Predicting Age Groups from Touch Patterns based on Neuromotor Models

Javier Hernandez-Ortega, Aythami Morales, Julian Fierrez, Alejandro Acien

BiDA Lab, Universidad Autonoma de Madrid, Madrid, Spain
(javier.hernandezo, aythami.morales, julian.fierrez, alejandro.acien)@uam.es

Keywords: User profiling, age prediction, neuromotor model, touchscreen, children.

Abstract

This paper proposes a new method for user classification into children and adults according to their interaction with touchscreen devices. The method is based on the Sigma-Lognormal theory of rapid human movements. The experimentation is conducted on a publicly available database with samples obtained from 89 children between 3 and 6 years old and 30 adults. The neuromotor skills of the users are first described using Sigma-Lognormal parameters from a single drag and drop task. We have then used Support Vector Machines to classify the resulting features into age groups. The experiments include single-sensor and cross-sensor scenarios using data from various smartphones and tablets. The results, with correct classification rates over 96%, show the discriminative ability of the proposed neuromotor-inspired features to classify age groups according to the interaction with touch devices.

1 Introduction

Touchscreen panels have changed the way users interact with new devices. The touchscreen enables an intuitive experience of use that allows a direct interaction with what is being displayed. There are billions of touch devices all over the world and new applications and services appear every day.

Touchscreen devices provide mobile access to an unlimited number of digital contents and services (e.g. more than a half of YouTube visits come from mobile devices and this percentage is increasing [1]). Digital services are used by people from everywhere, all ages, all ethnicities and all socioeconomic status. In this context, the classification of users according to geographic and demographic attributes is crucial for service personalization (e.g. recommender systems, parental control, security) [2]. Some of these attributes can be obtained from metadata associated to the device (e.g. IP address, language selection, GPS location) or can be inferred from the user behavior (e.g. browsing history, social network contents, and keystroke dynamics) [3]. We want to highlight the spread of the use of this kind of devices by young children. The study in [4] reveals that 97% of US children under the age of four use mobile devices, regardless of family income.

In this paper we analyze a way to classify users of touch panels according to two age groups (children and adult). The age is a key attribute in user profiling with direct application on several automatic systems (e.g. parental control, recommender systems, advertising). Three examples of use cases are: i) locking content and/or applications: locking some services in tablets and smartphones when children are using them, i.e. buying new applications or sensitive content; ii) user's age study by service providers: this way service providers could develop new content that fits better to their actual audience; iii) real-time interface adapting: as children have worse control of their fine movements than adults, changing default interfaces to special tailored ones could be beneficial.

The most popular method to reveal the age of the user is based on an online questionnaire in which the user directly answers questions about his age. However, this solution assumes: i) honesty on the response of the users, and ii) users can read. Both assumptions cannot be guaranteed because of many practical reasons. Besides the fact that people lie, nowadays children start to use digital platforms and services before learning to read.

In the existing literature, there are many experiments exploring the use of technology by children, seeking how to improve the design of adapted interfaces and applications [5]. However, modeling and characterizing mathematically how children interact with touch devices and how their conduct differs from the adult's one is a field that has not been studied deeply enough. A work related to this topic is [6] where they analyzed different types of touching tasks like tap, rotate or drag and drop, and they found that children have different success rates when trying to perform different tasks. Simple tasks - for example tapping - can be done by all children without any problem, but the more complex ones are very difficult to be completed by very young children.

In [7], they measured the touch patterns of children and compared it to patterns from adults. They discovered that children have a larger miss rate than adults when trying to hit small targets. In [8] tap tasks are used to extract time and precision-based features. They designed two different approaches, using only one tap for classification and using 7 consecutive tasks. They get high accuracy rates: 86.5% in the one tap approximation, and 99% of accuracy using 7 consecutive taps to combine their scores. Even though they get good results using tap tasks, we decide to use drag and drop tasks because

Parameter	Description
D_i	Input pulse: covered distance when executed isolated.
t_{0i}	Initialization time. Displacement in the time axis.
μ_i	Logtemporal delay.
σ_i	Impulse response time of the neuromotor system.
θ_{si}	Initial angular position of the stroke.
θ_{ei}	Final angular position of the stroke.

Table 1: Sigma-Lognormal parameters description

the differences between the neuromotor development of users can be manifested in a better way. The direct comparison between approaches is not fair because we are using different tasks/information to classify users. However, in our work we demonstrate that using a very common and fast action (e.g. unlock screen based on drag and drop) we can achieve higher classification rates that those achieved in [8] for the one task approach (the second approach have not been implemented yet). In our opinion, both approaches are complementary, have very different nature, and can be combined to achieve higher performances.

The difference between adults and children is mainly caused by the different maturity of their anatomy and neuromotor system. These are less mature in children, so they have worse manual dexterity causing rougher movements [9] [10].

In order to characterize the interaction of children and adults with touchscreen devices, we propose to use a model of the human neuromotor system. The Sigma-Lognormal theory of rapid human movements represents complex movements with an analytic model that describes some physical and cognitive features of human beings [11] [12]. This model has been applied successfully to handwriting analysis and synthesis [13].

The Sigma-Lognormal model decomposes the complex signals that describe the speed of muscular movements into simpler ones that can be explained by a few parameters. These parameters contain information about the activity itself and about the neuromotor skills of the person [14]. Studies like [14] have proved that the Sigma-Lognormal model can be used to characterize children handwriting. They conclude that there are two main groups of children separable by looking at their learning stage. Children's neuromotor skills become more similar to the adults' skills when they grow up, namely, when they finish their preoperational stage. At age 10 children know how to activate each little muscle properly to produce determinate fine movements [15]. As they are based on the same neuromotor skills, the principles applied to handwriting models can be also used to model touchscreen patterns.

In this paper the application of the Sigma-Lognormal model to touchscreen interaction patterns is analyzed. The present work is focused in age classification of users into two groups: children under 6 years old and adults. We use information of simple touch tasks collected from 119 people (89 children and 30 adults) using two different types of devices: a smartphone and a tablet. Single-sensor and cross-sensor scenarios have been evaluated. The results show accuracies over 90% in several scenarios with top correct classification rate of 96% for the data obtained from tablets.

The main contributions of this work can be summarized as follow: i) a novel user age-group classification approach based on neuromotor characteristics obtained from simple and fast drag and drop tasks of approximately 2 seconds; ii) our approach obtains the highest accuracy reported in the literature for individual actions based on touch tasks; iii) to the best of our knowledge we present the first cross-sensor performance analysis for different scenarios based on information obtained from smartphones and tablets.

The rest of the paper is organized as follows: Section 2 presents the Sigma-Lognormal model and its application to the age classification problem. Section 3 presents the experiments performed and analyzes the results, and Section 4 presents the conclusions and goals achieved in this work.

2 Sigma-Lognormal Model for Touch Patterns

The Sigma-Lognormal model states ([12] [16]) that the velocity profile of human hand movements can be decomposed into strokes. Moreover, the velocity of each of these strokes, i, can be described with a speed signal $v_i(t)$ that has a lognormal shape:

$$|v_i(t)| = \frac{D_i}{\sqrt{2\pi\sigma_i(t-t_{0i})}} \exp\left(-\frac{(\ln(t-t_{0i})-\mu_i)^2}{2\sigma_i^2}\right) \quad (1)$$

where each of the parameters are described in Table 1. The complete velocity profile is modelled as a sum of the different individual stroke velocity profiles as:

$$v_r(t) = \sum_{i=1}^{N} v_i(t) \quad (2)$$

where N is the number of lognormals of the entire movement. A complex action, like a handwritten signature or touch task, is a summation of these lognormals, each one characterized by different values for the six parameters in Table 1.

In Figure 1, the speed profile of an example of touchscreen pattern (numerical) is shown. This numerical signal is input to the proposed model described in [10]. The analytical signal is calculated using the Sigma-Lognormal parameters extracted from the numerical signal.

The parameters of the Sigma-lognormal model are used to calculate the 18 different features described in [11]. These features can be classified into two groups: space-based and time-based. Space-based features are those that give information about the spatial distribution of the strokes, such as D_i, μ_i, σ_i, and other features based in θ_{si} and θ_{ei} (see Table 1). Time basedfeatures are composed by the values of speed at some relevant points of the strokes like their maximum or inflexion points; and the time-offsets between those points. The task time and the number of lognormals in each task have been added as additional features.

Regarding to the age classification task, quite often it is possible to differentiate between children and adults simply looking at the velocity profile of a touch screen task. In Figure 2, an example of these types of profiles is presented, consisting in performing a drag and drop task in both cases.

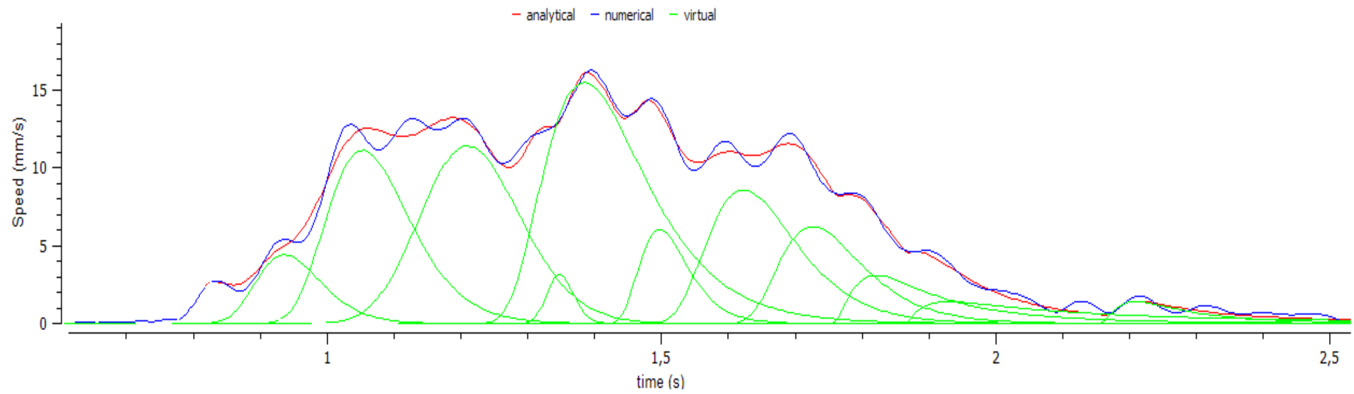
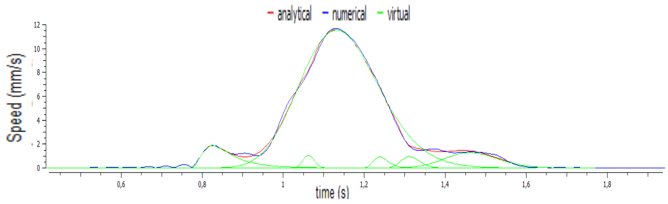
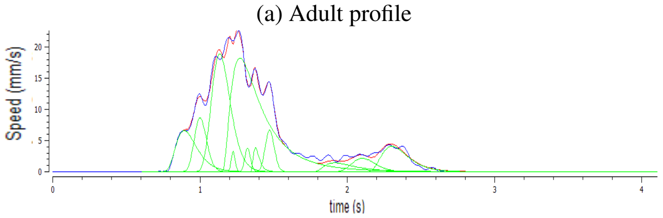


Figure 1: **Speed profile from a child's touchscreen pattern.** Numerical: is the captured velocity signal $|v(t)|$ of the touch activity (input of the model). Analytical: is the reconstructed Sigma-Lognormal velocity $v_r(t)$ profile (output of the model). Virtual: is the decomposition in individual strokes of the model.



(a) Adult profile



(b) Child profile

Figure 2: **Comparison between Sigma-Lognormal speed profiles.** The task type (drag and drop) is the same in both profiles.

A visual comparison between children and adults speed profiles have showed that children's signals are usually composed by a higher number of strokes than the adults' ones. The higher maturity on the neuromotor skills produces soft velocity profiles that reveal a fine control of the movements. Every task is composed of at least one lognormal having each one its own values for the features. In this work, the values of the features have been combined doing an arithmetic mean from the parameters obtained from each $|v_i(t)|$.

Figures 3a and 3b show the histograms of two features (Covered distance f_1, and Logtemporal delay f_2) for children and adults. These two features are highly discriminative as their histograms are clearly separated, showing differences between both classes and therefore suggesting the potential for the classification task.

3 Experiments

3.1 Database and Experimental Protocol

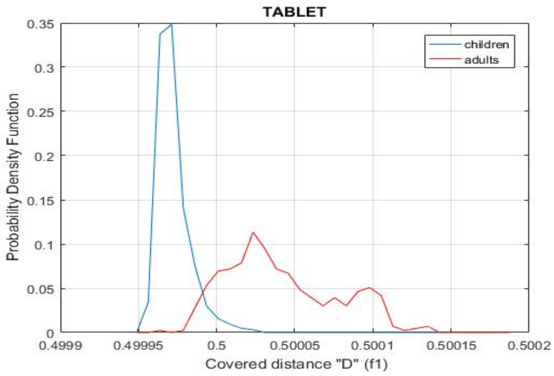
The experimental protocol is based on the publicly available database presented in [15].It is a database with touchscreen activity of both children and adults performing predesigned tasks in an ad-hoc app.

The database comprises samples from different guided activities such as tap, double tap and drag and drop tasks. In the present work, we have used the data from singletouch and multitouch drag and drop activities. Drag and drop activities consist of picking one object on the device screen and moving it to a target area. This task has been selected because it is simple and common. Multidevice information is available as the users have completed the tasks both in a smartphone and in a tablet. This allows to test our solution in both single-sensor (training with data from sensor A and testing with data from sensor A) and cross-sensor (training with data from sensor A and testing with data from sensor B).

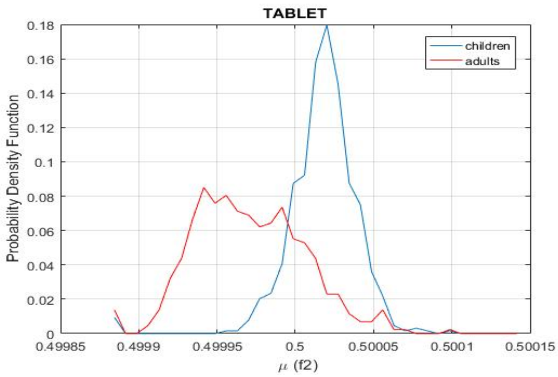
The dataset is composed by 89 children between 3 and 6 years old and 30 young adults under 25 years old. The mean age of the children is 4.6 years. The total number of drag and drop tasks is 2912 for children and 1157 for adults (see [15] for more details).

The main issue when acquiring data from children activity is to maintain the kid's attention during a long time period. The authors of the database have adapted the activities' interfaces to make the tasks more interesting to children. Thanks to this, they managed to obtain a completion rate near to 100% in tap tasks and above 90% in all types of tasks.

Drag and drop tasks fit into the Sigma-Lognormal model as they are composed by precision movements over time. The main aim of this work is to define how well this model can distinguish between children and adults. The parameters of the model (see Table 1) for all tasks were extracted using the Script Studio software [10]. This software provides $6 \cdot N$ parameters for each drag and drop task. N is the number of lognormals and it is automatically calculated by Script Studio according to the input data (coordinates x,y and their respective timestamps). In



(a) Covered distance D_i



(b) Logtemporal delay μ_i

Figure 3: **Probability Density Functions.** These are highly discriminative features as histograms are separated.

this work, we process the parameters of the model to remove the smallest lognormals. This post-processing is intended to discard all the small lognormal signals that do not have the same importance as the big ones in order to differentiate between children and adults. Any lognormal with a maximum amplitude (which occurs in t_3) under a specified threshold (in this work we use the mean value of v_3 across all the lognormals of the task as threshold) will be removed.

Is important to remark that all the small lognormal signals would be necessary to reconstruct the original stroke accurately, but not to perform a distinction between two types of users, as the smallest signals will have a negligible impact in the mean values of each task.

After reducing the number of lognormal signals, the 18 features from [8] are computed for each stroke, and each parameter is averaged across strokes. The 18 averaged parameters are augmented with the task time and the number of strokes to generate the final feature vector of size 20.

As a classifier we use a SVM (Support Vector Machine) with a RBF (Radial Basis Function) kernel because of its good general performance in binary classification tasks and the few number of parameters to configure [17]. The data have been separated randomly into training (60%) and testing (40%). The random selection was repeated 50 times and the final performance is presented in terms of averaged correct classification accuracy.

Singletouch and multitouch drag and drop tasks are available. Multitouch tasks have data from two fingers, one of each hand. Tests to determine if there is any difference between the performances of the fingers can be accomplished, since each finger could be used as one independent singletouch task to see if differences exist. Experiments have been done transforming the multitouch data into singletouch tasks and then using all the available data to train and test the SVMs.

Two different scenarios have been tested. First, a singledevice or single-sensor scenario in which training and testing is performed with data from the same type of device, namely, smartphone or tablet. Multidevice or cross-sensor is the other scenario, in which training is done with data from one type of device and testing is done with data from the remaining type.

3.2 Results

This section of the paper presents the results of the age classification into children and adults applying the neuromotor model to touch patterns.

After aplying the reduction of lognormals for both children and adults (see Section 3.1), the number of lognormals has been reduced in both classes, going from 4.36 to 1.55 lognormals per task in adults case and from 16.45 to 6.27 in children case. It can be seen that the effect has been more notable in the case of the children. This may be caused by the fact that children have to do a larger number of small corrections due to their underdeveloped neuromotor skills. However, despite of the deleting process, differences between children and adults continue being significant, so the number of lognormals remains a highly discriminative feature.

Following the experimental protocol specified in the previous section, Table 2 shows the accuracies obtained according to the different scenarios. They are presented in terms of correct classification accuracy (percentage of samples from both classes correctly classified).

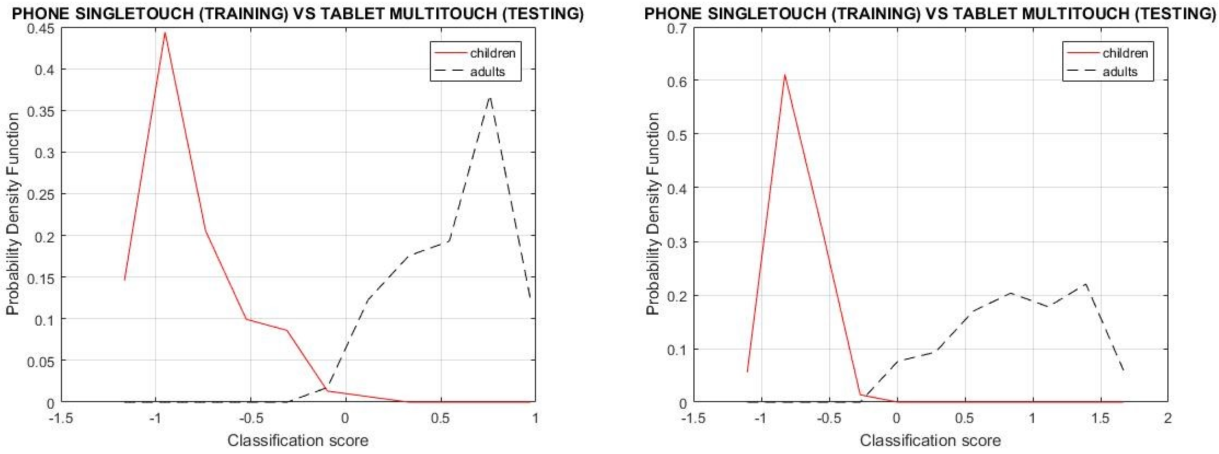
The mean value of accuracy having into account all the evaluated scenarios is 92.8%. The classification rates are over 96% in a single-sensor setting and over 95% in a cross-sensor scenario. The best results are obtained with tablets as sensors, while using smartphone's data slightly degrades the results.

Compared with [8] where they get an accuracy rate of 86.5% using one tap task for classification and with a single-sensor aproximation (using smartphone's data), our system performs better, getting a 93.6% of accuracy using only data from smartphones, and over 96% using data from tablets. These results show that our approach, based on drag and drop tasks used to model the neuromotor system of the users, is more discriminative than the scheme from [8].

Another conclusion that can be extracted of Table 2 is that the data obtained from multitouch tasks get worse results than the singletouch cases. The best multitouch scenario is obtained using tablet's data for both training and testing, with a 94.6% of accuracy, compared with its singletouch counterpart that gets a 96.3%. This may be caused by the less developed control of the left hand by right-handed people and vice versa. The main reason for using the Sigma-Lognormal model is that adults have a better control of fine movements than children, what is trans-

		Testing samples			
		Phone Singletouch	Tablet Singletouch	Phone Multitouch	Tablet Multitouch
Traning samples	Phone Singletouch	93.6%	95.0%	88.0%	92.1%
	Tablet Singletouch	93.7%	96.3%	88.9%	94.0%
	Phone Multitouch	94.1%	95.9%	88.0%	92.8%
	Tablet Multitouch	93.0%	96.3%	87.9%	94.6%

Table 2: Accuracy results for the 20 lognormal features. The accuracy is measured as the rate of correct classifications considering both classes.



(a) Single-sensor scenario

(b) Cross-sensor scenario

Figure 4: **Scores histograms using the Sigma-Lognormal model for classifying.** Left figure represents the scores for single-sensor scenario, using tablet singletouch data for both training and testing. Right figure shows the scores histogram for a cross-sensor scenario, using phone singletouch data for training and tablet multitouch data for testing the classifier.

lated to different values for the model parameters [15].

However, using the less skilled hand diminishes the differences between people of different age, as they all have the neuromotor skills of their unskillful hand underdeveloped.

The cross-sensor scenarios get results not too far from the single-sensor scenarios. The results obtained using smartphone singletouch data for training, and tablet singletouch data for testing (95.9% of accuracy) are quite similar to those obtained using only tablet singletouch data (96.3% of accuracy). This fact makes this type of systems very suitable for real applications due to its high independence of the device used.

Due to the higher number of children in the database compared to adults, selecting a percentage of the total users make the two scenarios unbalanced. Experiments balancing the number of both classes in training and testing have been made. The results show small variations around 1% of accuracy with respect to the presented results.

Figure 4 shows some histograms of the scores calculated in the classification process. It can be seen that the scores from children and adults are visibly separated into two different zones, making possible to get high accuracy rates (over 96%). There are also other zones where the scores distributions overlap. These regions are the source of incorrect classifications. Combining scores from several tasks of the same user could make possible to reduce the overlap areas, increasing even more the accuracy rate.

4 Conclusion

In this paper, we have presented a user classifier into children and adults according to their interaction with touchscreen devices like smartphones and tablets. We have proposed a feature extraction based on the parameters of the Sigma-Lognormal model. These features illustrate the neuromotor skills of users, making it capable to discern between children and adults.

An evaluation of its performance has been made, using a public database with touchscreen activity of both children and adults. The classification rates are over 96% in a single-sensor setting and over 95% in a cross-sensor scenario.

There must be said, that our approach presents important limitations related to the age range (younger/older than 6 years old). A reliable and fast classifier of users from all ages based on their interaction should be developed as a combination of different expert systems. The main drawback of other methods like using the browsing history or social network profiles, is that they need a high amount of data. Our system allows to classify users using data from simple and short (2 seconds) tasks. These features make our solution suitable for applications that require classification on the fly.

As future improvements of this system, three main aspects can be taken into account. First, the patterns used in this work are too simple: single and short strokes. Better classification rates may be achieved if the information comes from more complex tasks or from continuous monitoring. Second, cross-sensor results are quite similar to those obtained in single-sensor scenarios. Studying how the degradation changes with more heterogeneous device variation is another interesting line of future work. Finally, this study includes the analysis of touch patterns from children under 6 years. However, how to recognize users with mature neuromotor skills (from 10 years old onwards) is a challenging task and new models and methodologies should be proposed for that purpose in the future.

The classification of older users using the Sigma-Lognormal model is a possibility since it is demonstrated that the neuromotor abilities decay with the age. Note that the number of public databases is very scarce. We encourage researchers to work in this field.

Another possibility to work in, is to combine the features from [8], based on tap tasks, with our neuromotor features to improve the final accuracy rate. The features from both works have different nature, so they might be complementary. In [8] a combination of several consecutive tasks is performed to achieve a higher accuracy rate, reaching values of 99%. This kind of approximation can be done with the drag and drop tasks and the neuromotor features presented in this paper. We think that even better results can be achieved because the tasks we use are more complex and the features extracted have higher discerning capabilities.

Acknowledgements

This work was funded by the project CogniMetrics (TEC2015-70627-R) and JdC contract (JCI-2012-12357) from MINECO/FEDER.

References

[1] Youtube, “Youtube statistics.” https://about.twitter.com/company. Retrieved September 13, 2016.

[2] BBA, “Mobile phone apps become the UK’s number one way to bank.” https://www.bba.org.uk/news/press-releases/mobile-phone-apps-become-the-uks-number-one-way-to-bank. Retrieved September 13, 2016.

[3] R. Daza-Garcia, J. Hernandez-Ortega, A. Morales, J. Fierrez, M. Gonzlez Barrero, and J. Ortega-Garcia, “KBOC: Plataforma de evaluacion de tecnologias de reconocimiento biometrico basadas en dinamica de tecleo,” in *XXXI Simposium Nacional de la Union Cientifica Internacional de Radio*, 2016.

[4] H. K. Kabali, M. M. Irigoyen, R. Nunez-Davis, J. G. Budacki, S. H. Mohanty, K. P. Leister, and R. L. Bonner, “Exposure and use of mobile media devices by young children,” *Pediatrics*, vol. 136, no. 6, pp. 1044–1050, 2015.

[5] B. Cassidy and L. McKnight, “Children’s interaction with mobile touch-screen devices: Experiences and guidelines for design,” *Int. J. Mob. Hum. Comput. Interact.*, vol. 2, pp. 1–18, Apr. 2010.

[6] N. A. A. Aziz, F. Batmaz, R. Stone, and P. W. H. Chung, “Selection of touch gestures for children’s applications,” in *Science and Information Conference (SAI), 2013*, pp. 721–726, IEEE, 2013.

[7] L. Anthony, Q. Brown, J. Nias, B. Tate, and S. Mohan, “Interaction and recognition challenges in interpreting children’s touch and gesture input on mobile devices,” in *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, pp. 225–234, ACM, 2012.

[8] R.-D. Vatavu, L. Anthony, and Q. Brown, “Child or adult? inferring smartphone users age group from touch measurements alone,” in *Human-Computer Interaction*, pp. 1–9, Springer, 2015.

[9] J. Piaget and B. Inhelder, *The psychology of the child*, vol. 5001. Basic books, 1969.

[10] C. O’Reilly and R. Plamondon, “Development of a sigma–lognormal representation for on-line signatures,” *Pattern Recognition*, vol. 42, no. 12, pp. 3324–3337, 2009.

[11] A. Fischer and R. Plamondon, “A dissimilarity measure for on-line signature verification based on the sigma-lognormal model,” in *17th Biennial Conference of the International Graphonomics Society*, 2015.

[12] R. Plamondon, C. O’Reilly, C. Rémi, and T. Duval, “The lognormal handwriter: learning, performing, and declining,” *Frontiers in psychology*, vol. 4, p. 945, 2013.

[13] M. A. Ferrer, M. Diaz-Cabrera, and A. Morales, “Static signature synthesis: A neuromotor inspired approach for biometrics,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 667–680, 2015.

[14] T. Duval, C. Rémi, R. Plamondon, J. Vaillant, and C. OReilly, “Combining sigma-lognormal modeling and classical features for analyzing graphomotor performances in kindergarten children,” *Human Movement Science*, vol. 43, pp. 183–200, 2015.

[15] R.-D. Vatavu, G. Cramariuc, and D. M. Schipor, “Touch interaction for children aged 3 to 6 years: Experimental findings and relationship to motor skills,” *International Journal of Human-Computer Studies*, vol. 74, pp. 54–76, 2015.

[16] M. Djioua and R. Plamondon, “A new algorithm and system for the characterization of handwriting strokes with delta-lognormal parameters,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 2060–2072, 2009.

[17] D. Meyer, “Support vector machines: The interface to libsvm in package e1071,” *Paper available at http://cran. r-project. org/web/packages/e1071/vignettes/svmdoc. pdf*, 2014.