DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016 *Int. J. Exp. Res. Rev.*, Vol. 44: 185-195 (2024)

Original Article | Peer Reviewed | Open Access

International Journal of Experimental Research and Review (IJERR)

ISSN: 2455-4855 (Online)
www.iaph.in

Feature Ranking Using Novel Consistency Measure by Normalized Standard Deviation and Proposal of Three Novel Global Features for Online Signature Verification

Biswajit Kar*, Bhimraj Prasai Chetry and Sayan Das

Department of Instrumentation Engineering, Central Institute of Technology Kokrajhar, Assam, India

E-mail/Orcid Id:

BK, b.kar@cit.ac.in, https://orcid.org/0000-0003-2686-5814;
BPC, bp.chetry@cit.ac.in, https://orcid.org/0009-0002-5072-4739; *SD*, sayan.eie@gmail.com, https://orcid.org/0009-0006-4644-9438

Article History:

Received: 26th April, 2024
Accepted: 21st Oct., 2024
Published: 30th Oct., 2024

Keywords:

Consistency estimation, Feature extraction, Feature normalization, Feature ranking, Feature selection, mRMR

How to cite this Article:
Biswajit Kar, Bhimraj Prasai Chetry and Sayan Das (2024). Feature Ranking Using Novel Consistency Measure by Normalized Standard Deviation and Proposal of Three Novel Global Features for Online Signature Verification. *International Journal of Experimental Research and Review*, *44*, 185-195.
DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016

Abstract: Signature is a behavioral biometric that evolves throughout a person's life. Feature extraction and ranking are very important steps towards online signature verification in order to achieve high efficiency. In our case, we have extracted 48 global features. A novel feature ranking technique based on consistency measure using normalized standard deviation is proposed here and is compared with well-established mRMR based ranking. The use of normalized standard deviation formula is the novel approach for feature ranking. Moreover, we have proposed three novel global features for consistency measure and mRMR based ranking. These three features has shown its importance in consistency based ranking as well as in mRMR based ranking. Consistency estimation of global features is important in one-class classification framework, where only the genuine signatures are available. All the features that are ranked by consistency measures and its weighting factors of feature vector are computed for every signer. More consistent features has given more weight for verification. Feature ranking shows the importance of each features for a particular system. Also it helps in feature selection to select the more discriminating features for a particular system and removing all irrelevant features. It saves the computational time and size of the model. Different global features has different scaling factor. Generally, normalization is done before ranking but in our technique we have ranked our features without normalization. Because normalization disrupt the statistical consistency. Therefore, we have used min-max normalization where all the features are converted to (0-1) range after ranking. In our proposed system, the proposed phase related global features shows more consistency in our ranking process. Similarly well-established mRMR feature ranking has also ranked our proposed novel features number 48 (Standard deviation of the phase), 43 (Entropy of shape signature function) and 47 (Mean of phase) within top 13 features. It is seen that both our proposed novel feature ranking technique based on consistency measure and well-established mRMR based ranking has shown almost similar performance. The proposed algorithms are verified with SVC 2004 database.

Introduction

In online signature verification systems, the data is captured while the signature is being written. Acquisition in this form requires a special pen or digitizing pressure-sensitive tablet. These devices can capture both the static and dynamic features of the signature. Static features (attributes) are the visible properties of the signature (e.g., shape, size, position, etc.), while dynamic features are the invisible properties (e.g., timing, pressure, velocity, acceleration and other derivatives). From these signals, functional features and global features are calculated.

***Corresponding Author:** b.kar@cit.ac.in

Functional features are instantaneous measurement values, and these are vectors. Global features are average values, and these are scalar quantities. Functional features are pressure variation throughout the signature, speed of writing throughout the signature, speed variation towards positive x direction, etc. Global features such as average signing speed, average acceleration, duration of the signing process, the ratio of pen-down and pen-up duration, mean pressure, the standard deviation of the pressure, etc., are scalar quantities (Kar and Dutta, 2012).

Global features are divided into two categories: static features and dynamic features. Static features include the number of strokes, aspect ratio of signature boundary, number of loops, number of turning points, etc., whereas dynamic features include the position and sequence of the strokes, number of samples, total duration of signing, pen up time, etc. In addition to signing altitude, azimuth can also be recorded with the help of appropriate equipment. The invisible signature functions gathered by online signatures make them more reliable because timing and pressure attributes are much harder to imitate than the static information of a signature since a more significant number of features can be extracted in addition to the signature shape.

Approaches to online signature verification are generally classified into two groups: one based on a global features-based approach and the other often referred to as function-based approaches (Plamondon and Srihari, 2000; Leclerc and Plamondon, 1994). In the parametric algorithms, a set of parameters is selected to describe a signature pattern and the parameters of the reference and test signatures are compared to decide whether the signature is genuine or not. In this approach, signatures can be described compactly, so the enrollment data size is typically tiny. More importantly, this approach is expected to be more stable against the variations in local regions, which are common in signatures. For users sensitive to privacy problems, enrolling only parameters may also be considered an advantage because the original pattern cannot be reconstructed. The major limitation of this approach lies in its discriminative ability (Zhang et al., 2000; Kashi et al., 1996). An averaging effect arises when calculating the parameters over the whole pattern. Although this effect is the reason for the stability mentioned above, at the same time, the exact impact inevitably blurs the distinction between genuine and forge patterns. Furthermore, (Zhang et al., 2000) indicated difficulty in characterizing unpredictable forgeries in advance. The parameters selected with a small set of signers may need to work better on a more extensive set of signers (Guru and Prakash, 2009).

Proposed System

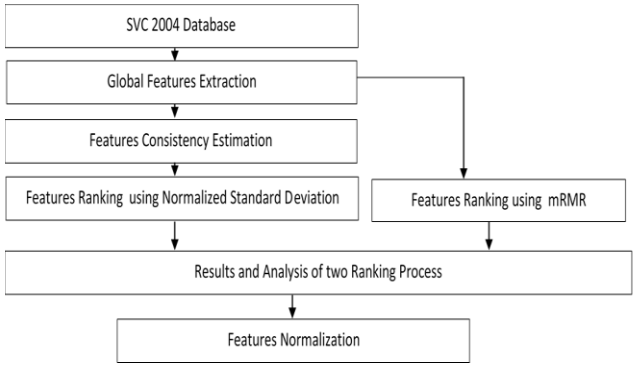

Figure 1. Block diagram of the proposed system.

Here in the proposed system as shown in Figure 1 above 48 global features are extracted from SVC 2004 Database. Out of 48 global features extracted three are novel global features proposed here namely, Entropy of shape signature function, Mean of phase and Standard deviation of the phase. After feature extraction our proposed consistency based features ranking is done using normalized standard deviation. And for comparison and its effectiveness, analysis our proposed feature ranking is compared with the same extracted features ranked using well-established mRMR ranking algorithm. Normalization disrupts the statistical consistency. Therefore, we have used min-max normalization after feature ranking where all the features are converted to (0-1) range.

Description of Database

The First International Signature Verification Competition (SVC 2004) (Yeung et al., 2004) was held as a step towards establishing common benchmark databases and benchmarking rules. Sample signatures from the database SVC 2004 is shown below in Figure 2. For each of the two tasks of the competition, a signature database involving 40 sets of signature data was created, with 20 genuine signatures and 20 skilled forgeries for each set (Najda and Saeed, 2022).

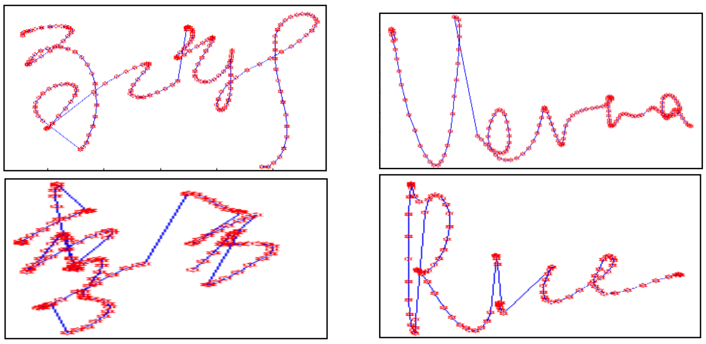

Figure 2. Sample signatures from the SVC 2004 database.

DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016

Global Features Extraction

Verification with global features of a signature has several advantages. It is simple to compute. Once the features are extracted, the original signature does not need to be retained, eliminating the privacy problem (Xia et al., 2018; Marana et al., 2010). This makes it ideal as an inexpensive technique that can be used to catch a majority of forgeries without hampering privacy. It can be seen that, with a small number of global features, this technique can classify signatures with approximately moderate accuracy.

Table 1. Extracted forty-eight global features of online signature.

Global features	Symbol	Features name and mathematical definition
g_1	N_s	Number of continuous pen-down sequences (strokes)
g_2	T_t	Duration of the complete writing process in ms
g_3	N_{sp}	Total number of samples
g_4	N_{ex}	Total number of local minima and maxima (extreme points) in $x\ (t)$
g_5	N_{ey}	Total number of local minima and maxima in $y\ (t)$
g_6	w	Width of the signature $= absolute\ deviation\ of\ x = mean\{\lvert(x-\mu(x))\rvert\}$
g_7	h	Height of the signature $= absolute\ deviation\ of\ y = mean\{\lvert(y-\mu(y))\rvert\}$
g_8	r_a	$Aspect\ Ratio = \frac{Width}{Height} = \frac{mean\{(x-\mu(x))\}}{mean\{(y-\mu(y))\}}$
g_9	T_{pu}	Pen up duration
g_{10}	T_{pd}	Pen down duration
g_{11}	r_{dt}	$\frac{\text{Pen down duration}}{\text{Total time duration}}$
g_{12}	r_{ut}	$\frac{\text{Pen up duration}}{\text{Pen down duration}}$
g_{13}	d_x	Total distance traveled in x direction (normalized) $\frac{\sum\lvert\Delta x\rvert}{w} = \frac{\sum\lvert x(n)-x(n-1)\rvert}{w}$
g_{14}	d_y	Total distance travelled in y direction (normalized) $\frac{\sum\lvert\Delta y\rvert}{h} = \frac{\sum\lvert y(n)-y(n-1)\rvert}{h}$
g_{15}	d_{xy}	Total distance travelled (normalized) $\frac{\sum\left\lvert\sqrt{\Delta x^2+\Delta y^2}\right\rvert}{(w+h)}$
g_{16}	μ_{vx}	x-directional average speed $\frac{d_x}{T_{pd}}$
g_{17}	μ_{vy}	y-directional average speed $\frac{d_y}{T_{pd}}$
g_{18}	μ_{vxy}	Average speed $\frac{d_{xy}}{T_{pd}}$
g_{19}	σ_{vx}	Standard deviation of x-directional speed $std\{\lvert\Delta X./\Delta T\rvert\} = std\left\{\left\lvert\frac{x(n)-x(n-1)}{t(n)-t(n-1)}\right\rvert\right\}$
g_{20}	σ_{vy}	Standard deviation of y-directional speed $std\{\lvert\Delta Y./\Delta T\rvert\} = std\left\{\left\lvert\frac{y(n)-y(n-1)}{t(n)-t(n-1)}\right\rvert\right\}$
g_{21}	σ_{vxy}	Standard deviation of speed $std\left\{\left\lvert\frac{\sqrt{\Delta x^2+\Delta y^2}}{\Delta t}\right\rvert\right\}$

DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016

g_{22}	μ_{ax}	x-directional average absolute acceleration $$\frac{\sum\lvert A_X\rvert}{T_{pd}}$$
g_{23}	μ_{ay}	y-directional average absolute acceleration $$\frac{\sum\lvert A_Y\rvert}{T_{pd}}$$
g_{26}	σ_{ay}	Standard deviation of y-directional acceleration $$std\{\lvert A_Y\rvert\}$$
g_{27}	σ_{axy}	Standard deviation of acceleration $$std\left\{\sqrt{\lvert A_X\rvert^2+\lvert A_Y\rvert^2}\right\}$$
g_{28}	μ_p	Average pen pressure in the entire signature= $mean\{p\}$
g_{29}	p_{max}	Maximum of pen pressure = $max\{p\}$
g_{30}	p_{med}	Median of writing pen pressure = $median\{p\}$
g_{31}	σ_p	Standard deviation of the pen pressure= $std\{p\}$
g_{32}	μ_θ	Average turn angle= $mean\{abs(\theta)\}$
g_{33}	$\sigma_{a\theta}$	Standard deviation of the absolute turn angle= $std\{abs(\theta)\}$
g_{34}	σ_θ	Standard deviation of the turn angle= $std\{\theta\}$
g_{35}	v_{+xmax}	Maximum velocity in positive x direction
g_{36}	v_{+ymax}	Maximum velocity in positive y direction
g_{37}	v_{-xmax}	Maximum velocity in negative x direction
g_{38}	v_{-ymax}	Maximum velocity in negative y direction
g_{39}	r_{dx}	$$\frac{\text{Total distance travelled in positive } x \text{ direction}}{\text{Total distance travelled in negative } x \text{ direction}}$$
g_{40}	r_{tx}	$$\frac{\text{Total time travelled in positive } x \text{ direction}}{\text{Total time traveled in negative } x \text{ direction}}$$
g_{41}	r_{dy}	$$\frac{\text{Total distance traveled in positive } y \text{ direction}}{\text{Total distance traveled in negative } y \text{ direction}}$$
g_{42}	r_{ty}	$$\frac{\text{Total time traveled in positive } y \text{ direction}}{\text{Total time traveled in negative } y \text{ direction}}$$
g_{43}	E_{SS}	Entropy of shape signature function $$E_{SS}=-\sum p_i(x).log_2(p_i(x))$$
g_{44}	fd_2	2nd Fourier descriptor co-efficient
g_{45}	fd_3	3rd Fourier descriptor co-efficient
g_{46}	fd_4	4th Fourier descriptor co-efficient
g_{47}	μ_{ph}	Mean of phase
g_{48}	σ_{ph}	Standard deviation of the phase = $std\{Ph\}$

DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016

Table 2. The feature and its weighting factor selection based on consistency: Global Features and their weighting factor for user 1 in the SVC 2004 database.

Rank	Features No.	$Mean(\mu)$	$Std\ (\sigma)$	$\frac{1}{\sigma_n} = \frac{\mu}{\sigma}$	$w\ (10^{-2})$
1	47	0.74	0.01	60.11	14.74
2	6	5225.05	172.63	30.27	7.42
3	7	3554.05	211.92	16.77	4.11
4	8	1.47	0.09	15.59	3.82
5	33	47.42	3.39	14.00	3.43
6	34	50.01	3.97	12.61	3.09
7	11	0.58	0.05	12.25	3.01
8	30	642.85	57.37	11.21	2.75
9	29	879.35	78.65	11.18	2.74
10	15	2.04	0.18	11.05	2.71
11	28	623.12	58.91	10.58	2.59
12	48	0.21	0.02	10.43	2.56
13	13	1.68	0.18	9.27	2.27
14	5	9.65	1.09	8.85	2.17
15	3	95.55	12.45	7.67	1.88
16	31	158.25	21.62	7.32	1.80
17	14	3.55	0.50	7.10	1.74
18	18	21.83	3.20	6.82	1.67
19	17	15.29	2.30	6.65	1.63
20	1	6.55	1.00	6.55	1.61
21	4	11.85	1.82	6.53	1.60
22	32	16.58	2.61	6.36	1.56
23	10	836.05	136.33	6.13	1.50
24	20	18.95	3.14	6.03	1.48
25	2	1456.15	244.03	5.97	1.46
26	44	109848.55	18762.09	5.85	1.44
27	16	10.73	1.94	5.54	1.36
28	35	49.53	9.22	5.37	1.32
29	19	13.90	2.59	5.37	1.32
30	38	54.38	10.21	5.33	1.31
31	43	0.32	0.06	5.23	1.28
32	12	0.75	0.15	5.05	1.24
33	21	13.13	2.66	4.94	1.21
34	46	38699.42	8506.76	4.55	1.12
35	9	620.10	139.21	4.45	1.09
36	41	0.40	0.10	4.12	1.01
37	42	0.80	0.20	4.00	0.98
38	26	1.78	0.45	3.94	0.97
39	23	0.13	0.03	3.92	0.96
40	27	1.36	0.35	3.84	0.94
41	24	0.18	0.05	3.81	0.93
42	45	39075.45	10443.37	3.74	0.92
43	40	1.82	0.49	3.71	0.91
44	39	5.75	1.57	3.66	0.90
45	36	32.44	9.13	3.55	0.87
46	37	15.37	4.37	3.52	0.86
47	22	0.10	0.03	3.47	0.85
48	25	1.49	0.43	3.47	0.85

Table 3. All the global features with common rank using consistency measures.

Rank	Average Rank	Feature No.	Rank	Average Rank	Feature No.
1	2.40	(47)	25	25.33	31
2	8.60	6	26	26.75	19
3	9.25	15	27	26.80	16
4	9.78	11	28	27.30	24
5	10.45	(48)	29	27.60	4
6	10.90	3	30	27.83	42
7	11.30	7	31	28.03	(43)
8	11.38	29	32	29.63	41
9	14.83	10	33	30.18	45
10	15.18	14	34	31.50	23
11	15.98	18	35	33.83	22
12	16.10	2	36	33.85	39
13	17.20	28	37	34.10	40
14	17.23	8	38	34.23	26
15	18.05	13	39	34.30	38
16	18.65	33	40	35.40	32
17	19.65	30	41	36.58	25
18	19.80	17	42	37.70	46
19	20.08	20	43	38.78	27
20	20.70	21	44	38.93	36
21	21.13	44	45	39.13	35
22	21.68	34	46	40.30	9
23	22.90	5	47	40.55	37
24	23.58	1	48	40.68	12

Table 4. All the global features with common rank using mRMR method.

Common rank	Average rank	Feature No.	Common rank	Average rank	Feature No.
1	09.45	22	25	24.82	31
2	13.85	23	26	24.87	35
3	15.10	25	27	25.07	38
4	15.70	24	28	25.95	42
5	16.62	28	29	26.20	37
6	18.07	(48)	30	26.30	09
7	18.45	30	31	26.60	46
8	18.57	(43)	32	27.67	11
9	18.65	27	33	28.75	41
10	18.77	02	34	29.45	12
11	18.80	03	35	29.82	04
12	19.07	29	36	30.07	07
13	19.40	(47)	37	30.07	40
14	19.72	10	38	30.55	14
15	19.97	26	39	31.50	05
16	20.22	18	40	31.52	06
17	21.25	19	41	31.55	01
18	21.25	44	42	31.60	15
19	21.62	21	43	32.80	39
20	22.15	45	44	33.37	08
21	22.20	17	45	34.60	13
22	22.22	16	46	34.60	33
23	22.87	20	47	34.82	32
24	23.55	36	48	35.82	34

As such, these features could be extracted in real-time as the signature is captured, thus eliminating the need to store the raw data at all if this is the only technique to be used. This is treated as a two-class classification problem when a new signature is presented for verification.

Forty-eight global features (parametric) are extracted for verification. These 48 global features are listed in Table 1 (Kar et al., 2018).

The global feature vector consists of the 48 global features as listed in Table 1.

$$G = [g_1, g_2, \dots, g_{48}] \quad (1)$$

Where, G is the global feature vector, and $g_1, g_2, etc.$, are the global features. In this paper, we propose three novel global features, g_{43}, g_{47} and g_{48}, as shown above in Table 1.

Novel Technique of Consistency Measure of the Global Features and its Ranking

Consistency estimation of global features is essential in a one-class classification framework, where only the genuine signatures are available. Consistency measures are used to select user-specific feature in a one-class classification framework (Lei and Govindaraju, 2005). The weighting factors of the feature vector are computed for every signer. Weighting factors of the features for user 1 in SVC 2004 are shown in Table 2. In statistics, consistency is measured by the inverse of the variance or the standard deviation. This technique does not apply directly to our feature set. Different features have different decimal scales, which are unknown and difficult to estimate. Consistency information will be lost if we normalize the features using the conventional normalization technique. A method of measuring consistency is proposed here. This technique is applicable when all the features are positive (or the same sign). The features are estimated such that all the features should be positive.

The normalized standard deviation of a feature is defined as the standard deviation divided by the mean of that particular feature. The weighting factor is estimated as the inverse of the variation of that feature. Therefore, a more consistent feature has more weight. Table 2 is sorted by descending order of the weighting factor. The calculation steps for the weighting factor are shown below. The following equation estimates the normalized using standard deviation. Features are formulated such that the value of all the features should be positive.

$$\sigma_n = \left[\frac{\sigma + \epsilon}{\mu}\right] = \left[\frac{standard\ deviation}{mean}\right] \quad (2)$$

The weighting. The following equation obtains the weighting factor.

$$w = \left[\frac{(1/\sigma_n)}{\sum(1/\sigma_n)}\right] \quad (3)$$

The weighting factor is normalized such that the sum of all weighting factors equals one. The standard deviation might be zero. To avoid division by zero, ϵ is added, a small quantity. ϵ is considered $0.001 \times \mu$. Global feature and its consistency based on the above derivation for user1 in the SVC 2004 database is shown in Table 2.

Table 3 lists all the global features with their rank and average rank. Features are ranked by consistency measure. Features are ranked differently for 40 different users. The average rank of a particular feature is estimated by considering all the ranks of those 40 users in SVC 2004. Average rank signifies the overall performance of a specific feature across the users.

The global feature vector consists of the 48 global features in Table 1.

$$G = [g_1, g_2, \dots, g_{48}] \quad (4)$$

G is the worldwide feature vector, and g_1, g_2, etc., are the global features.

mRMR based Global Features Ranking

The mRMR algorithm ranks the whole feature set according to its importance. To perform this, it evaluates the relevance of a feature and penalizes redundancy. The objective is to find the maximum dependency between the set of features X and class C, taking mutual information (I) (Chetry and Kar, 2024). Among the 48 global features, all common global features with common rank using minimum Redundancy Maximum Relevance (mRMR) method (Wang et al., 2018; Hanchuan Peng et al., 2005) are shown in Table 4. The average of these ranks are estimated across 40 users. Ten highest average ranks across the users selected features are dynamic features, which have more sensitivity and discrimination ability.

Global Feature Normalization

A critical process during signature verification is normalizing the signature signals before attempting to match them. Different signatures have different scales, even from the same user. Several techniques may be used to normalize this type, such as min-max, z-score, decimal scaling, mean absolute deviation, Median Absolute Deviation (MAD_1), double sigmoid function, and $tanh$ function. $Min - max$ and $z - score$ normalization schemes are efficient, whereas median absolute deviation, double sigmoid and $tanh$ methods are robust (Jain et al., 2005; Cabello-Solorzano et al., 2023). In the context of raw signature normalization, researchers most commonly

DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016

use min-max and z-score normalizations (Anggoro, 2019; Herwanto et al., 2021). The minimum and maximum coordinates of the signature signal are inconsistent, and it is susceptible to the part of the signature that provides the minimum and maximum conditions. A little distortion on that part changes the minimum or maximum. Normalization by $min - max$ is as follows:

$$X' = \frac{X - \min(X)}{\max(X) - \min(X)}, \qquad Y' = \frac{Y - \min(Y)}{\max(Y) - \min(Y)}. \tag{5}$$

Table 5. First ten global features of first ten genuine signatures for user 1 in SVC 2004

		Feature ID (Feature number in global feature list)									
		1	2	3	4	5	6	7	8	9	10
ID for genuine signatures	1	6	1202	83	11	9	5206	3354	1.552	482	720
	2	6	1182	84	10	9	5327	3434	1.551	452	730
	3	6	1252	89	13	9	5687	3578	1.589	472	780
	4	6	1192	81	9	8	5117	3548	1.442	491	701
	5	8	1342	91	11	11	5405	3644	1.483	581	761
	6	8	1242	81	9	9	5286	3994	1.323	581	661
	7	8	1262	90	11	9	5506	3726	1.478	511	751
	8	8	1392	85	12	12	5145	3642	1.413	692	700
	9	8	1522	92	17	11	5118	3583	1.428	750	772
	10	8	1382	82	14	11	5241	3793	1.382	711	671
ID for forged signatures	1	6	1673	114	11	9	5175	3331	1.554	643	1030
	2	6	1592	118	8	9	5971	3331	1.793	522	1070
	3	6	3465	185	14	9	5397	3238	1.667	1724	1741
	4	6	3114	180	16	14	5715	3134	1.824	1424	1690
	5	6	2994	189	12	10	6204	3246	1.911	1214	1780
	6	6	3235	214	14	9	5255	3136	1.676	1205	2030
	7	5	2384	162	11	11	5150	3503	1.470	843	1541
	8	5	2153	158	10	9	5019	3777	1.329	663	1490
	9	6	2433	180	10	12	6082	3515	1.730	743	1690
	10	6	2403	177	10	10	5539	3559	1.556	743	1660

Table 6. Min-max normalization on the data of Table 5 is shown here.

		Feature ID (Feature number in global feature list)									
		1	2	3	4	5	6	7	8	9	10
ID of genuine signatures	1	0.333	0.009	0.015	0.333	0.167	0.448	0.256	0.461	0.047	0.043
	2	0.333	0.000	0.023	0.222	0.167	0.515	0.349	0.459	0.024	0.050
	3	0.333	0.031	0.060	0.556	0.167	0.714	0.516	0.517	0.039	0.087
	4	0.333	0.004	0.000	0.111	0.000	0.398	0.481	0.295	0.054	0.029
	5	1.000	0.070	0.075	0.333	0.500	0.558	0.593	0.357	0.123	0.073
	6	1.000	0.026	0.000	0.111	0.167	0.492	1.000	0.117	0.123	0.000
	7	1.000	0.035	0.068	0.333	0.167	0.614	0.688	0.349	0.069	0.066
	8	1.000	0.092	0.030	0.444	0.667	0.414	0.591	0.251	0.208	0.028
	9	1.000	0.149	0.083	1.000	0.500	0.399	0.522	0.275	0.252	0.081
	10	1.000	0.088	0.008	0.667	0.500	0.467	0.766	0.205	0.223	0.007
ID of forge signatures	1	0.333	0.215	0.248	0.333	0.167	0.431	0.229	0.463	0.170	0.270
	2	0.333	0.180	0.278	0.000	0.167	0.871	0.229	0.822	0.078	0.299
	3	0.333	1.000	0.782	0.667	0.167	0.553	0.121	0.633	1.000	0.789
	4	0.333	0.846	0.744	0.889	1.000	0.729	0.000	0.868	0.770	0.752
	5	0.333	0.794	0.812	0.444	0.333	1.000	0.130	1.000	0.609	0.817
	6	0.333	0.899	1.000	0.667	0.167	0.475	0.002	0.646	0.602	1.000
	7	0.000	0.527	0.609	0.333	0.500	0.417	0.429	0.337	0.324	0.643
	8	0.000	0.425	0.579	0.222	0.167	0.344	0.748	0.125	0.186	0.606
	9	0.333	0.548	0.744	0.222	0.667	0.932	0.443	0.728	0.247	0.752
	10	0.333	0.535	0.722	0.222	0.333	0.632	0.494	0.467	0.247	0.730

If the same amount of distortion occurs in the central location and far from the central location, the latter affects the standard deviation more. So, the normalization scale will be different. However, distortion should be considered equally throughout the signature in the case of signature verification. Here, $min - max$ normalization techniques, as shown in equation 5, are used for global feature normalization.

Results of Normalization of the Global Feature

Different global features have distinct scaling factors. For example, several strokes are an integer, and generally, it is in one or two decimal digits, whereas the length of the signature is a number of pixels, and it is in the order of thousands. The mean, absolute deviation and standard deviation of all the 48 global features are shown in Table 2 for user 1 in the SVC 2004 database; 40 persons and their corresponding 20 genuine signatures are considered to calculate these features. These global features must be normalized before selecting and verifying results between two result sets. Different signatures have different scales, even from the same user. In the context of feature normalization, researchers most commonly use min-max and z-score normalizations. Here, in our case, we have used min-max normalization. In min-max normalization, all the features are converted to the range of (0-1). Let it be assumed that F_{jk}^{i} is k^{th} global feature of the j^{th} signature of a i^{th} person. i.e., person ID signifies by i variable, signature ID signifies by j variable and feature ID signifies by k variable. After normalization if it is denoted by $F_{jk}^{i}{}'$, the equation for min-max normalization can be written as

$$F_{jk}^{i}{}' = \frac{F_{jk}^{i} - \min_{j}(F_{jk}^{i})}{\max_{j}(F_{jk}^{i}) - \min_{j}(F_{jk}^{i})} \quad (6)$$

Table 5 shows an example of raw features without normalization. These sample features are taken from "user 1" in the SVC 2004 database. Here, only 10 global features out of 48 global features are shown for the first 10 genuine signatures. The result of the min-max normalization on the features in Table 5 is shown in Table 6.

Results and Discussion

The global features are considered to characterize an online signature. Here 48 global features are extracted as shown above in Table 1. Three novel features are proposed here viz. Entropy of shape signature function, Mean of phase and Standard deviation of the phase. A novel technique of consistency measure is studied here by computing the normalized standard deviation for feature ranking. This is done for signature of a particular user and across all the users. The most consistent features with their common rank are shown in Table 3. It is seen that for user1 geometric features are mostly consistent across the user as shown in Table 2. However, there is lot of redundancy in these features, and therefore redundancy also needs to be tested to achieve greater discriminating ability. Feature ranking shows the importance of each features for a particular system and helps in feature selection to select the more discriminating features for a particular system. It saves the computational time and size of the model. In ranking, it is pertinent to mention here that our proposed novel feature number 47 and 48 are ranked 1st and 5th respectively among all the 48 global features using consistency measure as shown in Table 3. which depicts its importance in signature verification purpose. If we compare our consistency, based ranking with mRMR based ranking technique it is observed that mRMR also ranks our proposed novel features number 48 and 43 as 6th and 8th respectively as shown in Table 4 among all 48 global features. Our novel feature ranking technique based on consistency measure as well as well-established mRMR techniques both ranks our two out of three novel features among the top ten features giving justification the importance of the three novel features extracted and the importance of novel feature ranking technique itself. Consistency Measure based all extracted feature ranking and mRMR based all extracted feature ranking can be seen in Table 3 and Table 4 respectively for various comparisons and analysis. Different global features has different scaling factor. Min-max normalization on the data of Table 5 is shown in Table 6 where all the features are converted to (0-1) range. The proposed algorithms are verified with SVC 2004 database.

Conclusions

Three novel extracted global features shows its importance in both the proposed ranking method as well as in well-established mRMR based ranking process. Our proposed consistency measure based feature ranking and well-established mRMR based ranking shows almost similar performance. If we consider to take best 32 features out of 48 features which are commonly ranked by both the ranking process then it is observed that almost 60% i.e. 19 features out of 32 are commonly selected. Min-max normalization on the features are successfully implemented where all features having different scaling factor are converted to (0-1) range. However, comparisons of verification accuracy of online signatures using various numbers of extracted features ranked using both the techniques is a subject of further research.

DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Anggoro, D. A. (2019). Improving accuracy by applying Z-score normalization in linear regression and polynomial regression models for real estate data. *International Journal of Emerging Trends in Engineering Research*, *7*(11), 549-555. https://doi.org/10.30534/ijeter/2019/247112019

Cabello-Solorzano, K., Ortigosa de Araujo, I., Peña, M., Correia, L., & J. Tallón-Ballesteros, A. (2023). The impact of data normalization on the accuracy of machine learning algorithms: A comparative analysis. *18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023)*, 344-353. https://doi.org/10.1007/978-3-031-42536-3_33

Chetry, B. P., & Kar, B. (2024). Kruskal Wallis and mRMR Feature Selection based Online Signature Verification System using Multiple SVM and KNN. *International Journal of Experimental Research and Review*, *42*, 298–311. https://doi.org/10.52756/ijerr.2024.v42.026

Guru, D., & Prakash, H. (2009). Online signature verification and recognition: An approach based on symbolic representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(6), 1059-1073. https://doi.org/10.1109/tpami.2008.302

Hanchuan Peng, Fuhui Long, & Ding, C. (2005). Feature selection based on mutual information criteria of Max-dependency, Max-relevance, and MIN-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1226-1238. https://doi.org/10.1109/tpami.2005.159

Herwanto, H. W., Handayani, A. N., Wibawa, A. P., Chandrika, K. L., & Arai, K. (2021). Comparison of Min-Max, Z-score and decimal scaling normalization for zoning feature extraction on Javanese character recognition. *2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*. https://doi.org/10.1109/iceeie52663.2021.9616665

Jain, A., Nandakumar, K., & Ross, A. (2005). Score normalization in multimodal biometric systems. *Pattern Recognition*, *38*(12), 2270-2285. https://doi.org/10.1016/j.patcog.2005.01.012

Kar, B., & K. Dutta, P. (2012). SVM based signature verification by fusing global and functional features. *International Journal of Computer Applications*, *60*(16), 34-39. https://doi.org/10.5120/9778-4359

Kar, B., Mukherjee, A., & Dutta, P. K. (2018). Stroke point warping-based reference selection and verification of online signature. *IEEE Transactions on Instrumentation and Measurement*, *67*(1), 2-11. https://doi.org/10.1109/tim.2017.2755898

Kashi, R. S. (1996). On-line handwritten signature verification using stroke direction coding. *Optical Engineering*, *35*(9), 2526. https://doi.org/10.1117/1.600857

Leclerc, F., & Plamondon, R. (1994). Automatic signature verification: The state of the art—1989–1993. *International Journal of Pattern Recognition and Artificial Intelligence*, *08*(03), 643-660. https://doi.org/10.1142/s0218001494000346

Lei, H., & Govindaraju, V. (2005). A comparative study on the consistency of features in on-line signature verification. *Pattern Recognition Letters*, *26*(15), 2483-2489. https://doi.org/10.1016/j.patrec.2005.05.005

Maiorana, E., Campisi, P., Fierrez, J., Ortega-Garcia, J., & Neri, A. (2010). Cancelable templates for sequence-based biometrics with application to on-line signature recognition. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *40*(3), 525-538. https://doi.org/10.1109/tsmca.2010.2041653

Najda, D., & Saeed, K. (2022). Impact of augmentation methods in online signature verification. *Innovations in Systems and Software Engineering*, 20, 477-483. https://doi.org/10.1007/s11334-022-00464-4

Plamondon, R., & Srihari, S. (2000). Online and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(1), 63-84. https://doi.org/10.1109/34.824821

Wang, X., Tao, Y., & Zheng, K. (2018). Feature selection methods in the framework of mRMR. *2018 Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC)*. https://doi.org/10.1109/imccc.2018.00307

Xia, Z., Shi, T., Xiong, N. N., Sun, X., & Jeon, B. (2018). A privacy-preserving handwritten signature verification method using Combinational features and secure KNN. *IEEE Access*, *6*, 46695-46705. https://doi.org/10.1109/access.2018.2866411

Yeung, D., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., & Rigoll, G. (2004). SVC2004: First international signature verification competition. *Biometric Authentication*, 16-22. https://doi.org/10.1007/978-3-540-25948-0_3

Zhang, K., Pratikakis, I., Cornelis, J., & Nyssen, E. (2000). Using landmarks to establish a point-to-point correspondence between signatures. *Pattern Analysis & Applications*, *3*(1), 69-75. https://doi.org/10.1007/s100440050007

How to cite this Article:

Biswajit Kar, Bhimraj Prasai Chetry and Sayan Das (2024). Feature Ranking Using Novel Consistency Measure by Normalized Standard Deviation and Proposal of Three Novel Global Features for Online Signature Verification. *International Journal of Experimental Research and Review*, *44*, 185-195.

DOI: https://doi.org/10.52756/ijerr.2024.v44spl.016