

ABSTRACT

In last years Hidden Markov Models (HMMs) have received significant attention in the task off-line handwritten text recognition (HTR). As in automatic speech recognition (ASR), HMMs are used to model the probability of an observation sequence, given its corresponding text transcription. However, in contrast to what happens in ASR, in HTR there is no standard set of local features being used by most of the proposed systems. In this thesis we propose the use of raw binary pixels as features, in conjunction with models that deal more directly with the binary data. In particular, we propose the use of Bernoulli HMMs (BHMMs), that is, conventional HMMs in which Gaussian (mixture) distributions have been replaced by Bernoulli (mixture) probability functions. The objective is twofold: on the one hand, this allows us to better modeling the binary nature of text images (foreground/background) using BHMMs. On the other hand, this guarantees that no discriminative information is filtered out during feature extraction (most HTR available datasets can be easily binarized without a relevant loss of information).

In this thesis, all the HMM theory required to develop a HMM based HTR toolkit is reviewed and adapted to the case of BHMMs. Specifically, we begin by defining a simple classifier based on BHMMs with Bernoulli probability functions at the states, and we end with an embedded Bernoulli mixture HMM recognizer for continuous HTR. Regarding the binary features, we propose a simple binary feature extraction process without significant loss of information. All input images are scaled and binarized, in order to easily reinterpret them as sequences of binary feature vectors. Two extensions are proposed to this basic feature extraction method: the use of a sliding window in order to better capture the context, and a repositioning method in order to better deal with vertical distortions. Competitive results were obtained when BHMMs and proposed methods were applied to well-known HTR databases. In particular, we ranked first at the *Arabic Handwriting Recognition Competition* organized during the *12th International Conference on Frontiers in Handwriting Recognition* (ICFHR 2010), and at the *Arabic Recognition Competition: Multi-font Multi-size Digitally Represented Text* organized during the *11th International Conference on Document Analysis and Recognition* (ICDAR 2011).

In the last part of this thesis we propose a method for training BHMM classifiers using

discriminative training criteria, instead of the conventional Maximum Likelihood Estimation (MLE). Specifically, we propose a log-linear classifier for binary data based on the BHMM classifier. Parameter estimation of this model can be carried out using discriminative training criteria for log-linear models. In particular, we show the formulae for several MMI based criteria. Finally, we prove the equivalence between both classifiers, hence, discriminative training of a BHMM classifier can be carried out by obtaining its equivalent log-linear classifier. Reported results show that discriminative BHMMs clearly outperform conventional generative BHMMs.