ABSTRACT Inkball models provide a tool for matching and comparison of spatially structured markings such as handwritten characters and words. Hidden Markov models offer a framework for decoding a stream of text in terms of the most likely sequence of causal states. Prior work with HMM has relied on observation of features that are correlated with underlying characters, without modeling them directly. This paper proposes to use the results of inkball-based character matching as a feature set input directly to the HMM. Experiments indicate that this technique outperforms other tested methods at handwritten word recognition on a common benchmark when applied without normalization or text deslanting.

Inkball Models as Features for Handwriting Recognition

Nicholas Howe, Smith College Andreas Fischer, University of Fribourg & HES-SO Baptiste Wicht, University of Fribourg & HES-SO





IDEA: Use the inkball model fit as input to a standard HMM. The model fit captures in one number a host of complicated structural details. While a single character model on its own is not

Baseline System



T64

The fit score at any point corresponds to a particular configuration of the model in

Inkball models represent characters as disks of ink distributed along a pen trace. Setting a distribution on the relative locations of neighboring disks yields a generative model. The likelihood of an observation under the model corresponds to the configuration with minimal collective displacement.

discriminating enough to serve for character recognition, the collection of responses to a full set of character models carries meaningful patterns of information that the HMM can utilize.



Hidden Markov Models (HMM) are a standard statistical framework for sequence recognition tasks. Originally introduced for speech, they have also been widely used for handwriting recognition based on different handwriting features. Well-known examples include Marti & Bunke's geometric features and the SIFT-like gradient features proposed by Rodriguez & Perronin and Terasawa & Tanaka, respectively. None of these directly model the character structure.

For any given horizontal coordinate, the HMM receives as input the best fit of the character model over all vertical positions.



prototyp

character |

ndividual

HMM states correspond to horizontal positions. At each state, the HMM sees the best fit of every character model.





| Maybe feature values are sensitive to scale variations? | How are the prototypes selected for the baseline algorithm? | Is there any other way to lower the error rate? |
|--|--|--|
| Experiment S3 triples the number of features, with prototypes at 80% | They are picked arbitrarily. | Maybe. Modifying the deformation score to use truncated Gaussians has shown promising results before. |
| Maybe fitting a part-structured boundary model gives better features | That doesn't sound very rigorous. Maybe choosing more representative character prototypes will improve the results? | Experiments T4 through T64 test varying truncation levels. Error Rate for Robust Fitting |
| than an inkball model? Experiment BM uses part-structured boundary models [Howe, HIP 2015] | Experiments KM1-KM5 cluster character samples using <i>k</i> -medoids and take the cluster centers as prototypes. | 40 30 20 10 |
| information to distinguish its position in each character? | Are there any other ways to select good prototypes? | 0 BL T4 T8 T16 T32 T64 Fold 1 Fold 2 Fold 3 Fold 4 Mean |
| Experiment DC adds a set of derivative features to the originals | Experiments IG1-IG3 greedily select prototypes for information gain | Does robust fitting lower the error? |
| Experiment WC adds a set of features comparing the local feature value to the minimum in a character- sized window | Error Rate for Prototype Selection 40 | Yes! The best results appear at T8. Comparison with Prior Work |
| Error Rate for Feature Variants | 30 | How does this method compare with |

We compare to Marti & Bunke (M01),

All use the 20-page George Washington dataset, divided into four folds.







| Letters orders and Instanctions Frombane , | Ligno anders and Instructions Dumber 195 | 306 Letter, Orders and Instructions Derin me | Leffers Biders and Instructions Burn in Bas | ses letters bodens and Anstructions timber the |
|---|--|--|---|--|
| la sene al colonneste al gran desena consta la la desena de la colonia de la grana de la de la de la colonia de la grana de la de la de la colonia de la dela desena de la colonia de la dela dela dela dela dela la dela dela dela dela dela dela dela Colonia Retariano de grana de fundare Colonia Retariano de grana de fundare | but the performance of the the constraint of the second se | data berra et al per barriel et for mangement ann es es dete data, i fan mandet de Barriel ann es es dete data es a fan mandet de data esta esta esta esta esta data data data esta esta esta esta esta data data data data data esta esta data data data data data data data esta esta data data data data data data data da | Constantine Sameta Margar Second and Sameta Sameta Margar Second and a second a second and second a second | extraction in the other of the constant of performance of the theory of the other of the theory of theory of the theory of theory of the theory of the theory of theory of theory of the theory of the theory of the |
| however, because the part of a second field dataset and the part of the second data and the second part of the second data and the second part of the second dataset of the second part of the part of the second dataset of the second part of the second dataset of th | Manghan 2 and a starting of the Springer and more one of the Springer Methy, the have a start of the springer Methy of the start and the springer springer and the springer and the springer springer and the springer of the springer springer and the springer of the springer springer and the springer of the springer sp | Fold 4 | ages a dar mande og som for at the public of a dar dar her det som en her en state det an en som en som en state and en som en som en som en som en som en som en som en state at det som etter at som en som en som en state at det som etter at som en som en som etter at som etter at som en som en som etter at som etter at som en som en som etter at som etter at som en som etter to etter at som etter at som en som etter at at som etter at som ett | Hence the least of a set of the second of the set of the second of the s |
| $\label{eq:second} \left\{ \begin{array}{llllllllllllllllllllllllllllllllllll$ | Harry or you to see the set have been there to see the set have the set of the second set of the second of the second second second second second to the second second second second second the second | "All March Labor " Jones of Labor " and the strength of the de- order of the Samati and the Marchata of go the produces of the strength of the Satahan the strength of the strength of the Satahan and the strength of the strength of the strength of the strength of the strength of the strength of the strength of the strength of the strength of strength of the strength of the strength of the strength of the strength of the strength of the strength of the strength of strength of the strength of the strength of the strength of the strength of the strength of the strength of the strength of the | you have be able to partice the second of th | 2. M. Darter, Henrich Levik, et d., Hagener, Sognand, S., Starter, S., Starter, S., Sandar, J., Sansa, et al. Staff of an et al. Sansa de field approximation of the software space of the software starter of the software space of the software starter of the software space of the software software space on some starter of the software on Staff of the software some starter of the software on Staff of the software some starter of the software on Staff of the software some starter of the software on the Staff of the software some starter of the software on the Staff of the software some starter of the software on the Staff of the software some starter of the software on the Staff of the software some starter of the software some starter of the Staff of the Staff of the software some starter of the Staff of Staff. |

Are the images preprocessed in any way?

> Page images are binarized and split into lines.



Do any of the variant feature selections improve on the baseline method (BL)?

30

No. Even though all except **BM** include the baseline features as a subset, none outperform the baseline. There seems to be a penalty for adding features.

Further experiments will stick to the baseline method.

BL KM1 KM2 KM3 KM4 KM5 IG1 **IG2** IG3 ■ Fold 1 ■ Fold 2 ■ Fold 3 ■ Fold 4 ■ Mean

Do any of the prototype selection methods improve on the baseline?

> Not significantly. The IG variants do the best, with **IG2** showing the lowest mean error. But the improvement is so small it may not be worth the extra effort required for the selection process.

The IG2 method may be useful in some situations. But the simplicity of the baseline method is still attractive.



Error Rates for Prior Work & Current 80 60 40 20 T09 IG2 M01 **R08 T**8 ■ Fold 2 ■ Fold 3 ■ Fold 4 ■ Mean Fold 1

The results look suspiciously good.

prior work?

All methods are run here without preprocessing & cleaning, which handicaps the prior work.

Don't you need to deslant the lines?

No. With this method neither deslanting nor noise cleanup is required.

CONCLUSION Inkball character models fitted to observations serve as excellent input to a Hidden Markov Model for the character recognition task. These results must be seen as preliminary since all experiments are carried out for just one data set (GW20). Future work should look at performance on additional standard data sets, including multiwriter text with a diversity of handwriting styles. A more strict comparison with prior work would include image deslanting and cleaning for the algorithms that depend upon it.