SPATIAL ENHANCEMENT OF ECG USING TRANSFORM DOMAIN MODELS

A

Thesis submitted

for the award of the degree of

DOCTOR OF PHILOSOPHY

By

Jiss J Nallikuzhy

DEPARTMENT OF ELECTRONICS AND ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781 039, INDIA

August 2018

TH-1915_11610237

Certificate

This is to certify that the thesis entitled "**Spatial Enhancement of ECG using Transform Domain Models**", submitted by **Jiss J Nallikuzhy** (11610237), a research scholar in the *Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati*, for the award of the degree of **Doctor of Philosophy**, is a record of an original research work carried out by him under my supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the Institute and in my opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Dated:
Guwahati.

Dr. Samarendra Dandapat
Professor
Dept. of Electronics and Electrical Engg.
Indian Institute of Technology
Guwahati - 781 039
India.

TH-1915_11610237

TH-1915_11610237

To my Wife

Asha

for her love, support, sacrifice and patience

&

To

Amma, Achachan, Appu, Tom and Rinku

for their love and support

TH-1915_11610237

TH-1915_11610237

Acknowledgments

I express my deepest and most sincere gratitude to my Ph.D advisor Prof. Samarendra Dandapat for his guidance, help and encouragement throughout my research work. I greatly admire his attitude towards research, creative thinking and enthusiasm for work. His insightful feedbacks have helped me greatly in improving myself scientifically and personally. I could not have imagined having a better advisor for my research work.

I am thankful to the members of my doctoral committee, Prof. H. B. Nemade, Prof. R. Sinha and Dr. Salil Kashyap for their support, encouragement, and suggestions rendered during my research work. I would also like to thank Prof. S. R. M. Prasanna, my ex-doctoral committee chairman for his support and encouragement.

I would like to thank Dr. A Rajesh, Dr. Tony Jacob, and other faculty members of the Department of Electronics and Electrical Engineering, IIT Guwahati, for their care and support. I would also like to acknowledge the help and support given by Dr. Benny George K and Dr. Ganesh Natarajan.

I would like to express my sincere gratitude to Dr. L. N. Sharma, senior technical officer and all other technical and office staff members of the department. I am also grateful to Mr. Uday Shankar Uzir and Mr. Mukut Baruah for their enormous help whenever required.

I am thankful to my friends Vivek Lukose, Vasudevan M.S., Anoop P., Sajith S., Hrishikesan V. M., Dileep P., Vijith Kumar K.P., Arjun P., Sooraj Chacko, Piyoosh P., Arun Mathew, Vishnu Venugopal T., Ardhendu Shekhar Chaudhury, and Kulkarni V. V. I am also thankful to Dr. Sandeep P., Dr. Haris B. C., Dr. Govind D., Dr. Syed Shahnawazuddin, Dr. Anurag Singh, Dr. Sibasankar Padhy, Dr. Nagaraj Adiga and to all the research scholars of EMST and Signal Informatics laboratory for their help and support. I also acknowledge the help and support shown by Eedara Prabhakara Rao and Vineeta Das in proofreading my manuscripts.

My deepest gratitude goes to my wife, my parents and my brothers. Without their love, support and sacrifice it wouldn't have been possible for me to complete my Ph.D.

Jiss J Nallikuzhy

TH-1915_11610237

TH-1915_11610237

Abstract

Electrocardiogram (ECG) is the most preferred non-invasive tool used by cardiologists to detect and diagnose various cardiovascular diseases. Electrodes, arranged in a specific order over the human torso are used for recording the signal. The lead signals acquired via these electrodes improves the spatial view of the heart. In other words, recording the signal with more electrodes improves the spatial resolution of ECG. The number of electrodes cannot be increased randomly as the comfort level of the patient is compromised, along with an increase in cost and complexity of the system. Another scenario arises in the case of ambulatory monitoring where it becomes extremely difficult to capture the lead signals. If fewer electrodes are used for recording, then the spatial resolution of the ECG is affected. A solution is to record ECG using fewer electrodes and use a system for enhancing the spatial resolution. Such systems are often recognized as the derived ECG systems and have many applications in personalized healthcare, telemonitoring, ambulatory monitoring, etc.

Among the various available electrode arrangements, the standard twelve-lead ECG system which uses ten electrodes is widely accepted by the cardiologists. The objective of this thesis is to derive the standard twelve-lead ECG from its subset without losing the diagnostic quality. The intra-lead and inter-lead correlations of the standard twelve-lead ECG in the transform domain is exploited for learning the derived ECG models. Three approaches are proposed using various machine learning techniques for enhancing the spatial resolution of ECG. The first two approaches exploit the inter-lead correlations in the wavelet and the sparse domain respectively, while the third approach simultaneously exploits both inter-lead and intra-lead correlations in the hidden feature space.

The first approach takes advantage of the improved inter-lead correlation in the wavelet domain. The proposed model learns a linear mapping between the same wavelet sub-bands of different lead signals. The model also selects the optimal predictor leads for a given patient using a lead selection algorithm. The lead selection algorithm is based on a new diagnostic similarity score, which computes the diagnostic closeness between the original and the derived lead signals. Analysis of the model using diagnostic and non-diagnostic distortion measures indicate that the diagnostic quality is maintained in the derived lead signals. The proposed model also performs better than the existing linear models in preserving the diagnostic information in the lead signals.

In the second approach, a new technique for utilizing the inter-lead correlation using a personal-

TH-1915_11610237

ized dictionary is explored. This approach integrates the dictionary learning and the sparse domain framework for learning the derived ECG model. Two models that aim to derive the standard twelve-lead ECG (high-resolution ECG) from its subset (low-resolution ECG) are proposed. In the first model, two over-complete dictionaries corresponding to the high and low-resolution ECG are jointly learned in the training phase. A conversion function is also learned for mapping the high and low-resolution sparse coefficients. During the reconstruction phase, an estimate of the high-resolution sparse coefficients is obtained by mapping the low-resolution sparse coefficients using the conversion function. This estimate of the high-resolution sparse coefficients is then used to obtain the derived ECG using the high-resolution dictionary. For the second model, the signal is divided into multiple segments using a thresholding technique, and joint dictionaries and conversion functions are learned for each segment. A similar procedure as that of the first model is followed during the reconstruction phase. Diagnostic and non-diagnostic distortion measures are used for evaluating the models. The analysis of the results shows that the proposed models are capable of capturing the diagnostic content in the lead signals.

The third approach captures the inter-lead and the intra-lead correlations simultaneously in the hidden feature space using the recurrent neural network (RNN) architecture. Three special cases of RNN, i.e., a simple RNN, a long short-term memory (LSTM) network and a gated recurrent unit (GRU) are used to derive the standard twelve-leads from its subset. The performance of these models is evaluated using the standard distortion measures as well as the diagnostic similarity tests to account for the loss of diagnostic information. These models are compared between each other to study and assess its ability to derive the standard twelve-leads without compromising significant diagnostic information. The results show that the diagnostic quality of the spatially enhanced ECG is comparable to the existing methods for simple RNN and GRU models and is much better for the LSTM model.

Keywords: Electrocardiography, ECG Synthesis, Derived ECG, Spatially enhanced ECG, Multi-scale analysis, Dictionary learning, Deep Learning, Recurrent neural network.

TH-1915_11610237

Contents

TH-1915_11610237

TH-1915_11610237

TH-1915_11610237

TH-1915_11610237

List of Figures

TH-1915_11610237

TH-1915_11610237

TH-1915_11610237

TH-1915_11610237

List of Tables

TH-1915_11610237

TH-1915_11610237

List of Acronyms

ADSM	Average Diagnostic Similarity Score
ALT	Antero-Lateral
ANN	Artificial Neural Network
ANT	Anterior
ASP	Antero-Septal
BBB	Bundle Branch Block
BPTT	Back-Propagation Through Time
BSPM	Body Surface Potential Map
CDM	Cardiomyopathy
DNN	Deep Neural Network
DSM	Diagnostic Similarity Measure
DSS	Diagnostic Similarity Score
DWT	Discrete Wavelet Transform
DYS	Dysrhythmia
ECG	Electrocardiogram
ERBF	Exponential Radial Basis Function
FN	False Negative
FP	False Positive
GRU	Gated Recurrent Unit
HC	Healthy Control
HR	High Resolution
HT	Hypertrophy

TH-1915_11610237

ICA	Independent Component Analysis
ILT	Infero-Lateral
INF	Inferior
IPL	Infero-Postero-Lateral
LA	Left Arm
LinR	Linear Regression
LL	Left Leg
LR	Low Resolution
LSLinR	Lead Selective Linear Regression
LSTM	Long Short-Term Memory
LSV	Least Squares Value
MECG	Multi-channel Electrocardiogram
MI	Myocardial Infarction
M-L	Mason-Likar
OMP	Orthogonal Matching Pursuit
PC	Principal Component
PCA	Principal Component Analysis
PTB	Physikalisch-Technische Bundesanstalt
RA	Right Arm
RBF	Radial Basis Function
RL	Right Leg
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SAECG	Signal Averaged Electrocardiogram
SS	State Space
SVEC	Stereovector Electrocardiograph
SVM	Support Vector Machine
SVR	Support Vector Regression
TN	True Negative

TH-1915_11610237

TP	True Positive
VCG	Vectorcardiogram
WDD	Weighted Diagnostic Distortion
WEDD	Wavelet Energy-based Diagnostic Distortion
WPRD	Weighted Percentage Root Mean Square Difference

TH-1915_11610237

TH-1915_11610237

1

Introduction

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी • Indian Institute of Technology Guwahati •

Contents

TH-1915_11610237

Electrocardiogram (ECG), even after a century of its inception, is the most preferred non-invasive tool used by cardiologists to detect and diagnose cardiac disorders. The ECG signal is produced by the electrical activities in heart. Hence, electrodes arranged in a specific order over the human torso can be used to capture the signal. The ECG signal consists of diagnostically significant features which differ among themselves in shape, duration and amplitude [2]. Any abnormality in the heart will be reflected as a change of these features in ECG. These changes are analyzed by the cardiologists for further procedures. Electrodes used for capturing ECG can be arranged in various configurations. Among these configurations, the standard twelve-lead ECG system which uses ten electrodes is the gold standard [3]. This is because limb leads in standard twelve-lead ECG provide information about the frontal plane and precordial leads provide information about the transverse plane of the heart.

The volume-conductor theory by Burger and van Milaan states that the electrical activities of the heart can be modeled using a heart vector or a heart dipole. The projection of this heart vector onto the human torso derives lead signals [2, 4, 5, 6]. An increase in the number of lead signals can improve the spatial resolution of ECG. This helps the cardiologists to analyze the heart in a better way. In other words, an increase in the number of electrodes can increase the spatial resolution of ECG as the lead signals are recorded using electrodes. As the number of electrodes increases, the comfort level of the patient is compromised along with an increase in the cost and the complexity. Also, it is strenuous to precisely place electrodes over the human torso at the time of ambulatory recording. If fewer electrodes are used for recording, then the spatial resolution of ECG is reduced which may make the diagnosis difficult. A trade-off can be achieved by recording the signal with fewer electrodes and improving the spatial resolution using information available from these acquired leads. This is possible as the information about the heart vector can be obtained from these few lead signals, which can then be used to derive other lead signals. Systems that utilize this concept to improve the spatial resolution are often identified as derived ECG systems [7, 1]. These systems record ECG signal using a specific set of leads known as the predictor leads and transforms it into derived leads or response leads using a previously learned model. These systems have a wide range of applications in recording, monitoring and/or processing of ECG [8]. This thesis investigates the capability of linear and non-linear models in capturing the heart vector and deriving the lead signals without losing diagnostic information.

TH-1915_11610237

Three major contributions are proposed in this thesis. The models discussed in this thesis are based on volume-conductor theory, and they exploit the spatio-temporal relationships between lead signals in the transform domain. The three approaches are as follows:

In the first approach, the model is learned by exploiting the wavelet domain characteristics of the signal. The inter-lead correlation of a twelve-lead ECG in the wavelet sub-band is higher compared to that of the original signals in the time domain [9]. This high correlation between different leads in the same scale can be exploited to learn a linear model. Also, the optimal predictor lead-set for a given patient is selected using a new diagnostic similarity score.

The second approach explores the applicability of dictionary learning and sparse domain representation in improving the spatial resolution of ECG. This model is learned in the sparse domain after obtaining the sparse representation of the standard twelve-leads and its subset. The ECG signal is divided into multiple segments and they are transformed into the sparse domain using a joint dictionary learning approach. The relationship between the sparse coefficients of the twelve standard leads and its subset are learned using a conversion function. Such an approach ensures a good reconstruction quality by preserving clinically significant diagnostic features.

In the third approach, the relationship between the standard twelve-lead ECG and its subset is learned in a high dimensional hidden feature space. A deep neural network is utilized for transforming the signal into the hidden feature space. The recurrent neural network (RNN) architecture of the deep learning framework is used to perform this transformation. The RNN architecture can simultaneously learn the spatio-temporal correlation between the lead signals and hence learn a better model. Three variations of RNN, simple RNN, long short-term memory (LSTM) unit, and gated recurrent unit (GRU), are used to learn the model. The three approaches are evaluated using the standard and the diagnostic closeness measures, and are discussed in the respective chapters.

This chapter is organized as follows. The standard twelve-lead ECG system and the multi-channel ECG system are discussed in section 1.1. In section 1.2, the morphological features of ECG is presented. The general approach used for enhancing the spatial resolution of ECG is described in section 1.3. The scope of the present work is discussed in 1.4 and the organization of the thesis is presented in section 1.5.

TH-1915_11610237

1.1 Standard twelve-lead ECG

Acquiring ECG using electrodes placed at different locations over the human torso helps in viewing the heart from different directions. Increasing the number of electrodes with the simultaneous recording of the signal implies a better spatial view of the heart. In other words, increasing the number of electrodes can increase the spatial resolution of ECG. Many systems are available that can record ECG using large number of electrodes. Systems that simultaneously record ECG using 24 leads or more are commonly termed as multi-channel ECG (MECG) or body-surface potential maps (BSPM) [10, 11].

Willem Einthoven, also regarded as the father of ECG, developed the first practical ECG recording system [4]. He used three limb leads, $Lead\ I$, $Lead\ II$ and $Lead\ III$, to form a triangle known as the Einthoven's triangle. These limb leads are captured from the body limbs and are linearly related to each other as

$$Lead\ III = Lead\ II - Lead\ I \ . \tag{1.1}$$

Many types of recording systems came into existence following Einthoven's work. The introduction of three augmented limb leads (aVR, aVL and aVF) and six precordial or chest leads (V_1 to V_6) improved the diagnosis of heart related diseases [12, 13, 14]. The augmented limb leads are related to the limb leads as

$$aVR = -\left(\frac{Lead\ I + Lead\ II}{2}\right) \ . \tag{1.2}$$

$$aVL = Lead\ I - \frac{Lead\ II}{2} \ . \tag{1.3}$$

$$aVF = Lead\ II - \frac{Lead\ I}{2} \ . \tag{1.4}$$

Later, the standard positions for placing the electrodes came into existence and is known as the standard twelve-lead ECG system [3, 11]. It consists of $Lead\ I$, $Lead\ II$, $Lead\ III$, aVR, aVL, aVF, V_1, V_2, V_3, V_4, V_5 and V_6. These twelve-leads are divided into eight independent leads and four derived leads. Derived leads consist of $Lead\ III$, aVR, aVL and aVF and can be derived from limb leads using the relationship given in equations 1.1, 1.2, 1.3 and 1.4. Hence, the eight independent leads recorded using ten electrodes can generate the standard twelve-leads. The limb leads, also known as the bipolar leads are recorded using three electrodes in addition to an optional electrode

TH-1915_11610237

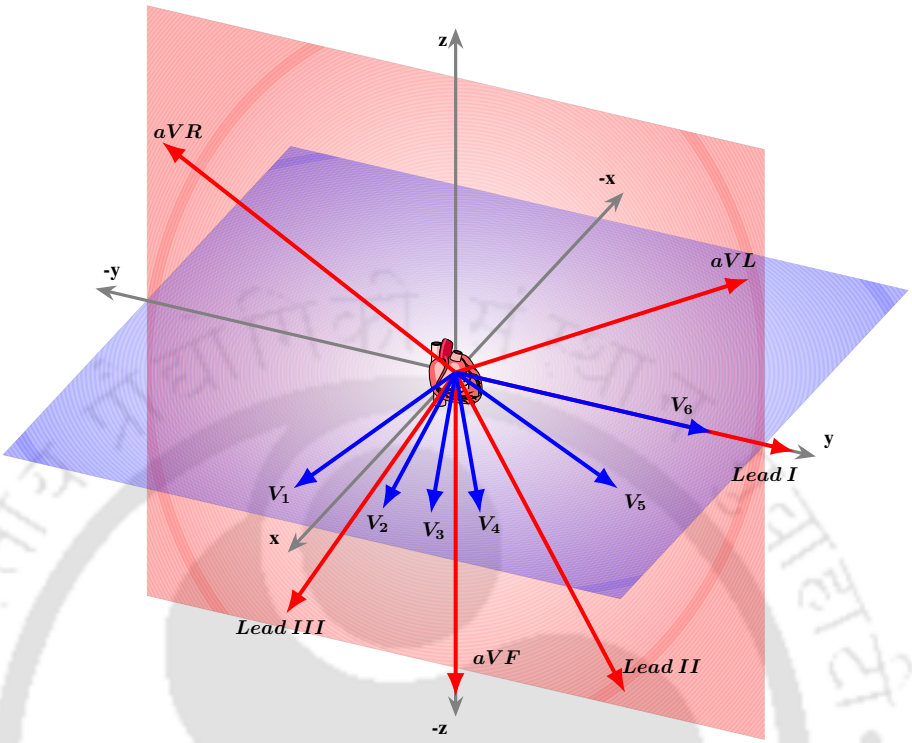


Figure 1.1: Schematic representation of the orientation of the lead vectors in a standard twelve-lead ECG based on volume-conductor theory [1].

for ground potential. Electrodes placed in right hand (RA), left arm (LA) and left leg (LL) form the Einthoven's triangle and the potential between electrodes form the limb leads. Electrode placed in right leg (RL) is the reference/ground electrode. The reference point for the augmented limb leads aVR, aVL and aVF is a central terminal known as the Wilson's central terminal. Wilson's central terminal is created by combining the terminals of the limb leads and represents the average of limb potentials [15]. The precordial leads are unipolar leads as it records the direct potential from the chest. There are standardized locations for placing the precordial leads. Leads V_1 and V_2 are placed at fourth intercostal space just to the right and the left of the sternum respectively. Lead V_4 is placed at fifth intercostal space in the midclavicular line. Leads V_5 and V_6 are placed next to V_4 at anterior axillary line and midaxillary line respectively. Lead V_3 is placed between V_2 and V_4.

1.1.1 Clinical significance of standard twelve-lead ECG

As discussed earlier, the volume-conductor theory enables to visualize the heart as a three dimensional organ [5, 6, 16]. Electrodes placed at different locations can provide important information

TH-1915_11610237

about the condition of the heart. The standard twelve-leads can capture the electrical activities of the heart from different directions. The limb leads and the augmented limb leads provide information in the frontal plane and the precordial leads provide information in the transverse plane. The schematic representation of the orientation of the lead vectors in a standard twelve-lead ECG based on volume-conductor theory is shown in Fig. 1.1. The twelve-standard leads can also be categorized into four groups based on specific anatomical areas of the heart. These areas are inferior, lateral, septal, and anterior. The inferior leads consisting of $Lead\ II$, $Lead\ III$ and aVF help in viewing the heart from the inferior or the diaphragmatic surface. Viewing the heart from the lateral wall of left ventricle can be achieved with the lateral leads $Lead\ I$, aVL, V_5 and V_6. Septal leads, V_1 and V_2 provide septal surface (inter-ventricular septum) view of the heart. Anterior view (sternocostal surface) of the heart is provided by the anterior leads V_3 and V_4. Thus the inferior, lateral, septal and anterior leads can capture hearts electrical activities in a better way and provide significant clinical information.

1.2 Morphological features of ECG

ECG is produced by the electrical activity of cardiac muscles [17]. ECG signal can thus indicate the nature and the physical condition of the heart. The signal can be recorded using electrodes placed at different locations over human torso to obtain a three dimensional view of the heart. This helps cardiologists in diagnosing various cardiovascular diseases. The ECG signal is characterized by $PQRST$-complex since it contains all relevant diagnostic information. The diagnostic features existing within the $PQRST$-complex can be divided into duration, shape and amplitude features [18]. A variation from the normal behavior of these features can be considered as an abnormality. Fig. 1.2 shows a single channel ECG depicting clinically significant features. The detailed description of these three features are discussed below. It should be noted that the diagnostic description of these features provided in this thesis are given as a technical aid and cannot be used for clinical diagnostics.

P-wave: The P-wave is generated when the electrical impulses generated at the sinoatrial (SA) node, travel from the atria towards the ventricles. The P-wave in the ECG signal represents the atrial depolarization. The duration of the P-wave for a normal person varies from 80 to 100 ms and possesses an amplitude of 0.25 mV. The atrial abnormalities are visible in the P-wave and are

TH-1915_11610237

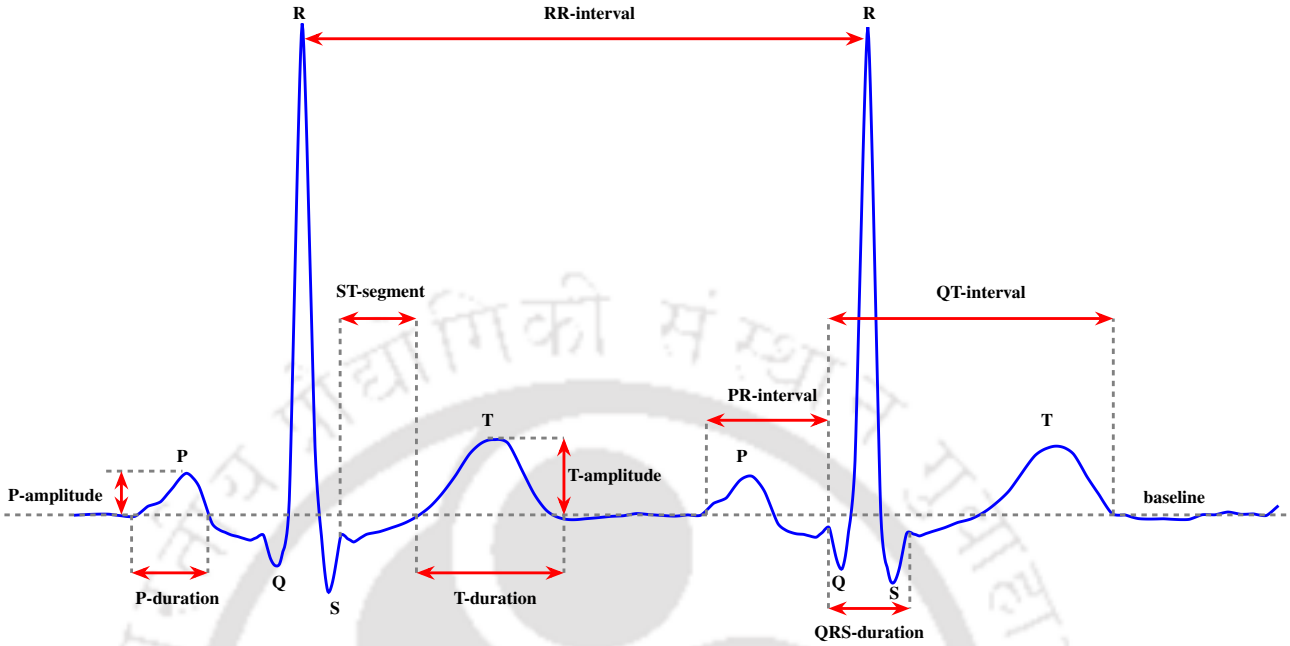


Figure 1.2: Morphological features of ECG

prominently visible in $Lead\ II$, $Lead\ III$, aVF and V_1 of the standard twelve-lead ECG.

QRS-complex: The QRS-complex in the ECG represents the ventricular depolarization. The QRS-complex is composed of Q-wave, R-wave and S-wave. The QRS-complex has the highest amplitude feature in ECG signal. This high amplitude results from the large electric potential caused by the ventricular muscles during depolarization. Common abnormalities visible in the QRS-complex include left and right bundle branch block (BBB), ventricular tachycardia, etc.

QRS-duration: The QRS-duration is the duration of the QRS-complex calculated from the onset till offset of the QRS-complex. The duration of the QRS-complex of a normal person can vary between 60 to 100 ms. A QRS-complex of duration above 100 ms is considered as an abnormality.

Q-wave: The Q-wave is the first wave in QRS-complex. It is identified as the first negative deflection before the R-peak. It represents the depolarization of inter-ventricular septum with a duration of 40 ms and depth of 0.2 mV for normal persons. Small Q-waves are observed in $Lead\ I$, aVL, V_5 and V_6. Pathological Q-waves are observed in leads V_1 to V_3 and they can indicate the presence of past or present cases of myocardial infarction.

R-wave: The R-wave is the first upward deflection after the P-wave and represents the early stages of ventricular depolarization. Common abnormalities include dominant R-wave in lead V_1 and aVR

and poor R-wave progression. These abnormalities can indicate the presence of right or left ventricular hypertrophy, bundle branch block or other myocardial diseases.

S-wave: The S-wave occurs as a downward deflection after the R-wave. It represents the later stages of ventricular depolarization.

PR-interval: The PR-interval is the duration between the start of atrial depolarization and the start of ventricular depolarization. The duration is typically between 120 ms and 200 ms. It is the segment starting from the beginning of the P-wave till the beginning of the Q-wave. It is also measured till R-wave in some special cases where the Q-wave is missing. Long PR-interval indicates first or second degree heart block and short PR-interval indicates pre-excitation syndromes like Wolff-Parkinson-White syndrome.

RR-interval: The duration between two R-peaks is known as the RR-interval.The RR-interval is commonly used to estimate the heart rate or the duration of a cardiac cycle. The heart rate is usually indicated in beats per minute (bpm) and varies between 60 to 120 bpm.

T-wave: The repolarization of the ventricles begins after the completion of ventricular depolarization. This give rise to the T-wave in ECG. It normally occurs as a positive deflection after the QRS-complex except in leads aVR and V_1. The T-wave duration in the normal ECG can vary between 100 ms and 250 ms. In the limb leads, the amplitude of the T-wave is less than 0.5 mV. The amplitude is less than 1.5 mV in the precordial leads. An inverted T-wave is an indication of myocardial infarction.

QT-interval: The total time taken for ventricular depolarization and ventricular repolarization is known as the QT-interval. It is measured from the starting of Q-wave till the end of T-wave. The duration of the QT-interval is between 350 ms and 450 ms. A prolonged QT-interval can be associated with the ventricular arrhythmias and short QT-interval can point towards tachycardia.

ST-segment: Duration between the ventricular depolarization and the ventricular repolarization is known as the ST-segment. The duration varies between 80 and 120 ms. It is represented as an iso-electric line for a normal ECG. An elevation or depression of this line occurs in case of abnormalities which includes myocardial infarction or myocardial ischaemia.

Table 1.1: Characteristics of waveforms in frontal plane leads

Leads	P-Wave	QRS-complex	T-wave	Views
Lead I	upright	upright	upright	lateral
Lead II	upright	upright	upright	inferior
Lead III	upright	upright	upright	inferior
aVR	negative	negative	negative	right atrium/left ventricle
aVL	upright	upright	upright	lateral
aVF	upright	upright	upright	inferior

Table 1.2: Characteristics of waveforms in transverse plane leads

Leads	P-Wave	QRS-complex	T-wave	Views
V_1	upright/biphasic	small R wave/QS	upright	septum
V_2	upright/biphasic	small R wave/QS	upright	septum
V_3	upright	equiphasic/QRS upright	upright	anterior
V_4	upright	equiphasic/QRS upright	upright	anterior
V_5	upright	upright	upright	lateral
V_6	upright	upright	upright	lateral

1.2.1 Characteristics of ECG signals in different leads

The characteristics of P-wave, QRS-complex and T-wave varies across the twelve-leads. Any variation from its normal characteristics can be considered as an abnormality. Characteristics of waveforms in the frontal plane leads are given in Table 1.1. The P-wave is always inverted in aVR and upright in all other frontal plane leads for the normal case. The QRS-complex is positive in all leads except aVR. In aVR, the QRS-complex is negative. Similarly, the T-wave is upright in $Lead\ I$ and $Lead\ II$, and inverted in aVR. It varies in other leads depending upon the QRS direction and the age of the patient.

Characteristics of waveforms in the transverse plane leads are given in Table 1.2. In the case of the transverse plane leads, the P-wave is upright in leads V_3, V_4, V_5 and V_6 for the normal ECG. It can be biphasic or upright in V_1 and V_2. The QRS-complex is negative with large S-wave in leads V_1 and V_2 and is equiphasic in lead V_3. It can also appear without a Q-wave in V_1 and V_2. In leads V_4, V_5 and V_6, the QRS-complex is positive with large R-wave. The T-wave is upright in precordial leads V_3 to V_6 for normal ECG, and that of other leads depend upon the QRS direction and the age of the patient.

TH-1915_11610237

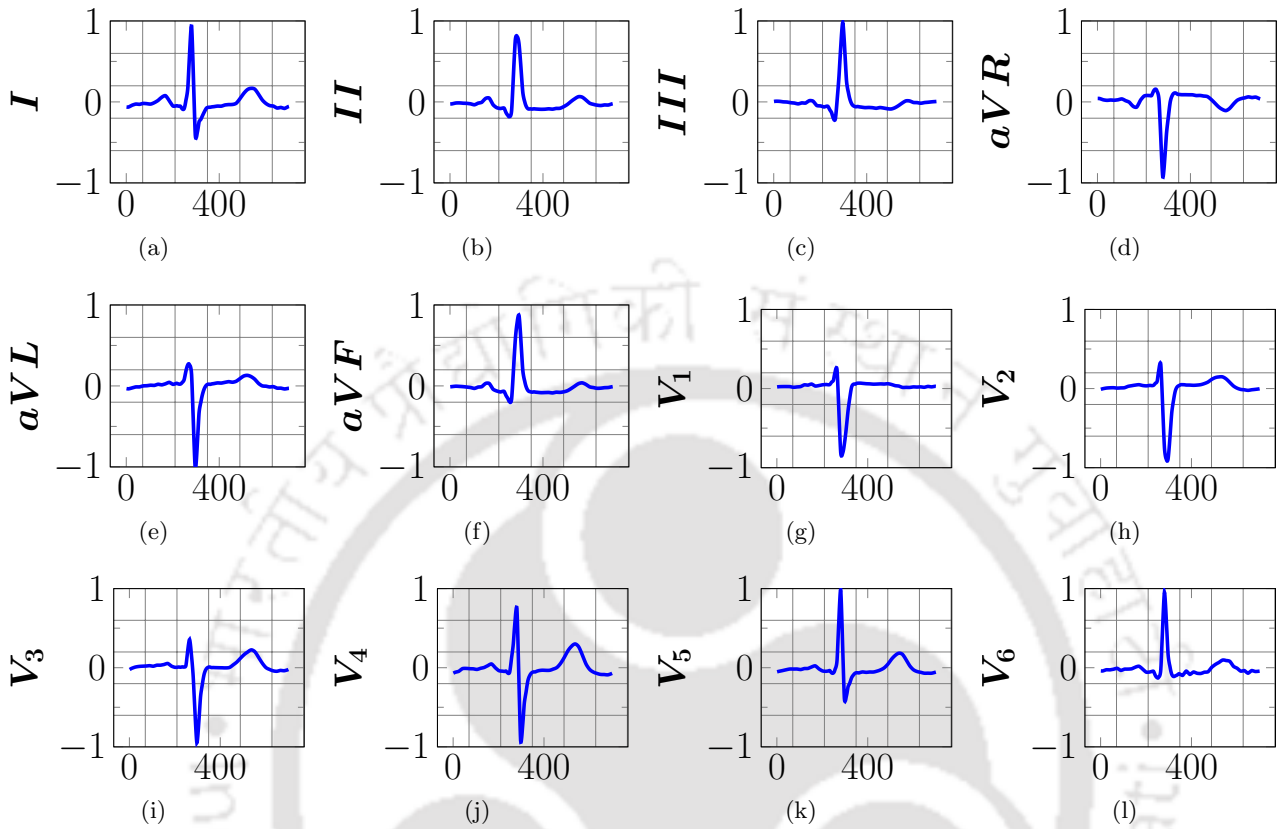


Figure 1.3: The standard twelve-leads of Healthy record 's0460_rem'. (a) $Lead\ I$; (b) $Lead\ II$; (c) $Lead\ III$; (d) aVR; (e) aVL; (f) aVF; (g) V_1; (h) V_2; (i) V_3; (j) V_4; (k) V_5; (l) V_6.

The standard twelve-lead ECG waveforms of a normal record 's0460_rem' is shown in Fig. 1.3 [19, 20]. The P-wave may get inverted in $Lead\ I$ for abnormal atrial rhythm conditions and is absent in junctional rhythm. In cases like the left atrial enlargement, duration of the P-wave changes and can be clearly observed in $Lead\ I$, $Lead\ II$ and aVF. Important pathological characteristics can be observed from the height and the width of QRS-complex, especially in case of conduction diseases like bundle branch block. It appears in the form of wide, slurred S-wave in $Lead\ I$, V_5 and V_6 and secondary R-wave in V_1 or V_2. Secondary R-wave (R') is an M-shaped wave with a R-wave appearing before the S-wave. Abnormal QRS-complex can appear in case of rapid ventricular tachycardia and ventricular fibrillation as well. Inverted T-wave in $Lead\ I$, $Lead\ II$, V_3 to V_6 along with an ST-segment depression indicates probable myocardial ischaemia. Severe myocardial ischaemia can cause a taller T-wave. Generally, inverted T-wave points to myocardial infarction, cardiomyopathy, etc. Further details regarding various cardiac pathologies can be found in [21]. Also, more details related to the

anatomy and physiology of the heart can be found in [22].

1.3 Spatially enhanced ECG systems

The spatially enhanced ECG systems are those systems that are used to improve the spatial resolution of ECG. According to the volume-conductor theory by Burger and van Milaan, leads are spatially oriented in a three-dimensional space [5, 6, 16]. The information regarding the electrical activities of the heart can be improved by increasing the number of leads since the leads are nothing but the projection of the heart vector. In other words, the resolution is improved when more number of leads are used. The term spatial resolution can be used here as the leads are spatially oriented in the three-dimensional space. This means, if a lead system is available which can record multiple leads at the same time, the spatial resolution of that system is more. The spatial resolution of the ECG obtained by a particular lead system depends upon the number of electrodes used for its acquisition. Hence, recording the signal with more electrodes implies a better spatial resolution for the acquired ECG. But the bottleneck arises when an increase in the number of electrodes compromises the comfort level of the patient. Similarly, a mechanism to simultaneously process signals from a large number of electrodes increases the complexity level of the whole procedure along with the cost. Another scenario arises in the case of ambulatory monitoring where it becomes extremely difficult to obtain even the standard twelve-leads. If less number of electrodes are used for recording, then the spatial resolution of the recorded ECG is affected.

A solution to the above issues is to obtain ECG using less number of electrodes and use a system which can enhance its resolution. This can be achieved by utilizing the inter-lead and the intra-lead correlation between the lead signals. The systems that can increase the spatial resolution of ECG recorded from few leads are commonly identified as the derived ECG systems [7, 1]. These derived ECG systems consist of a previously learned linear or non-linear model, which accepts inputs from a specific set of leads known as the predictor leads, and outputs the derived leads [1]. The output of these derived ECG systems can also be termed as spatially enhanced ECG as it improves the spatial resolution. In this thesis, the derived ECG system or a spatially enhanced ECG system refers to an approach or method used to improve the spatial resolution of ECG. Derived ECG systems posses many advantages over the conventional recording systems such as reduced complexity, cost

TH-1915_11610237

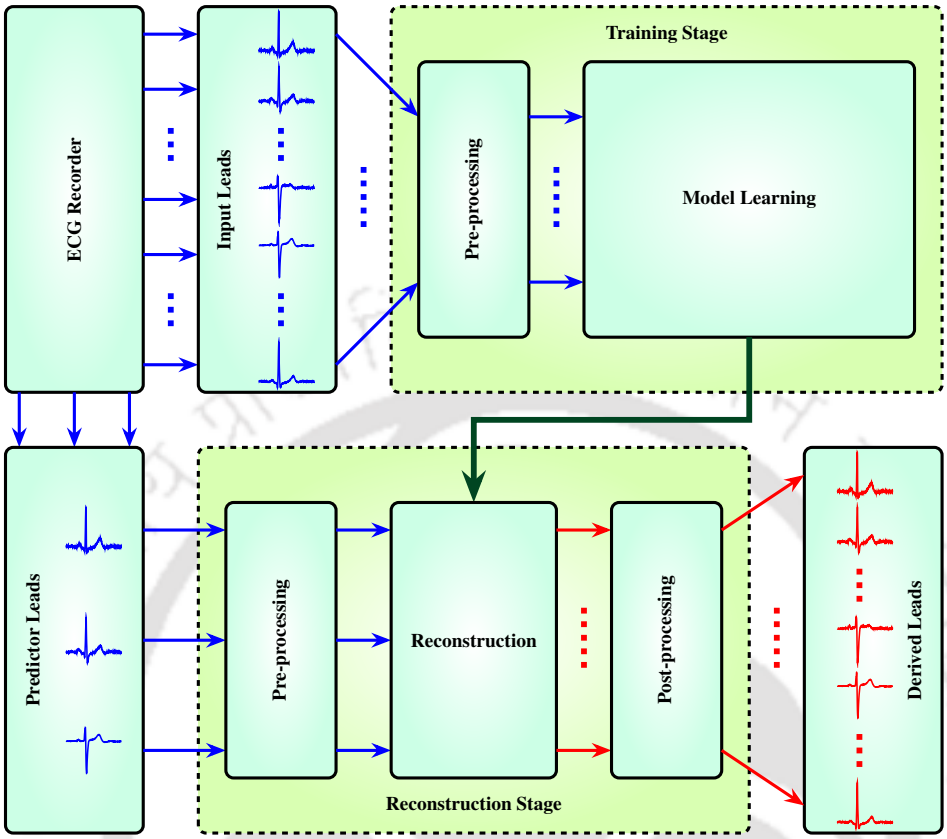


Figure 1.4: General structure of a patient-specific system that is used for deriving the standard twelve-lead ECG from its subset.

and an improved comfort level for the patients. The application of a derived twelve-lead ECG comes into the picture in those areas where recording, monitoring and/or processing of a standard twelve-lead ECG are difficult. This includes ambulatory monitoring, remote health care, personalized health care, continuous monitoring, etc. The difficulty in placing ten electrodes for recording the standard twelve-lead ECG during ambulatory monitoring can be made easier by using fewer electrodes. In applications like remote health care, a reduction in the number of leads can be utilized for data compression, transmission, etc. [8]. Also, in continuous monitoring as well as personalized health care, the patient's comfort is improved by reducing the number of electrodes for recording. The real challenge in the development of a derived ECG system lies in its capability to capture significant diagnostic information [7, 1].

The standard twelve-lead ECG system recorded using ten electrodes is the most popular lead system among cardiologists. Because of its popularity, systems that can transform fewer leads to

standard twelve-lead ECG are developed. One of the most popular derived ECG approach is to generate the standard twelve-lead ECG from its subset. The general procedure followed in such an approach is shown in Fig. 1.4. The twelve-leads are recorded using a twelve-lead ECG recorder and the whole signal is passed into the training stage. The training stage consists of the pre-processing block and a model learning block. The signal is pre-processed for removing baseline noise and power line interference, amplitude normalization, data arrangement, etc. The model learning block learns the model parameters, either in time domain or transform domain. The learned parameters are then passed into the reconstruction block in the reconstruction stage. In the reconstruction stage, input data is acquired only from few leads. These leads are pre-processed using a similar approach as that of the training stage. The pre-processed signal is then passed into the reconstruction block which, by using the previously learned model parameters, derives the missing lead signals. A post-processing block can be used to improve the quality of the signal. The block diagram shown in Fig. 1.4 follows a patient-specific approach. The accuracy of the derived lead signals depends upon the capability of the model in capturing the intra-lead and inter-lead correlations.

1.4 Scope of the present work

The standard twelve-lead ECG provides the cardiologists with a better view of the electrical activities in the heart. However, during continuous monitoring or ambulatory monitoring, acquisition of the twelve-lead signals using ten electrodes become extremely difficult. A standard twelve-lead ECG generated using information available from fewer lead signals can be a better alternative. It is a well known fact that standard ECG posses both inter-lead and intra-lead correlations [23]. These correlations can be exploited for generating a standard twelve-lead ECG from fewer leads. The objective of this thesis is to exploit various machine learning techniques that can utilize the intra-lead and inter-lead correlations in generating derived ECG models that can capture the diagnostic information.

Several approaches utilizing the inter-lead correlation are available for deriving a spatially enhanced ECG. Most of these models are learned using linear approaches by exploiting the spatial correlation in time domain or transform-domain. The quality of the derived ECG is a major factor of consideration as it contains significant clinical information. The question is whether the existing

TH-1915_11610237

models are capable of exploiting these correlations in such a way as to preserve the significant diagnostic information. This opens up the scope of the present work. A simple approach to exploit the spatial correlation is to localize the morphological features by decomposing ECG signal into different frequency sub-bands using discrete wavelet transform (DWT) [9, 24]. Since localization of the morphological features increases the correlation between leads, it is expected that learning models for each sub-band could improve the quality of the signal. The capability of the derived models in preserving diagnostic information depends upon how accurately it can represent the morphological features. A personalized dictionary can be used to map the ECG morphologies into another domain where it can represent the features in a better way. Such an approach when used with models that are customized for a single patient is expected to improve the reconstruction accuracy. One such approach is to learn the model in the sparse domain by transforming the signal using a personalized joint dictionary. The standard twelve-lead ECG possesses both spatial and temporal correlations and the previous two approaches focused only on spatial correlations. A model which can simultaneously learn the spatio-temporal correlations is expected to capture the diagnostic features. This can be achieved by using recurrent neural network (RNN) based models as it can learn the spatial and temporal features simultaneously.

Another aspect is the patient-specific approaches for learning the models. The morphological features can differ from patient to patient and hence the patient-specific approach is much better than a global model. As the objective is to capture significant diagnostic information while deriving lead signals, a patient-specific model is expected to improve the correlation in transform-domain than in time-domain. By using a suitable model, these correlations can be exploited much better for preserving the significant diagnostic information in the derived ECG leads. This can improve the quality of derived ECG systems which has a wide range of applications.

1.5 Organization of the thesis

The thesis is organized as follows:

In this chapter , i.e. **Chapter 1**, a general introduction to ECG, its clinical significance and necessity of the derived ECG system are discussed. A review of the approaches used for a derived ECG system along with its theoretical background is presented in **Chapter 2**.

TH-1915_11610237

Chapter 3 presents a novel approach to utilize the inter-lead correlation in wavelet domain for deriving the standard twelve-lead ECG. This approach learns the model in the wavelet domain using a linear model by selecting the best predictor lead. The selection of predictor lead is performed by a lead selection algorithm which is based on a novel diagnostic similarity score. The performance of the model is evaluated and compared with the existing linear approaches.

Chapter 4 proposes a new approach for exploiting the dictionary learning and the sparse domain framework in deriving standard twelve-lead signals. A personalized dictionary is learned for transforming the ECG into the sparse domain and a model is learned by mapping the sparse coefficients of predictor and response leads. This is then transformed back to obtain the derived signals. The importance of segmentation is also explored in improving the reconstruction accuracy of the derived signals. The performance evaluation of the model and its comparison with the existing models are also carried out.

Chapter 5 investigates a new approach to simultaneously exploit the intra-lead and inter-lead correlations in ECG. This approach uses RNNs for learning the spatio-temporal correlations between lead signals. The capabilities of the three different RNNs consisting of simple RNN, long short-term memory (LSTM) network, and gated recurrent unit (GRU) in enhancing the spatial resolution of ECG are analyzed. The performance of these models are evaluated and are compared with the existing methods.

The performances of the present approaches are summarized, and conclusions are drawn in **Chapter 6**. The future direction of research for improving the derived ECG systems are also presented.

TH-1915_11610237

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

2

Spatial Enhancement of ECG - A Review

Contents

TH-1915_11610237

A general introduction to ECG and its clinical significance are discussed in Chapter 1. In Section 1.3, a detailed description about the structure of a spatially enhanced ECG system is presented. The concept of a spatially enhanced ECG system relies on the volume-conductor theory, the heart-vector and its projection onto the human torso [5, 6, 16, 25]. Several models that utilize this concept using linear or non-linear transforms are available for deriving lead signals. Application of such a system comes into the picture in those areas where recording, monitoring and/or processing of ECG with a large number of electrodes are difficult. This includes ambulatory monitoring, remote health care, personalized health care, continuous monitoring, etc.

The main focus of this chapter is to discuss the theoretical background of derived ECG systems along with its implementation using linear and non-linear models. The chapter is organized as follows. In section 2.1, the lead theory, which is the backbone of the derived ECG system is described along with the different lead systems. The models that are used to enhance the spatial resolution of ECG are presented in section 2.2. Performance evaluation of a model is as important as its implementation, and is computed using standard metrics. Hence in section 2.3, the database and the performance evaluation metrics used for evaluating the models are presented. The motivation for the methods and problems discussed in this thesis work is explained in section 2.4 which is followed by the plan of the work in section 2.5.

2.1 Lead theory and lead systems in electrocardiography

In this section, lead theory and various lead systems currently available to capture ECG are discussed. The theoretical background behind the acquisition of lead signals using electrodes placed on human torso is explained using the lead theory [2]. Various lead systems that are used to capture ECG are discussed in the second part of this section.

2.1.1 Lead theory in electrocardiography

Lead theory in electrocardiography is used to define the relationship between the potential differences over the body surface and the electrical activity in the heart. There are two popular lead theories: (1.) The classical lead theory by Einthoven [4, 25] and (2.) The volume-conductor theory developed by Burger and van Milaan [5, 6, 25]. The basis of Einthoven's classical lead theory is established

TH-1915_11610237

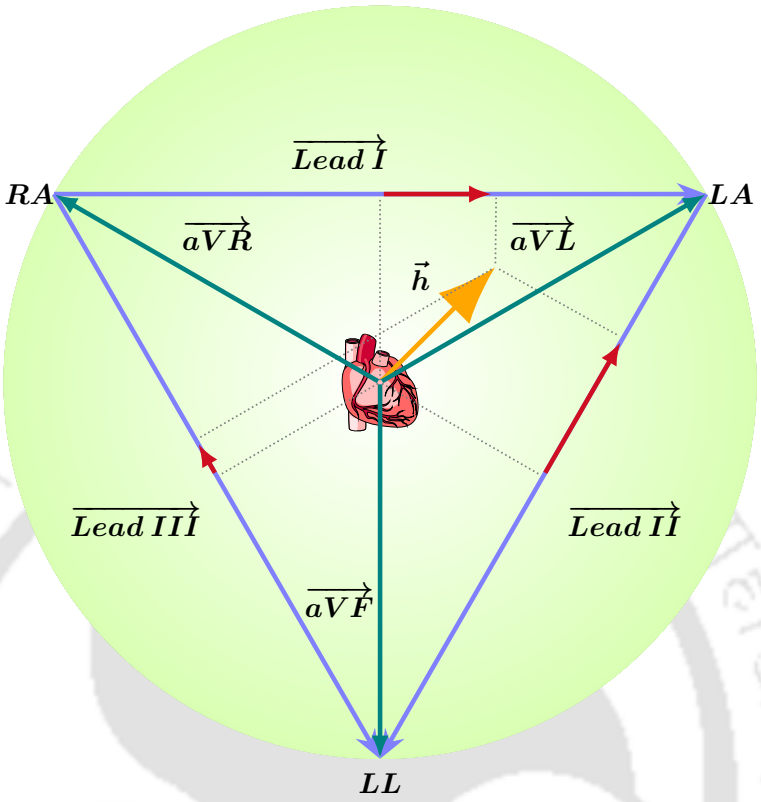


Figure 2.1: Classical lead theory: The projection of heart vector for generating limb leads and augmented limb leads.

over the assumption that the human body is a homogeneous conductor and electrical activities of the heart can be represented by a dipole [2]. The dipole is termed as the heart dipole or the heart vector and its location is fixed at the center of an equilateral triangle known as the Einthoven's triangle [4, 25]. The Einthoven's triangle is oriented in the frontal plane and the projection of the heart vector onto the edges of this triangle derives the limb leads. In Einthoven's classical theory, the dipole is depicted as a vector in two dimensions. This is shown in Fig. 2.1. The lead vectors $\overrightarrow{aVR}$, $\overrightarrow{aVL}$ and $\overrightarrow{aVF}$ are unipolar vectors and are oriented in the direction of electrodes positioned at RA, LA and LL respectively. The potential of the lead vector is the projection of the heart vector $\vec{h}$ onto the unipolar vectors. The bipolar lead vectors $\overrightarrow{Lead\ I}$, $\overrightarrow{Lead\ II}$ and $\overrightarrow{Lead\ III}$ are measured between two electrodes and the potential is the projection of the heart vector $\vec{h}$ onto the lead vectors.

The volume-conductor theory also assumes that the heart vector is originated from the center of the heart. It is also assumed that the body is an irregularly shaped three dimensional inhomogeneous

TH-1915_11610237

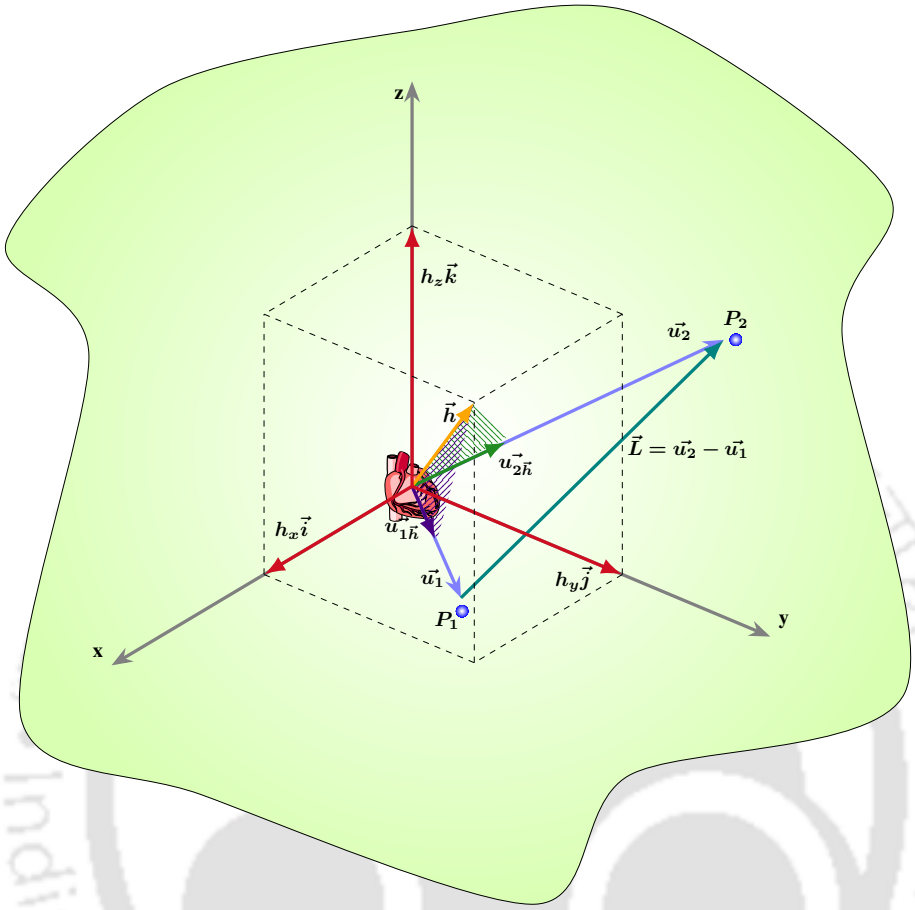


Figure 2.2: Volume-conductor theory: The projection of heart vector in three dimensional space for generating unipolar and bipolar leads.

volume-conductor. Electrical activity of the heart is again represented by the heart dipole or the heart vector $\vec{h}$. By projecting this vector onto any lead vector in three dimensional space, potential anywhere on the human torso can be determined. This is shown in Fig. 2.2 and can be represented mathematically using the Burger equation which is given in equation 2.1. The equation is derived based on the assumption that, the body is a linear physical system and the potential on a given lead can be obtained by projecting the heart vector onto the lead vector. The lead vector in case of volume-conductor theory represents the cardiac activity in a specific direction in three dimensional space.

$$L = |\vec{L}|.|\vec{h}|.\cos(\theta) \quad . \tag{2.1}$$

In equation 2.1, L is the potential difference between two points in three dimensional space, $\vec{h}$ is

the heart vector, $\vec{L}$ is the lead vector and θ is the angle between $\vec{L}$ and $\vec{h}$. The heart dipole $\vec{h}$ can be represented as $h_x\vec{i}+h_y\vec{j}+h_z\vec{k}$, where $\vec{i}$, $\vec{j}$ and $\vec{k}$ are the unit dipoles in a three dimensional coordinate system. In the case of unipolar electrodes such as the precordial leads, $\vec{L}$ is oriented between the origin and a point in three dimensional space. The point in the three-dimensional space is the electrode position on the human torso. Thus, the lead vector can be interpreted as the monitored spatial direction of the cardiac electrical activity [1]. This is shown in Fig. 2.2. The unipolar lead vectors at points P_1 and P_2 are represented by $\vec{u_1}$ and $\vec{u_2}$. The projections of the heart vector $\vec{h}$ onto the unipolar lead vectors are given by $\vec{u}_{1\vec{h}}$ and $\vec{u}_{2\vec{h}}$ and their magnitudes provide the potentials at points P_1 and P_2 respectively. For bipolar leads, the potential difference is calculated by taking the difference between two different points in three dimensional space. This is shown in Fig. 2.2 as $\vec{L} = \vec{u_2} - \vec{u_1}$. In these two lead theories, the lead vector depends purely on the location of the dipole, the positioning of lead electrodes and the electrical characteristics and shape of the conductor [2, 26].

Burger and van Milaan also introduced the concept of image surface [16]. This can be obtained by determining the potential at arbitrary number of points on the surface of the model [16]. These potentials, which are nothing but tips of unipolar lead vectors on the surface, can be connected together to form a cross-section of an imaginary torso. This is known as the image surface and the line connecting these points represent the electrical activity of the heart. Deviation of image points on the imaginary torso from that of the physical torso can be seen as distortion of the dipole behavior on the surface of the torso [27]. Frank studied the effect of the surface of the body on the leads by determining the image surface experimentally [28, 2, 26]. He concluded that inhomogeneities in the model and shape of the torso are less important and the location of the dipole inside the model is the major factor which influences the surface potentials [29]. He also concluded that a fixed-location dipole model can be used to explain almost all electrical activities of the heart [29, 30].

Assuming that the exact location of the dipole is known, there exists different volume-source models such as the dipole model, moving dipole model, multiple dipole model, etc. [31, 32]. But, it has been shown that infinite number of electrical heart sources are available internally and each of these sources are capable of inducing three dimensional potential distribution [29, 32]. This makes it difficult to find the exact location of the source of the volume-conductor models. Thus, to find the unknown source, an inverse problem has to be solved. The inverse problem seeks to find the source,

TH-1915_11610237

given the field and the conductor [32]. Various approaches are available to solve the inverse problem of which the simplified modeling method is a popular one [32]. In this approach, separate models are created for volume-source and volume-conductor. The volume-source models, for example, the dipole models, should have a limited number of independent variables and accuracy of the conductor model (in this case the human torso) should be good or better than source models [32]. Now, by taking as many independent measurements as the number of independent variables, a solution can be obtained using the least squares approach [32]. Miller and Geselowitz proposed a source model which considers the interconnected cells in the myocardium [33]. They showed that the current density in the intracellular and extracellular medium are similar. Hence the current density in the myocardium and the spatial gradient of intracellular potential distribution is proportional to each other. Geselowitz also showed that in case of a homogeneous conductor, the equivalent of the heart dipole is nothing but the integral of current density over the heart volume [34]. The equivalent of the heart dipole can be considered as a single dipole, and it can be justified by interpreting the current density as dipole density or dipole moment per unit volume [35]. Thus it is safe to assume that the sum of all individual dipole contributions over the entire volume of the heart generates the heart vector. Geselowitz, in his study of dipoles, indicated that any complex source can be approximated as a dipole if it has zero net current [30]. He also found that an improvement in the approximation can be achieved if the observation distance become greater than the largest dimension of the source. The condition that the algebraic sum of currents being zero is fulfilled by the heart, as the net charge generated at any instant is zero. Also, when distance increases, the deviation of the heart vector from the centroid of the heart becomes negligible. Therefore, the location of the heart vector is usually fixed at the center of the heart [1].

From the previous discussions, it can be shown that by measuring any three independent leads, the dipole can be evaluated. Thus, a derived twelve-lead ECG system can be obtained by measuring any three independent leads. Linear transformation is one of the most popular approaches used to generate a derived ECG. This is based on the assumption that the volume-source models are linear and quasi-static. It has been shown that by considering the body tissues as a resistive media and by neglecting the electromagnetic propagation effect, the volume-source models can be quasi-static [1, 36, 37, 38]. As the single dipole location model assumes the heart vector $\vec{h}$ at the center of the

TH-1915_11610237

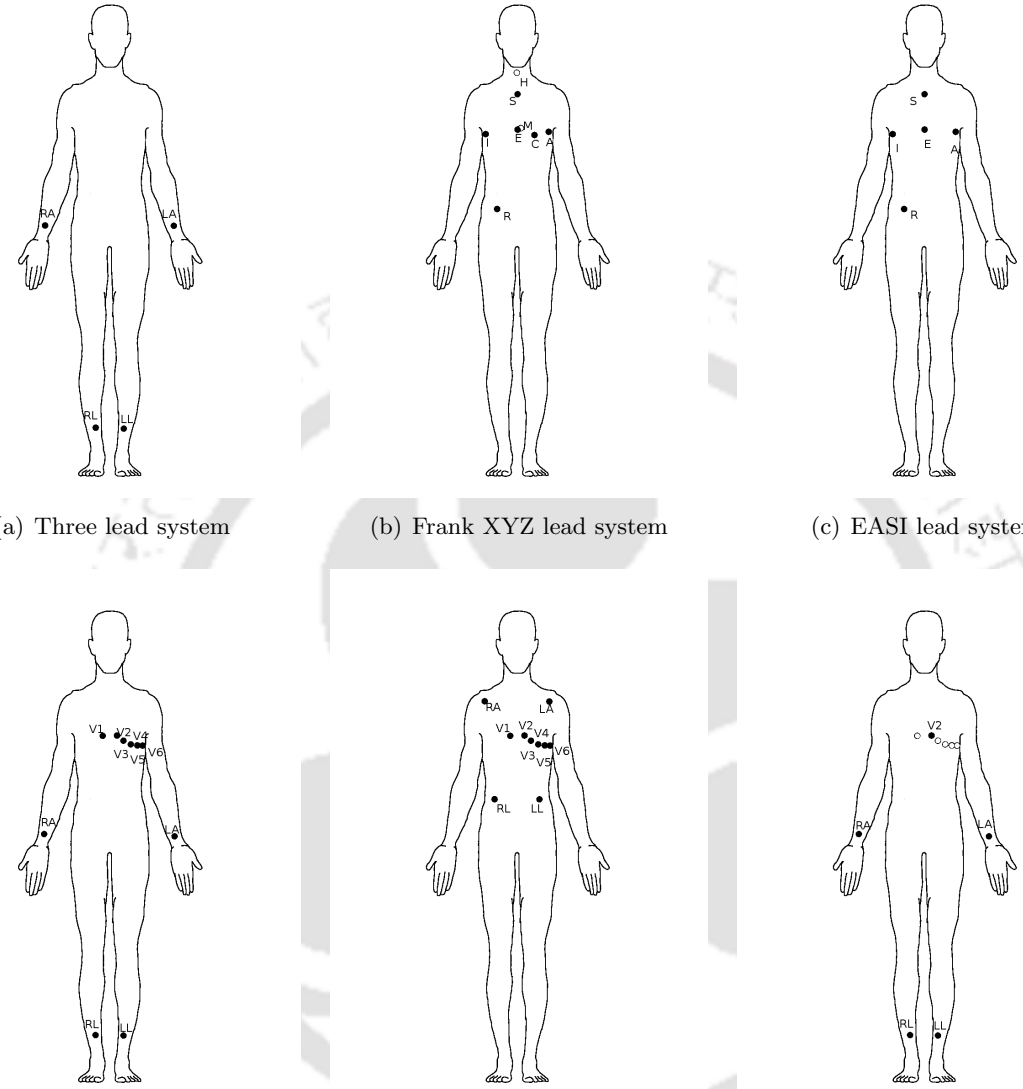
heart and by using equation 2.1, the potential of any lead on the surface of the torso can be estimated. Thus, with a prior knowledge about the three independent leads, and measured potential values with no noise present, three equations can be solved to obtain a unique solution for the components of $\vec{h}$ [2]. Thus, it can be shown that by using a minimum of three independent leads, the heart vector can be obtained which can then be used to generate a twelve-lead ECG or any other lead system. It can also be shown that three orthogonal leads produce better results compared to other leads [25]. It is also found that there is no universal location for the dipole as it varies from person to person, depending upon the body surface [29].

2.1.2 Lead systems in electrocardiography

There are various methods available for recording electrocardiogram. These methods vary from each other in the number of electrodes required for recording as well as the positioning of the electrodes placed over the body. An increase in the number of electrodes used for recording improves ECG diagnosability as it enhances the spatial resolution. It can vary from a few electrodes to hundreds of electrodes. Various electrode placing standards are available and the most common ones are discussed.

Einthoven developed the first practical system to record ECG where he used the three leads $Lead\ I$, $Lead\ II$ and $Lead\ III$, for analyzing the electrical activity of the heart [4, 39, 40]. These leads, commonly referred as the limb leads are obtained from the Einthoven's triangle. This system is commonly known as the three-lead system and is shown in Fig. 2.3(a). It requires three electrodes in addition to an optional electrode for ground potential. The three-lead system has been considered as the standard electrode configuration until its inadequacy to evaluate patients with different cardiac diseases was found [39]. The concept of unipolar leads that utilized a central terminal was introduced by Wilson [12]. The central terminal, popularly known as Wilson's central terminal, is a reference point that is created by connecting the limb leads together. A unipolar lead is obtained by measuring the potential difference between any location over the body and the central terminal. It is also reported that Wilson showed the importance of using unipolar leads for the detection of myocardial infarction [39]. Later, six standard locations for placing the unipolar leads, commonly known as precordial leads or chest leads were introduced and they were named as V_1 to V_6 [13]. Goldberger introduced the three

TH-1915_11610237



(a) Three lead system (b) Frank XYZ lead system (c) EASI lead system

(d) Standard twelve-lead system (e) Mason-Likar lead system (f) Subset of twelve-leads

Figure 2.3: Electrode placement in different lead systems.

augmented limb leads, aVR, aVF and aVL by taking the potential between the central terminal and each limb electrodes [14]. The standardization of the lead system introduced the standard twelve-lead ECG system, which later become the most common method for recording the ECG [3]. This configuration requires ten electrodes for recording and it consists of Einthoven's standard limb leads, precordial leads and augmented limb leads. It is shown in Fig. 2.3(d).

After the introduction of volume-conductor theory, a new system for recording the ECG came into

the picture known as the vectorcardiogram (VCG) [41]. This was introduced by Frank, commonly known as Frank XYZ system and it requires a minimum of four electrodes [41]. But he used eight electrodes (six electrodes on the front side and two electrodes on the rear side of the body) in order to improve the accuracy and is shown in Fig. 2.3(b) [41]. Since the standard twelve-lead ECG is the most favored system by the cardiologists, systems that derive the standard twelve-lead ECG became popular. The EASI lead system is one such system which is used to derive the standard twelve-lead ECG. This was developed by Dower and used five electrodes for recording [42]. This is shown in Fig. 2.3(c). Three approximately orthogonal leads, the E-S, A-S and A-I are formed from these electrodes [1]. The location of E, A and I electrodes are obtained from the Frank XYZ lead system and the S location is the manubrium. The leads thus obtained are then transformed to the standard-twelve-lead ECG using a linear combination of three leads [42].

A system to record exercise ECG was developed by Mason and Likar, known as the Mason-Likar (M-L) lead system and is shown in Fig. 2.3(e) [43]. This system is very similar to the standard twelve-lead ECG and the main difference is the way in which the electrodes are placed to measure the limb leads. The motive was to develop a system in order to reduce the effect of wrist and ankle motion while recording the exercise ECG [43]. Another popular approach is to use a subset of the standard twelve-lead ECG to derive the other leads [44, 45, 46, 47]. It uses two limb leads and a precordial lead to generate the twelve-lead ECG. In the next section, a detailed analysis of those systems which derive the standard twelve-lead ECG are discussed.

2.2 Systems for enhancing the spatial resolution of ECG

As discussed in the previous section, different lead systems are available for recording ECG. Some systems require fewer electrodes, while others give a better spatial resolution which helps in analyzing cardiac activities in a better way. Various methods are developed for transferring one lead system to another so as to reduce the disadvantages of some systems. In this section, three different approaches for generating ECG lead signals are discussed. The first approach generates one lead system from another using linear methods in time-domain or transform-domain. The standard twelve-lead ECG is derived using linear or non-linear models from its subset in second and third approaches. The methods used for generating global models are discussed in the last part of this

TH-1915_11610237

section. In linear models, the mapping function between the predictor and the response leads is linear, while it is non-linear in case of non-linear models.

2.2.1 Transformation between lead systems: Linear approach

Most of the transformations performed in the early stages mainly concentrated upon deriving the precordial leads from VCG. A device named “resolver” was used for lead derivation which was based on the work by Schmitt [48, 49]. Milnor *et al.* created a device to derive precordial leads using VCG [50]. This study showed that the generated precordial leads posses good similarity with the original leads. Another custom device to generate lead signals was developed by Helm, which linearly transformed the XYZ leads to one of thirteen leads oriented in frontal, sagittal or transverse planes [51]. Later he modified this system using custom electrodes and derived the precordial leads using Frank image space [52]. A resolver capable of rotating any XYZ lead system was developed by Pipberger and Wood [53]. The stereovector electrocardiograph (SVEC) III system was used to record the orthonormal leads [49]. Their study compared the derived system with the standard twelve-leads and showed that derived leads could be used to reveal clinically significant details which were not present in the XYZ system [54]. A similar method was utilized by Okada *et al.* to derive the precordial leads from the SVEC III vectorcardiographic system [55]. Their study showed that the wave shape of the derived precordial leads with the original leads are visually equivalent. Abildskov and Wilkinson derived precordial leads from orthogonal leads which were obtained using the McFee-Parungao system [56, 57, 58]. They also derived the precordial leads and compared it with that obtained from the standard twelve-lead ECG. In all these methods,VCG was recorded initially to obtain the XYZ leads, which were then rotated in order to derive the precordial leads.

Dower derived the standard twelve-leads from Frank XYZ leads using a linear transformation [59, 60]. He designed a simulator to derive the standard twelve-leads from Frank XYZ leads. This linear transformation is a universal transform, termed as the Dowers transform, which is based on the Frank torso model. The simulated twelve-lead generated using Dowers transform produced good resemblance with the standard twelve-leads. Later, Dawson presented a linear affine transform to derive the standard twelve-leads from Frank XYZ leads and vice versa [61]. This method gave a more consistent and accurate transform compared to the Dowers transform. Similar effort to transform

TH-1915_11610237

Frank XYZ lead system to standard twelve-lead was performed by Maheshwari *et al.* using least-squares fit method [62]. This was a patient-specific model which produced better accuracy than the previous methods. Efforts are also made to transform the standard twelve-lead ECG to VCG [63, 64, 65].

After the introduction of the EASI lead system by Dower, EASI derived twelve-lead ECG came into existence [42]. These conversions mainly utilized the Dower conversion box as the transformation system [1]. Klein *et al.* used the Dower conversion box to derive standard twelve-leads from EASI leads and observed a good similarity with original twelve-leads in case of different cardiac arrhythmias [66]. Similar study was performed for finding the accuracy of the EASI derived twelve-leads with the standard twelve-leads by Drew *et al.* [67, 68, 69]. Their study showed that EASI derived ECG is diagnostically equivalent to standard twelve-leads in detecting cardiac arrhythmias. Horáček *et al.* presented another study to transform EASI to standard twelve-lead ECG using universal linear regression and observed good similarity for the derived leads with the original leads [70]. Sejersten *et al.* performed a comparative study between the EASI-derived twelve-lead and paramedic acquired twelve-lead ECG [71]. They concluded that the EASI system can be a good alternative to the standard twelve-lead ECG.

In 2002, Dowers conversion box was replaced by Feild *et al.*, by introducing the universal transformation coefficient to transform the EASI system to other systems [72]. This study used 983 adults with 120 leads of ECG to achieve this and it improved the original EASI coefficients. These coefficients were later used by Rautaharju *et al.* in transforming EASI system to standard twelve-leads [73]. They classified acute myocardial ischaemia and old myocardial infarction for both cases and compared the results. Their results also showed that EASI derived ECG is a good alternative for standard twelve-leads. Wehr *et al.* also showed that the EASI derived twelve-lead ECG is as good as the standard twelve-leads for detecting myocardial ischaemia [74]. Sejersten *et al.* performed a comparative study for detecting acute ischaemia from EASI derived twelve-lead and standard twelve-lead [75]. They also used the same coefficients of Feild and showed that the use of EASI derived twelve-leads do not degrade the quality of signal and preserves vital information [75]. The assessment of the ST-segment deviation and cardiac rhythm for EASI derived twelve-leads was conducted by Chantad *et al.* [76]. Their results show that EASI derived system posses accurate and reliable

TH-1915_11610237

information for assessing the ST-segment deviation and cardiac rhythm in critical patients.

The accuracy of EASI derived twelve-leads in children was studied by Pahlm *et al.* and Welinder *et al.* [77, 78]. In these studies, it was observed that EASI derived system is a good alternative for standard twelve-lead ECG. A mobile transtelephonic system using derived twelve-lead ECG was proposed by Hadžievski *et al.* [79]. The method was a two step process where the first step learns the transform coefficients using a stationary calibration device and the second step generates the twelve-leads using a portable device. They used personalized linear regression to learn the coefficients to transform EASI system to twelve-lead ECG. The results showed good accuracy but was sensitive to the positional variations of the portable device.

The concept of eigenleads was proposed by Finlay *et al.* in which they reconstructed a standard twelve-lead ECG and a multichannel ECG [80]. They recorded a 117-lead unipolar multichannel ECG and calculated the principal components of these leads. The first three principal components, termed as eigenleads was then used to generate 117-lead ECG as well as standard twelve-leads and the transform was learned using linear regression. The results showed comparable reconstruction accuracy while reconstructing the twelve-leads. They also showed that better accuracy was obtained for the precordial leads than the limb leads.

The standard twelve-lead ECG was derived from wireless differential leads by R. Trobec and I. Tomašić [81]. They used the key idea that for best derivation, both position and transformation matrices should be personalized [81]. They developed a custom ECG device with 35 electrodes to find the optimal lead positions with the help of various combinations of differential leads [81]. The results showed that for reliable reproduction of standard twelve-lead ECG, a small number of bipolar measurements from the proximal electrodes are sufficient [81].

2.2.2 Transformation from twelve-lead subset: Linear approach

Linear methods for transforming one lead system to another is discussed in 2.2.1. This method has a disadvantage that it requires simultaneous recording from two different lead system. This process will increase the cost as well as complexity of the system. A solution to this problem is to record a standard twelve-lead ECG and use a subset of it to derive the rest. This sub-section intends to discuss various linear approaches which are used to derive standard twelve-lead ECG

TH-1915_11610237

from its subset. Two types of twelve-lead configurations, Mason-Likar twelve-leads and standard twelve-leads are discussed.

The Mason-Likar system was introduced in 1966 by Mason and Likar for recording exercise ECG. Later, studies showed that ECG recorded by the M-L system is different from standard twelve-lead ECG and it differs in diagnostic interpretation [82, 83]. Thus, methods to derive standard twelve-leads from the Mason-Likar system came into the picture. Reconstruction of twelve-lead ECG from reduced lead-sets was published by Nelwan *et al.* [45, 46]. They used linear regression for obtaining derived leads from a subset of the set of leads $\{Lead\ I, Lead\ II, V_1\text{-}V_6\}$ with $Lead\ I$, $Lead\ II$ and any precordial lead are always selected [45, 46]. The experiment was conducted for patient-specific and general reconstruction and the result showed that general reconstruction allows accurate reconstruction of one or two leads whereas patient-specific reconstruction allows three to four precordial leads [46]. The best three leads used for prediction were $Lead\ I$, $Lead\ II$ and V_2 and the best four leads were $Lead\ I$, $Lead\ II$, V_2 and V_5. Five lead subset was also used which gave better reconstruction accuracy [45, 46]. A similar attempt to derive the standard twelve-leads from the M-L system was presented by Wei [84, 85] where he used $Lead\ I$, $Lead\ II$, V_1 and V_6. He used least-squares method to learn the model and the transform was universal. Drew *et al.* conducted a comparative study of reduced lead-set ECG with standard ECG for diagnosis of cardiac arrhythmias and myocardial ischaemia [86]. They used a set of six electrodes (four limb and two precordial) to construct interpolated ECG. Precordial leads V_1 and V_5 were selected to derive the other precordial leads. Their statistical validation results show the best match for lead V_6 and poor match for lead V_3 [86]. The results show that interpolated ECG is comparable to standard ECG in diagnosing cardiac abnormalities. Two new systems using bipolar leads were introduced by Robertson *et al.* [87]. The bipolar leads in the first system are obtained by taking the potential difference between the M-L right arm electrode with leads V_2, V_4 and V_6. The potential differences between the M-L right arm electrode with leads V_1, V_2 and V_4 were used to obtain the second system. They used a universal linear transformation to derive the standard twelve-lead ECG from both these systems. The results indicated that the first system performed better than the second system in detecting ischaemia and myocardial infarction. Transforming the M-L System to standard twelve-lead using multiple linear transform was presented by Man *et al.* [88]. They learned an 8×8 matrix using multiple linear regression for achiev-

TH-1915_11610237

ing this. Unlike previous methods which used a subset of M-L leads, this method intends to transform the whole M-L leads to the standard twelve-lead ECG. Their study concluded that the diagnostic errors along with the distortion errors generated during the transformation of the M-L system can be reduced by using this approach.

Derivation of a twelve-lead ECG from a three-lead semi-orthogonal subset using the linear transformation arrays was proposed by Scherer and Nicklas [44]. This was based on the study that leads of the standard twelve-lead ECG are redundant and smaller subsets of electrodes can be used to construct the rest of the leads. The model was constructed using multiple linear regression for entire waveform and for the segmented waveforms (PR, QRS and ST-segments). Linear transformation coefficient was derived for each patient and was applied over $Lead\ I$, $Lead\ II$ and V_2 to obtain the derived leads. Verification was performed by taking the correlation between the actual and the derived leads. The method was patient-specific and it was evaluated only on two patients. Horáček *et al.* discussed the designing and testing of a transformation for deriving twelve-leads from reduced predictor leads [89]. The coefficients for lead transformation was obtained by regression analysis. They have used different combinations of dual precordial leads for deriving the twelve-lead ECG. Results showed several near equivalent choices of dual precordial leads.

Application of the derived ECG for a personalized remote health care application was proposed by Maheshwari *et al.* [90]. They transformed the Frank XYZ system and the reduced lead system to the standard twelve-lead system using least square fit approach. They studied the effect of using different precordial leads for reconstruction from reduced lead ECG. They also studied the variations in derived ECG for different cardiac abnormalities as well. Their results showed that the Frank XYZ system was more consistent in reconstructing the precordial leads than the reduced lead system, even though better results are observed for latter in case of twelve-lead reconstruction. Tsouri and Ostertag studied the use of independent components for deriving ECG from Frank leads and subset of twelve-leads [91]. They calculated independent components for three leads and used it to calculate a 12$\times$3 transformation matrix which was then used to derive the standard twelve-leads. For reduced lead-set, they used leads $Lead\ I$, $Lead\ II$ and V_2 and their results showed good correlation between the original and the reconstructed leads. A similar method was adopted by Maheshwari *et al.* using principal component analysis [8]. They calculated principal components of eight independent leads,

TH-1915_11610237

took the first three principal components and applied least-square fit to derive twelve-leads. Their results also showed that good reconstruction can be achieved using a subset of twelve-leads. Lee *et al.* used a state space model to derive standard twelve-lead ECG from its subset [92]. They used different combinations of signal averaged ECG (SAECG) leads consisting of $Lead\ I$, $Lead\ II$ and $Lead\ III$ as the source leads to derive the precordial leads using a state space model. Their results showed that the method is comparable with linear regression methods and posses good correlation between the original and the reconstructed leads.

Gregg *et al.* conducted an investigation about the failure of derived precordial leads [93]. They reconstructed the precordial leads using the combination of two leads consisting of (V_1, V_4) and (V_2, V_5). Compared to the standard twelve-leads, V_2-V_5 configuration showed weakness in the derivation of categories where V_1 is important such as bundle branch blocks. Similarly, V_1-V_4 configuration showed weakness in detection of anterior myocardial infarction. They concluded that the derived precordial leads are not equivalent to standard twelve-leads while diagnosing morphological abnormalities.

Derivation of orthogonal leads from twelve-lead ECG was performed by Guillem *et al.* [94]. They tested the accuracy for derivation of P-wave and compared it with a transform optimized for derivation of P-wave by least squares value (LSV) optimization method. The work was motivated by the fact that there were no optimized transforms for deriving P-waves. Individual transform was created for P-wave and QRScomplex. The work shows that the use of a newly optimized transform for P-wave increases the performance of Dowers transform.

Sejersten *et al.* compared EASI derived twelve-leads and M-L derived twelve-leads [95]. Their study showed that the M-L configuration has more resemblance to the standard ECG than that of the EASI-derived twelve-lead ECG. They also stated that derived ECG from both methods are not equivalent but diagnostic information obtained are similar and can be used as a substitute for ECG monitoring. Nelwan *et al.* also compared three derived ECG systems (EASI, patient-specific from subset of twelve-lead and universal from subset of twelve-lead) with standard twelve-lead ECG [47]. The results shows that patient-specific derived leads shows better performance than the other two. Maheshwari *et al.* compared the M-L system and the standard twelve-lead system which are derived from their respective subsets [96]. Their motive was to generate a derived ECG system that can be used for remote health care applications. They selected the predictor leads, $Lead\ I$, $Lead\ II$ and V_2

TH-1915_11610237

for both cases. Their findings concluded that these predictor leads outperform others in accurately reconstructing the precordial leads.

Challenges and directions for ECG lead reconstruction were published by Field *et al.* [97]. Challenges include selection of basis leads based on technical and nontechnical issues, database selection, optimized coefficient generation for different segments and evaluation of reconstructed ECG.

2.2.3 Transformation from twelve-lead subset: Non-linear approach

Existing non-linear approaches for deriving the standard twelve-lead ECG from its subset are discussed in this sub-section. The non-linear approach is superior compared to the linear methods in terms of reconstruction accuracy. But the cost in terms of time and complexity will increase which can be brought down by modern computing mechanisms.

Atoui *et al.* suggested an artificial neural network (ANN) model for deriving the twelve-lead ECG from its three-lead subset [98]. The model was a three layer ANN with the input layer having three inputs from leads $Lead\ I$, $Lead\ II$ and V_2 to the derive precordial leads. Their results showed better accuracies for the derived leads compared to the existing linear models. Later, they extended their study to compare universal as well as patient-specific cases [99]. The comparison of both the methods were performed with the linear regression based model and the results showed better performance. Also, better waveform similarity was observed for patient-specific models than the universal models.

The application of support vector regression (SVR) for deriving precordial leads was proposed by Yodjaiphet *et al.* [100]. They segmented the input signal consisting of $Lead\ I$, $Lead\ II$, V_1 and V_6 into four beats and selected twenty locations spread over the P, QRS and T waves. This was used as the input feature vector, and it ensured that important information is preserved in each beat. A radial basis function (RBF) kernel was used for the model and they claimed to achieve good reconstruction quality. A similar approach was followed by Kaewfoongrungsi and Hormdee [101]. They used leads $Lead\ I$, $Lead\ II$, $Lead\ III$, aVR, aVL,aVF and V_1 as input and derived other precordial leads. Two kernels, RBF and exponential radial basis function (ERBF) were used and the results showed that ERBF performs better than RBF in deriving the precordial leads.

TH-1915_11610237

2.2.4 Patient-specific models and global models

Various models are currently available, which uses linear or non-linear methods for enhancing the spatial resolution of ECG. These can either be patient-specific or global model of which the former is learned using data from a single patient while, the latter requires data from a group of patients. Among the models discussed in sub-sections 2.2.2 and 2.2.3, only few approaches are in the direction of global derived ECG systems. This includes approaches by Nelwan *et al.*, Robertson *et al.*, and Atoui *et al.* [45, 46, 87, 99]. The first two approaches used universal linear transforms while the latter used non-linear transform using ANN. In all these cases, the patient-specific models outperformed the global models.

2.3 Database, tools and performance evaluation of the models

In this section, a detailed description of the database, implementation platform, simulation tools and various performance evaluation methods used to asses the models are discussed.

2.3.1 Physikalisch-Technische Bundesanstalt (PTB) database

Physikalisch-Technische Bundesanstalt (PTB), which is the National Metrology Institute of Germany, has compiled a set of digitized ECGs, commonly known as the PTB diagnostic ECG database [19, 20]. This was collected from a total of 290 healthy and unhealthy subjects. The number of records per subject varies from one to five. A total of 549 records were obtained from 290 subjects with each record consisting of 15 simultaneously recorded lead signals (i.e., standard twelve-lead ECG and three Frank orthogonal leads). Each lead signal is digitized by sampling it at 1 kHz with 16-bit resolution. Clinical summary is missing for 22 subjects and the rest 268 subjects are categorized as: myocardial infarction - 148 subjects, cardiomyopathy/heart failure - 18 subjects, bundle branch block - 15 subjects, dysrhythmia - 14 subjects, myocardial hypertrophy - seven subjects, valvular heart disease - six subjects, myocarditis - four subjects, healthy control - 52 subjects and miscellaneous - four subjects. The models discussed in the following chapters are evaluated using the PTB database. Total duration of the signal varies across the records and hence for uniformity in evaluation, first thirty seconds of all records are selected.

TH-1915_11610237

2.3.2 Implementation platform and simulation tools

The models discussed in this thesis are implemented on a desktop computer which uses an Intel i5 processor with 8 gb RAM. The proposed models discussed in Chapter 3 and 4, and the models discussed in the literature are simulated in Matlab. Various Matlab toolboxes are used for simulating the models and will be discussed with the relevant chapters. The proposed model discussed in Chapter 5 is simulated in keras using the tensorflow backend [102, 103].

2.3.3 Performance evaluation methods

The performance of the models can be assessed by evaluating the quality of the output generated. In the case of derived ECG systems, original lead signals ($\boldsymbol{l}$) are compared with derived lead signals ($\hat{\boldsymbol{l}}$) for evaluating their performances. Various quality assessment measures are available which can be divided into subjective measures and objective measures. The subjective measures are computed using visual perception, usually by an expert person. The expert person, usually a cardiologist evaluates the record and a score is assigned based on his interpretation. A mean opinion score is obtained by taking the average of the scores from different experts. The objective measures are calculated using standard mathematical formulations which can quantify the quality of the signal. Objective measures can be divided into non-diagnostic measures and diagnostic measures. The objective measures used for assessing the performance of the models are discussed in this sub-section.

2.3.3.1 Non-diagnostic distortion measures

One of the most popular and frequently used distortion measure is the Pearson's correlation coefficient (r_x) and is given in equation 2.2. The value of correlation coefficient varies between -1 and 1 with -1 indicating strong negative correlation and 1 indicating a strong positive correlation.

$$r_x = \frac{\sum_{k=1}^{K}[\boldsymbol{l}(k) - \mu_o]\sum_{k=1}^{K}[\hat{\boldsymbol{l}}(k) - \mu_r]}{\sqrt{\sum_{k=1}^{K}[\boldsymbol{l}(k) - \mu_o]^2}\sqrt{\sum_{k=1}^{K}[\hat{\boldsymbol{l}}(k) - \mu_r]^2}} \ . \tag{2.2}$$

In equation 2.2, K denotes the total number of samples and μ_o and μ_r denotes the mean value of the original and the reconstructed lead signal respectively. Another popular similarity measure is the

TH-1915_11610237

root mean square error ($RMSE$), which is defined as

$$RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{l}(k) - \hat{\boldsymbol{l}}(k)]^2} \ . \tag{2.3}$$

where K denotes the number of samples. The goodness-of-fit of the model can be evaluated using the R^2 statistics which is defined as

$$R^2 = 1 - \frac{\sum_{k=1}^{K}[\hat{\boldsymbol{l}}(k) - \boldsymbol{l}(k)]^2}{\sum_{k=1}^{K}[\boldsymbol{l}(k)]^2} \times 100 \ . \tag{2.4}$$

2.3.3.2 Diagnostic distortion measures

As the name indicates, diagnostic distortion measures are used to evaluate the capability of the model in preserving clinically significant information in the lead signals while reconstructing. The weighted diagnostic distortion (WDD) measure is one such measure that computes the diagnostic similarity [18]. The WDD given in equation 2.5 calculates the distortion between different features of ECG. These features are based on duration, amplitude and shape of various ECG morphologies.

$$WDD(\boldsymbol{\beta}, \hat{\boldsymbol{\beta}}) = \boldsymbol{\Delta\beta^T}.\frac{\boldsymbol{\Lambda}}{\boldsymbol{tr}\,[\boldsymbol{\Lambda}]}.\boldsymbol{\Delta\beta} \times 100 \ . \tag{2.5}$$

where $\boldsymbol{\Delta\beta} = \begin{bmatrix}\Delta\beta_1 & \Delta\beta_2 & \cdots & \Delta\beta_u\end{bmatrix}^T$ is the normalized difference vector for u number of features. The weight matrix $\boldsymbol{\Lambda}$ is a diagonal matrix which determines the weights for different features.

Wavelet energy based diagnostic distortion ($WEDD$) is another diagnostic distortion measure [104] defined as:

$$WEDD = \sum_{j=1}^{D+1} w_j.WPRD_j \ . \tag{2.6}$$

where D is the number of wavelet decomposition level. The weight w_j is defined as

$$w_j = \frac{\sum_{k=1}^{K_j} \boldsymbol{c_j^2}(k)}{\sum_{m=1}^{D+1}\sum_{k=1}^{K_j} \boldsymbol{c_m^2}(k)}, \quad j = 1, 2, ..(D+1) \ . \tag{2.7}$$

where K_j denotes the number of coefficients in the j^{th} wavelet sub-band and D denotes the number of wavelet decomposition levels. The $WPRD_j$ is the wavelet percentage root mean square difference

TH-1915_11610237

of the j^{th} sub-band and is defined as

$$WPRD_j = \sqrt{\frac{\sum_{k=1}^{K_j}[\boldsymbol{c_j}(k) - \hat{\boldsymbol{c}}_{\boldsymbol{j}}(k)]^2}{\sum_{k=1}^{K_j}[\boldsymbol{c_j}(k)]^2}} \times 100 \ . \tag{2.8}$$

where K_j represents the number of coefficients in the j^{th} wavelet sub-band, $\boldsymbol{c_j}(k)$ and $\hat{\boldsymbol{c}}_{\boldsymbol{j}}(k)$ represents the coefficients in j^{th} wavelet sub-band of the original and the derived lead signal respectively.

In addition to the above diagnostic measures, the capability of models in capturing diagnostic information is evaluated using ECG diagnosability. This is performed by assessing the quality of useful information that is preserved in the derived lead signals. The detailed description of ECG diagnosability is presented in chapter 3.

2.4 Motivation for the present work

A lot of emphasis has been given recently on derived ECG systems due to the increasing demand of personalized healthcare applications [8]. A reduction in the number of leads can be utilized for data compression and transmission in remote health care applications. Patient comfort is another factor that has to be taken care of during continuous monitoring. Also, acquisition of ECG is difficult in the case of ambulatory monitoring. This opens up the necessity of using derived ECG or spatially enhanced ECG system. Several models are available that can improve the spatial resolution from fewer leads. The major challenge is to derive the lead signals without losing significant diagnostic information. This can be achieved by using models that can exploit inter-lead and intra-lead correlations between the standard twelve-leads.

The inter-lead correlation of a twelve-lead ECG in wavelet sub-bands is higher compared to that in time domain [9]. This high correlation between different leads in the same scale can be exploited to learn a linear model. Also, the predictor leads used to derive other lead signals may not serve as the best leads for prediction. A patient-specific model that can exploit the correlation in the wavelet domain using the best predictor leads can capture significant diagnostic information. This has motivated in developing the lead selective multi-scale linear model. This model not only exploits the frequency domain characteristics of the ECG signal, but also take advantage of the patient-specific approach.

TH-1915_11610237

Models that transform ECG signal into another domain for learning are discussed in the literature. These transforms use fixed dictionaries for transformation and may not always represent the morphological features of ECG. Also, the morphological features can vary from patient to patient. A solution is to use a personalized dictionary that can adapt to the characteristics of ECG signal. Such an approach could benefit the patient-specific derived ECG systems in improving the overall quality of the signal. These findings have motivated in using the joint dictionary and the sparse coding approach which was successfully applied in image processing, bandwidth extension, etc. [105, 106, 107, 108]. Such an approach has not been explored for ECG signals and has many applications in personalized health care, compressive sensing, telemonitoring, etc. [109, 110, 111, 112].

The derived ECG approaches discussed in literature mainly exploited the spatial or inter-lead correlations. None of these models are capable of simultaneously exploiting spatial and temporal correlations. A model capable of simultaneously learning the spatio-temporal correlation could represent the intra-lead and the inter-lead correlations in a better way. Such a model can improve the reconstruction accuracy thereby preserving the diagnostic information in lead signals. This has motivated to utilize the learning capabilities of the recurrent neural network (RNN) in creating a derived ECG model. Among various models discussed in literature, the ANN based model outperforms all other models by exploiting the spatial correlation. The advantage of using RNN is that it can use the learning capabilities of ANN and could further improve it by representing the spatio-temporal correlations.

2.5 Plan of the thesis

This thesis addresses three approaches for utilizing the spatial and temporal correlations of a twelve-lead ECG in the transformed domain for enhancing its spatial resolution. The first and the second approaches investigate the spatial correlation in wavelet domain and sparse domain respectively. Simultaneous utilization of spatial and temporal correlations are investigated in the third approach. In short, the proposed investigations in this thesis are planned as:

- To develop a model that can exploit the spatial correlations of a twelve-lead ECG in the wavelet domain. The inter-lead correlation between the predictor and response lead signals in the wavelet sub-band is found to be higher than that in the time-domain. This high correlation

TH-1915_11610237

can be exploited to learn linear models for the same sub-bands of different lead signals. Also, selecting the best predictor lead can improve the diagnostic quality of the derived leads.

- To explore the spatial relationship between different leads of twelve-lead ECG in the sparse domain. The clinically significant features and the spatial correlation between the lead signals of ECG can be represented in a better way by using a personalized dictionary. This approach is expected to preserve the clinically significant features in the reconstructed signal. The reconstruction quality of the derived ECG is expected to improve further by segmenting the signal and learning a personalized dictionary for each segment. This approach could segregate the clinical features in ECG, thereby improving the representation of these features in the dictionary.
- To examine the capability of different RNN models to simultaneously represent intra-lead and inter-lead correlations. The RNN is capable of combining the present and previous information in determining the current output. It is expected that a proper representation of the spatio-temporal correlations in ECG while learning the model can improve the diagnostic quality of the derived signal. The mapping of the predictor and the response lead signal is performed in high dimensional hidden feature space. Three different RNN models, which are simple RNN, LSTM and GRU, are studied and their capability for preserving the diagnostic quality in the derived ECG are evaluated using the standard and diagnostic distortion measures.

TH-1915_11610237

3

Lead Selective Multi-scale Linear Model

Contents

TH-1915_11610237

The concept of spatially enhanced ECG, its applications and different models used to derive the lead signals are discussed in chapters 1 and 2. It is observed that the models discussed in chapter 2 did not explore the possibility of segmenting the signal or exploiting frequency domain characteristics of ECG. Also, the performance of these models are determined by distortion measures that do not quantify the variations in diagnostic information. Most importantly, in all these models, it is observed that the precordial lead used for prediction is fixed for all patients. The quality of a derived ECG system depends upon the quality of the lead signals used for prediction. Hence, the positioning of electrodes is extremely important since the diagnosis depends upon the quality of ECG. But, the positioning of precordial electrodes can vary from patient to patient for each recording and hence could affect the quality of ECG. Even though this variation is less, a fixed predictor lead-set may not serve as the best lead-set for deriving ECG. Thus, a patient-specific model which exploits frequency domain characteristics of ECG signal along with the best predictor lead-set can improve the diagnosability of derived leads. In this chapter, a patient-specific multi-scale linear model that make use of the optimal predictor lead-set for enhancing the spatial resolution is explored.

The inter-lead correlation of a twelve-lead electrocardiogram in wavelet sub-band is higher compared to that of the original signals in time domain [9]. This high correlation between different leads in the same scale can be exploited to learn linear models for each sub-band. The optimal predictor lead-set for a given patient is selected using a lead selection algorithm which is based on a new diagnostic similarity score. The optimal predictor leads can personalize the model, thereby preserving diagnostic features in ECG. Performance of the proposed model is evaluated by standard distortion measures like correlation coefficient, $RMSE$ and R^2 statistics. The loss of diagnostic information is quantified using $WEDD$ measure and WDD measure [113, 18, 104]. The repeatability and diagnosability of the model are also evaluated and is compared with the existing linear models.

This chapter is organized as follows: The multi-scale linear model for enhancing the spatial resolution of ECG and a new approach for selecting optimal predictor lead are presented in section 3.1 and 3.2 respectively. The proposed approach for generating spatially enhanced ECG is described in 3.3. The procedure for computing the ECG diagnosability of the derived ECG model is defined in section 3.4. In section 3.5, results of the proposed method and its analysis are discussed, and in section 3.6, the ECG diagnosability is evaluated. The performance of the proposed model is compared

TH-1915_11610237

with that of the existing models in 3.7 and the chapter is summarized in section 3.8.

3.1 Multi-scale linear regression

ECG signal contains many diagnostic features of which the major ones are the P-wave, the QRS-complex and the T-wave. These features differ among themselves in terms of frequency, shape and amplitude and hence can be localized [2]. This localization can be performed by decomposing the ECG signal into various frequency scales or sub-bands using the discrete wavelet transform (DWT) [9][114]. Since the features are localized, a high inter-lead correlation can be achieved between the sub-bands of the predictor leads and the response lead [9][114]. It is reported in [9] that there is good correlation between the sub-band coefficients of different leads for a given low-frequency sub-band. To demonstrate this, in Fig. 3.1, the signal is recreated by using the low-frequency sub-bands for $Lead\ I$, $Lead\ II$, V_1, V_2 and V_6 of healthy record 's0460_rem'. A seven level wavelet decomposition is performed on these lead signals and is recreated using the inverse wavelet transform by making all but one sub-band coefficients zero. It is evident from Fig. 3.1 that a good correlation can be achieved between the sub-band coefficients of different leads for a given low-frequency sub-band. For e.g., the correlation between V_1 and V_2 in approximation sub-band $\boldsymbol{ac_7}$ is 0.96 and that in time domain is 0.94. Similarly, the correlation between $Lead\ II$ and V_6 in time domain is 0.96 and that in $\boldsymbol{dc_6}$ is 0.99. This high correlation between different leads in the same scale can be exploited to learn a linear model [114, 115, 116]. Thus, by learning a linear model between the same frequency scales of the predictor leads and the response lead, these localized features can be represented in a better way. The sub-band coefficients of the response lead can be estimated using a linear transformation which can be defined as

$$\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\boldsymbol{l_i}) = \boldsymbol{F_j}\left(\boldsymbol{c_j}(\boldsymbol{l_1}), \boldsymbol{c_j}(\boldsymbol{l_2}), \ldots, \boldsymbol{c_j}(\boldsymbol{l_k})\right) \quad . \tag{3.1}$$

The linear transformation $\boldsymbol{F_j}$ corresponds to j^{th} frequency sub-band. The leads $\boldsymbol{l_i}$'s and $\boldsymbol{l_k}$'s represent response lead and predictor leads respectively, where $i, k \in \{1, 2, , \ldots, 12\}$ and $\{i\} \neq \{k\}$. The j^{th} sub-band coefficients of i^{th} lead is denoted as $\boldsymbol{c_j}(l_i)$. The estimate of i^{th} response lead for j^{th} sub-

TH-1915_11610237

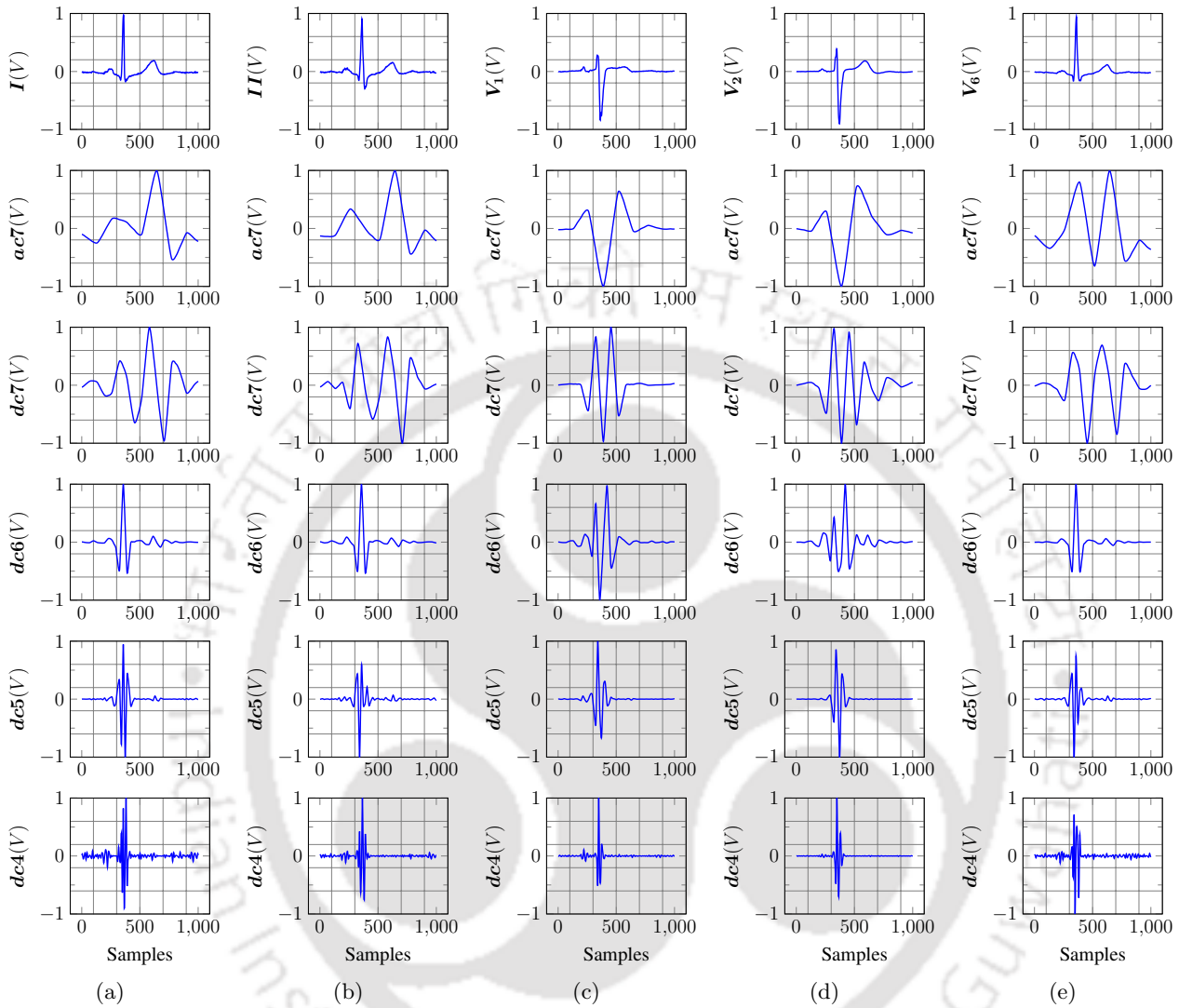


Figure 3.1: Correlation between sub-bands of different leads. First row shows the signal in time-domain and the other rows show the signal recreated using the low-frequency sub-bands for a seven level decomposition in the case of Healthy record 's0460_rem' is shown. (a) *Lead I*; (b) *Lead II*; (c) V_1; (d) V_2; (e) V_6.

band coefficients is given by $\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\boldsymbol{l}_{\boldsymbol{i}})$. The model defined in equation 3.1 can be represented as

$$\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\boldsymbol{l}_{\boldsymbol{i}}) = b_0 + \sum_{q=1}^{k} b_q . \boldsymbol{c}_{\boldsymbol{j}}(\boldsymbol{l}_{\boldsymbol{q}}) \quad . \tag{3.2}$$

where the regression coefficients are represented as b_q's. The precordial leads are indicated as $\boldsymbol{l}_{\boldsymbol{i}} = V_{i-6}$ with $i \in \{7, \ldots, 12\}$. The squared error between $\boldsymbol{c}_{\boldsymbol{j}}(\boldsymbol{l}_{\boldsymbol{i}})$ and $\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\boldsymbol{l}_{\boldsymbol{i}})$ is minimized using least square minimization technique to learn the model. The cost function corresponding to j^{th} sub-band

is defined as

$$\boldsymbol{J}(b) = \left(\boldsymbol{c_j}(\boldsymbol{l_i}) - \left(b_0 + \sum_{q=1}^{k} b_q . \boldsymbol{c_j}(\boldsymbol{l_q}) \right) \right)^2 \quad . \tag{3.3}$$

In 3.3, a subset of the eight independent leads consisting of $\boldsymbol{l_1} = Lead\ I$ and $\boldsymbol{l_2} = Lead\ II$ along with any precordial lead, $\boldsymbol{l_i} = V_{i-6}$, is used as the predictor lead-set. The frontal plane leads $Lead\ I$ and $Lead\ II$ are always preferred since the redundant leads $Lead\ III$, aVR, aVL and aVF can be perfectly recreated. The precordial lead V_{i-6} for prediction is selected in the lead selection module and is described in 3.2. The problem of deriving the leads using a predictor lead set can be posed as a linear regression case and the regression coefficients, b_q's can be learned using the least square minimization technique. The model defined in equation 3.2 can be represented in matrix form as

$$\boldsymbol{c_{rj}} = \boldsymbol{C_{pj}} \boldsymbol{b_j} \quad . \tag{3.4}$$

where $\boldsymbol{C_{pj}} = \begin{bmatrix} \boldsymbol{1}^T & \boldsymbol{c_j}(\boldsymbol{l_1}) & \boldsymbol{c_j}(\boldsymbol{l_2}) & \boldsymbol{c_j}(\boldsymbol{l_3}) \end{bmatrix}$ is an $K_j \times 4$ matrix containing j^{th} sub-band coefficients of the predictor leads. The vector $\boldsymbol{1}^T$ is a vector of ones with length K_j which is the same length as that of $\boldsymbol{c_j}(\cdot)$ sub-band. The j^{th} sub-band coefficients of response lead, $\boldsymbol{c_{rj}} = \boldsymbol{c_j}(\boldsymbol{l_i})$, is a $K_j \times 1$ matrix. The regression coefficients, $\boldsymbol{b_j} = \begin{bmatrix} b_0 & b_1 & b_2 & b_3 \end{bmatrix}^T$ is a 4×1 matrix corresponding to the j^{th} sub-band. The regression coefficients can be obtained by solving for $\boldsymbol{b_j}$ in equation 3.4 and is given by

$$\boldsymbol{b_j} = \left(\boldsymbol{C_{pj}^T} \boldsymbol{C_{pj}} \right)^{-1} \boldsymbol{C_{pj}^T} \boldsymbol{c_{rj}} \quad . \tag{3.5}$$

These regression coefficients can be used to linearly combine the sub-bands of the predictor leads to produce the derived lead. A total of six models corresponding to each precordial lead are learned using the method discussed above. Also, the required number of decomposition levels are determined from the sampling frequency using the method discussed in [117]. It is given as $level = [log_2(F_s - 2.96)]$ where $level$ is the decomposition level and F_s is the sampling frequency.

3.2 Lead selection using diagnostic similarity score

A diagnostic similarity score (DSS) is defined to select the best precordial lead to be used along with $Lead\ I$ and $Lead\ II$, and is given in equation 3.6. The diagnostic similarity score is obtained by

TH-1915_11610237

combining various diagnostic similarity measures (DSM).

$$\text{DSS}_i = d_1.\text{DSM}_1 + d_2.\text{DSM}_2 + \cdots + d_n.\text{DSM}_n \ . \tag{3.6}$$

The DSM's indicate different diagnostic similarity measures for the precordial predictor lead 'i'. The total number of DSM's used to calculate DSS is given by n. Since there are n DSM's, $2^n - 1$ different DSM combinations are possible to generate $2^n - 1$ DSS values for a given predictor lead 'i'. The weights d_n takes the value '1'or '0' to select or reject the n^{th} DSM. The number of DSM's, $n = 3$ and hence equation 3.6 can be modified as

$$\text{DSS}_i = d_1.\text{DSM}_1 + d_2.\text{DSM}_2 + d_3.\text{DSM}_3 \ . \tag{3.7}$$

The d_n values for combining the three DSM's can be given as $\begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix}$ = $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$, ..., $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. This will generate $2^3 - 1$ DSS values for a given predictor lead 'i'. The DSM's used to the compute DSS are defined in equations 3.8, 3.11 and 3.12. These DSM's are derived by modifying various diagnostic distortion measures as well as some standard measures.

$$\text{DSM}_1 = \frac{1}{\sum_{j=1}^{D+1} w_j.WPRD_j} \ . \tag{3.8}$$

The first diagnostic similarity measure, DSM_1, is obtained by modifying wavelet energy based diagnostic distortion ($WEDD$) measure described in [104]. In equation 3.8, w_j is the weight for j^{th} sub-band and is defined in equation 3.9. The wavelet percentage root mean square difference for j^{th} sub-band, $WPRD_j$, is defined in equation 3.10. Higher values of DSM_1 indicates a better waveform similarity between the original and the derived leads.

$$w_j = \frac{\sum_{k=1}^{K_j} \boldsymbol{c_j}^2(k)}{\sum_{m=1}^{D+1}\sum_{k=1}^{K_j} \boldsymbol{c_m}^2(k)}, \quad j = 1, 2, ..(D+1) \ . \tag{3.9}$$

$$WPRD_j = \sqrt{\frac{\sum_{k=1}^{K_j}[\boldsymbol{c_j}(k) - \hat{\boldsymbol{c}}_{\boldsymbol{j}}(k)]^2}{\sum_{k=1}^{K_j}[\boldsymbol{c_j}(k)]^2}} \ . \tag{3.10}$$

where, K_j is the number of wavelet coefficients in j^{th} sub-band. The k^{th} wavelet coefficient in j^{th} sub-band for the original and the estimated bands are given by $\boldsymbol{c_j}(k)$ and $\hat{\boldsymbol{c}}_{\boldsymbol{j}}(k)$ respectively [104]. Bi-orthogonal 9/7 wavelet filters are used to decompose the pre-processed ECG signal to seven levels. Bi-orthogonal 9/7 wavelet filters are selected as it can represent the ECG signal in a better way [118].

TH-1915_11610237

The second DSM, i.e. DSM_2, is the correlation coefficient and is given in equation 3.11. The higher values of DSM_2 means better similarity between the original and the derived leads.

$$\text{DSM}_2 = \frac{\sum_{k=1}^{K}[\boldsymbol{l}(k) - \mu_o]\sum_{k=1}^{K}[\hat{\boldsymbol{l}}(k) - \mu_r]}{\sqrt{\sum_{k=1}^{K}[\boldsymbol{l}(k)]^2}\sqrt{\sum_{k=1}^{K}[\hat{\boldsymbol{l}}(k)]^2}} \quad . \tag{3.11}$$

The third DSM, i.e. DSM_3, is nothing but the R^2 statistics and is given in equation 3.12. In this case also, the higher value implies a good similarity between original and derived leads.

$$\text{DSM}_3 = 1 - \frac{\sum_{k=1}^{K}[\hat{\boldsymbol{l}}(k) - \boldsymbol{l}(k)]^2}{\sum_{k=1}^{K}[\boldsymbol{l}(k)]^2} \quad . \tag{3.12}$$

In equation 3.11 and equation 3.12, $\boldsymbol{l}(k)$ represents the original lead, $\hat{\boldsymbol{l}}(k)$ represents the derived lead and k denotes the k^{th} sample. The total number of samples in the signal is denoted by K. The mean values of the original lead and the derived lead are represented as μ_o and μ_r respectively. In this experiment, since the mean is removed during pre-processing, the values of μ_o and μ_r are zero. Thus, a higher value of DSS indicate a better similarity between the original and the derived lead.

The patient-specific lead selection algorithm calculates the DSS of each model using equation 3.7. Based on the DSS value, the algorithm selects the best model and the best precordial lead. A precordial lead is fixed as the predictor lead and by using the corresponding model, the rest of the precordial leads are predicted. For example, V_1 is fixed as the predictor lead and the model corresponding to it is used to derive the other five precordial leads, i.e. V_2 to V_6. Various Diagnostic Similarity Measures (DSM's) defined in equation 3.8, equation 3.11 and equation 3.12 are calculated for these derived precordial leads. The average value of each DSM for the derived leads is then computed. The process is repeated until all six models corresponding to the six precordial leads are used. Thus, for a given predictor precordial lead and the corresponding model, three average DSM values are obtained. Now, for this model, different combinations of the average DSM values are used to compute the DSS. Seven combinations of DSM's ($2^n - 1$ with $n = 3$) are possible and hence seven DSS values are obtained for the given model. This is repeated for other models as well. Now, for a given combination of DSM, the model which gives the maximum DSS value is selected. This process is repeated for all combinations of DSM's and the most frequent model is selected as the best model. The precordial lead corresponding to the selected model is the best precordial lead for enhancing the spatial resolution for a given patient. The following sub-section gives a detailed description about the

TH-1915_11610237

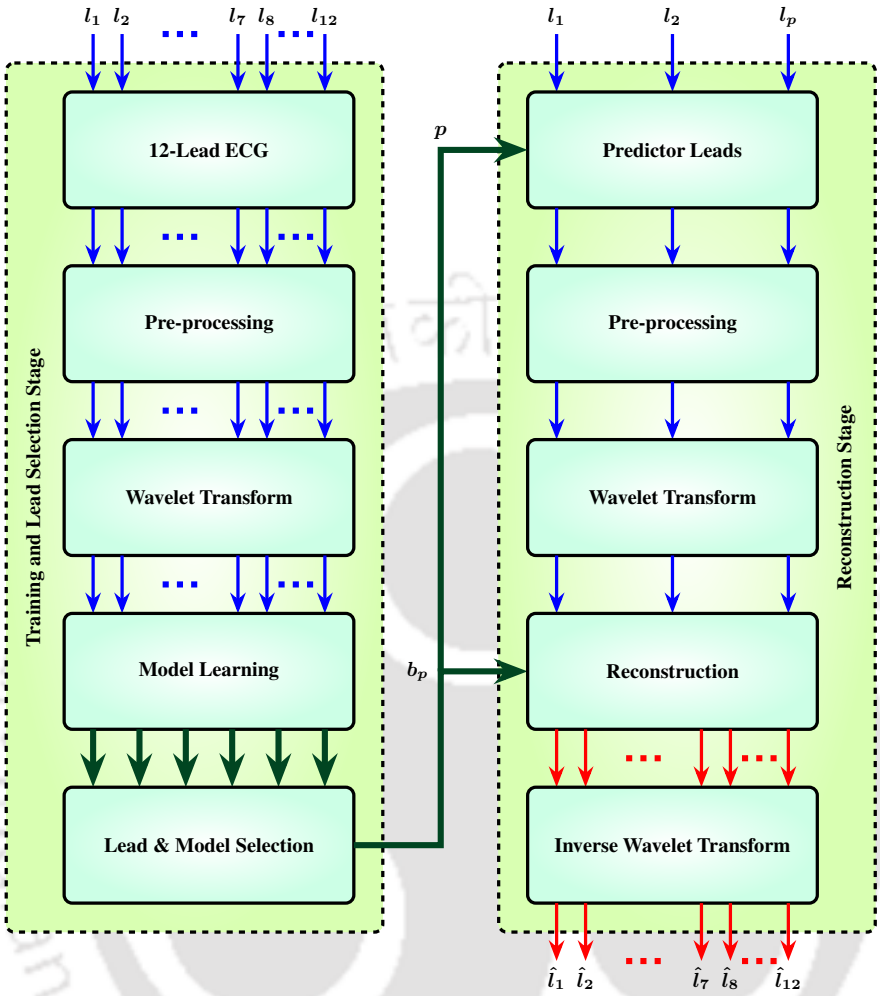


Figure 3.2: Block diagram representation of the proposed model.

proposed DSS based lead selective multi-scale linear regression algorithm.

3.3 Proposed DSS based lead selective multi-scale linear regression

The proposed DSS based lead selective multi-scale linear regression consists of the learning module, the DSS based lead selection module and the reconstruction module. Block diagram representation of the proposed approach is shown in Fig. 3.2. The twelve-lead ECG is represented as $\boldsymbol{l_1, l_2, \ldots, l_{12}}$ and the input signal is pre-processed before passing into any of these modules. The baseline wandering effect is removed from all leads using the following procedure. First, a low-pass filter is designed with a cut-off frequency of 0.67 Hz. It is assumed that there is no significant information below

Algorithm 1 Lead Selective Multi-scale Linear Regression Algorithm - Model Learning

$level \leftarrow$ wavelet decomposition level.
s_1 samples of $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_7}, \ldots, \boldsymbol{l_{12}}\}$.
Input: $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_7}, \ldots, \boldsymbol{l_{12}}\}$, $level$.
for $i \leftarrow 7$ to 12 **do**
  for $\forall\ k \neq i$, i,$k \in \{7, \ldots, 12\}$ **do**
    for $m \leftarrow 1, 2, i, k$ **do**
      $[\boldsymbol{c_1}(\boldsymbol{l_m}) \ \ldots\ \boldsymbol{c_{level+1}}(\boldsymbol{l_m})] \leftarrow \mathbf{wavTfm}(\boldsymbol{l_m}, level)$.
    end for
    for $j \leftarrow 1, \ldots, level + 1$ **do**
      $\boldsymbol{C_{pj}} \leftarrow [\mathbf{1}^T\ \boldsymbol{c_j}(\boldsymbol{l_1})\ \boldsymbol{c_j}(\boldsymbol{l_2})\ \boldsymbol{c_j}(\boldsymbol{l_i})]$.
      $\boldsymbol{c_{rj}} \leftarrow [\boldsymbol{c_j}(\boldsymbol{l_k})]$.
      $\boldsymbol{b_i}(\boldsymbol{c_j}(\boldsymbol{l_k})) \leftarrow (\boldsymbol{C_{pj}^T C_{pj}})^{-1} \boldsymbol{C_{pj}^T c_{rj}}$.
    end for
  end for
end for
Output: $\boldsymbol{b_i}$, $\forall\ i \in \{7, \ldots, 12\}$

0.67 Hz in the signal and the frequency of the baseline wandering effect falls below this limit [119]. Each lead of the twelve-lead ECG is passed through the low-pass filter and is then subtracted from the original lead. The amplitude of each lead is then normalized and the mean is removed. This pre-processed electrocardiogram is divided into the learning set consisting of first s_1 samples of the twelve-leads, the lead selection set with the next s_2 samples and the reconstruction set containing the remaining s_3 samples of the twelve-leads. The detailed description of these three modules are as follows.

The learning module is summarized in Algorithm 1. The learning set consists of leads $\boldsymbol{l_1}, \boldsymbol{l_2}$ and $\boldsymbol{l_7}, \ldots, \boldsymbol{l_{12}}$, which are the eight independent leads. Each lead in the learning set consists of s_1 number of samples. The value of $level$ is used to set the number of decomposition level, and it remains same for the next two modules. In this experiment, bi-orthogonal 9/7 wavelet filters are used to decompose the mean free ECG signal to seven levels since it represents the ECG signal in a better way [118]. In the learning module, the eight independent leads along with the decomposition level is passed into the system. The leads $\boldsymbol{l_1}$, $\boldsymbol{l_2}$ and $\boldsymbol{l_i}$ are set as the predictor leads and $\boldsymbol{l_k}$ is set as the response lead s.t. $i, k \in \{7, \ldots, 12\}$ and $i \neq k$. The leads are then decomposed into approximation and detail sub-bands using a wavelet function ($\mathbf{wavTfm}$). This is stored in $\boldsymbol{c_j}(\boldsymbol{l_m})$ where $\boldsymbol{c_j}(\boldsymbol{l_m})$ represents j^{th} sub-band of m^{th} lead with $m \in \{1, 2, i, k\}$. Thus a total of $level$ + 1 sub-bands are obtained from these

TH-1915_11610237

Algorithm 2 Lead Selective Multi-scale Linear Regression Algorithm - Lead Selection using DSS

$level \leftarrow$ wavelet decomposition level.
$n \leftarrow$ number of DSM's.
s_2 samples of $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_7}, \ldots, \boldsymbol{l_{12}}\}$.
Input: $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_7}, \ldots, \boldsymbol{l_{12}}\}$, $level$, n and $\boldsymbol{b_i}$.
for $i \leftarrow 7$ to 12 **do**
  for $\forall\ k \neq i$, i,$k \in \{7, \ldots, 12\}$ **do**
    for $m \leftarrow 1, 2, i, k$ **do**
      $[\boldsymbol{c_1}(\boldsymbol{l_m}) \ \ldots\ \boldsymbol{c_{level+1}}(\boldsymbol{l_m})] \leftarrow \mathbf{wavTfm}(\boldsymbol{l_m}, level)$.
    end for
    for $j \leftarrow 1, \ldots, level + 1$ **do**
      $\boldsymbol{C_{pj}} \leftarrow [\mathbf{1}^T\ \boldsymbol{c_j}(\boldsymbol{l_1})\ \boldsymbol{c_j}(\boldsymbol{l_2})\ \boldsymbol{c_j}(\boldsymbol{l_i})]$.
      $\hat{\boldsymbol{c}}_{\boldsymbol{rj}} \leftarrow \boldsymbol{C_{pj}} \boldsymbol{b_i}(\boldsymbol{c_j}(\boldsymbol{l_k}))$.
      $[\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}})] \leftarrow \hat{\boldsymbol{c}}_{\boldsymbol{rj}}$.
    end for
    $\hat{\boldsymbol{l}}_{\boldsymbol{k}} \leftarrow \mathbf{invWavTfm}([\hat{\boldsymbol{c}}_{\mathbf{1}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}}) \ \ldots\ \hat{\boldsymbol{c}}_{\boldsymbol{level+1}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}})])$.
    $[DSM_k(1) \ \ldots\ DSM_k(n)] \leftarrow \mathbf{dsm}(\boldsymbol{l_k}, \hat{\boldsymbol{l}}_{\boldsymbol{k}})$.
  end for
  $[ADSM_i(1) \ \ldots\ ADSM_i(n)] \leftarrow \mathbf{mean}(\mathbf{DSM})$.
end for
for $w \leftarrow 1$ to $2^n - 1$ **do**
  $DSS(w) \leftarrow \mathbf{dss_w}(\mathbf{ADSM})$.
end for
$p \leftarrow \mathbf{freq}([DSS(1) \ \ldots\ DSS(2^n - 1)])$, $p \in \{7, \ldots, 12\}$.
Output: $\boldsymbol{l_p}$ and $\boldsymbol{b_p}$.

leads. The model is obtained by applying the least square minimization technique on the sub-bands according to equation 3.5. This model is stored in $\boldsymbol{b_i}(\boldsymbol{c_j}(\boldsymbol{l_k}))$ which represents i^{th} model of j^{th} sub-band for k^{th} lead. The same process is repeated in order to obtain six different models corresponding to each precordial lead. Each of these six models consists of $level$ + 1 sub-models corresponding to each sub-band.

The next module is the lead selection module which is summarized in Algorithm 2. The best possible predictor lead to generate the response leads is selected using the DSS. The lead selection set with s_2 samples of each lead is used. As described in the learning module, the wavelet function is used to decompose the predictor and the response leads. The next step is to compute the sub-band coefficients for response lead $\boldsymbol{l_k}$. The model $\boldsymbol{b_i}(\boldsymbol{c_j}(\boldsymbol{l_k}))$ is then selected to predict $\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}})$, i.e., j^{th} sub-band estimate of k^{th} response lead using i^{th} model. This is performed using equation 3.4 and is repeated until all sub-band estimates of k^{th} lead are computed. These estimated sub-bands

TH-1915_11610237

Algorithm 3 Lead Selective Multi-scale Linear Regression Algorithm - Signal Derivation

$level \leftarrow$ wavelet decomposition level.
s_3 samples of $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_p}\}$.
Input: $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_p}\}$, $level$ and $\boldsymbol{b_p}$.
for $\forall\ k \neq p$, $k, p \in \{7, \ldots, 12\}$ **do**
  for $m \leftarrow 1, 2, p$ **do**
    $[\boldsymbol{c_1}(\boldsymbol{l_m}) \ \ldots \ \boldsymbol{c_{level+1}}(\boldsymbol{l_m})] \leftarrow \mathbf{wavTfm}(\boldsymbol{l_m}, level)$.
  end for
  for $j \leftarrow 1, \ldots, level + 1$ **do**
    $\boldsymbol{C_{pj}} \leftarrow [\mathbf{1}^T\ \boldsymbol{c_j}(\boldsymbol{l_1})\ \boldsymbol{c_j}(\boldsymbol{l_2})\ \boldsymbol{c_j}(\boldsymbol{l_p})]$.
    $\hat{\boldsymbol{c}}_{\boldsymbol{rj}} \leftarrow \boldsymbol{C_{pj}} \boldsymbol{b_p}(\boldsymbol{c_j}(\boldsymbol{l_k}))$.
    $[\hat{\boldsymbol{c}}_{\boldsymbol{j}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}})] \leftarrow \hat{\boldsymbol{c}}_{\boldsymbol{rj}}$.
  end for
  $\hat{\boldsymbol{l}}_{\boldsymbol{k}} \leftarrow \mathbf{invWavTfm}([\hat{\boldsymbol{c}}_{\mathbf{1}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}}) \ \ldots \ \hat{\boldsymbol{c}}_{\boldsymbol{level+1}}(\hat{\boldsymbol{l}}_{\boldsymbol{k}})])$.
end for
Output: $\{\boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{l_p}, \hat{\boldsymbol{l}}_{\boldsymbol{k}}\}$, $k \in \{\{7, \ldots, 12\}$ - $p\}$.

are then passed through an inverse wavelet function (**invWavTfm**) which generates an estimate of the derived lead $\hat{\boldsymbol{l}}_{\boldsymbol{k}}$. The DSM's defined in equations 3.8, 3.11 and 3.12 are used to compute the diagnostic similarity between $\boldsymbol{l_k}$ and $\hat{\boldsymbol{l}}_{\boldsymbol{k}}$. Since three DSM's are used , the value of n is set as three. The function **dsm** computes the n DSM's between $\boldsymbol{l_k}$ and $\hat{\boldsymbol{l}}_{\boldsymbol{k}}$. These n DSM's for k^{th} lead is stored as a row matrix. The process is repeated for all $k \neq i$ leads and a $5 \times n$ matrix (**DSM**) corresponding to model i is formed by appending the row matrices. The function **mean** is used to calculate the column wise mean of the **DSM** matrix and is stored as row matrix. The process is repeated for all six models and by appending these row matrices, the average DSM (**ADSM**) matrix which is a $6 \times n$ matrix is formed. The **ADSM** matrix is then passed into the function $\mathbf{dss_w}$, which computes the DSS values and returns the selected lead for a particular combination of DSM's. Thus, a total of $2^3 - 1$ selected models are obtained and by using the function **freq**, the most frequent model is selected. Thus, the best model ($\boldsymbol{b_p}$) and corresponding predictor lead ($\boldsymbol{l_p}$) is selected for enhancing the spatial resolution.

The third module is the reconstruction module where the reconstruction dataset with s_3 samples of each lead is used. This is summarized in Algorithm 3. The precordial lead corresponding to the selected model, i.e. $\boldsymbol{l_p}$, is used along with $\boldsymbol{l_1}$ and $\boldsymbol{l_2}$ as the predictor leads. The predictor leads are then decomposed into approximation sub-band and detail sub-bands using the wavelet function. The model corresponding the selected lead $\boldsymbol{l_p}$, i.e. $\boldsymbol{b_p}$, is then used to derive the approximation and

TH-1915_11610237

detail sub-bands. These sub-bands are then passed into the inverse wavelet function to generate an estimate of the response lead, i.e. $\hat{l}_k$. This process is repeated until all the precordial leads are estimated, thereby enhancing the spatial resolution of a three lead ECG.

3.4 Diagnosability of the derived ECG model

A derived ECG model must be capable of preserving the diagnostic information in derived lead signals. This is normally evaluated using diagnostic and non-diagnostic distortion measures discussed in section 2.3. Another approach to evaluate the model is to classify the original and the derived lead signals of different classes and compare their classification accuracies. This classification can be performed with the help of cardiologists or using machine learning techniques. Classification using machine learning techniques is preferred due to the recent advancements in this field as well as practicality. The term diagnosability is defined for evaluating the capability of the model in preserving the diagnostic information in the derived lead signals. Diagnosability is defined as the ratio of classification accuracy of the derived ECG (Accuracy$_d$) to that of the original ECG (Accuracy$_o$). This is given as

$$\text{Diagnosability(\%)} = \frac{\text{Accuracy}_d}{\text{Accuracy}_o} \times 100 \ . \tag{3.13}$$

where Accuracy = (TP + FP)/(TP + TN + FP + FN), where TP is true positive, TN is true negative, FP is false positive and FN is false negative. By using diagnosability, an assessment regarding the quality of useful information that is preserved by derived lead signals can be performed. In this way the capability of the model in preserving the diagnostic information can be evaluated. Diagnosability is calculated between two classes and hence a support vector machine (SVM) based binary classifier is used to compute the classification accuracies [120]. The feature vector for classification is extracted from all twelve-leads by calculating the energy and the eigenvalue features over different frequency sub-bands [121]. The features for training the classifier are extracted from the original samples of the twelve-lead ECG. The testing is performed using the original and the reconstructed samples separately to obtain Accuracy$_o$ and Accuracy$_d$ respectively.

TH-1915_11610237

3.5 Results and discussions

The Physikalisch-Technische Bundesanstalt (PTB) diagnostic ECG database [19][20] is used in this study for evaluating the proposed model. The details of the database is given in section 2.3. The standard twelve-leads are pre-processed before conducting the experiment using the method discussed in section 3.3. A total of 30,000 samples are used of which first 10,000 samples are used in the training set s_1, next 7000 samples are used in the lead selection set s_2 and the rest of the samples i.e. 13,000 samples are used in the reconstruction set s_3.

3.5.1 Performance evaluation

Original precordial leads are compared with the derived leads for evaluating the performance. Commonly used performance evaluation measures such as correlation coefficient and $RMSE$ are used to compute the closeness of the original lead with the derived lead [104]. The goodness-of-fit of the model is evaluated using R^2 statistics. The diagnostic closeness of the derived lead with the original lead is measured using $WEDD$ measure and WDD measure [18, 104]. The details of these distortion measures are discussed in section 2.3. In addition to these measures, repeatability and ECG diagnosability of the model are also computed. Repeatability is evaluated by measuring the $WEDD$ between the original and the derived leads while keeping the same model for different instances. Diagnosability is evaluated by following the method described in section 3.4.

3.5.2 Results of lead selective multi-scale linear regression

For analyzing the results, the proposed model is assessed using the standard evaluation metrics described in 3.5.1. Experiments are conducted using different decomposition levels to evaluate the reconstruction quality and is shown in Fig. 3.3. It is evident from Fig. 3.3 that the minimum value of $WEDD$ is observed at $level$ = 7 which validates the method in [117]. It is also observed that increasing the decomposition level does not guarantee a good reconstruction quality. As the number of decomposition levels is increased, the low-frequency sub-bands fail to represent the diagnostic features, thereby degrading the quality.

The performance of the model for different performance evaluation metrics is shown in Table 3.1. The evaluation is shown only for six derived precordial leads. Distortion measures are calculated for

TH-1915_11610237

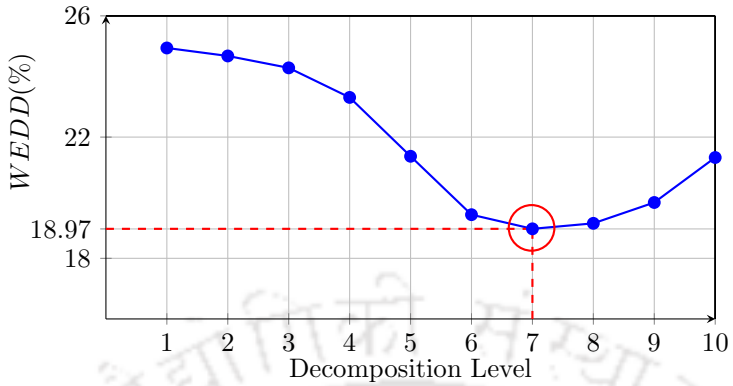


Figure 3.3: Variation in $WEDD(\%)$ to the change in decomposition level.

Table 3.1: Performance evaluation of the Proposed Model for all patient records.

	r_x	$RMSE$ (mV)	$WEDD$ $(\%)$	R^2 $(\%)$	WDD $(\%)$
V_1	0.97	70.97	20.96	93.00	13.55
V_2	0.98	61.05	16.90	95.26	9.83
V_3	0.98	61.35	16.78	95.42	8.99
V_4	0.98	64.42	18.75	94.36	9.80
V_5	0.97	64.76	20.75	93.51	10.48
V_6	0.98	60.70	19.70	93.82	9.41
Avg.	0.98	63.88	18.97	94.23	10.34

all 549 records except WDD. Only 518 records are evaluated for WDD due to difficulty in extracting the features automatically. The average values of different performance metrics given in Table 3.1 show a good similarity between the original and the derived leads. The correlation and R^2 statistics are high and $RMSE$, $WEDD$ and WDD values are low. The higher average correlation value of 0.98 and the lower average $RMSE$ value of 63.88 mV indicate that the spatially enhanced ECG has a very good similarity with the original signal. Similarly, the average R^2 value of 94.23% shows that the proposed model fits better for the evaluated set of data. The lower average values of 18.97% for $WEDD$ and 10.34% for WDD indicate that the derived leads preserve diagnostic information.

Considering the individual precordial leads, it can be observed from Table 3.1 that all precordial leads possess a very high correlation. It can also be observed that the lower values for $RMSE$ are given by V_2, V_3 and V_6. The R^2 values are higher in the case of leads V_2 and V_3. In the case of $WEDD$, leads V_2 and V_3 give lower values. From the values of WDD, it can be observed that the

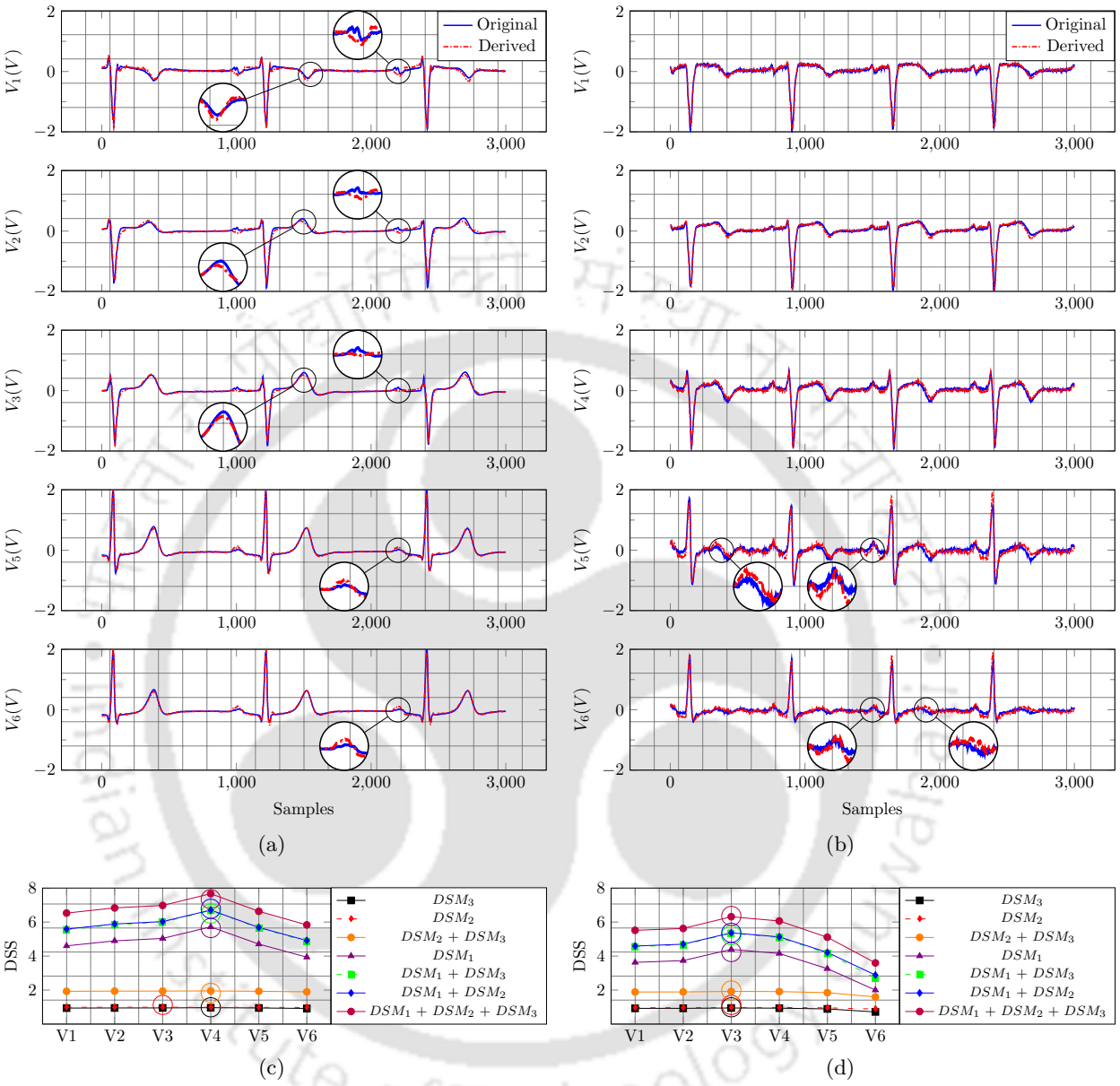


Figure 3.4: (a) Original (red) and reconstructed (blue) precordial leads in the case Healthy record 's0460_rem'; (b) Original (red) and reconstructed (blue) precordial leads in the case Anterior MI record 's0027lrem'; (c) DSS values for different combinations of DSM's for record 's0460_rem'; (d) DSS values for different combinations of DSM's for record 's0027lrem'.

derived leads fall in the good category [18]. It is also observed that, from 549 recordings of 290 patients, lead V_3 is the most frequently selected predictor precordial lead and lead V_6 is the least selected one. Lead V_3 is selected as a predictor lead for 256 records, followed by V_2 for 158 records, V_4 for 88 records, V_1 for 24 records, V_5 for 18 records and V_6 for 5 records.

TH-1915_11610237

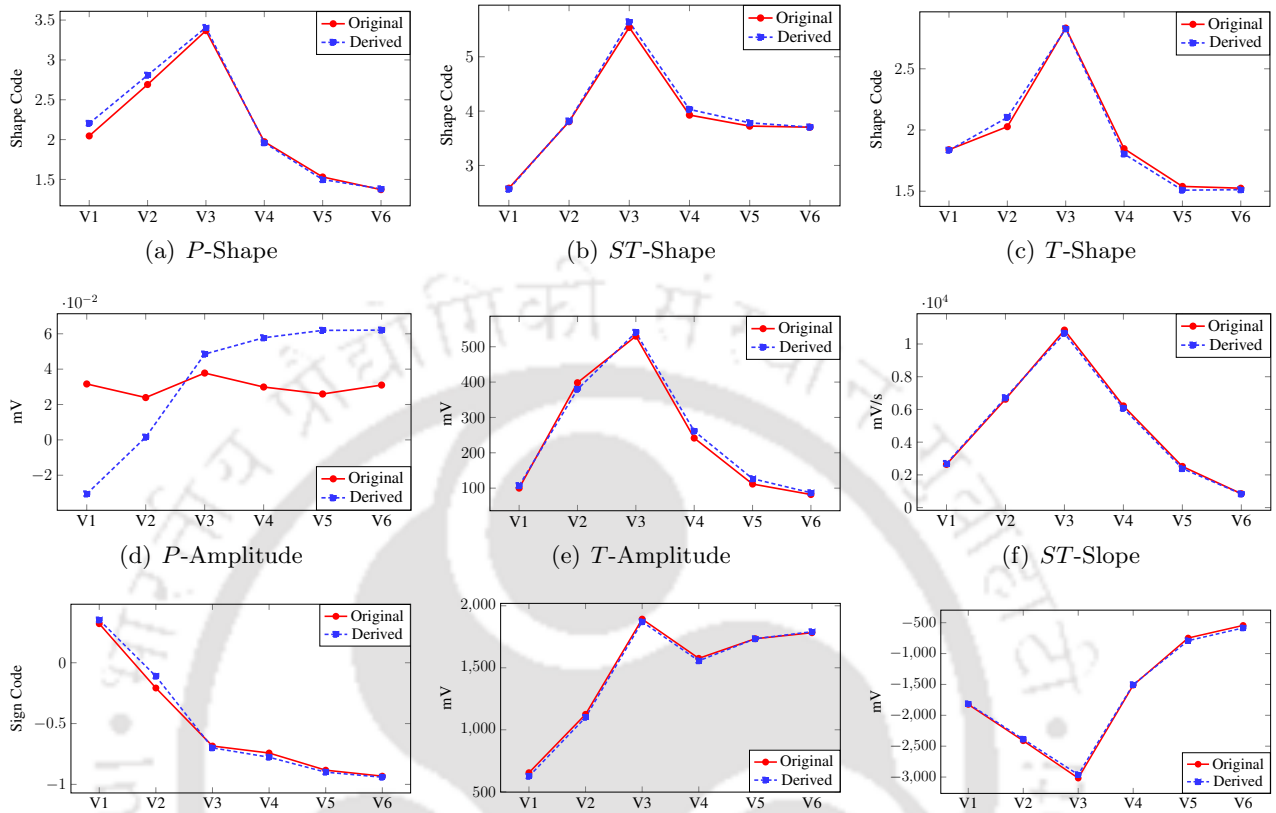
The Fig. 3.4(a) and 3.4(b) show the original leads and the derived leads of normal and anterior myocardial infarction (MI) data. The figures are selected based on the average value of $WEDD$ for these classes. Fig. 3.4(a) shows the original and the derived precordial leads of the patient record 's0460_rem' which is a healthy control data. In this case, lead V_4 is selected as the predictor precordial lead based on DSS values. The DSS values of precordial leads for different combinations of DSM's are shown in Fig. 3.4(c). The lead selected for each combination of DSM's is marked in the figure. It is evident from Fig. 3.4(c), that the most frequently selected lead is V_4. For this case, the average correlation is 0.99, $RMSE$ is 47.42 mV, $WEDD$ is 15.87% and R^2 statistics is 96.93%. From the figure, it can be observed that in lead V_1, V_2 and V_3, the amplitude of the T-wave is slightly distorted but preserves the shape. Also, the P-wave of the reconstructed signal does not follow the original signal in all leads. The QRS-complex in all leads are least affected.

The anterior MI record 's0027lrem' is shown in Fig. 3.4(b), with an average correlation value of 0.98, $RMSE$ value of 69.40 mV, $WEDD$ value of 19.75% and R^2 value of 94.50%. In this case, selected precordial lead for prediction is lead V_3. From the DSS values given in Fig. 3.4(d), it is evident that all cases selected the same lead. From Fig. 3.4(b), it can be noticed that for all these leads, the reconstructed signal preserves the shape feature. A slight deviation in the amplitude of these features is also observed for this case. In leads V_5 and V_6, P and T-waves show slight distortion. In V_4, V_5 and V_6 a variation of 5 to 10 ms is observed in the corrected QT-interval (calculated using Bazett's formula [122]). It is also evident from the figure that QRS-segment preserves shape and amplitude features.

In the above cases, it is evident that the P-wave is distorted in most of the leads. The P-wave is a low amplitude wave compared to the T-wave and the QRS-complex and the model fails to capture this variation. Insufficient atrial information in the predictor leads may be another cause for this distortion. The T-wave and QRS-complex preserves most of the information. In Fig.s 3.4(c) and 3.4(d), it can be observed that some combinations of DSM's give a very low DSS. This is because the values of DSM_2 and DSM_3 is less than one. Also, the value of DSM_1 is above one and hence higher values are observed.

For further quantification of distortion, an analysis at the feature level is conducted and is shown in Fig. 3.5. A total of nine features consisting of shape and amplitude features are analyzed. The

TH-1915_11610237



(a) P-Shape (b) ST-Shape (c) T-Shape

(d) P-Amplitude (e) T-Amplitude (f) ST-Slope

(g) QRS-Sign (1 or -1) (h) QRS-Positive Amplitude (i) QRS-Negative Amplitude

Figure 3.5: Comparison of shape and amplitude features between original leads and reconstructed leads.

average values of these nine features are calculated for each precordial lead. Only 518 records out of 549 records are used for analysis. Also, the duration features are excluded from the analysis assuming very low variations between original and derived leads. This can be quantified by observing the waveforms given in Fig. 3.4. The shape features used for analysis consists of P-shape, ST-shape and T-shape. While calculating WDD, the shape features are given a particular code according to its shape [18]. The average value of shape code is calculated for all leads. In the case of P-shape given in Fig. 3.5(a), a slight variation from the original shape is observed for V_1 and V_2. Derived leads V_3 to V_6 shows a good similarity with the original leads. The ST-shape given in Fig. 3.5(b) shows a very good similarity for leads V_1, V_2 and V_6. Slight variation is observed for leads V_3 and V_4. Leads V_2, V_4 and V_5 shows variation from original in the case of the T-shape given in Fig. 3.5(c).

Amplitude features consist of P-amplitude, T-amplitude, ST-slope, the sign of the first peak in the QRS-complex (1 or -1), QRS-positive amplitude and QRS-negative amplitude. These features are

TH-1915_11610237

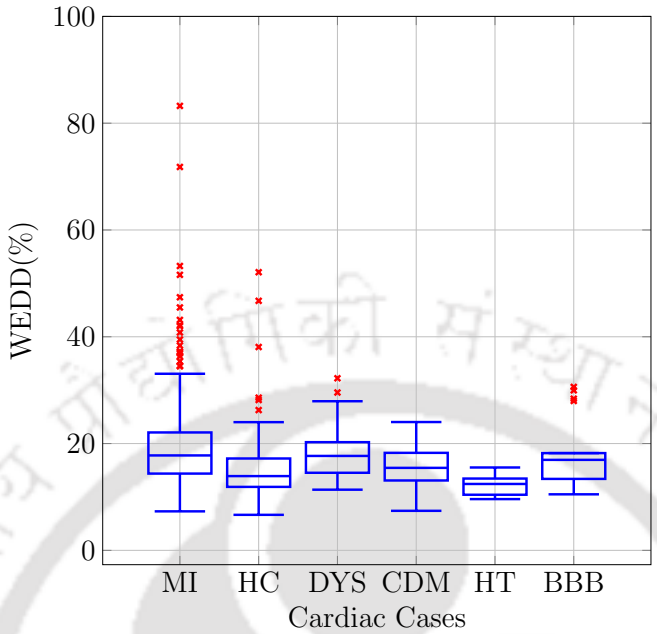


Figure 3.6: Box Plot showing the variation in $WEDD(\%)$ for the precordial leads of myocardial infarction (MI), healthy control (HC), dysrhythmia (DYS), cardiomyopathy (CDM), hypertrophy (HT) and bundle branch block (BBB).

given in Fig. 3.5(d) to 3.5(i). In Fig. 3.5(d), which shows the amplitude of the P-wave, a variation is observed for all precordial leads. Even though the variation is in the range of a few millivolts, the reconstructed lead signals fails to follow the original feature. This variation in the amplitude of the P-wave is also evident from the plots given in Fig. 3.4. The T-wave amplitude given in Fig. 3.5(e) shows the variation in the amplitude for leads V_4 and V_5 and slight variations in the amplitude for leads V_2 and V_3. The ST-slope feature given in Fig. 3.5(f) shows a very good similarity between the original and the derived leads. A variation in the sign of the first peak in the QRS-complex for lead V_2 is observed in Fig. 3.5(g). The features, QRS-positive amplitude and QRS-negative amplitude in Fig. 3.5(h) and 3.5(i) respectively show a very good similarity for all precordial leads.

From the analysis of shape features, it is evident that the proposed model preserves P, ST and T-shapes. Also, from the analysis of amplitude features, it is observed that the model fails to preserve the P-wave amplitude. Even though the proposed model preserves the shape of the P-wave, it fails to preserve the amplitude. As the P-wave is a low amplitude wave, the model fails to capture this low dynamic range. Also the atrial information present in the predictor leads may not be enough for proper P-wave reconstruction.

TH-1915_11610237

The variations in $WEDD$ for the precordial leads of selected cardiac diseases is evaluated and is shown in Fig. 3.6. The cardiac cases include records of myocardial infarction (MI), healthy control (HC) data, dysrhythmia (DYS), cardiomyopathy (CDM), hypertrophy (HT) and bundle branch block (BBB). The mean and the median (in bracket) $WEDD$ values for these cases are: MI = 19.90% (17.78%), HC = 15.85% (13.91%), DYS = 18.92% (17.68%), CDM = 15.84% (15.45%), HT = 12.19% (12.43%) and BBB = 18.53% (19.95%). From Fig. 3.6 it can be observed that the third quartile is below 18% for all cardiac cases except MI and DYS. The low value of the median and the third quartile indicate that the proposed model can preserve diagnostic information in majority of the leads. The mean and median $WEDD$ for all 549 records are 18.97% and 17.02% with 32 records marked as outliers.

The time taken for the execution of a single patient is also computed experimentally. A model is simulated five times for the same patient record and the average time taken for its execution is calculated. The average time taken for pre-processing, model learning and signal reconstruction using the proposed model is found to be 1.54 seconds.

3.5.3 Repeatability of the proposed model

In this experiment, the ability of the model to reconstruct the leads of the same patient recorded at different instances is analyzed. In the PTB diagnostic ECG database, there are 27 patients with two records, 32 patients with three records, 50 patients with four records, three patients with five records and one patient with seven records. In the case of the patient with seven records, record three and six are not used for analysis due to high distortion. The average value is calculated for each precordial lead for all cases. This is shown in Table 3.2. Algorithm 1 is used to learn the model using the first record. Lead selection is performed on all records using Algorithm 2. This is the reason behind $WEDD$ being zero for different leads in the case of seven records. The selected model is then used to reconstruct the leads using Algorithm 3.

From Table 3.2, it is observed that the first record, which is used for learning the model gives lower values of $WEDD$. In the case of two, three and four records, a deviation in $WEDD$ value from the first record is observed. The variation in $WEDD$ value in the case of the same patient is due to less correlation between sub-band coefficients at different instances. This less correlation may be due to

TH-1915_11610237

TH-1915_11610237

Table 3.2: Repeatability of the model for patients with more than one record.

	Two Records		**Three Records**			**Four Records**				**Five Records**					**Seven Records**				
	1^{st}	2^{nd}	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}	4^{th}	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	1^{st}	2^{nd}	4^{th}	5^{th}	7^{th}
V_1	25.78	29.09	20.96	27.12	32.75	19.29	27.46	34.97	40.12	19.30	25.30	29.33	45.29	50.53	27.44	0	0	0	29.10
V_2	21.15	26.51	13.80	22.37	27.13	16.18	28.31	31.90	31.98	13.16	16.21	19.69	22.82	27.05	0	29.15	33.16	29.40	0
V_3	17.83	24.23	16.00	25.70	27.79	17.35	31.71	36.60	35.22	7.73	9.09	17.72	23.93	27.88	14.25	46.03	18.24	32.72	24.09
V_4	27.66	33.49	15.84	29.57	32.68	17.36	35.71	36.24	37.96	14.90	23.77	25.41	27.68	18.15	24.34	26.93	24.83	33.72	27.37
V_5	30.72	23.78	19.82	35.75	41.67	21.67	38.15	37.74	40.74	17.62	26.92	53.24	59.63	33.77	23.02	24.67	25.43	34.75	25.90
V_6	31.23	28.32	20.83	33.14	34.78	20.09	30.11	31.71	34.16	16.09	23.22	26.66	35.12	35.55	19.38	22.88	22.62	32.21	19.40
Avg.	25.73	27.57	17.88	28.94	32.80	18.66	31.91	34.86	36.70	14.80	20.75	28.68	35.75	32.16	21.68	29.93	24.86	32.56	25.17

Table 3.3: ECG diagnosability of the proposed model for Myocardial Infarction, Dysrhythmia, Cardiomyopathy and Bundle branch block.

		TP	FP	FN	TN	Diagnosability(%)
MI	Original	74	12	6	68	98.59
	Derived	73	13	7	67	
DYS	Original	12	2	4	14	96.15
	Derived	11	2	5	14	
CDM	Original	13	1	4	16	96.55
	Derived	13	2	4	15	
BBB	Original	15	1	2	16	96.77
	Derived	15	2	2	15	

the intra-person variability. Also, the placement of electrodes at different instances, as well as distortion in predictor leads, can affect the derived signal. It is also observed that from the second record onwards, the variation in $WEDD$ value is less. This is due to the lead selection algorithm, which selects the best model which minimizes $WEDD$ for the selected record. From this analysis, it can be concluded that the repeatability of the model is low and hence a model should be learned for each instance to have a proper reconstruction

3.6 ECG diagnosability of the proposed model

The diagnosability is defined and evaluated using the method described in section 3.4. Table 3.3 shows the ECG diagnosability of the proposed model for some selected cardiovascular abnormalities like myocardial infarction, dysrhythmia (DYS), cardiomyopathy (CDM) and bundle branch block (BBB). These classes are distinguished from healthy data. For MI, a total of 160 records, with 80 records each of MI and healthy data are used for training. The ECG diagnosability in the case of MI data is 98.59%. In the case of DYS, 16 records each of healthy and unhealthy cases are used to train the model. In this case, a diagnosability of 96.15% is observed. Similarly, 17 records, each of healthy and unhealthy data are used in the case of CDM and BBB. The diagnosability is 96.55% and 96.77% for CDM and BBB respectively. The ECG diagnosability statistics show that the classifier can distinguish between the healthy and the unhealthy signal with very good accuracy using reconstructed signal. It is evident in Table 3.3 that, for the evaluated classes, the TP and TN values of the derived

TH-1915_11610237

Table 3.4: Comparison with the existing models

	LR[46]	**ICA**[91]	**PCA**[8]	**SVR**[100, 101]	**Proposed Model**
No. of Leads	3	3	8	4	3
r_x	0.96	0.93	0.95	0.97	0.98
$RMSE$	88.69	95.56	53.66	94.74	63.88
R^2	88.61	86.72	90.21	-	94.23
$WEDD$	26.03	26.99	16.07	26.64	18.97

dataset is very much closer to that of the original dataset. It is also evident from this evaluation that the slight variations obtained in the clinically significant features of derived ECG is negligible. ECG diagnosability is slightly higher for MI compared to the other classes since it uses more data at the time of training. The good accuracy is because of the improvement in the quality of lead signals when using the optimal predictor lead. Thus, the model can generate good quality derived leads signals that can classify healthy and unhealthy data with accuracy as good as that of the original lead signals.

3.7 Comparison with the existing models

The objective of the proposed model is to derive the standard twelve-lead ECG from its subset. Hence, for fair comparison, the models discussed in [59, 63, 64, 61, 62, 65, 42, 88, 81], which do not derive the standard twelve-leads from its subset are not compared. In this work, the comparison is performed with the linear regression model [46], the independent component analysis (ICA) based model [91], the principal component analysis (PCA) based model [8] and the support vector regression (SVR) based model [100, 101]. This is shown in Table 3.4. Performance evaluation methods described in 3.5.1 are used for analyzing different models. For standardization of results, all these models are implemented, trained and tested on the same platform. Training of these models are performed using the learning set with s_1 samples. Testing of these models are performed using the reconstruction dataset of s_3 samples. Since the SVR-based model is a non-linear model, evaluation using R^2 statistics is not performed. It can be noticed from Table 3.4 that the proposed model outperforms the other models in case of correlation and R^2 values. Also, the proposed model gives low $RMSE$ and $WEDD$ values than the other models except the PCA model. Since the PCA based model requires all eight leads for calculating the principal components (PC), it captures more infor-

TH-1915_11610237

mation which automatically improves the results. The $RMSE$ and $WEDD$ values of the proposed model is comparable with the PCA model despite using only three leads for derivation. It is evident from the comparison that the performance of the proposed model is better than the other models.

3.8 Summary

This chapter presents a novel method for enhancing the spatial resolution of a three-lead ECG. A lead selective algorithm which maximizes the diagnostic similarity of derived lead signals with the original lead signals is proposed. A linear model is learned by utilizing the high inter-lead correlation in the wavelet domain. Various performance evaluations are conducted on the proposed model using different performance evaluation metrics such as correlation, $RMSE$, R^2 statistics, $WEDD$ and WDD. A comparison of the proposed model is performed with the existing models. Analysis such as repeatability and ECG diagnosability is also performed to evaluate the model. From the performance evaluation results, it is evident that the proposed model preserves diagnostic information in derived leads than the existing models. The advantage of the proposed linear model is that it can be computationally efficient but may over-fit and may fail to capture the minute diagnostic details.

TH-1915_11610237

TH-1915_11610237

4

Sparse Domain Models using Joint Dictionary Learning Framework

Contents

TH-1915_11610237

In the previous chapter, a linear model which exploits the frequency domain characteristics of ECG signal is discussed. A linear model can be computationally efficient, but may fail to capture the minute diagnostic details in the signal. Also, a linear model is susceptible towards over-fitting. In this chapter, a new technique for improving the spatial resolution of ECG by integrating sparse domain and dictionary learning framework is discussed. Since the spatial resolution is high for a standard twelve-lead ECG, it can be identified as the high-resolution (HR) ECG and its subset as the low-resolution (LR) ECG. It should be noted that the term resolution refers to the spatial resolution and hence should not be misunderstood for time or amplitude resolution. Two over-complete personalized dictionaries corresponding to high and low-resolution ECG are jointly learned and high-resolution ECG is recovered from its low-resolution counterpart. Similar joint dictionary learning approaches can be found in applications like image processing [105, 106, 107], bandwidth extension [108], etc., but have not been explored for ECG signals. The main advantage of using a personalized dictionary over traditional dictionaries is that it can adjust to the characteristics of the signal.

The different leads of ECG can be obtained using the heart vector by projecting it onto the human torso. This ensures a very good correlation between different lead signals in ECG. As the LR ECG is nothing but a subset of the standard twelve-lead signals, joint learning of dictionaries may ensure similar sparse representation for LR and HR ECG. This similar sparse representation can be utilized to reconstruct the HR ECG from LR ECG. The sparse coding of the electrocardiogram signal using a personalized dictionary has many applications in areas like personalized health care, compressive sensing, telemonitoring, etc. [109, 110, 111, 112]. The objective of this study is to explore the applicability of dictionary learning and sparse domain representation in improving the spatial resolution of ECG.

In this chapter, two models that generate HR ECG from its LR subset is discussed and evaluated. The first model learns a joint dictionary consisting of high and low-resolution dictionary pairs. Once the joint dictionary is learned, the sparse coefficients of low and high-resolution ECG can be obtained. The low and high-resolution sparse coefficients are then mapped using a conversion function. The conversion function ensures a good similarity between the low and the high-resolution sparse coefficients. Such an approach can improve the reconstruction accuracy of derived lead signals. This model uses a four-lead subset of the standard twelve-lead ECG as the LR ECG. Once the model is

TH-1915_11610237

learned, it can be used to generate the high-resolution ECG from a four-lead subset. In the second model, the signal is divided into multiple segments and a similar learning process as that of the first model is followed for each segment. Segmenting and learning the model ensures that the clinically significant features are preserved while reconstructing. The second model uses a three-lead subset of the standard twelve-lead ECG as the LR ECG. The performance of the proposed methods are assessed by comparing the spatially enhanced ECG with the original ECG. This is performed using standard distortion measures such as correlation coefficient and $RMSE$. The changes in diagnostic information of the spatially enhanced ECG is evaluated using $WEDD$ and WDD measures [18, 104]. The comparison of the proposed models are also performed with similar existing models.

This chapter is organized as follows: In section 4.1, dictionary learning and sparse coding of ECG is discussed. Conversion function in the sparse domain is described in 4.2 and segmentation of the signal is presented in 4.3. In section 4.4, the detailed description of the proposed methods are presented. Results and its analysis are discussed in 4.5. In section 4.6, ECG diagnosability of the proposed model is evaluated. The performance of the proposed model is compared with the existing methods in 4.7 and the chapter is summarized in section 4.8.

4.1 Dictionary learning and sparse coding of ECG signal

A personalized over-complete dictionary can be used to obtain the sparse representation of a standard twelve-lead ECG. The main advantage of using a personalized dictionary is that it can adapt to the signal characteristics, which is required in many applications involving signal recovery [106]. The personalized dictionary learning problem can be posed as

$$\min_{D,s} \|E - Ds\|_F^2 \text{ subject to, } \|d_i\|_2 \leq 1 \ \ \|s_i\|_0 \leq T_0 \ \ \forall i \ . \tag{4.1}$$

where the twelve-lead ECG data is represented as E, its sparse representation as s the personalized dictionary as D and its atoms as d_i. Many over-complete dictionary learning approaches are available of which the K-SVD algorithm is selected as it is a simple and flexible dictionary learning method [123]. Since the K-SVD algorithm can be used with any pursuit algorithm, the sparse coding of the signal is carried out using the orthogonal matching pursuit (OMP) algorithm [124]. The OMP algorithm is selected due to its simplicity and fast implementation. Two dictionaries corresponding to

TH-1915_11610237

HR and LR ECG are learned simultaneously. This is performed by a joint dictionary learning approach which is inspired from the popular image super-resolution problem [105]. The joint learning approach is used to ensure a similar sparse representation of both LR and HR ECG. Thus, the problem defined in equation 4.1 can be updated as

$$\min_{D,s} \left\| \begin{bmatrix} E^h \\ E^l \end{bmatrix} - \begin{bmatrix} D^h \\ D^l \end{bmatrix} s \right\|_F^2 \quad \text{subject to,} \;\; \|d_i\|_2 \leq 1 \;\; \|s_i\|_0 \leq T_0 \;\; \forall i \;\; . \tag{4.2}$$

where D^h and D^l represent dictionaries corresponding to HR and LR ECG, E^h and E^l represent the twelve-leads and its subset, and $s = s^h \approx s^l$. The LR ECG is spatially aligned with the HR ECG by stacking it below the latter. The HR and LR ECG, i.e. E^h and E^l can be related as $E^l = AE^h$, where A is a matrix that transforms HR ECG to LR ECG. As the LR ECG is a subset of the HR ECG and the dictionaries are learned jointly, it can be assumed that $s^h \approx s^l$. In other words, the low-resolution ECG $E^l = AD^h s^h \approx AD^h s^l$. This assumption can be further exploited to generate the HR ECG from the LR ECG. Once the joint dictionary D consisting of D^h and D^l is learned, they can be separated out to obtain the HR and the LR dictionaries.

4.2 Learning the conversion function in the sparse domain

In section 4.1, it is assumed that the sparse representation of both LR and HR ECG with their corresponding dictionaries are similar. However, in case of signals like ECG, this assumption is too rigid and fails to address the variations in clinically significant features. A relaxation over this assumption is possible if a stable mapping exists between the HR and the LR sparse representations [107]. This stable mapping can address the variations in clinically significant features and hence an accurate reconstruction of high-resolution ECG is possible. In this chapter, the mapping is termed as the conversion function C_f and is learned in the sparse domain to transform the LR sparse coefficients to its HR counterpart. This can be posed as

$$\min_{C_f} \|s^h - C_f s^l\|_F^2 + \lambda_{C_f} \|C_f\|_F^2 \;\; . \tag{4.3}$$

where λ_{C_f} is the regularization parameter. Equation 4.3 is a ridge regression problem which is solved using the *cvx* optimization toolbox [125, 126]. Once the conversion function C_f is learned, the HR

TH-1915_11610237

sparse coefficients can be derived from LR sparse coefficients.

4.3 Fine tuning the model by segmentation

The ECG signal contains several clinically significant features with varying amplitude, shape and frequency. It is reported in [24, 114] that segmenting these features and learning a model improves the reconstruction quality. The segmentation of both high and low-resolution ECG is performed so that the clinically significant features are preserved. Segmentation of the signal presented in this chapter is based on thresholding and is performed during the learning and the reconstruction stages. The different steps followed for calculating the segments by thresholding the signal is as follows

- Calculate the absolute value of each samples in $Lead\ I$.
- Sort these values in ascending order.
- Divide this into *k*-segments of equal width.
- Calculate the upper and the lower bounds of these *k*-segments.

The absolute value of each sample is selected so that samples with similar amplitudes can be grouped together. This ensures grouping of temporal samples with similar characteristics thereby improving the dictionary for that particular segment. Since thresholding is required in learning and reconstruction stage, $Lead\ I$ is used as it is present in both HR and LR ECG. Once the segments are calculated, the samples are assigned into corresponding segments based on these bounds. The samples of all leads at the same instant as that of the $Lead\ I$ sample is passed into the same segment as they are spatially aligned. This method ensures grouping of the spatially aligned samples of all leads with similar characteristics. It also ensures equal number of data samples for all segments at the time of learning. Once the segmentation is completed, a joint dictionary representing the HR and the LR ECG, and the conversion function are learned for each segments. In the reconstruction stage, the absolute value of each sample in $Lead\ I$ is calculated and by using the previously calculated upper and lower bounds, the corresponding segment is selected.

TH-1915_11610237

4.4 Proposed joint dictionary learning models

In this chapter, two models that use the dictionary learning framework to generate HR ECG are discussed. The HR ECG is the standard twelve-leads and the LR ECG is the subset of the twelve standard leads. The volume-conductor theory by Burger and van Milaan states that a minimum of three independent leads are required to reconstruct the twelve-leads [5, 6]. This is based on the assumption that the heart vector or the heart dipole can be modeled using a single dipole fixed at the center of the heart. In this way, the heart vector can be expressed as the sum of three orthogonal components and is known as the Burger equation [5, 6]. Once the heart vector is obtained, any leads can be generated by projecting the heart vector onto the human torso.

A pre-processing is performed on all twelve-leads before the learning and the reconstruction stage for both models. This will remove any artifacts in the signal and will ensure good reconstruction quality. Standard pre-processing methods like removal of the baseline wandering effect, normalization of the amplitude and removal of the mean is performed for all lead signals. In order to remove the baseline wandering effect, the signal is passed through a low-pass filter with cut-off frequency set at 0.67 Hz. Assuming that there are no important diagnostic details below 0.67 Hz, the output is subtracted from the original signal [119].

4.4.1 Joint dictionary learning for spatially enhanced ECG

The proposed joint dictionary learning model for generating the spatially enhanced ECG is discussed. In this approach, the LR ECG uses four electrodes consisting of $Lead\ I$, $Lead\ II$, V_1 and V_3. The HR-LR dictionary pair is learned using K-SVD algorithm and the corresponding sparse coefficients are obtained using OMP. Once the sparse coefficients are obtained, the conversion function is learned so as to map the LR sparse coefficients to the HR sparse coefficients. In the reconstruction stage, the LR dictionary is used to obtain the LR sparse coefficients using OMP. This LR sparse coefficients and the conversion function C_f is then used to obtain an estimate of the HR sparse coefficients. This is represented as $\hat{s}^h = C_f.s^l$, where $\hat{s}^h$ is the estimated HR sparse coefficients. The estimate of HR sparse coefficients $\hat{s}^h$ is multiplied with the HR dictionary D^h to produce an estimate of the HR ECG, i.e., $\hat{E}^h = D^h.\hat{s}^h = D^h.C_f.s^l$. The post-processing of all twelve-leads in the estimated HR ECG is performed using a moving average filter of length five, to obtain the final estimate of HR ECG ($\hat{E}^h$).

TH-1915_11610237

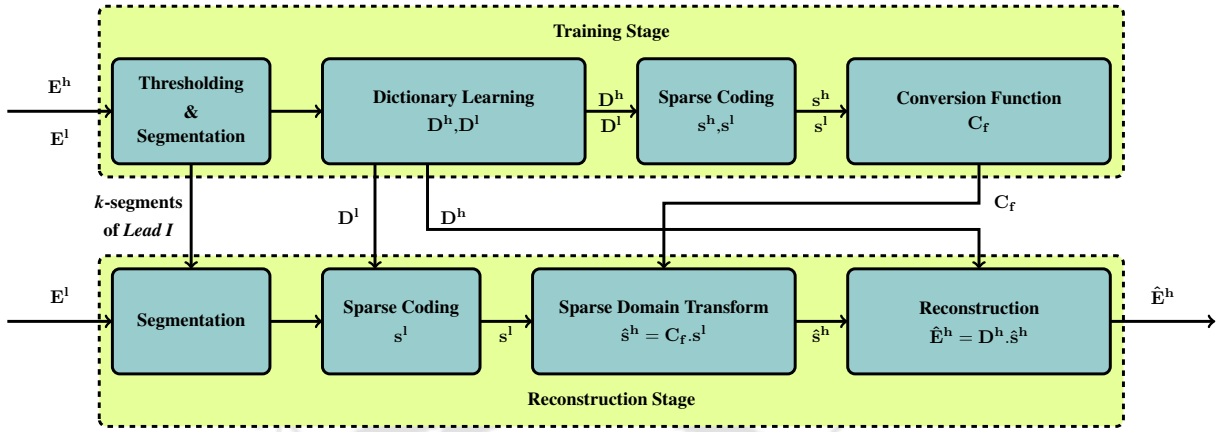


Figure 4.1: The block diagram representation of multiple joint dictionary learning model.

4.4.2 Multiple joint dictionary learning for spatially enhanced ECG

The second approach for enhancing the spatial resolution of ECG is by using multiple joint over-complete dictionaries. This approach also aims to recover the high-resolution ECG from its low-resolution counterpart. For this model also, the HR ECG is the standard twelve-lead ECG and LR ECG is a subset of standard twelve-leads. As compared to the previous model, the low-resolution subset consists of only three independent leads which are $Lead\ I$, $Lead\ II$ and V_3. The limb leads $Lead\ I$ and $Lead\ II$ are selected for ensuring a good reconstruction quality of the frontal plane leads as they are linearly related with these two [2]. The precordial lead V_3 is selected based on experimental results and also from the observations made in [24]. The ECG signal is divided into multiple segments using the method discussed in section 4.3 so as to capture various diagnostic features and joint dictionaries are learned for each segments. This approach preserves clinically significant diagnostic features and ensures a good reconstruction quality. Fig. 4.1 shows the block diagram representation of the proposed model. As shown in the block diagram, the complete process can be divided into two stages, the learning stage and the reconstruction or the testing stage. In the learning stage, the pre-processed signal is segmented as per the procedure described in section 4.3. The model is learned as follows. A joint dictionary is learned using the K-SVD algorithm to obtain the HR and the LR dictionaries. Once the joint dictionary is obtained, the HR-LR dictionaries are separated out and by using the OMP algorithm, the sparse representation of the data is obtained. These HR and LR sparse coefficients are then used to learn the conversion function. This procedure

TH-1915_11610237

is followed for all segments so that a model is obtained for each segment.

In the reconstruction stage, the input LR ECG is allotted to the previously defined segments using the procedure described in section 4.3. The corresponding low-resolution dictionary is then used to obtain the sparse coefficients of the input LR ECG. This LR sparse coefficients when applied on the conversion function produces an estimate of the HR sparse coefficients, i.e., $\hat{s}^h = C_f.s^l$. This estimate $\hat{s}^h$ is then multiplied with the HR dictionary D^h of the corresponding segment to generate an estimate of HR ECG, i.e., $\hat{E}^h = D^h.\hat{s}^h = D^h.C_f.s^l$. The final estimated HR ECG, i.e. $\hat{E}^h$ is obtained after the post-processing of all derived twelve-leads using a moving average filter of length five. This ensures smoothening of small discontinuities that may arise in the lead signals, especially in the high amplitude regions. The discontinuity occurs when the dictionary fails to estimate HR samples and may appear as a notch in the signal. The length of this notch is very small and hence to smoothen it, a five point moving average filter is selected. Thus the proposed method ensures good reconstruction quality along with minimal loss of significant diagnostic information.

4.5 Results and discussion

In this study, the most commonly accessible PTB (Physikalisch-Technische Bundesanstalt) diagnostic ECG database [19][20] is used. The details of the database is given in section 2.3. Each record of length thirty second is divided further into two sets with the first twenty seconds used for training and the last five seconds used for reconstruction. The five seconds between training and reconstruction set is removed to introduce a discontinuity between both. For the first model, the training data E is arranged as a 16$\times$N$_q$ matrix (12$\times$N$_q$ HR ECG and 4$\times$N$_q$ LR ECG). The dictionary size used in this model is fixed as 32 and the number of atoms used in each linear combination is fixed as two. In the case of the second model, the training data E is arranged as a 15$\times$N$_q$ matrix (12$\times$N$_q$ HR ECG and 3$\times$N$_q$ LR ECG). The number of samples used for learning the dictionary in the q^{th} segment is defined as N$_q$. For the first model q = 1 as the signal is not segmented. In the case of the second model, the data is divided into twenty five segments and the size of the dictionary for each segment is forty. For this model also, the number of atoms used in each linear combination is fixed as two. The conversion function for both cases is learned using the *cvx* optimization toolbox [125, 126].

TH-1915_11610237

Table 4.1: Performance evaluation of the multiple joint dictionary for all patient records.

	r_x	$RMSE(mV)$	$WEDD(\%)$	$WDD(\%)$
LeadI	**0.99**	**28.69**	**7.13**	**2.69**
LeadII	**0.99**	**29.90**	**7.22**	**2.66**
LeadIII	0.98	53.76	14.81	5.62
aVR	0.99	28.94	7.39	2.96
aVL	0.98	45.47	12.49	4.87
aVF	0.99	45.50	11.78	3.96
$\boldsymbol{V_1}$	**0.99**	**29.50**	**8.10**	**3.96**
$\boldsymbol{V_2}$	0.98	64.20	16.11	5.47
$\boldsymbol{V_3}$	**0.99**	**25.41**	**6.99**	**2.46**
$\boldsymbol{V_4}$	0.97	84.79	23.67	6.41
$\boldsymbol{V_5}$	0.95	96.09	29.01	9.10
$\boldsymbol{V_6}$	0.95	88.65	27.01	9.20
Avg. of All leads	0.98	51.74	14.31	4.95
Avg.of Precordial leads	0.96	83.43	23.95	7.53

4.5.1 Performance evaluation

The reconstructed signal is compared with the original signal for evaluating the proposed model. Popular distortion measures like correlation coefficient (r_x), and $RMSE$ are used to evaluate the performance of the models. The variation in diagnostic content between the original and the reconstructed leads are assessed using $WEDD$ and WDD measures [18, 104]. The details of these measures are discussed in section 3.4. The distortion measures except WDD are computed for all 549 records. Only 524 records are evaluated using WDD due to the difficulty in automatically extracting the features. For the second model, a comparison of diagnostically significant features between the original and the reconstructed lead signals are also performed. Additionally, the ECG diagnosability of the second model is computed using the method described in section 3.4 [24]. The relative error is also calculated by taking the difference between original and derived ECG's accuracies.

4.5.2 Results of joint dictionary learning model

The HR ECG generated using the first model is assessed by comparing the derived leads with the original leads. The performance of the model in converting the LR ECG to the HR ECG is presented in Table 4.1. The predictor leads $Lead\ I$, $Lead\ II$, V_1 and V_3 are the least distorted leads as they are supplied as the LR ECG input to the model. The average value of correlation coefficient, $RMSE$,

TH-1915_11610237

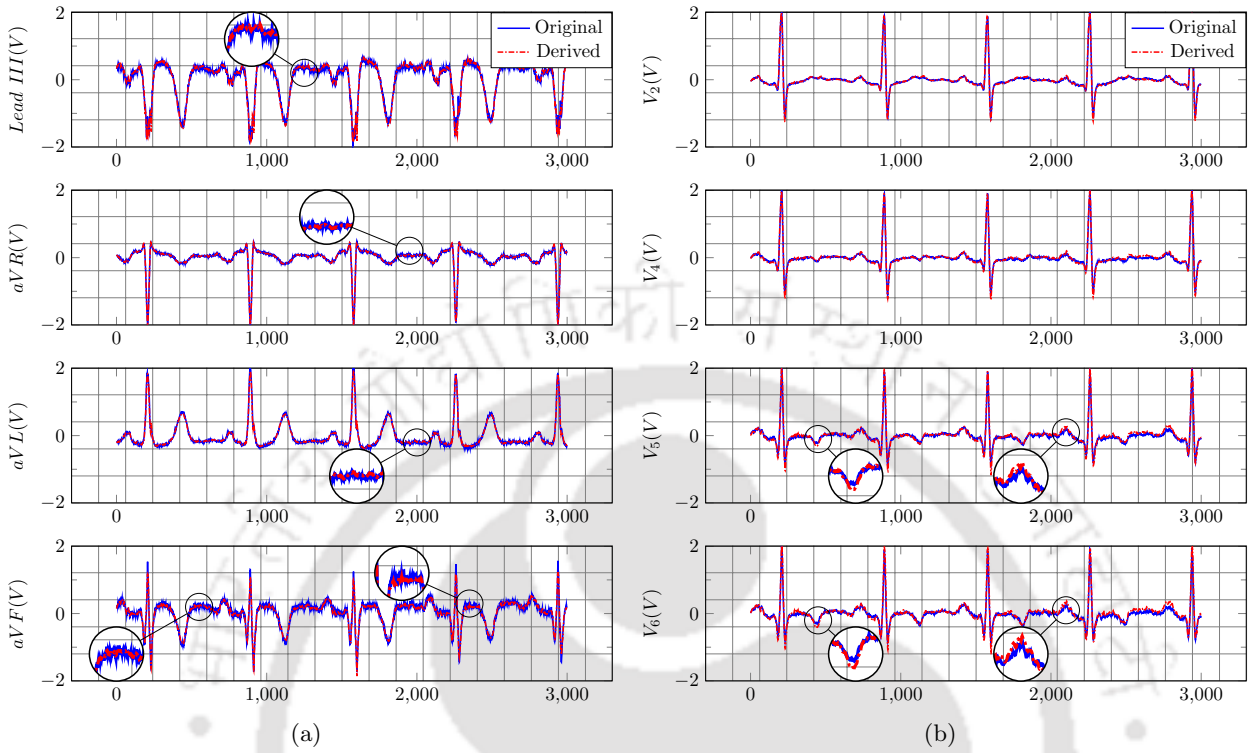


Figure 4.2: (a) Original (blue) and reconstructed (red, dash-dotted) frontal plane leads for inferior MI record 's0114lrem'; (b) Original (blue) and reconstructed (red, dash-dotted) precordial leads for inferior MI record 's0114lrem'.

$WEDD$ and WDD for all leads are 0.98, 51.74 mV, 14.31% and 4.95% respectively. It is also evident from the table that the model successfully derived the frontal plane leads with minimal reconstruction error. This is because of the linear relationship between $Lead\ I$ and $Lead\ II$ with the other frontal plane leads. Among the derived precordial leads, the model generated lead V_2 with high values of correlation, and low $RMSE$, $WEDD$, and WDD values. It is also evident from the table that the reconstruction accuracy is less for leads V_4, V_5 and V_6. Thus, the leads nearer to the precordial predictor leads reconstructed successfully when compared to leads farther from it, and this effect is termed as the proximal lead effect [46]. The overall value of correlation is 0.96, $RMSE$ is 83.43 mV, $WEDD$ is 23.95% and WDD is 7.53% for the precordial leads. The average value of correlation coefficient indicates that all derived leads maintain a very good similarity with the original HR ECG. The low $RMSE$ value indicates less distortion in the spatially enhanced HR ECG. The $WEDD$ values falls in the good category for most of the frontal plane leads. Although the precordial leads posses

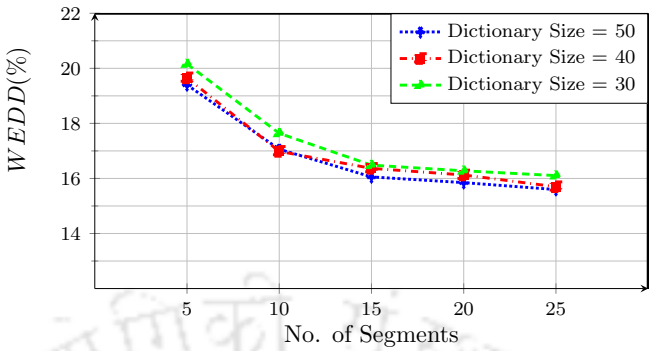


Figure 4.3: Variation in $WEDD(\%)$ with the number of segments and dictionary size for derived precordial leads.

good correlation, the $WEDD$ values falls in the bad category, but is better than similar models. The WDD values fall in the good category for all twelve-leads. The low values of $WEDD$ and WDD show that the estimated HR ECG preserves diagnostic information. Also, the median $WEDD$ value of all leads for the whole record is 8.94%, and that of the precordial leads is 20.02%. This shows that the proposed model reproduced the standard twelve-lead ECG with good reconstruction accuracy for half of records.

The original and the reconstructed ECG signal of the inferior MI record 's0114lrem' is displayed in Fig. 4.2(a) and 4.2(b). The $WEDD$ value for this record is 14.31%. In Fig. 4.2(a), it can be observed that the original and derived frontal plane leads maintain a good similarity. Also, the noise present in the original leads are reduced in the derived leads, which is an advantage of using a personalized dictionary. In Fig. 4.2(b), which shows the original and derived precordial leads, the P-wave and the T-wave are slightly overestimated in leads V_5 and V_6. It is also evident from Fig. 4.2(a) and 4.2(b) that the reconstructed leads retain the shape of the P-wave, the QRS-complex and the T-wave for all the leads. Also, the original and derived lead signals are time-aligned. It is apparent from the figures that the spatially enhanced ECG preserves the shape, the amplitude and the duration features.

4.5.3 Results of multiple joint dictionary learning model

The performance of the second model in converting a LR ECG to a HR ECG using the multiple joint dictionary learning technique is discussed. Experiments are conducted by varying the number of segments with different dictionary sizes. The variation in $WEDD$ with the change in dictionary size

TH-1915_11610237

Table 4.2: Performance evaluation of the multiple joint dictionary learning model for all patient records.

	r_x	$RMSE(mV)$	$WEDD(\%)$	$WDD(\%)$
$LeadI$	**0.99**	**25.55**	**5.91**	**2.27**
$LeadII$	**0.99**	**27.53**	**6.58**	**2.29**
$LeadIII$	0.99	42.28	11.60	4.44
aVR	0.99	28.15	6.93	2.77
aVL	0.99	36.49	9.93	3.82
aVF	0.99	37.81	9.64	3.12
V_1	0.97	61.84	17.87	9.83
V_2	0.99	48.40	12.63	6.04
V_3	**0.99**	**18.99**	**5.15**	**1.85**
V_4	0.98	49.97	14.45	5.55
V_5	0.98	56.07	17.04	7.45
V_6	0.97	55.02	16.50	7.84
Avg. of All Leads	0.99	40.68	11.19	4.77
Avg.of Precordial Leads	0.98	54.26	15.70	7.34

for the derived precordial leads and the number of segments is shown in Fig. 4.3. It is observed from Fig. 4.3 that an increase in the size of the dictionary triggers a decrease in $WEDD$. But an increase in the size of the dictionary increases the computational burden. It is also observed that an increase in the number of segments causes a decrease in $WEDD$. As the number of segments is increased further, the effect of dictionary size starts diminishing which can be observed in Fig. 4.3. In all these cases, the number of atoms used in each linear combination is fixed at two. Based on the observations from Fig. 4.3, a dictionary of size forty is used for each segment with the number of segments fixed at twenty-five. An advantage of using the low size dictionary is that the learning time can be reduced.

Table 4.2 shows the performance of the model in converting a LR ECG to a HR ECG. The predictor leads of the LR ECG are $Lead\ I$, $Lead\ II$ and precordial lead V_3. Table 4.2 shows very high values of correlation for all leads. The average value of correlation, $RMSE$, $WEDD$, and WDD for all leads are 0.99, 40.68 mV, 11.19% and 4.77% respectively. The limb leads and augmented limb leads posses a correlation of 0.99 with the original signal. Among the precordial leads, the correlation is low for V_1 and V_6 at 0.97 followed by V_4 and V_5 at 0.98. The highest value of correlation is shown by V_2 and predictor lead V_3. It can also be observed from Table 4.2 that limb leads and augmented

TH-1915_11610237

limb leads show low values of $RMSE$, $WEDD$ and WDD. This is because of the linear relationship between the six frontal plane leads. The $RMSE$, $WEDD$ and WDD values are higher for precordial leads except for predictor lead V_3. The precordial lead V_1 shows highest $RMSE$, $WEDD$ and WDD values at 61.84 mV, 17.87% and 9.83% respectively. It can also be observed from Table 4.2 that the leads closer to the predictor precordial lead V_3 posses very good correlation, and low $RMSE$, $WEDD$ and WDD values because of the proximal lead effect. Considering the precordial leads other than V_3, the values of correlation, $RMSE$, $WEDD$ and WDD are 0.98, 54.26 mV and 15.70% and 7.34% respectively. The average value of correlation coefficient indicates that all derived leads posses a very good resemblance with the original HR ECG. Low $RMSE$ indicates less distortion in the spatially enhanced HR ECG. The low values of $WEDD$ and WDD show that the estimated HR ECG maintains the diagnostic information.

The original and the reconstructed precordial leads of two different records belonging to two different classes are shown in Fig. 4.4(a) and 4.4(b). The median of $WEDD$ values of the precordial leads for these classes are used to select the records. The original and the derived precordial leads of a healthy control data record 's0487_rem' is shown in Fig. 4.4(a). This record gives an average $WEDD$ of 9.74%. In Fig. 4.4(a), a minor overestimation of T-wave and a minor underestimation of S-wave is observed for leads V_4 and V_5. In leads V_1 and V_2, the T-wave is slightly underestimated, but retains the shape. A slight dip in R-peak is observed in leads V_1, V_2 and V_4. In Fig. 4.4(b), the original and the reconstructed precordial leads of the anterior MI record 's0027lrem' is shown. In this case, the average $WEDD$ is 12.72%. In Fig. 4.4(b), it can be observed that the reconstructed signal retains the shape of P-wave, QRS-complex and T-wave for all precordial leads. It can also be observed that distortion present in the original signal is missing in the derived leads, which is an advantage of the dictionary learning approach. In V_2, the T-wave amplitude is slightly overestimated while in V_4, V_5 and V_6 it is underestimated. It is evident from Fig. 4.4(a) and 4.4(b) that spatially enhanced ECG retains the shape and amplitude features in the majority of leads.

The feature level analysis of the second model is performed for further quantification and is displayed in Fig. 4.5. The amplitude and shape features of original and derived precordial lead signals are compared. The variation in the duration features are negligible as the original and the derived lead signals are spatially aligned. The shape features used in the analysis consist of the P-shape,

TH-1915_11610237

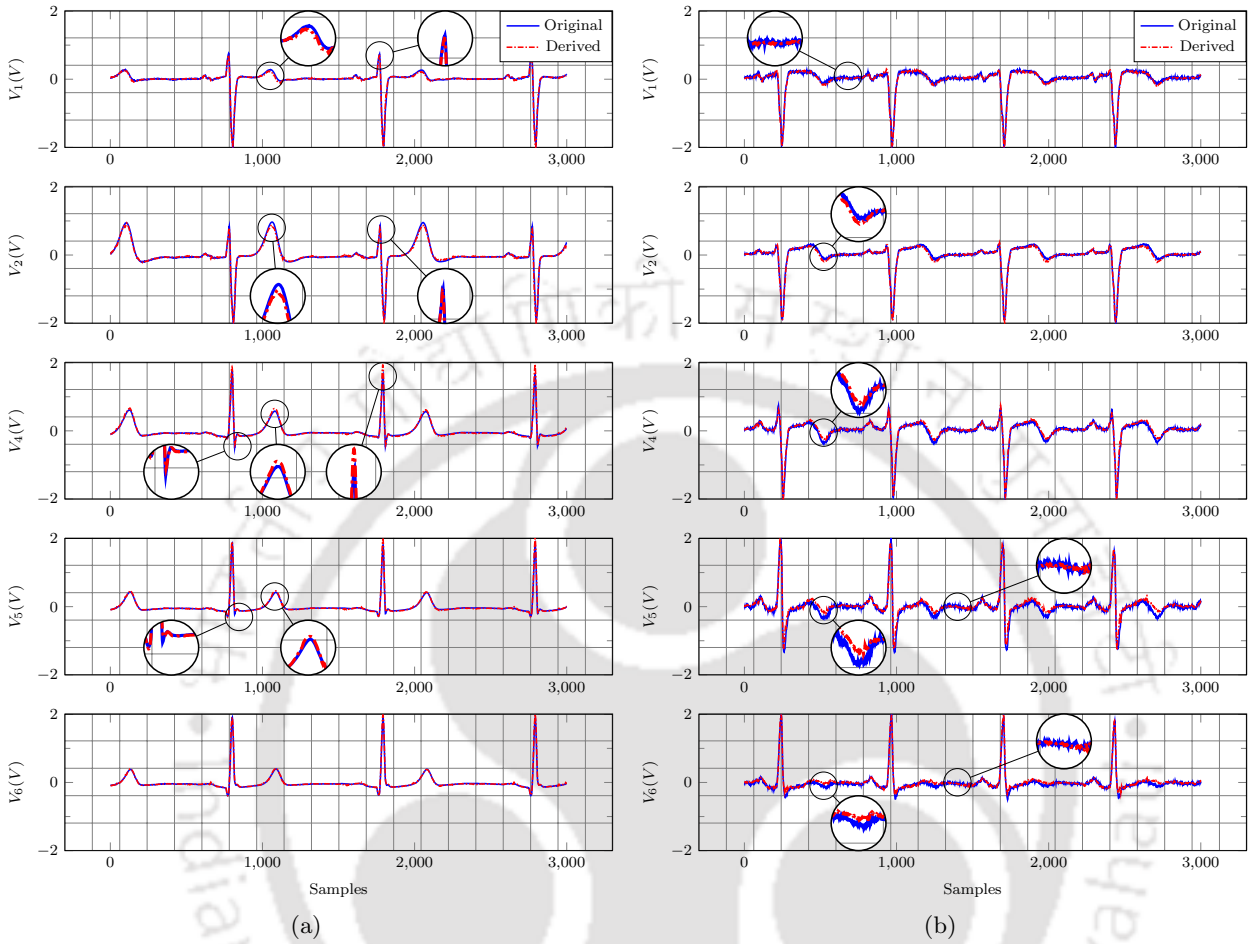
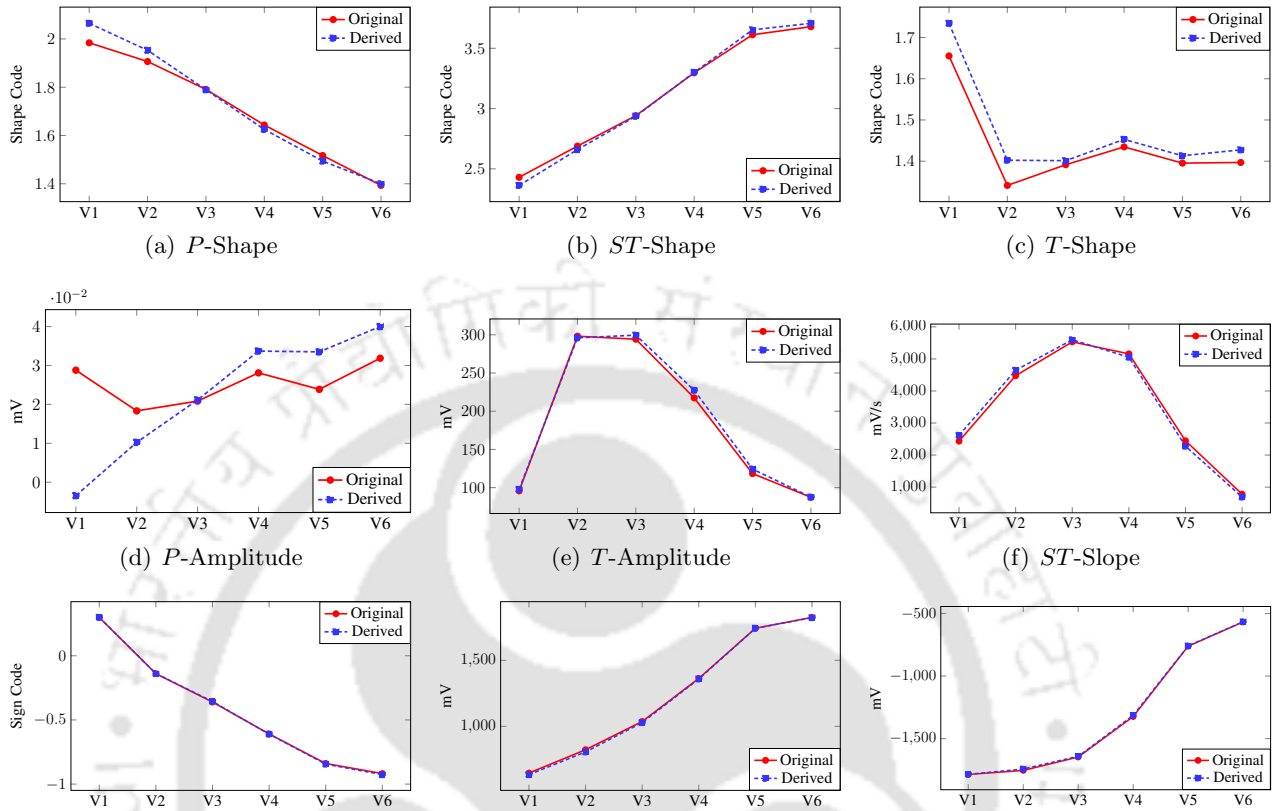


Figure 4.4: (a) Original (blue) and reconstructed (red, dash-dotted) precordial leads for healthy control record 's0487_rem'; (b) Original (blue) and reconstructed (red, dash-dotted) precordial leads for anterior MI record 's0027lrem'.

the ST-shape and the T-shape. The WDD of the shape features are assigned a shape code and the same shape code is used for this analysis [18]. A slight variation in the P-shape of the derived leads V_1 and V_2 is observed in Fig. 4.5(a). For other precordial leads, the P-shape of the derived leads follows the original leads. The ST-shape of the derived leads follow the original leads with minimum error and is evident in Fig. 4.5(b). A slight variation in the derived T-shape is observed in Fig. 4.5(c) for all precordial leads. The amplitude features of the derived leads such as the T-amplitude, the ST-slope, the sign of the first peak in the QRS-complex (1 or -1), the QRS-positive amplitude, and the QRS-negative amplitude follows the original leads with minimum reconstruction error. This is evident in Fig. 4.5. A variation from the original feature is observed in the derived P-amplitude for



(a) P-Shape (b) ST-Shape (c) T-Shape

(d) P-Amplitude (e) T-Amplitude (f) ST-Slope

(g) QRS-Sign (1 or -1) (h) QRS Positive Amplitude (i) QRS Negative Amplitude

Figure 4.5: Comparison of shape and amplitude features between original leads and reconstructed leads.

all precordial leads in Fig. 4.5(d). This variation is small and is in the range of a few millivolts. This is because, the model fails to represent the small variations in P-wave since its amplitude is low when compared to the other features. Also, the lack of sufficient atrial information in the predictor leads might have influenced the reconstruction accuracy of P-wave. From the analysis of these features shown in Fig. 4.5, it can be conclude that the proposed model is capable of capturing diagnostically significant features.

The variations in $WEDD$ for the precordial leads of selected cardiac diseases is evaluated and is shown in Fig. 4.6. The selected cardiac cases include 368 records of myocardial infarction (MI), 80 records of healthy control (HC) data, sixteen records of dysrhythmia (DYS), seventeen records of cardiomyopathy (CDM) seven records of hypertrophy (HT) and seventeen records of bundle branch block (BBB). The mean and the median (in bracket) $WEDD$ values for these cases are: MI = 15.32% (12.59%), HC = 14.01% (9.72%), DYS = 13.00% (12.02%), CDM = 11.45% (12.24%), HT = 7.79%

TH-1915_11610237

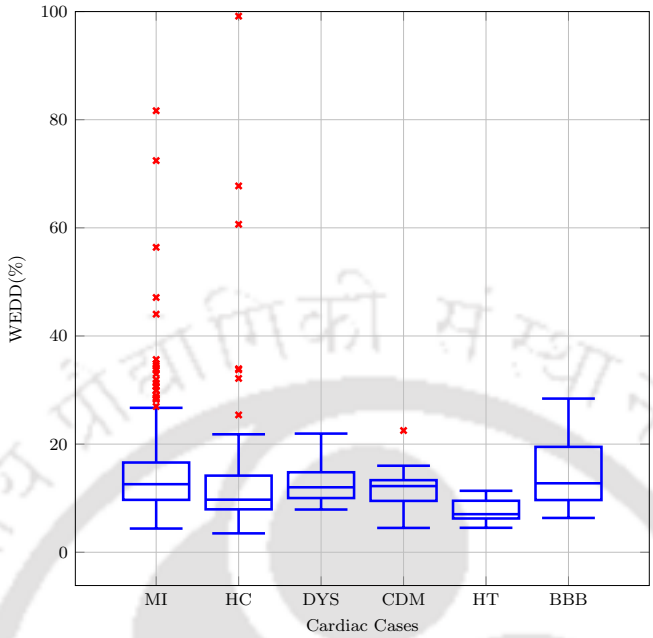


Figure 4.6: Box Plot showing the variation in $WEDD(\%)$ for the precordial leads of myocardial infarction (MI), healthy control (HC), dysrhythmia (DYS), cardiomyopathy (CDM), hypertrophy (HT) and bundle branch block (BBB).

(7.04%) and BBB = 14.51% (12.76%). Also, it can be observed from Fig. 4.6 that in case of MI, HC, DYS, CDM and HT, the third quartile is below 17% and for BBB, it is below 20%. The low value of the median and the third quartile indicate that diagnostic information is preserved in the majority of selected cardiac cases when reconstructed using the proposed model. The mean and median of precordial leads for all 549 records are 15.70% and 12.22% respectively with 43 records marked as outliers. These outliers are observed for poor quality records in the database which degrades the performance of the model.

The time taken for the execution is computed experimentally for a single patient record. The average execution time is computed by simulating the model for five times. The total time taken from pre-processing till final signal generation is found to be 58.95 seconds.

4.6 ECG diagnosability of multiple joint dictionary learning model

The capability of the model to generate the HR ECG with useful information is evaluated using diagnosability. The diagnosability is computed by taking the percentage of classification accuracy of

TH-1915_11610237

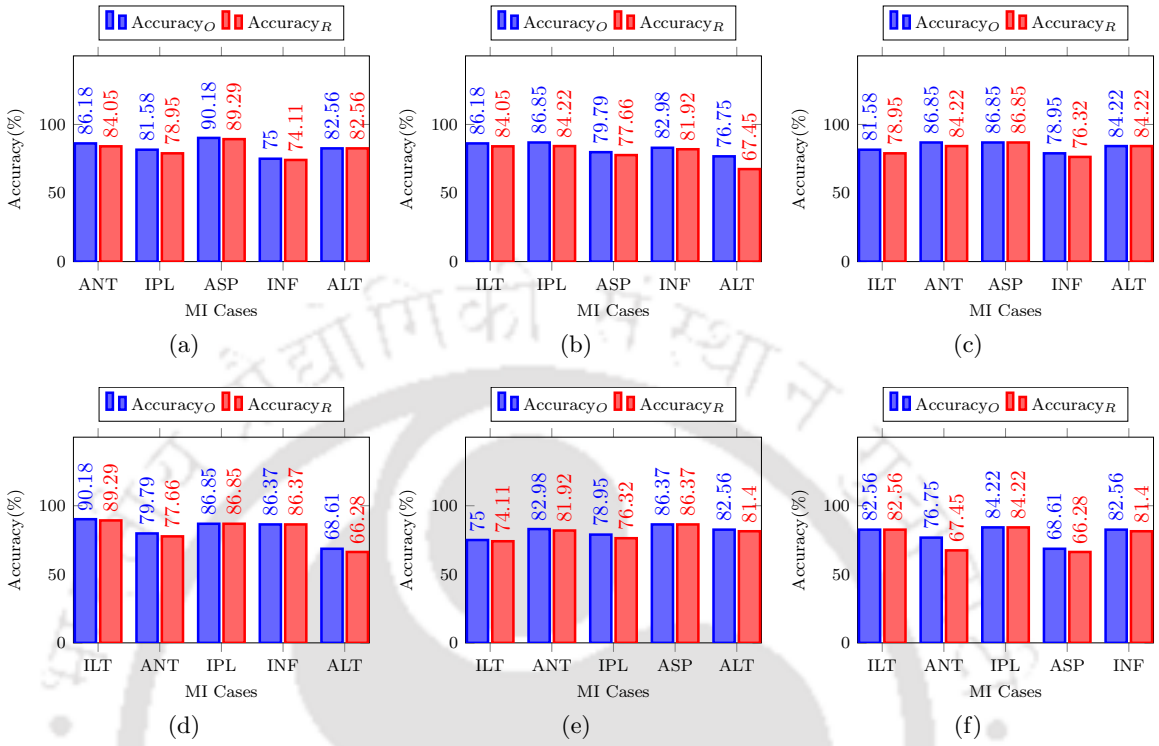


Figure 4.7: Comparing the accuracy of original and derived leads for (a) infero-lateral (ILT) MI with other MI cases; (b) anterior (ANT) MI with other MI cases;(c) infero-postero-lateral (IPL) MI with other MI cases;(d) antero-septal (ASP) MI with other MI cases;(e) inferior (INF) MI with other MI cases;(f) antero-lateral (ALT) MI with other MI cases.

the spatially enhanced HR ECG with that of the original HR ECG. Selected MI cases such as infero-lateral (ILT) MI, anterior (ANT) MI, infero-postero-lateral (IPL) MI, antero-septal (ASP) MI, inferior (INF) MI and antero-lateral (ALT) MI are classified between each other. The ECG diagnosability is computed by following the method described in section 3.4. An SVM based binary classifier is used with equal number of records for both cases. The model is trained with the original signal of length twenty seconds. The last five seconds of the thirty second samples are then used to obtain classification accuracy of original HR ECG. Similarly, the five second samples of spatially enhanced HR ECG is used to obtain its classification accuracy. In Fig. 4.7, classification accuracy of the original and the spatially enhanced HR ECG between different MI cases is shown. The classification between ILT and ASP shows a very good accuracy for both original and spatially enhanced ECG. The least classification accuracy is observed between ASP and ALT cases. It can be observed from Fig. 4.7 that most of the cases posses a good classification accuracy between the original and the reconstructed ECG.

TH-1915_11610237

Table 4.3: ECG diagnosability (%) and relative error (%) (inside brackets).

	ANT	**IPL**	**ASP**	**INF**	**ALT**
ILT	97.53 (2.13)	96.78 (2.63)	99.01 (0.89)	98.81 (0.89)	100 (0)
ANT	- -	96.97 (2.63)	97.33 (2.13)	98.72 (1.06)	87.88 (9.30)
IPL	- -	- -	100 (0)	96.67 (2.63)	100 (0)
ASP	- -	- -	- -	100 (0)	96.60 (2.33)
INF	- -	- -	- -	- -	98.59 (1.16)

The ECG diagnosability of the proposed model for various MI cases along with relative error is given in Table 4.3. For these MI cases, it is observed that, the diagnosability is above 96% except between ANT and ALT which is 87.88%. The relative error is also very low for all cases except between ANT and ALT. The ECG diagnosability is 100% and relative error is zero between ILT-ALT, IPL-ASP and IPL-ALT. It is evident from Fig. 4.7 and Table 4.3 that the diagnostic quality of spatially enhanced HR ECG is preserved while using the proposed model. Also, the ECG diagnosability is not affected by the variation in P-amplitude and T-shape for the MI cases discussed in this section.

4.7 Comparison with the existing models

Comparison of the proposed models are performed with the existing models and is shown in Table 4.4. The comparison is performed only with those models that use a subset of the standard twelve-leads as the predictor leads. This includes linear regression (LinR) model [46], independent component analysis (ICA) based model [91], principal component analysis (PCA) based model [8], support vector regression (SVR) based model [100, 101], artificial neural network (ANN) based model [99], state space (SS) model [92] and the lead selective linear regression (LSLinR) model [24]. All models used for comparison are pre-processed using the procedure described in section 4.4, implemented on the same platform and are evaluated on the same database. It is visible from Table 4.4 that the first method gives improved performance over the most popular linear regression model. This model

TH-1915_11610237

Table 4.4: Comparison with the existing models

Method	No. of Leads Used	Leads Used for Prediction	r_x	$RMSE(mV)$	$WEDD(\%)$
LinR[46]	3	*Lead I*, *Lead II*, V_2	0.96	88.69	26.03
ICA[91]	3	*Lead I*, *Lead II*, V_2	0.93	95.56	26.99
PCA[8]	8	*Lead I*, *Lead II*, V_1 to V_6	0.95	53.66	16.07
SVR[100, 101]	4	*Lead I*, *Lead II*, V_1 and V_6	0.97	94.74	26.64
ANN[99]	3	*Lead I*, *Lead II*, V_2	0.98	45.72	15.01
SS[92]	1 to 3	Best of *Lead I*, *Lead II*, *Lead III* or its combination	0.83	185.45	51.40
LSLinR[24]	3	*Lead I*, *Lead II* and best of V_1 to V_6	0.98	63.88	18.97
Proposed1	4	*Lead I*, *Lead II*, V_1 and V_3	0.96	83.43	23.95
Proposed2	3	*Lead I*, *Lead II*, V_3	0.98	54.26	15.70

TH-1915_11610237

posses an average correlation of 0.96, $RMSE$ of 83.43 mV and $WEDD$ of 23.95%. Also, from Table 4.4, it can be observed that correlation coefficient for the second model is 0.98 which is same as ANN and LSLinR models, and better than the other models. The $RMSE$ of the proposed model is 54.26 mV which is slightly higher than ANN and PCA based models. A similar result is observed with $WEDD$ values as well. In the case of the PCA based model, eight independent leads are used at the time of learning and reconstruction which automatically improves the results. From Table 4.4, it is observed that the ANN based model performs better among all models, but is time-consuming at the time of learning. As both the proposed methods uses low dictionary size, learning time can be reduced. It is also evident that the values of correlation coefficient, $RMSE$ and $WEDD$ of the proposed models are comparable with the existing models. Thus from Table 4.4, it is clear that the performance of the proposed models in preserving diagnostic information is analogous or better than that of the prevalent models.

4.8 Summary

Two novel approaches for enhancing the spatial resolution of a LR ECG are proposed and evaluated. Both approaches utilize a joint dictionary mechanism to convert a LR ECG to a HR ECG. In case of the first approach, a conversion function is learned which is then used to transform the LR sparse coefficients to the HR sparse coefficients. For the second approach, the ECG signal is partitioned into multiple segments and a joint dictionary is learned for each segment. In this case, the conversion function is learned for each segment which transforms the LR sparse coefficients of one segment to its HR version. The analysis of the proposed approaches using correlation coefficient, $RMSE$, $WEDD$ and WDD show that segmenting the data ensures good reconstruction quality. The evaluation of ECG diagnosability using selected MI cases shows that the proposed multiple joint dictionary learning approach helps capture significant diagnostic information. Rating of the proposed approaches is performed by comparing it with existing techniques. The capability of the proposed approach to create a HR ECG from its LR subset without compromising diagnostic information content is evident from the analysis of the results. The proposed approach improves the reconstruction accuracy along with an increase in the execution time.

TH-1915_11610237

5

Exploiting Spatio-Temporal Correlations using RNN Models

Contents

TH-1915_11610237

The approaches discussed in chapters 3 and 4 utilized only the inter-lead or spatial correlation between the standard twelve-leads in the transformed domain for learning the models. The twelve standard leads of ECG posses not only inter-lead correlation but also intra-lead correlation [23]. Learning a model that can exploit this spatio-temporal information in ECG could generate lead signals without losing significant diagnostic information. In this chapter, an approach for enhancing the spatial resolution of ECG by exploiting its spatio-temporal correlation in the hidden feature space is discussed.

Preserving the quality of spatially enhanced ECG is important as it contains significant diagnostic information. Among various available models, artificial neural network (ANN) based model by Atoui *et al.* provides the best reconstruction accuracy [99]. The derived ECG approaches discussed in chapter 2 mainly exploits the spatial or inter-lead correlations only. A model capable of simultaneously learning the spatio-temporal correlation could represent the intra-lead and inter-lead correlations in a better way. Such a model might improve the reconstruction accuracy thereby preserving the diagnostic information in lead signals. Advancements in machine learning have opened up a whole new level of learning technique using the deep learning framework. The learning capability of deep neural network (DNN) models is found to be superior to traditional ANN-based models. Hence, the reconstruction quality of derived ECG can be improved further in the deep learning framework. The recurrent neural network (RNN) is a deep learning architecture specifically designed for predicting time series data and hence can be used for generating the spatially enhanced ECG. RNN can combine information from the present and previous inputs to decide the present output. This capability of RNN can be utilized to capture the intra-lead and inter-lead correlation in the twelve-leads. In this chapter, the capability of RNN architecture in improving the spatial resolution of ECG is explored. The models are learned using different variations of RNN such as simple RNN, long short-term memory (LSTM) unit and gated recurrent unit (GRU) [127, 128, 129]. These methods are compared to each other using standard diagnostic closeness measures and are also compared with the existing models.

This chapter is organized as follows: In section 5.1, the spatio-temporal correlation in ECG signal and methods for exploiting it are discussed. The proposed approach for enhancing the spatial resolution of ECG is presented in section 5.2. In section 5.3, the results and its analysis are discussed.

TH-1915_11610237

ECG diagnosability of the proposed model is evaluated in section 5.4. The performance of the proposed models are compared with the existing methods in section 5.5 and the summary of the work is discussed in section 5.6.

5.1 Spatio-temporal correlation in ECG

A new patient-specific approach for predicting the standard twelve-lead ECG from its subset using deep learning techniques is discussed in this section. The basis for this approach is the volume-conductor theory by Burger and van Milaan [5, 6]. According to this theory, a single dipole source fixed at the center of the heart can be used to model the heart dipole or the heart vector. Using Burger's equation, the heart vector can be approximated as a linear combination of three orthogonal components [5, 6, 16]. Ideally, any three orthogonal leads can be used to obtain the heart vector and by using it any leads can be generated by projecting the vector onto the human torso.

The standard twelve-lead ECG consists of eight independent leads of which any three orthogonal or semi-orthogonal leads can be used to obtain the heart vector. The ECG is composed of clinically significant diagnostic features which are important in diagnosing cardiac disorders. This makes it extremely important to preserve these features for applications involving lead reconstruction. The clinically significant features in ECG consist of P-wave, QRS-complex and T-wave. These features are different from each other in terms of shape, amplitude and frequency [2]. These features in ECG posses good intra-lead and inter-lead correlation [2, 24, 23]. The lead signals are spatially aligned and are generated from the same source. Hence, good inter-lead correlation can be found between spatially aligned samples of the lead signals. The cardiac cycle is quasi-periodic in nature and hence good intra-lead correlations can be established between different morphological features. This intra-lead and inter-lead correlations can be exploited to learn models that can improve the spatial resolution of ECG. Incorporating temporal and spatial information into the models that generate standard twelve-lead ECG from its subset is hence beneficial. RNN is a popular deep learning architecture that is capable of embedding spatial and temporal information. In this work, three variations of RNN consisting of simple RNN, LSTM and GRU are used to learn the spatio-temporal correlation in ECG for improving its spatial resolution.

TH-1915_11610237

5.2 Proposed approach for exploiting spatio-temporal correlations using RNN

Diagnostic quality of derived leads using neural network based model is reported as the best among similar approaches [99]. This is achieved by exploiting the spatial or inter-lead correlation only. A model that can use the learning capabilities of ANN and represent the spatio-temporal correlations simultaneously can further improve the diagnostic quality of derived ECG. RNN models are capable of capturing the intra-lead and inter-lead correlation of ECG and can utilize this information to improve its spatial resolution. This is achieved by connecting present inputs with previous decisions using a feedback mechanism. This facilitates the RNN with memory which differentiates it from other feed-forward architectures. The memory is updated with new information and helps the network to accumulate information over time. This is achieved with the help of hidden states that share weights over time. Weights are updated using a method that allows the error to propagate from one time-step to another. This is known as back-propagation through time (BPTT) and is an extended version of the back-propagation algorithm. Using BPTT, the temporal and spatial correlations in different leads of ECG can be integrated into the model.

A subset of the standard twelve-lead ECG is provided as input into the RNN model for obtaining the derived leads. The selection of the subset is based on lead-theory and volume-conductor theory discussed in chapter 2 [5, 6, 16]. The three-lead subset used in this work consists of the independent leads $Lead\ I$, $Lead\ II$ and V_3. The frontal plane leads can be perfectly modeled by linearly combining $Lead\ I$ and $Lead\ II$ and are hence selected [2]. Based on the results observed in [24] and other experimental results, lead V_3 is selected as the third lead in the subset. As the frontal plane leads can be perfectly reconstructed, the aim is to derive the precordial leads V_1, V_2, V_4, V_5 and V_6. Each lead signal is pre-processed before learning and testing the model for removing signal artifacts. The pre-processing includes elimination of the baseline wandering effect, amplitude normalization and mean removal. The baseline wandering effect is eliminated by subtracting the low-pass filtered signal from its original. The cut-off frequency of low-pass filter is set at 0.67 Hz over the assumption that the diagnostic information is negligible below this frequency and hence can be discarded [119].

Fig. 5.1 shows the schematic representation of the proposed approach and can be divided into four stages. The first stage deals with the arrangement of input data before feeding it into the network.

TH-1915_11610237

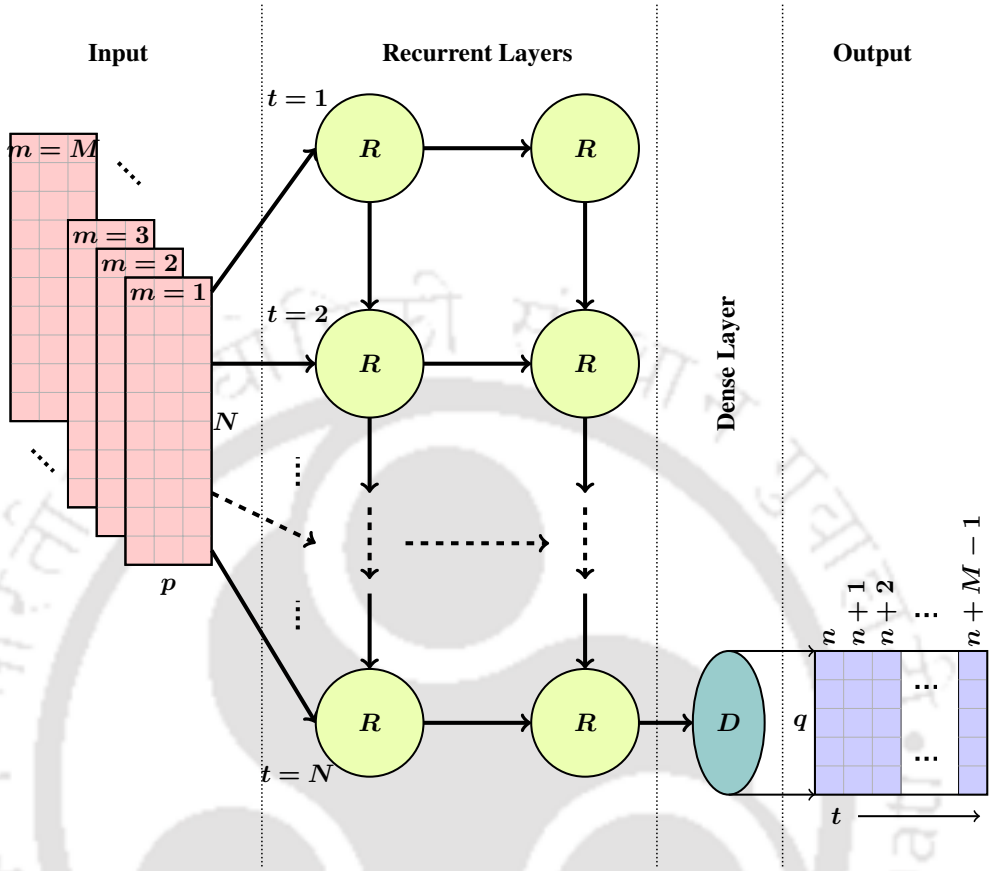


Figure 5.1: Schematic presentation of a multi-layer RNN for enhancing the spatial resolution of ECG.

Input consisting of three predictor leads is divided into mini-batches for training. A window of length equal to the number of time-steps is moved sample by sample to create the mini-batches. The number of mini-batches depends upon the number of samples used for training. As shown in the first stage of Fig. 5.1, the input data can be arranged in the shape of the number of mini-batches (M) $\times$ number of time-steps (N) $\times$ size of input vector (p). Size of input vector p is three since three predictor leads are used. Input data is fed into the second stage which consists of two RNN layers with one stacked over the other. This improves the capability of the model in learning the intra-lead and the inter-lead correlations of ECG. Input, divided into mini-batches from the previous stage is fed into the first recurrent layer one by one. In Fig. 5.1, $\boldsymbol{R}$ indicates the RNN unit. Dimensionality of the first RNN layer is set at fifteen. This means that input vector is mapped into a higher dimensional hidden space of size fifteen in the first layer. The output of the first layer is then mapped to into the next hidden space of size five. The RNN can operate in three different modes which are, one to many mode, many

TH-1915_11610237

to one mode and many to many mode. The first two layers are connected in many to many mode for preserving the temporal information. The output of the second RNN layer is connected to the next stage which is the fully connected layer or the dense layer. The dense layer is used to connect the RNN feature space with the output. In this case, many to one mode is used for connecting the second RNN layer with the dense layer. In this way, the temporal dimension is removed before passing the features into the third stage. The fully connected layer learns a non-linear mapping between output features of the recurrent layer and actual output. The last part is the output stage where the spatially enhanced ECG is generated. As the temporal dimension is removed before the third stage, the size of the output vector (q) is five, which corresponds to the five precordial leads. During training, the first mini-batch is mapped to the output vector at time N, second to N + 1, third to N + 2, etc. Since the RNN works in many to one mode, it requires N time-steps as input to predict the first output. This causes a delay in generating the first output vector. Hence during the reconstruction phase, the first output vector is predicted at time N, second at N + 1, third at N + 2, etc. It is to be noted that the delay does not shift the output vector and N^{th} output vector is time-aligned with the N^{th} input vector.

Three variations of RNN are used to model the input-output relationship between the predictor and the spatially enhanced ECG. The first model is a simple RNN model which uses a truncated version of BPTT for learning. Since multiple matrix multiplications are involved in relating layers and time-steps, truncated BPTT is preferred to reduce the computational complexity. In the simple RNN model, the preserved information is combined with the present input for determining the response to the current input. This can be mathematically represented as

$$\boldsymbol{g_t} = \boldsymbol{\phi}\left(\boldsymbol{W}\boldsymbol{x_t} + \boldsymbol{U}\boldsymbol{g_{t-1}}\right) \quad . \tag{5.1}$$

where $\boldsymbol{g_t}$ and $\boldsymbol{x_t}$ are the hidden state and input at time t, $\boldsymbol{g_{t-1}}$ is the hidden state at $t-1$, $\boldsymbol{W}$ is the weight matrix and $\boldsymbol{U}$ is the hidden state to hidden state transformation matrix. The function $\boldsymbol{\phi}$ can be sigmoid or tanh depending on the requirement. It can be noted from equation 5.1 that the hidden state can have elements of information with respect to the past hidden states as well. The weight matrices determine the amount of present and previous information required for determining the current output. The weights are updated using truncated BPTT. The downside of using truncated BPTT is that long-term dependencies may not be properly represented. Also, in simple RNN model,

TH-1915_11610237

the gradient can diminish exponentially and can vanish after few time-steps. This is a common problem for simple RNN known as the vanishing gradient problem and can cause the model to cease learning the long-term dependencies [127].

The second model used is the the long-short term memory (LSTM) network which solves the vanishing gradient problem [128]. In an LSTM network, the error information is preserved and is back-propagated through the layers and times-steps which essentially helps the model to learn over a large number of time-steps. This is achieved with the help of a gated cell which allows reading, writing and storage of information. With the help of interacting layers inside the gated cell, it can decide on what to read, write, store or delete. The gates are implemented using point-wise multiplication using a sigmoid. There are three main gates that controls the LSTM cell state: input gate, forget gate and output gate. The input gate decides on the information to be stored, the forget gate decides on the information to be deleted and the output gate decides on the information to be passed. At time t, the LSTM cell receives information about the previous cell state ($\boldsymbol{S^c_{t-1}}$), the previous hidden state ($\boldsymbol{g_{t-1}}$) and the current input ($\boldsymbol{x_t}$). The working of an LSTM cell is as follows. The first step is to decide on the amount of information to be thrown out of the cell state. This is decided by the forget gate which is defined as

$$\boldsymbol{f_t} = \boldsymbol{\sigma}\left(\boldsymbol{W_f x_t} + \boldsymbol{U_f g_{t-1}}\right) \ . \tag{5.2}$$

where $\boldsymbol{x_t}$ and $\boldsymbol{g_{t-1}}$ are the input and hidden state at time t and $t-1$ respectively. The matrices $\boldsymbol{W_f}$ and $\boldsymbol{U_f}$ are the weight matrices for the forget gate and $\boldsymbol{\sigma}$ is a sigmoid in the range 0 and 1. The forget gate $\boldsymbol{f_t}$ contributes a value between 0 and 1 to the cell state suggesting the percentage of information to be kept. The cell state is updated by multiplying the output of forget gate $\boldsymbol{f_t}$ with previous cell state $\boldsymbol{S^c_{t-1}}$. The next step is to select the new information to be stored and is decided by the input gate. This is a two stage process consisting of an input gate ($\boldsymbol{i_t}$) and a hidden state known as the candidate state ($\boldsymbol{\tilde{S}^c_t}$). Both $\boldsymbol{i_t}$ and $\boldsymbol{\tilde{S}^c_t}$ depends upon the present input and the previous hidden state and is defined as

$$\boldsymbol{i_t} = \boldsymbol{\sigma}\left(\boldsymbol{W_i x_t} + \boldsymbol{U_i g_{t-1}}\right) \ . \tag{5.3}$$

$$\boldsymbol{\tilde{S}^c_t} = \tanh\left(\boldsymbol{W_c x_t} + \boldsymbol{U_c g_{t-1}}\right) \ . \tag{5.4}$$

TH-1915_11610237

where $\boldsymbol{W_i}$ and $\boldsymbol{U_i}$ are the weight matrices corresponding to the input gate and $\boldsymbol{W_c}$ and $\boldsymbol{U_c}$ are the weight matrices corresponding to the candidate state. The input gate decides over update values and candidate state creates a new set of values which may be added to the cell state. These two are combined together by multiplying and is used for updating the current cell state $\boldsymbol{S_t^c}$. The new cell state is defined as

$$\boldsymbol{S_t^c} = \boldsymbol{f_t} \circ \boldsymbol{S_{t-1}^c} + \boldsymbol{i_t} \circ \boldsymbol{\tilde{S}_t^c} \quad . \tag{5.5}$$

The output $\boldsymbol{f_t}$ decides upon the amount of previous information in the cell state to forget by multiplying it with the previous cell state $\boldsymbol{S_{t-1}^c}$. This is added with $\boldsymbol{i_t} \circ \boldsymbol{\tilde{S}_t^c}$ which is nothing but the new information to be updated which generates the current cell state $\boldsymbol{S_t^c}$. The operator $\circ$ indicates point-wise multiplication. The output gate is used to pass the updated information to the next cell. The cell state is scaled down using the output gate to obtain the present state which is defined as

$$\boldsymbol{o_t} = \boldsymbol{\sigma}\left(\boldsymbol{W_o x_t} + \boldsymbol{U_o g_{t-1}}\right) \quad . \tag{5.6}$$

$$\boldsymbol{g_t} = \boldsymbol{o_t} \circ \tanh\left(\boldsymbol{S_t^c}\right) \quad . \tag{5.7}$$

where $\boldsymbol{W_o}$ and $\boldsymbol{U_o}$ are the weight matrices for the output gate. The output gate $\boldsymbol{o_t}$ decides on the amount of cell state information to be passed as output. This is multiplied with cell state $\boldsymbol{S_t^c}$ to obtain $\boldsymbol{g_t}$ which is the hidden state output at time t. In this way the gated cell structure of LSTM network act as a memory cell which maintains the cell sate over time.

The third model used is a variation of LSTM network known as the gated recurrent unit (GRU) [129]. The GRU can also capture long-term dependencies like LSTM. The difference is that GRU consists of only two gates, an update gate and a reset gate when compared with LSTM. The update gate determines the amount of previous information to be kept and the reset gate combines the previous information with new information. Since only two gates are involved, GRU can be trained faster than LSTM. The update and reset gates are defined as

$$\boldsymbol{z_t} = \boldsymbol{\sigma}\left(\boldsymbol{W_z x_t} + \boldsymbol{U_z g_{t-1}}\right) \quad . \tag{5.8}$$

$$\boldsymbol{r_t} = \boldsymbol{\sigma}\left(\boldsymbol{W_r x_t} + \boldsymbol{U_r g_{t-1}}\right) \quad . \tag{5.9}$$

TH-1915_11610237

where $\boldsymbol{W_z}$, $\boldsymbol{U_z}$, $\boldsymbol{W_r}$ and $\boldsymbol{U_r}$ are the weight matrices for the update gate and the reset gate. The reset gate combines the functions of input and forget gate in the LSTM. The GRU does not maintain a cell state like LSTM, but instead a current memory state is computed using the reset gate. This step decides on the amount of relevant information to be stored and is defined as

$$\tilde{\boldsymbol{g}}_{\boldsymbol{t}} = \tanh\left(\boldsymbol{W_h x_t}\right) + \boldsymbol{U_h}\left(\boldsymbol{g_{t-1}} \circ \boldsymbol{r_t}\right) \quad . \tag{5.10}$$

where $\boldsymbol{W_h}$ and $\boldsymbol{U_h}$ are the weight matrices of the current memory state. The current state $\boldsymbol{g_t}$ is defined as

$$\boldsymbol{g_t} = (1 - \boldsymbol{z_t}) \circ \boldsymbol{g_{t-1}} + \boldsymbol{z_t} \circ \tilde{\boldsymbol{g}}_{\boldsymbol{t}} \quad . \tag{5.11}$$

Thus, the GRU is able to store and use relevant information over time. All models are implemented following the same procedure as shown in Figure 5.1.

5.3 Results and discussions

The Physikalisch-Technische Bundesanstalt (PTB) diagnostic ECG database is used for evaluating the performance of the models [19][20]. The details of the database is given in section 2.3. The first thirty seconds of each record is used in this evaluation. Training is performed using the first twenty seconds and testing/reconstruction is performed using the last five seconds. All three models are implemented using the same frame structure as shown in Fig. 5.1. The training data is divided into training and validation set following an 80:20 ratio. The number of time-steps can be selected depending upon the amount of temporal information required for a good prediction. The bottleneck is that increasing the number of time-steps increases complexity of the model and it requires huge computational power. Considering these facts, the number of time-steps (N) is fixed at 16 for all the three models. Also, the size of the hidden layers are selected experimentally using trial and error method by considering the computational complexity and reconstruction accuracy. The models are trained by minimizing the mean absolute error using adam optimizer. The models are implemented in keras using the tensorflow backend [102, 103].

TH-1915_11610237

5.3.1 Performance evaluation

The output of the proposed models are compared with the original signals for evaluating their performance. Standard measures such as correlation coefficient and $RMSE$ are used to compute the similarity. The $WEDD$ and WDD measures are used to evaluate the diagnostic closeness between the original and the derived lead signals [18, 104]. Selected diagnostically significant features of the original and the reconstructed lead signals are also compared. ECG diagnosability of the proposed models is also computed by following the procedure described in section 3.4. The relative error between Accuracy$_o$ and Accuracy$_d$ defined in section 3.4 is also computed for the proposed models.

5.3.2 Results of proposed RNN models

The correlation coefficient (r_x), $RMSE$ and $WEDD$ of simple RNN, LSTM and GRU based models are shown in Table 5.1. The three models used the same network structure for fair comparison. It can be observed that the predicted leads using all three models show a very good correlation with the original leads. Highest average correlation of 0.98 is observed with the LSTM model. Leads V_1 and V_5 of simple RNN and GRU models show the lowest correlation of 0.96 and 0.95 respectively. The average $RMSE$ is 70.89 mV for the simple RNN model, 52.95 mV for the GRU model and 42.11 mV for the LSTM model. Lead V_1 provides the highest $RMSE$ values for all three models. The average $WEDD$ is highest for the simple RNN model at 20.65%, second highest for the GRU model at 16.06% and lowest for the LSTM model at 12.50%. It is also observed from Table 5.1 that lead V_1 shows maximum $WEDD$ for all models. The $RMSE$ and $WEDD$ is low for leads nearer to the predictor lead V_3 and is high for those away from it. A similar observation can be made for correlation coefficient as well. This is because of the proximal lead effect where the leads nearer to the predictor lead are reconstructed better than the ones away from it. From Table 5.1, it is evident that the LSTM model performs much better than the other two models. Performance of the simple RNN model is inferior to LSTM and GRU models. As explained earlier, simple RNN models are vulnerable to vanishing gradient problem which affects the performance of the model. Performance of the GRU model is satisfactory and require less time for learning when compared to LSTM model. The model parameters are fixed across all models. It can be optimized for individual models which may further improve its performance.

TH-1915_11610237

Table 5.1: Performance assessment of the proposed models for all patient records.

	Simple RNN			**LSTM**			**GRU**		
	r_x	$RMSE(mV)$	$WEDD(\%)$	r_x	$RMSE(mV)$	$WEDD(\%)$	r_x	$RMSE(mV)$	$WEDD(\%)$
V_1	0.96	81.83	23.67	0.98	51.35	15.42	0.95	65.09	19.68
V_2	0.97	69.25	17.56	0.98	43.74	11.48	0.98	45.83	12.24
V_4	0.97	65.17	18.36	0.99	35.78	10.01	0.98	42.22	12.22
V_5	0.96	71.94	22.64	0.98	39.88	12.75	0.95	58.79	19.08
V_6	0.97	66.27	21.04	0.98	39.79	12.85	0.97	52.82	17.10
Avg.	0.97	70.89	20.65	0.98	42.11	12.50	0.97	52.95	16.06

TH-1915_11610237

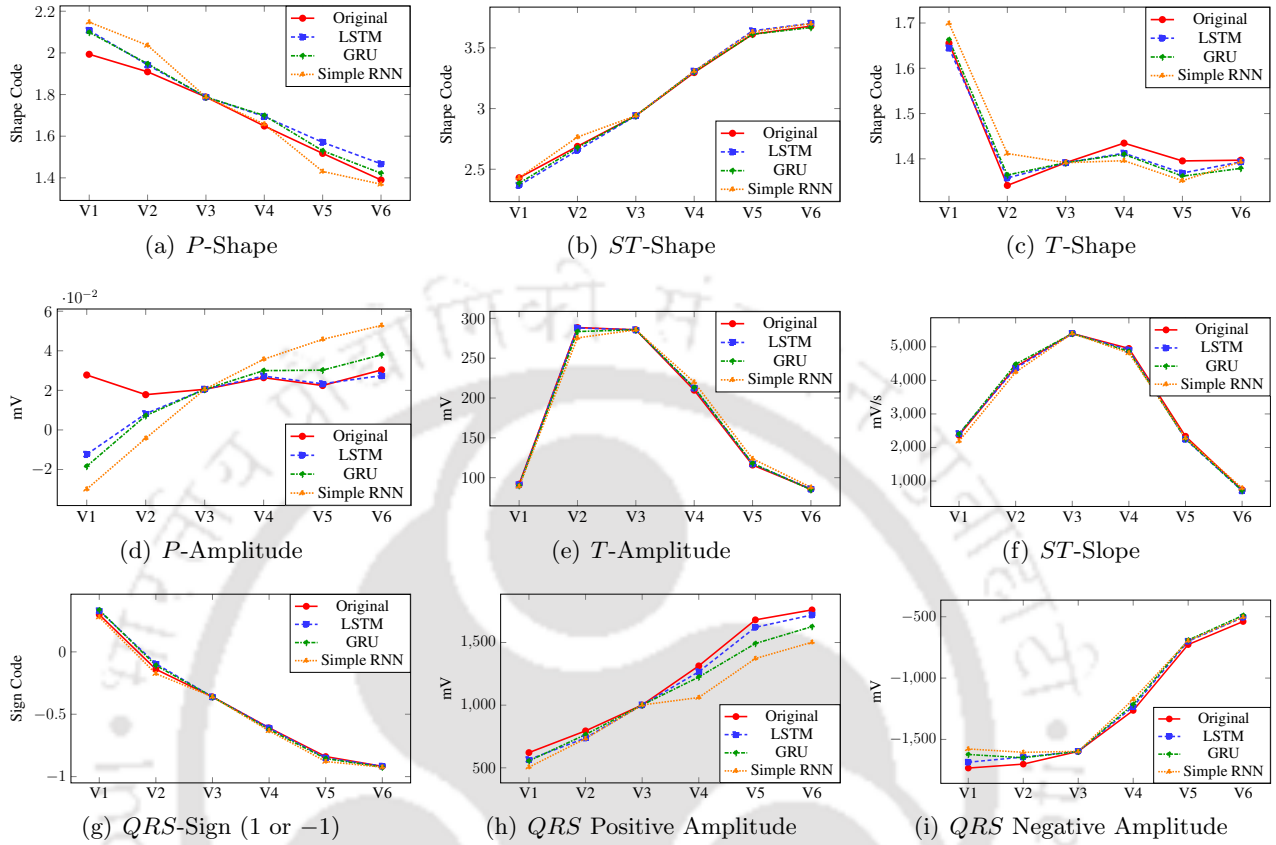
Table 5.2: WDD of the proposed models.

	Simple RNN	**LSTM**	**GRU**
$\boldsymbol{V_1}$	13.85	11.02	12.15
$\boldsymbol{V_2}$	8.70	6.72	6.63
$\boldsymbol{V_4}$	8.22	5.04	5.69
$\boldsymbol{V_5}$	9.61	6.87	8.57
$\boldsymbol{V_6}$	9.58	7.32	8.30
Avg.	9.99	7.39	8.27

The WDD is calculated for selected number of data and is shown in Table 5.2. The total number of records used are 524 and average WDD values for all leads are computed. It is evident from Table 5.2 that the LSTM model performs much better than the other two models. The WDD values of all leads except V_1 show that reconstructed leads fall in the good category. Lead V_1 is slightly distorted for all models because of the proximity lead effect. From the WDD values shown in Table 5.2, it is clear that the models preserve the diagnostically significant features while reconstructing the lead signals.

An analysis at the feature level is also performed for further quantification and is shown in Fig. 5.2. Nine features consisting of the shape and the amplitude features are analyzed for all three models. The average values of these nine features are calculated for each precordial lead. Also, 524 of 549 records are used for analysis. The reconstructed lead signals are temporally aligned with the original lead signals. Hence the duration features between the original and the derived lead signals are excluded from the analysis. The shape features consisting of P-shape, ST-shape and T-shape are used for analysis. The shape features are assigned a code while calculating WDD and the same is used for this analysis [18]. In Fig. 5.2(a), the P-shape shows variation from the original lead signal in V_1 for all three models. The LSTM and GRU models show good similarity and follows the shape in other precordial leads, while the simple RNN model show variation in leads V_1, V_2 and V_5. The LSTM and GRU models show very good similarity for all leads in case of ST-shape given in Fig. 5.2(b). A slight deviation is shown by lead V_2 by the simple RNN model. The LSTM and GRU models reproduced the T-shape in leads V_1 and V_2 and is shown in Fig. 5.2(c). Slight deviation is observed in leads V_4, V_5 and V_6. The simple RNN model showed maximum variation from the original shape for all leads.

TH-1915_11610237



(a) P-Shape (b) ST-Shape (c) T-Shape

(d) P-Amplitude (e) T-Amplitude (f) ST-Slope

(g) QRS-Sign (1 or -1) (h) QRS Positive Amplitude (i) QRS Negative Amplitude

Figure 5.2: Comparison of shape and amplitude features between original leads and reconstructed leads.

The amplitude features selected for this analysis consist of P-amplitude, T-amplitude, ST-slope, the sign of the first peak in the QRS-complex (1 or -1), QRS-positive amplitude and QRS-negative amplitude. In Fig. 5.2(d), it is observed that all models fail to capture the P-amplitude in leads V_1 and V_2. The simple RNN and the GRU models failed to capture this information in other leads as well, while the LSTM model reproduced this feature in leads V_4 to V_6. It is also evident from Fig. 5.2(d) that the variation is less because of the low dynamic range of P-amplitude. Since the dynamic range of other morphological features are high, the models fail to capture the low amplitude of P-wave. Another factor that might have affected the reconstruction is the amount of atrial information available in the leads used for prediction. All models preserved the T-amplitude given in Fig. 5.2(e) with LSTM and GRU models performing better than the simple RNN model. The ST-slope and QRS-sign features shown in Fig. 5.2(f) and 5.2(g) shows the capability of all three models in preserving these features. The QRS-positive amplitude and the QRS-negative amplitude show good similarity

TH-1915_11610237

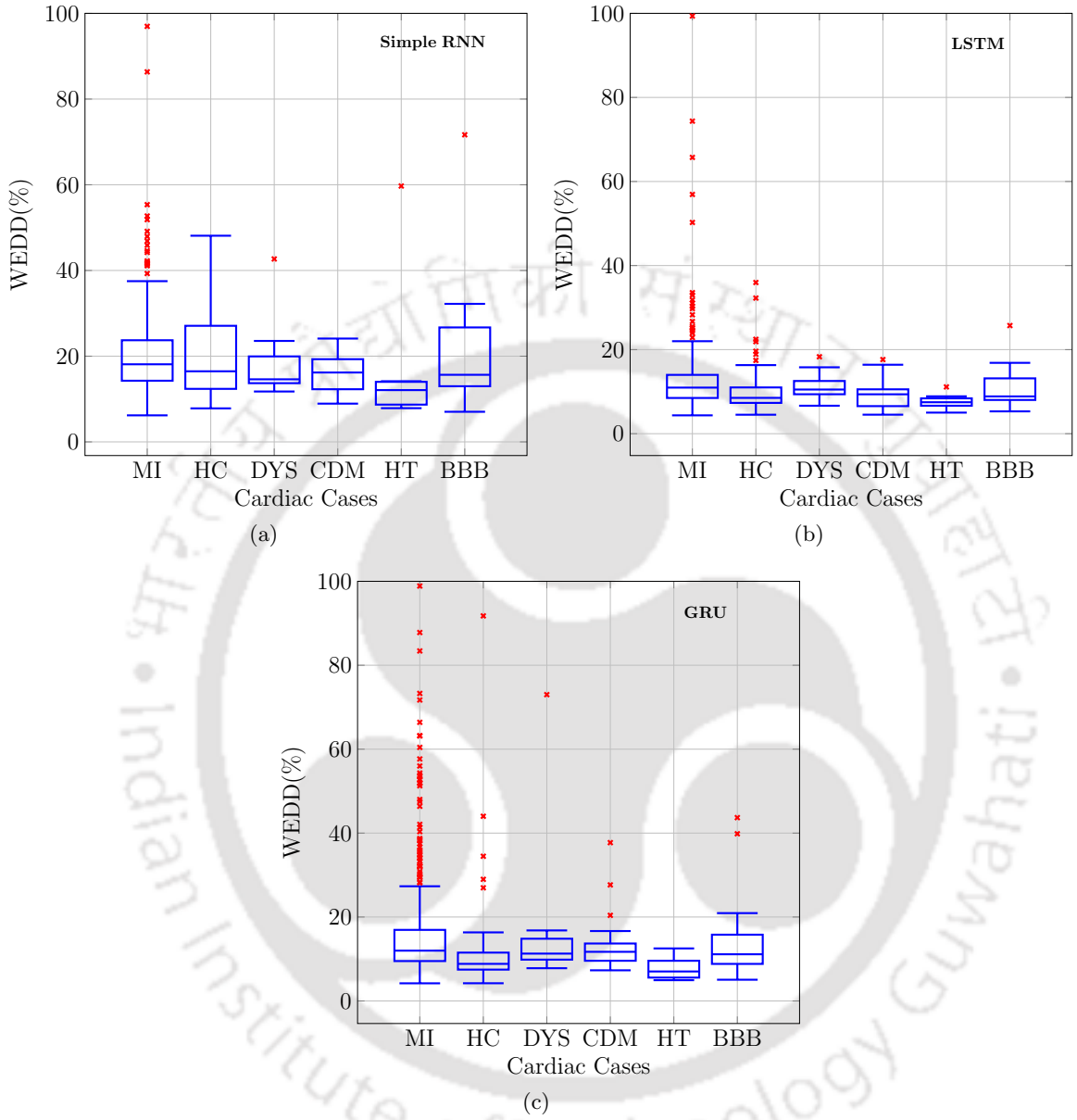


Figure 5.3: Box Plot showing the distribution of $WEDD(\%)$ for the precordial leads of myocardial infarction (MI), healthy control (HC), dysrhythmia (DYS), cardiomyopathy (CDM), hypertrophy (HT) and bundle branch block (BBB) for (a) Simple RNN model; (b) LSTM model;(c) GRU model.

with the original lead signals for LSTM model followed by the GRU model. The simple RNN model shows variation in QRS-positive amplitude for leads V_4 to V_6. From the analysis of these features, it can be concluded that the LSTM model perform much better than the other two models in capturing the diagnostic information. Also, the performance of the GRU model is comparable with the LSTM model.

TH-1915_11610237

The distribution of $WEDD$ for the three models in case of selected classes are shown in Fig. 5.3. Selected classes consist of myocardial infarction (MI), healthy control (HC), dysrhythmia (DYS), cardiomyopathy (CDM), hypertrophy (HT) and bundle branch block (BBB). The distribution of $WEDD$ for the simple RNN model is shown in Fig. 5.3(a). The mean and median (in brackets) $WEDD$ values of these classes for simple RNN model are: MI = 20.69% (18.11%), HC = 20.35% (16.47%), DYS = 17.65% (14.60%), CDM = 16.06% (16.19%), HT = 17.87% (12.08%) and BBB = 20.41% (15.66%). It can be observed from Fig. 5.3(a) that the third quartile is below 25% for MI, HC and BBB classes. In Fig. 5.3(b), the $WEDD$ distribution for LSTM model is shown. For the LSTM model, the mean and the median (in brackets) $WEDD$ values are: MI = 12.69% (10.96%), HC = 10.30% (8.52%), DYS = 11.16% (10.51%), CDM = 9.76% (9.37%), HT = 7.68% (7.48%) and BBB = 10.91% (8.87%). Similarly from Fig. 5.3(c), the mean and median (in brackets) $WEDD$ values for the GRU model is observed as: MI = 16.97% (11.98%), HC = 11.47% (8.84%), DYS = 15.52% (11.28%), CDM = 14.17% (11.71%), HT = 7.83% (7.02%) and BBB = 14.84% (11.12%). The mean and median of all records are 20.65% and 17.82% for the simple RNN model, 12.50% and 10.44% for the LSTM model and 16.06% and 11.48% for the GRU model. It can be observed from Fig. 5.3(a) that for the simple RNN model, the median and third quartile $WEDD$ values are on the higher side, especially for MI, HC and BBB classes. This implies that the capability of simple RNN model in capturing significant diagnostic information is less for half of the records. The third quartile value of $WEDD$ for the GRU model is less than the median $WEDD$ value of simple RNN model. This is evident in Fig. 5.3(c) and hence the GRU model captures the diagnostic information much better than the simple RNN model for the majority of the records. The best performance is observed for the LSTM model which is clear from Fig. 5.3(b). The third quartile $WEDD$ value is below 13% for these classes and is 13.46% for all records. This indicates that the significant diagnostic information is preserved in the majority of the records for the LSTM model.

In Fig. 5.4(a) and 5.4(b), original and derived precordial leads of healthy control and anterior myocardial infarction (MI) records derived using the LSTM model are shown. Records are selected based on median of $WEDD$ for the respective classes. Fig. 5.4(a) shows the original and the derived precordial leads of healthy record 's0305lrem'. A common observation from the figure is that the noise is reduced for all reconstructed leads. The QRS-complex is slightly underestimated for leads V_1, V_2

TH-1915_11610237

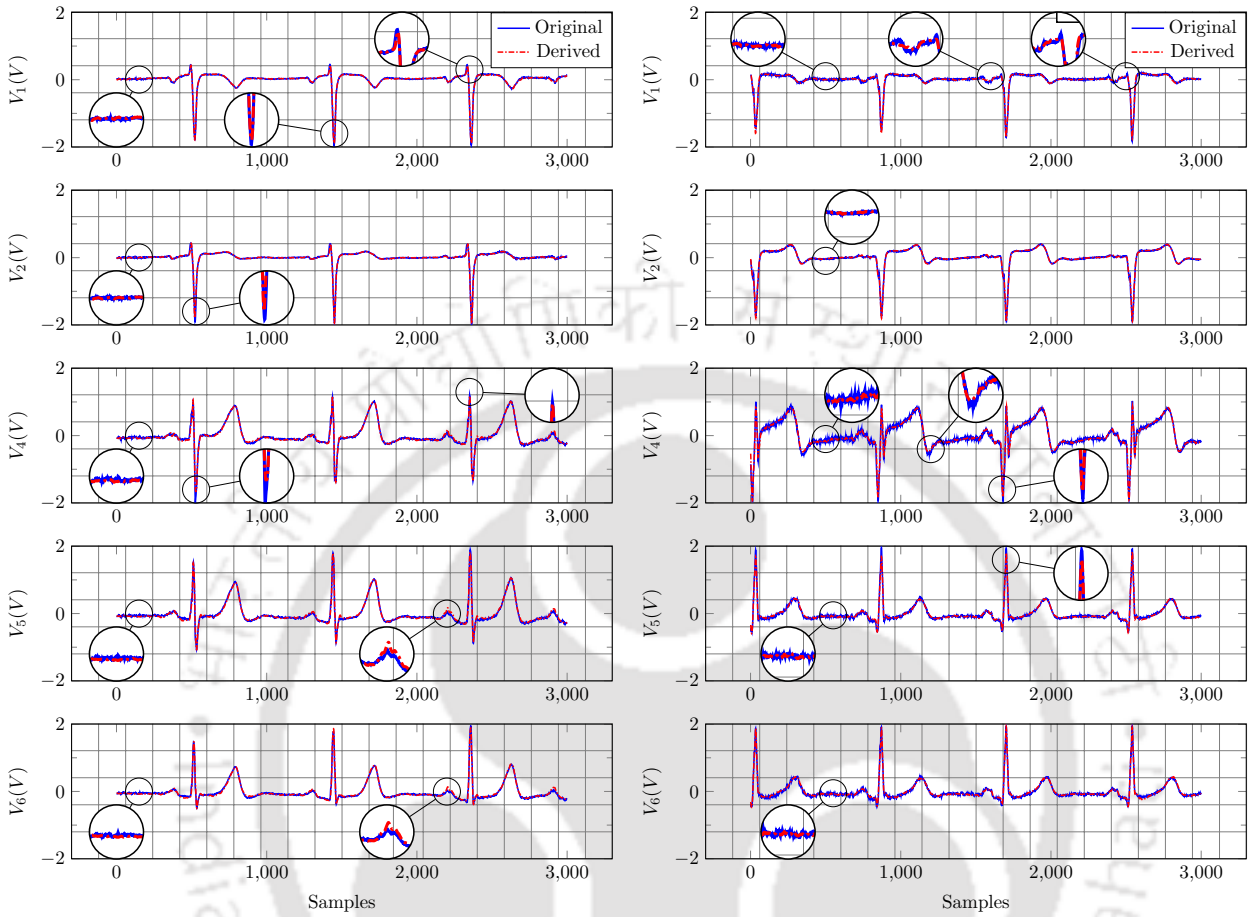


Figure 5.4: (a) Original (blue) and reconstructed (red, dash-dotted) precordial leads of healthy control record 's0305lrem' for LSTM model; (b) Original (blue) and reconstructed (red, dash-dotted) precordial leads of anterior MI record 's0112lrem' for LSTM model.

and V_4, especially the R-peak and the S-peak. The P-wave is slightly overestimated for leads V_5 and V_6. The overall quality of the reconstructed leads are observed to be satisfactory and follows the original signal. The original and the reconstructed leads of anterior MI record 's0112lrem' is shown in Fig. 5.4(b). In this case also, it can be observed that the model reduces the noise while reconstructing the leads. The QRS complex is slightly underestimated for leads V_1, V_4 and V_5. Also, a slight variation in P-wave and T-wave is observed for leads V_1 and V_4 respectively. It can also be observed from Fig. 5.4(b) that the clinically important leads (V_2, V_4 and V_5) for anterior MI cases preserve diagnostically significant features. The leads nearer to lead V_3 is observed to provide good reconstruction quality because of the proximal lead effect. From Fig. 5.4(a) and 5.4(b), it is evident

that the spatially enhanced ECG preserves diagnostically important features. The analysis of results suggest that proposed models are capable of preserving significant diagnostic information. Among the three models, the LSTM model showed the best performance followed by the GRU model and the simple RNN model. For fair comparison, these models used the same network structure and are not optimized individually. The performance of the simple RNN and GRU model may improve further by optimizing the model parameters. Also, these results are achieved with only 16 time-steps using the limited processing power of a normal cpu. The execution time is computed experimentally and is found to be 4.72 minutes for simple RNN model, 14.46 minutes for GRU model and 18.31 minutes for LSTM model. The performance of the proposed models are expected to improve if the number of time-steps are increased. Also, the execution time can be brought down by using systems with high processing power.

5.4 ECG diagnosability of the proposed models

ECG diagnosability and relative error of the three models are shown in Tables 5.3, 5.4 and 5.5. Diagnosability evaluates the capability of the models in capturing diagnostic information which are present in the original leads while reconstructing. ECG diagnosability is computed using the method described in section 3.4. Six different MI classes consisting of infero-lateral (ILT) MI, anterior (ANT) MI, infero-postero-lateral (IPL) MI, antero-septal (ASP) MI, inferior (INF) MI and antero-lateral (ALT) MI are classified from one another. The classification accuracy between these classes using the original ECG varies between 75% to 91% except for the ALT-ASP case. The accuracy obtained for ALT-ASP classification is 67.44%.

Table 5.3 shows the ECG diagnosability and the relative error for the simple RNN model. It is observed that the diagnosability is above 93% and the relative error is below 6% for all the cases. The diagnosability is 100% and the relative error is zero for ILT-INF, ILT-ALT, ANT-ALT, IPL-ALT and INF-ALT. It is evident from Table 5.3 that classification using spatially enhanced ECG obtained from the simple RNN model can produce very good accuracies.

In Table 5.4, the ECG diagnosability and relative error of the LSTM model is shown. In this case, the diagnosability is above 96% and the relative error is below 3%. The ECG diagnosability and the relative error between ILT-ANT, ILT-ALT, ANT-IPL, ANT-ASP, ANT-INF, IPL-ALT and INF-ALT are 100%

TH-1915_11610237

Table 5.3: ECG diagnosability (%) and relative error (%) (inside brackets) for Simple RNN model.

	ANT	**IPL**	**ASP**	**INF**	**ALT**
ILT	98.81 (1.06)	93.10 (5.26)	94.17 (5.36)	100 (0)	100 (0)
ANT	- -	96.88 (2.63)	97.44 (2.13)	95.06 (4.26)	100 (0)
IPL	- -	- -	100 (0)	96.67 (2.63)	100 (0)
ASP	- -	- -	- -	99.24 (0.65)	94.83 (3.49)
INF	- -	- -	- -	- -	100 (0)

Table 5.4: ECG diagnosability (%) and relative error (%) (inside brackets) for LSTM model.

	ANT	**IPL**	**ASP**	**INF**	**ALT**
ILT	100 (0)	96.55 (2.63)	99.03 (0.89)	98.85 (0.89)	100 (0)
ANT	- -	100 (0)	100 (0)	100 (0)	97.06 (2.33)
IPL	- -	- -	100 (0)	96.67 (2.63)	100 (0)
ASP	- -	- -	- -	100 (0)	96.55 (2.33)
INF	- -	- -	- -	- -	100 (0)

and zero respectively. The capability of LSTM model to generate spatially enhanced ECG that can classify between different MI classes with good accuracy is evident from Table 5.4.

The ECG diagnosability and relative error of the GRU model is shown in Table 5.5. The diagnosability is above 95% and the relative error is below 4% except that between IPL-INF. For IPL-INF, the ECG diagnosability is 86.67% and the relative error is 10.53%. The diagnosability is 100% and the relative error is zero for ILT-IPL, ILT-ALT, IPL-ASP, IPL-ALT and ASP-ALT. From Table 5.5, it is evident that very good classification accuracies can be obtained using spatially enhanced ECG generated by the GRU model. From Tables 5.3, 5.4 and 5.5, it is clear that maximum ECG diagnosability is produced by the LSTM model for most of the cases. It is also observed that the simple RNN model

TH-1915_11610237

Table 5.5: ECG diagnosability (%) and relative error (%) (inside brackets) for GRU model.

	ANT	**IPL**	**ASP**	**INF**	**ALT**
ILT	98.81 (1.06)	100 (0)	97.09 (2.68)	98.85 (0.89)	100 (0)
ANT	- -	96.88 (2.63)	98.72 (1.06)	96.30 (3.19)	98.53 (1.16)
IPL	- -	- -	100 (0)	86.67 (10.53)	100 (0)
ASP	- -	- -	- -	96.18 (3.25)	100 (0)
INF	- -	- -	- -	- -	95.77 (3.49)

shows better ECG diagnosability than the LSTM model for ILT-ANF and ANT-ALT classification. The simple RNN also performs better than the GRU model for ILT-INF, ANT-ALT, IPL-INF and ASP-INF classification. Similarly, the GRU model shows improved diagnosability values than the LSTM models for ILT-IPL, ANT-ALT and ASP-ALT classification. These observations indicate that the simple RNN model captures the ILT and ALT classes better than the LSTM and the GRU models. Also, the GRU model captures IPL and ALT classes better than the LSTM model.

5.5 Comparison with the existing models

Table 5.6 shows the comparison of the proposed RNN models with existing methods. Comparison is performed only with those models that generate twelve standard leads from its subset. This includes linear regression (LinR) model [46], independent component analysis (ICA) based model [91], principal component analysis (PCA) based model [8], support vector regression (SVR) based model [100, 101], artificial neural network (ANN) based model [99], and state space (SS) model [92]. The comparison is also performed with those models proposed in chapters 3 and 4. This includes lead selective linear regression (LSLinR) model [24], joint dictionary learning model (JDL) and multiple joint dictionary learning model (MJDL). The models shown in Table 5.6 are implemented and tested on the same database. The twelve-leads are pre-processed using the approach described in section

TH-1915_11610237

Table 5.6: Comparison with the existing models

Method	No. of Leads Used	Leads Used for Prediction	r_x	$RMSE(mV)$	$WEDD(\%)$
LinR[46]	3	*Lead I*, *Lead II*, V_2	0.96	88.69	26.03
ICA[91]	3	*Lead I*, *Lead II*, V_2	0.93	95.56	26.99
PCA[8]	8	*Lead I*, *Lead II*, V_1 to V_6	0.95	53.66	16.07
SVR[100, 101]	4	*Lead I*, *Lead II*, V_1 and V_6	0.97	94.74	26.64
ANN[99]	3	*Lead I*, *Lead II*, V_2	0.98	45.72	15.01
SS[92]	1 to 3	Best of *Lead I*, *Lead II*, *Lead III* or its combination	0.83	185.45	51.40
LSLinR[24]	3	*Lead I*, *Lead II* and best of V_1 to V_6	0.98	63.88	18.97
JDL	4	*Lead I*, *Lead II*, V_1 and V_3	0.96	83.43	23.95
MJDL	3	*Lead I*, *Lead II*, V_3	0.98	54.26	15.70
Simple RNN	3	*Lead I*, *Lead II*, V_3	**0.97**	**70.89**	**20.65**
GRU	3	*Lead I*, *Lead II*, V_3	**0.97**	**52.95**	**16.06**
LSTM	3	*Lead I*, *Lead II*, V_3	**0.98**	**42.11**	**12.50**

TH-1915_11610237

5.2 before training and testing. The same set of pre-processed training and testing data is used for evaluating the existing and the proposed models. The correlation and the $RMSE$ values of the GRU and the LSTM models are equal to or less than that of the PCA and ANN models. The proposed models are better than the PCA based model as they require only three leads compared to eight leads. The $WEDD$ value of the LSTM model at 12.50% indicates that it outperforms all the existing models in preserving the diagnostic information. The $WEDD$ value of the GRU based model is 16.06% and is comparable with the PCA and ANN models. The simple RNN model shows better $WEDD$ value than LinR, ICA, SVR and SS models and is comparable with the LSLinR model. From Table 5.6 it is evident that the LSTM model is the best model for enhancing the spatial resolution of ECG. Also, the simple RNN and the GRU models are comparable or better than the existing models.

5.6 Summary

A new method using RNN is proposed for enhancing the spatial resolution of ECG. The RNN framework incorporates the temporal and the spatial correlations between the twelve-leads into the model thereby improving the reconstruction accuracy. Three types of RNN models, simple RNN, LSTM, and GRU, are implemented and their performances are evaluated. Analysis of the results shows that the RNN models are capable of preserving the diagnostically significant features of ECG. The comparison of the proposed models with the existing methods indicates improved performance in enhancing the spatial resolution of ECG. The lowest $WEDD$ value of the LSTM model places it as a better alternative for existing linear and non-linear models. The proposed models are capable of simultaneously exploiting the spatio-temporal correlation for improved reconstruction accuracy, but posses higher execution time. The performance of the models may improve further by using more processing power and time-steps, and also by optimizing the individual model parameters.

TH-1915_11610237

Indian Institute of Technology Guwahati

6

Conclusions

Contents

TH-1915_11610237

In this thesis, three approaches that can enhance the spatial resolution of ECG are investigated. These approaches utilized the inter-lead and intra-lead correlations of ECG signal in the transformed domain for learning the models. The models learned using different machine learning techniques maps the transformed domain features of predictor and derived lead signals. The lead signals generated using the models proposed in this thesis are found to preserve the clinically significant diagnostic features of ECG.

Chapter 1 provides a brief introduction about ECG, its clinical importance and various morphological features. The concept of the derived ECG system and the importance of preserving clinically significant features in such a system are discussed along with the applications. In Chapter 2, the lead theory and its effective utilization for developing derived ECG systems are presented along with different types of lead systems. A review of various derived ECG systems are discussed by splitting it into three categories. In the first category, linear approaches for transforming one type of lead system into another are presented. The linear and non-linear approaches employed for deriving standard twelve-lead ECG from its subset are discussed in second and third categories respectively. Salient observations made from these two chapters are: a) standard twelve-lead ECG posses good spatial and temporal correlation b) necessity of deriving lead signals by preserving clinically significant features c) exploiting spatio-temporal correlations in transformed domain can improve the reconstruction accuracy.

A new patient-specific approach for enhancing the spatial resolution of ECG is proposed in Chapter 3. The model exploits the spatial correlation between the leads in wavelet domain. The high correlation between the predictor and the response lead signals in the same scale are used to learn the linear mapping. The optimal predictor lead set is also obtained by using a lead selection algorithm. The algorithm is based on the diagnostic similarity score which ensures that the model preserves significant diagnostic features in ECG. The bi-orthogonal 9/7 wavelet filters are used to decompose the mean free ECG signal to seven levels. The level is selected by computing the variation in $WEDD$ with respect to the change in decomposition level. The performance of the proposed model is evaluated using standard distortion measures as well as diagnostic distortion measures. The average correlation and $RMSE$ values are 0.98 and 63.88 mV which indicates that the spatially enhanced ECG posses good similarity with original ECG. The average $WEDD$ and WDD values are 18.97%

TH-1915_11610237

and 10.34% respectively. The lower values of $WEDD$ and WDD indicates that the derived leads are capable of preserving the diagnostic information. Feature level analysis shows that the shape and amplitude features are preserved in derived lead signals. From the analysis of repeatability, it is found that the model is able to preserve the diagnostic quality for the same patient at different instances. Diagnosability of the proposed model is performed by computing the classification accuracy between healthy and pathological records. The analysis of diagnosability shows that the derived ECG can classify healthy and patient data with accuracy as good as that of the original ECG. The comparison of the proposed approach is performed with the existing linear models and is found to be better than the other linear models in preserving the diagnostic information.

The second approach discussed in Chapter 4 utilizes the sparse domain and dictionary learning framework for improving the spatial resolution of ECG. The predictor and the response lead signals are transformed into the sparse domain before learning the model. Two models following similar approaches are discussed in this chapter. In this approach, the standard twelve-lead ECG is considered as the high-resolution (HR) version of its low resolution (LR) subset. For the first model, two dictionaries corresponding to the HR and LR ECG are jointly learned using K-SVD algorithm. These dictionaries are then used to obtain the HR and LR sparse coefficients using OMP algorithm which is then mapped using a conversion function. Once the learning phase is completed, the sparse coefficients of LR ECG obtained using the LR dictionary, which is then mapped into the HR sparse coefficients using the conversion function. These estimated HR sparse coefficients are then transformed to the HR ECG using the HR dictionary to obtain the derived leads or spatially enhanced ECG. The reconstruction quality of the spatially enhanced ECG is improved further by segmenting the signal and following the same approach for individual segments. This is the second model used in this chapter and it is termed as multiple joint dictionary learning approach. The first model used four leads as predictor leads, while the second model required only three leads. Both models are evaluated using the standard and diagnostic distortion measures. Despite using a four lead subset, the first model failed to capture significant diagnostic information in predictor leads. This is evident from the high $RMSE$ and $WEDD$ values of 83.43 mV and 23.95 % respectively. It is also evident that the model preserved the diagnostic features in the frontal plane lead. The comparison of this model with the existing approaches shows that it can derive leads with quality as good as that of

TH-1915_11610237

linear approaches.

The segmentation of signal improved the overall reconstruction quality of derived precordial leads as it represented the diagnostic features in a better way. The correlation coefficient, $RMSE$ and $WEDD$ for the precordial leads in this case are 0.98, 54.26 mV and 15.70% respectively. Improvements from the previous model in case of frontal plane leads are also evident from the results. The median $WEDD$ value of precordial leads at 12.22 % indicates that the model could generate leads without losing diagnostic quality for half of the records. Feature level analysis showed good correlation between the features of original and derived leads. Diagnosability of the second model is also evaluated by computing the classification accuracy between selected MI cases. The analysis of diagnosability indicates the capability of the second model in preserving the clinically significant information. The comparison of this model with the existing approaches show that it can preserve the diagnostic features in ECG much better than the most of the models.

The approach discussed in Chapter 5 uses RNN for enhancing the spatial resolution of ECG. In this approach, both intra-lead and inter-lead correlations are simultaneously exploited for learning the model. The capability of RNN in connecting the present and the past information is utilized to represent the inter-lead and the intra-lead correlations in lead signals. In this approach, the mapping between the predictor and the response lead signals is learned in a high dimensional hidden feature space. This ensures a better representation of the clinically significant diagnostic features in ECG. Three types of RNNs, simple RNN, LSTM and GRU, are used to learn the models. The performance of these models is analyzed using standard and diagnostic distortion measures. The analysis of distortion measure shows that the LSTM model outperforms the other two models. The correlation coefficient, $RMSE$, and $WEDD$ for the LSTM model are 0.98, 42.11 mV and 12.50 % respectively. The second best model is GRU followed by the simple RNN model. The WDD values of LSTM is 7.39 %, that of GRU is 8.27 % and that of simple RNN is 9.99 %. This indicates the better capability of LSTM model in preserving diagnostic features compared to GRU and simple RNN model. The features level comparison of original and derived leads are performed, and the results showed good similarity for LSTM based model, followed by GRU and simple RNN models. The median $WEDD$ values are 17.82% for simple RNN model, 10.44% for LSTM model and 11.48% for GRU model. This shows that the LSTM and GRU models are more capable of preserving the diagnostic quality than

TH-1915_11610237

the simple RNN model. The analysis of the diagnosability of these models suggests that the best model can be different for different classes. The comparison shows that the performances of simple RNN and GRU models are comparable with the existing models and that of the LSTM model.

This thesis can be summarized as follows. The approaches discussed in Chapter 3 and Chapter 4 exploited the spatial correlations and that discussed in Chapter 5 used the spatio-temporal correlations. It is also observed that the three approaches discussed in this thesis are capable of capturing the clinically significant diagnostic information. The first approach uses a linear method for capturing the diagnostic information and its execution time is low. Even though the accuracy of the generated ECG is low, this method can be used in areas where short execution time is expected. The second and third methods are non-linear methods and hence require more execution time. The dictionary learning approach can be useful in areas where a medium accuracy and execution time is required. The third approach that uses RNN require more execution time, but generates spatially enhanced ECG with the best accuracy. This approach can be used in areas where high processing power is available. The execution time can be brought down by using systems with good processing capacity. Hence RNN models may be preferred among the three models as it posses the best reconstruction accuracy.

6.1 Summary of contributions

The major contributions of the work reported in this thesis are:

(i) Exploitation of inter-lead correlations in wavelet domain for improving the reconstruction quality of derived ECG.

(ii) A novel diagnostic similarity score based lead selection algorithm for selecting the best predictor precordial lead.

(iii) Proposed and evaluated a method for segmenting the signal and learning joint dictionaries for each segment.

(iv) A new approach for enhancing the spatial resolution of ECG by the spatial mapping of the segmented sparse coefficients.

TH-1915_11610237

(v) Proposed and evaluated a method to simultaneously exploit inter-lead and intra-lead correlations for capturing diagnostic features of derived ECG.

(vi) Investigated and evaluated the capability of different RNN architectures like simple RNN, LSTM and GRU in improving the spatial resolution of ECG.

(vii) Evaluated the capacity of a derived ECG model in representing clinically significant diagnostic information using ECG diagnosability.

6.2 Scope for the future work

The thesis point towards some interesting future directions.

- Learning models that can utilize the intra-beat or intra-feature correlation along with intra-lead and inter-lead correlations could preserve the diagnostically significant features in a better way. The RNN models can be utilized for this approach and might improve the reconstruction accuracy.
- The reconstruction accuracy of the derived ECG depends upon the quality of predictor leads and how well it is correlated with the derived leads. Selection of predictor leads based on specific diseases might improve the accuracy of those leads in which the disease is prominently visible. This approach can be a two stage process which require classification to select adequate leads and derive leads using the selected leads in the second stage.
- Most of the derived ECG models are patient-specific models because of its high reconstruction accuracy due to the personalized learning process. The global models currently available are not capable of preserving the clinical information. Developing global models that can preserve diagnostic quality, thereby improving reconstruction accuracy can be an interesting area for future research.
- Current diagnostic distortion measures are designed to evaluate ECG compression algorithms. Such an approach might be less suitable for evaluating the derived ECG systems. Developing distortion measures that are suited for derived ECG systems can be another area to explore.

TH-1915_11610237

Bibliography

[1] I. Tomašić and R. Trobec, "Electrocardiographic systems with reduced numbers of leads - synthesis of the 12-lead ecg," *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 126–142, 2014.

[2] J. Malmivuo and R. Plonsey, *Bioelectromagnetism : Principles and Applications of Bioelectric and Biomagnetic Fields*, 1st ed. Oxford Univ. Press, USA, Jul 1995. [Online]. Available: http://www.bem.fi/book/index.htm

[3] F. N. Wilson, C. E. Kossmann, G. E. Burch, E. Goldberger, A. Graybiel, H. H. Hecht, F. D. Johnston, E. Lepeschkin, and G. B. Myers, "Recommendations for standardization of electrocardiographic and vectorcardiographic leads," *Circulation*, vol. 10, no. 4, pp. 564–573, 1954.

[4] W. Einthoven, "The different forms of the human electrocardiogram and their signification." *The Lancet*, vol. 179, no. 4622, pp. 853 – 861, 1912, originally published as Volume 1, Issue 4622.

[5] H. C. Burger and J. B. Van Milaan, "Heart-vector and leads," *Br. Heart J.*, vol. 8, no. 3, pp. 157–161, 1946.

[6] ——, "Heart-vector and leads. part ii," *Br. Heart J.*, vol. 9, no. 3, pp. 154–160, 1947.

[7] D. D. Finlay, C. D. Nugent, J. G. Kellett, M. P. Donnelly, P. J. McCullagh, and N. D. Black, "Synthesising the 12-lead electrocardiogram: Trends and challenges," *Eur. J. Intern. Med.*, vol. 18, no. 8, pp. 566 – 570, 2007.

[8] S. Maheshwari, A. Acharyya, M. Schiariti, and P. E. Puddu, "Personalized reduced 3-lead system formation methodology for remote health monitoring applications and reconstruction of standard 12-lead system," *Int. Arch. Med.*, vol. 8, no. 62, pp. 1–15, 2015.

[9] L. Sharma, S. Dandapat, and A. Mahanta, "Multichannel ecg data compression based on multiscale principal component analysis," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 730–736, July 2012.

[10] R. Trobec, "Computer analysis of multichannel ecg," *Comput. Biol. Med.*, vol. 33, no. 3, pp. 215 – 226, 2003, human Heart in the Focus of Computer Power.

[11] E. Trägårdh, H. Engblom, and O. Pahlm, "How many ecg leads do we need?" *Cardiology Clinics*, vol. 24, no. 3, pp. 317 – 330, 2006, advanced 12-Lead Electrocardiography.

[12] F. N. Wilson, A. Macleod, and P. S. Barker, "The potential variations produced by the heart beat at the apices of einthoven's triangle," *Am. Heart J.*, vol. 7, no. 2, pp. 207 – 211, 1931.

[13] "Standardization of precordial leads: Joint recommendations of the j. am. heart assoc. and the cardiac society of great britain and ireland," *Am. Heart J.*, vol. 15, no. 1, pp. 107 – 108, 1938.

[14] E. Goldberger, "A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads," *Am. Heart J.*, vol. 23, no. 4, pp. 483 – 492, 1942.

[15] J. Malmivuo and R. Plonsey, "12-lead ecg system," in *Bioelectromagnetism : Principles and Applications of Bioelectric and Biomagnetic Fields*, 1st ed. Oxford Univ. Press, USA, Jul 1995, ch. 15. [Online]. Available: http://www.bem.fi/book/index.htm

TH-1915_11610237

[16] H. C. Burger and J. B. Van Milaan, “Heart-vector and leads.” *Br. Heart J.*, vol. 10, no. 4, pp. 229–233, 1948.

[17] J. E. Hall and A. C. Guyton, “Cardiac muscle; the heart as a pump and function of the heart valves,” in *Guyton and Hall*, 12th ed. Saunders Elsevier, 2011, ch. 9.

[18] Y. Zigel, A. Cohen, and A. Katz, “The weighted diagnostic distortion (wdd) measure for ecg signal compression,” *IEEE Trans. Biomed. Eng.*, vol. 47, no. 11, pp. 1422–1430, Nov 2000.

[19] R. Bousseljot, D. Kreiseler, and A. Schnabel, “Nutzung der ekg-signaldatenbank cardiodat der ptb ber das internet,” *Biomed. Tech.*, vol. 40, no. 1, pp. 317 – 318, 1995.

[20] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[21] in *Chou’s Electrocardiography in Clinical Practice (Sixth Edition)*, sixth edition ed., B. Surawicz and T. K. Knilans, Eds. Philadelphia: W.B. Saunders, 2008, pp. 721 – 732.

[22] J. E. Hall and A. C. Guyton, *Guyton and Hall*, 12th ed. Saunders Elsevier, 2011.

[23] S. Padhy and S. Dandapat, “Third-order tensor based analysis of multilead ecg for classification of myocardial infarction,” *Biomed. Signal Process. Control*, vol. 31, pp. 71 – 78, 2017.

[24] J. J. Nallikuzhy and S. Dandapat, “Spatial enhancement of {ECG} using diagnostic similarity score based lead selective multi-scale linear model,” *Comput. Biol. Med.*, vol. 85, pp. 53 – 62, 2017.

[25] J. Malmivuo and R. Plonsey, “Theoretical methods for analyzing volume sources and volume conductors,” in *Bioelectromagnetism : Principles and Applications of Bioelectric and Biomagnetic Fields*, 1st ed. Oxford Univ. Press, USA, Jul 1995, ch. 11. [Online]. Available: http://www.bem.fi/book/index.htm

[26] B. M. Horáček, “Lead theory,” in *Comprehensive Electrocardiology*, 2nd ed., P. W. Macfarlane, A. van Oosterom, O. Pahlm, P. Kligfield, M. Janse, and J. Camm, Eds. Springer-Verlag London, 2011, vol. 4, ch. 7, pp. 348–376.

[27] E. Frank, “General theory of heart-vector projection,” *Circ. Res.*, vol. 2, no. 3, pp. 258–270, 1954.

[28] ——, “The image surface of a homogeneous torso,” *Am. Heart J.*, vol. 47, no. 5, pp. 757 – 768, 1954.

[29] ——, “Spread of current in volume conductors of finite extent,” *Ann. N. Y. Acad. Sci.*, vol. 65, no. 6, pp. 980–1002, 1957.

[30] D. B. Geselowitz, “Dipole theory in electrocardiography,” *Am. J. Cardiol.*, vol. 14, no. 3, pp. 301 – 306, 1964.

[31] ——, “Multipole representation for an equivalent cardiac generator,” *Proc.IRE*, vol. 48, no. 1, pp. 75–79, Jan 1960.

[32] J. Malmivuo and R. Plonsey, “Volume source and volume conductor,” in *Bioelectromagnetism : Principles and Applications of Bioelectric and Biomagnetic Fields*, 1st ed. Oxford Univ. Press, USA, Jul 1995, ch. 7. [Online]. Available: http://www.bem.fi/book/index.htm

[33] W. T. Miller and D. B. Geselowitz, “Simulation studies of the electrocardiogram. i. the normal heart.” *Circ. Res.*, vol. 43, no. 2, pp. 301–315, 1978.

[34] D. B. Geselowitz, “On bioelectric potentials in an inhomogeneous volume conductor,” *Biophys. J.*, vol. 7, no. 1, pp. 1 – 11, 1967.

[35] J. Malmivuo and R. Plonsey, “Sourcefield models,” in *Bioelectromagnetism : Principles and Applications of Bioelectric and Biomagnetic Fields*, 1st ed. Oxford Univ. Press, USA, Jul 1995, ch. 8. [Online]. Available: http://www.bem.fi/book/index.htm

[36] H. P. Schwan and C. F. Kay, “Capacitive properties of body tissues,” *Circ. Res.*, vol. 5, no. 4, pp. 439–443, 1957.

TH-1915_11610237

[37] D. B. Geselowitz, "The concept of an equivalent cardiac generator." *Biomed. Sci. Instrum.*, vol. 1, pp. 325–30, 1963.

[38] R. Plonsey and D. B. Heppner, "Considerations of quasi-stationarity in electrophysiological systems," *Bull. Math. Biophys.*, vol. 29, no. 4, pp. 657–664, Dec 1967.

[39] W. Fye, "A history of the origin, evolution, and impact of electrocardiography," *Am. J. Cardiol.*, vol. 73, no. 13, pp. 937 – 949, 1994.

[40] S. S. Barold, "Willem einthoven and the birth of clinical electrocardiography a hundred years ago," *Card. Electrophysiol. Rev.*, vol. 7, no. 1, pp. 99–104, Jan 2003.

[41] E. Frank, "An accurate, clinically practical system for spatial vectorcardiography," *Circulation*, vol. 13, no. 5, pp. 737–749, 1956.

[42] G. E. Dower, A. Yakush, S. B. Nazzal, R. V. Jutzy, and C. E. Ruiz, "Deriving the 12-lead electrocardiogram from four (easi) electrodes," *J. Electrocardiol.*, vol. 21, no. Supplement, pp. S182 – S187, 1988, computerized Interpretation of the Electrocardiogram XIII.

[43] R. E. Mason and I. Likar, "A new system of multiple-lead exercise electrocardiography," *Am. Heart J.*, vol. 71, no. 2, pp. 196 – 205, 1966.

[44] J. A. Scherer and J. M. Nicklas, "Synthesis of the 12 lead electrocardiogram from a 3 lead semi-orthogonal subset using patient-specific linear transformation arrays," in *Proc. Comput. Cardiol., 1988*, Sep 1988, pp. 449–451.

[45] S. P. Nelwan, J. A. Kors, and S. H. Meij, "Minimal lead sets for reconstruction of 12-lead electrocardiograms," *J. Electrocardiol.*, vol. 33, no. Supplement 1, pp. 163 – 166, 2000, research and Technology Transfer in Computerized Electrocardiology.

[46] S. P. Nelwan, J. A. Kors, S. H. Meij, J. H. van Bemmel, and M. L. Simoons, "Reconstruction of the 12-lead electrocardiogram from reduced lead sets," *J. Electrocardiol.*, vol. 37, no. 1, pp. 11 – 18, 2004.

[47] S. P. Nelwan, J. A. Kors, S. W. Crater, S. H. Meij, T. B. van Dam, M. L. Simoons, and M. W. Krucoff, "Simultaneous comparison of 3 derived 12-lead electrocardiograms with standard electrocardiogram at rest and during percutaneous coronary occlusion," *J. Electrocardiol.*, vol. 41, no. 3, pp. 230 – 237, 2008.

[48] O. H. Schmitt, "Cathoderay presentation of threedimensional data," *J. Appl. Phys.*, vol. 18, no. 9, pp. 819–829, 1947.

[49] S. Oh and S. E, "Symposium on electrocardiography and vectorcardiography: The present status of vectorcardiography," *AMA Arch. Intern. Med.*, vol. 96, no. 5, pp. 574–590, 1955.

[50] W. R. Milnor, S. A. Talbot, and E. V. Newman, "A study of the relationship between unipolar leads and spatial vectorcardiograms, using the panoramic vectorcardiograph," *Circulation*, vol. 7, no. 4, pp. 545–557, 1953.

[51] R. A. Helm, "The scalar presentation of orthogonal leads," *Am. Heart J.*, vol. 54, no. 1, pp. 89 – 97, 1957.

[52] ——, "A universal system of electrode placement for electrocardiography and spatial vectorcardiography," *Am. Heart J.*, vol. 58, no. 1, pp. 71 – 87, 1959.

[53] H. V. Pipberger and C. R. Wood, "A simplified method for the resolution of the orthogonal electrocardiogram," *Circ. Res.*, vol. 6, no. 3, pp. 239–243, 1958.

[54] H. V. Pipberger, S. M. Bialek, J. K. Perloff, and H. W. Schnaper, "Correlation of clinical information in the standard 12-lead ecg and in a corrected orthogonal 3-lead ecg," *Am. Heart J.*, vol. 61, no. 1, pp. 34 – 43, 1961.

[55] R. H. Okada, P. H. Langner, and S. A. Briller, "Synthesis of precordial potentials from the svec iii vectorcardiographic system," *Circ. Res.*, vol. 7, no. 2, pp. 185–191, 1959.

TH-1915_11610237

[56] J. A. Abildskov and R. S. Wilkinson, "The relation of precordial and orthogonal leads," *Circulation*, vol. 27, no. 1, pp. 58–63, 1963.

[57] R. McFee, A. Parungao, and W. Mueller, "An electronic coordinate transformer for electrocardiography," *IRE Trans. Bio-Med. Electron.*, vol. 8, no. 1, pp. 52–54, Jan 1961.

[58] R. McFee and A. Parungao, "An orthogonal lead system for clinical electrocardiography," *Am. Heart J.*, vol. 62, no. 1, pp. 93 – 100, 1961.

[59] G. E. Dower, "A lead synthesizer for the frank system to simulate the standard 12-lead electrocardiogram," *J. Electrocardiol.*, vol. 1, no. 1, pp. 101 – 116, 1968.

[60] G. E. Dower, H. B. Machado, and J. A. Osborne, "On deriving the electrocardiogram from vectorcardiographic leads," *Clin. Cardiol.*, vol. 3, no. 2, pp. 87–95, 1980.

[61] D. Dawson, H. Yang, M. Malshe, S. T. Bukkapatnam, B. Benjamin, and R. Komanduri, "Linear affine transformations between 3-lead (frank {XYZ} leads) vectorcardiogram and 12-lead electrocardiogram signals," *J. Electrocardiol.*, vol. 42, no. 6, pp. 622 – 630, 2009.

[62] S. Maheshwari, A. Acharyya, P. E. Puddu, E. B. Mazomenos, M. Schiariti, and K. Maharatna, "Robust and accurate personalised reconstruction of standard 12-lead system from frank vectorcardiographic system," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 4, no. 3-4, pp. 183–192, 2016.

[63] J. A. Kors, G. Van Herpen, A. C. Sittig, and J. H. Van Bemmel, "Reconstruction of the frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods," *Eur. Heart J.*, vol. 11, no. 12, pp. 1083–1092, 1990.

[64] L. Edenbrandt, A. Houston, and P. W. Macfarlane, "Vectorcardiograms synthesized from 12-lead ecgs: A new method applied in 1792 healthy children," *Pediatr. Cardiol.*, vol. 15, no. 1, pp. 21–26, 1994.

[65] S. Maheshwari, A. Acharyya, M. Schiariti, and P. E. Puddu, "Frank vectorcardiographic system from standard 12 lead ecg: An effort to enhance cardiovascular diagnosis," *J. Electrocardiol.*, vol. 49, no. 2, pp. 231 – 242, 2016.

[66] M. D. Klein, I. Key-Brothers, and C. L. Feldman, "Can the vectorcardiographically derived easi ecg be a suitable surrogate for the standard ecg in selected circumstances," in *Proc. Comput. Cardiol., 1997*, Sep 1997, pp. 721–724.

[67] B. J. Drew, M. M. Scheinman, and G. Evans, "Comparison of a vectorcardiographically derived 12-lead electrocardiogram with the conventional electrocardiogram during wide qrs complex tachycardia, and its potential application for continuous bedside monitoring," *Am. J. Cardiol.*, vol. 69, no. 6, pp. 612 – 618, 1992.

[68] B. J. Drew, M. G. Adams, M. M. Pelter, S.-F. Wung, and M. A. Caldwell, "Comparison of standard and derived 12-lead electrocardiograms for diagnosis of coronary angioplasty-induced myocardial ischemia," *Am. J. Cardiol.*, vol. 79, no. 5, pp. 639 – 644, 1997.

[69] B. J. Drew, M. M. Pelter, S.-F. Wung, M. G. Adams, C. Taylor, G. Evans, and E. Foster, "Accuracy of the easi 12-lead electrocardiogram compared to the standard 12-lead electrocardiogram for diagnosing multiple cardiac abnormalities," *J. Electrocardiol.*, vol. 32, no. Supplement 1, pp. 38 – 47, 1999, research and Technology Transfer in Computerized Electrocardiology.

[70] B. M. Horáček, J. W. Warren, P. Šťovíček, and C. L. Feldman, "Diagnostic accuracy of derived versus standard 12-lead electrocardiograms," *J. Electrocardiol.*, vol. 33, no. Supplement 1, pp. 155 – 160, 2000, research and Technology Transfer in Computerized Electrocardiology.

[71] M. Sejersten, O. Pahlm, J. Pettersson, P. M. Clemmensen, F. Rautaharju, S. Zhou, C. Maynard, C. L. Feldman, and G. S. Wagner, "The relative accuracies of ecg precordial lead waveforms derived from easi leads and those acquired from paramedic applied standard leads," *J. Electrocardiol.*, vol. 36, no. 3, pp. 179 – 185, 2003.

TH-1915_11610237

[72] D. Q. Feild, C. L. Feldman, and B. Horáček, "Improved easi coefficients: Their derivation, values, and performance," *J. Electrocardiol.*, vol. 35, no. 4, Part B, pp. 23 – 33, 2002.

[73] P. M. Rautaharju, S. H. Zhou, E. Hancock, B. Horáček, D. Q. Feild, J. M. Lindauer, G. S. Wagner, O. Pahlm, and C. L. Feldman, "Comparability of 12-lead ecgs derived from easi leads with standard 12-lead ecgs in the classification of acute myocardial ischemia and old myocardial infarction," *J. Electrocardiol.*, vol. 35, no. 4, Part B, pp. 35 – 39, 2002.

[74] G. Wehr, R. J. Peters, K. Khalif, A. P. Banning, V. Kuehlkamp, A. F. Rickards, and U. Sechtem, "A vector-based, 5-electrode, 12-lead monitoring ecg (easi) is equivalent to conventional 12-lead ecg for diagnosis of acute coronary syndromes," *J. Electrocardiol.*, vol. 39, no. 1, pp. 22 – 28, 2006.

[75] M. Sejersten, G. S. Wagner, O. Pahlm, J. W. Warren, C. L. Feldman, and B. M. Horáček, "Detection of acute ischemia from the easi-derived 12-lead electrocardiogram and from the 12-lead electrocardiogram acquired in clinical practice," *J. Electrocardiol.*, vol. 40, no. 2, pp. 120 – 126, 2007.

[76] D. Chantad, R. Krittayaphong, and C. Komoltri, "Derived 12-lead electrocardiogram in the assessment of st-segment deviation and cardiac rhythm," *J. Electrocardiol.*, vol. 39, no. 1, pp. 7 – 12, 2006.

[77] O. Pahlm, J. Pettersson, A. Thulin, C. L. Feldman, D. Q. Feild, and G. S. Wagner, "Comparison of waveforms in conventional 12-lead ecgs and those derived from easi leads in children," *J. Electrocardiol.*, vol. 36, no. 1, pp. 25 – 31, 2003.

[78] A. Welinder, D. Q. Feild, J. Liebman, C. Maynard, G. S. Wagner, G. Wettrell, and O. Pahlm, "Diagnostic conclusions from the easi-derived 12-lead electrocardiogram as compared with the standard 12-lead electrocardiogram in children," *Am. Heart J.*, vol. 151, no. 5, pp. 1059 – 1064, 2006.

[79] L. Hadžievski, B. Bojović, V. Vukčević, P. Beličev, S. Pavlović, Z. Vasiljević-Pokrajčić, and M. Ostojić, "A novel mobile transtelephonic system with synthesized 12-lead ecg," *IEEE Trans. Inf. Technol. Biomed.*, vol. 8, no. 4, pp. 428–438, Dec 2004.

[80] D. D. Finlay, C. D. Nugent, M. P. Donnelly, and R. L. Lux, "Eigenleads: Ecg leads for maximizing information capture and improving snr," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 1, pp. 69–78, Jan 2010.

[81] R. Trobec and Tomašić, "Synthesis of the 12-lead electrocardiogram from differential leads," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 4, pp. 615–621, July 2011.

[82] M. Papouchado, P. R. Walker, M. A. James, and L. M. Clarke, "Fundamental differences between the standard 12lead electrocardiograph and the modified (masonlikar) exercise lead system," *Eur. Heart J.*, vol. 8, no. 7, pp. 725–733, 1987.

[83] D. C. Sevilla, M. L. Dohrmann, C. A. Somelofski, R. P. Wawrzynski, N. B. Wagner, and G. S. Wagner, "Invalidation of the resting electrocardiogram obtained via exercise electrode sites as a standard 12-lead recording," *Am. J. Cardiol.*, vol. 63, no. 1, pp. 35 – 39, 1989.

[84] D. Wei, "Deriving the 12-lead electrocardiogram from four standard leads based on the frank torso model," in *Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2001*, vol. 1, 2001, pp. 381–384 vol.1.

[85] ——, "Deriving the 12-lead electrocardiogram from four standard leads using information redundancy in the 12-lead system," *Int. J. Bioelectromagn.*, vol. 4, pp. 127 – 128, 2002.

[86] B. J. Drew, M. M. Pelter, D. E. Brodnick, A. V. Yadav, D. Dempel, and M. G. Adams, "Comparison of a new reduced lead set ecg with the standard ecg for diagnosing cardiac arrhythmias and myocardial ischemia," *J. Electrocardiol.*, vol. 35, no. 4, Part B, pp. 13 – 21, 2002.

[87] A. J. Robertson, E. Clark, A. van Oosterom, and P. W. Macfarlane, "Ecg monitoring using a limited lead set," *Int. J. Bioelectromagn.*, vol. 4, pp. 295 – 296, 2002.

TH-1915_11610237

[88] S.-C. Man, A. C. Maan, E. Kim, H. H. Draisma, M. J. Schalij, E. E. van der Wall, and C. A. Swenne, "Reconstruction of standard 12-lead electrocardiograms from 12-lead electrocardiograms recorded with the mason-likar electrode configuration," *J. Electrocardiol.*, vol. 41, no. 3, pp. 211 – 219, 2008.

[89] B. M. Horáček, J. W. Warren, and J. J. Wang, "On designing and testing transformations for derivation of standard 12-lead/18-lead electrocardiograms and vectorcardiograms from reduced sets of predictor leads," *J. Electrocardiol.*, vol. 41, no. 3, pp. 220 – 229, 2008.

[90] S. Maheshwari, A. Acharyya, P. E. Puddu, and M. Schiariti, "Reduced lead system selection methodology for reliable standard 12-lead reconstruction targeting personalised remote health monitoring applications," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 2, no. 2, pp. 107–120, 2014.

[91] G. Tsouri and M. Ostertag, "Patient-specific 12-lead ecg reconstruction from sparse electrodes using independent component analysis," *IEEE J. Biomed. Health. Inf.*, vol. 18, no. 2, pp. 476–482, March 2014.

[92] J. Lee, M. Kim, and J. Kim, "Reconstruction of precordial lead electrocardiogram from limb leads using the state-space model," *IEEE J. Biomed. Health. Inf.*, vol. 20, no. 3, pp. 818–828, May 2016.

[93] R. E. Gregg, S. H. Zhou, J. M. Lindauer, D. Q. Feild, and E. D. Helfenbein, "Where do derived precordial leads fail?" *J. Electrocardiol.*, vol. 41, no. 6, pp. 546 – 552, 2008.

[94] M. S. Guillem, A. V. Sahakian, and S. Swiryn, "Derivation of orthogonal leads from the 12-lead ecg. accuracy of a single transform for the derivation of atrial and ventricular waves," in *Proc. Comput. Cardiol., 2006*, Sept 2006, pp. 249–252.

[95] M. Sejersten, O. Pahlm, J. Pettersson, S. Zhou, C. Maynard, C. L. Feldman, and G. S. Wagner, "Comparison of easi-derived 12-lead electrocardiograms versus paramedic-acquired 12-lead electrocardiograms using mason-likar limb lead configuration in patients with chest pain," *J. Electrocardiol.*, vol. 39, no. 1, pp. 13 – 21, 2006.

[96] S. Maheshwari, A. Acharyya, P. Rajalakshmi, P. Puddu, and M. Schiariti, "Accurate and reliable 3-lead to 12-lead {ECG} reconstruction methodology for remote health monitoring applications," {*IRBM*}, vol. 35, no. 6, pp. 341 – 350, 2014, healthcom 2013.

[97] D. Q. Feild, S. H. Zhou, E. D. Helfenbein, R. E. Gregg, and J. M. Lindauer, "Technical challenges and future directions in lead reconstruction for reduced-lead systems," *J. Electrocardiol.*, vol. 41, no. 6, pp. 466 – 473, 2008.

[98] H. Atoui, J. Fayn, and P. Rubel, "A neural network approach for patient-specific 12-lead ecg synthesis in patient monitoring environments," in *Proc. Comput. Cardiol., 2004*, Sept 2004, pp. 161–164.

[99] ——, "A novel neural-network model for deriving standard 12-lead ecgs from serial three-lead ecgs: Application to self-care," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 883–890, May 2010.

[100] A. Yodjaiphet, N. Theera-Umpon, and S. Auephanwiriyakul, "Electrocardiogram reconstruction using support vector regression," in *2012 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Dec 2012, pp. 000 269–000 273.

[101] P. Kaewfoongrungsi and D. Hormdee, "Support vector regression-based synthesis of 12-lead ecg system from the standard 5 electrode system using lead v1," *KKU Engineering Journal*, vol. 43, no. S3, pp. 494–498, 2016.

[102] F. Chollet *et al.*, "Keras," https://github.com/keras-team/keras, 2015.

[103] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[104] M. S. Manikandan and S. Dandapat, "Wavelet energy based diagnostic distortion measure for {ECG}," *Biomed. Signal Process. Control*, vol. 2, no. 2, pp. 80 – 96, 2007.

TH-1915_11610237

[105] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov 2010.

[106] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3467–3478, Aug 2012.

[107] S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 2216–2223.

[108] J. Sadasivan, S. Mukherjee, and C. S. Seelamantula, "Joint dictionary training for bandwidth extension of speech signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, March 2016, pp. 5925–5929.

[109] B. Mailhé, R. Gribonval, F. Bimbot, M. Lemay, P. Vandergheynst, and J. M. Vesin, "Dictionary learning for the sparse modelling of atrial fibrillation in ecg signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, April 2009, pp. 465–468.

[110] L. F. Polania and K. E. Barner, "Multi-scale dictionary learning for compressive sensing ecg," in *2013 IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, Aug 2013, pp. 36–41.

[111] Y. M. Lin, Y. Chen, H. C. Kuo, and A. Y. A. Wu, "Compressive sensing based ecg telemonitoring with personalized dictionary basis," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Oct 2015, pp. 1–4.

[112] K. Xu, Y. Li, and F. Ren, *An energy-efficient compressive sensing framework incorporating online dictionary learning for long-term wireless health monitoring*. United States: Institute of Electrical and Electronics Engineers Inc., 5 2016, vol. 2016-May, pp. 804–808.

[113] Y. Zigel, A. Cohen, and A. Katz, "A diagnostic meaningful distortion measure for ecg compression," in *Electrical and Electronics Engineers in Israel, 1996., Nineteenth Convention of*, Nov 1996, pp. 117–120.

[114] J. J. Nallikuzhy and S. Dandapat, "Enhancement of the spatial resolution of ecg using multi-scale linear regression," in *Communications (NCC), 2015 Twenty First National Conference on*, Feb 2015, pp. 1–6.

[115] J. Nallikuzhy, L. Sharma, and S. Dandapat, "Projection based approach for super-resolution ecg," in *Condition Assessment Techniques in Electrical Systems (CATCON), 2013 IEEE 1st International Conference on*, Dec 2013, pp. 270–274.

[116] J. Nallikuzhy and S. Dandapat, *Lecture Notes in Electrical Engineering - Advances in Communication and Computing*. Springer: Springer, 2015, vol. 347, ch. DCT-Based Linear Regression Approach for 12-Lead ECG Synthesis, pp. 211–220.

[117] A. S. Al-Fahoum, "Quality assessment of ecg compression techniques using a wavelet-based diagnostic measure," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 182–191, Jan 2006.

[118] M. Hilton, "Wavelet and wavelet packet compression of electrocardiograms," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 5, pp. 394–402, May 1997.

[119] P. Kligfield, L. S. Gettes, J. J. Bailey, R. Childers, B. J. Deal, E. W. Hancock, G. van Herpen, J. A. Kors, P. Macfarlane, D. M. Mirvis, O. Pahlm, P. Rautaharju, and G. S. Wagner, "Recommendations for the standardization and interpretation of the electrocardiogram: Part i: The electrocardiogram and its technology a scientific statement from the american heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the american college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology," *J. Am. Coll. Cardiol.*, vol. 49, no. 10, pp. 1109 – 1127, 2007.

[120] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.

[121] L. Sharma, R. Tripathy, and S. Dandapat, "Multiscale energy and eigenspace approach to detection and localization of myocardial infarction," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1827–1837, July 2015.

TH-1915_11610237

[122] H. C. Bazett, "An analysis of the time-relations of electrocardiograms," *Ann. Noninvasive Electrocardiol.*, vol. 2, no. 2, pp. 177–194, 1997.

[123] M. Aharon, M. Elad, and A. Bruckstein, "rmk-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov 2006.

[124] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, July 2011.

[125] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.

[126] ——, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[127] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural. Netw.*, vol. 5, no. 2, pp. 157–166, Mar 1994.

[128] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[129] K. Cho, B. van Merrienboer, Ç. Gülçehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *CoRR*, vol. abs/1406.1078, 2014.

TH-1915_11610237

LIST OF PUBLICATIONS

Journals:

1. **Jiss J. Nallikuzhy** and S. Dandapat, "Spatial enhancement of ECG using diagnostic similarity score based lead selective multi-scale linear model", *Computers in Biology and Medicine, Elsevier*, vol. 85, pp. 53-62, 2017, ISSN 0010-4825.

Conferences:

1. **Jiss J. Nallikuzhy** and S. Dandapat, "Enhancement of the spatial resolution of ECG using multi-scale Linear Regression", *2015 Twenty First National Conference on Communications (NCC), Mumbai, 2015*, pp. 1-6.

2. **Jiss J. Nallikuzhy** and S. Dandapat, "Projection based approach for super-resolution ECG, *2013 IEEE 1st International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), Kolkata, 2013*, pp. 270-274.

Book Chapter:

1. **Jiss J. Nallikuzhy** and S. Dandapat, "DCT-Based Linear Regression Approach for 12-Lead ECG Synthesis", *Lecture Notes in Electrical Engineering-Advances in Communication and Computing, Springer, 2015*, vol. 347, pp. 211-220.

Manuscripts Under Preparation:

1. **Jiss J. Nallikuzhy** and S. Dandapat, "Spatial Enhancement of ECG using Multiple Joint Dictionary Learning", *to be submitted in Computers in Biology and Medicine, Elsevier*.

2. **Jiss J. Nallikuzhy** and S. Dandapat, "Spatially Enhanced ECG by Exploiting Spatio-Temporal Correlations", *to be submitted in Biomedical Signal Processing and Control, Elsevier*.

TH-1915_11610237

TH-1915_11610237

TH-1915_11610237