

Probably the most relevant prior work for our method are two compelling approaches to produce synthetic stroke gestures: GPSR [58] and G3 [27]. GPSR is strongly focused on rapid UI prototyping, is computationally efficient, and adds minimal coding overhead. However, GPSR does not synthesize timestamps, which precludes a fine-grained analysis of handwriting behavior [34]. In contrast, G3 relies on the Kinematic Theory and, consequently, takes a more generic approach to gesture synthesis. G3 creates a model of a user-provided gesture example to which it adds local and global perturbations. Although resulting gestures are human-like [28], G3 employs a set of generic $\Sigma\Lambda$M parameters acquired from people without disabilities [18] and, consequently, it is unlikely that these variability ranges would also account for the actual variability attributed to people with various gesture articulation abilities.

OVERVIEW OF THE KINEMATIC THEORY

Many models have been proposed to study human movement production, such as models relying on neural networks [9], behavioral models [59], and models exploiting minimization principles [16]. Among these, the Kinematic Theory [46] provides a solid and well-established framework to study human movement production [47] and previous work showed that it outperforms many other approaches [49]. $\Sigma\Lambda$M is the latest instantiation of this framework [48], which was recently adopted for gesture synthesis and recognition. Leiva et al. [27] showed that synthesized gestures achieve similar recognition accuracy as their human counterparts and Plamondon et al. [48] showed that $\Sigma\Lambda$M generalizes to any type of human movements, including wrist movements and eye saccades.

Mathematical formulation

At a high level, $\Sigma\Lambda$M assumes that a complex handwritten trace (e.g., a character, word, signature, or gesture) is composed of a series of primitives[1] connecting a sequence of virtual targets. This series of primitives form the "action plan" of the user, which is fed through the neuromuscular network to produce a trajectory that leaves a handwritten trace on the supporting surface, such as a touchscreen; see Figure 2.

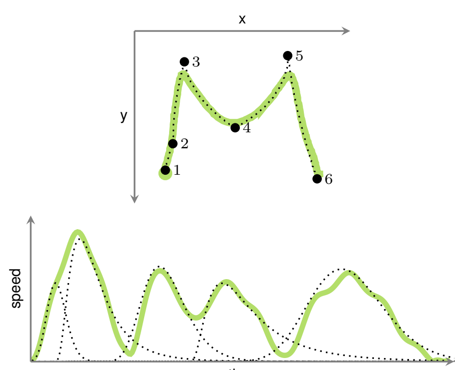


Figure 2. Top: A gesture stroke (solid green line) is described by a series of primitives (dotted arcs) that connect virtual targets (numbered dots). Bottom: each primitive is described by a lognormal velocity profile.

[1]In the gesture recognition literature, a "stroke" denotes the trajectory between two consecutive touch down and touch up events. For the Kinematic Theory, a "stroke" is what we call a "primitive" in this paper.

The magnitude of the velocity of the i-th primitive is described by a lognormal function scaled in amplitude by a command parameter D_i and time-shifted by the time occurrence t_{0_i} of the command:

$$\|\vec{v}_i(t)\| = D_i \Lambda(t; t_{0_i}, \mu_i, \sigma_i^2) = \frac{D_i}{\sigma_i \sqrt{2\pi}(t - t_{0_i})} \exp\left(\frac{-[\ln(t - t_{0_i}) - \mu_i]^2}{2\sigma_i^2} \right) \quad (1)$$

where μ_i and σ_i encode the variability of the neuromuscular execution of the i-th motor command. The trajectory that produces the human movement $\vec{v}(t)$ is computed as the temporal overlap of each primitive's velocity $\vec{v}_i(t)$:

$$\vec{v}(t) = \sum_{i=1}^{N} \vec{v}_i(t) = \sum_{i=1}^{N} \begin{bmatrix} \cos\phi_i(t) \\ \sin\phi_i(t) \end{bmatrix} D_i \Lambda(t; t_{0_i}, \mu_i, \sigma_i^2) \quad (2)$$

where the angular position $\phi_i(t)$ is given by:

$$\phi_i(t) = \theta_{s_i} + \frac{\theta_{e_i} - \theta_{s_i}}{2} \left[1 + \text{erf}\left(\frac{\ln(t - t_{0_i}) - \mu_i}{\sigma_i \sqrt{2}} \right) \right] \quad (3)$$

with θ_{s_i} and θ_{e_i} representing the start and end angles of the i-th primitive. The reconstruction of the original gesture trajectory is computed using the following compact notation [45]:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \sum_{i=1}^{N} \frac{D_i}{\theta_{e_i} - \theta_{s_i}} \begin{bmatrix} \sin\phi_i(t) & - & \sin\theta_{s_i} \\ -\cos\phi_i(t) & + & \cos\theta_{s_i} \end{bmatrix} \quad (4)$$

Human-like gesture synthesis

Previous work has demonstrated the connection between the distortion of the $\Sigma\Lambda$M parameters and the intra-variability of human handwriting [14], which enables generation of realistic, human-like synthetic samples [28]. Once the gesture primitives have been extracted and modeled, perturbations can be introduced to the model's parameters [27,35]:

$$p_i^* = p_i + n_{p_i} \quad (5)$$

where $p_i = \{\mu_i, \sigma_i, D_i, \theta_{s_i}, \theta_{e_i}\}$ denote the $\Sigma\Lambda$M parameters and $n_{p_i} = \mathcal{U}(-n_i, n_i)$ the noise applied to each primitive according to a uniform distribution centered around that particular $\Sigma\Lambda$M parameter (which we discuss later in the 'Gesture Synthesis across user populations' section). Variations in μ and σ mimic peripheral noise, like a writer who instantiates the same gesture intention, but executes it slightly differently each time. Variations in D, θ_s, and θ_e refer to central fluctuations that occur in the position of the virtual targets of the action plan from one execution to another. We leave the t_{0_i} parameter unmodified, because it is very sensitive even to small perturbations [27].

Until now, researchers have relied on a predefined set of distortion values for $\Sigma\Lambda$M that were estimated from a population of users without disabilities [18]. However, it is unlikely that these values can be used to generate accurate samples for users with various gesture articulation abilities and, therefore, a new method to estimate distortion values for $\Sigma\Lambda$M is needed.

EVALUATION

We conducted three experiments (i) to understand how people with and without visual impairments articulate stroke gestures by using concepts and tools from the Kinematic Theory and (ii) to evaluate how accurately gestures produced by our two groups of participants can be recognized.

1. **Gesture modeling.** In the first experiment, we analyze the motor control aspects of our participants' gesture articulations (e.g., the velocity of the finger touching the screen) by using the Kinematic Theory, and we report and discuss *significant differences between people with and without visual impairments in terms of their gesture velocity profiles*.
2. **Gesture articulation and synthesis.** In this experiment, we employ state-of-the-art gesture accuracy measures [65] to reveal even more differences between the gesture articulations of people with and without visual impairments. We also generate and analyze *synthetic gestures for people with visual impairments*.
3. **Gesture recognition.** In this final experiment, we compare the recognition performance of two gesture recognizers and we show that *using training samples from people without visual impairments increases the accuracy of recognizing gestures produced by people with visual impairments.*

Participants

We recruited a group of 10 participants (3 female) with visual impairments and another group of 10 participants (7 female) without visual impairments. Both groups had approximately the same average age: $M = 37.4$ years ($SD = 9.6$) for participants with visual impairments and $M = 33.0$ years ($SD = 12.2$) for participants without impairments. Visual impairments consisted in congenital nystagmus (7 participants), chorioretinal degeneration (1), astigmatism (2), amblyopia (1), macular dysplasia (1), microphthalmus (1), and macular choroiditis (1), with most of the participants with visual impairments having more than one eye condition. Nine participants had *moderate* or *severe* myopia (diopters ranged from -4.0 to -18.0, $M = -11.3$, $SD = 4.3$) and one participant had *severe* hyperopia ($+6.0$ diopters in both eyes).

Dataset

We collected gesture samples for the following unistroke and multistroke gesture types:

- Four directional flicks: left, right, down, and up.
- Six multistroke gestures: spiral, circle, square, star, letter "M" and the "stick figure" symbol.

We chose these gesture types because they represent a good mixture of geometrical shapes and symbols of various shape complexity levels, complexity that ranged from 1 to 11 according to Isokoski's definition [20]. Figure 3 depicts several gestures performed by our participants.

Each participant was asked to perform 10 repetitions of each gesture on a Samsung Galaxy Tab 4 with a touchscreen display of 10.1 inches and resolution of 1280×800 px (149 dpi). Gestures were shown onscreen in a large size (5×5 cm) and were communicated verbally to the participants. A training

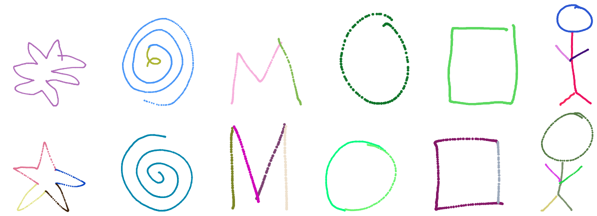


Figure 3. Examples of gestures collected from our participants with visual impairments (top row) and without visual impairments (bottom).

phase took place before the actual experiment so that participants would familiarize themselves with the device and the task. Participants were instructed to draw gestures as fast and accurately as possible. The order of gestures was randomized across participants. Overall, we collected 2 groups $\times$ 10 participants per group $\times$ 10 gesture types $\times$ 10 repetitions per gesture type = 2,000 gesture samples.

Experiment 1: Gesture modeling

Given that stroke primitives are "hidden" in the gesture shape, a $\Sigma\Lambda$M parameter extractor [36] is required to automatically detect them. To this end, we used the G3 web service [27], which computed the lognormal equations (represented by their $\Sigma\Lambda$M parameters) for the gesture samples in our dataset. Then, we assessed the gesture models with the standard performance criteria from the Kinematic Theory literature [50], as follows:

1. **Signal-to-noise ratio (SNR)** between the original and the reconstructed velocity profiles of a gesture. This measure accounts for the reconstruction quality of a gesture model.
2. **Number of extracted primitives (nbLog)**, i.e., N in Equations (2) and (4). This measure accounts for the user's handwriting and gesture articulation abilities.
3. **The lognormality principle**, measured with the ratio SNR/nbLog, which acts as a global indicator of a given user's motor control skills [50]. The lognormality principle states that users who are in perfect control of their movements produce the minimum number of ideal lognormals for their handwriting movement. In contrast, when users experience difficulties in producing a movement, the resulted primitives will not be ideal lognormals or their number will be considerably larger.

Tables 1 to 3 summarize the results of this experiment.

User group	Mdn	Mean	SD
without visual impairments	25.90	26.28	3.36
with visual impairments	25.50	25.26	2.99

Table 1. Signal-to-noise ratio values (SNR) expressed in dB. Note: larger values indicate better performance.

User group	Mdn	Mean	SD
without visual impairments	7.00	10.71	8.61
with visual impairments	12.00	19.32	21.71

Table 2. Number of primitives (nbLog). Note: smaller values indicate better performance.

User group	Mdn	Mean	SD
without visual impairments	17.20	17.28	5.53
with visual impairments	14.65	14.22	5.94

Table 3. The lognormality principle (SNR/nbLog). Note: larger values indicate better performance.

The average SNR values are above 25 dB; see Table 1. This result indicates that gestures from both groups can be successfully modeled using $\Sigma\Lambda$M. In practice, the $\Sigma\Lambda$M parameters for stroke gestures are considered to be well estimated when SNR $\geq$ 15 dB; see [27]. However, Table 2 shows that the average number of lognormals (nbLog) is higher for participants with visual impairments, which indicates a larger deviation from lognormality for this group. This rationale is also evidenced by the differences in the SNR/nbLog ratio between people with and without visual impairments; see Table 3.

Overall, the values of SNR, nbLog, and the SNR/nbLog ratio indicate better performance for participants without visual impairments: SNR $\left[t_{(193.39)}{=}4.01,\ p{<}.001,\ d{=}0.28\right]$, nbLog $\left[t_{(159.74)}{=}{-}3.18,\ p{<}.002,\ d{=}0.24\right]$, and the SNR/nbLog ratio $\left[t_{(196.56)}{=}3.73,\ p{<}.001,\ d{=}0.26\right]$. However, effect sizes suggest a small to moderate practical importance of these differences, which shows that gestures produced by the two user groups *are statistically different* according to the Kinematic Theory, but eventually *not that different* to avoid reusing gesture samples from people without visual impairments to generate gestures with the articulation characteristics of people with visual impairments. We rely on these findings in the next section, when we introduce our gesture synthesis method that works *across user populations*.

We also performed a correlation analysis between the quality of gesture reconstruction and participants' numbers of diopters, as an indicator of their visual acuity. We found significant, yet small correlations for nbLog $\left[r(98){=}{-}0.22,\ p{=}.027\right]$ and SNR/nbLog $\left[r(98){=}0.19,\ p{=}.049\right]$. These results show that, although a significant association was detected, the effect size is small, which gives hope that we can generate gestures for a wide range of visual acuity loss (specifically, between -4.0 and -18.0 diopters for the case of our participants). No significant correlation was found with SNR $\left[r(98){=}0.04,\ p{=}.638,\ n.s.\right]$, which corroborates that $\Sigma\Lambda$M can be used to successfully model gestures produced by people with visual impairments.

Experiment 2: Gesture articulation

We used GREAT [65] to compute the geometric, kinematic, and articulation accuracy of the stroke gestures produced by our two groups of participants. GREAT computes twelve gesture descriptors relative to a reference template called the "gesture task axis." In our experiments, we used the k-medoid as the reference gesture, i.e., the closest user-articulated sample to the median gesture.[2] The GREAT measures are grouped in the following categories:

1. **Geometric measures** or *shape*-related descriptors evaluate the deviation of a candidate gesture from the task axis in terms of the shape distance and capture users' tendencies to stretch and bend their gesture strokes during articulation.
2. **Kinematic measures** or *time*-related descriptors evaluate accuracy in the time domain and capture how fluent or smooth gestures are in terms of production time and speed.
3. **Articulation measures** or *consistency*-related descriptors measure how consistent users are in producing the individual strokes of their gestures.

We refer the reader to Vatavu et al. [65] for a detailed description of these measures. All gestures were uniformly resampled into 32 points to speed up computation time without sacrificing accuracy [61]. We also generated synthetic gestures for participants with visual impairments by following the traditional synthesis approach; see 'Overview of the Kinematic Theory' and [27,28]. Figure 4 shows the results of this experiment.

To understand the differences between our three experimental conditions (i.e., gestures produced by people with visual impairments, gestures produced by people without visual impairments, and synthesized gestures for people with visual impairments), we ran a one-way ANOVA test (Greenhouse-Geisser corrected to control for deviations in sphericity), followed by pairwise comparisons (Bonferroni corrected) as post-hoc tests of significance, if applicable. We observed statistically significant differences for most of the GREAT measures, marked with an asterisk in Figure 4 $\left[F_{(2,177)}{>}4,\ p{<}.001,\ \eta_p^2{<}0.2\right]$.

Post-hoc tests revealed that gestures produced by people with visual impairments were less accurate than those produced by people without visual impairments ($p < .01$). However, we found no significant difference between gestures produced by people with visual impairments and synthesized gestures ($p > .05, n.s.$). This result confirms previous findings that gestures synthesized with $\Sigma\Lambda$M look similar to gestures produced by humans [18,27,28].

Experiment 3: Gesture recognition

In this experiment, we evaluated the recognition performance of the popular Nearest-Neighbor classification approach working with the $P recognizer [64] and Dynamic Time Warping (DTW) [57]. $P represents gestures as clouds of 2D points to achieve articulation-invariant gesture recognition. DTW performs an elastic matching between two gestures, regardless their number of strokes, by computing a warping matrix of point-wise Euclidean distances.

We computed *user-independent* recognition rates by following a k-fold leave-one-out procedure: for each participant, we created a testing set containing all their gesture samples, which were classified against a training set composed of T training samples for each gesture type that were selected at random from all the remaining participants. The number of training samples T was varied from 1 to 5 and 10. Each gesture from the dataset was treated as a candidate gesture at least once. All gestures were resampled into 32 points prior to recognition. Figure 5 shows the results of this experiment.

We found that training recognizers with gestures produced by people with visual impairments did not achieve sufficient levels of recognition accuracy for practical use; for example,

[2]We modified GREAT to compute the k-medoid gesture.

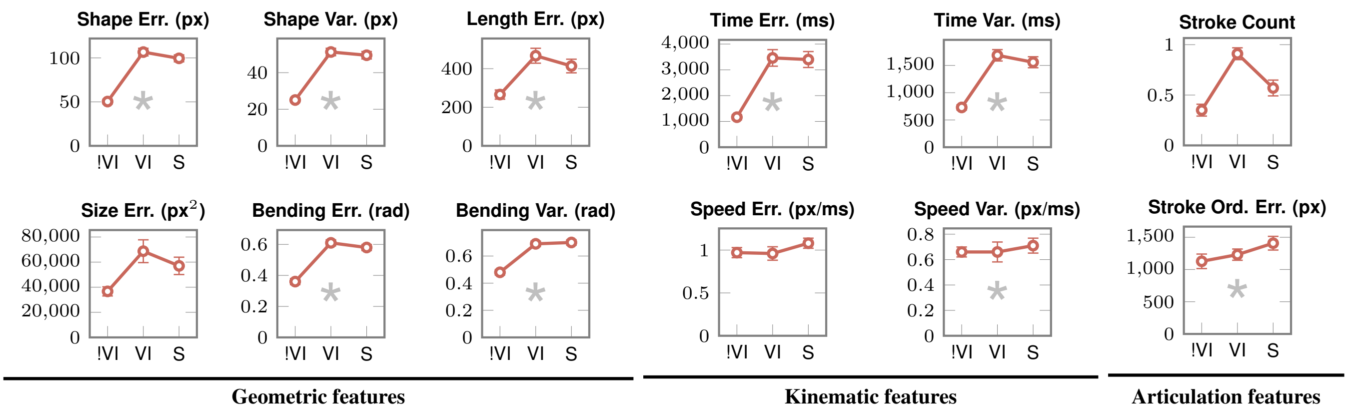


Figure 4. Articulation performance of gestures performed by people with (VI) and without (!VI) visual impairments and synthetic gestures (S) generated from people with visual impairments. Error bars denote 95% confidence intervals. An asterisk denotes statistically significant differences.

the average classification error for DTW was above 40%. In contrast, selecting training samples from the gestures produced by people without visual impairments decreased classification error on average, from 45% to 36% for DTW and from 36% to 27% for $P. The traditional approach to synthesizing gestures within the same population [18] by using examples collected from people with visual impairments achieved better results for both recognizers, and this was so when using any number of training samples.

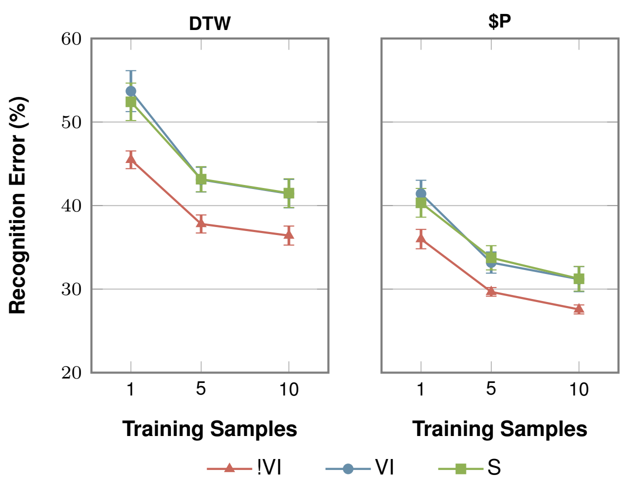


Figure 5. Recognition error rates for gestures produced by participants with visual impairments using different training sets composed of gestures from people with (VI) and without (!VI) visual impairments and synthetic gestures (S). Error bars show 95% confidence intervals.

We used the χ^2 test to attempt rejection of the null hypothesis that "the three training conditions lead to equal recognition performance." The test revealed statistical significance $\left[\chi^2_{(2,N=1000)}{=}13.99,\ p{<}.001,\ \phi{=}0.11\right]$, showing that at least one training condition was significantly different from the others. Post-hoc tests showed that training with gesture samples from people without visual impairments (!VI in Figure 5) significantly improved recognition performance for both recognizers using any number of templates $(p < .001)$.

Evaluation summary

Our first experiment showed that gestures produced by people with and without visual impairments are different in terms of their velocity profiles, but the small effect sizes indicate that gesture synthesis across the two user groups may be possible. In the second experiment, we found that synthetic gestures generated for people with visual impairments have the same articulation characteristics as original, human-generated gestures. These results support our motivation to introduce gesture synthesis *across user populations*. We also know that recognition error rates are higher when using training samples from people *with* visual impairments, but they decrease when training *across populations*. Consequently, we expect that by synthesizing gestures for people with visual impairments using templates collected from people without visual impairments will result in a boost of recognition accuracy.

GESTURE SYNTHESIS ACROSS USER POPULATIONS

Informed by the results of our evaluation, we introduce a new method to estimate the human variability of ΣΛM values for different user populations. According to previous work [19,38], when a human produces a very rapid stroke, the trajectory is nearly straight and there might be up to two reversals in the direction of the motion (known as "glitches"), either at the beginning or at the end of the trajectory. Therefore, the velocity profile can have up to three primitives (each described with a lognormal), with one dominating the others in terms of amplitude. Moreover, when a user repeats the same rapid movement many times, some variability is expected and observed, although each individual trajectory still has a dominant primitive as long as there is no trembling or hesitation. These observations are key to our method.

To delve into the principle of our approach, we focus on *directional flick gestures* (which are very rapid movements) to determine the expected variability ranges of a given user population. Flick gestures can be easily aligned at the stroke level since they are performed similarly by all users.

Estimating gesture variability

We start with the ΣΛM reconstruction of each directional flick gesture using the G3 web service [27,32]. We discard all gesture reconstructions that contain more than three primitives (16% of the gestures in the dataset), given that those cannot be considered rapid movements, according to the previous discussion. Next, we normalize the $\Sigma\Lambda$M parameters of each reconstructed gesture to make them scale and orientation independent [14], as follows:

$$\begin{aligned} D_i &\rightarrow D_i/D_{\text{max}} \\ \theta_{s_i} &\rightarrow \theta_{s_i} - \theta_{s_1} \\ \theta_{e_i} &\rightarrow \theta_{e_i} - \theta_{s_1} \\ \mu_i &\rightarrow \bar{\mu} = \frac{1}{N}\sum_{i=1}^{N}\mu_i \\ \sigma_i &\rightarrow \bar{\sigma} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i \end{aligned} \tag{6}$$

Each control parameter D_i is normalized by the maximum amplitude of all gestures (D_{max}). The start and end angles (θ_{s_i}, θ_{e_i}) are normalized by the initial angle θ_{s_1}. The peripheral parameters (μ_i, σ_i) are reduced to their mean values ($\bar{\mu}, \bar{\sigma}$). Then, to measure the variability of the $\Sigma\Lambda$M parameters, we focus on the dominant primitive of each gesture and, consequently, we discard the parameters pertaining to the glitches, which are potentially noisy. We identify the dominant primitive as the one with largest amplitude. Finally, we aggregate the parameters and compute their distributions.

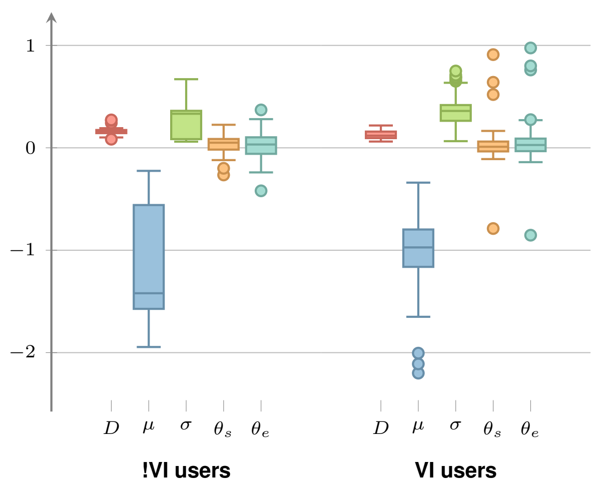


Figure 6. Distributions of the $\Sigma\Lambda$M parameters for participants with (VI) and without (!VI) visual impairments. Error bars show 95% CIs.

Figure 6 shows the estimated $\Sigma\Lambda$M parameter distributions for our two user populations. Among all $\Sigma\Lambda$M parameters, differences between medians are mostly important for μ (!VI *Mdn* = -0.32 vs. VI *Mdn* = -0.17). With respect to σ, we observed that medians were approximately the same (!VI *Mdn* = 0.33 vs. VI *Mdn* = 0.36). An increase in either μ or σ reflects slowing down caused by the neuromuscular units generating the response. After all, the generation of human movements is a complex neuromotor skill requiring the interaction of many cognitive processes, among which eyesight plays an important role. We also found that the control parameter D (associated with the command amplitude) was affected by the presence of visual impairments (!VI *Mdn* = 0.16 vs. VI *Mdn* = 0.12). This effect is more obvious if we remove the normalization (!VI *Mdn* = 67 vs. VI *Mdn* = 43). The decrease in D was compensated by participants with visual impairments by an increase in the nbLog measure; see Table 2. In fact, visual impairments may cause hesitation during handwriting, which eventually leads to a larger number of small lognormals (larger nbLog values, smaller D's). We also observed that the distributions of θ_s and θ_e do not differ greatly. These parameters represent the start and end angles of each primitive and, as long as the execution of a directional flick remains the same, so will the values of these parameters. The unpaired t-test (two-tailed) revealed a statistically significant difference between the two user populations for D $\left[t_{(75.50)}{=}2.445,\ p{=}.017,\ d{=}0.27\right]$ and for both peripheral parameters: μ $\left[t_{(69.43)}{=}{-2.863},\ p{=}.006,\ d{=}0.32\right]$ and σ $\left[t_{(77.91)}{=}{-3.243},\ p{=}.002,\ d{=}0.34\right]$.

Synthesizing human-like gesture samples

Once the empirical distributions of the $\Sigma\Lambda$M parameters are available, we can estimate the appropriate range in which they vary for a particular user population. Concretely, we estimated the range for each parameter as half of the interquartile range (IQR). We used this statistic because it is robust and resilient to outliers (see Figure 6) in contrast to other options, such as the mean or the variance. Table 4 shows the amount of noise (Equation 5) to apply to the $\Sigma\Lambda$M parameters to synthesize gestures for people with and without visual impairments.

User group	D	μ	σ	θ_s	θ_e
without visual impairments	0.01	0.47	0.14	0.05	0.08
with visual impairments	0.03	0.19	0.09	0.04	0.06

Table 4. Range of the $\Sigma\Lambda$M parameters for both user populations.

Now we can synthesize gestures with the expected variability of a particular user population. We do this with the same procedure as described in the 'Human-like gesture synthesis' section, but this time using our specific set of distortions for the $\Sigma\Lambda$M parameters. The next section validates this approach.

VALIDATION

We validate our method by synthesizing and evaluating gesture samples for people *with* visual impairments using gesture templates from people *without* visual impairments. To this end, we conducted two validation experiments:

1. **Gesture articulation.** In this experiment, we show that the cross-population synthetic gestures generated from samples produced by people without visual impairments *are similar to the gestures produced by people with visual impairments.*
2. **Gesture recognition.** In this experiment, we show that using our new synthetic gesture samples for training *improves gesture recognition accuracy significantly.*

Note that the four directional flick gestures were not considered for validation, as they were used for estimation and fine-tuning of the $\Sigma\Lambda$M parameter distributions. Therefore,

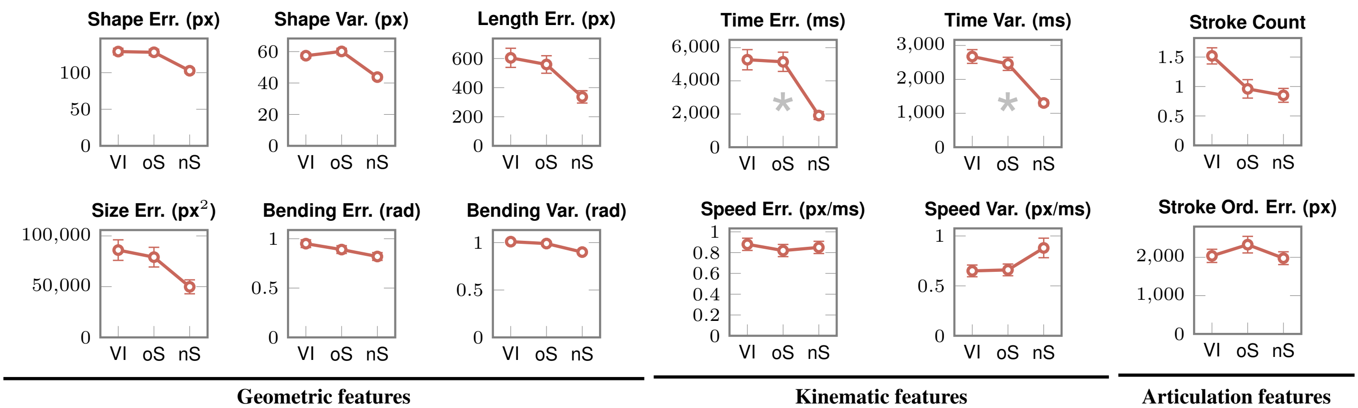


Figure 7. Articulation performance of gestures performed by people with visual impairments (VI) and samples synthesized with both the traditional approach (oS) and our new method (nS). Error bars denote 95% confidence intervals. An asterisk denotes statistically significant differences.

only six multistroke gesture types from our set (see the Dataset section) were considered for evaluation.

Experiment 4: Gesture articulation

We used the same experimental design as in the 'Experiment 2: Gesture articulation' section to analyze gestures produced by people with visual impairments and their synthetic, cross-population counterparts. We compared the traditional approach to synthesizing gestures [18,27] against our method, which transfers the variability ranges of people with visual impairments to the gesture templates of people without visual impairments. Figure 7 shows the results of this experiment.

We conducted a one-way ANOVA (Greenhouse-Geisser corrected to control for deviations in sphericity) to understand the differences between our three conditions (i.e., gestures produced by people with visual impairments, synthetic gestures generated from gesture examples produced by people with visual impairments, and synthetic gestures generated from gesture examples produced by people without visual impairments). Out of the twelve gesture articulation accuracy measures that we used for evaluation (Figure 7), we observed statistically significant differences only for the Time Error $\left[F_{(2,177)}{=}4.32,\ p{=}.014,\ \eta_p^2{=}0.05\right]$ and Time Variability measures $\left[F_{(2,177)}{=}6.39,\ p{<}.01,\ \eta_p^2{=}0.07\right]$. However, because effect sizes show small practical importance, we can conclude that our method successfully transfers the articulation characteristics of people with visual impairments (VI) to the gesture templates performed by people without impairments (!VI) and that synthetic samples produced with our method (nS) look the same as their human counterparts (VI).

Experiment 5: Gesture recognition

In this experiment, we evaluated the effect of synthetic training samples generated with our method on the recognition accuracy of gestures produced by people with visual impairments. We used the same recognizers and setup described in the 'Experiment 3: Gesture recognition' section. The results of this experiment are shown in Figure 8.

Results confirmed that training samples either from people with visual impairments or using the traditional synthesizing approach [18] does not achieve sufficient accuracy for practical use. Actually, results were similar to those achieved in the previous evaluation; see Figure 5. Our approach was on par with the other training conditions for the DTW recognizer, but delivered much better results for $P. For example, the recognition errors decreased from 31% to 17% when using 5 training templates and to 15% when using 10 training templates.

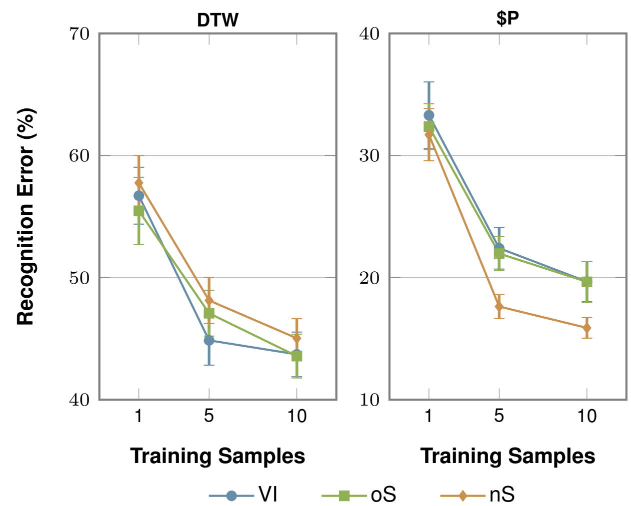


Figure 8. Recognition errors for gestures produced by people with visual impairments under different training conditions: human gestures from people with visual impairments (VI), synthetic gestures generated using the traditional approach (oS), and synthetic gestures generated with our new method (nS). Error bars denote 95% confidence intervals.

As in the previous evaluation, we used the χ^2 test to assess the differences between the three training conditions. The test did not reveal statistical significance for the DTW recognizer, but it was so for $P $\left[\chi^2_{(2,N=600)}{=}37.13,\ p{<}.001,\ \phi{=}0.24\right]$. Post-hoc tests confirmed that training with 5 or 10 gesture samples synthesized with our approach significantly improved recognition performance ($p < .001$). In other words, our method produces synthetic gestures for people with visual impairments that are sufficiently similar to explain their human variability, but not too similar to degrade recognition rates.

DISCUSSION

Our results showed that gesture synthesis *across user populations* is viable and, moreover, that gestures possess similar articulation characteristics as actual gestures created by users of the target population. We should note that our approach requires some examples of the target population in order to generate synthetic gestures. However, once a particular user population has been analyzed, their models can be reused for future studies. In fact, all the previous research on $\Sigma\Lambda$M gesture synthesis have relied on the results of just one study [18], which calculated the expected human variability ranges for people without disabilities. In short, our technique can be the seed to understand and create models of specific populations, which enables many application scenarios to assist the design of gesture user interfaces. In this section, we discuss such application scenarios and we point to opportunities for future work, directed at researchers and practitioners who wish to apply our method to other user populations.

Application scenarios

Our method enables free access to virtually unlimited gesture data for a specific user population by starting with an estimation of the gesture articulation parameters of the target population and just a few gesture examples produced by a person outside that population, e.g., the user interface designer. Specifically, we only need to collect one or two gesture examples from the user, as it has been shown to be enough for synthesizing samples that account for the variation required for high-quality training [27,35].

Cross-population gesture synthesis with G3

Informed by previous results on gesture synthesis and the $\Sigma\Lambda$M model of the Kinematic Theory, we estimate that the G3 web application [27,32] can generate about 100 unique synthetic gesture samples starting from just one example. Such gesture examples are available to generate and download from the G3 home page: **`https://g3.prhlt.upv.es`**. Our mehod has been implemented in G3, under the 'advanced options' menu (Figure 9). Therefore, the user only has to follow the usual procedure (see [27] for a guided example) and select a desired target population. At the moment, only two target populations are available for selection (*generic* and *visual impairments*), but other populations will be added in the future.

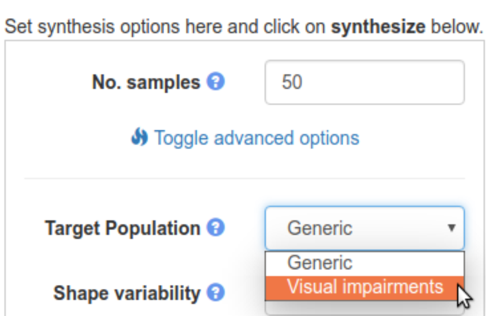


Figure 9. The G3 web application [27,32] now can generate gesture samples reflective of the gesture articulation characteristics of users with visual impairments.

Improving gesture recognition accuracy

More training templates increase the accuracy of gesture recognizers by providing them access to a much larger selection of gesture examples to cope with the variation in gesture articulation of the target user population. We know from previous work that template-based recognizers fare very well with as few as 5 samples per gesture type [41,64], either human or synthetic [27]. Overall, high recognition rates make users more effective with touchscreen input, increasing their task performance (less errors) and, potentially, user satisfaction with the interface.

Making gesture recognizers more robust

For user populations with considerable variability in gesture articulation, gestures that are more generic provide a better template for recognition, possibly because they are more representative or *prototypical*. Actually, the gestures that we synthesized in our experiments by using templates from people without disabilities provided better "average" training samples for the two gesture recognizers that we evaluated.

Another important aspect that affects recognition accuracy is the *within-group variability* in gesture articulation, which in our case was higher for participants with visual impairments than for participants without impairments. Higher within-group variability for gestures articulated by people with visual impairments may also explain why using synthetic gestures or gestures from people without visual impairments for training delivers better recognition performance than the actual gestures captured from people with visual impairments.

Designing gesture sets

The process of designing gesture sets is complex, as it involves many motor and cognitive aspects that the designer must consider, such as good discriminability with respect to other gestures in the set [1], ease of execution [43,68], ease of learning and memorability [42], good fit to application functions [39,71], etc. This process usually involves a lot of trial and error, where gestures go in and out of the gesture set while the designer optimizes the structure of the set with respect to the above criteria. Having fast access to actual gesture samples for new gestures that the designer might come up with during this process, without actually collecting them from the target user population, would have a positive effect on the designer's work, saving considerable time.

Supporting ability-based design

Our method connects with the concept of ability-based design that consists in *"focusing on ability throughout the design process in an effort to create systems that leverage the full range of human potential"* [70]. Fitting the gesture training set of a recognizer to the gesture articulation abilities of a specific user or user population is an implementation of ability-based design for gesture recognition. Furthermore, this process can now be fully automatized and launched by the application when needed, e.g., when the user adds a new gesture type to the gesture set. Consequently, new gesture samples with the particularities of the gesture articulation of that specific user will be automatically available at no cost.

Addressing other user populations

In this work, we focused specifically on people with visual impairments, for which touch interaction pose many challenges, because touchscreens rely almost exclusively on visual input [22,23,54]. However, we formalized our method in a way that is independent of the characteristics of the target population so that it would be easy to apply for synthesizing gesture sets for other user populations as well. For example, touch input remains largely inaccessible to people with motor impairments who need to adopt workaround strategies to be able to access content on touchscreen devices [4,37] and who need specific touch interaction techniques [40]. We also know from the literature of touch interaction for children that small children between 3 and 6 years old experience difficulties with touch and multitouch input [67] and that the touch gestures of children between 7 and 10 years old are recognized with lower accuracy rates than the same gestures produced by adults [3]. Because our method is able to transfer the articulation characteristics of gestures produced by a few users to a particular user population, we believe that addressing other user groups, such as those mentioned above, is viable and we leave these interesting exploration opportunities for future work.

Further application areas

In this work, we touched on a subject that may have implications in HCI and accessibility research beyond touch gesture input, and we would like to take this opportunity to mention a few of these future work opportunities. By doing this, we hope to draw the community's attention to an exciting line of work: transferring the characteristics of one user population to the input data generated by another population in order to e.g. synthesize practical templates, test cases, and accurate simulation results representative of the target user population. This approach is particularly useful and relevant when designing for people with disabilities, because it removes the need to expressly recruit and involve people with disabilities in long, time-consuming data collection experiments.

Interesting future work directions may look at the applicability of synthesizing data *across user populations* for mouse input [15,17,33], voice input [21,29], whole-body movement [53,62,63], and even EEG input [25,56]. While all these directions are definitely interesting, they are nevertheless challenging, but worth exploring in order to advance our theoretical and practical knowledge of simulating input data across different user populations toward better interface designs for users will all abilities. Looking forward, we believe that our work already forms a good demonstration of how *simulation can be used to refine the gesture design process within HCI and accessibility research* and we are eager to see how the community will pick up these ideas and use them for other application areas.

CONCLUSION

We presented a principled method to generate gesture samples for people with visual impairments using gestures collected from people without visual impairments. Our method is based on the foundations of the Kinematic Theory of Rapid Human Movements and its associated Sigma-Lognormal model. The software implementing our method is publicly available at `https://g3.prhlt.upv.es`, while the gestures dataset can be downloaded from `http://www.eed.usv.ro/~vatavu`.

We showed that our method can synthesize gestures across user populations that hold the same statistical characteristics as human gestures while improving recognition accuracy. Our method will benefit UI designers who wish to prototype gesture-driven applications tailored to users with different gesture articulation abilities, without having to expressly recruit them. Altogether, these are valuable advancements that open new opportunities for future efforts in this direction, and we look forward to see our method applied to other user populations as well. It is our hope that this work will enable better user interface designs, making touch gesture interaction more accessible to people with all abilities.

ACKNOWLEDGMENTS

We thank the anonymous CHI reviewers for their constructive and useful feedback. R.-D. Vatavu acknowledges support from the project PN-II-RU-TE-2014-4-1187 financed by UEFISCDI, Romania.

REFERENCES

1. Jr. Allan C. Long, James A. Landay, and Lawrence A. Rowe. 1999. Implications for a gesture design tool. In *Proc. CHI '99*.
2. Abdullah Almaksour, Eric Anquetil, Réjean Plamondon, and Christian O'Reilly. 2011. Synthetic handwritten gesture generation using Sigma-Lognormal model for evolving handwriting classifiers. In *Proc. IGS '11*.
3. Lisa Anthony, Quincy Brown, Berthel Tate, Jaye Nias, Robin Brewer, and Germaine Irwin. 2014. Designing smarter touch-based interfaces for educational contexts. *Pers. Ubiquit. Comput.* **18**(6).
4. Lisa Anthony, YooJin Kim, and Leah Findlater. 2013a. Analyzing user-generated youtube videos to understand touchscreen use by people with motor impairments. In *Proc. CHI '13*.
5. Lisa Anthony, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2013b. Understanding the consistency of users' pen and finger stroke gesture articulation. In *Proc. GI '13*.
6. Caroline Appert and Shumin Zhai. 2009. Using strokes as command shortcuts: Cognitive benefits and toolkit support. In *Proc. CHI '09*.
7. Shiri Azenkot, Kyle Rector, Richard Ladner, and Jacob Wobbrock. 2012. PassChords: Secure multi-touch authentication for blind people. In *Proc. ASSETS '12*.
8. Rachel Blagojevic, Samuel Hsiao-Heng Chang, and Beryl Plimmer. 2010. The power of automatic feature selection: Rubine on steroids. In *Proc. SBIM '10*.
9. Daniel Bullock and Stephen Grossberg. 1988. The VITE model: a neural command circuit for generating arm and articulator trajectories. In *Dynamic Patterns in Complex Systems*.
10. Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Loredana Martusciello. 2011. Making visual maps accessible to the blind. In *Proc. UAHCI '11*.

11. Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. 2015. Exploring visually impaired people's gesture preferences for smartphones. In *Proc. CHItaly '15*.
12. Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. 2016. Analyzing visually impaired people's touch gestures on smartphones. *Multimed. Tools Appl.*. In Press.
13. Baptiste Caramiaux, Nicola Montecchio, Atau Tanaka, and Frederic Bevilacqua. 2014. Adaptive gesture recognition with variation estimation for interactive systems. *ACM TiiS* **4**(4).
14. Moussa Djioua and Réjean Plamondon. 2009. Studying the variability of handwriting patterns using the Kinematic Theory. *Hum. Mov. Sci.* **28**(5).
15. Leah Findlater, Jon E. Froehlich, Kays Fattal, Jacob O. Wobbrock, and Tanya Dastyar. 2013. Age-related differences in performance with touchscreens compared to traditional mouse input. In *Proc. CHI '13*.
16. Tamar Flash and Neville Hogan. 1985. The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* **5**(7).
17. Clifton Forlines, Daniel Wigdor, Chia Shen, and Ravin Balakrishnan. 2007. Direct-touch vs. mouse input for tabletop displays. In *Proc. CHI '07*.
18. Javier Galbally, Réjean Plamondon, Julián Fierrez, and Javier Ortega-García. 2012. Synthetic on-line signature generation. Part II: Experimental validation. *Pattern Recogn.* **45**(7).
19. Stan Gielen, Kees van den Oosten, and Albert F. Pull ter Gunne. 1985. Relation between EMG activation patterns and kinematic properties of aimed arm movements. *J. Mot. Behav.* **17**(4).
20. Poika Isokoski. 2001. Model for unistroke writing time. In *Proc. CHI '01*.
21. Jiepu Jiang, Wei Jeng, and Daqing He. 2013. How do users respond to voice input errors?: Lexical and phonetic query reformulation in voice search. In *Proc. SIGIR '13*.
22. Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. 2009. Freedom to roam: A study of mobile device adoption and accessibility for people with visual and motor disabilities. In *Proc. ASSETS '09*.
23. Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable gestures for blind people: Understanding preference and performance. In *Proc. CHI '11*.
24. Daniel Kohlsdorf and Thad E. Starner. 2013. MAGIC Summoning: Towards automatic suggesting and testing of gestures with low probability of false positives during use. *J. Mach. Learn. Res.* **14**(1).
25. Kavita Krishnaswamy and Ravi Kuber. 2012. Toward the development of a BCI and gestural interface to support individuals with physical disabilities. In *Proc. ASSETS '12*.
26. Luis A. Leiva, Vicent Alabau, Verónica Romero, Alejandro H. Toselli, and Enrique Vidal. 2014. Context-aware gestures for mixed-initiative text editing UIs. *Interact. Comput.* **27**(1).
27. Luis A. Leiva, Daniel Martín-Albo, and Réjean Plamondon. 2016. Gestures à Go Go: Authoring synthetic human-like stroke gestures using the kinematic theory of rapid movements. *ACM T. Intel. Syst. Tec.* **7**(2).
28. Luis A. Leiva, Daniel Martín-Albo, and Réjean Plamondon. 2017. The kinematic theory produces human-like stroke gestures. *Interact. Comput*. In Press.
29. Silvia Lovato and Anne Marie Piper. 2015. "Siri, is this you?": Understanding young children's interactions with voice input systems. In *Proc. IDC '15*.
30. Hao Lü, James A. Fogarty, and Yang Li. 2014. Gesture Script: Recognizing gestures and their structure using rendering scripts and interactively trained parts. In *Proc. CHI '14*.
31. I. Scott MacKenzie, Tatu Kauppinen, and Miika Silfverberg. 2001. Accuracy measures for evaluating computer pointing devices. In *Proc. CHI '01*.
32. Daniel Martín-Albo and Luis A. Leiva. 2016. G3: bootstrapping stroke gestures design with synthetic samples and built-in recognizers. In *Proc. MobileHCI '16*.
33. Daniel Martín-Albo, Luis A. Leiva, Jeff Huang, and Réjean Plamondon. 2016b. Strokes of insight: User intent detection and kinematic compression of mouse cursor trails. *Inform. Process. Manag.* **56**(6).
34. Daniel Martín-Albo, Luis A. Leiva, and Réjean Plamondon. 2016a. On the design of personal digital bodyguards: Impact of hardware resolution on handwriting analysis. In *Proc. ICFHR '16*.
35. Daniel Martín-Albo, Réjean Plamondon, and Enrique Vidal. 2014. Training of on-line handwriting text recognizers with synthetic text generated using the kinematic theory of rapid human movements. In *Proc. ICFHR '14*.
36. Daniel Martín-Albo, Réjean Plamondon, and Enrique Vidal. 2015. Improving sigma-lognormal parameter extraction. In *Proc. ICDAR '15*.
37. Kyle Montague, Hugo Nicolau, and Vicki L. Hanson. 2014. Motor-impaired touchscreen interactions in the wild. In *Proc. ASSETS '14*.
38. Pietro Morasso. 1981. Spatial control of arm movements. *Exp. Brain Res.* **42**(2).
39. Meredith R. Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding users' preferences for surface gestures. In *Proc. GI '10*.
40. Martez E. Mott, Radu-Daniel Vatavu, Shaun K. Kane, and Jacob O. Wobbrock. 2016. Smart Touch: Improving touch accuracy for people with motor impairments with template matching. In *Proc. CHI '16*.
41. Martez E. Mott and Jacob O. Wobbrock. 2014. Beating the bubble: Using kinematic triggering in the bubble lens for acquiring small, dense targets. In *Proc. CHI '14*.
42. Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. In *Proc. CHI '13*.

43. Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2003. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In *Proc. GW '03*.
44. Uran Oh, Stacy Branham, Leah Findlater, and Shaun K. Kane. 2015. Audio-based feedback techniques for teaching touchscreen gestures. *ACM T. Access. Comput.* **7**(3).
45. Christian O'Reilly and Réjean Plamondon. 2009. Development of a Sigma-Lognormal representation for on-line signatures. *Pattern Recogn.* **42**(12).
46. Réjean Plamondon. 1995. A kinematic theory of rapid human movements. Part I: Movement representation and control. *Biol. Cybern.* **72**(4).
47. Réjean Plamondon, Adel M. Alimi, Pierre Yergeau, and Franck Leclerc. 1993. Modelling velocity profiles of rapid movements: a comparative study. *Biol. Cybern.* **69**(1).
48. Réjean Plamondon and Moussa Djioua. 2006. A multi-level representation paradigm for handwriting stroke generation. *Hum. Mov. Sci.* **25**(4–5).
49. Réjean Plamondon, Christian O'Reilly, Javier Galbally, Abdullah Almaksour, and Eric Anquetil. 2014. Recent developments in the study of rapid human movements with the Kinematic Theory: Applications to handwriting and signature synthesis. *Pattern Recogn. Lett.* 35.
50. Réjean Plamondon, Christian O'Reilly, Céline Rémi, and Thérésa Duval. 2013. The lognormal handwriter: learning, performing, and declining. *Front. Psychol.* **4**(1).
51. Yosra Rekik, Radu-Daniel Vatavu, and Laurent Grisoni. 2014. Understanding users' perceived difficulty of multi-touch gesture articulation. In *Proc. ICMI '14*.
52. Dean Rubine. 1991. Specifying gestures by example. *Proc. SIGGRAPH* **25**(4).
53. Jaime Ruiz and Daniel Vogel. 2015. Soft-constraints to reduce legacy and performance bias to elicit whole-body gestures with low arm fatigue. In *Proc. CHI '15*.
54. Kristen Shinohara and Josh Tenenberg. 2007. Observing Sara: A case study of a blind person's interactions with technology. In *Proc. ASSETS '07*.
55. Steve Smithies, Kevin Novins, and James Arvo. 2001. Equation entry and editing via handwriting and gesture recognition. *Behav. Inform. Technol.* **20**(1).
56. Seshadri Sridharan, Yun-Nung Chen, Kai-Min Chang, and Alexander I. Rudnicky. 2012. Neurodialog: An eeg-enabled spoken dialog interface. In *Proc. ICMI '12*.
57. Charles C. Tappert. 1982. Cursive script recognition by elastic matching. *IBM J. Res. Dev.* **26**(6).
58. Eugene M. Taranta, II, Mehran Maghoumi, Corey R. Pittman, and Joseph J. LaViola, Jr. 2016. A rapid prototyping approach to synthetic data generation for improved 2D gesture recognition. In *Proc. UIST '16*.
59. Arnold J.W.M. Thomassen, Paul J.G. Keuss, and Gerard P. van Galen. 1983. Motor aspects of handwriting. *Acta Psychol.* **54**(1–3).
60. Huawei Tu, Xiangshi Ren, and Shumin Zhai. 2012. A comparative evaluation of finger and pen stroke gestures. In *Proc. CHI '12*.
61. Radu-Daniel Vatavu. 2011. The effect of sampling rate on the performance of template-based gesture recognizers. In *Proc. ICMI '11*.
62. Radu-Daniel Vatavu. 2012. User-defined gestures for free-hand TV control. In *Proc. EuroiTV '12*.
63. Radu-Daniel Vatavu. 2015. Audience silhouettes: Peripheral awareness of synchronous audience kinesics for social television. In *Proc. TVX '15*.
64. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures as point clouds: a $P recognizer for user interface prototypes. In *Proc. ICMI '12*.
65. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2013. Relative accuracy measures for stroke gestures. In *Proc. ICMI '13*.
66. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2014. Gesture heatmaps: Understanding gesture performance with colorful visualizations. In *Proc. ICMI '14*.
67. Radu-Daniel Vatavu, Gabriel Cramariuc, and Doina Maria Schipor. 2015. Touch interaction for children aged 3 to 6 years: Experimental findings and relationship to motor skills. *Int. J. of Hum.-Comput St.* **74**(1).
68. Radu-Daniel Vatavu, Daniel Vogel, Géry Casiez, and Laurent Grisoni. 2011. Estimating the perceived difficulty of pen gestures. In *Proc. INTERACT '11*.
69. Paolo Viviani and Tamar Flash. 1995. Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *J. Exp. Psychol.* **21**(1).
70. Jacob O. Wobbrock, Shaun K. Kane, Krzysztof Z. Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. *ACM T. Access. Comput.* **3**(3).
71. Jacob O. Wobbrock, Meredith R. Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proc. CHI '09*.
72. Shumi Zhai, Per Ola Kristensson, Caroline Appert, Tue H. Anderson, and Xiang Cao. 2012. Foundational issues in touch-surface stroke gesture design — an integrative review. In *Foundations and Trends in Human-Computer Interaction*. Vol. 5.