Télécom SudParis - Diplôme d'Ingénieur

Méthodes statistiques pour la segmentation d'images

MAT5011: Modèles de Markov en Signal et Image

Réalisé pour le 22 octobre 2021

par

Lorenzo HERMEZ

Coordinateurs : Wojciech PIECZYNSKI & Clément FERNANDES

Table des matières

Introduction

Dans ce TP, on s'intéresse à la segmentation d'images. On se limitera toutefois à la segmentation d'images à deux classes pour faciliter les calculs.

1 Bruitage des images

Afin de brouiller les images, nous les avons bruitées avec des bruits gaussiens. Nommons les bruits proposés par l'énoncé :

- Bruit 1 pour $\mathcal{N}(m_1 = 1, \sigma_1^2 = 1)/\mathcal{N}(m_2 = 4, \sigma_2^2 = 1)$
- Bruit 2 pour $\mathcal{N}(1, 1)/\mathcal{N}(2, 1)$
- Bruit 3 pour $\mathcal{N}(1, 1)/\mathcal{N}(1, 9)$

Voici à quoi ressemblent nos images une fois bruitées selon ces bruits gaussiens :

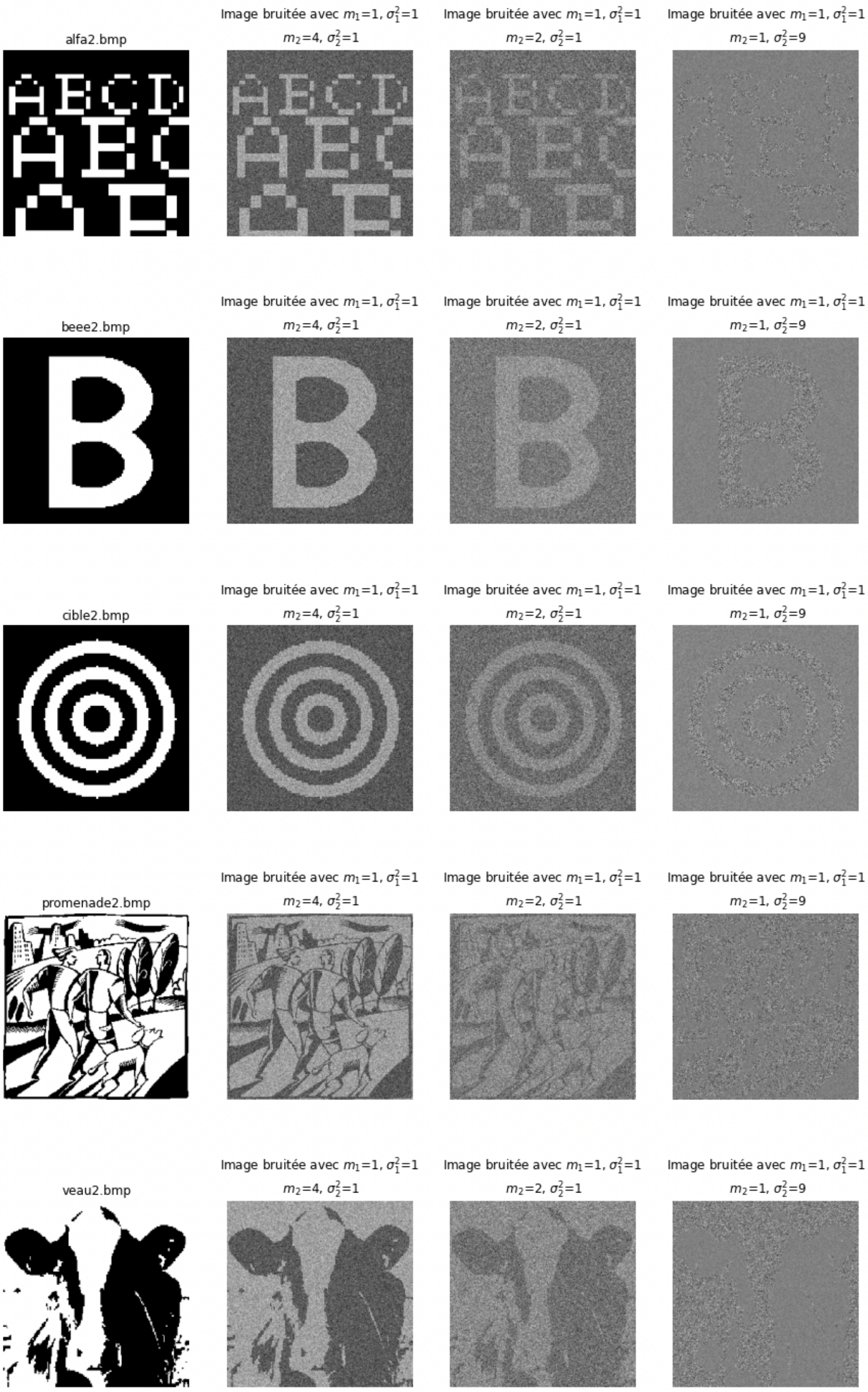

FIGURE 1 – Bruitage des images

On peut voir que, plus les moyennes des bruits sont éloignées, moins l'image paraît (visuellement) bruitée. Egalement, plus l'écart-type est important, plus le bruitage est sévère. On qualifie de *fort* un bruit qui va beaucoup dégrader le signal, si bien que l'erreur entre l'image bruitée soit totalement différente que l'image originale. La segmentation du signal va donc être complexe et les stratégies de restauration vont commettre un nombre plus important d'erreurs que pour un bruit *faible*.

2 Segmentation K-Means

La méthode de segmentation que nous allons utiliser dans cette partie peut paraître comme l'une des plus naïves. En effet, nous allons utiliser l'algorithme K-Moyennes (ou plus couramment appelé K-Means) afin de séparer les classes en les regroupant autour de leur position moyenne (centroïdes). L'algorithme K-Means n'est qu'une implémentation de la minimisation de la fonction de coût :

$$\sum_{S_i} \sum_{j \in S_i} ||x_j - \mu_i||^2$$

où S_i est le cluster n°i et μ_i son barycentre.
Cette fonction de coût s'écrit également de la manière suivante pour n individus et K clusters :

$$\sum_{j}^{n} \sum_{i}^{K} z_j^i ||x_j - \mu_i||^2$$

où z_j^i vaut 1 si l'individu j $\in$ i et 0 sinon.
La fonction de coût est donc caractérisée par $\mu_1, ..., \mu_K$ moyennes des clusters et les $z_j^i, \forall (i,j) \in \{1, ...n\} \times \{1, ..., K\}$ (Ici, K=2 et n=256x256). Le problème majeur dans la minimalisation de cette fonction est qu'il faut tester toutes les valeurs. Or, tester toutes les valeurs de moyennes est impossible car les valeurs sont continues. Afin de contourner ce problème, on peut le voir comme étant une minimalisation de la fonction de coût $J(Z, \mu)$.
Cette minimisation est effectuée de manière itérative en suivant le modèle :

1. Initialisation des centroïdes.
2. Mise à jour des clusters en affectant les données au cluster dont le centroïde est le plus proche. $\forall j \in \{1, ..., n\}$, on choisit $z_j^{k_j} = 1$ tel que $\text{Argmin}_i ||x_j - \mu_i|| = k_j$.
3. Réévaluation des centroïdes. À Z donné, on minimise par rapport à μ : on choisit alors $\mu_i = \frac{\sum_{j=1}^n z_j^i x_j}{\sum_{j=1}^n z_j^i}$.
4. Itérer les étapes 2. et 3. jusqu'à stabilisation des centroïdes.

Les défauts de l'algorithme K-Means sont les mêmes que tous les algorithmes d'optimisation alternée :

— Convergence vers un minimum local.
— Dépend de l'initialisation.

Comme on travaille avec des images à deux classes, on peut affecter un cluster à chaque pixel et ainsi prédire la classe à laquelle le pixel appartenait avant bruitage.
Après avoir vu la longue théorie, passons à la pratique (puisque c'est ça qui nous intéresse !). On bruite alors les images avec les différents bruits présentés plus tôt et on compare avec l'image d'origine.
Voici tout d'abord un résultat visuel de segmentation par KMeans sur l'image nommée "beee2.bmp" ainsi que le tableau récapitulatif des erreurs :

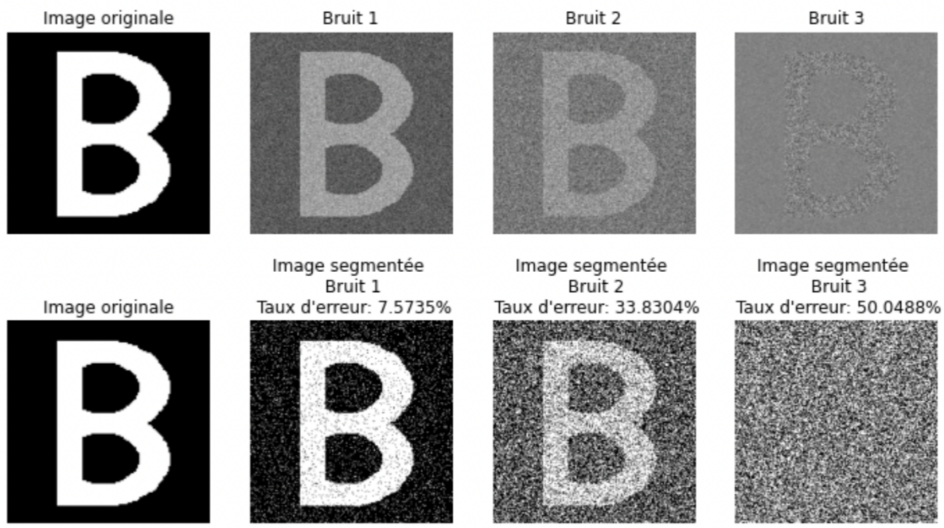

FIGURE 2 – Segmentation de l'image Beee par KMeans

Bruit	Alfa	Beee	Country	Veau	Zèbre
1	0.083	0.076	0.071	0.067	0.068
2	0.353	0.338	0.323	0.309	0.311
3	0.5	0.501	0.501	0.499	0.5

Tableau récapitulatif des taux d'erreur moyens (sur 100 itérations) obtenus suite à la segmentation de 5 images bruitées par 3 bruits par la méthode du K-Means.

On peut voir que, plus les moyennes des bruits sont éloignées, moins l'image paraît (visuellement) bruitée. Egalement, plus la différence d'écart-type est important, plus le bruitage est sévère comme on l'avait présenti précédemment.

Notons que, comme l'algorithme dépend énormément de l'initialisation. Pour se faire, on initialise 30 fois l'algorithme K-Means et on répète cette procédure 100 fois afin d'avoir un taux d'erreur pertinent.
On remarque aussi que l'algorithme K-Means n'est pas un algorithme permettant de segmenter précisément une image. Cela signifie que le clustering n'apporte pas suffisamment d'informations, et donc le fait de regrouper les classes autour de leur position moyenne ne permet pas de faire face à un bruit fort.
Il nous faudra donc recourir à une certaine modélisation statistique afin de segmenter en 2 classes nos images.

3 Segmentation aveugle supervisée d'images de synthèse

Désormais, on se donne deux processus aléatoires $X = (X_s)_{s\in S}$, $Y = (Y_s)_{s\in S}$. Pour tout $s \in S$, X_s prend ses valeurs dans un espace fini $\Omega = \{\omega_1, ..., \omega_k\}$ et Y_s dans $\mathbb{R}$. Le problème de la segmentation est celui de l'estimation de $X = x = (x_s)_{s\in S}$, à partir de l'observation $Y = y = (y_s)_{s\in S}$, i.e. l'image numérique à segmenter. Rappelons, toutefois, que cette segmentation n'est pas aisée car les $X = (X_s)_{s\in S}$ sont inobservables. On considérera les couples (X_s, Y_s) indépendants pour une modélisation plus simple.
On se place dans le cas où $X \in \Omega^S = \{\omega_1, \omega_2\}^S$. On définit également la fonction de perte du MPM L par :

$$\boxed{L(\omega_i, \omega_j) = \mathbb{I}_{\{\omega_i \neq \omega_j\}}}$$

$L(\hat{s}(y), \omega)$ désigne alors la valeur, au point (ω, y), de la fonction indicatrice du sous-ensemble de $\Omega \times \Upsilon$ sur lequel $\hat{s}$ se trompe et $\mathbb{E}[L(\hat{s}(Y), X)]$ la probabilité que $\hat{s}$ se trompe. Dans ce cas la stratégie bayésienne $\hat{s}_B$ est donc définie par :

$$\boxed{\hat{s}_B(y) = \begin{cases} \omega_1 & \text{si} \quad \mathbb{P}(\omega_1|y) \geqslant \mathbb{P}(\omega_2|y) \\ \omega_2 & \text{si} \quad \mathbb{P}(\omega_2|y) \geqslant \mathbb{P}(\omega_1|y) \end{cases}}$$

Notons que $\hat{s}_B$ peut aussi s'écrire :

$$\boxed{\hat{s}_B(y) = \begin{cases} \omega_1 & \text{si} \quad \mathbb{P}(\omega_1)\mathbb{P}(y|\omega_1) \geqslant \mathbb{P}(\omega_2)\mathbb{P}(y|\omega_2) \\ \omega_2 & \text{si} \quad \mathbb{P}(\omega_2)\mathbb{P}(y|\omega_2) \geqslant \mathbb{P}(\omega_1)\mathbb{P}(y|\omega_1) \end{cases}}$$

ce qui permet de faire les calculs à partir de la loi de X a priori et des densités du bruit.

Passons à la pratique, on se place dans la configuration précédente i.e. où l'on a 5 images bruitées avec 3 bruits différents. Désormais, on ne segmentera plus l'image bruitée de manière naïve à l'aide de l'algorithme K-Means mais on la segmentera suivant le critère du MPM dont la théorie se trouve ci-dessus. Comme précédemment, on effectue l'algorithme 100 fois afin de récupérer une erreur moyenne significative. On trouve alors les résultats suivants :
Voici tout d'abord un résultat visuel de segmentation suivant le critère du MPM sur l'image nommée "beee2.bmp".

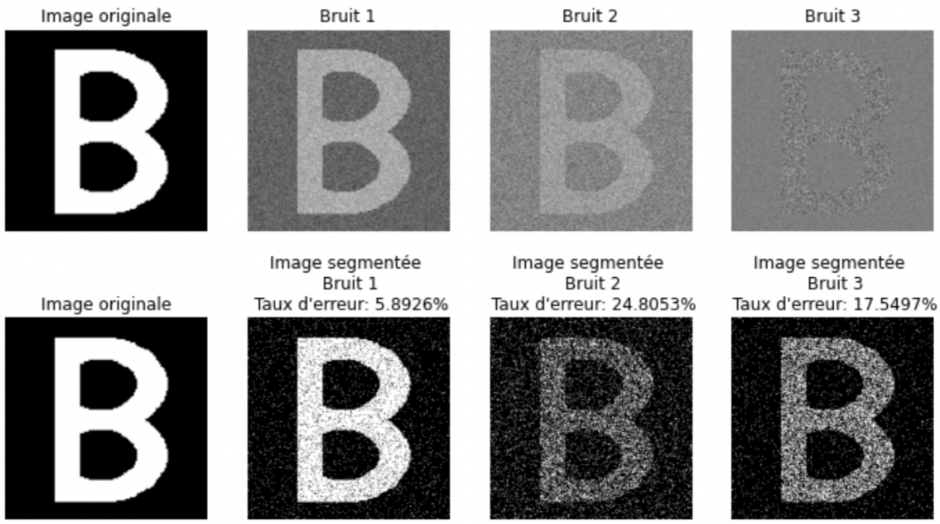

FIGURE 3 – Segmentation de l'image Beee par le critère du MPM

Bruit	Alfa	Beee	Country	Veau	Zèbre
1	0.055	0.059	0.063	0.067	0.066
2	0.22	0.248	0.278	0.307	0.302
3	0.154	0.175	0.204	0.265	0.236

Tableau récapitulatif des taux d'erreur moyens (sur 100 itérations) obtenus suite à la segmentation de 5 images bruitées par 3 bruits par la méthode du MPM.

On remarque dans un premier temps que les taux d'erreur semblent plus faibles que ceux obtenus par la segmentation K-Means. De plus, une fois la connaissance de $\mathbb{P}(\omega_1)$ et de $\mathbb{P}(\omega_2)$ acquise, on aurait pu prédire le taux d'erreur en fonction des différents bruits. La probabilité de se tromper lors de la segmentation par le critère du MPM correspond alors à calculer l'intégrale sous l'intersection des deux gaussiennes :

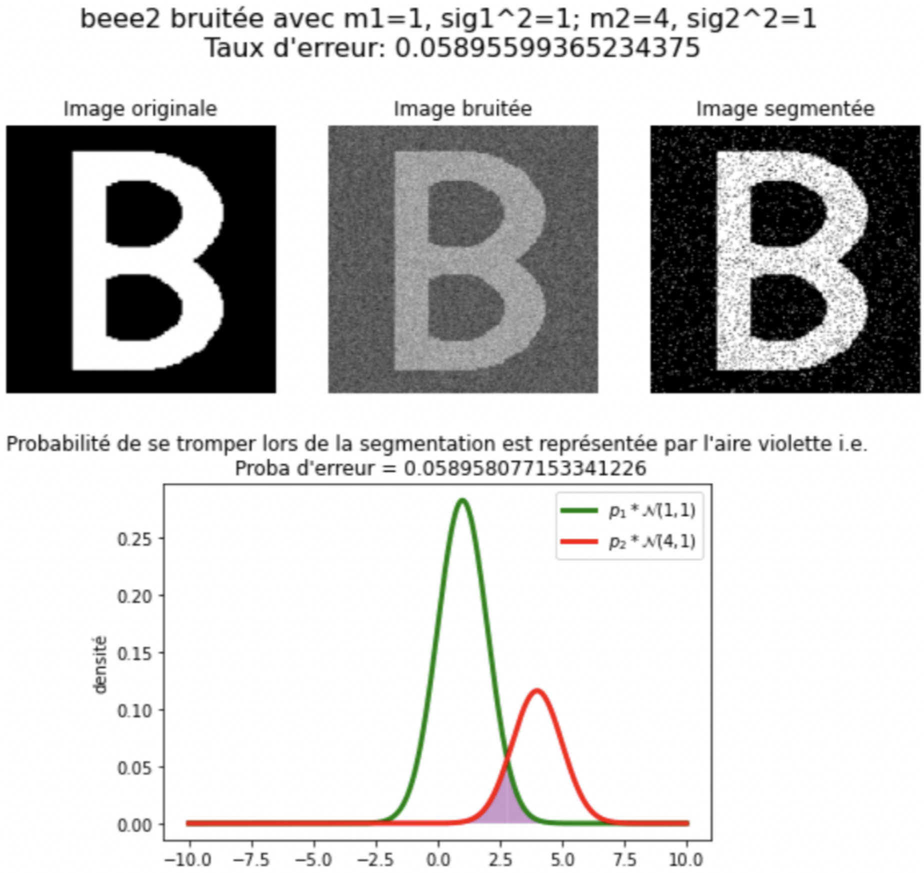

FIGURE 4 – Segmentation de l'image Beee par le critère du MPM et "prédiction" du taux d'erreur

Regroupons toutes ces informations dans un seul tableau, on y verra beaucoup plus clair :

Bruits	alfa2	beee2	country2	veau2	zebre2	Méthode utilisée
$\mathcal{N}(1, 1^2)$ et $\mathcal{N}(4, 1^2)$	0.083	0.076	0.071	0.067	0.068	K-Means
	0.055	0.059	0.063	0.067	0.066	MPM
$\mathcal{N}(1, 1^2)$ et $\mathcal{N}(2, 1^2)$	0.353	0.338	0.323	0.309	0.311	K-Means
	0.22	0.248	0.278	0.307	0.302	MPM
$\mathcal{N}(1, 1^2)$ et $\mathcal{N}(1, 3^2)$	0.5	0.501	0.501	0.499	0.5	K-Means
	0.154	0.175	0.204	0.265	0.236	MPM

TABLE 1 – Tableau récapitulatif des taux d'erreur moyens (sur 100 itérations) entre la méthode du K-Means et le critère du MPM . La meilleure erreur moyenne est colorée en vert et l'autre en rouge, ou en orange si les deux sont du même ordre de grandeur.

Ce tableau est très intéressant et nous apporte plusieurs informations :

1. La méthode du K-Means pour le dernier bruit est identique à tirer aléatoirement un pixel en fonction de la probabilité des classes $\mathbb{P}(\omega_1)$ et de $\mathbb{P}(\omega_2)$ d'où le taux d'erreur de 50%. Il est donc normal que le critère du MPM s'en sorte mieux dans ce cas.
2. Pour les autres bruits, on se rend compte que le critère du MPM fonctionne mieux que l'utilisation de l'algorithme K-Means lorsque les images ne sont pas très complexes. Cependant, pour des images complexes comme veau2 et zebre2 (i.e avec beaucoup de détails), les deux méthodes se comportent de la même manière.

Conclusion

Nous avons mis en oeuvre deux méthodes différentes pour la segmentation des images. Ces méthodes ne sont pas utilisées dans la pratique pour deux raisons :

1. Ici, on travaille à la segmentation des images en ayant connaissance des paramètres des bruits. Ce n'est pas le cas en pratique et la mise en place de l'algorithme EM ou SEM est souvent de rigueur.
2. L'hypothèse des couples indépendants est très limitante dans notre segmentation. Bien que facilitant grandement les calculs, l'étude d'un pixel est réalisée de manière isolée, sans tenir compte des pixels aux alentours. L'utilisation des chaînes de Markov cachées ou champs de Markov cachés peut aider dans ce sens, et c'est ce que nous verrons plus tard.

Statistics is Everywhere