Sign H3re: Symbol and X-Mark Writer Identification Using Audio and Motion Data from a Digital Pen

Maximilian Schrapel  
HCI Group, Leibniz University  
Hannover  
Hannover, Germany  
schrapel@hci.uni-hannover.de

Dennis Grannemann  
HCI Group, Leibniz University  
Hannover  
Hannover, Germany

Michael Rohs  
HCI Group, Leibniz University  
Hannover  
Hannover, Germany  
rohs@hci.uni-hannover.de

ABSTRACT
Although in many cases contracts can be made or ended digitally, laws require handwritten signatures in certain cases. Forgeries are a major challenge with digital contracts, as their validity is not always immediately apparent without forensic methods. Illiteracy or disabilities may result in a person being unable to write their full name. In this case x-mark signatures are used, which require a witness for validity. In cases of suspected fraud, the relationship of the witnesses must be questioned, which involves a great amount of effort. In this paper we use audio and motion data from a digital pen to identify users via handwritten symbols. We evaluated the performance of our approach for 19 symbols in a study with 30 participants. We found that x-marks offer fewer individual features than other symbols like arrows or circles. By training on three samples and averaging three predictions we reach a mean F1-score of $F_1 = 0.87$, using statistical and spectral features fed into SVMs.

CCS CONCEPTS
- Human-centered computing → Accessibility technologies; Gestural input.

KEYWORDS
digital pens, signature authentication, signing documents, pattern recognition, handwriting recognition, motor impairments, accessibility

ACM Reference Format:

1 INTRODUCTION
Even today, handwritten signatures are an important means of identifying a person. Although electronic signatures offer a digital alternative [20], the legal situation defines several cases in which a physical handwritten signature with a pen on paper is necessary. According to the National Telecommunications and Information Administration (NITA), cases include, for example, wills and testamentary dispositions, family law cases (e.g. divorce or adoption), insurance policy cancellations, products with significant health risks, all documents for the transport of hazardous materials and utility cancellations [87]. German law requires that every document in which rights are exercised must be signed. However, unless otherwise specified (e.g. testamentary dispositions [28]), documents can be signed digitally [27]. In addition, the finance and business sectors remain subject to physical signature requirements. Even today, businessmen travel around the world for notarized signatures in order to conclude valid contracts. In the financial sector, paper checks and debit cards require signatures for validity. Decentralized finance (DeFi) aims to offer a future alternative [22]. For instance, in contrast to banks, smart contracts on the Ethereum blockchain are not subject to human involvement. They can be seen as a type of account in the network with a balance and functions defined by code. Other accounts can irreversibly submit transactions to execute these functions [23]. Smart contracts offer various advantages, including anonymity, efficiency, transparency, and providing financial services to people without access to bank accounts [22]. However, legal enforceability is still unclear [29], implying that signatures are likely to remain important. Handwriting is unique because it is influenced by training, physiology, and further behavioral factors over a lifetime [26, 83]. Verifying the validity of a signature poses a major challenge. The mismatch of signatures on mail ballots with signatures on file led to the rejection of thousands of votes in the past 2012 and 2016 U.S. presidential elections [80]. In the same time period cases of signature fraud in Canada tripled between 2014 and 2016 [21].

These examples demonstrate the need for security precautions due to legal requirements and the risk of forged signatures. The verification of x-mark signatures in particular, which must be witnessed for legal validity, presents a major challenge. Illiteracy or disabilities may result in an individual not being able to write his or her full name. The witness of a x-mark signature serves to verify and confirm the identity of the writer. Hence, x-mark signatures do not provide sufficient security and validity for online signed contracts when no witness is involved. In cases of suspected fraud, witnesses of an x-mark signature must be investigated. Forensic methods on the x-mark handwriting are limited to stroke features like length, direction, orientation and connection as well as features related to pen and paper, such as line quality and the crossing point position [39, 43, 63].

In this paper, we present an approach to measure the individuality of handwritten symbols with predefined writing instructions for user identification. Recorded data of a digital pen when signing
a contract can support in cases of suspected fraud to identify the writer. We use measured motion data and scratching sounds from a digital pen [78]. We conducted a user study with 30 participants as a proof-of-concept and to collect a dataset of 19 different symbols for authentication. On this dataset different feature sets, which are fed into SVMs, are analyzed for their applicability for writer identification. In our tests, we focus on small training sample sizes, as it would be inconvenient for a person to write his or her x-mark signature several times before the proposed system is able to identify the handwriting. We show that by combining audio data – of the scratching sound that the pen tip makes when it moves across the paper – with 3D acceleration and gyroscope data of the pen, just three training samples of a symbol are sufficient to reach an average F1-score of over 87% on writing three test symbols. In addition, alternative symbols were identified that are easy to draw and, with our approach, reach higher F1-scores than commonly used x-marks.

2 RELATED WORK

Signature verification can be divided into offline [84] and online approaches [12]. Offline verification usually requires signature image analysis for proving one’s identity [37, 57]. Therefore, it only relies on features from the signature image. Skilled forgers imitate the style of a signature to claim a person’s identity. Online approaches use data recordings of the writing process to identify user-specific features. The larger number of features compared to offline approaches usually results in higher accuracies [45] and makes it more difficult for forgers to mimic a signature. As digital pens usually implement online-verification [89], we here briefly review online approaches related to digital devices and handwriting as well as digital pens.

Smartphones are plausible candidates for e-signatures. Beside their integrated authentication features, signatures drawn with the finger on the display have been proposed as a feasible alternative [76]. In addition to the visual representation, features such as sound and vibration can also be measured via the embedded motion sensors on the input device [92, 93]. Several classification methods have been applied on pictorial and force-related features including Dynamic Time Warping (DTW) [24, 70, 95], Hidden Markov Models (HMM) [25, 47, 66], Neural Networks [3, 46, 53, 81] as well as Support Vector Machines (SVM) [10, 34, 35] and Random Forests (RF) [64]. More recently, smartwatches gained attention for authentication purposes. For instance, drawing gestures [58] or signatures in the air [5, 14, 36, 56, 61, 65, 68, 86] can offer an unobtrusive authentication method [13]. The motion of a pen can also be tracked by wrist-worn devices for authentication [33, 60] and for recognizing text [75] or gestures [11]. In addition, other smartwatch measurements, such as blood volume changes, can increase the reliability of a signature belonging to a certain person [73]. The bending of fingers contains user-specific features [2] as well as arm movements [74], hand gestures [99] and even simple button presses can differentiate users [72]. In the vicinity of the pen, the scratching sound that the pen tip makes when it moves across the paper can serve as a distinguishing feature [19, 59, 100]. Via inaudible sound signals from a smartphone, vertical movement variations of the pen can be estimated from the reflected sounds [15]. However, it should be mentioned that audio alone is not secured against replay attacks [93] and sensor data alone are not sufficient for a valid signature according to the legal situation [18].

Digital pens emit ink on the paper for legally valid contracts and measure related handwriting features, mostly related to the motion of the pen [32, 42, 69, 90]. Companies like Anoto [1], Wacom [16], and Stabilo [30] offer digital pens for education and industrial applications. In addition, 3D Systems introduced a haptic pen input device that can be utilized to write signatures in the air [17]. Modular concepts for digital pens have been proposed to allow users to create their own applications [88] with different input and output modalities [52, 55, 91, 97]. Integrating fingerprint sensors [85] provides additional security mechanisms, e.g. for biometric authentication in pharmaceutical digital audits [1]. The pressure on the pen grip is another source for individual handwriting features [7, 9, 40, 51] and also pressure applied to the surface is a relevant feature [8, 67]. Pictorial features can not only be obtained from cameras [82] but also via magnet tracking, which has been used for sketching on the back of the hand [77] and signatures [49, 50]. Combinations of audio and motion data have been applied for pen-based handwritten digit recognition [78].

In contrast to related work, we explore the applicability of simple and easy to draw symbols for user authentication. Along with predefined writing trajectories, authentication becomes particularly challenging, because forgers do not have to fully learn another person’s handwriting. Furthermore, as few samples as possible should be used for training in order not to stress users who are only able to produce simple x-mark signatures. As previously applied for handwritten digit recognition [78], we combine audio and motion of a digital pen for our approach.

3 PROTOTYPE

Our prototype is based on Pentelligence by Schrapel et al. [78]. Data is transmitted via USB to the PC for further analysis. The integrated contact microphone samples audio data at a rate of 7.1 kHz and 3D accelerometer as well as 3D gyroscope measurements at a rate of 800 Hz. Ambient noise is attenuated by the housing, thus mainly scratching sounds are measured during writing. Symbols and strokes are separated by detecting when the pen tip touches the writing surface. For this purpose, a pin is soldered on the spring, which is connected to electrical ground on the PCB. When the pen...
We analyzed the collected dataset in order to investigate the individuality of the handwritten symbols. Here, we focus on the use of a small number of training samples with SVM classifiers, as it would be inconvenient, e.g. in the case of signing contracts, to give one’s own (symbol-) signature many times before an identification system can be used. Our total dataset consists of 11,400 samples including 380 samples per user with 20 samples for each of the 19 symbols.

4 USER STUDY

We conducted a study to collect a dataset of various easy to draw symbols. In order to adhere to a skilled forgery setup, we gave instructions regarding the order and direction of the strokes of each gesture. Figure 2 on the right, shows our gesture set with predefined writing instructions. Each of our selected symbols just consists of a few strokes and taps, since we aimed to analyze the individuality of symbols that are comparable in complexity to x-mark signatures. For x-marks we use the most frequently occurring writing trajectory in signing tasks [63]. In addition, common gestures from related work [38, 42, 48, 62] were selected to analyze their applicability for writer identification. In total 19 gestures were selected for our study.

We invited 30 volunteers aged 18 to 56 years (M = 25.6 years, SD = 10.3 years) including 5 female and 25 male individuals. Two of the male participants were left-handed writers. All participants reported feeling well and rested, which is important because fatigue affects handwriting [4]. Our study procedure was adapted from Schrapel et al. [78]. The study was conducted in a quiet office environment to avoid interfering the internal audio measurements and to allow the participants to focus on their tasks. We began by introducing our prototype and the study procedure. We stated that we intended to collect and analyze gestures with the digital pen. The participants were instructed to repeatedly write a symbol on a 21.0 × 29.7 cm white squared paper with a box size of 11 × 14 mm and 15 boxes per row. The symbol and its writing instruction was displayed on a nearby screen. To simplify the selection of individual symbols, the writing of a symbol was confirmed with a short beep and the current count was displayed with a green background. Symbols were cropped in time from the data stream by storing the data from 20 ms before to 20 ms after the pen tip touched the paper. Between two strokes more than 500 ms had to elapse to be recognized as two different symbols. Figure 2, left, shows a participant during the study procedure. All 19 symbols were collected in random order. After writing a symbol 20 times, another symbol was displayed until all symbols were written. When a participant accidentally wrote a symbol before hearing the soft beep sound, used the wrong writing instruction or wrote a wrong symbol, the samples were discarded and repeated. This ensured that a clean and balanced data set was created.

5 RESULTS

We analyzed the collected dataset in order to investigate the individuality of the handwritten symbols. Here, we focus on the use of a small number of training samples with SVM classifiers, as it would be inconvenient, e.g. in the case of signing contracts, to give one’s own (symbol-) signature many times before an identification would be inconvenient, e.g. in the case of signing contracts, to give one’s own (symbol-) signature many times before an identification.

In contrast to the earlier prototype of Schrapel et al. [78] we improved the case by adding a pen grip to constrain the ways the pen can be held. This reduces one source of intra-user variability and helps to capture the way the pen is held in the motion data. Figure 1 shows the prototype printed with a Keyence AGILISTA and the corresponding 3D model with the position of the internal PCB.

Figure 2: Study setup and gestures. The photo shows a participant writing symbols on squared paper. The screen displayed symbols with writing instructions. The set of symbols is visualized on the right. The arrows and numbers represent the writing direction and stroke sequence, respectively.

Figure 3: The block diagram depicts the algorithm for calculating a feature vector. Each sensor data is processed separately and then concatenated to a feature vector.
axes and gyroscope axes separately. For each summed sensor signal FFTs and statistical features are calculated using the same methods as for audios. The most relevant feature components for audio in handwriting recognition are located below 1 kHz [78]. To correctly measure the relevant features, the signal must be sampled with at least twice this cutoff frequency [79, 94]. Thus, we also added the downsampled audio signal to 2 kHz to the analyzed feature set. For motion we downsampled the signal to 300 Hz. Downsampling attenuates high-frequency noise and can may support a classifier to focus on user-specific features. Figure 3 visualizes the process of generating the feature vector sets.

5.2 Feature Analysis

To analyze the symbols on their suitability for writer identification, we first optimize the SVM parameters and then continue by cross validating on each given feature vector set. The SVM hyperparameters kernel (linear or rbf) and cost $C$ (1, 10, 100, 100) are optimized via grid search [71]. The goal of this grid search process is to recognize the 30 writers on all symbols as accurately as possible for each previously calculated feature set using the established SVM parameters on a given preprocessed dataset, we randomly split the data of each writer into 70% for training and 30% for testing. Hence, for each training iteration the training set consists of 266 samples per writer. The SVM is then trained using in total randomly picked 7980 samples to identify the 30 users. We repeat the training and test procedure for 25 times on each corresponding feature set and note down the resulting metrics, i.e. accuracy, precision, recall, and F1-score. As proposed by Levy et al. [57], for each feature set we use the same set of random seeds to obtain comparable results. We performed the procedure on 41 different feature sets resulting in a total of 1,025 tests (41 feature sets $\times$ 25 training and testing).

Figure 4 on the right shows the results of our feature vector sets tested by the maximum average F1-score. We obtained an average (macro) F1-score of $F1 = 0.78$ ($SD = 0.05$) with comparable recall ($R = 0.78; SD = 0.12$) and precision ($P = 0.79; SD = 0.12$) on the best performing feature set. Stating the average AUC score of $AUC = 0.96$ ($SD = 0.03$) and equal error rate $EER = 0.09$ ($SD = 0.04$) is misleadingly high in our case, because for each of the 30 classes (writers) there exist 114 positive and 3,306 negative test samples. Thus, we will continue reporting F1-scores. In general we derive from the tests that the statistical features have a substantial contribution to the distinguishability of users, and by splitting the audio signal into three parts further improvements are achieved. Combining audio and motion measurements from pens for user identification ($\bar{F1} = 0.74; SD = 0.12$) clearly outperforms audio alone ($\bar{F1} = 0.53; SD = 0.12$) and motion alone ($\bar{F1} = 0.46; SD = 0.11$). On all our 41 SVM hyperparameter optimizations, linear kernels showed better results than Gaussian kernels (rbf). In 31 cases a cost parameter of $C = 10$ was selected, followed by five cases using $C = 1000$ (selected by the SVMs trained exclusively on statistical features) and three cases using $C = 100$. A larger C value corresponds to a smaller margin hyperplane and defines how strongly a sample is penalized inside the margin. The best performance is achieved when the raw audio of the samples is splitted into three equal parts where of each statistical features and FFTs are added to the feature vector. Data of accelerometer axes and gyroscope axes are downsampled to 300 Hz and summed up respectively. The FFT and statistical metrics are calculated from each of the two summed signals and added to the feature vector. Splitting of the summed motion signals is not applied. The grid search on this best performing feature set resulted in using a linear kernel with a cost $C = 1$ for the trained SVMs. For the best performing feature set we repeated the grid search test on three, four, and five-degree polynomial kernels and a sigmoid kernel, again obtaining the best results with a linear kernel with a $C = 1$.

5.3 Symbol Analysis

We now analyze symbols with respect to their suitability for writer identification by using the best-performing feature set and corresponding SVM hyperparameters. The cross validation test is repeated by using one, three, five, and ten training samples. For each training size on each symbol we perform 25 tests by using a fixed set of random seeds to pick training and test samples. In total, 1,900 multi-class SVMs were trained and tested (4 training sizes $\times$ 19
which still indicates an acceptable tradeoff between the two measures.

Cross ANOVA revealed a statistically significant difference of the measures. A subsequent Tukey test found no significant difference between recall and pre-determined acceptance of the classifier's performance, as one would expect. In addition, we wanted to find out whether the precision and recall measures have comparable performances. A one-way ANOVA indicated statistically significant differences between both measures (F(8, 475) = 3125.71, p < 0.001). A subsequent post-hoc Tukey test found no significant difference between recall and precision when ten training samples are used (p = 0.1 > 0.05). On average the precision is slightly higher by Δ = 1.2% (SD = 0.38%), which still indicates an acceptable tradeoff between the two measures.

Ideally, the classifier would reject all other writers and only accept true samples of an individual. A higher precision means that the SVM is more accurately rejecting other writers. While a higher recall results in a more accurate identification of the writer. Therefore, precision is more important in authentication tasks. The F1-score is the harmonic mean of both measures and reaches an ideal case a value of 1.

5.4 Majority Voting

As contracts might require a person to sign a document several times, we analyze the impact of averaging classification results. For this purpose, we used the predicted results of the previous tests and average all possible combinations without duplicates by the formula: \( \binom{n}{k} = \frac{n!}{k!(n-k)!} \). Where \( n \) is the number of predictions and \( k \) is the number of voting samples. Using each previous training sample sizes \( t = \{1, 3, 5, 10\} \) and the corresponding test sample sizes \( n = \{19, 17, 15, 10\} \), we analyzed \( k = \{1, 3, 5, 7, 9\} \) voting test samples. The user is then determined in a voting result based on the maximum occurring class.

In Figure 6, left, the impact of majority voting on the F1-score and training sample size is plotted. The lines are the determined mean values of the F1-scores for each training sample size. The colored shades indicate the standard deviation around the mean values. For each training size a one-way ANOVA revealed significant differences between at least two groups of majority voting F1-scores (\( F(5, 475) = [727.5, 1805.9, 2405.0, 2263.5], p < 0.001 \)). A subsequent Tukey test found significant differences for each training size between all corresponding numbers of voting samples. There was no significant difference between 7 and 9 voting samples using 10 training samples (\( p = 0.127 \)). However, this result can be related to the limited amount of 20 samples per symbol and user. To exemplify the impact of majority voting, Figure 6, right, depicts voting results on the symbols Arrow, Cross and Right. On average three voting samples increase the F1-score by \( \Delta = 8.9\% (SD = 3.3\%) \). We derive that majority voting positively effects writer identification and that signing a contract with three symbols allows a more reliable user identification than with just one symbol. Symbols such as Arrow were found to have a significantly higher potential of accurately identifying users than Cross with our approach.

6 DISCUSSION

Identifying writers by handwritten symbols is a major challenge in cases of suspected fraud. So far, the focus has been on ordinary
signatures, but in cases of illiteracy or disabilities, a writer may be limited to signing with x-marks or similar easy to draw symbols. At least three people have to be involved in the process of a valid x-mark signature [18]. This includes the writer, the contractor, and a witness. Cases of suspected fraud are associated with a high financial expense, as the signatures have to be analyzed using forensic methods and the relationships of witnesses have to be investigated. Additional security mechanisms can contribute to identify forgeries. One challenge in collecting samples of handwritten symbols is that writers of x-mark signatures may tire faster. Thus, a classifier must be able to reliably distinguish writers with as few samples as possible.

Our results show that combining audio and motion data of digital pens can more reliably identify writers by handwritten symbols than single sensors. This confirms the finding by Schrapel et al. [78] that both sensors contain relevant and specific features of a symbol signature. In addition, this is in agreement with the results of Muramatsu et al. that combined sensor approaches increase the reliability of online verification approaches [67]. Splitting the incoming data stream is advantageous for scratching sounds in order to integrate a time component into the features. In contrast, higher recognition rates were achieved when motion data of a symbol were not split. In addition, downsampling the measurements to 300 Hz was beneficial on motion data, while using the raw audio signal at 7.1 kHz resulted in higher scores. Compared to previous approaches that used spectral analysis [54, 98], our results show the importance of the time component in signature verification tasks. Additional statistical features of both sensors can further increase the accuracy of the classifier. Hence, the time component of the signals is more important for audio measurements, while the high-frequency components of motion data contain less information. This can be attributed to the writing speed. Vibrations on the paper are more likely to occur at lower frequencies (up to 150 Hz). Produced scratching sounds are modulated by the writing speed [59]. Small irregularities of the paper surface act as a carrier that influence the pitch of the produced scratching sounds. In addition, the volume is influenced by the pressure on the paper [44]. Together with loop-back sounds that propagate via the bones of the hand, resonance properties of the writing surface and the way the pen is held, unique features are produced. This allows relatively accurate identification of a writer with just a few training samples.

Our analysis of the individual symbols shows that certain symbols include a higher diversity of individual characteristics for our participants. This can be attributed to the individual writing style and the way of holding the pen. While the x-marks achieved a uniform performance over the analyzed symbols, the F1-scores of the three best performing symbols Arrow, Circle Left, Question Mark exhibit a significantly higher reliability in identifying writers. However, arrows on contracts are commonly used to indicate signing fields and question marks may also be misinterpreted. In addition, according to the legal situation in Germany, a contract must be signed with three crosses for validity [18]. With three crosses as a training set, the F1-Score already increases to about 75 % in comparison to 55 % with only one training cross. In the case of suspected fraud, it would be required to write at least three crosses again and select a user based on majority voting. With this approach the F1-Score increased to over 80 %. The more training and test samples are given for majority voting, the more reliably the 30 users were identified. With 7 test majority samples, we already reached an F1-score of over 90 % on x-marks by using three training samples. It should be noted that x-mark signatures are particularly challenging due to their structural simplicity. A skilled forger can easily learn the trajectories of a symbol [39, 43, 63]. However, the scratching sounds that propagate via the loop-back channel over the hand bones contain individual features that make forgeries more challenging.

Compared to previous research on handwritten digit recognition with audio and motion data [78], we showed that combining both sensors allows identifying writers. Since these features are writer specific, generalization of handwriting recognition classifiers is challenging. Schrapel et al. circumvent this problem by retraining their neural networks on samples of the writer. This assumes that the current writer is known and individualized classifiers already exist. However, it could not be determined whether the current writer is already known to the system. Our research has closed this gap and shown that the writer can be detected by individual symbols. This detection allows automatic selection of a user-specific classifier and support for multiple users of a pen. In addition, the presented approach opens up new application scenarios.

We will now discuss the components and challenges of such a possible future x-mark signature verification application. In this context, our prototype may serve as an additional security element in cases of suspected fraud. The system components involved are illustrated in Figure 7. A full implementation would have to be integrated into a suitable public-key and document archive infrastructure. When signing a contract the pen would encrypt and cryptographically sign the measured data along with a unique pen hardware identifier and timestamps. The date and time on the physical contract allows assigning a digital timestamp of the pen to a signing event. Furthermore, the metadata of the contract can serve to enable unique identification. To additionally prevent subsequent modification of the stored pen data, each signing event could be irreversibly stored in a blockchain or on a trusted, secured server. Although both a blockchain and a server structure would be possible on the backend [96], there is a major challenge in making the stored data immutable with a central solution. An attacker could try to delete or modify the data to prevent later verification.

When signing a contract, only the person authorized to sign with crosses uses the digital pen. In case of suspected fraud, all persons involved (e.g. contractor, witnesses and writer) would write at least 3 samples of the symbol, e.g. the x-mark, with the digital pen that was used when signing the contract. These samples are used to verify the signature on the contract. As shown in Figure 6, the more samples are written, the more reliably a person can be identified by majority voting. To prevent the writing of invalid samples during this observation, another neutral person (investigator) monitors the process. The neutral person ensures that all writers give samples of the symbol with the same writing instructions as found on the physical contract in the same environment, or at least with comparably low ambient noise level in a similar environment. While the pen is not in use, the ambient noise level may be measured so that comparable conditions can be created later. Likewise, signatures could be automatically rejected in noisy environments. In order to verify a symbol signature, data previously stored in the blockchain
while filling out forms. Since fingerprint sensors in mobile devices easily identify individual features of the writers. We assume that contract partners are known and therefore a classifier can more must be limited to small numbers of writers. In our use case, all should be further investigated. Furthermore, our approach sounds. Therefore, scratching sounds on different surfaces and pa-
ing a contract can have an influence on the measured scratching noises. However, in case of intermediate hand surgeries or diseases those characteristics might change. Thus we do not consider such cases. In addition, it may happen that a writer merely claims not to have signed a document. In these cases, the person would try to fool the system by, for example, holding the pen differently or pressing the pen tip harder on the sheet of paper. It is likely that the resulting x-mark signatures differ from other previous contracts of the writer. Differences could also be determined by forensic analysis. Furthermore, stored signatures may already exist in the backend from other valid contracts that can be used for analysis. Since the handwriting changes in the course of life [26, 83] a time limit of used signatures for training an SVM would be necessary. Besides this application scenario, an additional security mechanism could be implemented, e.g. for biometric authentication in pharmaceutical digital audits [1]. In addition to the fingerprint sensor [85], our approach would recognize x-marks while filling out forms. Since fingerprint sensors in mobile devices can be easily fooled by fake fingers, e.g., made of silicone [31], our mechanism could provide additional confidence that the pen is held by an authorized user.

The results of our study and the example application scenario show limitations of the presented prototype. First, the acoustic signals might also include sounds from the environment. Even though the scratching sounds are considerably louder than environmental noises while writing, our approach is limited to quiet office environments. The influence of environmental noise has to be further investigated in future studies. In addition, the paper used for signing a contract can have an influence on the measured scratching sounds. Therefore, scratching sounds on different surfaces and papers should be further investigated. Furthermore, our approach must be limited to small numbers of writers. In our use case, all contract partners are known and therefore a classifier can more easily identify individual features of the writers. We assume that in a usage scenario with a significantly larger number of writers, there may be more false positive classifications. Hence, we limit our approach to contracts which are likely to involve a limited number of contract partners. To further increase the reliability with a significant higher number of users the pen could be extended to measure the force of holding the pen in the hand [6, 41] and a fingerprint sensor [85].

7 CONCLUSION

This paper presented an approach to identify users via handwritten symbols using a digital pen that measures audio and motion data. A study with 30 participants was conducted to evaluate the approach on 19 candidate symbols. The collected dataset consists of 11,400 samples resulting from 380 samples per user with 20 samples for each of the 19 symbols. The proposed method has three main contributions: (1) a preprocessing algorithm for combining audio and motion data from digital pens in the frequency domain to identify writers via handwritten symbols, (2) an analysis of different feature sets fed into SVMs on their suitability for writer identification, and (3) an examination of various symbols including x-marks on their ability to identify writers based on small training sets in an online-verification task.

We found that combining scratching sounds of a pen tip when moving on a sheet of paper with the motion of a pen outperforms single sensor approaches to identify writers via handwritten symbols. Signal representations in the frequency domain were fed into SVMs to establish our results. To the best of our knowledge, we are the first to unite audio and motion data for writer identification via handwritten symbols. The approach achieves an average F1-Score of 76 % when using only three training samples and 17 test sam-

The proposed system may be used as an additional security mechanism for illiterate persons or persons with motor disabilities to support investigators in cases of suspected fraud. In this application scenario, the challenges of writer identification are particularly high. As few symbol samples as possible must suffice to identify the writer. In the ideal case, the signature on the contract should be sufficient to train a classifier. The presented approach allows this, but is limited to low-noise environments, such as offices. Therefore, further research is required in the future on how external noises can be suppressed. A long-term study is also required to investigate the change of handwritten x-mark signatures of single persons as handwriting changes in the course of life.

REFERENCES

Jihoon Suh. 2016. Veri-Pen: A Pen-Based Identification Through Natural Bio-
metrics Extraction. In Proceedings of the 2016 CHI Conference Extended Ab-
stracts on Human Factors in Computing Systems (San Jose, California, USA) (CHI
EA ’16). Association for Computing Machinery, New York, NY, USA, 134–139.
https://doi.org/10.1145/2851581.2809379

Ziwen Sun, Yao Wang, Gang Qu, and Zhiping Zhou. 2016. A 3-D Hand Gesture
Signature Based Biometric Authentication System for Smartphones. Security and
Communication Networks 9 (07 2016), 157–164. https://doi.org/10.1002/sec.1422

National Telecommunications and United States Department of Commerce
Information Administration (NITA). 2002. Product Recall Exception to the
03.

Multimodal Interactions with a Programmable Modular Pen. In Proceedings of the
2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems
(Denver, Colorado, USA) (CHI EA ’17). Association for Computing Machinery,

Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia.
10.1109/TBIOM.2021.3054533

Jeen-Shing Wang and Fang-Chen Chuang. 2012. An Accelerometer-Based
Digital Pen With a Trajectory Recognition Algorithm for Handwritten Digit and
https://doi.org/10.1109/TIE.2011.2167895

Qinglong Wang, Xiaoshen Ren, Sayan Sarcar, and Xiaoying Sun. 2016. EV-Pen:
Leveraging Electrovibration Haptic Feedback in Pen Interaction. In Proceedings of
the 2016 ACM International Conference on Interactive Surfaces and Spaces
(Niagara Falls, Ontario, Canada) (ISS ’16). Association for Computing Machinery,
New York, NY, USA, 57–66. https://doi.org/10.1145/2992154.2992161

Zhengjie Wang, Nainsheng Zhou, Fang Chen, Xiaoxue Feng, Fei Liu, Yinjing Guo,
and Da Chen. 2021. Smart_Auth: User Identity Authentication Based on Smart-
phone Motion Sensors. In 2021 6th International Conference on Image, Vision and
9526964

Zhixiang Wei, Song Yang, Yadong Xie, Fan Li, and Bo Zhao. 2021. SVSV: Online
handwritten signature verification based on sound and vibration. Information

Edmund Taylor Whittaker. 1915. XVIII.—On the functions which are represented
by the expansions of the interpolation-theory. Proceedings of the Royal Society

Proceedings of 3rd International Conference on Document Analysis and Recognition,
Vol. 1. 179–182 vol.1. https://doi.org/10.1109/ICDAR.1995.598971

Karl Wies and Arthur Gervais. 2018. Do you need a blockchain?. In 2018
Crypto Valley Conference on Blockchain Technology (CVCBT). 45–54. https:
//doi.org/10.1109/CVCBT.2018.00011

Songlin Xu, Zhiyuan Wu, Shunhong Wang, Rui Fan, and Nan Lin. 2020. Hy-
drauo: Extending Interaction Space on the Pen through Hydraulic Sensing
and Haptic Output. In Adjunct Publication of the 33rd Annual ACM Sympo-
sium on User Interface Software and Technology (Virtual Event, USA) (UIST ‘20
Adjunct). Association for Computing Machinery, New York, NY, USA, 43–45.
https://doi.org/10.1145/3379350.3416180

Berrin Yanikoglu and Alisher Kholidatov. 2009. Online Signature Verification

Authentication by Smartwatch using Simple Hand Gestures. In 2020 IEEE In-
ternational Conference on Pervasive Computing and Communications (PerCom).
IEEE Computer Society. Los Alamitos, CA, USA, 1–10. https://doi.org/10.1109/
PerCom5495.2020.9127367

Run Zhao, Dong Wang, Qian Zhang, Xueyi Jin, and Ke Liu. 2021. Smartphone-
Based Handwritten Signature Verification Using Acoustic Signals. Proc. ACM
10.1145/3488544