

Hierarchical Hidden Markov Models in Image Segmentation

M. Ameer¹, C. Daoui², N. Idrissi³

University Sultan Moulay Slimane, Beni Mellal, Morocco

¹ ORCID: 0000-0003-0117-0055, ameurmeryem@gmail.com

² ORCID: 0000-0001-5435-6414

³ ORCID: 0000-0003-0038-2988

Abstract

Hidden Markov Models have been extensively used in various fields, especially in speech recognition, biology, image and signal processing and digital communication. They are well known by their effectiveness in modeling the correlations between adjacent symbols, domains or events, but they often suffer from high dimensionality problems. In this work, we propose two approaches to reduce the execution time of Hidden Markov Chain with Independent Noise used in image segmentation. The first one consists of dividing the image into blocks, each of them is treated independently of other. In the second approach, we have divided the observations into blocks, but the treatment of each block depends on its previous one. The obtained results, show that our approaches outperform standard one, and contribute efficiently to reduce the execution time and the number of iterations ensuring the convergence.

Keywords: HMC-IN, ICE algorithm, MPM estimator, divide and conquer technique, execution time, image segmentation.

1. Introduction

Image processing [1] is a scientific domain of research, which includes many mathematical applications such as filtering technique, classification of images, recognition [17], segmentation and restoration [21]. Among the aims of image processing is to improve the quality of images. It is used in many fields like medicine, chemistry. Image segmentation [20] is the most complicated task in image processing used implicitly in the majority of applications of image processing. Sometimes, it can be confused with classification. This operation allows to extract the objects and the components from the images in order to help humans to analyze the different regions of image and to extract the useful information from this image. It consists in distinguishing the pixels of the image, which have similar characteristics like color, texture, intensity into regions.

Many segmentation techniques exist in the literature [1,20,39]. Among these, we find thresholding and histogram techniques [38], that are the simplest methods of segmentation. Thresholding is based on indexing any similar region of image by a threshold, for example in binarization of images, we define two thresholds one for white color pixels and other for black color pixels. Furthermore, region approach [9] and edge approach [24] are two well known segmentation approaches. The first consists in dividing the image into homogeneous regions according to the similarity between the regions components, on the contrary, the second approach divides the image into regions according to discontinuity between the regions components. Many methods based on these approaches exist. In region approach we note region growing [9], merge and split [20]. For edge approach we have active contour [42], sobel and canny contours [15]. Some recent techniques combine region and edge approaches like mathematical morphology [4, 29,31] and watershed [6], these techniques provide better results than using region approach or edge approach only.

K-means [40] is also another segmentation method, based on clustering technique [56] that consists segmenting the image into regions by classifying the pixels of image in a suitable membership class satisfying some criteria such as distance between pixels, degrees of membership, number of neighbor pixels. The choice of a segmentation method depends on many selection criteria for example the nature of the image, dominant information, characteristics of objects to be extracted.

In this work, we consider a hidden Markov model [32], to segment images this model is an unsupervised statistical method of segmentation [22], based on the clustering technique. It is used in many applications such as medical images [35], sonar images [10,11], satellite and teledetection images [30]. We distinguish three type of Markovian models: Fields of Markov, Markov Chains and Trees of Markov. Each Markov model has its own procedure to model the image before segmenting it.

Hidden Markov Fields(HMF) models [50] have the principle to decompose the image into blocks of pixel neighbor, each block must contain four or eight pixel neighbor at the least. Each block represents a set of click of different orders. This Markovian model gives good results of segmentation, but it requires an important execution time and execution unit(CPU) compared to other Markovian models [8,50]. Markov chain transforms the image 2D shape in 1D shape respecting that each two pixel neighbor in the image 2D must remain neighbor in the chain 1D. There exist many paths to transform the image 2D in chain 1D, like Hilbert Peano path[23], zigzagging, line by line path. The trees of Markov are a particular case of Markov chains, we can use quadtree representation [13], or bitree representation, or other tree representations to convert an image to tree shape. Additionally to these models, we found the recently models of Markov, which are a generalization of classical models[49], such as pairwise models [8,33], triplet models [50] and evidential models [28]. These models are used in the modeling problems that, the classical models are unrealizable to resolve them. Another representation of these models is a fuzzy version [7,14], this model is used to segment the fuzzy images.

The main problem of hidden Markov models in image segmentation is the numerical complexity. Normally, the dimensions of segmented image are very large and so the classical methods are impracticable. For this reason, we have proposed two approaches based on the technique of divide and conquer. The first approach is a traditional method of dividing the image into blocks, each sub-block is segmented independently of other using HMC-IN. We consider that each block composing the image is treated as a sub-image. The second proposition consists in dividing the image into a set of blocks, the treatment of each sub-block depends on the result of its previous block. In order to reduce the execution segmentation time, we have also proposed two variants of the former approach, each variant has its principle and procedure. We have used ICE (Iterative Conditional Estimation) algorithm and MPM (Marginal Posteriori Mode) algorithm [18] to estimate the parameters of the HMC model.

The outline of this paper is as follows : Section 2 presents related works. Section 3 presents the standard hidden Markov chain HMC-IN, ICE and MPM algorithms. Section 4 explains our proposed approaches. Section 5 shows the conducted experiments and the obtained results. And finally, the last section gives a conclusion and addresses some open questions.

2. Related Works

Recently several techniques and researches have been carried out to segment images semantic and fast. Among these, it exist **graph cut** and **Convolutional Neural Networks(CNN)** these methods have been used to segment many different types of images like color, grey level, 2D, 3D, satellitefrom different dataset. In [54] the authors have described a novel framework for efficient object extraction form N-D image data using s/t graph cut. The same author have proposed in [53] a new implementation of Max-Flow/ Min-Cut algorithm using by graph cut to segment, restore and stereo images, to evaluate the performance of this proposition, the new algorithm has been compared by others graph algorithms in term of

running time, from the results this proposition minimizes energy faster than another algorithm. In [45] the author has proposed a new version of recently Generative Adversial Neural Network (GAN) to segment some multispectral satellite images. This version is called by conditional GAN. The same author has developed in another work [46] a new GAN network called HydroGAN, this model has used to labeling the hydrographic region in satellite imagery. The developed HydroGAN is capable to labeling the water objects in different seasons. In another work the authors [24] have proposed a new version of mobile architecture MobileNetV2 basing on inverted residuals and linear bottlenecks to segment images to detect objects. [16] have studied the high resolution representation of High Resolution Network (HRNet) by introducing a simple modification, the authors here have augmented the high resolution representation by aggregating the representations form all the parallel convolutions. This developed network is named HRNetV2, this network is applied to segment facial landmark detection. Another approach in [44] has proposed a new framework of deep convolutional encoder/decoder architecture for images segmentation named SegNet[34], here the authors have compared this new contribution with other existing architectures of Neural Network. Moreover, the authors of [43] have presented a new parsing task Unified Perceptual Parsing improving a multitasking network called UPENet with hierarchical structure, this new framework has been applied to segment some heterogeneous images.

3. Hidden Markov model in image segmentation

In this section, we present a classical Markovian model used to segment images, called Hidden Markov Chain with Independent Noise (HMC-IN). This model doesn't take into account noisy observations, but it is too effective to segment denoised images. Moreover, we expose the procedure of the ICE estimator, and we explain the principle of estimating the resulted segmented image according to MPM algorithm.

3.1. HMC-IN segmentation principle

In Markovian segmentation, the image to be segmented is represented by two random variables $X = \{x_1, \dots, x_N\}$ and $Y = \{y_1, \dots, y_N\}$, N is the total number of pixels. We consider that $Y \in \mathbb{R}$ is the observations. And $X \in \Omega = \{\omega_1, \dots, \omega_K\}$ is the result of an image segmentation. Where Ω is the set of membership classes and K is a number of membership classes, it's initialized by the user. Generally, hidden Markov model estimates the hidden image X from the observations Y , for that, Markovian model calculates the *a posteriori* probabilities of X knowing the observations Y using Bayes theorem [5]:

$$P(X | Y) = \frac{P(Y | X)P(X)}{P(Y)} \quad (1)$$

Where :

- $P(Y | X)$ is the probabilities of observations Y conditionally to X .
- $P(X)$ is the *a priori* probabilities of X .
- $P(Y)$ is a normalization constant $P(Y) = 1$.

HMC-IN assumes that the hidden process X is a Markov chain, it's homogenous and stationary of order 1, its law is:

$$P(X) = P(X_1 = x_1) \cdot \prod_{n=1}^{N-1} P(X_{n+1} = x_{n+1} | X_n = x_n) \quad (2)$$

The process X has two parameters θ_x the initial law $PI(i) = p(x_1 = i)$, $i \in \Omega$ and the matrix of transition between classes i and j $A(i, j) = p(x_{n+1} = j | x_n = i)$ $i, j \in \Omega$.

Also, HMC-IN assumes that the observations Y are conditionally independent of X , each observation $y_n, n \in N$ depends only on its hidden class of membership x_n .

$$P(Y_n = y_n | X) = P(Y_n = y_n | X_n = x_n) \quad (3)$$

The parameters of observations θ_y depend on the law of probability followed. In this work, we assume that our observations follow the gaussian law so, the process Y is defined by a gaussian density f in each class i :

$$P(Y_n = y_n | X = i) = f_i(y_n) = \frac{1}{\sqrt{2\pi(\sigma_i)^2}} \exp\left[-\frac{(y_n - \mu_i)^2}{(\sigma_i)^2}\right] | i \in \Omega. \quad (4)$$

with the mean $\mu^{(i)}$, and the variance $(\sigma(i))^2$.

HMC-IN has two type of parameters $\theta = (\theta_x, \theta_y)$ the parameters of the process X : $\theta_x = (PI, A)$ and the parameters of observations Y : $\theta_y = (\mu, (\sigma)^2)$.

To obtain the resulting image X , it should follow three phases: Initialization phase, Iterative estimation phase and Final decision phase. In the first phase, we initialize the initial configuration of X^0 process using K-means algorithm or FCM algorithm [55] or another algorithm of segmentation based on the clustering technique. Also, we initialize the parameters of each process X and Y . In the second phase, we estimate the parameters of X and Y iteratively until convergence using estimator iterative algorithms like EM(Expectation-Maximization) [3], ICE [48], SEM(Stochastic Expectation-Maximization) [51]. In the thrid phase, we estimate the final configuration of the segmented image X using Bayesian decision strategies as Viterbi algorithm [12] or MPM(Marginal Posteriori Mode) estimator. In the following paragraph, we briefly present Baum Welch algorithm, ICE and MPM algorithms.

3.2. Estimator algorithms

The ICE algorithm, introduced in [47], is an iterative method of estimation based on the Monte Carlo approximation method [27]. In each iteration, ICE simulates the hidden process X conditionally to the observations Y . This algorithm uses the deterministic strategy to calculate the parameters of the hidden process X , and the stochastic strategy to estimate the parameters of observations Y .

After parameters initialization $\theta^0 = (PI^0, A^0, \mu^0, (\sigma^0)^2)$, ICE algorithm calculates the parameters of each process $\theta^q = (PI^q, A^q, \mu^q, (\sigma^q)^2)$ for a number of iterations Q until convergence. The parameters estimation can be stopped according to a chosen stopping criterion adapted to each case. In the particular situation, this criterion may be based, for example, on the convergence of one of the estimated parameters. ICE algorithm uses Baum Welch algorithm[26] to estimate the parameters.

Baum Welch algorithm proceeds as follows:

1- Calculating the Forward probabilities $\alpha_n(i) = p(y_1, \dots, y_n, x_n)$ using Forward algorithm:

Algorithm1: Forward algorithm

$$\text{Initialization: } (n = 1) \quad \alpha_1(i) = \frac{PI(i)f_i(y_1)}{\sum_{j \in \Omega} PI(j)f_j(y_1)}, i \in \Omega \quad (5)$$

$$\text{Induction: } (n > 1) \quad \alpha_{n+1}(i) = \frac{f_i(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j)A(i, j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j)A(i, j)}, i \in \Omega \quad (6)$$

2- Calculating the Backward probabilities $\beta_n(i) = p(y_{n+1}, \dots, y_N | x_n)$ following the steps of Backward algorithm:

Algorithm2: Backward algorithm

Initialization: ($n = N$) $\beta_N(i) = 1, i \in \Omega$ (7)

$$\text{Induction : } (n < N) \quad \beta_n(i) = \frac{\sum_{j \in \Omega} A(i, j) f_j(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)}, i \in \Omega \quad (8)$$

3- After calculating $\alpha_n(i)$ and $\beta_n(i)$, Baum Welch estimates the joint *a posteriori* probabilities $\gamma_n(i, j) = p(x_n = i, x_{n+1} = j | y_n)$ and the Marginal *a posteriori* probabilities $\xi_n(i) = p(x_n = i | y_n)$ from $\alpha_n(i)$ and $\beta_n(i)$ using the following formulas :

$$\gamma_n(i, j) = \frac{\alpha_n(i) \cdot A(i, j) \cdot f_j(y_{n+1}) \cdot \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \cdot \sum_{l \in \Omega} \alpha_n(l) \cdot A(l, k)}, i, j \in \Omega \quad (9)$$

$$\xi_n(i) = \alpha_n(i) \cdot \beta_n(i), i \in \Omega \quad (10)$$

The procedure of ICE algorithm is explained in the following algorithm:

Algorithm 3: ICE algorithm

For each iteration $q \in Q$:

- Calculating the probabilities $\alpha^q, \beta^q, \gamma^q$ and ξ^q using Baum Welch algorithm;
- Simulating the process X^q for one random simulation according to the *a posteriori* probabilities $P(X | Y)$ calculated in the iteration q ;
- Estimating the parameters θ_x^{q+1} of hidden process X by the deterministic strategy:

$$PI^{q+1}(i) = \xi_1^q(i). \quad (11)$$

$$A^{q+1}(i, j) = \frac{\sum_{n=1}^N \gamma_n^q(i, j)}{\sum_{n=1}^N \xi_n^q(i)}. \quad (12)$$

- Estimating the parameters θ_y^{q+1} of observations Y for each class $i \in \Omega$ by the stochastic strategy based on the simulated process X^q :

$$\mu_i^{q+1} = \frac{\sum_{n=1}^N y_n 1[x_n^q = i]}{\sum_{n=1}^N 1[x_n^q = i]}. \quad (13)$$

$$(\sigma_i^{q+1})^2 = \frac{\sum_{n=1}^N (y_n - \mu_i^{q+1})^2 1[x_n^q = i]}{\sum_{n=1}^N 1[x_n^q = i]}. \quad (14)$$

- Calculating the gaussian density f^{q+1} of observations using the equation(4);
-

After the convergence of the algorithm 2.2, HMC-IN estimates the final segmented image X using MPM estimator. This estimator calculates the marginal *a posteriori* probabilities $\xi_n(i)$ from the final obtained parameters, basing on Forward Backward algorithm. To find the membership class of each observation, MPM maximizes these probabilities, by the following equation:

$$\bar{\mathbf{x}} = \arg \max_{\mathbf{i} \in \Omega} ((\xi_n^Q(\mathbf{i}))), \mathbf{i} \in \Omega \quad (15)$$

The following schema describes the different steps to estimate parameters and image result of segmentation, that we have followed in this work:

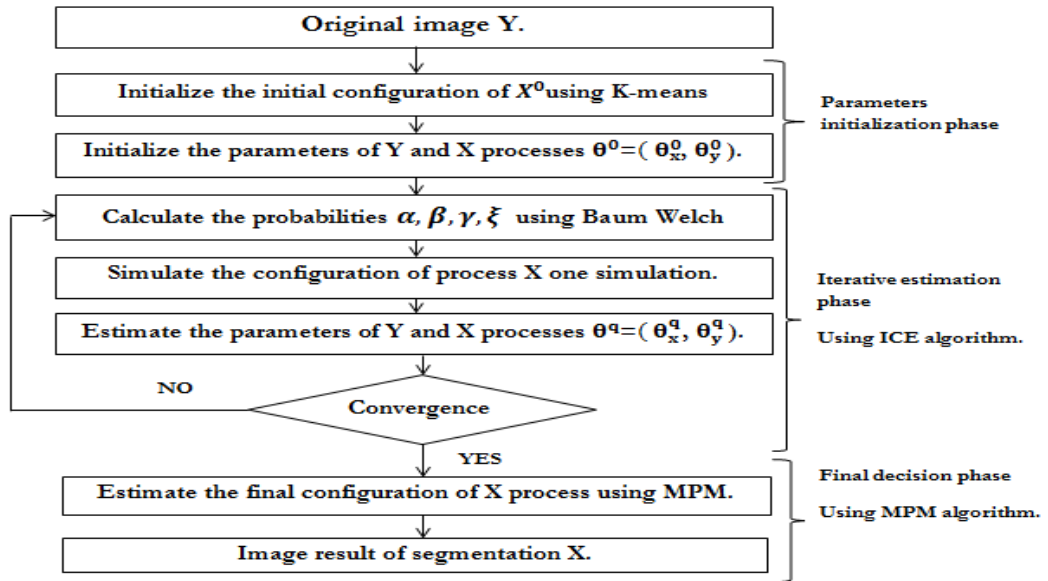


Fig.1. HMC-IN segmentation Phases

4. The proposed approaches

As cited above, hidden Markov chain is a robust model among Markovian models for image segmentation, but its drawbacks is in the high complexity of calculating parameters, especially when the size of a given image is very large. In fact, working in the real time with large size of images and large number of images in database makes the execution time of segmentation image task more and more slowly. So, we propose some improved HMC models for reducing time execution based on the technique of divide and conquer [36,37]. The main idea of these approaches is to divide the process of estimating and calculating HMC parameters into several sub-estimating processes resolved independently. Then, we combine all sub-optimal solutions to generate the global final solution. In the first one, we divide the observations Y into M blocks of the same size, each block is treated independently of the others. In the second one, also we divide the observations into M blocks of the same size, but the treatment of a block depends on the previous results, the solution of each current block is obtained by use the solution of its previous block. This technique permits to reduce the complexity of calculus compared to the first approach. The following sub-sections explain the process of the proposed approaches.

4.1. The First Approach : Independent Estimation

As a traditional approach of division, the independent estimation approach divides the global data to be treated into blocks(sub-processes). Each sub-block is solved independently of the others, and then we combine the solutions of blocks in order to determine the global solution of the initial problem.

Let $Y = \{Y_1, Y_2, \dots, Y_M\}$ is a set of blocks and $S_G = \{S_1, S_2, \dots, S_M\}$ is a set of solutions of blocks, where : S_G is the global solution, M is the number of blocks, and div is the size of a block. The figure 2 shows the scenario of the independent approach.

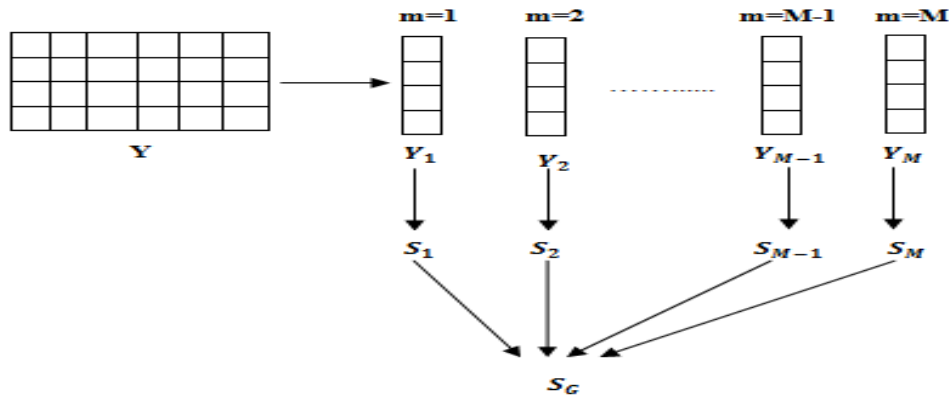


Fig.2. Scenario of independent estimation approach

The observations Y are divided into a set of equal sub-blocks. The number of blocks M is initialized by the width of image or the length of image or another divider of the observations size. To estimate the parameters of each block, we follow the same procedure as the standard HMC approach. Each block is treated like an independent HMC. Finally, we combine the final configurations of blocks to build the global final configuration of $X = \bigcup_{m=1}^M X_m$. The steps followed in the first approach are shown in the following algorithm:

Algorithm 4: Independent estimation approach steps

- 1- Initializing the hidden process X^0 using K-means method;
 - 2- Initializing the parameters of each process $\theta^0 = (\theta_x^0, \theta_y^0)$;
 - 3- Transforming the image Y into Markov chain using line by line path method;
 - 4- Dividing the obtained chain into M equals sub-blocks;
 - 5- For each block, calculating the probabilities of Baum Welch, estimating the parameters and simulating the process X until convergence, estimating the final configuration;
 - 6- Combining the final configurations of blocks and building the segmented image X .
-

Algorithm 5 illustrates the different steps of the independent approach; we have just added a loop to the original algorithm of estimation, that allows executing all blocks independently.

Algorithm 5: Independent approach

Input : $\theta^0 = (PI^0, A^0, \mu_i^0, (\sigma_i^0)^2)$

Out put : the global final configuration X

Initialization : $m = 1$ ($((m \leq M))$)

Initialization : $q = 0$

For each iteration $q \in Q$:

Calculating $\alpha, \beta, \xi, \gamma$ using Baum Welch algorithm :

Initialization: ($n = 1$)

$$\alpha_1^q(i) = \frac{PI^q(i)f_i^q(y_1)}{\sum_{j \in \Omega} PI^q(i)f_j^q(y_1)} \forall i \in \Omega \quad (16)$$

Induction: ($(n > 1)$ and $(n < div)$)

$$\alpha_{n+1}^q(i) = \frac{f_i^q(y_{n+1}) \sum_{j \in \Omega} \alpha_n^q(j) A^q(i, j)}{\sum_{k \in \Omega} f_k^q(y_{n+1}) \sum_{j \in \Omega} \alpha_n^q(j) A^q(i, j)} \forall i \in \Omega \quad (17)$$

Initialization: ($N = div$)

$$\beta_N^q(i) = 1 \forall i, j \in \Omega \quad (18)$$

Induction : ($n < div$)

$$\beta_n^q(i) = \frac{\sum_{j \in \Omega} A^q(i, j) f_j^q(y_{n+1}) \beta_{n+1}^q(j)}{\sum_{k \in \Omega} f_k^q(y_{n+1}) \sum_{j \in \Omega} \alpha_n^q(j) A^q(i, j)} \forall i, j \in \Omega \quad (19)$$

$$\xi_n^q(i) = \alpha_n^q(i) \beta_n^q(i) \forall i, j \in \Omega \quad (20)$$

$$\gamma_n^q(i, j) = \frac{\alpha_n^q(i) A^q(i, j) f_j^q(y_{n+1}) \beta_{n+1}^q(j)}{\sum_{k \in \Omega} f_k^q(y_{n+1}) \sum_{l \in \Omega} \alpha_n^q(l) A^q(l, k)} \forall i, j \in \Omega \quad (21)$$

Simulating the hidden process X_m for the current block Y_m one simulation; Calculating the parameters θ_x^{q+1} by:

$$PI^{q+1}(i) = \xi_1^q(i) \forall i \in \Omega \quad (22)$$

$$A^{q+1}(i, j) = \frac{\sum_{n=1}^N \gamma_n^q(i, j)}{\sum_{n=1}^N \xi_n^q(i)} \forall i, j \in \Omega \quad (23)$$

Calculating the parameters θ_y^{q+1} of observations Y by :

$$\mu_i^{q+1} = \frac{\sum_{n=1}^N y_n 1[x_n^q = i]}{\sum_{n=1}^N 1[x_n^q = i]} \forall i \in \Omega \quad (24)$$

$$(\sigma_i^{q+1})^2 = \frac{\sum_{n=1}^N (y_n - \mu_i^q)^2 1[x_n^q = i]}{\sum_{n=1}^N 1[x_n^q = i]} \forall i \in \Omega \quad (25)$$

For ($i = 1; i = K; i++$)

For ($n = 1; n = div; n++$)

Calculating the gaussian density $f_i^q(y_n)$ for all pixel belongs to current block using the equation4;

$q = q + 1$

Estimating the final configuration of current sub-block X_m using MPM algorithm.

$div = m * div$

$m = m + 1$

combining the final configurations of sub-block X_m to build the final global segmented X

Remark:

1. Note that, the division of observations Y into blocks is done after, transforming the observations $2D$ shape into vector $1D$ (chain) using line by line path. Then, we have divide the vector $1D$ into M blocks of the same size.
2. Since each block is treated independently of the others, parallelism technique can be used to further alleviate the computations.

4.2. The Second Approach : Dependent Estimation

Contrary to the first approach, the second approach considers that the solution of each data block depends on the solution of its previous block. The solutions of blocks are combined to build the final solution of the global problem. The following figure illustrates the scenario of the dependent approach:

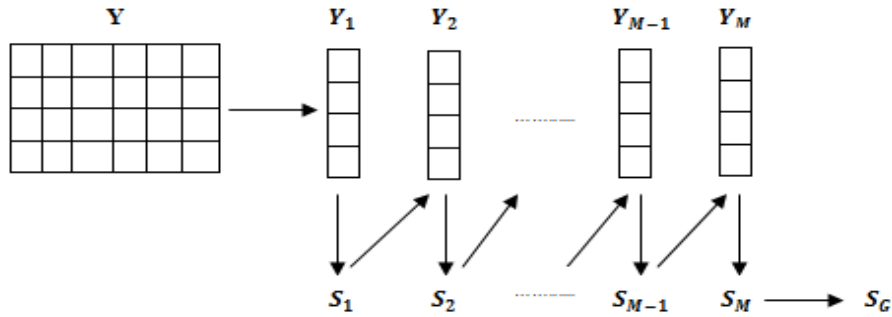


Fig.3. Scenario of dependent estimation approach

In this approach, we start by dividing the image Y into blocks(levels), each level contains one block. The calculus of probabilities of a block depends on the result of its previous block. After calculating the probabilities of all blocks, we recombine these probabilities to calculate the parameters of each HMC's process, we repeat this calculus until the convergence of the ICE algorithm. Finally, we combine the obtained final parameters to estimate the final configuration of the global process X using MPM algorithm.

Algorithm 6 describes the steps followed by the second approach:

Algorithm 6 : Dependent estimation approach steps

- 1- Initializing the hidden process X^0 using K-means method;
 - 2- Initializing the parameters of each process $\theta^0 = (\theta_x^0, \theta_y^0)$;
 - 3- Transforming the image Y into Markov chain using line by line path method;
 - 4- Dividing the obtained chain into M equals sub-blocks;
 - 5- Calculating the probabilities of each current sub-block m from the probabilities of the
-

previous sub-block $m-1$.

6- Combining these probabilities, and, estimating the parameters of the HMC-IN, repeating these phases, until convergence;

7- Estimating the final global configuration of X .

Considering that the execution time of the parameters estimation phase is the most costly, we have tried to make a modification in this phase. To do this, we have proposed an adapted process of estimation that consists in estimating the parameters α , β , γ , ξ of each block, and at the same time simulating the process X_m .

The procedure of estimation is illustrated in the following algorithm:

Algorithm 7: Dependent approach

Input : M is the number blocks;

m is the current block number ;

Initialization : $m = 1$

While ($(m \leq M)$) :

If ($m == 1$) :

Initialization: ($n = 1$)

$$\alpha_1(i) = \frac{PI(i)f_i(y_1)}{\sum_{j \in \Omega} PI(i)f_j(y_1)} i \in \Omega \quad (26)$$

Induction: ($(n > 1)$ and $(n < div)$)

$$\alpha_{n+1}(i) = \frac{f_i(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)} i, j \in \Omega \quad (27)$$

Initialization: ($N = div$)

$$\beta_N(i) = 1 i \in \Omega \quad (28)$$

Induction : ($n < div$)

$$\beta_n(i) = \frac{\sum_{j \in \Omega} A(i, j) f_j(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)} i, j \in \Omega \quad (29)$$

$$\xi_n(i) = \alpha_n(i) \beta_n(i) i \in \Omega \quad (30)$$

$$\gamma_n(i, j) = \frac{\alpha_n(i) A(i, j) f_j(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{l \in \Omega} \alpha_n(l) A(l, k)} i, j \in \Omega \quad (31)$$

Simulating the hidden process X_1 for the first block one simulation;

$m++$;

If ($m > 1$)

Initialization: ($n = div + 1$)

$$\alpha_n(i) = \frac{f_i(y_n) \sum_{j \in \Omega} \alpha_{div}(j) A(i, j)}{\sum_{k \in \Omega} f_k(y_n) \sum_{j \in \Omega} \alpha_{div}(j) A(i, j)} \quad i \in \Omega \quad (32)$$

Induction: $((n > (div+1)) \text{ and } (n < (m * div)))$

$$\alpha_{n+1}(i) = \frac{f_i(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)} \quad i \in \Omega \quad (33)$$

Initialization: $(n = (m * div))$

$$\beta_n(i) = \frac{\sum_{j \in \Omega} A(i, j) f_j(y_n) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_n) \sum_{j \in \Omega} \alpha_n(j) A(i, j)} \quad i, j \in \Omega \quad (34)$$

Induction : $(n < (m * div))$

$$\beta_n(i) = \frac{\sum_{j \in \Omega} A(i, j) f_j(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{j \in \Omega} \alpha_n(j) A(i, j)} \quad i, j \in \Omega \quad (35)$$

$$\xi_n(i) = \alpha_n(i) \beta_n(i) \quad i, j \in \Omega \quad (36)$$

$$\gamma_n(i, j) = \frac{\alpha_n(i) A(i, j) f_j(y_{n+1}) \beta_{n+1}(j)}{\sum_{k \in \Omega} f_k(y_{n+1}) \sum_{l \in \Omega} \alpha_n(l) A(l, k)} \quad i, j \in \Omega \quad (37)$$

Simulating the hidden process X_m of the block m one simulation;
 $div = m * div$
 $m++$

After calculating the probabilities α , β , ξ , γ and simulating the sub-blocks of hidden process $X = \cup_{m=1}^M X_m$ using the algorithm 7, we combine these probabilities to calculate the global parameters of HMC $\theta^q = (\theta_x^q, \theta_y^q)$. We test the convergence of ICE algorithm, if it achieves the convergence, we take the final obtained parameters to estimate the final configuration of the global process X using MPM algorithm, else we recalculate the probabilities and the parameters until convergence. The procedure of the second approach is explained in figure 4:

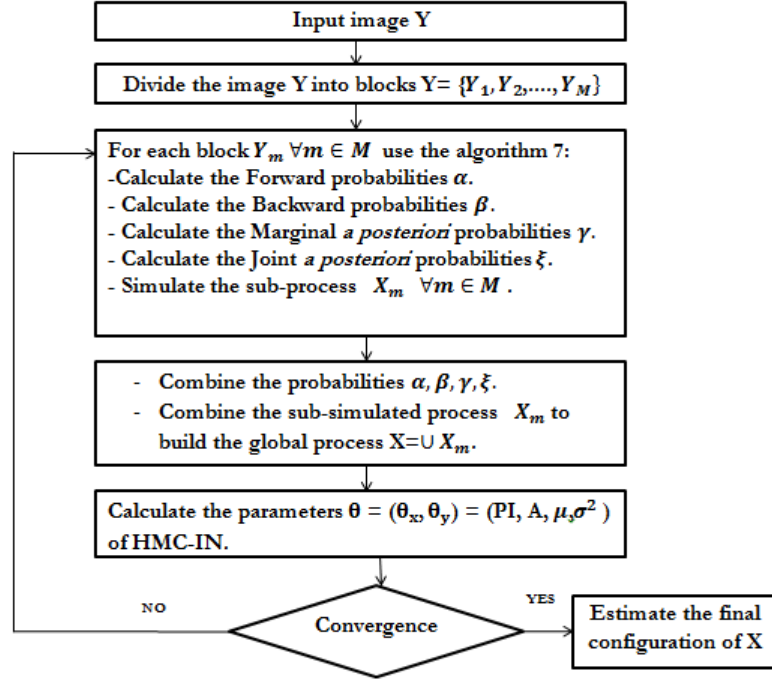


Fig.4. Dependent estimation process : Approach 2

We have proposed two variant methods of Approach 2 (Approach 2-1, Approach 2-2). The following subsections explain in detail the process of these variants.

4.2.1. Approach 2-1

The approach 2-1 has the same procedure as approach 2, but it calculates just the probabilities α , β , ξ , γ without simulating the process X_m of each block. Then, it combines these probabilities to simulate the global process X and to estimate the parameters $\theta^q = (\theta_x^q, \theta_y^q)$ of the HMC. it repeats these procedures until ICE convergence. Finally, it estimates the global configuration of X process. The approach 2-1 proceeds as follows:

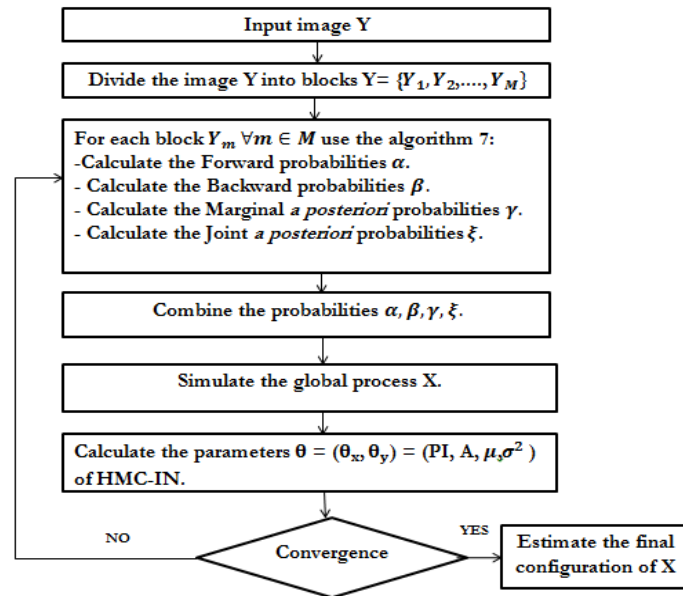


Fig.5. Dependent estimation process : Approach 2-1

4.2.2. Approach 2-2

In the approach 2-2, we have used the algorithm 7 to calculate quickly the probabilities of Forward α , and Backward β of each block. Then, we combine these probabilities to calculate the Marginal *a posteriori* probabilities γ and the Joint *a posteriori* probabilities ξ , to simulate the global process X , and to calculate the parameters $\theta^q = (\theta_x^q, \theta_y^q)$ of HMC. We repeat this calculus until ICE convergence. After that, we calculate the Forward α^0 and Backward β^0 probabilities for the global process, and we estimate the final configuration of X . The figure 6 shows the procedure of the approach 2-2.

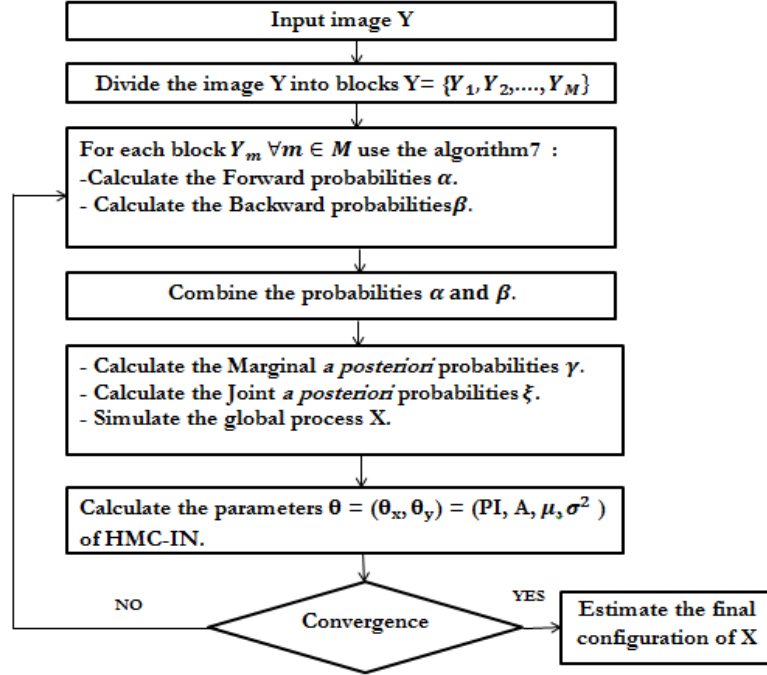


Fig.6. Dependent estimation process : Approach 2-2

Remark:

1. We note that the algorithm 7 can be applied with any estimator like SEM, MCEM(Monte Carlo Expectation-Maximization). These estimators have a similar estimation process than ICE. For EM algorithm, we can use the algorithm 7 without simulating the process X in each iteration.
2. The difference between these variant approaches lies in the way that the BaumWelch probabilities are calculated and the simulation of the process X , before or after combining.

5. Experimental Results

In this section, we evaluate the effectiveness of the proposed approaches compared to the standard approach(HMC without decomposition). The comparison is made in level of quality measures[52] such as PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Measure), error rate, execution time, and the number of iterations to reach convergence. We have realized six experiment image segmentations, in each experiment we have defined the number of classes K , the number of blocks M . For all experiments, we have followed this procedure to initialize the parameters of each process $\theta^0 = (\theta_x^0, \theta_y^0)$.

The initial configuration X^0 of the process X is estimated by K-means algorithm. The parameters θ_x^0 of the hidden process X are estimated using the following formulas:

$$PI^0(i) = \frac{1}{K} \quad (38)$$

$$A^0(i, j) = \begin{cases} A(i, j) = \frac{1}{2} & \text{if } i = j \\ A(i, j) = \frac{1}{2(K-1)} & \text{else} \end{cases} \quad (39)$$

The initial mean $\mu^0(i)$ of the observed process Y is calculated from the initial process X^0 obtained by K-means using this formula:

$$\mu^0(i) = \frac{\sum_{n=1}^N y_n \cdot 1[x_n^0 = i]}{N}. \quad (40)$$

The variance σ^0 is calculated from the mean μ^0 using this formula:

$$\sigma^0(i) = \frac{\sum_{n=1}^N (y_n - \mu^0(i))^2 \cdot 1[x_n^0 = i]}{N}. \quad (41)$$

Note that for color image segmentation, we calculated the mean and variance of each color level Red, Green, Blue (RGB).

5.1. Experiments

We have segmented five images with different sizes and kinds. We have segmented, in experiment 1 a cervical medical image, in experiment 2 a color image of PASCAL VOC2010 dataset[25], in experiment 3 we have a color image of ADE 20k dataset available at [2], and in experiment 4 a normalized satellite image from ISPRS Potsdam dataset [41]. Table 1 indicates the size of images, the number of classes K and the number of blocks M resulting from the division. The choose of number of classes depends on the level color in image, that we consider each color is a class.

Table.1. Experiences characteristics of the experiments.

Experiments	Size of images	Number of classes	Number of blocks
1	344*344	4	344
2	500*324	4	1620
3	2000*1500	10	30000
4	6000*6000	6	360000

5.2. Visual Results

The following figures describe the different conducted experiments and the results of segmentation carried out by the proposed approaches.

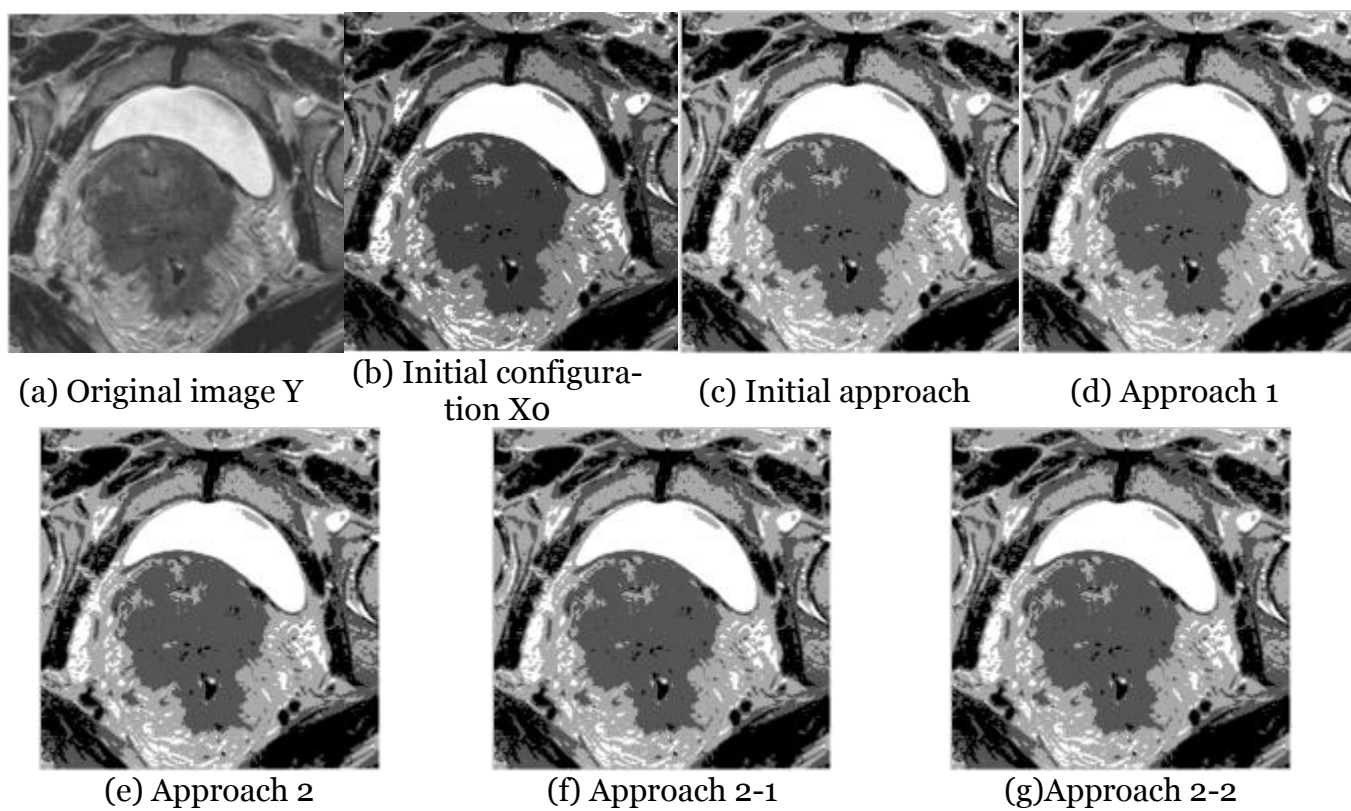
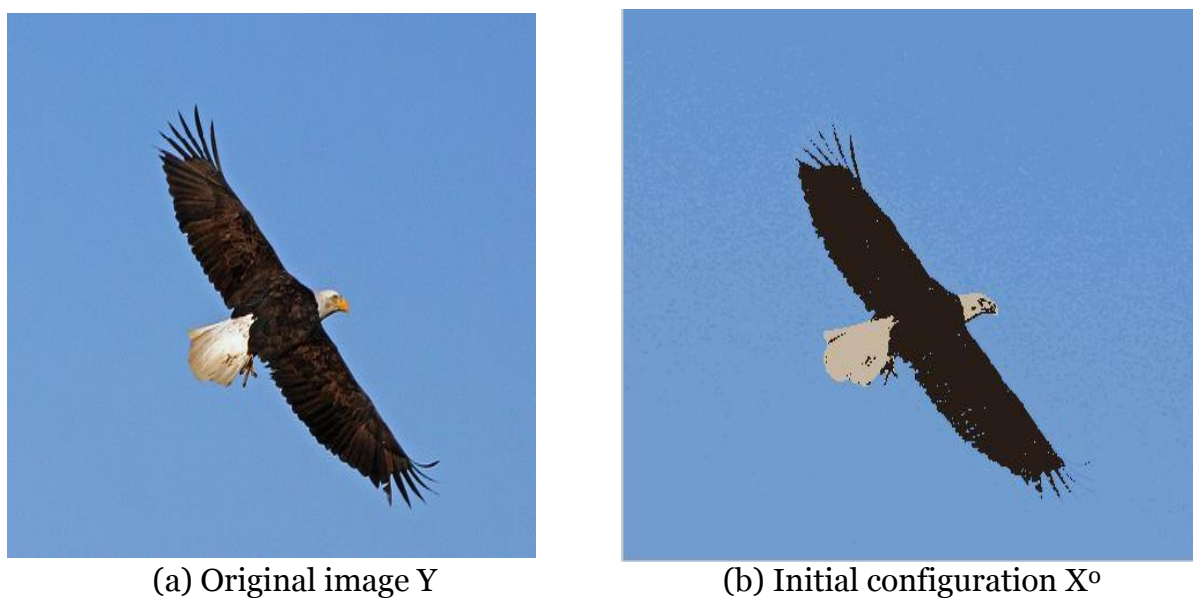


Fig.7.Results of segmenting cervical medical image





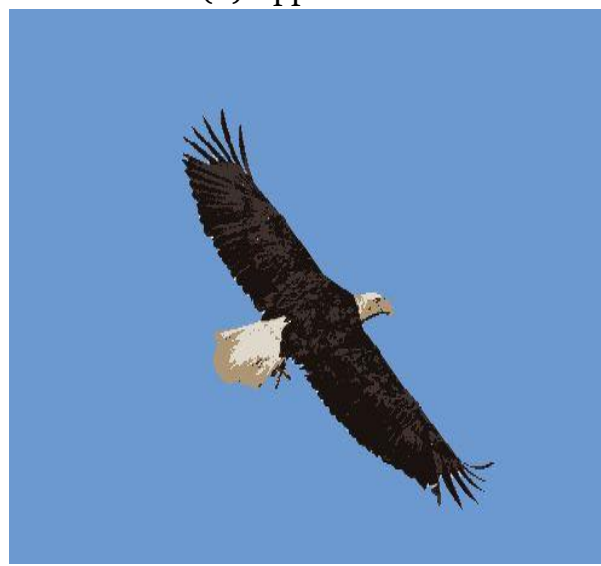
(c)Initial approach



(d) Approach 1



(e) Approach 2



(f) Approach 2-1



(g) Approach 2-2

Fig.8. Results of segmenting color image



(a) Original image Y



(b) Initial configuration X^0



(c) Initial approach



(d) Approach 1



(e) Approach 2

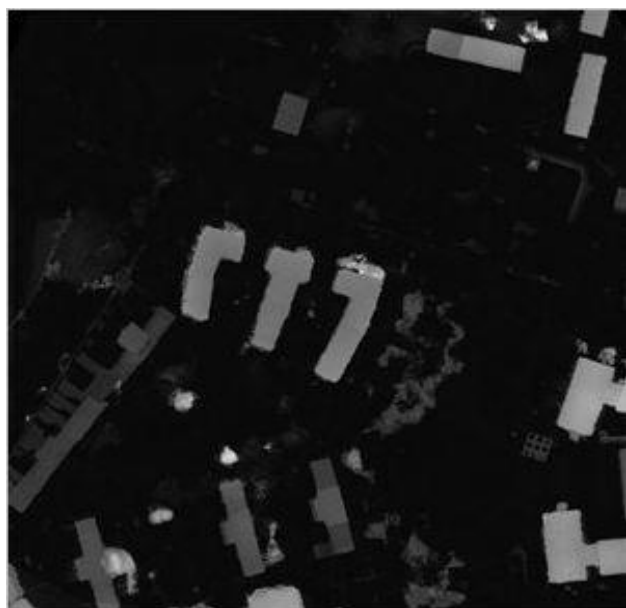


(f) Approach 2-1

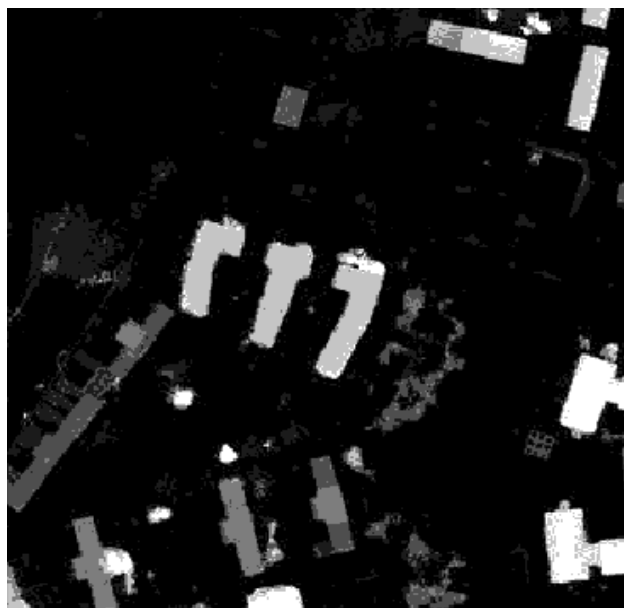


(g) Approach 2-2

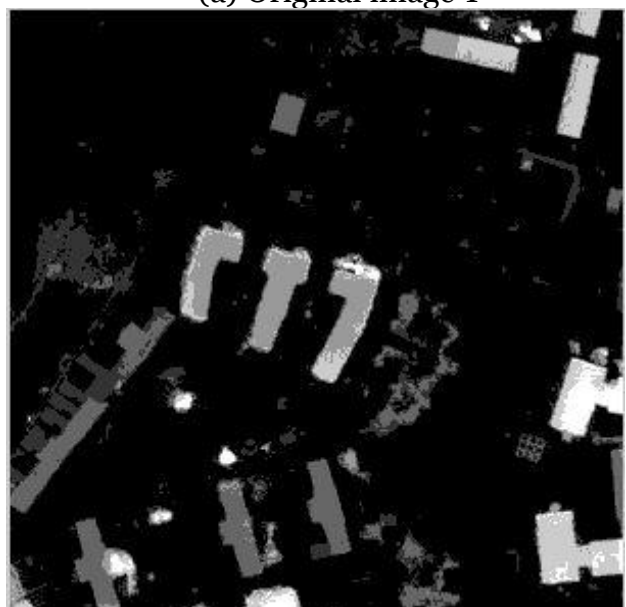
Fig.9. Results of segmenting color image



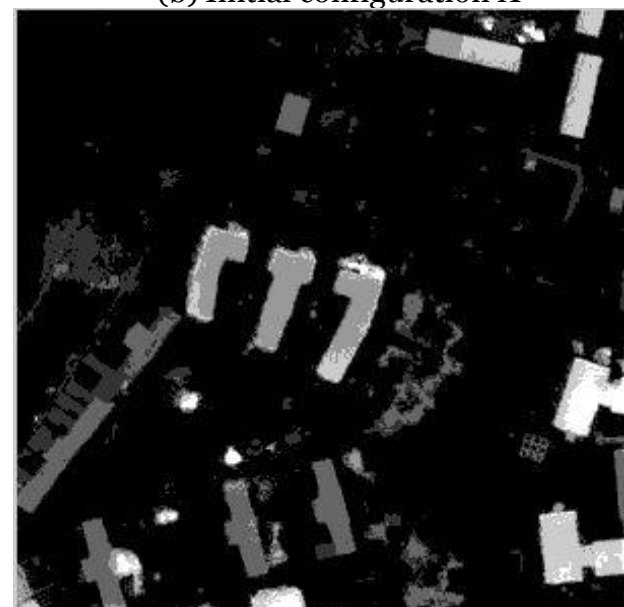
(a) Original image Y



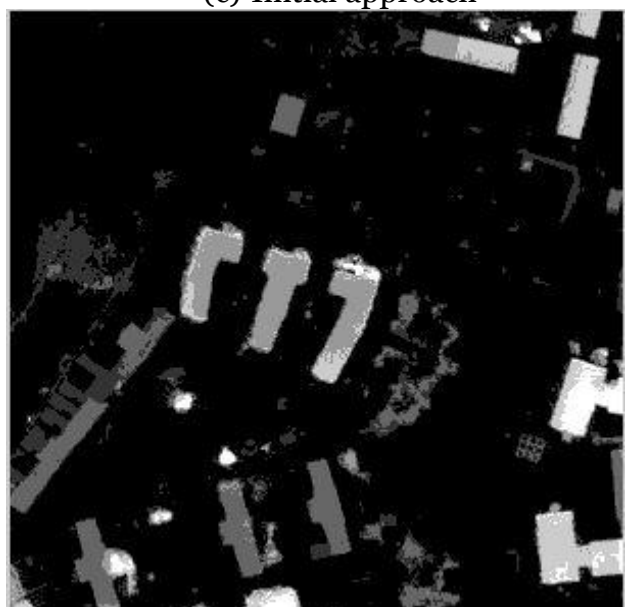
(b) Initial configuration X^0



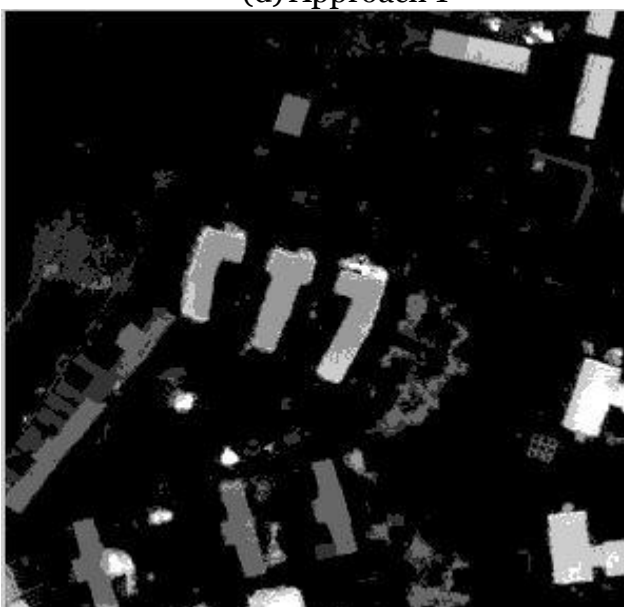
(c) Initial approach



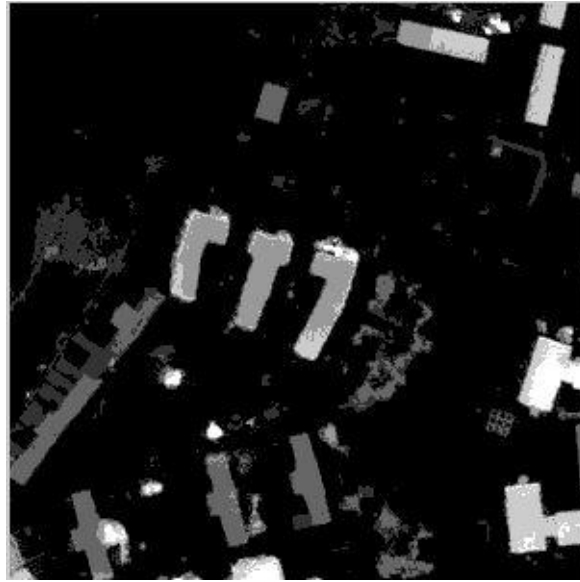
(d) Approach 1



(e) Approach 2



(f) Approach 2-1



(g) Approach 2-2

Fig.11.Results of segmenting satellite image

From these experiments, we observe that the quality of segmented images obtained by the original HMC approach are similar to those of the proposed approaches. As a first good conclusion, our approaches are competitors of the classical HMC approach in term of segmentation quality.

5.3. Quantitative Results

To confirm the visual obtained results, we have calculated some evaluated criteria. The comparison is made in term of quality measures : PSNR index, SSIM index and Error Rate. The obtained results for each experiment are resumed in the following tables: Table2 represents the PSNR index, Table3 shows the index of similarity SSIM and Table 4 presents the error rate.

Table.2. PSNR index values.

Experiments	Initial Approach	Approach 1	Approach 2	Approach2-1	Approach 2-2
1	29,6172	29,6172	29,6172	29,6172	29,6172
2	34,2024	34,2024	34,2024	34,2024	34,2024
3	35,7621	35,7621	35,7621	35,7621	35,7621
4	35,8622	35,8622	35,8622	35,8622	35,8622

Table.3. SSIM index values.

Experiments	Initial Approach	Approach 1	Approach 2	Approach2-1	Approach 2-2
1	0,8196	0,8196	0,8196	0,8196	0,8196
2	0,8934	0,8934	0,8934	0,8934	0,8934
3	0,9145	0,9145	0,9145	0,9145	0,9145
4	0,9350	0,9350	0,9350	0,9350	0,9350

Table.4. Error rate values.

Experiments	Initial Approach	Approach 1	Approach 2	Approach2-1	Approach 2-2
1	15,6690	15,6690	15,6690	15,6690	15,6690
2	26,9111	26,9111	26,9111	26,9111	26,9111
3	21,3230	21,3230	21,3230	21,3230	21,3230
4	23,7645	23,7645	23,7645	23,7645	23,7645

The indexes values (PNSR, SSIM and error rate) illustrated in the previous tables, demonstrate that for all the experiments, these parameters are almost identical, this confirms our remarks concerning the visual obtained results. We can explain; this by the fact that we used the same segmentation procedure, as well as the same estimators and initial parameters values for all approaches.

Additionally, we have calculated the execution time of the proposed approaches and we have compared it with the execution time of the HMC approach. The graphic in figure 12 represents the execution time per second of all approaches under each presented experiment.

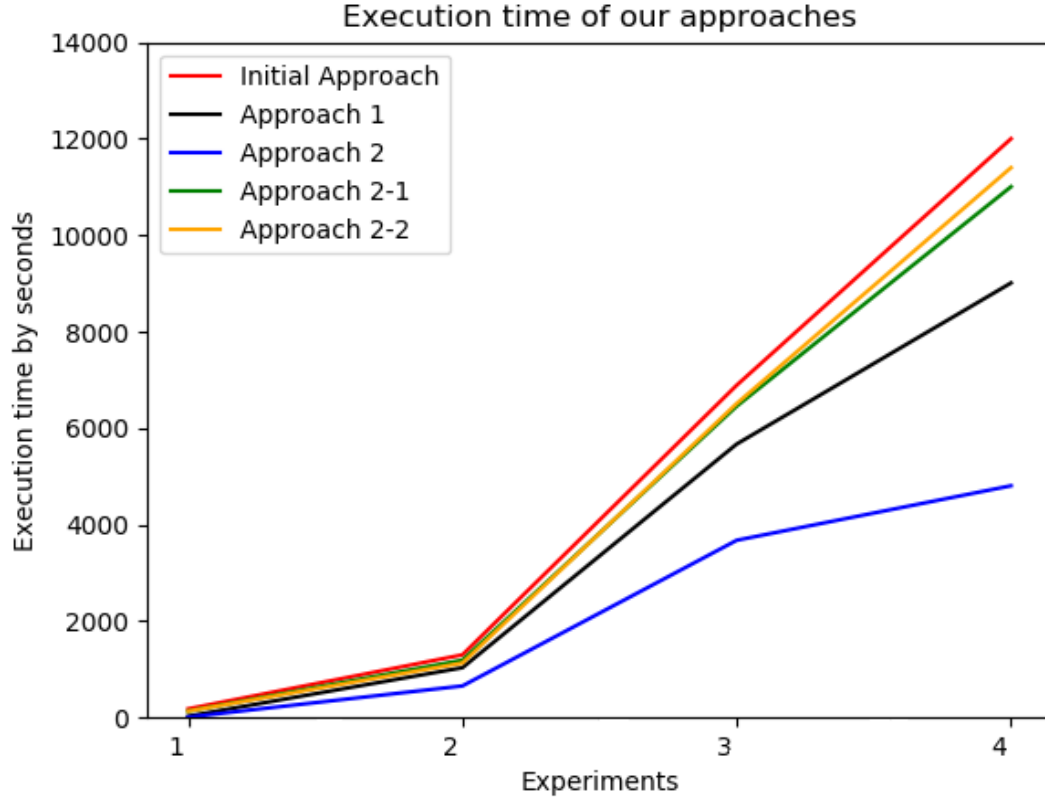


Fig.12. Execution time by approach per seconds

As a first observation, we note that the proposed approaches give better results than the initial approach; in fact, the execution time has decreased considerably. The dependent estimation approach (approach 2) is very efficient compared to the others, in fact, it reduces the execution time in the order of 70% compared to the HMC approach, 30% compared to the independent approach (approach 1), and about 50% compared to the variants of approach 2 (approach 2-1 and approach 2-2). Note that calculating the probabilities of Baum Welch and simulating the process X are the most costly phases in the estimation process. The execution time depends on many factors and experimental conditions such as memory space, execution unit (CPU), speed of microprocessors, the size of data, the number of membership classes, the type of images (grey, color), and the convergence of parameters estimators. From the obtained results and the evaluated parameters, we notice that the standard approach HMC provides better results of segmentation with many types of denoised images (medicals, colors, textured, satellites), but it requires an important time of execution, especially when the size of data is very large.

In this work, we have also studied the convergence rapidity of the proposed approaches. For that, we have compared the number of iterations needed to reach convergence of the dependent estimation approaches with the initial HMC approach. It is difficult to compute the number of iterations of the independent estimation approach, because each block converges independently of other and the number of resulted blocks can be very large. The following figure

shows the number of iterations needed to reach convergence for each approach by experiment.

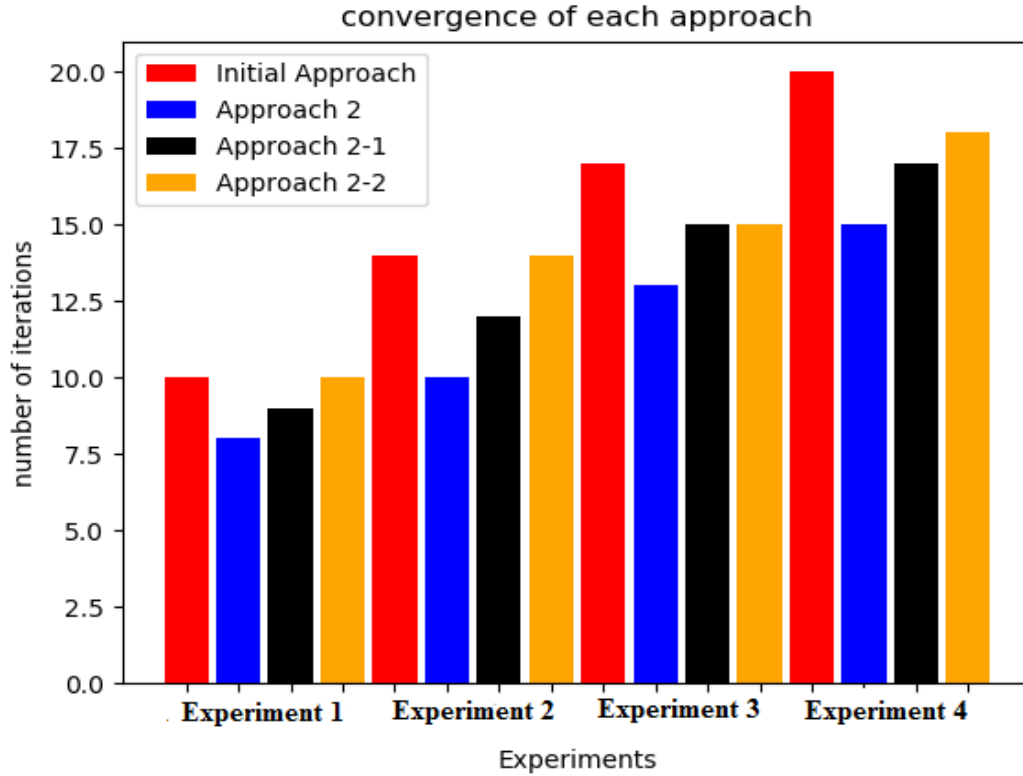


Fig.13. The number of iterations to achieve convergence by approach

From the figure13, approach 2 requires the fewest number of iterations compared to other approaches. This further confirms that this approach is the fastest.

Finally, we can conclude that, all proposed approaches reduce the execution time and the number of iterations needed for convergence, and they keep the quality of segmentation. Approach 2 gives the best results in terms of segmentation quality, execution time, and convergence.

6 .Conclusion

For fast and reliable image segmentation, we have proposed, in this article, some approaches to estimate the parameters of the standard HMC-IN model. The first approach is an evident method of division, which consists of dividing the image into a set of sub-blocks (sub-images) with the same size, each sub-image is segmented independently of the others (independent approach). The second approach consists of dividing the image into a number of blocks, each block is executed using the results of its previous block (dependent approach). Approach 2 uses algorithm 7 to calculate the probabilities of Baum Welch, and to simulate the sub-processes of blocks. From approach 2, we have established two variants, the first variant (approach 2-1) uses algorithm 7 to estimate only the probabilities of Baum Welch of each block. And, the second variant (approach 2-2) uses it just to calculate the probabilities of Forward Backward of each block. To demonstrate the performance of the proposed approaches, we have compared them with the initial HMC approach. Visually, there is no difference between the segmented images obtained by all approaches. These results are confirmed by the parameters evaluated PSNR index, SSIM index and error rate, which are significantly identical for all experiments. However, the proposed approaches provide encouraging results by reducing the execution time and the number of iterations to achieve convergence. In conclusion, approach 2 is the best in terms of segmentation quality, execution time, and convergence. Our work comes up to be closed with some open questions that we address in future

works, one of these queries is to use parallelism techniques in the independent estimation approach, since the blocks are executed independently. Another one is to apply the decomposition technique in the image segmentation process based on Pairwise Markov chains and Triplet Markov chains.

7. Acknowledgment

The authors acknowledge support of Institut Henri Poincaré (UMS 839 CNRS-Sorbonne University), LabEx CARMIN (ANR-10-LABX-59-01) and CIMPA.

References

- [1] A. A. Aly, S. Bin Deris, N. Zaki, Research Review for Digital Image Segmentation Techniques, International Journal of Computer Science and Information Technology(IJCSIT) Vol 3, No 5, Oct 2011.
- [2] ADE 20k dataset : <https://groups.csail.mit.edu/vision/datasets/ADE20K/>
- [3] A. Dempster, al, Maximum likelihood from imcomplete data via the EM Algorithm, Journal the Royal Statistic Society , serie B(Methodological),1977.
- [4] A.M.Raid, al, Image Restoration Based on Morphological Operations, International Journal of Computer Science, Engineering and Information Technology (IJCSIT), Vol. 4, No.3, June 2014.
- [5] Bayes theorem : [https://www.ucd.ie/t4cms/Bayes Theorem.pdf](https://www.ucd.ie/t4cms/Bayes%20Theorem.pdf)
- [6] C.Benson, al, Brain tumor extraction from MRI brain images using marker based watershed algorithm, pp. 318–323, 2015.
- [7] C. Carincotte, Unsupervised image segmentation based on a new fuzzy HMC model, ICASSP'04, Montreal, Canada, May 2004.
- [8] E. Monfrini,al, Image and Signal Restoration using Pairwise Markov Trees, IEEE Workshop on Statistical Signal Processing (SSP 2003), Saint Louis, Missouri, Sep-Oct 2003.
- [9] E. Roura, al, multispectral adaptive region growing algorithm for brain extraction on axial MRI, Computer methods and programs in biomedicine 113(2), pp.655–673, 2014.
- [10] El- H.Guerrou, al, Combination of Hidden Markov Random Field and Conjugate Gradient for Brain Image Segmentation, arXiv:1705.04823v2[cs.CV], 16 May 2017.
- [11] F. Girard-Ardhuin, al, Oil slick detection by SAR imagery : potential and limitation, in Oceans, pp. 22–26, San Diego, USA, september 2003.
- [12] G. D Fornay , the Viterbi algorithm, Proceeding of the IEEE, vol.61, no.3, pp. 268-277, 1970.
- [13] G. R. C. Marquez, H. J. Escalante, L. E. Sucar, Simplified Quadtree Image Segmentation for Image Annotation, AIAR2010: Proceedings of the first Automatic Image Annotation and Retrieval Workshop, pp. 24-34, vol.1, issue.1, 2010.
- [14] H. Caillol, A. Hillion, W. Pieczynski, Fuzzy random Markov field and unsupervised image segmentation, IEEE transactions,Geosci, Remote sensing, 801-810,1993.
- [15] J. A. Canny, computational approach to edge detection, Pattern Analysis and Machine Intelligence, IEEE Transactions on (6), pp.679–698, 1986.
- [16]K. Sun et al. High-Resolution Representations for Labeling Pixels and Regions //arXiv preprint arXiv:1904.04514, 2019.
- [17] L.A.D.S.A. Molligoda,P.G Wijayarathna, Applicability of hidden Markov model approach for sinhala speech recognition? a systematic Review, International Research Symposium on Engineering Advancements(RESA 2015, 2015.
- [18] M. L. Corner and E. J. delp, The EM/ MPM Algorithm for segmentation of textured images, pp.1731-1744, October 2000.
- [19] M. Sandler et al. Mobilenetv2: Inverted residuals and linear bottlenecks //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,pp. 4510-4520, 2018.
- [20] M. Waseem Khan, A Survey: Image Segmentation Techniques, International Journal of Future Computer and Communication, vol. 3, no. 2, April 2014.

- [21] N. Brunel, W. Pieczynski, Signal restoration using hidden Markov chains with copulas, *Signal Processing*, pp. 2304-2315, 2005.
- [22] N. Giordana, W. Pieczynski, Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 5, pp. 465-475, May 1997.
- [23] N. J. Rose, Hilbert-Type Space-Filling Curves, 2000.
- [24] N. Senthilkumaran, al, Edge detection techniques for image segmentation—a survey of soft computing approaches, *International journal of recent trends in engineering*, 2009.
- [25] PASCAL VOC2010 dataset : <http://host.robots.ox.ac.uk/pascal/VOC/>
- [26] P. Devijver, Baum's forward backward algorithm revisited, *Pattern Recognition Letters*, 3, pp. 369-373, 1985.
- [27] P. Korese Drik, Monte Carlo Methodes, Course University of Queensland, 2011.
- [28] P. Lanchantin, W. Pieczynski, Unsupervised Restoration of Hidden non stationary Markov Chains Using Evidential Priors, *IEEE Transactions on Signal Processing*, vol. 53, no. 8, august 2005.
- [29] P. Maragos, A Representation Theory for Morphological Image and Signal Processing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. II, no. 6, June 1989.
- [30] P. Masson, W. Pieczynski, SEM algorithm and unsupervised statistical segmentation of satellite images, *Transaction on Geosciences and Remote Sensing*, vol. 31, no. 3, May 1993.
- [31] R. Haralick, Stanley Stenberg, Xinhua Zhuang, Image Analysis Using Mathematical Morphology, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 4, July 1987.
- [32] R. Van Hadel, Hidden Markov models, pp 51-64, July 28 2008.
- [33] S. Derrode, W. Pieczynski, Unsupervised data classification using pairwise Markov chains with automatic copulas selection, *Computational Statistics and Data Analysis* 63 (2013) 81-98.
- [34] SegNet : <http://mi.eng.cam.ac.uk/projects/segnet/>.
- [35] S. Faisan, L. Thoraval, J. P. Armspach and F. Heitz, Hidden semi-Markov event sequence models : application to brain functional MRI sequence analysis, in *IEEE Int. Conf. Image Processing*, vol. 1, pp. 880-883, Rochester, New York, USA, September 22-25 2002.
- [36] S. Lou, X. Jiang, J. Paul Scott, An efficient divide-and-conquer algorithm for morphological filters, 12th CIRP Conference on Computer Aided Tolerancing, Science Direct Procedia CIRP 10, pp. 142-147, 2013.
- [37] S. Paira, S. Chandra, S. Safikul Alam, and P. Sarthi Dey, Review Report on Divide and Conquer Sorting Algorithm, IEEE Kolkata Section.
- [38] S. S. Al-amri, N. V. Kalyankar, Khamitkar S.D, Image Segmentation by Using Threshold Techniques.
- [39] S. Saini, K. Arora, A Study Analysis on the Different Image Segmentation Techniques, *International Journal of Information and Computation Technology*, pp. 1445-1452, 2014.
- [40] S. tatiraju, al, Image segmentation using K-means clustering EM and normalized cuts, 2008.
- [41] SPRS Potsdam : <http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html>
- [42] T.F Chan, al, Active contours without edges, *Image processing, IEEE Transactions on* 10(2), pp. 266-277, 2001.
- [43] T. Xiao et al, Unified perceptual parsing for scene understanding // *Proceedings of the European Conference on Computer Vision (ECCV)*, – pp. 418-434, 2018.
- [44] V. Badrinarayanan, A. Kendall, R. Cipolla Segnet: A deep convolutional encoder-decoder architecture for image segmentation // *IEEE transactions on pattern analysis and machine intelligence*. – T. 39, N°. 12, pp. 2481-2495, 2017.
- [45] V. Vladimir Kniaz, Conditional GANs for semantic segmentation of multispectral satellite images, *Proc. SPIE 10789, Image and Signal Processing for Remote Sensing XXIV*, 107890R 9 October 2018, <https://doi.org/10.1117/12.2325601>.

- [46] V. Vladimir Kniaz, Deep learning for dense labeling of hydrographic regions in very high resolution imagery, Proc. SPIE 11155, Image and Signal Processing for Remote Sensing XXV, 111550W, 7 October 2019, <https://doi.org/10.1117/12.2533161>.
- [47] W. Pieczynski, Convergence of the iterative conditional estimation and application on the mixture proportion identification, IEEE Statistical Signal Workshop, SSP(2007) Madison, WI, USA, August 26-29, 2007.
- [48] W. Pieczynski, EM and ICE in hidden and triplet Markov models, Stochastic Modeling Techniques and analysis, International Conference, June 8-11, 2010.
- [49] W. Pieczynski, Pairwise Markov chains, IEEE Transactions Pattern Analysis Machine Intelligence, vol. 25, no. 5, May 2005.
- [50] W. Pieczynski, D. Benboudjema, P. Lanchantin, Statistical image segmentation using triplet Markov fields, International Symposium on Remote Sensing, SPIES, Crete, Greece, pp. 22-27, 2002.
- [51] Wei, Tanner, A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithm, Journal of the American Statistical Association 85, pp. 699-704, 1987.
- [52] Y. Al-Najjar, D. Chen Soong. August 2012. Comparison of Image Quality Assessment: PSNR, HVS, SSIM, UIQI. International Journal of Scientific & Engineering Research, Volume 3, Issue 8, August-2012.
- [53] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision // IEEE Transactions on Pattern Analysis & Machine Intelligence. – No. 9. pp. 1124-1137, 2004.
- [54] Y. Boykov, G. Funka-Lea, Graph cuts and efficient ND image segmentation // International journal of computer vision T. 70. No. 2, pp. 109-131, 2006.
- [55] Y. Yang, Image Segmentation By fuzzy C-means clustering algorithm with a novel penalty term, Computing and Informatics, vol. 26, pp. 17-31, 2007.
- [56] Z. Wu, R. Leahy, An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation, Pattern Analysis and Machine Intelligence, IEEE Transactions on 15(11), pp. 1101-1113, 1993.