論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Intelligent methods for capnogram feature extraction applied to asthma classification
著者(和文)	PomaresBetancourt
Author(English)	Janet Pomares
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第9673号, 授与年月日:2014年9月25日, 学位の種別:課程博士, 審査員:廣田　薫,室伏　俊明,宮下　英三,木賀　大介,小野　功
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第9673号, Conferred date:2014/9/25, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Intelligent Methods for Capnogram Feature Extraction Applied to Asthma Classification

Janet Pomares Betancourt

Supervisor: Prof. Kaoru HIROTA

Doctoral Thesis

Tokyo Institute of Technology
Interdisciplinary Graduate School of Science and Engineering
Department of Computational Intelligence and Systems Science

September 2014

Intelligent Methods for Capnogram Feature Extraction Applied to Asthma Classification

喘息分類のためのカプノグラム特徴量抽出における知的手法

Janet Pomares Betancourt
ジャネト　ポマレス　ベタンクルト

Supervisor: Prof. Kaoru HIROTA

Doctoral Thesis

Tokyo Institute of Technology
東京工業大学
Interdisciplinary Graduate School of Science and Engineering
大学院総合理工学研究科
Department of Computational Intelligence and Systems Science
知能システム科学専攻

September 2014

Abstract

A method for ECG and capnogram signals classification is proposed based on fuzzy similarity evaluation, where shape exchange algorithm and fuzzy inference are combined. It aims to be applied to quasi-periodic biomedical signals and has low computational cost.

On the experiments for atrial fibrillation (AF) classification using two databases: MIT-BIH AF and MITBIH Normal Sinus Rhythm, values of 100%, 94.4%, and 97.6% for sensitivity, specificity, and accuracy respectively, and execution time of 0.6 s are obtained. As for the capnogram database, experiments for asthma classification produce 83.3%, 80.9%, and 81.5 for sensitivity, specificity, and accuracy respectively, and execution time of 0.3 s.

The proposal is capable of been extended to classify different diseases, from ECG and capnogram signals, such as: Brugada syndrome, AV block, hypoventilation, and asthma among others to be implemented in low computational resources devices.

To improve the previous classification results, a method for feature extraction of the capnogram based on wavelet decomposition for asthma classification is proposed. The method requires low computational cost and shows adequate performance for a real time classification of asthma severity. The validation experiments on 23 capnograms, collected from an Asthma Camp in Cuba, suggest 97.39% of accuracy as best classification result using a SVM classifier. An estimation of the execution time for a

physiological multiparameter monitor is obtained, showing an average of 8 sec to determine the suitable features. The proposal aims to be a part in the decision support system for asthma classification that is under development by research group of Tokyo Institute of Technology (TITECH) and Tokyo Dental and Medical University (TMDU).

To integrate the two proposals in this investigation and evaluate their influence in the classification result, two different architectures for the decision support system are evaluated. Both architectures show similar classification results in terms of correct rate. Execution time represents the most significant difference between architectures and could define the kind of application where to be employed.

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Kaoru Hirota, for accepting me in his lab. I am very grateful for his teaching and dedication to me on the way to become a Ph.D. I would like to thank Professor Toshiaki Morofushi, Professor Eizo Miyashita, Professor Daisuke Kiga, and Professor Isao Ono for their advices and recommendations that served as a guide for my research.

I would like to thank Dr. Fang-Yan Dong, assistant professor at Hirota Lab, for her advices and encouragement that were important since the very beginning of my research. Thank to Hoshino-san for her invaluable help in crucial moments.

I would also like to thank Professor Yasunari Miyasaki and Dr. Meiyo Tamaoka for their precious time and insightful discussions.

My grateful goes also to Dr. Marianela Otaño for providing the capnogram data and her passion for the asthma-related research.

Thank to my goods friends at Hirota Lab, Dr. Chastine Fatichah, Dr. Martin Tangel, Mr. Adrian Garcia, and Mr. Yan Fei for their advices and research collaboration. I would also like to thank all Hirota Lab members for their support and assistance.

I am also grateful to Mr. Andres Ballester, Dr. Kimura, and Ms. Alis Chacon for their crucial help to achieve that I could finish the doctorate.

Finally, I would like to thank to my family; thank to my parents, Marina and Guido, my sister Mabel, and my grandmother Lucia for their support, encouragement, and confidence to me. Thank to my love, Alejandro, for his advices, insightful talks, his constant support, and love to me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Asthma is a chronic inflammatory disease of the respiratory airways. It is a problem worldwide with an estimated of 300 million affected individuals of all ages. When asthma is controlled, there should be merely sporadic recurrence of symptoms. However, if it is uncontrolled, can place severe limitations on daily life and is sometimes fatal [1]. Moreover, the asthmatic patient undergoing surgery is at risk for perioperative morbidity and mortality. Even though the incidence of bronchospasm is about 2%, it is considered a life-threatening event in anesthesia practice, since the 90% of these patients involved severe brain injury or death [2]. Thus, objective diagnosis and assessment of the severity of asthma is important for controlling the disease. Besides that, early detection of the incidents during the use of general anesthetic is essential to prevent further reaction or a fatal outcome.

In Cuba, asthma is considered a common illness. It affects about 10% of the population and the prevalence was 92.2 individuals per 100 000 habitants in 2011. In addition to that, 49% of the population suffers some type of allergy which increases the possibility of been affected by asthma. Other risk factors of particular importance in Cuba are the environmental factors, e.g. the meteorological changes. A study reports the occurrence of peaks in the number of asthmatic patients that go to the hospital during the transition between seasons [3]. Also it has been reported that in central areas which are more exposed to air pollution the frequency of respiratory symptoms and diseases is higher [4].

Asthma is a burden for the Cuban government which dedicate millions USD per year for treatments and hospitalizations of asthmatic patients, e.g. this amount was 61 million USD in 2011, according to the Institute of Hygiene and Epidemiology of Cuba. To this must be added that, over 2500 people with this condition is absent from school or work daily, around 3500 attend the emergency medical services, and about fifty are admitted to hospitals.

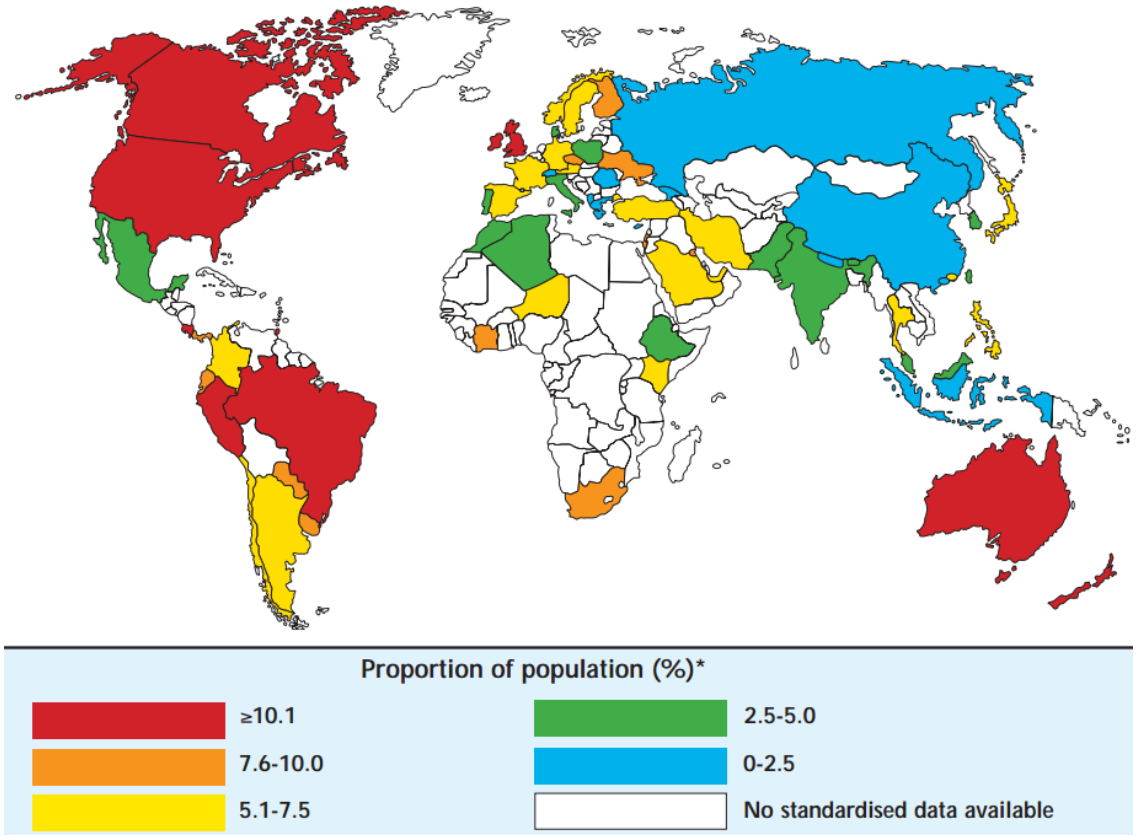


Fig. 1.1. World map of the prevalence of clinical asthma

1.1 Asthma Physiology

The respiratory system is involved in the intake and exchange of oxygen (O_2) and carbon dioxide (CO_2) between the human and the environment. Its primary function is

to obtain oxygen to be used by the body's cells and eliminate carbon dioxide that cells produce.

Respiration takes place in the respiratory organs called lungs. The passage of air into the lungs to supply the body with oxygen is known as inspiration, and the passage of air out of the lungs to expel carbon dioxide is known as expiration; the complete process is called breathing or ventilation. The anatomical features of the respiratory system include trachea, bronchi, bronchioles, lungs, and diaphragm (see Fig. 1.2).

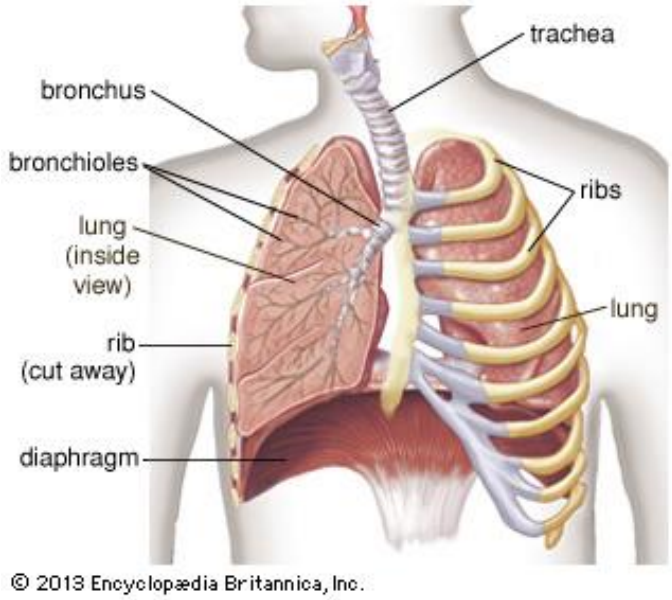


Fig. 1.2. Anatomical features of the respiratory system.

The air enters into the body through the nose, the gateway into the nasal cavity. The cavity surface has a mucus membrane which heats and moistures the air, and also trapped the particles. Leaving the nasal cavity behind, the air enter the pharynx (throat), continuing through the larynx that is the passage way from the pharynx into the traquea. At the bottom, the traquea divides into two tubes, the right and the left bronchus. The bronchus branches forming a tree of airway passages within each lung. From the smallest bronchus branches the bronchioles which end in clusters of microscopic sacs called the alveoli. Alveoli are surrounded by capillaries within which blood travels. The O_2 molecules diffuse across the alveoli's membrane entering a capillary where it

matches the blood cells entering the circulatory system. Meanwhile, the CO_2 wasted by the cells diffuses into the bloodstream. When rich the capillaries surrounded the alveoli, CO_2 diffuses and start the way out of the respiratory system using same way the O_2 came in. This process of gases exchange (O_2 and CO_2) between the gaseous external environment and the blood is called diffusion [5].

Asthma is a respiratory disease which inflames and narrows the airways. Bronchioles expand when the air is warm, moist, and free of irritants and allergy causing substances called allergens. When air is cold or dry, or contents irritants or allergens, bronchioles contract. Certain substances can cause the inflamed airways to overreact even more resulting in an asthma attack. During an asthma attack or bronchospasm, the muscles surrounded the airways tighten and airway wall becomes more swollen. Airways also produce thick mucus that fill up them making difficult to breathe. Fig. 1.3 shows the differences between a normal bronchiole and an asthmatic bronchiole.

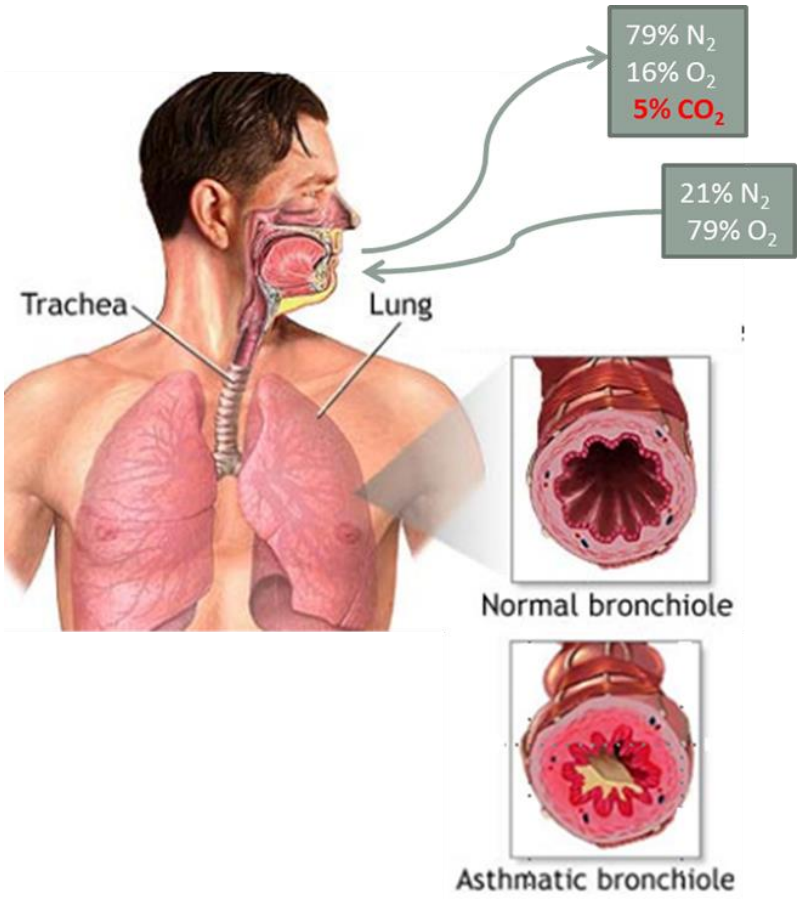


Fig. 1.3. Cross-section of the bronchiole. This structure inside the lungs inflames causing an asthma attack.

1.2 Current Methods for Asthma Diagnosis

A correct diagnosis of asthma is crucial if appropriate treatment is to be given. The established methods for diagnosis, assessment of the severity of asthma and of the degree of improvement following treatment are: clinical, such as auscultation and respiratory frequency estimation, and the measurements of lung function.

Clinical methods based on the presence of symptoms, are usually used for the diagnosis of asthma. Asthma symptoms, for example, wheezing, cough, breathlessness,

and chest tightness may be episodic and their significance may be disregarded by patients and physicians. Because they are non-specific, they may result in misdiagnosis. In addition, in some people with asthma, the most usual abnormal finding, wheezing, may be absent. Assessment of symptoms by physicians may be inaccurate [6].

The measurements of lung function provide confirmation of the diagnosis of asthma, assessment of the severity of the airflow limitation, its reversibility, and its variability. Two methods of lung function measurement have gained the widespread acceptance for it use in patients over 5 years of age. These are the spirometry, involving two major parameters, FEV_1 (Forced Expiratory Volume in 1 sec) and FVC (Forced Vital Capacity), and PEF (Peak Expiratory Flow).

FEV_1 and FVC require that the patient powerfully expire and airflow is measured using a spirometer (Fig. 1.4). Spirometry is effort-dependent; therefore the correct completion of this task not only depends on the patient's capacity to do so, but on the clinician's ability to guide the procedure as well.

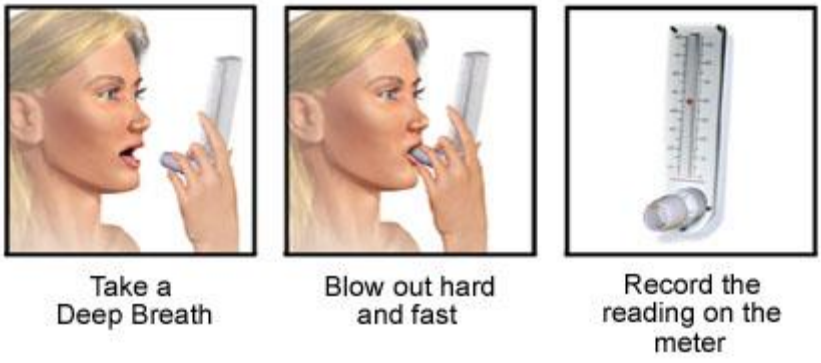


Fig. 1.4. Illustration of a spirometer and way of use.
Source: http://www.drugs.com/cg/how-to-use-a-peak-flow-meter.html

Peak expiratory flow is measured using a peak flow meter and can be important for diagnosis and monitoring of the asthmatic condition. PEF can underestimate the degree of airflow limitation particularly as it worsens. Because values of PEF, obtained from different peak flow meters, vary and the range of predicted values is too wide, should be preferable to compare the patient's values with his own previous best measurements

made using his own peak flow meter. As spirometry, PEF is effort-dependent and requires instruction to be reliably measured. Fig. 1.5 shows two types of peak flow meter USB PC-based spirometer (a), ultrasound based desktop spirometer (b), and an illustration on how the measurement is performed (c).

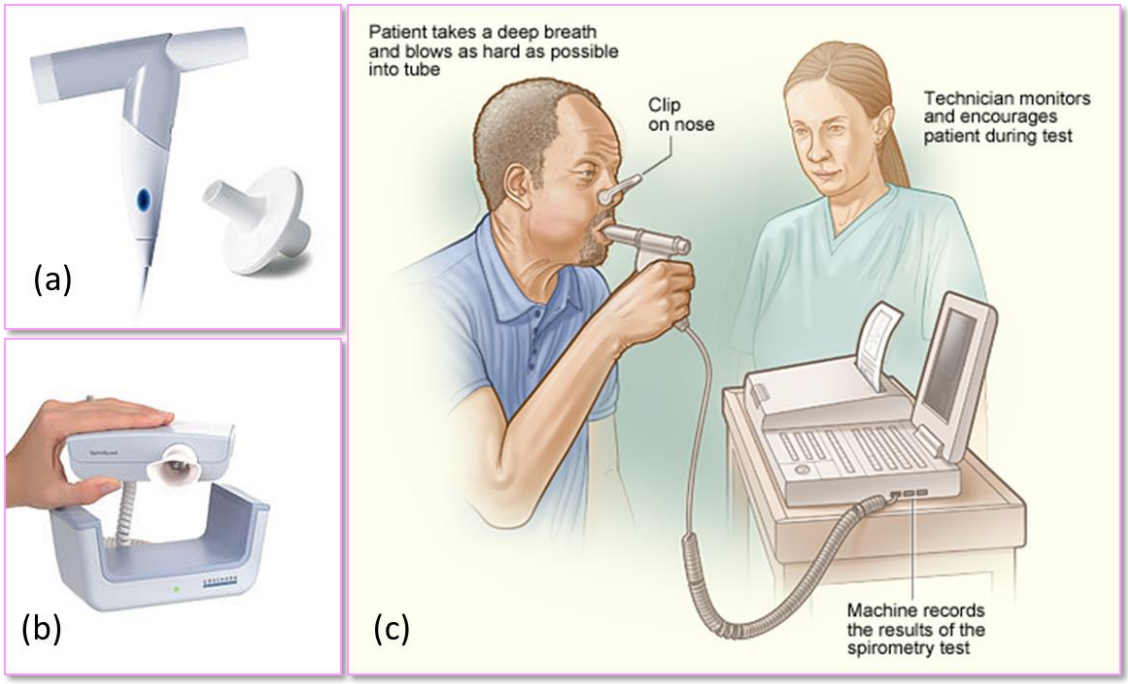


Fig. 1.5. Two types of peak flow meter: USB PC-based spirometer (a), ultrasound based desktop spirometer (b), and an illustration on how the spirometry is performed (c).
Source: (a, b) http://en.wikipedia.org/wiki/Spirometer
(c) http://www.nhlbi.nih.gov/health//dci/Diseases/lft/lft_types.html

1.3 Diagnostic Challenges

Several challenges exist for diagnosing and monitoring asthma using the established methods. These are related with the patient's group of age, its physical

condition, and the medical procedure the patient is undergoing. Physical condition here refers to the degree of airway obstruction.

The diagnosis of asthma in children 5 years old or younger relies mainly on the evaluation of symptoms. But, some symptoms of asthma, as wheezing and cough, are very common in children, particularly in those under age 3. Moreover, the use of spirometry is difficult due to the forced nature of expiration maneuvers.

In older children and adults, the established methods, in most of cases, will confirm the diagnosis of asthma. However, alternative diagnosis should be considered because asthma can be found in association with other common disease, e.g. upper airway obstruction and inhaled foreign bodies, vocal cord dysfunction, other obstructive lungs diseases, etc. This complicates the diagnosis as well as the assessment of the asthma severity and control.

In elderly individuals, the diagnosis of asthma is also complicated because of different factors. Frequently ancients overlook symptoms; they accept them as normal in old age. Also, comorbid diseases are commonly present. On the other hand, the use of spirometry is not recommended in old patients.

The monitoring of the very dyspnoeic patient relies, most of cases, on the clinical judgment since they cannot undertake the effort-dependent lungs function tests [7].

For the case of the asthmatic patient under a general anesthesia procedure, either spirometric or clinical methods are not suitable. In this situation, is very difficult to detect or evaluate the asthma attack since the symptoms may be hidden under the effect of the anesthesia.

1.4 Novel Methods for Asthma Diagnosis

Impulse Oscillometry (IOS) is an alternative modality to the conventional pulmonary function tests. It measures the impedance of the respiratory system by superimposing short pulses of air pressure on the subject's tidal breathing. In contrast to forced spirometry and PEF, it minimizes the demands to the patient and only requires the subject breaths normally through the mouth [8, 9]. IOS has demonstrated its utility in the early diagnosis of asthma especially on children under 5 years old. Fig. 1.6 shows a child undertaken the IOS test. Moreover, although the usefulness of IOS in assessing the severity of asthma has been reported, there is little data on the potential of IOS measurements to classify asthmatic airway obstruction stages systematically [10].

Fig. 1.6. A child is undertaken the IOS test performing a normal breathing through the mouth. Source: http://www.acp-asia.com.hk/web/product.php?mid=3&actid=menu_prod_77_20_20_bu&lang=1

Capnography is an essential tool for monitoring the cardiorespiratory function providing a real-time assessment of ventilatory status [11]. This non-invasive and accurate method of monitoring directly shows the elimination of CO_2 by the lungs and

indirectly reflects the production of CO_2 by tissues and CO_2 circulatory transport to the lungs [12]. Information obtained through capnography includes: CO_2 production, lung perfusion, alveolar ventilation, respiratory patterns, and elimination of CO_2 from the anesthesia breathing circuit and ventilator.

Capnography is usually presented as a graph of expiratory CO_2 plotted against time named capnogram. This waveform is generated by a device called capnograph. Meanwhile, capnometry is the measurement and numerical display of maximum inspiratory and expiratory CO_2 concentrations during a respiratory cycle. The device that performs the measurement and displays the reading is the capnometer. However, the capnograph or capnogram is vastly preferable in anaesthetic practice where breath-by-breath waveform needs to be displayed to permit continuous monitoring and analysis [13].

There are several methods to measure CO_2 levels including infrared spectrography, Raman spectrography, mass spectrography, photoacoustic spectrography, and chemical colorimetric analysis. The infrared method (IR) is most widely used and most cost-effective.

Infrared rays are given off by all warm objects and are absorbed by non-elementary gases (i.e. those composed of dissimilar atoms), while certain gases absorb particular wavelengths producing absorption bands on the IR electromagnetic spectrum. The intensity of IR radiation projected through a gas mixture containing CO_2 is diminished by absorption; this allows the CO_2 absorption band to be identified and is proportional to the amount of CO_2 in the mixture. This can be explained by the Beer-Lambert law which relates the absorption of the radiation flowing through a gas with certain properties of it, as provided in equation (1).

$$I = I_0 \cdot e^{-\alpha \cdot l \cdot C} \tag{1}$$

where:

I_0*:* is the intensity of the incident radiation in the gas,

I: is the intensity of the radiation emerging the gas,

l: is the optical path length traversed by the radiation,

α: is the absorption coefficient of the gas, and

C: is the gas concentration in the medium traversed by the incident radiation.

The value of the absorption coefficient α depends on the wavelength of the incident energy, and for the case of CO_2 is a maximum for a wavelength of 4.26 μm. From the above, it can be inferred that in a gaseous mixture CO_2 absorbs radiation of a wavelength of 4.26 μm which will not be affected by other gases in the mixture. Determining the intensity of the radiation absorbed by the mixture is possible to calculate the concentration of CO_2 in it [15]. CO_2 concentration in the respiratory gases is measured in % or as a function of the partial pressure in mmHg.

There are two types of capnographs: sidestream capnographs and mainstream capnographs. In sidestream capnography, the CO_2 sensor is located in the main unit itself (away from the airway) and a tiny pump aspirates gas samples from the patient's airway through a capillary tube into the main unit. Sidestream gas analyzers utilize a long sampling plastic tube connected to an adapter in the breathing circuit or a nasal catheter.

Sidestream capnographs have a major advantage: they allow monitoring of spontaneously breathing non intubated subjects, as sampling of the expiratory gases can be obtained from the nasal cavity using nasal adaptors. This feature enables monitoring of expired CO_2 in subjects receiving simultaneous oxygen administration using nasal cannulae.

In the mainstream capnograph, an airway adapter containing the CO_2 sensor is located in line with the respiratory gas stream. The IR rays traverse the respiratory gases to an IR detector within the airway adapter. Thus there is no need for gas sampling and scavenging. The response time is faster: no gas is subtracted from the breathing circuit, no sampling pumps or other suction devices add complexities to the mechanical system,

and there is no uncertainty caused by the rate of gas sampling. Fig. 1.7 shows schemes of mainstream and sidestream capnographs.

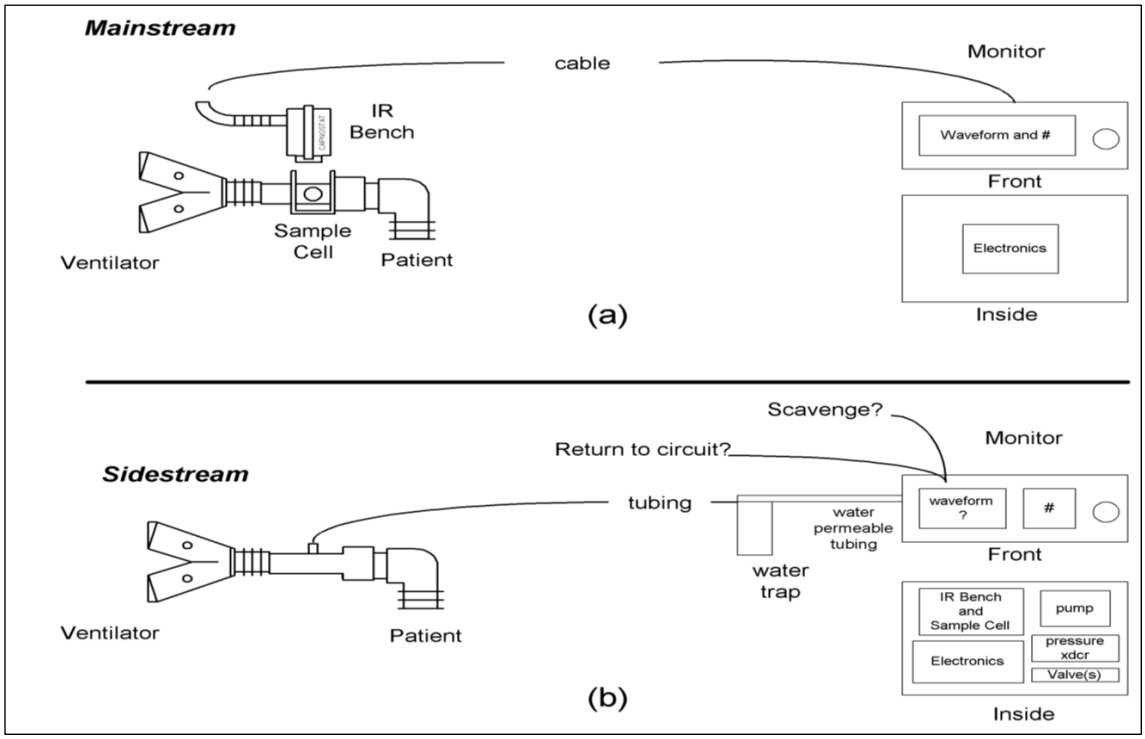


Fig. 1.7. Mainstream capnograph in (a) shows its simplicity compared with sidestream capnograph in (b).

The measurement system comprises a CO_2 detector including a source of infrared radiation (IR) which can be composed of a LED emitter, for an incandescent lamp, or an IR emitter which employs the principle of emission of a black body, all with a suitable reflector to allow focusing the radiation on the surface of IR energy detector [15].

Between emitters of the same type the characteristics of the emitted radiation are not identical, in the same way that no two identical reflectors of same type exit. Therefore, an individual calibration of each IR emitter-detector assembly is essential to obtain a reliable determination of the CO_2 concentration measurement.

Clinical Applications of Capnography

CO_2 is produced in the body by cellular metabolism, conveyed by the circulatory system to the lungs, excreted by the lungs and transported by the breathing system.

Therefore, changes in respired CO_2 may reflect alterations in metabolism, circulation, respiration, the airway, or breathing system.

An increase in end-tidal CO_2 is a reliable indicator of increased metabolism only in mechanically ventilated patients. Some metabolic causes of the CO_2 rising include: shivering, convulsions, administration of blood, and others.

Related to the circulatory system, applications of capnography are various. In general, a decrease in end-tidal CO_2 is seen with a decrease in cardiac output if ventilation remains constant. In the same way, a sudden increase of end-tidal CO_2 during resuscitation is an early sign that spontaneous cardiac output has been repaired.

The major application of capnography involves the respiratory system. The capnographic signal is mainly used in anaesthesia practice for the immediate verification of tracheal intubation beyond doubt by the immediate and continuous presence of metabolic CO_2 in the expired gas.

A CO_2 monitor is used to monitor respiratory rate of exhaled CO_2 in nonintubated patients breathing spontaneously. Apnoea or airway obstruction can be detected.

Another major use of capnography has been to determine the correct ventilatory needs during controlled ventilation. This use is easily extended to continuous monitoring of spontaneous ventilation and to monitoring of appropriate ventilation during titration of anaesthetic agents that depress ventilation.

Nowadays, capnography is also used to monitor the asthmatic condition. Unlike spirometry and peak expiratory flow, capnogram is taken as the patient is breathing as comfortably as able without instructions and without interfering the treatment (Fig. 1.8) [6, 8, 17, 18, 19]. Capnography can be used on subjects that are either sleeping or awake, and allows various new applications to be foreseen such as detection of nocturnal attacks, evaluation of the duration of action for bronchodilator drugs, intra- and postoperative monitoring of asthmatic patients, and dynamic bronchial provocation tests, especially in children.

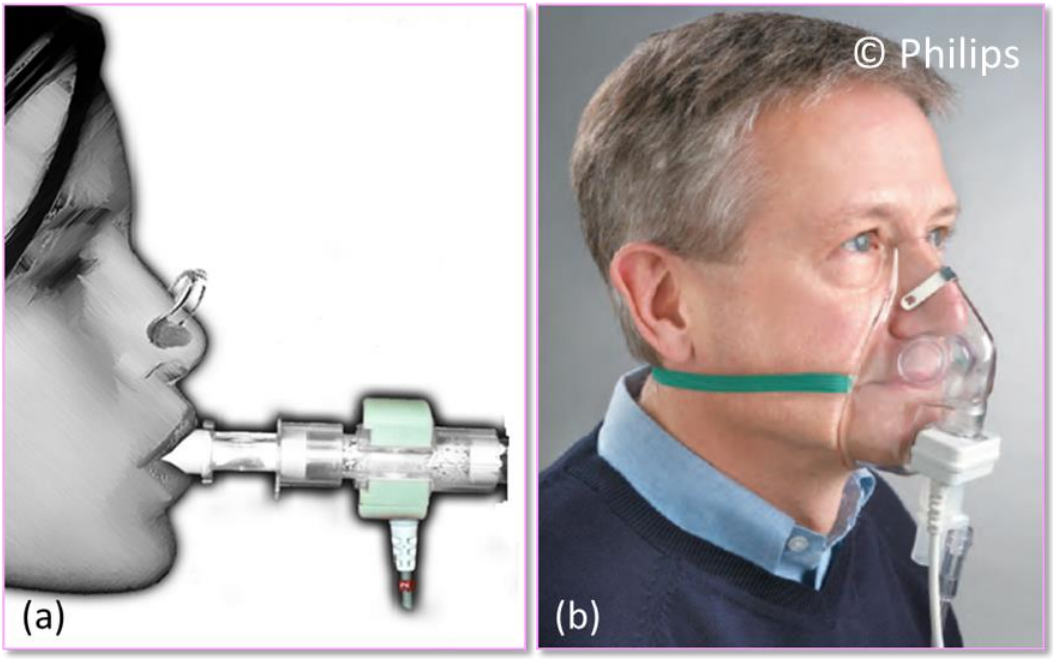


Fig. 1.8. The capnogram is taken from two patients. In (a), a nasal clip is placed on the patient's nose in order to limit the airway through the mouth [15]. In (b), the patient is using a mask that allows him to breathe normally while receiving O_2.

Source: (b) http://capno2mask.respironics.com/

1.5 Capnogram Signal Characterization

The normal capnographic waveform (Fig. 1.9) shows a square wave pattern [7], marked by alternating inspiratory and expiratory periods. Expiratory period consists of three successive phases. Phase I is the beginning of the expiration where the gas (free of CO_2) is expired from the anatomical dead space. Phase II, marked by a rapid rise of the CO_2 concentration, corresponds to the expiration of the mixed gas from the alveoli. Phase III or plateau phase, reflecting the elimination of alveolar air, results in a peak at the end of tidal expiration which immediately precedes the start of the inspiration.

These three phases are separated by two transitions angles: α and β. The α is the angle between phase II and phase III and it increases as the increases of the slope of phase III that is dependent on the V(ventilation)/P(perfusion) status of the lung. Normally ranges between 100 and 110 degrees. The β is the almost 90 degrees angle between phase III and the limb of the inspiration.

It is known that in patients undergoing asthma the shape of the waveform changes. The deformation of the normal curve is marked by the loss of verticality of phase II, opening of α angle and the increased inclination of phase III. In severe cases, the capnogram takes on a "shark fin" appearance, as shown in Fig. 1.9.

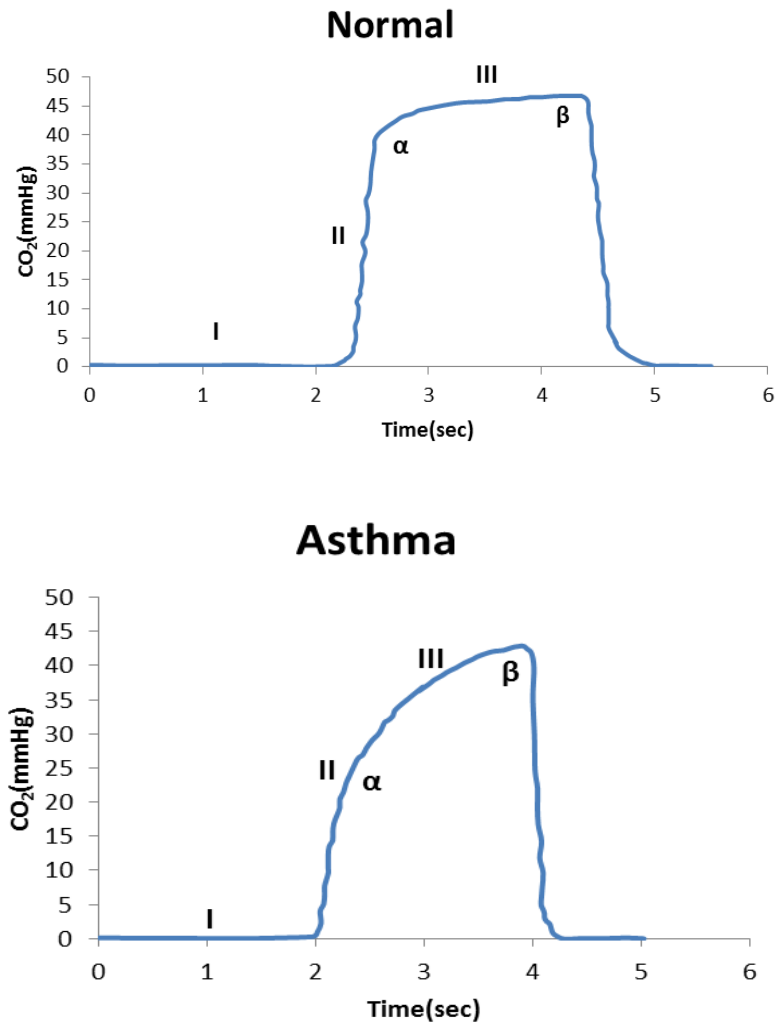


Fig. 1.9. Normal (top) and asthma (bottom) capnografic waveform.

1.6 Classification of Asthma

There are different types of classification of asthma: by etiology, by phenotype, classification of asthma control, and classification of asthma severity [1]. The classification according etiology is limited because there are patients in whom no environmental causes can be identified. The diversity of manifestations of asthma and its response to treatment is often described in terms of phenotypes. Several different clinical prototypes are recognized and the search of factors that could explain these patterns continues.

Asthma control classification refers to the control of the manifestation of the disease. Several standardized measures for assessing clinical control of asthma have been developed, which evaluate the effectiveness of the treatment and provides numerical values to distinguish different levels of controls. They have been promoted for use not only in research but also for patient care. Some are suitable for self-assessment of asthma control by patients. Their value in clinical use has yet to be demonstrated but will be expected to become evident in the near future. There is interest in controlling not only the manifestations of asthma, but also the inflammatory and patho-physiological features of the disease. In addition, these features may be predictors of future risk of exacerbations and deterioration in lung function, independent of the level of clinical control of the patient. Fig. 1.10 presents the levels of asthma control for assessment of clinical control and future risk of exacerbations.

Figure 2-4. LEVELS OF ASTHMA CONTROL

A. Assessment of current clinical control (preferably over 4 weeks)

Characteristic	Controlled (All of the following)	Partly Controlled (Any measure present)	Uncontrolled
Daytime symptoms	None (twice or less/week)	More than twice/week	Three or more features of partly controlled asthma*†
Limitation of activities	None	Any	
Nocturnal symptoms/awakening	None	Any	
Need for reliever/ rescue treatment	None (twice or less/week)	More than twice/week	
Lung function (PEF or FEV_1)‡	Normal	<80% predicted or personal best (if known)	

B. Assessment of Future Risk (risk of exacerbations, instability, rapid decline in lung function, side-effects)

Features that are associated with increased risk of adverse events in the future include:

Poor clinical control, frequent exacerbations in past year*, ever admission to critical care for asthma, low FEV_1, exposure to cigarette smoke, high dose medications

*Any exacerbation should prompt review of maintenance treatment to ensure that it is adequate
† By definition, an exacerbation in any week makes that an uncontrolled asthma week
‡ Without administration of bronchodilator.
Lung function is not a reliable test for children 5 years and younger

Fig. 1.10. Levels of asthma control. There are two types of asthma control: control of the manifestations of asthma and control of future risk of exacerbations [1].

Classification of asthma severity is a subjective assessment, so there is potential for divergence of opinion between doctors, i.e., if a patient with symptoms of asthma goes to different doctors, he may be diagnosed different in terms of the disease degree. In addition, severity is not static, but may changes over time. Severity is measured more easily in a patient not receiving long-term control therapy. During the patient's initial presentation severity is classified as mild intermittent, mild persistent, moderate persistent, and severe persistent based on symptom frequency and either spirometric (FEV_1) or PEF [1, 16]. From this initial visit the appropriate medication and other therapies are decided. Once the therapy is initiated, the importance for clinical management is changed to assessment of asthma control.

In our research, we are focus on the classification of the severity of asthma based on the severity of the asthma attack according to the classification recommended by the Asthma Prevention and Management Guideline from Japan (see Fig. 1.11). Since our proposal analyses the capnogram for achieving the classification task, it can be used in most of the patients, from children under 5 years to elderly. In addition, it can be applied in different scenarios, such as in emergency rooms, surgery rooms for monitoring the asthmatic patient, and in the initial visit in a patient with newly diagnosed asthma. Moreover, it is desired to be employed by the patients for self-assessment of the disease at home.

Fig. 1.11. Asthma classification according to Asthma Prevention and Management Guideline, Japan, 2012.

1.7 Structure of the thesis

This research contributes to address the problem of automatic diagnosis and classification of asthma based on capnogram analysis. Two approaches for the study of this signal are realized, in the time domain and in the frequency domain. This dissertation is organized into three chapters as is explained next.

In chapter 2, the proposed method for capnogram classification based on fuzzy similarity is presented, which address the analysis of the signal in the time domain. The

extracted similarity features performs as inputs of a fuzzy inference to achieve the classification of the signals in *normal* or *disease*. In order to demonstrate the validity of the method, experiments are conducted using a capnogram database provided by an Asthma Camp in Cuba. In addition, two ECG databases Normal Sinus Rhythm and MIT-BIH Atrial Fibrillation are also used for testing. Since the method takes advantage of the quasi-periodic nature of the signal, it has potential to be applied to biomedical signals. Moreover, the proposal is capable of been extended to classify different diseases, from ECG and capnogram signals, such as Brugada syndrome, AV block, hypoventilation, and asthma among others, to be implemented in low computational resources devices.

In chapter 3, a different approach on analyzing the capnogram is presented. A method for feature extraction for asthma classification is developed based on wavelet decomposition. Classification experiments are performed in order to evaluate the proposal using support vector machine. Experiments results show the extracted features are suitable for the classification of asthma. An estimation of the execution time for a physiological multiparameter monitor is obtained, showing adequate performance for real time applications. The proposal is part of the decision support system for asthma classification that is under development by research group of Tokyo Institute of Technology and Tokyo Medical and Dental University.

In chapter 4, a decision support system for the classification of asthma severity in six degree is presented. To integrate the proposals in this investigation and evaluate their influence in the classification result, two different architectures for the decision support system are evaluated in this chapter. Experiments demonstrate the proposal has computation time appropriated for assessing asthma in real time. The proposal aims to be implemented in multiparameter monitors developed by the Central Institute of Digital Research in Cuba. Finally, chapter 5 presents conclusions and future works. Fig. 1.12 illustrates the structure of the thesis.

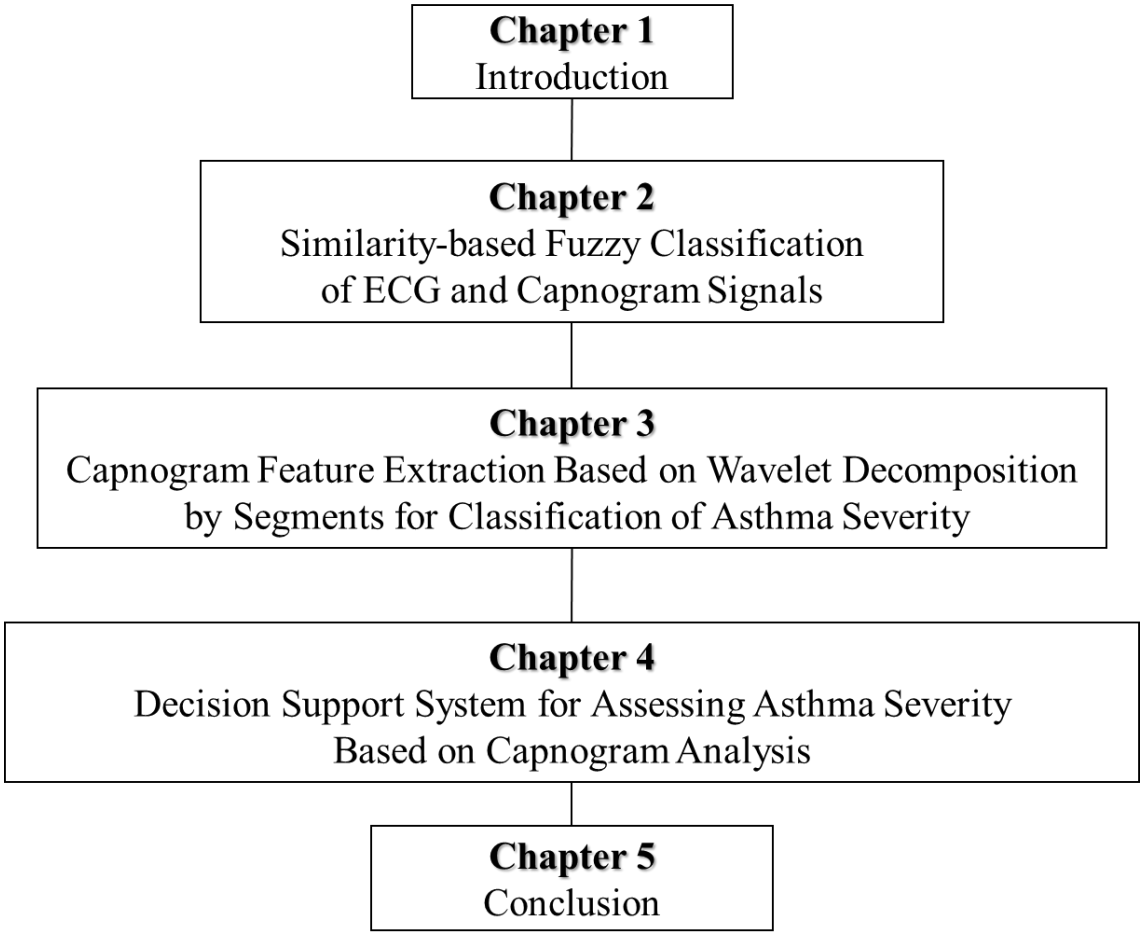


Fig. 1.12. Structure of the thesis.

Chapter 2

Similarity-Based Fuzzy Classification of ECG and Capnogram Signals

2.1 Introduction

Biomedical signals are usually measured in regular time intervals and repeat similar complexes cyclically. Therefore, they are considered quasi-periodic signals. The morphology analysis of these waveforms provides information related to diseases or the patient's condition [20].

Multiparameter monitors typically include the measurement of several biomedical signals such as pulse oximetry, electrocardiography, blood pressure, body temperature, and capnography. Since the mentioned monitors have limited computational resources, they require efficient and low computational cost algorithms for the analysis of these signals. Based on the fact that the measured biosignals are quasi-periodic in nature, to develop a general method, based on this feature, for classification and diagnosis of diseases is feasible and valuable.

Two representative examples of quasi-periodic biomedical signals, the electrocardiogram (ECG) and the capnogram, are studied in this research for the classification of atrial fibrillation (AF) and asthma, respectively. The ECG records the electrical activity of the heart. It is the most important tool for diagnosing heart diseases.

The capnogram is the continuous measurement of CO_2 concentration in the exhaled air. It is an essential tool for monitoring the cardiorespiratory function [20].

Algorithms specifically to classify the most common cardiac arrhythmia, named atrial fibrillation, are studied through [21–27]. In [21] and [27], an evaluation of the performance of various classifiers is discussed. References [22] and [23] use neural networks for the detection of AF; using frequency features and three physiological characteristics of the cardiac signal (electrical atrial activity, P wave absence, and heart rate irregularity) as inputs respectively. In [24], an algorithm based on three statistics of the heart beat interval is proposed. Two methods that compare density histograms of the test data with templates of standard density histograms of two physiological features (RR and ΔRR intervals) are presented in [25] and [26].

In the case of capnography for monitoring the asthmatic patient, the literature is reduced to a few studies analyzing the capnogram shape to differentiate asthmatic patients from normal patients [11, 17 - 19, 28]. One of the most relevant studies, [17], evaluates various shape indices of the capnogram, and demonstrates the possibility of asthma classification from the capnogram shape by using the manually calculated indices. References [18] and [28] statistically evaluate indices of the capnogram to differentiate the asthmatic and non-asthmatic capnograms. In [18], indices in [17] are automatically calculated and a new one is proposed. Reference [28] proposes a Linear Predictive Coding (LPC) analysis of the capnogram, and shows the extracted LPC coefficients can be used to distinguish the asthmatic conditions.

Some of the algorithms for AF classification, even if they could be easily implemented in real time applications, have classification accuracy from 60% to 75% [21]. Others present good classification results such as those proposed in [22, 23], but sometimes are not suitable for implementation in portable or embedded devices because of their computational complexity. Although the results for real time AF detection in [24–26] are promising, 128-beat segments for [24] and 100-beat segments for [25, 26] are required for the correct classification of the rhythm. For the classification of asthma,

previous studies suggest different indices for analyzing the shape of the capnogram. However, few utilize computational methods to classify the capnogram as asthma or normal. An efficient method to guarantee correct classification of quasi-periodic signals with low computational cost is required to overcome the drawbacks of the previous studies and to be implemented in multiparameter monitors that have limited computational resources.

In this chapter, a method for classifying ECG and capnogram signals is proposed from viewpoints of low computational cost and general quasi-periodic biomedical signals analysis. In the proposed method, the signal to be classified, first, is compared with two groups of patterns by means of the Shape Exchange Algorithm (SEA) [29]. One group is comprised of signals from normal patients and the other corresponding to the disease to be classified. As a result, two measures of similarity: normalized Percent Root Difference (PRD) and Correlation (Corr) are determined. Next, by applying fuzzy inference, the similarity between the input signal and each of the patterns is evaluated in two levels *similar* and *different*. Finally, applying an averaging approach, the signal is classified in *normal* or *disease*.

The combination of SEA and fuzzy similarity evaluation to classify the signal provides a general approach to be applied to other quasi-periodic biomedical signals such as oximetry and blood pressure. The proposed method has computational complexity of quadratic order; in terms of the size of the input signal and the pattern, and the number of patterns. Its computational cost is low. Moreover, it does not need a training process due to the use of fuzzy inference. Additionally, representation of PRD and Corr in a fuzzy similarity evaluation offers interpretability to the proposal.

The proposed method is tested for classifying atrial fibrillation using the MIT-BIH AF and MIT-BIH Normal Sinus Rhythm databases on PhysioNet [30, 31]. Segments of 30-second sample duration are selected from the 41 records in the databases; 23 with 100% AF episodes (according to the annotation rhythm) and 18 labeled as normal sinus rhythm. As regards the asthma classification, some preliminary experiments are realized.

The dataset comprises 22 asthmatic patients in a normal state and 4 suffering an asthma crisis during the measurement. From the dataset, 27 segments of 1200 samples, 21 from normal and 6 from asthma state are classified. The data have been collected during the School of Asthma held at the clinic "Julian Grimau" in Cuba. The school has conducted a training camp once a year for teaching patients and their relatives how to behave during and between crisis periods. As some other groups working on asthma research, they are seeking feasible and affordable tools for assessing the severity of asthma. For this purpose they collected a number of volunteers to be monitored. The method is implemented in MATLAB 7.9.0.529 and the experiments are carried out in a PC Intel (R) Core (TM) i7 CPU, 2.8 GHz and OS Windows 7 (32 bit) environment.

In subchapter 2.2, a detailed explanation of the proposal is presented. Starting from the original idea, then the introduction of the matching algorithm, the construction of the fuzzy inference engine, and to finish with the decision making process. The experiments performed for the validation of the proposal are presented in subchapter 2.3.

2.2 Similarity-Based Fuzzy Inference System for ECG and Capnogram Classification

The similarity based fuzzy classification is inspired by the analysis the physicians do of the biomedical signals for diagnosing diseases. They have special knowledge to understand and interpret the variations of the signals' shape and determine whether it belongs to a healthy person or not. This study proposes a method based on the similarity evaluation between biomedical signals of quasi-periodic nature. As two examples of this kind of signals, the ECG and the capnogram are used to test the method. Fig. 2.1 illustrates the differences between the ECG from a healthy patient and one from a

patient with atrial fibrillation. Notice that the irregularity of the heart rate and the absence of the P wave differentiate both signals. Fig. 2.2 shows how the shape of the capnogram changes from a normal person to a patient with asthma. Specifically the capnogram at the bottom represents a severe case of asthma where it takes a "shark's fine" appearance.

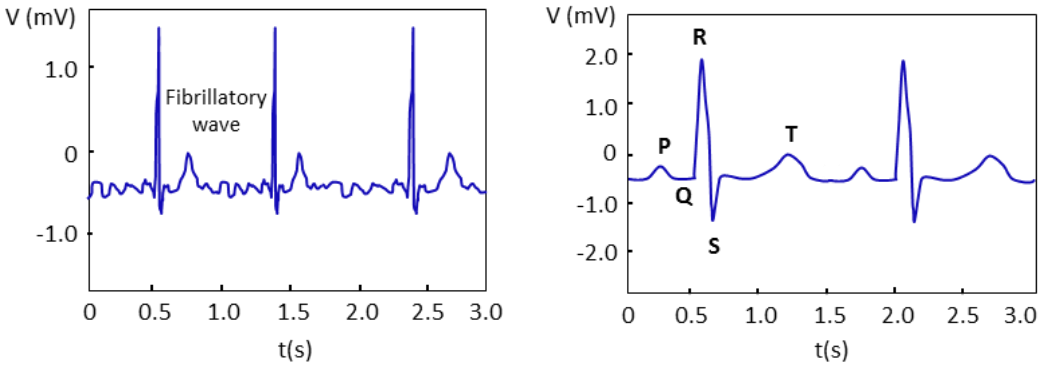


Fig. 2.1. Atrial fibrillation episode (left) and normal ECG (right).

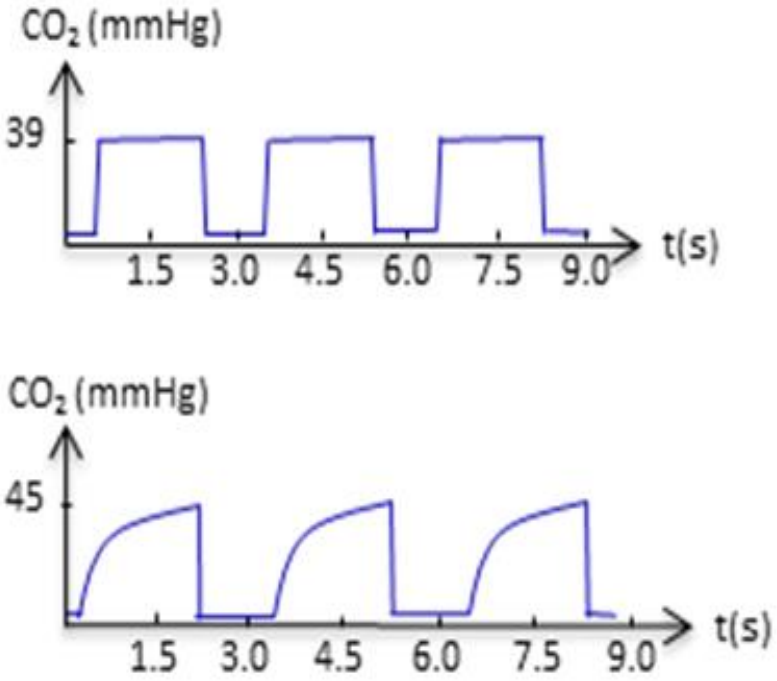


Fig. 2.2. Normal (top) and asthma (bottom) capnogram signals.

The proposed method comprises three fundamental steps: the first one is matching

the input signal with two groups of patterns (normal and disease) using the SEA [29]. The second step, by applying fuzzy inference, evaluates the similarity between the input signal and the patterns in two levels: *similar* and *different*. The third one is the classification itself, applying an averaging approach. In Fig. 2.3 a general scheme of the proposal is presented. The different steps are described in this subchapter.

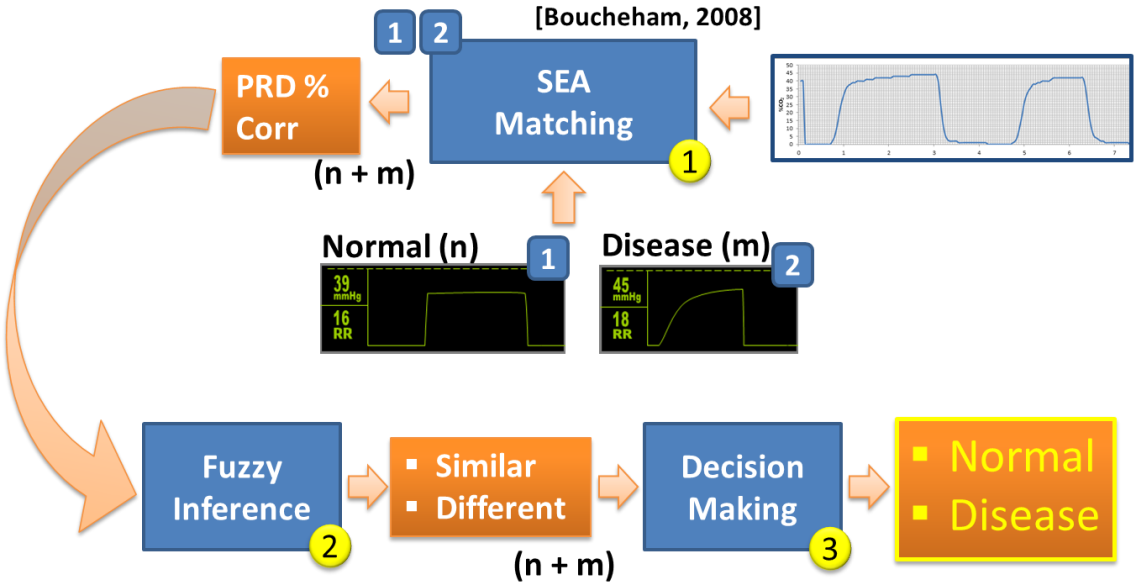


Fig. 2.3. Similarity-based fuzzy inference system for ECG and capnogram classification.

2.2.1 Shape Exchange Algorithm

To compare two quasi-periodic time series X and Y where $X = (x_i)_{i=1:N}$ and $Y = (y_i)_{i=1:N}$ two criteria of similarity-dissimilarity are used: the normalized percent root difference (PRD) (Eq. 2) and the correlation measure (Corr) (Eq. 3).

$$\mathrm{PRD}\,(\mathrm{X},\mathrm{Y}) = \sqrt{\frac{\sum_{i=1}^{N}|x_i - y_i|^2}{\mathrm{Max}(\sum_{i=1}^{N}|x_i-\bar{X}|^2,\ \sum_{i=1}^{N}|y_i-\bar{Y}|^2)}} \times 100\%\,, \qquad (2)$$

$$\mathrm{Corr}(\mathrm{X},\mathrm{Y}) = \mathrm{cov}(\mathrm{X},\mathrm{Y})^2/\mathrm{var}(\mathrm{X}).\mathrm{var}(\mathrm{Y})\,, \qquad (3)$$

with covariance (cov) of *X* and *Y* and variance (var) of *X* expressed as in Eq. 4 and 5

$$\mathrm{cov}(X,Y) = 1/N \sum_{i=1}^{N}(x_i - \bar{X})(y_i - \bar{Y}), \quad (4)$$

$$var(X) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{X})^2} \quad (5)$$

In Eq. 2, the numerator is the Euclidean distance between *X* and *Y*. The denominator under the root is a normalization of PRD to ensure values within 0-100% range, being 0% perfect match and 100% perfect mismatch. In Eqs. 2, 4, and 5 $\bar{X}$ and $\bar{Y}$ are the mean values of *X* and *Y* respectively. Note that the correlation measure range is -1 ≤ Corr(*X*,*Y*) ≤ +1, where the more *X* and *Y* are correlated the closer to +1 is Corr(*X*,*Y*).

The SEA method is composed of two main steps: signature establishment and shape exchange and comparison, which are explained below.

Signature establishment

A time series is composed by magnitude and time components; therefore, it can be represented using the notation $X = (x_i, i), i = 1{:}N$, where i is the temporal index of magnitude value x_i. Sorting of $(x_i, i), i = 1{:}N$ on the magnitude order outcomes a set of time indices and a trace of sorted magnitudes. This trace is referred by the signature of (*X*) and has two important features. First, it is a global descriptor, since it is the result of sorting the whole data. Second, it is stable, in the sense that sorting of segments of the same time series always produces quasi-similar signatures.

Shape exchange and comparison

The signature ignores the temporal component of the time series. Therefore, it cannot locate in time the existing dissimilarities between the time series. In the second step, the direct matching is carried out between time series through the exchange of the signatures preserving the temporal index. That is, *X* will take the shape of the pattern through its signature and vice versa. Finally, the reconstructed time series are obtained by ordering them on the temporal indices. The comparison is then performed for each

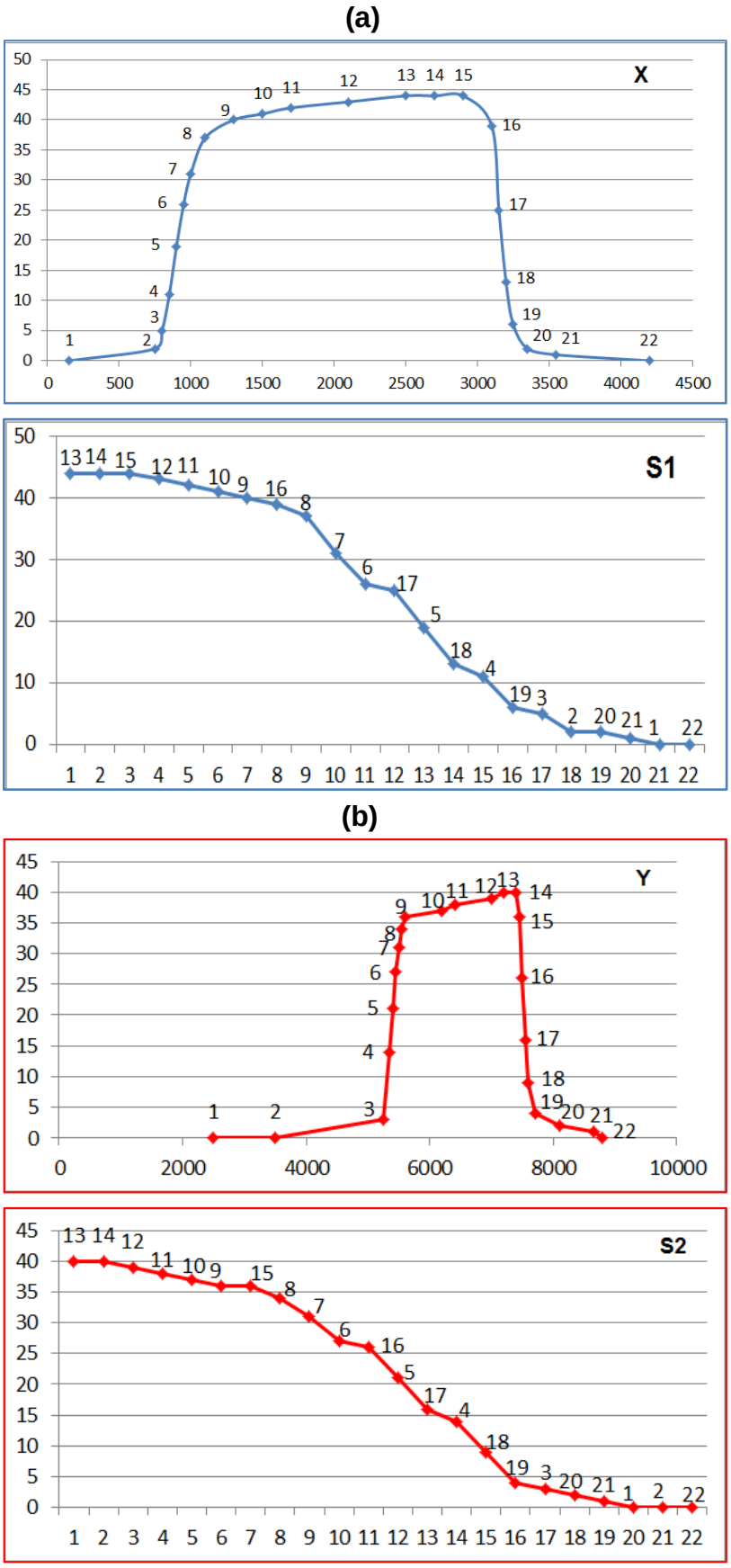
time series (e.g. X) with its reconstructed (e.g. X_{rec}) one using the PRD and Corr measures.

The main procedures of the SEA are sort-on-magnitude, sort-on-temporal-index, and the PRD and Corr calculations. Sorting procedure is Sort-by-Selection which has a theoretical O(k^2) temporal complexity. The PRD and Corr measures can be computed in a linear time. As a result, the temporal complexity of the SEA is of quadratic order; namely, O(2(n^2 + m^2) + 2(n + m)). Regarding the space complexity, the main data structure is two couples of vectors X (n), X_{rec}(n), Y(m), *and* Y_{rec}(m) then it is of linear order; namely, O(2(n + m)).

In [29], experiments are reported where the SEA is applied to ECG signals. The results suggest that the method may solve the problem of person identification using ECG and the problem of novelty detection in time series. Also, it is demonstrated that SEA is able to match the sub-patterns (P wave, QRS complex, and T wave) even when the time series are of different lengths, patterns, and number of periods.

The SEA is basically a quasi-periodic time series similarity measure which always induces a high correlation factor (Corr) and low distortion measure for similar time series. On this basis and perhaps with the combination of other criteria can be used in specific applications. It is a simple method to implement and requires no parameters to match the time series. It deals with the basic problems of time series matching: noise, offset translation, amplitude scaling and time axis scaling. In addition, it has a temporal complexity of quadratic order and space complexity of linear order, leading to a lesser execution time and lower memory requirements than the well-known DTW (Dynamic Time Wrapping) algorithm [29].

Based on the fact that the capnogram is also quasi-periodic in nature, the presence of the signature, as defined in [29], has been analyzed for this signal. Fig. 2.4 shows an example of capnogram signals and their signatures. Also, a reconstructed signal and the comparison with its original.


(a)
X
S1
(b)
Y
S2

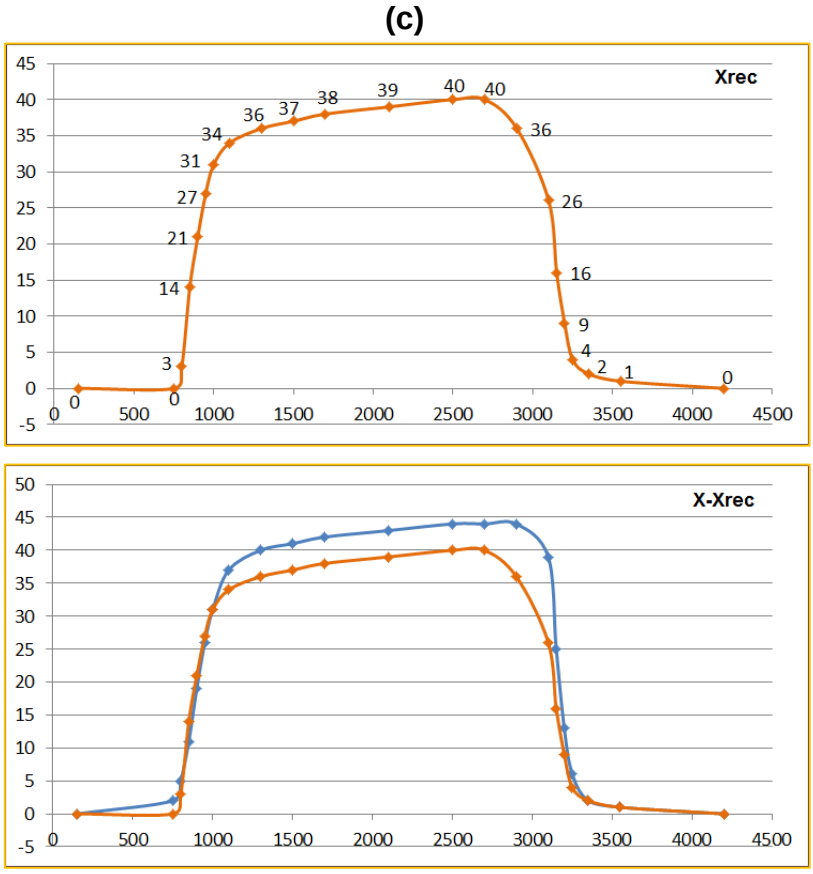


Fig. 2.4. SEA main steps. (a) X signal and its signature S1 (blue), (b) Y signal and its signature S2 (red), (c) Reconstructed signal of X (Xrec, brown) wear the magnitudes of Y. Comparison between X and its reconstructed Xrec.

Note that similar traces are obtained for different capnograms (a, b). The finding supports the use of the SEA as part of the solution to the problem of asthma classification based on the capnogram.

The first step of the proposed method is the matching where the SEA is applied. The matching inputs are the signal to be classified and two groups of patterns, which are signals already labeled by other methods. One group of patterns comprises normal signals and the other comprises signals representing the disease to be identified. The input signal is compared with both groups of patterns. The outcome is a pair of similarity measures PRD and Corr, one pair for each pattern. Next, the mean Corr and

PRD are calculated for each group of patterns.

Fig. 2.5 shows a scheme of the SEA used in the matching step of the proposed classification method. Let $X = (V_{1..N}, P_{1..N})$ the signal to be classified and $A = (W_{1..M}, Q_{1..M})$ the pattern to be X compared with, where:

$V_{1..N}$ is the set of magnitude values of X

$P_{1..N}$ is the set of time indices of X

$W_{1..M}$ is the set of magnitude values of pattern A

$Q_{1..M}$ is the set of time indices of pattern A

In the first step, the signal X and the pattern are sorting on the magnitude order to get the signature. Next, the signatures are exchanged keeping the temporal indices and X'' and A'' are obtained, then, X'' and A'' are reordered on their temporal indices resulting in the reconstructed signals. Finally, the original signal X and its reconstructed Y are matched and the two measures of similarity PRD and Corr are calculated.

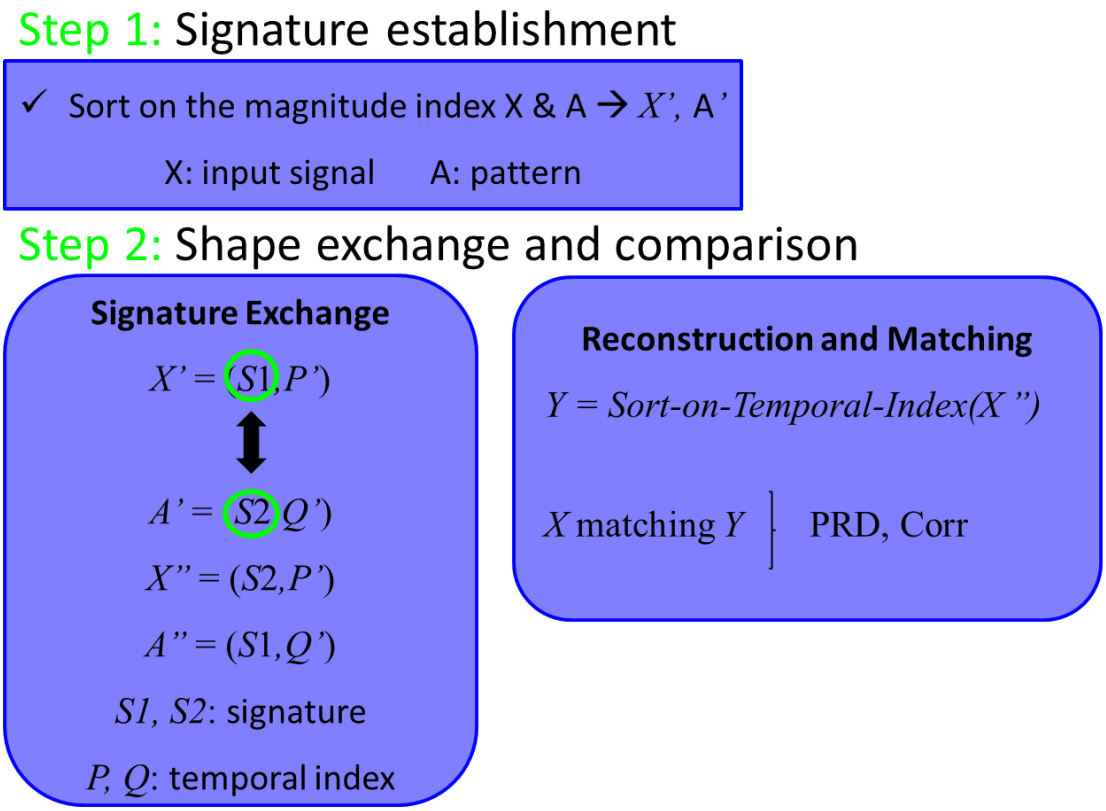


Fig. 2.5. Scheme of the SEA used in the matching step of the proposed classification method.

2.2.2 Fuzzy Similarity Evaluation

Fuzzy inference is built for assessing the similarity between the signals having as inputs two measures of similarity: PRD and Corr. From the analysis of the data, 3 membership functions are determined for both inputs: *low*, *medium*, and *high*. For the output, two membership functions, *similar* and *different*, are built. The membership functions' tuning and the fuzzy rules' construction are explained in ensuing sections.

Setting Corr, PRD, and Output Membership Functions

Similarity measures, PRD and Corr, are obtained first, by matching all the signals in the database to a number of patterns by applying the SEA. To determine the parameters of PRD and Corr membership functions (μ), the fuzzy c-means clustering method is used. Corr and PRD are clustered independently into 3 categories: *low*, *medium*, and *high*. From each category, after removing the outliers and extreme values (PRD = 0, Corr = 1), the minimum (min) and maximum (max) values are determined. Fig. 2.6 illustrates the process of setting the membership functions.

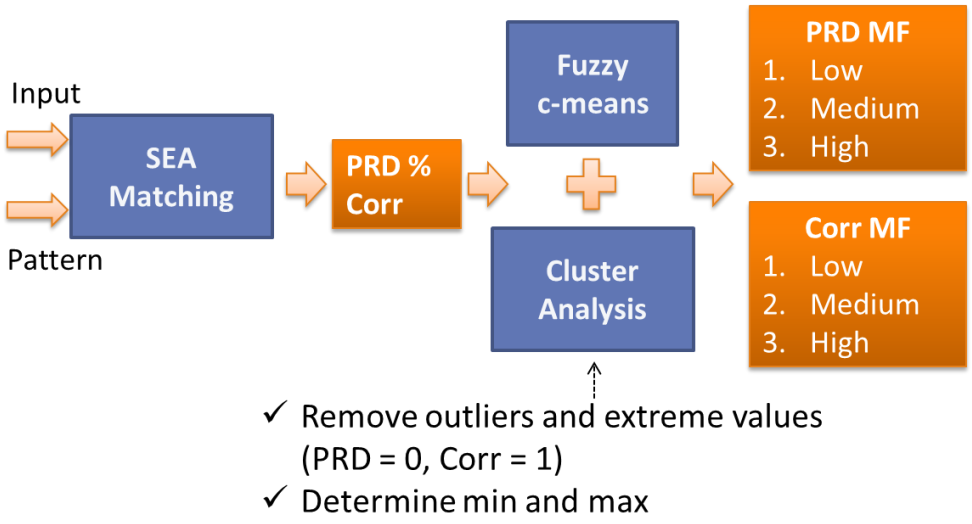


Fig. 2.6. Scheme of the process of setting PRD and Corr membership functions.

For membership function *low*, the degree of membership function 0 is assigned to the maximum value of the cluster *low*. The value *c* is equal to the minimum of this cluster and is assigned the degree of membership function as 1. In the case of membership function *high*, the procedure is similar but the assignation of the degree of membership function to the maximum and minimum values of the corresponding cluster are reversed. The degree of membership function 0 and 1 are assigned to *a* and *b* from *low*, also to *d* and *c* from *high*, respectively, and they are equal to 0. For *medium*, the degree of membership function 1 (*b* and *c*) is assigned to the minimum and maximum of

the cluster medium. To tune the parameters *a* and *d* the best intersection with *low* and *high* is found to get the best classification result using fuzzy inference. Fig. 2.7 illustrates the min and max of the cluster to the membership functions' parameters being assigned. The trapezoidal membership function is selected because it fits the Corr and PRD changes well along the evaluation of the similarity. It uses the formula

$$\mu(x,a,b,c,d) = \begin{cases} 0, & x \le a, x > d \\ \dfrac{x-a}{b-a}, & a < x \le b \\ 1, & b < x \le c \\ \dfrac{d-x}{d-c}, & c < x \le d \end{cases}, \quad (6)$$

where a, b, c, and d are real numbers.

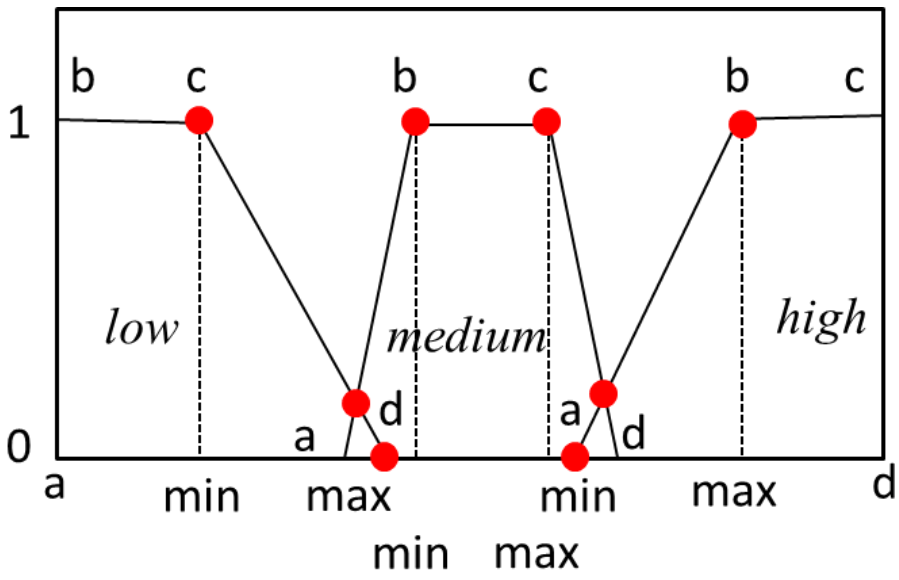


Fig. 2.7. Tuning the membership functions.

Figs. 2.8 – 2.11 illustrate the membership functions and Tables 2.1 and 2.2 show the values of the tuned parameters of Corr and PRD for the ECG and capnogram data, respectively.

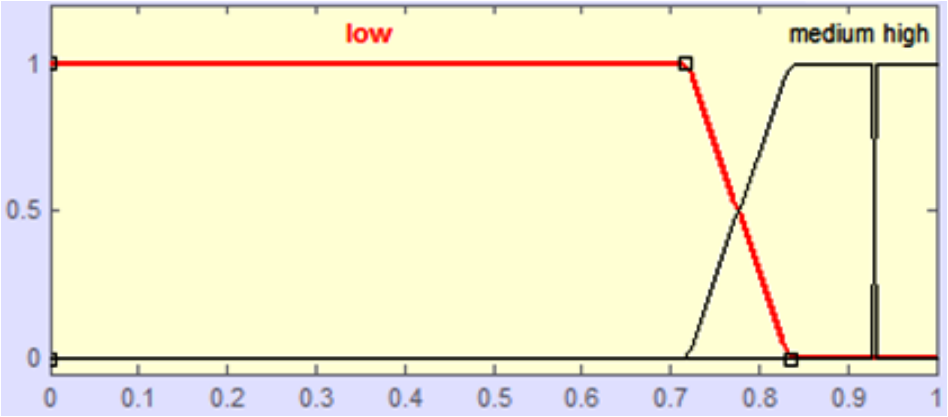


Fig. 2.8. ECG Corr membership functions.

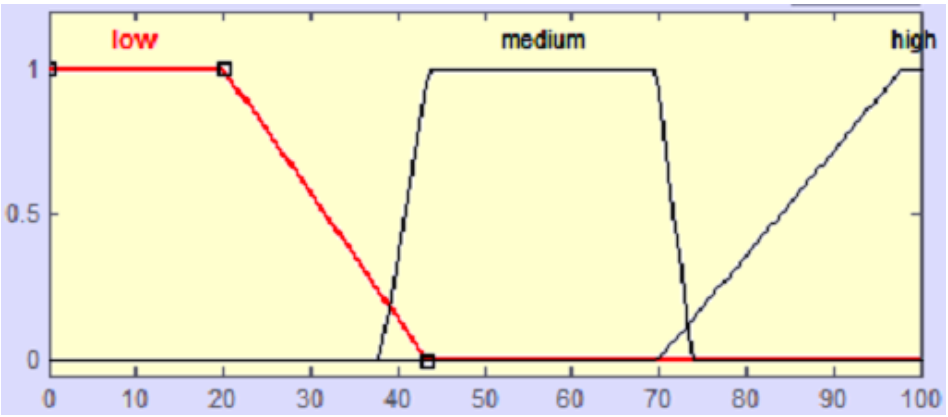


Fig. 2.9. ECG PRD membership functions.

Table 2.1. Parameters of Corr and PRD for ECG.

	μ_{low}		μ_{med}		μ_{high}	
	Corr	*PRD*	*Corr*	*PRD*	*Corr*	*PRD*
a	0	0	0.72	37.9	0.931	69.8
b	0	0	0.84	43.6	0.932	97.9
c	0.72	20	0.931	69.7	1	100
d	0.83	43.5	0.932	73.7	1	100

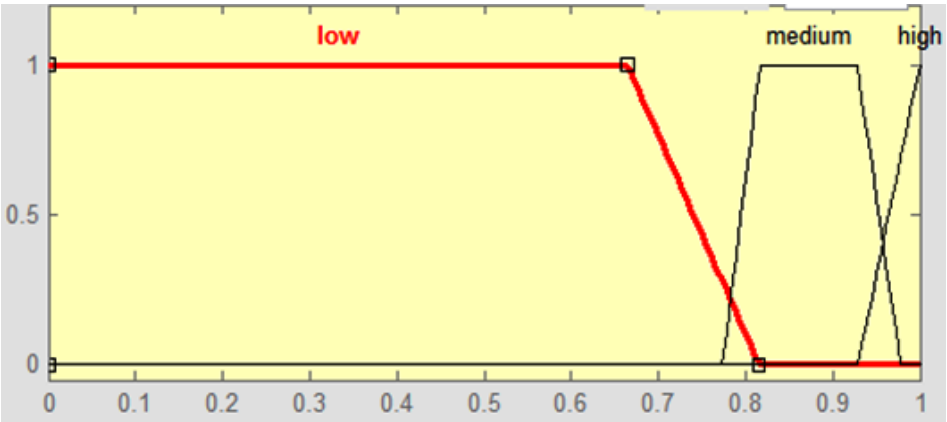


Fig. 2.10. Capnogram Corr membership functions.

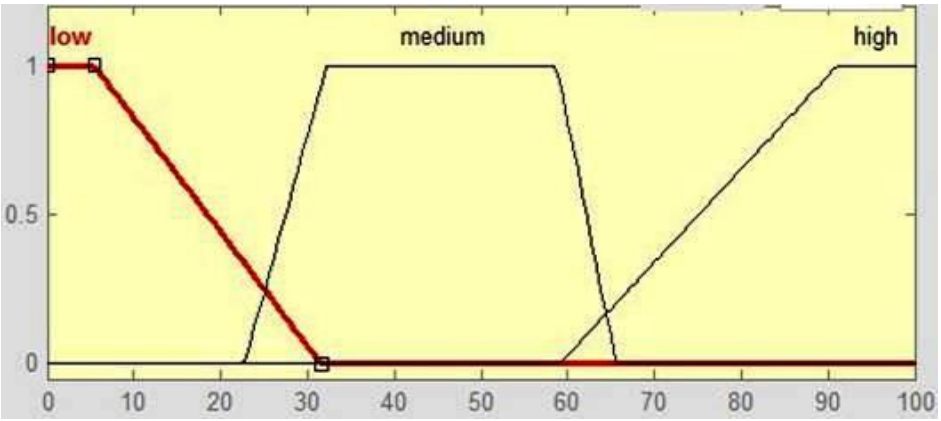


Fig. 2.11. Capnogram PRD membership functions.

Table 2.2. Parameters of Corr and PRD for capnogram.

	μ_{low}		μ_{med}		μ_{high}	
	Corr	*PRD*	*Corr*	*PRD*	*Corr*	*PRD*
a	0	0	0.77	22.7	0.93	59.0
b	0	0	0.82	32.2	0.99	91.1
c	0.66	5.41	0.93	58.7	1	100
d	0.82	31.58	0.98	65.6	1	100

Two trapezoidal membership functions are tuned for the output, corresponding with the two possible classifications of input, *similar* and *different*. Similar tuning is done for

ECG and capnogram databases. The tuning is performed manually by trying different parameters and the best result of classification is achieved for the configuration shown in Fig. 2.12.

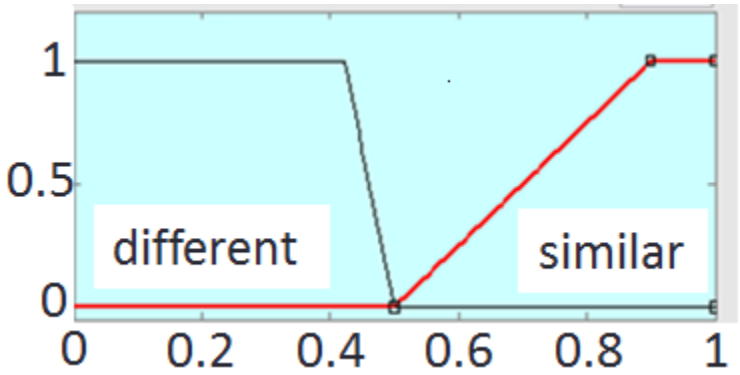


Fig. 2.12. Capnogram and ECG membership function for the output.

Fuzzy Rules for Assessing the Similarity

For assessing the similarity between the input signal and the pattern, fuzzy rules are manually constructed using the criteria of similarity: PRD and Corr; and from analyzing the experimental data obtained from matching signals and patterns. Table. 2.3 shows the constructed fuzzy rules. Notice, when Corr is low regardless of the PRD value the signals are considered different. Likewise, the evaluation result is different when PRD is medium or high. There are only 2 cases where the signals are similar: when the PRD is low and the Corr is medium or high. This is related with the fact that for signals to be considered similar it is mandatory to have low values of PRD which is a measure of low distortion between the signals; i.e., the closer the PRD is to 0%, the more the signals are similar. Alternatively, as the correlation between signals approaches to 1, the more similar they are.

Table 2.3. Fuzzy rules for assessing the similarity.

	Corr		PRD		Output
IF	low	AND	low	THEN	different
	low		medium		different
	low		high		different
	medium		low		similar
	medium		medium		different
	medium		high		different
	high		low		similar
	high		medium		different
	high		high		different

2.2.3 Classification from the Space of Similarity

For the classification of the signal an average approach is used. The results of the similarity evaluation of the signal with respect to the normal patterns are averaged. Likewise, it is done for the disease's patterns. The highest mean value determines the final classification of the signal. E.g., if the mean similarity value when matching with the normal patterns is higher than the result of matching with the disease patterns, then the input signal is classified as normal.

2.3 Classifying Atrial Fibrillation and Asthma

2.3.1 Measures for Performance Evaluation of the Proposal

To evaluate the performance of the classification method, sensitivity, specificity, and accuracy are calculated by means of the following equations

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}), \quad (7)$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}), \quad (8)$$

$$\text{Accuracy} = \text{TP} + \text{TN}/(\text{TP} + \text{TN} + \text{FP} + \text{FN}), \quad (9)$$

where TP is the positive labeled signal correctly classified as such; FN is the positive labeled signal misclassified as normal; TN is the negative labeled signal correctly classified as normal; and FP is the negative labeled signal misclassified as positive.

Sensitivity is defined as the percentage of positive labeled signals that are correctly classified as positive. The specificity is the percentage of negative labeled signals that are correctly classified as negative.

Additionally, the assessment of the classification method's performance is done via 2 fold cross-validation. Regarding the inputs and the patterns used for matching, those used for tuning the parameters of the membership functions are different from those used for classification. For each iteration, the accuracy is calculated. Finally, the accuracy is averaged.

2.3.2 Experiments on the ECG Databases

Two databases from PhysioNet are used for the experiments: MIT-BIH Atrial

Fibrillation and MIT-BIH Normal Sinus Rhythm [30, 31]. The MIT-BIH AF Database contains 23 long-term ECG recordings of individuals with atrial fibrillation containing two signals, each sampled at 250 Hz. 30-second sample duration segments of lead 2 with presence of AF episodes (according to rhythm annotation) in '.txt' format are selected from the records as input of the proposed method. The MIT-BIH Normal Sinus Rhythm comprises 18 long-term ECG recordings of subjects with no significant arrhythmias. Each one has two signals sampled at 125 Hz. Segments of 30-second sample duration of lead 2 in '.txt' format are selected. For the experiments environment, Matlab 7.9.0.529 is used in a PC Intel (R) Core (TM) i7 CPU, 2.8 GHz and OS Windows 7/32 bit.

The signals are filtered to eliminate the high frequency noise. The parameters of the membership functions are tuned, matching the 41 signals with 15 AF patterns and 15 normal patterns selected among them, 1230 PRD and 1230 Corr values are obtained. The 41 recordings from the databases mentioned above are classified based on the similarity measures between them and the same patterns used for tuning the parameters of the membership functions. Notice that some PRD and Corr values are not used for tuning the membership functions but rather for classification, i.e., 11 signals are not included as patterns for matching; representing 26.8 % of the total of PRD and Corr values if using the 41 signals as patterns. As a result, shown in Table. 2.4, from 18 normal signals, 17 are classified correctly and only 1 is misclassified. The 23 atrial fibrillation signals are also well classified; achieving values of sensitivity, specificity, and accuracy of 100%, 94.4%, and 97.6%, respectively.

For the 2 fold cross-validation experiments, only 5 AF patterns and 4 normal patterns are used for tuning the membership functions and the classification. The database of 41 signals is divided in two (21 and 20). Each of the partition comprises 9 normal signals and for the number of AF signals, one has 12 and the other 11. The experiment is performed two times; first, one partition is used for tuning the parameters of the membership functions and the other for classification and next, the partitions are

exchanged, such that, the one used for tuning is now used for classification and reverse. The accuracy is calculated for each of the experiment, being the mean accuracy of 92.6%.

Table 2.4. Results for ECG classification experiment.

	Normal	**AF**
Normal state	TN=17	FP= 1
AF state	FN= 0	TP= 23

Discussion on the Results of ECG Classification

The approach proposed for detecting atrial fibrillation is based on the analysis of the signal's shape not on the classification of each heart beat as in the classifiers compared in [21] or the classification of n-beat segments in [21, 23]. The method is closer to the analysis physicians perform for diagnosing; detecting the presence of AF in the ECG not every single AF beat. Compared with [24], which can detect atrial fibrillation using a minimum of 64-beat segments, the proposed method correctly classifies segments of 36 beats duration. The experiments validate the ability of the proposed method to classify an ECG signal as AF or normal with high sensitivity and specificity in two databases. The Venn diagram in Fig. 2.13 illustrates the performance of the classification method. The shadow area represents the set of AF signals and the circle containing it represents the classifier. Observe, there are no false negatives (FN), i.e., the probability of correctly classify AF is 100%. In medical applications, it means that all the actual cases of disease will be detected.

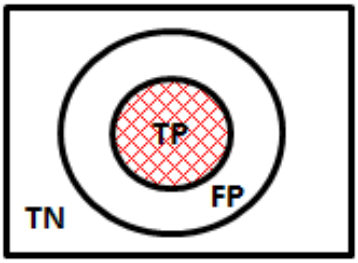


Fig. 2.13. Performance for ECG classification.

Compared with the results reported by S. Dash et al. for the MIT-BIH Atrial Fibrillation Database, the proposed method is more sensitive (100% vs. 94.4%) but slightly less specific (94.4% vs. 95.1%).The latter means it classifies correctly 17 of 18 normal signals. Because the size of the normal dataset is small, 1 misclassification represents 5.55% of the number of normal signals. In addition, there is not much mention of the number of normal beats classified or results for Normal Sinus Database in [24]. Compared with [22], which uses a dataset for testing 24 normal and 28 AF signals, which is quite similar to the size of that used for validating the proposed method, the values of specificity (94.4% vs. 100%) and accuracy (97.6% vs. 100%) are lower; however, [22] does not provide real-time detection.

Additionally, a 2 fold cross-validation experiment is realized which uses only 5 AF patterns and 4 normal patterns for tuning the membership functions and the classification. The database of 41 signals is divided in two (21 and 20). Each of the partition comprises 9 normal signals and for the number of AF signals, one has 12 and the other 11. The experiment is performed two times; first, one partition is used for tuning the parameters of the membership functions and the other for classification and next, the partitions are exchanged, such that, the one used for tuning is now used for classification and reverse. The accuracy is calculated for each of the experiment, being the mean accuracy of 92.6% which is similar to the result previously discussed.

2.3.3 Experiments on the Capnogram Database

Data from 26 asthmatic patients (18 women and 8 men, mean ± SD age equal 27.3 ± 24.7 years) suffering bronchial asthma based on usual criteria, have been collected during the Asthma Camp held at the clinic "Julián Grimau" in Cuba. 22 of the subjects were in a normal state and 4 suffering an asthma crisis during the measurement.

One normal signal and one of asthma have been excluded for the experiments because the expiration phase lasts more than 3 s [17]. Since the number of samples of the rest of the asthma records is large (6070, 6536, and 6146), each one has been split in two segments to enlarge the data for testing.

After filter the signal at 5 Hz, since only expiratory phase segments are considered from the respiratory cycles, the inspiratory phase is eliminated. The expiratory phase is the rising portion of each cycle starting from 0, until the maximum value achieved, called end-tidal CO_2, as illustrated in Fig. 2.14. The cycles with end-tidal CO_2 higher than 60 mmHg are removed.

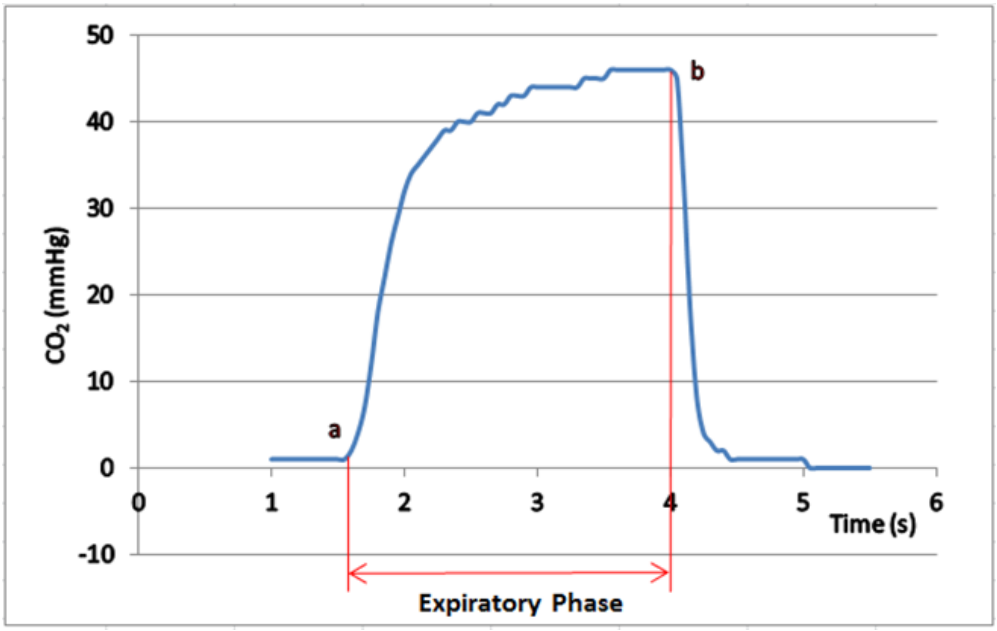


Fig. 2.14. Expiratory phase of the capnogram.

The capnogram has been recorded using a commercial sensor from Combiomed, named Capnosafe [15] at a sample frequency of 20 Hz. Fig. 2.15 shows pictures of the Cuban capnograph.

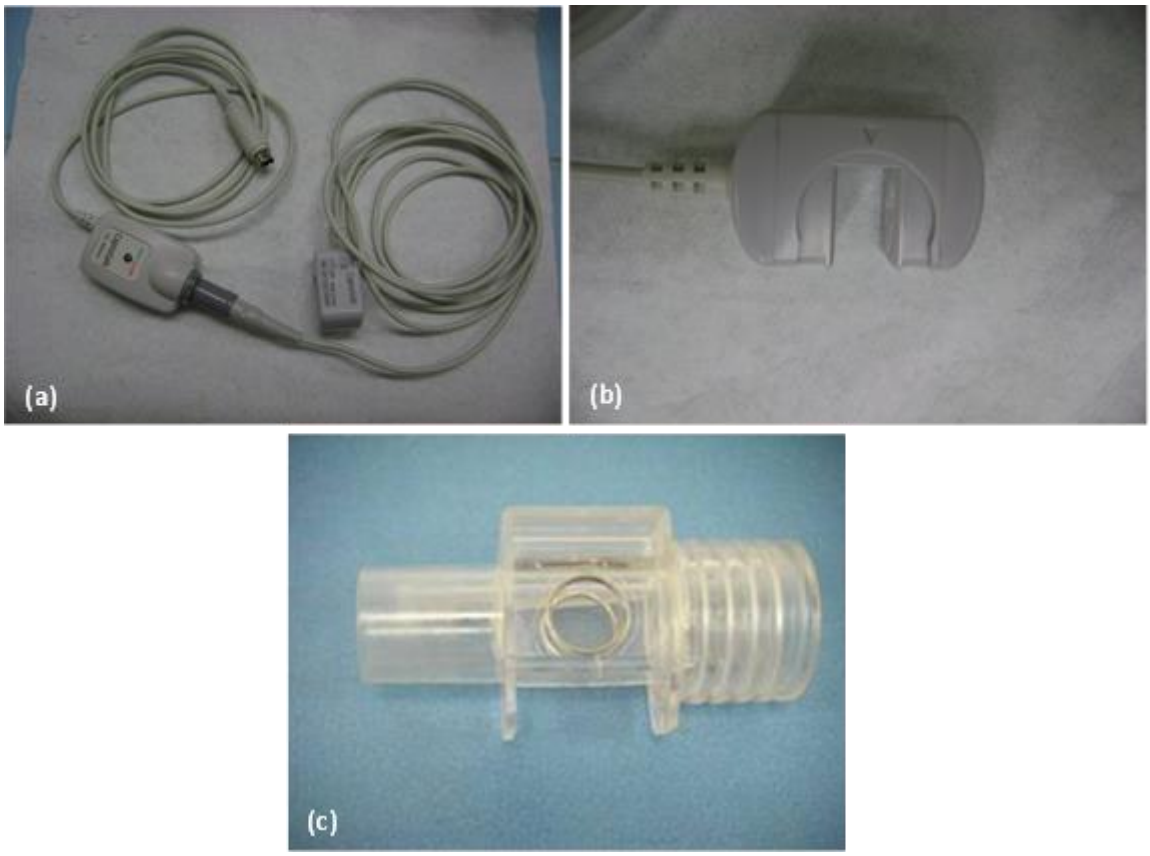


Fig. 2.15. Parts of capnograph Capnosafe: (a) IR sensor of CO_2 and processing module. (b) Detail view of the sensor. (c) Airway adapter.

The capnograph comprises an airway adapter through which the respiration gases flow, a sensor detecting the presence of CO_2 and a processor that carries out the calculations and outputs a digital signal to a physiologic parameters monitor, that displays the CO_2 waveform (capnogram), and the $EtCO_2$ value on the screen.

The same procedure as for the ECG is applied to determine the parameters of the membership functions, using, in this case, 6 patterns of asthma and 6 patterns of normal state. Meaning that from the 27 signals, 44.4% are used as patterns for tuning the

membership functions.

As result of the classification, presented in Table. 2.5, 1 of the asthma signals and 4 normal signals are misclassified. Values of 83.3%, 80.9%, and 81.5% for sensitivity, specificity, and accuracy, respectively are achieved.

Table 2.5. Results for capnogram classification experiment

	Normal	**Asthma**
Normal state	TN=17	FP= 4
Asthma state	FN= 1	TP= 5

Furthermore, a 2 fold cross-validation experiment is performed. In this case, 3 asthma patterns and 3 normal patterns are used for tuning the membership functions and the classification. The 27 signals database is divided in two (14 and 13). Each partition comprises 3 asthma signals and for the number of normal signals, one has 11 and the other 10. The experiment is performed two times; first, one partition is used for tuning the membership functions and the other for classification, then, the partitions are exchanged, i.e. the one used for tuning is now used for classification. The accuracy is calculated each time, leading to a mean accuracy of 81.8% which is similar to the previous experiment.

Discussion on the Results of Capnogram Classification

The method proposed for the classification of capnogram signals as asthma or normal is based on previous studies about the correlation between indices characterizing the shape of the capnogram and spirometric parameters commonly used for monitoring the asthmatic patient. As a first step for developing a method for automatic monitoring of the asthma condition, the proposed method is a valuable, pioneering approach, thus, there are currently no other works to compare the results with.

The database is small; especially the number of asthma signals. Therefore, the values indicating the performance of the classifier, such as sensitivity, specificity, and accuracy, (83.3%, 80.9% and 81.5% respectively) are affected. Notice that only one asthma signal is misclassified which represents 16.6% of the positive cases. In addition, the capnogram signals used for testing the method are from asthmatic patients classified as degree 0 of asthma (normal) and degree 1(asthma). These early stages of asthma are difficult to differentiate because the changes in the capnogram's shape are light. Fig. 2.16, in a Venn diagram, schematically represents the performance of the classifier.

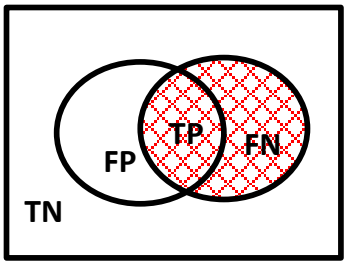


Fig. 2.16. Performance for capnogram classification.

The cross-validation experiments, as for ECG classification, show similar mean accuracy (81.8%) to the previous discussed. The variability of the accuracy between the partitions is high.

Experiments using different normal patterns and number of patterns are done. The classification results are altered each time, corroborating the importance of a good selection of the patterns.

Considering only 6 asthma signals from 3 different patients are available for testing, they are not representative of all the degrees of asthma. Adding more asthma data could enlarge the range of values of PRD, thus slightly modifying the boundaries of the membership functions; this is required for future works.

2.3.4 Execution Time and Memory Space for the Classification

The SEA has a temporal complexity of quadratic order given by $O(2(n^2 + m^2) + 2(n + m))$ [29], where n and m are the number of samples of the input signal and the pattern respectively. The fuzzy inference's main procedure; the fuzzy IF-THEN rules, as well as the averaging procedure can be computed in a linear time O(n) where n is defined by the number of patterns. As a result, the overall temporal complexity of the proposed method is of quadratic order. Additionally, it has no training process and it does not require the extraction of standard features. Due to these reasons, the computational cost of the method is low leading to an execution time for classifying one signal of 0.6 s for the ECG and 0.3 s for the capnogram. This time is acceptable because, in the medical practice, the analysis of the signals, usually takes several seconds.

A raw estimation of the execution time for a typical configuration of a multiparameter monitor assuming it has a 120 MHz microcontroller, is about 14 s for classifying one signal (0.6 s in the experimental environment).

The method requires the storage of the patterns to be compared with the input signal. For the capnogram signals of 1200 samples, considering 6 patterns, it needs about 12 KB of memory space. The ECG signal consumes a space of 28 KB to store the 30-second patterns. The memory space requirements are compatible with the features of the dsPIC microcontrollers from Microchip commonly employed by the multiparameter monitors, e.g., the dsPIC33E has a RAM memory up to 53 KB.

2.4 Chapter Summary

The results of this research show the capacity of the fuzzy similarity based system approach to accomplish the classification of ECG and capnogram signals as atrial fibrillation and asthma.

Experiments on classifying atrial fibrillation from MIT-BIH AF Database and Normal Sinus Rhythm Database from PhysioNet are done, showing high values of 100%, 94.4%, and 97.6% as sensitivity, specificity, and accuracy, respectively. The performance of the proposal is comparable with the previous ones studied while offering real time detection.

In the case of the experiments intended to classify asthma using capnogram data, a result of 83.3% for sensitivity, 80.9% for specificity, and 81.5% for accuracy is obtained. The performance of the method is affected by the small size of the dataset; 27 signals including only 6 records from asthmatic patients. The results are promising, as it is the first attempt at automatic classification of capnogram signals as asthma or normal.

The proposal transforms the input signal into a similarity space comprising two features, correlation and percent root difference; and by feeding a fuzzy classifier with these two features, the classification is performed. The main point is that the proposed method allows reducing the number of features to 2. In other words, no matter the features can be extracted from the input signal, the feature vectors can be matched with patterns, applying the SEA, to be transformed into similarity measures. The merit makes possible to extend this method to other quasi-periodic biosignals.

The proposal has a low computational cost exhibiting a low execution time, for classifying one signal, of 0.6 s for the ECG data, and 0.3 s for the capnogram data. The execution time is estimated for a typical configuration of multiparameter monitors to evaluate its feasibility in this scenario. The result of 14 s (worst case) is suitable for this

kind of device.

Chapter 3

Capnogram Feature Extraction Based on Wavelet Decomposition by Segments

This chapter introduces a different approach for the analysis of the capnogram. A feature extraction method from capnogram based on wavelet decomposition is presented. Fig. 3.1 shows a diagram of the proposal.

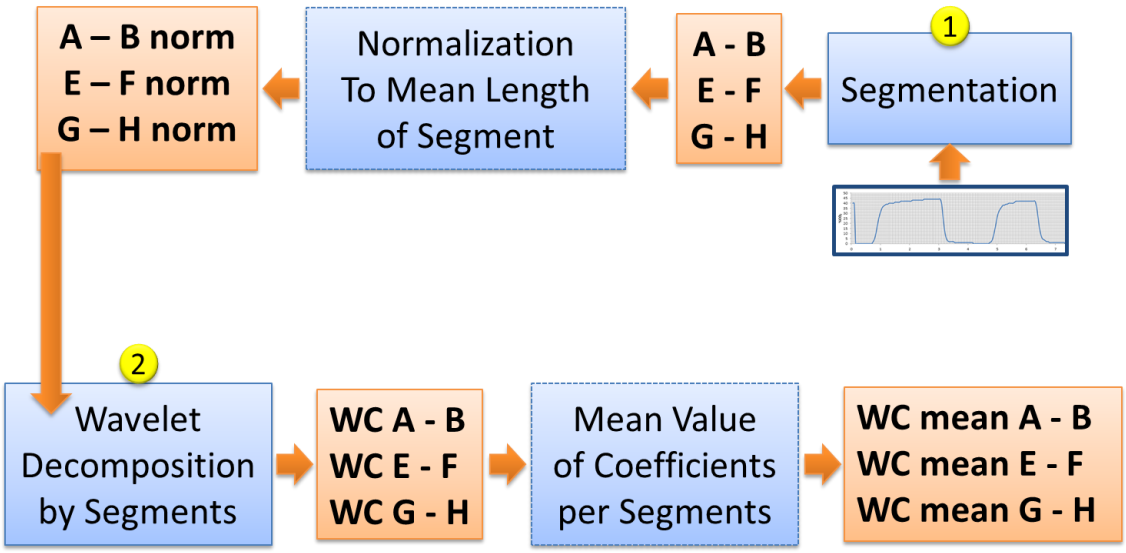


Fig. 3.1. Feature extraction based on wavelet decomposition by segments.

The ensuing subchapters explain the main steps of the proposed method: the segmentation and the wavelet decomposition by segments.

3.1 Capnogram Segmentation Based on Piecewise Linear Regression

Segmentation is considered in order to extract the segments which contain useful information about asthma. This process comprises, first, the identification of the cycles and then the extraction of the interest segments.

The capnogram phases (I, II, III) have been divided in linear segments and the corresponding intersegments that represent the transition between them as shown in Fig. 3.2. These segments are considered straight lines which make easier the initial approach of the capnogram analysis.

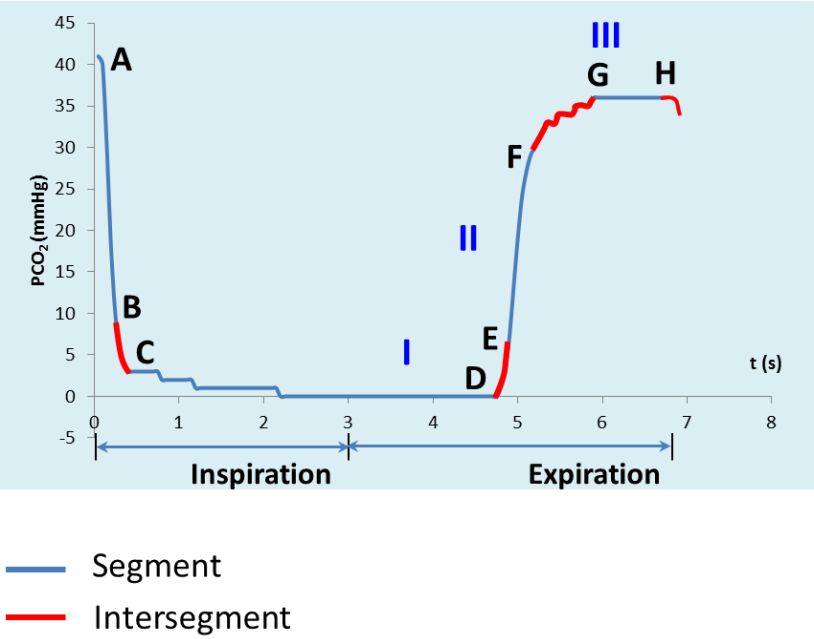


Fig. 3.2. Segments and intersegments in a capnogram cycle.

The phase I, in which the CO_2 concentration is considered zero, does not contain information about the status of the lungs. Therefore, it can be assumed that it is possible to exclude this phase from the analysis limiting the feature extraction to phases of the

capnogram that contains relevant information of the CO_2 variation.

In Fig. 3.3 it can be observed the differences between these segments. Notice that, for the patient with symptoms of asthma (bottom), the rate of change of CO_2 in segment E-F is slower comparing with that of the patient in normal condition (top). Also, it can be appreciated that segment G-H almost disappear for the asthmatic condition. Another difference is that the duration of the respiratory cycles is different for either the normal or asthma capnogram. This means that for an individual the duration of the respiratory cycles can vary. From the analysis of the signal's features that differentiate the asthmatic condition, we assume it is possible, using features in the frequency domain, to improve the results of previous research works based mainly on the study of the capnogram in the time domain.

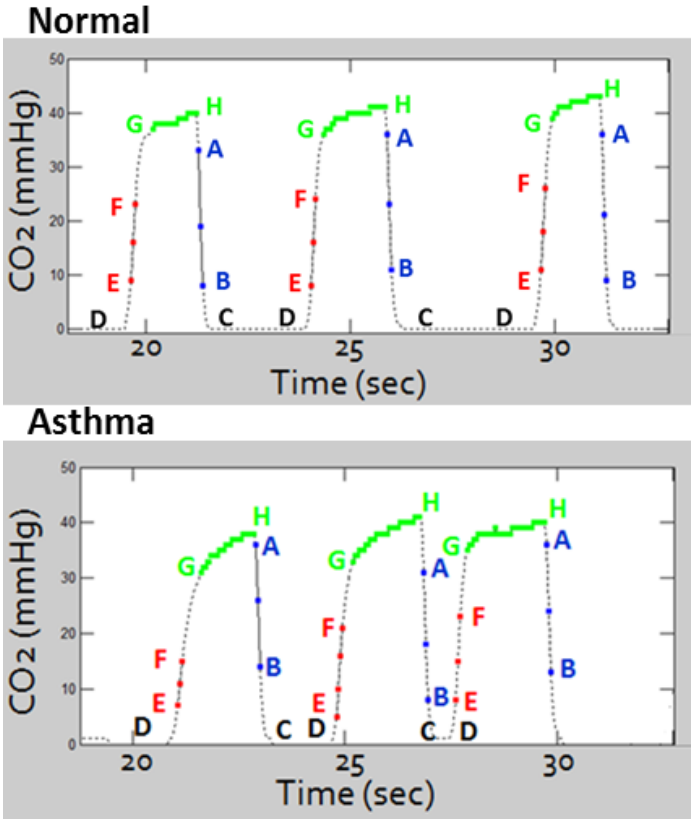


Fig. 3.3. Segments in real normal (top) and asthma (bottom) capnograms.

To accomplish segmentation, the capnogram signal is converted into a function of slopes. First, a sliding window is defined having a width equal to 3 samples, which is

the minimum number of segment samples of signals used in experiments. Second, using the Matlab polyfit function, samples in the window interval are fitted onto a line segment and its slope is calculated. The window is shifted one sample each time until the whole signal has been processed. Third, a curve of slope vs. slope index is constructed. From this curve, respiratory cycles are first determined by identifying points of minimum negative slope that correspond to the end of expiration. Segments A-B, E-F, and G-H are then identified. Segment A-B consists of points on the same slope starting from the minimum slope point. Segment E-F corresponds to points on the same slope starting from the maximum slope point. Segment G-H is identified as the line from the end of expiration to the point where the value of the slope changes above 0.01. Once segments are identified, intersegments are identified along with sets of samples between segments.

3.1.1 Capnogram Features in Asthma Classification

Capnographic analysis has several advantages compared to spirometric methods in that it is effort-independent and continuously monitors tidal respiration even while the subject is been treated [8, 15, 17, 18].

B. You et al. [17] evaluates various capnogram shape indices and demonstrates the possibility that asthma could be classified from capnogram shape by using manually calculated indices. They [17] refer also to differentiation among three subgroups of asthmatic subjects based on capnogram shape indices. These indices do not, however, differentiate very well between subgroup A1, with normal or subnormal spirometric values (FEV1 $\geq$ 80%), and subgroup A2, showing moderate signs of obstruction (40% $\leq$ FEV1 < 80%). In [28] a method for classifying the capnogram in normal or asthma based on the similarity between the input capnogram signal and normal and asthma

patterns is proposed as a first step in automatically classifying asthma severity.

Differentiating well among the 6 degrees of asthma severity thus requires finding new capnographic features. Due to the time-frequency localization of the wavelet transform, it is assumed in this paper that frequency analysis of the capnogram using wavelet transform may provide useful information for classifying asthma.

3.2 Capnogram Feature Extraction Based on Segmented Wavelet Decomposition

In this subchapter two well-known techniques for signal processing are evaluated for the analysis of the capnogram in the frequency domain: Fourier transform and Wavelet transform.

3.2.1 Fourier and Wavelet Transforms for Feature Extraction

Fourier transform uses complex sinusoids, being Fourier transform of a continuous signal: $X_F(\omega) = \langle e^{j\omega t}, x(t) \rangle$ [33]. The difficulty with this approach, already referred, is that because of the extent of the basis function it is no possible the time-localization of the frequency information [34].

The windowed Fourier Transform solves this problem by first windowing the signal and then taken its Fourier transform [33]. Therefore, the time frequency localization is possible. The discrete version of this transform is

$$X_{WF}(\omega,t)=\int_{-\infty}^{+\infty} f(s)g(s-nt_0)e^{-im\omega_0 s}ds\ , \tag{10}$$

where ω_0, $t_0 > 0$ are fixed and $m, n \in \mathbb{Z}$ (integers). The limitation here is that a single window (regardless the value of ω, the window width is the same) is used for all the frequencies, so no multiresolution analysis can be performed.

The wavelet transform gives an analogous representation of the frequency content of signals with a few differences. Continuous and discrete versions of wavelet transform equations, are

$$X_W(a,b)=a^{-1/2}\int_{-\infty}^{+\infty} f(t)\varphi(\frac{t-b}{a})dt\ , \tag{11}$$

$$X_W(a,b)=a_0^{-m/2}\int_{-\infty}^{+\infty} f(t)\varphi(a_0^{-m}t-nb_0)dt\ . \tag{12}$$

Equation (12) is obtained by constraining a and b to discrete values: $a=a_0^m$, $b=nb_0a_0^m$ with $a_0 > 1$, $b_0 > 0$ fixed and $m, n \in \mathbb{Z}$. Similar to the Fourier transform, the wavelet transform is obtained by the inner product of f and a set of functions but instead of sinusoidals, it uses functions $\varphi_{a,b}$ called "wavelets". The difference here is that the window width varies adapted to the frequency, meaning that small values of a correspond to high frequency and a narrow time scale, while large values of a correspond to low frequency and a wide time scale. In addition, changing the parameter b, the center of the wavelet can be translated [35]. The major advance of the wavelet transform is its time and frequency localization ability [36, 37]. As a result, the wavelet transform enables multiresolution analysis.

To achieve the analysis of the capnogram, the time localization of the frequency features is preferred. From viewpoint of clinical interpretation, it is important to know in which phase of the capnogram the differences among the diverse asthmatic conditions are located. In addition, Wavelet transform is less computational complex

than Fourier Transform which is important for the expected application of the method in real time applications.

3.2.2 Segmented Wavelet Decomposition for Feature Extraction

The main idea behind the proposed method is to localize features in capnogram segments containing information relevant to classifying asthma severity. In other words, extracted segments have information on physiological meaning so, restricting features to segments makes it possible to interpret them, in the same sense, more easily than when features are extracted from the complete signal including phases that do not contain information about lungs status.

The proposed method for extracting features is shown in Fig. 3.4. Three stages of multiresolution analysis are designed -- each for a segment type -- A-B, E-F, or G-H. In each stage, the signal passes through a filter bank and emerges as two signals, then downsampling halves the two signals obtaining detail and approximation coefficients (CD and CA). The filter bank consists of low-pass (LPF) and high-pass (HPF) filters. The decomposition process is repeated until a desired level of decomposition is obtained.

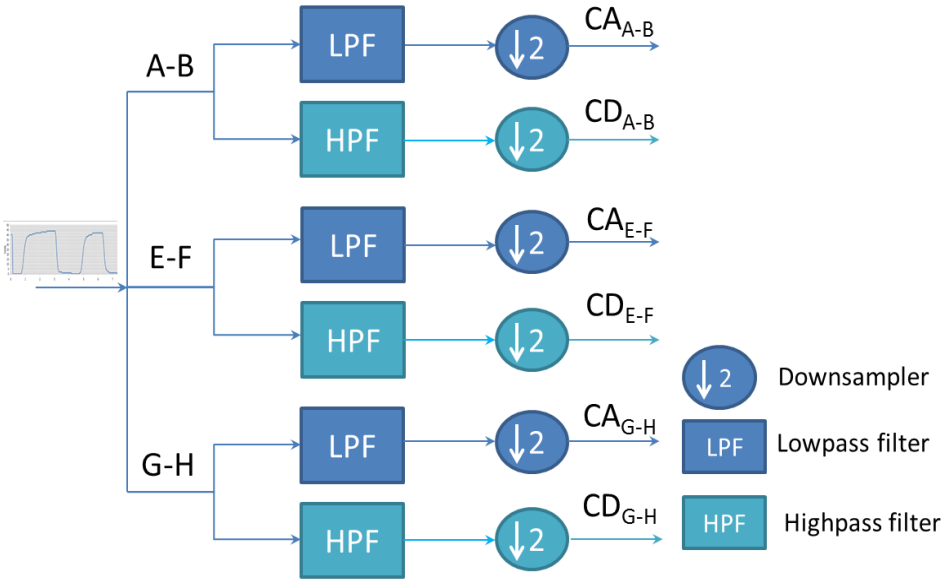


Fig. 3.4. Wavelet decomposition by segments.

3.3 Extracting Features from Real Capnogram Data

Data from 22 asthmatic subjects without symptoms and 4 subjects undergoing an asthma crisis during testing were collected during an asthma camp held at Julian Grimau clinic in Cuba. From 22 subjects without symptoms or classified by the physicians as asthma degree 0, 6 signals were excluded due to anomalies in morphology or low signal/noise (S/N) ratio, likely resulting due to involuntary mistakes by subjects while capnograms were being recorded. For subjects undergoing an asthma crisis, classified as degree 1, one signal was excluded from the study because expiration phases exceeded the limit established for a voluntary breath as reported in [17].

The number of samples of remaining asthma records is large -- 6070, 6536, and 6146 -- so they have been split to enlarge data for testing conveniently up to 7. In the

end, 23 capnograms were selected for experiments.

Capnograms were recorded using a commercial capnograph sensor (Capnosafe®, Combiomed, Cuba). The signal sampling frequency is 20 Hz and collected signals correspond to CO2 levels expressed in pressure units, i.e., mmHg, in the range of 0 - 100 mmHg [15].

The proposed method was implemented in MATLAB 7.9.0.529 and experiments were performed in a PC Intel (R) Core (TM) i7 CPU, 2.8 GHz and OS Windows 7 (32 bit) environment.

Experiments are conducted using the Haar wavelet transform. The number of wavelet coefficients is considerably reduced when wavelet decomposition is applied to the segments of interest, as shown in Tables 3.1 and 3.2. For 6803 capnogram samples, for example, detail coefficients of the first level of decomposition are reduced from 3402 to 2110, equivalent to a 40% reduction. The percentage of reduction is similar for remaining coefficients. This may improve the performance of feature selection and classification steps in decision support system.

Table 3.1. Number of wavelet coefficients for a complete and segmented capnogram of degree 1

	CA3	CD3	CD2	CD1	samples
Multilevel wavelet decomposition	**379**	**379**	**757**	**1513**	3026
Wavelet by segments					
A-B (31 segments)	1 x 31	1 x 31	1 x 31	2 x 31	96
E-F (30 segments)	1 x 30	1 x 30	2 x 30	4 x 30	108
G-H (30 segments)	16 x 30	16 x 30	31 x 30	62 x 30	1075

Table 3.2. Number of wavelet coefficients for a complete and segmented capnogram of degree 0.

	CA3	CD3	CD2	CD1	samples
Multilevel wavelet decomposition	851	851	1701	3402	6803
Wavelet by segments					
A-B (32 segments)	1 x 32	1 x 32	1 x 32	2 x 32	4 x 32
E-F (31 segments)	1 x 31	1 x 31	2 x 31	4 x 31	7 x 31
G-H (31 segments)	16 x 31	16 x 31	31 x 31	62 x 31	123 x 31

Because the main procedure for extracting features is based on the wavelet, it can be calculated in linear time. A multilevel bank requires a number of operations

$$\text{Operations} = fn + \frac{fn}{2} + \frac{fn}{4} + \frac{fn}{8} + \ldots \leq fn\left(1 + \frac{1}{2} + \frac{1}{4} + \ldots\right) = 2\,fn$$

where f is the filter length and n is the number of coefficients -- decreasing the number of coefficients thus reduces computational cost. The average execution time is 0.35 seconds for extracting features from one signal.

Execution time for a typical configuration of a physiological multiparameter monitor, assuming it has a 120 MHz microcontroller, is about 8 seconds. This time cost for extracting features of one capnogram signal is suitable in real-time applications.

3.3.1 Wavelets Coefficients by Segments as Features

In the procedure for obtaining capnogram features, individual cycles of signals were first segmented in segments of interest A-B, E-F, and G-H. Then, for each type of segment, the number of samples was normalized to an average calculated without

considering the minimum or maximum number of samples for that segment in the signal. Wavelet transform was applied afterwards to segments. This way, the same number of wavelet coefficients was obtained by segment type. Mean values were calculated last. After the foregoing procedure, mean values of approximation and detail coefficients of the third level of decomposition -- CA3, CD3, CD2, and CD1 -- were selected as features, being CA3 the approximation coefficients of the third level of decomposition and CD3, CD2, and CD1 the detail coefficients of decomposition levels 3, 2, and 1.

The following example for segment A-B illustrates how features were obtained from wavelet coefficients. In the equations below, CA_{A-Blij} and CD_{A-Blij} are detail and approximation wavelet coefficients of segment A-B. $\overline{CA_{A-Bl_J}}$ and $\overline{CD_{A-Bl_J}}$ are their average. Here, *l* is the level of decomposition, *i* is the number of wavelet coefficients, and *j* is the number of segments.

Wavelet coefficients for different decomposition levels 1, 2, and 3 were calculated by using Equations 15-18. Note from Equations 19-23 that the feature was finally obtained by averaging i^{th} wavelet coefficients over j^{th} segments.

$$CA_{A-Blij} = \{CA_{A-B3ij}\} \quad 1 \leq i \leq m,\ 1 \leq j \leq n\ , \tag{13}$$

$$CD_{A-Blij} = \{CD_{A-B3ij}, CD_{A-B2ij}, CD_{A-B1ij}\}, \tag{14}$$

$$CA_{A-B3ij} = \{\overline{CA_{A-B31_J}}\}, \tag{15}$$

$$CD_{A-B3ij} = \{\overline{CD_{A-B31_J}}\}, \tag{16}$$

$$CD_{A-B2ij} = \{\overline{CD_{A-B21_J}}\}, \tag{17}$$

$$CD_{A-B1ij} = \{\overline{CD_{A-B11_J}}, \overline{CD_{A-B12_J}}\}, \tag{18}$$

$$\overline{CA_{A-B31_J}} = \left(\sum_{j=1}^{n} CA_{A-B31j}\right)/n\ , \tag{19}$$

$$\overline{CD_{A-B31_J}} = \left(\sum_{j=1}^{n} CD_{A-B31j}\right)/n\ , \tag{20}$$

$$\overline{CD_{A-B21_J}} = \left(\sum_{j=1}^{n} CD_{A-B21j}\right)/n\ , \tag{21}$$

$$\overline{CD_{A-B11_J}} = \left(\sum_{j=1}^{n} CD_{A-B11j}\right)/n\ , \tag{22}$$

$$\overline{CD_{A-B12_J}} = \left(\sum_{j=1}^{n} CD_{A-B12j}\right)/n\ . \tag{23}$$

3.3.2 Classification Experiment

This section introduces classification experiments using support vector machines (SVMs) to evaluate the proposed feature extraction method.

Support vector machines (SVMs) are widely used to solve classification problems in different fields, such as bioinformatics, engineering, computer vision, and others. This supervised learning algorithm finds the optimal hyperplane that separates the data maximizing the distance from the decision boundary.

Sixteen capnograms of asthma degree 0 and 7 of asthma degree 1 are employed in experiments. Haar wavelet coefficients, i.e., CA3, CD3, CD2, and CD1, comprise the input vector of the classifier. Output is the degree of asthma severity, which has 2 classes, i.e., 0 no attack and 1 chest oppression.

The number of wavelet coefficients per segment varies among signals in the dataset. To build the input vector, the first m_0 wavelet coefficients per segment are selected where m_0 is the minimum number of wavelet coefficients per segment among signals from the two classes.

In the verification procedure, training data is first defined by applying cross-validation, i.e., by randomly selecting 90% of the dataset. The training of an SVM classifier with Gaussian radial basis function kernel (sigma = 1) is performed and the full dataset is then tested. Last, measures of sensitivity (Se), specificity (Sp), the correct rate (CR), and the error rate (ER) are calculated by using Equations 24-27. The procedure is repeated 10 times and performance evaluation measures are averaged:

$$Se = TP/(TP + FN)\,, \tag{24}$$

$$Sp = TN/(TN + FP)\,, \tag{25}$$

$$CR = (TP + TN)/(TP + TN + FP + FN)\,, \tag{26}$$

$$ER = (FP + FN)/(TP + TN + FP + FN)\,. \qquad (27)$$

where TP -- true positive -- is degree 1 of the asthma cases correctly classified as such, FN -- false negative -- is degree 1 of asthma cases incorrectly classified as degree 0, TN -- true negative -- is degree 0 of asthma cases correctly classified as such, and FP -- false positive -- is degree 0 of asthma cases incorrectly classified as degree 1.

Classification is performed for different feature vectors -- separately for wavelet coefficients of each segment of interest, combining two segments, combining three segments of interest, and using coefficients resulting from applying wavelet transform to the signal without segmentation. Table 3.3 lists results of experiments and Fig. 3.5 illustrates the correct rate for the different extracted features.

Table 3.3. Results of the classification experiment for validation of the extracted features. The performance evaluation of the SVM classifier includes measures of sensitivity, specificity, correct rate and error rate.

Feature	Se (%)	Sp (%)	CR (%)	ER (%)
A-B	97.50	37.14	79.13	20.87
E-F	96.25	52.86	83.04	16.96
G-H	99.38	55.71	86.09	13.91
A-B+E-F	**91.43**	**100.00**	**97.39**	**2.61**
A-B+G-H	99.38	72.86	91.3	8.70
E-F+G-H	99.38	77.14	92.61	7.39
A-B+E-F+G-H	99.38	75.71	92.17	7.83
standard WT	100.00	88.57	96.52	3.48

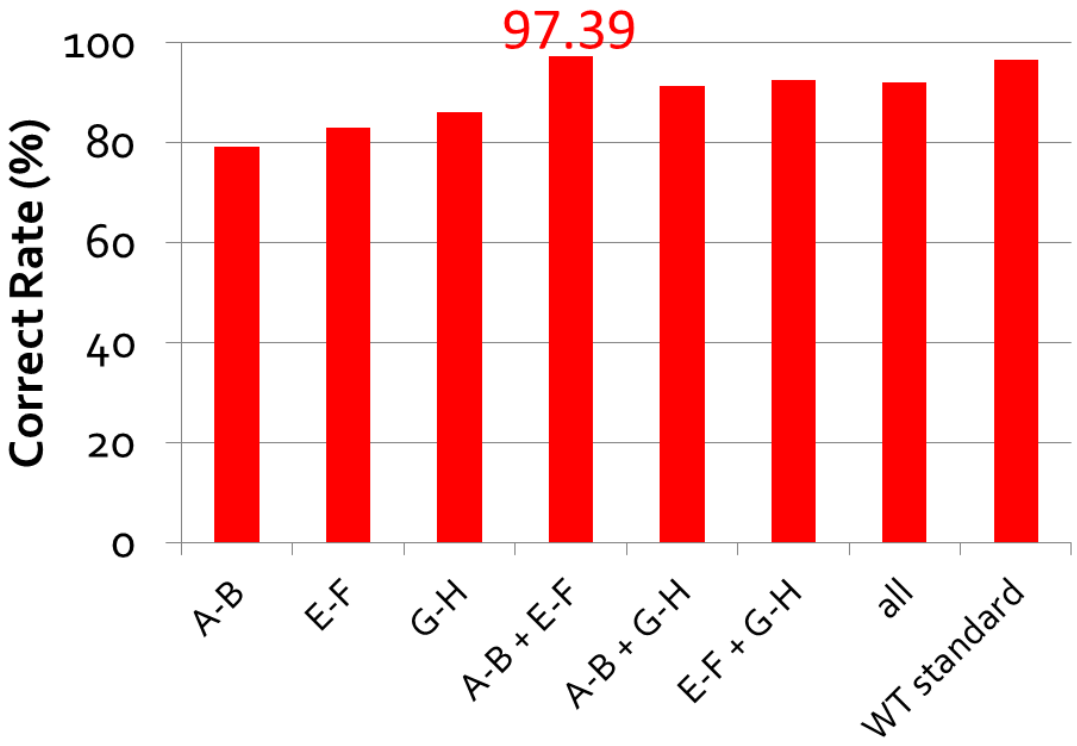


Fig. 3.5. Correct rate of the classification experiment using different features.

3.3.3 Discussion of the Results

The database includes cases of two classes from the six existent. These two classes correspond to degree 0 and degree 1 of the severity of asthma. It is well-known that, these two primary stages are the most difficult to be distinguish since the patient's symptoms are mild, intermittent and hard to be notice either by the patient or the physician. This fact corresponds to the findings of other groups that have approach the classification of asthma using the capnogram [17]. It is expected that the proposed features would be also able to differentiate higher degrees of asthma since all the symptoms, including the morphology of the capnogram, are significantly modified as the patient transits to higher stages of asthma severity.

The Haar wavelet transform is used to extract features because of its simplicity and

is a fast transformation algorithm [38]. In addition, it shows good results in analyzing signals with sudden transitions, as is the case with the capnogram. In general, for wavelet applications, the choice of the type of wavelet is still an unresolved question [39]. A common approach is to base the selection on the classification rates.

Comparing the performance of classifiers when wavelet coefficients from only one of the three segments is used, the best result is achieved for the G-H segment – *Se* = 55.71%, *Sp* = 99.38%, *CR* = 86.0%, and *ER* = 13.91%. In [17], it is suggested that terminal indices of the capnogram are highly sensitive to airway obstruction, as current results confirm.

Classifier performance is improved when wavelet coefficients of two or three segments of interest are used as features. The best result is achieved by combining segments A-B+E-F (*Se* = 91.43%, *Sp* = 100%, *CR* = 97.39%, and *ER* = 2.61%). Specifically, sensitivity (*Se* = 91.43%) is higher than in any other case. This finding suggests that differences exist in the inspiration phase (Segment A-B) of the capnogram from asthmatic subjects during an asthmatic crisis. As far as the authors know, no previous reports have been published in which the analysis of the capnogram's inspiratory phase has supported the assessment of asthma severity in subjects undergoing an asthmatic crisis.

The last classification experiment uses as features the wavelet coefficients of a signal without segmentation. The performance of the classifier is good (*Se* = 88.57%, *Sp* = 100%, *CR*=96.52%, and *ER* =3.48%), but does not outperform that when coefficients of segments A-B+E-F are used.

The performance of the classifier when G-H is used is good compared to that when A-B or E-F are used separately. Combining A-B+E-F, however, produce a better performance than G-H alone. When only feature AB+E-F is accounted for, for example, a *CR* of 97.39% is achieved, while for G-H alone, *CR* is 86.09%. When both features are accounted for, however, *CR* decreased to 92.17%. This may suggest that dependence

or interaction exists between feature A-B+E-F and feature G-H.

In addition to this, the SVM assumes features to be independent, i.e., it collects the evidence for a class from individual features separately, after which it aggregates all pieces of evidence. Assuming that there is interaction, this classifier may fail and thus degrade its performance.

This may have several interpretations from a physiological viewpoint. One is that dependence exits between CO_2 values on the plateau of the expiration phase (G-H) and CO_2 values of the expiration rise (E-F). This is a more consistent interpretation because the two expiration phases are closely related. A second interpretation may imply dependence also exists between the expiration plateau and CO_2 values of the previous inspiration cycle (A-B).

Further studies are needed to clarify these hypotheses. To the best of our knowledge, no report published in the literature has discussed dependence between expiratory phase of a respiratory cycle and inspiratory phase of the previous one. The study of interactions among the wavelet coefficients of different segments is an interesting topic and may be important for classifying asthma severity, so it could also add further perspective to this research.

Results generally show better specificity than sensitivity, i.e., the classifier does not predict the degree 1 class so well. This may due to the small number of cases of class in question (7) compared to that of the degree 0 class (16). For different features, the number of false positives is low, being zero in two cases. This is supported by the high specificity index value obtained for different feature vectors. In contrast, the classifier required that segments be combined for reducing the number of false negatives and improving the sensitivity index. These facts suggest that it is more complex to differentiate subjects without asthma symptoms from subjects having mild symptoms. One possible reason leading to this misclassification is that the dataset used in experiments contained only signals for the first two degrees of asthma, i.e., in which

symptoms are mild and thus the feature variation. An analysis of results makes it possible to summarize the following advantages of the proposed method over the standard wavelet transform:

1. A reduction in the number of features

The segmented wavelet transform implies segmentation of the capnogram in segments of interest and afterward the application of the wavelet transform to selected segments. This reduces the number of wavelet coefficients while improving the efficiency of subsequent steps in the classification of asthma severity.

2. Interpretability of features

Since features (wavelet coefficients) are limited to segments of interest that contain important information about asthma, it is possible to interpret them from a physiological viewpoint. This is important for better understanding the disease through computerized analysis of features.

3. Best performance for combined segments A-B+E-F

Comparing classification results when standard wavelet transform is used to that when segmented wavelet transform is used, combining segments A-B+E-F shows the best performance in sensitivity at 91.43% vs. 88.57%, in correct rate at 97.39% vs. 96.52%, and in error rate at 2.61% vs. 3.48%. The specificity value in both cases is the same, i.e., 100%.

3.3.4 Comparison with T.T.Kean's Best Features

In [18], 13 features are extracted automatically from the capnogram; 7 are proposed in [17] and the rest are newly introduced for the purpose of the capnogram analysis. Two of them, namely the slope ratio (SR) and the Hjorth parameter mobility (HP2-

mobility), show the best statistical significance ($p < 0.0001$ and $p = 0.0001051$ respectively). These two are extracted according with [18] and work as input features of an SVM classifier. Fig. 3.6 and Fig. 3.7 illustrate the extraction of SR and HP2-Mobility in a respiratory cycle of the database respectively. Training and testing of the classifier are performed alike the validation experiment.

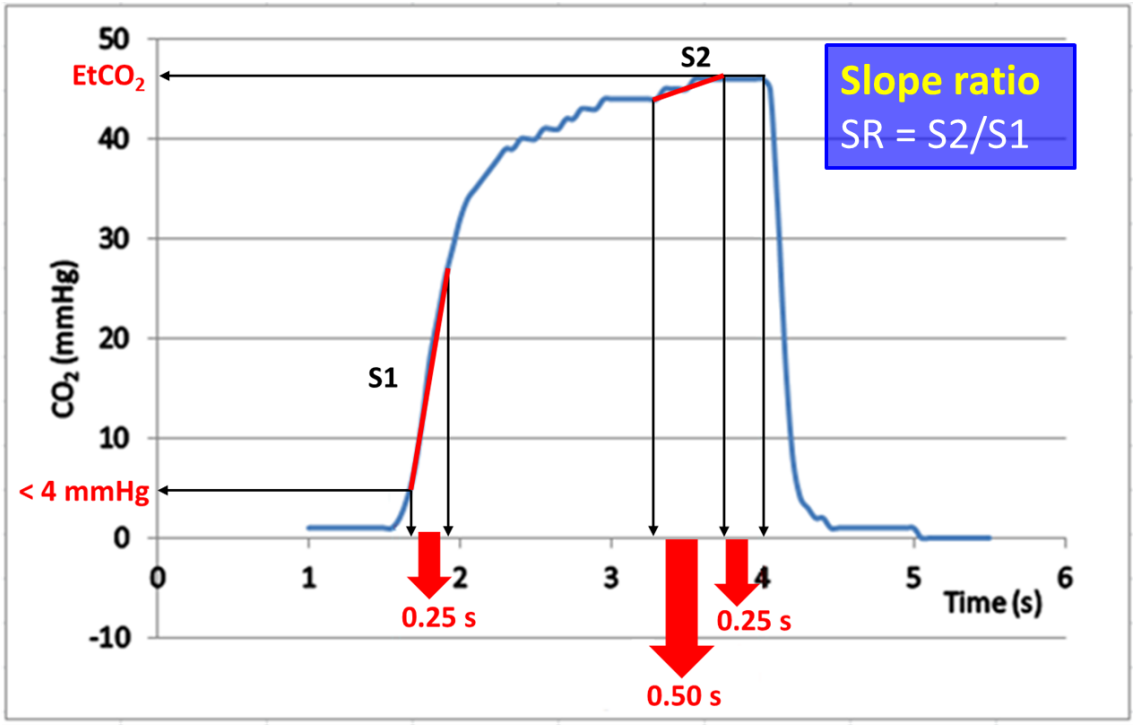


Fig. 3.6. Slopes S1 and S2 are extracted from a respiratory cycle and the slope ratio (SR) is determining according to [18].

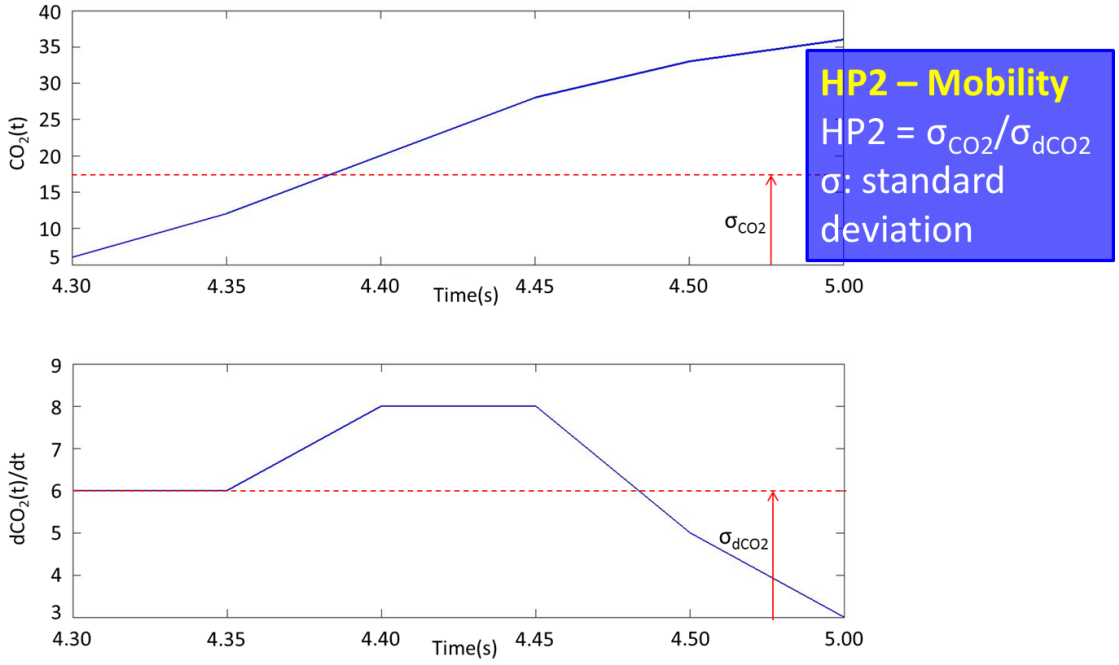


Fig. 3.7. The plot at the top shows how the standard deviation is calculated for the expiration phase of a respiratory cycle. The one at the bottom shows the calculation of the standard deviation of the derivative of the expiration phase. The ratio of this two deviation results the Hjorth parameter mobility.

Results are presented in Table 3.4 compared to the best classification results obtained when proposed features are used.

Table 3.4. Performance of the classifier when using the features with best statistical significance proposed by T.T. Kean et al. [18] and the wavelet coefficients by segments.

Feature	Se (%)	Sp (%)	CR (%)	ER (%)
T.T.Kean	20.00	96.25	73.04	26.96
Proposal	**91.43**	**100.00**	**97.39**	**2.61**

The proposal improves T. T. Kean's method in sensitivity, specificity, the correct

rate, and the error rate. The first three parameters increased by 71.43%, 3.75%, and 24.35% and the last parameter, i.e., error rate, decreased by 24.35%. Results of experiments thus confirm that extracted wavelet coefficients differentiate between degrees 0 and 1 of asthma severity with high performance.

3.4 Chapter Summary

A feature extraction method from capnogram based on wavelet decomposition is has been proposed. Extracted features are validated by classification experiments using 23 capnograms from asthma camp participants in Cuba and an SVM classifier.

Experimental results have demonstrated that features differentiate 6 degrees of asthma severity, especially for degrees 0 and 1. The best classification result is achieved by combining segments A-B + E-F (Se = 91.43%, Sp = 100%, CR = 97.39%, and ER = 2.61%). This is a new finding because segment A-B is part of the inspiration phase of the capnogram that was not considered in previous studies of asthma assessment using capnography. Further validation of this fact from a clinical point of view is necessary, however, to explain its correspondence to asthma pathology.

Execution time has been estimated for a typical configuration of multiparameter monitors in order to evaluate its feasibility in this scenario. The result, 8 seconds, is suitable for this kind of device. Experimental results have confirmed that the proposal has low computational cost and suitable performance for the real-time classification of asthma severity.

Classification results using an SVM classifier and the wavelet coefficients extracted by the proposed method are better than those obtained when the same classifier and features having the best statistical significance proposed by T. T. Kean are used.

Our proposal is expected to become part of a decision support system for asthma classification that is under development by research group at the Tokyo Institute of Technology (TITECH) and Tokyo Medical and Dental University (TMDU). It is also being planned for implementation in physiological multiparameter monitors that are being developed at the Central Institute of Digital Research in Cuba.

Chapter 4

Decision Support System for Assessing Asthma Severity Based on Capnogram Analysis

Computer-based decision support systems for assessing the severity of asthma are necessary to support clinicians in the evaluation of the patient. There are studies demonstrating that the least clinically experienced providers are those which evaluation differs from the outcome of a clinical decision support system (CDSS), representing true differences from the appropriate care [40]. In general, decision support systems improve care by providing intelligently filter patient-specific information and can give advice to clinicians in appropriate time [41].

A number of computer-based systems have been proposed in order to cope with the problem of misdiagnosing of asthma. Different classification techniques are involved in the solutions such as fuzzy logic [42, 43], neural networks [44, 45], decision tree [46], Naïve Bayes [46], etc. All of them evaluate symptoms, prognostic factors, epidemiological factors, lung function test, and others, alone or by combining some of them. The outcomes of the systems vary from binary classification of asthma to classification in three to five degrees of asthma.

In this research work, a decision support system is proposed based on the analysis of the capnogram signal. Spirometric parameters currently used for classifying asthma

severity have numerical scales to make it easier for the clinician to make an evaluation. The concept behind developing a decision support system for assessing asthma severity involves to "translating" capnogram shape changes into a numerical index supporting the physician in classifying asthmatic subjects and in evaluating improvements during treatment. Our proposal also give objective assessment of asthma been easily interpreted by physicians and patients. In addition, it can be used either in primary care setting, specialty clinic, and surgery rooms.

With the aim to integrate the two proposals in this investigation and evaluate their influence in the classification result, two different architectures for the decision support system are evaluated. The first architecture combines the segmented wavelet transformation method for feature extraction with a fuzzy classifier and the second one adds a step of similarity evaluation after the extraction of the wavelet coefficients.

In ensuing subchapters the two architectures are introduced and evaluations of them are presented.

4.1 Decision Support System Architecture 1

The architecture of the proposed system, shown in Fig. 4.1, consists of three main steps. First, the capnogram signal, which is the system input, is segmented into segments of interest as previously describe in Chapter 3 (see Fig. 3.2).

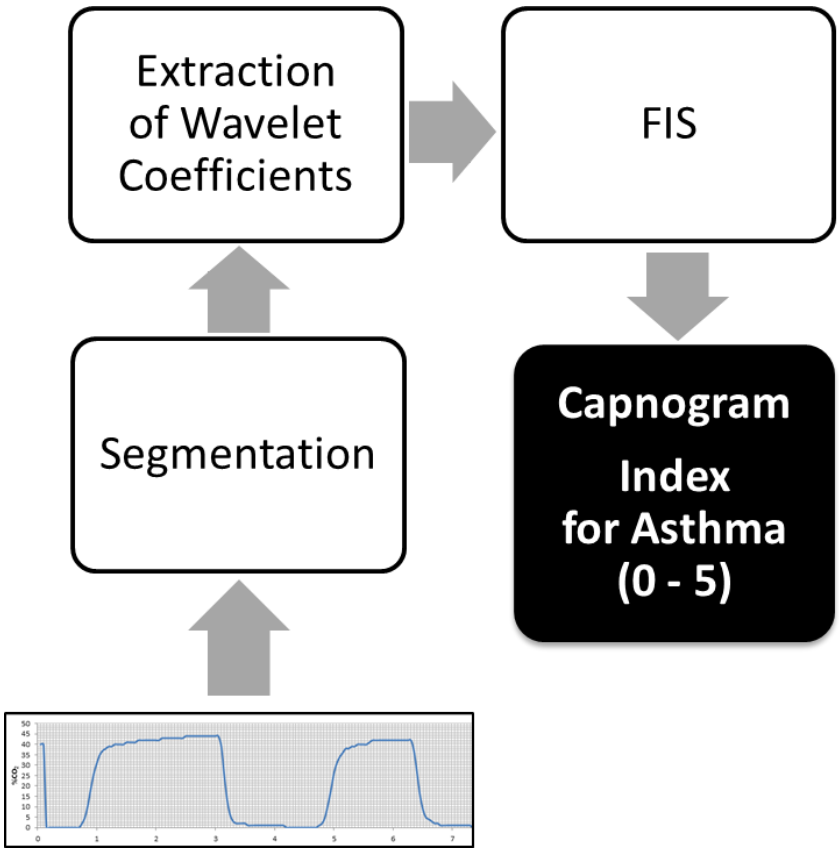


Fig. 4.1. Architecture 1 of decision support system for asthma classification.

Second, features are extracted from the capnogram signal. The proposed feature extraction method based on wavelet decomposition by segments is used [47]. A maximum number of 6 inputs, corresponding to the wavelet coefficients of the combination of two segments of interest, A-B and E-F, are the inputs of the fuzzy inference. These features offer the better classification results in [47].

Third, capnograms are classified on scale from 0 to 5 (6 degrees) corresponding to asthma attack severity. Severity itself is classified into 6 degrees: 0 no attack, 1 chest oppression, 2 mild attack, 3 moderate attack, 4 severe attack, and 5 severest attack, based on classification recommended by Asthma Prevention and Management Guideline, Japan. Fuzzy inference system is built for the classification of the capnografic signal with automatic tuning of the membership functions.

4.1.1 Fuzzy Inference for Classification based on Segmented Wavelet Coefficients

One of the most known and widely used methods for automatic generation of fuzzy rules from numerical data is the Wang and Mendel's method. It has been applied to a variety of problems and is one of the benchmark methods in the field [48].

This method has 5 main steps for generating the fuzzy rules and showing how to use these fuzzy rules to obtain a mapping from input space to output space. Step 1 divides the input and output spaces into fuzzy regions. Step 2 generates fuzzy rules from given desired input-output data pairs. Step 3 assigns a degree to each generated rule. Step 4 forms the combined fuzzy rule base. Finally, Step 5 presents a defuzzifying procedure for obtaining a mapping based on the combined fuzzy rule base [49].

Wang and Mendel's method is used in this research to build the membership functions and the fuzzy rules for classification of asthma degree. From the analysis of the data, three membership functions are designed for each of the inputs: *small*, *medium*, and *high*. For the output, two membership functions, *degree0* and *degree1* are built. The used procedure is hereinafter described.

Defining input and output fuzzy regions

Let's k be the number of samples (vector of wavelet coefficients) in the database and n the number of features (wavelet coefficients). We start from a set of k desired input-output data pairs:

$(c_1^{(1)}, c_2^{(1)}, \ldots c_n^{(1)}, y^{(1)}), (c_1^{(2)}, c_2^{(2)}, \ldots c_n^{(2)}, y^{(2)}), \ldots (c_1^{(k)}, c_2^{(k)}, \ldots c_n^{(k)}, y^{(k)})$

where c_1, c_2, ... , and c_n are the wavelet coefficients and y is the classification output.

First, the domain interval of c_1, c_2, ... c_n and y is defined as:

$[c_1^-, c_1^+]$, $[c_2^-, c_2^+]$,…, $[c_n^-, c_n^+]$, and $[y^-, y^+]$

where c_i^- (c_i^+) is the minimum (maximum) value of the *i* wavelet coefficient in the database, determining by the expressions:

$$c_i^- = min_j(c_i^{(j)}, c_i^{(j+1)}, \dots, c_i^{(k)}), \quad i = 1, \dots, n, \ j = 1, \dots, k \tag{28}$$

$$c_i^+ = \max_j(c_i^{(j)}, c_i^{(j+1)}, \dots, c_i^{(k)}), \quad i = 1, \dots, n, \ j = 1, \dots, k \tag{29}$$

and y^- (y^+) is the minimum (maximum) value of the output, in our case 0 and 1 respectively.

Second, four intermediate values (*l*) inside each domain interval are determining by the formula:

$$c_i^l = \frac{c_i^j + c_i^{j+1}}{2} \tag{30}$$

where $= l \times \frac{k}{l+1}$, $l = 1,\dots, m\text{-}1$ and *m* is the number of membership functions

Third, each domain interval of the inputs is divided in three fuzzy regions denoted by *small*, *medium*, and *high*; and the domain interval of the output is divided in two fuzzy regions *degree0* and *degree1*. To each fuzzy region a trapezoidal membership function is assigned as described below.

For membership function *small*, the membership function is 0 for the minimum value c_i^- of the wavelet coefficient *a* and for *d*, which is the intermediate value c_i^2. The value *c* is equal to the intermediate value $c_i^1$and is assigned the degree of membership function as 1. For *medium*, the degree of membership function 1 (*b* and *c*) is assigned to intermediates values c_i^2 and c_i^3. To tune the parameters *a* and *d*, these coincide with intermediate values c_i^1 and c_i^4 respectively. In the case of membership function *high*, the degree of membership function 0 is assigned to values *a* and *d* as for *small* but *a* is equal intermediate value c_i^3 and *d* is equal to maximum value c_i^+. The degree of membership function 1 is assigned to *b* and *c* where *b* is equal the intermediate value $c_i^4$and *c* is equal to maximum value c_i^+. For illustration about this procedure see Fig. 4.2.

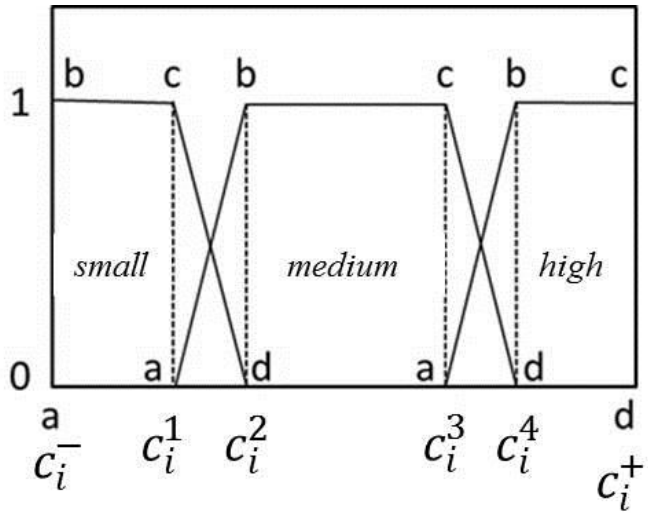


Fig. 4.2. Tuning the membership functions of the input.

For tuning the trapezoidal membership function of the output, maximum and minimum of the domain interval (y^-, y^+), which are 0 and 1 respectively, are assigned as degree of membership function 1. The same degree is assigned for values 0.4 and 0.6. Degree of membership function 0 is assigned to values *a* and *d* from *degree0* and *degree1*. Fig. 4.3 shows the membership functions of the output.

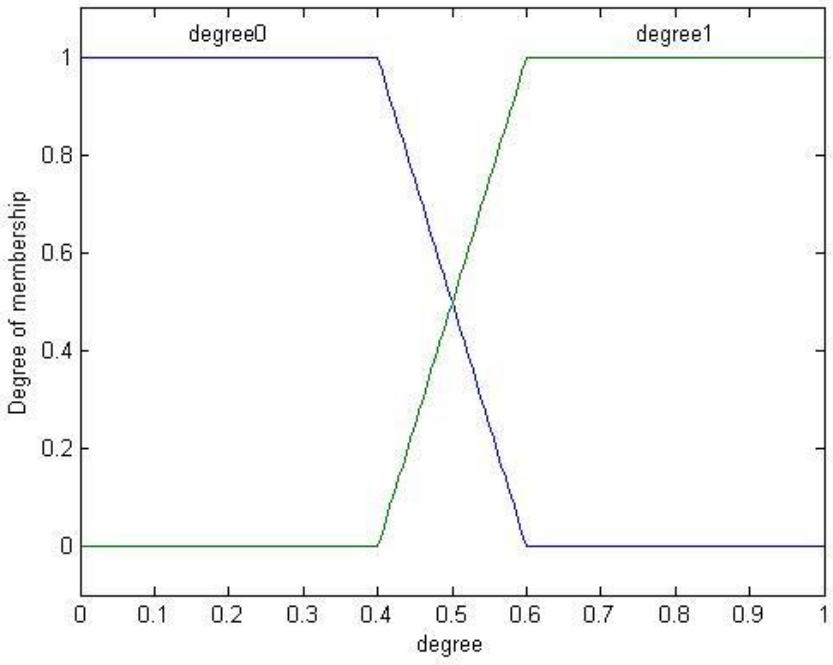


Fig. 4.3 Tuning of the output membership functions.

Constructing the fuzzy rules

First, membership function *small*, *medium*, or *high* is assigned to each wavelet coefficient of the input vector. We used 1, 2, and 3 for designating the membership functions. This procedure is repeated for all the samples in the training dataset. As a result, a matrix of rules with number of rows equal the number of samples (*k*) in the training dataset and number of columns equal number of wavelet coefficients (*n*) is obtained.

This matrix of rules has n+1+2 columns where the first *n* columns refer to the number of wavelet coefficients. Each column contains a number that refers to the index of the membership function for that variable. The next column refers to the outputs of the system. Each column contains a number that refers to the index of the membership function for that variable. The n + 1 + 1 column contains the weight that is to be applied to the rule. The weight must be a number between zero and one, and we left as one. The n + 1 + 2 column contains a 1 since the fuzzy operator for the rule's antecedent is AND.

Eliminating conflicting rules

Since there are many input-output pairs in the dataset, and each data pair generates a rule, it is probable that there will be some conflicting rules, i.e., rules that have the same IF part but a different THEN part. Since the output has only two possible values, from the pair of conflicting rules we select the one that produces better classification result. In this way not only the conflict is solved but also the number of rules is reduced.

An example of rule of our system is rule 3 = [1 3 3 1 2 2 2] which can be interpreted as: IF c_1 is *small* AND c_2 is *high* AND c_3 is *high* AND c_4 is *small* AND c_5 is *medium* AND c_6 is *medium* THEN output is *degree1*.

Defuzzification strategy to determine the output

To determine the defuzzified output we use the Matlab function *evalfis*, which given a matrix specifying the input values and a FIS structure performs fuzzy inference

calculations and computes the output variables.

4.1.2 Classification of Asthma Severity

Once obtained the defuzzified output, it is evaluated in the membership functions of the output. The results of the evaluation for the two membership functions are compared and the highest value determines the classification of the input.

4.2 Decision Support System Architecture 2

In architecture 2 of the decision support system, after extract the wavelet coefficients from the segments, a step of similarity evaluation is performed in order to reduce the number of wavelet coefficients to the two measures of similarity as in [32]. The rest of this architecture is similar to the previously described in 4.1. Fig. 4.4 shows the architecture 2 of the decision support system. In ensuing subsections steps of this architecture are detailed.

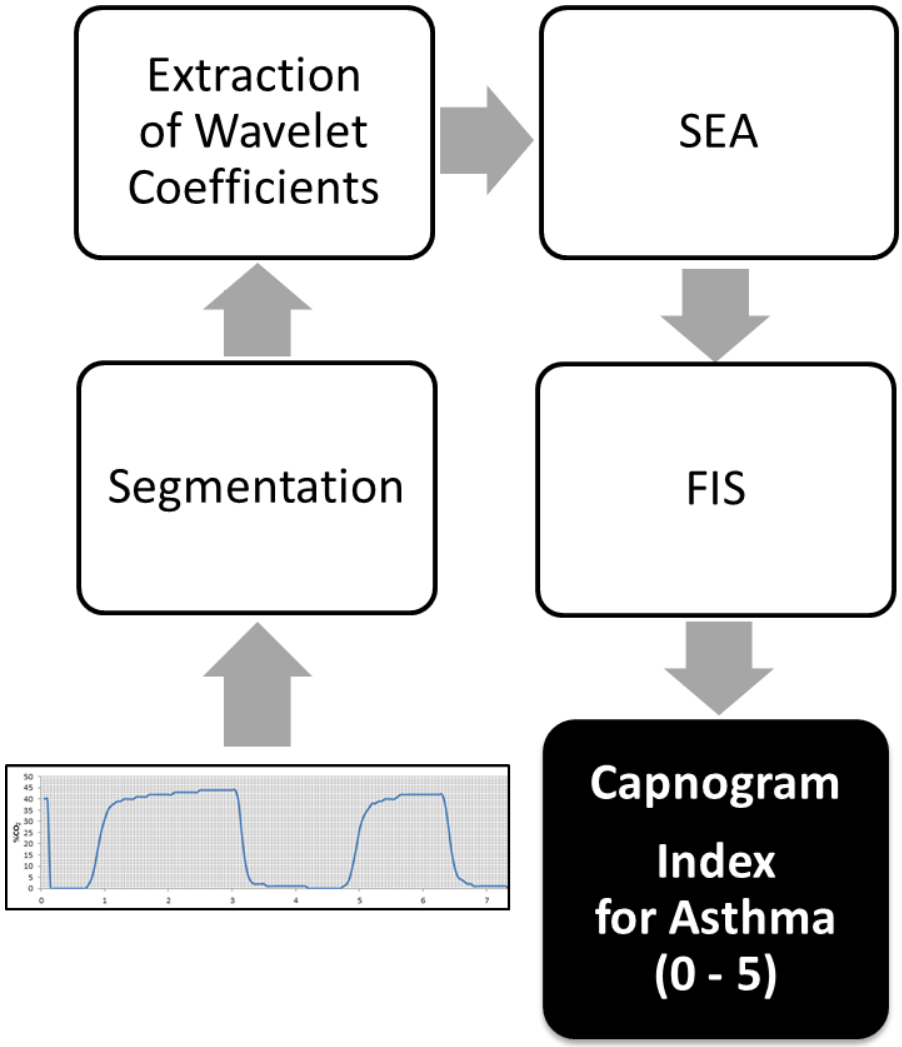


Fig. 4.4. Architecture 2 of decision support system for asthma classification

4.2.1 SEA of the Segmented Wavelet Coefficients

The wavelet coefficients from segments of interest A-B and E-F of 23 capnograms are the input signals of this step of the decision support system. The patterns are 6 of degree 0 and 6 of degree 1 signals. They are selected randomly from the input signals. After applying the SEA as described in Section 2.2.1, 276 values of each measure of

similarity, PRD (SD = 0.02) and Corr (SD = 10.3) are obtained.

4.2.2 Classification Using Fuzzy Similarity Evaluation

Fuzzy similarity evaluation is performed using similar membership functions of inputs and the output to those in the first proposal in this thesis [32] and the same fuzzy rules. An average approach is used for classification. The results of fuzzy similarity evaluation of each signal respect to each type of pattern is average and the highest mean value determine the final classification. For instance, if the mean similarity value when matching with the degree 0 patterns is higher than the result of matching with the degree 1 patterns, then the signal is classified as *degree 0*.

4.3 Validation of the Decision Support System \

Twenty three capnograms from 26 asthmatic subjects collected during an asthma camp held at clinic Julian Grimau in Cuba are used for the experiments. Description of the dataset, the capnogram sensor used for the measurements, and the experimental setup is presented in Section 3.3.

4.3.1 Experiments on Capnogram Data Using Architecture 1

From now let's explain the procedure that was followed in this validation experiments. First, training set is defined randomly selecting 6 degree 0 and 6 degree 1 signals, meaning that 52% of the data is used for tuning the membership functions. Meanwhile, testing set is defined as the rest of the signal of both type, in our case, 10 signals of degree 0 and 1 of degree 1. Second, wavelet coefficients from segments of interest A-B and E-F until third level of decomposition are extracted using Haar wavelet transform. Analyzing the coefficients it was observed that some of them, CD_{12} and CD_{31} (detail coefficient 2 of first level of decomposition and detail coefficient 1 of third level of decomposition), were 0 in A-B segments for all the signals and for E-F were 0 for the majority of the signals. These coefficients were removed. Third, the selected wavelet coefficients are used for building the membership functions (*small, medium,* and *high*) and the rules of the fuzzy inference using the procedure described in Section 4.1.1. Figs. 4.5 – 4.11 illustrate the membership functions of inputs and output, Table 4.1 shows values of the tuned parameters of the membership functions, and Table 4.2 shows the constructed fuzzy rules. Fourth, testing set is classified as described in Section 4.1.2.

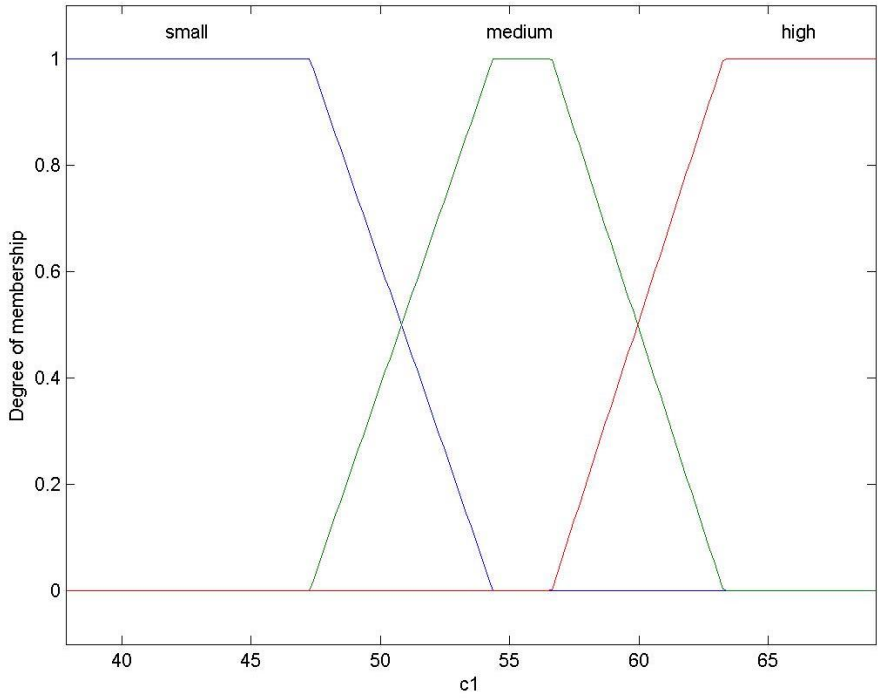


Fig. 4.5. Membership functions of wavelet coefficient c1

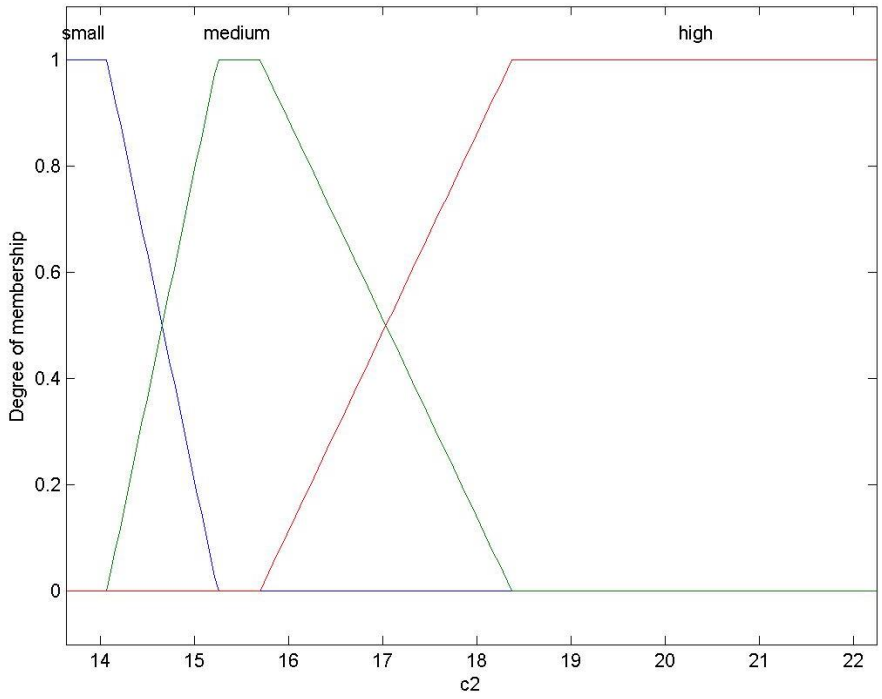


Fig. 4.6. Membership functions of wavelet coefficient c2

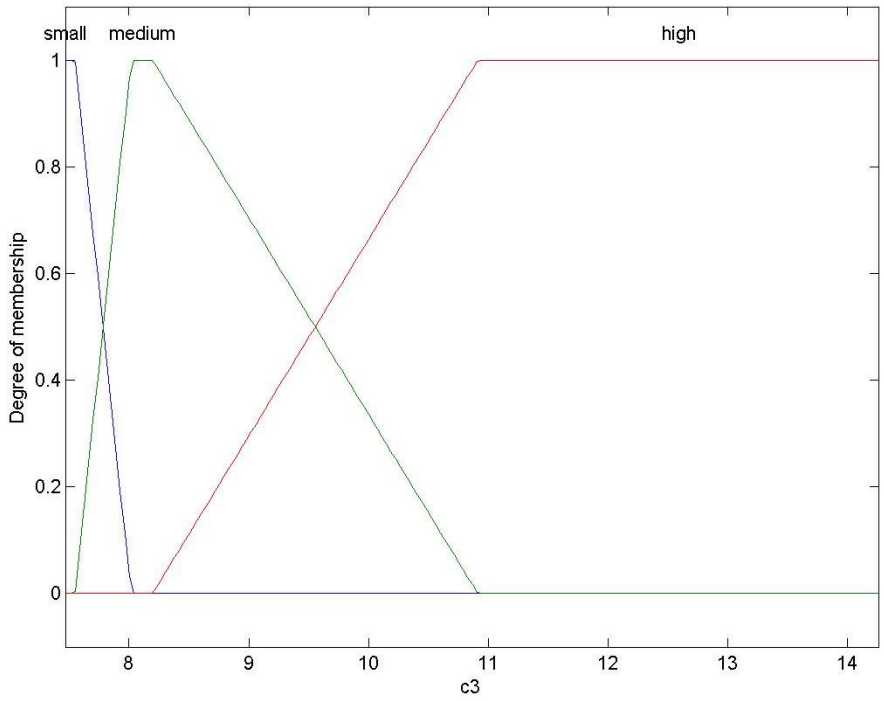


Fig. 4.7. Membership functions of wavelet coefficient c3

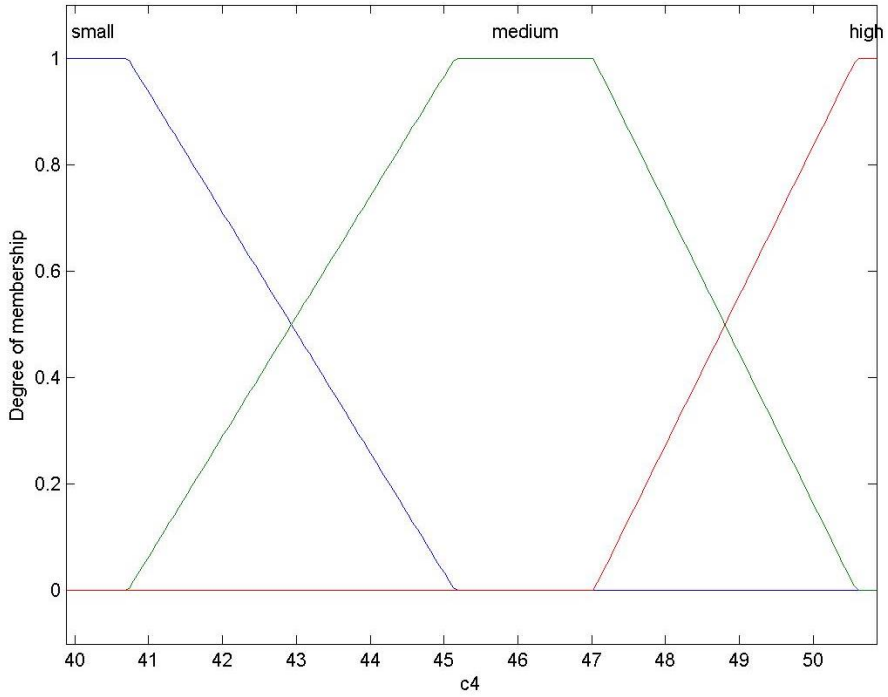


Fig. 4.8. Membership functions of wavelet coefficient c4

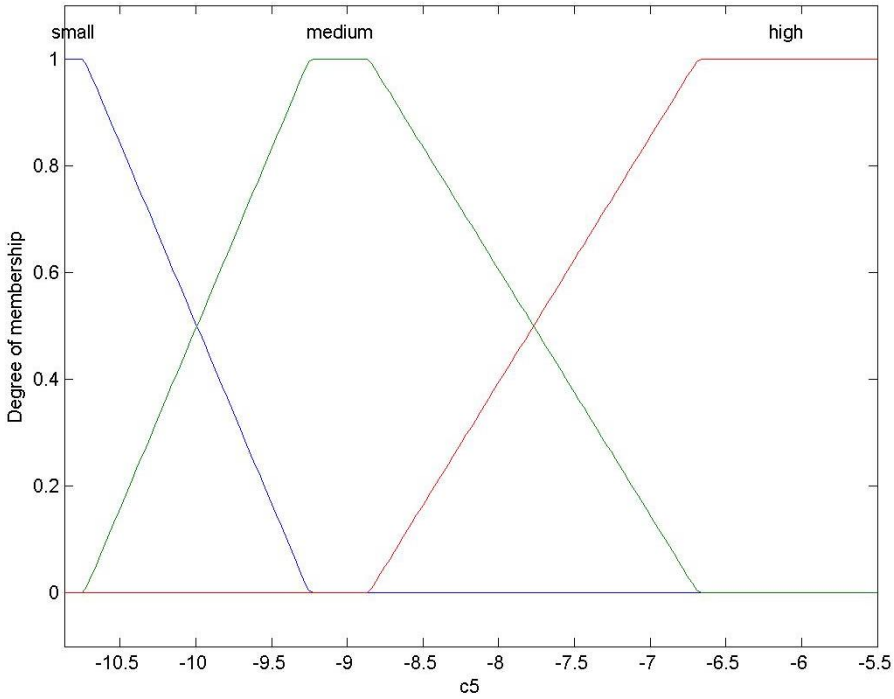


Fig. 4.9. Membership functions of wavelet coefficient c5

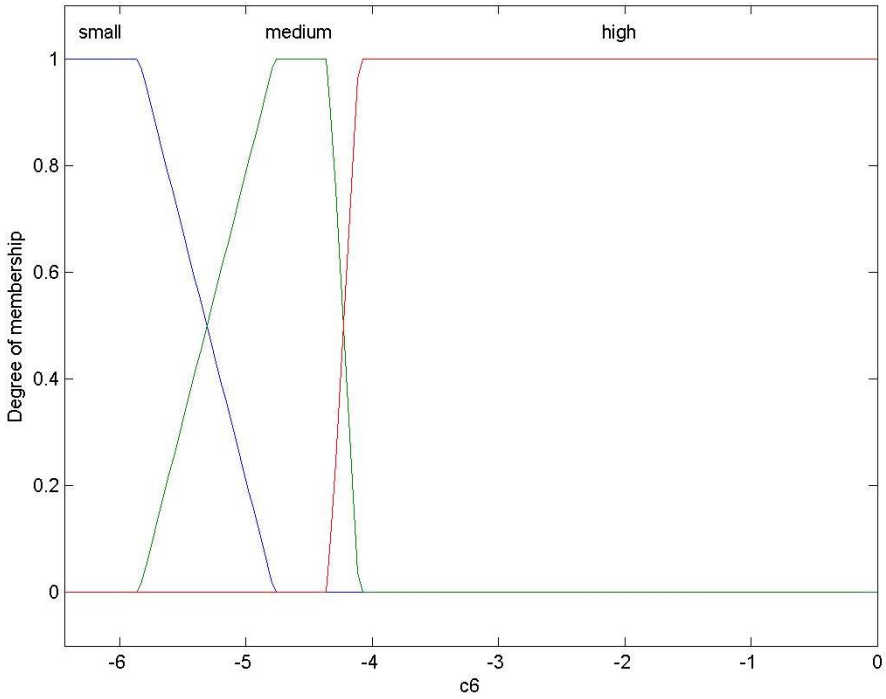


Fig. 4.10. Membership functions of wavelet coefficient c6

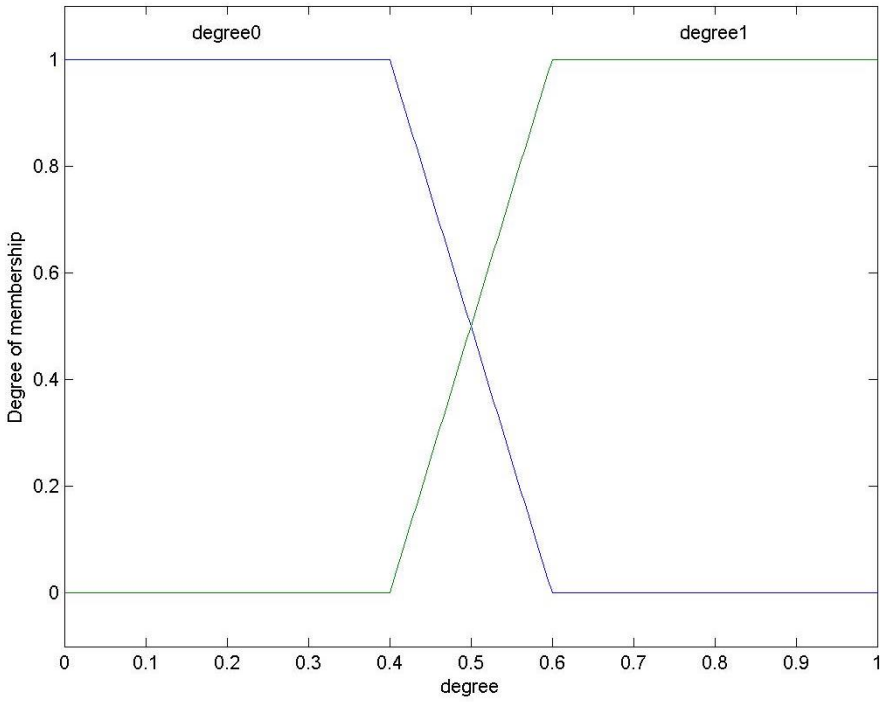


Fig. 4.11. Membership functions of the output

Table 4.1. Parameters of membership functions of wavelet coefficients

	μ_{small}				μ_{med}				μ_{high}			
	a	*b*	*c*	*d*	*a*	*b*	*c*	*d*	*a*	*b*	*c*	*d*
C1	37.8	37.8	47.2	54.3	47.2	54.3	56.6	63.2	56.6	63.2	69.2	69.2
C2	13.6	13.6	14.1	15.2	14.1	15.2	15.6	18.4	15.6	18.4	22.3	22.3
C3	7.5	7.5	7.6	8.0	7.6	8.0	8.2	10.9	8.2	10.9	14.3	14.3
C4	39.9	39.9	40.7	45.1	40.7	45.1	47.0	50.6	47.0	50.6	50.9	50.9
C5	-10.9	-10.9	-10.7	-9.2	-10.7	-9.2	-8.9	-6.7	-8.9	-6.7	-5.5	-5.5
C6	-6.4	-6.4	-5.8	-4.8	-5.8	-4.8	-4.4	-4.1	-4.4	-4.1	0.0	0.0

Table 4.2. IF-THEN rules obtained used Wang and Mendel method

Rule	C1	C2	C3	C4	C5	C6	Out
1	1	1	1	2	2	2	2
2	1	2	2	1	2	1	1
3	1	3	3	1	2	2	2
4	2	1	1	3	1	1	2
5	2	1	2	2	1	3	1
6	2	2	2	3	1	1	1
7	2	3	3	3	2	2	2
8	3	1	1	2	2	2	1
9	3	2	2	1	3	3	1
10	3	2	2	2	3	3	2
11	3	3	2	2	3	2	2

When the defuzzied output is evaluated in the membership functions *degree0* and *degree1*, in some cases, mainly for degree 0 signals, the evaluation value is 0.5 for both membership functions. In these cases the signal was classified as degree 0. As result of the classification presented in Table 4.3, 1 of the degree 1 signals and 2 of the degree 0 signals were misclassified. Values of 85.7, 87.5, 86.9, and 13.0 are achieved for sensitivity, specificity, correct rate, and error rate respectively.

Table 4.3. Results of classification using architecture 1 of decision support system

	Degree 0	**Degree 1**
Degree 0 state	14	2
Degree 1 state	1	6

4.3.2 Experiments on Capnogram Data Using Architecture 2

For these experiments some adjustments to the tuning parameters of the membership functions of the input are done based on differences in the boundaries of the measures of similarity in respect of those obtained in [32]. Table 4.4 shows the adjusted tuning parameters. The parameters of the membership function of the output are the same as in [32].

Table 4.4. Tuning parameters of similarity measures result of SEA of wavelet coefficients

	μ_{low}		μ_{med}		μ_{high}	
	Corr	*PRD*	*Corr*	*PRD*	*Corr*	*PRD*
a	0	-26.3	0.927	2.196	0.990	26.26
b	0	0	0.969	13.33	0.999	47.74
c	0.927	2.196	0.990	26.16	1	100
d	0.969	13.09	0.999	47.74	1	100

The experiment is repeated using different patterns and the performance of the classifier changed each time. Values of 100%, 75%, 82.6%, and 17.4% for sensitivity, specificity, correct rate, and error rate are achieved as best result.

4.3.3 Comparative Analysis of Experimental Results in Validation of Decision Support System Architectures

Table 4.5 Summarize the results of validation experiments of architectures 1 and 2 of the proposed decision support system for classifying asthma severity.

Table 4.5. Validation experiments results of the two decision support system architectures.

	Se (%)	Sp (%)	CR (%)	ER (%)	Execution Time (s)
Architecture 1	85.7	87.5	86.9	13.0	1.41
Architecture 2	100.0	75.0	82.6	17.4	6.09

In general, both architectures give similar good results in terms of correct rate, with slight better value for architecture 1 (85.7%). Architecture 2 has 100% sensitivity, this means that all the degree 1 signals are well classified, being cero the number of false negative. For this architecture, however, specificity is low meaning that the false positive cases are high. These results are consistent with those obtained in the similarity fuzzy classification in which also the sensitivity (83.3%) was higher than specificity (80.9%).

Execution time presented in the table was measured for the experimental setup. This parameter differentiates the most the evaluated architectures. This parameter represents the most significant difference between the evaluated architectures and could define the application for which to use each of them. For example, for monitoring the asthmatic patient during general anesthesia use, architecture 1 is the most suitable. Considering a

typical configuration of multiparameter monitor could has a 200-470 MHz microprocessor (ARM926) this time would be in the range 8-20 sec which is acceptable for a real time monitoring. Meanwhile, architecture 2 can be used for ambulatory monitoring of the asthmatic patient where the execution time is not so important as the correct classification of the degree 1 patient it is.

4.5 Chapter Summary

Two different architectures for the decision support system are evaluated. Architecture 1 combines the segmented wavelet transformation method for feature extraction with a fuzzy classifier and Architecture 2 includes a step of similarity evaluation after the extraction of the wavelet coefficients.

Experiments testing the architectures of the decision support system on 23 capnogram data collected during an asthma camp in Cuba achieve similar results of correct rate (86.9% vs. 82.6%).

The time for accomplishing the classification of one capnogram signal is 1.41 seconds and 6.09 seconds for Architecture 1 and Architecture 2 in the experimental setup. The time estimated for a typical configuration of multiparameter monitor would be in the range 8-20 seconds (for Architecture 1) which is acceptable for a real time monitoring.

The proposal intends to support the clinicians in the complex task of asthma classification while provides algorithms which may improve those developed in previous research works. Important applications are foreseen for the proposal. Since it is based on capnogram analysis, can be used in children under 6 years old. In addition, it can be employed in surgery rooms in the monitoring of patients under anesthesia. Moreover, the proposed decision support system is planned to be implemented in multiparameter monitors that are developed in the Central Institute of Digital Research in Cuba.

.

Chapter 5

Conclusions and Future Perspectives

5.1 Conclusions

Similarity-Based Fuzzy Classification of ECG and Capnogram Signals

This research is a first attempt for the classification of asthma based on the capnogram signal. As result, a method based on fuzzy similarity evaluation is proposed. It is demonstrated the ability of the method for classification purposes by experiments using databases of capnogram and ECG signals. The performance of the classifier for the capnogram data is acceptable showing 81.5 % of accuracy. The combination of the shape exchange algorithm (SEA) and fuzzy inference reduces the number of features to two measures of similarity (PRD and Corr). Additionally, representation of PRD and Corr in a fuzzy similarity evaluation offers interpretability to the proposal. The proposal can be applied to the classification of quasi-periodic biomedical signals. Moreover, it can be extended for the classification of other cardiac and respiratory diseases.

Capnogram Feature Extraction Based on Wavelet Decomposition by Segments for Classification of Asthma Severity

To allow assessment of asthma severity, a feature extraction method based on wavelet decomposition by segments is proposed. The proposal is tested in 23

capnograms signals and the best classification result suggest 97.39 % accuracy when using as features the wavelet coefficients of segment A-B and segment E-F of the capnogram. The proposed method is able to extract appropriate features from the capnogram for accomplishing the capnogram classification in 6 degrees of asthma with low computational cost. Moreover, the proposed feature extraction method is used as part of the decision support system for asthma classification.

Decision Support System for Assessing Asthma Severity

A decision support system for asthma classification based on the capnogram analysis is proposed. Two architectures are evaluated on 23 capnograms from an asthma camp in Cuba. Architecture 1 combines the segmented wavelet transformation method for feature extraction with a fuzzy classifier. Architecture 2 adds a step of similarity evaluation after the extraction of the wavelet coefficients. Both architectures show similar classification results in terms of correct rate. Execution time represents the most significant difference between architectures and could define the kind of application where to be employed.

5.2 Future Perspectives

Automatic Diagnosis of Asthma Severity

The automatic diagnosis of asthma severity is the most evident application that is foreseen may use the proposals on the thesis. Since 2010, in the Central Institute of Digital Research in Cuba, a capnograph was designed and produced for the National Health System and exportation to countries in Latin America area. The real application

of the proposals is to be embedded on this low computational resources device. The expected outcome is to support clinical experts in the classification of asthma in the different clinical interventions.

Classification of Respiratory and Cardiac Diseases

There are other cardiac and respiratory diseases that also can be identified by analyzing the capnogram and the ECG signals, e.g., Brugada syndrome, AV block, and hypoventilation. The proposals are expected to be extended to achieve this goal.

Multimodal Personal Identification System

A multimodal personal identification system integrates biometrics, such as fingerprint, iris, dental, palm, biosignals, and others, to identify an individual. There is one research proposal in 2012 about the development of a system called SMaRTIS: Profile-Tracking Based Identification System Using Encoded Multimodal Biometric Data. This research proposal is a collaboration work between King Saud University/Salman Bin Abdulaziz University and Tokyo Institute of Technology. The research done in this thesis can contribute to the development of the module of personal identification based biosignals.

Redesign of Electronic Biomedical Management Kits (BMK)

BMK assist in providing day-to-day support for the patients suffering intractable diseases such as asthma, diabetes, or epilepsy in Saudi Arabia. BMKs are easily affected by minor factors such as weather and environment usage being susceptible to random failure or damage, thereby putting the patients at further risk of fatal attacks. The research collaboration between King Saud University/Salman Bin Abdulaziz University and Tokyo Institute of Technology in 2014 aims to redesign BMKs to better suit the

environmental conditions in the Kingdom while contributes towards the development of the Kingdom's industrialization in the area of producing electronic devices. The proposals in this thesis can contribute to improve the performance of the kits especially for asthma disease.

Bibliography

[1] From the Global Strategy for Asthma Management and Prevention, Global Initiative for Asthma (GINA) 2012. Available from: http://www.ginasthma.org/.

[2] B. D. Woods, R. N. Sladen, "Perioperative Considerations for the Patient with Asthma and Bronchospasm", British Journal of Anaesthesia, Vol. 103, Suppl 1, doi:10.1093/bja/aep271, 2009.

[3] T. De La Vega ; V. Perez,; A. Alerm, and L. Lecha,. "El asma bronquial y su asociación con los cambios de tiempo". Rev Cubana Med Gen Integr [online]., Vol.26, No. 4 , pp. 665-672, 2010. (in Spanish). (Available in: <http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21252010000400009&lng=es&nrm=iso>. ISSN 1561-3038).

[4] M. Brown, A. Prieto, and B. Cuéllar. "Crisis de asma y enfermedades respiratorias agudas: Contaminantes atmosféricos y variables meteorológicas en Centro Habana". Rev Cubana Med Gen Integr, Vol. 17, No. 1, 2001. Available in: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21252001000100002&lng=es.

[5] A. Maton et al., Human Biology and Health. Englewood Cliffs: Prentice Hall. pp. 108–118. ISBN 0134234359, 2010

[6] N.A.R. Nik Hisamuddin, A. Rashidi, K. S. Chew, J. Kamaruddin, Z. Idzwan, A. H. Teo, "Correlations between capnographic waveforms and peak flow meter measurement in emergency department management of asthma", International Journal of Emergency Medicine, Vol. 2, No. 2, pp. 83-89, 2009.

[7] M. Yaron, P. Padyk, M. Hutsinpiller, Ch. B Cairns, "Utility of the Expiratory Capnogram in the Assessment of Bronchospasm", Annals of Emergency Medicine, Vol. 28, No. 4, pp. 403-407, 1996.

[8] E.G. Meraz, H. Nazeran, C.D. Ramos, P. Nava, B.Diong, M.D. Goldman, "Analysis of Impulse Oscillometry Measures of Lung Function and Respiratory System Model Parameters in Small Airway-impaired and Healthy Children Over a 2-year Period", Biomedical Engineering OnLine, http://www.biomedical-engineering-online.com/content/10/1/212011, time: August, 2013.

[9] G. QI, et al., "Detection of the Airway Obstruction Stage in Asthma Using Impulse Oscillometry System," Journal of Asthma, Vol. 50, No.1, pp. 45–51, 2013.

[10] Y. H. Shin et al., "Use of Impulse Oscillometry System in Assessment of Asthma Severity for Preschool Children" , Journal of Asthma, 50(2): 198–203,2013

[11] N.A.R. Nik Hisamuddin, A. Rashidi , K.S. Chew, J. Kamaruddin, Z. Idzwan, T. A. Howe, "Correlations Between Capnografic Waveforms and Peak Flow Meter Measurement in Emergency Department Management Asthma ", International Journal of Emergency Medicine, Vol. 2, No. 2, pp. 83-89, 2009.

[12] Z. Nikolova-Todorova, "Clinical Applications of Capnography", Signa Vitae, Vol. 3, Suppl 1, pp. 44 – 45, 2008.

[13] J. D'Mello and M. Butani, "Capnography", Indian J. Anaesth., Vol. 46, No. 4, pp. 269– 278, 2002.

[14] S.Yamamori et al., Capnometer. United States Patent 6267928, julio 31 2001.

[15] A. E. Portela, A. Milanes, and J. Folgueras, "Desing of a transducer for Mainstream capnograph (Diseno de un transductor para capnografo Mainstream)," IFMBE Proc. of the VIII Latin American Congress of Biomedical Engineering, Havana, Cuba, T145, 2009 (in Spanish). ISBN: 978-959-212-531-5

[16] National Heart, Lung, and Blood Institute "Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma", National Asthma Education and

Prevention Program, pp.36 – 60, 2007.

[17] B.You, R. Peslin, C. Duvivier, V.Dang Vu, JP Grilliat,"Expiratory Capnography in Asthma", European Respiratory Journal, Vol. 7, No. 2, pp. 318-323, 1994.

[18] T.T. Kean, T.A. Howe, M.B. Malarvili, "Feature Extraction of Capnogram for Asthmatic Patient", Second International Conference on Computer Engineering and Applications, 2010.

[19] T.A. Howe et al., "The Use of End-Tidal Capnography to Monitor Non-intubated Patients Presenting with Acute Exacerbation of Asthma in the Emergency Department", The Journal of Emergency Medicine, Vol.41, No. 6, 581-589, 2011.

[20] J. G. Webster and J. W. Clark, "Medical instrumentation: application and design," Wiley, 1995.

[21] P. de Chazal and R. B. Reilly, "A comparison of the ECG classification performance of different feature sets," Proc. of Computer in Cardiology 2000 Conf., Boston, USA, Vol.37, pp. 327-330, 2000.

[22] S. Kara and M. Okandan, "Atrial fibrillation classification with artificial neural networks," Pattern Recognition, Vol.40, No.11, pp. 2967-2973, 2007.

[23] R. Couceiro et al., "Detection of Atrial Fibrillation Using Model based ECG Analysis," Int. Conf. in Pattern Recognition, pp. 1-5, Tampa, FL, Dec. 2008.

[24] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder, "Automatic real time detection of atrial fibrillation," Annals of Biomediacl Engineering, Vol.37, No.9, pp. 1701-1709, September 2009.

[25] K. Tateno and L. Glass, "Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and ΔRR intervals," Medical and Biological Engineering and Computing, Vol.39, pp. 664-671, 2001.

[26] K. Tateno and L. Glass, "A method for detection of atrial fibrillation using RR intervals," Proc. of Computers in Cardiology Conf., Boston, USA, Vol.27, pp. 391-394, 2000.

[27] S. Parvaresh and A. Ayatollahi, "Automatic atrial fibrillation detection using autoregresive modeling," Int. Conf. on Biomedical Engineering and Technology, Vol.11, pp. 105-108, Singapore, 2011.

[28] M. Kazemi and M. B. Malarvili, "Analysis of capnogram using linear predictive coding (LPC) to differentiate asthmatic conditions," J. Tissue. Sci. Eng., Vol.2, No.5, doi:10.4172/2157-7552.1000111, 2011.

[29] B. Boucheham, "Matching of quasi-periodic time series patterns by exchange of block-sorting signatures," Pattern Recognition Letters, Vol.29, No.4, pp. 501-514, 2008.

[30] G. B. Moody and R. G. Mark, "A new method for detecting atrial fibrillation using R-R intervals," Computers in Cardiology, No.10, pp. 227-230, 1983.

[31] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," Circulation, Vol.101, No.23, pp. e215-e220, 2000.

[32] J. Pomares, Ch. Fatichah, M.L. Tangel, Fei Yan, A.J. Garcia, Fang-Yan Dong, K. Hirota, "Similarity-Based Fuzzy Classification of ECG and Capnogram Signals", Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.17 No.2, pp. 302-310, 2013.

[33] M. Vetterli, C. Herley, "Wavelets and Filter Banks: Theory and Design", IEEE Transactions on Signal Processing, Vol. 40, No. 9, pp. 2207-2232, 1992.

[34] M. Sifuzzaman, M.R. Islam, and M.Z. Ali, "Application of Wavelet Transform and its Advantages Compared to Fourier Transform", Journal of Physical Sciences, Vol. 13, 121-134, 2009.

[35] I. Daubechies, "Ten Lectures on Wavelets", CBMS-NSF Regional Conference Series in Applied Mathematics, pp. 1-16, 1992.

[36] G. Icke , S.V Kamarthi, "Feature extraction trough discrete wavelet transform coefficients", Proceedings of SPIE –The International Society for Optical

Engineering, Vol. 5999, 599903-1, 2005.

[37] A. Adib, M.A.Haque, "ECG beat classification using discrete wavelet coefficients", 5th International Symposium on Health Informatics and Bioinformatics, 2010.

[38] Hui Zhang, Tu Bao Ho, Mao-Song Lin, Xuefeng Liang, "Feature Extraction for Time Series Classification Using Discriminating Wavelet Coefficients", Lecture Notes in Computer Science, Vol. 3971, pp. 1394-1399, 2006.

[39] Li Deqiang, W. Pedrycz, N.J, Pizzi, "Fuzzy Wavelet Packet Based Feature Extraction Method and Its Application to Biomedical Signal Classification", IEEE Transactions on Biomedical Engineering, Vol. 52, No. 6, pp. 1132-1139, 2005.

[40] L. J. Hoeksema et al., "Accuracy of a computerized clinical decision-support system for asthma assessment and management" Journal of the American Medical Informatic Association, Vol. 18, No. 3, pp. 243-250. doi:10.1136/amiajnl-2010-000063, 2011.

[41] E. A. Lomotan, L. J. Hoeksema, D. E. Edmonds, G. Ramírez-Garnica, R. N. Shiffman, and L. I. Horwitz, "Evaluating the use of a computerized clinical decision support system for asthma by pediatric pulmonologists", Journal of the American Medical Informatic Association, Vol. 81, No. 3, pp.157-165, doi: 10.1016/j.ijmedinf.2011.11.004, 2011.

[42] A. Patel, J. Choubey, Sh. K. Gupta, M. K. Verma, R. Prasad, and Q. Rahman, "Decision Support System for the Diagnosis of Asthma Severity Using Fuzzy Logic", Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol I, Hong Kong, 2012.

[43] M.H. Fazel Zarandi, M. Zolnoori, M. Moin, and H. Heidarnejad, "A Fuzzy Rule-Based Expert System for Diagnosing Asthma", Transaction E: Industrial Engineering, Vol. 17, No. 2, pp. 129-142, 2010.

[44] C. Chakraborty, T. Mitra, A. Mukherjee, and A. Kumar Ray. "CAIDSA: Computer-aided intelligent diagnostic system for bronchial asthma", Expert System.

Applications, Vol. 36, No.3, pp. 4958-4966, 2009.

[45] E. A. Chatzimichail, A.G. Rigas, E.N. Paraskakis, "An Artificial Intelligence Technique for the Prediction of Persistent Asthma in Children", IEEE International Conference on Information Technology and Applications in Biomedicine, pp. 1-4, 2010.

[46] D. O'Sullivan et al., "Using Secondary Knowledge to Support Decision Tree Classification of RetrospectivevClinical Data" Third International Workshop on Mining Complex Data, pp. 37 – 48, 2007.

[47] J. Pomares, M. L. Tangel, Fei Yan, M. Otaño, A. E. Portela, Fangyan Dong, and K. Hirota, "Capnogram Feature Extraction Based on Wavelet Decomposition by Segments for Classification of Asthma Severity ", The Third International Workshop on Advanced Computational Intelligence and Intelligent Informatics, Shanghai, 2013.

[48] L.X. Wang, "The WM Method Completed: A Flexible Fuzzy System Approach to Data Mining", IEEE Transactions on Fuzzy Systems, Vol. 11, No. 6, pp. 768 – 782, 2003.

[49] L.X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 22, No. 6, pp. 1414-1427, 1992.

Related Publications

Journal Papers

[J3] **Janet Pomares B.,** Martin Leonard Tangel, Fei Yan, Meiyo Tamaoka, Yasunari Miyazaki, Fangyan Dong, and Kaoru Hirota, "Decision Support System for Assessing Asthma Severity based on Capnogram Analysis", *Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII)* (In preparation).

[J2] **Janet Pomares B.**, Martin Leonard Tangel, Fei Yan, Marianela Otaño, Alejandro E. Portela, Fangyan Dong, and Kaoru Hirota, "Segmented Wavelet Decomposition for Capnogram Feature Extraction", *Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII)*, Vol. 18, No. 4, 2014. (Accepted).

[J1] **Janet Pomares B.**, Chastine Fatichah, Martin Leonard Tangel, Fei Yan, Jesus Adrian Garcia Sanchez, Fangyan Dong, and Kaoru Hirota, "Similarity-based Fuzzy Classification for ECG and Capnogram Signals", *Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII)*, Vol. 17, No. 2, 2013.

Conference Papers

[C2] **Janet Pomares B.**, Martin Leonard Tangel, Fei Yan, Marianela Otaño, Alejandro E. Portela, Fangyan Dong, and Kaoru Hirota, "Capnogram Feature Extraction Based on Wavelet Decomposition by Segments for Classification of Asthma Severity ", *Proceedings of the Third International Workshop on Advanced Computational Intelligence and Intelligent Informatics*, ISSN2185-758X, Shanghai, 2013.

[C1] **Janet Pomares B.**, Chastine Fatichah, Martin Leonard Tangel, Fei Yan, Jesus Adrian Garcia Sanchez, Fangyan Dong, and Kaoru Hirota, "Similarity-based Fuzzy Inference System for ECG and Capnogram Signals Classification", *Proceedings of the International Symposium of Soft Computing*, Tokyo, 2012.