



HAL
open science

Temporal and Hierarchical HMM for Activity Recognition Applied in Visual Medical Monitoring using a Multi-Camera System

Arnaud Ahouandjinou, Eugène C. Ezin, Cina Motamed

► **To cite this version:**

Arnaud Ahouandjinou, Eugène C. Ezin, Cina Motamed. Temporal and Hierarchical HMM for Activity Recognition Applied in Visual Medical Monitoring using a Multi-Camera System. *Revue Africaine de Recherche en Informatique et Mathématiques Appliquées*, 21, pp.49-66, 2015, Special issue - CARI 2014, 10.46298/arima.1999 . hal-01304733

HAL Id: hal-01304733

<https://inria.hal.science/hal-01304733v1>

Submitted on 20 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

.....

Temporal and Hierarchical HMM for Activity Recognition Applied in Visual Medical Monitoring using a Multi-Camera System

Arnaud Ahouandjinou^{1,2}, Eugène. C. Ezin², Cina Motamed¹

¹Laboratoire d'Informatique Signal Image de la Côte d'opale (LISIC)

Université du Littoral de la Côte d'Opale (ULCO)

Bat 2, 50 Rue F. Buisson , 62228 Calais Cedex, France

ahouandjinou@lisis.univ-littoral.fr, motamed@lisis.univ-littoral.fr

²Institut de Mathématiques et de Sciences Physiques (IMSP)

Unité de Recherche en Informatique et Sciences Appliquées(URISA)

Université d'Abomey-Calavi (UAC)

BP 613 Porto-Novo, Bénin

eugene.ezin@imsp-uac.org

.....

RÉSUMÉ. Nous proposons dans cet article une solution pour améliorer le système actuel de surveillance médicale en Unité de Soins Intensifs (USIs) cardiologique grâce à un système de reconnaissance automatique d'activités humaines. Une approche de vidéo surveillance multi-caméras est proposée à cet effet et permet l'acquisition des données pour l'analyse et l'interprétation automatique de la scène. Cette dernière est basée sur le Modèle de Markov Caché (MMC) avec une durée d'état explicite et intégrant une gestion de la structure hiérarchique interne des scénarios. Plusieurs séries d'expérimentations sont effectuées sur le nouveau système de surveillance proposé en USIs et démontre ainsi la nécessité d'une surveillance assistée par ordinateur des patients afin d'aider les médecins surveillants et les cliniciens dans le processus de prise de décision. De plus, le MMC temporel offre une solution très adaptée pour la reconnaissance automatique des événements en USIs. Enfin, les résultats obtenus avec le modèle de MMC temporel et hiérarchique ont été comparés à ceux des MMC classiques.

MOTS-CLÉS : Système de surveillance en USIs, Reconnaissance d'activités humaines; Analyse et interprétation vidéo; MMC classique et à durée d'état explicite, MMC hiérarchique.

ABSTRACT. We address in this paper an improved medical monitoring system through an automatic recognition of human activity in Intensive Care Units (ICUs). A multi camera vision system approach is proposed to collect video sequence for automatic analysis and interpretation of the scene. The latter is performed using Hidden Markov Model (HMM) with explicit state duration combine at the management of the hierarchical structure of the scenario. Significant experiments are carried out on the proposed monitoring system in a hospital's cardiology section in order to prove the need for computer-aided patient supervision to help clinicians in the decision making process. Temporal and hierarchical HMM handles explicitly the state duration and then provides a suitable solution for the automatic recognition of temporal events. Finally, the use of Temporal HMM (THMM) based approach improves the scenario recognition performance compared to the result of standard HMM models.

KEYWORDS: Monitoring System in ICUs, Human Activities Recognition (HAR); Video Analysis and interpretation; Classic HMM and HMMs with explicit state duration.

.....

1. Introduction

Human Activity Recognition (HAR) in video scenes is an important topic in computer vision which has had a significant upward in term of research. In high dependency environment, like intensive care units, all clinical, economic and sociological studies tend to demonstrate the need for computerized assistants that help clinicians in the decision making process. A patient's monitoring in Intensive Care Units (ICUs) in medical services such as cardiology, anesthesia is a key challenge for the efficiency of health care for the sick. In these ICUs, the medical staff number to monitor patients is very limited and the patients may even be in separate wards. It is well known that standard monitoring systems based on alarm thresholds for each individual parameter are very primitive since they introduce noise pollution instead of real informative events [13, 16]. Improving the performance of the monitoring system in intensive care units is a major issue about the recent research topic on medical devices. Therefore, an automatic and intelligent real-time monitoring system is necessary to increase quality of medical-service in intensive care units. We feel that video monitoring based on multi-camera perception enables to supplement the existing monitoring system in ICU because it is quite obvious that hearing and vision are complementary. This video layer can provide visual-information in real-time for pro-active actions in medical emergency, and design an image database that we term Medical Black Box (MBB) for future diagnosis and analysis. The proposed video surveillance system must be able to analyze and interpret the medical video scenes for recognizing significant events of a patient's hospital stay and to help physicians in the monitoring tasks. This automatic recognition of clinical scenes or retrieval of specific behavior of patients as we have developed is fundamental for computer-aided in the patient supervision. In this work, we also tackle the recognition problem of the physician's activities and the behavior or risk situations for the patients in the multi-camera perception approach. The recognition method is based on Hidden Markov Models with explicit state duration. The latter enables us to quantify the duration spent by an HMM in a state through the probability computation. The aim is to recognize scenarios in the form of temporal sequences. The proposed model is based on Markov models. This model allows one hand, hierarchical modeling for managing sub models integrated into higher-level models and also the semi-Markov modeling to control the duration of each state of the Markov model. These two extensions have already been exploited (in isolation) in the field of recognition of temporal sequences but their joint or simultaneous operation is a new and major contribution of this work. The remaining part of this paper is organized as follows. In section 2, we describe the proposed architecture of the ICUs monitoring system and to make paths towards the establishment of Medical Black Box. Section 3 focuses on the proposed approach for human activities recognition in ICUs. Section 4 is devoted to experiments and the discussion about results.

2. Proposed ICU Monitoring System for Building Medical Black Box

2.1. Proposed ICU Monitoring System

The intensive care units are a specialized section of a hospital containing medical equipment, physicians and the nursing staff, monitoring and life support devices necessary to provide continuous care to the patients that are severely ill and medically unstable [14]. We can find in a typical ICU environment a ventilator which assists the patient's ventilation, pumps to supply of various drugs, and various devices to the heart or respiratory functions [3, 11]. The set of these devices is handled by a multidisciplinary team that may consist, intensivists (clinicians who are specialized in critical illness care), clinical pharmacists, nutritionists and others. These facilities are connected to the data collector which displays the collected data [14]. The vital sign information acquired from bedside monitoring ICU system is analyzed by supervising physicians for decision making. ICUs alarms were designed to pay attention to a patient, to alert a change in the patient's physiology or to alert the staff to a device problem [4]. Moreover, alarms do not always match the criticality of the patient's condition, hindering the nurses' ability to react rapidly with the appropriate clinical intervention [16]. The current monitoring system in ICU is very alarmist and worse, suffers from an acute lack of visual data such as images of the basic actions, the behavior and the clinical status in real-time of patients as well as those of the physicians during the hospital stay of the patients. For this, we propose to improve the current monitoring system in ICU by supplementing it with an automatic and intelligent video-surveillance system in a multi-camera framework. To achieve this, the proposed system of monitoring in ICU is an audio-visual system. To this end, The ICU is monitored using a multi-camera vision system based-approach: in wards, the cameras are attached to the ceiling and the others faced at the bedside of the patient bed. In our application context, the topology of the proposed video surveillance system is composed of six cameras (two cameras mounted to the ceiling with one facing each bed) with a self-calibration approach in order to perform several pre-treatments of image acquired before by the automatic recognition framework [2]. This multi-camera vision system provides overlapping views that will be used by information fusion to improve the quality of recognition. The main functionality of the proposed ICU video-surveillance system is the automatic recognition of human activities in medical environment. Before the recognition stage, the scenario is automatically first learned by the system. Afterwards, we have the step of scenario recognition which enables to make the summary of the clinical scenes of the physician and important or critical events of the patient during his hospital stay through the video selection. The latter computer vision task helps us to design an especial database in ICU, that we term Medical Black Box (MBB) in order to store interesting images of clinical scenes and important behavior of the patient. Figure 1 is an illustration of the proposed architecture for the visual-monitoring system in ICUs.

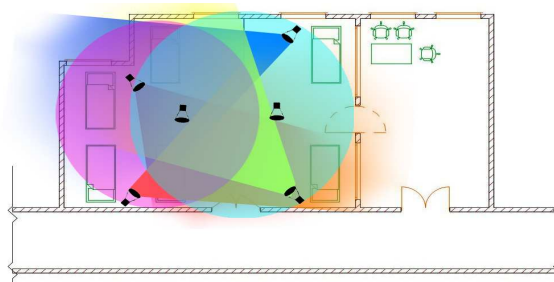


Figure 1: Proposed architecture of the visual-monitoring system multi-camera based approach in ICUs.

2.2. Towards the Establishment of Medical Black Box

Due to the advancement of instrumentation technology and more specifically, in medical fields, such as intensive care units, developing a medical black box, like aircraft black box, is a genuine challenge and an original recent idea. Working on this topic is motivated by the need to create a support system for medical decision. From the automatic proposed monitoring system, enormous collections of medical videos are available with several applications in various aspects like medical imaging, medical diagnoses, the training the medical professionals, a decision support for forensic physicians, the uses of evidence by health insurers, medical research and education. The intelligent video monitoring system helps to manage, to analyze, to index and retrieve the information from the IP video stream. The aim of this system is to extract good visual content representative key-frames. In order to achieve this, we used an efficient method for human activity recognition in medical videos which leads to the extraction of better key-frames as representative (interest scenario) for important events summary. The MBB is a database containing interesting images obtained by video selection which represents the summary of critical or important events from hospital stay. The concept of the medical black box is an original new idea that presents double interest : firstly, it is a decision support for medical monitoring in intensive care units and secondly, it allows, after analysis and diagnosis of MBB's content, to contribute to the development a new intervention protocol which improves the performance of medical care. The indexing method of the big data and retrieval information contained in the medical black box is a new research issue that we will tackle this in our future work. The Medical Black Box is a very promising solution to improve medical service performance in ICU.

3. Proposed Approach for Human Activities Recognition in ICUs

The proposed multi-camera scenario recognition approach is based on a Temporal HMM (THMM) inspired from many related work proposed in [1, 6, 7, 8, 9]. It is performed on two main levels of processing: the low-level processing and the high-level processing. An illustration for the proposed architecture for human action recognition is given in Figure 2:

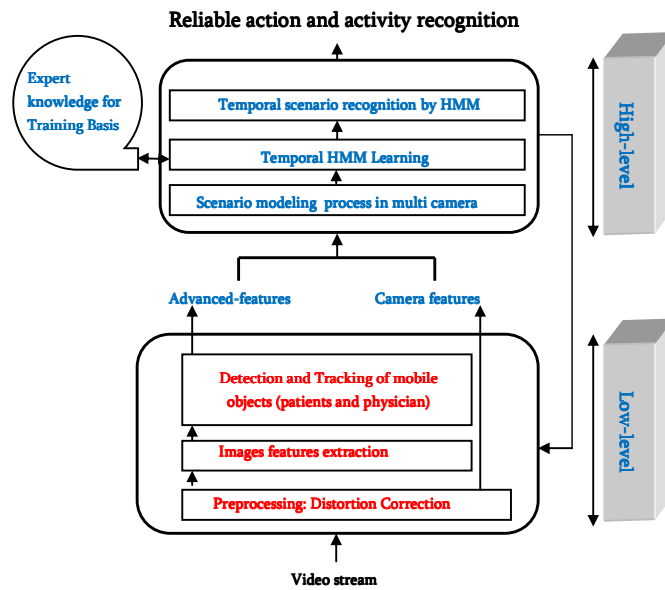


Figure 2: Proposed architecture for human action recognition in medical videos.

3.1. Features extraction

One of the basic tasks of the low-level processing is the features extraction. We used many feature descriptors in the video for defining some areas of interest and for detecting mobile objects in the scene. Each detected object is classified based on morphological features by combining object color and texture. We have defined a codebook where each character corresponds to different predefined regions of interest in the video. The input states of HMM [18] are interest area classes labeled symbolically by the codebook letters (A, B, C, D, E, F, G, H, I) and the output states are basic events establishing high level scenarios. The motion detection of mobile objects is based on a simple and efficient algorithm. It uses the reference image approach proposed in [5]. Then, moving objects are classified using the labels defined for the recognition phase. The objects tracking module is based on a robust real-time object tracker using improved mean shift by integrating texture and color features. The position of a subject's feet is approximated by taking the centre of the bottom of the tracker supplied bounding box. The position was used to estimate the proximity of the person to defined areas of interest. Figure 3 shows the modeling result of the mask of interest regions and the various trajectories.

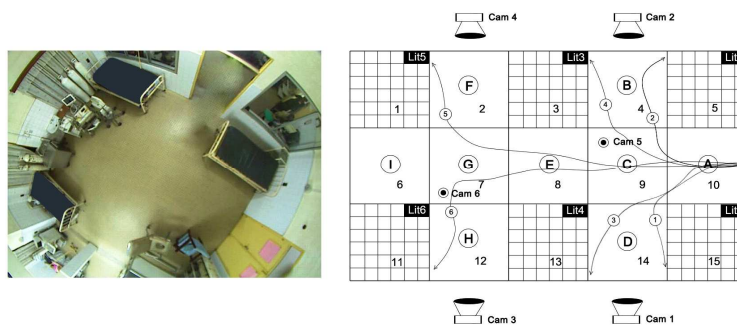


Figure 3: Proposed of the mapping of the trajectories architecture using areas of interest for low-level processing.

3.2. Temporal and Hierarchical HMM for activity recognition

In the related work on activity recognition, generally approaches are based on two steps, firstly, feature extraction and representation, then learning the temporal characteristics in activity sequences. While the HMM [10] is a mere and efficient model for learning sequential data, its performance tends to degrade when the range of activities becomes more complex, or the activities exhibit long-term temporal dependency that is difficult to deal with under the Markov assumption. To get around these limitations, two classes of extension to the HMM have been proposed [6, 7]: Hierarchical HMM (HHMM) which includes the Abstract HMM, the Hierarchical HMM, and the Layered HMM and Semi-Hidden Markov Model (SHMM), including explicit duration HMMs and segmental HMMs. We described in the following subsection the first abstraction layer which concerns the temporal event evolution and the second layer which is hierarchical internal structure of events.

3.2.1. Temporal evolving management in videos sequences

In the classical HMM, state duration is implied as a function of a self transition probability of the state. The likelihood of the remaining in a unique state is exponential. More extensively, the duration of human activity is an important consideration if we are to accurately model a person's behavioral patterns. We show that duration modelling enables us to differentiate between activity sequences in which the order of the observations is identical and which the duration of the activities is varied. We believe that its application to the domain of human activity recognition remains very novel and original [8].

Duration model. An explicit state duration in a HMM was introduced in 1985 as part of an application of speech recognition by [15, 12, 10]. This time variable is measured by means of a probability not $p_i(d)$ such that $1 \leq d \leq D$ where D constrains the maximum duration that can be estimated by one analytical discrete distribution (i.e., Poisson) and two other continuous (i.e., Normal and Gamma) for modeling state duration. These probabilities are defined by equation (1) for Poisson distribution and equation (2) for Gamma distribution. For more details, the reader may refer to [1] for computation parameters:

$$P_j(d) = e^{-l_j} \times \frac{(l_j)^d}{d!} \quad (1)$$

$$P_j(d) = \frac{1}{\sigma_j(2\pi)^{1/2}} \times e^{-\frac{(d-m_j)^2}{2\sigma_j^2}} \quad (2)$$

Model parameter re-estimation. i) given the sequences of observations O obtained over the time $t \in [1, T]$ $O = \{O_1, O_2, \dots, O_{t-1}, O_t\}$ produced, ii) given the path variable such as $\alpha_t(i)$ and $\beta_t(i)$ estimated in equations 3 and 5 the model re-estimation formulas are presented in equations 4 and 6. The model is based on a temporal HMM noted λ ($\lambda = (A, B, \Pi, P_j(d))$) with an explicit state duration probability distribution is defined by the following parameters: $A, B, \Pi, P_j(d)$ that are respectively state transition probability matrix, output probability matrix, initial state probability vector, and a state duration probability vector [8, 9].

$$\alpha_t(j) = P(O_1, O_2, O_3, \dots, O_t, S_j \text{ ends at } t/\lambda) \tag{3}$$

$$\alpha_t^*(j) = P(O_1, O_2, O_3, \dots, O_t, S_j \text{ begins at } t+1/\lambda) \tag{4}$$

$$\beta_t(j) = P(O_{t+1}, \dots, O_T / S_i \text{ ends at } t, \lambda) \tag{5}$$

$$\beta_t^*(j) = P(O_{t+1}, \dots, O_T / S_i \text{ begins at } t+1, \lambda) \tag{6}$$

The model likelihood $P(O/\lambda)$ of an HMM λ with explicit state duration, for a discrete observation sequence O , can be computed by a generalized forward-backward algorithm as follows:

$$P(O / \lambda) = \sum_{i=1}^N \sum_{j=1, i \neq j}^N \sum_{d=1}^T \alpha_{t-d}(i) a_{ij} P_j(d) \prod_{s=t-d+1}^t b_j(o_s) \beta_t(j) \tag{7}$$

$$\pi_i^* = \frac{\pi_i \beta_0^*(i)}{P(O|\lambda)} \tag{7}$$

$$a_{i,j}^* = \frac{\sum_{t=1}^N \alpha_t(i) a_{i,j} \beta_t^*(j)}{\sum_{j=1}^N \sum_{t=1}^T \alpha_t(i) a_{i,j} \beta_t^*(j)} \tag{8}$$

3.2.2. Hierarchical structure management with temporal constraint

In the classical HMM, Hierarchical structure modeling in the classical HMM enables managing efficiently the recognition decision of the sub-models embedded in models of the highest level. In doing so, a hierarchical HMM has the ability to effectively recognize several scenarios with different organizational structures but all being located in the same category. Managing the hierarchical organizational and structure of the scenarios is based on the definition of new version of HMM. This last can be viewed as a tree with a single root at the top of the hierarchy. Issuers statements are the sheets and the links between the states are the branches. His approach is a purely hierarchical form of assignment within a framework of multi-level decision. A hierarchy is a tree having a root associated with Ω , level (composed of a set of elements) is associated with a partition of Ω , a sheet associated with a singleton element. By convention, the first level is the level consists of Ω and singletons last level is Ω compound. Let $n \in \{1, \dots, N\}$ n denotes the level of the hierarchy; where N represents the number of levels of hierarchy. The proposed approach is based on Hidden Semi Markov Model (HSMM) and includes an explicit management of state duration model by estimating a probabilistic variable through the discrete and continuous distributions. To this end, the proposed model is composed of a sequence of time slot, wherein each time is actually a variable activation of the end of the current state, and the same network structure is replicated as and as time progresses. Temporal variables are interconnected by arcs of how the current states are triggered by the end of the previous variable. This model allows to recognize activities falling at the same level of the structure and at the same time, offers the possibility of estimating the time spent by the automaton in a given state to discriminate temporally these scenarios come together at the same level of the

hierarchy. Figure 4 shows an example graphic for the hierarchical organization management in the classic HMM.

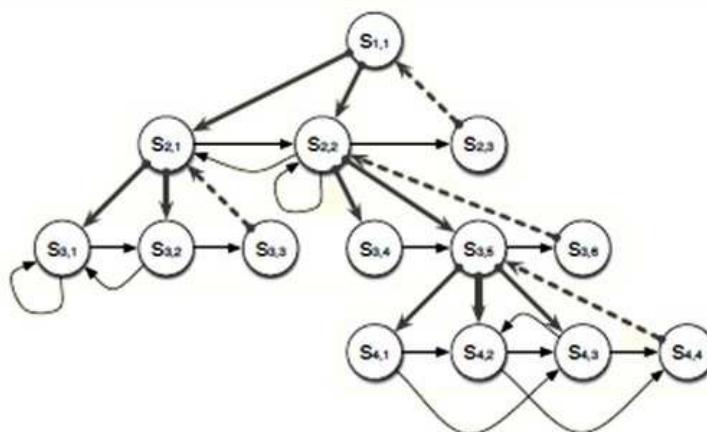


Figure 4: An example of hierarchical organization representation.

We define the model of THMM by $\lambda^n = (A^S, B^S, \pi^S, \Gamma^S, P_S(d), N)$ where $P_S(d)$ is the probabilistic framework the dwell time in the transmitter S state and d the duration of the state of the model [22]. The components of THMM are defined as in a standard case of HMM version. For each sub-automaton (represented by a standard HMM), the process of the transitions from one state to another is governed by a distribution A , and the probability distribution for the initial state is governed by a probabilistic distribution law noticed B . However, there is a restriction on the number of terminal states is set to 1. The only authorized horizontal transitions are those between states of the same sub-model and in other words the same sub-automaton, in other way, none direct transition can occur between the different states of automata in the hierarchy.

For instance, in the context in our application, we used a hierarchical organization like a tree structure to model the category of scenario "The doctor provides a patient care" which is divided into six sub-scenarios grouped into six classes "doctor provides with medical care for patient at different beds" such as, the bed number one, the bed number two, so on to the bed number 6.

4. Experiments and Results

Significant tests are carried out on dataset that are gathered from the proposed experimental monitoring system in intensive care units. Our Algorithms have been implemented using Matlab, C/C++ and OpenCV library). The database is characterized by its heterogeneity with a high variation of illumination (day and night) and several people such as physicians and patients (male, female with different actions, skin colors) other moving people as the visitors among others, may appear at defined and known instants. Furthermore, the database has been collected from the experimental video-surveillance system that we installed in the cardiology department at the hospital (CNHU: Centre National Hospitalier Hubert Manga) in Benin.

4.1. Used Video Dataset Description

We have collected roughly 50 videos sequences for three activities with 1500 frames/sequences. An observation sequence is recorded every one and a half second from the 25fps video. The duration of the video sequences is 300 seconds with an average length of circa 90 seconds. In these videos sequences, we distinguish two main classes of activities including: Activities of Doctor (AD) and Activities of Patient (AP) which are composed of two main types of scenarios each. The activities of a patient are derived from these elementary actions, their behavior and their clinical status. As for as the activities of a doctor, they are formed from these elementary actions and these clinical scenes. For each class, two scenarios are recognized in two scenario categories which are: For AD class "Activities of Doctors" have been tested during the experimental phase, and then we have: *"The doctor provides a patient with medical care"*, *"The doctor carries out a general inspection of ward"* and for the AP class "Activities of the Patient": *"The patient convulses"*, *"The patient feels pain"*. Figure 5 shows an example of frames for medical videos.



Figure 5: An example for ICU's monitoring video data base

4.2. Description of the scenario modeling using SHMM

4.2.1. Modeling the scenario of the class "doctor activities"

Concerning the category scenario of the class AD "Activities of Doctors" we have, for instance, the scenario *"The doctor makes a patient care at bed 1"* recognition is performed by the HSMM with three states. Indeed, the time spent by the physician to provide a health care for a patient is first defined and known under normal conditions. This time varies depending on the type of care and therefore differs finally from the time spent by the physician during his visit to the patient in the normal case. See figure 6 for an illustration for HMM topology for recognizing the scenario *"The doctor makes a patient care at bed 1"*.

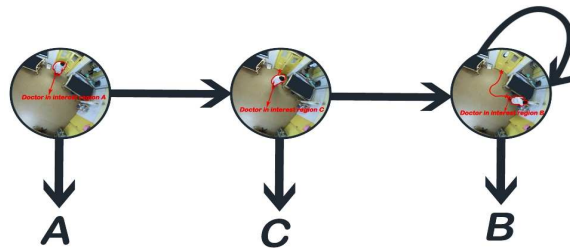


Figure 6: An illustration for HMM topology to recognize the scenario "*The doctor makes a patient care at the bed 1*"

As in the previous case of scenario recognition, the input states of the HMM are made through the results of detecting and tracking moving objects and those in output symbols representing the codebook to recognize interesting scenarios. A ward inspection by the doctor can begin with any bed. Figure 7 shows an illustration of HMM topology to recognize the scenario "*The doctor carries out a general inspection of visit in the ward*".

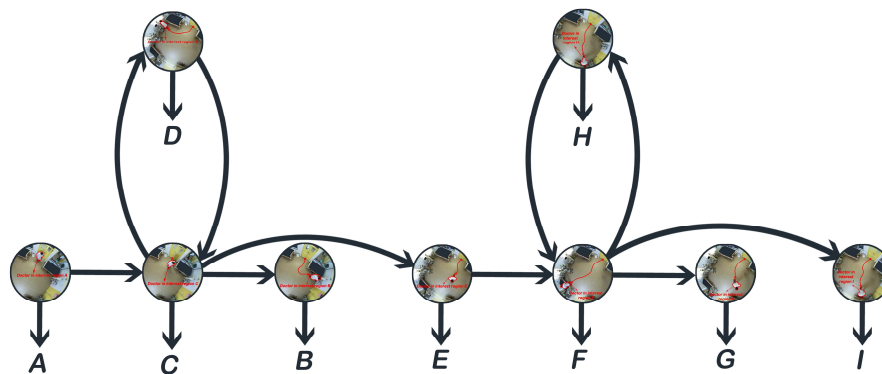


Figure 7: an illustration for HMM topology to recognize the scenario "*The doctor carries out a general inspection of visit in the ward*".

Concerning the hierarchical layer abstraction that we used to model the internal organization structure of scenario, the Figure8 shows the hierarchical structure exploited by THHMM to recognize the scenarios of the class "*the doctor visits a patient*" and "*doctor makes medical care to a patient*". The intensive care unit consists of a surveillance room where supervisors or intensivists physicians and the care unit where are staying patients. The care unit as we have previously modeled in the above section consists of regions or areas of interest. Each area of interest is itself decomposed into basics events, with which they are linked in order to generate medical activities with a time parameter differentiator. Indeed, the structure and number of THHMM states are computed from a priori knowledge available about the environment. The structure of THHMM presented on the above Figure 8, can be viewed as a tree with a single root at the top of the hierarchy. Emitters states are the leaves and the links between the states are the branches [19], [20]. To differentiate properly, these two scenarios, it took explicitly model the model status-lived. We introduce thus an explicit modeling of the time spent by the hierarchical MMC in a given state by exploiting a mixture of statistical distributions (Poissonian, Gaussian). The structure of THHMM presented on the above

Figure 8, can be viewed as a tree with a single root at the top of the hierarchy. Emitters states are the leaves and the links between the states are the branches [22] [21]. To differentiate properly, these two scenarios, it took explicitly model the model status-lived. We introduce thus an explicit modeling of the time spent by the hierarchical HMM in a given state by exploiting a mixture of statistical distributions (Poissonian, Gaussian). The structure of THHMM used is composed of three main hierarchical levels which are: the root, the interests-area level (*forward-areas and backward-areas*) and the basis events or actions. Three hierarchical levels are defined in our proposed THHMM model: “*a scene level*”, “*actions of interest-area level*” and finally, we have the level of “*actions or basics events*”. We choose n the level of the hierarchy, and so each level is described as follows:

At the “*scene*” level, where $n = 1$, we used a THHMM consisted of a single state K noted A . This is the root of THHMM. In an application case, the state called R is composed of N sub-levels “interest areas” (for instance, in the case of the studied scenario, $N = 3$).

At the “*actions of interests-areas*” level, where $n = 2$, we used a THHMM composed of four sub-levels of interests area which help us to define the forward-areas and backward-areas. The first sub-level consists of three interests zones noted (zone A, zone B, zone C) in which the presence of the physician is satisfied. The second sub-level is composed of two interests zones called (zone E and zone D) and the third sub-level also includes three interests zones noted (zone F, zone G, zone H) and finally the last sub-level zone is associated with a single interest area called Zone I.

At the “*action and basics events*” level, where $n = 3$, we used a THHMM composed four actions or basics events, like shown in Figure 8. Each action is associated with a type of event presence of a doctor facing a bed for the medical visit or patient care. Action noted “action 1” is generated by the presence of physician at the bed 1 beforehand and reflects the fact that the doctor is close to or facing the bed 1 of the patient. Depending on the time spent by the doctor at this level (the area of interest), we can define another event that is the doctor makes a specific medical type of patient care.

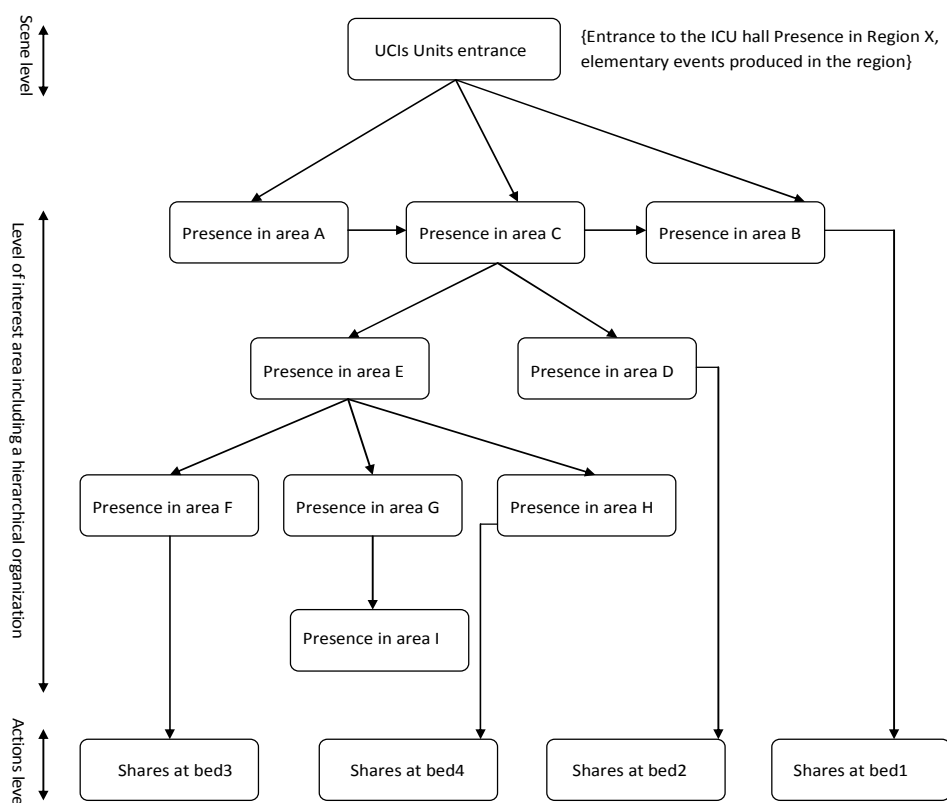


Figure 8: An illustration for HMM hierarchical topology used in order to recognize the sub-scenario in the class of scenario "the physician visit or makes a medical care is a patient from a bed n".

4.2.2. Modeling the scenario of the class "patient activities"

With the scenario "The patient convulses", the convulsion is characterized by several physiological signs in the patient (medical expert information). Therefore, the efficient recognition of this scenario requires a more complex modeling and also sophisticated image acquisition sensors. However, we simplified the modeling and carried out a high-level video scene analysis and interpretation for the recognition of behavior of the patient. We chose to model this scenario based on the limbs sequence motion (hands and feet) of the patient but also his whole body while highlighting the rhythm, cadence and speed of motion of the latter. An HMM with four states was used to recognize this scenario. Each state of HMM enables the modeling and the detecting of these two types of actions (elementary events) motion (slight and/or jerky motion) of the foot (left and/or right), motion (slight and/or jerky motion) hands (left and/or right). Depending sometimes on the state and the speed of the limbs motion by the patient's body, the scenario "the patient restlessness" is characterized and detected. Concerning, the motion of the hands left and right respectively is symbolized by " H_L " called "Left Hand Motion" and " H_R " called "Right Hand motion". With regard to the motion of the left foot and right foot respectively, it is symbolized " F_L " called "Left Foot motion" and " F_R " called "Right Foot motion". Figure 9 (a) shows an illustration of HMM topology

for recognizing the scenario *"the patient convulses"*. The scenario *"the patient feels pain"* is not only based the motion of the limbs (hands and feet) of the patient, but also based on the motion of his head and even more effectively by combining facial expressions. The latter is also quite difficult to model using a sensor and needs the use of sophisticated sensors for the patient's face acquisition. We simplified by just combining the limbs motion to those of the heads to in order to detect if the patient feeling of pain. Two HMM were used to recognize this scenario, one with four states for limbs motion recognition (i.e. the scenario patient convulses) and the other with two states to recognize head motion (head left : H_L and head right: H_L). Combining the results of basics actions recognition with each of the two HMMs, we detect if the patient feels the pain. According to our experiments results which are reported in Table I, we have noticed a marked improvement in the scenario recognition rate, on average by 30%, compared to the classic HMM. Figure 9a and Figure 9b show an illustration for HMM topology to recognize the scenario *"the patient feels pain "*.

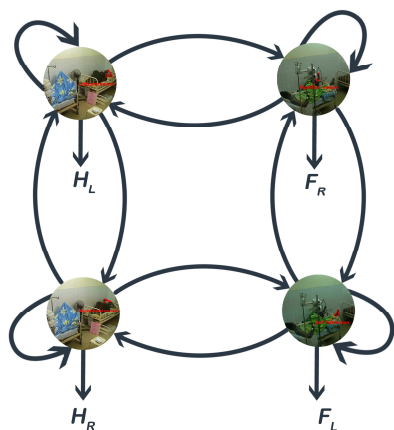


Figure 9 (a): an illustration for HMM topology to recognize the scenario *"the patient feels pain"*.

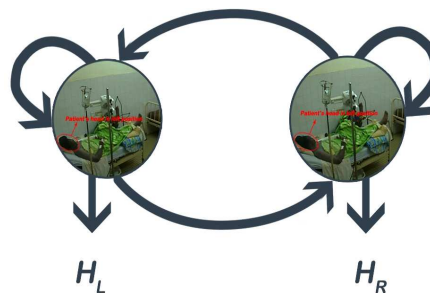


Figure 9(b): an illustration for HMM topology to recognize the scenario *"the patient convulses"*.

4.3. Learning technique: results and analysis

4.3.1. Training scenario by THHMM

In this section, we present experimental results in relation to learning of the three scenarios and identified by the class to which they belong. Not only does this work focus on automatic scenarios recognition issues, however it also tackles a challenging problem which is the learning and distinguishes the scenario of the same category mainly based on differences in the durations of their sub-activities. We experimented by using the THHMM variants in learning and recognizing three scenarios for two main classes and compared them with the standard HMM. In these THHMM variants, different kinds of distribution were used for modeling state durations, including the Poisson distributions and the Gaussian distribution (normal and inverse case). The Poisson distribution as an example for discrete distribution was chosen because of its simplicity and its good results in modeling state durations for the SHMM in speech recognition, e.g., [12] and did not need a lot of data for the learning before becoming stable. The Gaussian

distribution was selected as an example of continuous distributions for duration modeling because it was restricted to the positive domain and the inverse Gaussian was used to model patients' staying time in hospital with successful results [17]. In this work, a deep analysis of the impact of the temporal parameter introduction has not been made and we intend to do it in our future work. However, in the next section, we tried to make a less detailed analysis of our learning technique.

4.3.2. Performance Analysis of the Learning Step using THHMM

This section is devoted to succinct performance analysis of the learning step based on temporal and hierarchical hidden Markov model. The analysis of the learning curve of the interest scenario training by the THHMM shows the learning performance and the efficiency of this HMM version which integrates two reasoning abstractions layers like the management of temporal constraint and the internal organization of event for the sub-scenarios management. The analysis of the learning curve of the interest scenario training by the THHMM shows the learning performance and the efficiency of the HMM. It's helped us to see that around eighteen and twenty, the learning curve peaked and that remained constant until the end of learning process. We noted that the Gaussian law requires a large amount of data to estimate its parameters, which is not the case of the Poisson distribution. In this work, a deep analysis of the impact of the temporal parameter introduction has not been made. We intend to do it in our future work. Meanwhile, Figure 10 shows the results of learning scenarios by the temporal HMM for scenario *"The doctor visit the patient at the bed 1"*.

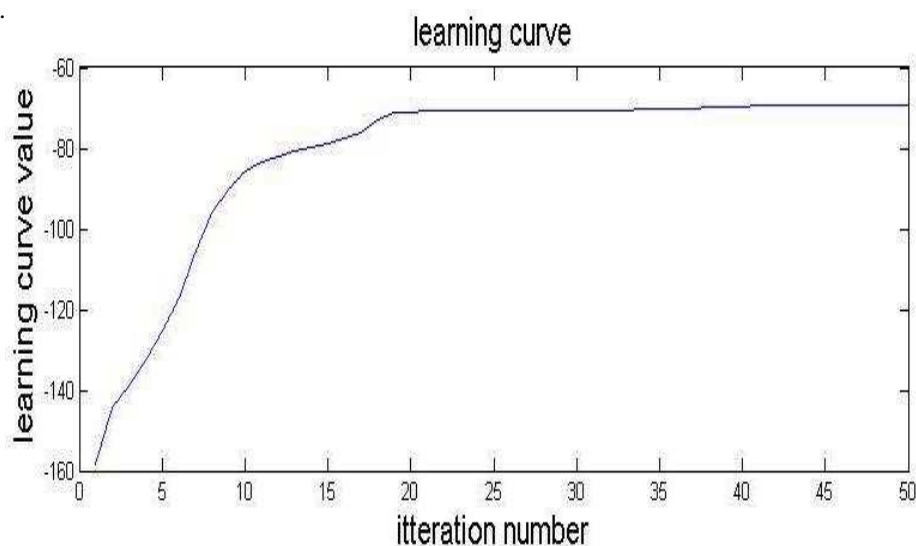


Figure 10: An example of THHMM learning curve for scenario *"The doctor visit the patient at the bed 1"*.

Figure 11 and Figure 12 show the results of the statistical estimation of the parameters distribution for modeling the amount of time spent by the model in a state e . The experiments allowed us to see that the estimation of Poisson law parameters requires less data for learning while the Gaussian distribution requires a larger amount

of data. The advantage of the Gaussian is that these parameters are more stable once estimated on a large amount of data while the Poisson law parameter can rapidly diverge in the presence of data not having the same or similar data structure used for the learning step.

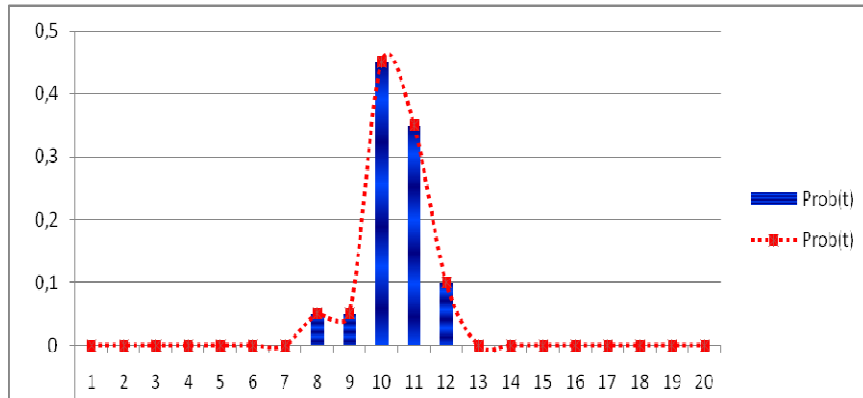


Figure 10: An example of THHMM learning curve for scenario “The doctor visit the patient at the bed 1”.

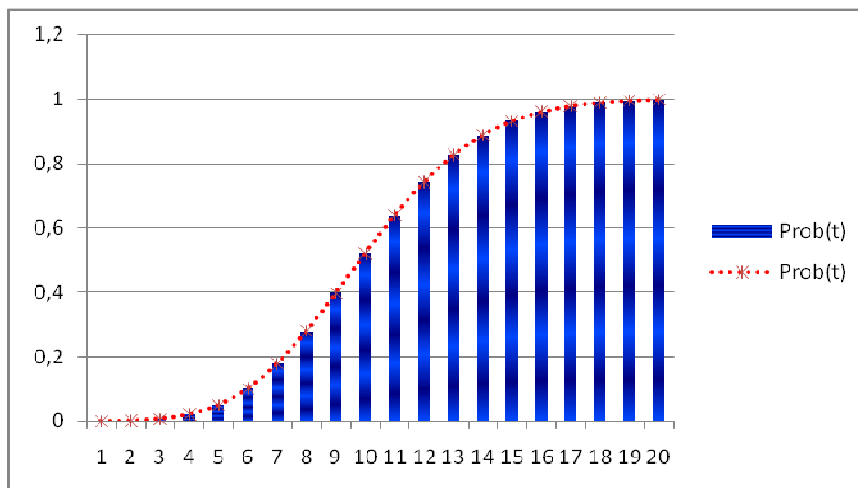


Figure 11: An example of THHMM learning curve for scenario “The doctor visit the patient at the bed 1”.

4.4. Scenario Recognition: results and analysis

4.4.1. Recognition of the Scenario "Activities of a Doctor"

We defined two categories scenario of the class AD "*Activities of a Doctor*" have been tested during the experimentation phase: The category scenario "*The doctor provides a patient care*" is divided into six sub-scenarios which are summarized in the scenario "*doctor provides with medical care at different beds*". For instance, we have "The doctor makes a patient care at bed 1, or 2, ..., 6". The scenario "*the doctor makes a general round visit in the ward*" is recognized through an HMM with nine states.

Referring to statistics results of scenarios recognition tests performed on our system as presented in Table 1, these results show that HMMs with explicit state duration are more efficient for modeling scenario recognition in a medical environment compared to classical HMMs. The performance gain is about 85.00% for the different beds (e.g. " the attendance of doctor at bed 1") with Poisson distribution (the best recognition rate for the scenario attendance of the physician in front of the bed1 is 80.80%). The preeminence of Gamma distribution for state duration can be attributed to its statistical proprieties and to the appropriateness of the data used for estimating its parameters. The performance with Gamma distribution gain is about 90.50% for the different beds (the best recognition rate at bed one is 94.99%). Find attached the reporting of the results of the recognition of the rate reached following an HMM model and the distribution of the explicit state duration in the table in appendix.

4.4.2. Recognition of the Scenario "Activities of a patient"

We tested two categories of scenarios of the class AP "*Activities of the Patient*" during the phase of experimentation stage. The scenario "*the patient convulses*", is based on the *restlessness detection* of the patient which is modeled and recognized through an HMM with four states. In the case of the scenarios related to the activities of the patient we have the scenario "*The patient feels pain*" which is not only based on the motion of the limbs (hands and feet) of the patient, but also on the motion of his head and even more effectively by combining his facial expressions (emotions recognition). Two HMM were used to recognize these scenarios, one with four states for limbs motion recognition (ie scenario "*the patient convulses*") and the other with two states to recognize head motion (Head Left : H_L and Head Right: H_R). Combining the results of basic actions recognition with each of two HMM, we have detected if the patient feels the pain scenario.

Scenario types	Classic HMM	THHMM Using Poisson Distribution	THHMM Using Gaussian Distribution)
The doctor makes a patient care	65,00	85,00	90,50
The doctor makes a round visit in the ward	60,60	80,00	90,00
The patient convulses	50,50	75,80	85,00
The patient feels pain	55,00	80,60	85,00

Table I. Some results of scenario recognition rate with temporal hierarchical HMM (%).

5. Conclusion and Future Works

In this paper, we propose a novel monitoring system for the recognition of the human action in medical videos of the intensive care units applied to the video selection in order to design MBB (Medical Black Box). Our approach is based on temporal reasoning by using an extended HSMM. To this end, the temporal Baum-Welch and the temporal forward-backward algorithm of the Semi-Hidden Markov Models are used respectively for learning the important scenarios and recognize them. Several tests were carried out by altering on medical videos gathered by the visual monitoring system at the ICU at cardiology section. They showed promising results. In our future work, a thorough analysis for the learning phase and the impact of duration constraints introduced on the state of the HMM during the training step. We will develop a futuristic monitoring system in ICU using multi sensors in order to achieve information fusion and multisensors video data processing (network of cameras) and contact sensor (physiological sensor, medical monitoring equipment medical). Finally, cooperation and collaboration between different types of sensors will also be explored.

8. Acknowledgments

We thank the cardiology section team at CNHU (Centre National Hospitalier Universitaire) who provided the ICUs video dataset from the visual-surveillance system in hospital. We thank also Professor Hippolyte AGBOTON in charge of cardiology section research team.

6. References

- [1] A. Benouareth, A. Ennaji, and M. Sellami, "HMMs with Explicit State Duration Applied to Handwritten Arabic Word Recognition", *ICPR*, 2006.
- [2] Ahouandjinou A. S. R. M., E. C. Ezin, C. Motamed, and P. Gouton, "An approach to correcting image distortion by self calibration stereoscopic scene from multiple views", *SITIS*, 2012.
- [3] A-S Silvent., M. Dojat and C. Garbay, "Multi-level temporal abstraction for medical scenario construction", *Special Issue of thrd Internationnal Journl of Signal Processing and Adaptatove Control*, 2004.
- [4] C. Meredith and J. Edworthy, "Are there too many alarms in the intensive care unit? An overview of the problems", *Journal of Advanced Nursing*, vol. 21, pp. 15-20, 1995.
- [5] C. Motamed, "Motion detection and tracking using belief indicators for an automatic visual-surveillance system", *Image Vision Comput.* 24(11): 1192-1201, 2006.
- [6] Duong, Thi, Phung, Dinh, Bui, Hung and Venkatesh, Svetha, "Efficient duration and hierarchical modeling for human activity recognition", *Artificial intelligence*, vol. 173, no. 7-8, pp. 830-856, Elsevier BV, Amsterdam, Netherlands, 2009.
- [7] Duong, Thi V., Bui, Hung H., Phung, Dinh Q. and Venkatesh, Svetha, "Activity recognition and abnormality detection with the switching hidden semi-Markov model", *Proceedings of the IEEE Computer Society Conference CVPR*, 2005.
- [8] Luhr, Sebastian, Venkatesh, Sventha, West, Geoff and Bui, Hung H., "Duration abnormality detection in sequences of human activity", *Curtin University of Technology, Perth*, 2004.

- [9] Luhr Sebastian, Venkatesh, Svetha West, Geoff A. W. and Bui Hung H., "Explicit state duration HMM for abnormality detection in sequences of human activity", *PRICAI 2004*.
- [10] L. R. Rabiner, "A Tutorial on HMMs and Select Applications in Speech Recognition" *IEEE, Vol 77, N° 2, pp: 257-286*, 1989.
- [11] M. Dojat and L. Brochard., "Knowledge-based Systems for Automatic Ventilatory Management," *Saunders: Philadelphia, 379-396*, 2001.
- [12] M.J. Russel, and R.K Moore, "Explicit Modeling of State Occupancy in Hidden Markov Models for Speech Recognition" , *ICASSP*, 1985.
- [13] M. Stacey, and C. McGregor, "Temporal abstraction in intelligent clinical data analysis: A survey", *Artificial Intelligence in Medicine 39, 1-24, Elsevier*, 2007.
- [14] R. Fonseca1, and P. P. Rodrigues, "the importance of system integration in intensive care units: A review", *International Conference on Health Informatics, "HEALTHINF*, 2011.
- [15] S. E. Levinson, "Continuously Variable Duration Hidden Markov Models for Automatic Speech Recognition. Computer", *Speech & Language, Vol 1, N° 1, pp: 29-45*, 1986.
- [16] S. Nicholas, A. R. Giannareas, V. Kern, A.V. Trevino, F-M. Margaret and A. L. King, "Insup Lee: Smart alarms: multivariate medical alarm integration for post CABG surgery patients", *IHI: 533-542*, 2012.
- [17] V. Seshadri, *The Inverse Gaussian Distribution: A Case Study in Exponential Family*, Oxford Science Publications, 1993.
- [18] Y. Bengio and P. Frasconi, "Input/Output HMMs for sequence processing" , *IEEE Transactions on Neural Networks*, 1995.
- [19] H.H. Bui, Q. Phung, and S. Venkatesh., Hierarchical Hidden Markov Models with General State Hierarchy, *In Proceedings of the Nineteenth National Conference on Artificial Intelligence*, San Jose, California, pp. 324-329, 2004.
- [20] M. Duong, T. Vu., H.H. Bui, H. Phung, Q. Dinh and V. Svetha Efficient duration and hierarchical modeling for human activity recognition, *Artificial intelligence*, vol. 173, no.7-8, Elsevier BV, Amsterdam, Netherlands, pp. 830-856, 2009.
- [21] S. Fine, Y. Singer and N. Tishby. The Hierarchical Hidden Markov Model: Analysis and Applications, *Machine Learning*, Vol 32, N°1, pp.41-62, 1998.
- [22] A. S. R. M. Ahouandjinou. Reconnaissance de scénario par les Modèles de Markov Cachés Crédibilistes. Application à l'interprétation automatique de séquences vidéo médicales. *PhD thesis, Université du Littoral de la Côte d'Opale*, Calais-France, 2014.