# Signature features with the visibility transformation (ICPR2020)

## Definition: signature [7]

Given $X : [a, b] \to \mathbb{R}^d$, an $\mathbb{R}^d$-valued path mapping from $[a, b]$. The *signature* of a path $X$ is the infinite collection of all iterated integrals of $X$. That is,

$$S(X)_{a,b} := (1, S(X)^1_{a,b}, \ldots, S(X)^d_{a,b}, S(X)^{1,1}_{a,b}, S(X)^{1,2}_{a,b}, \ldots), \quad (1)$$

where the superscripts of the terms after the 0th term run along the set of all multi-index $\{(i_1, \ldots, i_k) | k \geq 1, i_1, \ldots, i_k \in [d]\}$. The finite collection of all terms $S(X)^{i_1,\ldots,i_k}_{a,b}$ with the multi-index of fixed length $k$ is termed as the *kth level of the signature*. The truncated signature up to the $p$th level is denoted by $\lfloor S(X)_{a,b} \rfloor_p$.

## Signatures as features

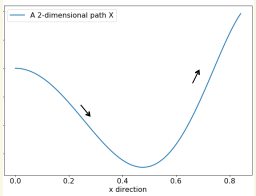
- Determining solutions for controlled differential equations.
- Invariance under time-reparameterisations.
- Unwrapping nonliearity.
- Fixed dimension under length variations, vectorisation.
- Faithful representation: tree-like equivalence.
- **Only capturing the effect of pattern change and not ones depending on the absolute position.**

$$\text{Path X} \overset{\text{Signature}}{\Longrightarrow} \text{Incremental effects of path X.}$$

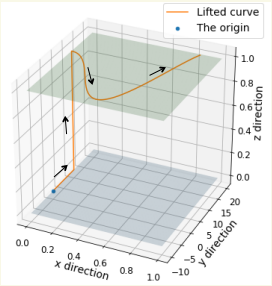
## The goals of our study

- To introduce some transform that can preserve effects of both increments and positions of the original streamed data within signature features.
- To offers a new way of preparing datasets and does not need to change the pipeline.

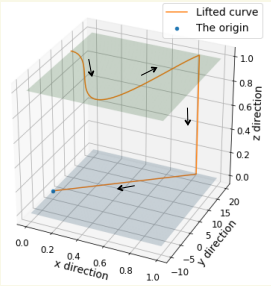
## Definition and property: visibility transformation (VT)




**(a)** Original 2-dimensional path X.




**(b)** $\mathbf{I}$-visibility transformation ($\gamma_{\mathbf{I}}$).




**(c)** $\mathbf{T}$-visibility transformation ($\gamma_{\mathbf{T}}$).

**Figure:** The paths (in subfigure **(b)** and **(c)**) generated from original path $X$ (subfigure **(a)**) via $\mathbf{I}$-visibility transformation and $\mathbf{T}$-visibility transformation.

$$\text{Path X} \overset{\text{VT}}{\Longrightarrow} \text{Path } \bar{\text{X}} \overset{\text{Signature}}{\Longrightarrow} \text{Positional and incremental effects of path X}$$

## Discrete VT

Given streamed data $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, $\mathbf{x}_i \in \mathbb{R}^d$. Because of the availability of Python packages to calculate signature such as *iisignature* [9] from data points, one only need to input data without having to transfer it to a path by

$$\text{data } \mathbf{X} \overset{\text{Python packages}}{\Longrightarrow} \text{signatures.}$$

To apply VT in practice, one needs a discrete transform such that

$$\text{data } \mathbf{X} \overset{\text{discrete VT}}{\Longrightarrow} \text{data } \bar{\mathbf{X}} \overset{\text{Python packages}}{\Longrightarrow} \text{signatures } \& \text{ data } \bar{\mathbf{X}} \underset{\text{into a path}}{\overset{\text{turned}}{\Longrightarrow}} \text{path } \bar{\text{X}}.$$

Take a discrete data with two 2-dimensional observations $[1, 2]^T$, $[3, 4]^T$ for example, the discrete $\mathbf{I}$-visibility transformation (IVT) would give

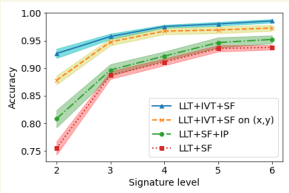
$$\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix} \overset{\text{IVT}}{\Longrightarrow} \begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}3\\4\\1\end{bmatrix},$$

## Pipeline

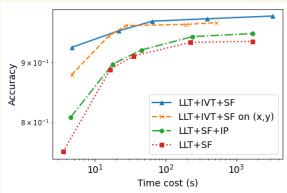

**Figure:** The workflow of the pipeline for feature extraction using visibility transformation: the input is the original data stream on the left end, and one may utilise different transforms on the data stream for data cleaning and scaling; this is followed by applying the visibility transformation on the cleaned data; finally one will use package to translate the transformed data to signature.

## Applications: character trajectories dataset

- The data [12] consists of 2858 instances for 20 different characters, and was captured using a WACOM tablet at 200Hz. Each character sample is a 3-dimensional pen tip velocity trajectory, namely $(x, y, p)$. The original handwriting data contains training set (50%) and testing set (50%).
- 4 Different features: truncated signatures with the lead lag transform (LLT), truncated signatures with LLT and being prefixed by the explicit initial position (LLT+IP), and truncated signatures with LLT and the discrete $\mathbf{I}$-visibility transformation (LLT+IVT). We also extracted truncated signature features with LLT and IVT on the trajectory, namely the $(x, y)$ path (LLT+IVT on $(x, y)$). In the experiment, the signature features were truncated to levels $\{2, 3, 4, 5, 6\}$.
- lightGBM Classifer: hyperparameter tuning implemented via grid search with cross validation.




**(a)** Signature level VS average accuracy with standard deviation.




**(b)** CPU time VS average accuracy (log-log scale).

**Table:** Comparison of different methods on character trajectories dataset.

| Method | Accuracy | Method | Accuracy |
|---|---|---|---|
| $\phi$(O,HMM)+SVM [8] | 92.91% | TK [11] | 93.67% |
| LLT+IVT+SF on $(x, y)$ | **97.27%** | SDD [2] | 98.00% |
| MCDS [3] | 98.25% | LLT+IVT+SF | **98.54%** |

## An application: Chalearn 2013 data

- The ChaLearn 2013 multi-modal gesture dataset [1] contains 23 hours of Kinect data of 27 subjects performing 20 Italian gestures.
- Liao et al [5] proposed a log-signature-based recurrent neural network model.

**Table:** Comparison of different methods on the Chalearn 2013 data.

| Method | Accuracy | Method | Accuracy |
|---|---|---|---|
| Deep LSTM [10] | 87.10% | Two-stream LSTM [13] | 91.70% |
| ST-LSTM + Trust Gate [6] | 92.00% | Three-stream net TTM [4] | 92.08% |
| PT-Logsig-RNN | 92.21% | Modified PT-Logsig-RNN | **92.89%** |

## References


[1] Escalera, S., Gonzàlez, J., Baró, X., Reyes, M., Lopes, O., Guyon, I., Athitsos, V. and Escalante, H. *Multi-modal gesture recognition challenge 2013: Dataset and results*, In Proceedings of the 15th ACM on International conference on multimodal interaction, (2013): 445-452.

[2] Grabocka, J., Wistuba, M. and Schmidt-Thieme, L., Fast classification of univariate and multivariate time series through shapelet discovery. *Knowledge and information systems*, 49.2 (2016): 429-454.

[3] Iosifidis, A., Tefas, A. and Pitas, I., Multidimensional sequence classification based on fuzzy distances and discriminant analysis. *IEEE Transactions on Knowledge and Data Engineering*, 25.11 (2012): 2564-2575.

[4] Li, C., Zhang, X., Liao, L., Jin, L. and Yang, W., Skeleton-based gesture recognition using several fully connected layers with path signature features and temporal transformer module. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 33 (2019): 8585-8593.

[5] Liao, S., Lyons, T., Yang, W. and Ni, H., Learning stochastic differential equations using RNN with log signature features. *arXiv preprint*, arXiv:1908.08286.

[6] Liu, J., Shahroudy, A., Xu, D., Kot, A.C. and Wang, G., Skeleton-based action recognition using spatio-temporal lstm network with trust gates. *IEEE transactions on pattern analysis and machine intelligence*, 40.12(2017):3007-3021.

[7] Lyons, T., and Qian, Z., *System control and rough paths.* Oxford University Press, 2002.

[8] Perina, A., Cristani, M., Castellani, U. and Murino, V., A new generative feature set based on entropy distance for discriminative classification. *International Conference on Image Analysis and Processing*, (2009): 199-208, Springer, Berlin, Heidelberg.

[9] Reizenstein, J., The iisignature library: efficient calculation of iterated-integral signatures and log signatures. *arXiv preprint*, arXiv:1802.08252.

[10] Shahroudy, A., Liu, J., Ng, T.T. and Wang, G., Ntu rgb+ d: A large scale dataset for 3d human activity analysis. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016):1010-1019.

[11] Tsuda, K., Kawanabe, M., Rätsch, G., Sonnenburg, S. and Müller, K.R., A new discriminative kernel from probabilistic models. *Advances in Neural Information Processing Systems*, (2002): 977-984.

[12] Williams, B.H., Toussaint, M. and Storkey, A.J., Extracting motion primitives from natural handwriting data. *International Conference on Artificial Neural Networks*, (2006): 634-643, Springer, Berlin, Heidelberg.

[13] Wang, H. and Wang, L., Modeling temporal dynamics and spatial configurations of actions using two-stream recurrent neural networks. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 499-508.



**Yue Wu (yue.wu@maths.ox.ac.uk) Mathematical Institute, University of Oxford**

**Co-authors: Hao Ni, Terry Lyons, Robin Hudson**

Department of Mathematics, University College of London; Mathematical Institute, University of Oxford; Department of Mathematical Sciences, Loughborough University


DataSig

A rough path between mathematics and data science