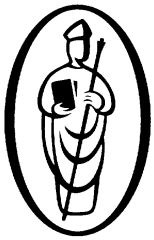
Signal Analysis of Out-of-Hospital Cardiac Arrest Electrocardiograms for Decision Support During Cardiopulmonary Resuscitation

by

Trygve Eftestøl

Dissertation submitted in partial fulfillment
of the requirements for the degree of
DOKTOR INGENIØR



Electrical and Computer Engineering
Stavanger University College
Norway

2000

Abstract

This thesis focuses on signal analysis of electrocardiograms (ECG) from out-of-hospital cardiac arrested patients. The application of such methods may eventually contribute in guiding therapy towards improved survival rates which in general are dismally low, but varies among different ambulance systems depending on time from arrest to first electrical defibrillation given to the patient.

One of the possible reasons for this is that a large part of the valuable therapy time is wasted in futile attempts to restart the heart by electrical defibrillations. Using this time to provide precordial compressions and ventilations to establish and keep up an artificial supply of oxygenated blood would serve the patient better. It would improve the heart condition and thus increase the chance of successful defibrillation outcome.

By predicting defibrillation outcome, the ratio of failured defibrillations can be decreased. The ability to monitor therapeutic efficacy can help the rescuer to optimise performance. A decision support system involving ECG signal analysis to extract descriptive features has been investigated for these purposes in earlier work. We propose to use a pattern recognition framework for the decision support system. In contrast to most earlier work with one-dimensional features, this allows analysis of multivariate information. In our experiments we demonstrate that performance both in outcome prediction and monitoring is better when the analysis is based on combined rather than one-dimensional features. We also propose and experimentally verify methods firstly to control performance and secondly to ensure that performance results are reliable.

Another problem also has to be resolved. The precordial compressions and ventilations cause artefacts in the ECG so that treatment has to be stopped during analysis. We propose using adaptive filters for removing such artefacts. These filters use reference signals providing information correlated to components in the artefacts. In one of our experiments we mix human ECG with artefacts from animal ECG and show that the adaptive filter is successful in restoring the original human ECG signal.

Preface

This dissertation is submitted in partial fulfillment of the requirements for the degree of *doktor ingeniør* at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway. Associate professor Sven Ole Aase and professor John Håkon Husøy of Stavanger University College (HiS), Stavanger, Norway, have been the supervisors.

The work, including compulsory courses corresponding to one year full time studies, has taken place in the years from 1996 to 2000 at the Electrical and Computer Engineering Department at HiS and at the Research and Development Department at Laerdal Medical AS, Stavanger, Norway. In 1998 the author spent four months at the Research and Development Department, Cardiology Products Division, Medical Products Group, Hewlett-Packard Company, Andover, Massachusetts, USA. Funding for the work has been provided through a scholarship from the Norwegian Research Council (grant number 111491/320) and by financial support from Laerdal Medical AS.

The work described in the main chapters in this book involving the analysis of human data has been conducted in cooperation with medical professionals from Ullevaal University Hospital, Department of Anesthesiology and Institute of Experimental Medical Research. This work has been or is in the process of being published in international journals [100, 38, 39, 1, 63] and as abstracts from national and international conference presentations [99, 101, 40, 102]. Furthermore, articles have been published in national and international conference proceedings [34, 36, 35, 37, 41].

Acknowledgments

I would like to thank my main supervisor, associate professor Sven Ole Aase for giving invaluable support sharing his insight and providing helpful ideas and criticism on pattern recognition and signal processing details. Likewise I would like to express my deep gratitude to my other supervisor and project coordinator at HiS, professor John Håkon Husøy for sharing ideas and giving constructive criticism.

I would also like to provide thanks to Helge Myklebust and Helge Fossan at the Research and Development Department at Laerdal Medical for involving me in their deeply inspiring projects. For the cooperation and guidance on medical issues I am deeply grateful for the help from Dr. Kjetil Sunde, Audun Langhelle and professor Petter Andreas Steen at Ullevaal University Hospital.

Additional thanks goes to Al Languth, George Diller, Dr. Sophia Zhou and Cheuw Lee Yeoh at HP, for making my stay in USA possible and interesting.

I also wish to express thanks to professor Fritz Sterz at Universitatsklinik fur Notfallmedizin Allgemeines Krankenhaus der Stadt Wien for sharing data and advice on the analysis of the long-term VF data. Likewise, a sincere thank you to professor Hans Ulrich Strohmenger, Department of Anesthesiology Critical Care and Emergency Medicine, The Leopold-Franzens-University Innsbruck for advice and support throughout the years and for providing the short-time VF data.

It was a great inspiration for me to supervise the students, Eli Vatland and Bjørn Terje Holten.

I would also like to thank the rest of my colleagues at both Stavanger University College and Laerdal Medical.

I am grateful to my parents for supporting and encouraging me both in early life and in education.

Last but not least I would like to thank my wife, Ingrid, and children, Kristina and Eivind, for love and support.

Contents

Chapter 1

Introduction

The quest to reverse sudden death goes back to biblical times. The first account of *artificial respiration* appeared in 1744, while the first documented report of *artificial circulation* by chest compressions is dated 1891. Various techniques for resuscitation based on either of these principles were used until recent times. Supporting ventilation to the lungs (pulmo) and circulation to the heart (cardio) are thus the key elements of resuscitation, and in the period between 1958 and 1961 they were recognized as parts of a whole and complete approach to resuscitation. *Cardiopulmonary resuscitation* (CPR) as we know it was thus established.

The use of electricity to terminate fatal rhythm in an arrested animal heart by *defibrillation* appeared as a footnote in a paper in 1899. The first human patient was successfully defibrillated in 1947. This was followed by developments in the technique to make the defibrillators portable. By the early 1970s defibrillators and CPR were established as key elements in prehospital care. Programs to train emergency medical technicians (EMTs) in the use of defibrillators began to appear in 1980. Since then advances in technology has allowed for automated external defibrillators (AEDs). The AEDs advises the EMT to apply a defibrillators shock if it recognizes the rhythm appearant in the patient's electrocardiogram (ECG) to correspond to cardiac arrest. This allowed shorter training time for the EMTs as they do not have to interpret the cardiac rhythm. Since then we have seen a gradual advancement in technology to make AEDs smaller and easier to use which leads to a more wide-spread use of AEDs.

In the conclusion of [42] it is remarked that

> The last 20 years has not seen a fundamental breakthrough. CPR is fundamentally unchanged since it was first invented. Defibrillation is also unchanged, although there are now sophisticated machines capable of automatically interpreting the rhythm and guiding the operator through the resuscitation. The pharmacologic management of cardiac arrest has not fundamentally changed in decades, and emergency medical services are not fundamentally different in the 1990s compared to the 1970s. The quest to reverse sudden death has not stopped in the last 2 decades, but it has also not led to a new scientific or historical breakthrough.

In the light of these remarks it is worth noting that the manufacturers of AEDs focus on defibrillation alone. In the development of the defibrillators a lot of attention is directed towards the shape of the defibrillation waveform, size of equipment and new user groups. CPR is provided according to a predefined protocol. "Early defibrillation" is the key element in resuscitation. A recent study [24] indicates the importance of reestablishing artificial circulation and ventilation as there had been no improvement of overall survival rates in the community after the introduction of AEDs in the emergency systems. Focus in parts of cardiac arrest research is shifting towards the possibility to improve outcomes by individualising therapy for each patient. The characteristics of the cardiac arrest rhythm, *ventricular fibrillation* (VF)is pointed out as being an important factor in this respect. The challenge of VF analysis is the core problem discussed in this thesis.

In the remaining part of this introductory chapter we will discuss the prehospital scenario, identify the key problems, discuss the salient features of pattern recognition, give an outline of the contents of this thesis and point out the major contributions of this work.

1.1 Prehospital scenario

Cardiac arrest is a common cause of sudden death. An incident may be witnessed by a bystander. One of the crucial factors for survival is whether this bystander, the *lay rescuer*, knows the procedures of basic life support. Absence of consciousness and signs of life is verified and chest compressions in combination with mouth-to-mouth ventilation is provided until an ambulance arrives with a *professional rescue team*. They assess the situation and provide advanced life support by clearing airways and providing defibrillation, chest compressions, ventilation and drug therapy. Finally, the patient is brought to

the intensive care unit of a nearby hospital. If the outcome of the incident is happy the patient is admitted to the hospital and gets out alive with normal life functioning.

1.2 Key problems and earlier research

The description of this scene is simplified. There are many variable factors: Previous heart disease, presence of witnesses, bystander basic life support, response time of the ambulance system etc [5]. Even if the heart is restarted, the brain might be damaged.

When the professional rescuers provide life support in a cardiac arrest setting, denoted *cardiopulmonary resuscitation* (CPR), they follow a standardised protocol [81], which varies in different countries all over the world. Three successive electroshocks (*defibrillations*) are given if the heart rhythm is identified as VF which is electrically treatable. Further, they provide repeated sequences of chest compressions. At the end of each sequence they attempt to defibrillate the heart. This procedure is interrupted if *return of spontaneous circulation* (ROSC) is verified.

The probability of survival from cardiac arrest is quite low, commonly between 10 and 20 percent [51]. This work studies cardiac arrest episodes from the ambulance system in Oslo and data from animal experiments. The objective is the designing a VF analysis decision support tool as an add-on to the AED in the prehospital setting. Eventually such a system may contribute in improving the number of survivals from out-of-hospital cardiac arrests.
In essence:

- Today the protocol decides how to treat the patient.

We want to focus on the potential for increased survival in the alternative formulation:

- Let knowledge of the physiologic state of the individual patient obtained through the analysis of ECG measurements determine the treatment.

Thus, the treatment procedure should be decided through a continuous evaluation of diagnostic information provided through physical examinations of the patient, the effect of treatment and most important, *electrocardiograms*

(ECGs). These recordings of the electric activity in the heart, provides information on the *myocardial metabolism* [1] These physiological parameters constitute the "patient's voice". They provide the rescuers with the information needed to decide how to proceed with the treatment in order to have the best chance of a successful outcome.

In this work we will focus on the following questions:

- Is it possible to decide the optimal timing of defibrillation?
- Is it possible to monitor the effect and quality of chest compressions and ventilations?

We attempt to answer these questions through a thorough and comprehensive investigation of the basic constituents of our proposed decision support system. We base our analysis on some of the methods established through medical research discussed in the last two paragraphs of this section. The experiments are performed on our own data which represent a realistic prehospital setting. Through careful design of our decision support system we show how improvement in performance can be achieved. Based on the findings in our study we propose methods for combining ECG measurements for prediction and monitoring and a scheme for removing CPR artefacts from the ECG.

VF analysis in medical research on out-of-hospital cardiac arrest goes back to 1985 and has been evolving since then mostly through animal experiments and retrospective studies of prehospital cardiac arrest ECGs. The following paragraph gives a representative, but not exhaustive list of previously reported research on VF analysis.

In 1985 the *amplitude of ventricular fibrillation* was reported to be indicative of defibrillation outcome after cardiac arrest [109]. Further on, the *median frequency's* potential for guiding therapeutic interventions during CPR was discussed in [33]. Later work [11, 12, 68, 9, 92, 14, 62, 72, 97, 98, 96, 95, 10, 94, 70, 73] has focused on variations of these two measurements when applying ECG measurement techniques to study CPR related problems. Promising results, especially in animal experimental studies, have been reported. It has been pointed out that further studies need to be done to establish these or related ECG measurements as diagnostic tools. We give a more detailed presentation of medical background, main problems and previous research in chapter 2.

[1] The myocardial metabolism are the physical and chemical processes, making energy available to the myocardium (muscular tissue of the heart) [30]. Efficient CPR improves the myocardial metabolism as it increases the myocardial blood flow.

1.3 A decision support system for CPR guidance

The work referred to in the previous section has been done from a medical viewpoint. Most of the previous work, with some exceptions [70, 73] have used univariate measures. We think this is a limitation, and will further investigate the use of multivariate measures. As we will discuss in chapter 3 the analysis methods applied to univariate measures have shortcomings as to the analysis of multivariate measures. We will develop an analysis system that handles these problems. We introduce a *decision support system* using a *pattern recognition* framework for this purpose. In this framework signal analysis methods are used to structure the data and to extract *features* carrying diagnostic information. Methods from classification theory is used to establish *decision regions* pertaining to different diagnostic outcomes. The understanding of the medical aspects is of crucial importance for the success in applying these methods. Therefore we give a more thorough presentation of the medical aspects of this work in chapter 2.1 and 2.2 so that we can integrate our decision support system into the medical research tradition on VF analysis of out-of-hospital cardiac arrest ECG(chapter 3).

1.4 Thesis outline

A medical discussion of the mechanisms and treatment effects of cardiac arrest are presented along with a medical formulation of the key questions addressed in this work. Following this, we give a brief presentation of the key tools in the decision support system framework: signal analysis and classification methods. The questions are reformulated in engineering terms. After this follows a presentation and discussion of the data material used before the novel methods actual to the particular problem addressed here are proposed and discussed in a pattern recognition and signal analysis theoretic framework. Following this, experimental results are presented and discussed. Finally we will summarise and conclude the work.

Major contributions of this work The material in the experimental part of this work has been or is in the process of being published in international journals [100, 38, 39, 1, 63] and as abstracts from national and international conference presentations [99, 101, 40, 102]. Further, articles has been published in national and international conference proceedings [34, 36, 35, 37, 41].

- A method is proposed for integrating the decision support system with classification and signal analytic methods in the medical research field of out-of-hospital cardiac arrest.
- A database of ECGs and demographics of cardiac arrest patients have been established in cooperation with an acute medical research group in Oslo [100].
- Data analysis of ECG rhythm annotations and Utstein style demographics [17] have been applied to study the mechanisms of resuscitation in practice [100].
- A method for removing CPR artefacts from human ECG by using an adaptive filtering technique is proposed [63, 1].
- A method for predicting defibrillation outcome by combining several ECG derived measurements applying classification methods is proposed. The classifier has been adapted to the problem at hand so that the user can control the ratio of successful outcomes to be correctly predicted. Furthermore we ensure reliability in the performance results. This is achieved by tuning the classifier so that it performs similarly on an independent data test set and on the data set used to train the classifier [38].
- A general, feature independent method, is proposed for monitoring the *Probability of defibrillation success* in a human cardiac arrest patient [38].

The following major parts constitute the thesis:

Introductory part

Chapter 1 provided an introduction to the problem of treating cardiac arrest to a successful outcome. Previous work has shown that VF-analysis methods has a potential for providing useful therapy guidance. In this work, we will focus on the classification, signal analytic and CPR artefact removal aspects of VF analysis in the design of a decision support system for guidance of CPR therapy.

Theoretic backround part

Chapter 2 gives the background for the medical aspects of CPR, ECG and VF analysis as was sketched out in section 1.2in this chapter. The problems focused upon in this thesis are formulated. A review of previous work on VF analysis is given.

Chapter 3 Formulates a VF analysis decision support system for guiding therapy. The key elements of this system are described, including the classification and feature extraction techniques and methods for CPR artefact removal.

Methodological and experimental part

Chapter 4 discusses the data material used in the experimental part of this work including both human and animal cardiac arrest data.

Chapter 5 illustrates the main aspects of this work. We demonstrate the use of artefact filtering methods and further how ECG features can be used to predict outcome and monitor the effect of therapy to animals in cardiac arrest.

Chapter 6 provides a detailed analysis of important time intervals in the treatment of out-of-hospital cardiac arrest patients.

Chapter 7 shows how adaptive filtering techniques can be used to reduce CPR artefacts in human ECG.

Chapter 8 demonstrates how the techniques described in chapter 3 can be used to predict the outcome of defibrillation attempts in a retrospective analysis of human cardiac arrest ECG.

Chapter 9 describes and demonstrates a method for monitoring the probability of defibrillation success.

Conclusive part

Chapter 10 summarises the major contributions and conclusions from this work and finally suggests problems for further research.

Chapter 2

Background

In this chapter we first describe some basic medical and engineering nomenclature regarding the interaction of the cardiovascular system, ECG recordings and the procedures and mechanisms involved in providing CPR to out-of hospital cardiac arrested patients.

We continue by identifying the main issues of VF analysis. The objectives, methods and possible benefits to CPR therapy are sketched out. VF analysis is broken down into the problems of predicting defibrillation outcome, monitoring CPR efficacy and removal of CPR artefacts.

Finally we review the earlier work on these issues.

2.1 CPR and ECG

ECG recordings play an important role in assessing the patient's status during CPR. The circulatory, or *cardiovascular*, system is vital to the upkeeping of life as it serves the exchange of oxygen and carbondioxide to body cells. The *heart* is the most important part of this system, serving as a pump distributing and collecting the blood throughout the body. During cardiac arrest, the heart stops pumping. Death will occur in a matter of minutes if no treatment is provided to reverse this process and make the heart start pumping again.

The rescuer has to make important decision on how to proceed during the treatment of cardiac arrest. Defibrillation of the heart is the ultimative intervention [57], but it is also the most drastic as it is harmful to the heart [115]. Other less dramatic elements in therapy are precordial chest compressions [79] and ventilations [53] to maintain artificial circulation. Such intervention can

slow down the degradation in heart condition [57] which ultimatively results in death. It is also speculated that the degradation can be reversed [24]. These techniques, in some cases with medication for added stimulation [75], does not carry the potential of effecting *return of spontaneous circulation* (ROSC) without defibrillation. The choice of therapy depends on several factors.

- *Cardiac aetiology*: Prior heart diseases like infarction or arrhythmias affect the heart's condition unfavorably to resuscitation.
- *Duration of cardiac arrest*: As time goes, the energy resources of the heart are gradually emptied making successful resuscitation less probable.
- *Bystander CPR*: In some cases a witness plays the role of *lay rescuer* by performing basic CPR of precordial compressions and ventilations to the patient. This may have a positive effect depending on the quality of the treatment.

All these factors affect the heart condition, or myocardial metabolism, during cardiac arrest. If the level of myocardial metabolism is to low, defibrillation is futile and harmful. It may be improved through establishing a certain amount of *myocardial perfusion*[1] through adequately performed CPR. If this level is high enough, resuscitation through defibrillation may be successful. Therefore the key issues for the rescuer to resolve is whether to defibrillate or provide CPR. Of course, the rescuer also has to check the patient's pulse. If there is no pulse, treatment is continued, otherwise stabilising treatment is provided.

It is evident that information of the level of myocardial metabolism would be an extremely useful guidance for the rescuer in deciding whether defibrillating or giving CPR will be the optimal treatment for ensuring a successful outcome for the patient [24].

In this respect we seek some physiological measurements which may provide valid information on the myocardial metabolism or degree of resuscitability. Examples of such are measurements of *myocardial bloodflow* (MBF) [2], *coronary perfusion pressure* (CPP)[3] and *end-tidal* CO_2 ($ETCO_2$)[4]. It has been experienced [60] that CPP correlates well with MBF during CPR and also with the

[1] Supplying of nutrients and oxygen to the middle layer of the walls of the heart, composed of cardiac muscle.

[2] The blood supplying the muscular tissue of the heart with oxygen.

[3] The difference in pressure between the aortic and right atrial chamber necessary to forward MBF

[4] The amount of carbon dioxide in the patient's expiratory gas. $ETCO_2$ reflects the amount of blood flow to the lungs and thus indirectly reflects cardiac output.

outcome of resuscitation. CPP has been shown to correlate with ROSC and survival in both animal [58] and human studies [76], but the measurement requires placement of arterial and right atrial or central venous catheters, which are hardly ever practical in the cardiac arrest situation. $ETCO_2$ varies with cardiac output [112, 48, 54, 19, 61], but it is also influenced by the general metabolic status [112, 48, 54, 19, 61], ventilation [54], the use of epinephrine [19] and time [61], and should therefore be used with caution when evaluating myocardial perfusion and resuscitability during CPR. Therefore, other means of assessing the myocardial metabolism should be sought.

2.1.1 Electrical activity in the myocardium - ECG measurements

Recordings of MBF and CPP involve flow and pressure measurements. Recording of ECGs is a means for non-invasive measurement of the cardiac activity. The pumping activity of the heart is regulated by electrical stimulus of the cardiac cells. A cell respond to such stimuli by contracting and further stimulating of neighboring cardiac cells. The electrical activity propagates along conduction pathways in the myocardium. Electrodes placed on the patient's chest measure changes in body potential due to this electrical activity. The salient feature of ECG recordings is that each rhythm has its characteristic waveform.

Heart rhythms deviating from sinus rhythm (SR),the rhythm characteristic of a normally operating heart, are denoted arrhythmias. There exists a lot of different arrhythmias. Figure 2.1 shows some examples of the typical rhythms which are experienced by an out-of-hospital defibrillator. The ECG tracings in part a) ,b) and c) in the figure corresponds to pulse rhythms, while the ones in part d),e) and f) are examples of rhythms without pulse. For the rescuer operating the defibrillator in an out-of-hospital cardiac arrest situation, the pulse rhythms of a), b) and c) corresponds to no CPR or defibrillation to be given (no-treat rhythms).

The rhythms without pulse in part d),e) and f) in the figure should be treated with CPR and/or defibrillation. The *ventricular tachycardia* (VT) in d) should be defibrillated immediately. VT is driven by an impulse generation originating from an impulse carousel (reentry). This causes one impulse to stimulate the heart successively. VT is often the initiating rhythm of cardiac arrest. The chance of successful resuscitation is very good.

In most out-of-hospital situations, however, the patients are past the initial phase of VT and well into VF (part e) in the figure). The activity in the

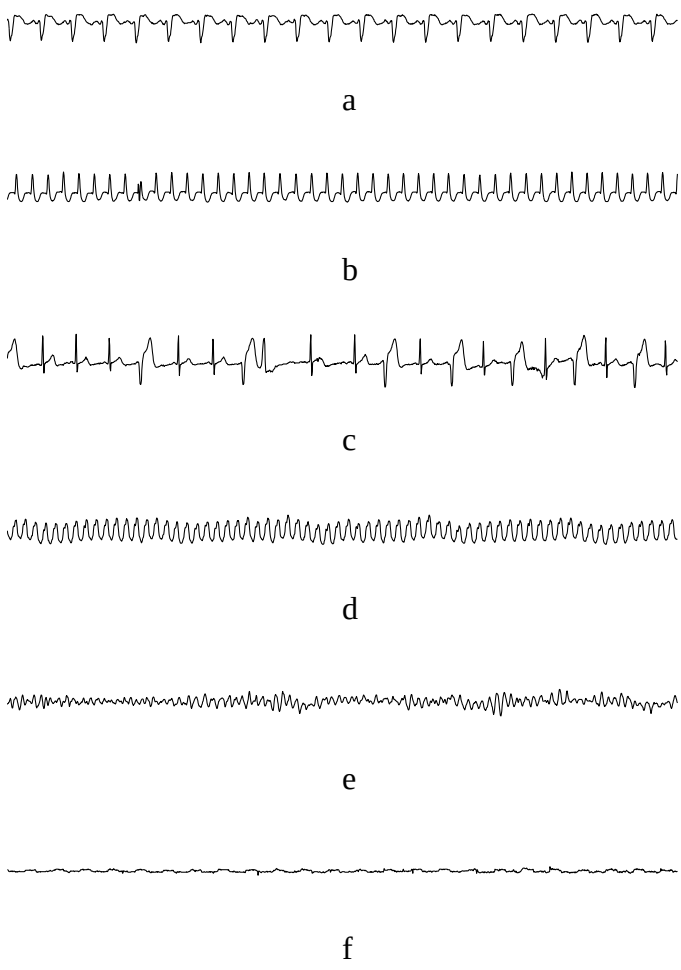


Figure 2.1: Example ECG tracings corresponding to rhythms typically recorded by defibrillators. a) sinus rhythm (SR), b) supraventricular tachycardia (SVT), c) premature ventricular contraction (PVC), d) ventricular tachycardia (VT), e) ventricular fibrillation (VF) and f) asystole

myocardium gradually becomes less organised with no cardiac output. VF is treated by defibrillation, but the probability of defibrillation success diminishes as time goes and the rhythm gradually changes [57].

In the final stage preceding death, the electrical activity has nearly disappeared. This is termed asystole, and is seen as a near flat line in the ECG tracing (part f) in the figure). The rescue team has to provide CPR with medications, heart compressions and ventilations to increase circulation of blood and oxygen to the ischemic[5] myocardium. This might increase the perfusion to the heart, resulting in a conversion to VF which can be defibrillated [67]. *Electromechanical dissociation* (EMD) denotes another pulseless rhythm where there is electrical activity in the myocardium without any corresponding mechanical activity. EMD often appears in the transition from VF to asystole and is treated with medications, compressions and ventilations to obtain VF. However, EMD can also have a other causes which can be specifically treated. Both EMD and asystole are no-treat rhythms.

The defibrillators do not give any distinction between the no-treat rhythms with and without pulse. The rescuer has to check for signs of life (a time-consuming operation), and determine whether CPR should be given or not. VF and VT are defibrillated and thus termed as *treat rhythms*.

In terms of decisions to be made, the defibrillators handle the treat/no-treat distinction automatically and advice the rescuer to defibrillate or not. In the case of no-treat, the rescuer has to determine the presence of pulse which is a time consuming operation. In the case of a treatable rhythm, defibrillation is recommended.

The current protocol for treatment of cardiac arrest [81] advices three successive defibrillations before sequences of artificial circulation is resumed by precordial compressions, ventilation and drug therapy. Each round of artificial circulation is ended by a defibrillation attempt. This procedure is kept going until ROSC is achieved or the patient is given up. Following each defibrillation, the pulse is checked to determine further treatment.

The semi-automatic defibrillators analyse the ECG, and a shock is delivered if VF is detected. No artificial circulation is provided in the period from start of analysis until the defibrillation. The consequence of this is a drop in the MBF and consequently a decrease in the probability of defibrillation success.

[5] Blood supply is deficient.

2.1.2 Improving therapy of cardiac arrest

As discussed previously, a means for monitoring a parameter correlated to resuscitability during CPR would be helpful in determining how to proceed with the treatment to assure a high probability of defibrillation success. It is important to note that VF is a dynamic rhythm changing with time. The evolution of VF can be segmented into four phases as follows [105]:

1. *The undulatory stage* which is characterised by a few large wavefronts of excitation in the myocardium. This phase lasts for 1–2 seconds.

2. *The convulsive stage* with a larger number of smaller wavefronts activating the ventricles more rapidly lasting from 10 to 30 seconds.

3. *The tremulous stage* with a large number of segments of reentry [6] lasting for several minutes.

4. *The final stage* characterised by contractions becoming gradually weaker because of reduced supply of oxygen to the myocardium.

Time is one of the determinant factors for the outcome of cardiac arrest [57], and the differences in the nature of VF listed above are dependent on time and thus correspond to different outcomes. The number (complexity) and speed (frequency) of wavefronts exciting the myocardium as well as the strength of the contractions (energy) affect the morphology of the recorded ECG. Figure 2.2 shows ECG tracings corresponding to different defibrillation outcomes. We have plotted four second tracings corresponding to VF recordings immediately prior to defibrillations. The tracing in a) corresponds to ROSC conversion while those in b), c) and d) were converted to the No-ROSC rhythms EMD, VF and asystole respectively. Note the important differences in signal characteristics showing the great variability in the VF curveform morphologies. The successful outcome trace demonstrates a higher degree of complexity, frequency and energy than the tracings corresponding to different types of unsuccessful outcomes. VF analysis tries to quantify such characteristics. In the following we identify the possible benefits to CPR by use of VF analysis.

One of the key questions addressed here is *whether it is possible to predict the outcome of defibrillation, the prediction based on analysis the patient's VF characteristics.*

[6] Reexcitation of a region of cardiac tissue by a single impulse, continuing for one or more cycles... [30]

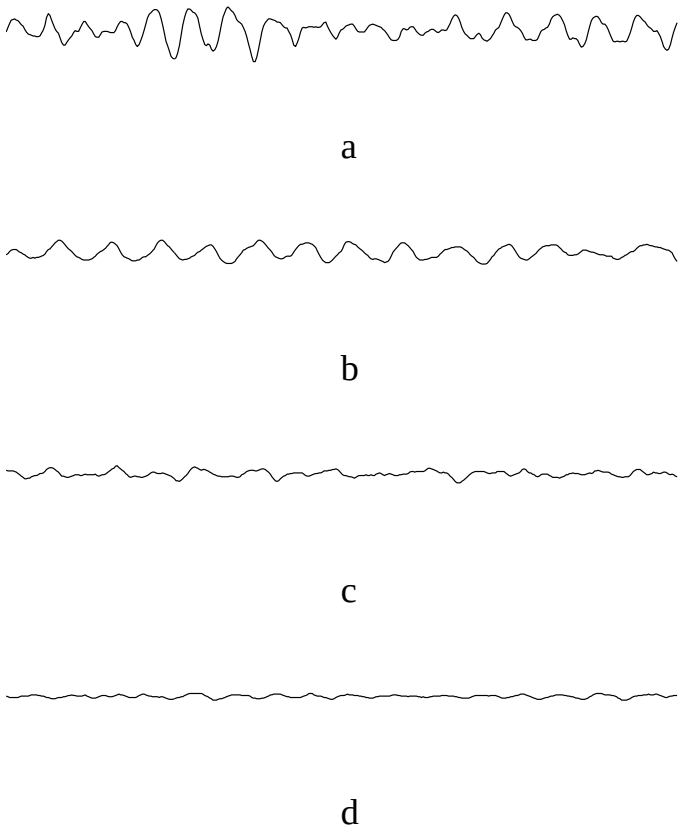


Figure 2.2: Example ECG tracings corresponding to four second VF tracings prior to defibrillations resulting in conversion to a) ROSC, b) EMD, c) VF, d) asystole.

A second related question is whether the effect of CPR can be monitored using VF characteristics.

The third question raised in this work is *whether rhythm analysis and defibrillation can be performed during CPR, thus preventing a decrease in the myocardial and cerebral blood flow levels?*

We discuss these questions further in the following section.

2.2 Problem formulations

The main problems we investigate in later chapters are presented in some detail in this section.

Although defibrillation of ventricular fibrillation (VF) is the most important intervention in the treatment of cardiac arrest patients [64, 93, 50, 6, 47, 100], the majority of the individual defibrillation shocks do not result in return of spontaneous circulation (ROSC) [50, 6, 47, 100].

State-of-the-art defibrillators use ECG rhythm analysis to determine the patient's heart rhythm. The device advices the rescuer whether to shock or not. VF-analysis as an add-on to rhythm analysis would provide further information by which to optimise treatment. The ECG subject to rhythm and/or VF analysis has to be free of CPR artefacts as the presence of these would interfere with the analysis [95]. Today, treatment is stopped for analysis and further preparing for an eventual shock. These pauses for analysis are thus pauses in treatment. Valuable time has been wasted if the following defibrillation is unsuccessful.

Two major and closely related problems stand out: Is it possible to *predict the outcome of defibrillation* to reduce the number of unsuccessful shocks and furthermore to *monitor the effectiveness of CPR*? Can signal analysis methods be applied successfully to these two problems and possibly guide therapy towards improved survival rates in out-of-hospital cardiac arrested patients?

If the defibrillator can perform the analysis and charging during CPR less time will pass letting the metabolism change unfavorably. In this context it is necessary to determine the possibility of splitting the ECG during CPR into two channels, one reflecting the myocardial metabolism and the other containing the CPR artefacts. The key elements of CPR are precordial compressions in combination with ventilation and drug therapy. The *removal of CPR artefacts* stands out as the third major problem to be addressed.

In the experimental part of this thesis we analyse both human and animal ECG to study these problems.

Removal of CPR artefacts from animal ECG is demonstrated in chapter 5.1. The effect of various drugs during CPR are studied in chapter 5.2 and 5.3 to illustrate and evaluate methods for both outcome prediction and monitoring of CPR efficacy.

In chapter 7 we mix CPR artefacts in animal ECG with human ECG to simulate CPR artefacts in human VF. This lets us evaluate the quality of the different CPR artefact removal techniques.

ECG segments prior to defibrillations in human cardiac arrested patients are grouped according to the following outcome. Furthermore methods for outcome prediction are evaluated in chapter 8. Monitoring analysis is also done on these data in chapter 9.

We discuss the three problems in more detail in the following subsections.

2.2.1 Predicting defibrillation outcome

A recent study from Cobb et al [24] indicates that patients with VF may have a better chance of return of spontaneous circulation (ROSC) after a period with chest compressions and ventilation before the first defibrillation attempt. The cardiopulmonary resuscitation (CPR) induced myocardial perfusion can cause changes in the power spectrum of the VF with an attendant increase in the probability of ROSC [13, 95, 73]. Futile defibrillation attempts are in themselves detrimental as tissue damage and post-resuscitation myocardial dysfunction may be caused by the shock itself[115], and by the lack of tissue perfusion from chest compressions during the shock period (analysis, charging, defibrillation and outcome evaluation). It would therefore be important if it could be predicted whether a shock will cause ROSC or not. If the rescuer can predict the outcome of defibrillation, he can avoid wasting valuable time and applying useless and harmful energy to the patient's heart. Much time is spent while the defibrillator analyses the ECG and charges its capacitor.

20-80% of defibrillation attempts in clinical studies are previously reported to cause discontinuation of VF[50, 6, 47, 100] (the great variability depends on different time definitions of VF reoccurrence and/or different shock waveforms).

We illustrate the potential benefit of outcome prediction in light of some of the findings presented and discussed in chapter 6. Approximately 10% of the 883 shocks given to the patients in the Oslo database succeeded in converting the

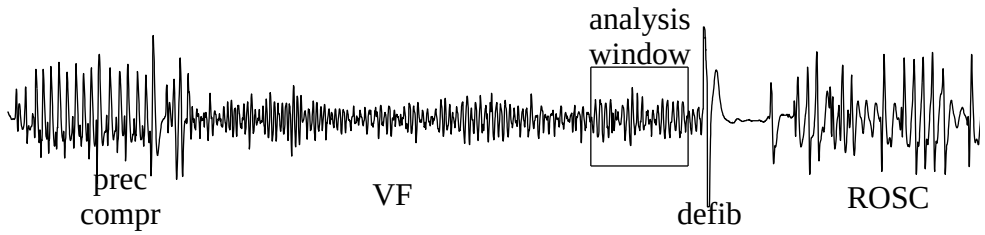


Figure 2.3: ECG tracing illustrating precordial compressions (prec compr), VF, defibrillation (defib) and conversion to pulse rhythm (ROSC).

heart to produce a pulse rhythm. This is similar to reports from Milwaukee and Iowa [50, 6]. The median time spent preparing for each defibrillation was 20 seconds in the Oslo data. A median of six shocks were given to each patient. Thus, for the average patient this means nearly two minutes passing with no treatment provided. In addition to the time spent preparing for a shock comes the time spent after an unsuccessful shock to determine outcome, which in many cases will be at least 20 seconds.Providing efficient CPR during these minutes would unquestionably be of great value to the patient rather than wasting it preparing for useless shocks.

Figure 2.3 shows 50 seconds of ECG recordings illustrating the typical events preceding and succeeding a defibrillation. During the first ten seconds, precordial compressions are performed. The ECG is dominated by the CPR artefacts. In the following period, preceding the defibrillatory shock, VF is seen. During the VF period rhythm analysis and capacitor charging is done to prepare for the shock. The window containing the VF analysis block is placed to show the point of analysis. In the given illustration the shock is followed by an *isoelectric* period signifying that the myocardium has been electrically reset. In the following period the heart reorganises into the outcome rhythm. In this case a pulse rhythm results. Other conversion rhythms might be EMD, asystole or VF. A non-reset shock is a special case where the shock is unable to reset the heart. After such a defibrillation VF reappears within five seconds. All these non-pulse outcomes are considered unsuccessful.

The VF analysis blocks prior to all shocks in the data material are grouped according to outcome. The following scheme for grouping successful versus unsuccessful shocks is used:

- *Successful*: A palpable pulse is present in the post shock period.
- *Unsuccessful*: A palpable pulse is not present in the post shock period. This class of data consists of the EMD, Asystole, VF and non-reset shocks.

The preshock feature values corresponding to the post shock pulse group should fall in a region distinguishable from the region corresponding to the post shock non-pulse group. The prediction outcome analysis will be evaluated in the pattern recognition framework presented in chapter 3.

2.2.2 Monitoring the effect of CPR

A certain amount of myocardial perfusion is required for successful resuscitation [58, 76]. Cobb et al have pointed to the usefulness of a variable enabling monitoring of CPR efficacy and thus providing a means for the rescuer to improve treatment and patient outcome [24].

The rescuer has to decide the placement, depth and frequency of the precordial compressions, ventilation force and drug dosages in addition to when to shock. If the rescuer can be guided in performing the therapy, this will cause faster and more successful resuscitation.

Monitoring the effect of CPR involves detecting changes in the myocardial metabolism caused by these therapies. In the human data we will monitor the effect of foregoing CPR in the artefact free preshock ECG tracings. Thus, as an example the monitoring variable extracted from first preshock tracing gives information on the resuscitability before the first defibrillation. The following sequence of extracted monitoring variables should convey information about the evolution of effect of the therapy. The monitoring of patients who have been successfully resuscitated after a long period of CPR and shocking should show a positive development (increased resuscitability) in the monitoring parameter(s). In a patient which has not been resuscitated after a similar treatment period, the parameter(s) should remain unchanged or have a negative development.

The evaluation of monitoring methods is difficult because many parameters vary in the patient database. Some of these parameters are the time duration of the sequences, the initial myocardial metabolism, the cardiac aetiology, the frequency at which the precordial compressions are given and the dosage and timing of drug therapy.

During animal research the treatment is provided according to a well defined protocol so that the ECG may be divided into phases according to treatment given. This will illustrate the effect of monitoring the ECG measurements during the different treatment phases and to evaluate how well the measurements discriminate between these phases. Our proposed method for monitoring (chapter 9) is based on principles from statistical pattern recognition presented in chapter 3.

2.2.3 CPR artefact removal

In the current operation of automated external defibrillators, substantial time is consumed in the "hands off" interval during which precordial compressions are discontinued to allow for automated rhythm analysis before delivery of the electric counter-shock. Current guidelines for cardiopulmonary resuscitation require that chest compression and ventilation must be interrupted prior to any shock, to avoid the effects of artifacts on the ECG analysis [43, 81]. During this period there is no circulation of the heart muscle or brain with a rapid deterioration of the metabolic state of the tissues [29, 69]. This detoriation can be partially reversed by chest compressions and ventilations [72, 59]. In agreement with this, Sato et al found that the interruption of chest compressions before a defibrillation attempt reduced the defibrillation success rate. 24 hour survival was significantly reduced with a 20-sec delay [84]. They concluded that automated defibrillators are likely to be maximally effective if they are programmed to secure minimal "hands-off" delay before delivery of the electric counter-shock.

CPR artefacts interfere with the VF part of the ECG signal. Strohmenger et al showed that this interference could be reduced by frequency selective filtering to achieve reliable VF analysis [95].

The successful use of frequency selective filtering (illustrated in chapter 5.1) relies on the fact that the frequency components of CPR artefacts and VF are non-overlapping in pig data. In human ECG, these components are overlapping. Strohmenger et al pointed this out as a factor detoriating the predictive power of VF analysis in human ECG [94]. We will design a method (chapter 3.4) for removing CPR artefacts from human ECG that handles this problem (chapter 7.

A good filter should not effect VF analysis, and the performance of the defibrillator algorithm should remain unchanged. The removal of cardiopulmonary resuscitation (CPR) artefacts in VF would thus make it possible to assess the CPR effect on the myocardium as indicated by VF changes.

2.3 Previous VF analysis research

In the following section we summarise the key elements of previous work in VF-analysis. The work has mainly been concerned with features describing amplitude characteristics of the VF tracings and studies of frequency parameters derived from the fast Fourier transform of the VF.

2.3.1 Objectives

The optimal ECG derived feature should reflect the patient's degree of resuscitability. In the works referred to in the following several different strategies has been applied for this purpose. Brown et al discussed the most important of these in [9]:

- Estimate the duration of VF, i.e. the time passed since the onset of VF. This is also referred to as VF downtime.
- Monitor the myocardial perfusion during CPR, i.e. the effect of CPR on the metabolism. In this case, the ECG features ability are gauged through correlation analysis to additional recordings of CPP, MBF, $ETCO_2$ or other established invasive techniques for monitoring the effect of CPR.
- Determine the optimal time to defibrillate during VF. The features predictive ability is determined by calculating the sensitivity and specificity (definition given in chapter 3.2).

These are the main objectives of VF analysis as it is described in the literature. VF analysis by ECG features have traditionally been dominated by features measuring the VF amplitude and features identifying the dominant frequency components from the fast Fourier transform of the VF signal. From reading the articles referred to in the following it is clear to us that there are several examples of contradictory findings.

2.3.2 Research findings

In [116] it was demonstrated in animals that the probability of successful defibrillation of VF varies inversely with downtime, i.e. the duration of time between the onset of VF and the initiation of CPR.

Amplitude parameters Fibrillation amplitude was pointed out as a powerful indicator of outcome after cardiac arrest in [109]. It was reported that a detailed inspection of the VF curveform reveals variation in both signal amplitude and frequency, amplitude differences being most appearant. Initial VF amplitude was measured in 394 human patients, and it was shown that the group of patients presenting VF with low signal energy had a lower rate of successful defibrillations compared to the group presenting VF with high signal

energy. The group with coarse VF [7] had the higher proportion of witnessed arrests with initiated bystander CPR and lower downtime before initiation of CPR and defibrillation.

In [14] it was reported that VF amplitude was predictive of countershock outcome in a study of 265 human patients. VF amplitude was also shown to be positively related to bystander CPR, postshock rhythm, electrically stable conversion, inpatient admission and hospital discharge but unrelated to response and defibrillation intervals. The VF amplitude was highly predictive of hospital outcome.

In [62] a measurement of VF amplitude in animals associated with successful outcomes of defibrillations was shown to be different from those associated with unsuccessful outcomes.

The potential of VF amplitude as a monitor of the effectiveness of CPR was studied in [72]. In a rodent model a rise in VF voltage was correlated to a rise in CPP during CPR. It was also confirmed that VF amplitude, serves as a quantitative predictor of defibrillation outcome.

Frequency parameters The median frequency was introduced in [33] as a measurement for estimating downtime. The study revealed the potential of the median frequency as a downtime estimator. It also pointed out that VF amplitude is an unreliable estimator of downtime due to uncontrollable variations in anatomy, physiology and instrumentation. A further study of median frequency was presented in [12] and in [68] where the downtime estimated from human ECG was analysed and compared to comparative animal data from [33, 13]. In [13] the median frequency was reported as being very well suited for prediction of defibrillation outcome in an animal experiment.

Comparing amplitude and frequency parameters In [92] the dominant frequency and amplitude of VF was compared for ability to distinguish between primary and secondary VF in 41 human patients . Primary VF had a higher probability of successful countershock outcome than secondary VF. The study suggested that low frequency VF indicates a poor chance of successful resuscitation. The peak-to-trough amplitude of VF showed no significant difference.

In [97] the median frequency was studied in 20 human patients undergoing aortocoronary bypass grafting. It was reported that there was a high degree

[7] The medical literature denotes VF with low signal energy by "fine VF" while VF with high signal energy is denoted "coarse VF".

of correlation between successful defibrillation and the median frequency, and that the observations support the hypothesis that measurement of the median frequency reflects the overall metabolic state of the myocardium.

The interference by precordial compressions on VF analysis was assessed in [95] to determine which of several spectral analysis parameters reflected best myocardial blood flow and resuscitation success during CPR. The median frequency extracted from the frequency area corresponding to 4.3-35 Hz differentiated best between resuscitated and non-resuscitated animals. There was no significant difference between the measurements taken immediately before CPR compared to the ones taken immediately after start of CPR. It was also noted that the experiments revealed no correlation between VF amplitude and myocardial blood flow.

A study was performed by Strohmenger et al on 154 defibrillations in ECG recordings from 26 cardiac arrested out-of-hospital patients [94]. The predictive ability of both spectral and amplitude parameters were evaluated. In this study, the dominant frequency was identified as the single best predictor of defibrillation outcome. The predictive power of the amplitude parameter was comparable to that of the median frequency. In this study, features were not combined. Strohmenger et al also evaluated the effect of CPR artefacts on VF analysis in this study and found that the predictive power of the parameters was detoriated even after the artefacts had been eliminated.

Combining amplitude and frequency parameters In a similar study predating the one by Strohmenger et al [94] 128 defibrillations administered to 55 human patients were analysed in [10]. Both frequency and amplitude parameters immediately prior to countershock were analysed. The parameters were grouped according to countershock outcome and compared. The median frequency was shown to be the single best predictor of countershock outcome, while the dominant frequency and amplitude parameters were poor predictors. Brown et al proposed using the combination of the median and dominant frequency. Thus, improved predictive power was achieved as compared to using the median frequency alone.

Momsieurs et al combined manually derived variations of energy and frequency parameters to distinguish survivors from non-survivors in [70]. The parameters were extracted from VF tracings at the start of the initial rhythm analysis period in the ECG record for each of the patients. The combined features distinguished fairly well between both ROSC and No-ROSC as well as between survivors and non survivors. Performance was improved by adding age to the parameter set.

In [73] the monitoring and predictive capability of both frequency and amplitude ECG measurements were investigated in an animal experiment. This work identified the combination of VF amplitude and dominant frequency as being superior in this respect.

Other recent methods Recently other methods for characterising ECG has been introduced to the field of VF analysis.

In the recent publications from [117, 90] nonlinear methods were used to characterise the dynamics of the initial phase of VF in animal experimental data. The usefulness of wavelets giving a time-frequency representation of VF were indicated in [108]. These methods were not evaluated with regard to any of the objectives stated in the beginning of this section. It will be interesting to see if these more complex methods will outperform the frequency and amplitude parameters.

2.3.3 Comments

In [97] it was remarked in the conclusion that the findings are specific for certain clinical settings and that further studies on a broader range of both animal experimental and human patient settings is necessary to determine the general utility of the median frequency.

It is our impression that the previous work has focused on either amplitude or frequency parameters. Both these types of parameters have been found to reflect myocardial perfusion, myocardial energy metabolism, defibrillation success and outcome after cardiac arrest. The main critisism against the amplitude parameter is that it depends on the direction of the main fibrillatory vector and interindividual variability is considerable.

There are only a few articles describing work were parameters has been combined rather than compared [10, 73]. The results are encouraging. There might be information reflecting resuscitability in an amplitude parameter that is not contained in the frequency parameter. If the frequency parameter outperforms the amplitude parameter when evaluated independently we expect the combined paramter to outperform both. One of our main objectives is to adapt methods to VF analysis that allows the combination of several parameters.

The monitoring aspect has been investigated in several differently structured studies. High values in median/dominant frequency has been assosciated with high resuscitability. In chapter 3 we will have a closer look at the methods used in some of the articles we have referred to in this section. We will point

out some weak spots and propose our own methodology designed to handle these problems. One of the key issues is the methods used for the prediction of defibrillation outcome. The prognostic criteria are mostly developed and tested on the same data set (except for Noc et al [73] and Monsieurs et al [70]). We will use the concept of training and testing in the design of our decision support system to control performance and ensure more reliable results.

2.4 Summary

This chapter has presented an engineer's perspective on how CPR affects the basic mechanisms of the circulatory system. We have emphasised the relation to the recorded ECG.

We have provided a careful definition of the problems of predicting defibrillation outcome, monitoring the effect of CPR and removing CPR artifacts. These problems are addressed in later chapters.

An overview has been given of earlier work on VF analysis. This work has demonstrated promising results based on analysis of ECG from animal experiments. Results from research on human data are not as encouraging.

Chapter 3

Decision Support

In this chapter we identify the modules in a decision support system for guiding therapy using VF-analysis. It is shown how the concepts of prediction of defibrillation outcome, monitoring of CPR-efficacy and removal of CPR artefacts fits into this system.
In the following sections we motivate for our choice of methods for the different modules in the decision support system and present the theoretic foundation necessary to justify our choices.

3.1 A VF analysis based medical decision support system for guiding therapy during CPR

The system we are discussing is a medical decision support system for guiding therapy [107]. It is convenient to split the data processing for decision making into modules as follows:

1. Recording of ECG and related information.
2. ECG is preprocessed and CPR artefacts are removed.
3. Characteristic features are extracted from the data.
4. A decision support module guides therapy learning its decision rules from annotated data.

In the following we recapitulate how the problems related to modules 2-4 has been dealt with in the previous work we discussed in section 2.3.
Developing good features and selecting the best ones are the key issues for the design of the decision support system. The features should contain different aspects of the sought-after information - resuscitability.

Methods for CPR artefact removal

The second module of the system involves artefact filtering. In animal experiments, the features have been extracted continuously during the different therapeutic stages. When these stages have involved precordial compressions and ventilations, the features has been shown to have a lower correlation to the measurements associated to resuscitability [95]. This effect of CPR on the features has been identified to be caused by noise on the ECG, the CPR-artefacts. The importance of being able to do ECG analysis during CPR has been stated by several investigators [24, 100]. In [95] it was shown that the interference of the CPR-artefacts on the VF-analysis was reduced by using frequency selective analysis corresponding to removing the three lowest frequency harmonics of the CPR-artefacts prior to analysis. Similarly, a digital low-pass filter was applied by Noc et al in their animal experiment [73]. The effect of this for analysis purposes is mainly the same as achieved by Strohmenger et al [95]. Strohmenger et al used the frequency selective method in a study of the effect of CPR artefacts on VF-analysis in human ECG [94]. It was shown that the three lower harmonics of the artefact had to be removed to remove the artefact interference on the VF-analysis. This was shown to deteriorate the predictive power of the features in question.

We see weaknesses related to the artefact-filtering methods used by Strohmenger et al [95] and Noc et al [73]. The success of the frequency selective methods depend on the frequency components of the VF and artefact signals being separable. This criterion is fullfilled by animal ECG (pigs), but in human ECG, the signal components of these two signal sources overlap in the frequency domain. This means that the application of the frequency band limiting methods to human ECG will remove information from VF. To retain the VF information, more of the CPR-artefact will be mixed into the filtered signal. These problems were identified and discussed by Strohmenger et al in [94]. We demonstrate this problem in chapter 5.1.

Evaluation of features for monitoring and outcome prediction

Typically, in the previous work of VF-analysis, the features of interest are extracted and evaluated individually except for for the cases where Monsieurs et al and Noc et al [70, 73] where regression techniques [107] were applied to combine two features into a single valued unit. The evaluation methods applied vary according to the problem addressed.

In the case where the aspect of *monitoring* of the myocardial perfusion has been investigated, the evaluation has been done by *correlation* to established

measurements like MBF or CPP which relates to the myocardial perfusion [109, 92, 73]. These measurements are only available in animal experimental settings. The parameters monitoring capabilities has been investigated on an individual basis except for the cases mentioned above.

To determine the features' capability for *outcome prediction*, the features have been grouped according to successful and unsuccessful outcomes in resuscitation attempts. The evaluation has been done both by analysis of statistical distributions and a retrospective predictive analysis. Various aspects of *statistical hypothesis* testing techniques were applied to determine which features are significantly different when compared between groups [62, 95, 10, 73].

For further design of decision support rules, prediction analysis by studying *Receiver Operating Characteristics* (ROC) has been applied. In the previous work this has basically involved the design of decision models by calculating several linear decision functions. Each of these functions are made so that a given proportion of the successful outcomes are correctly predicted. This is the sensitivity. The corresponding proportion of correctly predicted unsuccessful outcomes, the specificity, is calculated. In the single-feature cases, the ROC-analysis is easily performed to produce sensitivities ranging from 0-1 with corresponding specificities. This enables the designer to choose the decision function satisfying a specific performance criterion. For the multiple feature cases [10, 70, 73] the ROC analysis has been limited to selecting a decision function corresponding to a sensitivity equal to one [10] except for the cases where multiple regression techniques were applied [70, 73]. Interestingly, in both [70, 73] the use of feature combinations resulted in decision functions incorporating both a frequency and an energy feature.

To evaluate the *reliability* of the resulting decision model, the prognostic criteria were defined from one data set (training) and evaluated on another data set (testing) in the study of Noc et al [73]. The results from the two data sets did not match. A more sophisticated variation of this, cross-validation, where the decision function applied in the prediction of one case is derived from all other but that case [70]. In other work, the prognostic criteria has been evaluated on the training sets alone [10, 94]. Typically, the results obtained from prediction in animal data outperforms the results obtained from human data. We believe this to be du to the animal models are simple in the sense that all animals are healthy prior to the treatment which follow the same timeline for all animals. Thus, the animal material does not represent the great variability in factors like age, health, prior treatment etc as found in human cardiac arrest patients.

To summarise, we find several weaknesses in the methodology applied in the development of the decision functions in previous work.

Most important is the limited analysis of higher-dimensional features. In our opinion a decision function can be improved by adding features with new information. This means that both feature selection and ROC analysis needs to be done on higher dimensional data. The regression techniques applied by Monsieurs et al [70] and by Noc et al [73] handles this. A closer inspection of their decision functions reveals them to be either a linear or logistic plane in 3-D with unlimited expansion. This is another crucial point we would like to point out: The decision function should not map areas of feature space not inhabited by data in the training set. Such areas should be invalid for prediction, resulting in *rejection* of testing features inhabiting these areas.

Introducing a VF-analysis decision support system using pattern recognition methods

In conclusion, we want to design a decision support system for VF analysis satisfying the following criteria:

1. Allow multi dimensional feature handling.
2. Allow sensitivity specification (ROC) for multi dimensional feature classification.
3. Allow evaluation concerning reliability of system model.
4. Allow artefact filtering for human ECG.

As we will show in the following sections the three first issues can be handled by a pattern recognition framework as depicted by Duda and Hart [32]. Our methods are based on the more recent publications by Schurmann [87] and Ripley [80]. We approach the fourth issue by using adaptive filtering techniques [113] with references obtained through measuring signals correlated to associated artefact components.

The medical decision support system we design will consist of the modules illustrated in figure 3.1.

Pattern source The failing heart of the patient generates electrical potentials registered in an ECG. Our working hypothesis is that this signal contains information correlated to the resuscitability of the patient. This is the information we want to find using feature extraction methods.

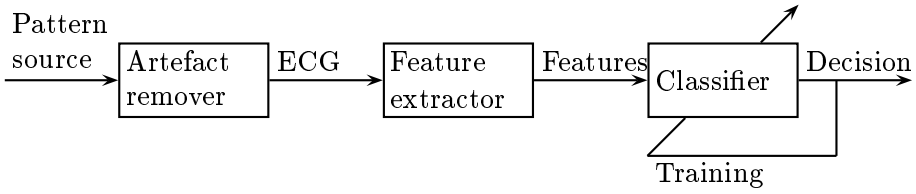


Figure 3.1: The modules in a medical decision support system for VF analysis.

CPR artefact remover In this module, the CPR-artefacts and other possible noisy components are removed from the ECG signal.

Feature extractor A feature vector is extracted from the ECG.

Classifier The classifier processes the feature vector and determines the outcome based on a minimum risk analysis. This decision is based on a training set of feature vectors where the correct outcomes are known. The training is done by adjusting the parameters of the classifier so that its outcome predictions meets a performance criteria for ROC-analysis. To avoid the problem of overtraining (too good fit to training data) an independent test set is used to assure generality.

In the remaining parts of this chapter we describe the classification and feature extraction techniques and the CPR-artefact removal methods used in this work. As the first three of the four criteria mentioned at the beginning of this section are being satisfied through proper design of the classification module, we present the decision support system starting with this very module. Presentations of the two other modules shown in figure 3.1.

3.2 Classification

We present the parts of pattern recognition theory necessary to formulate the classification module satisfying the criteria given in section 3.1.

In pattern recognition the objective is to identify the belonging of a *pattern* to one of several possible *classes*. The field of pattern recognition is generally

subdivided into two main groups, according to whether syntactic or statistical methods are applied.

3.2.1 Statistical pattern recognition

In this work we use the statistical method where a pattern is represented by a number of numerical features which are combined in a *feature vector*, $\mathbf{v}$. We also know the number, K, of distinct classes,$\omega_i, i = 1, 2, \ldots, K$, and have access to a number of example objects labelled according to class membership. We want to design a *classifier* that assigns unlabelled objects to one of the classes.

Statistical functions

The statistical distributions of the class features are defined by the the class specific *a priori probabilities* and the *class specific probability density functions* (PDF). In the following we discuss these statistical functions and show how Baye's rule is used to derive the *a posteriori probability functions*. In this and the following section we assume the statistical functions to be known unless otherwise stated.

The probability of occurrence of patterns belonging to class ω_i is the a priori probability denoted as.

$$P(\omega_i), i = 1, 2, \ldots, K. \tag{3.1}$$

The a priori probability is the only knowledge we have of the occurrences of the patterns before any observation is done. When a measurement, as realised by the feature vector $\mathbf{v}$, is observed the situation is changed. Our knowledge of the class identity is influenced by the observation as given by the a posteriori probability which is the probability of occurrence of class ω_i given the knowledge of the observation:

$$P(\omega_i|\mathbf{v}), i = 1, 2, \ldots, K. \tag{3.2}$$

The distribution of the measurement vectors are statistically described by the probability density function (PDF) given by equation 3.3.

$$p(\mathbf{v}) \tag{3.3}$$

This compound PDF is the weighted sum of the class specific PDFs as given in equation 3.4.

$$p(\mathbf{v}|\omega_i), i = 1, 2, \ldots, K \tag{3.4}$$

As the compound PDF, $p(\mathbf{v})$ involves the relations shown in Equation 3.5

$$p(\mathbf{v}) = \sum_{j=1}^{K} P(\omega_j) p(\mathbf{v}|\omega_j) \tag{3.5}$$

we get the following relation between the a posteriori probability, the a priori probability, the compound PDF and the class specific PDF as given by Baye's rule in equation 3.6

$$P(\omega_i|\mathbf{v}) = \frac{p(\mathbf{v}|\omega_i)P(\omega_i)}{p(\mathbf{v})}, i = 1, 2, \ldots, K. \tag{3.6}$$

Decision functions

In the classifier each feature vector,$\mathbf{v}$, is considered as belonging to one of K classes, $\omega_i, i = 1, 2, \ldots, K$. In our setting the classes may correspond to different outcomes of resuscitation attempts, or different therapeutic interventions corresponding to different degrees of resuscitability.

We want to split the feature space into K decision regions, $R_i, i = 1, 2, \ldots, K$, by assigning costs for the possible wrong decisions. Each cost, $C(\omega_i, \omega_j)$, expresses the risk associated with classifying a pattern of the true class, ω_i, as belonging to the decided class, ω_j. As we show later in this section, the costs are selected to satisfy a given performance criterion. A reject class, ω_{K+1}, is added to handle out-of-range patterns.

Each R_i is calculated by selecting the minimum component of the risk vector [87].

$$\mathbf{r} = \mathbf{C}^T\mathbf{p} = \begin{bmatrix} \sum_{j=1}^{K} C(\omega_j, \omega_1)P(\omega_j|\mathbf{v}) \\ \sum_{j=1}^{K} C(\omega_j, \omega_2)P(\omega_j|\mathbf{v}) \\ \vdots \\ \sum_{j=1}^{K} C(\omega_j, \omega_K)P(\omega_j|\mathbf{v}) \\ \sum_{j=1}^{K} C(\omega_j, \omega_{K+1})P(\omega_j|\mathbf{v}) \end{bmatrix} \tag{3.7}$$

where

$$\mathbf{p} = \begin{bmatrix} P(\omega_1|\mathbf{v}) & P(\omega_2|\mathbf{v}) & \cdots & P(\omega_K|\mathbf{v}) \end{bmatrix}^T. \tag{3.8}$$

and

$$\mathbf{C} = \begin{bmatrix} C(\omega_1, \omega_1) & C(\omega_1, \omega_2) & \cdots & C(\omega_1, \omega_K) & C(\omega_1, \omega_{K+1}) \\ C(\omega_2, \omega_1) & C(\omega_2, \omega_2) & \cdots & C(\omega_2, \omega_K) & C(\omega_2, \omega_{K+1}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C(\omega_K, \omega_1) & C(\omega_K, \omega_2) & \cdots & C(\omega_K, \omega_K) & C(\omega_K, \omega_{K+1}) \end{bmatrix} \tag{3.9}$$

is the loss matrix where ω_{K+1} is the *reject* class. Thus, the classification, or decision rule, corresponds to selecting the class number equal to the minimum component index of $\mathbf{r}$. This represents the optimal decision in the sense that it minimises the expected overall classification risk [87].

To summarise, we want to estimate the a posteriori probability, $P(\omega_i|v), i = 1, 2, \ldots, K$, of a given outcome based on the knowledge of the a priori probability of occurrence of the corresponding class given the measurements of the features. Using these probability functions and the assigned costs, the risk vector, $\mathbf{r}$, is formed. The components, $r_i, i = 1, \ldots, K$, are the class specific decision functions. For a given outcome,$\mathbf{v}$ the estimated class corresponds to the minimum r_i.

Performance evaluation

We have to determine how discriminable the class patterns are. The classifier performance has to be evaluated, and for this purpose, the *confusion matrix*, $\mathbf{P}_R$, is very useful for deriving several performance measures. It is given as

$$\mathbf{P}_R = \begin{bmatrix} P(R_1|\omega_1) & P(R_1|\omega_2) & \cdots & P(R_1|\omega_K) \\ P(R_2|\omega_1) & P(R_2|\omega_2) & \cdots & P(R_2|\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ P(R_{K+1}|\omega_1) & P(R_{K+1}|\omega_2) & \cdots & P(R_{K+1}|\omega_K) \end{bmatrix} \tag{3.10}$$

where $P(R_j|\omega_i), i = 1, \ldots, K+1, j = 1, \ldots, K$ is the probability of classifying a pattern of the true class ω_i as belonging to ω_j. $R_i, i = 1, \ldots, K+1$ are the class specific decision regions. These are dependent of the choice of values in the cost matrix $\mathbf{C}$.

Figure 3.2 shows how the decision regions in a two class problem of discriminating between 1-D feature vectors are found, given equal costs for all decision errors. This corresponds to the regular Bayes classifier [87]. Class ω_1 is identified by the solid lines in the plot while dotted lines are used for class ω_2. The class specific PDFs weighted by their respective a priori probabilities are shown in part a) of the figure. The distributions used are gaussian with unity standard deviation and different means. The a priori probabilities are identical. The corresponding a posteriori probability functions are shown in part b) of the figure. For the Bayes classifier, the decision rule is to select the class corresponding to the maximum a posteriori probability. This means that the decision region, R_1, for class ω_1 corresponds to the range of feature values where $P(\omega_1|\mathbf{v}) \geq P(\omega_2|\mathbf{v})$ (and vice versa for the decision region, R_2, for class ω_2).

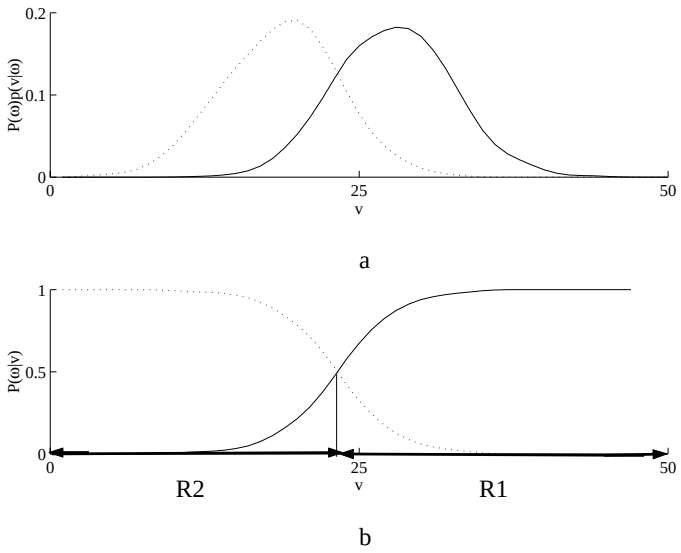


Figure 3.2: Two-class statistical functions. Solid line = class ω_1 and dotted line = ω_2. a) Class specific PDFs b) a posteriori probability functions with decision regions R_1 and R_2 for classes ω_1 and ω_2 respectively.

3.2.2 Feature observations

Until now we have assumed the statistical functions to be known. As we will discuss later this is not the case, so the functions have to be estimated from observed feature vectors with known class identity. In the following we define how the class specific observed feature vectors are organised into feature sets.

A feature vector $\mathbf{v}$ consists of D elements so that

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_D \end{bmatrix}^T . \tag{3.11}$$

We define the *class specific feature set* of observed feature vectors corresponding to observations associated with pattern class ω_i according to

$$V_i = \{\mathbf{v}|''\mathbf{v}\epsilon\omega_i''\}. \tag{3.12}$$

The total number of classes being K, we define the *total feature set*, V to be

$$V = \cup_{i=1}^{K} V_i. \tag{3.13}$$

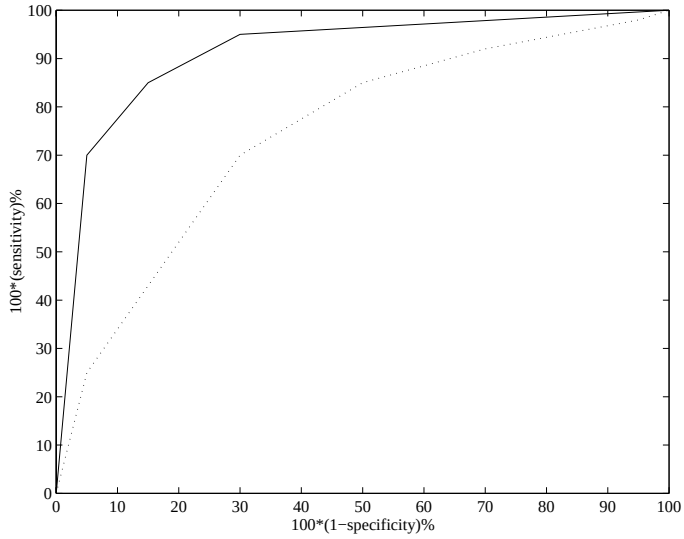


Figure 3.3: ROC curves for two classifiers. The predictive power of a classifier is defined as the area beneath its ROC curve. The dotted and solid lines represent the classifiers with the lowest and highest predictive power respectively.

3.2.3 Performance control

As we discussed in section 3.1, we want to design our classifier to meet specific performance criteria defined by the user. For example the clinicians might demand that the decision support system we design should be expected to predict 95% of all successful shocks correctly. Thus, we would have to train our classifier to meet this demanded *sensitivity* or *true positive rate*. A value closely connected to the sensitivity is the *specificity* which in our example would be the proportion of correctly predicted unsuccessful shocks. Receiver Operator Characteristics (ROC) analysis studies the effect of varying the sensitivity on the specificity. Figure 3.3 illustrates this. The plot shows the ROC curve with the sensitivity on the y-axis and $100(1-specificity)$ (*false positive rate*) on the x-axis. In the following we will show how we can integrate ROC analysis into the pattern recognition framework presented above by showing how the loss matrix of equation 3.9 can be used to control the sensitivity.

First we need to define the performance parameters used in ROC analysis. Second, we will show how these performance parameters can be expressed by the equations given in section 3.2.1.

In medical diagnostics, *sensitivity and specificity* are two important performance characteristics and we start off by defining some quantities according

to the definitions of sensitivity and specificity given in [30].

n	total number of observations
n_{ω_i}	number of observations belonging to class ω_i
$n_{R_{i\|\omega_i}}$	number of observations correctly classified as belonging to class ω_i
$n_{R_j\|\omega_i}$	number of observations wrongly classified as belonging to class ω_j, the correct class being ω_i

Letting the number of observations approach infinity we get

$$P(\omega_i) = lim_{n\to\infty}\frac{n_{\omega_i}}{n}$$
$$P(R_j|\omega_i) = lim_{n\to\infty}\frac{n_{R_j|\omega_i}}{n_{\omega_i}}$$

In the following class ω_i is the state we want to identify, and $\omega_j, j = 1, \cdots, K, j \neq i$ are the other classes, and

$n_{R_{i\|\omega_i}}$	number of true positive results
$\sum_{j\neq i} n_{R_j\|\omega_i}$	number of false negative results
$\sum_{j,k\neq i} n_{R_j\|\omega_k}$	number of true negative results
$\sum_{j\neq i} n_{R_i\|\omega_j}$	number of false positive results

We now show how these performance characteristics can be expressed in the pattern recognition framework.

Sensitivity The sensitivity is defined as *the number of true positive results divided by the number of true positive and false negative results* and may be expressed

$$P_{sns}(\omega_i) = \lim_{n\to\infty}\frac{n_{R_{i|\omega_i}}}{n_{\omega_i}} = P(R_i|\omega_i) \tag{3.14}$$

Specificity The specificity is defined as *the number of true negative results divided by the number of true negative and false positive results* and may be expressed

$$\begin{aligned} P_{spc}(\omega_i) &= \lim_{n\to\infty} \frac{\sum_{j,k\neq i} n_{R_j|\omega_k}}{\sum_{j,k\neq i} n_{R_j|\omega_k} + \sum_{j\neq i} n_{R_i|\omega_j}} \\ &= \lim_{n\to\infty} \frac{1}{1 - n_{\omega_i}/n} \sum_{j,k\neq i} \frac{n_{\omega_k}}{n} \frac{n_{R_j|\omega_k}}{n_{\omega_k}} \\ &= \frac{1}{1 - P(\omega_i)} \sum_{j,k\neq i} P(\omega_k) P(R_j|\omega_k) \end{aligned} \tag{3.15}$$

Controlling sensitivity

Equations 3.7 and 3.10 illustrate two crucial points in the design of the classifier.

1. Equation 3.7 defines the decision rule which relies on the statistics of the data given by the a posteriori probability functions of the classes and the cost values which should be set by the designer.
2. All performance parameters are ultimatively derived from the confusion matrix expressed in equation 3.10 as we showed in equations 3.14 and 3.15.

In the following we show how the cost values in equation 3.9 may be calculated to meet a sensitivity criterion for class ω_1. [1] It is important to remember the result of equation 3.14 which tells us that

$$P_{sns}(\omega_1) = P(R_1|\omega_1), \tag{3.16}$$

so that what we want to do is to control the number of feature vectors in the class specific feature set V_1 going into R_1.

As we are dealing with resuscitability (ω_1) versus nonresuscitability (ω_2), we are in principle handling a two-class problem. In the cases were we use more than two classes, these may be considered as subclasses of the two main classes. The risk vector will thus be

$$\mathbf{r} = \begin{bmatrix} 0 + C(\omega_2, \omega_1) P(\omega_2|\mathbf{v}) \\ C(\omega_1, \omega_2) P(\omega_1|\mathbf{v}) + 0 \end{bmatrix}, \tag{3.17}$$

[1] In our work we relate class ω_1 to successful outcome (high resuscitability). The sensitivity criterion defines the proportion of patterns corresponding to successful outcomes we want to identify.

where $C(\omega_i, \omega_i) = 0$ and the confusion matrix will be given by

$$\mathbf{P}_R = \begin{bmatrix} P(R_1|\omega_1) & P(R_1|\omega_2) \\ P(R_2|\omega_1) & P(R_2|\omega_2) \end{bmatrix}. \tag{3.18}$$

Considering the feature set formulation of equations 3.12 and 3.13, the training feature vectors in V_1 and V_2 will be distributed between R_1 and R_2 upon classification. For the training data this is true, but in testing, outliers will fall in R_3, the reject class. For the feature vectors (training) of V_1 classified to R_1, the following is true

$$C(\omega_2, \omega_1)P(\omega_2|\mathbf{v}) \leq C(\omega_1, \omega_2)P(\omega_1|\mathbf{v}), \tag{3.19}$$

or equivalently

$$\frac{C(\omega_2, \omega_1)}{C(\omega_1, \omega_2)} \leq \frac{P(\omega_1|\mathbf{v})}{P(\omega_2|\mathbf{v})}. \tag{3.20}$$

After the normalisation $C(\omega_1, \omega_2) = 1$ we finally get

$$C(\omega_2, \omega_1) \leq \frac{P(\omega_1|\mathbf{v})}{P(\omega_2|\mathbf{v})} = \frac{P(\omega_1|\mathbf{v})}{1 - P(\omega_1|\mathbf{v})}. \tag{3.21}$$

The a posteriori probabilities of the feature vectors in V_1 are organised into the vector $\mathbf{C}_{ROC}$ in descending order according to

$$\mathbf{C}_{ROC} = \begin{bmatrix} \frac{P(\omega_1|\mathbf{v}_1)}{1-P(\omega_1|\mathbf{v}_1)} & \frac{P(\omega_1|\mathbf{v}_2)}{1-P(\omega_1|\mathbf{v}_2)} & \cdots & \frac{P(\omega_1|\mathbf{v}_{M_1})}{1-P(\omega_1|\mathbf{v}_{M_1})} \end{bmatrix}^T, \tag{3.22}$$

where M_1 is the number of training feature vectors in V_1.

We want at least $100P_{desired}\%$ of the elements in V_1 to be mapped into R_1. $\mathbf{C}_{ROC}(i)$ corresponds to the element with highest value satisfying $P_{desired} \leq i/M1$. Selecting the cost value corresponding to

$$C(\omega_2, \omega_1) = \mathbf{C}_{ROC}(i), i = \lceil P_{desired} M_1 \rceil \tag{3.23}$$

ensures that $P_{sns}(\omega_1) \geq P_{desired}$.

3.2.4 Reliable function approximation

We do not know the true properties of the statistical functions in equations 3.1 and 3.4 necessary to calculate the a posteriori probability given by equation 3.2. Thus, finding approximations for these functions is one of the key issues when we design the classifier in our medical decision support system. As mentioned in section 3.1, we want our system to be *reliable* which in [107] is defined as follows:

> Reliable means that the measurement method produces similar results, irrespective of who measures it or when - assuming that the quantity being measured is static.

We seek to make our system reliable by instigating the following criteria on our classifier:

- *Generality*: The system performance obtained on our data should be reproducible in similar data obtained elsewhere.
- *Restriction*: New data out of the range defined by the data used to design the classifier should be rejected by the classifier.

In the remainder of this section, we will present and discuss methods by which to achieve fulfillment of the above given criteria. We will show how this can be done by using cross-validation techniques to evaluate the system performance. In the discussion of the technique for function approximation, we will focus on identifying the parameters having influence on generality and restriction.

Training and testing

We discuss a method by which to evaluate the degree of generality of a classifier.

The set of feature vectors in V (equation 3.13) is split into a *training set*, V_{train}, and an independent *test set*, V_{test}. The functional approximations for the a posteriori probability functions given by equation 3.6 are derived from the training set. A reclassification, which is a performance evaluation of the classifier based on the training set, gives the *training performance* of the classifier. Classification of the test set gives the *test performance*. Generality is achieved if the test performance approaches the training performance.

There are parameters in the function approximation scheme that affects generality. Therefore different parameter settings are investigated to identify the setting which corresponds to the highest performing general classifier; i.e. select the highest performing classifier among those satisfying the criterion that the test performance meets the training performance to within a given tolerance.

One of the problems in the training and testing of classifiers, is the limited amount of data. We want our functional approximation to be accurate. The feature space should be densely populated by the training set. One might

consider using valuable data for testing as wasteful. On the other hand, we wish to know the performance to some degree of resolution to be able to distinguish reliably between different classifiers. Cross-validation (sometimes referred to as jack-knifing) offers a method by which to get the most out of a scarce data set. And our data are indeed scarce.

Cross-validation

Performance evaluation by cross-validation [87, 80] handles the problem of scarce data sets by dividing the vector set in V into S_t subsets by dividing each of the K class specific sets into S_t equally sized subsets $V_{i,j}, i = 1, \ldots, K, j = 1, \ldots, S_t$ and using $S_t - 1$ of these subsets for training and 1 for testing. This is done S_t times, thus getting the performance characteristics for each classifier. In the evaluation of the kth classifier,

$$\begin{array}{l} V_{test} = \cup_{i=1}^{K} V_{i,k} \\ V_{train} = \cup_{i=1}^{K} \cup_{j=1, j\neq k}^{S_t} V_{i,j}. \end{array} \tag{3.24}$$

We define the performance matrices of the cross-validated training and testing results to be

$$\mathbf{P}_{train} = \begin{bmatrix} P_{sns}(\omega_1)_1 & P_{spc}(\omega_1)_1 \\ P_{sns}(\omega_1)_2 & P_{spc}(\omega_1)_2 \\ \cdots & \cdots \\ P_{sns}(\omega_1)_{St} & P_{spc}(\omega_1)_{St} \end{bmatrix} \tag{3.25}$$

and

$$\mathbf{P}_{test} = \begin{bmatrix} P_{sns}(\omega_1)_1 & P_{spc}(\omega_1)_1 \\ P_{sns}(\omega_1)_2 & P_{spc}(\omega_1)_2 \\ \cdots & \cdots \\ P_{sns}(\omega_1)_{St} & P_{spc}(\omega_1)_{St} \end{bmatrix} \tag{3.26}$$

respectively. These performance matrices contain the performance characteristics by which we want to evaluate our classifier performance (the kth row represents the performance of the kth classifier in the cross validation. We will typically represent the performance of a classifier by estimates of the mean and standard deviation of its test performance matrix:

$$\mathbf{M}_{test} = \frac{1}{St} \sum_{i=1}^{St} \begin{bmatrix} P_{sns}(\omega_1)_i & P_{spc}(\omega_1)_i \end{bmatrix} = \begin{bmatrix} \mu_{sns} & \mu_{spc} \end{bmatrix} \tag{3.27}$$

and

$$\mathbf{S}_{test} = \left[\sqrt{\frac{1}{St-1}\sum_{i=1}^{St}(P_{sns}(\omega_1)_i - \mu_{sns})^2} \quad \sqrt{\frac{1}{St-1}\sum_{i=1}^{St}(P_{spc}(\omega_1)_i - \mu_{spc})^2} \right]. \tag{3.28}$$

$\mathbf{M}_{train}$ and $\mathbf{S}_{train}$ are similarly defined from the training performance matrix.

Classifier generality

When the difference between the average performance of training sets and corresponding testing sets is significant, this is caused by *over-training*. The classifier has to be *generalised*. This is done by adjusting the parameters of the classifier and choosing the parameter setting so that generality is achieved. If the parameter set controlling generality and restriction is denoted α, a classifier is tested for generality for each possible combination of parameter elements. If the total number of such combinations are G, the generality train and test matrix is the concatenation of the mean performance matrices:

$$\mathbf{G}_{train} = \begin{bmatrix} \mathbf{M}_{train_1} \\ \mathbf{M}_{train_2} \\ \cdots \\ \mathbf{M}_{train_G} \end{bmatrix}. \tag{3.29}$$

and

$$\mathbf{G}_{test} = \begin{bmatrix} \mathbf{M}_{test_1} \\ \mathbf{M}_{test_2} \\ \cdots \\ \mathbf{M}_{test_G} \end{bmatrix}. \tag{3.30}$$

A generality test matrix is defined as

$$\mathbf{T}_{gen} = |\mathbf{G}_{train} - \mathbf{G}_{test}|. \tag{3.31}$$

The classifier we seek corresponds to the ith parameter combination corresponding to the highest overall test performance according to

$$\arg\max_i \left(\frac{1}{2}(\mathbf{M}_{test_{i,1}} + \mathbf{M}_{test_{i,2}})\right), \tag{3.32}$$

satisfying the tolerance criterion

$$(\mathbf{T}_{gen_i,1} \leq \tau) \cap (\mathbf{T}_{gen_i,2} \leq \tau), \tag{3.33}$$

τ being the generality tolerance. Thus the classifier chosen satifies the following criteria:

1. The test and training performances are equal to within the tolerance limit for both the averaged sensitivity and specificity.

2. Amongst the classifiers satisfying the first criterion, the one with the maximum overall test performance is chosen.

Uncertainty in estimates of performance characteristics

The number of observations, M determines the uncertainty in the resolution of the performance characteristics. In [80] the accuracy of a given error rate P is given with a resolution of ΔP (95% confidence interval) according to

$$2\sqrt{P\frac{0.95}{M}} \approx \Delta P. \tag{3.34}$$

In our experiments with human out-of-hospital data in chapter 8 we have a total of 868 observations, 87 in the successful group and 781 in the unsuccessful group. We want to calculater the sensitivity to around 95% which corresponds to an error rate of 5% for the successful group. Thus the accuracy for the sensitivity, ΔP_{sns} is around 5%. For the unsuccessful group, the specificities varies from around 10 to 45% with the equivalent error rate varying from 55 to 90%. Thus, ΔP_{spec}, for this group lies in the 5 to 7% range.

In the following discussion of function approximation techniques we will identify the parameters influencing generality.

3.2.5 Approximation techniques

The statistical functions of equations 3.6 have to be estimated to determine the decision rules for the classifiers. Different means for doing this involves a diversity of methods amongst which are histogram-, neural network- and radial basis approaches. In our work we have used a histogram approach. As mentioned in section 3.2.4 we will identify the parameters influencing generality and restriction of the classifier, so that we will have established the means for designing a reliable decision support system for VF-analysis.

Smooth histogram

The multidimensional smooth histogram approach discussed here estimates the class specific PDFs of $\mathbf{v}$. Each feature axis is partitioned in n_b equidistant segments, thus defining n_b^D equal sized bins in the D-dimensional feature space. First a local PDF is estimated for each class within each bin. Secondly, multidimensional gaussian kernel functions are applied to each bin to interpolate a continuous estimate from the available training data. A local estimate for the class specific PDF is given as

$$\hat{p}(\mathbf{v}|\omega_j) = \frac{n_{i_{\omega_j}}}{n_{\omega_j}} \tag{3.35}$$

$n_{i_{\omega_j}}$ denotes the number of measurements $\mathbf{v}$ belonging to class ω_j contained in histogram bin number i. n_{ω_j} represents the total number of measurements belonging to class ω_j.

The a priori probability estimates are calculated according to

$$\hat{P}(\omega_j) = \frac{n_{\omega_j}}{n}, \tag{3.36}$$

where n is the total number of measurements. The a posteriori probability estimate $\hat{P}(\omega_i|v)$ is derived by using equations 3.35 and 3.36 according to Baye's rule given in equation 3.2.

Using the estimates described above, the performance characteristics estimates, $\hat{P}_{sns}(\omega_i)$ and $\hat{P}_{spc}(\omega_i)$ are computed according to Equations 3.14 and 3.15.

The gaussian kernel functions controls the interpolation, and in the following we will discuss how adjusting the parameters of the kernel functions will influence the *smoothness* of the estimate. As we will show, we will control generality through varying the smoothness of the estimate.

The estimate smoothing is performed with a gaussian kernel function. But first, the feature vector values are normalised to integers in the range 1 tp n_b. This is done to let the kernel function operate symmetrically on a multidimsional matrix. We start by defining some quantities to determine the range to be used of each of the feature axes: $vmin_i$ and $vmax_i$ define the midpoints of the first and last bin of feature axis number i. $\mathbf{v}_n$ is normalised from $\mathbf{v}$ according to calculating each of the elements $v_i, i = 1 \ldots D$ according to

$$v_{n_i} = round(\frac{n_b - 1}{vmax_i - vmin_i} v_i + 1 - \frac{n_b - 1}{vmax_i - vmin_i} vmin_i). \tag{3.37}$$

If we consider the local bin PDF as a volume with probability mass P, we want to distribute this probability mass into its neighborhood bins so that the probability decreases with increasing distance to the volume in the central bin. The probability of the distributed mass should be equal P. One way to achieve this is to apply a gaussian kernel function. We let $\mathbf{v}_n$ and $|\mathbf{d}|$ denote the neighborhood bin and its distance from the central bin respectively. The probability distributed to $\mathbf{v}_n$ from the central bin is computed according to

$$\mathbf{z}(\mathbf{d}) = C \exp(-\frac{|\mathbf{d}|^2}{2\sigma_{v_n}^2}), \tag{3.38}$$

where C is a constant and σ_{v_n} is the kernel width governing the degree by which each bin is spread out. If we choose C equal to

$$\frac{1}{\sqrt{2\pi^D \sigma_{v_n}^D}}, \tag{3.39}$$

equation 3.38 corresponds to the normal distribution function with probability mass 1 distributed all over the feature space. The contribution from the central bin $\mathbf{c}_n$ with probability mass $P = \hat{p}(\mathbf{c}_n|\omega)$ into $\mathbf{v}_n$ is

$$\hat{p}(\mathbf{c}_n|\omega)\mathbf{z}(\mathbf{d}). \tag{3.40}$$

Thus, the PDF estimate in every bin is spread out into feature space. We define the use of $\sigma_{v_n} = 0$ to correspond to direct use of the histograms without smoothing. A restriction may be set, so that a bin does not spread further than a restrition factor ρ. We have set this value so that the gaussian kernel function is reduced to 95 % of its total volume. The constant C is accordingly divided by ρ so that the restricted kernel function retains probability mass 1.

Thus, we have adopted histogram smoothing so that generality in our decision support system is governed by the number of bins n_b and the kernel width, σ_{v_n}, and restriction is affected by ρ.

In our experiments in chapter 8 we will focus on generality, thus making n_b and σ_{v_n} the key parameters of the classifier when utilizing histogram smoothing. As we will see, a small number of bins provides low histogram resolution, while a large number of small bins provides high histogram resolution. Each feature axis of the PDF estimate is divided into n_b intervals. Thus, if the feature dimension is D for a specific feature combination, the feature space is divided into n_b^D bins of equal volume. Smoothness is governed by the kernel function width. A narrow kernel function provides a high resolution estimate with high variance, while a wide kernel function provides a smoother low resolution estimate with low variance.

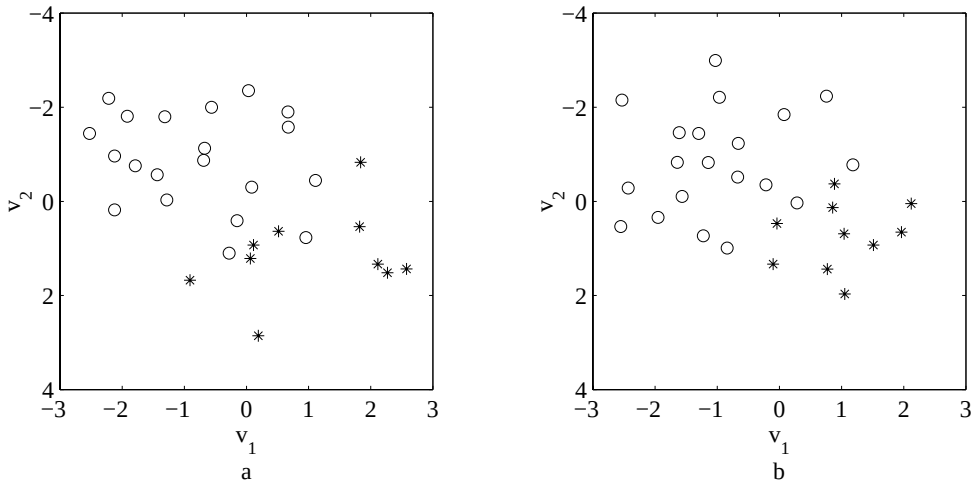


Figure 3.4: Data used in design of example classifiers. The data from classes ω_1 and ω_2 are plotted with stars and circles respectively. a) Training data, b) Test data.

Example

To illustrate the working principles of our classifier we have designed a simple experiment. We have sampled a small number of feature vectors from two 2-D gaussian distributions and split the data into a training and testing set. The data in the training and test sets is shown in figure 3.4 a) and b) respectively. The data from classes ω_1 and ω_2 are plotted with stars and circles respectively. The first step in the smooth histogram technique is to divide the feature space into equal sized bins and compute the class specific PDF estimates within each of these bins from the training data. The number of bins affect the estimate and thus the performance and generality of the resulting classifier. In figure 3.5 we have tried to illustrate this by using many and few bins in classifier designed using the training data shown in figure 3.4.

In part a) of the figure we have used $n_b^2 = 4$ bins. Each bin contains data, but some of the bins are partly uninhabited. The resulting decision regions are shown in part b) of the figure. Performance for this classifier corresponds to a sensitivity for class ω_1 of 100% both in training and testing with corresponding specificities of 55and 60%. From these results the classifier seems to be general, but as is evident from the plot, the decision border runs parallell to the feature axes. It would be easy to manually draw a border which would give better performance. Using as few bins as four provides poor resolution,

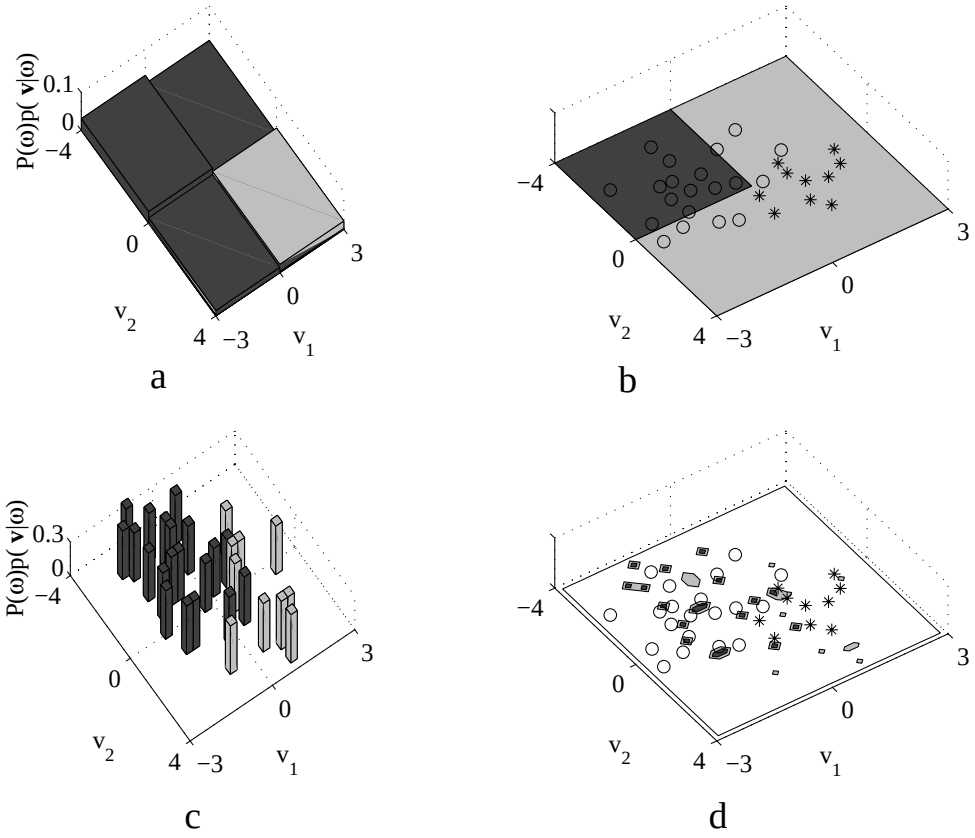


Figure 3.5: Designing histogram classifiers without smoothing ($\sigma_{v_n} = 0$). The PDF estimates and decision regions corresponding to classes ω_1 and ω_2 are plotted with light and dark gray respectively, while the test data are shown with stars and circles as in figure 3.4. a,c) Class specific PDF estimates using $n_b = 4$ (a) and $n_b = 25$ (c). b,d) Resulting decision regions for using $n_b = 4$ (b) and $n_b = 25$ (d) with testing data superimposed.

with the borders between the two decision regions being piecewise parallell to the coordinate axes. This causes an unnecessary drop in the specificity with increased sensitivity.

In part c) of the figure, $n_b^2 = 25^2 = 625$ bins. As the data points in feature space is sparsely populated, the number of vectors within each bin are scarce (Typically less than two). Therefore the estimate consists of isolated histogram columns. The effect of this on the resulting classifier is shown in part d) of the figure. The decision regions become small isolated regions around each of the histogram columns. The classification performance corresponding to this is near ideal in reclassification; 100 % in both sensitivity and specificity. That this classifier is over-trained becomes evident when we consider the testing sensitivity of 20 %. This is caused by a lot of testing data falling outside the decision region bins and into the rejection area (white), thus giving poor generality.

We use smoothing on a larger number of bins to improve on this situation. Two smoothed class specific PDF estimates and corresponding decision regions (trained for 95% sensitivity) are shown in figure 3.6. The PDFs are smoothed from the histogram with 625 bins shown in figure 3.5 c) with varying kernel widths. Part a) of the figure shows how smoothing with a narrow kernel width $\sigma_{v_n} = 1.5$) affects the estimate. The most isolated histogram colums are smeared into restricted surfaces with gaussian bell shapes. In the areas of feature space were the histogram columns are placed more densely, the smoothed bells are summed, providing a smoother surface. The corresponding decision region is shown in part b) of the figure. Performance for this classifier corresponds to a sensitivity of 100% in training and 60% in testing with corresponding specificities of 87and 90%. If we compare to the corresponding classifier without smoothing, generality in sensitivity has improved although not satisfactorily with a deviation of 40 % between testing and training. The classifier is still over-trained, although not as much as without smothing.

Further smoothing by larger σ_{v_n} affects the PDF and decision region as illustrated by part c) and d) in the figure. The isolated training data points are no longer evident from the smoothed PDF. This is caused by the the larger kernel width causing contributions from several training data sources on most places on the function surface. The classifier corresponding to these parameter settings evidences a higher degree of generality and better performance as is evident from a sensitivity of 96% in training and 90% in testing with corresponding specificities of 80and 85%. It is interesting to note the border between the decision regions appearing less training data specific.

The effect of using a restriction factor is illustrated in the PDF plots by the fact that the columns are spread into a limited area of its neighborhood. Therefore,

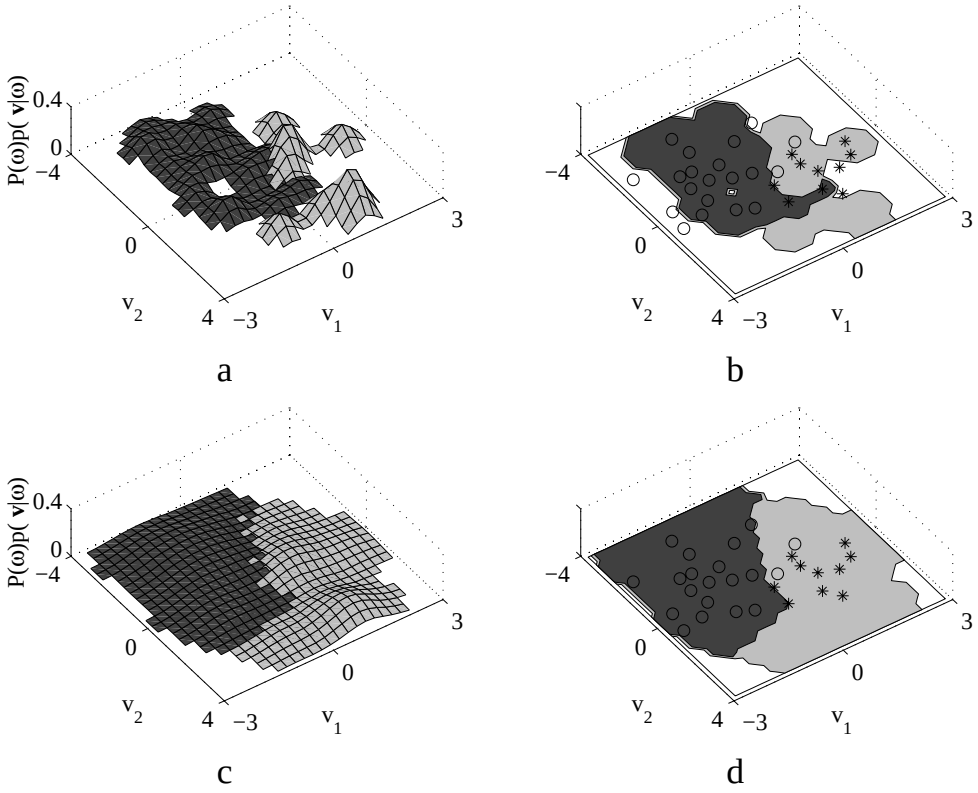


Figure 3.6: Designing histogram classifiers ($n_b = 25$) with varying kernel widths. The PDF estimates and decision regions corresponding to classes ω_1 and ω_2 are plotted with light and dark gray respectively, while the test data are shown with stars and circles as in figure 3.4. a,c) Class specific PDF estimates using kernel widths of $\sigma_{v_n} = 1.25$ (a) and $\sigma_{v_n} = 5$ (c). b,d) Resulting decision regions using $\sigma_{v_n} = 1.25$ (b) and $\sigma_{v_n} = 5$ (d) with testing data superimposed.

the parts of the feature space not inhabited by feature observations are not extrapolated. In the plots of the decision regions these areas appear as white, thus defining the reject class.

Having defined this important part of our system we proceed by discussing the feature extraction part.

3.3 Feature extraction

As we briefly discussed in section 3.1, developing good feature extraction methods is one of the key issues in the design of our VF-analysis decision support system. To put it bluntly, we seek a set of ECG-derived features expressing the resuscitability of the subject from whom the ECG was measured. In other terms, we want our system to be *valid* which in [107] is defined as follows:

> Validity assesses how much the measurement result reflects what it is intended to measure, for example, diagnostic accuracy. To be sure that a measurement is valid, it must be compared to a gold standard,...

For high validity we need features yielding high performance classification results. In the following we will after a brief discussion of feature extraction by methods for signal representation, discuss methods by which we will try to meet this end. At the end of this section we will discuss strategies to select the best features from a larger feature set.

3.3.1 Feature extraction in general

In the present setting the ECG is the primary measurement. We start off by establishing nomenclature:

- *record*: The entire ECG sequence recorded from a patient.
- *block*: The record is split into blocks which may be overlapping. For each block one feature vector is calculated.
- *segment* The block can be further subdivided into segments which may be overlapping.

In the case of ECG, a block is typically of very high dimension (several hundred). This makes the formation of a feature vector directly from the ECG block impractical. Therefore we seek to restructure the signal so that it will be better suited for feature extraction for classification.

3.3.2 Spectral feature extraction

Features based on estimates of the PSD are central in some of the earlier work discussed in section 2.3. We look into some of these methods and also introduce some new methods or variations of older ones.
Signals with varying properties may be characterised in the frequency domain by the *power spectral density* (PSD). Often, the PSD is estimated by averaged periodograms,

$$\hat{P}_{AV}(f) = \frac{1}{K}\sum_{m=0}^{K-1}\frac{1}{L}\mid\sum_{n=0}^{L-1} w(n)x_m(n)e^{-j2\pi fn}\mid^2, \tag{3.41}$$

where $x_m(n)$ denotes sample n in segment m in a block of ECG. Each block is divided into K segments, each of length L. Both the segments as well as the blocks may overlap. Each segment is weighted by a window function $w(n)$. The properties of the averaged periodogram is discussed in detail in [56].
Studies of the spectra for varying myocardial metabolism indicate changes in shape and placement in the frequency region. A smeared spectrum indicates high myocardial perfusion while a peaky spectrum indicates low myocardial perfusion. This suggests that a parameter that carries information about the shape of the PSD may be used as a monitor to detect changes in the myocardial metabolism [12].
In one of the later experimental chapters 5.2, the spectrograms from some animals give an impression of how this might seem intuitively correct. These kinds of observations might have inspired the introduction of some of the spectrally derived features we discuss in the following.

The centroid frequency

The observed shift of the centroid of the area under the PSD curves due to the therapy can have been the motivation for the introduction of the centroid frequency (CF) as an indicator for changes in the myocardial metabolism.
The *centroid frequency* (CF), commonly referred to as the median frequency [12], indicates placement by using the frequency bisecting the area under the PSD and is given as

$$CF = \frac{\int_{f_l}^{f_u} f\hat{P}_{AV}(f)df}{\int_{f_l}^{f_u}\hat{P}_{AV}(f)df}, \tag{3.42}$$

where $f_l \leq f \leq f_u$. f_l and f_u are the lower and higher frequency band limits respectively. By varying these limits, we can study the effect of extracting features from different frequency bands.

The peak power frequency

The shift in the placement of frequency components can also be captured by the frequency corresponding to the maximum power in the PSD as given by

$$PPF = \arg\max_{f}(\hat{P}_{AV}(f)). \tag{3.43}$$

Again, the frequency band wherein PPF is to be determined is limited by f_l and f_u.

The spectral flatness measure

Motivated by observations of changes of spectral shape in accordance with with changes in resuscitability, we introduce of a spectral shape parameter. The *spectral flatness measure* SFM [55] is suitable due to its ability to describe the flatness of the spectrum. The SFM is given by

$$SFM = 2\frac{e^{\int_{f_l}^{f_u} \ln \hat{P}_{AV}(f)df}}{\int_{f_l}^{f_u} \hat{P}_{AV}(f)df}, \tag{3.44}$$

SFM attains a value in the region between zero and one. A flat spectrum renders a spectral flatness measure of 1, while peakier spectra attains values closer to 0.

The spectrum energy

Various time domain measurements of signal amplitude characteristics have been investigated [109, 62, 73], some of these efforts are described in chapter 2.3. In this work we investigate an alternative frequency band limited energy measurement (ENRG) as given by

$$ENRG = \int_{f_l}^{f_u} \hat{P}_{AV}(f)df. \tag{3.45}$$

As we will see in the later analysis of animal experimental data, trend curves of the spectral features depicts how the ECG measurements change along with alterations in therapy affecting the myocardial metabolism.

3.3.3 Feature selection

The dimension of a feature vector may be larger than necessary because the individual features are correlated. It is sometimes convenient to reduce the dimension before presenting the feature vector to the classifier. The problem is to determine the subset or combination of features giving the best classifier performance. There are many strategies by which to determine this subset, but in this work we only use the straightforward optimal method. This is a time-consuming method, but we only handle feature vectors of limited dimensions in this work.

Optimal feature selection

A lot of methods have been proposed to determine the best subset. The essence of these methods is that features are evaluated individually and ranked according to a class separability criteria. The straightforward method is by comparing all possible combinations and determining the one best suited for the classification task. Of course, if the dimension is large, this is quite exhaustive.

Feature projection

Alternatively by selecting the best subset, projection pursuit, determines the best dimension reduction $N \mapsto M$ by combining the features by some criterion. As discussed in [44], principal component analysis using a criterion of best representation has advantages as compared to methods based on alternative criteria. Principal component analysis (or principal axis representation) is presented in detail in appendix A.

3.4 Filtering of CPR artefacts

The appearance of CPR artefacts in the ECG affects the validity of the feature measurements. If these artefacts are present during VF-analysis, the extracted information will be infected with noise. Other investigators have demonstrated that CPR artefacts causes an interference problem during VF-analysis on human ECG [94]. Therefore it is necessary to handle this problem to perform reliable VF-analysis during CPR.

In chapter 3.1 we discussed the use of filters in reducing artefacts caused by precordial compressions and/or ventilations. In the following we briefly discuss

the artefact removal methods used in earlier work. As these methods has serious limits in the case of human VF-analysis, we end this section proposing a method for CPR artefact removal using adaptive filters whereby VF-analysis is made possible during CPR in humans.

3.4.1 Filtering CPR artefacts in animal ECG

Frequency selective analysis was applied by Strohmenger et al in an animal experiment [95] to study the interference of CPR-artefacts to VF-analysis. Noc et al used a frequency selective high pass filter for artefact removal prior to VF analysis in an animal study [73].

Frequency-band limiting

The method applied by Strohmenger et al in [95] corresponds to setting the lower and upper frequencies, f_l and f_u, of the frequency region to be analysed . The least interference to VF-analysis was achieved by setting the frequency band to be analysed according f_l set between the third and fourth harmonic of the CPR artefact and f_u set higher than the highest significant frequency components of the VF signal. This effectively removed the dominant artefact components in the spectrum. As we will show later, some of our experiments on animal data will be performed according to this method.

Digital filtering with fixed coefficients

Noc et al used a digital high pass filter which similarly to the frequency-band limiting method of Strohmenger et al removed the three lowest harmonic components of the artefact prior to VF-analysis. For the digital filtering we will be performing on animal experiment ECG to split the ECG we will apply high pass and low pass digital FIR filters. The high pass and low pass filters will be used for the VF and CPR channels respectively. A general form for a digital filter is given by:

$$H(z) = \frac{b_0 + b_1 z^{-1} + \ldots + b_{n_b-1} z^{-(n_b-1)}}{1 + a_1 z^{-1} + \ldots + a_{n_a-1} z^{-(n_a-1)}} \tag{3.46}$$

The constellation of the coefficients define the type of filter [77] in this work we will use finite impulse response FIR filters given by $n_a = 1$ and $n_b \geq 0$. Thus the filter's behavior is governed by the choice of coefficients $b_i, i = 0, \ldots, n_b - 1$.

3.4.2 Adaptive filtering of CPR artefacts in human ECG

As we discussed in section 3.1, the success of the frequency-selective methods depends on the frequency components of the VF signal and the CPR artefact signal being non-overlapping. This criterion is satisfied in the animal data analysed by Strohmenger et al [95] and Noc et al [73].

In an attempt to solve this problem, we have put forward the following strategy[63, 1]: The ECG artefact is modeled as a sum of several sources; the main source being the mechanical stimulation of the heart itself. For each source the idea is to obtain a *reference signal* correlated to the associated artefact component. In the case of artefacts due to chest compressions one such reference is the compression depth. The reference signal facilitate removal of the artefact component using an adaptive, digital filter. We investigated the potential usefulness of this model in [63] and [1]. The fundamentals of the filter solution developed by Aase et al [1] is described below.

Filter considerations

In designing the best possible filter strategy for removing artefacts we make use of the following observations:

- The reference signals are correlated to the associated artefact components.
- The system should combine information inherent in the multiple references in an optimal manner.
- It should be expected that the causal relationship between cause (chest compression/ventilation) and the associated artefact components will change over time. It follows that the filters used should be adaptive, using *short-time signal analysis* when computing the current filter coefficients.

Letting P denote the number of artefact-causing sources, the artefact signal can be modeled as the sum of the effect of each source v_p:

$$y(n) = \sum_{p=1}^{P} \sum_{k=0}^{K_p-1} h_p(k) v_p(n-k). \tag{3.47}$$

In this expression the source signals $\{v_p(n)\}$ are identical to the reference signals, and $h_p(n)$ is the unit pulse response of the Finite-Impulse-Response

(FIR) filter modeling the impact of v_p on y. K_p is the corresponding filter length. $y(n)$ is an estimate of the actual artefact $a(n)$.

If we set $P = 2$, the structure of the proposed multi-channel filter algorithm is as shown in Figure 3.7.

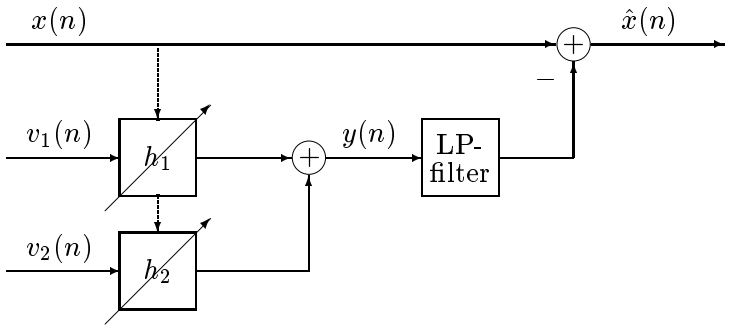


Figure 3.7: Structure of artefact removal system when using 2 reference signals. $\hat{x}(n)$ is the restored VF/VT signal.

The proposed algorithm also includes a simple low-pass filtering of the reconstructed artefact. This is a practical way of removing spurious frequency components above 10 Hz. Such components may arise due to the sample-to-sample adaptivity of the channel filters h_1 and h_2. The dotted arrows indicate that $x(n)$ is used for analysis when computing the filter responses h_1 and h_2.

Further details on the derivation of this filter solution is given in appendix B, while experimental results are presented and discussed in chapter 7.

3.5 Summary

We have discussed and presented methods for CPR artefact removal both in animal and human ECG. Frequency selective filters are applicable to animal ECG because the frequency components of VF and CPR are separable. This assumption is not valid for human data, and we propose a method for removing CPR artefacts from human ECG using adaptive filters with references containing information about the cause of the CPR artefacts.

We have also discussed the feature extraction methods to be used in later chapters.

Finally, we have shown how a statistical classifier can be fitted into our decision support system, with special focus on the ability to perform ROC-analysis.

We have presented methods for measuring performance by cross-validation techniques, and we have presented the function approximation technique we will be using and have identified the parameters governing *generality* and *restriction*.

Chapter 4

ECGs, demographics and annotations

The data considered in this work includes ECG and related information. The EMS database is our primary object for prediction and monitoring analysis, while the AED rhythm library is used to evaluate CPR artefact removal. To illustrate the basic concepts of CPR artefact removal, shock outcome prediction and monitoring analysis we study data from animal studies. Two of these, the short-term and long-term drug studies were originally done to evaluate the effect of drug therapy during CPR. In the third animal study from which we have data, one of the study objectives were the application of precordial compressions at different frequencies.

4.1 Human data

The human data includes the EMS database of ECGs, annotations and demographics from cardiac arrested patients and the AED rhythm library of ECGs covering a diversity of cardiac arrhythmias usually encountered by AEDs.

4.1.1 The EMS database

The EMS database was collected in Oslo in the two-year period between February 19th 1996 and February 18th 1998. The author of this thesis joined forces with researchers working closely with the EMS system to facilitate the establishment of a database of ECG recordings and demographic data from out-of-hospital cardiac arrested patients. Approval for this study was obtained

through the Regional Committee for Research Ethics, health region III, and the Norwegian Data Inspectorate.

The data were collected as part of an observational prospective study of what actually occurred during defibrillation and advanced life support (ALS) in the Oslo emergency medical system [100]. This subsection is adapted from the materials description of [100].

The study was conducted with a target land area of 427 square km and a population of approximately 500000; 1170 persons per square km. 48% of the population are men and 16% are aged 65 or more. The death rate was 1096 per 100000 (1997) with 301 of these being of cardiac origin (1995).

A one-tiered centralised community run system responds to all medical emergencies on a designated telephone number. All ambulances are dispatched from one centre staffed with specially trained paramedics and nurses.

There are eight to 11 emergency ambulances between 07:00 and 22:00 h and six ambulances at night. They are all fully equipped for advanced life support (ALS) and from 07:30 to 22:00 on weekdays one of the teams includes an anesthesiologist. The paramedics perform ALS according to the ERC guidelines [16, 81] and are certified yearly.

If cardiac arrest is suspected, the dispatch centre will, if available, send the M.D. manned ambulance. If a regular ambulance is dispatched, the paramedics, after confirming cardiac arrest, request the assistance of a second ambulance team. The ambulances are equipped with Heartstart 3000 defibrillators. Maximum emergency cross-town driving time is about 30 min.

For the study, the following was collected from cardiac arrests where ALS was attempted :

- The log file in the Medical Control Module (MCM) of the defibrillator (Heartstart 3000 (Laerdal Medical)).
- Regular Utstein registration of demographics 'Report of confirmed prehospital cardiac arrest' [17].

The MCM stores up to 20 min of the digitised ECG signals and all the logdata describing the use of the defibrillator. If a resuscitation attempt exceeds 20 min, new prioritised ECG segments (analysis, defibrillation and monitoring once each min) are recorded overwriting the nonprioritised segments last recorded. The ECGs were reassembled in its correct time frame from the information in the MCM.

Printouts of the ECG tracings along with the MCM log data were generated for each patient episode for systematic review by the medical team providing rhythm annotations. The ECG is sampled at 100 samples per second and 8 bit resolution. The EMS database thus consists of ECG, defibrillator log data, rhythm annotations and demographics from 156 out-of hospital cardiac arrest treated patients. Figure 4.1 gives an example of a printout of ECG with corresponding defibrillator log data.

Relevance

This EMS database provides data from the treatment of cardiac arrested out-of-hospital patients. This database is used in the analysis of how ALS is performed (chapter 6) and in the evaluation of shock outcome prediction (chapter 8) and CPR efficacy monitoring methods (chapter 9).

4.1.2 The AED rhythm library

The AED library (proprietary of Laerdal Medical AS, Stavanger, Norway) consists of 481 ECG segments each of 15 seconds duration. This material represents a variety of rhythms, representing what an AED would be expected to see (table 4.1).

Rhythm	Number of segments
Asystole (ASY)	26
Sinus rhythm (SR)	26
Ventricular tachycardia (less than 180 bpm) (NVT)	26
Pacemaker rhythm (PAC)	26
Premature ventricular complex (PVC)	26
Sinus tachycardia (S-T)	26
Supraventricular tachycardia (SVT)	27
Sinus bradycardia (less than 60 bpm) (V60)	27
Ventricular fibrillation (VF)	200
Ventricular tachycardia (VT)	71

Table 4.1: Representation of cardiac arrhythmias in the AED rhythm library

The ECG strips were selected from field data recorded by various Heartstart models (Laerdal Medical) and have been further annotated by three cardiologists. The data format is 100 samples per seconds, 8 bits per sample.

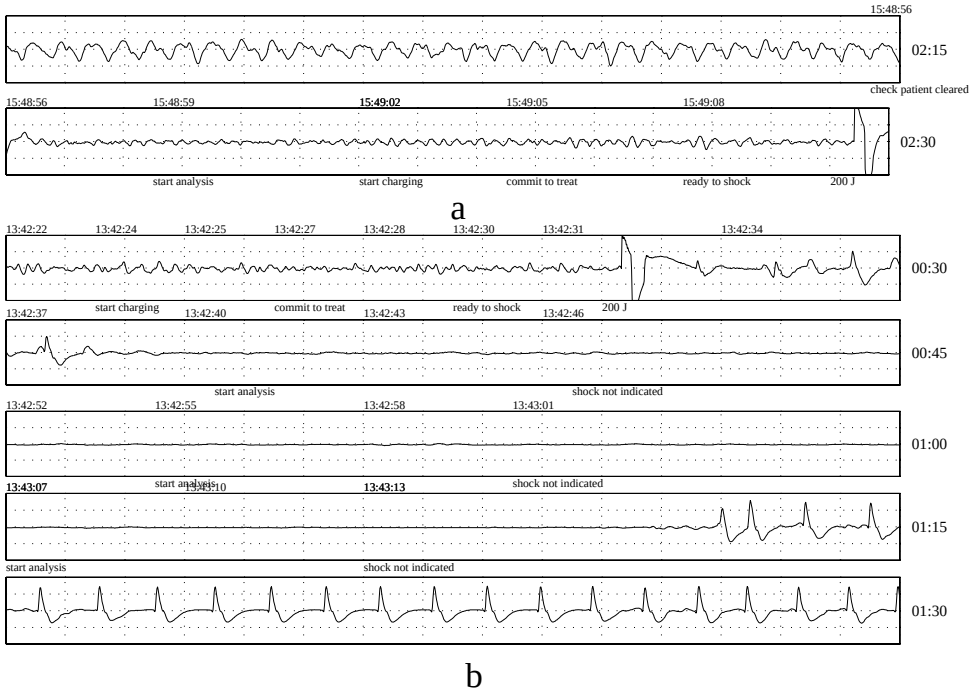


Figure 4.1: Example of MCM printouts of ECG and defibrillator log data translated from the MCM file. The minimum and maximum values on the y-axis correspond to -2.45 and 2.45 mV respectively. On the right side of each strip the time from initiation of monitoring is registered. a) In the first half of the strip 30 chest compression artefacts can be counted. The compressions are stopped at the start of the second part of the strip, the patient has VF and a shock is given 14 sec thereafter. b) In the first strip the patient has VF and a shock is given at 25 sec. Thereafter some electrical activity occurs before the patient has a long isoelectric period lasting for 37 sec. During this period the paramedics analyse the rhythm of the patient three times and do not initiate precordial compressions. Despite this, the patient obtains a pulsegiving rhythm 46 sec after the shock.

Relevance

In chapter 7 we use the AED library to evaluate the effect of CPR artefact removal on the recognition rate of VF and VT in the AED algorithm. The VF and VT segments are mixed with artefacts which we attempt to remove. We evaluate how well this is done by calculating

- Signal to noise ratio (SNR) improvement
- Sensitivity of the AED algorithm

for both filtered and uncontaminated segments.

4.2 Animal data

The animal ECG includes data from the short-time and long-term drug effect studies and the chest compression frequency study.

4.2.1 The short-term drug effect study

The animal short-term drug effect experiment ECGs originate from a porcine model used in an experiment by Strohmenger et al. The interference by CPR on VF was studied in [95].

Animal preparation

The animal preparation is described in detail in [95]. In short, the animals were anesthetised, intubated for ventilation and instrumentated for measurement of ECG and myocardial blood flow. Ventilation was performed at 20 breaths/minute. The ECG was monitored with standard lead II. The ECG sampling rate is 100 Hz with 8 bit resolution.

Experimental design

VF was induced in the pigs with ventilation stopped at the same time. At four minutes into VF, closed-chest CPR was started and continued for ten minutes using an automatic piston device. The compression rate was 80 per minute. At the same time ventilation was resumed.

After three minutes of CPR, the animals were randomly allocated to receive either vasopressin or saline without addition of any drug (placebo). Myocardial blood flow was measured at 90 seconds after initiation of CPR (before drug administration) and at 90 seconds and 5 minutes after drug administration.

After approximately 13 minutes of cardiac arrest, the animals were defibrillated to restore spontaneous circulation.

The time line of the protocol is illustrated in figure 4.2.

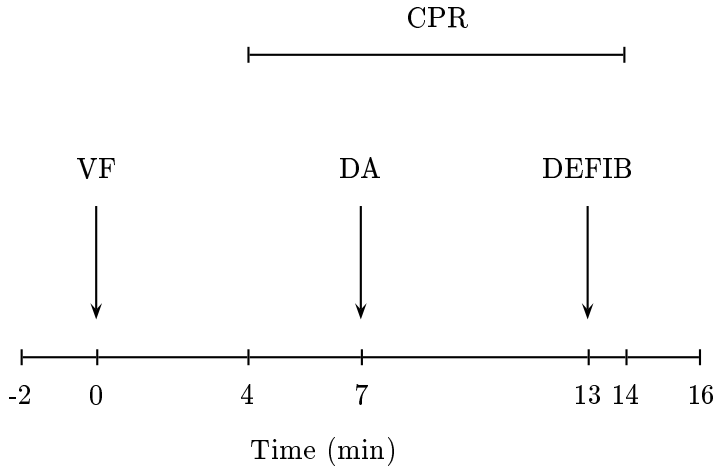


Figure 4.2: Protocol timeline showing the timings for activities during animal experiments. VF, inducing ventricular fibrillation; CPR, cardiopulmonary resuscitation; DA, drug administration; DEFIB, defibrillation.

Results

Resuscitation was 100% successful in the vasopressin group and 0 % successful in the placebo-treated group.

The measurements of the myocardial blood flow (taken from [95]) are shown in table 4.2. MBF was significantly higher in the vasopressin group than in

Time of measurement	Myocardial blood flow (ml/min/100g)	
	vasopressin	placebo
before drug administration (DA)	14.7 ± 2.1	17.5 ± 0.9
90 s after DA	74.5 ± 6.9	15.5 ± 0.9
5 min after DA	59.9 ± 9.6	13.0 ± 1.0

Table 4.2: Myocardial blood flow measurements

the control group after drug administration. The MBF was also significantly higher in the vasopressin group when comparing the measurements before drug administration to those after drug administration.

Relevance

In chapter 5.2 we analyse the ECG segments corresponding to the specific treatment phases. These segments may contain information reflecting the state of the myocardial metabolism. The myocardial blood flow (MBF) during CPR is indicative of the outcome of defibrillation. The higher MBF, the higher probability of a successful outcome following defibrillation (P(success)) [13, 98].

The relations between the phases of ECG and MBF of the animal experiments are as follows. In the start phase where ventricular fibrillation is induced and no treatment is given, the MBF stops immediately and prolonged periods of untreated ventricular fibrillation are associated with low defibrillation success (very low P(success)). With administration of CPR, the MBF increases immediately but hardly reaches the threshold of myocardial perfusion that is crucial to succeed in defibrillation therapy. This indicates a somewhat higher P(success) in this phase. With the administration of vasopressin, adequate myocardial perfusion is achieved for restarting the arrested heart and therefore a much higher P(success) can be expected. In [60] the relationship between MBF and ease of resuscitation is given according to table 4.3. The MBF measurements in this study and the MBF levels in table 4.2 indicates a link between the phase of ECG and the P(success) as given in table 4.4. We group the data into classes according to treatment phases given in table 4.4. The ECG features are evaluated according to their ability to discriminate between these classes.

Ease of resuscitability	MBF (ml/min/100g)
Not possible	0–15
Difficult	15–30
Easy	30–70

Table 4.3: Relationship between MBF and ease of resuscitation

Treatment	P(success)	MBF (ml/min/100g)
No treatment	very low	0
Vent. and compr. with no drugs	low to medium	13.0 ± 1.0–17.5 ± 1.0
Vent. and compr. with drugs	medium to high	$59.9. \pm 9.6$-74.5 ± 6.9

Table 4.4: Class division scheme

4.2.2 The long-term drug effect study

The ECGs from the long-term drug effect study ECGs originate from a porcine model. The data were kindly provided by Fritz Sterz [1]. The purpose of the study was to analyse the effects of different dosages of endotheline as compared to those of epinephrine and a placebo.

Experimental design

The animals were cardiac arrested, CPR started at five minutes into VF and continued for 20 minutes until defibrillation was attempted. Drugs were given during CPR at 10,13,16,19 and 22 minutes into VF. Six animals were given a placebo, six received epinephrine. Another six animals were treated with a small dosage of endotheline, and seven with a medium sized dosage, while five animals were given a high dosage. These animals received endotheline at 10 minutes and a placebo later on.

The time line of the protocol is illustrated in figure 4.3.

[1] Dr.med. Fritz Sterz, Stellvertreter des Vorstandes, a.o. Univ.Prof. und Facharzt fur Innere Medizin (Intensivmedizin), Notarzt, Universitatsklinik fur Notfallmedizin, Allgemeines Krankenhaus der Stadt Wien, Wahringergurtel 18-20/6/D, 1090 Wien, Austria. The analysis of these data discussed in this thesis was done in cooperation with M. Holzer, W. Behringer, F. Sterz, E. Oschatz, J. Kofler, P. Eisenburger and A. N. Laggner.

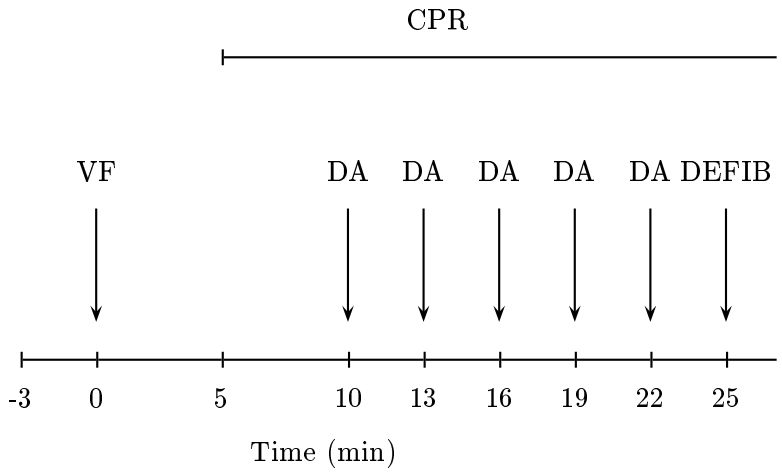


Figure 4.3: Protocol timeline showing the timings for activities during animal experiments. VF, inducing ventricular fibrillation; CPR, cardiopulmonary resuscitation; DA, drug administration; DEFIB, defibrillation.

Results

In the groups with low and medium dosages of endotheline, 50 % and 71 % of the animals where converted to ROSC respectively. In the other groups, none of the animals were successfully converted.

Relevance

In chapter 5.3, we analyse ECGs from this model to illustrate the effect of prolonged VF on ECG features. It is meant to be a supplement to the study of the short-term drug effect data (chapter 5.2). The analysis results show how the parameters used with good performance in the short-term study fails to reflect the myocardial metabolisms in long term VF. We also show how the introduction of a new parameter into the feature combination amends this problem to some degree.

4.2.3 The compression frequency study

The data from the animal compression frequency study was recorded in the study by Langhelle et al [63]. The purpose was to provide data for use in experiments with CPR artefact removal. The original intent was to provide data enabling us to evaluate candidate methods for artefact removal.

Animal preparation

The study was approved by the Norwegian Council for Animal Research. We give a brief description of the animal preparation. A more detailed description is given in [63]. Twelve domestic healthy pigs were anesthetised. Ventilations were given at 16 min^{-1}. The ECG signals were registered at 2000 samples per second at 32 bits per sample both through regular electrodes attached to the thorax and through electrodes attached to the limbs. These two channels are denoted the defibrillation (DEF) channel and the monitor (MON) channel respectively.

CPR was performed with a mechanical chest compression device with changeable compression rate. Aside from the ECG channels, signals measuring the compression displacement and for the last three animals in the study, the thoraic impedance, were recorded (sampling rate 25 samples per second).

Experimental protocol

Ventricular fibrillation (VF) was induced with ventilation discontinued at the same time. After 3 min of VF, chest compressions were initiated choosing one of three compression rates; 60, 90 or 120 min^{-1}, picked at random from a list. After 3 min. of CPR in each method, a 30 seconds pause with VF without CPR followed. Ventilation was performed after every fourth, every sixth and every eighth compression with 60, 90 and 120 min^{-1}.

The time line of the protocol is illustrated in figure 4.4.

After 180 sec the compression rate was changed to one of the other two patterns, and all measurements were repeated with the same time schedule. After 360 sec the rate was changed to the last pattern, and all measurements were repeated again. The randomisation list was written so that the different compression rates were spread uniformly between being performed as the first, second or third method. Thereafter asystole was initiated, and the same randomised procedure as with VF was repeated. After completion of the experiment CPR and ventilation were stopped.

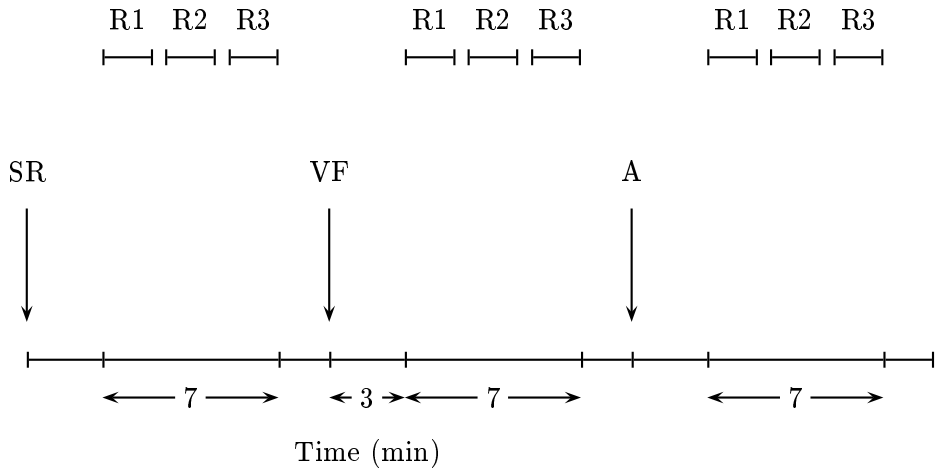


Figure 4.4: Protocol timeline showing the timings for activities during animal experiments. SR, sinus rhythm; R1, precordial compressions given at rate R1; R2, precordial compressions given at rate R2; R3, precordial compressions given at rate R3; VF, inducing ventricular fibrillation; A, inducing asystole.

Relevance

These data are used in our CPR artefact removal experiments. Observations on these data as well as simple filtering experiments are described in chapter 5.1. The recordings of asystole with CPR artefacts provide us with realistic data which we will use to simulate artefacts in human data. As we will see in chapter 7 the ECG recordings of the asystole sequences with artefacts from precordial compressions and ventilations provide data to be mixed with human ECG. In addition to this we have the information in the reference signals which will be used in the adaptive filtering technique described in chapter 3.4.

Chapter 5

Animal studies

In this chapter we report and discuss ECG observations from animal experimental data. Furthermore the basic concepts of CPR artefact removal are illustrated. We also perform simple experiments with shock outcome prediction and CPR efficacy monitoring. CPR artefact removal is discussed in section 5.1, while the potential of some of the ECG features for prediction and monitoring from chapter 3 are discussed in section 5.2 and 5.3.

5.1 CPR artefacts in animals

As discussed in chapter 3.1, CPR artefacts in ECG interfere with both outcome prediction and monitoring if left unhandled. In the following we report some observations regarding the effect of CPR on various rhythms. It is also illustrated how CPR artefacts in animal ECG can be separated from the VF part of the ECG.

5.1.1 Methods

The observations and experiments described in this section originate from analysis on compression frequency study data described in chapter 4.2.3. The human data originates from the AED rhythm library (chapter 4.1.2). Parts of this section are adapted from [63].

Observational data

From the animal ECG recordings from the study the following tracings were extracted:

- Three six second tracings recorded via the MON channel showing the transition from CPR to no CPR for SR, VF and asystole.
- Two 15 second tracings simultaneously recorded via both MON and DEF channel showing artefacts not simultaneously present in the two channels.
- From the VF sequences recorded from two of the animals (8 min duration), spectrograms were recorded using short time Fourier transforms computed from 3 second ECG blocks padded to 512 samples.

Filtering example

An attempt at splitting the ECG signal from one of the animals into a CPR channel and a VF channel was performed. A straightforward frequency selective filtering method was applied to split the ECG signal into a VF channel and a CPR channel. As was noted in [95], CPR interference may be obviated in spectral analysis if only the frequency band corresponding to 4.5-35 Hz is submitted to the analysis. This corresponds to the removal of the first three harmonics of the CPR artefact.

A high-pass filter was designed with the passband in the region above 5.5 Hz and a stopband attenuation of 20 dB and applied to the ECG to estimate the VF channel. A lowpass filter with passband in the region below 4.5 Hz and a attenuation of 20 dB in the stopband was also applied to estimate the CPR channel. The output signals were aligned with the original signal.

Comparing animal and human spectra

Frequency spectra were computed from ten 15 second ECG recordings of CPR artefacts (90 compressions per minute) randomly extracted from our animal experimental data. Likewise, spectra were computed from ten animal VF records and from ten human VF records from the AED rhythm library. All VF spectra were computed from periods when no CPR was performed, thus the VF spectra are without artefacts. All records were normalised to unity variance. The estimates were computed as the mean of the individual PSD estimates of ten 15-second ECG records.

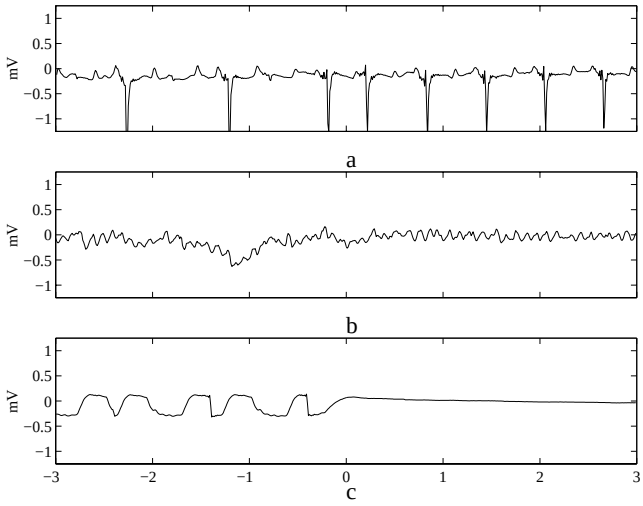


Figure 5.1: Different degree of disturbance in a) sinus rhythm (SR), b) ventricular fibrillation (VF) and c) asystole (A). The first and last three second periods are with and without ongoing CPR respectively.

5.1.2 Results

Observations on artefacts from compression frequency study

The CPR artefacts appear different during SR, VF and asystole (figure 5.1). During SR the only observed changes in the ECG with CPR are RR-interval modifications in the QRS-complexes during chest compressions. During VF and asystole, the magnitude of the artefact increases as the spontaneous electrical activity drops.

The simultaneous MON and DEF channel ECG recordings are shown in figure 5.2. We observe that there tends to be differences in artefact magnitudes and shapes between the monitoring channel and the defibrillator channel. During the first 7.5 seconds there are distinct artefacts present in the defibrillation channel that do not appear in the monitoring channel.

The spectrograms of the VF sequences are shown in (figure 5.3). In the two parts of figure 5.3, the spectrograms represent the period after onset of VF until right before asystole starts in two different animals. In figure 5.3 a), the frequency components corresponding to the harmonics of the CPR artefacts given at the three rates appear as distinct horizontal red lines in the 0-5Hz

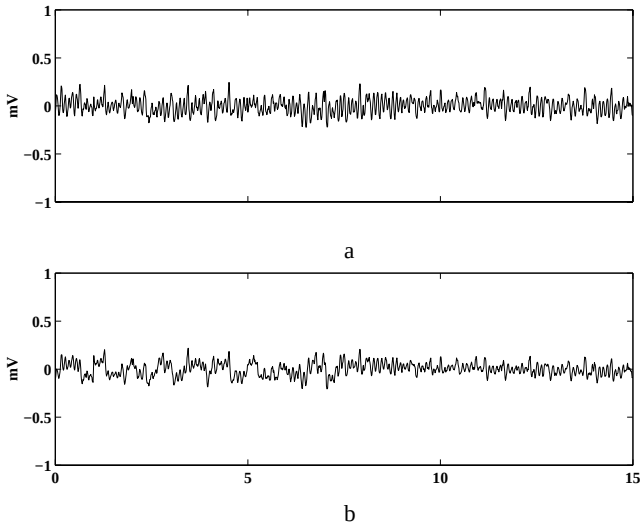


Figure 5.2: Example of difference in artefacts presented by a) MON channel and b) DEF channel. The artefact components present in the DEF channel are not simultaneously present in the MON channel.

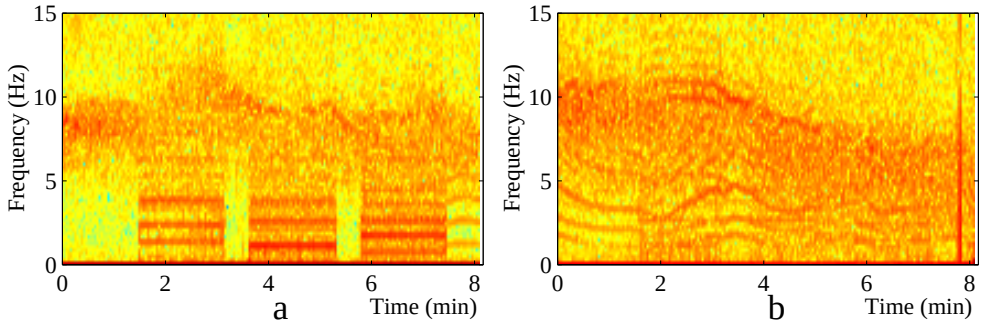


Figure 5.3: Time frequency representations (spectrograms showing the scaled logarithm of the magnitudes of the time dependent Fourier transform) of the VF period in two animals with a) low degree of spontaneous activity with prominent artefact components and b) high degree of spontaneous activity with less distinct artefact components. The colours identify the magnitudes as follows: low (dark blue), below middle (light blue), above middle (yellow), high (red). The frequency components of VF and CPR artefacts (and spontaneous activity) are identified by high magnitudes (red) in the frequency areas 5-15 Hz and 0-5 Hz respectively.

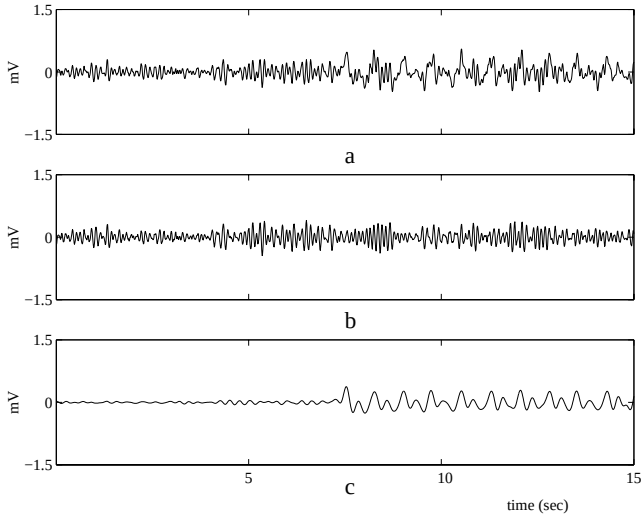


Figure 5.4: ECG showing transition from VF to VF with CPR artefacts. a) Original ECG, b) VF channel being the output high-pass filtered ECG c) CPR channel being the output of the low-pass filtered ECG

frequency band. The frequency components corresponding to VF appear as a time-varying composition of frequency components in the frequency area above 5 Hz. In the pauses between the first two compression periods the direct effect of CPR disappears as indicated by the yellow to green colouring without distinct bands.

In some animals spontaneous periodic activity appeared during VF and continued after the end of CPR. This is illustrated in figure 5.3 a), where there is evidence of low frequency activity in the 0-5 Hz areas after the last compression period. In figure 5.3 b) such spontaneous activity is evident from the onset of VF. The compression periods are less evident than in the corresponding parts of figure 5.3 a. The frequency components of the artefacts become less distinct when spontaneous activity is present.

Filtering example

In figure 5.4 a) the original ECG signal is shown. CPR by precordial compressions is introduced at seven and a half seconds, and resulting artefacts are clearly visible.

The VF channel output from the high-pass filter is shown in figure 5.4 b) while the CPR channel output from the low-pass filter is shown in figure 5.4 c).

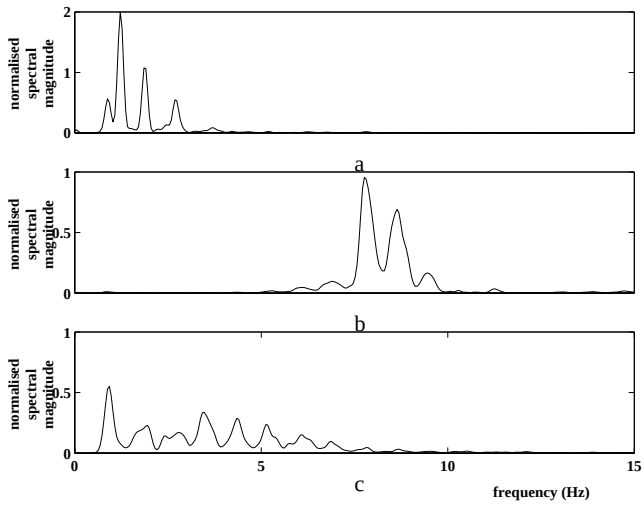


Figure 5.5: Averaged spectral estimates for (a) CPR-artefact in asystole, (b) Ventricular fibrillation (VF) in animals and (c) VF in humans. The area under the spectrum curve represents the total power of the signal.

Looking at the part of the plots corresponding to the period before introduction of CPR, it seems like the VF morphology is unchanged in the VF channel when comparing to the corresponding waveforms in the original ECG. Some minor low frequency components have leaked into the corresponding part of the CPR channel which is near isoelectric as expected.

If we proceed studying the part of the plots illustrating the period from the start of CPR, the start and proceeding of the precordial compressions is clearly evident in the CPR channel. It is interesting to note a change in the morphology of the VF channel in this period. There is clearly a more distinct amplitude modulation.

Comparing human and animal VF spectra

The spectra comparing human and animal ECG are shown in figure 5.5. The spectral components of the CPR artefacts are in the frequency area below five Hz (figure 5.5 a), of animal VF in the area above five Hz (figure 5.5 b), and of the human VF in the area from zero to 10 Hz (figure 5.5 c).

5.1.3 Discussion

In this section, observations from the ECG recordings from the animal compression frequency study has been presented. The illustrations showed how different rhythms are affected by precordial compressions and how artefacts appear in ECG recorded via different lead configurations. A simple experiment was performed to illustrate the principles of CPR artefact removal. Furthermore spectra computed from bot animal and human VF were compared with reference to animal CPR artefacts.

The use of frequency selective filters (figure 5.4) to remove CPR artefacts gave visually encouraging results, but as we was demonstrated by the spectra comparing human and animal data (figure 5.5), the problem of removing CPR artefacts from human ECG is not solved as straightforward as compared to animal ECG. This shows that the separation of CPR artefacts from the myocardial ECG signal during VF in man is more difficult than the corresponding problem in pigs where it can be successfully achieved by applying digital filters with fixed coefficients as illustrated in figure 5.4 and demonstrated in [95, 73]. These filters work by suppressing fixed frequencies, and their success thus require that the major frequency components of the CPR artefacts and the signal reflecting the cardiac rhythm are separable. This is the case in pigs with the major artefact components below 5 Hz and the VF band in the area above 5 Hz as demonstrated in this section (figure 5.3 and 5.5 a and b). It is not the case in humans where the frequency components of VF are lower with much more overlap with the frequency components of the artefacts as seen in the present study (figure 5.5 c) and reported by Strohmenger et al [94]. Therefore, other techniques are required to remove CPR artefacts from human ECG. We apply the adaptive filtering technique proposed in chapter 3.4 in our attemp at removing CPR artefacts from artificial mixes of human VF and animal CPR artefacts (chapter 7).

5.2 Effects of CPR in a short term VF animal model

In this section we study the potential of some of the features described in chapter 3 for monitoring and outcome prediction animal model. As we discussed in chapter 2, the level of MBF during CPR is indicative of the outcome of defibrillation. The higher MBF, the higher probability of a successful outcome following defibrillation (P(success)) [13, 98].

We chose two very different groups of animals as regards treatment to emphasise how well the changes in MBF levels caused by changes in therapy may be

expressed by ECG derived features. We studied five animals having received an effectful drug and compared to five animals having received a placebo. We wanted to show how changes in therapy with documented beneficial increases in the MBF level is accompanied by distinct chances in the time-frequency domain as shown in spectrogram plots of the transformed ECG.

We wanted to demonstrate how the changes evident in the time frequency representations can be captured by the features CF and SFM thus illustrating these features' potential for monitoring the effect the applied therapy.

Finally we grouped the features according to different levels of resuscitability. This was done to demonstrate the use of these two spectral features for outcome prediction.

5.2.1 Methods

We used the short term drug study pig data described in detail in chapter 4.2.1.

The PSD was estimated according to equation 3.41 using non-overlapping blocks of 300 samples with $K = 3$ segments each of length $L = 128$ with a segment overlap of 32 samples. Further, the features ,SFM and CF, were extracted sequentially according to equations 3.44 and 3.42, respectively, with the frequency range $f_l - f_u$ set to $4 - 20$ Hz to reduce the effects of the CPR artefacts in the VF analysis. The sequences of SFM and CF were smoothed to remove short time variations.

Monitoring

The feature sequences were aligned so that the sequence from each animal started at the onset of VF ($t = 0$) minutes and stopped approximately at the end of CPR ($t = 13$) minutes. Thus the sequences $SFM_{i,j}(n)$ and $CF_{i,j}(n)$ where $n = 1 \ldots 240$, $i = 1, 2$ and $j = 1, \ldots, 5$ for animal number j in group number i were obtained. For each of the two animal groups a within group averaged feature was made by computing the mean feature value for every third second according to $SFM_{AV_i}(n) = \frac{1}{5}\sum_{j=1}^{5} SFMi, j(n)$ and following a similar procedure for the computation of $CF_{AV_i}(n)$. $i = 1$ and $i = 2$ correspond to the vasopressin and placebo group respectively.

Prediction of expected outcome

To consider the ability of the measurements to distinguish between ECG from different phases of treatment, the measurements were grouped according to the class division scheme given in table 5.1. The reason for this division scheme is best explained by referring to tables 4.4 and 4.2 which shows the MBF levels in the two groups. Further, the discussion in section 4.2.1 shows how the classes may be associated with varying levels of resuscitability (increasing with class number).

Expected outcome	Class	Class criteria	Numbers
ROSC	ω_1	CPR+VASO	600
No-ROSC	ω_2	CPR	600
No-ROSC	ω_2	CPR+PLAC	600
No-ROSC	ω_3	VF	800

Table 5.1: Class division scheme analysis of predictive capability of features in short term drug effect study. VF: VF without artefacts or effects from drugs; CPR: ongoing CPR; VASO: vasopressin has been applied; PLAC: placebo has been applied.

The estimates of the a priori probabilities and class-specific PDFs are computed for each of the three classes according to equations 3.36 and 3.35 using the histogram method of section 3.2 with histogram resolution set according to $n_b = 50$ and no smoothing used. This is done because of the large number of features in each class. The performance is measured by the sensitivity and specificity of class ω_1 given by equations 3.14 and 3.15, respectively. Five-fold cross-validation is used in the evalution of each of three classifiers using the following feature vectors: $\mathbf{v} = \begin{bmatrix} v_{SFM} \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} v_{CF} \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_{SFM} & v_{CF} \end{bmatrix}$.

5.2.2 Results

Two time frequency plots, of the PSD sequences computed from two animals, one from the vasopressin group and the other from the placebo group, are shown in figure 5.6 a) and b). It must be noted that the frequency components (in red) in the low frequency region starting at 4 minutes are artifacts due to the chest compressions performed on the animals. The ridge of frequency components (in red) in the frequency area $4 - 18$ Hz roughly corresponds to the electrical activity in the heart. As can be seen from both spectra, the

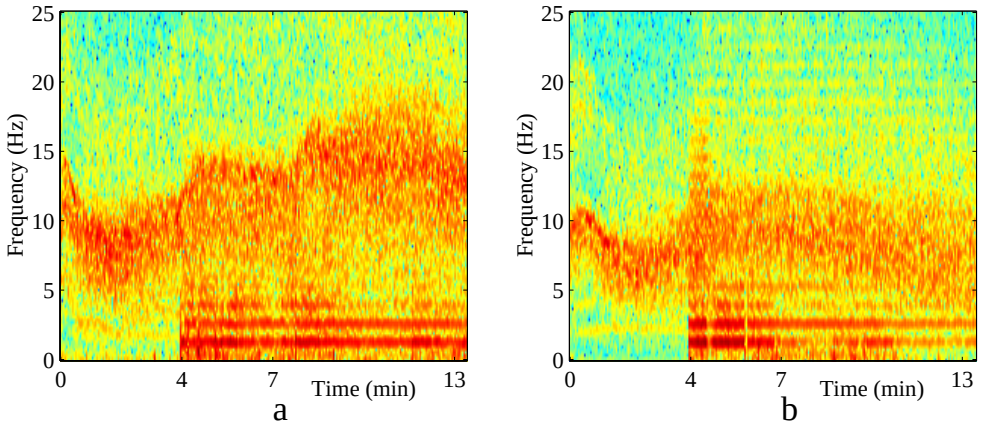


Figure 5.6: Time frequency representations of the VF period in two animals from a) vasopressin group b) placebo group. The colours identify the magnitudes as follows: low (dark blue), below middle (light blue), above middle (yellow), high (red). The frequency components of VF and CPR artefacts (and spontaneous activity) are identified by high magnitudes (red) in the frequency areas 5-15 Hz and 0-5 Hz respectively.

effect of treatment is much the same until drug administration. The effect of introducing VF is evident in both spectra by a transient period the first 2–3 minutes where the VF rhythm is organised after defibrillation. At the introduction of CPR the centroid of the area under the PSD curves shift markably to a higher frequency level. This corresponds to the increase in MBF (and resuscitability) with the introduction of artificial circulation. At the introduction of drugs at 7 minutes a further shift of centroid is evident in the spectrogram of the animal receiving vasopressin. In the spectrogram corresponding to placebo, no such change is evident.

Monitoring

Figure 5.7 show plots of the individual (dotted lines) and averaged feature sequences (solid lines) illustrating how the ECG measurements change during the drug study experiments. Note that significant events are indicated on the time axis. As can be seen from the plots the features attain distinguishable levels for the different treatment phases.

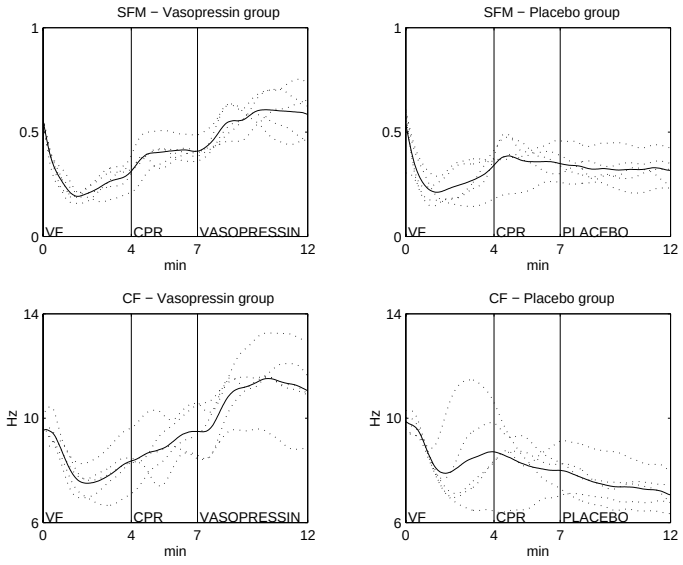


Figure 5.7: Time Domain Parameter Plots: individual animals - dotted lines, averaged within group - solid line.

Prediction

In figure 5.8 the plots for the estimates of the class specific PDFs are shown for the combined SFM and CF.

The estimated performance characteristics for the classifiers aiming to distinguish between ECG measurements originating from different treatment phases are shown in table 5.2.

5.2.3 Discussion

We have analysed ECG from 10 pigs having received CPR with effectful drug therapy and compared to similar analysis results from 10 pigs which did not receive effectful drug therapy. Our results demonstrate how structure relevant to therapy is revealed in time frequency transformation as shown in figure 5.6. This is useful for capturing information relevant to the changes in myocardial metabolism.

The plots of SFM and CF shown in figure 5.7 clearly shows a distinction between the vasopressined animals as compared to the placeboed animals after the drug administration. The way the features change according to each of the

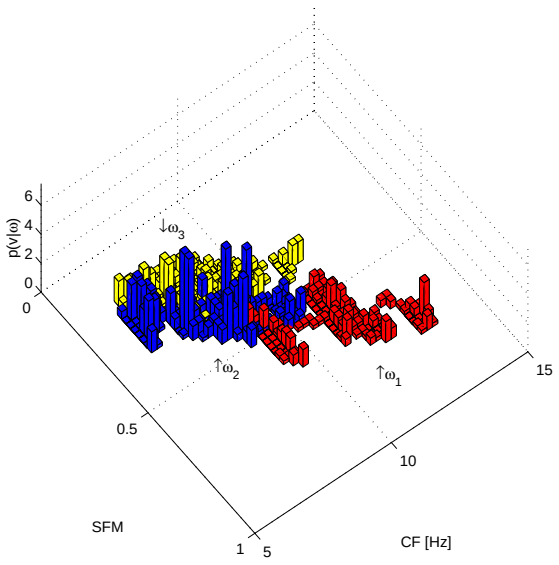


Figure 5.8: Estimated class-specific probability density functions. The subclasses are indicated with labels: $\omega_i, i = 1, 2, 3$.

v	$\hat{P}_{sns}(\omega_1)$	$\hat{P}_{spc}(\omega_1)$
SFM	79%	98%
CF	86%	95%
$[SFM, CF]$	93%	99%

Table 5.2: Performance characteristics. The sensitivity and specificity for predicting ROSC for classifier using centroid frequency (CF) and spectral flatness measure (SFM) either alone or in combination. The standard deviations for the performance characteristics were all less than 2.2 %.

animals' response to treatment indicates that these features have a potential for being used as monitoring parameters.

The features capture the information relevant to resuscitability(all animals in the vasopressin group were successfully resuscitated as compared to none in the other group). This, and the performance of the classification experiments shown in table 5.2 indicate the capability of the features as shock outcome predictors. It is worth noting that the highest performance is achieved when the two features are combined.

These results are indeed encouraging, and seems to be in accordance with other equally encouraging results from other animal experimental studies using VF analysis for purposes similar to ours. But, these results are not representative for what one might expect to achieve in a human study of cardiac arrested patients as we indeed will demonstrate in chapter 8. This, we assume, is due to the simple animal model we have studied in this section. All animals had more or less the same age, weight and were healthy. Treatment followed exactly the same timeline for all the animals. This may explain the lack of variation in how the animals responded to the treatment.

5.3 Effects of CPR in a long term VF animal model

In the previous section we demonstrated the potential usefulness of CF and SFM for monitoring the effect of CPR and further in predicting the outcome of defibrillation.

Here we will show that the specific animal experimental setting is an important factor when we want to analyse the discriminatory capability of features extracted from ECG.

We studied the ECG from an animal experimental setting similar to the one studied in the previous section. More drugs were studied, but most important, the time to defibrillation was much longer than in the study referred to in the previous study. We only considered the features' monitoring capability.

5.3.1 Methods

The data described in chapter 4.2.2 were used in this analysis.

The feature extraction was performed according to the procedures described in the previous section, except for the fact that ENRG was extracted according equation 3.45 in addition to SFM and CF. No periodogram averaging was used, and the features were not smoothed.

5.3.2 Results

Figure 5.9 show plots of SFM, CF and ENRG depicting how the ECG measurements change during the drug study experiments. The solid black lines are the averaged within group curves while the within group standard deviations are plotted with dotted grey lines.

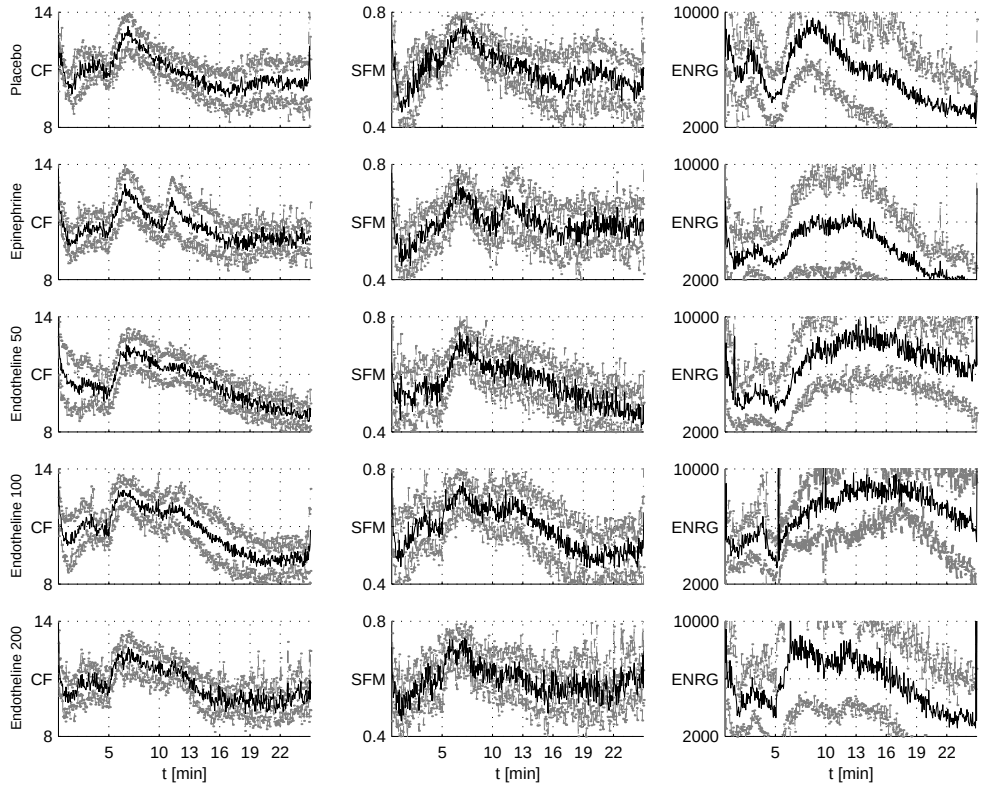


Figure 5.9: Time Domain Parameter Plots: averaged within group - solid black line. Within group standard deviations - dotted grey lines.

5.3.3 Discussion

Looking at figure 5.9, we clearly see the CPR effect reflected in the part of the plots corresponding to the early phases of treatment. If we look at SFM and CF their behaviour correspond to that seen in the plots in section 5.2.

The more effectful drugs (endotheline 50 and 100) upholds a positive effect longer than the other drugs (until 16 minutes). After 16 minutes we do not see any clear distinction in the CF or SFM plots indicating which animals are successfully converted. These features seem to loose their discriminant capability relevant to resuscitability in the late stages of VF.

In section 5.2, our results showed an improvement in discriminatory power when SFM and CF was combined as compared to using the features individually. If we consider the parameter plots of ENRG for the different drug groups we see that the average ENRG seems to contain information capable of discriminating the drug groups with higher success from those with zero success in the late stages of VF, where SFM and CF fails. In the earlier stages of VF, however, the ENRG plots show less detail than SFM and CF. To us this indicates that the different features capture different aspects of the metabolic condition of the heart. The ideal feature should fully reflect the resuscitability and nothing else. We suspect that the features we are investigating only partly reflects resuscitability and holds other irrelevant information. Our results suggest to us that the combination of features containing information partly relevant to resuscitability will perform better than single feature systems.

Chapter 6

Quality of ALS in the Oslo EMS

The study presented in this chapter investigated what actually occurred during the two last links in the "Chain of survival": defibrillation and advanced life support (ALS) in the Oslo EMS system. To save more lifes after cardiac arrest, the current guidelines [16, 81] recommend that the four links in the "Chain of survival":early recognition and call, basic life support (BLS), defibrillation and advanced life support (ALS), be performed as early as possible [28, 64] and with the best possible quality [114]. The material in this chapter is adapted from [100].

In analysing the Oslo data described in chapter 4.1.1 we sought to evaluate important factors concerning defibrillation and ALS procedures. These are factors like the timing of the shocks, which rhythms occurs after the shocks, actual chest compression rates and duration of "hands off" intervals. If this proved successful, the results could be used for feedback to the EMS system, both on an individual and EMS system level. This was the primary purpose in [100]. In this thesis, with prediction of defibrillation outcome, monitoring of CPR efficacy and CPR artifact removal in mind, we wanted our discussion to be relevant to these issues, thus to get an idea of what potential these concepts could have for strengthening the two last chains in the "Chain of survival".

6.1 Methods

We analysed the data described in chapter 4.1.1.

The annotations of the raw data was logged into a database and read into the Matlab environment and linked to the MCM information and the demographic data. The database was organised as a matrix where each row entry described the timing of a treatment sequence of defibrillation, the following rhythms and the times of initiation and frequency of precordial compressions. The initiation of precordial compressions following initial asystole or electromechanical dissociation (EMD) was also logged. The events in row "no i" happened before row "no i + 1" and so on. The information was arranged in columns so that all defibrillation timings appears in the same column, isoelectric onsets in another etc. This enabled the computation of time intervals between significant events in the treatment.

Return of spontaneous circulation (ROSC) included sustained ROSC and shorter periods of spontaneous circulation (pulsegiving rhythm). Thus, more than one episode of a pulsegiving rhythm could appear in a patient record. Whether a rhythm was pulsegiving or not was decided from the Utstein- and paramedic record.

The end-points of the analysis are described as follows:

- *The total number of shocks for each patient*
- *The number of successful shocks*
  - defined as a non-VF/VT rhythm for at least the first 5 sec after the shock.
  - or defined as return of pulsegiving activity regardless of duration.
  - or defined as sustained ROSC (lasting until admittance at the hospital).
- *The duration of isoelectric ECG (IE) after a shock*
  - all IE periods.
  - IE periods ending in VF/VT.
  - IE periods ending in EMD.
  - IE periods ending in pulsegiving electrical activity regardless of duration.
- *Time intervals:*
  - from initiation of monitoring until the first shock was given in cases with VF/VT as initial rhythm.

  - from the discontinuation of precordial compressions during VF/VT until the next shock was given.
  - from initiation of monitoring until the precordial compressions were initiated in cases of asystole and EMD as initial rhythm.
  - from initiation of monitoring until sustained ROSC.

- *The precordial compression rate* The mean rate is calculated from the compression artefacts of two randomly picked 15-sec periods for each patient.

Analysis results are sometimes presented as medians with lower and upper quartiles in parenthesis (q25,q75). In the time analysis the 95% confidence interval is also presented. For testing of hypotheses the Students t-test is used if the data are normally distributed, and the Mann-Whitney-Rank-Sum Test if the normality test fails. For comparisons of different proportions we use the chi-square test with Yates correction. A P-value < 0.05 is regarded as significant in all tests [31]. Counts are denoted by n.

6.2 Results

In the two-year study period 453 confirmed cardiac arrests were considered for resuscitation. 573 cases were confirmed dead, and ALS was not initiated. The survival rate with cardiac aetiology was 9%, and 26 (87 %) of these 30 survivors had no or moderate disability.

Of the 328 cases with cardiac aetiology, we have on line computer registration of 201 cases. 45 of these cases were excluded due to problems with incomplete registrations, sampling failures and difficulties in visually separating the CPR artefacts from the electrical activity of the heart. Thus, the final data analysis was performed on 156 cardiac arrests of cardiac aetiology corresponding to the data material described in section 4.1.1 (table 6.1). 46 of these patients stayed in a non-shockable rhythm. The survival rate (6%) and the other variables in the group of 156 patients are not different from the same variables in the cardiac aetiology group of the overall Utstein results [100].

There are no cases of initial VT, and only a few runs of VT during CPR, thus VT and VF episodes are grouped together.

Bystander CPR was initiated in 82 patients (53 %). In 48 (59 %) cases the bystander effort was recorded as adequate, in 17 (21 %) as not adequate and in 17 (21 %) the effort was not assessed.

Total number	156
Witnessed (bystander)	118 (76 %)
Initial VF	78 (50 %)
Bystander CPR	74 (46 %)
Sustained ROSC	40 (26 %)
ICU/Ward	38 (24 %)
Discharged alive	10 (6 %)
Alive 3 months	9 *
Alive 6 months	9 *
Alive 12 months	9 *

Table 6.1: Utstein Style Data of the 156 patients included in the final analysis. * 1 unknown - left country

Endotracheal intubation was performed in 143 patients (92 %) with a median duration of time of 2 min (1, 3) from after arrival at the scene.

The median time interval from the reception of the call until the ambulance crew was with the patient is 8 min (5, 11) with a significant difference between survivors (5 min (1, 5)) and non-survivors (7 min (5,11)) (P=0.01)

The median time interval from the estimated collapse time to first shock during initial VF was 11 min (7, 14), again with the same significant difference between survivors (6 min (6, 8)) and non-survivors (11 min (9, 15)) (P=0.003).

110 of the 156 patients received a total of 883 shocks. Median 6 shocks were given to each patient (3,12) with a significant difference between survivors (1 (1, 1)) and non-survivors (7 (3,12)) (P$<$ 0.001) (figure 6.1).

In 329 shocks (37 %) the patient was still in VF 5 sec after the shock. 310 (94 %) of these shocks has VF continuing directly after the shock.

554 shocks (63 %) were successful when defined as a non-VF/VT rhythm 5 sec after the shock. There is no significant difference in the success rate for the first, second, third and fourth shocks when required: 82 out of 110 (75 %), 64 out of 92 (70 %), 69 out of 87 (79 %) and 56 out of 75 (75 %), respectively (figure 6.2). The successrate for all later shocks combined is significantly lower, 283 out of 519 (55 %) both compared to the first, second, third and fourth or the four combined (P$<$0.001).

90 shocks (10 %) in 51 patients were successful when defined as a pulsegiving rhythm after the shock regardless of duration, while 35 shocks (4 %) were successful when defined as sustained ROSC after the shock. 14 ROSCs occurred

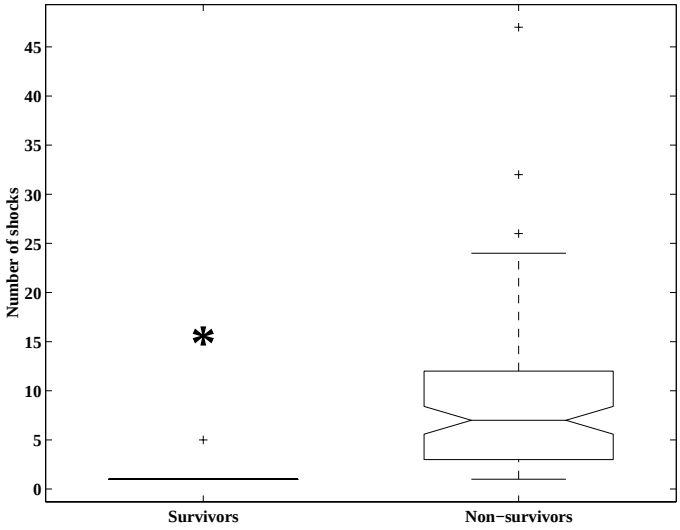


Figure 6.1: Shocks in survivors and non-survivors. The box illustrates q25, median and q75. The notch reflects uncertainty about the mean. The whiskers and outliers (+) show the data range. * P< 0.001.

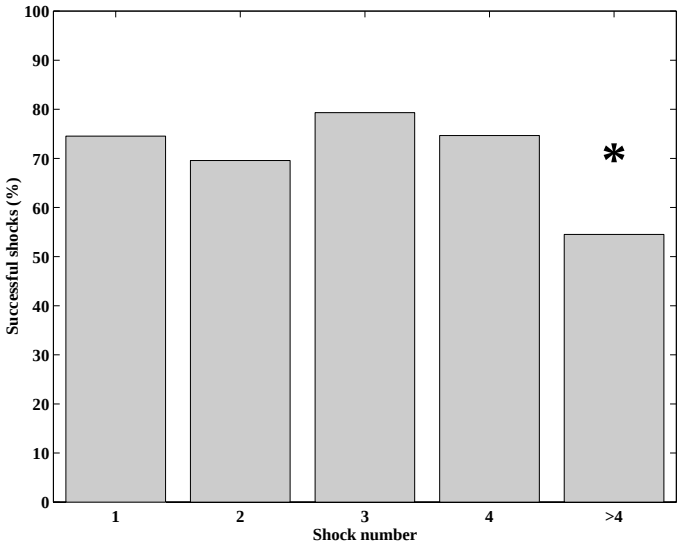


Figure 6.2: Successrate (non-VF/VT within the first 5 sec after the shock) after the first, second, third, fourth and later shocks. * P< 0.001 compared to 1, 2, 3, 4.

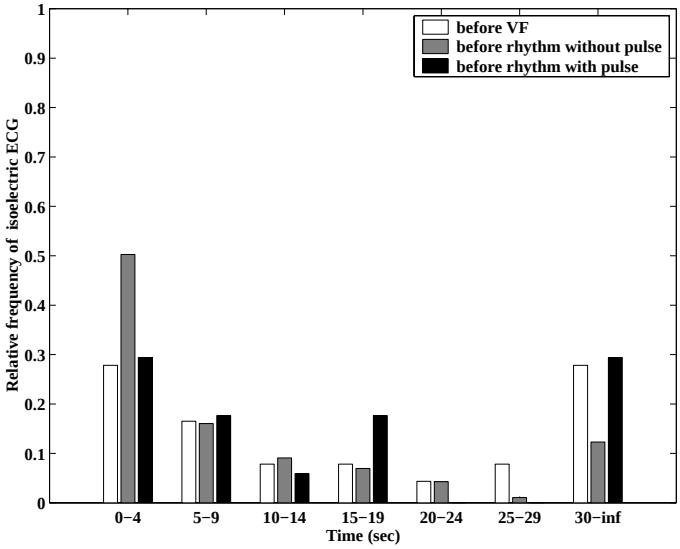


Figure 6.3: Duration of isoelectric ECG after a shock when the episode ended in a new VF/VT, rhythm without pulse (electromechanical dissociation, EMD) or rhythm with pulse.

after the first shock, 2 after the second shock (P=0.012 compared to the first shock), 3 after the third (P=0.041 compared to the first shock) and fourth shock (P=0.079 compared to the first shock) and 13 after more than four shocks (P<0.001 compared to the first shock). More patients had sustained ROSCs after four or less shocks than after >4 shocks (P=0.013). Five patients obtained ROSC with a non-shockable rhythm, but none of them survived.

332 shocks, 38 % of all shocks, were followed by an isoelectric period with a median duration of 8 sec (3,22). Figure 6.3 presents the different isoelectric time periods dependent of which rhythm occurs after the shock.

25 (8 %) of the isoelectric episodes lasted between 20 and 30 sec, 64 (19 %) lasted more than 30 sec. In 5 episodes a pulsegiving electrical activity occurred after more than 30 sec of IE. 13 patients (2 %) remained in isoelectricity.

Important time intervals are presented in table 6.2 The median time from initiation of ECG monitoring until the first shock was given (19 sec, n=78) and from discontinuation of precordial compressions until the next shock was given during VF/VT (20 sec, n=180) are not different (table 6.2.

The median time from initiation of monitoring during asystole (n=49) until precordial compressions was initiated (29 sec), is significantly shorter than

	n	median (q 25, q 75)	Confidence interval
a) From initiation of monitoring to first shock in primary VF	78	19 (16, 25)	22-34
b) From end of precordial compressions to next shock (VF)	180	20 (18, 24)	23-28
c) From initiation of monitoring to start of precordial compressions (asystole)	49	29 (21, 51)*	33-50
d) From initiation of monitoring to start of precordial compressions (EMD)	29	109 (34, 196)* **	92-204
e) From initiation of monitoring to rhythm with pulse	95	627 (283, 998)	576-755
f) From initiation of monitoring to sustained ROSC	40	559 (278, 978)	503-814

Table 6.2: Time intervals (sec). * P<0.001 compared to a), ** P<0.001 compared to c)

during EMD (n=29,109 sec) (P<0.001). These times are both significantly longer than the median time from initiation of monitoring until the first shock was given in cases with VF (P<0.001).

The median time from the initiation of monitoring until sustained ROSC is present is significantly shorter in survivors (163 sec, n=10) than for non-survivors (796 sec, n=30) (P=0.011) (table 6.2).

The precordial compression rate is 108 min^{-1} (100,120)

6.3 Discussion

The Utstein data from the 156 patients studied in detail (table 6.1) here are not different from the same data for the total patient population with out-of-hospital cardiac arrest in the two-year study period. The demographics of the study population, and the call-response interval in the survivors and non-survivors, are also comparable to other reports [86, 91, 110, 27, 111, 26, 6, 46, 88, 89, 25, 50, 2]. In [100], we concluded that the study population represents a normal group of prehospital cardiac arrested patients. The 9 % survival rate after arrests of cardiac aetiology in Oslo is in the lower range compared to those in a recent survey of European cities picked for their potential good results [51].

The median six shocks (3, 12) per patient is more than in some other reports with monophasic waveform devices (like in the Heartstarts) [110, 6, 88, 89, 50].

The median of 11 min from collapse to first shock during initial VF in the present study is longer than in Seattle [110], Iowa [6] and Milwaukee [50], and will also probably negative influence the total shock number [50] as the survival rate [28, 64, 86, 91, 110, 27, 111, 26, 6, 46, 88, 89, 25, 50, 106, 2].

58 % of the successful shocks (defined as a non-VF/VT rhythm 5 sec after the shock), or 38 % of all shocks, were followed by an isoelectric period with a median duration of 8 sec before a new VF/VT or an organised rhythm with or without a pulse. The isoelectricity is usually due to a transient period of electrical and/or myocardial "stunning" [81]. 19 % of the isoelectric periods lasted more than 30 sec, and five of these episodes ended in pulsegiving electrical activity. The paramedics should initiate precordial compressions during this time interval according to the guidelines [16].

Only 10 % of the 883 shocks resulted in a pulsegiving rhythm, which is similar to previous reports. With the same waveform Behr et al found an organised rhythm after 13 % of all shocks (with self adhesive pads) [6], Hargarten et al after 10 % of the first five shocks [50]. From these data we postulate that it might be important to optimise the myocardial perfusion and resuscitability before delivery of a countershock. Few shocks indicate a higher survival rate [6, 46, 88, 89, 25, 50, 106, 2] and each unsuccessful shock damages the heart and can increase the post-shock dysfunction [115].

Only 4 % of the shocks resulted in sustained ROSC, after a relatively long median time interval of 490 sec, over 8 min, from initiation of monitoring. Others have also found that ROSC often occurs after several min and several countershocks [50], and Bonnin et al reported ROSC up to 25 min after the arrival of the paramedics in a series of 1461 consecutive cardiac arrests [8]. This indicates the importance of well performed CPR to optimise the cerebral and myocardial perfusion before ROSC to improve the rate of survival and the chance of a normal neurological status. In the present study the ten survivors had a significantly shorter time interval, but still 163 sec, almost three min.

The compression rate of 106 min^{-1} is close to the recommended 80-100 min^{-1} in the 1992 guidelines [16] or 100 min^{-1} in the 1998 guidelines [81]. There are to our knowledge no human data showing the optimal compression rate. In pigs and dogs the haemodynamics seem to be optimal with a rate of 100-120 min^{-1}, depending also on the compression time of the duty cycle [65].

The ERC 1992 guidelines [16] were followed during the study, and chest compressions were given for periods consisting of 10 sequences of five compressions to one breath. As almost all patients were intubated early during CPR, no compression pauses were needed to ventilate the patients. With a median

compression frequency of 106 min^{-1}, the patients thus received chest compressions for less than 30 sec between non-compression "hands off" intervals for ECG analysis, defibrillation attempts etc., while the latter lasted 45sec if three defibrillation attempts were needed. It is questionable if this gives adequate perfusion to the brain and the heart. Even optimally performed non-interrupted compressions give only 30 % of normal blood flow [48, 18]. It is important to emphasise that the quality of manual chest compressions are not optimal when evaluated over time [103, 52], which certainly further reduces the myocardial and cerebral perfusion. Thus, the myocardial condition might not have been optimised for return of a pulsegiving rhythm after this short compression interval. This might partly explain the high shock rate required in the present study. Sanders et al found an inverse relationship between rate of successful resuscitation and duration of inadequate coronary perfusion pressure during standard CPR in dogs [83], and Sato et al have recently reported adverse effects of interrupting precordial compression during cardiopulmonary resuscitation in rats [84]. The rate of spontaneous circulation decreased significantly when the shock was delayed 30 sec after discontinuation of precordial compressions, and 24-hour survival was significantly reduced with a 20-sec delay.

There are some limitations in the present study. Firstly, as in other studies the time of arrest is unreliable, also when based on a witnessed account, and times reported by the ALS performers and not automatically reported, could be incorrect. Secondly, the analysis of the chest compressions was based on manual evaluation of the MCMs. Thus a subjective influence on those results cannot be excluded. There were on the other hand three persons with continuous checks on the evaluations, which were performed over an intensive, short time period. To reduce the chance of errors to a minimum, patients were excluded if there were difficulties in the evaluations. Thus, 22 % of the patients were excluded. The evaluation was concentrated around the initial monitoring period, periods around the shocks and when changes in the ECG signals occurred.

In conclusion this study demonstrates the importance of quality control of ALS performance, which can be further improved by optimising the work of the ambulance personnel and the physicians. In the Oslo EMS system, the ratio between chest compression periods and the "hands-off" intervals should be reduced. Time is a decisive factor in CPR, and, besides the important early defibrillation, myocardial and cerebral perfusion must be optimised during CPR. The information stored in the computer chips of the defibrillator can be very useful in evaluating what occurs during the ALS attempt. Together with the information recorded on the "Utstein style template", this can be used to

evaluate the total EMS system performance and for individual feedback to the personnel after a specific CPR attempt.
The results computed from the defibrillator data indicate that the ALS performance probably could be improved by reducing the duration of "hands-off" intervals and thereby enabling an increase in the duration of the compression/ventilation periods. From the analysis of our data, it is evident that many of these intervals are related to shocks. The median time from discontinuation of precordial compressions for rhythm analysis and charging until the next shock was given during VF/VT was 20 sec in our material The importance of minimising this "hands off" interval was clearly demonstrated by Sato et al [84], as earlier described. As we see it, this highly encourages applying a reliable method for CPR artefact removal. This would enable rhythm analysis during CPR, thus reducing these intervals by at least typically 9-12 seconds (chapter 7).
Another issue observed in our analysis is the high number of unnecessary shocks. As we have pointed out each of these shocks damages the heart and increases the overall "hands-off" interval. Thus our data material clearly bears evidence of the potential benefit a shock outcome predictor may add to the ALS performance. This is further explored in chapter 8.
Finally, we have considered the quality of CPR measured by how much provided and at which frequency. But there is no reference as to what is optimal performance. Larsen et al found that early and effective ALS would increase the survival when it is added to early defibrillation and early BLS [64]. In addition to the use of drugs and advanced control of the airways, it is possible that monitoring of a parameter reflecting CPR efficacy, could increase the quality of CPR and thereby optimise the myocardial (and cerebral) perfusion and resuscitation probability. We will describe a method for monitoring this probability in chapter 9.

6.4 Summary

In this chapter we have analysed the data from the Oslo EMS system. This has allowed us to assess the quality of defibrillation and ALS using data from the medical control module of the defibrillator.
We have pointed out that the ALS quality can be improved by decreasing "hands-off" intervals, reducing the number of unnecessary shocks and by monitoring CPR performance. We claim that these improvements to some degree can be achieved by CPR artefact removal, prediction of defibrillation outcome and monitoring of resuscitation probability. These issues will be investigated in the following chapters.

Chapter 7

CPR artefact removal

In chapters 8 and 9 we demonstrate how VF-analysis techniques are applicable in a CPR decision support system. In the VF-analysis described in those chapters, we work with ECG free from CPR artefacts. In chapter 3, we emphasised the importance of being able to remove CPR artefacts from the ECG signal for valid rhythm and/or VF-analysis. This would allow continuous monitoring of CPR efficacy during precordial compressions and thus decrease the "hands-off" intervals which as we showed in chapter 6 is one of the important factors for improving CPR. This chapter is adapted from [1].

In this study we wanted to find out if artefact removal by filtering would allow the maintenance of precordial compressions during automatic rhythm analysis, without interrupting the sensitivity of simultaneous rhythm classification.

In an attempt to solve this problem we utilise the adaptive filtering strategy put forward in chapter 3.4.

7.1 Methods

The data used in this study originate from the animal experiment described in section 4.2.3.

The purpose of using an animal model is to provide a controlled environment where the distinction between artefact and signal is clear. To simulate CPR artefacts in human ECG, we manually added animal artefacts to human VF and VT rhythms.

In this study we used signals collected from 2 animals. We decided to use the worst case artefact as seen in asystole having the highest amplitude and

frequency components over a broader band as compared with the artefacts which could be isolated from pigs in VF.

We focused on artefacts due to 2 different sources:

1. Chest compressions: These were done using an instrumented mechanical compression device, which had a reference signal output that was proportional to the compression depth.
2. Ventilation: Manual bag valve ventilation with 100% O_2 was performed without interruption of the compression-relaxation cycle. To provide a reference for the artefact induced by the bag ventilation the thorax impedance, as seen between the defibrillator pads, was recorded.

The animal preparation and the experimental protocol was described in chapter 4.2.3. For further details see [63].

Preprocessing and artefact addition

The human ECG data (VF/VT) were filtered with a 0.5 Hz high-pass filter for removal of offset components. No other filtering or normalization is done to the human data in order to simulate a real-life situation. The average signal amplitude variance is 0.027 and 0.34 for the VF and VT records, respectively.

Similarly to the human ECG data, the animal asystole data are high-pass filtered at 0.5 Hz. In addition, they are limited to 10 Hz using a low-pass filter – this is discussed in section 7.3. After filtering, each of the 15 sec asystole records are normalized to unit variance. In order to simulate a wide range of noise conditions, we add the animal data, now considered artefact noise, to the human VF/VT data using an adjustable scaling factor. Denote by $x_h(n)$ the human VF/VT signal and $a_n(n)$ the normalized artefact signal. Setting

$$a(n) = C \cdot a_n(n), \tag{7.1}$$

where C is a chosen constant, the noisy signal $x(n)$ is modeled as

$$x(n) = x_h(n) + a(n). \tag{7.2}$$

Given a target SNR defined as

$$\text{SNR} = 10 \log_{10}(\frac{\sigma^2_{x_h}}{\sigma^2_a}), \tag{7.3}$$

the constant C is found as

$$C = \sqrt{\frac{\sigma_{x_h}^2}{10^{\frac{\text{SNR}}{10}}}}. \tag{7.4}$$

By individually computing C for the construction of every 15 s noisy signal record, a fixed SNR is ensured for the whole ensemble.

The signal records used in the experiments are as follows:

- Artefacts: Animal asystole data from 2 animals are used to model artefacts in human ECG. The data records are 15 sec each, and the settings are:

  ECG recording: As seen through the monitor (MON) or defibrillation (DEF) pads.

  Compression rate: 60, 90, or 120 min^{-1}.

- Human ECG: The data records are 15 sec each and of type:

  VF: 200 records.

  VT: 71 records.

- Noise levels: Model artefacts are added to the human data on a fixed SNR basis using equations 7.2 and 7.4. The SNR range is from -10 dB to 10 dB.

The filtering is performed according to the adaptive filtering methods described in chapter 3.4. The derivation of the optimal filter solution and further tuning of filter parameters is described in [1] and given in appendix B.

7.2 Results

In the following experiments described in the following a we use a 1-tap filter for both channels, i.e. when using compression depth and thorax impedance as reference signals, combined with a window size of 301 and 161 samples in the MON and DEF case, respectively.

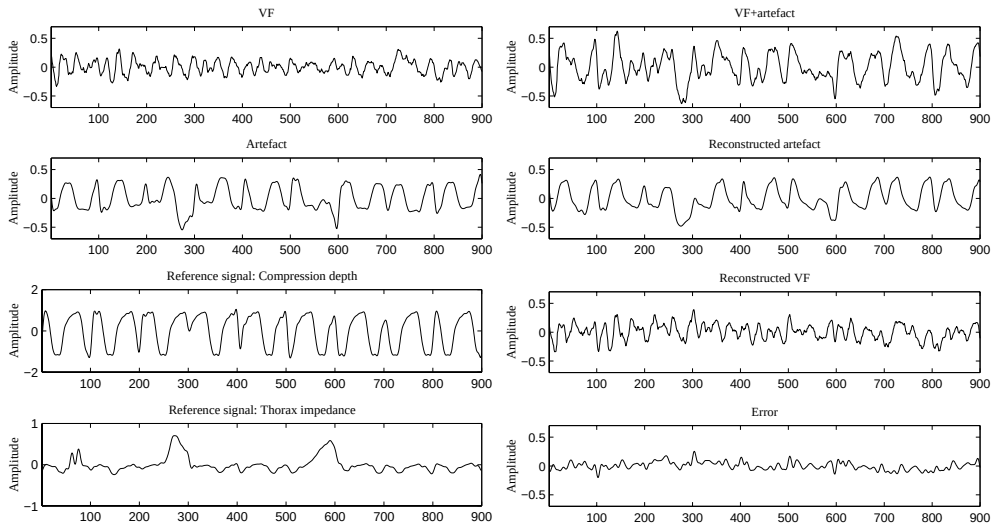


Figure 7.1: VF filtering example using artefact as seen through the monitor pads.

Filtering examples

Before presenting an overall evaluation we visualize the effect of the proposed system with a filtering examples. We have selected a 15 sec human VF record, to which an animal artefact is added, with the constant C in equation 7.2 chosen to give an SNR of -5 dB. The artefact was originally measured through the monitor pads. Figure 7.1 shows a selection of the resulting waveforms, where only the center 9 sec are included in order to avoid filter edge effects.

Comparing the artefact signal with the two reference signals we observe that in this example both references are negatively correlated to the artefact. Using the proposed algorithm the reconstructed artefact resembles the original, but some errors are still present in the reconstructed VF. This is due to the heavy noise conditions in the example (-5 dB).

SNR evaluation

Using all the data records summarized in section 7.1 we evaluated the proposed artefact removal system in terms of SNR improvement. The results are shown in figure 7.2.

There is little difference between the results obtained for VF and VT. Again we observe that the artefacts, as seen through the defibrillator pads, are more

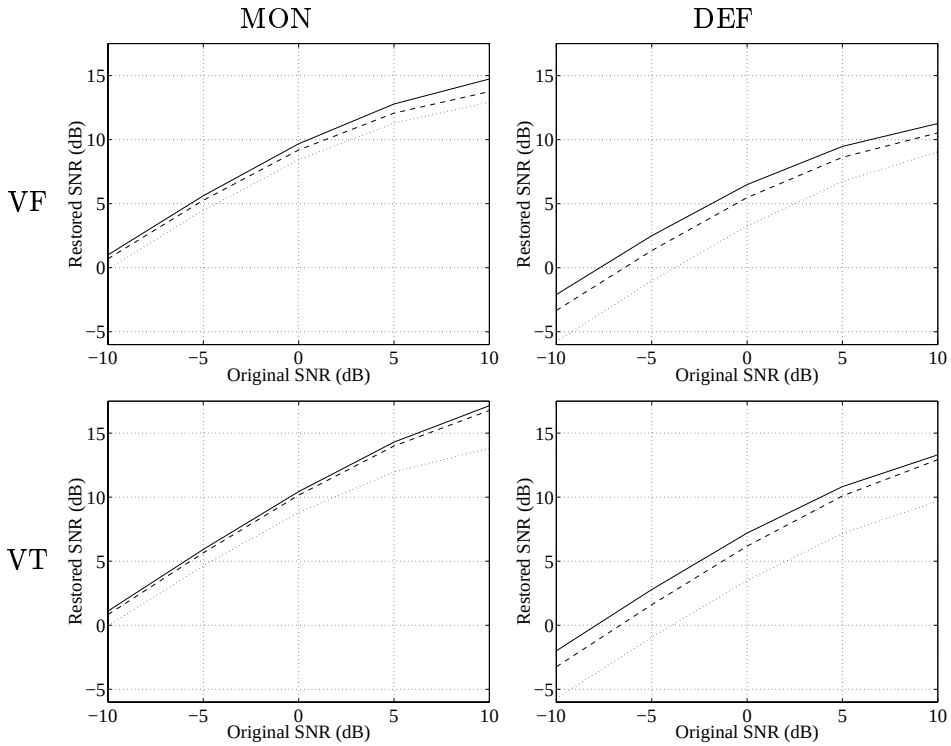


Figure 7.2: Average SNR performance of the artefact removal system at varying compression rates: 60 (solid), 90 (dashed), and 120 chest compressions per minute (dotted).

difficult to remove than those seen through the monitor pads. The difference is about 3 dB at a compression rate of 60 min^{-1} going up to about 5 dB at 120 min^{-1}.

The artefact removal becomes increasingly more difficult when the compression rate goes from 60 to 90 and to 120 min^{-1}. This is due to a higher degree of overlap in the frequency domain. A compression rate of 120 min^{-1} corresponds to a fundamental frequency of 2 Hz.

Sensitivity evaluation

Although the SNR increase is a qualified measure of the success of the artefact removal system, the applicability rests on it's ability to facilitate accurate automatic classification of human ECG rhythms in noisy environments.

The VF and VT rhythms used here are both considered rhythms to treat, i.e. rhythms for which delivery of electric countershock is considered beneficial to cardiac resuscitation. Given a number of classification outcomes, let T and N denote the number of treat and non-treat classifications, respectively. The *sensitivity* is then defined as

$$S = \frac{T}{T+N}. \tag{7.5}$$

The sensitivity expresses the ability of the classification system to correctly identify treat rhythms. In figure 7.3 we compare the obtained sensitivities using a computer version of the Laerdal Heartstart classification algorithm on noisy and restored human ECG signals. When computing the results only the 9 sec center part of each 15 s record is used due to filter edge effects. The classification algorithm is based on an analysis of 3 sec block, thus giving 3 classification outcomes for each 9 sec block.

The interesting observations related to figure 7.3 are when the classification breaks down. As a reference the obtained sensitivities when no artefacts were added were 0.95 and 0.97 for VF and VT, respectively.

We observe that the classification of VT rhythms fails more often than for VF rhythms. This applies to monitor as well as defibrillator data. One of the reasons is the higher variance of the VT records, see section 7.1. In our setting a high signal variance implies high variance on the added artefact in order to obtain a desired SNR level. This gives overflow in the signal representation used by the Laerdal Heartstart classifier. Removing a significant part of the artefact help reducing the extent of the overflow. As an example, in the VT/MON case with a compression rate of 120 min^{-1}, 16.5% of the noisy signal samples where out of range when the SNR was -10 dB. This was reduced to 4.2% when removing the artefact before doing the signal classification.

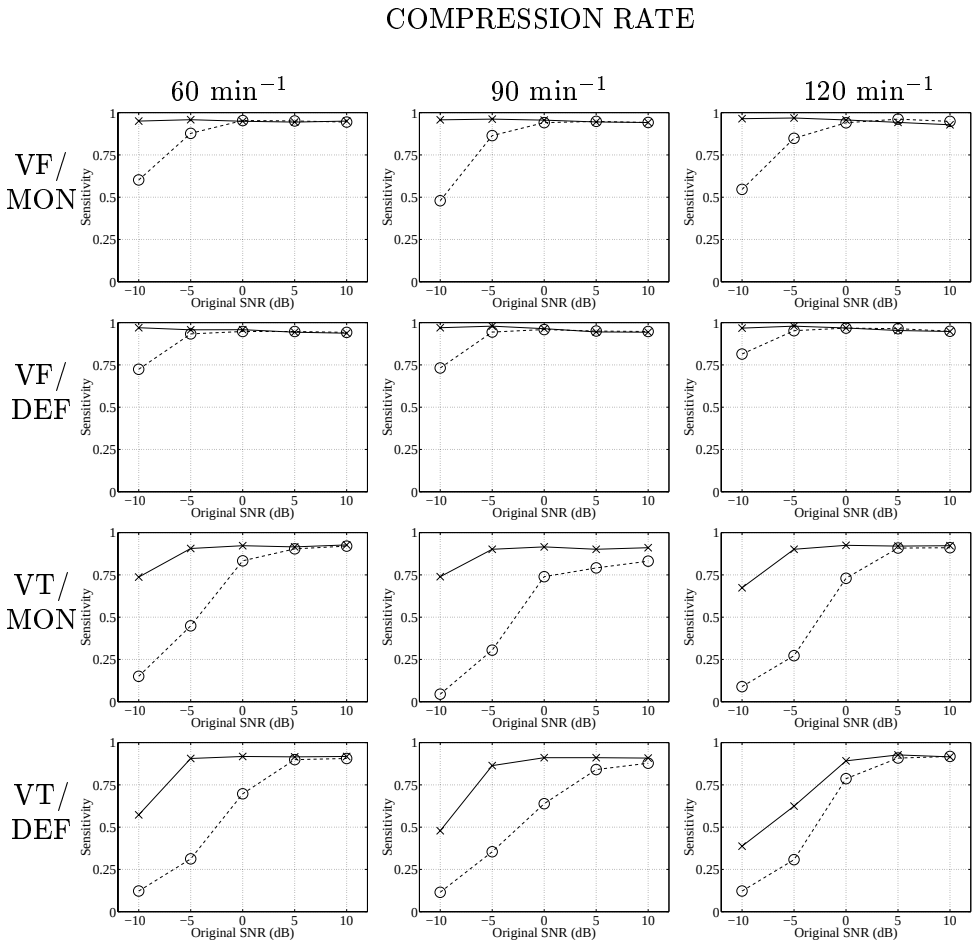


Figure 7.3: Sensitivity results using the Laerdal Heartstart classification algorithm. The classification is performed on ECGs having artefacts ('$\circ$') and the corresponding restored ECGs ('$\times$').

7.3 Discussion

In the presented work all filtering was performed at a 100 Hz sampling rate. At this rate the causal relationship between references (compression depth/ventilation) and resulting artefact was effectively instantaneous in the sense that (adaptive) 1-tap filters provided the best model. The use of a higher sampling frequency would allow for more detailed modeling of the sought relationship using longer filter responses and may possibly lead to better artefact removal.

A limitation of the current work was the low time resolution of the reference signals. They were originally sampled at 25 Hz, then up-sampled to 100 Hz. Recognizing the futility of modeling high resolution artefact signals with low resolution reference signals, a decision was made to band-limit the artefact signals. Before any experiments were conducted, all artefact signals were low-pass filtered with 10 Hz cut-off frequency. Future work should use the same sampling frequency on all signals to be used, and possibly higher than 100 Hz.

The sensitivity results in figure 7.3 clearly shows the benefit of the proposed method for artefact removal. Looking at the VF sensitivity results, it is evident that the unfiltered ECGs with artefacts are problematic for the classification algorithm at low SNR levels. For VF the performance results indicate that the ECGs recorded with the monitoring electrodes cause the algorithm to fail in more cases than the ECG recorded with the defibrillation electrodes. Knowing that the algorithm searches for organized rhythms in the ECG to satisfy the criterion for non-treat classification, this makes sense as the ECG recorded from the monitoring electrodes appears more organized than the corresponding recordings from the defibrillation electrodes. In our results performance degradation for the unfiltered VF with artefacts only starts at very low SNR levels – both for the monitor and the defibrillator case. This indicates that the artefacts themselves do not appear as clearly organized rhythms to the classification algorithm, which in consequence appears to be robust to this kind of artefact noise in VF. The explanation for this can be found by looking at the "VF+artefact" plot in figure 7.1, where the effect of mixing the organized artefact with the disorganized VF is a random cancellation/amplification of the periodical components of the artefact. This makes the mixture less organized than the artefact.

For the unfiltered VT performance, the difference in the corresponding VF cases are striking. The performance drops at a much higher SNR level. We consider two possible reasons for this. Firstly, the signal energy of VT is typically higher than for VF. Thus the artefacts are amplified to a higher factor than for VF to achieve a given SNR level. The classification algorithm operates

within a restricted dynamic area, and signals exceeding this area are clipped and classified as non-treat. Secondly, VT is an organized rhythm considered a treat rhythm for rates exceeding 180 beats per minutes (3 Hz). Otherwise, VT is considered as non-treat. The appearance of organized artefacts at lower rates (1, 1.5, and 2 Hz) causes the algorithm to interpret the rhythm as non-treat.

It is evident from both the SNR- and sensitivity evaluation that artefacts read from the defibrillation electrodes present a more difficult problem than artefacts read from the monitoring electrodes. In a realistic situation the defibrillation electrodes would be used for reading the ECG submitted to analysis. In this work we have shown how the artefact problem can be solved satisfactorily in a situation where the signals representing the main causes of the problem, i.e. the sources contributing the significant part of the artefact, are known. In our opinion, a solution to the problem with the defibrillation electrodes will be through identification of the cause of the part of the artefacts not caused by the precordial compressions and ventilation and further seek to represent this by a suitable number of reference signals to be incorporated as additional channels of our multichannel filter.

In conclusion the present performance results indicate that our proposed method for artefact removal might allow for rhythm analysis during CPR.

This would in effect allow CPR to continue, reducing the deterioration of the metabolic state of the tissues. This has the potential of dramatically changing the outcome of some cardiac arrest patients.

7.4 Summary

In this chapter we demonstrated how the adaptive filtering technique described in section 3.4 can be used to reduce the effect of CPR artefacts in ECG. By using the signals measuring compression depth and ventilation as references both individually we saw how the artefacts components caused by these factors are reduced. Applying these signals in combination, caused both components to be reduced. This indicates an optimal solution to the CPR artefact removal problem with signals representing the causes of the artefact components being used as references.

In this study, the effect of CPR artefact removal was evaluated for rhythm analysis. We strongly believe that the proposed filtering method would reduce interference during VF-analysis equally well.

Chapter 8

Predicting outcome of defibrillation

In this chapter we provide a method for predicting the outcome of defibrillation. We apply the principles for designing a decision support system as sketched out in chapter 3. In chapter 6 we indicated the potential benefit for resuscitation performance in the appliance of such a predictor, and in this chapter a quantification of this gain is given. This chapter is adapted from [38]. Brown et al. reported that they could predict ROSC with a sensitivity of 100 % and a specificity of 47.1% when applying centroid frequency (CF) and peak power frequency (PPF) of the VF in combination [10]. We question the reliability of their results both due to the study design and the small data set with only nine successful shocks (ROSCs) out of 128 shocks in 55 patients. The reliability of their results could have been confirmed if the prognostic criteria had been defined from one data set ("training set"), and the sensitivity and specificity derived from a new data set ("testing set") instead of both being determined from the same dataset.
In the present attempt to predict defibrillation outcome in human cardiac arrest by combining features of spectral characterization, we therefore split the data in training and testing sets and used classifier generalization techniques (as described in chapter 3.2) in an attempt to increase the degree of expected reliability. Among others, one of the combinations studied is that reported by Brown et al [10].

8.1 Methods

As in chapter 6 we used the data from Oslo already discussed in detail in section 4.1.1.

Characterisation

The ECG segments prior to shock were grouped according to shock outcome. Outcome was defined as ROSC if a palpable pulse was present in the post shock period (independent on duration). The rest of the shocks corresponded to No-ROSC and included conversions to electromechanical dissociation (EMD), asystole, VF (VF starting more than five seconds after the shock) or non-reset shocks (VF starting before five seconds after the shock). If the initial post shock rhythm was present for more than 10% of the duration of the interval, it was defined as the post-shock rhythm. Otherwise, the next rhythm was considered. This was done by an automated procedure, which handled all but 15 shocks. These failures were caused by illogics in the annotation structure.

The shock outcome prediction analysis was performed in two stages applying the feature extraction and pattern recognition methods described in chapter 3. Firstly, the ECG was spectrally characterized (feature extraction) and secondly, decision regions for shock outcome prediction were determined and evaluated.

Feature Extraction

The characterizing features were computed from the estimated power spectral density (PSD) of each ECG segment. As in previous studies of spectral characteristics of VF [13, 95, 73, 10, 12, 92, 97] we used the periodogram method for estimating the PSD as given by equation 3.41.

We attempted to discriminate between preshock ECG segments corresponding to ROSC and No-ROSC outcome. Based on previous VF-analysis with extraction of features carrying shape and placement information of the PSD to detect changes in myocardial metabolism [12], we computed the following features from the ECG segment PSD estimates:

- The centroid or median frequency (CF) (equation 3.42)
- The peak power frequency (PPF) (equation 3.43)
- The spectral flatness measure (SFM) (equation 3.44)
- The frequency band limited energy measurement (ENRG) (equation 3.45)

An alternative decorrelated feature set was generated by principal component analysis (PCA) transformation (chapter 3.3 and appendix A) [87]. The

Class	Subclass	Conversion to	Numbers	Prior ALS
ROSC	ω_1	Pulse rhythm	87	74%
No-ROSC	ω_2	Pulseless rhythm	337	86%
No-ROSC	ω_3	Isoelectric	98	66%
No-ROSC	ω_4	VF	35	77%
No-ROSC	ω_5	No conversion	311	87%

Table 8.1: Class division scheme. 110 of the 156 patients received shock treatment. The 87 shocks in the ROSC class were given to 46 of the patients. 35 of these shocks converted to sustained pulse rhythm, while 52 converted to pulse rhythm of limited duration (>20sec). 19 ROSCs were achieved after the first shock (four with prior ALS).

features were projected onto the eigenvectors which best represented the entire data set. Thus, the new decorrelated feature set was represented by the magnitudes of the projections along the eigenvectors. Before classification, a combination from either the original or the decorrelated feature set was placed into a feature vector, $\mathbf{v}$.

Classification

In the classifier each feature vector, $\mathbf{v}$, was considered as belonging to one of the K classes, $\omega_i, i = 1, \ldots, K$, which corresponds to the five shock outcome rhythms ($K = 5$). As shown in table 8.1, ω_1 and ω_{2-5} correspond to the ROSC and No-ROSC group respectively. As described in chapter 3.2, classification theory provides methods to retrospectively calculate decision regions for these defined classes from annotated data. The class membership of new data is decided by prospective comparison to these decision regions.

K decision regions, $R_i, i = 1, 2, \ldots, K$, were computed by assigning costs for the possible wrong decisions. A reject class, ω_{K+1}, was added to handle ambiguous or out-of-range patterns. Each R_i was calculated by selecting the minimum component of the risk vector $\mathbf{r}$ given in equation 3.7.

The classifier performance characteristics were expressed by the sensitivity (probability of positive prediction of ROSC outcome) and specificity (probability of negative prediction of No-ROSC outcome) given by equations 3.14 and 3.15 respectively.

We apply the principles given in chapter 3.2 to ensure that the classifier met the desired performance criterion, allowing us to specify a sensitivity for recognition of ROSC outcome.

For estimating the underlying statistics we applied the smooth histogram approach described in chapter 3.2, where the histogram bin resolution and kernel width governed by n_b and σ_{v_n} are the two key parameters for controlling the generality of the classifier.

Training and testing was done by applying the cross validation technique described in section 3.2.

Experimental setup

The ECG was sampled at 100 Hz with 8 bit resolution, and PSD estimated from segment lengths $L = 400$ zero padded to 512 samples. A hamming window was used for $w(n)$, while no averaging was used. Three feature sets were extracted with frequency ranges ($f_l - f_u$ Hz) 0-50 Hz, 0-25 Hz and 0-12.5 Hz. The spectral features produced in each of these experiments were v_{SFM}, v_{ENRG}, v_{CF} and v_{PPF}. The PCA transformation of these features gave the corresponding decorrelated feature set of v_{PCA_1}, v_{PCA_2}, v_{PCA_3} and v_{PCA_4}. The ECG immediately prior to defibrillation was analysed and the measurements grouped according to the post-shock rhythm for classifier design (table 8.1).

Classifiers were designed and tested using all possible combinations of spectral features and decorrelated features (table 8.2). The statistical functions were estimated using multidimensional histograms. The training sensitivity was set to $P_sns(\omega_1) = 0.95$ according to performance control technique described in chapter 3.2.3. Resolutions were adjusted according to setting n_b equal to 4, 8, 16, 32, 64 and 128 bins. For each of these resolutions the smoothness was varied by setting,σ_{v_n} equal to 0, 1, 5, 10, 15 and 20. The classifier parameters n_b and σ_{v_n} were discussed in chapter 3.2.5. This combination of changing bin size resolution and smoothness enabled a search for the classifier meeting the generality criterion which we defined to be that the test sensitivities and specificities should approach the training sensitivities and specificities to within a 5% tolerance range ($\tau = 0.05$, equation 3.32). In figure 8.1 the training and testing specificities and sensitivities are shown as functions of resolution and kernel width. Training with high bin resolution and narrow kernel width generates a classifier with 100% performance in both sensitivity and specificity as the result of overtraining as verified by the large deviation in sensitivity in test performance. Generality in sensitivity is achieved either by increasing the kernel width or using lower bin resolution, both resulting in lower specificity.

A full scale evaluation with respect to generality of the classifiers corresponding to all possible feature combinations (table 8.2) was performed. Finally, for a

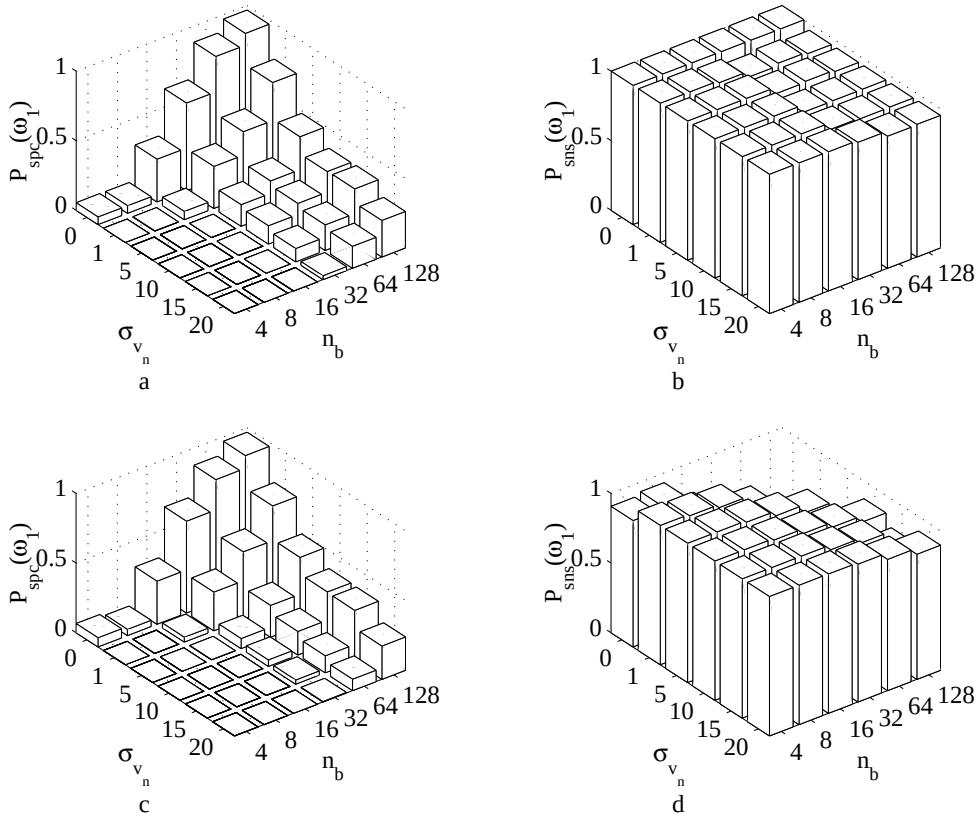


Figure 8.1: Effects of generalization. The training performances are shown in: a) Specificity increase with number of bins and decrease with increased kernel width. b) All classifiers are trained to 95% sensitivity. The corresponding testing performances are shown in: c) The specificities corresponds roughly with that of training. d) The sensitivities do not correspond at high bin numbers and low kernel widths. Generality is achieved at low bin numbers and high kernel widths.

Combination	Spectral	PCA
	v	**v**
1	v_{SFM}	v_{PCA_1}
2	v_{ENRG}	v_{PCA_2}
3	v_{CF}	v_{PCA_3}
4	v_{PPF}	v_{PCA_4}
5	v_{SFM} v_{ENRG}	v_{PCA_1} v_{PCA_2}
6	v_{SFM} v_{CF}	v_{PCA_1} v_{PCA_3}
7	v_{SFM} v_{PPF}	v_{PCA_1} v_{PCA_4}
8	v_{ENRG} v_{CF}	v_{PCA_2} v_{PCA_3}
9	v_{ENRG} v_{PPF}	v_{PCA_2} v_{PCA_4}
10	v_{CF} v_{PPF}	v_{PCA_3} v_{PCA_4}
11	v_{SFM} v_{ENRG} v_{CF}	v_{PCA_1} v_{PCA_2} v_{PCA_3}
12	v_{SFM} v_{ENRG} v_{PPF}	v_{PCA_1} v_{PCA_2} v_{PCA_4}
13	v_{SFM} v_{CF} v_{PPF}	v_{PCA_1} v_{PCA_3} v_{PCA_4}
14	v_{ENRG} v_{CF} v_{PPF}	v_{PCA_2} v_{PCA_3} v_{PCA_4}
15	v_{SFM} v_{ENRG} v_{CF} v_{PPF}	v_{PCA_1} v_{PCA_2} v_{PCA_3} v_{PCA_4}

Table 8.2: Tested combinations of candidate features. Vectors for spectral flatness (SFM), frequency band limited energy (ENRG), centroid frequency (CF) and peak power frequency (PPF) and their respective decorrelated features by principal component analysis (PCA) transformation.

given feature combination, the classifier with the best general performance was defined as the one corresponding to the highest average test performance, lowest bin resolution and narrowest kernel width requiring that the training and test performances satisfied the generality criteria (equations 3.32 and 3.33).

Statistical analysis

The comparisons between ROSC and No-ROSC was tested with the Wilcoxon Rank Sum test and presented as median (25th and 75th percentiles). A P-value less than 0.05 was regarded as significant. The classifier performance results are presented as the $mean \pm SD$ of the cross-validated sensitivities and specificities. These values were computed according to equations 3.27 and 3.28.

8.2 Results

The results for the spectral parameters and the corresponding decorrelated parameters are summarised for the different frequency ranges in table 8.3.

	0-50 Hz			0-25 Hz			0-12.5 Hz		
	ROSC	No-ROSC	P	ROSC	No-ROSC	P	ROSC	No-ROSC	P
SFM	0.3 (0.26,0.35)	0.32 (0.28,0.36)	*	0.43 (0.38,0.46)	0.42 (0.37,0.46)	NS	0.52 (0.48,0.55)	0.52 (0.48,0.56)	NS
$ENRG$ $x10^{-4}$	2 (1.5,2.5)	1.2 (0.81,1.7)	***	1.9 (1.4,2.3)	1.1 (0.73,1.6)	***	1.7 (1.2,2)	0.94 (0.64,1.4)	***
CF	8.8 (8.1,9.7)	8.7 (7.8,9.6)	NS	6.9 (6.3,7.4)	6.3 (5.7,6.9)	***	5.3 (4.9,5.6)	4.7 (4.2,5.2)	***
PPF	3.9 (3.2,5.1)	3.1 (2.4,4.3)	***	3.9 (3.2,5.1)	3.1 (2.4,4.3)	***	3.9 (3.2,5.1)	3.1 (2.4,4.3)	***
PCA_1 $x10^{-3}$	5.9 (0.74,11)	-2.2 (-6.1,2.7)	***	5.5 (0.72,10)	-2.1 (-5.8,2.7)	***	5.2 (0.24,8.9)	-2.1 (-5.2,2.5)	***
PCA_2	0.99 (-0.23,1.8)	-0.21 (-1.4,1)	***	0.7 (-0.47,1.9)	-0.29 (-1.3,0.92)	***	0.62 (-0.36,1.8)	-0.32 (-1.3,0.94)	***
PCA_3	0.02 (-0.51, 0.64)	-0.13 (-0.65, 0.45)	NS	0.19 (-0.26, 0.65)	-0.04 (-0.45, 0.33)	**	0.16 (-0.09, 0.45)	-0.02 (-0.31, 0.26)	**
PCA_4	-0.01 (-0.02, 0.01)	0 (-0.01, 0.01)	**	0 (-0.03, 0.01)	0 (-0.02, 0.02)	*	-0.01 (-0.04, 0.02)	0 (-0.02, 0.03)	*

Table 8.3: Distribution of features. Spectral flatness measure (SFM), frequency band limited energy (ENRG), centroid frequency (CF) and peak power frequency (PPF) and their respective decorrelated features by principal component analysis (PCA) transformation. P-values NS: P$\geq$0.05; *: P$<$0.05; **: P$<$0.01; ***: P$<$0.0001.

Feature combination	Freq. Range	$\hat{P}_{sns}(\omega_1)$	$\hat{P}_{spc}(\omega_1)$
$\begin{bmatrix} v_{CF} & v_{PPF} \end{bmatrix}$	0-12.5 Hz	$92 \pm 2\%$	$27 \pm 2\%$
$\begin{bmatrix} v_{PCA_1} & v_{PCA_2} \end{bmatrix}$	0-25 Hz	$92 \pm 2\%$	$42 \pm 1\%$

Table 8.4: Performance characteristics.The sensitivity and specificity for predicting ROSC for the reference classifier [10]: centroid frequency (CF) and peak power frequency PPF at 0-12.5 Hz and the highest performing classifier based on principle component analysis (PCA) decorrelation at 0-25 Hz.

The test performance results of the classifiers meeting the generality criterion are shown in figure 8.2. The highest classifier performance is achieved with decorrelated feature combination no 5 with the frequency range set to 0-25 Hz (table 8.1, 8.3 and 8.4). The highest performance using the original spectral features corresponded to the combination of all four features in frequency range 0-12.5 Hz.

The performance of the reference classifier, $\mathbf{v} = \begin{bmatrix} v_{CF} & v_{PPF} \end{bmatrix}$, for comparison with earlier work, and the highest performing classifier $\mathbf{v} = \begin{bmatrix} v_{PCA_1} & v_{PCA_2} \end{bmatrix}$ are listed in table 8.4 and the class specific PDFs with corresponding decision regions for these two classifiers are shown in figure 8.3.

The highest performing classifier $\mathbf{v} = \begin{bmatrix} v_{PCA_1} & v_{PCA_2} \end{bmatrix}$ shows a clearer distinction between ROSC and No-ROSC than the reference classifier $\mathbf{v} = \begin{bmatrix} v_{CF} & v_{PPF} \end{bmatrix}$

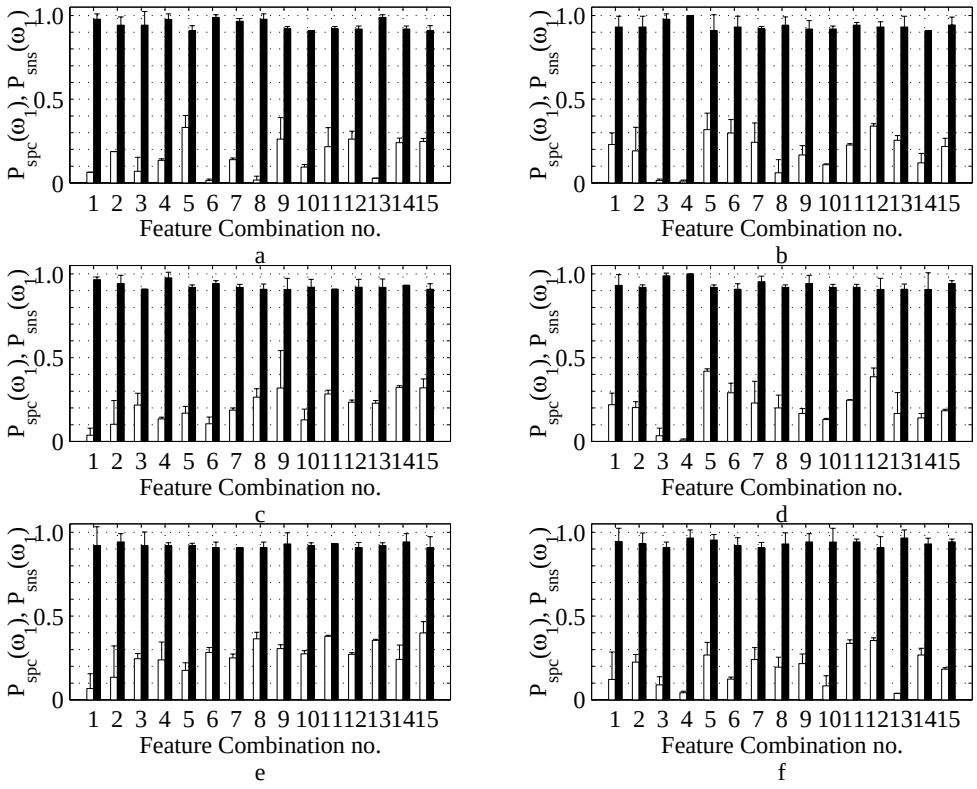


Figure 8.2: Test performances for generalised classifiers for all feature combinations (table 8.2) and frequency ranges. Black bars = sensitivities. White bars = specificities. Performance of the different frequency ranges: Combinations of original spectral feature sets, a) 0-50 Hz, c) 0-25 Hz and e) 0-12.5 Hz. Combinations of decorrelated spectral feature sets, b) 0 -50 Hz, d) 0 -25 Hz and f) 0 -12.5 Hz

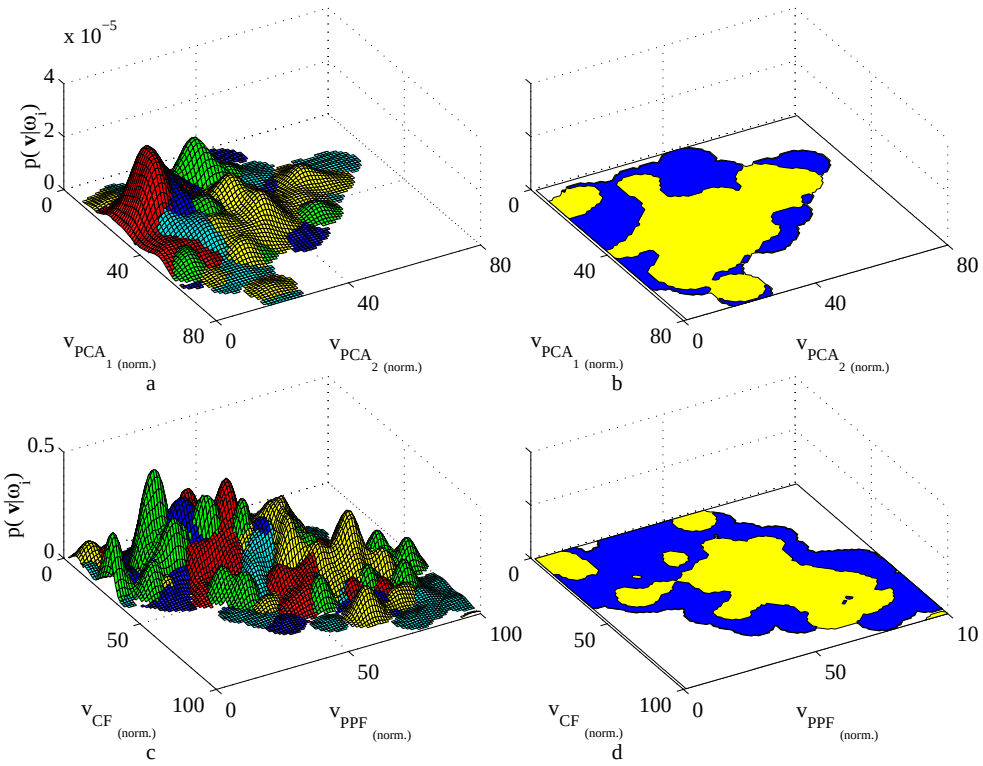


Figure 8.3: Estimated class-specific probability density functions (PDF) and corresponding decision regions. The feature values have been normalised according to equation 3.37. The subclass PDFs are identified by the following color scheme: yellow=ROSC, blue=EMD, red=asystole, green=VF reappearing >5 sec after the shock, cyan=VF reappearing within five seconds after the shock (non-reset shocks), white=reject. The coloring in the decision region corresponds to yellow=ROSC, blue=No-ROSC, white=reject. Highest performing classifier: a) probability densities, b) decision regions Reference classifier: c) probability densities, d) decision regions

	No-Shock	Shock	Total
No-ROSC	328	453	781
ROSC	7	80	87
Total	335	533	868
	Negative predictive factor	Positive predictive factor	
	$\frac{328}{328+7} = 0.98$	$\frac{80}{80+453} = 0.15$	

Table 8.5: Positive and negative predictive values. The positive and negative predictive values for achieving ROSC using the highest performing classifier $\mathbf{v} = [v_{PCA_1} v_{PCA_2}]$ based on principal component analysis transformation at 0-25 Hz.

were there is more intermingling of the classes. The highest performing classifier is based on PCA decorrelation and dimension reduction to two features and achieves a sensitivity of $92 \pm 1\%$ and specificity of $42 \pm 1\%$ in testing, or a positive predictive value of 0.15 and a negative predictive value of 0.98 (table 8.5).

The frequency ranges 0-25 Hz and 0-12.5 Hz are best suited for discriminating ROSC from No-ROSC outcomes (table 8.3),and the spectral flatness measure is the least suitable individual feature for all frequency ranges. Although the three other spectral features seem promising , the results of the decorrelated features indicate only two features as being significantly different when grouped according to ROSC and No-ROSC outcome. This indicates that there is redundant information in the original feature set. For example, which we will discuss in the following, PPF does not add much relevant information that is not already expressed by CF.

The highest performing single feature spectral classifiers are centroid frequency (CF) and peak power frequency (PPF) in the low frequency range, but combining these two features does not improve the results. For the single decorrelated feature classifiers, the two principal ones give the best results for all frequency ranges. Combining decorrelated features improves the performance significantly when the two mid frequency range principal features are combined. Including more than two decorrelated features does not improve the performance further.

8.3 Discussion

In this study of 868 shocks in 156 patients it is possible to partly predict the outcome of the shock to either ROSC or No ROSC by analyzing four

spectral features of the preshock ECG with improved results by combining, decorrelating and reducing the features.

We have further demonstrated how classification methodology allows the combination of features with an increase in classifier performance compared to classifying features individually. We also have showed how decorrelation by principal component analysis allows dimensional reduction of the feature set with no decrease in performance compared to a combination of the complete feature set.

Human studies have shown that the percentage of shocks resulting in ROSC is very low [50, 6, 100]. In the recent study from Oslo described in chapter 6 only 10% of all shocks resulted in a pulsegiving rhythm. Most shocks are thus individually futile. Based on the present results, 42% of the unsuccessful shocks (328 of 781) could have been avoided and a period of chest compressions, ventilations and vasoactive drugs could have been administered before a new defibrillation attempt. Animal studies have shown that this may be favorable [71, 7], and the recent study from Cobb et al indicated that this can improve the outcome of patients [24]. It would minimise the detriment of the "hands-off" intervals, where the vital organs are without perfusion thus reducing the possibility to restore spontaneous circulation and/or a recovery with intact neurologic status. In addition, the number of shocks should be kept to a minimum as repetitive shocks and total electric power are injurious to the already ischemic myocardium, and increase the severity of postmyocardial dysfunction [115]. Moreover, as the spectral characteristics of the VF have been reported to reflect myocardial perfusion [13, 95, 73], the defibrillator might also potentially guide the CPR attempt as the myocardial perfusion depends on compression force, rate and duration [66, 49, 74, 45]. We explore these monitoring aspects further in chapter 9.

On the other hand, seven shocks that resulted in a pulse giving rhythm would not have been administered. These shocks presumably would have been given later, if CPR changed the characteristics of the VF. The effects of this could not be evaluated.

The present study demonstrates how a general classifier can be designed by cross-validation which allows training and testing on independent data sets in combination with different resolutions and kernel widths in the estimation of the statistics describing the features according to the principles given in section 3.2. This method gives an indication on how well the classifier will perform when challenged with new data in the future.

In a similar study of 128 shocks in 55 patients with only nine successful shocks (defined as a conversion of VF to a supraventricular rhythm with a palpable

pulse or blood pressure of any duration within 2 min of the shock without ongoing CPR), Brown et al [10] extracted four parameters from the recorded ECG (centroid and peak power frequency and average segment and wave amplitude). The combination of CF and PPF was reported as having the best predictive potential with a sensitivity of 100% and a specificity of 47.1% [10]. We had a poorer predictive potential of combining the same two features with a sensitivity of $92 \pm 2\%$ and specificity $27 \pm 2\%$. We believe the results achieved with our generalized classifier are more realistic due to a much larger database and the use of independent testing and generalization not done by Brown et al [10]. They generated the sensitivity and specificity from the same data that the threshold values were computed from with no independent evaluation.

Noc et al [73] recently reported in pigs that maximum and mean VF amplitude and dominant VF frequency were all acceptable shock outcome predictors. Noc et al did derive the threshold values in one group and tested these in a separate validation group, but had different results in the two groups, indicating that the results might not be reliable [73]. Our results indicate that the results from Brown et al would have experienced the same if their threshold values had been tested on a new data set [10].

Our method includes both independent testing and generality to avoid these problems. To ensure reliability of the results the data are randomly split into two sets of equal size with equal representation of each class. Decision regions and training performance for ROSC and No-ROSC prediction are computed from half the data while the other half is used to reevaluate the decision regions by computing the corresponding test performance.

We further compute a set of decision regions for each feature combination by varying parameters in the classifiers and seeking the decision region, corresponding to a match in training and testing performance within a tolerance of $\pm 5\%$.

There are some limitations in the present study. Firstly, the number of observations in the ROSC group is low. Secondly, in the cross-validating processing of the data, the test performances are considered in the design of the classifiers to choose the generalizing parameters. Ideally, a final evaluation should have been performed on yet another data not influencing the design process. Thirdly, we only use one type of classifier, the histogram method. To get even more reliable results, the experiments should be repeated with other types of classifier.

8.4 Summary

As we have pointed out in chapter 2.2 and confirmed in chapter 6, designing a method whereby to reduce the number of shocks given in the resuscitation of out-of-hospital cardiac arrest patients, would be of benefit to the patient.

In this chapter we have quantified this benefit by designing a defibrillation outcome predictor and evaluating a variety of possible classifiers on the Oslo data. Our results clearly indicates a benefit.

Spectral characterization of VF can be of clinical importance in the treatment of cardiac arrested patients if it can be incorporated in the software of defibrillators. We have demonstrated a method for developing an outcome predictor for defibrillation attempts in patients with out-of-hospital cardiac arrest although the sensitivity of $92 \pm 1\%$ and specificity of $42 \pm 1\%$ in the present study are not satisfactory for clinical use. Other features should therefore also be investigated to add discriminative power to the feature set.

Chapter 9

Monitoring the probability of defibrillation success

In chapter 2.2 we discussed the need for a parameter monitoring the effectiveness of CPR. In this chapter we suggest a method for computing such a parameter, the probability of defibrillation success $P_{ROSC}(\mathbf{v})$. This chapter is adapted from [39].
In this study, the main objective was to establish a method for devising a variable for monitoring of CPR efficacy. The monitoring concept is closely related to the prediction of defibrillation outcome. A good predictor discriminates between different levels of resuscitability, which also should be salient in a monitoring variable. In chapter 8 we showed how the power of a classifier for predicting defibrillation outcomes can be improved by decorrelating and combining ECG derived features. These results indicate that more than one feature may be needed for the prediction. This is a problem because direct use of multiple prediction features would give a multidimensional monitoring variable. The use of many features would add complexity to the treatment situation. Therefore, we suggest a method for expressing this multi-dimensional information in a single reproducible variable reflecting the probability of defibrillation success.

9.1 Methods

As in chapters 6 and 8 we used the data from Oslo already discussed in detail in chapter 4.1.1.
We termed the class (table 8.1) corresponding to the 87 shocks which caused ROSC by ω_{ROSC}. The other classes corresponding to the shocks failing to cause ROSC was commonly termed by $\omega_{No-ROSC}$.

Characterization

We applied the methodology described in the chapter 8 for selecting the optimal combination from a set of ECG characterising features. We used the combination of the two most expressive PCA features,$\mathbf{v} = \begin{bmatrix} v_{PCA_1} & v_{PCA_2} \end{bmatrix}$,in the frequency area 0-25 Hz, which in chapter 8 was identified to achieve the highest classifier performance.

The Probability of Defibrillation success

The methodology we used for prediction and the results reported in chapter 8 suggest that more features should be added to improve the performance. For the highest predictive power, the information from each and all of these features should therefore be presented to the rescuer. One of the drawbacks of presenting such a combination of features to the rescuer is that each added feature would represent added complexity in an already complex treatment situation. We therefore propose a general methodology for expressing the information in the shock outcome predictor feature vectors of higher dimensions into one single meaningful variable suitable for monitoring. This variable expresses the probability of defibrillation success as seen through the feature vectors.

D features are combined into a feature vector characterising the ECG prior to defibrillation. The observed features are divided into the sets V_{ROSC} and $V_{No-ROSC}$ which correspond to observations from ω_{ROSC} and $\omega_{No-ROSC}$, respectively. Thus we have n_{ROSC} D-dimensional feature vectors in V_{ROSC} and $n_{No-ROSC}$ feature vectors in $V_{No-ROSC}$. These sets are the training data representing the knowledge of the relationship between the information inherent in the preshock ECG and shock outcome. We use this knowledge in the construction of the variable, which is supposed to monitor resuscitability, the probability of defibrillation success. Each feature axis is split into n_b parts so that the D-dimensional histogram representing the feature space is divided into n_b^D bins, $\{b_i\}_{i=1,\ldots,n_b^D}$, of equal volume. Let n_{ROSC_i} and $n_{No-ROSC_i}$ correspond to the number of observations from V_{ROSC} and $V_{No-ROSC}$ respectively, in b_i. We define the local estimate of the probability of defibrillation success for a D-dimensional feature vector $\mathbf{v}$ observed in bin b_i as

$$P_{ROSC}(\mathbf{v}) = \frac{n_{ROSC_i}}{n_{ROSC_i} + n_{No-ROSC_i}}. \tag{9.1}$$

According to [85], this corresponds to estimating the a posteriori probability function $P(\omega_{ROSC}|\mathbf{v})$. This signifies that

$$P_{ROSC}(\mathbf{v}) = \hat{P}(\omega_{ROSC}|\mathbf{v}). \tag{9.2}$$

This is very useful as it allows us to use the established techniques for estimating the a posteriori probability functions, given an observation $\mathbf{v}$.

The number of observations available for forming the classifiers is limited, and we have to extrapolate the part of the feature space not covered by the observations to get a continuous estimate. We use the smooth histogram method of chapter 3.2.

Experimental setup

In chapter 8, $\mathbf{v} = \begin{bmatrix} v_{PCA_1} & v_{PCA_2} \end{bmatrix}$ was identified as the feature combination corresponding to the highest performing classifier for shock outcome prediction with generality satisfied. Using this feature combination and the corresponding classifier settings, we computed $P_{ROSC}(\mathbf{v})$. Instead of using the local estimate method proposed above, $P_{ROSC}(\mathbf{v})$ was calculated from estimates of the a priori class probabilities and class-specific PDFs according to Bayes rule given by equation 3.6. These were identical to the functions used for outcome prediction in chapter 8.

For outcome prediction based on one dimensional feature vectors, analysis of the results in chapter 8 identified the ones corresponding to the most expressive PCA feature in the $0-25$ Hz frequency area and the centroid frequency in the frequency area $0-12.5$ Hz features to be the best feature vectors for decorrelated and original features respectively. In a monitoring device using one-dimensional outcome predictor feature vectors directly, these features would be natural candidates, so each of these was recalculated for individual comparison to $P_{ROSC}(\mathbf{v})$.

We also used the preshock derived $P_{ROSC}(\mathbf{v})$ function to monitor two complete patient records to be able to study its behaviour during treatment. ECG blocks of four seconds duration and one second overlap were extracted from the ECG so that the monitoring variable was updated every third second. From these blocks we computed feature vectors, $\mathbf{v}$, and further calculated $P_{ROSC}(\mathbf{v})$. From these recordings of $P_{ROSC}(\mathbf{v})$, we selected to study the ones corresponding to the last 15 seconds of ECG prior to each defibrillation.

Statistical analysis

The comparisons between ROSC and No-ROSC were tested with the Wilcoxon Rank Sum test. A P-value <0.05 was regarded as significant.

9.2 Results

The function surface of $P_{ROSC}(\mathbf{v})$

In figure 9.1, the resulting $P_{ROSC}(\mathbf{v})$, as based upon the preshock feature vectors of the data material in V_{ROSC} and $V_{No-ROSC}$, is shown. The function surface attains its highest values in the area where the feature vectors from V_{ROSC} discriminates the best from those in $V_{No-ROSC}$. The low function values of $P_{ROSC}(\mathbf{v})$ correspond to the areas dominated by feature vectors from $V_{No-ROSC}$. The surface covers only a limited part of the feature space. This is the effect of interpolating the associated probability density functions with a restricted gaussian kernel function.

Comparison to single feature representations from outcome prediction

The parameter distributions are illustrated in the box plots of figure 9.2. The ROSC distribution is compared to the No-ROSC distribution for each parameter. The distributions for v_{CF}, v_{PCA_1} and $P_{ROSC}(\mathbf{v})$ are shown in part a), b) and c) of the figure, respectively. For each of these three variables the parameter values corresponding to ROSC was found to be significantly higher than for the No-ROSC values ($P<0.0001$ for all three comparisons). The P-value corresponding to the test of the $P_{ROSC}(\mathbf{v})$ values was actually around 10^{-300} times smaller than either of the two other P-values. This indicates that $P_{ROSC}(\mathbf{v})$ is a significantly better monitor variable than the other two individual feature vectors: v_{CF} computed from the $0-12.5$ Hz frequency range and v_{PCA_1} computed from the $0-25$ Hz frequency range.

Monitoring patient records using the probability function

The retrospective use of $P_{ROSC}(\mathbf{v})$ as a monitor of CPR efficacy is illustrated in figure 9.3 and figure 9.4. Figure 9.3 shows the last 15 seconds of analysis and charging periods prior to each of 8 shocks in a patient who finally converted to ROSC. The ECG tracings show how the VF is "coarsened" as an effect of the CPR provided between the shocks, and the probability function values (*100%) reflect this coarsening by an increase in value towards the points of successful conversions.

A similar plot shows five unsuccessful shocks in another patient (figure 9.4). There is no evidence of VF "coarsening" in these tracings. The function values decrease throughout the episode.

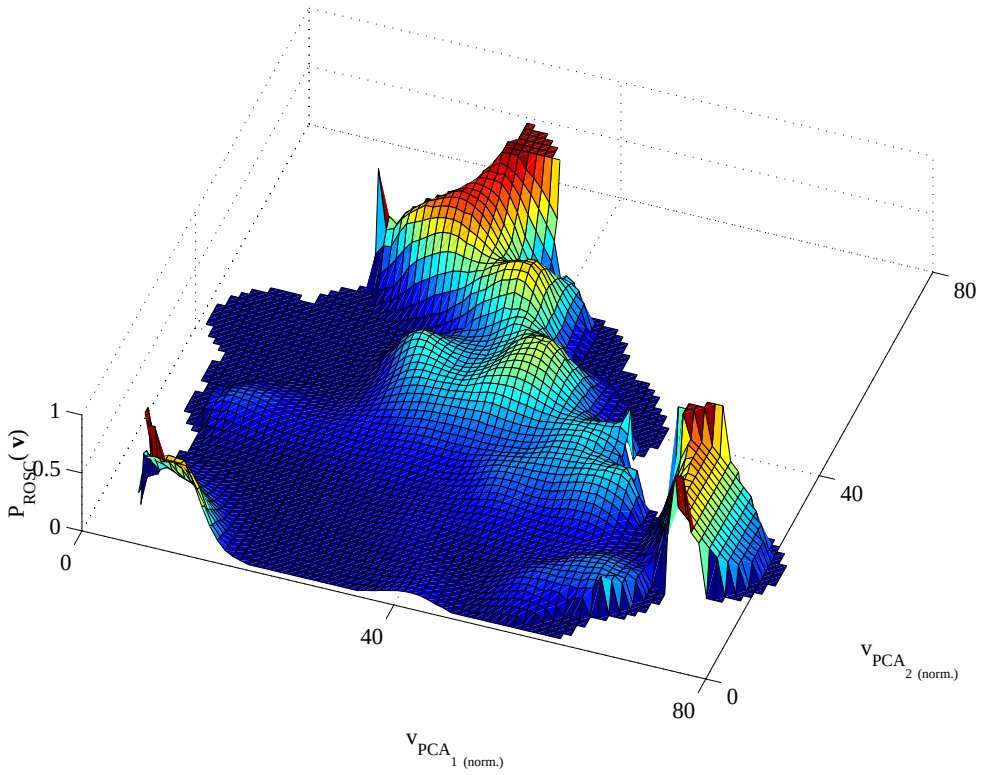


Figure 9.1: The function "Probability of ROSC", $P_{ROSC}(\mathbf{v})$, derived from 883 defibrillation attempts in 156 patients. $P_{ROSC}(\mathbf{v})$ is calculated from the knowledge of defibrillation outcome, ROSC or No-ROSC, and the preshock VF spectral feature observations of $\mathbf{v} = [v_{PCA_1}, v_{PCA_2}]$. In each bin the amplitude on the function surface corresponds to the number of ROSC outcomes divided by the total number of defibrillation attempts in the bin. The amplitude represents the estimate of the probability of defibrillation success for a feature vector, $\mathbf{v}$, observed in that bin. The colour scheme for the function values are approximately as follows: dark blue=0-0.25, light blue=0.25-0.5, yellow=0.5-0.75, red=0.75-1.

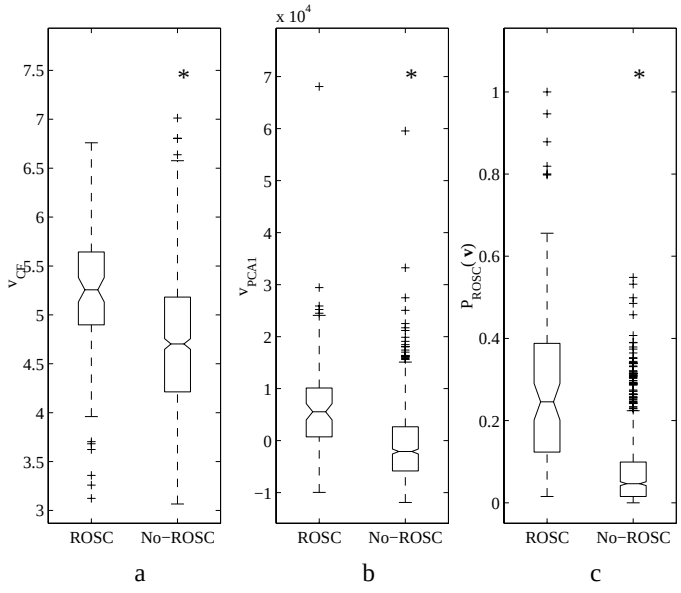


Figure 9.2: The statistical distributions of the preshock ECG blocks represented by a) v_{CF} computed from the 0-12.5 Hz frequency range, b) v_{PCA_1} computed from the 0-25 Hz frequency range and c) $P_{ROSC}(\mathbf{v})$. All feature sets were significantly different for ROSC and No-ROSC measurements. *)$P < 0.0001$. Boxplot description is presented in figure 6.1.

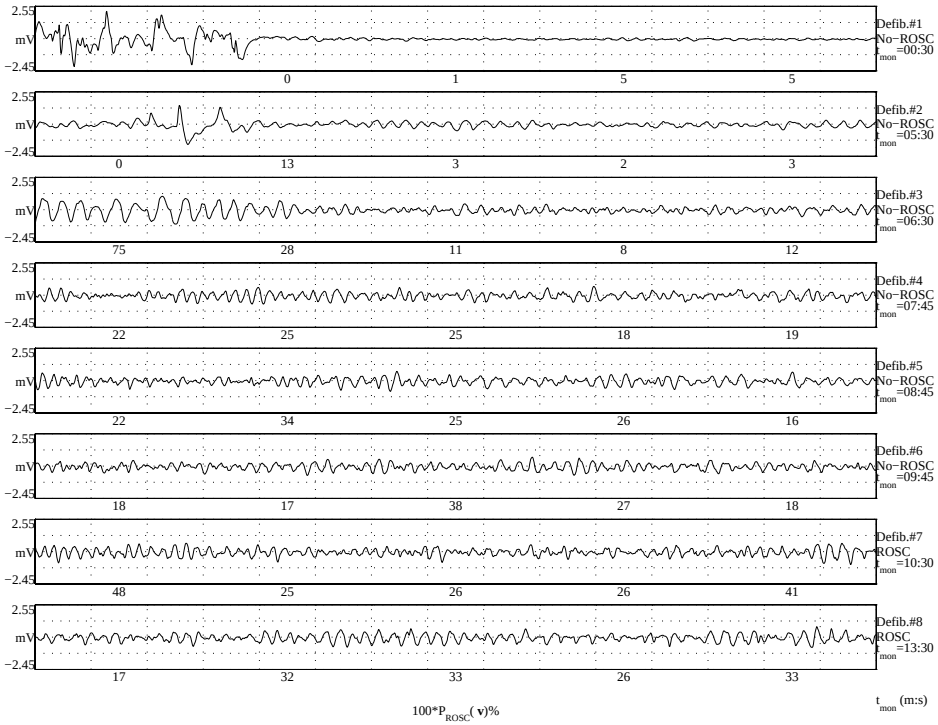


Figure 9.3: 15 second ECG tracings prior to eight shocks in a patient who was successfully defibrillated with attempts #7 and #8. After shock # 7 the pulsegiving rhythm was not sustained, but sustained after shock # 8. $P_{ROSC}(\mathbf{v})$ was retrospectively calculated and presented here as ($100P_{ROSC}(\mathbf{v})$ %) updated every three seconds. t_{mon} : time in minutes:seconds from the defibrillator was connected until a defibrillation attempt.

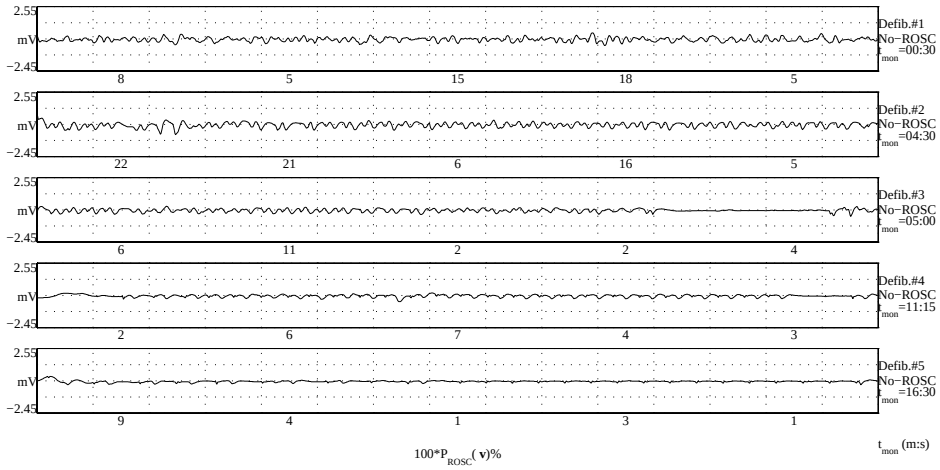


Figure 9.4: 15 second ECG tracings prior to five shocks in a patient who did not convert to ROSC. $P_{ROSC}(\mathbf{v})$ was retrospectively calculated and presented here as ($100P_{ROSC}(\mathbf{v})$ %) updated every three seconds. t_{mon} : time in minutes:seconds from the defibrillator was connected until a defibrillation attempt.

9.3 Discussion

In this study we have proposed and demonstrated a method for representing multivariate information as a single variable - the probability of successful defibrillation ($P_{ROSC}(\mathbf{v})$) - suited for monitoring the effectiveness of CPR. This indicates a way to guide CPR towards delivery of shock at a time optimal for successful defibrillation outcome.

As already mentioned, the methodology for calculating $P_{ROSC}(\mathbf{v})$ is general. The feature vector,$\mathbf{v}$, can be of any dimension and content, which means that there are a multitude of possible $P_{ROSC}(\mathbf{v})$ functions. Which features to incorporate depend on the capability of the feature vector to predict the shock outcome. The first step on the way to constructing a $P_{ROSC}(\mathbf{v})$ function is to find a set of feature candidates. The next step is to determine which features to select for $\mathbf{v}$. A bad choice for $\mathbf{v}$ will consequently provide a worthless $P_{ROSC}(\mathbf{v})$.

In chapter 8 we determined the combination within a set of features with the highest prediction sensitivity and specificity, and used this as the basis for the methodology in the present study. To improve the usefulness of $P_{ROSC}(\mathbf{v})$

further, one should test the addition of new features to the existing set and search for the combination giving the highest prediction performance. $\mathbf{v}$ is, in a sense, the glasses through which the ECG is seen.

Another aspect to consider is that $P_{ROSC}(\mathbf{v})$ is an empirical function based on actual defibrillation attempts in patients. The reliability and scope of the function is dependent on the data available. In figure 9.1 the function appears to behave strangely in areas with few data points. This emphasises the importance of a large representative database for computing $P_{ROSC}(\mathbf{v})$.

We applied the smooth histogram technique described in chapter 3.2.5 for computing $P_{ROSC}(\mathbf{v})$. Each feature axis was divided into 128 bins so that the function was represented by a 2-dimensional matrix with $128^2 = 16384$ elements. The effect of adding features is an increase in the number of elements needed to represent the function. If D feature axes are divided into n_b bins, the number of elements needed is n_b^D. This problem, the curse of dimensionality, makes the present technique for representing $P_{ROSC}(\mathbf{v})$ impractical. To avoid this problem, one should consider using other techniques for function estimation like multilayer perception neural networks or radial basis function networks.

It would have been much simpler if single features such as centroid frequency or amplitude alone could reliably predict the success rate and thus provide a good monitoring variable as previously indicated in some studies [12, 109]. Brown et al [13] thus reported that centroid frequency alone predicted ROSC with 92 % specificity and 100% sensitivity in pigs. In most of these studies there is a varying combination of factors that limits their usefulness in patients:

1. The frequency spectra of VF in animals such as pigs [12, 13, 73] differ significantly from those in humans, making an extrapolation difficult.

2. In many studies [13, 95, 10, 14] the specificity and sensitivity were determined from the same data set that had been used to determine the predicting feature initially. To increase the validity and reliability the prognostic criteria should be defined from one data set and the specificity and sensitivity derived from a new data set.

3. To be an effective guide during CPR the predictor must give the clinician advice before each individual shock.

Point one is illustrated by the results from Brown et al [10] in 56 patients with out-of-hospital cardiac arrest, where the specificity of approximately 20 % and sensitivity of 100 % for centroid frequency as a predictor was much lower than in pigs.

From 265 out-of-hospital cardiac arrest patients, Callaham et al [14] reported that the maximum peak-to-through amplitude (AVF) of VF during the initial 3-7 seconds of pre-defibrillation analysis predicted ROSC with a sensitivity of 89 % and a specificity of 33 %. This prediction was based only on the initially determined VF, and not directly related to each individual shock as in our prediction analysis described in chapter 8. They also reported that the patients developing a pulse had an increase in the development of AVF with refibrillations compared to a decrease for those who did not [14].

When Brown et al tried to predict the outcome of each countershock independent of any clinical variable [10], as we did in our prediction study described in chapter 8, AVF only achieved 10 % specificity at 90 % sensitivity when each countershock was considered independently. The predictive value of the combination of centroid frequency (CF) and peak power frequency (PPF) for ROSC was higher with a specificity of 47.1 % and a sensitivity of 100 %.

Noc et al [73] designed a defibrillation predictor in domestic pigs from a stepwise multiple regression on AVF and PPF to coronary perfusion pressure (CPP). This correlated better to the CPP than the individual features and had greater predictive power on shock outcome with 80 % specificity at 100 % sensitivity. When the predictive value was validated on another independent animal group with corresponding threshold values, however, the specificity fell to 44 % at 100% sensitivity [73].

In our outcome prediction analysis (chapter 8), ENRG, a somewhat different energy measurement than AVF indicated a sensitivity of 94% ($\pm$ 5%) and a specificity of 13% ($\pm$ 19%), while CF alone achieved a sensitivity of 92% ($\pm$ 8%) and a specificity of 25% ($\pm$ 3%). The combination of CF and the energy measurement achieved a sensitivity of 91% ($\pm$ 3%) and a specificity of 36% ($\pm$ 4%). All these features were extracted from the 0-12.5 Hz frequency range. One of our conclusions was that increased predictive accuracy could be obtained by combining features.

The information reflected in the ECG features can be clinically useful if the predictive value is high with high reliability. As time is critical in the cardiac arrest situation, and an additional monitor represents one extra element that the rescuer must relate to, the information relayed must be simple. By selecting feature combinations and evaluating the performance, the reliability of the classifier can be improved (chapter 8), and we have here demonstrated how the number of variables in a monitoring device can be reduced to one manageable unit. We have also indicated how $P_{ROSC}(\mathbf{v})$ could be used as a CPR monitor by visualising the computed function value retrospectively in a human patient. The values were updated every third second as illustrated in

figure 9.3 and figure 9.4. We think these figures illustrate how our monitoring unit reflects the positive changes with ongoing CPR in the patient with ROSC, while there was a decrease in the probability of defibrillation success in the No-ROSC patient. It is important to emphasise that these examples are used for illustration, and are not intended as scientific "proof".

Based on findings from 100 cardiac arrest patients Paradis et al [76] reported that a certain level of CPP was needed for ROSC. In accordance with this, animal studies have indicated that chest compressions prior to the first defibrillation attempt can improve the myocardial condition and increase the resuscitability in prolonged VF [71, 7]. In 639 patients with prehospital cardiac arrest found in VF [24], Cobb et al recently reported that chest compressions prior to defibrillation appeared to improve the outcome. This was true for the whole material and for the subset of patients with more than four minutes of cardiac arrest before the EMS team arrived, but not for 0- 4 minutes compared to a historic control group[24]. They speculated that this was due to an improvement in the myocardial condition with coronary blood flow secondary to the chest compressions. We further speculate that the outcome might improve if the timing of a defibrillation attempt in the individual patient is guided by information on the myocardial condition. This should enable a defibrillation attempt as soon as the myocardial condition is favourable and at the same time limit the number of no-flow periods during unsuccessful defibrillation attempts.

It has been reported that 80 % of patients who survive are defibrillated by one of the first three shocks [50, 25, 2, 106]. In agreement with others [50, 6] we have recently reported from Oslo (chapter 6) that only 10 % of 883 individual shocks resulted in a pulsegiving rhythm, and only 4 % of the shocks resulted in sustained spontaneous circulation. We therefore speculate that most of these shocks are administered to a myocardium that is not in an optimal condition for successful resuscitation. In these situations chest compressions with resulting coronary perfusion should increase the resuscitability [76]. The use of $P_{ROSC}(\mathbf{v})$, incorporated into the defibrillator, could potentially improve the quality of the chest compressions by direct feedback, and thereby the myocardial perfusion, and indicate the optimal timing of defibrillation. This would also reduce the number of unsuccessful shocks damaging the myocardium [115], and the duration of "hands off" intervals without vital organ perfusion during resuscitation attempts. We do not know at present if and for how long chest compressions should be performed before delivery of a shock in an individual patient, or between shocks in prolonged VF, but there are probably large inter- and intraindividual differences, which could be demonstrated by the $P_{ROSC}(\mathbf{v})$.

9.4 Summary

We have demonstrated by analysing the ECG prior to 883 shocks from 156 patients with prehospital cardiac arrests, that the principle of monitoring $P_{ROSC}(\mathbf{v})$ might be useful as guidance during CPR to reflect the myocardial condition and optimise shock timing for the patient. This could enable individualised CPR, the avoidance of unnecessary shocks, and increase the time of vital organ perfusion during cardiac arrest treatment.

Chapter 10

Conclusion

Here follows a summary of the major contributions and conclusions from this work.

10.1 Major contributions of this work

We have formulated a decisions support system for VF analysis to guide therapy during treatment of cardiac arrest in out-of-hospital patients.

1. In chapter 3.1 a framework for a decision support system is proposed. We emphasise the importance of reliability and validity. In chapter 3.2 we emphasise the importance of combining features for improved performance as compared to using features individually. Further we propose the use of cross-validation techniques as an instrument to ensure reliability in the performance results in the design of the decision support system. The applicability of these principles are demonstrated in the experiments presented in chapter 8 (to be published [38]).

2. We use the principles laid down in classification theory to formulate a method to control the sensitivity performance (Receiver Operator Characteristics analysis) by using the a posteriori probability function values of our training data features in section 3.2.

3. A solution for CPR artefact removal using adaptive filters with signals representing the cause of artefact components as references to the filter is proposed in chapter 3.4. In chapter 7, experiments were we apply references corresponding to combining the compression depth and ventilation

references indicates the potential of this method were fixed coefficient filter methods fail in separating CPR artefacts from the VF part of the recorded ECG (to be published [63, 1]).

4. We propose a method to design a parameter to monitor CPR efficacy in chapter 9 by showing that the a posteriori probability function of ECG extracted features prior to successful defibrillation reflects the probability of defibrillation success. In the experiments performed in that chapter, the applicability of this monitoring parameter is demonstrated (to be published [39]).

5. We have participated in the establishment of a database of human ECG and demographics which is described in chapter 4.1.1. In chapter 6 we demonstrated how the information stored in the medical control module of the defibrillator can be used to evaluate the events occurring during ALS (published in [100]).

10.2 Major conclusions of this work

Several conclusions presents themselves from the work presented in this thesis. To summarise:

1. The design of a decision support system for VF analysis incorporating classification techniques and feature extraction methods is a powerful tool with potential for guiding CPR in an out-of-hospital setting. The crucial elements for the ability of improvement in performance as compared to earlier proposed methods is the use of classifiers that use feature combinations and nonlinear decision borders.

2. CPR artefact removal in human ECG may be performed by using adaptive filtering techniques with reference signals reflecting the cause of the artefact components.

3. CPR efficacy can be monitored by a parameter reflecting the probability of defibrillation success. This parameter reduces the complexity of multi-dimensional features into a one-dimensional parameter which is intuitively related to resuscitability. This may be extremely useful for guidance of CPR therapy.

4. This work provides methods to increase the portion of CPR given during therapy by reducing hands-off intervals. This is done by applying methods allowing rhythm analysis during CPR by CPR artefact removal and

reducing the number of unnecessary shocks by predicting shock outcome using VF analysis.

10.3 Suggestions for further research

We see several directions for further research.

1. More data should be collected for further development and reliable evaluation of VF-analysis decision support systems. These should be collected from several ambulance systems. Furthermore additional data (reference signals for adaptive filtering of CPR artefacts) should be included in the databases.

2. New feature extraction techniques should be investigated. We believe this to have the greatest potential for performance improvement.

   The results in chapter 8 showed that performance depends on the frequency range on which the features are calculated. In the experiments we used a frequency response corresponding to a rectangular window to select the frequency range with all components given equal weighting. Another technique might be to use filters for greater freedom in selecting the frequency response prior to extracting the features. One might consider optimising the response towards classifier performance. These principles has been applied in texture segmentation [78].

   Other approaches could be to use other representations than the short-time Fourier transform. Various time-frequency representations for the onset of in-hospital VF has been studied by Clayton et al [21, 20]. Wavelets were considered for representing VF in pigs in a study CPR in cardiac arrested pigs by Watson et al in [108]. One might evaluate the performance of such time-frequency representation for prediction outcome in a manner similar to what was done for classification using time-frequency distributions by Engleheart [44]. In chapter 2.1 the stages of VF were discussed in relation to morphologic characteristics in VF ECG recordings. We used the terms "energy", "frequency" and "complex" to describe the traces. In this thesis we have studied features related to the two first terms. It might be worthwhile to add new features relating to the third term, "complex". VF characteristics has been related to the number and organisation of wavefronts activating the myocardium by Casaleggio et al [15], Ideker et al [3, 82, 4] and with the recent work by Clayton et al [22, 23] pointing towards the use of nonlinear methods for characterising the complexity of the VF waveform.

It would also be interesting to evaluate which length of the preshock VF blocks corresponds to highest performance.

3. Other classifier techniques could be investigated as alternatives to the smooth histogram technique. As we discussed in chapter 9, the number of histogram bins grow exponentially with feature dimension. We experienced that having feature dimension equal to four caused an increase in computing time for the classifier evaluation. One of the conclusions discussed in chapters 8,5.2 and 5.3 was that more features should be combined. This motivates the replacement the smooth histogram classifier. The radial basis function approach to classification [87] is a strong candidate. The smooth histogram approach computes a histogram for all the training vectors which are represented by histogram columns. Furthermore the contributions from these columns into the the whole of the feature space is represented by a multidimensional matrix. The radial basis function approach, on the other hand, only stores the training vectors to represent the statistical distribution in feature space. The contribution from a training vector to a test vector is computed directly using a gaussian weighting of the distance between the two vectors. Using this approach allows handling feature vectors of higher dimensions than what we have investigated in the present work.

4. The effect of artefact removal by adaptive filtering on outcome prediction should be investigated. In chapter 7 we evaluated the filtering method by improvement in SNR and rhythm detection. Strohmenger et al evaluated their frequency selective approach to artefact reduction on VF analysis outcome prediction [94]. A similar approach can be used to evaluate our outcome prediction system. We could use an artificial mix of animal CPR artefacts (as in chapter 7) and mix with preshock data similar to those used in the outcome prediction analysis in chapter 8. Strohmenger et al did not evaluate their filtering method on preshock VF corrupted with CPR artefacts [94].

Appendix A

Linear Algebra

In this appendix we give some details on how estimates of mean and correlation and covariance matrices and principal axis (PCA) transforms can be obtained from a set of observed feature vectors [104].

A.1 Some linear algebra concepts

In the following some basic concepts of linear algebra is described.

As we discussed in chapter 3.2.2 a feature vector $\mathbf{v}$ consists of D elements so that

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_D \end{bmatrix}^T. \tag{A.1}$$

A.1.1 Orientation and spread

A measure for the *mean* or expectancy of $\mathbf{v}$ is

$$\mu_v = E\{\mathbf{v}\}. \tag{A.2}$$

A measure for the orientation and spread of the measurements is given by the *correlation matrix* which is given as

$$\mathbf{R}_v = E\{\mathbf{v}\mathbf{v}^T\}. \tag{A.3}$$

The spread about the mean is given by the *covariance* matrix given by

$$\mathbf{C}_v = E\{(\mathbf{v} - \mu)(\mathbf{v} - \mu)^T\} = \mathbf{R}_v - \mu\mu^T. \tag{A.4}$$

A.1.2 Estimating from observed data

Defining the *data matrix*

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_K \end{bmatrix}^T \tag{A.5}$$

as having K rows consisting of measurement vectors and columns of random variables. Estimates of the mean and correlation matrix may be given as

$$\hat{\mu}_v = \frac{1}{K} \sum_{i=1}^{K} \mathbf{v}_i \tag{A.6}$$

and

$$\hat{\mathbf{R}}_v = \frac{1}{K} \sum_{i=1}^{K} \mathbf{v}_i \mathbf{v}_i^T = \frac{1}{K} \mathbf{V}^T \mathbf{V} \tag{A.7}$$

respectively.

A.2 Principal axis transformation

A symmetric matrix $\mathbf{A}$ may be diagonalised using its orthonormal eigenvectors $\mathbf{e}_i$ and corresponding eigenvalues λ_i where $i = 1 \dots \dim(A) = N$. Using the properties

$$\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i, \tag{A.8}$$

and

$$\mathbf{e}_i^T \mathbf{A} \mathbf{e}_j = \lambda_j \mathbf{e}_i^T \mathbf{e}_j = \begin{cases} \lambda_j & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{A.9}$$

which leads to

$$\mathbf{E}^T \mathbf{A} \mathbf{E} = \Lambda \tag{A.10}$$

where $\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_N \end{bmatrix}$ and $\Lambda = \operatorname{diag}\begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_N \end{pmatrix}$, a unitary transform may be expressed as

$$\mathbf{w} = \mathbf{E}^T \mathbf{v}. \tag{A.11}$$

If $\mathbf{E}$ is the eigenvector matrix of $\mathbf{R}_v$, $\mathbf{w}$ has a diagonal correlation matrix as can be shown by

$$\mathbf{R}_w = E\{\mathbf{w}\mathbf{w}^T\} = E\{\mathbf{E}^T\mathbf{v}\mathbf{v}^T\mathbf{E}\} = \mathbf{E}^T\mathbf{R}_v\mathbf{E} \tag{A.12}$$

meaning that the values of $\mathbf{w}$ are orthogonal. The covariance matrix may be similarly diagonalised, thus giving transformed data $\mathbf{w}$ with uncorrelated elements. The transformation in equation A.13 rotates the axis system so that the axes are aligned with the principal directions of the data set. The axis system is further translated to the mass center of the data set by subtracting the mean.

$$\mathbf{w} = \mathbf{E}^T(\mathbf{v} - \mu) \tag{A.13}$$

Appendix B

CPR artefact removal filter derivations

This material has been taken from [1]. The mathematical details of the derivation of the optimal solution for the multi-channel filter described in chapter 3.4.2 is given in section B.1. The experiments performed to tune the filter parameters are described in section B.2.

B.1 Filter derivations

The objective is now to find the P FIR filters that collectively minimize the object function defined as

$$J = \sum_{n=-\infty}^{\infty} w(n, n_0) \left[x(n) - y(n)\right]^2 . \tag{B.1}$$

Limited to the model in equation 3.47, the synthesized artefact $y(n)$ should be tuned to maximum similarity to the noisy part of the ECG signal $x(n)$. The role of the window function $w(n, n_0)$ is to define the analysis region used when computing the optimal filters. Letting n_0 denote the current signal sample, the following cases can be emulated by choosing the appropriate window function:

Stationary solution: Setting $w(n, n_0) = 1, -\infty < n < \infty$ the minimization of equation B.1 yields the best possible *fixed* filter solution.

Causal solution: If $w(n, n_0) = 0, n > n_0$ the filter solution will only use knowledge about the relationship between $x(n)$ and the reference signals using past samples.

Non-causal solution: If we are to compute the best possible filter solution to use at time n_0, the region used for the analysis should be short-term. The best way of providing this adaptivity is to allow the region of support of $w(\cdot, n_0)$ to extend on both sides of n_0. In a real-time setting this will introduce a time-delay equal to $n_1 - n_0$, where n_1 is the last nonzero value of the window function. In our experiments we use a Hamming window [77] centered around n_0. This is illustrated in figure B.1.

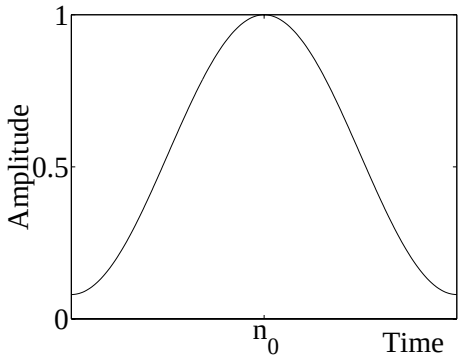


Figure B.1: A Hamming window defining the short-term analysis region. n_0 denote the "now" point in time.

Setting $\frac{\partial J}{\partial h_i(j)} = 0$ we obtain:

$$\sum_{p=1}^{P} \sum_{k=0}^{K_p-1} h_p(k) \underbrace{\sum_{n=-\infty}^{\infty} w(n, n_0) v_p(n-k) v_i(n-j)}_{r_{v_p v_i}(k,j)} = \underbrace{\sum_{n=-\infty}^{\infty} w(n, n_0) x(n) v_i(n-j)}_{r_{x v_i}(j)}, \tag{B.2}$$

where $i = 1, \ldots, P$ and $j = 0, \ldots, K_i - 1$. Note that the time window defining the short-term analysis region is included in the auto- and cross-correlation functions $r_{v_p v_i}(\cdot, \cdot)$ and $r_{x v_i}(\cdot)$.

equation B.2 defines a linear set of $\sum_{p=1}^{P} K_p$ equations with the same number of unknowns. Alternatively, the system of equations can be written as a single

matrix equation:

$$\underbrace{\begin{bmatrix} \mathbf{R}_{11} & \cdots & \mathbf{R}_{P1} \\ \mathbf{R}_{12} & \cdots & \mathbf{R}_{P2} \\ & \vdots & \\ \mathbf{R}_{1P} & \cdots & \mathbf{R}_{PP} \end{bmatrix}}_{\mathbf{R}} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_P \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{xv_1} \\ \mathbf{r}_{xv_2} \\ \vdots \\ \mathbf{r}_{xv_P} \end{bmatrix}, \tag{B.3}$$

where

$$\mathbf{R}_{st} = \begin{bmatrix} r_{v_s v_t}(0,0) & \cdots & r_{v_s v_t}(K_s-1,0) \\ & \vdots & \\ r_{v_s v_t}(0,K_t-1) & \cdots & r_{v_s v_t}(K_s-1,K_t-1) \end{bmatrix} \tag{B.4}$$

$$\mathbf{h}_i^T = [h_i(0), \ldots, h_i(K_i-1)] \tag{B.5}$$

$$\mathbf{r}_{xv_i}^T = [r_{xv_i}(0), \ldots, r_{xv_i}(K_i-1)]. \tag{B.6}$$

It should be noted that since

$$\begin{aligned} r_{v_s v_t}(k,j) &= r_{v_t v_s}(j,k) \\ r_{v_s v_t}(k,j) &\neq r_{v_s v_t}(k+1,j+1) \end{aligned}$$

it follows that $\mathbf{R}$ is symmetric but not Tőplitz. This can also be seen as a result of the way the autocorrelation functions are defined in equation B.2. This is commonly referred to as the *covariance* method in linear prediction [55].

The solution of equation B.2 or, equivalently, equation B.3 gives the optimal set of P FIR filters for removing the artefact component in the ECG signal $x(n)$ at time $n = n_0$. This is commonly referred to as the Wiener solution [77].

The price paid for using the optimal filter solution is the computational burden. For each signal sample, the solution of equation B.3 requires the inversion of a square matrix with dimensions equal to the total number of FIR filter coefficients. However, our experiments will demonstrate that very few filter coefficients are necessary, thus reducing the complexity of the system.

B.2 Filter tuning

The data used in the tuning experiments are the ones described in chapter 7.

In the proposed filter structure in figure 3.7 the following parameters have to be chosen: The filter lengths K_1 and K_2 for the chest compression and

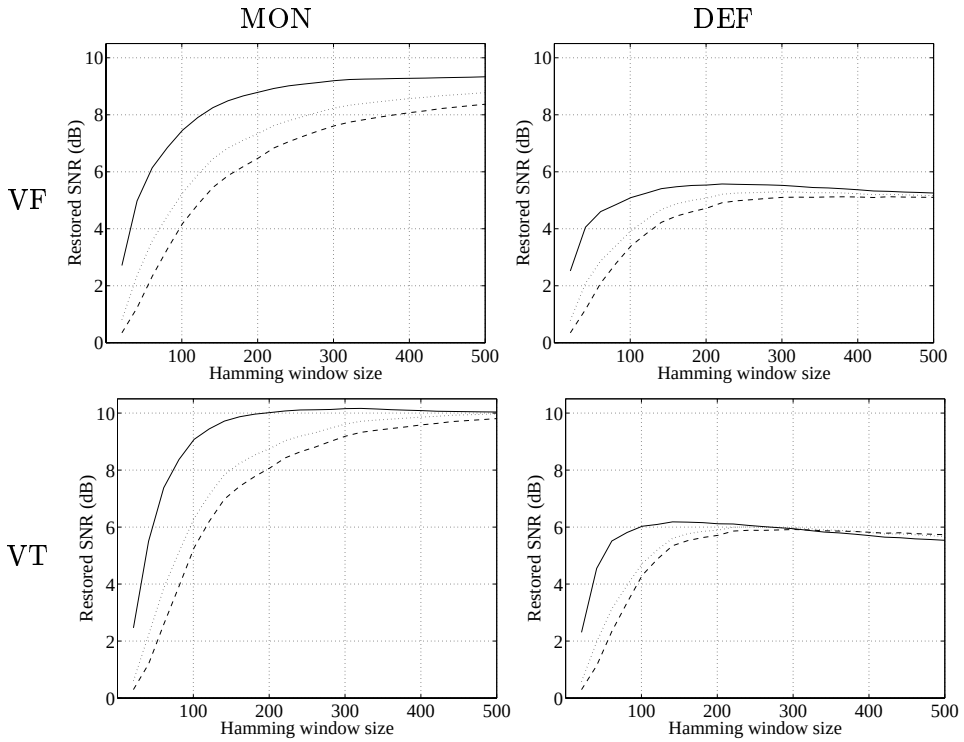


Figure B.2: Artefact removal performance as a function of the window length and the number of filter taps (K) in each channel: 1-tap filters (solid), 2-tap filters (dotted), and 3-tap filters (dashed). The compression rate is 90 min^{-1} and the input SNR is 0 dB.

the ventilation channels, respectively. The size (the support) of the Hamming window function $w(\cdot, n_0)$ also has to be decided.

The filter construction is based on the minimization of a quadratic error term. Accordingly, we tune the parameter settings for maximum SNR improvement. In order to limit the search for the best parameter setting we adapt the following strategy: Experiments are performed using a limited data set. We have chosen to evaluate the artefact removal system at 90 chest compressions per minute, at 0 dB SNR. The channel filter lengths are set to be identical, i.e. $K_1 = K_2 = K$.

Using the scheme outlined above, we average the obtained SNR after restoration over the human ECG ensemble (VT or VF) and over the 2 added animal artefact records. The obtained results are shown in figure B.2.

In all 4 cases (VF/MON, VF/DEF, VT/MON, and VT/DEF) we observe the superior results obtained using only one filter tap in each channel filter, the

exception being the VT/DEF case where the 1-tap filter is outperformed when the window size is well above 300 samples. This conclusion may not hold for window sizes larger than 500 samples – the range used for these plots – but this hardly matters since the use of larger window sizes would result in a filter time delay larger than 2.5 seconds, which is unacceptable. We therefore conclude that 1-tap filters are suitable for our purposes. Note that this does not imply a simple filtering scheme identical to multiplication by a constant. The filters are adaptive on a sample-to-sample basis allowing the "constant" to vary.

The artefact removal is more successful on the MON measurements than on the DEF measurements, with the former achieving up to 10 dB and the latter 6 dB in the VT case. The VF performance is 1–2 dB lower.

The best choices of window lengths are about 300 samples in the MON case and about 160 samples in the DEF case, when using 1-tap filters. A probable cause for this difference is the higher degree of signal stationarity obtained when measuring the ECG through the monitor pads.

Bibliography

[1] S. O. Aase, T. Eftestøl, J. H. Husøy, K. Sunde, and P. A. Steen. CPR artifact removal from human ECG using optimal multichannel filtering. *IEEE Transactions on Biomedical Engineering*. In press 2000.

[2] A. A. J. Adgey. The belfast experience with resuscitation ambulances. *American Journal of Emergency Medicine*, 2:193–199, 1984.

[3] P. V. Bayly, E. E. Johnson, P. D. Wolf, W. M. Smith, and R. E. Ideker. Predicting patterns of epicardial potentials during ventricular fibrillation. *IEEE Transactions on Biomedical Engineering*, 42(9):898–907, 1995.

[4] P. V. Bayly, B. H. KenKnight, J. M. Rogers, E. E. Johnson, R. E. Ideker, and W. M. Smith. Spatial organization, predictability, and determinism in ventricular fibrillation. *Chaos*, 8(1):103–115, 1998.

[5] L. B. Becker. The epidemiology of sudden death. In N. A. Paradis, H. R. Halperin, and R. M. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 2, pages 28–47. Williams & Wilkins, Baltimore, 1996.

[6] J. C. Behr, L. L. Hartley, D. K. York, D. D. Brown, and R. E. Kerber. Truncated exponential versus damped sinusoidal waveform shocks for transthoracic defibrillation. *The American Journal of Cardiology*, 78:1242–1245, 1996.

[7] R. A. Berg, R. W. Hilwig, K. B. Kern, E. V. Corso, J. W. Heidenreich, and G. A. Ewy. Optimal treatment of prolonged VF in a swine model of prehospital cardiac arrest (abstract). *Circulation*, 100:I–91, 1999.

[8] M. J. Bonnin, P. E. Pepe, K. T. Kimball, and P. S. Clark Jr. Distinct criteria for termination of resuscitation in the out-of-hospital setting. *Journal of American Medical Assosciation*, 270:1457–1462, 1993.

[9] C. G. Brown and R. Dzwonczyk. Non-invasive monitoring during ventricular fibrillation. *Applied Cardiopulmonary Pathophysiology*, 4:293–299, 1992.

[10] C. G. Brown and R. Dzwonczyk. Signal analysis of the human electrocardiogram during ventricular fibrillation: Frequency and amplitude parameters as predictors of successful countershock. *Annals of Emergency Medicine*, 27(2):184–188, February 1996.

[11] C. G. Brown, R. Dzwonczyk, J. Jenkins, and H. Werman. The median frequency of ventricular fibrillation - A useful guide to therapeutic interventions during CPR (abstract). *Annals of Emergency Medicine*, 17:436–437, 1988.

[12] C. G. Brown, R. Dzwonczyk, H. A. Werman, and R. L. Hamlin. Estimating the duration of ventricular fibrillation. *Annals of Emergency Medicine*, 18(11):1181–1185, November 1989.

[13] C. G. Brown, R. F. Griffith, P. V. Ligten, J. Hoekstra, G. Nejman, L. Mitchell, and R. Dzwonczyk. Median frequency - a new parameter for predicting defibrillation success rate. *Annals of Emergency Medicine*, 20(7):787–789, July 1991.

[14] M. Callaham, O. Braun, W. Valentine, D. M. Clark, and C. Zegans. Prehospital cardiac arrest treated by urban first-responders: Profile of patient response and prediction of outcome by ventricular fibrillation waveform. *Annals of Emergency Medicine*, 22(11):21–34, 1993.

[15] A. Casaleggio, A. Corona, R. Ranjan, and N. V. Thakor. Dimensional analysis of the electrical activity in fibrillating isolated hearts. *International Journal of bifurcation and Chaos*, 6(8):1547–1561, 1996.

[16] D. Chamberlain, L. Bossaert, P. Carli, E. Edgren, L. Ekstrom, S. Hapnes, S. Holmberg, R. Koster, K. Lindner, V. Pasqualucci, N. Perales, M. von Planta, C. Robertson, and P. Steen. Guidelines for advanced life support. A statement by the Advanced Life Support Working Party of the European Resuscitation Council. *Resuscitation*, 24:111–121, 1992.

[17] D. Chamberlain, R. O. Cummins, et al. Recommended guidelines for uniform reporting of data from out-of-hospital cardiac arrest: the "Utstein style". *Resuscitation*, 22:1–26, 1991.

[18] N. C. Chandra and M. F. H. red. Basic life support for healthcare providers. *Dallas: American Heart Association*, pages 4.1–4.23, 1997.

[19] P. B. Chase, K. B. Kern, A. B. Sanders, C. W. Otto, and G. A. Ewy. Effects of graded doses of epinephrine on both noninvasive and invasive measures of myocardial perfusion and blood flow during cardiopulmonary resuscitation. *Critical Care Medicine*, 21:413–419, 1993.

[20] R. H. Clayton, R. W. F. Campbell, and A. Murray. Characteristics of multichannel ECG recordings during human ventricular tachyarrhythmias. *IEEE Engineering in Medicine and Biology Magazine*, 17(1):39–44, 1998.

[21] R. H. Clayton and A. Murray. Comparison of techniques for time-frequency analysis of the ECG during human ventricular fibrillation. *IEEE Proc. Sci. Meas. Technol.*, 145(6):301–306, 1998.

[22] R. H. Clayton and A. Murray. Linear and non-linear analysis of the surface electrocardiogram during human ventricular fibrillation shows evidence of order in the underlying mechanism. *Medical & Biological Engineering & Computing*, 37:354–358, 1999.

[23] R. H. Clayton, D. J. Yu, M. Small, V. N. Biktashev, R. G. Harrison, and A. V. Holden. Linear and nonlinear characteristics of ECG signals produced by simulations of ventricular tachyarrhythmias. *Computers in Cardiology*, 26:479–482, 1999.

[24] L. A. Cobb, C. E. Fahrenbruch, T. R. Walsh, M. K. Copass, M. Olsufka, M. Breskin, and A. P. Hallstrom. Influence of cardiopulmonary resuscitation prior to defibrillation in patients with out-of-hospital ventricular fibrillation. *Journal of American Medical Assosciation*, 281:1182–1188, 1999.

[25] S. M. Cobbe, M. J. Redmond, J. M. Watson, J. Hollingworth, and D. J. Carrington. "Heartstart Scotland" - Initial experience of a national scheme for out of hospital defibrillation. *British Medical Journal*, 302:1517–1520, 1991.

[26] R. O. Cummins, M. S. Eisenberg, A. P. Hallstrom, T. R. Hearne, J. R. Graves, and P. E. Litwin. What is a "save"?: Outcome measures in clinical evaluations of autmatic external defibrillators. *American Heart Journal*, 110:1133–1138, 1985.

[27] R. O. Cummins, M. S. Eisenberg, P. E. Litwin, J. R. Graves, T. R. Hearne, and A. P. Hallstrom. Automatic external defibrillators used by emergency medical technicians: A controlled clinical trial. *Journal of American Medical Assosciation*, 257:1605–1610, 1987.

[28] R. O. Cummins, J. P. Ornato, W. H. Thies, and P. E. Pepe. Improving survival from sudden cardiac arrest: The "Chain of Survival" concept. a statement for health professionals from the Advanced Cardiac Life Support Subcommittee and the Emergency Cardiac Care Committee, American Heart Association. *Circulation*, 83:1832–1847, 1991.

[29] R. V. Ditchey, Y. Goto, and J. Lindenfield. Myocardial oxygen requirements during experimental cardiopulmonary resuscitation. *Cardiovascular Research*, 26:791–797, 1992.

[30] *Dorland's illustrated medical dictionary.* W. B. Saunders Co., Philadelphia, USA, 28th edition, 1994.

[31] E. R. Dougherty. *Probability and statistics for the engineering, computing, and physical sciences.* Prentice-Hall International, Inc, Englewood Cliffs, NJ, 1990.

[32] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* John Wiley and Sons, Inc., New York, 1973.

[33] R. Dzwonczyk, C. G. Brown, R. B. Taylor, and J. Ashton. Frequency analysis of the ECG during ventricular fibrillation. *Proc. 9th Annu. Conf.IEEE EMBS*, 1:147–148, November 1987.

[34] T. Eftestøl, S. O. Aase, and J. H. Husøy. Characterisation of changes in cardiac arrhythmias using spectral parameters. In *Proc. NORSIG '97*, pages 129–133, Tromsø, Norway, May 1997.

[35] T. Eftestøl, S. O. Aase, and J. H. Husøy. A flexible pattern recognition system for analysis of ECG and related demographics and annotations. In *Proc. EMBS '98*, volume 20, pages 135–138, Hong Kong, October 1998.

[36] T. Eftestøl, S. O. Aase, and J. H. Husøy. Spectral flatness measure for characterising changes in cardiac arrhythmias. In *Proc. NOBIM '98*, Oslo, Norway, June 1998.

[37] T. Eftestøl, S. O. Aase, and J. H. Husøy. Spectral characterisation of ECG in out-of-hospital cardiac arrest patients. In *Proc. Computers in Cardiology '99*, volume 26, pages 707–710, Hannover, Germany, September 1999.

[38] T. Eftestøl, K. Sunde, S. O. Aase, J. H. Husøy, and P. A. Steen. Predicting outcome of defibrillation by spectral characterization and nonparametric classification of ventricular fibrillation in out-of-hospital cardiac arrested patients. *Circulation*. In press 2000.

[39] T. Eftestøl, K. Sunde, S. O. Aase, J. H. Husøy, and P. A. Steen. "Probability of successful defibrillation" as a monitor during CPR in out-of-hospital cardiac arrested patients. *Resuscitation*. In press 2000.

[40] T. Eftestøl, K. Sunde, S. O. Aase, J. H. Husøy, and P. A. Steen. Predicting defibrillation outcome from VF-characteristics in out-of hospital cardiac arrest. *Resuscitation*, 2000. Abstract, to be published in june issue.

[41] T. Eftestøl and E. Vatland. A method for detection of signal harmonics in the spectrum. In *Proc. Computers in Cardiology '99*, volume 26, pages 547–550, Hannover, Germany, September 1999.

[42] M. S. Eisenberg. The quest to reverse sudden death: A history of cardiopulmonary resuscitation. In N. A. Paradis, H. R. Halperin, and R. M. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 1, pages 1–27. Williams & Wilkins, Baltimore, 1996.

[43] Emergency Cardiac Care Committee and Subcommittees, American Heart Association. Guidelines for cardiopulmonary resuscitation and emergency cardiac care. *Journal of American Medical Assosciation*, 268:2171–2302, 1992.

[44] K. Engleheart. *Signal representation for classification of the transient myolectric signal*. PhD thesis, University of New Brunswick, Fredericton, New Brunswick, Canada, October 1998.

[45] M. P. Feneley, G. W. Maier, K. B. Kern, J. W. Gaynor, S. A. Gall, A. B. Sanders, K. Raessler, L. H. Muhlbaier, S. Rankin, and G. A. Ewy. Influence of compression rate on initial success of resuscitation and 24 hour survival after prolonged manual cardiopulmonary resuscitation in dogs. *Circulation*, 77:240–250, 1988.

[46] B. E. Gliner, D. B. Jorgensen, J. E. Poole, R. D. White, K. G. Kanz, T. D. Lyster, K. W. Leyde, D. J. Powers, C. B. Morgan, R. A. Kronmal, and G. H. Bardy. Treatment of out-of-hospital cardiac arrest with

a low-energy impedance-compensating biphasic waveform automatic external defibrillator. *Biomedical Instrumentation & Technology*, 32:631–644, 1998.

[47] B. E. Gliner and R. D. White. Electrocardiographic evaluation of defibrillation shocks delivered to out-of-hospital sudden cardiac arrest patients. *Resuscitation*, 41:133–144, 1999.

[48] C. V. Gudipathi, M. H. Weil, J. Bisera, H. G. Deshmukh, and E. C. Rackow. Expired carbon dioxide : a noninvasive monitor of cardiopulmonary resuscitation. *Resuscitation*, 77:234–239, 1988.

[49] H. R. Halperin, J. E. Tsitlik, A. D. Guerci, E. D. Mellits, H. R. Levin, A. Y. Shi, N. Chandra, and M. L. Weisfeldt. Determinants of blood flow to vital organs during cardiopulmonary resuscitation in dogs. *Circulation*, 73:539–550, 1986.

[50] K. M. Hargarten, H. A. Stueven, E. M. Waite, D. V. Olson, J. R. Mateer, T. P. Aufderheide, and J. C. Darin. Prehospital experience with defibrillation of coarse ventricular fibrillation: A ten-year review. *Ann Emerg Med*, 19:157–162, 1990.

[51] J. Herlitz, J. Bahr, M. Fischer, M. Kuisma, K. Lexow, and G. Thorgeirsson. Resuscitation in Europe: A tale of five European regions. *Resuscitation*, 41:121–131, 1999.

[52] D. Hightower, S. H. Thomas, C. K. Stone, K. Dunn, and J. A. March. Decay in quality of closed-chest compressions over time. *Annals of Emergency Medicine*, 26:300–303, 1995.

[53] A. H. Idris, O. G. Florete Jr, R. J. Melker, and N. C. Chandra. Physiology of ventilation, oxygenation, and carbon dioxide elimination during cardiac arrest. In N. A. Paradis, H. R. Halperin, and R. M. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 21, pages 382–419. Williams & Wilkins, Baltimore, 1996.

[54] A. H. Idris, E. D. Staples, D. J. O'Brien, R. J. Melker, W. J. Rush, K. D. D. Duca, and J. L. Falk. Effect of ventilation on acid-base balance and oxygenation in low blood-flow states. *Critical Care Medicine*, 22:1827–1834, 1994.

[55] N. S. Jayant and P.Noll. *Digital coding of waveforms: Principles and applications to speech and video.* Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1984.

[56] S. M. Kay. *Modern spectral estimation: Theory & application.* Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1988.

[57] R. E. Kerber and C. E. Robertson. Transthoracic defibrillation. In N. A. Paradis, H. R. Halperin, and R. M. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 20, pages 370–381. Williams & Wilkins, Baltimore, 1996.

[58] K. B. Kern, G. A. Ewy, W. D. Vorhees, C. F. Babbs, and W. A. Tacker. Myocardial perfusion pressure: A predictor of 24-hour survival during prolonged cardiac arrest in dogs. *Resuscitation*, 16:241–250, 1988.

[59] K. B. Kern, H. S. Garewal, A. B. Sanders, W. Janas, J. Nelson, D. Sloan, W. A. Tacker, and G. A. Ewy. Depletion of myocardial adenosine triphosphate during prolonged untrated ventricular fibrillation: Effect on defibrillation success. *Resuscitation*, 20:221–229, 1990.

[60] K. B. Kern and J. T. Niemann. Coronary perfusion pressure during cardiopulmonary resuscitation. In N. A. Paradis, H. R. Halperin, and R. M. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 14 A, pages 270–285. Williams & Wilkins, Baltimore, 1996.

[61] K. B. Kern, A. B. Sanders, W. D. Vorhees, C. F. Babbs, W. A. Tacker, and G. A. Ewy. Changes in expired end-tidal carbon dioxide during cardiopulmonary resuscitation in dogs: A prognostic guide for resuscitation efforts. *J Am Coll Cardiol*, 13:1184–1189, 1989.

[62] K. W. Kuelz, P. Hsia, R. M. Wise, R. Mahmud, and R. J. Damiano. Integration of absolute ventricular fibrillation voltage correlates with successful defibrillation. *IEEE Transactions on Biomedical Engineering*, 41:782–791, 1994.

[63] A. Langhelle, T. Eftestøl, H. Myklebust, M. Eriksen, B. T. Holten, and P. A. Steen. Reducing CPR artefacts in ventricular fibrillation in vitro. *Resuscitation.* In press 2000.

[64] M. P. Larsen, M. S. Eisenberg, R. O. Cummins, and A. P. Hallstrom. Predicting survival from out-of-hospital cardiac arrest: a graphic model. *Annals of Emergency Medicine*, 22:1652–1658, 1993.

[65] G. W. Maier, J. R. Newton, J. A. Wolfe, G. S. Tyson Jr, C. O. Olsen, D. D. Glower, J. A. Spratt, J. W. Davis, M. P. Feneley, and J. S. Rankin.

The influence of manual chest compression rate on hemodynamic support during cardiac arrest: high-impulse cardiopulmonary resuscitation. *Circulation*, 74 (suppl IV):51–59, 1986.

[66] G. W. Maier, G. S. Tyson Jr, C. O. Olsen, K. H. Kernstein, J. W. Davis, E. H. Cohn, D. C. Sabiston, and J. S. Rankin. The physiology of external cardiac massage: high-impulse cardiopulmonary resuscitation. *Circulation*, 70:86–101, 1984.

[67] J. E. Manning and P. M. Zoll. Therapy of bradyasystolic arrest. In N. Paradis, H. Halperin, and R. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 33, pages 621–640. Williams & Wilkins, Baltimore, 1996.

[68] D. R. Martin, C. G. Brown, and R. Dzwonczyk. Frequency analysis of the human and swine electrocardiogram during ventricular fibrillation. *Resuscitation*, 22:85–91, 1991.

[69] J. D. Michenfelder and R. A. Theye. The effects of anesthesia and hypothermia on canine cerebral ATP and lactate during anoxia produced by decapitation. *Anesthesiology*, 33:430–439, 1970.

[70] K. G. Monsieurs, H. D. Cauwer, F. L. Wuyts, and L. L. Bossaert. A rule for early outcome classification of out-of-hospital cardiac arrest patients presenting with ventricular fibrillation. *Resuscitation*, 36:37–44, 1998.

[71] J. T. Niemann, C. B. Cairns, J. Sharma, and R. J. Lewis. Treatment of prolonged ventricular fibrillation: Immediate countershock versus high-dose epinephrine and CPR preceding countershock. *Circulation*, 85(1):281–287, January 1992.

[72] M. Noc, M. H. Weil, R. J. Gazmuri, S. Sun, J. Bisera, and W. Tang. Ventricular fibrillation voltage as a monitor of the effectiveness of cardiopulmonary resuscitation. *Journal of Laboratory and Clinical Medicine*, 124(3):421–426, 1994.

[73] M. Noc, M. H. Weil, W. Tang, S. Sun, A. Pernat, and J. Bisera. Electrocardiographic prediction of the success of cardiac resuscitation. *Critical Care Medicine*, 27(4):708–714, 1999.

[74] J. P. Ornato, R. L. Levine, D. S. Young, E. M. Racht, A. R. Garnett, and E. R. Gonzalez. The effect of applied chest compression force on systemic arterial pressure and end-tidal carbon dioxide concentration during CPR in human beings. *Annals of Emergency Medicine*, 18:732–737, 1989.

[75] N. A. Paradis, E. Koscove, K. H. Lindner, C. B. Cairns, and C. M. Little. Vasopressor therapy during cardiac arrest. In N. Paradis, H. Halperin, and R. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 26, pages 497–527. Williams & Wilkins, Baltimore, 1996.

[76] N. A. Paradis, G. B. Martin, E. P. Rivers, M. G. Goetting, T. J. Appleton, M. Feingold, and R. M. Nowak. Coronary perfusion pressure and the return of spontaneous circulation in human cardiopulmonary resuscitation. *Journal of American Medical Assosciation*, 263:1106–1113, 1990.

[77] J. G. Proakis and D. G. Manolakis. *Digital signal processing: Principles, algorithms and applications.* Macmillian Publishing Company, New York, 2nd edition, 1992.

[78] T. Randen. *Filter and Filter Bank Design for Image Texture Recognition.* PhD thesis, Norwegian University of Science and Technology, Oct. 1997.

[79] B. K. Rayburn and H. R. Halperin. External chest compression - standard and alternate techniques. In N. Paradis, H. Halperin, and R. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 22, pages 420–438. Williams & Wilkins, Baltimore, 1996.

[80] B. D. Ripley. *Pattern recognition and neural networks.* Cambridge University Press, Cambridge, UK, 1996.

[81] C. Robertson, P. Steen, J. Adgey, L. Bossaert, P. Carli, D. Chamberlain, W. Dick, L. Ekstrom, S. A. Hapnes, S. Holmberg, R. Juchems, F. Kette, R. Koster, F. J. de Latorre, K. Lindner, and N. Perales. The 1998 European Resuscitation Council guidelines for adult advanced life support. A statement from the Working Group on Advanced Life Support, and approved by the executive committee of the European Resuscitation Council. *Resuscitation*, 37:81–90, 1998.

[82] J. M. Rogers, M. Usui, B. H. KenKnight, R. E. Ideker, and W. M. Smith. Recurrent wavefront morphologies: A method for quantifying the complexity of epicardial activation patterns. *Annals of Biomedical Engineering*, 25(5):761–768, 1997.

[83] A. B. Sanders, K. B. Kern, M. Atlas, S. Bragg, and G. A. Ewy. Importance of the duration of inadequate coronary perfusion pressure on resuscitation from cardiac arrest. *J Am Coll Cardiol*, 6:113–118, 1985.

[84] Y. Sato, M. H. Weil, S. Sun, W. Tang, J. Xie, M. Noc, and J. Bisera. Adverse effects of interrupting precordial compression during cardiopulmonary resuscitation. *Critical Care Medicine*, 25:733–736, 1997.

[85] R. J. Schalkoff. *Pattern recognition: Statistical, structural and neural approaches.* John Wiley & sons, New York (NY), 1992.

[86] T. Schneider, D. Mauer, P. Diehl, B. Eberle, and W. Dick. Quality of on-site performance in prehospital advanced cardiac life support (ACLS). *Resuscitation*, 27:207–213, 1994.

[87] J. Schürmann. *Pattern Classification: A Unified View of Statistical and Neural Approaches.* John Wiley & Sons, Inc., New York, 1996.

[88] M. L. Sedgwick, J. Watson, K. Dalziel, D. J. Carrington, and S. M. Cobbe. Efficacy of out of hospital defibrillation by ambulance technicians using automated external defibrillators. The Heartstart Scotland project. *Resuscitation*, 24(73–87), 1992.

[89] M. L. Sedgwick, J. Watson, K. Dalziel, D. J. Carrington, and S. M. Cobbe. Performance of an established system of first responder out-of-hospital defibrillation. The results of the second year of the Heartstart Scotland project in the "Utstein style". *Resuscitation*, 26:75–88, 1993.

[90] M. Small, D. J. Yu, R. G. Harrison, C. Robertson, G. Clegg, M. Holzer, and F. Sterz. Characterizing nonlinearity in ventricular fibrillation. In *Proc. Computers in Cardiology '99*, pages 17–20, Hannover, Germany, September 1999.

[91] D. W. Spaite, T. D. Valenzuela, H. W. Meislin, E. A. Criss, and P. Hinsberg. Prospective validation of a new model for evaluating emergency medical services systems by in-field observation of specific time intervals in prehospital care. *Annals of Emergency Medicine*, 22:638–645, 1993.

[92] A. J. Stewart, J. D. Allen, and A. A. J. Adgey. Frequency analysis of ventricular fibrillation and resuscitation success. *Quarterly Journal of Medicine*, 85:761–769, October 1992.

[93] I. G. Stiell, G. A. Wells, V. J. DeMaio, D. W. Spaite, B. J. Field, D. P. Munkley, M. B. Lyver, L. G. Luinstra, and R. Ward. Modifiable factors associated with improved cardiac arrest survival in a multicenter basic life support/defibrillation system: OPALS study phase I results. *Annals of Emergency Medicine*, 33:44–50, 1999.

[94] H.-U. Strohmenger, K. H. Lindner, and C. G. Brown. Analysis of the ventricular fibrillation ECG signal amplitude and frequency parameters as predictors of countershock success in humans. *Chest*, 111(3):584–589, March 1997.

[95] H.-U. Strohmenger, K. H. Lindner, A. Keller, I. M. Lindner, and E. G. Pfenninger. Spectral analysis of ventricular fibrillation and closed-chest cardiopulmonary resuscitation. *Resuscitation*, 33(2):155–161, 1996.

[96] H.-U. Strohmenger, K. H. Lindner, A. Keller, I. M. Lindner, E. G. Pfenninger, and U. Bothner. Effects of graded doses of vasopressin on median fibrillation frequency in a porcine model of cardiopulmonary resuscitation: Results of a prospective randomized, controlled trial. *Critical Care Medicine*, 24(8):1360–1365, 1996.

[97] H.-U. Strohmenger, K. H. Lindner, K. G. Lurie, A. Welz, and M. Georgieff. Frequency of ventricular fibrillation as a predictor of defibrillation success during cardiac surgery. *Anasthesia and Analgia*, 79:434–438, 1994.

[98] H.-U. Strohmenger, K. H. Lindner, A. W. Prengel, E. Pfenninger, U. Bothner, and K. G. Lurie. Effects of epinephrine and vasopressin on median fibrillation frequency and defibrillation success in a porcine model of cardiopulmonary resuscitation. *Resuscitation*, 31:65–73, 1996.

[99] K. Sunde, T. Eftestøl, C. Askenberg, and P. A. Steen. Kvalitetsvurdering av avansert hjerte-lunge-redning i ambulansetjenesten i Oslo ved analysering av defibrillatorens datamodul. *NAForum*, 12(3):34, 1999. Abstract 27, in Norwegian.

[100] K. Sunde, T. Eftestøl, C. Askenberg, and P. A. Steen. Quality assessment of defibrillation and advanced life support using data from the medical control module of the defibrillator. *Resuscitation*, 41:237–247, 1999.

[101] K. Sunde, T. Eftestøl, and P. A. Steen. Prehospital hjertestans - Resultater etter to års Utstein-registrering ved ambulansetjenesten i Oslo. *NAForum*, 12(3):34, 1999. Abstract 28, in Norwegian.

[102] K. Sunde, T. Eftestøl, and P. A. Steen. Survival rate depends on place of collapse in patients with out-of-hospital cardiac arrest in Oslo. *Resuscitation*, 2000. Abstract, to be published in june issue.

[103] K. Sunde, L. Wik, and P. A. Steen. Quality of mechanical, manual standard and active compression-decompression CPR on the arrest site

and during transport in a manikin model. *Resuscitation*, 34:235–242, 1997.

[104] C. W. Therrien. *Discrete random signals and statistical signal processing.* Prentice Hall signal processing series, Englewood Cliffs, NJ, USA, 1992.

[105] G. F. Tomaselli. Etiology, electrophysiology and mechanics of ventricular fibrillation. In N. Paradis, H. Halperin, and R. Nowak, editors, *Cardiac arrest: The science and practice of resuscitation medicine*, chapter 16, pages 301–319. Williams & Wilkins, Baltimore, 1996.

[106] H. Tunstall-Pedoe, L. Bailey, D. A. Chamberlain, A. K. Marsden, M. E. Ward, and D. A. Zideman. Survival of 3765 cardiopulmonary resuscitations in british hospitals (the BRESUS study): Methods and overall results. *British Medical Journal*, 304:1347–1351, 1992.

[107] J. H. van Bemmel and M. A. Musen, editors. *Handbook of medical informatics.* Springer-Verlag, Heidelberg, Germany, 1997.

[108] J. N. Watson, P. S. Addison, G. R. Clegg, M. Holzer, F. Sterz, and C. E. Robertson. A novel wavelet transform based analysis reveals hidden structure in ventricular fibrillation. *Resuscitation*, 43:121–127, 2000.

[109] W. D. Weaver, L. A. Cobb, D. Dennis, R. Ray, A. P. Hallstrom, and M. K. Copass. Amplitude of ventricular fibrillation waveform and outcome after cardiac arrest. *Annals of Internal Medicine*, 102:53–55, January 1985.

[110] W. D. Weaver, M. K. Copass, D. Bufi, R. Ray, A. P. Hallstrom, and L. A. Cobb. Improved neurologic recovery and survival after early defibrillation. *Circulation*, 69:943–948, 1984.

[111] W. D. Weaver, J. S. Martin, M. J. Wirkus, S. Morud, S. Vincent, P. E. Litwin, and C. Morgan. Influence of external defibrillator electrode polarity on cardiac resuscitation. *Pacing Clin Electrophysiol*, 16:285–290, 1993.

[112] M. H. Weil, J. Bisera, R. P. Trevino, and E. C. Rackow. Cardiac output and end-tidal carbon dioxide. *Critical Care Medicine*, 13:907–909, 1985.

[113] B. Widrow and S. D. Stearns. *Adaptive Signal Processing.* Prentice-Hall, Englewood Cliffs, 1985.

[114] L. Wik, P. A. Steen, and N. G. Bircher. Quality of bystander cardiopulmonary resuscitation influences outcome after prehospital cardiac arrest. *Resuscitation*, 28:195–203, 1994.

[115] J. Xie, M. H. Weil, S. Sun, W. Tang, Y. Sato, X. Jin, and J. Bisera. High-energy defibrillation increases the severity of postresuscitation myocardial dysfunction. *Circulation*, 96:683–688, 1997.

[116] R. W. Yakaitis, G. A. Ewy, C. W. Otto, D. L. Taren, and T. E. Moon. Influence of time and therapy on ventricular defibrillation in dogs. *Critical Care Medicine*, 8:157–163, March 1980.

[117] D. Yu, M. Small, R. G. Harrison, C. Robertson, G. Clegg, M. Holzer, and F. Sterz. Measuring temporal complexity of ventricular fibrillation. *Physics Letters A*, 265:68–75, 2000.