Workbook for e-Education: Children's Online Handwriting Quality Analysis.

Hanen Akouaydi, Yahia Hamdi, Houcine Boubaker, Faouzi Alaya Cheikh and Adel M. Alimi

Abstract—This paper describes an innovative e-learning project which is the development of a digital workbook that helps teaching handwriting at school. In this work, we propose a new qualitative and quantitative analysis process of cursive handwriting. This process detects automatically mistakes, gives a real-time feedback and helps teachers evaluate children's writing skills. The main aim of this digital workbook is to help children learn how to write correctly. The proposed process is composed of five main criteria: shape, direction, stroke order, position respect to the reference lines and kinematics of the trace. It analyzes the handwriting quality and gives automatically feedback based on the Beta-Elliptic Model using similarity detection (SD) and dissimilarity distance (DD) measure. Our work apprehends dynamic and visual representation of the acquired traces and selects efficient features adapted to various handwriting styles and multiple script languages such as Arabic, Latin, digits, and symbols drawing. It demonstrates that beta-elliptic is not only a model for segmentation and recognition but also a tool to evaluate handwriting. Our application offers two interactive interfaces respectively dedicated to learners, and experts or teachers who can adapt it easily to the specificity of each child. The validation of the proposed system is done on a database collected in Tunisia primary schools with 400 children. Experimental results show that the efficiency and robustness of our suggested framework that do help teachers and children by offering positive feedback throughout the handwriting learning process using tactile digital devices.

Index Terms—Online handwriting, Analysis, Evaluation, Beta-elliptic model, Mobile Device, Segmentation, Cursive Handwriting.

1 INTRODUCTION

WITH emergence of tablets and smartphones, the existent kind of tablets allows writing on screen with a sensation very similar to writing on paper. In our application, children's writings are mainly analyzed according to shape, direction and order. With pen-based tablet or finger, various criteria can be studied or analyzed in the light of shape, pressure and direction. Digital notebook gives immediately and personalized feedback to children and also to teachers or trainers who teach kids how to write. Another advantage of the notebook is the ability to adopt the educational progression to each child. The notebook is currently composed of Letter Writing for Arabic/French/English, Word Writing Arabic/French/English, Digit Writing, Symbol Recognition and Drawing writing. All evaluated writings are cursive. The evaluation implies dealing with the segmentation of acquired traces into basic shapes and the extraction of beta-elliptic parameters. We define models and we compare the child's trace by applying some rules. We have to make difference between recognition and analysis of handwriting tasks. The analysis evaluates the target symbol by comparing it with a model of reference. However, the recognition task is to identify the written symbol independently to its quality and style of writing. This paper describes a handwriting analyser for

educational systems which proposes finer analysis based on three criteria : shape, direction and order together with their combination of them. The assessor system is mainly based on beta-elliptic model and the perceptual presentation. The structure of this paper is as follows: Section 2 presents an overview of systems for handwriting analysis in literature. Section 3 illustrates the architecture of mobile application for online handwritten script analysis. Section 4 describes the different on-line scripts treated on our digital workbook. Section 5 reports the rate recognition performed on online handwriting and draws concluding remarks of future works.

2 RELATED WORK

Handwriting quality is related to legibility and kinematics [8]. Three main skills are required in the writing process: a complete visual representation of each acquired script, a recognition of elementary strokes in an online trace, and the ability to reproduce a script as a sequence of elementary strokes respecting direction [9]. In the literature, detection of letter or a word as a part of a sentence or a text has been extensively studied in the field of handwriting recognition [1], [2], [3], [4]). The challenge of evaluation is more complex than recognition because as you know we have not only evaluated the shape but also direction and order. Children frequently make several kinds of errors: write correct shape in wrong direction or unversed, add addition strokes, letter or word transformation... So, we have to be able to detect the kind of errors related to shape, order or direction. Handwriting recognition is the task of transforming language from graphical marks into symbolic representation. Humans have skills to read or write. To evaluate handwriting forms, we proceed by the detection of common properties, then we

- H.Akouaydi, Y.Hamdi, H.Boubaker and Adel M.Alimi are with REGIM-Lab.: REsearch Groups in Intelligent Machines, University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, Sfax 3038, Tunisia. E-mail:(hanen.akouaydi.tn, yahia.hamdi.tn, adel.alimi)@ieee.org, boubaker.houcine@yahoo.fr
F.A.Cheikh is with The Norwegian Colour and Visual Computing Laboratory, Department of Computer Science, Norwegian University of Science and Technology, POBox 191, 2802 Gjøvik, Norway, E-mail:faouzi.cheikh@ntnu.no

Manuscript received February 14, 2021; revised 2021.

gather them according to some perceptual laws like shape, direction, and order... The goal of this part is to mention the theoretical approaches used for handwriting evaluation. We mark the existence of different methods treating the online handwriting evaluation in the literature.

Simonnet et al. [5] proposed a multi-criteria approach for Latin handwriting quality analysis. In this work, the handwriting children are evaluated with regards to reference models using intra-class and inter-class scores. Indeed, a multi-criteria score describes the legibility (shape) and kinematic (order and direction) aspects for children, according to the teacher expectations. Likewise, in 2018, the authors [6] introduced an explicit segmentation approach for handwritten cursive word evaluation. First, the authors start by extracting the primary segmentation hypotheses to reduce error propagation by adding a verification step through supervision. Next, they extract the letter hypotheses based on scoring and finally word hypothesis extraction and evaluation by combining elastic matching and writing analysis scores. Accadro et al. [7] analyze handwriting kinematics of kids on tablets by dealing with different types of features: number of strokes per letters, pen lift duration... and shows that the handwriting is based on four kinematics characteristics: velocity, spatial arrangement and motor planning plan. Only one work [8] deals with Chinese characters that identify 3 types of errors related to sequence stroke, stroke relationship and stroke production. Falk et al. [9] uses five primitives to evaluate handwriting proficiency of children: legibility, form, alignment, size and space.

Generally medical applications mainly deal with the kinematics features of handwriting in contrast with the educational systems that focus to legibility criteria and use only the features of velocity and acceleration.

Jolly et al. [10] presented a comparative study about the acquisition of the training between digital devices and paper. Results show a significant improvement of children trained on paper in terms of fluency. The study tackles the problem of handwriting quality analysis. More specially, this system tries to take advantages of the digital devices in primary schools and preschools. It aims to develop digital workbooks to help teachers and children from three to seven.

Factors do not generalize very well to a one so it makes the task of evaluation difficult. Even though Arabic is the fourth most spoken language in the world after Chinese, English and Spanish, to the best of our knowledge, there is any work in the field of Arabic text evaluation, which makes this area of research ripe for exploration. The main objective of this work is to offer an advanced digital writing experience at schools by using tablets and tactile digital devices (with finger touch and stylus). The validation of this research project is based on experiments performed in primary schools and preschools of Tunisian children.

The paper represents results of the first experimentation in 20 primary schools with 900 children. During three-hour sessions, groups of six to eight kids as illustrated in Fig.1, were writing letters, symbols, words and having automatic feedback.

- It allows children to work in autonomy with online and relative feedback;

Fig. 1. (a) Beta-elliptic Evaluation project with tablet tactile devices about the analysis handwriting (b) First in-class experiment

- It proposes automatically pedagogical exercises that are adapted to each child ability to write correctly or his difficulties based on its automatic writing analysis;
- It provides precise analysis of kids writing (i.e shape, order, direction) to aid teachers on the automatic examination of kids writing skills and difficulties.
- The constraints imposed to the writer before the process of writing;
- Our system allows teachers write their own models: letters or words samples and secondly by consolidating with children samples collected during the remediation to adjust the model to teachers exceptions.
- Update models will allows various configurations of correct orders and directions.
- The work demonstrates that beta-elliptic is not only a model for segmentation and recognition but also a tool to evaluate handwriting.

3 LANGUAGES SPECIFICITIES

All works focus on Latin scripts and we found some approaches analyzing the .Chinese Language. For the Arabic Language, we have not found any work dealing with this cursive script. Many different languages have a script based upon to the Arabic alphabet, such as Pashto, Urdu, Persian, Jawi, Uyghur, Kurdish. The ability to recognize Arabic handwriting is crucial and has many applications which would be beneficial to millions of people. The Arabic script is one. with a number of challenges that add to the complexity of handwriting recognition:

- **Cursive Nature:** The Arabic script comprises of 28 characters. Arabic is always written cursively, regardless of whether it is handwritten or typed.
- **Changing letter shapes:** Each of the 28 characters can change shape according to its position in the word. This is a direct result of the cursive nature of the script and adds to the complexity of the script. Other cursive scripts often have a fixed letter shapes with small variations. However, Arabic letters may change drastically depending on their location in the word.
- **Joining and non-joining characters:** Of the 28 standard characters, 6 are non-joining which means that although the script is cursive, they do not join to the following letter. This will result in a small space within a word without the word actually ending.
- **Ligatures:** we find a number of combinations of letters that may form ligatures. One ligature that always occurs, as opposed to being part of different

Fig. 2. Flowchart of the proposed system.

handwriting styles. As a sample, J is followed by I , instead of becoming U , it becomes V .

The Arabic script is used by over 400 million people worldwide and with a history spanning over 1800 years. It also remains one of the most important languages in the world. Latin script is however an easier script. It serves the largest family of a total of 59 languages [1]. It is composed of 26 characters. Our work is based on the perceptual Theory with the fact that all we perceive is a form including handwriting scripts. In fact, handwriting is a visual scene perceived in order to decode the contained message and as any visual scene it presents illusions. The existence of perceptual illusions in handwriting makes its analysis, identification and evaluation difficult. In our work, we analyze both Latin and Arabic script: cursive and printed.

4 PROPOSED SYSTEM

Our system consists of evaluating children's writing based on the criteria of shape, order and direction, which give an online feedback. Our system, as a mobile application, is the combination of visual regeneration by beta-elliptic model for characterized each stroke of online trace and a beta-classifier for evaluation. The application offers two interactive interfaces respectively dedicated to learners, educators, experts or teachers and allows them to adapt it easily to the specificity of their disciples. The main originality of the developed algorithm for childrens handwriting quality assessing, consists of the adoption of multiple analysis criteria and the application of Beta-elliptic model to extract complementary and efficient feature vectors used for handwriting evaluation (see Fig.2). The aim of our work is to help teachers with handwriting evaluation and return corrective feedback to childrens schools during the learning process. Indeed, the handwritten sequences are online signals captured with digital devices which have undergone pre-treatment step. Next, five main criteria are utilized to evaluate the handwriting quality based on Beta-elliptic model. Finally, a meaningful confidence score about

Fig. 3. The online handwriting analysis application that enables children to be autonomous during the learning and analysis tasks of letters, words, and symbols.

each criterion is established for test samples depending on a principle method: comparison engine method based on similarity detection (SD) and measuring distance (MD) technique. We present them in further details below.

4.1 Criteria of Evaluation

In this section, we explain the list of criteria used for handwriting analysis.

4.1.1 Shape

Handwriting shape represents the degree of writing produced by taking account of the size and alignment on a line. It is the principal component of the handwriting quality evaluation process. The proposed shape characteristic corresponds to some geometric parameters such as the coordinate of starting and ending points of the trace. The graphical shape means the overall geometry of the trace of its curvature function. The latter is considered one of the properties by comparing it to a reference model. Fig.4.

TABLE 1
Basic Perceptual Codes

Number	Basic Perceptual Codes	Abbreviation	Shape
1	Valley	V	—
2	Left-oblique-shaft	L-O-S	/
3	Shaft	S	
4	Right-oblique-shaft	R-O-S	\

illustrates a bad shape of Arabic Word Koraton because of missing dots.

Fig. 4. Incorrect Shape of Arabic Word "Koraton".

4.1.2 Strokes order

The writing order of the sequence (letter, word, symbol) corresponds to the correct succession of the composed elementary strokes. Each script has its own order distinguished from others. For example, the Arabic language has several characters that contain dots (i.e: ش, ب) or small marks (ط, ع) and other without dots (ح, ع). We always start with the main corps before the diacritics.

4.1.3 Strokes direction

Direction corresponds to the correct succession of basic perceptual codes with same numbers as the models. To identify direction of each stroke, we use the inclination angle of the trajectory tangent at the point $M_i(x_i, y_i)$ with respect to the direction of the ellipse small axis. We go from 4 basic perceptual codes to 8 and we pay attention to direction of stroke. Fig.5.(b) represents the 8 basic perceptual codes. Moreover, the correct direction is identified using specifics features such as start point, endpoint.

4.1.4 Kinematics of the trace

The handwriting movement is created by hand. To extract the kinematic of the handwriting we concentrate most on its velocity profiles. In fact, we believe that with this crucial property we can evaluate the dynamism of children's writing. We notice a variation between fast and slow writing, which causes some diversity in the inflexion points of the trajectory.

4.1.5 Respect of the reference lines

Baseline is the alignment of the text lines and words with respect to the horizontal direction. It reveals a lot of accurate information about the writer. Baseline or reference line is necessary to identify the skill levels of children. It is an

import element use in handwriting analysis. The baseline detection is used in various pre-treatment stages such as skew correction, text normalization, character segmentation, feature extraction, etc. Based on this criterion, the online signal can be decomposed in three main zones: Median, Upper, and Lower (see Fig.6) which allows the child to distinguish between the different forms of letters: those written on, below and above the baseline.

Fig. 6. Respect of the reference lines.

4.2 Selection of the pertinent parameters

The evaluation of the quality of kids' writing tests depends on the five criteria presented above. The proportional progression of scores for all criteria during the learning process is not guaranteed since kids often tends to focus on one or two criteria at the same time. The diversification of the evaluation criteria needs diversification of the parameters involved in the step of evaluation in order to refine their estimation and gives an online feedback to the user and the teacher. The selection of the pertinent parameters to characterize each criterion is a very crucial step in the learning process. Indeed, the geometric parameters of the Beta-elliptic model are used to apprehend the graphical features of allographs or basic perceptual forms (between two successive extremums of velocity or curvature radius) composing the characters or words. Similarly, to assess to the criteria direction and order, we need to compare the succession of geometric and dynamic feature vectors of the strokes or segments making up the scripts (character or word). We adopted the beta-elliptic parameters which model the curvature function of the trace. Likewise, to evaluate the importance of the kinematics criterion in the handwriting learning process, we compare the velocity profile of the test script (character or word) executed by the child to the trace models recorded by an expert (voluntary teachers). To do this, we have decided to use the dynamic part of the beta-elliptic model to characterize this criterion that evaluates quantitatively the correctness of a criterion characteristics of the trajectory A stroke which is limited

Fig. 5. (a):Basic Perceptual Codes;(b):Basic Perceptual Codes respecting direction.

by extremum velocity. A stroke is characterized by 10 parameters. The following beta-elliptic parameters a , b , x_0, y_0 and θ reflect the geometric properties of the set of muscles and joints used in a particular handwriting movement and describe the static aspect of handwriting. With θ , each beta-stroke can be classified into these four basic codes: $-$, $/$, $|$, $\backslash$. Those four perceptual codes are expanded to eight in order to identify the direction of elementary stroke.

With the 10 beta-parameters and its corresponding perceptual code: we construct parametric-perceptual or dynamic-visual features of each stroke. Our system compares the traces produced by children with three correct traces named 'models', existing in the training database by using their features vector.

5 FEATURES EXTRACTION

Handwriting is a visual scene perceived in order to decode the contained message. The existence of perceptual illusions in handwriting makes its analysis difficult. As mentioned, the goal of this section is to illustrate extracted features used for handwriting evaluation, which is based on Beta-elliptic perceptual representation. The main hypothesis of the present system is that handwriting is a group concatenation of basic codes respected directions concatenated together in order to get a form, a character, a digit, a symbol or a word.

5.1 Preprocessing

Preprocessing is a crucial step for handwriting analysis in order to obtain better evaluation of children's trace. The two principal of preprocessing steps are used to eliminate trembles in writing, to reduce noise and to remove the hardware imperfections. The preprocessing operations used in our system are:

- **Removing noise:** we apply a Chebyshev low pass filter to eliminate the noise generated by spatial and temporal sampling;
- **De-hooking:** means to eliminate trembles in writing because of inaccuracies in rapid pen-down/up detection and tremble in writing;

In this analysis system we must avoid the **Interpolation** which adds missing points caused by the variation of writing velocity, because this step can affect the direction information and gives bad representation of child handwriting. To have a good representation, all acquired traces should have the same scale, as we apply **Normalization** on the acquired trace (x, y) coordinates. The algorithm of normalization will be described below:

Algorithm 1 Normalization

```

 $max_x = \max(data_x)$ 
 $max_y = \max(data_y)$ 
 $min_x = \min(data_x)$ 
 $min_y = \min(data_y)$ 
 $m_y = max_y - min_y$ ; Normalisation hauteur
 $m_x = max_x - min_x$ ; Normalisation largeur
if ( $m_x > 0$ ) then
  if ( $(m_y/m_x) > 1.6$ ) then
     $rapport\_norm_l = rapport\_norm_h * (m_x/m_y)$ 
  else if ( $(m_y/m_x) < 0.625$ ) then
     $rapport\_norm_l = rapport\_norm_h * (m_x/m_y)$ 
  else
     $rapport\_norm_l = rapport\_norm_h * (m_x/m_y)$ 
else
   $rapport\_norm_l = rapport\_norm_h * (m_x/m_y)$ 

```


5.2 Beta-elliptic Features

A movement is the sum of impulse signals as the response to the neuromuscular system which is [11] the Beta function [12]. The Beta-elliptic model consists of decomposing the online trajectory into elementary strokes based on the combination of the dynamic (beta ellipse) and geometric aspects (arc of ellipse) of online handwriting modeling. Furthermore, the model has been proven to be highly successful in various domains such as handwriting recognition [2], [3], [4], handwriting regeneration [13], [14], [15], [16], writer identification [17]. To our knowledge, the Beta-elliptic model has not yet used in handwriting evaluation. It proposes that handwriting can be reconstructed by the sum of impulse signals by the curvilinear velocity and can be also approximated by a sequence of elliptic shapes. [13], [14], [15] Thus, handwriting can be segmented into simple movements called strokes, which are the result of superposition of time overlapped velocity profiles. Therefore, each stroke with its curvilinear velocity obeys the Beta-elliptic approach and online script can be segmented into different ones by tracing Beta-elliptic parameters. [18], [19], [20], [21] The Beta function is presented in this equation(1)

$$\beta(t, p, q, t_0, t_1) = \left(\frac{t - t_0}{t_c - t_1}\right)^p \left(\frac{t_1 - t}{t_1 - t_c}\right)^q, IF t \in [t_0, t_1] \quad (1)$$

$$= 0, IF not$$

where t_0 : starting time of Beta function.

t_c : is the instant when the curvilinear velocity reaches the amplitude of the inflexion point.

t_1 : the ending time of Beta function.

With $(p, q, t_0 < t_1) \in R$, and

$$t_c = \frac{p * t_1 + q * t_0}{p + q}, p = q * \frac{t - t_0}{t_c - t_1}$$

p, q intermediate parameters which have an influence on the symmetry and the width of Beta shape.

As a result, a stroke is characterized by 10 parameters. The following beta-elliptic parameters a, b, x_0, y_0 and θ reflect the geometric properties of the set of muscles and joints used in a particular handwriting movement and describe the static aspect of handwriting. These parameters are presented as follows:

- x_0, x_1 : the coordinates of the ellipse center,
- a : big axe of the ellipse
- b : small axe of the ellipse
- θ : the angle of the deviation of the elliptic arc and the horizontal which is obtained by the following equation 2:

$$\theta = \arctan\left(\frac{Y_1 - Y_0}{X_1 - X_0}\right) \quad (2)$$

An online handwriting script can be segmented into segments and each segment is a group of strokes which could be classified into basic shapes. In fact, the number of strokes is identified automatically from the curvilinear velocity representation. In the trajectory domain, each stroke is limited by both M1 and M3 that correspond to the maximum curvature and minimum curvilinear velocity. M2 is defined by the maximum tangential velocity. Beta-elliptic arcs are presented in red in Fig.8.

Fig. 7. Segmentation and Detection of Beta-points of Arabic letter 'a'.

5.3 Basic Perceptual Codes (BPC)

Handwriting is a message which can be decoded into basic codes. The main assumption of our proposed model, consists in the fact that handwriting is a concatenation of visual codes grouped together so as to get a shape. The principal basic codes to write are: $-$, $/$, $\backslash$ [22], [23], which can compose any script illustrated in Table 1. The analysis of any language has demonstrated that a script is an arrangement of vertical and diagonal lines as shown in Fig.8. It demonstrates that a letter is a combination of two or more basic codes [2], [15], [20], [24].

Fig. 8. (a):Scripts with Basic Perceptual Codes.

Hence, in this paper we use 4 basic perceptual codes which can form any script: 4 basis codes, for which we must pay attention to their directions. In Fig.8, we notice that handwriting is an arrangement of perceptual codes in a special order, and we can form different scripts with the same group of basic code such as the two Arabic letters ع hamza ا . In this step, we extract the basic shape gathered together to identify a script. Hence, we use the beta-elliptic model as a model of segmentation. Basic codes are not directly detectable. Therefore, we use in the step of segmentation the beta-elliptic model.

The main idea of this part is to associate an BPC (Basic Perceptual Code) to each stroke. It means that we classify strokes of the handwritten script generated by the beta-elliptic model into BPCs. This classification is performed by using the parameter, corresponding to the deviation angle of each ellipse " θ ", by the horizontal axis. An BPC is limited by beta control points M1, M3 and goes through H which is the orthogonal projection of M2. The number of strokes is identified automatically by beta-elliptic model which is equal to the number of BPCs. The BPCs are similar

to Freeman's code, which are frequently used for online handwriting recognition, being able to express stroke directions. Based on this technique, features are extracted from each online trajectory segment. By means of analogy with Freeman's codes, an BPC is a whole region incorporating a group of directions as depicted in Fig.5.(b) showing the 8 directions of BPCs in their corresponding parts. In this work, we use only 4 BPCs and we take attention to direction: they become 8 BPCs. Initially, we choose to extract the starting points, the ending points and the inflection points. As an Arabic handwritten example, the letter *o* is composed of a set of beta points is shown in Fig.7. These characteristics formed the beta points which are grouped together to form a sequence of basic shapes. The more the handwriting dynamics increase, the more the number of basic shapes increases. The handwritten script is segmented into strokes limited by beta-points. A segment is a group of strokes, and stroke is composed of n BPCs. Each stroke is classified n BPCs, and a segment is a group of n strokes. Handwriting has an equation as our perceptual theory which is:

$$handwriting = \{Segment_1\} \quad (3)$$

handwriting is composing of one **segment**.

$$segment = \{BPC_{ij}, BPC_{ij}, \dots, BPC_{ij}\} \quad (4)$$

$$BPC = \{Stroke_{1j}, Stroke_{2j}, \dots, Stroke_{pj}\} \quad (5)$$

p : Number of **strokes** identified in the trace.

j : 1,2,3,4,5,6,7,8 correspond to **BPCs**. (see Table 2)

With an adequate arrangement of basic perceptual codes (BPCs),

$$Stroke/BPC = beta-point1 + beta-point2 + beta-point3 \quad (6)$$

The beta-points are three anchorage perceptual points [15], [20], [24], forming a stroke. The stroke goes from M1 to M2, which are the maximum curvilinear velocity and the minimum curvilinear velocity, respectively. It passes through H which is the orthogonal projection of M2, standing for the maximum tangential velocity or the inflexion point. As already mentioned, handwriting is a sequence of basic perceptual codes and the models can be different of combination Basic shapes. Numerous possible combinations of BPCs can form the sample script. In fact, there is a large number of combinations that form a similar letter or symbol.

6 COMPARISON ENGINE

Thanks to the beta-elliptic model for handwriting BPCs generation, we obtain an additional vector related to the dynamics domain. As already observed, we can generate as many segments as we gratitude to not only the beta-elliptic segmentation method but also the enhanced control points extracted. Our Beta-elliptic model is a best model for handwriting trace generation. It is extremely beneficial in terms of the generation of realistic stroke animations of motor actions and the study of dynamic criteria such as direction and order for handwriting learning tools. The problem we tackle here is to quantitatively evaluate a cursive trace with respect to a reference model [25]. In order to be able to teach children how to write, we must be able to analyse their

handwriting, to evaluate if the traces are correctly written, and to detail what aspects of the child scripts do not correspond to the teacher models. This problem is completely different from the classical task of handwriting recognition, where the challenge is to determine to which class data samples belongs. In our case, we already know data labels, because children were asked to draw a specific script, but we want to evaluate how close is this drawn trace to the reference model, and for which aspects it does not. Our objective is to be able to analyse and evaluate handwritten symbols, with regards to reference models, and for multiple aspects. A correctly handwritten script is characterised by several aspects: first its shape, but also its drawing direction and order, its speed and its fluidity for instance. For each these aspects of the analysis, we use a specific beta-feature set, specially designed to capture the desired aspect. In this paper, we present different beta-feature defining a priori to analyze three aspects: the shape, the order and the direction. It allows to easily define reference models from few data samples to customize the writing exercises to the children. Then, the analysis system can be used to evaluate drawn traces, regarding a specific feature set, and finally give a confidence score, regarding the specific aspect of the feature set. Our objective is to be able to analyze and evaluate handwritten symbols, with regards to reference models, and for multiple aspects. A correctly handwritten script is characterised by several aspects: first its shape, but also its drawing direction and order, its speed and its fluidity for instance. For each these aspects of the analysis, we use a specific beta-feature set, specially designed to capture the desired aspect. As mentioned previously, the Beta-elliptic model describes a set of dynamic and geometric features and perceptual representation designed for online handwriting trajectory model by breaking it into elementary stroke. Each stroke is classified into BPC and characterized by 10 beta-parameters. Based on this model, the comparison of two handwritten samples (test and model) is done by comparing their respective stroke by a measure of similarity detection and dissimilarity distance. This method is based on an algorithm that makes it possible to detect for each beta stroke trajectory of the test/ sample, the most similar stroke in the trajectory of the model sample using similarity technique. In addition, it calculates the dissimilarity between the two samples which would be the sum of the dissimilarity distance of all the beta stroke composing the test sample in comparison with those of the model sample. In the analysis, the trace is already known but the aim is to evaluate its quality by comparing it to against a reference model under various criteria that can be interdependent (e.g. order depends on shape). Two complementary measurements are related to this process: Similarity or an intra-class measure describing the closeness to the model of the expected letter, and Dissimilarity or an inter-class measure which quantifies its discriminatory quality with respect to other symbols. The analysis system is based on three criteria: shape, order and direction. Each criterion can be used by teachers during the remediation to identify errors and children difficulties. A refined multi-criteria confidence gives to children a score related to legibility and ductus aspects. More precisely, the elementary-analyser presented in Figure 7 consists in extracting features that are used by

that limit the substantial part of the distribution DD fixed empirically to 96% and 4% respectively. A first normalized score NS_1 is assigned to the test sample relying on the correct partition using Eq. 10.

$$\begin{cases} NS_1 = 1, & \text{if } DD_{T_M} < T_{CC} \\ NS_1 = 0, & \text{else if } DD_{T_M} > T_{CC} \\ NS_1 = \frac{T_{CW} - DD_{T_M}}{T_{CW} - T_{CC}} \end{cases} \quad (10)$$

Similarly, we calculated the DD_{T_W} distance separating the test sample from the set of wrong samples which represent the most common errors before converting them to a normalized score noted NS_2 between 0 and 1 analogously to Eq.10. As described in Eq. 11, the final score NS is computed as the average between NS_1 and the complement of NS_2 .

$$NS = \frac{NS_1 + (1 - NS_2)}{2} \quad (11)$$

7.3 Generation of a Ground-Truth with an Expert

In evaluation, it is often difficult to estimate the correctness of a sample as children are learning to write. Therefore an expert has manually created the ground truth with the following rules. First, a sample is correct with respect to shape if it is similar to the expected letter. It cannot be demented with another character and if there are no strokes altering the character aspect. Besides, as it exits different levels of deformation for the shape, the expert can qualify positive/negative samples with a strong or weak confidence. So, a sample is correct relatively to order and direction respectively if the order and direction of elementary strokes are correct.

Fig. 9. Distance Normalization.

7.4 Results

Our system considers three levels of evaluation

7.4.1 Type of Evaluation

7.4.1.1 Global Evaluation:

A global description level which allows us to classify the test script into six classes (Correct, wrong Shape, Wrong order, Wrong direction, Irregular kinematic, Reference line surpass). With the value of distance between the test trace and models and the value certitude distance, the test script is evaluated correct or wrong. Table 3 sums up the global

results achieved by our analyser system using the different feature extraction models. We can see from this table that our method is very effective for handwriting childrens evaluation. In fact, it gives good results for characterizing the three criteria. The specific temporal information, geometric parameters and perceptual representation provided by BEM allow us to analyze the three criteria (shape, direction, order). In fact, the final score assigned for each criterion is calculated as a weighted average of the scores assigned by each separate subsystem (dynamic system and perceptual system) weighted by its correct classification rate (CCR).

7.4.1.2 Quantitative Evaluation:

The second type of evaluation provides a qualitative description of the test script by attributing for each criterion a quantitative evaluation label. It assigns a confidence score ranging from [0,100] for each criterion as illustrated in Fig.12 (red circle). The evaluation of our framework is enhanced by a database collected in Tunisia primary schools with 400 children (letters, words and symbols). Experimental results show the efficiency and robustness of our suggested framework that helps teachers and children by offering positive feedback throughout the handwriting learning process using tactile digital devices. The obtained quantitative results of some observations are illustrated in Table 4. It shows that the best confidence score of shape criterion is achieved by using parametric features rather than perceptual features. This can be explained by the use of visual and curvature generic features generated by Beta-elliptic model. Moreover, the perceptual approach is more accurate for the analysis of order and direction criteria than parametric method. This robustness is due to the precise features to the change of direction and order of basic perceptual codes.

7.4.1.3 Qualitative Evaluation:

Fig. 10. Qualitative & Quantitative Evaluation of Arabic Letter Haa هـ .

The third evaluation method provides a qualitative description of the test script by attributing for each criterion a qualitative evaluation label: (very well (VW), W (well), medium (M), bad (B), very bad (VB)) attributing according to the final value of score provided at the input of the decision fusion stage (see Fig.7), on the intersected linguistic fuzzy subsets among the five labels (see Fig.9). In blue rectangle with emoticon in Fig.12 qualitative evaluations are illustrated.

7.4.2 Criteria of Evaluation

Before giving quantitative results about handwriting analysis, it is important to analyze the feedback of children about

TABLE 3
Overall CCR [%] using correct/incorrect samples of shape, direction and order criteria.

Models	Shape	Direction	Order	Kinematic	Position
Arabic Characters	96,56	98,16	98,75	98,75	99,00
Arabic Words	94,13	97,10	97,62	97,20	95,77
Latin Characters	96,56	98,16	98,75	98,90	96,51
Latin Words	94,13	97,10	97,62	97,75	98,50
Symbols	97,13	98,13	100	98,60	98,62

TABLE 4
Table 4. Quantitative results of some Arabic, scripts and symbols using shape, order and direction criteria.

Criteria	Shape		Direction		Order	
	Parametric	Perceptual	Parametric	Perceptual	Parametric	Perceptual
Arabic Letters						
ا	0,95	0,97	0,99	0,95	1	0,94
ب	0,97	0,97	0,96	0,94	0,91	0,94
ج	0,98	0,97	0,99	0,95	1	0,94
ع	0,93	0,97	1	0,95	0,97	0,94
س	0,94	0,97	0,99	0,95	0,92	0,94
Symbols						
Symbol -	0,98	0,97	0,99	0,97	1	1
Symbol /	0,95	0,97	0,99	0,95	1	1
Symbol ر	0,96	0,97	0,99	0,95	1	1
Symbol	0,98	0,97	0,99	0,97	1	1
Symbol \	0,95	0,97	0,99	0,95	1	1
Symbol ل	0,96	0,97	0,99	0,95	1	1
Symbol C	0,99	0,99	0,99	0,97	1	1
Latin Characters						
A	0,98	0,97	0,99	0,97	1	1
B	0,95	0,97	0,99	0,95	1	1
C	0,96	0,97	0,99	0,95	1	1
U	0,99	0,99	0,99	0,97	1	1
V	0,98	0,97	0,99	0,97	1	1
T	0,95	0,97	0,99	0,95	1	1
P	0,96	0,97	0,99	0,95	1	1
O	0,99	0,99	0,99	0,97	1	1
Arabic Words						
كرة	0,95	0,97	0,99	0,95	1	0,94
جمل	0,97	0,97	0,96	0,94	0,91	0,94
ثلج	0,98	0,97	0,99	0,95	1	0,94

this first experimentation with tactile 385 digital devices. In-class experiments conducted with the multi-criteria classifier have shown that children get quickly familiar with the application and try to improve the evaluation score that is well understood. Moreover, the personalised feedback allows children to progress with autonomy at their own speed, and to stay concentrated during the whole 20-minute session which is difficult for 390 young children. Some of them were so involved that they did not want to stop. Quantitative and qualitative results of all criteria are summed up in Table 5 and Fig.16. In all qualitative figures, evaluation results are displayed with an encompassing rectangle: green (true positive) and red (false positive). The rest of this section presents results per criterion as follows,

7.4.2.1 Shape:

Quantitative results in Table 4 show a significant results of shape criterion with the dynamic approach. The proposed shape features consist in selecting parametric beta-elliptic features that do not depend on order and direction. More precisely, qualitative results in Fig.13 shows a coherent

decrease of the score independently of the number of strokes and direction/order correctness. Nevertheless, some errors remain as they are as presented in Fig.13(c) and Fig.13(d). They are shapes with missing/additional strokes and contain strokes that do not respect relative stroke proportions.

Fig. 11. Results of the shape criteria.

7.4.2.2 Order:

The correct order of a trace corresponds to the correct order of the associated models. The identification of the correct order is based on our four basic perceptual strokes which are part of an elementary stroke that is invariant to direction changes. Thus, quantitative results in Table 4 show a significant results of direction criterion with perceptual approach. This is mainly due to the presented approach to identify perceptual elementary codes as depicted in Fig.14 which is more robust than a dynamic approach. Remaining errors cases in Fig.14(d) and Fig.14(e) correspond to characters with deformations at the limit of the shape correctness.

Fig. 12. Results of the order criteria.

7.4.2.3 Direction:

Quantitative results in Table 4 show a significant improvement of perceptual method comparing with dynamic method corresponding to a decreasing of respectively 70% and 99% . It is explained by the use of four basic perceptual codes which go from four to eight codes that is a finer representation of directions. It works with various numbers of strokes and independently of the order as presented in Fig.15 .Similarly to order, errors cases correspond to deformations at the limit of the shape correctness as depicted in Fig.15.

Fig. 13. Results of the direction criteria.

7.4.2.4 Multi-Criteria:

As presented in Fig. 16, qualitative assessments of the used criteria are presented respectively for some samples of the Arabic letter Tad . As shown in Fig.16, the test samples and their qualitative scores with five levels are considered VW (green), W (dark green), M (yellow), B (orange), and VB (red) corresponding to a uniform partition of confidence ranges.

Fig. 14. Qualitative Results of Arabic Letter "Tad" ط

The multi-criteria evaluation corresponds to the feedback given to the children. Therefore, this score depends on two main aspects: the legibility and the ductus on which the importance of each aspect is fixed by teachers based on the children's skills and the aim of the handwriting exercise. Fig.16 illustrates three possible strategies for the teacher to customize the feedback returned to the children. In Fig.16.(a), it corresponds to an early learning stage, shape, order and direction must correspond to the model. Therefore, errors in order Fig16.(d), direction Fig16.(b) and shape Fig16.(c) are penalized. Opposite to that, the strategy in Fig16.(c) corresponds to an advance learning stage, where children acquired their own writing style. The legibility is therefore the only evaluation criterion. Finally, in Fig.16.(b) and Fig.16.(d) respectively, there is no constraint of order and direction. These strategies are intermediate learning stages where a constraint about order or direction can be relaxed. the multi-criteria approach has been used with the shape and direction criteria as order is not meaningful for cursive writing.

8 DETECTION & IDENTIFICATION OF CHILDREN'S ERRORS

Beta-Evaluation application is able to detect the stroke(s) which contain(s) errors as shown in Fig.19, in it two strokes contain errors of the shape of the Arabic letter س . The existence of perceptual illusions in handwriting makes its analysis, and evaluation difficult. The challenge of evaluation is to identify the kind of errors. The children frequently make several kinds of errors: write correct shape in wrong direction or unversed, add addition strokes, letter or word transformation...Fig.18, Fig.19 and Fig.20 present some examples of handwriting errors and mistakes made by children. Those errors can be performed by our analysis language system. In Figure 18 : This handwriting shape is correct but the direction is wrong. We enrich our workbook with symbols in terms of the simple view of writing and instructional strategies to increase spelling writing quality in children with dyslexia. When learning to read, children first have to link the shape of the word on the page with the sound it makes. Then, when it comes to writing, they have to recreate that shape back onto paper. For children with dyslexia, decoding these patterns and making these links can often be very difficult. As a result, they frequently fail to develop the automatic flow of writing which will help them to express themselves clearly and easily in writing. So, we need to improve our workbook with simple symbols

Fig. 15. The Arabic letter Jim ج written in reverse

Fig. 17. Arabic Letter "Seen" س with added strokes.

Fig. 16. Wrong Arabic Letter "Tad" ط .

Fig. 18. Some symbols.

as illustrated in Fig.14, to short timed sessions so the kid is able to maintain concentration without becoming bored or uncomfortable. For some pupils with dyslexia, the difficulties associated with handwriting can mean that the only way that they can achieve the speed of writing needed for success in the education system is to use a tablet. If this is the case then learning to touch type will be beneficial.

Fig. 19. Examples of symbols.

9 CONCLUSION

A new system for online handwriting evaluation and analysis is presented in this paper which deals with both cursive and not cursive handwriting multi-script. It analyzes the handwriting quality of children's school and provide them real-time feedback based on five main criteria (shape, direction, order, kinematic, position). It uses the Beta-elliptic model for handwriting representation. The Beta-elliptic model offers the possibility of detecting dynamic and geometric aspects in online handwriting trajectory modeling and decoding the handwriting trace in a group of basic perceptual codes. We validate our system

with databases collected from pre-schools which contained digits, letters Arabic/Latin, symbols. Experimental results demonstrate the effectiveness of our proposed system that helps teachers and children by giving feedback throughout the handwriting learning process. The extracted features are universal and can be extracted from other scripts like Chinese, Urdu,... is interesting, so we plan it as future work. Adding deep neural networks is also a perspective of improvement of our application.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the grant agreement number LR11ES48.

REFERENCES

- [1] D. Keysers, T. Deselaers, H. A. Rowley, L. L. Wang, and V. Carbone, "Multi-language online handwriting recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1180–1194, June 2017.
- [2] H. Akouaydi, S. Abdelhedi, S. Njah, M. Zaied, and A. M. Alimi, "Decision trees based on perceptual codes for on-line arabic character recognition," no. 1-5, pp. 153–157, April 2017.
- [3] Y. Hamdi, A. Chaabouni, B. Houcine, and A. M. Alimi, "Off-line online arabic handwriting recognition using neural network," pp. 1–5, 2016.
- [4] H. Akouaydi, S. Njah, W. Ouarda, A. Samet, M. Zaied, and A. M. Alimi, "Convolutional neural networks for online arabic characters recognition with beta-elliptic knowledge domain," no. 6, 2019.
- [5] D. Simonnet, E. Anquetil, and M. Bouillon, "Multi-criteria handwriting quality analysis with online fuzzy models," *Pattern Recognition*, p. 31, 2017.
- [6] D. Simonnet, N. Girard, E. Anquetil, M. Renault, and S. Thomas, "Evaluation of children cursive handwritten words for e-education," *Pattern Recognition Letters*, p. 11, 2018.

- [7] A. P. Accardo, M. Genna, and M. Borean, "Development, maturation and learning influence on handwriting kinematics," *Human movement science*, p. 11, 2013.
- [8] C. W. Li-Tsang, A. S. Wong, H. W. Leung, J. S. Cheng, B. H. Chiu, F. Linda, and R. C. Chung, "Validation of the chinese 530 handwriting analysis system (chas) for primary school students in hong kong, research in developmental disabilities," p. 28722883, 2013.
- [9] T. H. Falk, C. Tam, H. Schellnus, and T. Chau, "On the development of a computer-based handwriting assessment tool to objectively quantify 535 handwriting proficiency in children," *Computer Methods and Programs in Biomedicine*, p. 102111, 2011.
- [10] C. Jolly, R. Palluel-Germain, and E. Gentaz, "Evaluation of a tactile training for handwriting acquisition in french kindergarten children: A pilot study, kindergartens: Teaching methods, expectations and current challenges," p. 161176, 2013.
- [11] R. Plamondon, "A kinematics theory of rapid human movements. part i: Movement representation and generation," pp. 295–307, 1995.
- [12] M. Alimi, "An evolutionary neuro-fuzzy approach to recognize online arabic handwriting," p. 382386, 199/.
- [13] H. Bezine, A. M. Alimi, and N. Derbel, "A developmental theory of the relation between geometry and kinematics in handwriting," *IEEE International Conference Systems, Man and Cybernetics*, p. 5, 2002.
- [14] M. Ltaief, H. Bezine, and A. M. Alimi, "A spiking neural network model for complex handwriting movements generation," *International Journal of Computer Science and Information Security* 14, p. 7, 2016.
- [15] S. Njah, H. Bezine, and A. M. Alimi, "A new approach for the extraction of handwriting perceptual codes using fuzzy logic," pp. 302–307, 2008.
- [16] H. Akouaydi, S. Njah, and A. M. Alimi, "Android application for handwriting segmentation using perthos theory," no. 1-5, 2016.
- [17] T. Dhieb, S. Njah, W. Ouarda, H. Boubaker, M. B. Ayed, and A. M. Alimi, "An online writer identification system based on beta-elliptic model and fuzzy elementary perceptual codes," p. 12, 2018.
- [18] M. Ltaief, S. Njah, H. Bezine, and A. M. Alimi, "Genetic algorithms for perceptual codes extraction," *International Journal of Intelligent Learning Systems and Applications, IJLSA*, pp. 256–265, 2012.
- [19] S. Njah, H. Bezine, and A. M. Alimi, "On-line arabic handwriting segmentation via perceptual codes: application to mayastroun database," pp. 1–5, March 2011.
- [20] S. Njah, M. Ltaief, H. Bezine, and A. M. Alimi, "The perthos theory for on-line handwriting segmentation," *IJCSI*, vol. 9, pp. 142–151, 2012.
- [21] S. Njah, H. Bezine, and A. M. Alimi, "A fuzzy genetic system for segmentation of on-line handwriting: Application to adab database," pp. 95–102, April 2011.
- [22] J. R. Pinales, "Reconnaissance hors-ligne de l'écriture cursive par utilisation de modèles perceptifs et neuronaux," *PhD, University of Paris*, 2002.
- [23] A. Malaviya, L. Peters, and P. Camposano, "A fuzzy online handwriting recognition system: Fohres," *Proceedings of the International Conference on fuzzy theory and technology*, pp. 1–15, 1993.
- [24] S. Njah, H. Bezine, and A. M. Alimi, "A new encoding system: Application to on-line arabic handwriting," pp. 451–456, 2010.
- [25] V. Kulesh, K. Schaffer, I. K. Sethi, and M. Schwartz, "Handwriting quality evaluation," *Proceedings of the Second International Conference on Advances in Pattern Recognition, ICAPR 01, London, UK*, pp. 157–165, 2001.

Yahia Hamdi is a PhD student at the National School of Engineering of Sfax, Tunisia. He obtained the master degree in computer science from the University of Sfax in 2012. His research focused on the area of online Handwriting recognition and handwriting trajectory modeling.

Houcine Boubaker Graduated from Electrical Engineering Department in 1995. He obtained his master degree in systems analyses and digital signal processing in 1997. He is a researcher and a PhD student in Electrical Computer Engineering Department at the University of Sfax. His research interest includes trajectory modeling and applications of intelligent methods to pattern recognition. He is an IEEE student member.

Faouzi Alaya Cheikh is currently a Professor at Norwegian University of Science and Technology (NTNU), Norway. His research interests include e-Learning, 3D imaging, image and video processing and analysis, video-surveillance, behaviour analysis, biometrics, video-guided intervention, image guided navigation and content-based retrieval. He is member of the Norwegian Colour and Visual Computing Laboratory (Colourlab).

Adel M. Alimi is currently a Professor at the National School of Engineers of Sfax. He received the Ph.D. degree and then the HDR both in Electrical and Computer Engineering in 1995 and 2000 respectively. His research interest includes applications of intelligent methods and vision systems. He was guest editor of several special issues of international journals like Soft Computing. He is IEEE Senior member since 2000 and he is the director of REGIM Laboratory.

Hanen Akouaydi is currently a PhD student in computer science at the National School of Engineers of Sfax. She obtained her engineering degree in Computer Science from Higher Institute of Multimedia Arts of Manouba in 2012. Her research interests include online handwriting recognition and deep learning. She is a member of the IEEE Student and the IEEE Computer Society and affiliates to REGIM Laboratory.