

# Third Year PhD Report

## Extracting Motion Primitives from Natural Handwriting Data

Ben H Williams  
Institute of Adaptive and Neural Computation  
School of Informatics  
University of Edinburgh  
5 Forrest Hill, Edinburgh UK  
*ben.williams@ed.ac.uk*

July 19, 2007

## 1 Project Summary

### 1.1 Last year's report

In last year's review, we described the motivation behind the Piano Model, the workings of the fHMM, along with its constraints, and the formalised integrate and fire timing model, as described in the IJCAI07 paper. Generative samples from the uncoupled model show the scribbling behaviour, generative samples from the IF model show it can be used in this uncoupled fashion to produce characters.

In the discussion section, we talked about possible Gaussian models, such as a mixture model to model the timing of the spikes.

The main aims were to make a generative model for handwritten characters, using the current fHMM model, and a simple timing model. Examination of the timing code would also be interesting. Development of primitives during a learning task could be investigated. Validation of the primitives themselves should be done, in terms of length and number. Also exploring style transformations from one person's primitives to another's would be interesting to look at.

### 1.2 Summary of project changes

The main change to the project this year has been the development of the HMM Timing model, which models the timing of the spikes derived from the fHMM primitive model. Using such a model allows the presence of a spike to be

dependent upon the presence of the previous spike from a particular primitive. Also it allows a convenient method of learning the model, using standard EM for HMMs. More importantly it allows a distribution of where spikes are likely to occur, given a particular character model, to be produced, and then used as a ‘spike prior’ for the fHMM model, effectively coupling the two models together. This gives a more informed prior for the onset probability of a primitive at time  $t$ . Instead of a flat prior, as was used before, this time-dependent prior takes into account the fact that the data came from character samples, allowing more appropriate inference of primitives. Details of this model can be found in the latest attached paper.

The shifting algorithm was modified to allow extensions and reductions to the length of the primitive, meaning the primitive length is no longer fixed, but determined by the shape of the primitive, as described in the thesis chapter attached.

Several other algorithmic changes were made to the fHMM, as detailed in the attached chapter, including changes to the constraints placed upon the parameters - zero state constraints, self transitions, rank shortening, flat priors, ht smoothing to avoid oscillations, merging. The heuristic methods such as shifting and merging were altered to take into account the variance of the primitive.

The treatment of the data was modified, so that the character samples are normalised to unity variance over time. The end section of the characters are extended, to alleviate pen up problems, as described in the attached chapter.

Several modifications to the generative script were made, trying different ways to improve the performance. Although the timing HMM captures the dependencies between the presence of neighbouring activations of a particular primitive, it does not capture them between different primitives. This could possibly be captured by using a larger, joint hidden state space, looking at all possible occurrences of all primitives. Alternatively, a higher level model, looking at the ‘mode’ of a character, capturing the hidden state progression of the timing model states could provide a further level of abstraction, and help with coupling between primitives in terms of their presence or absence. The current model assumes the spike times have independent Gaussian noise associated with them. It could also be possible to look at the correlations between the timings of the spikes, suggesting either some sort of common time offset, or time warping, or by using a single, multivariate Gaussian, thus capturing the covariance between different spikes. These correlations need to be examined further, and a simple model introduced to deal with them.

A recognition script has been implemented, that selects a character sample at random, then fits some pre-learnt primitives to it, then assesses the likelihoods of the various timing models for explaining the spike data. An alternative way would be to use the coupled models. So far this approach works for reduced datasets, but it appears as though certain characters are more ‘popular’ than others. Possibly the primitives are not fit as well as they could be, or the timing models are not discriminative enough. Looking at individual primitives, rather than the average over the whole model might be better. Also, using a coupled timing model might help. Further work needs to be done in the recognition area

of this model.

Currently the characters are labelled before the learning process, allowing different timing models to be learnt based upon which character is present. It would be better to have a discrete random variable modelling the character type, and for this to be inferred.

## 2 Achievements

The current uncoupled fHMM model works very well to extract primitives that represent the data, and a variance analysis shows that the variance of the data set is, to a large extent, modelled by the timing variation of these primitives.

These primitives can be shared amongst all characters, and if learnt from a varied dataset, will be used in virtually all character types.

A common output set of primitives can therefore be learnt by an individual, with these primitives potentially being generalised to every character. The timing model selection allows top down coordination of the primitives for character reproduction.

Using a timing model allows generative samples of characters to be reproduced, with an easily recognisable paper reproduction, and a close approximation to the variance profile of the original dataset.

Primitives can be learnt on reduced datasets, and then quickly fit to larger, datasets using a single E-step. They can also be quickly fit to novel data samples, as would be the case for character recognition. A simple character recognition script can assess the likelihood of different character models giving rise to the data, and therefore form a basis for a discriminative algorithm.

## 3 Necessary work

Analysis of which primitives are used in what characters, to show whether certain primitives are used more than others for particular characters.

Improvements to the generative samples, to include coupling between primitives, either by using a single combined HMM for the timing data, or by introducing a 'mode' model.

Correlations between timing and presence need to be thoroughly examined, and possibly a dynamic time offset model needs to be introduced, to deal with the correlations.

Analysis of state space size, and relationship to character reproduction performance, and variance profile modelling.

Literature review of biological papers on motor cortex and cerebellum, looking at time-extended primitives, such as Hatsopoulos2007.

## 4 Possible work

Recognition script could be improved in a number of ways, but whether it is worth attempting to compete in a field of heavy research such as character recognition is debatable. Probably best to leave current recognition script as it is. Possibly think about an online form of primitive fitting - ie. not to single characters, but to words, then using spike patterns to segment characters - certain patterns are commonly produced at the start of characters maybe?

Differences between primitives from different people. Recently acquired dataset from a student doing related research in Tehran university, gives many samples of his handwriting. One set in Latin script, one in Persian (arabic) script. Several studies looking at differences between primitive shapes, and timing variation are now possible using the uncoupled model.

Psychophysical experiments involving either learning new characters, altering the visual feedback of characters whilst drawing, memory experiments, or differentiation experiments could be performed, although these would involve a long data collection period.

Introducing a pen-up pen-down variable could help with characters that have separate strokes, such as a t or an i.

Learning primitives from natural writing rather than single characters may provide different results. Also more varied datasets such as drawing and sketching might be interesting to look at.

Examination of how natural handwriting could be modelled generatively, eg. a HMM for character type, then a linked state looking at transitions between characters.

Generalisation across characters - instead of a separate timing model for each character, it would be better to generalise somehow. Having a discrete random variable modelling the character type, and then the dependencies to different primitives from this could generalise in this fashion.

## 5 Time Plan

**July** Report, new dataset, log likelihood calculations

**August** Analysis of correlations, analysis of primitives, repeatability of algorithm, more thorough analysis of results, for the purposes of thesis results section

Timing model improvements - both joint HMM (already implemented, but some learning issues), and mode model (analysis of correlations will help)

**September** Journal paper?

Writing up and possible experiments

Timing model extensions

**October** Writing up and possible experiments

**November** Writing up

**December** Any alterations to model, or further experiments for justification of model

**January** Finalise thesis

## 6 Thesis Chapter Plan

A thesis for this project will be based upon the following chapter descriptions, allowing for small rearrangements or additional chapters should the content prove significant.

**Introduction** An overview of the project goals and achievements. A description of why the subject is interesting and useful, and an introduction to related research. An overview of the thesis chapters, and organisation of the model from a non-technical top-down perspective.

**Background** Motor Primitives have been studied in biology for over 10 years now, and are a hot subject of research in robotic control. This chapter will look at the evidence for biological motor primitives, focussing on the differences between definitions and characteristics of how the primitives operate.

Recently, robotic control research has started to use primitives rather than direct motor-task space calculations to plan and execute movements. Several approaches to this problem have been attempted, all making slightly different assumptions about what a primitive is. These differences will be contrasted in this chapter.

A related area of research to the above fields is that of how to extract primitives from data. As this mainly depends on how primitives are defined, details of this research will fit in to this chapter.

Another background area of research is that of generative handwriting models. There are a several that are learnt from data, and can be contrasted with our model. An overview of these models will be presented here, as the technical details will not help with the understanding of our own model.

**Assumptions and Definitions** How and why we define primitives as we do, and the implications for the modelling process. Differences between what we expect to see in biology, and the simplifications we make for the model. A formal definition of the motivating Piano Model will be described, along with related research.

**Probabilistic Model Overview** A top-down description of our model, detailing the coupling of the timing and the primitive components, along with a walk through of the generative operation, similar to that in the NIPS

paper. Possibly algorithmic pseudo-code, detailing the iterative loops, and shifting/numerical issues/cluster merging/initialisations etc. Also the structure of the inference procedure will be detailed.

**fHMM Model** As the models are naturally described separately, this chapter will focus on the detailed operation of the fHMM model, without the coupling prior from the Timing model. Much of this chapter will come from the ICANN paper, (and earlier parts of the ICML paper). **See attached**

**fHMM results** Uncoupled operation of the fHMM will be described here, along with the approximations that allow a timing code to be produced.

**Timing Model** Various timing models have been tried, and they will be detailed here, along with samples of the generative behaviour of each model. The bulk of the chapter will focus on the HMM based timing model, and will compare likelihoods of different models where possible, and show the convergence of the HMM model inference.

**Coupled model** The method of coupling the models together (by introducing a non-stationary primitive onset prior) will be described in this chapter. Results of coupling the models, showing the suppression of outliers, and improved clustering of the spikes will be demonstrated.

**Generative results** A more complete treatment of the results will be presented here, with different datasets, and multiple characters. A comparison of the variance of the generative samples is useful for model assessment, and will be shown here, along with reconstruction errors, likelihoods of various models, and repeatability of the inference.

**Applications** Several applications are possible, the most obvious being that of character recognition, and writer identification. A worked example of how these might be possible using this model will be detailed here. Other applications taking advantage of the generative nature of the model such as personalised letter writing, or web pages will also be described.

**Experiments** Several experiments looking at the effect of primitive corruption, or timing noise on perception and reproduction would be interesting to perform. If time permits, results of these will be detailed here.

**Discussion** The usefulness of this approach, and benefits to understanding biological systems, designing robotics, modelling handwriting, and other biologically produced data will be discussed here. Also possible extensions such as feedback will be addressed.