

En els últims anys els models ocults de Markov (HMMs) han tingut un paper destacat en el camp del reconeixement de text manuscrit off-line (HTR). Al igual que ocorre en el reconeixement automàtic de la parla (ASR), els HMMs s'empren per a modelar la probabilitat d'una seqüència d'observacions donada la seua transcripció. En canvi, a l'inrevés del que ocorre en ASR, en HTR no hi ha un conjunt de característiques estàndard que estiga sent utilitzat per la majoria de sistemes existents. En aquesta tesi proposem utilitzar directament com a característiques els píxels originals binaritzats, conjuntament amb models específicament dissenyats per a tractar amb dades binàries. Concretament, proposem l'ús de Bernoulli HMMs (BHMMs), es a dir, HMMs convencionals on les distribucions de (mixture de) gaussianes són reemplaçades per (mixture de) funcions de probabilitat de Bernoulli. L'objectiu és doble: per una banda açò ens permet modelitzar la naturalesa binària de les imatges amb text (lletres/fons) utilitzant BHMMs. Per una altra banda, açò garanteix que no és perdrà informació discriminant en el procés d'extracció de característiques (la majoria de les base de dades d'HTR disponibles poden ser fàcilment binaritzades sense una pèrdua rellevant d'informació).

Tota la teoria d'HMMs necessària per a desenvolupar un toolkit basat en HMMs es revisada i adaptada per al cas dels BHMMs. Concretament, comencem definint un classificador senzill basat en BHMMs amb funcions de probabilitat de Bernoulli als estats, i acabem amb un reconeixedor de HTR continuu amb mixtures de Bernoulli embegudes. Respecte a les característiques, proposem un procés senzill d'extracció de característiques binàries amb poca pèrdua d'informació. Totes les imatges són escalades i binaritzades per a què es pugen reinterpretar fàcilment com a seqüències de vectors de característiques binàries. Es proposen dos extensions a aquest mètode bàsic d'extracció de característiques: l'ús d'una finestra relliscant per a capturar millor el context, i un mètode de reposicionament per a tractar millor les distorsions verticals. S'han obtingut resultats competitiu quan els BHMMs i els mètodes proposats s'han aplicat a bases de dades ben conegudes en el àmbit del HTR. En particular, hem quedat primers en la competició *Arabic Handwriting Recognition Competition* organitzada en el *12th International Conference on Frontiers in Handwriting Recognition (ICFHR 2010)*, i en la competició *Arabic Recognition Competition: Multi-font Multi-size Digitally*

---

*Represented Text* organitzada en el *11th International Conference on Document Analysis and Recognition (ICDAR 2011)*.

En la part final d'aquesta tesis proposem un mètode per a entrenar classificadors basats en BHMMs fent ús de criteris d'entrenament discriminatius, en comptes del criteri Maximum Likelihood Estimation (MLE). Concretament, comencem proposant un classificador log-lineal per a dades binàries basat en un classificador de BHMMs. L'estimació dels paràmetres d'aquest model es pot dur a terme utilitzant qualsevol criteri d'entrenament discriminatiu per a models log-lineals. En particular, presentem les fórmules per diversos criteris basats en el criteri MMI. Per a acabar, demostrarem l'equivalència entre ambdós classificadors, i per tant, demostrarem que l'entrenament discriminatiu d'un classificador de BHMMs pot dur-se a terme obtenint el model log-lineal equivalent. Els resultats reportats demostren clarament que els BHMMs discriminatius milloren als BHMMs convencionals.