

## PREPROCESSING APPROACHES IN DTW BASED ONLINE SIGNATURE VERIFICATION

<sup>1</sup> Mohammad SALEEM\*, <sup>2</sup> Bence KOVARI

<sup>1,2</sup> Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Muegyetem rkp 3  
1111 Budapest, Hungary, e-mail: <sup>1</sup>msaleem@aut.bme.hu, <sup>2</sup>kovari@aut.bme.hu

Received 17 March 2019; accepted 8 April 2019

**Abstract:** Amongst different approaches, dynamic time warping has shown promising results during the online signature verification competitions of previous years. To improve the results of dynamic time warping, different preprocessing steps may be applied and different dimensions of the samples may be compared. The choice of preprocessing steps and comparing dimensions may significantly influence the results. Thus, to aid researchers with these decisions, a comparison made between the results of promising preprocessing algorithms as horizontal scaling, vertical scaling and alignment using dynamic time warping in different dimensions and their combinations on two datasets (SVC2004 and MCYT-100). The results showed that preprocessing methods made a very promising improvement in the verification accuracy.

**Keywords:** Biometrics, Pattern recognition, Online signature verification, Dynamic time warping, Preprocessing

### 1. Introduction

Online signature verification is an important field in forensic examination. Although fingerprint or face-based identification methods are more secure, signatures are still used in several fields around the world. Furthermore, based on the input data, there are two types of signatures: offline and online. Accordingly, in offline signatures, data is an image of the signature like bank checks and documents), and in online signatures, data is captured using a special device, for example a pressure-sensitive tablet. The data in

---

\* Corresponding Author

online signatures includes many features, for instance strokes, speed, pen pressure, pen position, etc. These features make the signature unique and hard to forge.

In this work, online signature verification systems have been focused. The main steps of this work were similar in most of the verification systems. The first step was data acquisition where data was collected from one or different databases. Signature Verification Competition 2004 (SVC2004) [1] and the database of the Spanish Ministry of Science and Technology (MCYT-100) [2] databases were used in this work. In addition, other databases were available, for instance Sabanci University Signature database (SUSIG) [3], Biometric recognition group database [4] and the Malaysian signatures' database (SIGMA) [5].

The next step was feature extraction. Since features like position, pressure and velocity describe the signature and make it unique, they were used to compare signatures and identify the signer. X and Y positions of the signatures were used to apply the verification method.

Data preprocessing is very important step to improve the results, not only on signature verification systems but also in different fields [6], [7]. Following the above, preprocessing steps were applied to improve the accuracy of the similarity measurement. Accordingly, the three main methods: alignment, horizontal scaling and vertical scaling were applied on the 3 dimensions of Dynamic Time Warping (DTW). Alignment and scaling can be used in different methods (see *Table I*) but here, alignment to the origin used, which was done by subtracting the mean value from each point of the signature to re-align it so that it started from the origin. Subsequently, signature scaling was done by multiplying the signature points through a specific ratio. Thus, scaling could be applied horizontally, vertically or in both directions. Alignment, scaling and other normalization preprocessing methods are shown in *Table I*.

*Table I*

Some normalization preprocessing methods used in signature verification

Preprocessing	References
Alignment, alignment of the center of gravity, translation, location normalization, alignment based on GMM, mean removal	[8], [9], [10], [11]
Scaling, size normalization, horizontal scaling, vertical scaling, length normalization	[12], [13], [14]
Rotation, rotation normalization, removing the angle using DTW	[15], [16], [17]
Time normalization	[14]
Min. normalization	[18]
Max. normalization	[18]

The last step was verification where the preprocessed extracted features used to match the signatures and classify them into original and forged categories. Accordingly, distances between the reference signatures were used in the comparison step. One of the most popular methods of distance measurement is DTW since it finds the best non-linear alignment of two vectors where the overall distance between them is minimized. DTW has shown promising results in the signature verification field [19], [20].

Resultantly, the signer provided a set of original and forged signatures and a set of the original signatures was used as reference for the matching measurements. In addition, they were grouped into training and testing sets, which will later be explained in detail. The results of applying all these steps using DTW on different dimensions, alignment, and scaling for preprocessing are discussed and compared in the results section.

## 2. Related work

There are many popular function-based verification methods, such as DTW [21] and hidden Markov model [22]. The DTW algorithm was established in the 70's when it was mainly applied in word recognition [23]. Subsequently, it was used for clustering and classification in many different domains, including different recognition systems, for instance speaker recognition [24] and handwritten recognition [25]. In signature verification, DTW matches tested signatures directly with reference samples of the user signatures. Moreover, the preprocessing step was very important in verification systems since different image enhancement methods were applied in [26] and [27] as a preprocessing step. Additionally, a normalization method proposed in [28] for image resizing based on the multi-rate filter theory to reduce size-invariance. The preprocessing step was required before applying DTW to improve the accuracy of the system. Consequently, alignment to the mass center, scaling, and Fourier transform methods were commonly used in this step [29]. In accordance with the above, DTW showed great results in the previous signature and handwritten recognition competitions [19] where most of the winners applied it.

## 3. Contribution

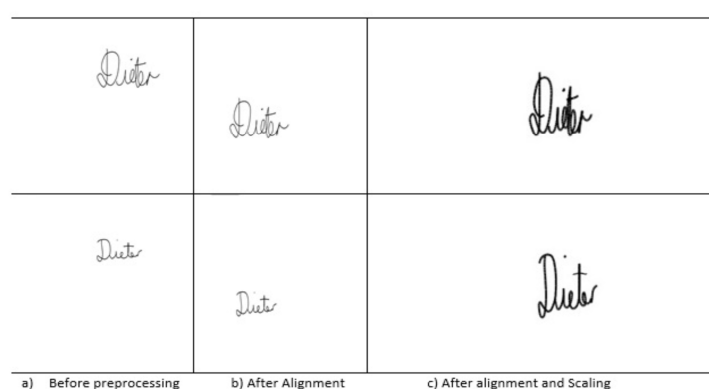
The verification of the signature was based on different processing steps. Accordingly, the data in this work consists of original and forged signatures for each signer. After applying the preprocessing algorithms for each signature, DTW was used to calculate the distances between each pair of signatures in the training set. Moreover, the same preprocessing algorithms were applied to the tested signatures and compared to a threshold calculated from the training data and defined as an accepted or rejected original signature. Subsequently, the results of the accepted and rejected signatures were compared to decide the best preprocessing algorithm with the least error rate using two different databases. These results helped the signature verification field researchers choose the most significant and useful approaches for the preprocessing step of their works in the field.

## 4. Methodology

Alignment was an important preprocessing algorithm. Consequently, although a signer may create similar signatures every time, there is always some difference between each signature. Furthermore, one of the differences is between the starting and ending points of the signature, which may result in non-accurate measurements. Thus,

the alignment method was applied for each signature before implementing any calculations. Accordingly, signature alignment can be applied for both X and Y coordinates since the signature was aligned to the origin of the coordinate by subtracting the mean value of the coordinate values from each point.

Furthermore, the scale may be different between the signatures. In this case, the scaling preprocessing method was needed for more accurate measurements. In addition, the scaling could be applied horizontally or vertically. In horizontal scaling, the X values were scaled by multiplying each point to a ratio that keeps the point in a specific range of values. For vertical scaling, the same method was applied to the Y values. *Fig. 1* shows an example of the preprocessing steps applied on the signature, the first and second columns shows the signatures before alignment and scaling, and the last columns shows the result.



*Fig. 1.* Signatures before and after the preprocessing step

Subsequently, after applying the preprocessing algorithms on the signature, the dynamic time warping algorithm was used to calculate differences between the signatures. Since DTW is a well-known technique to find an optimal alignment between two given time sequences under certain restrictions, it finds the best nonlinear alignment and a 'warping path' between 2 vectors to find the minimum distance between them. The algorithm applied on different dimensions was X, Y and XY.

The distances between the signatures were used for classification. The average of the distances in the original training signature group was used as a threshold for classification. The test group of signatures consisted of the same number of original and forged signatures. Accordingly, each signature of the test group was tested by the system and assigned as original and forged. Since the accuracy of any verification system depends on the error rate, there were two types of errors in the system:

- 1) when the system accepted a forged signature as original; and
- 2) when it rejected an original signature.

Furthermore, the original test group was used to calculate the False Rejection Rate (FRR) and the forged test group was used to calculate the False Acceptance Rate (FAR). Subsequently, the average of the FAR and FRR was used to evaluate the accuracy of the system, which is known as the Average of Error Rate (AER).

In this work all the combinations of preprocessing and DTW for different dimensions were tested. In the beginning, the algorithms were tested without the preprocessing work. Following this, each algorithm was tested separately for different dimensions. The alignment was applied on the X, Y and XY dimensions with and without scaling (on different dimensions). Consequently, all the previous combinations were applied using DTW algorithms for the different dimensions of X, Y and XY.

## 5. Experimental results

In this work, the system was applied using two different available databases. The SVC2004 database was freely accessible and consisted of 40 signatures for each of the 40 signers. The signature files consisted of data about the pen position, stroke points, pen pressure, etc. The strokes were then used to draw and compare the signatures. Consequently, each signer had 20 original signatures and 20 forged signatures. The MCYT-100 database was obtained through 100 signers, 25 client signatures, and 25 highly skilled forgeries each [2]. Additionally, the pen trajectory, pressure and pen azimuth acquired during the signing process using the WACOM pen tablet, model INTUOS A6 USB [30].

Furthermore, signatures were grouped into training and testing sets. Subsequently, a group of original signatures were used as references for distance calculations. The same sets were then used as a training set along with the same number of forged signatures. In addition, the rest of the original and forged signatures were used for testing. In the SVC2004, 10 original signatures were used as references: 10 original and forged signatures for training and another 10 original and forged for testing. Moreover, for the MCYT-100, the same number of signatures used in SVC2004 was used as a training set and references, but 15 original and forged signatures were used for testing.

The distance measurements were applied for the first 10 signatures and the results were used to calculate the threshold for classification. Resultantly, the threshold used was a summation of the average and standard deviations of DTW distances between the signatures. In addition, the test set of the original signatures was used to calculate the false rejection rate. Subsequently, the average of the distances was calculated and compared to the threshold. If equal or less than the threshold, it was accepted, and rejected if otherwise. The rejected signatures of this set were falsely rejected since the ideal system would accept them all. Thus, the number of rejected signatures in this set was divided by the total number of the signatures in the set to calculate the FRR.

The test set of the forged signatures was used to calculate the false acceptance rate. The average of the distances for each signature was calculated and compared to the threshold. Accordingly, the signature would be accepted if the average of the distances was less than or equal to the threshold. This set of signatures should be rejected in the ideal system since any accepted signature would be considered as false acceptance. Therefore, the number of the falsely accepted signatures of this set was divided by the

total number of signatures in the set in order to calculate the FAR. Subsequently, to evaluate the system, both FAR and FRR were used and the average error rate was used for the evaluation, which was the average of FAR and FRR. The AER was calculated for each compilation of the preprocessing and distance measurement methods. Resultantly, the best system was the system with the less AER.

### 5.1. SVC2004 results

Table II shows the result data for all the combinations. According to the results, applying the DTW algorithm on the XY dimension gave the best results compared to the X and Y dimensions (see Fig. 2). Thus, the average of the AER was 0.16734 for XY, 0.259 and 0.1818 for X and Y respectively.

Table II

The verification results for all the combinations

DTW	Alignment	Scale horizontally	Scale vertically	AER (SVC2004)	AER (MCYT_100)
X				0.286	0.375
X	*			<b>0.216</b>	<b>0.052</b>
X	*	*		0.268	0.084
X	*		*	<b>0.216</b>	<b>0.052</b>
X	*	*	*	0.268	0.084
X		*		0.268	0.084
X			*	0.268	0.084
X			*	0.286	0.375
Y				0.264	0.337
Y	*			<b>0.114</b>	<b>0.034</b>
Y	*	*		<b>0.114</b>	<b>0.034</b>
Y	*		*	0.193	0.067
Y	*	*	*	0.196	0.067
Y		*		0.196	0.067
Y		*	*	0.183	0.067
Y			*	0.196	0.067
XY				0.220	0.359
XY	*			<b>0.091</b>	<b>0.019</b>
XY	*	*		0.136	0.038
XY	*		*	0.134	0.039
XY	*	*	*	0.173	0.051
XY		*		0.171	0.051
XY		*	*	0.171	0.051
XY			*	0.243	0.036

Moreover, the experiment showed that the alignment method gave better results. Thus, the average of the AER when applying the alignment method was 0.17645 compared to 0.2292 without alignment. Furthermore, scaling produced less improvement in the work and the horizontal scaling results were slightly better than vertical scaling.

Based on the previous results, the best result was achieved by using the DTW algorithm on the XY dimension with a preprocessing method of alignment without scaling. Accordingly, the best result was 0.09125.

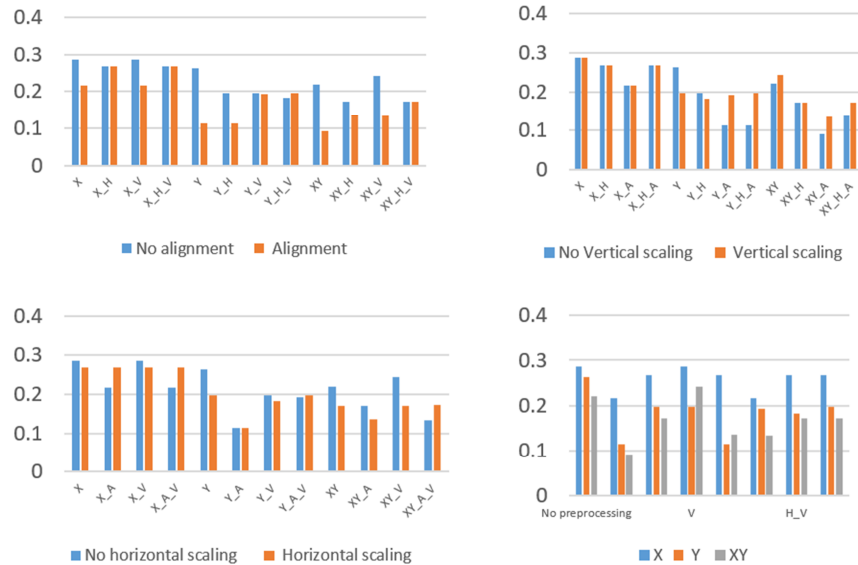


Fig. 2. A comparison between the effect of applying each method on the results (where  $A$ = alignment,  $H$ =horizontal scaling,  $V$ =vertical scaling)

### 5.2. MCYT-100 results

The results of MCYT-100 database show almost the same effect of the preprocessing steps on the results but with better improvement. Therefore, applying the DTW algorithm on the XY dimension gave the best result compared to the X and Y dimensions (see Fig. 3). The average of the AER was 0.0806 for XY, 0.1486 and 0.0923 for X and Y respectively.

Furthermore, the experiment showed that using the alignment method gave better results. The average of the AER when applying the alignment method was 0.05182 compared to 0.1626 when no alignment was applied. Moreover, scaling gave less improvement than alignment in this database as well, but with good improvement compared to not using any preprocessing methods.

Thus, based on the previous results, we can see that the best result was achieved by using the DTW algorithm on the XY dimension with a preprocessing method of alignment without scaling. The best result was 0.0193.

## 6. Conclusion

The accuracy of the DTW-based classifier can be improved by using preprocessing algorithms as peer the results. In addition, the 3 different preprocessing methods

(alignment, horizontal scaling and vertical scaling) and their combination on different dimensions were applied to the signatures before verification. A total of 24 different simulations were applied on two sets of databases, which are SVC2004 and MCYT-100.

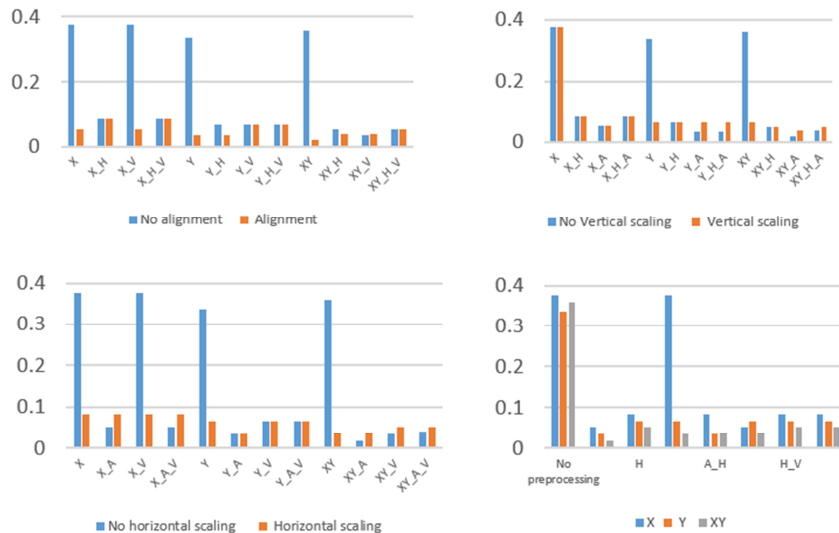


Fig. 3. A comparison between the effect of applying each method on the results (where A=alignment, H=horizontal scaling, V=vertical scaling).

These results showed the importance of preprocessing. Furthermore, the matching of DTW distances achieved the best results when it was applied on the XY dimension. Consequently, alignment showed better result improvement than horizontal and vertical scaling. The improvement was very promising and varied between the different methods. For the SVC2004 data, the best results were when applying DTW on XY dimension using only alignment for preprocessing where the AER reduced from 0.28625 to 0.09125 with 19.5% accuracy improvement. The same applied for the MCYT-100 database where the best results were achieved when applying DTW on XY dimension with only alignment preprocessing. Resultantly, the AER reduced from 0.359 to 0.0193 with 33.97% accuracy improvement.

### Acknowledgments

The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

This work was performed in the frame of FIEK\_16-1-2016-0007 project, implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the FIEK\_16 funding scheme.



### Open Access statement

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited, a link to the CC License is provided, and changes - if any - are indicated. (SID\_1)

### References

- [1] Yeung D. Y., Chang H., Xiong Y., George S., Kashi R., Matsumoto T. Rigoll G. SVC2004: First international signature verification competition, in: Zhang D., Jain A. K. (Eds) *Lecture Notes in Computer Science*, Vol. 3072, 2004, pp. 16–22.
- [2] Ortega-Garcia J., Fierrez-Aguilar J., Simon D., Gonzalez J., Faundez-Zanuy M., Espinosa V. Escudero, Statue M., Hernaez I., Igarza J. J., Vivaracho C., Escudero D., Moro Q. I. MCYT baseline corpus: a bimodal biometric database, *IEEE Proceedings-Vision, Image and Signal Processing*, Vol. 150, No. 6, 2003, pp. 395–401.
- [3] Kholmatov A., Yanikoglu B. SUSIG: an on-line signature database, associated protocols and benchmark results, *Pattern Analysis and Applications*, Vol. 12, No. 3, 2009, pp. 227–236.
- [4] Biometrics and data pattern analytics, <http://atvs.ii.uam.es/atvs/databases.jsp>, (last visited 1 April 2019).
- [5] Ahmad S. M. S., Shakil A., Ahmad A. R., Agil M., Balbed M., Anwar R. M., SIGMA-A Malaysian signatures' database, *IEEE/ACS International Conference on Computer Systems and Applications*, Doha, Qatar, 31 March - 4 April 2008, pp. 919–920.
- [6] Kahlout K., Csorba K., Ekler P. Categorizing of inhaling behaviors using signal processing and data mining techniques, *Pollack Periodica*, Vol. 13, No. 3, 2018, pp. 119–130.
- [7] Gocsál Á. Female listeners' personality attributions to male speakers: The role of acoustic parameters of speech, *Pollack Periodica*, Vol. 4, No. 3, 2009, pp. 155–165.
- [8] Tolosana R., Vera-Rodriguez R., Fierrez J., Ortega-Garcia J. Feature-based dynamic signature verification under forensic scenarios, *3<sup>rd</sup> International Workshop on Biometrics and Forensics*, Gjøvik, Norway, 3-4 March 2015, pages 1–6.
- [9] Xia X., Chen Z., Luan F., Song X. Signature alignment based on GMM for on-line signature verification, *Pattern Recognition*, Vol. 65, 2017, pp. 188–196.
- [10] Ahrabian K., Babaali B. On usage of autoencoders and siamese networks for online handwritten signature verification, *arXiv.org > cs > arXiv:1712.02781*, 2017,
- [11] Wirocius M., Ramel J. Y., Vincent N. Selection of points for on-line signature comparison, *Ninth International Workshop on Frontiers in Handwriting Recognition*, Kokubunji, Tokyo, Japan, 26-29 October 2004, pp. 503–508.
- [12] Mohammadi M. H., Faez K. Matching between important points using dynamic time warping for online signature verification, *J. Sel. Areas Bioinf*, 2012, pages 1–7.
- [13] Nilchiyan M. R., Yusof R. B., Alavi S. E. Statistical on-line signature verification using rotation-invariant dynamic descriptors, *10th Asian Control Conference*, Kota Kinabalu, Malaysia, 31 May-3 June 2015, pages 6.
- [14] López-García M., Ramos-Lara R., Miguel-Hurtado O., Cantó-Navarro E. Embedded system for biometric online signature verification, *IEEE Transactions on Industrial Informatics*, Vol. 10, No. 1, 2014, pp. 491–501.

- [15] Ibrahim M. T., Khan M. A., Alimgeer K. S., Khan M. K., Taj I. A., Guan L. Velocity and pressure-based partitions of horizontal and vertical trajectories for on-line signature verification, *Pattern Recognition*, Vol. 43, No. 8, 2010, pp. 2817–2832.
- [16] Rashidi S., Fallah A., Towhidkhah F. Feature extraction based DCT on dynamic signature verification, *Scientia Iranica*, Vol. 19, No. 6, 2012, pp. 1810–1819.
- [17] Jindal U., Dalal S., Dahiya N. A combine approach of preprocessing in integrated signature verification (ISV), *International Journal of Engineering & Technology*, Vol. 7, No. 1.2, 2018, pp. 155–159.
- [18] Ansari A. Q., Hanmandlu M., Kour J., Singh A. K. Online signature verification using segment-level fuzzy modeling, *IET Biometrics*, Vol. 3, No. 3, 2014, pp. 113–127.
- [19] Malik M. I., Ahmed S., Marcelli A., Pal U., Blumenstein M., Alewijnse L., Liwicki M. ICDAR2015 competition on signature verification and writer identification for on-and off-line skilled forgeries (SigWlcomp2015), *13th International Conference on Document Analysis and Recognition*, Tunis, Tunisia, 23-26 August 2015, pp. 1186–1190.
- [20] Malik M. I., Liwicki M., Alewijnse L., Ohyama W., Blumenstein M., FoundB. ICDAR 2013 competitions on signature verification and writer identification for on- and offline skilled forgeries, *12th International Conference on Document Analysis and Recognition*, Washington, DC, USA, 25-28 August 2013, pp. 1477–1483.
- [21] Kholmatov A. A. Biometric identity verification using on-line & off-line signature verification, *Master Thesis*, Sabanci University, 2003.
- [22] El-Yacoubi A., Gilloux M., Sabourin R., Suen C. Y. An HMM-based approach for off-line unconstrained handwritten word modeling and recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 8, 1999, pp. 752–760.
- [23] Fierrez J., Ortega-Garcia J., Ramos D., Gonzalez-Rodriguez J. HMM-based on-line signature verification: Feature extraction and signature modeling, *Pattern Recognition Letters*, Vol. 28, No. 16, 2007, p. 2325–2334.
- [24] Auckenthaler R., Carey M., Lloyd-Thomas H. Score normalization for text-independent speaker verification system, *Digital Signal Processing*, Vol. 10, No. 1, 2000, pp. 42–54.
- [25] Fischer A., Keller A., Frinken V., Bunke H. Lexicon-free handwritten word spotting using character HMMs, *Pattern Recognition Letters*, Vol. 33, No. 7, 2012, pp. 934–942.
- [26] Ulges A., Lampert C. H., Breuel T. M. Document image dewarping using robust estimation of curled text lines, *Proceedings of the Eighth International Conference on Document Analysis and Recognition*, Sydney, Australia, 31 August-1 September 2005, pp. 1001–1005.
- [27] Zhang L., Tan C. L. Warped image restoration with applications to digital libraries, *Eighth International Conference on Document Analysis and Recognition*, Seoul, South Korea, 31 August-1 September 2005, pp. 192–196.
- [28] Jayadevan R., Kohle S. R., Patil P. M., Pal U. Automatic processing of handwritten bank cheque, *International Journal on Document Analysis and Recognition*, Vol. 15, No. 4, 2012, pp. 267–296.
- [29] Al-Shoshan A. I. Handwritten signature verification using image invariants, *International Conference on Computer Graphics, Imaging and Visualization*, Sydney, Australia, 26-28 July 2006, pp. 173–176.
- [30] Wacom, <http://www.wacom.com>, (last visited 1 April 2019).