

Représentation et reconnaissance de caractères manuscrits par Réseaux Bayésiens Dynamiques

Laurence Likforman-Sulem, Marc Sigelle

GET-ENST/ Traitement du Signal et des Images et CNRS-LTCI (UMR 5141)
46, rue Barrault, 75013 Paris
{likforman|sigelle}@tsi.enst.fr

1 Introduction

Les approches stochastiques, tels que les modèles de Markov cachés (HMM), sont largement utilisées pour la reconnaissance de la parole et de l'écrit (Elms et al. 1998 ; Hallouli et al. 2002) pour leur capacité à s'adapter aux distorsions élastiques temporelles et spatiales. Cependant ces modèles sont mono-dimensionnels. Une adaptation doit donc être réalisée pour les images, par nature bi-dimensionnelles : celles ci sont converties en séquences 1D d'observations le long d'une direction. Une séquence admissible d'observations est par exemple la suite des colonnes de pixels en balayant l'image de gauche à droite. D'autres séquences sont possibles : vecteurs de caractéristiques sur des fenêtres glissantes, lignes de texte...

Les HMM font l'hypothèse que les observations sont indépendantes conditionnellement aux états cachés, ce qui n'est pas toujours réaliste pour les images. Des extensions des HMM permettant de mieux prendre en compte l'aspect bi-dimensionnel des images ont ainsi été proposées avec les modèles pseudo-2D (ou planar HMM) (Gilloux 1994). Plus récemment, des modèles 2D à base de champs de Markov ont été développés (Park et Lee 1998 ; Saon et Belaid 1999 ; Chevalier et al. 2003). En faisant apparaître les dépendances entre variables d'états ou observations, une modélisation plus fine de phénomènes peut être obtenue. Dans cette optique, des modèles probabilistes s'appuyant sur les réseaux bayésiens statiques sont apparus dans le domaine de la reconnaissance de l'écriture en-ligne (Cho et Kim 2003), l'analyse de documents (Souafi 2002) et l'authentification de signatures (Xiao et Leedham, 2002). Les réseaux bayésiens dynamiques sont une extension des réseaux statiques qui prennent en compte des séquences variables d'observations. On note ξ_t , l'ensemble des variables d'états et d'observations au temps t . Un réseau bayésien dynamique à deux pas de temps (2TBN) est défini par

- un réseau initial B_1 qui spécifie la distribution initiale des états et les distributions conditionnelles des états et des observations à $t=1$
- un réseau de transition B_t qui spécifie les distributions $P(\xi_{t+1} | \xi_t)$. Ces distributions sont supposées stationnaires, i.e. indépendantes de t .

Cette étude expérimente des modèles simples mono-flux de type HMM et des modèles couplés. Les structures couplées sont toutes construites par la mise en correspondance de deux réseaux simples mono-flux (ajouts de liens dans la structure graphique). Dans notre application, les états cachés sont des variables discrètes et les observations sont continues. Les observations sont soit les lignes, soit les colonnes normalisées de pixels d'un caractère, obtenues par balayage séquentiel, soit les deux à la fois. L'évaluation de ces modèles a été réalisée sur la base de chiffres MNIST (LeCun 1998).