

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة غرداية

Université de Ghardaia

كلية العلوم والتكنولوجيا

Faculté des Sciences et de Technologie

قسم الرياضيات و الاعلام الآلي

Département des Mathématiques et Informatique



MEMOIRE

Présenté pour l'obtention du **diplôme** de **MASTER**

En : Informatique

Spécialité : Systèmes Intelligents pour l'Extraction de Connaissances (SIEC)

Par : Sid ali HACINI et Mustapha SAADA

Sujet

Segmentation d'image par les chaines de Markov cachées

Soutenu publiquement, le 12/ Juin / 2017, devant le jury composé de :

M.Djelloul ZIADI	Professeur	Univ. Rouen-France	Président
M. Youcef MAHDJOUR	Maitre Assistant	Univ. Ghardaia	Encadreur
M. Slimane BELLAOUAR	Maitre Assistant	Univ. Ghardaia	Examineur
M.Mehamed BETKA	Maitre Assistant	Univ. Ghardaia	Examineur

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة غرداية

Université de Ghardaia

كلية العلوم والتكنولوجيا

Faculté des Sciences et de Technologie

قسم الرياضيات و الاعلام الآلي

Département des Mathématiques et Informatique



MEMOIRE

Présenté pour l'obtention du **diplôme** de **MASTER**

En : Informatique

Spécialité : Systèmes Intelligents pour l'Extraction de Connaissances (SIEC)

Par : Sid ali HACINI et Mustapha SAADA

Sujet

Segmentation d'image par les chaines de Markov cachées

Soutenu publiquement, le 12/ Juin / 2017, devant le jury composé de :

M.Djelloul ZIADI	Professeur	Univ. Rouen-France	Président
M. Youcef MAHDJOUR	Maitre Assistant	Univ. Ghardaia	Encadreur
M. Slimane BELLAOUAR	Maitre Assistant	Univ. Ghardaia	Examineur
M.Mehamed BETKA	Maitre Assistant	Univ. Ghardaia	Examineur

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dédicace

Je dédie ce travail:

*A ma chère mère qui a perdu sa jeunesse pour nous, et mon
père,*

A mes chères frères,

*A toute ma famille, mon binôme et sa famille, et mes amis,
A tous les enseignants qui m'ont suivis durant ma formation,*

A tous qui m'ont aidé à réaliser ce travail,

A tout qui porte le nom HACINI.

Sid Ali

Dédicace

Je dédie ce travail:

~~*A ma chère mère qui a perdu sa jeunesse pour nous, et mon père,*~~

~~*A mes chères frères,*~~

~~*A toute ma famille, mon binôme et sa famille, et mes amis,*~~

A tous les enseignants qui m'ont suivis durant ma formation,

A tous qui m'ont aidé à réaliser ce travail,

A tout qui porte le nom SAADA.

Mustapha

Table des matière

LISTE DES FIGURES	v
REMERCIEMENT	vi
Résumé	x
Notations et Acronymes	x
1 PRELIMINAIRES	3
1.1 Introduction	3
1.2 Image numérique	3
1.2.1 Types d'images	4
1.3 Segmentation d'image	4
1.4 Apprentissage automatique	5
1.4.1 Définitions	5
1.4.2 Types d'algorithmes d'apprentissage	5
1.5 Modèles de Markov	6
1.5.1 Paramètres d'un Modèle	7
1.5.2 Processus de Markov	7
1.5.3 Chaîne de Markov	8
1.5.4 Chaînes de Markov Cachées	9
1.5.5 Chaînes de Markov Couples CMCo	10
1.5.6 Chaînes de Markov Triplets CMT	11
2 Segmentation par les Chaines de Markov Cachées (CMC)	13
2.1 Introduction	13
2.2 Processus de segmentation d'image	13

2.2.1	Etape de Pré-traitement	14
2.2.2	Image en niveau de gris	14
2.2.3	Parcourt de HILBERT-PEANO	15
2.2.4	Construction de la courbe HILBERT-PEANO	15
2.2.5	Exemple de Construction de la courbe HILBERT-PEANO	16
2.3	Modélisation statistique	17
2.3.1	Décision Bayésienne	18
2.3.2	Maximum a posteriori (MAP)	19
2.3.3	Maximum A Posteriori Marginale(MPM)	19
2.4	Chaine de Markov cachée	20
2.4.1	Estimations des paramètres	21
2.5	Critère d'arrêt	24
2.6	Conclusion	24
3	Implémentation et évaluation	25
3.1	Introduction	25
3.2	Outils d'implémentation	25
3.3	Mesures et Critères d'évaluation	26
3.4	Jeu d'essai (Data set)	26
3.5	Etape de prétraitement	26
3.5.1	Image originale	26
3.5.2	Image en niveau de gris	26
3.5.3	Parcours d'Hilbert – Peano	27
3.5.4	Transformation en signale monodimensionnel	27
3.6	Exemples de segmentation d'images par modèles CMCs	27
3.6.1	Exemple 1	28
3.6.2	Exemple 2	30
3.7	Conclusion	31
	Conclusion Générale	32
	Bibliographie	33

Liste des Figures

1.1	Représentation de l'image binaire	4
1.2	Apprentissage supervisé	5
1.3	Apprentissage semi-supervisé	6
1.4	Apprentissage non supervisé	6
1.5	une petite version d'un véritable avion	7
1.6	un avion en papier	7
1.7	Andreyevich Markov	8
1.8	Représentation de la matrice transition par graphe	8
1.9	La représentation graphique de la structure de dépendance d'une chaîne de Markov cachée, où Y_k est le processus observable et X est la chaîne cachée	9
1.10	Représentation graphique de CMCo	11
1.11	La représentation graphique de la structure de dépendance d'une chaîne de Markov triplets	12
2.1	Processus de segmentation d'image	13
2.2	Exemple d'image en niveau de gris	14
2.3	Devisions espace image	15
2.4	Etape d'initialisation	16
2.5	Duplication courbe	16
2.6	étape intermédiaire	16
2.7	L'étape de parcoure final	17
2.8	Images en niveau de gris	17
2.9	Images après l'exécution du parcours d'HILBERT-PEANO	18
3.1	Images original	26
3.2	Images en niveau de gris	27

3.3	Parcours HILBERT-PEANO d'images exemple 1 et 2	27
3.4	Image et graphe initiaux en niveaux de gris exemple 1	28
3.5	Etapas segmentation EM, MPM exemple 1	29
3.6	Résultats de segmentation EM, MPM exemple 1	30
3.7	Graphe de l'image Pingouins en niveau de gris	30
3.8	Etapas segmentation EM, MPM exemple 2	30
3.9	Résultats de segmentation EM, MPM exemple 2	31

REMERCIEMENT

Au premier ~~temps~~, nous remercions Allah qui nous a aidé à réaliser ce travail, et qui a été avec nous en tout ~~le~~ moment.

Nous remercions également notre encadreur Monsieur **M. Youcef MAHDJOUB** pour l'aide et les conseils concernant les missions évoquées dans ce Projet de fin d'étude, qu'il nous a apporté lors des différents suivis.

Nous remercions également **Pr Djelloul ZIADI, M. Slimane BELLAOUAR, M. Slimane OULAD NAOUI, M. Mehamed BETKA** et tous les enseignants ~~de~~ département de Mathématiques et Informatique de notre université.

Enfin, nous adressons nos plus sincères remerciements à tous nos proches et amis, qui nous ont toujours soutenue et encouragée au cours de la réalisation de ce mémoire.

ملخص

في هذا العمل المتواضع ، تمت دراسة تجزئه الصورة بواسطة سلاسل ماركوف المخفية، حيث قدمنا بعض التعريفات للتصوير الرقمي، وعناصر التحليل المرفقة به، ودرسنا تجزئة الصورة في إطار التعلم الآلي تحت بدون الإشراف.

في البداية قدمنا التعلم الآلي وفقا لنموذج بايزي كلاسيكي ، حيث بينا مجموعه من أساليب التقدير للمعلومات مثال نموذجي : MAP (الحد الأقصى للاحق) ، MPM (الحد الأقصى لهامشييه لاحقه) في المرحلة الثانية قمنا بدراسة وفقا للنهج العشوائي لتجزئة الصورة ، ونحن نتحدث عن تجزئه بواسطة سلاسل ماركوف المخفية. قدمنا الأساليب المختلفة والخوارزميات لتقدير المعلومات مثل: EM) التقديرات العظمى)، SEM (الحد الأقصى للتقدير العشوائي)، ICE (التقدير المشروط التكراري) في النهاية ، أخذنا الخوارزميه الكلاسيكية (MPM) من النموذج البايزي ، و (EM) لسلاسل ماركوف، وكمثال على تطبيق تجزئه الصورة، ونتائج التطبيق، والملاحظات المتعلقة بالسلوك والأداء، أجرينا مقارنة بين النتائج التي تم الحصول عليها.

الكلمات المفتاحية : سلسله ماركوف، وسلسله ماركوف الخفية، تقسيم الصور ، MPM،MAP ، SEM،EM، هيلبرت-بيانو

Résumé

Dans ce modeste travail on a étudié le thème de segmentation d'image par les chaînes de Markov cachées, où on a donné quelques définitions sur l'imagerie numérique, et les éléments d'analyse lui en rattachées, on a étudié la segmentation d'image dans un cadre d'apprentissage.

Dans un premier temps nous avons présenté l'apprentissage selon le paradigme Bayésien classique, où on a cité un ensemble de méthodes d'estimation de paramètres d'un modèle exemple : MAP (**M**aximum **A** **P**osteriori), MPM (**M**aximum a **P**osteriori **M**arginal).

Dans un second temps on a entamé une approche stochastique pour la segmentation d'image, nous parlons de la segmentation par chaînes de Markov cachées. Nous avons présenté les différentes méthodes et algorithmes d'estimation de paramètres tels que : EM (**E**stimation **M**aximisation), SEM (**S**tochastique **E**stimation **M**aximisation), ICE (**I**tératif **C**onditionnelle **E**stimation).

A la fin, on a ~~pris~~ un algorithme classique (MPM) de l'approche Bayésienne, et un deuxième plus récent (EM) pour les chaînes de Markov, comme exemple d'application de la segmentation d'image, on a récupéré les résultats d'application, tirés des observations du comportement et de performance et nous avons fait une comparaison entre les résultats obtenus.

Mots-clés : Chaîne de Markov, Chaîne de Markov cachée, Segmentation d'image, [MAP, MPM, EM, SEM, ICE], HILBERT-PEANO.

Abstract

In this modest work, we have studied the subject of image segmentation using the hidden Markov chains, where we have given some definitions on digital image, and the elements of analysis attached to it, we have studied this subject in an unsupervised learning environment.

First, we presented the learning according to the classical Bayesian paradigm, in which a set of methods for estimating parameters of a model examples was given: MAP (Maximum a Posteriori), MPM (Maximum a Posteriori Marginal).

In a second step, a stochastic approach to image segmentation has been presented; we are talking about hidden Markov chains segmentation. We have presented various methods and algorithms for estimating parameters such as: EM (Expectation Maximization), SEM (Stochastic Expectation Maximization), ICE (Iterative Conditional Expectation).

At the end, a classical algorithm (MPM) of the Bayesian approach was taken, and a second more recent (EM) for the Markov hidden chain, as an example of the application of the image segmentation, we recover the results Of performance, and we have compared the results obtained.

Key words: Markov chain, Hidden Markov chain, Image segmentation, MAP, MPM, EM, SEM, ICE, HILBERT-PEANO.

Notations et Acronymes

MAP Maximum A Posteriori

MPM Maximum marginal A Posteriori

CMC Chaîne(s) de Markov cachée(s)

CMCo Chaîne(s) de Markov couple(s)

CMT Chaîne(s) de Markov triplet(s)

CMP Chaîne(s) de Markov Pairwise

EM Espérance-Maximisation (en anglais Expectation-Maximization)

SEM EM stochastique (en anglais Stochastic EM)

ECI Estimation conditionnelle itérative (en anglais Iterative conditional estimation ICE)

X,Y,U Variables aléatoires (en majuscule) des processus caché, observables et auxiliaire

x,y,u Leurs réalisations respectives en minuscules

INTRODUCTION GÉNÉRALE

Le traitement d'image est défini comme étant une discipline de l'informatique et des mathématiques appliquées dans le but d'extraction d'information et d'amélioration de la qualité d'image, ainsi que le traitement d'image peut se qualifier d'un sous-ensemble de traitement de signal. Didier aux images et vidéo.

Une image numérique est le résultat d'acquisition et de numérisation d'une image, dans le but d'aller vers une interprétation et manipulation dans des différents contextes notamment de l'intelligence artificielle et l'extraction de connaissance [1].

Pour traiter une image on doit comprendre le contexte d'une image, les modes et les conditions d'acquisition et de traitement pour aboutir à des résultats et extraire des connaissances.

C'est en 1920 que fut apparue les premières opérations de traitement d'image pour la transmission d'images de New-York à Londres par câble sous marin qui a nécessité un temps très élevé. Vers la fin de la seconde guerre mondiale et l'arrivée du radar le traitement de signal a connu plusieurs développements, mais c'est qu'après l'apparition des ordinateurs puissants et la découverte des transformées de Fourier rapide dans les années 1960 qu'a été les véritables essors du traitement d'images, alors que la première détection précise de contours n'était possible qu'en 1980 par DAVID MARR [2]. C'est la mode des systèmes experts !! Les années 1990 ont connus une grande explosion scientifique dans le domaine de traitement d'images où la recherche médicale a été le plus grand demandeur de techniques de traitement d'image, plusieurs applications ont vu le jour tel que Photoshop, Paintshop...

Une image numérique est souvent représentée sous forme d'un tableau à deux dimensions (3 dimensions pour les images 3D) chaque cellule est appelée pixel où en général une valeur que prend un pixel est un niveau de luminosité, de gris ou de couleur.

Plusieurs Operations peuvent être appliquées sur une image

- Modification pixel par un autre pixel
- Traitement de pixel (érosion, dilatation, squelettisation..)
- Calcule distances
- etc...

Il existe d'autres opérations qui consistent à segmenter et classifier des pixels d'une image selon des paramètres et des critères bien définies L'apprentissage automatique a ouvert des grandes perspectives pour le traitement et l'acquisition des connaissances, ainsi que des grande-avancées dans le domaine de traitement d'images

Plusieurs approches de traitement d'image existent voir, supervisé, non supervisé, semi-supervisé ...etc.

Dans ce qui suit en c'est intéressé beaucoup plus à la segmentation d'image, une opération qui consiste à regrouper des ensembles de pixel d'une image en des sous ensembles de pixel ayant une corrélation entres eux dans le but de définir les différents objets qui composent une image, de détecter des contours, extraire des connaissances...etc. Plusieurs travaux ont été menés sur la segmentation d'image et les modèles de Markov, des ouvrages plus adaptés au contexte de ce mémoire peuvent êtres consultés [3][4][5].

Nous avons adopté une approche d'apprentissage non supervisée, et on a utilisée une méthode stochastique (probabiliste) qui est les chaines de Markov cachées. Ce mémoire est organisé comme suite :

- Un premier chapitre préliminaires et définitions, sur l'imagerie numérique, la segmentation, l'apprentissage automatique et les modèles de Markov.
- Chapitre 2 segmentation par les chaine de Markov cachées ce chapitre débute par une approche plus généralisée pour le processus de modélisation et la segmentation statistique, et les différent méthodes et algorithmes associés a cette approche, en suite, une approche plus adaptée a la segmentation d'image et a la notion des phénomènes cachés (les chaines de Markov cachées).
- Le dernier chapitre contient une implémentation, discussion et comparaison des résultats.
- Et on a terminé notre travail par une conclusion général.

PRELIMINAIRES

1.1 Introduction

Le traitement d'images étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information.

Dans le contexte de la vision artificielle, le traitement d'images se place après les étapes d'acquisition et de numérisation, assurant les transformations d'images et la partie de calcul permettant d'aller vers une interprétation d'images traitées. Cette phase d'interprétation est d'ailleurs de plus en plus intégrée dans le traitement d'images, en faisant appel notamment à l'intelligence artificielle pour manipuler des connaissances, principalement sur les informations dont on dispose à propos de ce que représentent les images traitées (connaissance du domaine).

1.2 Image numérique

Image numérique est dite *digital image* en anglais, Numérique signifie discret

Les IMAGES numériques sont électroniques, instantanés ~~prise~~ d'une scène ou numérisées de documents, tels que photographies, manuscrits, textes imprimés et illustrations [6].

L'image numérique est échantillonnée et mappée comme une grille de points ou éléments d'image (pixels). Chaque pixel est assigné une valeur de teinte (noir, blanc, nuances de gris ou couleur), qui est représentée en code binaire (zéros et uns). Comme sur la Figure 1.1.

Les chiffres binaires bits pour chaque pixel sont stockés dans une séquence par un ordinateur et souvent réduites à une représentation mathématique (compressée). Les morceaux sont ensuite interprétés et lu par l'ordinateur pour produire une version analogique pour l'affichage ou l'impression.

1	1	1	0	1	1	1	1	1	1
1	1	0	1	0	1	1	1	1	1
1	0	1	1	1	0	1	1	1	1
0	1	1	1	0	1	0	1	1	1
1	0	1	0	1	1	1	0	0	0
1	1	0	1	1	1	0	1	1	1
1	1	1	0	1	0	1	1	1	1
1	1	1	1	0	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1

Figure 1.1: Représentation de l'image binaire

1.2.1 Types d'images

On peut coder une image de plusieurs manières, principalement, Images matricielles, images vectorielles.

1. **Image matricielle:** Elle est composée d'une matrice (tableau) de points à plusieurs dimensions, chaque dimension représentant une dimension spatiale (hauteur, largeur, Profondeur), temporelle (durée) ou autre (par exemple, un niveau de résolution)[1].
2. **Image vectorielle:** Le principe est de représenter les données de l'image par des formules géométriques qui vont pouvoir être décrites d'un point de vue mathématique. Cela signifie qu'au lieu de mémoriser une mosaïque de points élémentaires, on stocke la succession d'opérations conduisant par exemple, au tracé d'un dessin[1].

1.3 Segmentation d'image

La segmentation d'images est l'un des problèmes phares du traitement d'images.

Elle consiste à partitionner l'image en un ensemble de régions connexes. L'intérêt de ces régions est de pouvoir être manipulées ensuite, via des traitements de haut niveau pour extraire des caractéristiques de forme, de position, de taille, etc...[6].

Le problème peut être posé autrement, car on ne sait jamais dire quelle est la segmentation idéale. L'idée est bien sûr que la région se rapproche de la notion d'objet, au sens courant du terme.

Donc on peut définir la Segmentation d'une image comme étant une classification de données d'une image (pixels) et de les regrouper en sous ensemble en se basant sur certaines caractéristiques et critères sous contrainte que les classes ne sont pas connues au préalable, et qui fait appel à des méthodes heuristiques que d'utiliser des approches déterministes.

1.4 Apprentissage automatique

L'Apprentissage automatique ou artificiel (Machine Learning en anglais) est le sous-corps de l'informatique qui permet aux ordinateurs d'apprendre sans être explicitement programmé. C'est le résultat d'une évolution à partir de l'étude de la reconnaissance des formes et la théorie de l'apprentissage ~~computation intelligence artificielle~~ [7].

1.4.1 Définitions

Définition 1.4.1. :

Il est aussi défini comme étant l'exploration, l'étude et la construction des algorithmes qui peuvent apprendre et faire des prédictions sur les données [8].

Définition 1.4.2. :

L'Apprentissage automatique est étroitement lié au calcul statistique, qui se concentre également dans la fabrication de prédictions à l'aide d'ordinateurs.

Définition 1.4.3. :

Tom M. MITCHELL définit l'apprentissage automatique formellement comme [9] :

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E ".

1.4.2 Types d'algorithmes d'apprentissage

Il existe plusieurs types d'algorithmes d'apprentissage selon différents critères. Il inclut trois types fondamentaux d'apprentissage : supervisé, non supervisé, semi-supervisé

1. **Apprentissage supervisé:** Les données d'entrée sont appelées données d'apprentissage ayant un label connu ou un résultat comme par exemple spam/non-spam, voir exemple de la Figure 1.2. Un modèle est préparé grâce à un processus d'apprentissage dans lequel il est nécessaire de faire des prédictions et les corriger lorsque elles sont erronées.

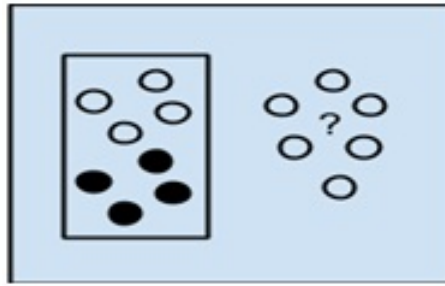


Figure 1.2: Apprentissage supervisé

Le processus de formation continue jusqu'à ce que le modèle permette d'obtenir un niveau de précision désiré sur les données d'apprentissage [10].

Plusieurs algorithmes utilisent l'apprentissage supervisé comme la régression logistique et le réseau de neurone.

2. **Apprentissage semi-supervisé:** Les Données d'entrée sont un mélange d'exemples marqués et étiquetés, Figure 1.3.

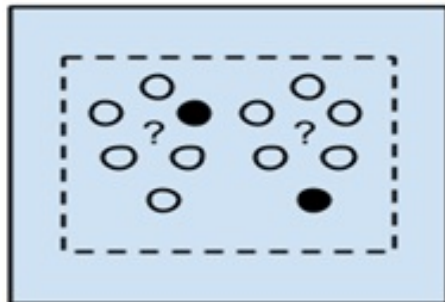


Figure 1.3: Apprentissage semi-supervisé

Il y a un problème de prévision désirée, mais le modèle doit apprendre les structures pour organiser les données mais aussi de faire des prédictions[10].

Les Algorithmes de l'exemple sont des extensions à d'autres méthodes souples, qui des hypothèses sur la façon de modéliser les données sans étiquette.

3. **Apprentissage non supervisé:** Les données d'entrée ne sont pas étiquetées et n'ont pas un résultat de classification connu voir Figure 1.4. Le modèle est établi en déduisant les structures présentes dans les données d'entrée.

Soit pour extraire les règles générales. Ou peut-être grâce à un processus mathématique à réduire systématiquement la redondance, ou il peut être d'organiser les données par similitude[10].

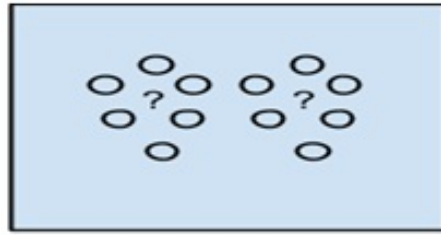


Figure 1.4: Apprentissage non supervisé

Exemples d'algorithmes utilisant l'apprentissage non supervisé comme l'algorithme Apriori et k-moyenne.

Dans notre cas la segmentation d'image rentre dans le cadre d'apprentissage non supervisé et les modèles et les approches qu'on va voir plus tard dans ce mémoire doivent être adaptés à cette approche.

1.5 Modèles de Markov

Un modèle est une représentation abstraite de la réalité.

Un modèle est une représentation abstraite de la réalité. Considérons par exemple les modèles d'avions. Certains modèles d'avions ressemblent beaucoup à une petite version d'un véritable avion comme sur la Figure 1.5, mais ne volent pas du tout. D'autres modèles d'avions (par exemple, un avion en papier) ne ressemblent absolument pas aux avions comme sur la Figure 1.6, mais volent très bien. Ces deux types de modèles représentent différentes caractéristiques de l'avion; Le premier représente son apparence extérieure, tandis que le second représente ses propriétés aérodynamiques (en partie). Ainsi, quel type de modèle est approprié d'utiliser dépend de l'objectif prévu.



Figure 1.5: une petite version d'un véritable avion

Les modèles mathématiques représentent un système et sont utilisés pour faire des prédictions sur ce système.



Figure 1.6: un avion en papier

1.5.1 Paramètres d'un Modèle

Chaque modèle se compose d'une structure, ainsi que les paramètres qui doivent être définies pour que le modèle soit significatif. La structure du modèle définit les dépendances entre les différentes parties du modèle. Les Paramètres sont souvent des valeurs, mais pas nécessairement les valeurs numériques qui sont requis par le modèle. Les Paramètres peuvent être fixés, dans ce cas constituent des hypothèses du modèle, ou ils peuvent être variable[11].

1.5.2 Processus de Markov

Un processus de Markov est un processus qui peut être dans plusieurs états, peut faire des transitions entre ces états, et dans lequel les états disponibles et les probabilités de transition dépendent seulement de l'état actuel, en d'autres termes, il n'y a pas de mémoire dans un processus de Markov [1].

Nous supposons que nous connaissons pour chaque paire d'états i et j , et pour chaque instant t , la probabilité $p_{ij}(t)$ que le processus soit dans l'état j à l'instant $t+1$ étant donné qu'il se trouve dans l'état i à l'instant t . De plus, la probabilité p_{ij} sera supposée dépendre de t , chaque processus est appelé chaîne de Markov

1.5.3 Chaîne de Markov

Une chaîne de Markov est un modèle statistique d'un système qui se déplace dans l'ordre d'un état à un autre, du nom de son inventeur Andrei Andreyevich Markov (1856-1922), dont vous voyez le photo ci-contre

Les probabilités de transition d'un état à un autre reposent uniquement sur l'état actuel. Généralement modélisé comme un processus stochastique. Une chaîne de Markov peut être décrite par une matrice de transition. On représente généralement P par un graphe orienté G dont les sommets correspondent aux m états et les arcs aux couples ordonnés d'états (i,j)

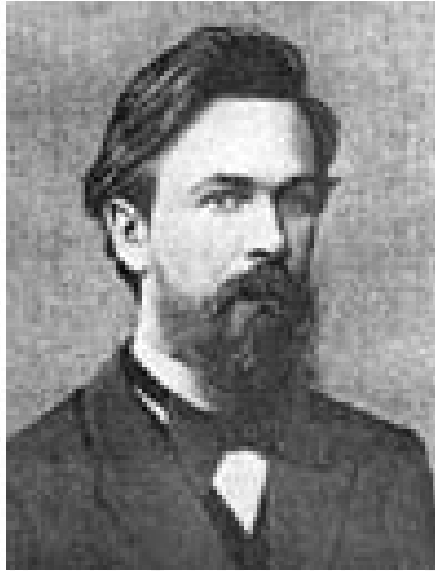


Figure 1.7: Andreyevich Markov

tels que $p_{ij} > 0$

Exemple 1.1. :

Pour représenter le passage d'une molécule de phosphore dans un écosystème, nous considérerons quatre états possibles:

1. *la molécule est dans le sol,*
2. *la molécule est dans l'herbe,*
3. *la molécule a été absorbée par du bétail,*
4. *la molécule est sortie de l'écosystème.*

La matrice de transition est la suivante:

$$p = \begin{pmatrix} \frac{3}{5} & \frac{3}{10} & 0 & \frac{1}{10} \\ \frac{1}{10} & \frac{2}{5} & \frac{1}{2} & 0 \\ \frac{3}{4} & 0 & \frac{1}{5} & \frac{1}{20} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

En remarque que la somme de chaque ligne vaut 1. Cette matrice correspond est représenté par le graphe dans la Figure au dessous:

Il existe plusieurs types de chaine de Markov selon différent principe d'algorithme, Nous allons commencer par présenter les modèles général des Chaines de Markov cachées (CMC) et couples (CMCo), Triplets (CMT)

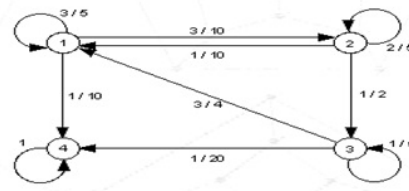
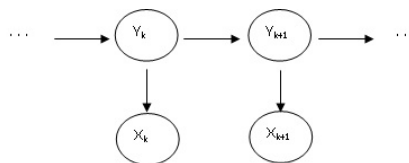


Figure 1.8: Représentation de la matrice transition par graphe

1.5.4 Chaînes de Markov Cachées

Une Chaîne de Markov Cachée (CMC) est un processus doublement stochastique avec un processus stochastique sous-jacent qui n'est pas directement observable (il est "caché") mais ne peut être observé que par un autre processus stochastique qui produit la séquence d'observations. La portée des CMC est grande et couvre une variété de situations. Pour tenir compte de ces modèles conceptuellement différents, nous définissons formellement une chaîne de Markov cachée dans la section suivante[12].

La structure de dépendance d'un CMC peut être représentée par un modèle graphique comme dans la Figure 1.9. Des représentations de ce type utilisent un graphique dirigé sans boucles pour décrire les structures de dépendance parmi les variables aléatoires. Les nœuds (cercles) dans le graphique correspondent aux variables aléatoires et les arêtes (flèches) représentent la structure de la distribution de probabilité conjointe, avec l'interprétation que ce dernier peut être pris en compte comme un produit des distributions conditionnelles de chaque nœud étant donné son Nœuds "parent" (ceux qui y sont directement connectés par une flèche).

Figure 1.9: La représentation graphique de la structure de dépendance d'une chaîne de Markov cachée, où Y_k est le processus observable et X est la chaîne cachée

Dans des cas simples tels que des modèles complètement discrets, il est fréquent de définir des chaînes de Markov cachées en utilisant le concept d'indépendance conditionnelle. En effet, une CMC a été défini comme un processus bi-varié $\{(Y_k, X_k)\}, k \geq 0$ tel que:

- $\{Y_k\}, k \geq 0$ est une chaîne de Markov avec le noyau Q de transition et la distribution initiale v
- Conditionnellement sur le processus d'état $\{Y_k\}, k \geq 0$, les observations $\{X_k\}, k \geq 0$ sont indépendantes, et pour chaque n la distribution conditionnelle de X_n dépend uniquement de Y_n .

Il s'avère que l'indépendance conditionnelle est mathématiquement plus difficile à définir dans des contextes généraux (en particulier, lorsque l'espace d'état Y de la chaîne de Markov n'est pas dénombrable), et nous adopterons un itinéraire différent pour définir les chaînes de Markov cachées généraux. La CMC est définie comme une chaîne de Markov bi-variée, mais seulement partiellement observée, dont le noyau de transition possède une structure spéciale. En effet, son noyau de transition devrait être tel que le processus de joint $\{Y_k, X_k\}, k \geq 0$ et la chaîne marginale non observable (ou cachée) $\{Y_k\}, k \geq 0$ sont Markoviens [12].

1.5.5 Chaînes de Markov Couples CMCo

Soit $Y = (Y_1, \dots, Y_n)$ et $X = (X_1, \dots, X_n)$ deux processus aléatoires, X_i est un processus caché qui prend ses valeurs à partir d'un ensemble fini de classes $\Omega = \{w_1, \dots, w_k\}$ et qui doit être estimé à partir d'un processus observable Y_i tel que $Y_i \in \mathbb{R}$. Soit $Z = (Z_1, \dots, Z_n)$, Le processus couple $Z_i = (X_i, Y_i)$ est dit chaîne de Markov couple (CMCo) . En remplace z_i par (x_i, y_i) et z par (x, y)

$$p(z) = p(z_1)p(z_2 | z_1) \dots p(z_n | z_{n-1}) \quad (1.1)$$

Sachant que les transitions (1.1) peuvent s'écrire

$$p(z_n | z_{n-1}) = p(x_n | x_{n-1}, y_{n-1})p(y_n | x_n, x_{n-1}, y_{n-1}) \quad (1.2)$$

Nous avons vu ci-dessus comment la plus grande généralité de CMCo par rapport aux CMC peut être appréhendée, dans le cas général, en comparant les transitions $p(z_i, z_{i+1})$.

Soit Z une CMCo stationnaire, donc les trois conditions suivantes

1. X est une chaîne de Markov

2. Pour tout $2 \leq i \leq n$, $p(y_i | x_i, x_{i-1}) = p(y_i | x_i)$
3. Pour tout $1 \leq i \leq n$, $p(y_i | x) = p(y_i | x_i)$

La condition (2) permet d'appréhender, dans le contexte des chaîne stationnaires, l'apport des CMC_o par rapports aux CMC localement : dès l'instant où l'on a supposé la ~~markovien~~ de processus caché X – ce qui est toujours le cas dans les différentes extensions classique des CMC, il est impossible de faire apparaître les différences éventuelles ~~des~~ $p(y_i | x_i, x_{i-1})$ lorsque x_{i-1} varie. La condition (3) permet d'aller encore plus loin dans cette voie : dans ~~un~~ CMC la loi $p(y_i | x)$ ne peut dépendre d'aucun x_j différent ~~de~~ x_i . ainsi la modélisation du bruit est bien plus complète dans les CMC_o où les lois $p(y_i | x)$ dépendent de toutes les composantes x_1, \dots, x_n de x [13].

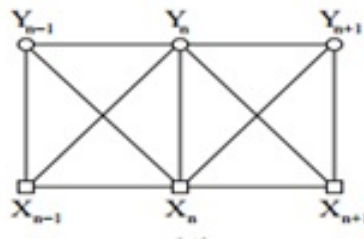


Figure 1.10: Représentation graphique de CMC_o

1.5.6 Chaînes de Markov Triplets CMT

Posant $Y = (Y_n) \quad 1 \leq n \leq N$ et $X = (X_n) \quad 1 \leq n \leq N$ deux processus stochastiques. X est caché (chaque X_n prend ses valeurs dans un ensemble fini $\Omega = \{w_1, \dots, w_k\}$), et Y est observé (chaque Y_n prend ses valeurs dans l'ensemble des nombres réels \mathbb{R}). Le problème est alors d'estimer $X=x$ de $Y=y$

Le modèle considéré s'appelle une chaîne Triplet de Markov s'il existe un processus stochastique $U = (U_n)_{n=1}^N$ Avec chaque U_n prend ses valeurs dans un ensemble ~~A =~~ $\{\lambda_1, \dots, \lambda_k\}$, tel que $T = (X, U, Y) = ((X_n, U_n, Y_n)) \quad 1 \leq n \leq N$ est une chaîne de Markov

L'idée de CMT est de considérer la distribution de $Z=(X,Y)$, qui modélise les interactions entre le processus observé et le processus recherché, en tant que distribution marginale d'une chaîne de Markov $T=(X,U,Y)$ comme sur la Figure 1.8. Les distributions $P(x_n, y_n)$, qui sont utilisés dans la restauration Maximum A posteriori Marginal MPM (voir la définition chapitre 2), sont alors calculables. En fait, ~~laissez~~ $V=(X,U)$. Comme T est

Markovien, le processus (V, Y) est une Chaîne de Markov Pairwise (CMP) et nous pouvons formuler et calculer la distribution de $P(V_n, Y)$ en utilisant les récurrences avant-arrière (Forward - Backward) [14]. Cela signifie que les distributions

$$p(x_n, y) = \sum_{u_n \in \Lambda} p(x_n, u_n, y) \sum_{u_n \in \Lambda} p(v_n, y)$$

Sont également calculables. Finalement, bien que la distribution de $Z = (X, Y)$ est Pas nécessairement Markoviennes, les distributions marginales $p(x_n | y)$ Sont calculables. Ce qui permet notamment l'utilisation de la méthode de restauration MPM Bayésienne

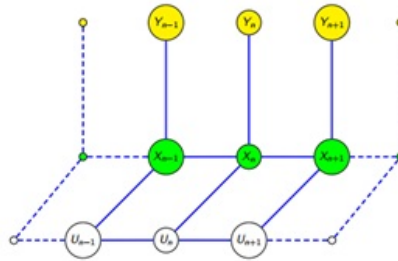


Figure 1.11: La représentation graphique de la structure de dépendance d'une chaîne de Markov triplets

Nous l'avons dit plus haut $Z=(X, Y)$ n'est pas nécessairement une chaîne de Markov. Plus précisément, le résultat suivant spécifie "localement" sur ce que la plus grande généralité de CMT à l'égard de PMC est [15] :

Proposition 1.5.1. :

Soit $T=(X, U, Y)$ un CMT vérifiant :

- a) $P(t_i, t_{i+1})$ Ne dépend pas de $1 \leq i \leq N - 1$.
- b) $P(t_i, t_{i+1}) = P(t_{i+1}, t_n)$. tel que $1 \leq i \leq N - 1$.

Ensuite, les trois conditions suivantes:

- I) $Z=(X, Y)$ est une chaîne de Markov
- II) Pour chaque $2 \leq n \leq N$, $p(u_n | z_n, z_{n+1}) = p(u_n | z_n)$
- III) Pour chaque $1 \leq n \leq N$, $p(u_n | z) = p(u_n | z_n)$

Sont équivalents.

Le modèle Chaîne de Markov Pairwise (CMP)[14],[16]), dans lequel Z est supposé être une chaîne de Markov, est alors un cas particulier du CMT obtenu en prenant $\Omega = \Lambda$ et $X = U$. En outre, selon la Proposition 1.5.1 ci-dessus, CMT est strictement plus général que

CMP, ce dernier étant strictement plus général que la CMC classique, ~~en ce qu'il~~ existe des chaînes de Markov T telles que Z ne sont pas des chaînes de Markov.

Dans cette étude le model de Markov adopté pour la segmentation d'image est la chaîne de Markov cachée.

Segmentation par les Chaines de Markov Cachées (CMC)

2.1 Introduction

Dans ce chapitre nous avons étudié quelques approches et méthodes permettant la définition d'un ~~model~~ pour segmentation d'image et l'estimation des paramètres de ce modèle et les techniques d'évaluations.

2.2 Processus de segmentation d'image

Le schéma général d'un processus de segmentation d'image suivant une approche statistique est donné par le diagramme de la Figure 2.1:

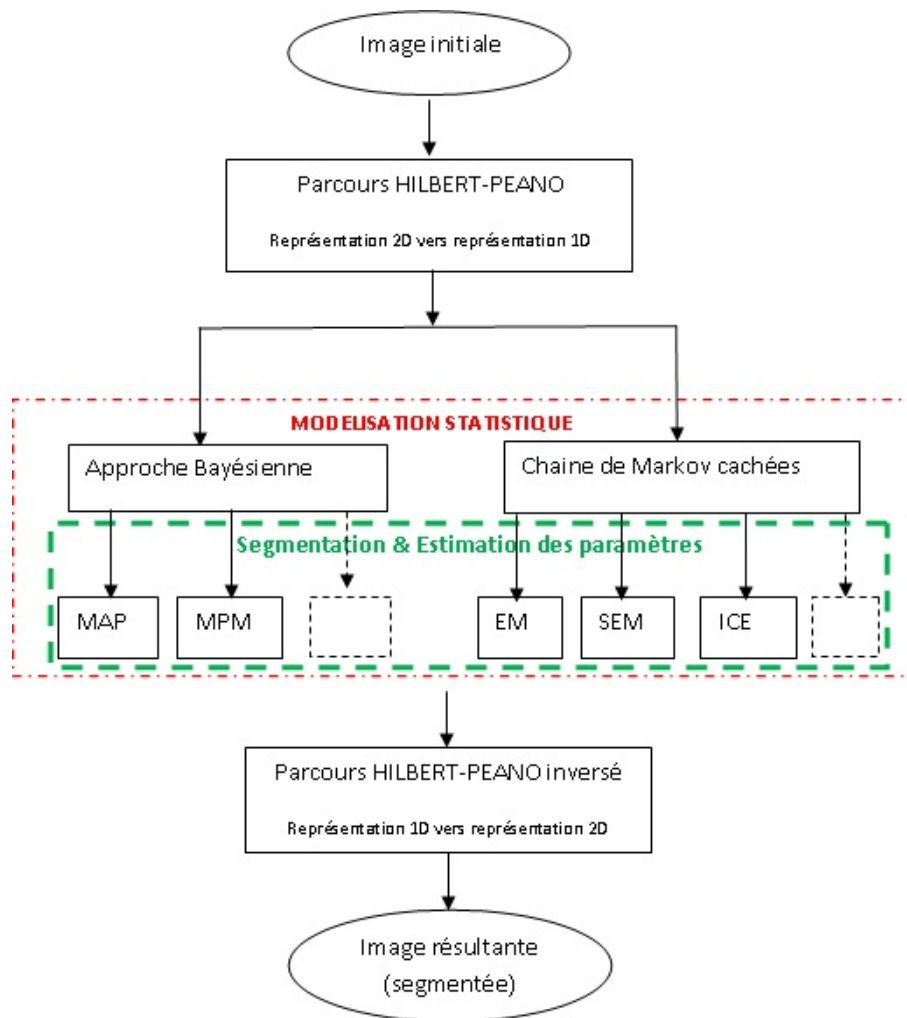


Figure 2.1: Processus de segmentation d'image

Dans un premier temps, La modélisation statistique et étudier dans un cadre général où nous avons exploré un ensemble de méthodes de modélisation et on a cité l'ensemble des paramètres de chaque ~~model~~ et les approches de définition de ces ~~paramètre~~, en parle d'une étude dans un cadre ~~Bayésienne~~.

En suite une approche plus adaptée a notre cadre d'étude (segmentation d'image) sera étudier explicitement où en va voir l'ensemble des conditions et d'hypothèses qui nous en amené à adopter cette approche.

2.2.1 Etape de Pré-traitement

Une image est stockée en mémoire pour qu'elle soit exploitée, et manipulée en appel ça numérisation d'image, plusieurs manières de codage existent où en travail toujours sur les pixels.

Parmi les types de codage en trouve :

- Matrices de points
- Codage en ligne
- Arbre quaternaire
- Codage de contours
- Palettes

2.2.2 Image en niveau de gris

Le passage de l'image originale vers une image en niveau de gris et une phase préparatoire pour la segmentation d'image, où chaque pixel aura une valeur de niveau de gris sur l'intervalle $[0-255]$ suivant l'intensité du couleur sur l'image originale.



Figure 2.2: Exemple d'image en niveau de gris

En s'intéresse beaucoup plus à la représentation d'une image (initialement sous forme d'un signal 2D) en un signal monodimensionnel 1D. chose qui est imposée par l'utilisation de chaînes de Markov, pour cela il est possible d'utiliser des techniques qui permettent ce genre de transformation, en parle ici de la courbe de remplissage de l'espace, en Anglais *space filling curve*, qui est une technique de parcourir d'un espace multidimensionnel en passant une et une seule fois par chaque point de cet espace (dans notre cas en parle d'espace 2D), il existe plusieurs méthodes de passage d'une représentation 2D vers une représentation 1D, on a utilisé le parcours d'HILBERT-PEANO comme technique de remplissage d'espace[17].

2.2.3 Parcours de HILBERT-PEANO

En 1890 Giuseppe-Peano [1858-1932] proposa une technique qui permet le parcours de l'ensemble de points d'un espace en passant par chaque point en préservant la notion de

voisinage, un an plus tard (1891) David-Hilbert[1862-1943] proposa une manière de tracer cette courbe[18].

Pour une image 2D La courbe d'HILBERT-PEANO est une courbe qui ~~parcourt~~ parcourt tout les pixels en passant une seul fois par chaque point de l'image.

La préservation de voisinage est une propriété importante assurée par la courbe d'Hilbert-Peano, le parcourt est aussi réversible c.à.d. qu'on peut revenir à l'image 2D initiale a partir de la représentation monodimensionnel, cela est nécessaire pour la reconstruction de l'image après que le traitement est achevé.

Ainsi, lorsqu'on parle d'une représentation d'une image à traité, on parle d'une représentation sous forme de chaine, Et non pas une matrice. A la fin d'un parcourt d'HILBERT-PEANO, l'image considérée comme une réalisation stochastique mono dimensionnelle $Y = (y_1, \dots, y_n)$, où N est le nombre Total de pixels dans une image et que $N = (L \times C)$ tel que L et le nombre de lignes et C représente le nombre de colonnes [18].

2.2.4 Construction de la courbe HILBERT-PEANO

Rappelons qu'un parcours d'HILBERT-PEANO est une courbe continue dans un espace de dimension n

Pour le cas d'une image 2D en parle d'un espace euclidien de dimension 2 (donc une courbe plane)

On considère un carré, divisible en quatre petits carrés isométrique. Pour construire la courbe en suit les étapes suivantes :

1. En devise l'espace (image) en quatre carrées identiques

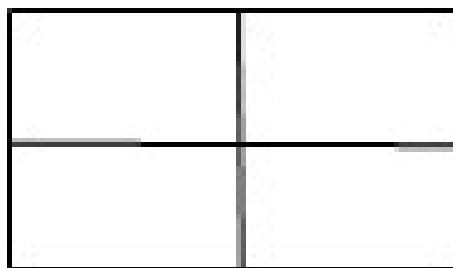


Figure 2.3: Devisions espace image

2. En trace une courbe reliant les centres des quatre carrés
3. En devise chacun des carrés en 4 autres carrés identique En copie la courbe tracées dans l'étape 2 dans chaque ensemble de quatre carrés voisins droite-haut, droite-bas

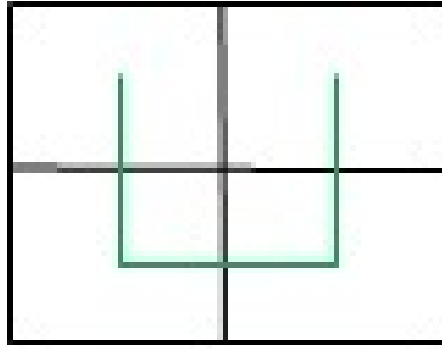


Figure 2.4: Etape d'initialisation

gauche-haut et gauche-bas mais cette fois en faisons pivoter la courbe dans les deux ensemble de carrés supérieur de tel façon qu'une extrémité coïncide avec celle de la partie inférieure

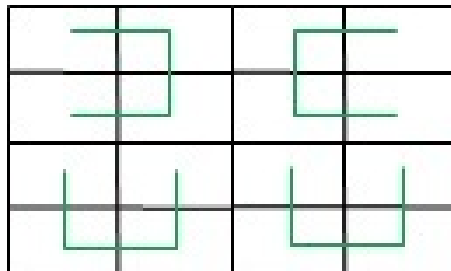


Figure 2.5: Duplication courbe

4. En relit les courbes pour qu'elles forment une nouvelle courbe continue ayant une -extrémité appelée entré E et une extrémité appelée sortie S

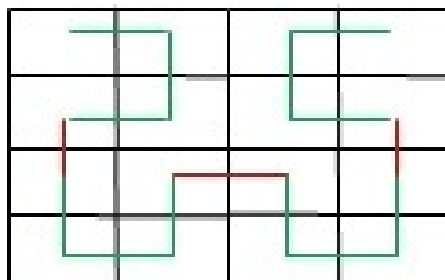


Figure 2.6: étape intermédiaire

5. On refait l'opération jusqu'a remplissement de tout l'espace pixel composant l'image 2D. Est voila en a pu représenter notre image 2D par une courbe monodimensionnel qui à chaque pixel de cordonné ligne-colonne (L, C) associe une abscisse n indiquant

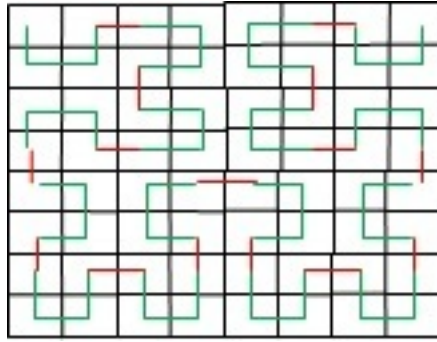


Figure 2.7: L'étape de parcoure final

le numéro d'ordre dans la courbe allons de la position initiale $(1,1)$ $n=1$ jusqu'à arriver a la position ou ce trouve l'extrémité finale $(1,n)$

Plusieurs manières de construction de la courbe existent, suivant le choix des coordonnées initiales et finales, ainsi que les algorithmes permettant l'implémentation de ces courbes.

2.2.5 Exemple de Construction de la courbe HILBERT-PEANO

- Soit l'image si dessous pour l'application du parcours d'HILBERT-PEANO



Figure 2.8: Images en niveau de gris

- après l'exécution des étapes du parcours d'Hilbert -Peano nous obtiendrons l'image suivante

Nous observant que la courbe dans la Figure 2.9 passe par tous les pixels qui composent l'image originale $(N \times N)$ de la Figure 2.8

L'opération de transformer la donnée bidimensionnelle en une autre monodimensionnelle est une phase préparatoire pour la manipulation et le traitement de la donnée qu'en va voir dans la partie suivante.



Figure 2.9: Images après l'exécution du parcours d'HILBERT-PEANO

2.3 Modélisation statistique

On peut définir Un modèle statistique comme étant une description mathématique approximative d'un mécanisme générant des observations, on parle ici d'un processus stochastique et non pas déterministe exprimé par un ensemble de distributions et d'hypothèses sur des variables aléatoires X_1, \dots, X_n . Chaque membre de la distribution est une approximation possible de F : ça consiste donc à déterminer les variables qui s'accordent le mieux avec les données.

Un modèle : est une représentation (modélisation) de la réalité faisant des hypothèses explicites sur les processus étudiés qui peuvent (les hypothèses) être fausses, et qui permet un raisonnement abstrait.

Un bon modèle est celui qui offre un bon compromis entre description juste de la réalité, la facilité de manipulation mathématique, et la production de solutions proches de l'observation. Le jugement qu'un modèle est bon ou pas nécessite une analyse de données et des tests d'hypothèses.

Une variable aléatoire est une abstraction de la donnée manipulée dans un modèle. Dans le cas d'une image on parle d'un critère relative au pixel, il peut designer l'intensité, la luminosité ou le niveau de gris dans le cas d'une image en niveau de gris, comme ça peut être un vecteur en trois composante représentant les valeurs en couleurs de base RGB (Red, Green, Blue) dans le cas d'une image couleur.

Dans la section suivante nous débutant par exposer l'approche statistique général Bayésienne pour la segmentation d'image et l'ensemble des modèles d'estimation des paramètres et par la suite nous passons vers le model Markovien, et nous exposant a la fin de cette partie les modèles d'estimation adéquate a ce model.

2.3.1 Décision Bayésienne

Segmenter une image revient à rechercher des caractéristiques cachées à partir d'un phénomène observé, et dans l'absence de liaisons déterministes entre phénomène caché et observé, on peut aboutir à plusieurs résultats et configurations de segmentation possibles pour la même observation initiale. Pour cela, la solution qui semble la plus appropriée est l'approche statistique.

L'approche Bayésienne pour la segmentation d'image modélise le lien entre le phénomène observé et le phénomène caché. Et à partir de la configuration observée une configuration optimale est reconstruite.

Soit l'image observée 'y', $y = (y_1, \dots, y_n)$ considérée comme une réalisation aléatoire Y , $Y = (Y_1, \dots, Y_n)$ où chaque pixel de l'image y est modélisé par une variable Y_i qui prend sa valeur dans \mathbb{R} . On doit estimer une réalisation dite cachée x , $x = (x_1, \dots, x_n)$ comme réalisation aléatoire de X (à construire), $X = (X_1, \dots, X_n)$ où chaque variable aléatoire X_i représentera un pixel de l'image en sortie cette variable X_i prend sa valeur dans un espace fini de cardinalité K identifié comme $\Omega = \{1, \dots, k\}$.

En utilisant le critère de maximum a posteriori, qui consiste à chercher une configuration x maximisant la probabilité de X conditionnellement à Y , notée $P(X = x/Y = y)$.

La règle de Bayes permet d'écrire

$$P(X = x/Y = y) = \frac{P(Y = y/X = x)P(X = x)}{P(Y = y)}$$

En supposons que les données (pixels) sont conditionnellement indépendantes. Alors on peut écrire :

$$P(Y = y/X = x) = \prod_s P(Y_i = y_i/X_i = x_i)$$

L'objectif est de déterminer une estimation \hat{X} qui optimise certains critères, \hat{X} est déterministe de y

Et

$$x = \phi(y) \quad \text{avec} \quad \phi : \Omega \rightarrow \Omega$$

Pour une estimation dans un cadre Bayésien :

Soit L une fonction de coût de $\Omega \rightarrow \mathbb{R}^+$

L représente le coût de remplacement de x par $\phi(y)$. (La distance entre x et $\phi(y)$)

Avec

$$L(x, x') \geq 0 \quad L(x, x') = 0 \Leftrightarrow x = x'$$

Le coût est toujours positif ou nul. Et l'objectif dans la théorie Bayésienne est d'avoir un coût minimum.

Si le coût est nul, c à d que $x = x'$.

La fonction ϕ optimale est celle minimisant l'erreur conditionnelle a y

Et l'estimateur est donnée par

$$\hat{X} = \phi(\text{opt}(y))$$

Plusieurs méthode d'estimation existent, en cite par exemple MAP (maximum a posteriori), MPM (maximum a posteriori marginale) qu'on va voir la section suivante.

2.3.2 Maximum a posteriori (MAP)

L'estimateur maximum a posteriori, ou encore appelé maximum de vraisemblances, Part du principe de pénaliser toute différence entre les deux configurations, la configuration observée et la configuration a posteriori optimal (à calculer).

Cette méthode est basée sur le principe du tout ou rien

La fonction de coût 'L' dans ce cas est définie comme suit:

$$L(x, x') = 1 \text{ si } x \neq x' \quad L(x, x') = 0 \text{ sinon}$$

La fonction ϕ_{opt} (minimise l'Esperance mathématique, et maximise la probabilité a posteriori $P(x/y)$)

$$\hat{X} = \phi_{opt}(y) = \text{Argmax}[P(x = \phi_{opt}(y)/y)]$$

En cherche alors la réalisation \hat{X} en fonction de y qui maximise la probabilité a posteriori $p(x/y)$

2.3.3 Maximum A Posteriori Marginale(MPM)

Si en part du principe de différences locales entre deux configurations (différence en chaque site d'une configuration) contrairement à la méthode MAP (tout ou rien).

La fonction coût est définie comme suit :

$$L(x, x') = \sum L(x_s, x'_s)$$

Donc c'est le nombre de différences locaux qui est pénalisé cette fois si.

En peut écrire alors :

$$L(x, x') = \sum L(x_s, x'_s) = \sum sL(x_s, x'_s)$$

La valeur optimale de $\phi(y)$ est de \hat{x} est :

$$\hat{x} = \phi_{opt}(y)_s = \text{Argmax} \phi[p(x_s = \phi(y)_s/y)]$$

En arrive donc à maximiser la marginale a posteriori $P(x_s = \phi(X_s/Y))$ locale en chaque site (pixel).

Le déroulement de l'estimation MPM est donné par l'algorithme suivant :

Algorithme MPM

Pour chaque élément de la chaîne

Pour chaque classe ω possible

1/ Forward

* Initialisation

$$\alpha_1^{(q)}(i) = \frac{\pi_i^{(q)} f_i^{(q)}(y_1)}{\sum_{1 \leq j \leq k} \pi_j^{(q)} f_j^{(q)}(y_1)} \text{ pour } : 1 \leq i \leq k$$

* Récurrence

$$\alpha_t^{(q)}(i) = \frac{f_i^{(q)}(y_t) \cdot \sum_{1 \leq j \leq k} \alpha_{t-1}^{(q)}(j) a_{ij}^{(q)}}{\sum_{1 \leq s \leq k} f_s^{(q)}(y_t) \cdot \sum_{1 \leq j \leq k} \alpha_{t-1}^{(q)}(j) a_{ji}^{(q)}} \text{ pour } : 1 \leq i \leq k, t=2, \dots, T$$

2/ Backward

* Initialisation

$$\beta_T^{(q)}(i) = 1 \text{ pour } : 1 \leq i \leq k$$

$$\varepsilon_T^{(q)}(i) = \frac{\alpha_T^{(q)}(i)}{\sum_{1 \leq j \leq k} \alpha_T^{(q)}(j)} \text{ pour } : 1 \leq i \leq k$$

* Récurrence pour $t=T-1, \dots, 1$

$$\beta_t^{(q)}(i) = \frac{\sum_{1 \leq j \leq k} a_{ij}^{(q)} f_j^{(q)}(y_{t+1}) \beta_{t+1}^{(q)}(j)}{\sum_{1 \leq s \leq k} f_s^{(q)}(y_{t+1}) \cdot \sum_{1 \leq j \leq k} \alpha_t^{(q)}(j) a_{ji}^{(q)}} \text{ pour } : 1 \leq i \leq k$$

$$\varepsilon_t^{(q)}(i) = \frac{\alpha_t^{(q)}(i) \beta_t^{(q)}(i)}{\sum_{1 \leq j \leq k} \alpha_t^{(q)}(j) \beta_t^{(q)}(j)} \text{ pour } : 1 \leq i \leq k$$

3/ Classification

Pour chaque pixel, on retient la classe qui maximise la probabilité a posteriori marginale :

$$X_t = \text{argmax}_{w_i} P(X_t = w_i / Y = y)$$

Remarque 2.3.1. :

Les estimateurs cités au dessous sont des estimateurs utilisés dans le cadre de l'apprentissage supervisé, du fait que les paramètres sont préalablement connus (nombre de classes, moyenne et matrice de covariance).

Dans un cadre d'apprentissage non supervisé, d'autres estimateurs existent. Dans ce qui suite en va étudier un ensemble d'estimateurs et voir dans quelle situation sont 'ils utilisés.

2.4 Chaîne de Markov cachée

Les chaîne de Markov caché (CMC) est un modèle statistique d'un processus consistant en deux variables aléatoires variables (dans notre cas, les variables sont discrètes) X et Y, qui changent leurs états séquentiellement. La variable X avec les états $\{x_1, \dots, x_n\}$ est appelée

”variable cachée”, car son état n’est pas directement observable. L’état de Y change séquentiellement avec un soi-disant dans notre cas, le premier ordre la propriété-Markov. Cela signifie que la probabilité de changement d’état de X dépend uniquement de son état actuel et ne change pas dans le temps. Nous écrivons sous la forme:

$$P(X_{t+1} = x_i | X_0 \dots X_t) = P(X_{t+1} = x_i | X_t) = P(X_2 = x_i | X_1).$$

La variable Y avec les états $\{y_1, \dots, y_m\}$ s’appelle la ”variable observable”, car son état peut être observé directement. Y n’a pas de propriété de Markov, mais sa probabilité d’état dépend de manière statique de l’état actuel de X [19].

Officiellement, un CMC est défini comme un tuple $M = (n, m, A, B, \pi)$, où n est le nombre d’états cachés, m est le nombre d’états observables, π est un vecteur n-dimensionnel contenant des probabilités d’état cachées initiales, A est la matrice de transition $n \times n$ dimensionnelle contenant les probabilités de transition telles que $A[i, j] = P(X_t = x_i | X_{t-1} = x_j)$ et B est la matrice d’émission $m \times n$ -dimensionnelle contenant Les probabilités d’observation tel-que :

$$B[i, j] = P(Y = y_i | X = x_j).$$

Une chaine de Markov caché (CMC) $M = (A, B, \pi)$ est défini par:

1. Un ensemble X de N états où $X = \{x_1, \dots, x_N\}$. Un état à un instant donné t est $q_t (q_t \in X)$.
2. - m symboles observables dans chaque état, $Y = \{y_1, \dots, y_m\}$. Un élément O_t de Y représente le symbole observé à l’instant t.
3. Une matrice A des probabilités de transition entre les états.

$$a_{ij} = A(i, j) = P(q_{t+1} = x_j | q_t = x_i), \quad \forall i, j \in [1..N], \quad \forall t \in [1..T]$$

4. Une matrice B de probabilités d’observation de symboles.

$$b_j(k) = P(O_t = v_k | q_t = s_j), \quad 1 \leq j \leq n, \quad 1 \leq k \leq m$$

5. Un vecteur de probabilités π représentant la probabilité initiale de chaque état.

$$\pi = \{\pi_i\}, \quad i = 1, 2, \dots, N. \quad \pi_i = P(q_1 = s_i), \quad 1 \leq i \leq n$$

2.4.1 Estimations des paramètres

L'estimation des paramètres sert à rendre les algorithmes de restauration non supervisés à savoir applicables selon les valeurs des paramètres définissant le modèle. Prenant le nombre de classes est k , donc les paramètres définissant la loi de Y sont au nombre de k . Lorsque les densités f_1, f_2, \dots, f_k sont gaussiennes, ça nous donne les paramètres définissant les lois de X conditionnelles à y sont au nombre de $2k$ (k moyennes et k variances). En principe, Ces paramètres varient avec le cas étudié, pour un type de problème donné, ils peuvent également prendre des valeurs significativement différentes.

Enfin, en notant l'ensemble des paramètres déterminant le modèle par θ , l'estimation de θ à partir de la seule réalisation $X = x$ qui est le problème préalable à la restauration. Ce problème engendré du problème général de l'estimation des composantes d'un mélange de distributions, où les dernières techniques de traitement consistent à utiliser des procédures itératives.

Dans cette section nous présentons des méthodes d'estimation des paramètres dans les chaînes de Markov cachées.

1) Algorithme Espérance-Maximisation (EM)

Expectation Maximisation (EM), proposé par Baum et al [20]. Connue sous le nom Forward—Backward Algorithm.

L'estimateur EM, est un algorithme itératif ayant le but de maximiser la vraisemblance d'observation, c'est un algorithme qui assure des bons résultats, mais l'inconvénient de cet algorithme c'est qu'il est coûteux en termes de temps d'exécution.

Dans l'algorithme EM les classes sont supposées connues et l'estimation est sensible au choix de l'initialisation **Principes** : Soit la fonction objective (les paramètres à estimer)

$$\hat{\phi} = \text{Argmax}_{\phi} P(Y = y / \phi)$$

X le champ d'étiquette caché

En considère la maximisation de l'espérance conditionnelle

$$Q(\phi / \phi^{(k)}) = E_{x/y, \phi} [\log P(X, y / \phi) / y, Q^{(k)}]$$

$\phi^{(k)}$ est le vecteur des paramètres estimés à l'étape K .

L'estimation EM est exécutée itérativement jusqu'à convergence :

→ Expectation : calcul de $Q(\phi / \phi^{(k)})$

→ Maximisation : $\phi^{(k+1)} = \text{argmax}_{\phi} Q(\phi / \phi^{(k)})$

Algorithme EM

- Initialisation des paramètres :
 $\pi_i^{(0)}, a_{ij}^{(0)}, f_i^{(0)}$ pour : $1 \leq i, j \leq k$
- A chaque itération q :
 - a. Etape E : calcul des probabilités suivantes :
 - $\alpha_t(i)$ et $\beta_t(i)$
 - Dédution des $\psi_t^{(q)}(i, j)$ et $\varepsilon_t^{(q)}(i)$ à partir de $\alpha_t(i)$ et $\beta_t(i)$ calculées sur la base des paramètres $a_{ij}^{(q)}, f_i^{(q)}$
 - b. Etape M : calcul des paramètres $\pi_i^{(q+1)}, a_{ij}^{(q+1)}, f_i^{(q+1)}$:

$$\pi_i^{(q+1)} = \frac{1}{T} \sum_{1 \leq t \leq T} \varepsilon_t^{(q)}(i)$$

$$a_{ij}^{(q+1)} = \frac{\sum_{1 \leq t \leq T} \psi_t^{(q)}(i, j)}{\sum_{1 \leq t \leq T} \varepsilon_t^{(q)}(i)}$$

$$\mu_i^{(q+1)} = \frac{\sum_{1 \leq t \leq T} y_t \cdot \varepsilon_t^{(q)}(i)}{\sum_{1 \leq t \leq T} \varepsilon_t^{(q)}(i)}$$

$$(\sigma_i^2)^{(q+1)} = \frac{\sum_{1 \leq t \leq T} (y_t - \mu_i^{(q+1)})^2 \cdot \varepsilon_t^{(q)}(i)}{\sum_{1 \leq t \leq T} \varepsilon_t^{(q)}(i)}$$

Remarque 2.4.1. Notant que cet algorithme est sensible à l'initialisation. Donc EM peut converger vers un point stationnaire (Maximum local).

En note aussi qu'il est de lenteur considérable (sauf pour des configurations de petites tailles).

Ils existent Plusieurs algorithmes (dérivant de l'algorithme EM) variantes de l'algorithme EM[21].

2) Stochastique Estimation Maximisation SEM

Stochastique Expectation Maximisation (SEM), proposé par Celeux et al [22][23]

L'algorithme SEM est une variante de l'algorithme EM, par l'adjonction d'une étape stochastique dans laquelle un tirage aléatoire est effectué. Pour résoudre les limites de l'algorithme EM où le nombre de classe supposé connu et la forte dépendance entre solution et initialisation, les améliorations qu'apporte SEM sont:

- il suffit de connaître une borne supérieure du nombre de classes K ;
- la solution est largement indépendante de l'initialisation.

L'algorithme SEM permet de résoudre le problème de convergence de l'algorithme EM, et permet d'éviter le blocage dans un maximum local en introduisant des perturbations à chaque itération par le tirage aléatoire. Ainsi que SEM converge plus rapidement par rapport à l'algorithme EM.

L'algorithme SEM réduit la difficulté de l'évaluation du terme $Q(\phi/\phi^{(k)})$ de l'étape de l'estimation

Le déroulement de l'algorithme SEM est comme suit :

- Initialisation des paramètres à estimer
- A chaque itération [k] :
 - Simulation d'une seule réalisation x de X selon la loi a posteriori basée sur les paramètres courants
 - Ré-estimation des paramètres $c_{ij} = p(s_{t+1} = w_j; s_t = w_i)$:

$$c_{ij}^{[q+1]} = \frac{1}{T-1} \mathbb{1}_{s_t=w_i, s_{t+1}=w_j}$$

(tout se passe comme si le résultat du tirage était la vraie réalisation de X) ;

- Ré-estimation des paramètres définissant la loi du bruit (μ et Γ)

3) Itératif conditionnel Expectation ICE

Itératif Conditionnel Expectation (ICE), ou 'Estimation Conditionnelle Itératif' proposé par Pieczynski [11][24][25] dont l'application aux chaînes de Markov a été décrite dans [26][27].

ICE, basé sur l'espérance conditionnelle des estimations à partir des données complètes (x, y), c'est un algorithme itératif produisant une suite d'estimation Q de paramètres, $Q \in \phi$. ICE est exécuté en trois étapes :

- L'initialisation
- L'induction
- L'arrêt

Dans la phase de l'initialisation, l'estimation de Θ est effectuée à partir des données (y, \hat{X}) , où \hat{X} est obtenu par l'application d'une méthode de classification tel que la méthode de K-moyenne classification.

2.5 Critère d'arrêt

Le critère d'arrêt d'un algorithme d'estimation de paramètre est définie selon les paramètres de la fonction d'objectifs, la qualité d'une segmentation va dépendre de cette fonction objectif. Plus la valeur d'un paramètre s'éloigne de la valeur vraie, plus la segmentation est moins correcte

2.6 Conclusion

Nous avons étudié le problème de segmentation d'image et on a devisé notre étude en deux grands volets, dans le premier nous avons présenté une étude générale suivant une approche Bayésienne où on a annoncé deux algorithmes classique d'estimation de paramètres, l'algorithme MAP, et MPM.

Dans le deuxième volet on a étudié une approche plus adaptée à la nature cachée liée à une image à prédire à partir d'une image observée, dans ce deuxième volet on a annoncé trois algorithmes EM, SEM et ICE.

Il faut mentionner que d'autres variantes de ces algorithmes et méthodes existent, ainsi que quelques méthodes combinent les deux grandes familles d'estimateurs cités au paravent exemple EM+MPM, ICE+MPM, Dont le but de cette combinaison est de donner plus de souplesse et de rapidité à l'opération d'estimation pour la segmentation d'image

Implémentation et évaluation

3.1 Introduction

Dans ce chapitre nous avons fait des expérimentations sur des images, dans le but d'aboutir à une segmentation d'image suivant quelque algorithme qu'on a déjà vu dans le chapitre précédant.

En a choisie deux algorithmes, MPM commettant un algorithme pour le paradigme Bayésienne, et l'algorithme EM pour l'estimation dans le cadre des chaines de Markov. À la fin nous allons fait une comparaison entre les résultats obtenus par les deux algorithmes.

3.2 Outils d'implémentation

Les exécutions sont faites sous MATLAB en utilisant des primitives et des fonctions de la bibliothèque de fonction MATLAB. Et on utilisé aussi quelque fragments de code open source [28].

MATLAB offre un environnement d'exécution et de test et de visualisation très riche pour le domaine de l'apprentissage automatique. Avec MATLAB, nous pouvons :

- Réduire le temps passé à accéder et pré-processer des données
- Utiliser des techniques de Machine Learning complexes sans connaissance préalable
- Avoir accès à des fonctions prêtes à l'emploi pour le traitement d'images et du signal
- Gérer des volumes de données importants, let faire du calcul parallèle.

3.3 Mesures et Critères d'évaluation

Plusieurs critères en été pris en considération pour l'évaluation de la performance et la robustesse d'un algorithme

L'observation reste toujours comme le premier et l'inévitable outil d'évaluation avant de passer aux mesures et paramètres.

On a utilisé une autre mesures d'évaluation basée sur le calcule de taux d'erreurs entre classification.

Le nombre d'itérations pour que le résultat de segmentation converge vers une valeur ou configuration stationnaire

3.4 Jeu d'essai (Data set)

Les données de test sont des simple images facile a interprété, par l'œil nu.

3.5 Etape de prétraitement

La premier phase dans l'implémentation et la phase de prétraitement comme nous l'avons déjà mentionné dans le chapitre 2 ; transformation en niveau de gris, parcours Hilbert-Peano, le passage de l'image 2D a l'image 1D (signale monodimensionnel)

3.5.1 Image originale



Figure 3.1: Images original

Les images originales qui sont initialement des images de différentes extensions (couleurs) (Figure 3.1) sont transformées en des images en niveau de gris

3.5.2 Image en niveau de gris

Nous observant la transformation des images de l'image couleur ver l'image en niveaux de gris (voir la Figure 3.2)



Figure 3.2: Images en niveau de gris

3.5.3 Parcours d'Hilbert – Peano

La Figure 3.3 montre le tracé de la courbe Hilbert- Peano pour les deux images pingouin et université de Ghardaïa



Figure 3.3: Parcours HILBERT-PEANO d'images exemple 1 et 2

Comme nous l'avons déjà vu dans le chapitre 2 (Construction de la courbe HILBERT-PEANO), ils existent plusieurs manières de construction de la courbe, et de choix des points initiale

3.5.4 Transformation en signale monodimensionnel

La Figure 3.3 montre des images où les pixels sont ordonnés en ligne et colonnes, et que la courbe tracée passe par tous ces pixels, en faisant étirer la courbe d'une façon qu'elle forme une ligne droite nous obtiendrons un vecteur où les deux extrémités coïncident avec les pixels initial et finale de l'image.

3.6 Exemples de segmentation d'images par modèles CMCs

Pour réaliser la segmentation par les chaînes de Markov Cachés CMCs, nous considérons deux images. Une première intitulée université Ghardaïa ,. L'autre appelée Pingouins , illustre bien entendue une Pingouins. C'est une image réelle. Les deux sont de taille 512x512 pixels, et sont traitées en considérant la caractéristique intensité. Signalons que la procédure de segmentation est réalisée en utilisant les deux algorithmes EM et MPM qui sera suffisamment détaillé dans le chapitre 2.

En fin une petite comparaison entre l'algorithme EM et MPM

Remarque 3.6.1. Img : l'image initiale de la segmentation.

K : défini le nombre de région de image 'img' qui segmenté.

step: Défini les nombres d'itérations on processus de segmentation.

3.6.1 Exemple 1

Segmentation l'image " université Ghardaïa " de Figure 3.2 par EM et MPM

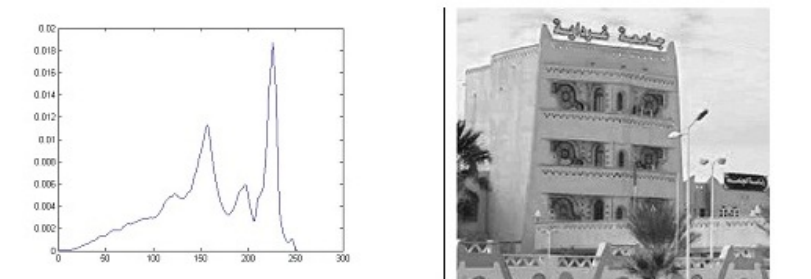


Figure 3.4: Image et graphe initiaux en niveaux de gris exemple 1

Le graphe de la Figure 3.4 décrit la répartition des pixels en niveau de gris, en remarque qu'il ya des valeurs ou sur le graphe montrant un saut de couleurs proche de la valeur 220 et la valeur 150.








K	EM	MPM
K=2 Step=5		
K=3 Step=5		
K=5 Step=5		
K=9 Step=5		
K=13 Step=5		

Figure 3.5: Etapes segmentation EM, MPM exemple 1

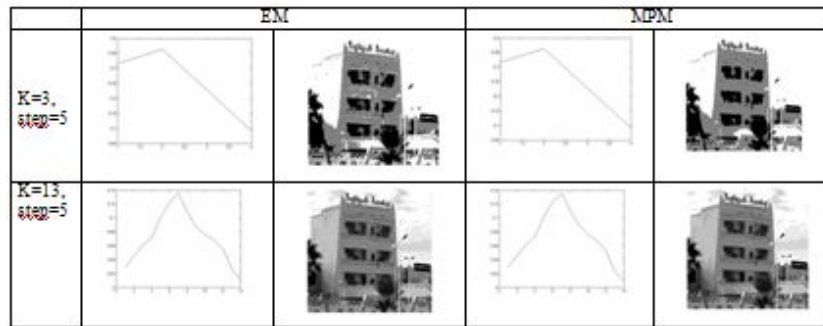


Figure 3.6: Résultats de segmentation EM, MPM exemple 1

Interprétation résultats exemple 1

Les graphes pour ($k=3, \dots, k=13$) des deux algorithmes ont donné des résultats proches, et les paramètres ont convergé rapidement vers un résultat équilibré de niveaux de gris.

Le nombre d'itération n'a pas une grande influence à partir de la valeur 5 et nous avons constaté qu'il y a pas de changement pour plus d'itérations des l'algorithme EM, MPM. Alors que le nombre de classe pour chaque exécution donne plus de détails de segmentation.

3.6.2 Exemple 2

Segmentation l'image " Pingouins " par EM et MPM

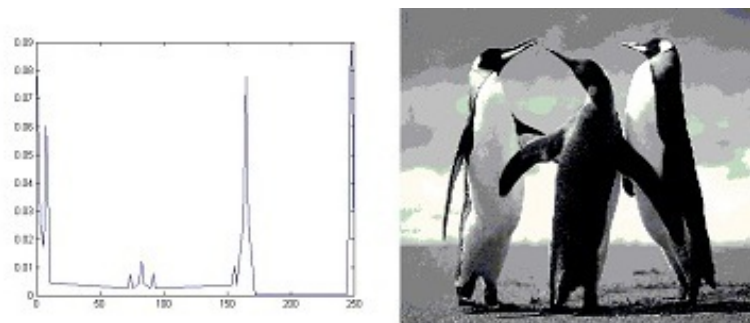


Figure 3.7: Graphe de l'image Pingouins en niveau de gris

Le graphe de la Figure 3.7 décrit la répartition des pixels en niveau de gris, en remarque qu'il ya des valeurs ou sur le graphe montrant un saut de couleurs proche de la valeur 160 et la valeur 10

Interprétation résultats exemple 2

Les graphes pour ($k=3, \dots, k=8$) les deux algorithmes ont donnés des résultats proches, et les paramètres ont convergé rapidement vers un résultat équilibré de niveaux de gris. Mais a partir de $k = 9$ nous constatons que l'algorithme MPM n'a pas donnés des résultats (image noir)

Alors que les résultats de l'algorithme EM sont acceptables Le nombre de classe, et le chevauchement des couleurs des régions sur l'image de l'exemple 2 n'ont pas été bien interprétés dans les résultats MPM

Dans cet expérimentation nous pouvons dire que l'algorithme EM a donnés un meilleur résultat de segmentation que l'algorithme MPM, mais généralement il faut prendre en considération que le nombre d'exemples n'été pas suffisent pour généraliser le jugement de performance de tel ou tel algorithme.

3.7 Conclusion

Notons que le but d'une segmentation est de prédire une optimale construction d'une image, et lorsque l'image initiale est d'une bonne qualité en observant que la segmentation que nous expérimentant essaye de revenir a chaque définition de nombre de classe vers l'image de niveau de gris initiale.

Les résultats peuvent être plus significatifs pour des images bruitées que pour des images de bonnes qualités.

Le choix du nombre de classe a affecté les résultats d'expérimentations, et un résultat meilleur nécessite la définition d'une optimale valeur pour le nombre de classe.

La contrainte de choix du nombre de classe est la principale caractéristique prise en charge dans les solutions et les améliorations apportées aux algorithmes EM, et MPM par l'introduction par exemple de calcul de k optimale par la méthode des k -means dans les algorithmes SEM et ICE

Conclusion Générale

Nous rappelons que l'objectif de ce travail été de mer une étude sur la segmentation d'image par les chaines de Markov cachées, cela nous a conduit a passer par les définitions de base de la théorie probabiliste et les fondements de la modélisation statistique chose qui nous a demandé de faire un effort plus au moins intéressent pour mieux comprendre le sujet de ce mémoire.

L'objectif sous-jacent, mais le plus intéressent dans cette étude, et d'appliquer les technique de l'apprentissage automatique dans le domaine de l'imagerie, et d'apprendre à faire une projection du monde réelle observé sur un monde discret et de savoir établir la modélisation la plus adéquate.

Nous avons réussis à connaitre et a rassembler les ficèles nécessaires pour une future compréhension plus profonde du problématique, et pour finaliser la phase d'implémentation qui reste incomplète a cause de la non maîtrise des critères d'évaluations et de mesure des performances.

Comme perspective, nous envisageons de reprendre les approches d'estimation et de segmentation SEM, ICE et de les implémenter, aussi de faire exploité d'autres solution qui existent et que nous les avons pas cité dans ce manuscrite, et d'autre algorithme hybrides pour formé une vision exhaustive sur l'apprentissage dans la segmentation d'image.










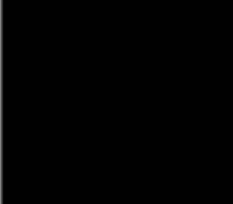
K	EM	MPM
K=2 Step=5		
K=3 Step=5		
K=5 Step=5		
K=9 Step=5		
K=13 Step=5		

Figure 3.8: Etapes segmentation EM, MPM exemple 2

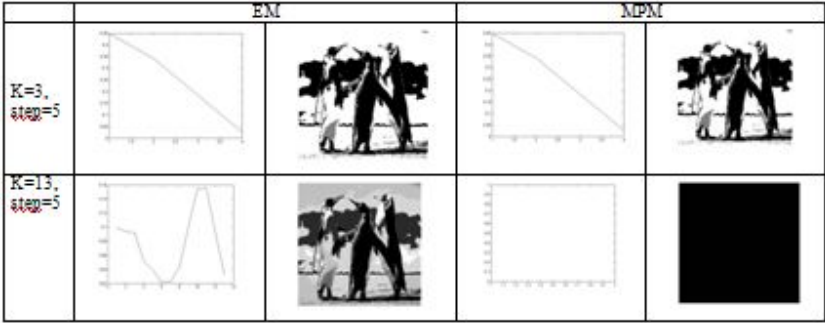


Figure 3.9: Résultats de segmentation EM, MPM exemple 2

Bibliographiques

- [1] Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing. s.l. : Pearson Education, Inc, 2008.
- [2] Théorie of Edge Détections. Hildreth, D. Marr et E. 1980, Proc. R. Soc. London, pp. 187-217.
- [3] A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes, Graphical Models and Image Processing. Heitz, C. Kervrann and F. 3, 1998, Vol. 60, pp. 173-195.
- [4] Modèles markoviens d'images et algorithmes d'estimation non linéaire sur le quadarbre, Traitement du Signal. J.-M. Laferté, F. Heitz, and P. Pérez. 3, 1998, Vol. 15, pp. 213-230.
- [5] PIECZYNSKI, Wojciech. Markov models in image processing. 2003. Vol. 20.
- [6] Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing. America : Pearson Education, Inc, 2008.
- [7] Kohavi, Ron. Glossary of Terms Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. Machine Learning. [En ligne] 1998.
<http://ai.stanford.edu/~ronnyk/glossary.html>.
- [8] Hosch, William L. Machine learning, ARTIFICIAL INTELLIGENCE. ENCYCLOPÆDIA BRITANNICA. [En ligne] 8 11 2009.
<https://www.britannica.com/technology/machine-learning>.
- [9] MITCHELL, T. M. Machine Learning. s.l. : McGraw-Hill Science, 1997.

- [10] Brownlee, Jason. A Tour of Machine Learning Algorithms. Master Machine Learning Algorithms. [En ligne] 25 11 2013.
<https://machinelearningmastery.com/master-machine-learning-algorithms/>.
- [11] Pieczynski, W. Champs de Markov cachés et estimation conditionnelle. 1994.
- [12] Olivier Cappé, Eric Moulines, Tobias Ryde'n. Inference in Hidden Markov Models. USA : Springer, 2005.
- [13] Multisensor triplet Markov chains and theory and theory of evidence. w.pieczynski. 2007, Approximate Reasoning , pp. 1-16.
- [14] Pairwise Markov chains. Pieczynski, W. s.l. : PAMI, 2003, PAMI, pp. 634-639.
- [15] Triplet Markov Chains and Theory of Evidence. Pieczynski, W. 2003, Approximate Reasoning.
- [16] SAR image segmentation using generalized Pairwise Markov chains. Pieczynski, S. Derrode and W. 2002, SPIE' s International Symposium on Remote Sensing, pp. 22-27.
- [17] K. Abend, T.J . Harley, L .N. Kanal. Classification of binary random patterns. s.l. : IEEE Transactions on Information Theory, 1965. Vol. IT-11.
- [18] Estimation conditionnelle itérative dans les chaînes de Markov cachées et segmentation statistique non supervisée d'images. B . Benmiloud, A. Peng, W. Pieczynski. 105-108, France : Juans-les-Pins, 1993, Vol. 93.
- [19] R.RABINER, LAWRENCE. A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition. s.l. : Proceedings of the Ieee, 1989.
- [20] L .E . Baum, T. Petrie, G . Soules, N . Weiss. A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. 1970.
- [21] G.Celeux, D.Chauvea and J.Diebolt. On Stochastique Version of the EM algorithm. s.l. : INRIA, 1995.
- [22] G. Celeux, J . Diebolt. L'algorithm SEM : un algorithme d'apprentissage probabiliste pour la reconnaissance de mélanges de densités. 1986.
- [23] A . Peng, W. Pieczynski. Adaptative mixture estimation and unsupervised contextual Bayesian image segmentation. 1995.

- [24] B . Braathen, W. Pieczynski, P . Masson. Global and local methods of unsupervised Bayesian segmentation of images . 1993.
- [25] Pieczynski, W. Statistical image segmentation. 1992.
- [26] B . Benmiloud, A. Peng, W. Pieczynski. stimation conditionnelle itérative dans les chaînes de Markov cachées et segmentation statistique non supervisée d'images . 1993.
- [27] M . Emsalem, H . Caillol, P. Obvié, G . Carnat, W. Pieczynski. Fast unsupervised statistical image segmentation. 1992.
- [28] Costa, Alceu. EM-MPM Image Segmentation Algorithm. MathWorks. [En ligne] 09 09 2013. [Citation : 02 05 2017.] <https://www.mathworks.com/matlabcentral/>.
- [29] Rec onnaissance de mélanges de densités par un algorithme d'apprentissage probabilistes. M. Broniatowski, G. Celeux, J. Diebolt. 1983, In Data Analysis and Informatics 3, pp. 359-373