

## Auteurs

Hugo Gangloff

IRISA, Université Bretagne

Katherine Morales

IMT/Télécom SudParis

Institut Polytechnique de Paris

Yohan Petetin

IMT/Télécom SudParis

Institut Polytechnique de Paris

## Introduction

Nous proposons un modèle de Chaînes de Markov Cachées (CMC) triplet à variable auxiliaire latente continue. Un tel modèle généralise la plupart des modèles de CMC existants. Nous utilisons l'inférence variationnelle pour la mise en œuvre pratique du modèle. Rappelons qu'une CMC triplet [1] s'écrit :

$$p_{\theta}(\mathbf{t}_K) = p_{\theta}(h_0) \prod_{k=1}^K p_{\theta}(h_k|h_{k-1}) \prod_{k=0}^K p_{\theta}(z_k)p_{\theta}(x_k|h_k, z_k).$$

Dans l'inférence variationnelle [2], nous estimons les paramètres  $\theta$  du modèle lorsque les lois *a posteriori*,  $p_{\theta}(z|x)$ , ne sont pas connues. Elles sont approximées par  $q_{\varphi}(z|x)$  où  $\varphi$  doit aussi être estimé. Dans l'inférence variationnelle, on maximise  $F(\theta, \varphi)$ , une borne inférieure de la log-vraisemblance appelée ELBO :

$$\log p_{\theta}(x) \geq \int q_{\varphi}(z|x) \log \frac{p_{\theta}(z, x)}{q_{\varphi}(z|x)} dz = F(\theta, \varphi),$$

où  $F(\theta, \varphi) = \mathbb{E}_{q_{\varphi}(z|x)}(p_{\theta}(x|z)) - D_{\text{KL}}(q_{\varphi}(z|x)||p_{\theta}(z))$ .

## Inférence variationnelle dans les CMC-BG

**Estimation des paramètres dans les CMC-BG.** Nous pouvons montrer que la ELBO se décompose dans le cas séquentiel des CMC-BG triplet en :

$$\mathcal{L}(\theta, \varphi) = \mathbb{E}_{q_{\varphi}^{\text{opt}}(\mathbf{h}_K, \mathbf{z}_K|\mathbf{x}_K)}(\log \prod_{k=0}^K p_{\theta}(x_k|h_k, z_k)) - \beta_1 D_{\text{KL}}(q_{\varphi}^{\text{opt}}(\mathbf{h}_K, \mathbf{z}_K|\mathbf{x}_K)||p_{\theta}(h_0)p_{\theta}(z_0) \prod_{k=1}^K p_{\theta}(z_k)p_{\theta}(h_k|h_{k-1})) + \beta_2 \mathbb{E}_{p^{\text{emp}}(\mathbf{h}_K)}(\log q_{\varphi}(\mathbf{h}_K|\mathbf{x}_K)),$$

- $q_{\varphi}^{\text{opt}}(\mathbf{h}_K, \mathbf{z}_K|\mathbf{x}_K) = p_{\theta}(\mathbf{h}_K|\mathbf{z}_K, \mathbf{x}_K)q_{\varphi}(\mathbf{z}_K|\mathbf{x}_K)$
- $p^{\text{emp}}(\mathbf{h}_K) = \delta_{\mathbf{h}_K^{\text{pe}}}(\mathbf{h}_K)$  fait intervenir une pré-segmentation déjà disponible.
- $p_{\theta}(\mathbf{h}_K|\mathbf{z}_K, \mathbf{x}_K)$  est calculable exactement par un algorithme de Forward-Backward étendu.
- La ELBO est pénalisée par deux scalaires ( $\beta_1$  et  $\beta_2$ ) afin de garantir l'interprétabilité des états cachés d'intérêt  $\mathbf{h}_K$ .

**Estimation des états cachés.** La distribution *a posteriori*, peut être calculée par échantillonnage séquentiel.

**Paramétrisation profonde des modèles.** Suivant [3], des paramétrisations profondes sont utilisées pour les lois de probabilités  $q_{\varphi}(\mathbf{z}_K|\mathbf{x}_K) = \prod_{k=1}^K q_{\varphi}(z_k|z_{k-1}, x_k)$  et  $p_{\theta}(x_k|h_k, z_k)$  : ces lois sont choisies comme des loi normales de moyennes paramétrées par un réseau à une couche cachée de 100 neurones + activation ReLU.

**Algorithme 1 :** Estimation des paramètres dans les CMC-BG.

**Entrée :**  $\mathbf{x}_K$ , les observations

**Résultat :**  $(\theta^*, \varphi^*)$ , les paramètres estimés

Initialiser  $(\theta^{(0)}, \varphi^{(0)})$

$t \leftarrow 0$

**tant que** convergence pas atteinte **faire**

Tirer  $\mathbf{z}_K^{(n)} \sim q_{\varphi}(\mathbf{z}_K|\mathbf{x}_K)$

Calculer  $p_{\theta}(h_{k-1}|h_k, \mathbf{z}_K^{(n)}, \mathbf{x}_K)$

Calculer la dérivée  $\nabla_{(\theta, \varphi)} \mathcal{L}(\theta, \varphi)$

Mettre à jour les paramètres

$$\begin{pmatrix} \theta^{(t+1)} \\ \varphi^{(t+1)} \end{pmatrix} = \begin{pmatrix} \theta^{(t)} \\ \varphi^{(t)} \end{pmatrix} + \epsilon \nabla_{(\theta, \varphi)} \mathcal{L}(\theta, \varphi) \Big|_{(\theta^{(t)}, \varphi^{(t)})}$$

$t \leftarrow t + 1$

$\theta^* \leftarrow \theta^{(t)}$

$\varphi^* \leftarrow \varphi^{(t)}$

## Segmentation d'images avec les CMC-BG

**Bruit multiplicatif.** On génère une observation  $\mathbf{x}$  selon :

$$x_k|h_k, z_k \sim \mathcal{N}(a_{h_k}; b_{h_k}^2) * z_k.$$

**Bruit non-stationnaire.** On génère une observation  $\mathbf{x}$  selon :

$$\begin{cases} x_k|h_k \sim \mathcal{N}(a_{h_k}; \sigma^2), & \text{if } k \in \{1, \dots, \lfloor K/2 \rfloor\}, \\ x_k|h_k \sim a_{h_k} + \mathcal{E}(\lambda), & \text{if } k \in \{\lfloor K/2 \rfloor + 1, \dots, K\}. \end{cases}$$

La Figure 1 illustre les résultats obtenus. Quelque soit le niveau du RSB considéré : les nouveaux modèles D-CMC-BG et CMC-BG améliorent les modèles existants CMC et CMC-BG-di.

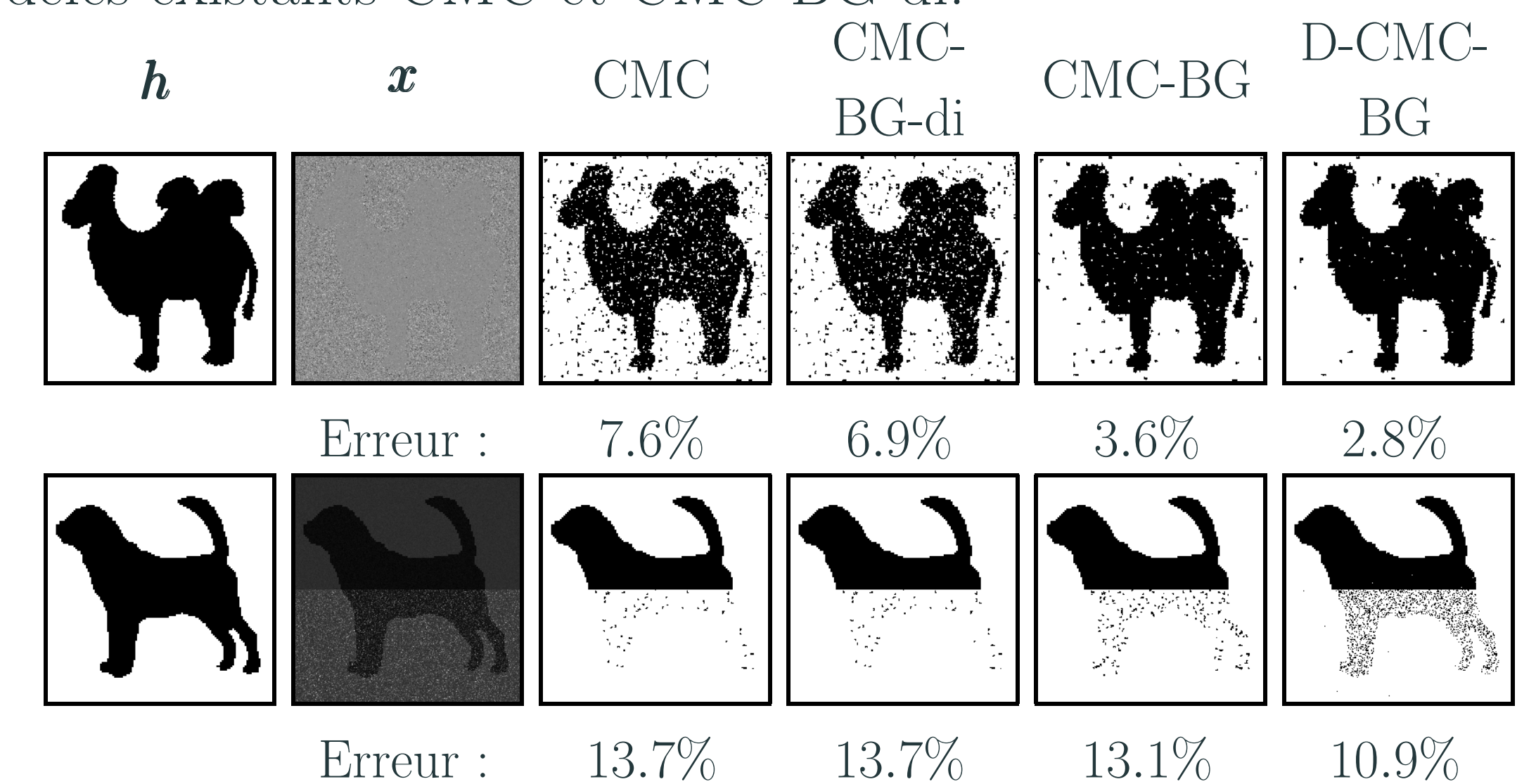


Figure 1: Extraits des résultats des deux expériences de segmentation d'images.

## Références

- [1] Ivan Gorynin, et al., *Assessing the segmentation performance of pairwise and triplet Markov models*, Signal Processing, 2018.
- [2] Diederik P. Kingma, et al., *Auto-encoding variational Bayes*, ICLR, 2014.
- [3] Hugo Gangloff, et al., *A general parametrization framework for pairwise Markov models: an application to unsupervised image segmentation*, MLSP, 2021

## Conclusions

- Introduction d'une variable auxiliaire latente continue qui généralise les modèles de chaînes de Markov cachées existants.
- Développement des algorithmes nécessaires à la préservation de l'interprétabilité et de temps de calculs raisonnables.
- Résultats expérimentaux en accord avec les attentes théoriques.