The generation of synthetic handwritten data for improving on-line learning

Marta Režnáková, Lukas Tencer, Réjean Plamondon, Mohamed Cheriet

The generation of synthetic handwritten data for improving on-line learning

Marta REŽNÁKOVÁ[a], Lukas TENCER[b], Réjean PLAMONDON[c] and Mohamed CHERIET[d]
[a,b,d] *École de technologie supérieure (ÉTS), Synchromedia Lab*
Montréal, Québec, Canada (514) 396-8800
[c] *École Polytechnique de Montréal, Laboratorie Scribens*
Montréal, Québec, Canada (514) 340-4711
marta.reznakova@gmail.com, lukas.tencer@gmail.com, rejean.plamondon@plymtl.ca,
mohamed.cheriet@etsmtl.ca

Abstract. In this paper, we introduce a framework for on-line learning of handwritten symbols from scratch. As such, learning suffers from missing data at the beginning of the learning process, in this paper we propose the use of Sigma-lognormal model to generate synthetic data. Our framework deals with a real-time use of the system, where the recognition of a single symbol cannot be postponed by the generation of synthetic data. We evaluate the use of our framework and Sigma-lognormal model by comparison of the recognition rate to a block-learning and learning without any synthetic data. Experimental results show that both of these contributions represent an enhancement to the on-line handwriting recognition, especially when starting from scratch.

1. Introduction

In the recent years, we have seen a growth in the use of smart devices which are often handled with handwritten symbols. In many systems, the use of such symbols requires some pre-definition, which disallows the user to use symbols he finds the best suited for each task. Thus, our work focuses on the on-line learning from scratch, where no pre-definition takes place and it is up to the user to define his own symbols.

However, such a freedom brings challenges from which a low number of samples stands out. Moreover, in this setup, new classes of symbols can be defined one after another and at any time of the use. This leaves a small or even no time for the system's improvement, making it to over-fit, as there is often only one sample per class, and keeps the initial recognition rate low. For this purpose, it is only natural to search for options to obtain more data, so that the system can become more robust sooner and avoid the over-fitting.

There have been several attempts to address the generation of synthetic handwritten data based on a number of models. From these, we list some more recent works, i.e. methods proposing to generate synthetic data based on behavioural models (Schmidt and Lee, 1999), kinematic models (Plamondon and Djioua, 2006), minimization principles (Neilson and Neilson, 2005; Tanaka, 2006) or neural networks (Gangadhar, 2007). Work in this paper is relies on Kinematic Theory describing rapid human movements by Sigma-lognormal (Plamondon et al, 2014; Djioua and Plamondon, 2009; Plamondon et al, 2003).

We apply this method on evolving models capable of on-line learning from scratch. There are some attempts to tackle on-line learning problem (Ditzler, 2010; Yao et.al, 2010; Leistner, 2009; Grabner and Bischof, 2006; Luughofer, 2008). However, very few methods handle also the learning from scratch. Nevertheless there are works also capable of learning from scratch, such as (Angelov, 2010; Almaksour, 2010; Reznakova, 2012; Angelov, 2004). In this work we use ARTIS (Reznakova, 2013), an evolving fuzzy model combined with ART network (Carpenter, 1991).

This paper is organized as follows. In section 2 the proposed framework for on-line learning using synthetic data is described. The Sigma-lognormal model is described in section 3 and the ARTIST model in section 4. Evaluation results are provided in section 5, concluded along with future works in section 6.

2. Framework for online real-time learning using synthetic data

Since in our work we focus on real-time and real-use on-line learning system for handwritten symbols recognition, we need to adjust the use of synthetic data to this problem. To use these data for the learning, a system needs to generate on the fly, just after a new sample is introduced. In the case of block learning, where each new sample waits for all synthetic samples generated from the previous real sample to be learned, the recognition of such new sample is naturally postponed. This happens in case of fast addition of new real samples to the system. To prevent this from happening and to support the real-time on-line learning, we propose a framework described in following.

Figure 1 displays the complete framework we propose in this section. It is divided into two major processes denoted with solid and dashed lines. At the beginning, one sample X is fed to both of these processes, original recognition and learning, and generation of synthetic data. Once a sample is processes by the synthetic data generator, new synthetic data are added to a buffer, waiting for their use. When no real sample is being fed to the system, samples from the buffer can start to be processed. This means that real data have a priority over synthetic data in the means of feeding them to the model. Once and until no real sample is needed to be processes, buffer incrementally supplies the model with synthetic data that are being picked randomly, i.e. the buffer is being shuffled.

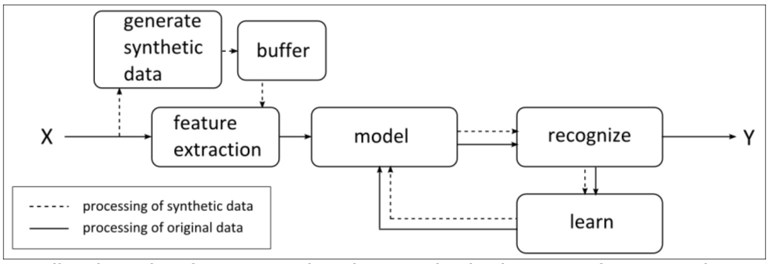

Figure 1. Framework for on-line learning from scratch using synthetic data. X refers to an input sample that is fed to the model and recognized. The information from the recognition process is then used for the learning and new data are added to the model (the model is evolving). The solid line refers to the processing of real samples, where based on recognition the system gives an output Y of a predicted label. The dashed line refers to the additional processes related to the generation of synthetic data.

Ideally, each time a new class is introduced, a number of samples are provided to the system to allow it to learn these classes faster. This framework is thus coping with this problem, avoiding the single-class block learning where it is not appropriate by shuffling its buffer. In the results section, we will note that even though we do not learn classes in blocks right after their occurrence, this framework is able to sustain the recognition rate at the beginning of the learning at the same level as the ideal case.

3. Log-normal model for artificial handwritten data

The Sigma-lognormal model is based on the Kinematic Theory of rapid human movements describing the impulse responses of a neuromuscular network (Plamondon et al, 2014). It is composed of a number of vectors with 6 parameters $\mathrm{p} = [\mathrm{t}_0, \mathrm{D}, \mu, \sigma, \theta_s, \theta_e]$. Details about the various equations of the model have been presented quite often and we refer the reader to the previous reference for more information.

Each time new sample composed of coordinates is introduced to the system, Sigma-lognormal model is derived based on its trajectory. After this, a noise $\mathrm{n}_\mu \in [-0.15\mu, 0.15\mu]$, $\mathrm{n}_\sigma \in [-0.15\sigma, 0.15\sigma]$ resulting into vector $[0, 0, \mathrm{n}_\mu, \mathrm{n}_\sigma, 0, 0]$ (Martin-Albo, 2014) is added to the lognormal model. Then, Sigma-lognormal model is translated back to the coordinates that are used for feature extraction. In this work we use geometrical features based on (Reznakova, 2012; Willems, 2008; Peura, 1997; Rubine, 1991). As we can notice in Figure 1, these samples, now represented by a feature vectors, are then supplied to the buffer awaiting their further use.

4. ARTIST

The ARTIST model is based on Takagi-Sugeno fuzzy model (Takagi and Sugeno, 1985). It contains three main parts, i.e. rule generation, antecedent part and consequent part. TS models are represented by rules, where each rule gives an output for every class. The output of one rule is derived from antecedent and consequent parts (1). The antecedent part states, how much an unlabeled sample is similar to samples that have created this rule ($x\ IS\ a$). Then, the consequent part returns the opinion of this rule about the label of the sample ($THEN\ \{y^i = \beta x \pi^i\}_{i=1..c}$).

$$IF\ x\ IS\ a\ THEN\ \{y^i = \beta x \pi^i\}_{i=1..c} \tag{1}$$

The generation of the rules is based on ART-2A network (Carpenter, 1991). This allows generating the rules in an incremental manner, where the classes do not need to be pre-defined and can be added on the fly. Moreover, rules are not generated with the influence of information about newly added class, but in a completely automatic way.

5. Results

In this work, we focus on real-time on-line learning from scratch. This means that no pre-processing has taken place before the initiation of recognition and learning process. Moreover, each class is introduced on the fly by random order, mostly creating no time for more than one-sample learning. This occurs especially at the beginning of the learning process, when only few samples are known to the system, and thus the recognition is weak. Thus, by using synthetic data, we try to solve this problem, and as we can see in following results, using the Sigma-lognormal model is helpful.

We perform our evaluation on a handwritten gestures dataset (Synchromedia) containing ~33k samples for 17 classes. In this work however, we use only a few samples to better demonstrate the enhancement that using synthetic data offers.

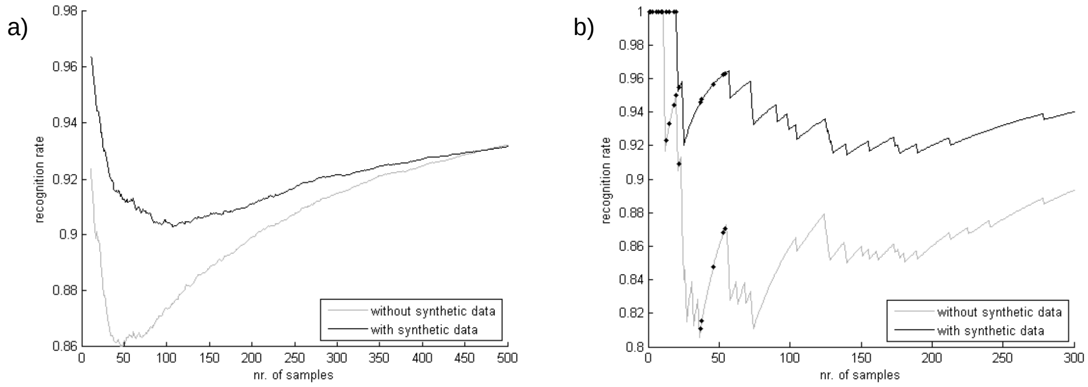

Figure 2. Comparison of recognition rate when using only original data to the use of synthetic data, without (a) and with (b) highlighted times of addition of new class (dots).

In Figure 2a, we compare the initial learning with the use of purely original data to the initial learning using also synthetic data based on our proposed framework. In real-use, there is no fixed time left for the learning of synthetic data, i.e. the time between two consecutive original samples. To mimic this, we let the system select a random number of samples from the buffer within a range [0, |buffer|], that are allowed to be learned in the time between two consecutive original samples are added. As it can be seen, the use of synthetic data helps the system to learn new classes much faster, which results into higher initial recognition rate, i.e. avoiding drop in the recognition rate after new class is introduced. In Figure 2b, we compare the two methods with highlighted time of when new classes are added to the system. As we can note, at the beginning of the learning process, the recognition rate seems to be very high. This is because when a new class is introduced to the system, no output is expected, and thus no error is recorded. As it is clear from the graph, most of the classes have been introduced at the beginning, which also influences this phenomenon, but here again the use of synthetic data is beneficial.

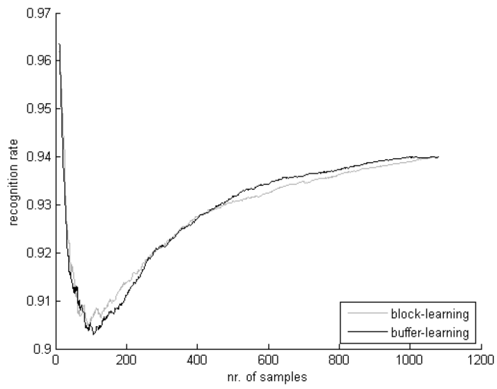

Figure 3. Comparison of recognition rate between block and buffer-based learning.

In Figure 3, we compare the initial recognition rate with block learning, i.e. the system learns all the synthetic data for each newly added sample, to the recognition rate when using buffer, i.e. random number of randomly picked synthetic samples from a buffer are allowed to be learned. As it can be seen, the buffer learning approach is as efficient as the block learning, with the benefit of working real time.

All results except Figure 2b are an average of 50 distinct runs. In all results, we can clearly note that using Sigma-lognormal model for generation of synthetic data improves the recognition rate in the initial learning. Each original sample is used to generate additional 10 samples. This whole process takes on average 12.96 s. The size of the buffer is unlimited, emptying at random occasions, which simulates the real use.

6. Conclusion and discussion

The main goal of this work was to solve the problem of low performance in recognition at the initial phase of the learning. The main causes of this problem are related to learning from scratch and on the fly. However, both of these features are necessary for user-friendly applications, where the user is let to keep his freedom of choice in what symbols he will use on his device.

In this work, we proposed to use synthetic data based on Sigma-lognormal model. Also, we proposed a new framework using shuffled buffer in order to follow the real-time use of the recognition system, when there is no room to postpone its responses. For evaluation, we have used a random subset of the samples from handwritten symbols database in order to show the performance of our propositions for the initial phase of the learning process.

We have shown that the use of synthetic data indeed helps to increase the performance which is so needed for on-line learning and recognition. We have also shown that the use of buffer helps to retain similar level of recognition rate as compared to the block learning.

In our future work, we will focus on the adjustment of noise added to the Sigma-lognormal model, to be more suited for the purpose of synthesis of handwritten symbols. Also we will work on adjusting the number of generated synthetic data, in order to enhance the system and at the same time to prevent big amounts of data. Another study of interest will be to explore the behaviour of the system when noisy data are being synthesized.

Acknowledgement

This work was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC) as well as the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Abdulah Almaksour, E. A. (2010). Evolving Fuzzy Classifiers: Application to Incremental Learning of Hadwritten Gesture Recognition Systems. Istanbul: International Conference on Pattern Recognition (ICPR).

Angelov, P., Filev, D. P., & Kasabov, N. (2010). Evolving Intelligent Systems: Methodology and Applications. In J. W. Sons. New York.

Carpenter, G., Grossberg, S., & Rosen, D. B. (1991). ART 2-A: An adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks*.

Ditzler, G., Muhlbaier, M. D., & Polikar, R. (2010). Incremental Learning of New Classes in Unbalanced Datasets: Learn + + .UDNC. *Multiple Classier Systems*, 33-42.

Djioua, M., & Plamondon, R. (2009). Studying the variability of handwriting patterns using the Kinematic Theory. *Human Movement Science*, 588-601.

Don J.M. Willems, a. R. (August 19, 2008). *Definition for Features used in Online Pen Gesture Recognition.*

Gangadhar, G., Joseph, D., & Chakravarthy, V. (2007). An oscillatory neuromotor model of handwriting generation. *Int. J. Doc. Anal. Recognit.*, 69–84.

Grabner, H., & Bischof, H. (2006). On-line Boosting and Vision. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, (pp. 260-267).

Leistner, C., & Bischof, H. (2009). On-line Random Forests. *IEEE 12th International Conference on*, (pp. 1393-1400).

Luughofer, E. D. (2008). FLEXFIS: A Robust Incremental LEarning Approach for Evolving Takagi-Sugeno Fuzzy Models. *IEEE Transactions on fuzzy systems, vol. 16, no. 6*, 1393-1410.

M. Peura, J. I. (1997). Efficiency of Simple Shape Descriptors. *Advances in Visual Form Analysis*, 443-451.

Martin-Albo, D., Plamondon, R., & Vidal, E. (2014). Training of On-line Handwriting Text Recognizers with Synthetic Text Generated Using the Kinematic Theory of Rapid Human Movements. *14th International Conference on Frontiers in Handwriting Recognition*, (pp. 543-548).

Neilson, P., & Neilson, M. (2005). An overview of adaptive model theory: solving the problems of redundancy, resources and nonlinear interactions in human movement control. *J. Neural Eng.*, 279–312.

Plamen P. Angelov, D. P. (February 2004). An Approach to Online Identification of Takagi-Sugeno Fuzzy Models. *IEEE Transactions on system, man, and cybernetics, vol. 34*, 484-498.

Plamondon, R., & Djioua, M. (2006). A multi-level representation paradigm for handwriting stroke generation. *Hum. Mov. Sci. 25*, 586–607.

Plamondon, R., Feng, C., & Woch, A. (2003). A kinematic theory of rapid human movement. Part IV: a formal mathematical proof and new insights. *Biol. Cybern. 89*, 1266-138.

Plamondon, R., O'Reilly, C., Galballyb, J., Almaksour, A., & Anquetil, E. (2014). Recent developments in the study of rapid human movements with the kinematic theory: Applications to handwriting and signature synthesis. *Pattern Recognition Letters*, 225-235.

Režnáková, M., Tencer, L., & Cheriet, M. (2012). Online handwritten gesture recognition based on Takagi-Sugeno fuzzy models. *11th International Conference on Information Science, Signal Processing and their Applications (ISSPA).*

Režnáková, M., Tencer, L., & Cheriet, M. (2013). ARTIST: ART-2A driven Generation of Fuzzy Rules for Online Handwritten Gesture Recognition. *12th International Conference on Document Analysis and Recognition,ICDAR '13*. .

Rubine, D. (1991, July). Specifying gestures by example. *SIGGRAPH Comput. Graph., 25*(4), 329--337.

Schmidt, R., & Lee, T. (1999). Motor Control and Learnng: A behavioral Emphasis. *3rd Ed. Human Kinetics*, (p. 493p.). Champaign Illinois.

Synchromedia. (n.d.). *http://www.synchromedia.ca/web/ets/gesturedataset.*

Takagi, T., & Sugeno, M. (1985). Fuzzy Identification of Systems and Its Appliations to Modeling and Control. *IEEE Transactions on systems, man, and cybenetics, vol. SMC-15, no. 1*, 116-132.

Tanaka, H., Krakauer, J., & Qian, N. (2006). An optimization principle for determining movement duration. *J. Neurophysiol.*, 3875–3886.

Yao, M., Qu, X., Gu, Q., Ruan, T., & Lou, Z. (2010, June). Online PCA with adaptive subspace method for real-time hand gesture learning and recognition. *W. Trans. on Comp., 9*(6), 583--592.