

Universum Learning for Semi-Supervised Signature Recognition from Spatio-Temporal Data

Lukas Tencer, Marta Režnáková, Mohamed Cheriet

► **To cite this version:**

Lukas Tencer, Marta Režnáková, Mohamed Cheriet. Universum Learning for Semi-Supervised Signature Recognition from Spatio-Temporal Data. Céline Rémi; Lionel Prévost; Eric Anquetil. 17th Biennial Conference of the International Graphonomics Society, Jun 2015, Pointe-à-Pitre, Guadeloupe. 2015, Drawing, Handwriting Processing Analysis: New Advances and Challenges. <hal-01165925>

HAL Id: hal-01165925

<https://hal.univ-antilles.fr/hal-01165925>

Submitted on 20 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Universum Learning for Semi-Supervised Signature Recognition from Spatio-Temporal Data

Lukas TENCER ^a, Marta REŽNÁKOVÁ ^a and Mohamed CHERIET ^a

^a *École de technologie supérieure (Département de Génie de la Production Automatisée)
1100 Rue Notre-Dame Ouest
H3C 1K3, Montreal, CANADA*

Abstract. We present a novel approach towards signature recognition from spatio-temporal data. The data is obtained by recording gyroscope and accelerometer measurements from an embedded pen device. The idea of Universum learning was previously presented by Vapnik and recently popularized in machine learning community. It assumes that the decision boundary of a classifier lies close to data with high uncertainty. The quality of the final classifier strongly depends on a way how to choose the Universum data and also on the representation of original data. In our paper we use a novel approach of Universum learning to classify signature data, also we present our novel idea how to sample the Universum data. At last, we also find more effective representation of the signature data itself compared to the baseline method. These three novelties allow us to outperform previously published results by 4.89% / 5.58%.

1. Introduction

Signature is a unique identifier which is used by millions of institutions and agencies to authenticate people. Over the ages, it served as one of the main ways to acknowledge a contract or an agreement. Despite modern advances in biometrics, signature is a prevailing means of unique identification. Because of this, to classify a signature is an essential task for biometrics and other security applications.

Besides visual information, a signature itself captures also a lot of additional data, which are not captured as a final result recorded on a paper or a digital medium. Specifically it is a way how the signature is drawn. Small aspects like tilt of the pen, acceleration over segments of the signature or order in which are elements of the signature drawn could provide essential insight into signee's identity. Because of this we rely on these rich data and we develop an algorithm, which can identify a user based on these data. Specifically we are talking about information captured from accelerometer and gyroscope devices. This approach is even more justifiable if we think about mass production and spread of mobile devices with embedded sensors to gather acceleration and orientation data, which enable collection of data for signature verification if the signature is made while holding the phone.

Once we are able to collect rich input data, we can further think also about algorithms for classification and how to improve them. Our aim is to use the popular semi-supervised learning methods, which can capture a structure hidden in the data and impose additional assumptions on the structure of the data. One of the very recent ideas was introduced by Vapnik and is called Universum learning. It relies on Universum examples and forces the decision boundary to be close to these examples (see Figure 1). Our approach can benefit from this additional assumption and basically gain “free” increase in performance. How “free” it really is, we will discuss further in the paper. Nevertheless, the improvement in performance is significant enough to be considered a serious improvement compared to the last published results on the same dataset (Griechisch, Malk, & Liwicki, 2013).

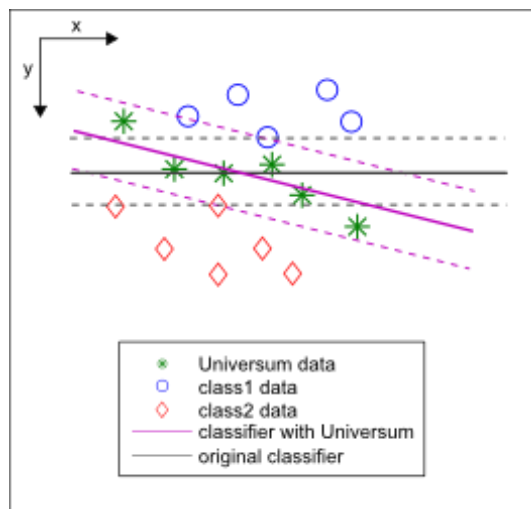


Figure 1: Illustration of Universum learning and change of the decision boundary based on Universum data (Dhar, 2014)

The general overview of works regarding signature classification could be find in (Impedovo & Pirlo, 2008) (Weiping, Xiufen, & Kejun, 2004) and (Wu, Jou, & Lee, 1997). Further works which regard specifically accelerometer and gyroscope data could be seen as early as (Plamondon & Parizeau, 1988) (Baron & Plamondon, 1989) or (Rohlik, Mautner, & Matousek, 2001) (Mautner, Rohlik, Matousek, & Kempf, 2002). More recent approaches which use spatio-temporal data are (Bashir, Scharfenberg, & Kempf) (Bunke, Csirik, Gingl, & Griechisch, 2011) (Shastry, Burchfield, & Venkatesan, 2011) and (Malik, Ahmed, Dengel, & Liwicki, 2012). One of the most recently published approaches, which uses Legendre series and SVM is (Griechisch et al., 2013) and is also the source of our dataset and our baseline technique. The Universum learning works were originally proposed by Vapnik and the theory was further expanded in (Weston, Collobert, & Sinz, 2006) (Sinz, Chapelle, Agarwal, & Schölkopf, 2008) (Cherkassky, Dhar, & Dai, 2011) and (Dhar, 2014). At last an overview of active learning techniques could be found in (Settles, 2010) and QbC is more specifically described in (Seung, Opper, & Sompolinsky, 1992)

This paper is structured as following: Section 2 describes our general methodology, Section 3 presents our experiments and achieved results and in Section 4 we conclude and discuss our work and outline future work.

2. Methodology

Our approach does have a structure of a standard pattern recognition pipeline, that means: feature extraction, learning and classification. One additional step which is present is the generation of Universum examples. This step comes during before the learning phase and a way how Universum examples are generated is essential for performance of Universum learning.

In a first step, we extract the features from the sequential signature data. Prior works (Griechisch et al., 2013) (Parodi, Gómez, & Liwicki, 2012) used Legendre series for the approximation. But based on our preliminary experiments we have selected Hermite polynomials which do have a form:

$$p(x) = c_0 + c_1 H_1(x) + \dots + c_n H_n(x); H_n(x) = n! \sum_{k=0}^{\frac{n}{2}} \frac{(-1)^k}{k! (n-2k)!} \frac{x^{n-2k}}{2^k} \quad (1)$$

This choice improved the performance of the algorithm by $\sim 0.5-0.7\%$. Other polynomials which we evaluated were Legendre and Chebyshev polynomials.

Once we extracted the features, we need to generate the Universum examples. The most common approach how to generate Universum examples is the random averaging. This method is designed for a binary case but could be easily generalized to multi-class case using one-against-all or one-against-one strategy. The random averaging selects a sample from the positive and negative class and then averages their respective feature values individually. This approach is feasible in cases, the linear transition along the features axes yield similar results to the input elements. But this is not always the case with polynomial approximation. A small change in the value of polynomial coefficients might result in quite a different example if reconstructed from these coefficients. Because of this we decided to use active learning method Query-by-Committee to select the relevant examples for the Universum class.

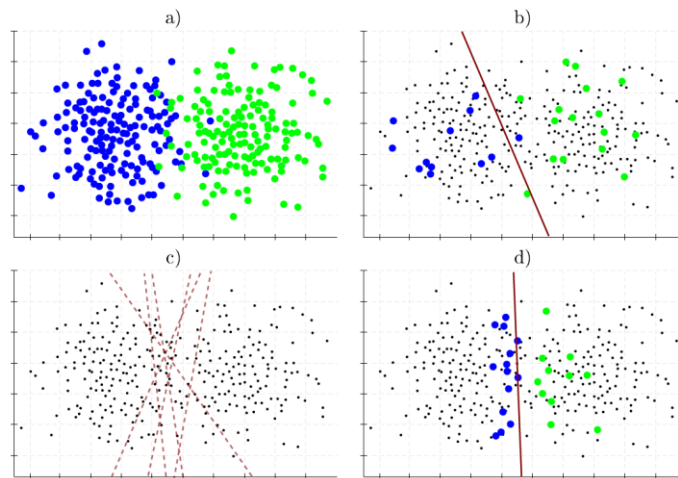


Figure 2: Illustration of QbC sampling a) Original data; b) Decision boundary based on original data; c) Members of the committee; d) The new decision boundary based on uncertain samples

As seen in Figure 2., the QbC samples examples with low confidence and inserts them into the Universum set. The QbC algorithm first constructs a collection (Committee) of sub-classifiers created from a fraction of original data. Then these sub-classifiers vote on generated Universum examples and examples with a highest number of disagreements are selected. After this improvement, the performance of the algorithm improved significantly, as you can see in Table 1.

Once we have obtained the Universum examples, we need to formulate the objective function. As mentioned earlier, Universum learning assumes that the decision boundary is close to uncertain examples. This gives a criterion which needs to be included in the cost function. In the case of SVM classifier then the cost function is:

$$\frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^M \varphi[y_i f_{w,b}(x_i)] + D \sum_{j=1}^N \rho[f_{w,b}(u_j)] \quad (2)$$

where φ and ρ are loss functions, $f_{w,b}$ is the discriminant function with parameters w and b for training points $\{x_i, y_i\}$ and Universum points $\{u_j\}$ with free parameters C and D . The Universum was used also in combination with Boosting (AdaBoost) algorithm (with hypothesis F and L1 regularization where D controls the weight) and in that case the cost function is:

$$\min_w \frac{1}{M} \sum_{i=1}^M \exp(-y_i F(x_i)) + \frac{C}{2N} \sum_{j=1}^N F(u_j)^2 + D 1^T w ; s. t. w \geq 0 \quad (3)$$

Once training is done, during the inference process is standard as for any regular SVM classifier.

As a part of our future work, we would like to explore the possibility of a “close miss” Universum data. Especially in cases like signature verification, where forgery is common. We would expect the forged examples to be quite close to the decision boundary. If we have a label for forged examples, we would like to have them the “forged” label, but at the same time be close to the decision boundary.

3. Experiments and Results

We have evaluated our method to the baseline method presented in (Griechisch et al., 2013), where we used the same evaluation protocol. The method of (Griechisch et al., 2013) achieves good results on the dataset, (composed of 300 signatures from 10 authors) but our technique could outperform it. At the same time we used a toy dataset to visualize the influence of Universum learning on a small dataset and as you can see in Figure 3, the decision boundary is much more viable and is much less overfitting than in case without Universum examples.

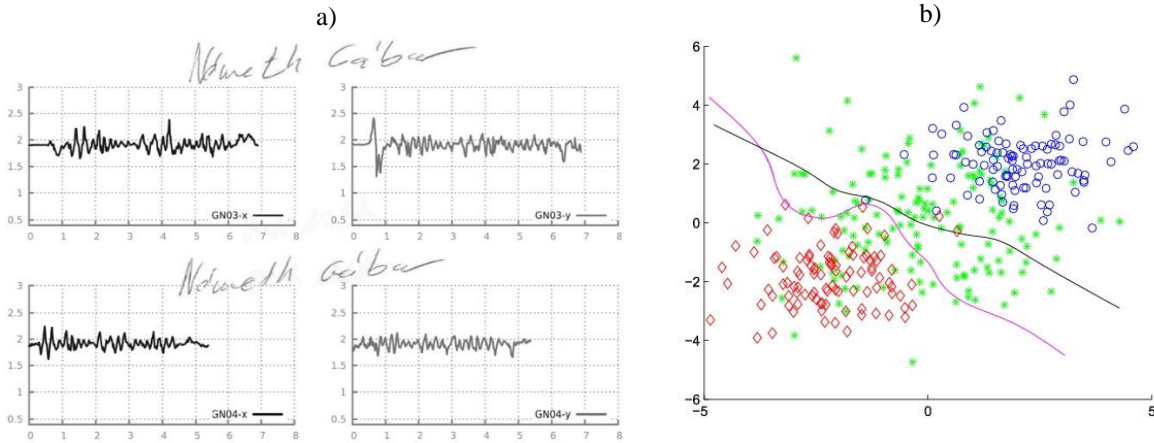


Figure 3: a) Examples from the dataset and qualitative results of the Universum learning b) [(pink) original classifier (black) classifier with Universum data]

Table 1: Performance of classifiers on Accelerometer and Gyroscope data

Method → Dataset ↓	baseline	USVM	USVM + Hermite	USVM + QbC	USVM + Hermite + QbC
AccSig	88.00	90.72	91.25	92.46	92.89
GyroSig	80.44	83.96	84.61	85.88	86.02

The qualitative results are presented in Figure 3 and the quantitative results are evaluated in Table 1. In both cases, we can see the improvement achieved thanks to Universum learning. Our method was able to improve

by 4.89% / 5.58% the baseline method due to various factors. If we considered each of the improvements separately, then the gain was as following: Hermite polynomial approximation: 0.53% / 0.65%, Universum Learning 2.72% / 3.46%, QbC sample selection for Universum learning 1.74% / 1.92%. As we can see the combined approach is higher than any of the individual approaches, but the final gain does not add to the sum of partial gains, what could be expectable, since some of the gains could actually overlap for different partial methods.

4. Conclusion, Discussion and Future Work

In this paper, we have presented a novel use of the Universum learning on task of classification of signatures from accelerometer and gyroscope data. At the same time, we introduced a novel approach for selection of Universum examples which outperforms the random averaging used in previous Universum learning classifiers. At last we also presented a novel approach for feature selection based on Hermite polynomials, which improves the performance of baseline technique which used the Legendre polynomial approximation. In the end we can conclude that our solution was able to outperform the baseline technique by 4.89% / 5.58% and that the combination of Universum learning, our new sample selection method, and the new features proves to be an efficient combination.

Also, as we can see, the concept of Universum learning is highly dependent on Universum set. The generation of this set still does not fully capture the uncertainty in examples and needs to be tuned and designed in task-specific manner. This is a topic which we would like to address in our future work. Especially we would like to build a model which can produce unlikely examples before the approximation happens. For this, we need to build a meaningful representation of the signatures themselves, such that the final result will be hard to classify even by human evaluators. Some of the methods which we would like to explore in this manner are generative density models. The authors would like to thank to SSHRC Canada and NSERC Canada for their financial support.

References

- Baron, R., & Plamondon, R. (1989). Acceleration measurement with an instrumented pen for signature verification and handwriting analysis. *IEEE Transactions on Instrumentation and Measurement*, 38(6), 1132–1138.
- Bunke, H., Csirik, J., Gingsl, Z., & Griechisch, E. (2011). Online signature verification method based on the acceleration signals of handwriting samples. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7042 LNCS, pp. 499–506).
- Cherkassky, V., Dhar, S., & Dai, W. (2011). Practical conditions for effectiveness of the universum learning. *IEEE Transactions on Neural Networks*, 22(8), 1241–1255.
- Dhar, S. (2014). Analysis and Extensions of Universum Learning. (pp. 37-116)
- Griechisch, E., Malk, M. I., & Liwicki, M. (2013). Online signature analysis based on accelerometric and gyroscopic pens and legendre series. In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*
- Impedovo, D., & Pirlo, G. (2008). Automatic Signature Verification: The State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(5), 609–635.
- Malik, M. I., Ahmed, S., Dengel, A., & Liwicki, M. (2012). A signature verification framework for digital pen applications. In *Proceedings - 10th IAPR International Workshop on Document Analysis Systems, DAS 2012* (pp. 419–423).
- Mautner, P., Rohlik, O., Matousek, V., & Kempf, J. (2002). Signature verification using ART-2 neural network. *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02.*, 2, 636–639.
- Parodi, M., Gómez, J. C., & Liwicki, M. (2012). Online signature verification based on legendre series representation. Robustness assessment of different feature combinations. In *Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR* (pp. 379–384).
- Plamondon, R., & Parizeau, M. (1988). Signature verification from position, velocity and acceleration signals: a comparative study. In *Pattern Recognition, 1988., 9th International Conference on* (pp. 260–265 vol.1).
- Rohlik, O., Mautner, P., & Matousek, V. (2001). A new approach to signature verification: digital data acquisition pen. *Neural Networks, IEEE Transactions on*, 2000.
- Settles, B. (2010). Active learning literature survey. University of Wisconsin, Madison.
- Seung, H., Opper, M., & Sompolinsky, H. (1992). Query by committee. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 287–294.
- Shastri, A., Burchfield, R., & Venkatesan, S. (2011). Dynamic Signature Verification Using Embedded Sensors. 2011 International Conference on Body Sensor Networks, 168–173.
- Sinz, F., Chapelle, O., Agarwal, A., & Schölkopf, B. (2008). An Analysis of Inference with the Universum. In *Advances in Neural Information Processing Systems 20* (pp. 1–8).
- Smart, B., Device, P., Bashir, M., Scharfenberg, G., & Kempf, J. (n.d.). Person Authentication by Handwriting in air using a 2 Biometric Smart Pen Device for Data Acquisition, 219–226.
- Weiping, H., Xiufen, Y., & Kejun, W. (2004). A Survey of Off-line Signature Verification. *International Conference on Intelligent Mechatronics and Automation*, (August), 536–541.
- Weston, J., Collobert, R., & Sinz, F. (2006). Inference with the universum. *Proceedings of the 23rd International Conference on Machine Learning*, 1009–1016.
- Wu, Q. Z., Jou, I. C., & Lee, S. Y. (1997). On-line signature verification using LPC cepstrum and neural networks. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society*, 27, 148–153.