Online signature verification using signature down-sampling and signer-dependent sampling frequency

Mohammad Saleem¹ · Bence Kovari¹

Received: 18 January 2021 / Accepted: 9 September 2021
© The Author(s) 2021

Abstract

Online signature verification considers signatures as time sequences of different measurements of the signing instrument. These signals are captured on digital devices and therefore consist of a discrete number of samples. To enrich or simplify this information, several verifiers employ resampling and interpolation as a preprocessing step to improve their results; however, their design decisions may be difficult to generalize. This study investigates the direct effect of the sampling rate of the input signals on the accuracy of online signature verification systems without using interpolation techniques and proposes a novel online signature verification system based on a signer-dependent sampling frequency. Twenty verifier configurations were created for five different public signature databases and a variety of popular preprocessing approaches and evaluated for 20–40 different sampling rates. Our results show that there is an optimal range for the sampling frequency and the number of sample points that minimizes the error rate of a verifier. A sampling frequency range of 15–50 Hz and a signature point count of 60–240 provided the best accuracies in our experiments. As expected, lower ranges showed inaccurate results; interestingly, however, higher frequencies often decreased the verification accuracy. The results show that one can achieve better or at least the same verification accuracies faster by down-sampling the online signatures before further processing. The proposed system achieved competitive results to state-of-the-art systems for different databases by using the optimal sampling frequency. We also studied the effect of choosing individual sampling frequencies for each signer and proposed a signature verification system based on signer-dependent sampling frequency. The proposed system was tested using 500 different verification methods and improved the accuracy in 92% of the test cases compared to the usage of the original frequency.

Keywords Online signature verification · Resampling · Dynamic time warping

1 Introduction

Signature verification is one of the oldest biometric verification methods with strong legal support and wide usage including bank checks, writer identification, face recognition, medical detection, and a wide range of other applications [6, 24, 31]. Therefore, automated signature verification has been in the focus of researchers for several decades.

Today, we can create systems that reach impressive error rates on different benchmark databases; however, the

possibilities for further improving these systems are always an open question [4]. In this paper, we discuss a specific aspect of online signatures that may allow researchers to make better choices when designing new verification systems or to improve existing ones.

In online signature verification, a device (like a tablet or a camera) is used to acquire the signature as a function of time for each feature captured. The capturing typically happens at frequencies between 75 and 200 Hz. The aim of this work was to study the direct effect of the sampling frequency and the number of sample points on a simple signature verification system. One of the expectations was that the error rate would decrease when the sampling frequency increases because more points should provide more information about the signatures. It was also expected that

✉ Mohammad Saleem
msaleem@aut.bme.hu

¹ Budapest University of Technology and Economics,
Budapest, Hungary

the error rate might reach a minimum level at a sufficiently high frequency.

As discussed in Sect. 3, while most resampling studies in signature verification are focused on single frequencies or composite approaches, this study provides a detailed and extensive landscape of the isolated effects of signature resampling and shows the importance of choosing the optimal sampling frequency on the accuracy of the signature verification system. Furthermore, this study proposes a novel approach for automatically calibrating a signer-dependent sampling frequency-based online signature verification system that achieved promising results. These unique insights and the proposed system should aid researchers in the field in choosing their resampling approach in future signature verification systems.

We conducted thousands of measurements on five different public signature databases, and the results of our experiments showed a different behavior from the mentioned expectations. The relation between the sampling frequency and the error rate was not monotonous in the majority of the cases; however, the error rate had a local minimum. Moreover, this local minimum was achieved in a similar range for several databases. These results will be explained in detail in the subsequent sections.

Further, the individual sampling frequency or signature point count for each signer is studied and used to build an online signature verification system that relies on the signer-dependent sampling frequency. A total of 500 tests applied in this work using several online signature verification systems assure the quality of our results. The signatures were down-sampled and tested for different sample rates in each verification test. The best sampling frequency for each signer was assigned in a testing verification system using only the reference signatures (provided by the signer) before using them later in the verification process. The results proved the strength of the proposed method, where the majority of the tests provided better results, as will be discussed in the results section.

2 Related work

Sampling is applied when a digital device is used to acquire an analog signal by recording it at a specific frequency. If this sampling frequency is sufficiently large, human perception cannot notice the difference between digitized and analog information.

2.1 Online signature verification

An online signature verification system typically consists of four main steps (Fig. 1). The first step is data

Fig. 1 The main steps of online signature verification systems

acquisition, where the signatures are acquired using dedicated devices (typically digitizing tablets).

The next step is preprocessing. Even when the signer provides the signatures under similar circumstances, there will always be some differences in size or location that may hinder their comparison. Thus, in the preprocessing step, methods such as scaling [14], alignment [1, 45], rotation [23], or z -normalization [3] can be applied to enhance the similarity measurement in the later steps. After that, feature extraction is applied, where several features can be collected, such as the position, speed, pressure, and azimuth.

In the verification step, a verification method is applied to decide whether a signature is genuine or forged. There are several approaches used for this purpose, such as dynamic time warping (DTW), neural networks [16], or hidden Markov models (HMMs) [7]. Among these methods, DTW has shown the most promising results [20].

2.2 Dynamic time warping

DTW is one of the most common algorithms used for signature matching [4], it has shown promising results for signature verification both in scientific experiments [44] and competitions, where the top winning systems used it for verification [19, 21]. Thus, we used DTW in our work to ensure that our results will have the widest applicability. DTW measures the distance between two (possibly multi-dimensional) time series; it calculates the distance between one point from the first series and several points from the other series and stores the minimum distance. In the classical DTW implementation, an $m \times n$ cost matrix is created whose elements (k th, j th) are the calculated Euclidean distances between the elements of the compared feature. In a comparison between two signatures (Q and R) for a certain feature F , the best alignment between the two signatures will be calculated by applying the following equation:

$$\psi(k, j) = \text{dist}(F(r_i^k), F(q_i^k)) + \min \begin{cases} \psi(k, j-1) \\ \psi(k-1, j-1) \\ \psi(k-1, j) \end{cases} \quad (1)$$

Where the first part of the equation represents the Euclidean distance between the points, and $\psi(k, j)$ is the cumulative distance up to the (k th, j th) element [29], r_i and q_i are the sampled points of the current reference (R_i) and questioned (Q_i) signatures respectively. $F(r_i^k)$ and $F(q_i^k)$ represent a feature of the current point of the reference and

questioned signatures, respectively. Then, the minimum warping path between the two matrices is measured, which represents the chosen distance. Figure 2 shows an example of a DTW calculation for two time series [38]. Several modifications to DTW have been introduced in the past, which usually aim to capture the relevant information content of the compared time series. Feng and Wah proposed a new warping technique called extreme points warping, which only warps the extreme points of the signals [10]. Sharma and Sundaram used an enhanced DTW by utilizing the code vectors generated from a vector-quantization (VQ) strategy [35]. Faundez-Zanuy proposed a method using a combination of VQ and DTW [8], whereas Parziale et al. used stability-modulated DTW for their verification system [29].

2.3 Resampling in signature verification

Although the input devices used for signature acquisition have sampling rates as high as 200 Hz, it does not mean that they will provide better verification performance. The first study on signature frequencies [30] stated that signature signals have a maximum frequency (f_s) of 20–30 Hz. Throughout the past three decades, several papers have been published that dealt with the subject. Another study suggested that, as the number of harmonics in handwriting is low, 5 Hz is sufficient to provide the most important frequency components and 10 Hz for all of them, whereas to be able to apply some filters for noisy data, one needs a frequency range of 10–37 Hz [39].

The Nyquist rate or frequency [17] is the minimum rate at which a finite bandwidth B signal needs to be sampled to retain all of the information. The sampling frequency should be at least twice the highest frequency contained in the signal. The Shannon–Nyquist sampling theorem [13] guarantees that any signal whose Fourier transform is supported on this bandwidth limit can be entirely reconstructed from the discrete-time signal as long as the frequency rate is at least twice this bandwidth limit, as illustrated by the following equation [32]:

$$f_s > 2B \quad (2)$$

In our case, a frequency range of 40–60 Hz should be sufficient to contain all the signature information without providing redundant information. Although this is a general

theory for signal sampling, it helps us understand and analyze the results and explain some behaviors.

Martinez-Diaz et al. studied the effects of sampling rate and interpolation in an HMM-based verification system on a single database [22]. The signatures were down-sampled to 25 and 50 Hz and then upsampled to 100 Hz (the original frequency of the input device) using Catmull-Rom [5] and linear interpolation schemes. Their results showed that the accuracy could be improved by using resampling and interpolation together.

The above results show the benefits of using resampling and interpolation but do not show the direct effect of the frequency itself, and they cannot be generalized to other databases and verification systems. That is why, in this work, we used different approaches for sampling and pre-processing and many different sampling rates on five different databases.

Vivaracho-Pascual et al. proposed a low-cost approach to online signature recognition based on length normalization. Although their work was about signature recognition, not verification, it showed that, for the Spanish Ministry of Science and Technology database (MCYT) database, it is possible to reduce the number of signature points without performance loss [42] as we will also show in our results. In offline signatures, image resolution is similar to sampling frequency for online signatures. Vargas et al. studied the effect of image resolution on the verification performance; they used images with a resolution of as high as 600 dpi to as low as 45 dpi. Their results showed that a resolution of 150 dpi offers a good trade-off between performance and image resolution [40].

Bin Li et al. proposed a new signature matching algorithm for online signatures. One of the work steps was to resample the signatures to a fixed number of points using equidistance spacing [18].

Vatavu studied the effect of sampling rate on the performance of template-based gesture recognizers using down-sampling and a DTW approach. His results showed that six sampling points are sufficient for Euclidean and angular recognizers to provide high-performance [41].

We need to note here that part of this work is an extension of our previous work [34] where the signer-dependent sampling frequency approach was introduced.

3 Methodology

The previous results were mostly obtained by using some kind of interpolation approach. Therefore, it cannot be clearly stated whether the changes in the verification or recognition accuracy should be attributed to the resampling or to the interpolation itself. In this work, we did not use interpolation to avoid its effect on the results. Therefore all

Fig. 2 Two time series (left), the DTW cumulative cost matrix (center), and the optimal alignment between the series (right) [38]

results (and improvements) introduced here can be directly attributed to the resampling itself. Interpolation techniques could still improve the resulting accuracies, but this is outside the scope of this paper. In this section, our verification system and experimental protocol will be discussed in detail.

3.1 Proposed verification system

We created a simple signature verification system and evaluated it with different preprocessing approaches on several databases to support our conclusions. This will provide a large variation in the experimental work and eliminate some other factors that may affect the system's accuracy.

3.1.1 Databases

Five different databases were used in this work to avoid any data-dependent results: The Signature Verification Competition 2004 database (SVC2004) [47], the MCYT-100 [27], the Dutch and Chinese subsets of the Signature Verification Competition 2011 database (SigComp'11) [19], and the German database of the Signature Verification Competition 2015 (SigComp'15) [20].

Each database consists of a specific number of signatures from different signers. They are divided into groups of genuine and forged signatures. These databases may differ in the sampling rate of the capturing device, resolution, or features captured. Table 1 summarizes and compares these properties.

3.1.2 Preprocessing

Preprocessing is important to improve the similarity measurement accuracy. We chose scaling, translation, and z -normalization methods for preprocessing purposes. Signature scaling may be used to resize all signatures by multiplying all the points by a certain ratio to keep the signature in a specific range (see Eq. 3).

$$\hat{x}(i) = x_{\text{newMin}} + \frac{x(i) - x_{\text{oldMin}}}{x_{\text{oldMax}} - x_{\text{oldMin}}} * (x_{\text{newMax}} - x_{\text{newMin}}) \quad (3)$$

where $x(i)$ is the old feature value, $\hat{x}(i)$ is the signature feature value after signature preprocessing, x_{newMin} and x_{newMax} are the new minimum and maximum values of the preprocessed feature, x_{oldMin} and x_{oldMax} are the old minimum and maximum values of the feature respectively.

In the case of translation, all signature points were shifted by a given vector. In this work, we used translation to move the center of gravity of the signatures to the origin using the following equation:

$$\hat{x}(i) = x(i) - \mu_x \quad (4)$$

where μ_x is the mean of the old values of the feature.

In our previous works, we studied the effect of scaling and translation preprocessing methods which showed that both are strong methods, especially when applied to both horizontal and vertical axes [33].

Normalization to zero mean and unit of energy (Z -normalization) aims to transform all elements of the input vector into an output vector where its mean (μ) is approximately 0, and the standard deviation (σ) is around 1. Here is the formula used in this work for the z -normalization:

$$\hat{x}(i) = \frac{x(i) - \mu}{\sigma}, \text{ where } i \in N \quad (5)$$

The three preprocessing methods and their combinations were used in our verification systems.

3.1.3 Feature extraction

The optimal selection of a set of features is the key to an effective and more accurate online signature verification system. Among several features available, we chose a compilation of the horizontal (X) and vertical positions (Y) and the pressure (P) as they were available in all the

Table 1 Databases used and their statistics

Database	Language	Signers/total signatures	Initial average points	Device sampling rate (Hz)	Best sampling rates (Hz)	Best average points
SVC2004	English/ Chinese	40/1600	198	100	14–50	30–104
MCYT-100	Spanish	100/5000	440	100	20–50	88–220
SigComp'11	Dutch	64/1905	978	200	16–50	81–244
SigComp'11	Chinese	20/1339	792	200	18–50	72–198
SigComp'15	German	30/750	125	75	37–75	63–125

databases used. Five different combinations of features were tested: X , Y , P , XY , and XYP .

3.1.4 Similarity measurement and verification

After the previous steps, the signatures were ready to be verified. For each signer, 10 genuine signatures were chosen to act as references. These signatures were used to calculate a similarity threshold for the verification using DTW. DTW can be used with different distance measurement algorithms; here, the Euclidean distance was used. According to the Euclidean distance formula, the distance ($dist$) between two points in the plane with the coordinates (a_1, b_1) , and (a_2, b_2) is given by

$$dist((a_1, b_1), (a_2, b_2)) = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2} \quad (6)$$

After doing the preprocessing for all signatures, we calculated a threshold for each signer. When a new signature was being tested, the average distance from the reference signatures was calculated and then compared with the threshold, and if it was equal to or lower than the threshold, the signature was classified as genuine; otherwise, it was classified as forged. Two types of errors may occur during this process. The first one occurs when a genuine signature is classified as forged (false rejection rate (FRR)), and the second one occurs when a forged signature is classified as genuine (false acceptance rate (FAR)). In the testing phase, ten genuine signatures (different from the signatures used as references) are used to calculate the FRR as follows:

$$FRR = \frac{\text{Genuine signatures classified as forged}}{\text{Genuine signatures in the testing set}} \quad (7)$$

while 20 forged signatures used to calculate the FAR as follows:

$$FAR = \frac{\text{Forged signatures classified as genuine}}{\text{Forged signatures in the testing set}} \quad (8)$$

Changing the threshold will have opposite effects on FAR and FRR. FRR will rise, and FAR will drop in case the sensitivity of the system increases. Contrariwise, if the sensitivity is reduced, FRR will drop, and FAR will rise. Therefore, there will be a point where the two values meet, which is called the equal error rate (EER), which is widely used and accepted in the field of signature verification to achieve a trade-off between both error types Impedovo and Pirlo [12]. In our benchmarks, we tuned the initial threshold until we reached this point.

3.2 Effect of sampling frequency

In this section, we focus on evaluating and analyzing the effect of using different sampling frequencies of each database on the accuracy of the online signature verification system. Several tests were applied to assure more accurate comparisons regardless of the effect of any other factor of using a specific verification system.

3.2.1 Signature sampling

To test the effect of the sampling rate and the number of sample points on the verification accuracy, we applied the verification steps using different sampling rates in each test. Thus, in each test, some points were skipped to reduce the sampling rate and signature points. We began the test with the initial sampling rate, and then, in each iteration, we skipped some of the points. The number of iterations depends on the average number of sample points in the database. For SVC2004, MCYT-100, and SigComp'15, 20 iterations were used, whereas for SigComp'11 (Dutch and Chinese), 40 iterations were used. These iterations provided the tests for sampling rates between 5 and 200 Hz. The ranges of the sampling rates and the average number of signature points tested were as follows: SVC2004, 5–100 Hz (10–208 points); MCYT-100, 5–100 Hz (21–440 points); SigComp'15, 3–75 Hz (6–125 points); SigComp'11 (Dutch), 5–200 Hz (24–978 points); and SigComp'11 (Chinese), 5–200 Hz (19–792 points).

3.2.2 Experimental protocol

Our signature verification system uses different databases, features, preprocessing, and verification methods, which provide us with several results that can help generalize the results regardless of the effect of the specific methods chosen. With five databases, four preprocessing methods, and five combinations of features, we were able to test the results using 100 different configurations (see Fig. 3).

Fig. 3 Combinations of the applied verifiers

The 20 different combinations of the verification system (for each database) were then applied using different sampling rates. As discussed earlier, the number of iterations used was different for the databases. Overall, 2800 different tests were applied in these iterations. For each configuration, the results were evaluated and visualized to study the exact effect in all cases. Algorithm 1 pseudo-code describes the experimental protocol of the work.

Algorithm 1 Pseudo-code of the experimental protocol

```

for database in Databases do
  while Combinations != NULL do
    x=1
    i = iteration number           ▷ 20 or 40
    s = d.freq                     ▷ device sample frequency
    for x<=i do
      s=d.freq/x
      resample the signatures using sampling rate s
      calculate EER
      x=x+1
    return EER
  end for
end while
end for

```

3.3 Signer dependent sampling frequency system

In this section, we focus on the results of using a signer-dependent sampling frequency rather than using the same sampling frequency for the entire database.

The proposed technique is based on choosing the best sampling frequency for each signer before starting the verification process. In real-life situations, the signer provides some signatures as references. These signatures are used to compare the tested signature and check if it is forged or genuine. Therefore, the proposed technique will use only these references for choosing the signer-dependent sampling frequency to evaluate the system in a real-life situation where no more genuine or forged signatures are available. The references were divided into training and testing sets, and only the false rejection rate (FRR) was calculated and used to choose the best sampling frequency. To evaluate the efficiency of the proposed technique, more genuine (not used in the previous step) and forged signatures are used later to calculate the error rate in the evaluation step.

Although the first step contains all the main stages of a signature verification process, it is used as a preprocessing step to improve the full system's quality later. The same verifier used in the first step, where we choose the best sampling frequency for each signer, is used in the system

evaluation step to obtain the actual improvement under the same circumstances.

Several inputs for each stage of the online signature verification process were used to obtain more results that help better understanding the effect of using a signer-dependent sampling frequency.

The variant verification steps were combined to provide many tests. These combinations were tested using 3–7 samples (n) as references. The total number of tests was 500. After choosing the current verifier, it is first applied using different sampling rates (f) from a group of sampling frequencies (F). In each iteration, the sampling rate is changed, and the average error rate (AER) is calculated for each signer, where AER is calculated as shown in Eq. (9). The number of iterations depends on the initial sampling frequency, 20–40 iterations for 100–200 Hz databases.

$$\text{AER} = \frac{\text{FRR} + \text{FAR}}{2} \quad (9)$$

In each iteration, the sampling frequency is changed, and the system is applied using it. Since n samples are used as references to calculate the threshold in the verification process, the rest of the signatures S_i ($10-n$) are used to test the accuracy when using the current sampling frequency. Only references are used in this phase, so only FRR of the ($10-n$) signatures are used to compare the results and assign the best sampling frequency to each signer.

After assigning each signer with his/her best sampling frequency (F_{optimal}), the same verification system with the same preprocessing methods, features, and classification algorithm is applied to each signer individually, and all error rates are calculated for them. The optimal sampling frequency for each signer s is calculated as the following:

$$F_{\text{optimal}} = \min_{\forall f \in F} \text{Frr}(S_i) \quad (10)$$

Later the average error rate for the signers is calculated. In order to evaluate the current verification system and its efficiency when using signer-dependent frequencies, we need to apply the same verification system for the same

Fig. 4 Proposed method

database but using the initial sampling frequency and then compare the results of both methods and calculate the accuracy improvement. Figure 4 provides a brief description of the procedure.

4 Experimental results

In this section, the results of the experiments are discussed and evaluated. Section 4.1 discusses the effect of using different sampling frequencies and Sect. 4.2 discusses the results of the proposed signer-dependent online verification system.

4.1 Sampling rate impact on EER

After applying all the previous combinations at different sampling rates and numbers of signature points, we analyzed and studied the results to determine the effect of using different sampling rates on the accuracy of a verification system. The best results of the experiments are shown in Table 2. We selected the sampling rate for each configuration where the lowest EER was achieved.

The results showed that we could obtain better results by decreasing the sampling rate and the average number of sample points of the databases in most of the combinations. The expected behavior was that the accuracy would increase when the sampling rate (and, thus, the sample count) increases because providing more information about the signatures will help differentiate them.

Our experiments showed that, in most cases, the accuracy started to increase until some point, and then it started to decrease. Figure 5 shows two examples for both cases of the effect of the average number of sample points on the EER from SVC2004. In fact, it follows the second behavior in 92.5% of the configurations of all databases acquired between 100 and 200 Hz, 100% for SigComp'11 (Dutch), in 95% for SigComp'11 (Chinese), 100% for SVC-2004, and 75% for MCYT-100; only a few combinations provided better results when using the initial sampling rate. Moreover, 91.25% of the best results were obtained in these databases using a sampling frequency of less than or equal to 50 Hz.

In the case of SigComp'15, the database was acquired using a relatively low frequency of 75 Hz; thus, even the initial sampling frequency provided good results without down-sampling, which also means that, for all the databases, 93% of the best results of the configuration were obtained using a sampling frequency of less than or equal to 75 Hz. Figure 6 shows some examples of the effect of sampling frequency on the EER.

4.1.1 The best frequency ranges

Our observations showed that, in most cases, the accuracy increased until some point and then started decreasing or stagnating. However, we also observed that there was a specific range where we could obtain the best results. As an example, Table 2 shows that, for the SigComp'11 (Dutch) database, 95% of the best results were obtained when using an average sampling frequency of less than or equal to 50 Hz. Similar observations were made with the other databases. In general, we can say that the majority of the best results were obtained using a range of around 15–50 Hz. These ranges are shown in Fig. 7.

4.1.2 The best sample count ranges

The effect of sample count followed the same behavior as that of the frequency rate as they are related. Table 1 shows the best results of the sample counts. We can see that the sample count ranges between 30 and 104 for SVC2004 and between around 60 and 240 for the other databases that provided the lowest error rate. Figure 8 shows the best sample counts for all the tests.

4.1.3 Sampling restrictions

Our results showed that there are three cases for the frequency range. The first one is where the frequency rate is low or under a certain range (undersampling). In this case, the error rate is high because it does not provide sufficient information about the signature that makes it unique and allows it to be distinguished from other signatures. In the second case, the best results can be obtained when the sampling rate and the signature points are in a specific range where they are neither low nor high. This makes sense because it was shown that the maximum frequency band limit for online signatures is 20–30 Hz and that a Nyquist rate of 40–60 Hz will be sufficient to provide adequate information about the signal without any redundant information. The third case is where the frequency is above a certain range where the accuracy of the results decreases again. We believe that the redundant data not only may provide redundant information that will not help in obtaining better results but may also worsen it (oversampling). These cases are related to the situation with analog signal sampling, where choosing the wrong sampling frequency may produce undersampling or oversampling issues.

Table 2 Best sampling frequencies, sample counts, and EER for each combination for each database

Combination		MCYT			SigComp'11 (Dutch)			SigComp'11 (Chinese)		
Scaling	Translation	Z-normalization	Features	Frequency	Average points	EER	Frequency	Average points	EER	Frequency
•	•		X	50	220	5.1%	16	81	3.0%	20
•	•		Y	33	147	2.7%	22	109	2.3%	50
•	•		P	20	88	5.9%	20	98	2.5%	66
•	•		XY	50	220	2.1%	28	140	1.4%	18
•	•		XYP	100	440	1.9%	22	109	0.6%	16
•			X	25	110	7.8%	22	109	3.8%	40
•			Y	20	88	5.9%	20	98	3.4%	33
•			P	25	110	7.6%	33	163	2.9%	50
•			XY	33	147	4.5%	10	51	1.9%	28
•			XYP	33	147	2.1%	50	244	1.0%	200
	•		X	50	220	4.2%	50	244	1.7%	20
	•		Y	50	220	3.8%	15	75	2.4%	22
	•		P	20	88	4.9%	100	489	2.0%	40
	•		XY	100	440	2.3%	9	44	1.4%	66
	•		XYP	100	440	1.8%	33	163	0.8%	25
		•	X	50	220	4.1%	25	122	2.5%	40
		•	Y	25	110	3.1%	22	109	1.8%	50
		•	P	20	88	5.9%	40	196	2.3%	66
		•	XY	100	440	2.1%	25	122	0.9%	22
		•	XYP	100	440	1.8%	22	109	0.8%	18

Combination		SVC2004			SigComp'15 (German)		
Scaling	Translation	Z-normalization	Features	Frequency	Average points	EER	Frequency
•	•		X	50	104	13.4%	75
•	•		Y	50	104	14.3%	75
•	•		P	33	69	10.6%	75
•	•		XY	33	69	9.3%	75
•	•		XYP	50	104	4.4%	75
•			X	33	69	17.8%	75
•			Y	20	42	16.3%	75
•			P	33	69	13.8%	75
•			XY	14	30	13.9%	75
•			XYP	16	35	8.0%	75
	•		X	14	30	19.0%	75

Table 2 (continued)

Combination	SVC2004				SigComp'15 (German)			
	Scaling	Translation	Z-normalization	Features	Frequency	Average points	EER	EER
	•	•	•	Y	25	52	16.2%	4.5%
				P	25	52	9.0%	17.7%
				XY	11	23	14.7%	3.3%
				XYP	11	23	14.7%	4.2%
				X	33	69	11.9%	7.3%
			•	Y	50	104	13.8%	7.2%
			•	P	25	52	10.3%	14.9%
			•	XY	33	69	3.5%	3.7%
			•	XYP	33	69	1.6%	5.0%
			•

Fig. 5 Expected (left) versus typical (right) behavior of the EER as a function of sample points on SVC2004

4.2 Signer dependent sampling frequency system results

A comparison study has been conducted to check whether the proposed method has improved the existing online signature verification system. Since sometimes other factors may affect the accuracy of the results, several tests will provide a more accurate evaluation by investigating the behavior of the majority of these systems to avoid any other factors that may impact the results. It also helps in choosing the best or optimal methods that can be used to provide the most accurate verification systems. In this section, all cases are discussed and evaluated to show the improvement in both randomly chosen systems and optimal systems.

Testing all scenarios using scaling, translation, and z-normalization for preprocessing and using n : 3–7 (500 tests) showed that the accuracy improved in 72% of the tests. The improvement in the accuracy reached up to 8.38%. However, 20 tests show a decrease in accuracy, and eight tests resulted in the same accuracy. Overall, 80% of the tests provided better or at least equal verification accuracy regardless of the preprocessing techniques

Fig. 6 Examples from other databases of the effect of sampling frequency on the EER

Fig. 7 Sampling frequencies of the best results for all the tests

Fig. 8 Sample counts of the best results for all the tests

applied or features selected or any number of samples used. Figure 9 shows the accuracy improvement for all cases.

Preprocessing methods have a significant impact on verification accuracy. The experiment showed that z-normalization provides the most accurate results. Therefore, the results of the test where z-normalization was used showed accuracy improvement in 88% of the cases, there was no change in 1% of the cases, and the accuracy decreased in 11% of the cases (see Fig. 10).

We have shown that choosing the preprocessing method will affect the accuracy. Also, the number of samples can significantly influence the results. The best results were achieved when using 3 or 6 samples. Combining these facts into one verification system will provide the most accurate system. Choosing six samples as references and z-normalization for preprocessing while using five different databases and five different features (25 tests) has led to only one negative result, one result with no change and 23 results (92%) with accuracy improvement up to around 8.4% (see Fig. 11).

Fig. 9 Percentage of experiments (vertical axis) where the accuracy improved when using the signer-dependent sampling frequency approach (all the tests)

Fig. 11 Percentage of experiments (vertical axis) where the accuracy improved when using the signer-dependent sampling frequency approach (using z-normalization and 6 samples)

Fig. 10 Percentage of experiments (vertical axis) where the accuracy improved when using the signer-dependent sampling frequency approach (using z-normalization)

4.3 Comparison

Although this study aimed to measure the effect of the sampling frequency on the accuracy, it is also worth mentioning that some of the verification systems applied here achieved competitive results to state-of-the-art systems. In Table 3, we show some of the best results achieved with down-sampling compared to other results for different databases.

5 Conclusion

In this work, we studied the effect of the sampling rate of the input devices used for signature acquisition and the number of sample points on the accuracy of online signature verification systems. We proposed an online signature verification based on signer-dependent sampling frequency and DTW. Several configurations of a DTW-based verification system were used to assess the achievable EER at different sampling rates. Altogether, we conducted 2800

Table 3 A comparison with recent results in the field

Database	EER%	References
MCYT-100	3.69	[46]
MCYT-100	3.05	[36]
MCYT-100	2.76	[37]
MCYT-100	1.81	[15]
MCYT-100	1.8	Proposed
MCYT-100	1.34	[25]
MCYT-100	1.28	[26]
SVC2004	6.65	[43]
SVC2004	3.41	[2]
SVC2004	2.73	[35]
SVC2004	1.8	[11]
SVC2004	1.6	Proposed
SVC2004	0.83	[9]
SigComp'11 (Dutch)	3.73	[19]
SigComp'11 (Dutch)	5.5	[28]
SigComp'11 (Dutch)	3.65	[19]
SigComp'11 (Dutch)	0.6	Proposed
SigComp'11 (Chinese)	14.68	[19]
SigComp'11 (Chinese)	8.93	[28]
SigComp'11 (Chinese)	6.83	[19]
SigComp'11 (Chinese)	0.1	Proposed

different experiments, which helped generalize the results regardless of the effect of other factors that may affect the system's accuracy. To our knowledge, these properties have never been studied within the scope of online signature verification.

The results showed that the majority of the best results could be obtained using a sampling frequency between 15 and 50 Hz and a sample count between 60 and 240 points. Using frequencies lower than these ranges greatly decreased the accuracy, whereas using higher frequencies

decreased or did not affect the accuracy in 92.5% of the configurations of all databases acquired between 100 and 200 Hz. For these databases, 91.25% of the best results were obtained using a sampling frequency of less than or equal to 50 Hz and 93% of less than or equal to 75 Hz.

As the sampling rate and the sample count are strongly correlated, it is too early to conclude which one of the two plays a more significant role in the observed relation; therefore, we presented our results by including both of them. Regardless of this fact, we can state that, in classic DTW-based signature verification, using sampling frequencies higher than 100 Hz will not improve the accuracy of the systems but will instead increase the computational cost of the verification. The results of the proposed system of using signer-dependent sampling frequencies also showed that in 80% of the 500 tests, the accuracy improved or at least did not change. Moreover, the ratio of improved results reached 92% when chosen the optimal preprocessing methods and number of samples. The results showed that using the optimal frequency provides competitive systems for online signature verification. These results are auspicious and suggest that DTW-based online signature verifiers can be improved in the future by using different criteria for choosing the best sampling frequency for each signer.

Acknowledgements The work presented in this paper has been carried out in the frame of Project No. 2019-1.1.1-PIACI-KFI-2019-00263, which has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the 2019-1.1. funding scheme.

Funding Open access funding provided by Budapest University of Technology and Economics. The work presented in this paper has been carried out in the frame of project no. 2019-1.1.1-PIACI-KFI-2019-00263, which has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the 2019-1.1. funding scheme.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahrabian K, BabaAli B (2019) Usage of autoencoders and siamese networks for online handwritten signature verification. *Neural Comput Appl* 31(12):9321–9334
- Alpar O (2018) Online signature verification by continuous wavelet transformation of speed signals. *Expert Syst Appl* 104:33–42
- Auckenthaler R, Carey M, Lloyd-Thomas H (2000) Score normalization for text-independent speaker verification systems. *Digit Signal Proc* 10(1–3):42–54
- Bibi K, Naz S, Rehman A (2020) Biometric signature authentication using machine learning techniques: current trends, challenges and opportunities. *Multimed Tools Appl* 79(1):289–340
- Dodgson NA (1997) Quadratic interpolation for image resampling. *IEEE Trans Image Process* 6(9):1322–1326
- Fairhurst M, Kaplani E (2003) Perceptual analysis of handwritten signatures for biometric authentication. *IEE Proc Vis Image Signal Process* 150(6):389–394
- Farimani SA, Jahan MV (2018) An hmm for online signature verification based on velocity and hand movement directions. In: 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS). IEEE, pp 205–209
- Faundez-Zanuy M (2007) On-line signature recognition based on VQ-DTW. *Pattern Recogn* 40(3):981–992
- Fayyaz M, Hajizadeh Saffar M, Sabokrou M, Fathy M (2015) Feature representation for online signature verification. *arXiv preprint arXiv:150508153*
- Feng H, Wah CC (2003) Online signature verification using a new extreme points warping technique. *Pattern Recogn Lett* 24(16):2943–2951
- Foroozandeh A, Hemmat AA, Rabbani H (2020) Online handwritten signature verification and recognition based on dual-tree complex wavelet packet transform. *J Med Signals Sensors* 10(3):145–157
- Impedovo D, Pirlo G (2008) Automatic signature verification: The state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38(5):609–635
- Jerri AJ (1977) The shannon sampling theorem-its various extensions and applications: A tutorial review. *Proc IEEE* 65(11):1565–1596
- Jindal U, Dalal S, Dahiya N (2018) A combine approach of preprocessing in integrated signature verification (iSV). *Int J Eng Technol* 7(1–2):155–159
- Lai S, Jin L (2018) Recurrent adaptation networks for online signature verification. *IEEE Trans Inf Forensics Secur* 14(6):1624–1637
- Lai S, Jin L, Yang W (2017) Online signature verification using recurrent neural network and length-normalized path signature descriptor. In: 2017 14th IAPR international conference on document analysis and recognition (ICDAR). IEEE, vol 1, pp 400–405
- Landau H (1967) Sampling, data transmission, and the nyquist rate. *Proc IEEE* 55(10):1701–1706
- Li B, Wang K, Zhang D (2007) Minimizing spatial deformation method for online signature matching. In: International conference on biometrics. Springer, pp 646–652
- Liwicki M, Malik MI, Van Den Heuvel CE, Chen X, Berger C, Stoel R, Blumenstein M, Found B (2011) Signature verification competition for online and offline skilled forgeries (sigcomp2011). In: 2011 international conference on document analysis and recognition. IEEE, pp 1480–1484
- Malik MI, Ahmed S, Marcelli A, Pal U, Blumenstein M, Alewijn L, Liwicki M (2015) Icdar2015 competition on signature verification and writer identification for on-and off-line

- skilled forgeries (sigwicom2015). In: 2015 13th International Conference on Document Analysis and Recognition (ICDAR), IEEE, pp 1186–1190
21. Malik MI, Liwicki M, Alewijnse L, Ohyama W, Blumenstein M, Found B (2013) Icdar 2013 competitions on signature verification and writer identification for on-and offline skilled forgeries (sigwicom 2013). In: 2013 12th international conference on document analysis and recognition. IEEE, pp 1477–1483
22. Martinez-Diaz M, Fierrez J, Freire M, Ortega-Garcia J (2007) On the effects of sampling rate and interpolation in hmm-based dynamic signature verification. In: Ninth international conference on document analysis and recognition (ICDAR 2007), vol 2. IEEE, pp 1113–1117
23. Mohammadi MH, Faez K (2012) Matching between important points using dynamic time warping for online signature verification. *J Sel Areas Bioinf (JBIO)*
24. Moos S, Marcolin F, Tornincasa S, Vezzetti E, Violante MG, Fracastoro G, Speranza D, Padula F (2017) Cleft lip pathology diagnosis and foetal landmark extraction via 3d geometrical analysis. *Int J Interact Des Manuf (IJIDeM)* 11(1):1–18
25. Okawa M (2019) Template matching using time-series averaging and DTW with dependent warping for online signature verification. *IEEE Access* 7:81010–81019
26. Okawa M (2020) Online signature verification using single-template matching with time-series averaging and gradient boosting. *Pattern Recogn* 102:107227
27. Ortega-Garcia J, Fierrez-Aguilar J, Simon D, Gonzalez J, Faundez-Zanuy M, Espinosa V, Satue A, Hernaez I, Igarza JJ, Vivaracho C et al (2003) Mcyt baseline corpus: a bimodal biometric database. *IEE Proc Vis Image Signal Process* 150(6):395–401
28. Parodi M, Gómez JC (2014) Legendre polynomials based feature extraction for online signature verification. consistency analysis of feature combinations. *Pattern Recogn* 47(1):128–140
29. Parziale A, Diaz M, Ferrer MA, Marcelli A (2019) Sm-dtw: stability modulated dynamic time warping for signature verification. *Pattern Recogn Lett* 121:113–122
30. Plamondon R, Lorette G (1989) Automatic signature verification and writer identification-the state of the art. *Pattern Recogn* 22(2):107–131
31. Rehman A, Naz S, Razzak MI, Hameed IA (2019) Automatic visual features for writer identification: a deep learning approach. *IEEE Access* 7:17149–17157
32. Romanov E, Ordentlich O (2019) Above the nyquist rate, modulo folding does not hurt. *IEEE Signal Process Lett* 26(8):1167–1171
33. Saleem M, Kovari B (2020) Preprocessing approaches in DTW based online signature verification. *Pollack Period* 15(1):148–157
34. Saleem M, Kovari B (2020) Online signature verification based on signer dependent sampling frequency and dynamic time warping. In: 2020 7th international conference on soft computing and machine intelligence (ISCMi). IEEE, pp 182–186
35. Sharma A, Sundaram S (2016) An enhanced contextual STW based system for online signature verification using vector quantization. *Pattern Recogn Lett* 84:22–28
36. Sharma A, Sundaram S (2016) A novel online signature verification system based on GMM features in a DTW framework. *IEEE Trans Inf Forensics Secur* 12(3):705–718
37. Sharma A, Sundaram S (2018) On the exploration of information from the DTW cost matrix for online signature verification. *IEEE Trans Cybern* 48(2):611–624
38. Silva DF, Giusti R, Keogh E, Batista GE (2018) Speeding up similarity search under dynamic time warping by pruning unpromising alignments. *Data Min Knowl Discov* 32(4):988–1016
39. Teulings HL, Maarse FJ (1984) Digital recording and processing of handwriting movements. *Hum Mov Sci* 3(1–2):193–217
40. Vargas J, Ferre M, Travieso C, Alonso J (2007) Off-line signature verification system performance against image acquisition resolution. In: Ninth international conference on document analysis and recognition (ICDAR 2007). IEEE, vol 2, pp 834–838
41. Vatavu RD (2011) The effect of sampling rate on the performance of template-based gesture recognizers. In: Proceedings of the 13th international conference on multimodal interfaces. ACM, pp 271–278
42. Vivaracho-Pascual C, Faundez-Zanuy M, Pascual JM (2009) An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances. *Pattern Recogn* 42(1):183–193
43. Wang K, Wang Y, Zhang Z (2011) On-line signature verification using wavelet packet. In: 2011 international joint conference on biometrics (IJCB). IEEE, pp 1–6
44. Wu X, Kimura A, Uchida S, Kashino K (2019) Prewarping siamese network: Learning local representations for online signature verification. In: ICASSP 2019–2019 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2467–2471
45. Xia X, Chen Z, Luan F, Song X (2017) Signature alignment based on GMM for on-line signature verification. *Pattern Recogn* 65:188–196
46. Xia Z, Shi T, Xiong NN, Sun X, Jeon B (2018) A privacy-preserving handwritten signature verification method using combinational features and secure KNN. *IEEE Access* 6:46695–46705
47. Yeung DY, Chang H, Xiong Y, George S, Kashi R, Matsumoto T, Rigoll G (2004) Svc2004: First international signature verification competition. In: International conference on biometric authentication. Springer, pp 16–22

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.