T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Ischemia Diagnosis based on Fuzzy Association Rules and its Monitoring Application in Smart Home Care Environment
著者(和文)	李 天宇
Author(English)	Tianyu Li
出典(和文)	学位:博士（工学）, 学位授与機関:東京工業大学, 報告番号:甲第9877号, 授与年月日:2015年3月26日, 学位の種別:課程博士, 審査員:廣田　薫,佐藤　誠,柴田　崇徳,室伏　俊明,小野　功
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9877号, Conferred date:2015/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Powered by T2R2 (Science Tokyo Research Repository)

Ischemia Diagnosis based on Fuzzy Association Rules and its Monitoring Application in Smart Home Care Environment

Tianyu Li

Supervisor: Prof. Kaoru HIROTA

Doctoral Thesis

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

March 2015

Ischemia Diagnosis based on Fuzzy Association Rules and its Monitoring Application in Smart Home Care Environment

Tianyu Li
李　天宇

Supervisor: Prof. Kaoru HIROTA

Doctoral Thesis

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

March 2015

Abstract

A method based on fuzzy association rule mining is proposed for myocardial ischemia diagnosis, and real time heart disease monitoring application is done in smart home care environment. The proposed method provides the rule mining results in an interpretable and understandable format to professional doctors as an assistant reference. ECG recordings from European ST-T Database are used in the fuzzy association rule mining experiment, in which 43 interpretable association rules are extracted. Afterwards, above mined rules are applied on test dataset in the classification experiment which is conducted on a PC via Matlab software. Classification results obtain adequate assistant reference with values of 83.4%, 80.7%, and 81.4% for sensitivity, specificity, and accuracy, respectively. In addition, a framework for classification using fuzzy association rules is elaborated to explore how to achieve accurate classification results with understandable rules. Despite of tons of proposed intelligent methodologies, the diagnosis of heart diseases in reality still requires the practical clinic experience from professionals. The proposal, that aims to be an effective, helpful assistant medium, is intended to be expanded to other heart disease diagnosis.

To avoid the situation that elderly people get permanent harm due to a sudden attack of myocardial ischemia when they are at home alone, an application that detects ischemia and monitors people's well-being via smart devices in real time is proposed and applied in smart home care environment. In the experiment, the application, "MonitoringApp", which is developed on MacBook via Xcode, shows ECG signal and

its corresponding parameters in real time, recognizes ischemic heartbeat episodes based on a Support Vector Machine (SVM) model which classifies single heartbeat with sensitivity of 82.1%. Once any abnormality is found based on real time classification results, the application is capable of warning people by vibration, sound, and sending alarms to professional doctors, family members instantly.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Professor Kaoru Hirota, for letting me enter his lab, being my mentor, and guiding me through the past four years. Thanks to his trust, encouragement, and patient advices, I am able to complete my study in master and doctoral programs. Also, his insightful talks, guidance, and wisdom will become my precious asset for the rest of my life.

I would like to thank Dr. Fangyan Dong, assistant professor in Hirota Lab, for believing in me and giving me courage. Her insightful guidance and enduring support become my strength to confront problems and difficulties. Also, many sincere thanks to Mrs. Harumi Hoshino, secretary in Hirota Lab, for her considerate help every day, that make me be able to focus on my study.

Additionally, I would like to express my genuine gratitude to Professor Makoto Sato, Professor Takanori Shibata, Professor Toshiaka Murofushi, and Professor Isano Ono for their precious time and helpful comments that have instructed me on my doctoral dissertation.

My honest gratitude also goes to Dr. Yongkang Tang, Dr. Fei. Yan, Dr. Martin. L. Tangel, Dr. Zhentao Liu, Dr. Janet. P. Betancourt, Mr. Jiaju Lu and Mr. Luefeng Chen. Thank you all for valuable advices, comments and inspiring discussions. I also would like to thank all Hirota Lab members for being supportive and friendly.

Finally, I would like to thank my parents for understanding me, supporting me, and letting me follow my heart. I am also grateful to all my family members and my dearest friends. Nothing is possible for me without your trust, encouragement and love.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Myocardial ischemia, also called cardiac ischemia, occurs when blood flow to heart muscle is decreased by a partial or complete blockage of heart's coronary arteries. The drop of blood flow reduces the oxygen supply to heart, which can damage the heart muscle, reducing its ability to pump efficiently. A sudden, severe blockage of a coronary artery may lead to a heart attack. Myocardial ischemia, as one of the cardiovascular diseases, is threatening the lives of millions of people. It usually causes a few symptoms, including shortness of breath, chest pressure or pain, neck or jaw pain, shoulder or arm pain, and fast heartbeats. In fact, a healthy lifestyle habit can help to prevent ischemia happening at the first place, because there are a set of lifestyle related factors that may increase the risk of developing myocardial ischemia, such as, obesity, diabetes, tobacco, and lack of physical activity.

Given its vast damage range and severity, besides the physical check of professional doctors, the research of intelligent ways to diagnose myocardial ischemia using Electrocardiogram (ECG) has been rapidly developing for a few years. ECG signal, as one of the most significant signals from human body, interprets the electrical activity of the heart. It is usually collected for the diagnosis of heart abnormalities. Within the ECG recording of single heartbeat, there are several important waves and intervals, including P wave, QRS complex, T wave, R-R interval, ST segment, and QT interval, etc. Yet, in a diseased heart, some of above waves and intervals may become abnormal. For instance, myocardial ischemia commonly cause ST segment deviation, flattened or inverted T waves. Therefore, the research for diagnosis of ischemia always focus on

detecting pathological changes of these parts.

Diverse artificial intelligent methodologies, models and algorithms have been performed to detection myocardial ischemia on ECG signals, such as, Support Vector Machines (SVMs), Support Vector Regression (SVR), different types of neural networks, Hidden Markov Models, Wavelet transformations, and kinds of hybrid methods. Most of these research proposals aims to directly provide automatic detection results as final diagnosis. However, in fact, the myocardial ischemia in reality still cannot only rely on the detection results of these research proposals, due to the complexity of ECG signals. According to Mayo Clinic, the current diagnosis of ischemia actually still mostly depend on the experience of professional doctors.

Besides, it becomes very worrying that the elderly people with heart disease threats is dramatically increasing. It is a huge problem to provide timely health care for such large amount of population. What is more troubling is that offering all-time nursing service or making the people who may has potential risk of heart diseases stay in hospital is impossible. As a result, such people usually cannot get treatment or rescue on time. Therefore, in order to eliminate such threat, smart home care environment, in which vital bio-information are recorded, analyzed by smart devices, then monitored in real-time so that people can get timely, appropriate care in their own home, is definitely necessary.

The thesis focuses on how to provide assistance to doctors that can help them to diagnose myocardial ischemia, and how the application in smart home care environment can provide people with timely care, aid and help via smart devices. This dissertation is organized as follows:

In chapter 2, a distance measure is proposed for time series data mining based on Symbolic Aggregate Approximation (SAX) with direction representation. Each subsequence of the time series data is mapped onto one of the three direction types: ‘convex’, ‘concave’, and ‘linear’. Not only is the original time series transformed into a series of symbols, but also into a series of direction representations. With direction

representation, the distance between two symbols in the SAX representation with different directions is overlooked. Therefore, the tightness of the lower bound is increased compared to that of the original SAX. Also, some unnecessary errors of time series data mining tasks can be avoided. The error rate of time series data mining tasks, for example, classification, could be decreased.

In chapter 3, a method based on fuzzy association rule mining is proposed to assist doctors to diagnose by providing interpretable and understandable information as a reference. Its implementation composes of fours steps: significant feature extraction from every single heartbeat, fuzzy transformation of above features, association rule mining on fuzzy itemsets, and automatic ischemia, normal beats classification via extracted rules. Features that are extracted include ST segment deviation, ST segment duration, ST segment area, T wave peak, T wave area, and T wave direction. Segmentation of above features is done by fuzzy c-mean clustering, and membership function parameters of each feature are determined based on above fuzzy c-means clustering results. Then, association rules are mined on fuzzy itemsets. At last, automatic heartbeat classification on ECG signal is performed. The proposal inherits the merit of association rule based method. First, the results of the proposal are interpretable information which makes doctors understand the underlying correlation before they make diagnosis. Second, to obtain meaningful rules, extracted features need to be segmented to different intervals before conducting association rule mining. In the proposal, fuzzy c-means clustering is applied to determine how to discretize features. In this way, it makes the feature segmentation more functional and effective.

In chapter 4, a monitoring application in smart home care environment is proposed to detect myocardial ischemia on ECG signals in real time and trigger alarms, send notifications when abnormality is found for the purpose of acquiring aid or rescue on time. On one hand, smart device is responsible for collecting ECG signals of monitored person in real time, and transmitting them to cloud server over Wi-Fi. On the other hand, a Support Vector Machine (SVM) classification model is first built offline on the cloud

server. Then, after receiving ECG signal from smart device, the classification of every heartbeat and the search of ischemic heartbeat episodes are conducted. The implementation process includes, real time QRS complex detection, feature extraction from each single heartbeat, classification using above heartbeat features. Heartbeat features that are used in SVM model training and the real time classification includes ST segment deviation, R-R interval, R wave peak, T wave peak, S wave peak, etc. Afterwards, if dangerous myocardial ischemic abnormality is discovered, smart device performs different actions, such as vibration, warning sound. Besides, the scale of necessary actions expand to notifying professional doctors, directly calling ambulance and sending messages to family members.

Finally, chapter 5 presents conclusions and future perspectives.

The research roadmap is illustrated in Fig. 1.1.

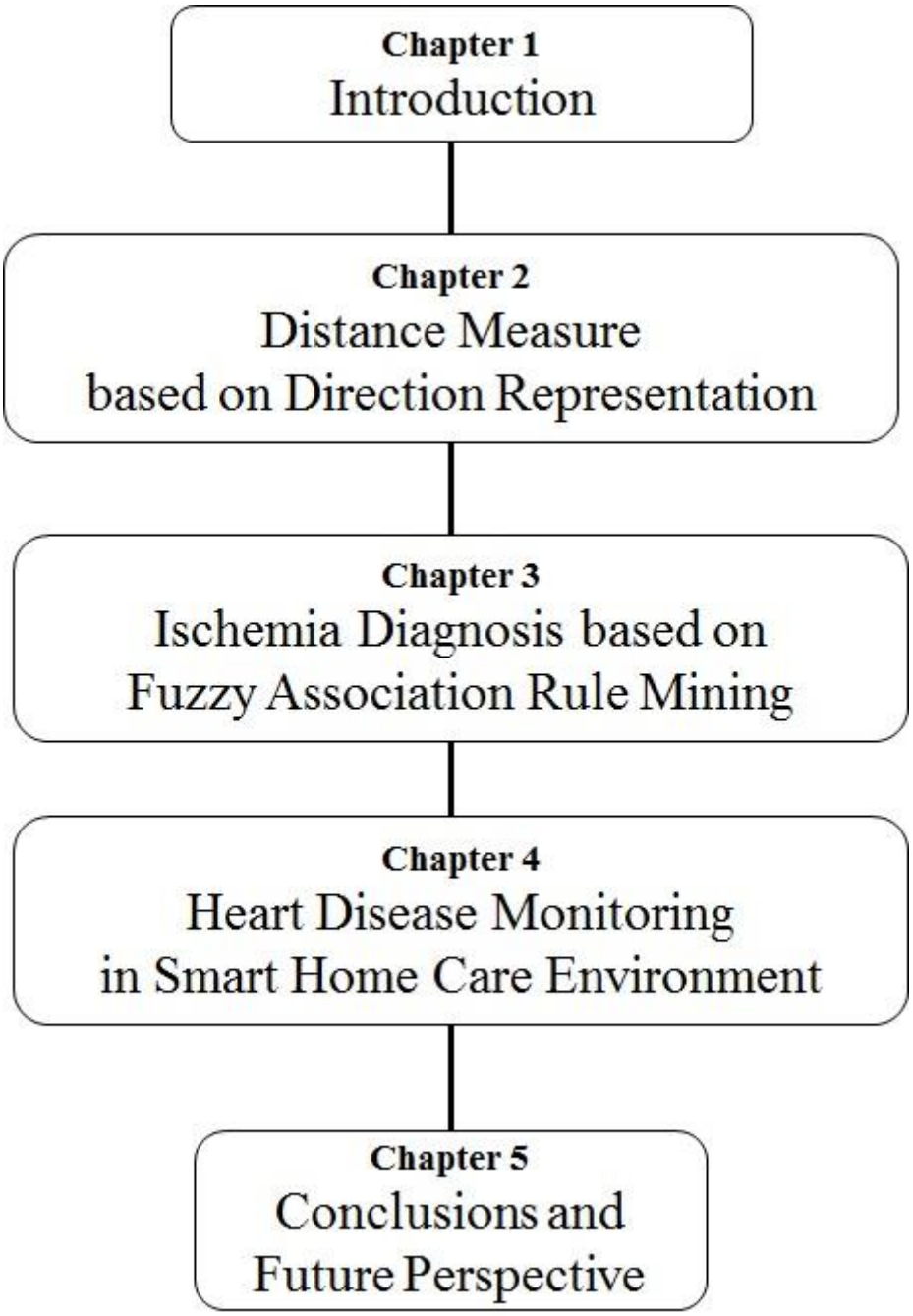

Fig. 1.1. Research roadmap

Chapter 2

Distance Measure based on Direction Representation

A time series is a collection of observations made sequentially in time [1]. They exist almost everywhere in our lives, such as stock price, sound, and meteorology information, like annual rainfall, temperature every day, etc. Besides, even our handwriting is also a form of time series data. In addition, there are several extremely significant time series data in medical field, for instance, Electrocardiogram (ECG). These time series data are so crucial that time series data mining has been attracting a lot of attentions. Typical time series data mining tasks, as shown in Fig. 2.1, include clustering, classification, motif discover, rule discovery, and novelty detection, etc. The first step of generic time series data mining algorithm is to create an approximation of original data. The reason why we do not execute algorithm directly on raw data is that raw data has too many dimensions, while very large portion of time series data mining algorithms would degrade their performance with high dimensionality. Therefore, dimension reduction is very important for data mining tasks. Most time series representations are shown in Fig. 2.2. There are many famous algorithms, such as, Wavelet Transform, Singular Value Decomposition, and Fourier Transform, etc. The research work in this chapter is based on one of the symbolic representations, Symbolic Aggregate Approximation (SAX). Since symbolic representation allows data miner to

use suffix tree, Hashing, Markov Models, text processing algorithms, and bioinformatics algorithms, SAX has been widely applied in many application areas, such as, financial, medical, and gesture recognition area, etc.

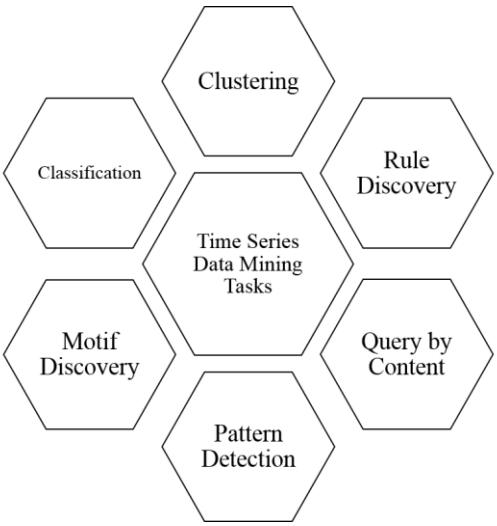

Fig. 2.1. Time series data mining tasks

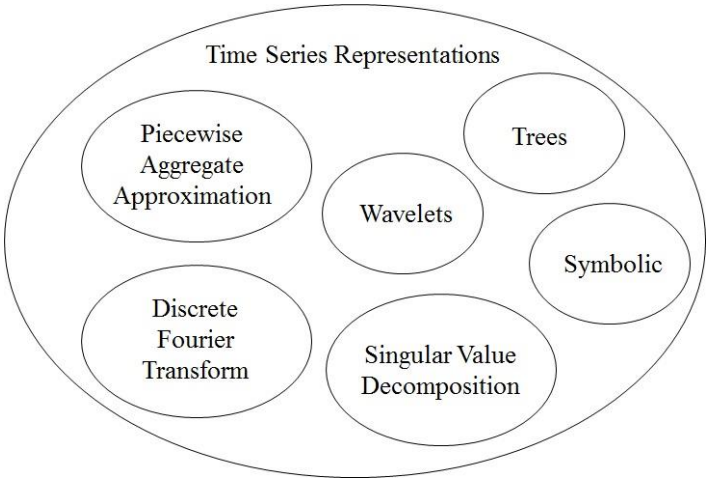

Fig. 2.2. Time series approximation representations

2.1 Symbolic Aggregate Approximation (SAX)

Symbolic Aggregate Approximation (SAX) [2] transforms original time series data into a series of symbols and effectively reduces the high dimensionality of time series data by producing the mean value of equal sized subsequence in time series data mining. The distance measure of SAX guarantees that the calculated distance between two SAX representations lower bounds to their Euclidean distance [2][3][4][5].

As one among symbolic representations, SAX allows time series data miners to apply a series of methods like hashing, Markov Models, suffix trees, and decision trees. There is, moreover, an enormous wealth of existing algorithms and data structures in text processing and bioinformatics communities that allow efficient manipulation of symbolic representation [2][6]. Dimensionality reduction characteristic of SAX that is obtained through Piecewise Aggregate Approximation (PAA) [7] avoids situations that the data mining algorithms' performance degrade with the high dimensionality of original time series data.

2.1.1 Dimensionality Reduction via PAA

In order to produce SAX representation, the original time series data is first transformed into PAA representation as is shown in Fig. 2.3. Each subsequence of time series data is divided into k segments with equal length and the average value of each segment is used as a coordinate of a k-dimensional feature vector.

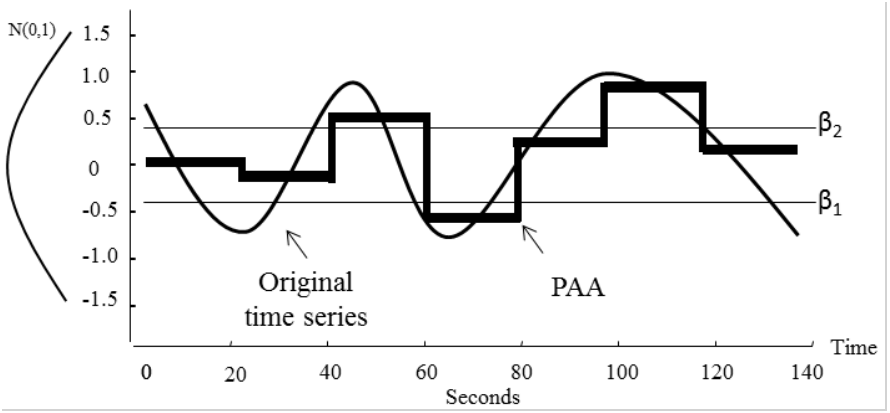

Fig. 2.3. PAA transformation. Breakpoints (β1=-0.43, β2=0.43) discretize PAA coefficients to 3 symbols with equal probability under normal distribution.

2.1.2 Discretization

Given that time series data from the UCR public data source are normalized and have Gaussian distribution, equal-sized areas under normal distribution can be determined by breakpoints [2]. After time series data is transformed into PAA, a breakpoint table, as is shown in Table 2.1 [2], is used to discretize PAA coefficients into symbols with equal probability.

Table 2.1. Lookup table that contains breakpoints that divide Gaussian distribution in an arbitrary number (from 3 to 6) of equal probability regions

α / β_i	3	4	5	6
β_1	-0.43	-0.67	-0.84	-0.97
β_2	0.43	0	-0.25	-0.43
β_3		0.67	0.25	0
β_4			0.84	0.43
β_5				0.97

As is shown in Fig. 2.4 [2], all PAA coefficients that are below the smallest breakpoint are mapped onto symbol a, all coefficients greater than or equal to the smallest breakpoint and smaller than the second smallest breakpoint are mapped onto b, etc. In this way, time series data is transformed to SAX representation.

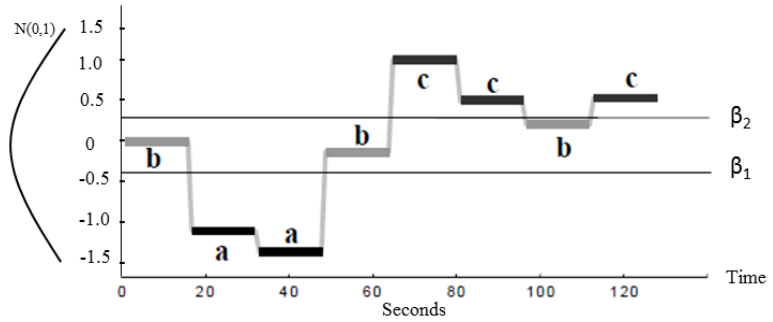

Fig. 2.4. Symbolic discretization. Breakpoints (β1=-0.43, β2=0.43) discretize PAA coefficients to 3 symbols with equal probability under normal distribution.

2.1.3 Distance Measure of SAX

The MINDIST function,

$$MINDIST(\hat{Q},\hat{C}) = \sqrt{\frac{n}{w}}\sqrt{\sum_{i=1}^{w}(dist(\hat{q}_i,\hat{c}_i))^2} \quad (2.1),$$

is defined to calculate the distance between two SAX representations, where n is the length of the original time series, and w is the length of symbolic representation. The function dist() is the distance between the two corresponding symbols of SAX strings. The computation of this function is based on a pre-computed lookup table, as is illustrated in Table 2.2.

Table 2.2. Lookup table used by distance measure function MINDIST. This table is for circumstances in which SAX alphabet cardinality is 4. The distance between every two symbols corresponds to each value in the table.

	a	b	c	d
a	0	0	0.67	1.34
b	0	0	0	0.67
c	0.67	0	0	0
d	1.34	0.67	0	0

The merits of SAX representation of time series data are dimensionality reduction and lower bounding. SAX achieves dimensionality reduction through its first transformation step, PAA representation. The computation process is very fast and easy. The dimensionality reduction characteristic of SAX avoids circumstances where data mining algorithms degrade their performance with the high dimensionality of original time series data.

2.1.4 SAX Related Research

Since the birth of SAX, it has been attracting a lot of attentions in time series data mining research community due to its own significant advantages compared to other already well-known time series approximation approaches. Enormous research about algorithms based on SAX and application that utilizing SAX have been proposed by its original proposer and huge amount of other time series data mining researchers.

SAX is applied to find time series discords in [12], and to create time series bitmaps in [13]. In [14], time series motif is found for the first time by SAX. A blinding fast probabilistic algorithm is proposed in [15]. Meaningful clustering is performed in [16]. Applications, such as shape mining [17], shape motif discovery [18], and image mining [19] are proposed, implemented in various ways. Also, several algorithms are proposed

based on SAX to highlight its superiority or avoid its deficiency [3][4][20][21]. Besides, SAX is also applied in many application fields, such as medical area [22][23], motion/gait detection [24], meteorology [25], and financial data mining.

2.2 SAX with Direction Representation

2.2.1 Direction Representation

The advantages of SAX have made it applicable to many application fields. If the subsequence directions of original time series data, however, could be combined into SAX distance measure, the weak lower bound phenomenon, which means that measured distance between two SAX representations is much smaller than the Euclidean Distance between their original time series data, would be avoided. As a result, the distance measure can eliminate some errors and improve the accuracy of time series data mining tasks, such as classification, clustering, and motif discovery.

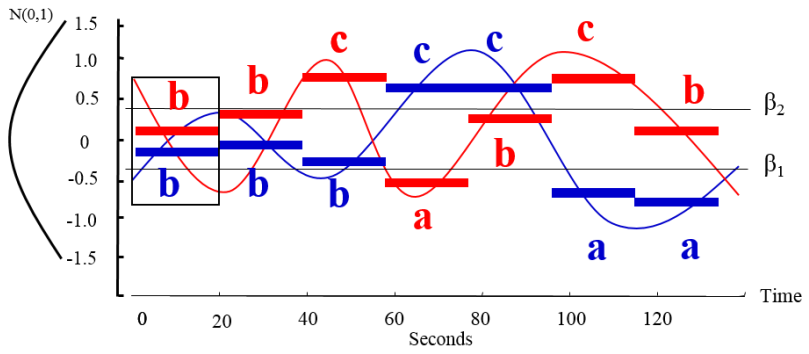

Fig. 2.5. False alarm example. Breakpoints (β1=-0.43, β2=0.43) discretize PAA coefficients to 3 symbols with equal probability under normal distribution.

For example, as shown in Fig. 2.5, although the first symbol for both the blue and the red time series data representations are ‘b’s, the corresponding subsequence of original

time series data has completely different directions. MINDIST distance between the first two ‘b’s is zero, yet the Euclidean Distance between original data is much larger than zero. Unnecessary mistakes appear in time series data mining tasks. For instance, if a query is conducted on the red SAX representation, some inappropriate subsequences from the blue SAX representation are possible to be returned as match results, even they have completely different directions and their Euclidean distance is rather large. Yet, such returned results can be hardly considered as a successful match.

A distance measure is proposed for time series data mining based on SAX with Direction Representation. Subsequence direction is also an important feature for original data and should not be ignored. It is considered and added to original SAX representation. Each subsequence of time series data is mapped onto one of three direction types: convex, concave, and linear. The way to get direction representation of each subsequence is to first calculate the number of positive/negative slopes between every two subsequence data points. The direction type of each subsequence is then obtained based on production rules. Not only is the original time series transformed to a series of symbols, but also to a series of direction representations.

With direction representation, distance between two symbols in SAX representation with different directions will not be overlooked. The tightness of the lower bound is thus increased compared with that of original SAX. Unnecessary error in time series data mining tasks can be avoided. The error rate of time series data mining tasks, for example, classification, could then be decreased.

As is shown in Fig. 2.6, each subsequence of original time series data is mapped onto one of three direction types: convex, concave, and linear.

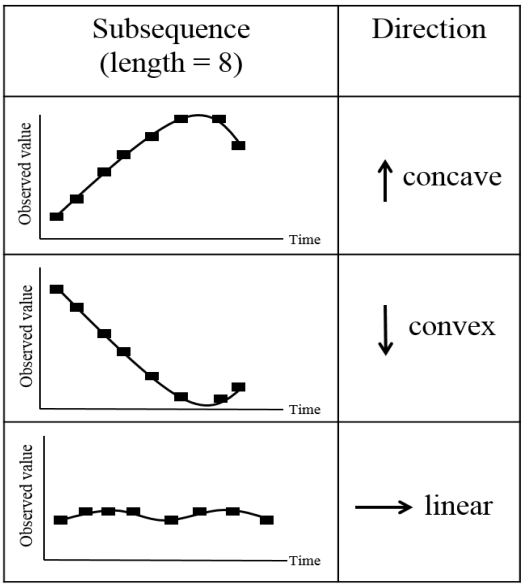

Fig. 2.6. Direction representation. In the above example, the subsequence length is 8.

The flowchart for calculating the direction type of each subsequence is described in Fig. 2.7.

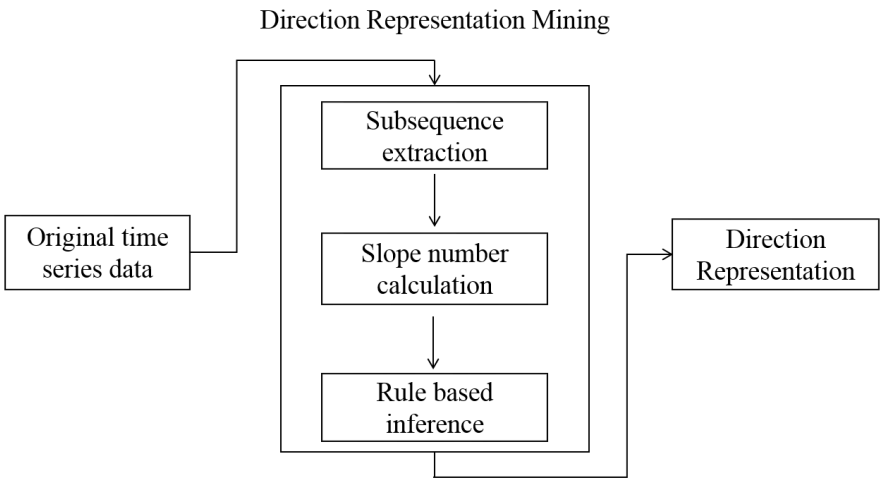

Fig. 2.7. Flowchart for producing direction representation

In order to mine direction representation, in the first step, subsequence is extracted from input time series data. Second, numbers of positive/negative slopes between every two data points of original time series subsequence are calculated in the slope number calculation step. In the last step, production rules, as is illustrated in Table 2.3, are applied to produce direction representation for each subsequence. Through the direction representation mining process, high dimensionality is effectively reduced and direction representation is acquired along with symbolic representation.

Table 2.3. Production rules to determine the subsequence direction

Production rules (subsequence length = 8)
IF positive slopes are more than 4 THEN the subsequence is mapped onto 'concave' IF negative slopes are more than 4 THEN the subsequence is mapped onto 'convex' IF positive slopes <=4 AND negative slopes <= 4 THEN the subsequence is mapped onto 'linear'

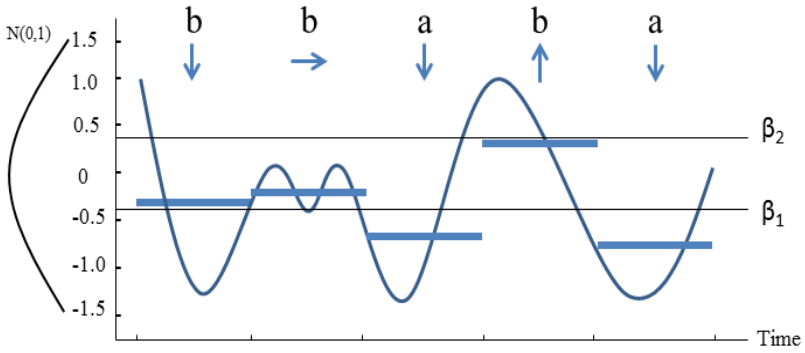

Fig. 2.8. Example of SAX with direction representation. Breakpoints (β1=-0.43, β2=0.43) discretize PAA coefficients to 3 symbols with equal probability under normal distribution.

In the proposal, not only is the original time series transformed into a series of symbols but the direction representation of each subsequence is also obtained along with SAX representation, as is shown in Fig. 2.8.

2.2.2 Distance Measure of SAX with Direction Representation

The distance measure of the proposed SAX with direction representation uses both the original SAX lookup table and the directional distance lookup table. The distance between two SAX symbols is still calculated through the MINDIST function by using the original SAX pre-computed lookup table, as is illustrated in Table 2 [2]. The DIRDIST function is defined in the proposal to calculate the distance between two different directions. This calculation is also based on a pre-computed lookup table, as shown in Table 2.4, which is calculated using three common values of their corresponding time series subsequence.

Table 2.4. Distance lookup table between different directions

	Convex ↓	Linear →	Concave ↑
Convex ↓	0	(VEXMAX – VEXMEAN)	(VEXMAX – VEXMEAN) + (CAVEMEAN – CAVEMIN)
Linear →		0	(CAVEMEAN – CAVEMIN)
Concave ↑			0

The DIRDIST(dir1, dir2) function,

$$DIRDIST(dir1, dir2) = \sqrt{\frac{n}{w}} \sqrt{\sum_{i=1}^{w} (dir(dir1_i, dir2_i))^2 / w} \tag{2.2}$$

is defined to compute the distance between direction representations of two SAX strings, where dir1 and dir2 are direction representations and calculation of function dir() is implemented using the pre-computed lookup table for distance measure between different directions.

The pseudo code for distance calculation of the proposed distance measure is shown in Table 2.5.

Table 2.5. Pseudo code of the proposed distance measure

Pseudo code
function dist = dist_direction(str1, str2, dir1,dir2,alphabet_size,compression_ratio) % Input: str1, str2 are two SAX strings % Input: dir1, dir2 are the direction % representations % Input: alphabet size % Input: compression ratio % Output: the distance between two SAX % strings with direction representation % string distance dist_str = MINDIST(str1, str2) % direction distance dist_dir = DIRDIST(dir1,dir2) return dist = dist_str + dist_dir

The proposed distance measure has two main advantages. First, the tightness of Euclidean Distance is increased. Second, error rate of time series data mining tasks is reduced. The reason for increased tightness is that the direction factor of each subsequence is not ignored. Instead, it is considered and added to original SAX as an

important feature. The distance between two symbols that have two different directions is calculated using the lookup table, as is shown in Table 2.4, and added to the distance computed from MINDIST. Compared with original SAX, the proposed distance measure avoids unnecessary errors. It is, therefore, possible to lower the error rate of time series data mining tasks, such as classification and clustering.

2.3 Experiments on UCR Time Series Datasets

The experimental datasets from the UCR Time Series public service data source [8] are used to evaluate the performance of the proposal. Detailed information is listed in Table 2.6.

Table 2.6. Experiment datasets from UCR Time Series Public Data Source

Dataset	Number of class	Size	Length	Type
50words	50	455	270	Real
CBF	3	900	128	Synthetic
ECG200	2	100	96	Real
FaceFour	2	150	150	Real
Gun_Point	6	242	427	Shape
OSULeaf	15	625	128	Shape
SwedishLeaf	4	100	275	Synthetic
Trace	4	4000	128	Synthetic
TwoPatterns	2	6164	152	Real
wafer	2	3000	426	Shape
yoga	4	88	350	Shape

This dataset contains very diverse time series data. The type of several datasets is synthetic. For instance, CBF was created by researchers to test some properties of certain time series classification algorithms. Some real datasets are recorded as natural time series from physical processes. Several datasets are also shape data, such as SwedishLeaf and OSULeaf, which are one-dimensional time series extracted by processing two dimensional shapes [9]. These public datasets have been widely accepted and applied by many time series data miners.

Experiments are conducted on a PC with a dual core processor (2.5GHz) and 2GB memory. Simulation software is from MATLAB.

2.3.1 Tightness comparison between the proposal and SAX

Tightness is the distance between two approximation representations divided by the Euclidean Distance between their original time series.

The tightness of the proposal and MINDIST of SAX are contrasted in the experiment. Comparison indicates that the proposed distance measure has a higher tightness than that of original SAX.

In the tightness experiment, the window size is set to 8, which means that each symbol in the SAX string contains 8 data points of original time series data. Alphabet sizes from 3 to 10 are tested separately. The tightness increase percentage of each alphabet size is shown in Fig. 2.9. The tightness of the proposed distance measure increases under all alphabet size scenarios compared with MINDIST of SAX. It is shown in Fig. 2.9 that since the direction distance between two symbols with different directions is added to the distance between symbols that is calculated by MINDIST, the tightness of the proposed distance measure is higher than MINDIST of SAX on each tested dataset. Specifically, the tightness of the proposed distance measure is 17.54% greater than MINDIST of original SAX.

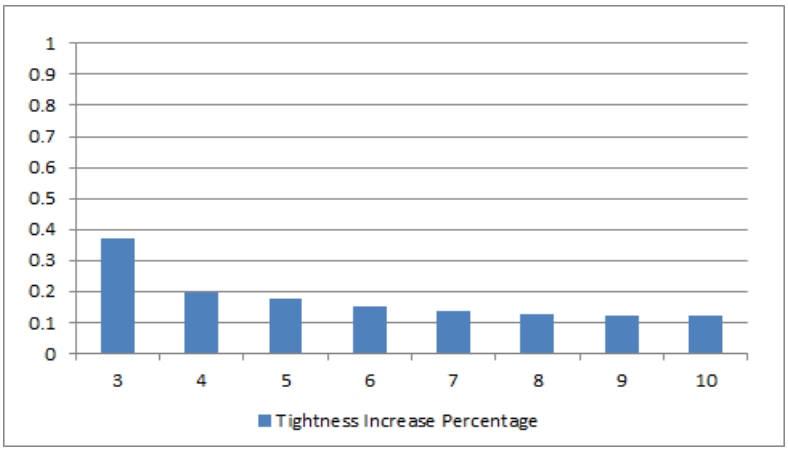

Fig. 2.9. Tightness increase percentage

Specific tightness experiment results in which alphabet sizes are set to 3 to 10 are illustrated in Fig. 2.10 to Fig. 2.17, individually.

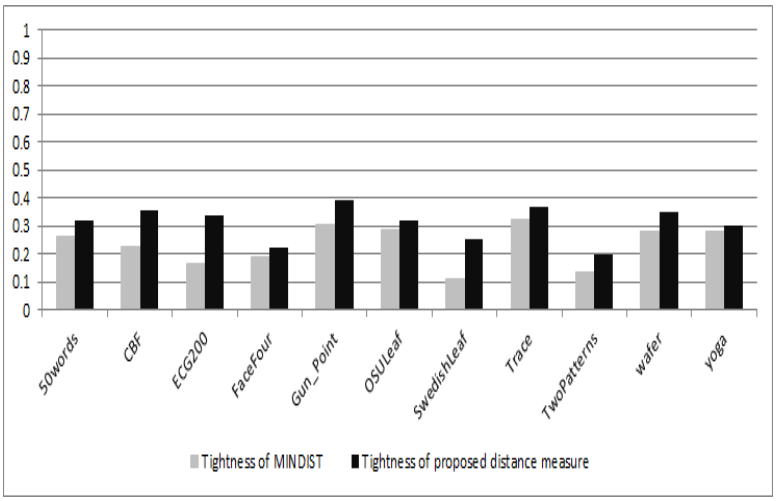

Fig. 2.10. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 3)

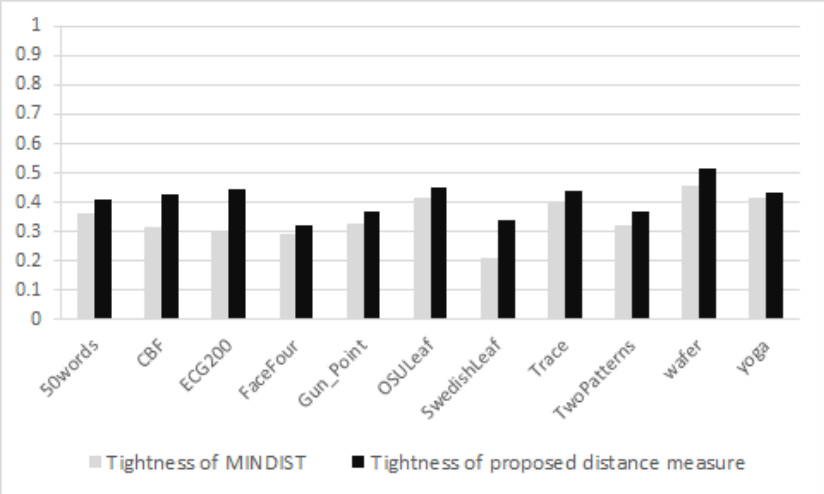

Fig. 2.11. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 4)

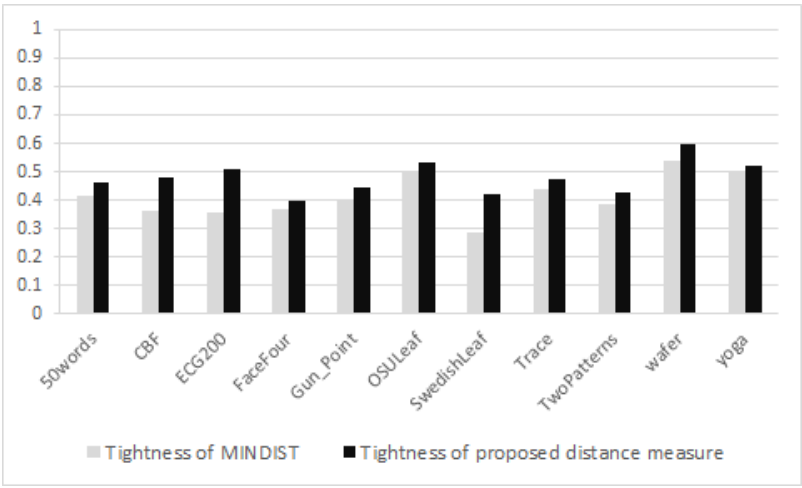

Fig. 2.12. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 5)

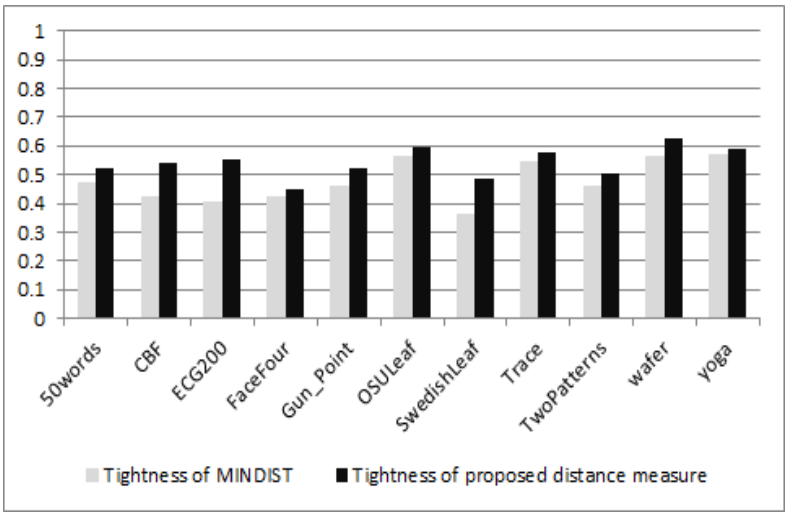

Fig. 2.13. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 6)

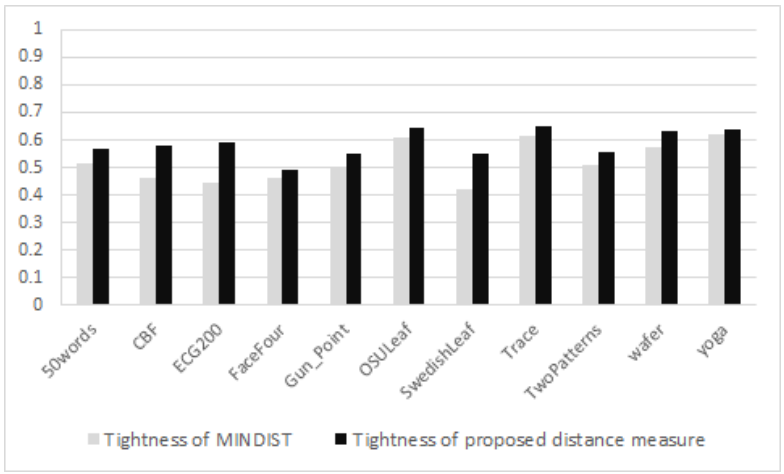

Fig. 2.14. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 7)

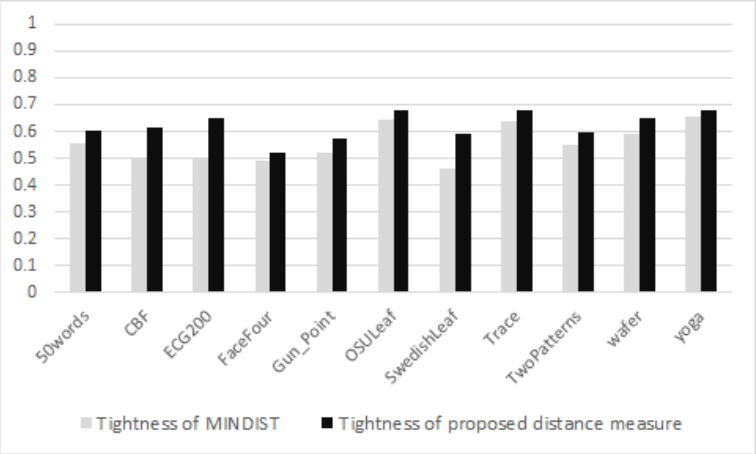

Fig. 2.15. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 8)

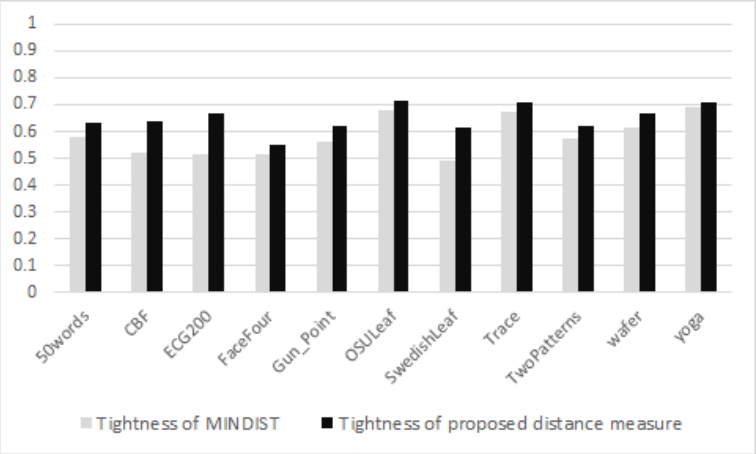

Fig. 2.16. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 9)

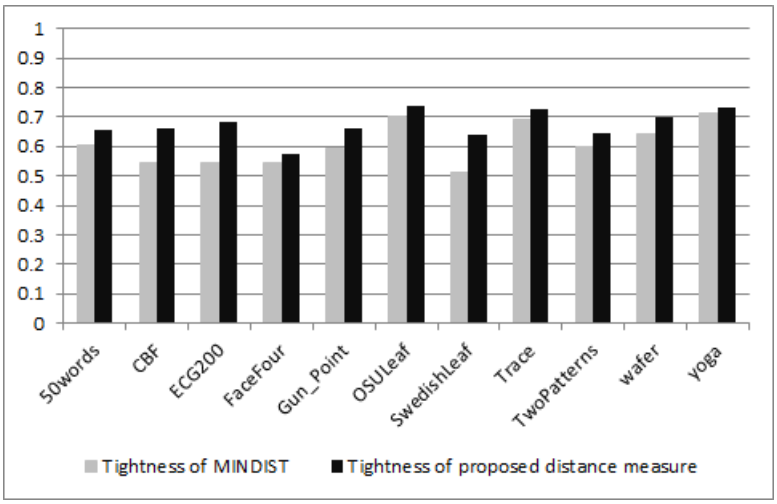

Fig. 2.17. Comparison of tightness between MINDIST and the proposed distance measure (alphabet_size = 10)

2.3.2 Classification on UCR Time Series Datasets

To confirm the proposal's advantage, the low classification error rate, the one-nearest-neighbor classification is performed to compare the error rate of the proposed distance measure with that of MINDIST of SAX.

In the classification experiment, the window size is still set to 8 and the alphabet size varies from 3 to 10. The error rate reduction percentage on each alphabet size setting is shown in Fig. 2.18. It indicates that the classification error rates of all alphabet size scenarios are significantly decreased. Statistically, the classification error rate of SAX with direction representation is reduced by 16.22% on average when compared with that of original SAX on UCR time series dataset.

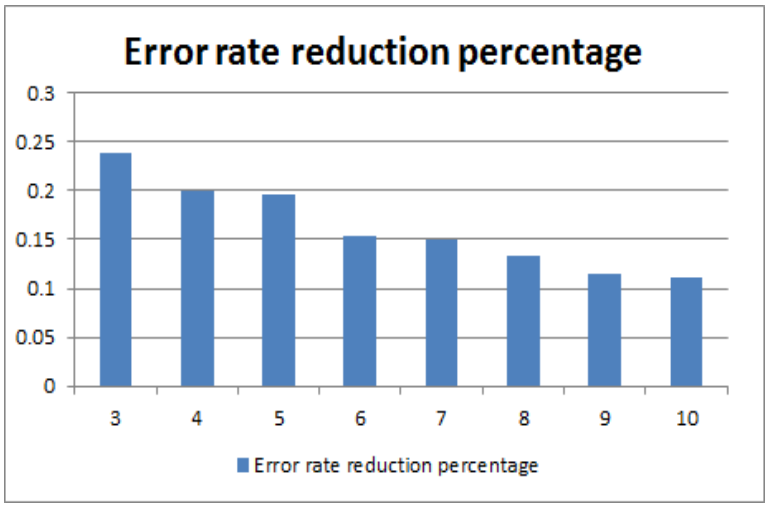

Fig. 2.18. Classification error rate reduction percentage

Detailed classification experiment results on UCR time series datasets, as is shown in Fig. 2.19 to Fig. 2.26, demonstrate that the error rates of the proposed distance measure on most different datasets are lower than that of MINDIST of original SAX.

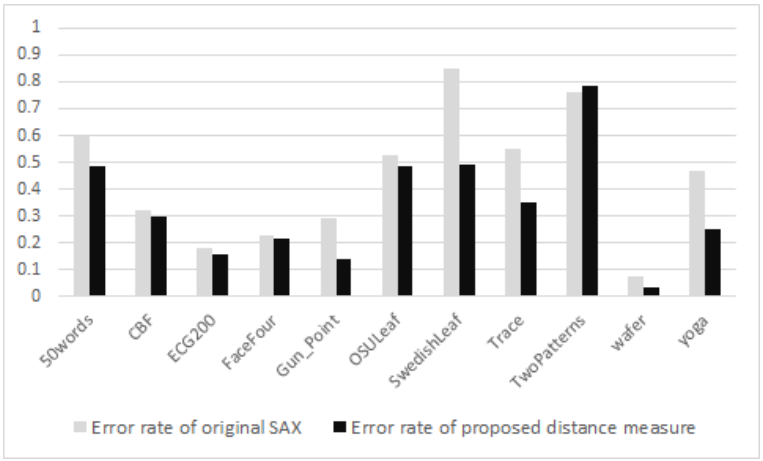

Fig. 2.19. Comparison of classification error rate between the proposal and original SAX (alphabet size = 3)

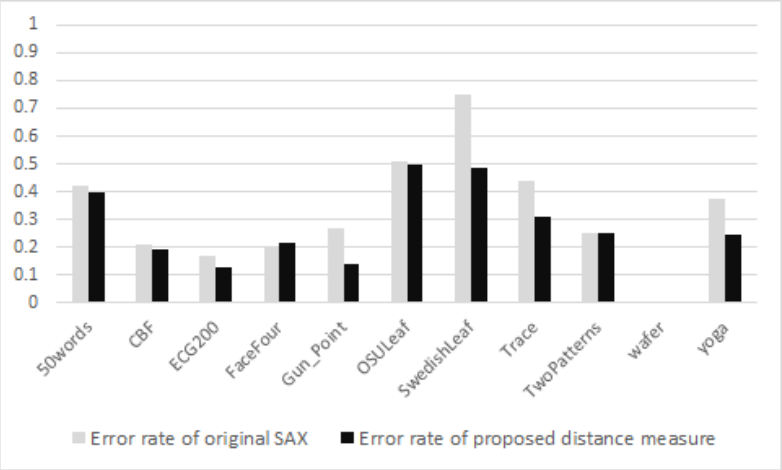

Fig. 2.20. Comparison of classification error rate between the proposal and original SAX (alphabet size = 4)

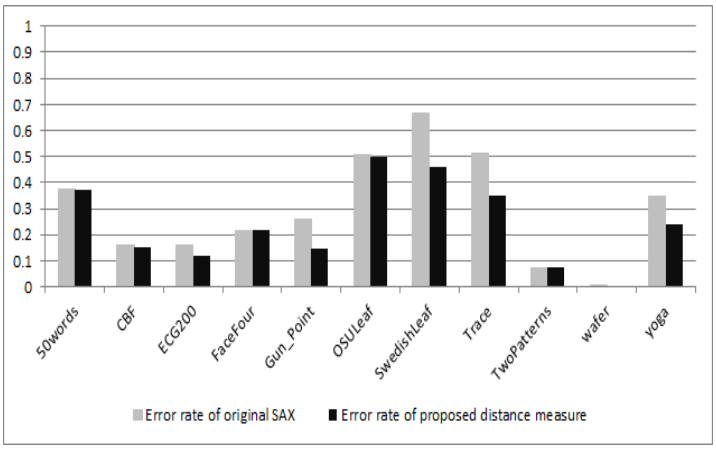

Fig. 2.21. Comparison of classification error rate between the proposal and original SAX (alphabet size = 5)

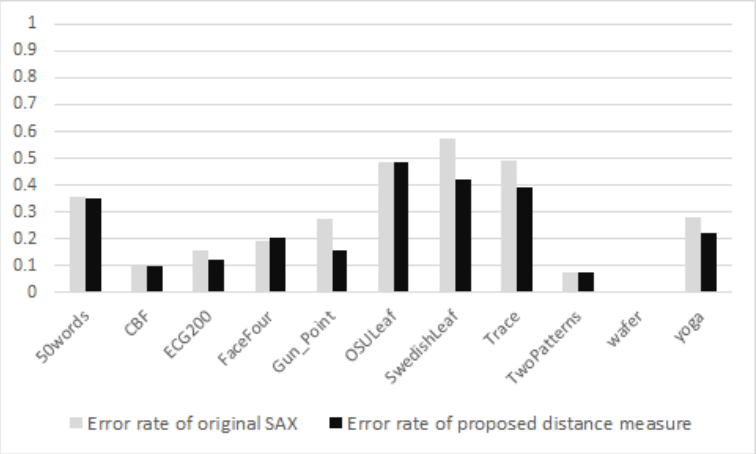

Fig. 2.22. Comparison of classification error rate between the proposal and original SAX (alphabet size = 6)

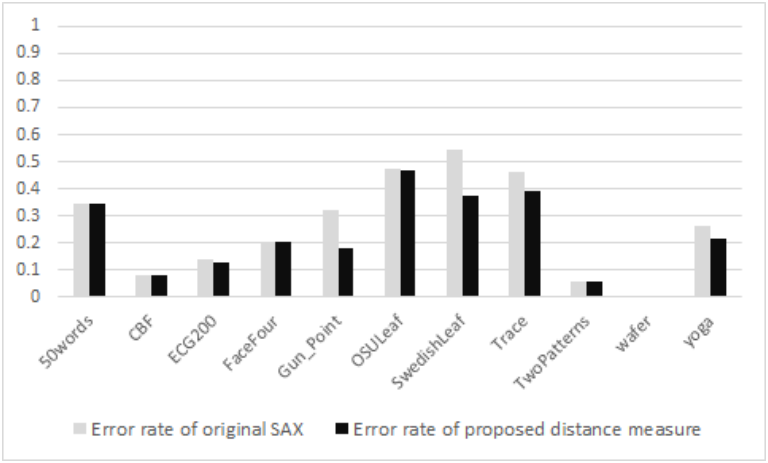

Fig. 2.23. Comparison of classification error rate between the proposal and original SAX (alphabet size = 7)

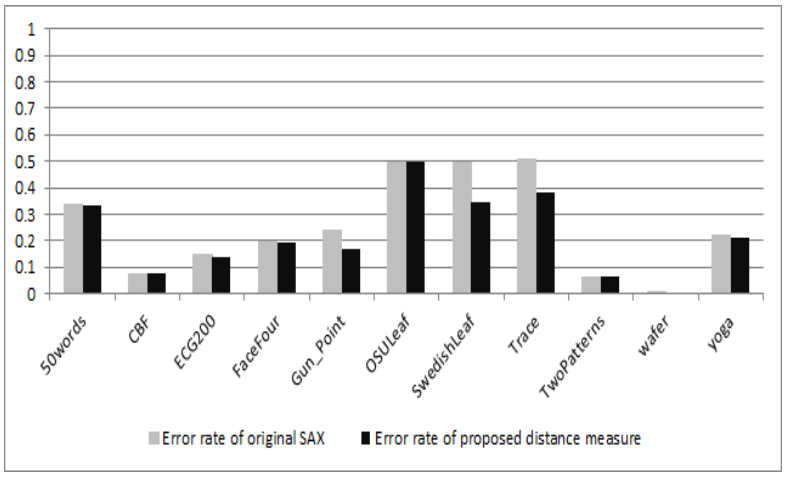

Fig. 2.24. Comparison of classification error rate between the proposal and original SAX (alphabet size = 8)

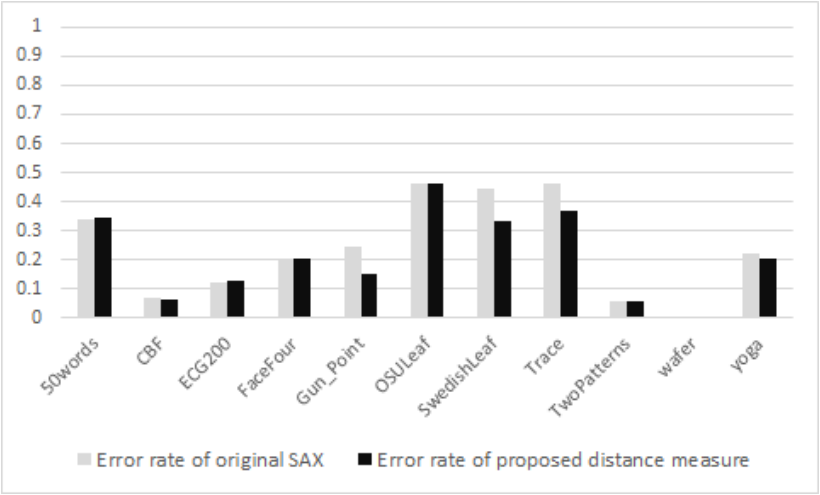

Fig. 2.25. Comparison of classification error rate between the proposal and original SAX (alphabet size = 9)

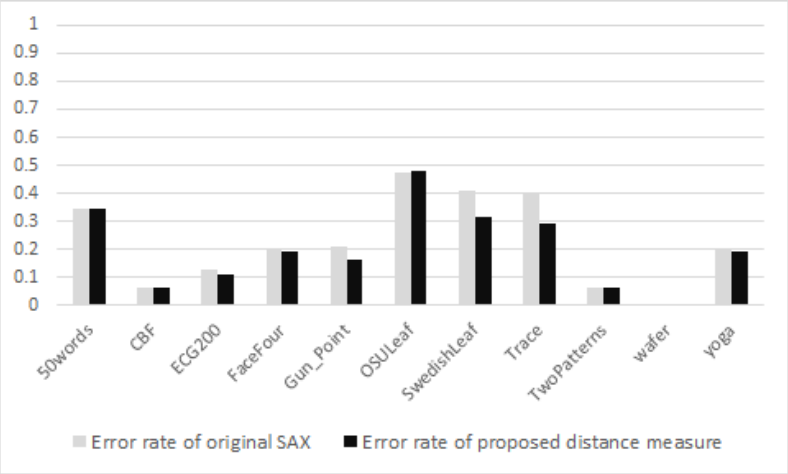

Fig. 2.26. Comparison of classification error rate between the proposal and original SAX (alphabet size = 10)

The classification error rate reduction percentage chart also indicates that the error rate reduction percentage decreases as the alphabet size increases. The reduction of the classification error rate when the alphabet size is 3 is much more effective when the alphabet size is 10. When the alphabet size is set to 10, the error rate of the proposed distance measure becomes close to that of original SAX.

2.3.3 Classification result difference analysis

In the classification experiment, the error rate is effectively reduced on most test datasets, and the average reduction is 16.22%. It is also noted, however, that the error rate is not effectively reduced on all test datasets under every alphabet size scenario. Conversely, there are a few unexpected exceptions, the error rates of the proposed distance measure on some datasets actually increased compared to that of SAX. Despite being a very small portion of all test dataset scenarios, it is worth questioning why this phenomenon is happening. The above question can be transformed into asking why the proposed distance measure produces better experimental results on some of the test datasets, such as dataset SwedishLeaf, and yoga, than some other datasets, like 50words.

In order to make the proposed distance measure more convincing, it is necessary to find out the reason for and explanation to the emergence of imperfect experiment results. Since a comparison with original SAX, the only difference in the proposed distance measure is that the time series subsequence direction factor is considered, and integrated into SAX. The direction factor of the original time series needs to be analyzed.

Datasets SwedishLeaf and 50words are chosen to be representative datasets that have relatively opposite results. The error rate of the proposed distance measure is nearly the same or higher than that of SAX on 50words, while no matter what the alphabet size is, all error rates on SwedishLeaf are effectively reduced. The original time series data of these two datasets may therefore have the most different features, and could help us to analyze the classification error rate difference. Ten objects are extracted from test and train datasets of SwedishLeaf and 50words. The extraction results are shown in Figs. 2.27 and 2.28.

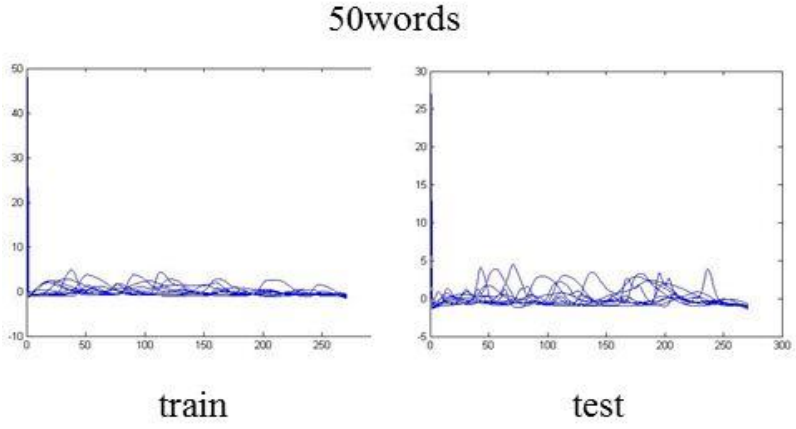

Fig. 2.27. Ten objects from 50words train and test datasets

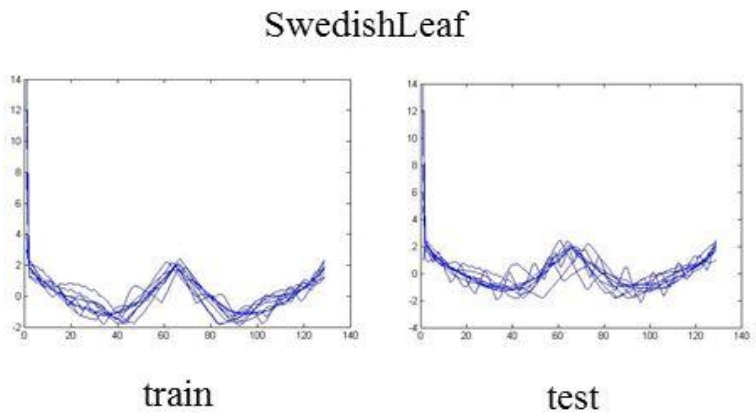

Fig. 2.28. Ten objects from SwedishLeaf train and test datasets

It is clearly shown from these two figures that the original time series data in the 50words dataset is much more smoothly and without so many ups and downs compared to SwedishLeaf. In contrast, the SwedishLeaf time series data changes more obviously and frequently and has greater amplitude. Since the subsequence direction factor is added into SAX in our proposal, the proposed distance measure is more effective on datasets that have more direction changes. The classification error rate should be reduced more when the proposed distance measure is performed on such datasets. The above analysis may explain the reason behind the phenomenon that there are experiment result differences on certain test datasets.

In order to confirm the statistical significance of the difference produced by the proposal, a paired t-test is conducted on the classification error rate results of both the proposal and SAX. The significance level is commonly used 0.05 in the t-test. The p value of the t-test is 7.9×10-9, which indicates that the null hypothesis that the proposed distance measure has no effect is rejected and the results produced by the proposal are statistically significant.

2.3.4 Classification computational cost

Since the proposed distance measure adds some additional calculations based on SAX, it is necessary to prove that the computational cost of the proposal does not increase dramatically in comparison with that of SAX. Otherwise, the proposal may be assumed to lack practical meaning in a real application. To eliminate such an unnecessary assumption, comparison on classification computational cost between the proposed distance measure and SAX is listed in Table 2.7.

Table 2.7. Classification computational cost comparison (alphabet size = 3)

Dataset name	Size of train dataset	Size of test dataset	Time per test object (SAX)	Time per test object (proposal)
50words	450	455	0.0**3**1579912	0.0**4**29612
CBF	30	900	0.01**2**707149	0.01**5**9514
ECG200	100	100	0.01**1**560170	0.01**8**7156
FaceFour	24	88	0.00**9**211625	0.01**0**7341
Gun_Point	50	150	0.0**0**6034607	0.0**1**14720
OSULeaf	200	242	0.02**1**589835	0.02**9**2014
SwedishLeaf	500	625	0.0**3**2101755	0.0**4**72058
Trace	100	100	0.0**1**2928440	0.0**3**23542
TwoPatterns	1000	4000	0.0**5**6277311	0.0**7**58873
wafer	1000	6164	0.0**5**5847929	0.0**7**06123
yoga	300	3000	0.0**2**7164482	0.0**5**17317

From Table 2.7, it is shown that on 7 out of 11 test datasets, the computational cost for each test subject increases at the one hundredth level, while every test subject's computational cost increases at one thousandth level on the other 4 datasets. It is obvious that the computational cost increase of the proposed distance measure is in an acceptable range.

2.4 Perspective in Pattern Recognition

Pattern recognition is the problem of identifying, given a query sequence and a database of sequences, the database subsequence that best matches the query sequence. Motivation applications include keyword-based query, DNA/protein matching, video surveillance, and query by humming, etc. The followings are common similarity measures used in time series data mining. For example, Euclidean Distance, Dynamic Time Warping (DTW), Edit distance with Real Penalty (ERP), Longest Common Subsequence (LCSS), and Edit Distance on Real sequence (EDR). They determine similarity of two time series by comparing their individual point values. LCSS finds the length of the longest matching subsequences. ERP is defined as the minimum number of edits, such as, insertion, deletion, and substitution, needed to transform a string to another. DTW searches for the best alignment between two time series, attempting to minimize the distance between them.

However, current similarity measures, like Euclidean Distance, even Dynamic Time Warping, that used in pattern recognition cannot provide high satisfaction in terms of speed, since the characteristic that calculating distance between all individual data points.

Direction representation can be applied in pattern recognition to help prune some impossible subsequence so that the detection speed is able to be accelerated promisingly. Direction-based approach is particularly useful in applications where domain experts compare signals based on the arrangement of morphological events present in signals, for instance, in radar signal detection and speech recognition. Therefore, the proposal is pattern recognition based on direction representation with warping characteristic.

The flowchart of pattern recognition using direction representation with warping characteristic is shown in Fig. 2.29.

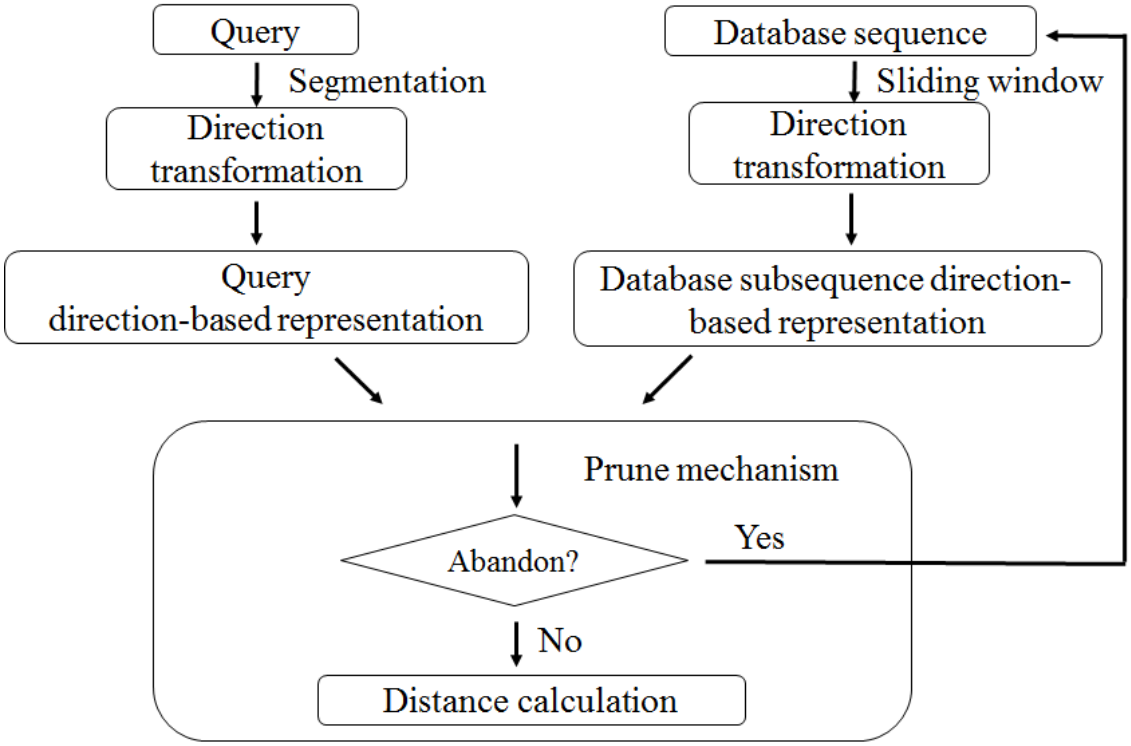

Fig. 2.29. Flowchart of direction representation with warping characteristic

Query sequence is first transformed to direction-based representation query through segmentation and direction transformation. Afterwards, sequence in database is similarly converted to subsequence via sliding window step and direction transformation. Then, to determine the similarity between query and database subsequence, the distance between their direction-based representations need to be calculated. Yet, before the calculation, there is a very important step. Some unnecessary calculations are going to be skipped by the proposed prune mechanism. This step is the key to speed up the pattern recognition. Through the prune step, if database subsequence is decided to be abandoned, then next subsequence is acquired to continue the recognition process. Otherwise, the distance between direction representations of query and database subsequence is computed.

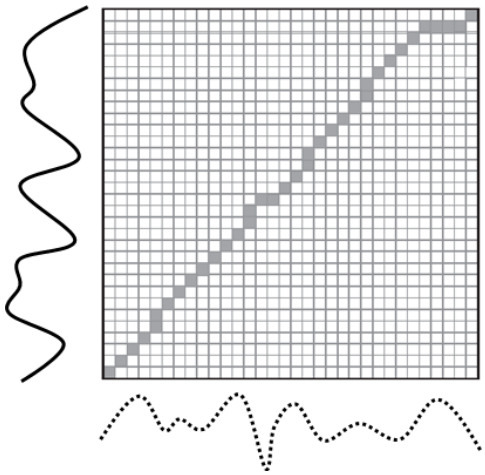

Fig. 2.30. Matrix construction of DTW

In the following paragraphs, the prune mechanism that supposed to be applied are discussed in detail. The warping feature in prune approach is derived from DTW. Dynamic time warping is another robust distance measure for time series data mining, allowing similar shape to match even they are out of phase in time axis. To align two sequences using DTW, a matrix, illustrated in Fig. 2.30, is built in which each element is the distance between two data points. The goal is to find a warping path that can minimize the warping cost of Equation 2.3. The path search is able to be completed by dynamic programming in Equation 2.4.

$$DTW(Q, C) = \min \left\{ \sqrt{\sum_{k=1}^{K} w_k} . \right.$$

(2.3)

$$\gamma(i, j) = d(q_i, c_j) + \min \{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\}.$$

(2.4)

Same as DTW, a warping path that minimizing the warping cost between two direction-based representations is expected to be found. Yet, the difference is that instead of distance of two data points, the matrix element in prune mechanism is the

distance of two direction based representations as shown in Fig. 2.31. If two direction based representations match, their distance will be calculated. Otherwise, the distance is assigned as positive infinity as Equation 2.5.

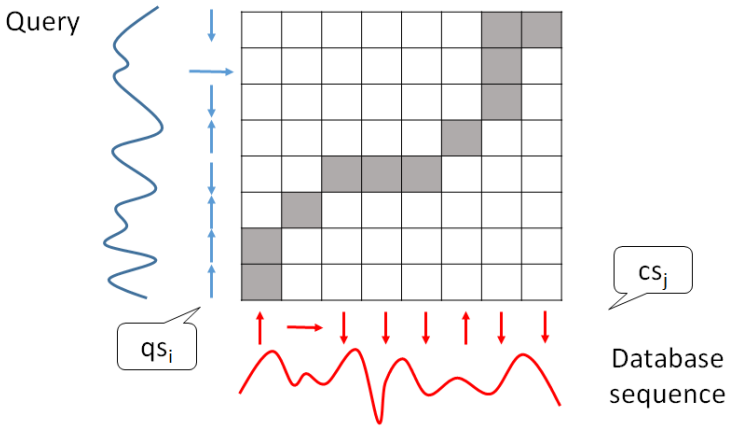

Fig. 2.31. Warping path in prune mechanism

$$\mathrm{dis}(qs_i, cs_j) = \begin{cases} Dis(qs_i, cs_j), \ qs_i \ and \ cs_j \ match \\ +\infty, otherwise \end{cases} \tag{2.5}$$

During the warping calculation, the prune mechanism is performed to avoid unnecessary computation. The principle is that, if there is any direction based representation that cannot find its own match within the constraint area as shown in Fig. 2.32, the distance calculation is going to be terminated. The next subsequence from database sequence is then performed direction representation transformation, and applied to conduct next distance measure. Therefore, the unnecessary computation of this pattern recognition process is eliminated so that the speed of pattern detection is dramatically increased.

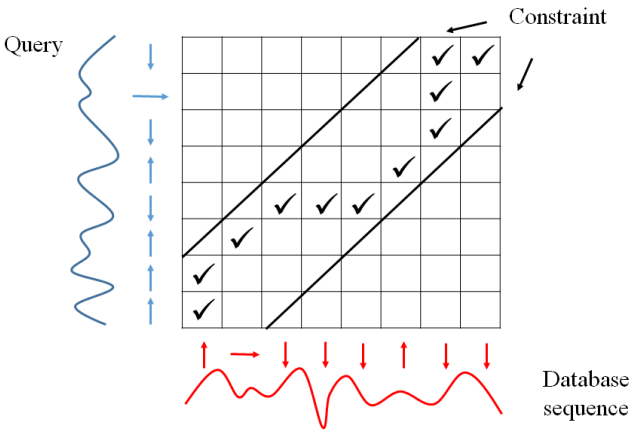

Fig. 2.32. Constraint in prune mechanism

2.5 Chapter Summary

In the tightness experiment, the compression rate of original time series data is 8 and alphabet sizes vary from 3 to 10. Experiment results intuitively show that under each circumstance, the tightness of the proposed distance measure is higher than that of original SAX. Statistically, the tightness of the proposed distance measure is on average 17.54% tighter than that of MINDIST on each experiment dataset.

In the classification experiment, the one-nearest-neighbor classification is conducted on both the proposal and original SAX. The error rate of the proposed distance measure and that of MINDIST of SAX are contrasted to confirm that the proposal produces a lower classification error rate. The compression rate of original time series data is 8 and the alphabet size is set from 3 to 10. Experiment results demonstrate that the proposal has a smaller classification error rate on most UCR time series datasets. Specifically, it is on average 16.22% lower than MINDIST of SAX on each tested dataset.

The reason for classification result differences is analyzed by comparing the original time series data from two datasets, 50words and Swedishleaf, which have relatively opposite classification results. Ten objects are extracted from both train and test datasets. It can be said that the proposed distance measure has more effective experimental

results on a dataset in which the original time series data has frequent direction changes and more up and downs. From the classification error rate reduction percentage figure, it is shown, moreover, that the reduction percentage decreases while alphabet size adding. The computational cost comparison between SAX and the proposal is also done. The comparison proves that the proposed distance measure does not cause enough additional computational cost to compromise its practical value.

The proposed distance measure increases the lower bound tightness and reduces the error rate of classification task. SAX with direction representation aims to be an effective approximation representation method for time series data mining tasks, such as clustering, query by content, and motif discovery, in distinct application areas. One of the promising application areas is motion recognition [10]. The proposed distance measure can be applied to classify different motion situations such as walking, jogging, and falling. Such applications could be deployed in a portable device to recognize a person's activity. Another prospective application field is the medical area [11]. Lots of biomedical signals are recorded as time series data, for instance, ECG data. The proposal may help to identify some diseases via conducting classification, motif discovery.

Chapter 3

Ischemia Diagnosis based on Fuzzy Association Rule Mining

3.1 Electrocardiography and Heart disease

3.1.1 ECG Signal

As one of the most important signals produced by our body, Electrocardiography (ECG) is a transthoracic interpretation of the electrical activity of the heart, as detected by electrodes attached to the surface of the skin and recorded by a device external to the body, as shown in Fig. 3.1. The ECG recording is a noninvasive procedure, and measures the rate and regularity of heartbeats. The common form of ECG signal is shown in Fig. 3.2. The diagnosis of heart abnormalities, or diseases are usually performed on ECG signals [26][27][28].

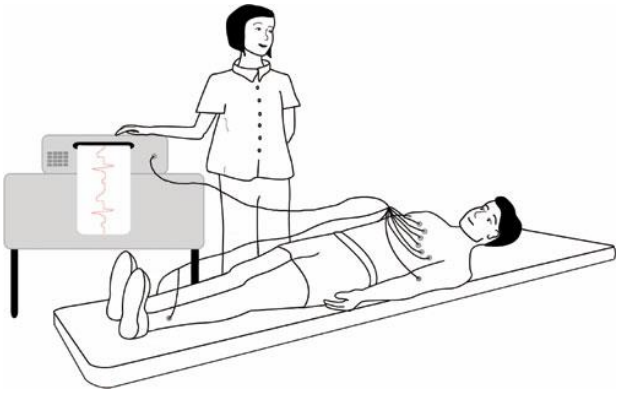

Fig. 3.1. ECG signal recording [26]

Fig. 3.2. ECG signal examples [28]

An ECG is used to measure the electrical conduction system of a heart and collect electrical impulses generated by the polarization and depolarization of cardiac tissue and translates into a waveform, as shown in Fig. 3.3. Sinoatrial node (SA node) is impulse-generating tissue located in the right atrium of the heart. Atrioventricular node, (AV node) is a part of the electrical control system of the heart that coordinates the top of the heart. It connects atrial and ventricular chambers. It is an area of specialized tissue between the atria and ventricles. The waveform is then used to measure the rate and regularity of heartbeats, as well as the size and position of the chambers, the presence of any damage, and the effects of drugs or devices used to regulate the heart [28].

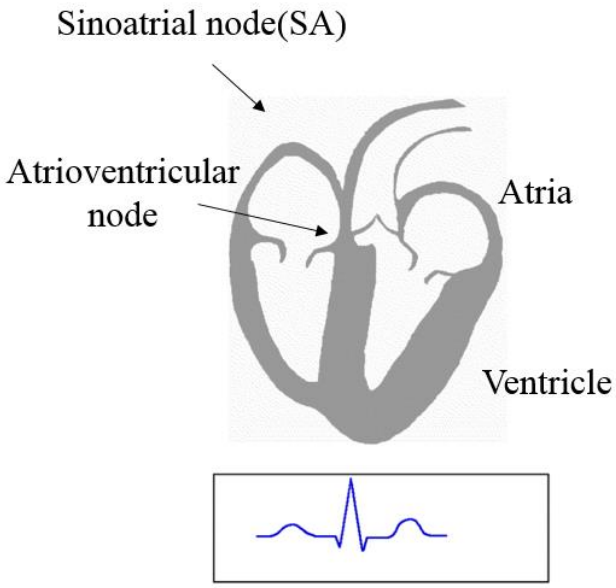

Fig. 3.3. Generation of ECG signal [28]

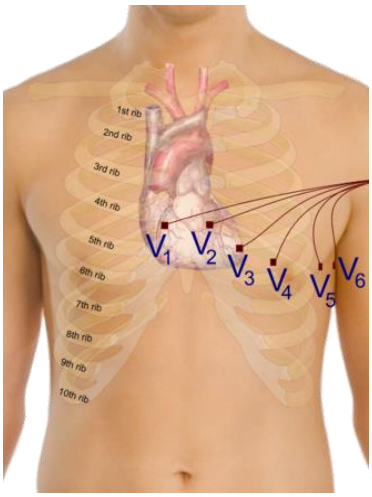

Fig. 3.4. Lead placement in ECG recording [28]

Several electrodes are often used in ECG recording, and they can be combined into a number of pairs. For instance, left arm (LA), right arm (RA), and left leg (LL) electrodes, in Fig. 3.4, form the three pairs LA+RA, LA+LL, and RA+LL. The output

from each pair is known as a lead. A 12-lead ECG is one in which 12 different electrical signals are recorded at approximately the same time and will often be used as a one-off recording of an ECG. The output of an ECG recorder is a graph (or sometimes several graphs, representing each of the leads) with time represented on the x-axis and voltage represented on the y-axis. The term "lead" in electrocardiography usually refers to the tracing of the voltage difference between two of the electrodes and is what is actually produced by the ECG recorder. Each will have a specific name. For example "lead I" is the voltage between the right arm electrode and the left arm electrode, whereas "Lead II" is the voltage between the right arm and the left leg.

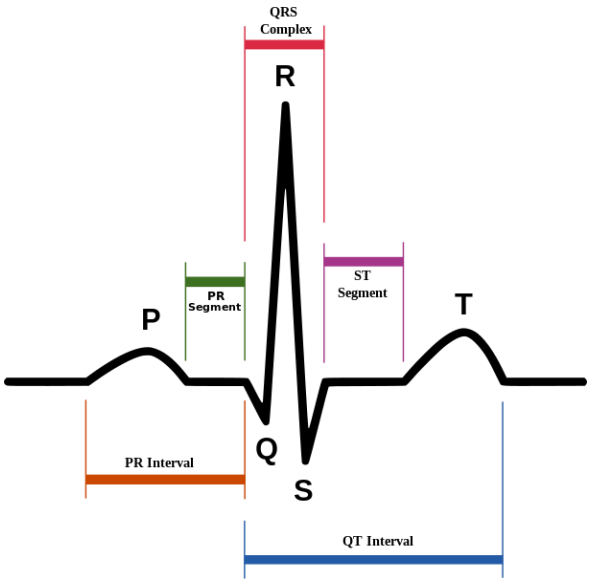

Fig. 3.5. Waves and Intervals within one heartbeat [27]

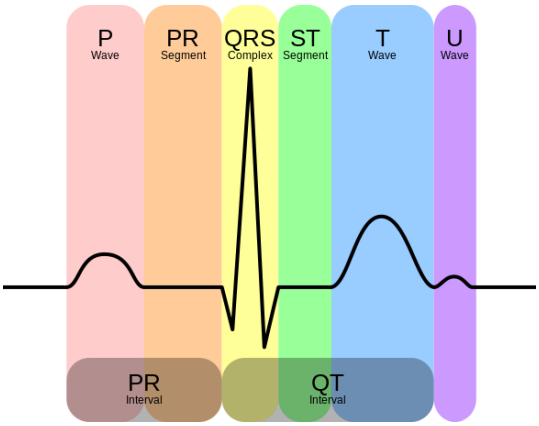

Fig. 3.6. Schematic representation of normal ECG [28]

There are a few extremely significant waves and intervals within ECG recording of one single heartbeat, as shown in Fig. 3.5 and Fig. 3.6. It consists of a P wave, a QRS complex, a T wave, and a U wave, which is usually invisible in more than half of ECGs, due to that it is hidden by the T wave and upcoming new P wave. The baseline of the electrocardiogram is labeled as the portion of the recording following the T wave and preceding the next P wave and the segment between the P wave and the following QRS complex. In a healthy heart, the baseline is equivalent to the isoelectric line (0 mV) and represents the periods in the cardiac cycle when there are no currents towards either the positive or negative ends of the ECG leads.

The following are detailed explanation for different waves and intervals. RR interval is the interval between an R wave and the next R wave. Its duration is normally from 0.6 second to 1.2 second. During a normal atrial depolarization, the main electrical vector is directed from SA node towards the AV node and spreads from the right atrium to the left atrium. This turns into the P wave on ECG signal. PR interval is from the beginning of the P wave to the beginning of the QRS complex. It reflects the time that the electrical impulse takes to travel from the sinus node through the AV node and entering the ventricles. Thus, PR interval is always measured for the estimation of AV node function. PR segment connects the P wave and the QRS complex. The impulse

vector is from the AV node to the Bundle of His, to the bundle branches and then to the Purkinje fibers. This electrical activity does not directly produce a contraction and rarely travel down to the ventricles. QRS complex shows the rapid depolarization of the right and left ventricles. The ventricles have a large muscle mass compared to the atria, therefore, QRS complex normally has a much larger amplitude than P wave. J point, a very important mark point, exists at where QRS complex finishes and ST segment begins. It is used to measure the degree of ST elevation or depression. ST segment connects QRS complex and T wave, and represents the period when the ventricles are depolarized. T wave represents the repolarization of the ventricles. The interval, from the beginning of QRS complex to the apex of T wave, is referred to as the absolute refractory period. The last half of the T wave is referred to as the relative refractory period. ST interval is from the J point to the end of T wave. QT interval, varying with heart rate, is from the beginning of QRS complex to the end of T wave. A prolonged QT interval is a risk factor for ventricular tachyarrhythmias and sudden death. U wave is hypothesized to be caused by the repolarization of the interventricular septum. It always follows T wave with a very low amplitude, or even more often, completely absent. On the other hand, if U wave is too prominent, some heart diseases, such as hypokalemia, hypercalcemia or hyperthyroidism should be suspected and checked.

3.1.2 Heart Disease Diagnosis on ECG Signal

However, in a diseased heart, the baseline may be depressed, such as cardiac ischemia or elevated due to myocardial infarction, relative to the isoelectric line because of injury currents during the TP and PR intervals when the ventricles are at rest. The ST segment typically remains close to the isoelectric line as this is the period when the ventricles are fully depolarized and thus no currents are in the ECG leads. Since most ECG recordings do not indicate where the 0 mV line is, baseline depression often gives the appearance of an elevation of the ST segment and conversely baseline elevation

gives the appearance of depression of the ST segment. To sum up, the waves and intervals on ECG signal become abnormal when someone's heart has certain disease.

Heart disease, also called cardiovascular disease, is a class of diseases that involve the heart, or blood vessels, such as arteries, capillaries, and veins, or both of them. They affect the cardiovascular system, primarily cardiac disease, vascular diseases of the brain and kidney, and peripheral arterial disease. There are diverse elements can cause heart disease, but atherosclerosis and hypertension are the most common ones. Besides, age is another main reason that one gets heart disease. With aging, a number of physiological and morphological changes happen, and alter the cardiovascular functions, then lead to increased risk of heart disease [29].

There are many types of cardiovascular disease. For instance, coronary artery disease, also known as coronary heart disease and ischemic heart disease; Cardiomyopathy, diseases of cardiac muscle; Hypertensive heart disease, diseases of the heart secondary to high blood pressure; Cardiac dysrhythmias, abnormalities of heart rhythm; Peripheral arterial disease, disease of blood vessels that supply blood to the arms and legs; Cerebrovascular disease, disease of blood vessels that supply blood to the brain such as stroke.

Cardiovascular disease is the leading cause of deaths worldwide, though, since the 1970s, cardiovascular mortality rates have declined in many high-income countries. At the same time, cardiovascular deaths and disease have increased at a fast rate in low- and middle-income countries [29].

Since heart disease are the main cause of death all over the world, also heart disease comes with the pathological changes of ECG signal all the time, for example, coronary ischemia cause flattened or inverted T waves, myocardial infarction comes with hyperacute T waves, Hypercalcemia is along with shortened QT interval and Peaked T wave, the research about how to detect abnormal changes on ECG signal is always developing and making a real difference.

The detection and diagnosis of arrhythmia, one of the main, troubling heart disease has been researched by diverse methodologies and algorithms. For example, artificial neural network is one of the conventional classifier used for ECG arrhythmias classification [30][31][32]. Discrete Wavelet Transform (DWT) and Continues Wavelet Transform (CWT) are often added to improve the performance of neural network algorithms [33][34][35][36]. Support Vector Machine (SVM) is also carried out by numerous research due to that SVM classifiers do not trap in local minima and do not require massive amount of training data [37][38][39][40]. Besides, Dynamic Time Warping (DTW) [41], Independent Component Analysis (ICA) [42], Principal Component Analysis (PCA) [43], anti-dictionary coding [44] are applied to different arrhythmias recognition, too.

On the other hand, ischemia heartbeat detection and recognition has drawn a lot of attentions in the time series data mining community as well. Support vector machine/regression is one of the most popular algorithms used in this area [45][46][47]. Neural networks are also very popular algorithms thanks to its powerful problem-solving ability [48][49][50]. Hidden Markov Models are widely and well applied for ECG analysis [51]. Besides, Wavelet theory has drawn a lot of attentions and has been implemented by many researchers [52] [53]. In order to take advantage of distinct models, hybrid methods are constantly used [54].

3.1.3 Motivation of the proposal

An ECG signal records a pattern reflecting the electrical activity of the heart and typically requires a trained clinician and doctor to interpret the context of the signs and symptoms presented by the patient. It can provide information regarding the rhythm of the heart. This includes whether the electrical impulse consistently arises from an area of the heart where expected and at the appropriate rate, whether the impulse is

conducted normally throughout the heart, and whether any area of the heart is contributing more or less than expected to the electrical activity. Current advanced ECG recording equipment often includes analysis software that attempts to interpret the pattern. However, the generated diagnostic results may not always be sufficiently accurate to be used as the only evidence for a heart disease diagnosis.

According to summary information from the Mayo Clinic [55], which is considered as one of the best medical research institutions in the world, the current diagnosis of myocardial ischemia is described in the following and illustrated in Fig. 3.7. To begin, the medical history of a potential patient is reviewed by a clinician and possibly the nurse will conduct the ECG recording. Then, the clinician inspects the recorded ECG data carefully to form an opinion as to whether there is any abnormality on the ECG signal caused by myocardial ischemia. If the abnormalities on the ECG signals are not sufficiently prominent to support an accurate diagnosis by the doctor and clinician, further tests such as a CT scan or coronary angiography are performed to collect additional detailed information for diagnosis. The practical clinical experience accumulated by clinician is critical for the diagnosis. The proposal aims to provide interpretable rule mining results, reveal strong connections between different segmentations or intervals on a single heartbeat, and finally assist professional clinicians in the forming of their myocardial ischemia diagnosis.

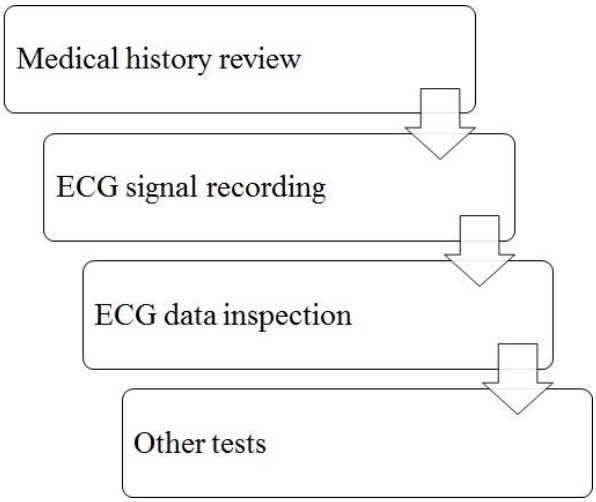

Fig. 3.7. Current heart disease diagnosis process

How to provide a real intelligent way to assist doctors making ischemia diagnosis faster and more accurate is the key issue for related research works. Diverse algorithms, techniques have been applied to evaluate the ST segment, T wave changes, and to detect ischemia on ECG. Almost all the proposals aim at directly providing automatic detection results to doctors. Although methodologies, algorithms about ischemia detection on ECG have developed for a few years, the diagnosis in reality still cannot only depend on these proposed methods due to the extreme complexity of ECG signals from different patients. The practical experience of doctors are always necessary for the ischemia diagnosis.

A fuzzy association rule mining based method is proposed for ischemia diagnosis on ECG signal. Its implementation composes of fours steps: significant feature extraction from every single heartbeat, fuzzy transformation of above features, association rule mining on fuzzy itemsets, and automatic ischemia, normal beats classification via extracted rules. The goal of the proposal is to provide doctors a set of fuzzy association rules that reveal a strong connections between significant heartbeat features and ischemia, as an assistant tool to help the diagnosis itself. Features to be extracted include ST segment deviation, ST segment duration, ST segment area, T wave peak, T

wave area, and T wave direction. Segmentation of above features is done by fuzzy c-mean clustering, and membership function parameters of each feature are determined based on above fuzzy c-means clustering results. Then, association rules are mined on fuzzy itemsets. At last, a validation algorithm using these extracted rules is proposed for automatic heartbeat classification on ECG signal.

The proposal inherits the merit of association rule based method. First, the results of the proposal are interpretable information which makes doctors understand the underlying correlation before they make diagnosis. The proposal aims at being a useful, crucial medium for the ischemia diagnosis process. Second, it reveals strong relationships between different feature types, which gives itself the possibility to achieve accurate results for ischemia detection. To obtain meaningful rules, extracted features need to be segmented to different intervals before conducting association rule mining. In the proposal, fuzzy c-means clustering is applied to determine how to discretize features. In this way, it makes the feature segmentation more functional and effective.

Experiments are conducted on a PC with a dual core processor (2.5 GHz) and 8 GB memory. Simulation software is from Matlab. Data used in the experiment are the ECG recordings from European ST-T Database [56] [57]. This database, collected by PhysioNet, is intended to be applied for evaluation of algorithms for ST segment and T wave changes which are the most common pathological changes of myocardial ischemia. All experiment data are divided to training dataset and test dataset. A set of significant features are first extracted from each single heartbeat in training dataset. Afterward, each extracted feature is segmented to several intervals by fuzzy c-means clustering algorithm. All training data are used to train fuzzy membership functions. Then, all the data are transformed to fuzzy values. At last, association rule mining algorithm is performed to extract fuzzy association rules. In the validation process, these mined rules are applied to test dataset for confirming the effectiveness of automatic ischemia classification.

3.2 Diagnosis Process based on Rule Mining

Every step of the proposed fuzzy association rule mining based method is elaborated in detail. The rule mining and classification process is shown in Fig. 3.8 as follows.

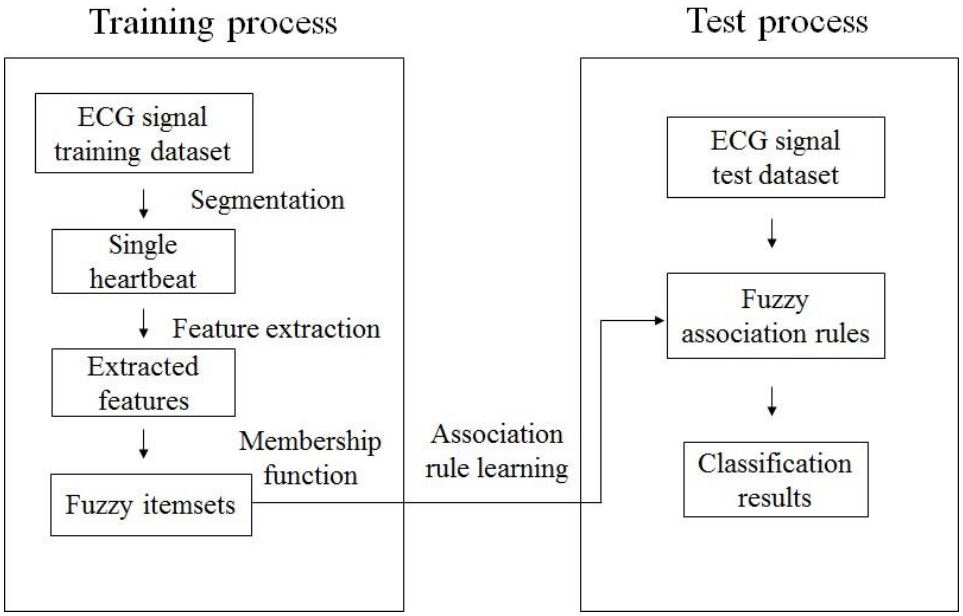

Fig. 3.8. Fuzzy association rule mining and classification on ECG signal

In the training process, each individual heartbeat is extracted from a long duration ECG signal according to the attribute documents provided by PhysioNet. Then, normal and ischemia heartbeats are separated based on their annotation files, where heartbeats are labeled individually via two cardiologists working independently. Following this, the J point detection is conducted and a set of features are extracted based on the J point from each heartbeat. Fuzzy c-means clustering is applied on every feature type, to determine the parameters for the feature's corresponding membership functions. After transforming features to fuzzy itemsets, the association rule mining is performed. A set of fuzzy association rules is obtained from the last mining process. These rules have

their individual realistic meaning and are intended to be helpful in the ischemia diagnosis process. Finally, in the test process, these extracted rules are applied to the ECG test dataset to confirm the effectiveness and accuracy for automatic classification of ischemia and normal heartbeats.

3.3 Feature Extraction of ECG signal

In order to get features from each heartbeat, noise elimination is not an ignorable task. Baseline wandering is removed by Infinite Impulse Response(IIR) high pass zero phase filtering in the proposal, since it is proved to be an effective algorithm for baseline wandering removal [58] [59].

Baseline wandering are primarily caused by respiration, electrode impedance change because of perspiration and increased body movement [60]. Baseline wander elimination, as a classical problem, is considered as an artifact which produces artifact data during ECG recording. There are mainly two categories of methods that applied to remove baseline wander: non-adaptive and adaptive filtering. Non- adaptive filtering approaches include Infinite Impulse Response (IIR) filer, Finite Impulse Response (FIR) filter, and notch filter [59]. The most essential part of high pass filter for baseline wander removal is the choice of cut-off frequency and phase response characteristic [61]. High pass FIR filter can be implemented by utilizing Kaiser Window [62]. The idea of window-based algorithms is to truncate a reasonable response with a finite length window. The dilemma for FIR filter is that as the filter order increases, the complexity increases, as well, while low filter order cannot guarantee the performance of the filter. On the other hand, although IIR filter can achieve a sharp transition region with fewer coefficients, high cut-off frequency-based IIR filter has a nonlinear phase response that distorts meaningful components of ECG signals. Therefore, to preclude

this distortion, bidirectional filters are created, in which the ECG recording is filtered at two directions: forward and backward. This approach is called zero-phase filtering [63]. Besides, wavelet approaches, in which ECG signal is decomposed to basic functions known as wavelets, are often implemented for the purpose of baseline wander removal. Also, moving average approach, which smooth data by replacing each data point with the average of the neighboring data points within a span, is considered as a well-applied algorithm for eliminating the baseline wander on ECG signal.

Due to that myocardial ischemia always causes morbid change of ST segment and T wave, several features of these two parts are extracted. After J point detection algorithm is applied, the following features are collected based on this key point for single heartbeat, also shown in Table 3.1.

(1)ST segment deviation: 80 milliseconds if heart rate does not exceeds 120 bpm or 60 milliseconds otherwise after J point.

(2)ST segment duration: the time duration of ST segment deviation.

(3)ST segment deviation area: sum of ST segment amplitude.

(4)T wave peak: amplitude of T wave peak.

(5)T wave area: sum of T wave amplitude.

(6)T wave direction: the abnormality of T wave sometimes includes reverse direction.

As illustrated in the following figures, the feature extraction in the proposal mainly focus on two most significant parts: ST segment and T wave. The deviation at certain point and the area of these two parts are extremely crucial for ischemia and normal detection. Besides, the direction of T wave is also obtained since it is one of the most obvious changes of ischemia beat [64].

Table 3.1. Feature extraction from single heartbeat

Features	Description
ST segment deviation	80 milliseconds if heart rate does not exceeds 120 bpm or 60 milliseconds otherwise after J point
ST segment duration	The time duration of ST segment deviation
ST segment deviation area	The sum of ST segment amplitude
T wave peak	The amplitude of T wave peak
T wave area	The sum of T wave amplitude
T wave direction	The abnormality of T wave sometimes includes reverse direction

These extracted features of each heartbeat are illustrated in detail from Fig. 3.9 to Fig. 3.13.

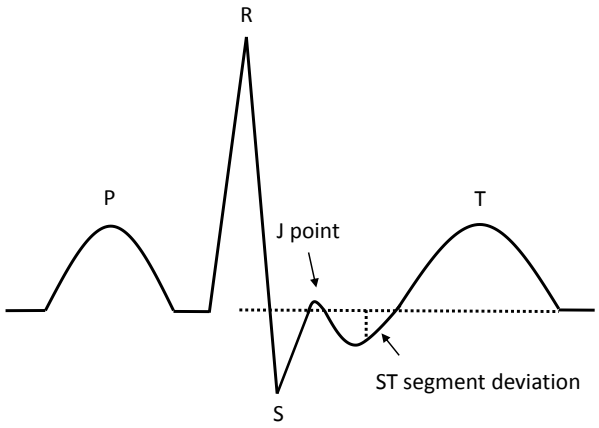

Fig. 3.9. Heartbeat feature: ST segment deviation

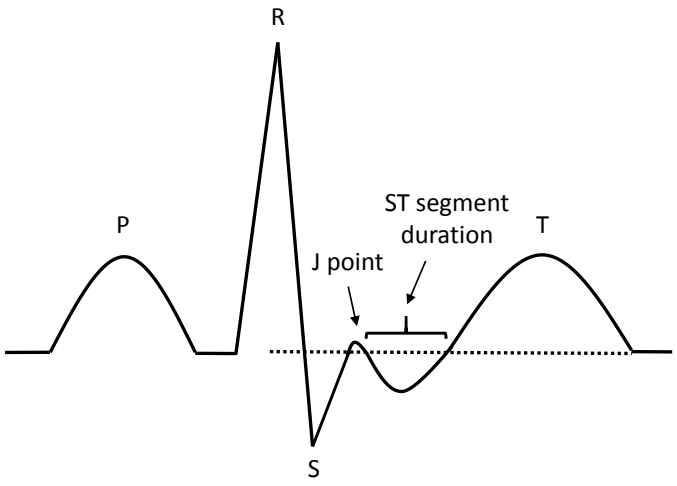

Fig. 3.10. Heartbeat feature: ST segment duration

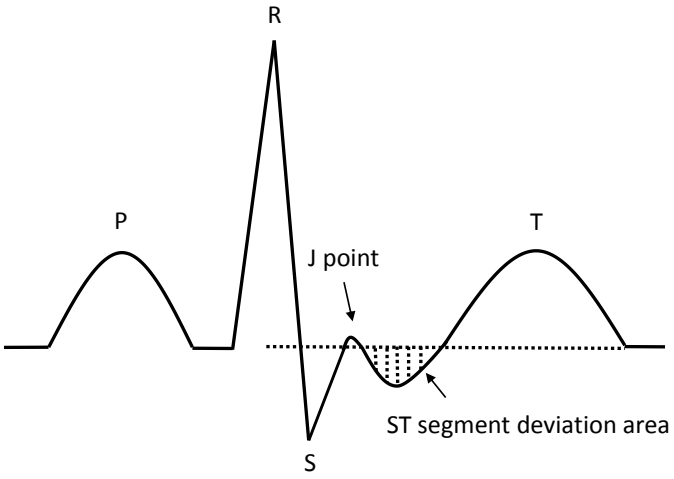

Fig. 3.11. Heartbeat feature: ST segment area

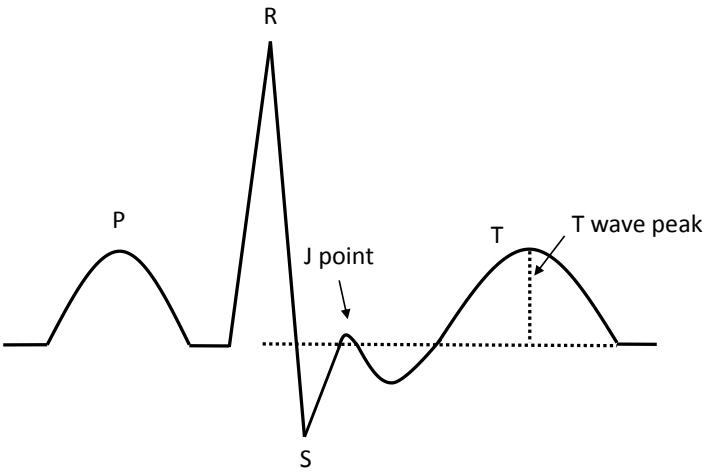

Fig. 3.12. Heartbeat feature: T wave peak

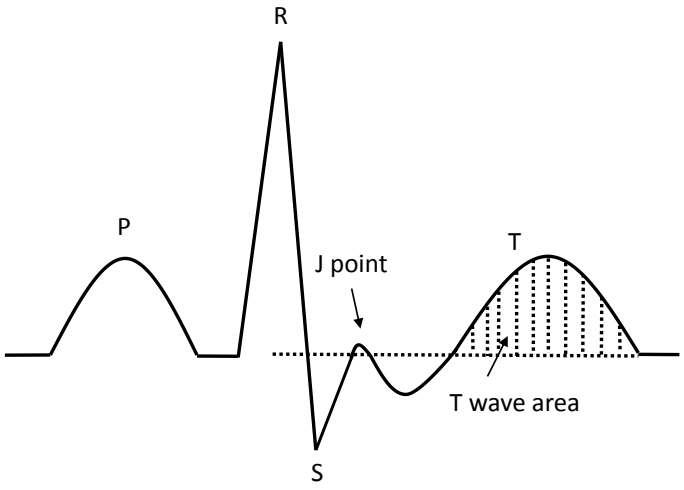

Fig. 3.13. Heartbeat feature: T wave area

The above features are crucial elements employed to distinguish between ischemia and normal heartbeats. The association rules in the proposal are expected to reveal strong relationships between these features allowing them to be meaningful and worthwhile for myocardial ischemia diagnosis.

3.4 Fuzzy Transformation of Heartbeat Features

Before conducting association rule mining task, continuous-valued features which are obtained in the last step should be segmented to intervals. In previous research, distinct discretization ways, such as equal depth binning algorithm and CT-Disc algorithm have been applied [65]. Instead of crisp intervals, fuzzy c-means clustering algorithm is performed to discretize those continuous-valued features to fuzzy itemsets in the proposal, as shown in Fig. 3.14.

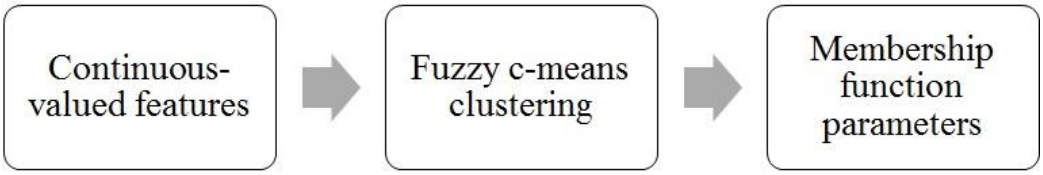

Fig. 3.14. Membership function construction

The advantage of conducting association rule mining on fuzzy itemsets instead of crisp set has already been comprehensively discussed and proved in [66]. It is possible to miss important values because of excluding values near the sharp boundary. The effect of sharp boundary problem is that, for association rule mining, the values within one interval may not satisfy the support threshold, yet, if the values near both boundaries are considered, the partition of certain discrete interval may become meaningful. Therefore, it indicates that the classical set theory has its own deficiency as least under the circumstance of association rule mining. However, in fuzzy set theory, an element is able to belong to a set with membership value in [0, 1]. For attribute x

and its domain D_x, the mapping of the membership function is $f_x(x): D_x \rightarrow [0,1]$ [66]. Fuzzy itemsets provide a smooth change between different intervals to solve above sharp boundary problem.

In the proposed method, fuzzy c-means clustering is applied to determine the membership function parameters for each extracted heartbeat feature. The fuzzy c-means clustering is widely applied in many application areas [67].

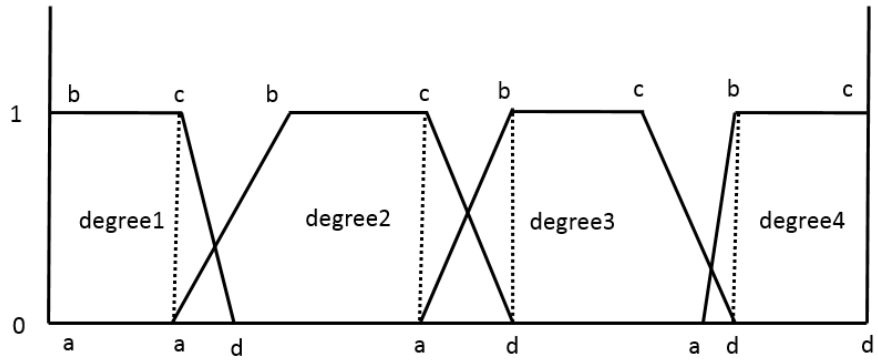

Fig. 3.15. Parameter setting of membership functions

An example in Fig. 3.15 is used to explain how the parameters of membership functions are determined by fuzzy c-means clustering algorithm. In this example, each extracted feature from 3.3 is clustered to 4 categories, which are called: degree1, degree2, degree3 and degree4. Every membership function has four parameters: a, b, c and d. Here, the setting of these parameters for the first, the last and the otherwise membership functions are explained, individually. The c and d of membership function degree1 are the minimum and maximum of cluster degree1 from results of fuzzy c-means clustering. For the last membership function, degree4 in above example, a and b of membership function degree4 are the minimum and maximum of cluster degree4, respectively. The degree of membership function 0 and 1 are assigned to a and b of first membership function, d and c of last membership function, respectively. For any other

membership functions, b and c are the minimum and maximum of its corresponding cluster, while a is the maximum of previous cluster and d is the minimum of the following cluster. Above explanation describes how the membership function parameters are determined.

The membership function value for each extracted feature is calculated using

$$\mu(x,a,b,c,d)=\begin{cases}0, & x\le a, x>d\\ \dfrac{x-a}{b-a}, & a<x\le b\\ 1, & b<x\le c\\ \dfrac{d-x}{d-c}, & c<x\le d\end{cases} \tag{3.1}$$

3.5 Association Rule Mining on Fuzzy Itemsets

3.5.1 Association Rule

The purpose of association rule learning to discover meaningful and crucial relations between variables in large database from the retail industry and business [68]. It is first proposed by R. Agrawal in [69] to search for regularities between products in large-scale transaction data in supermarkets. The goal is to find out whether the customer who buys certain goods is likely to also buy another good. Such market retail analysis through association rule mining is the foundation for further business decisions, such as merchandise placement, pricing adjustment, etc. The following example in Table 3.2 shows the usage of association rule learning in transaction data.

Table 3.2. Example of supermarket transaction data

ID	Meat	Onion	Milk	Beer
1	1	1	0	1
2	1	1	1	1
3	0	0	1	0
4	1	0	0	1

It is easy to discover an association rule $\{Meat, Onion\} \Rightarrow \{Beer\}$, which means when customer buys meant and onion together, there is a very high possibility that the same person buys beer, as well. According to such analysis result, corresponding business decisions can be done. Starting from business analysis area, association rule mining is expanded to a lot of other applications, such as, medical, financial fields.

The definition of association rule learning is as follows: Let $I = \{i_1, i_2, ...i_m\}$ be a set of m attributes called items. Let $T = \{t_1, t_2, ...t_n\}$ be a set of transactions called database. Each transaction in T contains a subset of items in I. A rule is defined as the form of $A \Rightarrow B$, where $A, B \subseteq I$ and $A \cap B = \phi$.

To discover rules from a set of all possible rules, some threshold values need to be determined. The best well-known constraints are support and confidence. The support $Supp(A)$ of an itemset A is the proportion of transactions in the database that contain the itemset A. The confidence of a rule is defined as $Conf(A \Rightarrow B) = Supp(A \cup B) / Supp(A)$. The confidence of rule $A \Rightarrow B$ indicates the percentage of the transactions that contains both A and B at the same time.

3.5.2 Association Rule Mining Algorithms

To be brief, there are two steps in the search of association rules. First, finding frequent itemsets in a database via minimum support threshold. Second, calculating confidence value for each possible rule that formed from acquired frequent itemsets, saving those that satisfy the confidence threshold value.

There are plenty of algorithms that can realize the rule mining task. The most known algorithm in association rule mining research community is the Apriori algorithm. It starts iteratively from frequent itemsets containing a single item, and uses breadth-first search and a Hash tree structure to count candidate itemsets efficiently, then generates candidate itemsets of length k from itemsets of length k-1, afterwards prunes the candidates that have an infrequent sub pattern. This algorithm is based on the A Priori Property that every subset of a frequent itemset must be a frequent itemset, as well. Eclat (Equivalence Class Transformation) algorithm is a depth-first search algorithm which applies set intersection method. Frequent Pattern growth algorithm, known as FP-growth algorithm, is another well applied methodology to mine association rules. This algorithm represents the transaction database as a prefix tree which is enhanced with links that organize the nodes into lists referring to the same item. The prefix tree is projected to carry out the search, working recursively on the result, and pruning the original tree. The implementation supports filtering for closed and maximal itemsets with conditional item set repositories as well, although the approach used in the program differs in as far as top-down prefix trees rather than FP-trees. It does not cover the clever implementation of FP-trees with two integer arrays. There are also some other efficient algorithms, such as AprioriDP, Node-set-based algorithms, OPUS search and Context based Association Rule Mining algorithm.

3.5.3 Mining on Fuzzy Itemsets

Yet, due to fuzzy itemsets in the proposal, above traditional rule mining algorithms are not directly suitable here. The association rule mining is conducted by combining the implementation of Apriori algorithm and the proposal in [70], which handles continuous-valued data to discover association rules between fuzzy itemsets. The applied algorithm in the proposal is discussed as follows and shown in Table 3.3.

Input: a set of continuous-valued data $T = \{t_1, t_2, ...t_n\}$, a set of membership functions f, a minimum support threshold value α, a minimum confidence threshold value β.

Output: a set of fuzzy association rules with confidence values.

Table 3.3. Fuzzy association rule mining algorithm

Step	Actions
(1)	Transform continuous-valued data to fuzzy itemsets via membership functions.
(2)	Calculate scalar cardinality of each fuzzy itemset, prune those support-unsatisfied candidate 1-itemsets to form large 1-itemsets. Set r=1.
(3)	Generate candidate (r+1)-itemsets from large r-itemsets, calculate fuzzy value for every itemsets pair by using the minimum operator, calculate mean of above fuzzy value in all tuples. Put them in large (r+1)-itemsets if their supports are larger than the predefined minimum support.
(4)	Go to next step, if large (r+1)-itemsets is null; otherwise, set r= r+1, repeat above (3).
(5)	Calculate confidence value for every possible association rules, compare to predefined minimum confidence value, then prune those unsatisfied ones.

The results of above algorithm are a set of fuzzy association rules with corresponding confidence value. The degrees that appear in the fuzzy association rules are called dominant feature degree in the proposal. The rules for ischemia and normal beats detection expect different dominant feature degrees which means that the proposed method can distinguish the significant discriminations between ischemia and normal.

3.6 Classification of Ischemia Heartbeats

Upon completion of the fuzzy association rule mining, a validation method is proposed for the automatic classification of ischemia and normal heartbeats on the test dataset. The scheme of the validation process is presented in Fig. 3.16.

The validation process of fuzzy association rule is quite different with rules on crisp sets or discrete segments. For example, the methods, CPAR [71] and CBA-CB [72], cannot be directly applied to the experiment. Therefore, a validation method is proposed based on these traditional algorithms. The explanation of the proposed method is as follows.

First, the fuzzy association rules that obtained from 3.5 are sorted by their confidence values. Next, k rules with descending confidence are selected by the criteria that the sum of their support is not less than 100%. Then, these selected rules are applied to each test heartbeat data which are previously transformed to fuzzy values via membership functions. Each test heartbeat data is calculated twice, by the set of rules for ischemia and the set of rules for normal, separately. The calculation is done by iterating every rule, doing the sum of their fuzzy values. Afterwards, the calculation results are compared between ischemia and normal. The test heartbeat is classified to the class which wins above comparison. Through above process, each single test data is classified to either ischemia or normal.

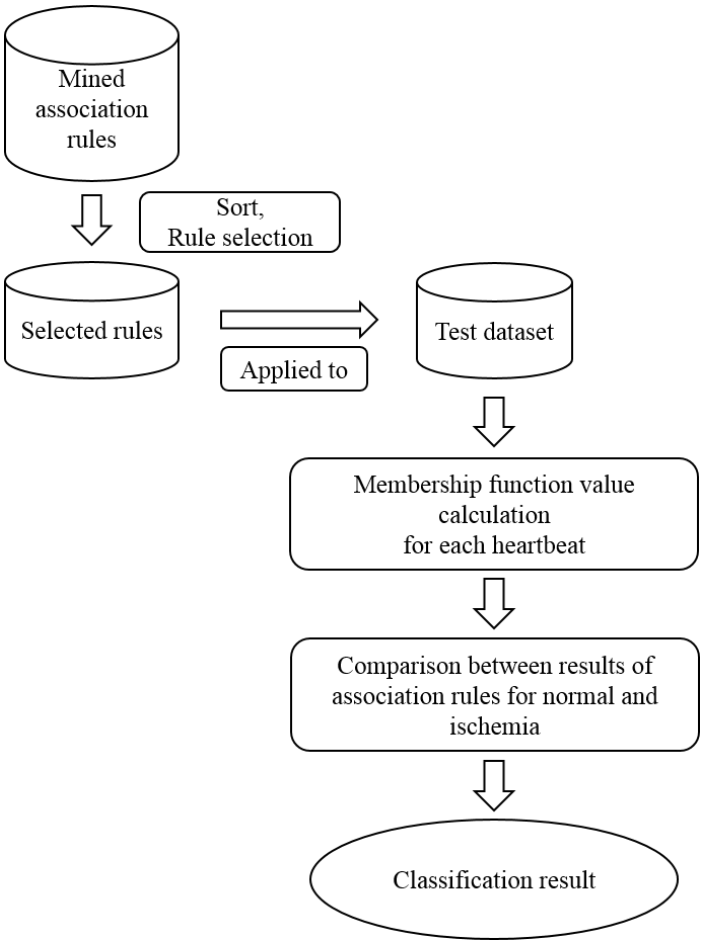

Fig. 3.16. Classification of ischemia and normal beat process

3.7 Experiments on European ST-T Database

3.7.1 European ST-T Database

This open data source is collected for the evaluation of ischemia detection algorithms. Each record in this database is two hours in duration and contains two signals, both sampled at 250 Hz with 12-bit resolution over a nominal 20 millivolt input range. As an

open data source, it has made a great contribution to the heart disease diagnosis research community. For each recording, there are two important data files. 'Signals' file contains recorded physiologic signals; 'Annotations' file describes the annotation information, such as ischemic, normal or artefact, for every single beat. The experiment data in the proposal is from the Modified Lead II(MLII) of recordings e0103, e0104, e0147, e0159, e0162 and e0206. Experiment dataset contains 39705 beats after artefacts and some other irrelevant abnormal beats are cleared away according to their corresponding 'Annotations' data files. 16318 beats are used as training dataset. Among it, 8567 are ischemic beats, while the rest are normal beats. On the other hand, there are 23387 beats in the test dataset, 6312 are ischemic beats and 17075 are normal beats, shown in Table 3.4.

Table 3.4. Experiment data from European ST-T Database

Training data		Test data	
16318 beats		23387 beats	
Ischemia	Normal	Ischemia	Normal
8567	7751	6312	17075

Extracted features are decided based on the explanations about ST segment and T wave episodes from the PhysioNet website. Since ST segment deviation is the most common standard change due to myocardial ischemia, it becomes the first necessary feature without any doubt. This feature is measured by taking 80 milliseconds after J point if heartbeat does not more than 120 bpm, or 60 milliseconds after J point otherwise. ST segment duration and ST segment area are also selected since they contains more detailed information about ST segment in each heartbeat. For T wave, its amplitude peak, area and direction are measured for the further evaluation because of the similar reason that they can provide concrete information about T wave change.

3.7.2 Fuzzy Association Rule Extraction Experiment

Before doing rule mining, the fuzzy c-means clustering algorithm is performed on extracted features. For each feature, they are clustered to eight categories in our experiment. Therefore, there are eight membership functions for every feature: from degree1 to degree8. According to the description in 3.4, four parameters are determined respectively for eight membership functions. The parameters of ST segment deviation membership functions are shown in Table 3.5 and Table 3.6 as examples here. The range of ST segment deviation is from -0.3151 to 1.6649. The parameter settings for other features follow the same principle.

Table 3.5. Parameters of ST segment deviation membership functions (1)

	Degree1	Degree2	Degree3	Degree4
a	-0.3151	-0.3102	0.1076	0.2231
b	-0.3151	0.0323	0.1077	0.2236
c	-0.3102	0.1076	0.2231	0.3505
d	0.0322	0.1077	0.2236	0.3507

Table 3.6. Parameters of ST segment deviation membership functions (2)

	Degree5	Degree6	Degree7	Degree8
a	0.3505	0.4884	0.6471	0.8010
b	0.3507	0.4892	0.6474	0.8014
c	0.4884	0.6471	0.8010	1.2667
d	0.4892	0.6474	0.8014	1.6649

After transforming all experiment data by the membership functions, there are also some parameters need to be decided for the rule mining process. As shown in Table 3.7, the settings of predefined support and confidence value are 5% and 0.7, respectively. Then, rule mining is performed using the algorithm elaborated in 3.5.

Table 3.7. General parameter setting of the rule mining experiment

Threshold settings	
Minimum support	5%
Minimum confidence	70%
Fuzzy itemset number	eight

About the results of association rule mining process, 43 rules are obtained. Among these rules, 19 rules are for normal, while 24 rules are for ischemia. Some extracted rules are listed in the following Table 3.8 and Table 3.9.

Table 3.8. Association rules for normal heartbeats

Rule	Confidence
ST segment deviation area is degree2 and T wave area is degree7 and T wave peak is degree6	92.97%
ST segment deviation area is degree2 and T wave peak is degree6	91.33%
ST segment duration is degree7	87.91%

and ST segment deviation area is degree2	
T wave area is degree7 and ST segment duration is degree7 and ST segment deviation area is degree2	86.92%

Table 3.9. Association rules for myocardial ischemia heartbeats

Rule	Confidence
ST segment deviation area is degree6 and ST segment deviation is degree7	96.61%
ST segment deviation area is degree6 and T wave peak is degree7	95.40%
ST segment deviation is degree7 and ST segment duration is degree3 and T wave peak is degree7	95.22%
ST segment deviation is degree7 and T wave area is degree4 and T wave peak is degree7	94.82%
ST segment duration is degree3 and T wave area is degree4 and T wave peak is degree7	92.51%

As highlighted previously, these obtained fuzzy association rules based on membership functions of each feature are meaningful, understandable to human being. Moreover, they succeed to reveal strong relationships between different significant heartbeat features. These association rules are expected to be used as assistant reference to accelerate the diagnosis of ischemia.

The association rule results in Table 3.8 and Table 3.9 shows that dominant degrees of each feature, which is emphasized in 3.5, are actually distinct from each other, which means that the significant discriminations between normal and ischemia beat are captured by rule mining process.

3.7.3 Classification Evaluation of Extracted Rules

Although association rules are applied as reference information to help the diagnosis process with speed and accuracy, it is still necessary to propose a validation method to detect ischemia on test dataset using these extracted rules. In this part of experiment, the validation is conducted as the method discussed in 3.6. In the rule selection step, first 14 association rules with best confidence values are picked for normal heartbeat classification, while first 12 rules are used for ischemia heartbeat detection. The classification result is shown in Table 3.10.

There are totally 23387 heartbeats in the test dataset, in which there exist 6312 ischemia beats and 17075 normal beats. In the classification experiment, 13777 normal beats are correctly detected as normal. On the other hand, 5266 heart beats are classified to ischemia without mistakes. The sensitivity and specificity of ischemia and normal classification are 83.4% and 80.7%, individually.

Table 3.10. Classification results of association rules

	Classified as normal	Classified as ischemia	Total
Normal	13777	3298	17075
Ischemia	1046	5266	6312

3.7.4 Discussion about Interpretability of Mined Fuzzy Association Rules

The rule mining experiment shows that the fuzzy association rules from the experiment are clear and meaningful to assist doctors with ischemia diagnosis. It should be also mentioned that, according to the extracted rules, it is more helpful that doctors are informed with background knowledge about the feature segmentation in the proposal. In addition, in order to provide mined association rules more instinctive and straightforward, a visualization user interface of these rules becomes necessary in the future research. And, such visualization user interface can be added in the current ECG signal recording machines, and serve as one of the important monitoring elements.

From the association rule extraction results, it is noticed that the dominant feature degrees in the rules for ischemia and normal are very different. This is the evident that the extracted rules are able to classify ischemia based on these distinct, meaningful rule elements. This fact makes the classification with extracted rules reliable and convincing. As expected, validation results demonstrate the competitiveness and efficiency of the proposal.

Except ischemia, there are also some other severe, threatening heart diseases that

make people worry all the time, the diagnosis of them are still relying on doctors' practice experience just like current situation of ischemia. Since the interpretable and understandable results are able to be very helpful for the diagnosis process, the proposal aims to be expanded to other heart disease diagnosis areas.

3.8 Classification via Fuzzy Association Rules

Associative classification, that is proposed to integrate classification and association rule mining, commonly first generates a set of association rules, then select a compact set of rules to perform classification on unknown features [71][72]. Afterwards, the attention goes to the methods that adopt fuzzy set concept instead of other partition ways that discretize quantitative attributes, based on the reason that fuzzy sets provide a smooth transition between member and non-member of a set [66]. Different algorithms and methods have been proposed to conduct classification task using fuzzy association rules [73][74][75][76][77]. These proposals have shown their superiority by comparing to other well-known algorithms, such as C4.5 and CPAR [71].

A method about how to use fuzzy association rules to do classification is proposed. In the proposal, the fuzzy itemsets transformation of continuous-valued quantitative attributes via fuzzy c-means clustering is illustrated. The association rule generation algorithm on fuzzy itemsets is, then, elaborated. Finally, classification using these generated fuzzy association rules is described in details. The goal of the proposal is to provide a general framework about classification using fuzzy association rules with a convincing classification accuracy.

Compared to other state-of-the-art methods, the proposal provides a more intuitive perspective to classification using fuzzy association rules, and aims to produce more accurate classification prediction for unknown attributes. Also, the proposal is not

biased toward any different size of rule's antecedent part by adopting a more plausible measurement to classification.

Experiment datasets are from UCI Machine Learning Repository [78]. This open data source is collected as a service to the machine learning community. Experiments are performed on a PC with a dual core processor (2.5 GHz) and 8 GB memory. Simulation software is from Matlab. For each experiment dataset, the features of training data are transformed to fuzzy itemsets at first. Afterwards, rule mining algorithm is conducted to generate a set of fuzzy association rules. Finally, the proposed classification method using above extracted rules is evaluated on the test data by comparing with a well-applied method, which utilize confidence degree to classify unknown data.

3.8.1 Fuzzy Itemsets Generation

Each data item in the training and test dataset has several attributes, which are normally continuous-valued. In order to take advantage of association rule merits, these continuous-valued attributes need to be properly segmented. In the proposal, every attribute is transformed to fuzzy itemsets instead of crisp intervals. The advantages of this way is illustrated in 3.4. Fuzzy c-means clustering is applied to determine the membership function parameters for every dataset attribute.

For example, Iris dataset, from UCI Machine Learning Repository, has three species classes. The first attribute of each tuple in the dataset is sepal length that is transformed to three fuzzy itemsets, short, medium, and long, as shown in Fig. 3.17. In such case, every quantitative value finds its corresponding fuzzy itemset. Yet, this is just the example for Iris dataset. For other datasets, the fuzzy segmentation step is also done accordingly, but with their own suitable explanations.

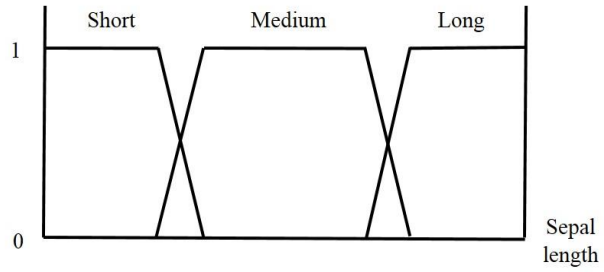

Fig. 3.17. Membership functions example of three fuzzy itemsets

3.8.2 Fuzzy Association Rule Mining

After transforming all attributes to fuzzy itemsets, the following goal is to extract the associated relationship between a set of fuzzy itemsets and classes. The mining results are supposed to be understandable and meaningful. For instance, for the Iris dataset, the association rule, which is fictional here, "if sepal length is long and petal length is also long, then it is Iris Setosa with 90% confidence" is found out. Such rule-based information aims to serve as helpful medium.

The applied algorithm in the proposal is comprehensively discussed as follows, and details of every step is illustrated in 3.5.3.

Input: a set of continuous-valued data $T = \{t_1, t_2, ...t_n\}$, a set of membership functions f, a minimum support threshold value α, a minimum confidence threshold value β.

Output: a set of fuzzy association rules with confidence values.

3.8.3 Classification using Fuzzy Association Rules

The classification algorithm that makes use of fuzzy association rules is elaborated as follows.

First, all the extracted fuzzy association rules are sorted descending by confidence values. Afterwards, for each instance with unknown class, the following measure, which is named as average proximity in the proposal, is done between the instance and every sorted rules. The rule that returns the minimum average proximity value is used to classify the unknown instance. If multiple minimums are found, the one with highest confidence value is chosen to do the classification.

Given a dataset D with n attributes, and for each attribute $i \in \{1,...,n\}$, let $j \in \{1,...,w\}$ be the predefined fuzzy itemset number. Given a rule r, of which the antecedent part A contains $rule\_n \subset i$ attributes, a subset of all attributes of the dataset.

For an instance d, its average proximity to rule r is defined as:

$$\sum_{rule\_i}^{rule\_n}(1-\mu_{rule\_i,j}(d))^2 / rule\_n .$$

The $\mu_{rule\_i,j}(d)$ is the membership function value of the attribute of instance d.

The logic behind the proposed average proximity is to find the closest status to certainty. On one hand, for crisp set, the value is either 0 or 1, which represents a definite status of non-existence and existence. On the other hand, for fuzzy itemset, the membership value is in [0, 1], varying from non-existence to existence. Therefore, the average proximity is able to match the unknown attributes to the nearest fuzzy association rule, and naturally it is reasonable to give the class of this rule to the unknown instance.

Also, the rule generation number is discreetly controlled by using support, confidence threshold. It is proper to adopt rules with relatively high support and confidence values.

Besides, the proposed classification algorithm is compared to a current well-applied method, which is called confidence degree, notated as “conf_deg”, here, in the experiments. The details of the comparison target is introduced as follows.

Given a fuzzy association rule , and a transaction d, the confidence of classifying d with r is:

$$\mu_F(d) * Dconf(r)$$ [76].

$Dconf(r)$ is the confidence of rule r, and the other variable is the match that d is compatible with fuzzy set F. The rule with highest confidence degree is selected to classify unknown instance.

3.8.4 Experiment Data Source: UCI Machine Learning Repository

Experiment datasets are from a commonly applied benchmarking database, UCI ML Repository. Also, the attribute value of each used dataset is quantitative, continuous-valued instead of discrete. The detailed information of each dataset is listed in Table. 3.11.

Table. 3.11. Experiment data from UCI Machine Learning Repository

Dataset	Number of instances	Training instances	Test instances	Number of attributes	Number of classes
Iris	150	105	45	4	3
Wine	178	122	56	13	3
Breast	683	480	203	9	2
Pima	768	500	268	8	2

3.8.5 Classification Experiments

In the experiments, each attribute is segmented to three fuzzy itemsets via fuzzy c-means clustering. For every experiment dataset, a set of fuzzy association rules are obtained using the algorithm described in 3.5.3. Afterwards, these rules are adopted to

classify unknown instances in the test dataset. Before the classification task, a rule selection step is performed, in which, unsatisfied association rules are abandoned. For example, the rules of which antecedent parts are from same attribute, etc.

Support threshold is set to control the appropriate number of generated rules, and proper classification results for both methods in the experiments. In addition, different confidence threshold values are set to compare the classification accuracy between confidence degree and the proposal.

For each dataset, two different support thresholds are set, while confidence threshold is set from 0.7 to 0.9 ascending by 0.05 under every support value. The experiment results are shown in Table. 3.12. From the experiments results comparison, it is shown that the proposal has better performance in terms of both maximum and average accuracy.

Table. 3.12. Accuracy results comparison of classification experiments

Dataset	Conf_deg		The proposal	
	Max	Average	Max	Average
Iris	95.6	90.2	95.6	90.2
Wine	96.4	87.0	96.4	95.5
Breast	95.1	91.5	99.5	99.3
Pima	69.0	68.2	76.5	70.4

3.9 Chapter Summary

In the fuzzy association rule mining experiment, fuzzy c-mean clustering algorithm is first performed on extracted heartbeat features, such as, ST segment deviation, T wave peak, individually. Each of these features is clustered to 8 categories, which are named from degree1 to degree8. Membership function parameters are decided based on

above fuzzy c-means clustering results. Then, all the experiment data are transformed to fuzzy values via these membership functions. Afterwards, rule mining algorithm is executed to obtain fuzzy association rules. This rule mining algorithm is based on the implementation of Apriori algorithm and the proposal in [70]. In the rule mining experiment, the minimum confidence threshold is set at 0.7. The mining results are that 43 rules are successfully extracted. Among these rules, 19 rules are for the normal beats, while 24 rules are for ischemia beats. Above association rules are first sorted by confidence value in the validation process. Then, several rules with descending confidence value are selected based on the criteria that the sum of their support value is not less than 1. Afterwards, for each test heartbeat, the sums of membership function values of both the set of ischemia and normal rules are calculated. Based on the comparison between ischemia and normal, the test heartbeat is classified to one of these two classes. 6312 ischemia beats and 17075 normal beats are used to validate the effectiveness of automatic ischemia detection using association rules. The sensitivity and specificity of ischemia and normal classification are 83.4% and 80.7%, respectively.

In the proposal, the focus is mining association rules on fuzzy itemsets. First, the characteristic of association rule based method brings the advantage that the results of the proposal are meaningful and useful. The high interpretability of these mined rules are actually beneficial to the acceleration of ischemia diagnosis. Compared to other discretization methods, the proposed way that using fuzzy c-means clustering before the rule mining task can produce more practical and feasible segmentation results. The classification results of automatic ischemia detection using association rules shows the effectiveness of the proposed method.

Despite that a lot of algorithms and methods for heart disease detection on ECG signal have been proposed by many researchers, the real diagnosis of heart disease, such as, myocardial ischemia, still cannot only rely on these proposals. The practical experience of doctors are always necessary for the heart disease diagnosis. The proposed fuzzy association rule mining based method provides helpful information, and

can be adopted as reliable medium to assist heart disease diagnosis. The advantages and practicality demonstrates the competitiveness of the proposal and will lead to its broader implementation in other cardiovascular disease diagnosis fields. The proposal aims to be a useful tool to accelerate the diagnosis of ischemia on ECG signal, and also to be applied to other heart disease diagnosis areas.

Chapter 4

Heart Disease Monitoring in Smart Home Care Environment

4.1 Smart Home Care Environment

The “Smart home”, “Intelligent home” concept is proposed to introduce the notion that devices, equipment in the house are connected by networks for certain purpose. These technical terms are adopted to highlight that home environment should be able to respond and modify itself all the time to accommodate its diverse residents and their constantly changing needs.

Since the advent of the smart home concept, a set of smart homes in distinct types have already been realized. According to the proposal in [79], with a focus on the functionality of smart homes, they can be classified to five hierarchical classes. First, homes containing intelligent objects: homes contain isolated, intelligent applications and objects; Second, homes containing intelligent, communicating objects: appliances and objects functioning in their own right and also exchanging information between each other to increase functionality; Third, connected homes: homes with external and internal networks which allow interactive and distant operation of systems; Fourth, learning homes: recording, analyzing the patterns of people’s activities, applying the learned knowledge to anticipate the needs of users and to control the technology; At last, attentive homes: the activity and location of people are constantly tracked, then as a feedback such information is applied to control technology for the satisfaction of users’

potential demands.

The technologies used in smart home have been developing constantly, and evolves from simple integration of electrical equipment in the home to a wider perspective, which includes diverse networks for information communications, sensors, microsystems, displays and software. These technologies become part of the smart home environment, in which user interfaces disappear gradually, interaction and communication between these intelligent devices, such as sensors, actuators and microcomputers, are based on speech, gestures, emotions and other vital information. A variety of technologies are required to realize smart home, such as, micro-electronics technology, embedded systems, micro systems technology, control technology, communication technology and software, web and network technologies.

According to different service purpose, the fields of smart home technology are divided to several categories. For example, multimedia and entertainment, security and safety, communication, and health care. Yet, the strict separation of service functionalities is unnecessary and nearly impossible, since the technologies and even user purpose sometimes overlap. As one of the main service areas of smart home, health care has been attracting a lot of attentions. It generally includes alarm systems and remote monitoring, and support service for nursing, doctor and family members. Personal health monitoring, that person can be monitored at home via kinds of medical devices, is continuously receiving an increasing amount of interest.

The increase of elderly people with kinds of health threats is becoming extremely dramatic all over the world. The global population of people over the age of 65 is due to rise to 20% of the whole population by the year 2025 [80]. How to provide timely, proper health care for such enormous amount of population is a huge problem for the family, government and society. The financial cost of caring also become a great burden and challenge. What is more troubling is that providing all-time nursing service or staying in hospital is impossible for the people who has potential risk of outburst of

certain diseases, such as stroke and heart attack. As a result, for those people who lives in his/her own home, and suddenly suffer from such emergent attack, there is a huge possibility that they cannot get treatment or rescue on time, which is a big life-threatening issue. The New England Journal of Medicine states that the chances of surviving a fall, heart attack or stroke are six times greater if people get help within one hour [81]. Therefore, in order to eliminate the situation that people get permanent harm or even die due to delayed or lack of help, the smart home care environment is needed, in which vital bio-information, such as ECG signal, are real-timely captured, analyzed, and the health condition are being closely monitored so that people can get personal health care in his/her own home.

Smart home care environment is defined, in our dissertation, as follows: the combination of technologies that record vital bio-signals, monitor the well-being of persons at home in real time via smart devices and appropriate actions, such as instant aid, care, and other reactions to sudden emergency.

In our dissertation, the research target is one of the dangerous heart diseases, myocardial ischemia. The ultimate objective is to monitor the potential ischemia attack in real time by analyzing the recorded ECG signal that collected by smart devices, and to trigger alarm, provide opportune, necessary care, so that the unnecessary risk can be avoided, reduced or even eliminated.

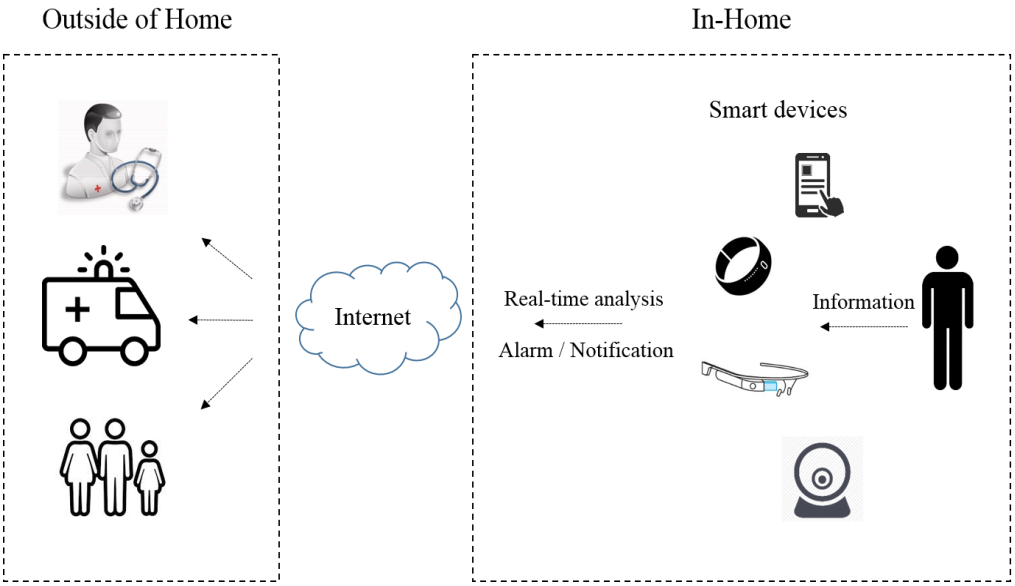

Fig. 4.1. Smart home care environment architecture

As shown in Fig. 4.1, within user's own home, the vital bio-information, such as, ECG signal, is recorded in real-time to track the physical condition of user's heart via smart devices. Afterwards, heart disease detection algorithm is applied to diagnose whether there is any abnormality of the heart. Based on the detection result, a decision is made to whether trigger alarm or any notification. If there is any discomfort, an alarm is triggered to call professional doctor or nursing staff outside of home, or even ambulance for emergency circumstances, also to notify family members. In such way, monitoring application in smart home care environment makes a solid contribution to people's daily life.

4.2 Smart Devices in Different Application Areas

A smart device generally means an electronic device which can connect to other devices or networks. Smart devices are designed to support a range of properties to ubiquitous computing and to be applied in physical world, human-centered

environments and distributed computing environments. Right now, smart device are definitely not just a distant phantasy to everyone, since they already take part in our life closely. For example, smartphones, such as Apple iPhone and others that running Android, and kinds of tablets are used by millions of people every day.

Wearable smart devices are clothing and accessories that are integrated with advanced computer or electronic technologies. Such devices, not just limited to smartphones and tablet computers, are constantly growing in a promising speed and playing a vital role in our daily life. Wearable technology relates to both ubiquitous computing and wearable computers. The development and implementation of wearable technology aim to interweave technology to the everyday life and to make technology pervasive. Such wearable technology tends to be more sophisticated because it can provide sensory and scanning features not only typically seen in mobile and laptop devices, such as tracking of bio-signals and feedback of physiological function. Wearable smart devices, generally speaking, have certain communications feature and can process information in real time. There are diverse categories of distinct wearable smart devices, such as glasses, contact lenses, watches, and bracelets, etc.

The implications and implementations of wearable smart devices and technologies are influentially promising and have a great potential in the fields of health and medicine, fitness, education, finance and gaming. The ultimate purpose of wearable technology in each application areas is to smoothly incorporate functional, portable electronics and computers into daily lives of each individual. For instance, in the field of gaming and entertainment, wearable technology and augmented reality can be combined to create a much more realistic, enjoyable and entertaining atmosphere in real time [82].

Along with the development in all aspects of technology, the prototypes of wearable smart devices are evolving from bulky technology to smaller, lightweight and more sophisticated systems. It can be expected that they are growing to much more advanced in the future without any doubt.

Fig. 4.2. Google Glass

Google Glass, as one of the most cutting-edge wearable devices, was released by Google in 2013, shown in Fig. 4.2. It aims to produce a mass-market ubiquitous computer, and displays information in phone-like hand-free format with an optical head-mounted display. The prototype of Google Glass resembles eyeglasses with lens replaced by a head-up display. Such smart device provides kinds of different features, for instance, its touchpad allows user to take control by swiping through a timeline interface displayed on screen. Also, it has camera to take photo or video. What makes Google Glass more powerful is the wide variety of applications on it. The applications, such as facial recognition, voice control, exercise tracking, and so on, lift up the participation level to the daily lives of users.

The attempt of developing health care applications on Google Glass has already begun. For instance, doctors in Wexner Medical Center used Google Glass to consult with a colleague in distant part of Ohio State, while students also observed the operation on their laptop computers in June, 2013 [83]. Besides, an application was developed to allow new-mothers to nurse their babies while viewing instructions about common breastfeeding issues, or calling a lactation consultant via the Glass in February 2014 in Australia [84].

In addition to glass, watch has been considered as another promising and challenging smart device because of its portability and convenience. A lot of companies are focusing on developing real smart watch.

Fig. 4.3. Apple Watch

Apple just announced Apple Watch, the most recent, eye-catching wearable smart device, as shown in Fig. 4.3, in September, 2014. Apple Inc. whose goal has always been make pioneer, powerful technologies more accessible, more relevant, and ultimately more personal look forward to that this Apple Watch represents a new chapter in the relationship between human and technology. In fact, Apple Watch should be considered as a multi-functional tool rather than just a watch. Apart from the basic features, such as, reading message, email, receiving alert, notification, and other functions via different applications, the more exciting feature is to be an intelligent, close companion in terms of health and fitness. Since as a watch it naturally touches the wrist, Apple Watch provides a more complete picture of users' physical activity. It measures not only the simple quantity of movement, like the distance users walk every day, but also the quality and frequency [85].

Fig. 4.4. Sensors on Apple Watch

More sensors and more edge-cutting technologies are integrated into wearable smart devices, more intelligent and powerful they become and evolve. As shown in Fig. 4.4, several distinct sensors are planted at the back of Apple Watch. The wrist is really a

convenient area to collect physical activity data. A designed sensor that uses infrared and visible-light LEDs and photodiodes to detect heart rate makes it literally a portable heart rate detector. Along with accelerometer, GPS and Wi-Fi, the movement and fitness can be recorded and monitored.

Except above Apple Watch, there are some other pioneer-developed smart watches at earlier period, as well. For instance, Pebble Smartwatch, in Fig. 4.5, and Moto 360, in Fig. 4.6. The Pebble Smartwatch is released in 2013, which features an e-paper display, an accelerometer, light sensors, a magnetometer, and a vibrating motor, enabling its application as an activity recorder. There are already over 1000 applications in the Pebble app store, providing varieties of application services, such as tracking of burned calories, remote control of smartphones, etc. What is more worthy of mentioning is that the release of open Pebble software development kit (SDK), platform for application development, enabling the broader, thriving possibility of the smart watch.

Fig. 4.5. Pebble Smartwatch

Fig. 4.6. Moto360

Moto 360, an Android smart watch, was announced in March, 2014, and released in September, 2014. It has a circular capacitive touch display, a removable wrist band, along with an ambient light sensor and a heart rate sensor.

The entertainment features and routine services of wearable smart device are, of course, very crucial to the daily life of users with no doubt. Yet, the new equipped sensors and more advanced technology endue these smart devices with larger possibility and more responsibility. These devices are able to be applied to construct a smart home care environment, and make it possible to provide timely, comprehensive health care to the people even in their own home.

4.3 Detection and Monitoring Process

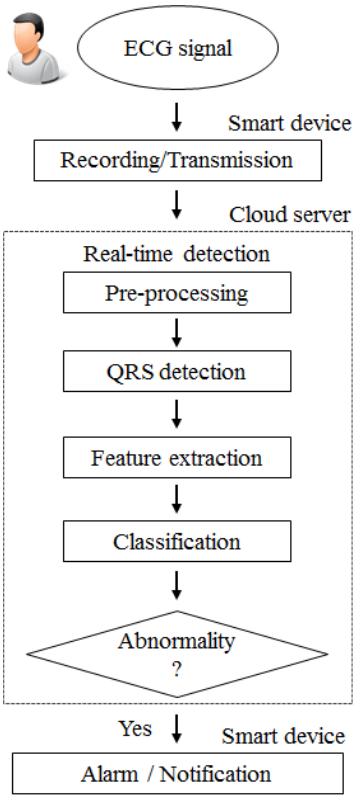

Fig. 4.7. Real-time detection and monitoring process

The myocardial ischemia detection and real-time monitoring for person within smart home care environment is illustrated in Fig. 4.7. Thanks to the rapid development of diverse intelligent functions on smart devices, the tracking and collection of ECG signal from monitoring target is not a troubling puzzle any more. First of all, user's ECG signal is recorded at a proper sampling rate via smart devices. Then, the pre-processing of collected ECG signal is conducted. In this step, noises, distortions, and baseline wanders need to be removed to make sure the following processing acquires classification result as accurate as possible. Afterwards, real-time ischemia detection is implemented on ECG signal. In order to classify each heartbeat to normal or ischemic beat, the reference point, such as J point and QRS complex of every heartbeat need to be found at the beginning. The following is to extract heartbeat features based on discovered reference point. These features should be based on medical evidence that they are different enough between normal and ischemic heartbeats. For instance, the ECG signal of patient who suffers from cardiovascular ischemia usually has abnormality at the T wave and ST segment parts. Finally, classification algorithm is performed on extracted heartbeat features to produce the detection result that whether an ischemic heartbeat is found or not. Generally speaking, the user is probably alone in the smart home care environment. Therefore, according to the real-time ischemia detection result, if ischemic heartbeat is repeatedly detected within a very short time period, like two minutes, alarm is triggered to contact nearby doctor or emergency personnel for timely treatment. On the other hand, notification is also sent to family members for their thoughtful care.

4.4 Real Time Ischemia Detection

4.4.1 Real Time QRS Detection

During ECG signal recording, it is almost inevitable that the recordings are contaminated by a few different types of artifacts, such as baseline wander, noises and artifacts. Baseline wander, as a severe problem, is mainly caused by perspiration and body movements. Baseline wander noise makes the accuracy of data mining result on ECG signal, especially the ST segment measurement, decrease dramatically. Therefore, kinds of algorithms have been proposed to deal with this tricky problem. For example, Butterwoth filter, Elliptic filter, Kalman filter are often applied [86]. Linear digital filter, adaptive filter, curve fitting, and multi-resolution analysis are also the common algorithms for baseline wander removal [87].

After eliminating noise and baseline wanders from recorded ECG signal, the reference point of each heartbeat need to be located in order to extract features for the last classification step. QRS complex, which reflects the rapid depolarization of the right and left ventricles, are the most striking part on ECG signal, as shown in Fig. 4.8. Since ventricle has larger muscle than atria, QRS complex normally appears with larger amplitude compared to P wave. Prior to any other tasks in real-time ischemia detection, QRS complex is detected at the first step in the feature extraction process, as shown in Fig. 4.9. After R peak and S point are found, the J point, which is the peak right after S point, is located. Based on J point, a set features are extracted from each heartbeat. Then, the classification algorithm uses these extracted features as input, and produce the output as normal or ischemic heartbeat.

There are a variety of real time QRS detection algorithms proposed by different researchers [88][89][90]. The real time QRS complex detection algorithm of the proposal is developed based on an online detector that uses state-machine logic to determine different peaks in an ECG via averaging and adaptive thresholds which are fluctuating in respect to the noise and the signal. It has the ability to confront noise by high pass filtering and baseline wander by low pass filtering [91]. Afterwards, J point is located by edge detection algorithm.

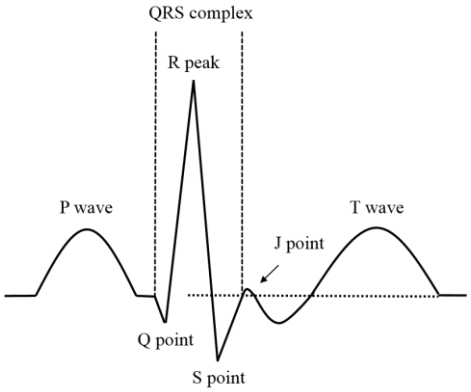

Fig. 4.8. QRS complex

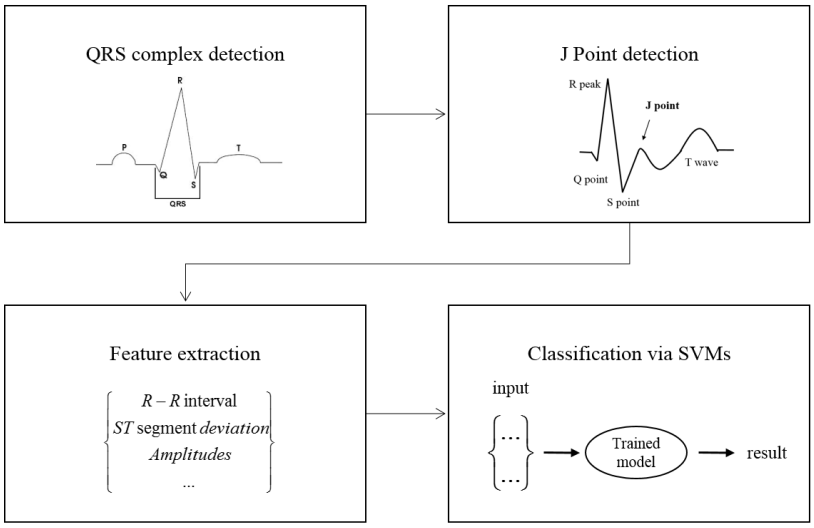

Fig. 4.9. Feature extraction process

4.4.2 Heartbeat Feature Extraction

The extracted features are decided based on the explanations about ST segment and T wave episodes from the PhysioNet website. ST segment deviation becomes the first target feature because of that it is the most common change caused by myocardial ischemia. This feature is measured by taking 80 milliseconds after J point if heartbeat does not more than 120 bpm, or 60 milliseconds after J point otherwise. ST segment duration and ST segment area are also selected as representative information of each heartbeat. For T wave part, its amplitude peak, area and direction are measured for the further evaluation, as well.

After J point detection algorithm is applied, the following features are collected based on this key point for single heartbeat.

(1) R-R interval : the time duration between two neighboring R peaks, as shown in Fig. 4.10.

(2) R peak : the amplitude of R wave peak.

(3) S peak : the amplitude of S wave peak.

(4) ST segment deviation : 80 milliseconds if heart rate does not exceeds 120 bpm or 60 milliseconds otherwise after J point.

(5) ST segment duration : the time duration of ST segment deviation.

(6) ST segment deviation area : the sum of ST segment amplitude.

(7) T wave peak : amplitude of T wave peak.

(8) T wave area : sum of T wave amplitude.

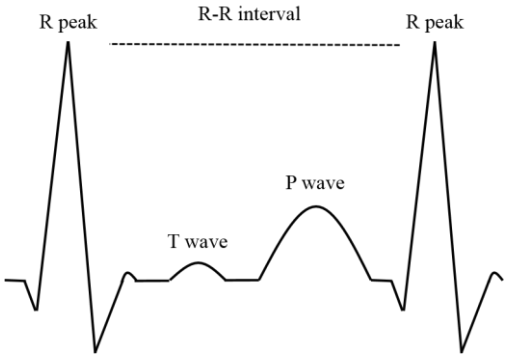

Fig. 4.10. R-R interval

As illustrated in the following figures, the feature extraction in the proposal mainly focus on two most significant parts: ST segment and T wave. The deviation at certain point and the area of these two parts are extremely crucial for ischemia and normal detection. These features of each heartbeat are described in Fig. 4.11 and Fig. 4.12.

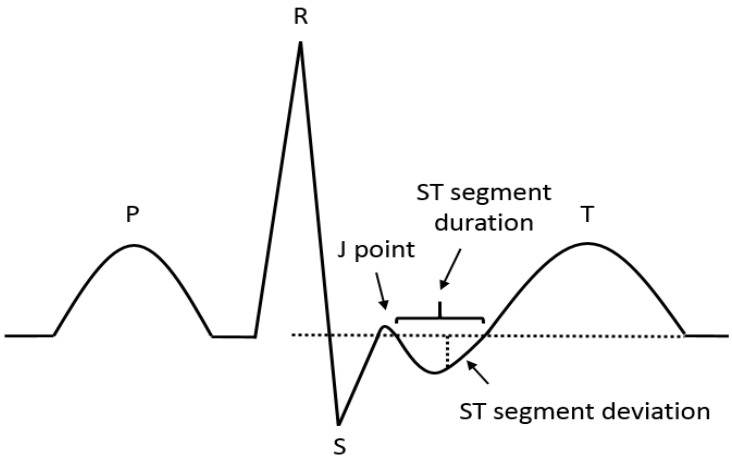

Fig. 4.11. Heartbeat features: ST segment part

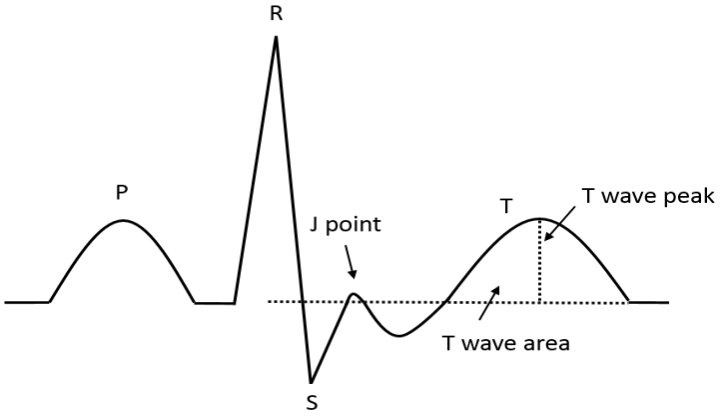

Fig. 4.12. Heartbeat features: T wave part

4.4.3 Classification based on Extracted Features

Support Vector Machine classification model is built offline on cloud server, and used to classify heartbeat to normal or ischemia based on extracted heartbeat features in the last step. SVMs have been widely applied for a variety of classification tasks, such as facial expression recognition [92], genetic analysis, and text classification [93]. The high accuracy on even small dataset and the highly variable generalization performance make SVMs especially suitable for real time processing [92].

Given training dataset $\{x_1,...,x_n\}$ in space $\mathrm{X} \subseteq \Re^d$, and their corresponding labels $\{y_1,...y_n\}$ where $y_i \in \{-1,1\}$, SVMs are hyperplanes which separate training dataset using a maximal margin. Support vectors refer to the training data that lie closest to the hyperplane. Generally speaking, training data in space X is projected to a higher dimensional feature space Γ via Mercer kernel operator K. The set of SVM classifiers are considered as the following form:

$$f(x) = \left(\sum_{i=1}^{n} \alpha_i \mathrm{K}(x_i, x) \right). \qquad (4.1)$$

When Mercer's condition is satisfied, $K(u,v) = \Phi(u) \cdot \Phi(v)$, where $\Phi : X \to \Gamma$ and "·" is inner product. Therefore, above $f(x)$ can be rewritten as:

$$f(x) = w \cdot \Phi(x), \text{ where } w = \sum_{i=1}^{n} \alpha_i \Phi(x_i). \qquad (4.2)$$

Hence, training data is projected into a higher dimensional feature space. SVMs then compute the maximal margin hyperplane which is defined as the sum of the distances of the hyperplane from the nearest data point in space Γ [93].

There are different types of SVM for both classification and regression, such as C-SVC, nu-SVC and epsilon-SVR, etc. Also, there are several kernel types for each SVM type. For example, linear kernel, polynomial kernel, and radial basis function kernel, and so on.

4.5 Experiments using European ST-T Database

4.5.1 Experiment Diagram

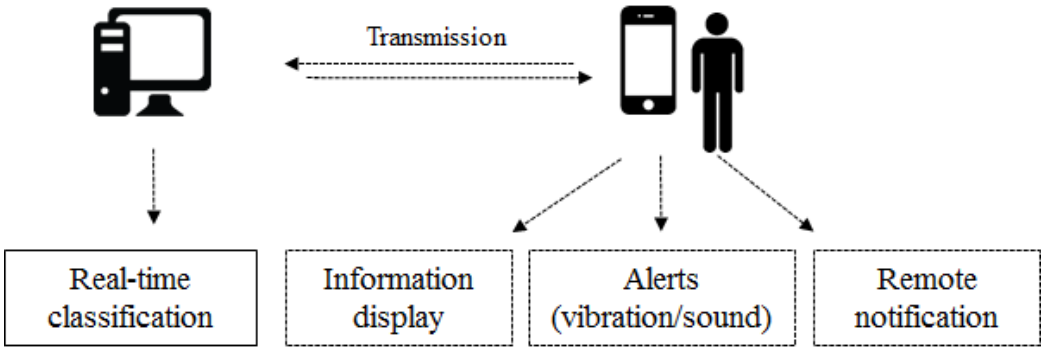

Fig. 4.13. Simulation experiment diagram

In the experiment, the monitoring application is developed as an iPhone App using Xcode on MacBook Air. The iPhone is simulated as a wearable smart heartbeat sensor, reads experiment data, transmits ECG signal to a PC which is the cloud server, where all the real time processing is programmed via Matlab software. The iPhone monitoring application is responsible for displaying important heartbeat information, giving alerts, including sound, vibration, also sending notification if necessary. The experiment diagram is shown in Fig. 4.13.

The open data source used in the experiment is from European ST-T Database. Each record in this database is two hours in duration and contains two signals, both sampled at 250 Hz with 12-bit resolution over a nominal 20 millivolt input range.

4.5.2 QRS Complex Detection Experiment

The recognition rate of applied online QRS complex detector, that bases on state-machine logic, is first tested using seven ECG recordings from experiment database. According to the 'Annotations' file of each recording, the total number of heartbeats for test is 58128. Each recording is processed through the online detector, the experiment results show that the R waves of 57227 heartbeats are successfully detected. The recognition rate is 98.45%, as shown in Table 4.1. Based on the detected QRS complex, heartbeat features are extracted for the following classification by SVM.

Table 4.1. Recognition rate of online QRS detector

Heartbeat number for test	Recognized R wave number	Recognition rate
58128	57227	98.45%

4.5.3 Classification Experiment

To achieve SVM classification outcomes in best effort, a practical and well accepted guide to SVM [94] is referred and followed in our experiments. A more productive procedure is proposed in this guide and show its improvements in some experiment examples. Also, a library for Support Vector Machines—LIBSVM [95] is applied to conduct SVM experiment. This library has been proved its effectiveness to SVM/SVR tasks, and widely used in many research and applications.

In order to obtain satisfactory result, cross validation on training dataset is an indispensable step as recommended in the guide [94]. The goal of cross validation is to identify relatively optimal parameters for SVM kernel so that the model can accurately predict unknown data. The over fitting problem that classification model has an excellent performance on training dataset, however, a poor performance on test dataset can be eliminated. In our experiment, the training dataset is divided into 5 subsets. Sequentially, one subset is tested using the trained model on the remaining 4 subsets. Then, the final obtained model is applied to classify on our test dataset. The SVM type used in experiment is C-SVC, and the kernel type is radial basis function.

In the experiment, the first recording of the European ST-T Database, e0103, is divided to training dataset and test dataset. Training dataset is utilized to train SVM classification model offline, while test dataset is applied to test the accuracy of built classification model and to be presented in the user interface of the monitoring application. There are 1172 heartbeats in the training dataset, among which 592 heartbeats are ischemia heartbeat. On the other hand, there are 6013 heartbeats in test dataset, 5868 are normal, the rest are ischemia heartbeats. The sensitivity and specificity of single heartbeat classification are 82.1% and 80.0%, respectively, as shown in Table 4.2. The proposed monitoring application dose not only rely on the classification of

each single heartbeat. In fact, it takes action according to whether an episode of ischemic heartbeat is found or not.

Table 4.2. Classification experiment results

	Classified as normal	Classified as ischemia	Total number
Normal	4755	1113	5868
ischemia	26	119	145

4.5.4 Application User Interface

The user interface is shown in Fig. 4.14 and Fig. 4.15. As indicated, the ECG signal is recorded and displayed in real time at the upper part of the application. Also, the parameters of the recorded ECG signal is shown in a tabular format. At the lower part, an alarm light is placed. When the real time recorded heartbeat is classified as Normal, all the parameters are displayed in green color meaning no danger found, and the alarm light stays green light, as well. However, once the heartbeat is classified as ischemic heartbeat, the alarm light is turned to red light with vibration and warning sound. Also, the parameters are presented in red color, indicating that something is not right. At the same time, the remote notification function is also turned on to give warnings to doctors and family members.

With the application like “MonitoringApp”, it is able to monitor the person who is under threat by myocardial ischemia in a very practical way. Moreover, due to that the sudden emergencies can be found on time, it becomes possible for people who stays home alone to get timely help, care, and rescue. The real time monitoring application creates a much safer and more reliable home atmosphere.

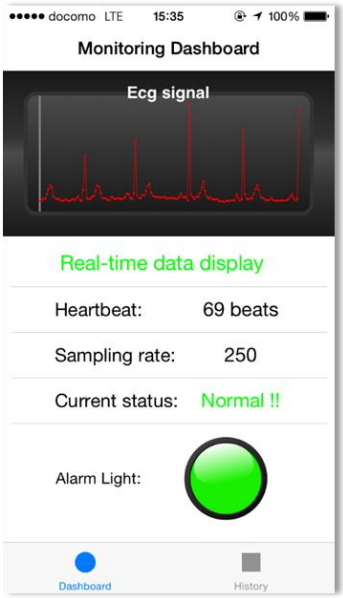

Fig. 4.14. "MonitoringApp": normal status

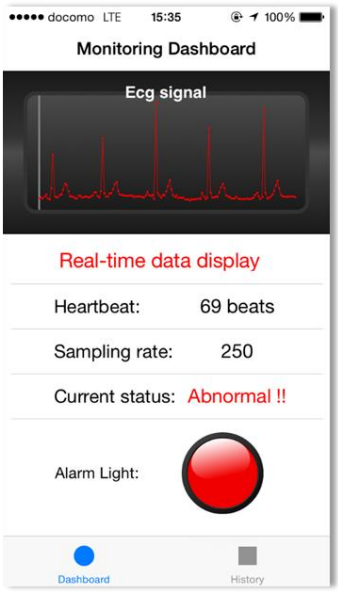

Fig. 4.15. "MonitoringApp": abnormal status

4.6 Chapter Summary

With the increase of aging population all over the world, it becomes more and more critical that how to provide proper care to such vast amount of people. The ischemia detection and real time monitoring in smart home care environment is proposed to avoid the circumstance that people get permanent harm or worse because of delayed help.

Due to the portability, ubiquity, and the rapidly developing intelligence of smart devices, it is utilized to record ECG signals, and perform real-time ischemia detection. The real time detection includes QRS complex discovery, J point detection, and the extraction of significant features, such as ST segment deviation, R-R interval, T wave amplitude, etc. Afterwards, SVMs are performed on these extracted features to classify the heartbeat signal to normal heartbeat or ischemic heartbeat.

In the experiment, ECG signal transmitted from iPhone application is first processed with a real time QRS complex detector. The recognition rate of the applied online detector is confirmed using 58128 heartbeats from seven recordings of European ST-T Database. The experiment results present the online detector has accuracy of 98.45%. Afterwards, a set of heartbeat features, such as R-R interval, ST segment deviation and T wave peak, etc, are obtained for classification. The used SVM classification model is built offline on cloud server. The SVM type used in experiment is C-SVC, and the kernel type is radial basis function. The accuracy, sensitivity and specificity of the classification experiment are 81.1%, 82.1% and 81.0%, separately. The monitoring application on iPhone receives real time classification results from cloud server, displays significant heartbeat information, and sends alerts, alarms and notifications if needed.

Given the fact that it is almost impossible for the people who suffers from kinds of heart disease get all-time nursing service or stay at hospital all the time, the integration of distinct technologies, like diverse smart devices, in smart home care environment is going to make a real difference. Although the proposal only focus on cardiovascular

ischemia at present, it definitely takes not very long to realize the more comprehensive and much closer monitoring of people's well-being, with the extremely fast development of various technologies on smart devices.

Chapter 5

Conclusions and Future Perspective

5.1 Conclusions

To deal with the fact that the population of elderly people is increasing very dramatically and the challenge that they cannot get timely treatment or rescue when they suffers from a sudden attack of myocardial ischemia at home, an application built in smart home care environment is proposed to provide real-time myocardial ischemia detection and well-being monitoring. And, a method based on fuzzy association rule mining is proposed as an assistant reference to help doctor to make ischemia diagnosis.

In chapter 2, the research focus on approximation representation for time series data mining. The proposal is based on one of the symbolic approximation representations, called Symbolic Aggregate Approximation (SAX). Despite of its vast implementation, it is noticed that if the subsequence direction can be considered and added to SAX, it might avoid false alarms of time series data mining tasks and reduce error rate of tasks, such as classification. According to the results of one-nearest-neighbor classification on UCR time series public service data source, the error rate is reduced by averagely 16.22% on each test dataset. Also, a paired t-test is conducted on the error rate results of both SAX and the proposal. The p-value of the t-test indicates that the null hypothesis that the proposal has no effect is rejected.

In chapter 3, the purpose of the proposal, which is based on fuzzy association rule mining, is to help doctor to make myocardial ischemia diagnosis, since the proposal is

able to provide interpretable and understandable information. Its implementation includes four steps: feature extraction from each single heartbeat, fuzzy transformation of extracted features, association rule mining on fuzzy itemsets, and automatic classification of ischemia and normal heartbeats. Data used in the experiment are the ECG recordings from European ST-T Database. The rule mining results are that 43 rule are extracted. Among them, 19 rules are for normal heartbeat while 24 rules are for ischemia heartbeats. Afterwards, the extracted rules are applied to classify the heartbeats in the test dataset. The classification results demonstrate that the sensitivity and specificity are 83.4% and 80.7%, respectively. Although diverse intelligent methods for myocardial ischemia detection have been proposed, the practical experience of professional doctors are still necessary for ischemia diagnosis all the time. The proposal provide interpretable information, and can be adopted as reliable medium to help the ischemia diagnosis.

In chapter 4, an application that detects ischemia using smart devices and provides real-time monitoring is proposed and applied to smart home care environment. At present, one of the biggest challenges for human society is that the population of elderly increases rapidly. Moreover, a large portion of this population suffers from different diseases, including myocardial ischemia, arrhythmia, etc. When they are left alone at home, there exists a great danger that they cannot get timely help when troubled by sudden burst of certain disease. Therefore, a smart home care environment that can detect abnormality and monitor in real-time becomes definitely indispensable. The proposal takes advantage of the portability and ubiquity of smart devices, utilize them to record ECG signals and perform real-time myocardial ischemia detection. The implementation of the proposal includes QRS complex detection, J point detection, feature extraction and classification via SVM. If classification results show any abnormality, the monitoring process begin to work. Through smart devices, alarms and notifications are sent to call professional doctors, ambulance, and notify their family members. With such application in smart home care environment, the tragedy that

people get permanent harm or even worse can be substantially reduced.

5.2 Future Perspective

Cardiovascular disease, commonly known as heart disease, has been the leading cause of deaths worldwide since the 1970s. Kinds of algorithms, methods have been proposed to detect and predict different heart diseases, such as, hypertensive heart disease, arrhythmia, and cardiomyopathy, etc. Yet, the current diagnosis for these heart diseases cannot be isolated from the years of professional, practical experience from doctors. The proposal based on association rule mining can provide interpretable, understandable results to professional doctors. It is able to be an effective and meaningful assistant medium and act a positive and helpful role in the diagnosis of heart disease. The proposed method aims to be applied to the diagnosis of other heart diseases, for example, arrhythmia which is another life-threatening, troubling heart disease. On the other hand, the proposed application in smart home care environment currently focus on the detection and monitoring myocardial ischemia. Similarly, other heart diseases are also available to become monitoring targets by implementing different real time detection algorithms.

In the future, more convenient and practical smart devices are expected to be applied in smart home care environment. For instance, as a common sense, smart watch is attaching to people's skin all the time, which makes it more accurate to collect real time information, and faster for people to feel the alarms sent by the smart devices. Besides, the fact that more and more sensors are being deployed in t-shirts make them the ideal smart devices. First, the necessity is guaranteed, because wearing shirts is definitely unavoidable for the most of people. In addition, smart shirt is not like smartphone or smartwatch, which have limitations on usage duration, and also requires constant battery recharging, and so on.

At present, only ECG signals are recorded by smart devices in the smart home care environment. Yet, with the rapidly developing of technologies that applied on smart devices, a lot of other significant information can be recorded and gathered. Through these important information, the functions of smart home care environment are able to be expanded to great quantity of aspects. For example, human motion gait at home can be tracked and logged by smart devices. According to motion recording, the quantity of exercise, and the amount of burned calorie every day can be calculated. Then, such results are used as reference for users to keep track of their life-style. The ultimate purpose is to urge users to live with a healthy life-style. Besides, the data analysis in the future may not only limit to individual. Big data can be built on cloud via gathering the information from a group of people instead of single person. Novel knowledge mining, abnormality discovery, and so on, are going to be applied on such big data. Users may get a variety of feedbacks regarding a wide range of aspects, such as advices to fitness or lifestyle, and predictions to many forms of abnormalities.

Bibliography

[1] http://en.wikipedia.org/wiki/Time_series, August, 2012.

[2] J. Lin, E. Keogh, "A Symbolic Representation of Time Series, with Implications for Streaming Algorithms", 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 2-11, 2003.

[3] B. Lkhagava, Y. Suzuki, K. Kawagoe, "Extended SAX: Extension of Symbolic Aggregate Approximation for Financial Time Series Data Representation", Data Engineering Workshop, 2006.

[4] M. M. M. Fuad, P.F. Marteau, "Enhancing the Symbolic Aggregate Approximation Method Using Updated Lookup Tables", Knowledge-Based and Intelligent Information and Engineering Systems, pp. 420-431, 2010.

[5] G. Li, L. P. Zhang, L. Q. Yang, "TSX: A Novel Symbolic Representation for Financial Time series", Trends in Artificial Intelligence, pp. 262-273, 2012.

[6] H. Ding, E. Keogh, "Querying and Mining of Time Series Data: Experimental Comparison of Representations and Distance Measures", Proceedings of VLDB Endowment, Vol. 1, No. 2, pp. 1542-1552, 2010.

[7] B. K. Yi, C. Faloutsos, "Fast Time Sequence Indexing for Arbitrary Lp Norms", In processing of the Very Large Database, pp. 385-394, 2000.

[8] E. Keogh, Q. Zhu, B. Hu, Y. Hao, X. Xi, L. Wei, C. A. Ratanamahatana. (2011). The UCR Time Series Classification/Clustering. URL: www.cs.ucr.edu/~eamonn/time_series_data/

[9] X. Wang, H. Ding, G. Trajcevski, P. Scheuermann, E. Keogh, "Experimental comparison of representation methods and distance measures for time series data",

In Proceedings of CoRR, 2010.

[10] P. Siirtola, H. Koskimaki, V. Huikari, P. Laurinen, J. Roning, "Improving the classification accuracy of stream data using SAX similarity features", Pattern Recognition Letters, Vol. 32, No. 13, pp 1659-1668, 2011.

[11] A. Sant'Anna, N. Wickstrom, A. Salarian, "A new measure of movement symmetry in early Parkinson's disease patients using symbolic processing of inertial sensor data", IEEE Trans. Biomed. Engineering, Vol. 58, No. 7, pp 2127-2135, 2011.

[12] E. Keogh, J. Lin and A. Fu, "HOT SAX: Efficiently Finding the Most Unusual Time Series Subsequence", In Proc. of the 5th IEEE International Conference on Data Mining (ICDM 2005), pp. 226 - 233., Houston, Texas, Nov 27-30, 2005.

[13] C. Ratanamahatana, E. Keogh, T. Bagnall, and S. Lonardi, "A Novel Bit Level Time Series Representation with Implications for Similarity Search and Clustering", PAKDD, 2005.

[14] B. Chiu, E. Keogh, and S. Lonardi, "Probabilistic Discovery of Time Series Motifs", the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 493-498, August 24 - 27, 2003. Washington, DC, USA.

[15] P. Patel, E. Keogh, J. Lin, and S. Lonardi, "Mining Motifs in Massive Time Series Databases", In proceedings of the 2002 IEEE International Conference on Data Mining, Maebashi City, Japan, Dec 9-12.

[16] E. Keogh, J. Lin, and W. Truppel, "Clustering of Time Series Subsequences is Meaningless: Implications for Past and Future Research", In proceedings of the 3rd IEEE International Conference on Data Mining, pp. 115-122, Melbourne, FL, Nov 19-22.

[17] L. Wei, E. Keogh, and X. P. Xi, "SAXually Explict Images: Finding Unusual Shapes", ICDM, 2006.

[18] X. P. Xi, E. Keogh, L. Wei, and A. Mafra-Neto, "Finding Motifs in a Database of Shapes", SIAM International Conference on Data Mining, 2007.

[19] B. Hu, T. Rakthanmanon, B. Campana, A. Mueen, and E. Keogh, “Image Mining of Historical Manuscripts to Establish Provenance”, SDM, pp. 804-815, SIAM / Omnipress, 2012.

[20] A. M. Ahmed, A. A. Bakar, and A. R. Hamdan, “Improved SAX time series data representation based on Relative Frequency and K-Nearest Neighbor Algorithm”, 2010 10th International Conference on Intelligent Systems Design and Applications (ISDA), Nov. 29 - Dec. 1, 2010.

[21] N. Q. V. Hung, D. T. Anh, “Combining SAX and Piecewise Linear Approximation to Improve Similarity Search on Financial Time Series”, International Symposium on Information Technology Convergence, ISITC 2007, Nov 23 -24, 2007.

[22] M. Fabri, G. Mascioli, G. Palonara, A. M. Perdon, and S. R. Viola, “Activation and delay in FMRI brain signals of selective attention”, In Proceedings of International IJCNN07 Workshop on Neurodynamics, Orlando, Florida, USA, August 17, 2007.

[23] J. D. Scheff, R. R. Almon, D. C. DuBois, W. J. Jusko, and I. P. Androulakis, “A New Symbolic Representation for the Identification of Informative Genes in Replicated Microarray Experiments”, Journal of Integrative Biology, Vol. 14, No. 3, pp. 239-248, 2010.

[24] Y. Tanaka, K. Uehara, “Motif Discovery Algorithm from Motion Data”, In proceedings of the 18th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI), Kanazawa, Japan, June 2-4, 2004.

[25] V. Ergovic, S. Tonkovic, and V. Medved, “Human Gait Data Mining by Symbol Based Descriptive Features”, World Congress on Medical Physics and Biomedical Engineering, Vol. 25, No. 9, pp. 460-463, Munich, Germany, September 7-12, 2009.

[26] http://www.sads.org.uk/cardiac_tests.htm, April, 2014.

[27] http://en.wikipedia.org/wiki/QT_interval, April, 2014.

[28] http://en.wikipedia.org/wiki/Electrocardiography, April, 2014.

[29] http://en.wikipedia.org/wiki/Cardiovascular_disease, April, 2014.

[30] H. Khorrami, M. Moavenian, “A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification”, Expert Systems with Applications, Vol. 37, No. 8, pp. 5751-5757, August 2010.

[31] Y. Ozbay, R. Ceylan, and B. Karlik, “A Fuzzy Clustering Neural Network Architecture for Classification of ECG Arrhythmias”, Computers in Biology and Medicine, Vol. 36, No. 4, pp. 376-388, April 2006.

[32] H. M. Rai, A. Trivedi, and S. Shukla, “ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier”, Measurement, Vol. 46, No. 9, pp. 3238-3246, November 2013.

[33] R. Ceylan, Y. Ozbay, “Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network”, Expert Systems with Applications, Vol. 33, No. 2, pp. 286-295, August 2007.

[34] T. Froese, S. Hadjiloucas, R. K. H. Galvao, V. M. Becerra, and C. J. Coelho, “Comparison of extrasystolic ECG signal classifiers using discrete wavelet transforms”, Pattern Recognition Letters, Vol. 27, No. 5, pp. 393-407, April 2006.

[35] S. Yu, Y. Chen, “Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network”, Pattern Recognition Letters, Vol. 28, No. 10, pp. 1142-1150, July 2007.

[36] R. Cervigon, C. Sanchez, F. Castells, J. M. Blas, and J. Millet, “Wavelet analysis of electrocardiograms to characterize recurrent atrial fibrillation”, Journal of the Franklin Institute, Vol. 344, No. 3-4, pp. 196-211, May-July 2007.

[37] E. Ubeyli, “Support vector machines for detection of electrocardiographic changes in partial epileptic patients”, Engineering Applications of Artificial Intelligence, Vol. 21, No. 8, pp. 1196-1203, December 2008.

[38] S. Abe, “Support Vector Machines for Pattern Classification”, NJ: Springer-Verlag, London, 2005.

[39] N. Acir, “A support vector machine classifier algorithm based on a perturbation

method and its application to ECG beat recognition systems", Expert Systems with Applications, Vol. 31, No. 1, pp. 150-158, July 2006.

[40] C. P. Shem, W. C. Kao, Y. Y. Yang, M. C. Hsu, Y. T. Wu, and F. P. Lai, "Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines", Expert Systems with Applications, Vol. 39, No. 9, pp. 7845-7852, July 2012.

[41] B. S. Raghavendra, D. Bera, A. S. Bopardikar, and R. Narayanan, "Cardiac arrhythmia detection using dynamic time warping of ECG beats in e-healthcare systems", Proceedings of the 2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, pp. 1-6, IEEE Computer Society, Washington DC, USA, 2011.

[42] C. Ye, M. T. Coimbra, and B. V. K. V. Kumar, "Arrhythmia Detection and Classification using Morphological and Dynamic Features of ECG Signals", 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.1918-1921, Buenos Aires, Argentina, August 31 – September 4, 2010.

[43] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health", Expert Systems with Applications, Vol. 39, No. 14, pp. 11792-11800, October 2012.

[44] T. Ota, H. Morita, "On Real-Time Arrhythmia Detection in ECG Monitors Using Antidictionary Coding", the 2012 International Symposium on Information Theory and its Applications, pp. 194-198, Honolulu, USA, October 28-31 2012.

[45] J. Park, W. Pedrycz, and M. Jeon, "Ischemia episode detection in ECG using kernel density estimation, support vector machine and feature selection", BioMedical Engineering OnLine, 2012.

[46] I. Babaoglu, O. Findik, and M. Bayrak, "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine", Expert

Systems with Applications, Vol. 37, No. 3, pp. 2182-2185, March 2010.

[47] C. A. Bustamante, S. I. Duque, A. Orozco-Duque, and J. Bustamante, "ECG Delineation and Ischemic ST-Segment Detection Based in Wavelet Transform and Support Vector Machines", Pan American Health Care Exchanges (PAHCE) Conference, Workshops and Exhibits Cooperation, Medellin, Colombia, pp. 1-7, April 29 – May 4, 2013.

[48] N. Maglaveras, T. Stamkopoulos, C. Pappas, and M. G. Strintzis, "An Adaptive Backpropagation Neural Network for Real-Time Ischemia Episodes Detection: Development and Performance Analysis using the European ST-T Database", IEEE Transactions on Biomedical Engineering, Vol. 45, No. 7, pp. 805-813, 1998.

[49] S. Papadimitriou, S. Mavroudi, L. Vladutu, and A. Bezerianos, "Ischemia Detection with a Self-Organizing Map Supplemented by Supervised Learning", IEEE Transactions on Neural Networks, Vol. 12, No. 3, pp. 503-515, 2001.

[50] C. Papaloukas, D. I. Fotiadis, A. Likas, L. K. Michalis, "An Ischemia Detection Method based on Artificial Neural Networks", Artificial Intelligence in Medicine, Vol. 24, No. 2, pp. 167-178, 2002.

[51] R. V. Andreao, B. Dorizzi, J. Boudy, and J. Mota, "ST-segment Analysis using Hidden Markov Model Beat Segmentation: Approach to Ischemia Detection", Computers in Cardiology, pp. 381-384, Sep 19-22, 2004.

[52] P. Ranjith, P. C. Baby, and P. Joseph, "ECG Analysis using Wavelet Transform: Application to Myocardial Ischemia Detection", ITBM-RBM, Vol. 24, No. 1, pp. 44-47, 2003.

[53] C. Papaloukas, D. Fotiadis, A. Likas, C. Stroumbis, L.K. Michalis, "Use of a Novel Rule-based Expert System in the Detection of Changes in the ST segment and the T wave in Long Duration ECGs", Journal of Electrocardiology, Vol. 35, No. 1, 2002.

[54] A. Bakhshipour, M. Pooyan, H. Mohammadnejad, and A. Fallahi, "Myocardial Ischemia Detection with ECG Analysis, Using Wavelet Transform and Support

Vector Machines", Proceedings of the 17th Iranian Conference of Biomedical Engineering, pp. 1-4, Isfahan, Nov 3-4, 2010.

[55] http://www.mayoclinic.org/diseases-conditions/myocardial-ischemia/basics/tests-diagnosis/CON-20035096, April, 2014.

[56] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg, and C. Marchesi, "The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography", European Heart Journal, Vol. 13, No. 9, pp. 1164-1172, 1992.

[57] A. L. Goldberger, L. Amaral, L. Glass, J. M. Hausdorff, P. 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. 215-220, 2000.

[58] M. Kaur, B. Singh, and Seema, "Comparisons of Different Approaches for Removal of Base Line Wander from ECG signal", Proceedings of International Conference & Workshop on Emerging Trends in Technology(ICWET), pp. 1290-1294, Mumbai, 2011.

[59] R. Dev, "Different Techniques to Remove Baseline Wander from ECG Signal: A Review", VSRD International Journal of Electrical, Electronics & Communication Engineering, Vol. 2, No. 7, pp. 532-537, 2012.

[60] B. Mozaffary, A. Tinati, "ECG baseline wander elimination using wavelet packets", World Academy of Science, Engineering and Technology, pp. 14-16, 2005.

[61] V. S. Chouhan, S. S. Mehta, "Total removal of baseline drift from ECG signal", International Conference on Computing: Theory and Applications, pp. 512-515, 5-7 March 2007.

[62] M. S. Chavan, R. Agarwala, and M. D. Uplane, "Use of Kaise window for ECG processing", Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation, pp. 285-289, Madrid, Spain, 15-17 February

2006.

[63] L. P. Harting, N. M. Fedotov, and C. H. Slump, “On baseline drift suppressing in ECG-recordings”, Proceedings of the 2004 IEEE Benelux Signal Processing Symposium, pp. 133-136, 2004.

[64] T. Li, F. Dong, and K. Hirota, “Distance Measure for Symbolic Approximation Representation with Subsequence Direction for Time Series Data Mining”, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 17, No. 2, pp. 263-271, 2013.

[65] T. P. Exarchos, C. Papaloukas, D. I. Fotiadis, and L. K. Michalis, “An Association Rule Mining-Based Methodology for Automated Detection of Ischemic ECG Beats”, IEEE Transactions on Biomedical engineering, Vol. 53, No. 8, pp. 1531-1540, 2006.

[66] C. M. Kuok, A. Fu, and M. H. Wong, “Mining Fuzzy Association Rules in Databases”, ACM SIGMOD, Vol. 27, No. 1, pp. 41-46, 1998.

[67] J. P. Betancourt, C. Fatichah, M. L. Tangel, F. Yan, J. A. G. Sanchez, F. Y. Dong, and K. Hirota, “Similarity-Based Fuzzy Classification of ECG and Capnogram Signals”, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 17, No. 2, pp. 302-310, 2013.

[68] M. Delgado, “Mining fuzzy association rules: an overview”, BISC Conference, Decemebr 2003.

[69] R. Agrawal, T. Imielinski, and A. Swami, “Mining association rules between sets of items in large databases”, Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, pp. 207-216, New York, USA, 1993.

[70] C. H. Chen, T. P. Hong, and V. S. Tseng, “Fuzzy Data Mining for Time-Series Data”, Applied Soft Computing, Vol. 12, No. 1, pp. 536-542, 2012.

[71] X. Yin, J. Han, “CPAR: Classification based on Predictive Association Rules”, Proceedings of the 2003 SIAM International Conference on Data Mining, San

Francisco, May 1-3, 2003.

[72] B. Liu, W. Hsu, and Y. Ma, "Integrating Classification and Association Rule Mining", KDD-98 Proceedings, New York, Aug 27-31, 1998.

[73] W. H. Au, K. C. Chan, "Mining Fuzzy Rules for Time Series Classification", Proceedings of IEEE International Conference on Fuzzy Systems, Budapest, Hungary, 25-29 July, 2004.

[74] A. Tajbakhsh, M. Rahmati, and A. Mirzaei, "Intrusion Detection using Fuzzy Assocation Rules", Applied Soft Computing, Vol. 9, No. 2, pp. 462-469, March 2009.

[75] J. A. Fdez, R. Alcala, and F. Herrera, "A Fuzzy Association Rule-based Classification Model for High-Dimensional Problems with Genetic Rule Selection and Lateral Tuning", IEEE Transactions on Fuzzy Systems, Vol. 19, No. 5, pp. 857-872, October 2011.

[76] Z. L. Chen, G. Q. Chen, "Building An Associative Classifier based on Fuzzy Association Rules", International Journal of Computational Intelligence Systems, Vol. 1, No. 3, pp. 262-273, August 2008.

[77] Y. Yi, E. Hullermeier, "Learning Complexity-Bounded Rule-based Classifier by Combining Association Analysis and Genetic Algorithms", 4th Conference of the European Society for Fuzzy Logic and Technology, pp. 47-52, 7-9 September, 2005.

[78] Bache, K. & Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[79] F. K. Aldrich, "Smart Homes: Past, Present and Future", Inside the Smart Home, pp. 17-39, 2003.

[80] E. Dishman, "Inventing Wellness Systems for Aging in Place", IEEE Computer, Vol. 37, No. 5, pp. 34-41, May 2004.

[81] P. Leijdekkers, V. Gay, and E. Lawrence, "Smart Homecare System for Health

Tele-monitoring", Proceedings of the Fisrt International Conference on the Digital Society, pp. 3-7, 2-6 January 2007.

[82] http://www.wearabledevices.com/, August, 2014.

[83] "First US surgery transmitted live via Google Glass", Medical Xpress, 27 August 2013.

[84] "http://www.inquisitr.com/1224638/google-glass-connects-breastfeeding-moms-with-lactation-help/", Inquisitr, 12 June 2014.

[85] http://www.apple.com/watch/, August, 2014.

[86] M. S. Chavan, R. A. Agarwala, and M. D. Uplane, "Suppression of Baseline Wander and Power Line Interference in ECG using Digital IIR Filter", International Journal of Circuits, Systems And Signal Processing, Vol. 2, No. 2, pp. 356-365, 2008.

[87] F. A. Afsar, M. S. Riaz, and M. Arif, "A Comparison of Baseline Removal Algorithms for Electrocardiogram (ECG) based Automated Diagnosis of Coronary Heart Disease", 3rd International Conference on Bioinformatics and Biomedical Engineering , pp. 1-4, Beijing, China, June 11-13, 2009.

[88] J. Pan, W. J. Tompkins, "A Real-Time QRS Detection Algorithm", IEEE Transactions on Biomedical Engineering, Vol. 32, No. 3, pp. 230-236, March 1985.

[89] H. H. So, "Development of QRS Detection Method for Real-time Ambulatory Cardiac Monitor", Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 282-292, 1997.

[90] P. Hamilton, "Open Source ECG Analysis", IEEE Computers in Cardiology, Vol. 29, No. 1, pp. 101-104, September 2002.

[91] Hooman Sedghamiz, "An online algorithm for R, S and T wave detection", December 2013. Linkoping university. URL: http://www.mathworks.com/matlabcentral/fileexchange/45404-ecg-q-r-s-wave-online-detector

[92] P. Michel, R. E. Kaliouby, “Real Time Facial Expression Recognition in Video using Support Vector Machines”, Proceedings of the 5th International Conference on Multimodal Interfaces, pp. 258-264, New York, USA, 2003.

[93] S. Tong, D. Koller, “Support Vector Machine Active Learning with Applications to Text Classification”, The Journal of Machine Learning Research, Vol. 2, pp. 45-66, March 2002.

[94] C. Hsu, C. Chang, C. Lin, “A Practical Guide to Support Vector Classification”, http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf , 2010.

[95] C. Chang and C. Lin, “LIBSVM : a library for support vector machines”, ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2011.

Related Publications

Journal Papers

[J1] **Tianyu Li**, Fang-Yan Dong, Kaoru Hirota, "Distance Measure for Symbolic Approximation Representation with Subsequence Direction for Time Series Data Mining", *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 17, No. 2, pp. 263-271, 2013.

[J2] **Tianyu Li**, Fang-Yan Dong, Kaoru Hirota, "Fuzzy Association Rule Mining based Myocardial Ischemia Diagnosis on ECG Signal", *Journal of Advanced Computational Intelligence and Intelligent Informatics*, to appear in Vol. 19 No. 2, 2015.

[J3] **Tianyu Li**, Fang-Yan Dong, Kaoru Hirota, "Myocardial Ischemia Monitoring Application based on Computational Intelligence in Smart Home Care Environment", *Applied Soft Computing*, Elsevier. (Submitted).

International Conference Papers

[C1] **Tianyu Li**, Fang-Yan Dong, Kaoru Hirota, "Distance Measure of Symbolic Aggregate Approximation with Direction Representation for Time Series Data Mining", *International Symposium on Soft Computing*, Tokyo, Nov 8-9, 2012.

[C2] **Tianyu Li**, Fang-Yan Dong, Kaoru Hirota, "Ischemia Diagnosis using Fuzzy Association Rule Mining on ECG Signal", *The Joint International Conference of ITCA2014&ISCIIA2014*, Changsha, Hunan, China, 16-21 September, 2014.