

Visual Study of ECG Signal Based on Variant Measurement Model

Yan Ji^{1*}, Jeffrey Zheng¹, Lu Shan²

¹School of Software, Yunnan University, Kunming Yunnan

²Kunming Phuket Hospital, Kunming Yunnan

Email: kunmingjy@sina.com

Received: Dec. 27th, 2015; accepted: Jan. 12th, 2016; published: Jan. 15th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

ECG signal is an important basis of clinical experts diagnosing heart disease. At present, it had become a fact through using electrocardiograph to record ECG data sequences of 1D and storage. However, it is difficult for medical personals to detect changing of ECG data itself from mass ECG data sequences. Because of this, the paper uses variant model and visualization method to process ECG data sequences through input parameters and getting 2D maps distribution characteristics. The paper introduces process model of ECG data sequences and 2D maps simple analysis. The visualization results provide reference for deeply resolution of distribution characteristics ECG data.

Keywords

ECG Signal, Variant Measurement, Visualization, 2D Maps

基于变值测量模型的心电信号可视化研究

吉 艳^{1*}, 郑智捷¹, 山 路²

¹云南大学软件学院, 云南 昆明

²昆明普济医院, 云南 昆明

Email: kunmingjy@sina.com

收稿日期: 2015年12月27日; 录用日期: 2016年1月12日; 发布日期: 2016年1月15日

*通讯作者。

摘要

心电信号是临床专家诊断心脏病的重要依据。目前通过心电图机读取1D的心电数据序列并进行存储已成普遍事实，但是，医务人员从存储的批量数据中很难观察到数据本身的变化情况。为了方便医务人员观察每种心电数据的变化情况，本文使用变值模型和可视化方法对心电数据序列进行处理，通过参数输入得到2D图谱的方式展示正常和异常心电的分布特征。本文重点介绍心电数据的处理模型以及对输出的2D图谱进行简单分析。这些分析和展示的结果为深入解析心电数据的分布特性和可视化研究提供参考。

关键词

心电信号，变值测量，可视化方法，2D图谱

1. 背景

心电信号是生命体最重要的生物电信息之一，也是临床医生诊断心血管疾病的重要依据[1]，对心电数据序列进行处理以及以散点图的形式输出处理结果具有十分重要的实用价值。心电散点图是一项以非线性分析方法为基础，运用计算机软件技术将较长时间记录的心电图数据转换为高密度散点数据，制成直观、简约图形表示的前沿技术[2]。心电散点图将传统的1D连续心电图信号序列表现为高维图示化散点表示。大数据时代，利用现代概率和统计分析工具对心电数据进行处理，有助于提取和挖掘隐含在大样本的心电数据中，难以被传统心电图识别技术识别出来的特殊类别的非线性生理和病理信息。在这个研究方向，目前可以看到的处理模型主要是基于近代混沌理论和法国科学家彭加莱100年前提出的图示处理方法。这类方法已经在判别不同的心律失常疾病中逐渐发挥作用[2]。

由于特定疾病类型可能具备相似的变化规律，例如：如何根据长时间心电监测数据，快速、精确地区分出正常心律和心律失常的心电图中蕴含的特殊差别，对探讨的长时间心电数据的计算机辅助诊断和分析模式是一个值得进行深度研究的课题。本文使用变值可视化特征分析技术处理心电数据序列，根据序列密码分析模式正在发展的一类原创型的连续信号分析技术，该类技术能有效地将1维的连续心电图信号，分解成为四基元的伪基因多路离散信号，从而利用一套有效地用于序列密码分析的前沿技术工具，将所处理的信号形成可视化分布图式。本文中的心电数据序列经过处理之后得到的输出结果是一系列的正常和异常心电的2D图谱。

2. 体系结构

处理心电数据序列的整个框架结构如图1所示：

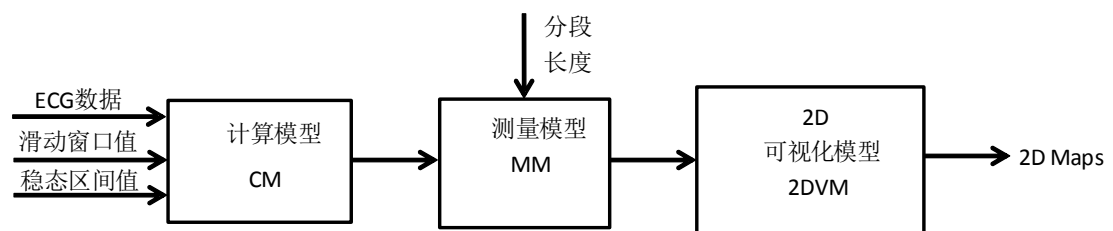


Figure 1. Architecture
图 1. 体系结构图

图 1 中的计算模型(CM)包括基值计算模型(BVCM)、范围值计算模型(RVCM)以及区间取值计算模型(RCM)。下面就这三个模型详细介绍。图 2 是基值计算模型(BVCM)，图 3 是范围值计算模型(RVCM)，图 4 是区间取值计算模型(RCM)。

图 1 中的测量模型(MM)可以使用图 5 进行详细介绍。

2.1. 计算模型(CM)

本文中所使用的输入参数有： w 表示滑动窗口值， s 表示稳态区间值， m 表示分段段长。

2.1.1. 基值计算模型(BVCM)

基值计算模型(BVCM)的作用是通过公式 1 对连续心电数据序列 $\{a_0, a_1, a_2, \dots, a_i, \dots, a_n\}$ 进行处理，从而获得相应的基值序列 $BV(\{bv_0, bv_1, bv_2, \dots, bv_j, \dots, bv_{n-w+1}\})$ 。



Figure 2. Basic value computation model (BVCM)

图 2. 基值计算模型(BVCM)



Figure 3. Range value computation model (RVCM)

图 3. 范围值计算模型(RVCM)

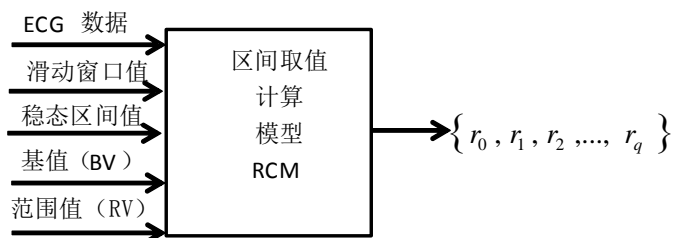


Figure 4. Region computation model (RCM)

图 4. 区间取值计算模型(RCM)

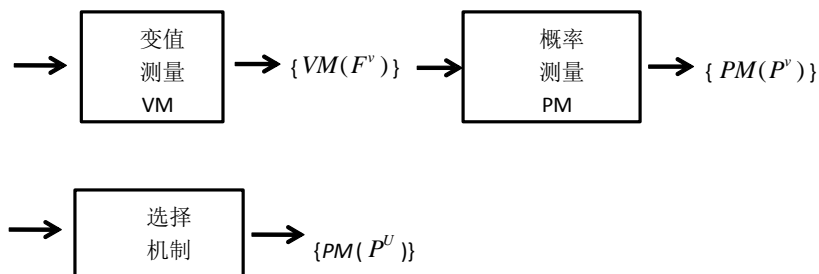


Figure 5. Measurement model (MM)

图 5. 测量模型(MM)

$$bv_j = \frac{1}{w} \sum_{i=j}^{j+w-1} a_i \quad (j=0,1,2,\dots,n-w+1; i=j, j+1, j+2,\dots,n) \quad (1)$$

输入:

心电数据序列: $\{a_0, a_1, a_2, \dots, a_i, \dots, a_n\}$;

滑动窗口值: w ;

输出:

基值序列: $\{bv_0, bv_1, bv_2, \dots, bv_j, \dots, bv_{n-w+1}\}$ 。

2.1.2. 范围值计算模型(RVCM)

范围值计算模型(RVCM)通过公式 2 计算范围值序列: $\{rv_0, rv_1, rv_2, \dots, rv_j, \dots, rv_{n-w+1}\}$ 。

$$rv_j = \frac{s}{2} \left(\max \{a_j, a_{j+1}, a_{j+2}, \dots, a_{j+w-1}\} - \min \{a_j, a_{j+1}, a_{j+2}, \dots, a_{j+w-1}\} \right) \quad (j=0,1,2,\dots,n-w+1) \quad (2)$$

$\max \{a_j, a_{j+1}, a_{j+2}, \dots, a_{j+w-1}\}$: 表示包含在滑动窗口内的心电数据序列的最大值;

$\min \{a_j, a_{j+1}, a_{j+2}, \dots, a_{j+w-1}\}$: 表示包含在滑动窗口内的心电数据序列的最小值。

输入:

心电数据序列: $\{a_0, a_1, a_2, \dots, a_i, \dots, a_n\}$;

滑动窗口值: w ;

稳态区间值: s ;

输出:

范围值序列: $\{rv_0, rv_1, rv_2, \dots, rv_j, \dots, rv_{n-w+1}\}$ 。

2.1.3. 区间取值计算模型(RCM)

区间取值计算模型(RCM)通过公式 3 计算心电数据序列的区间值: $\{c_0, c_1, c_2, \dots, c_j, \dots, c_{n-w+1}\}$ 。

$$c_j = \frac{a_j - bv_j}{s * rv_j} \quad (j=0,1,2,\dots,n-w+1) \quad (3)$$

经过以上三个关键计算模型的处理,连续的心电信号数据序列 $\{a_0, a_1, a_2, \dots, a_i, \dots, a_n\}$ 被转换为区间取值序列 $R(c)\{c_0, c_1, c_2, \dots, c_j, \dots, c_{n-w+1}\}$ 。现根据区间范围参数表示如下:

$$D = \begin{cases} A, & c_i \geq s; \\ G, & 0 \leq c_i < s; \\ C, & -s \leq c_i < 0; \\ T, & c_i < -s; \end{cases} \quad i=0,1,2,\dots,n-w+1; \quad (4)$$

经过上述比较,进一步将区间取值序列 $R(c)\{c_0, c_1, c_2, \dots, c_j, \dots, c_{n-w+1}\}$ 转换为由生物学上的四个碱基 A,G,C,T 组成的伪 DNA 序列。

输入:

心电数据序列: $\{a_0, a_1, a_2, \dots, a_i, \dots, a_n\}$;

稳态区间值: s ;

基值序列: $BV(bv)\{bv_0, bv_1, bv_2, \dots, bv_j, \dots, bv_{n-w+1}\}$;

范围值序列: $RV(rv)\{rv_0, rv_1, rv_2, \dots, rv_j, \dots, rv_{n-w+1}\}$;

中间参数:

区间取值序列: $R(c)\{c_0, c_1, c_2, \dots, c_j, \dots, c_{n-w+1}\}$

输出:

$\{IV(R^Q)\}$ 区间范围取值向量: $Q = \{A, G, C, T\}$

2.2. 测量模型(MM)

测量模型(MM)的主要作用是将伪 DNA 序列转换成 0-1 序列, 根据输入的分段段长 m 将 0-1 序列转换为逻辑空间的四元符号: $\perp, +, -, T$ 组成的序列[3], 最后分别运用归一化和非归一化方法计算四个基元的概率测度。

2.2.1. 变值模型(VM)

变值模型(VM)的作用是把伪 DNA 序列转换为由变值逻辑空间的四个基元组成的序列(图 6)。

输入:

$\{IV(R^Q)\}$ 区间范围取值向量: $Q = \{A, G, C, T\}$;

分段段长: m ;

中间参数:

0-1 序列: 由 0 和 1 组成的一串 0-1 序列。

输出:

$\{VM(F^V)\}$ 变值逻辑空间四元符号: $v \in V = \{\perp, +, -, T\}$ 。

在该模型中, 首先将区间取值模块输出的区间范围取值向量以映射的方式转换为 0-1 序列。假设 $IV(R^Q) = R_0 R_1 R_2 \dots R_i \dots R_{n-w+1}$, $0 \leq i \leq n-w+1$, $R_i \in Q$, 对于一个区间范围取值向量的每一位 R_i^q 存在映射[4]如下:

$$R_i^q = \begin{cases} 0, & R_i \neq q; \\ 1, & R_i = q; \end{cases} \quad q \in Q \quad (5)$$

那么一组输入向量可以对应四组 0-1 序列, 分别是对于“ A ”的映射 $IV(R^A)$, 对于“ B ”的映射 $IV(R^B)$, 对于“ C ”的映射 $IV(R^C)$, 对于“ D ”的映射 $IV(R^D)$ 。根据不同的映射策略输出对应的 0-1 序列。根据分段段长 m 将 0-1 序列进行分段处理, 每段段长为 m 。将前后两段序列进行如下公式 6 进行转换, 输出变值逻辑空间的四种变值状态, 分别为: $0 \rightarrow 0, 0 \rightarrow 1, 1 \rightarrow 0$, and $1 \rightarrow 1$ 。变值[5]表示如下:

$$V = \begin{cases} \perp, & X_i = 0, Y_i = 0; \\ +, & X_i = 0, Y_i = 1; \\ -, & X_i = 1, Y_i = 0; \\ T, & X_i = 1, Y_i = 1; \end{cases} \quad X_i, Y_i \in B_2 \quad (6)$$

经过上述变换, 得到的输出结果为: $\{VM(F^V)\}$ 。

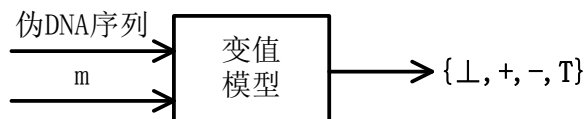


Figure 6. Variant model (VM)

图 6. 变值模型(VM)

2.2.2. 概率测量(PM)

概率测量模型(PM)的主要作用是使用归一化和非归一化方法来计算变值逻辑空间四种状态 $\{VM(F^v)\}$ 的概率测度 $\{PM(P^v)\}$ 。

本文中 N_0 表示 0 的测度, N_1 表示 1 的测度, N_{\perp} 表示 \perp 的测度, N_+ 表示 + 的测度, N_- 表示 - 的测度, N_T 表示 T 的测度。则有: $N = N_{\perp} + N_+ + N_- + N_T = N_0 + N_1$, 其中: $N_0 = N_{\perp} + N_+$, $N_1 = N_- + N_T$ 。

归一化方法如下表示:

$$P^v = N^{v_1}/N_0 \text{ or } N^{v_2}/N_1, v_1 \in \{\perp, +\}, v_2 \in \{-, T\}, v \in V \quad (7)$$

$$P_0 = N_0/N, P_1 = N_1/N \quad (8)$$

非归一化方法表示如下:

$$P^v = N^v/N, v \in V \quad (9)$$

根据以上两种方法输出得到 10 个概率测度, 分别是归一化方法输出的 6 个: $P^{\perp}, P^+, P^-, P^T, P^0, P^1$ 和非归一化方法输出的 4 个 P^{\perp}, P^+, P^-, P^T 。本文从非归一化方法所产生的 4 个中任意选取 2 个作为可视化模块的输入, 从而得到 2D 图谱。

3. 可视化结果

本文主要是比较正常和异常心电在不同的子段段长输入条件下的可视化分布特征, 所以, 文中使用的滑动窗口值 m 为 30, 稳态区间值 s 为 0.95。这两个输入值的选取能够使得可视化结果具有明显的分布特征。

图7是健康人的心电信号图示信息, 图8是室性失常病人的心电信号图示信息。

非归一化方法: 选择机制为 $\{P_{\perp}, P_T\}$ 。

归一化方法: 选择机制为 $\{P_{\perp}, P_T\}$ 。

从图 9 中可以观察到健康人的心电序列处理得到的 2D 图谱中的每一个子图与室性失常病人的心电序列处理得到的 2D 图谱中对应的子图具有明显的分布特征。其中健康人的心电该模型方法下所呈现的分布是: 主要集中于 2 个区域, 呈现四边形分布; 而室性失常病人的分布呈现三角形分布。

从图 10 中同理也能观察到健康人和室性失常病人的 2D 图谱中的对应分段段长的子图分布具有明显的差异。健康人的分布呈现的是复杂性的分布, 而室性失常病人所呈现的分布是比较混沌, 几乎布满整个平面。

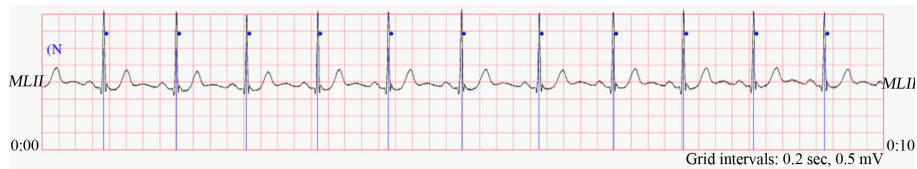


Figure 7. Healthy people

图 7. 健康人

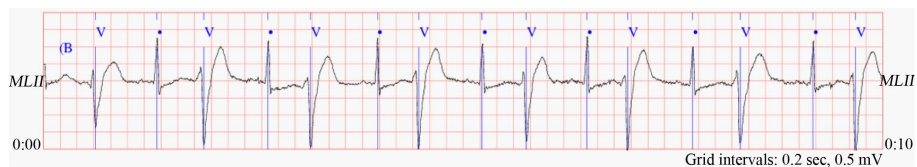


Figure 8. Ventricular arrhythmia

图 8. 室性失常病人

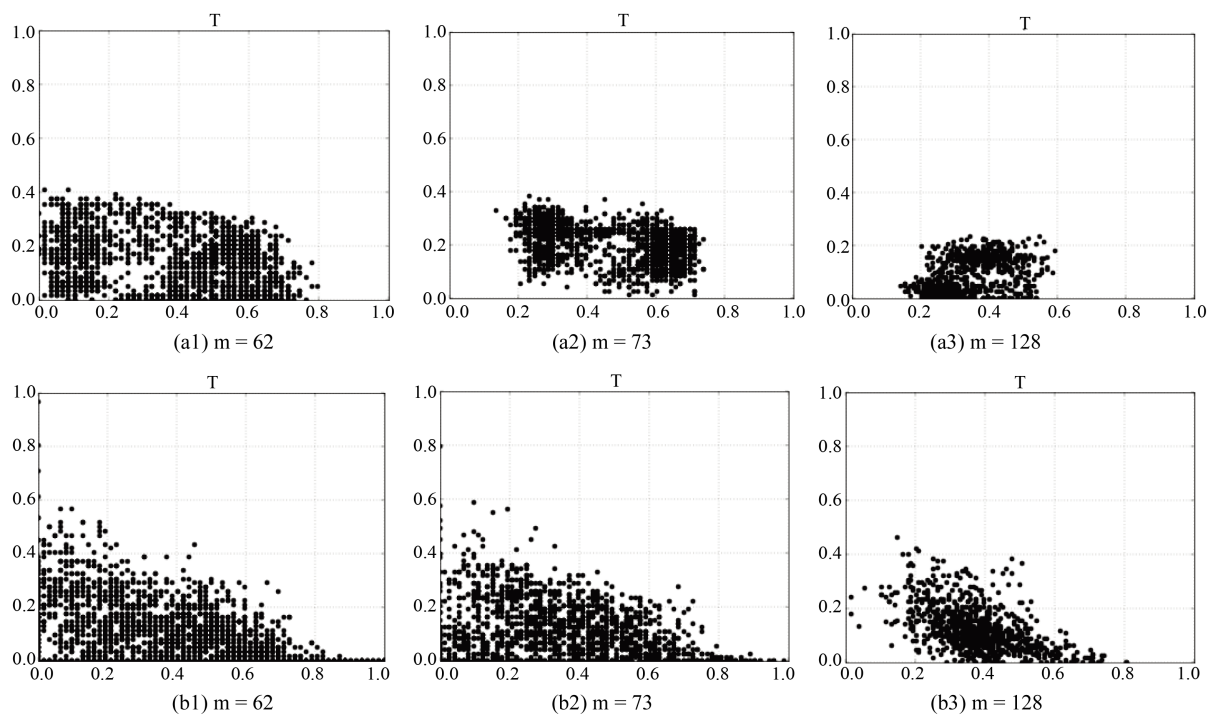


Figure 9. Non-normalized Model: (a1)-(a3) ECG of the Healthy People, Maps of component T; (b1)-(b3) ECG of the Ventricular Arrhythmia, Maps of component T

图 9. 非归一化模型：(a1)~(a3)是健康人的心电，针对分量 T 进行映射的图谱；(b1)~(b3)是室性失常病人的心电，针对分量 T 进行映射的图谱

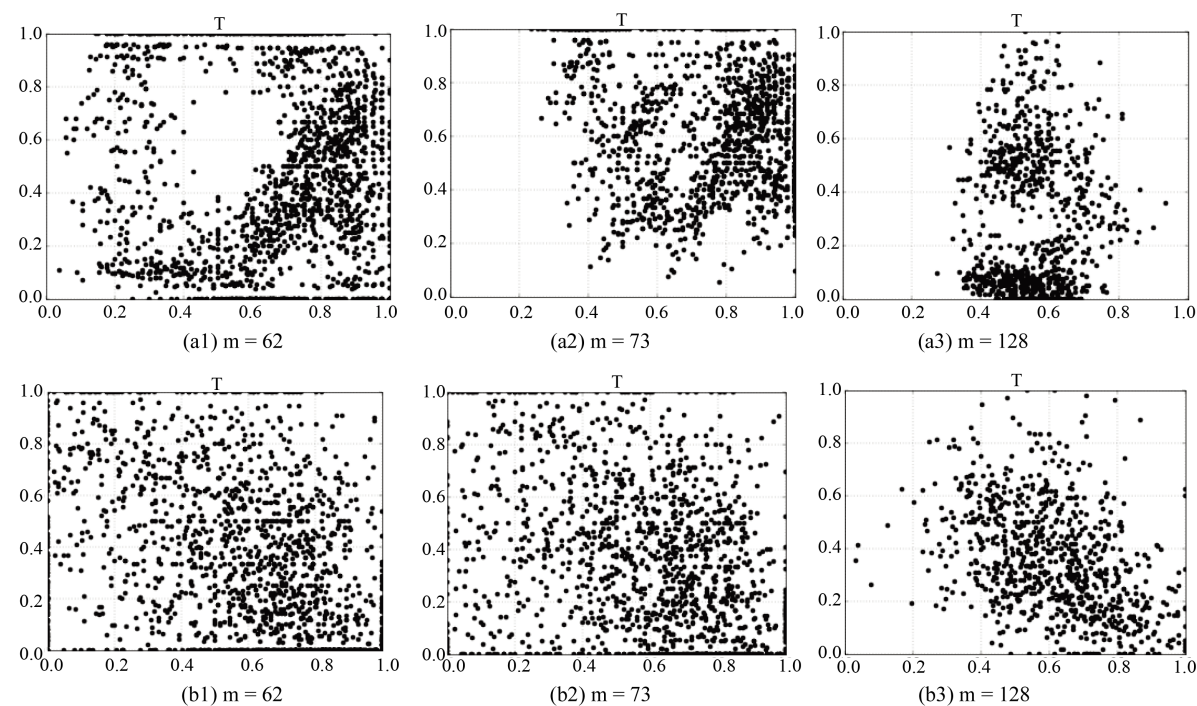


Figure 10. Normalized Model: (a1)-(a3) ECG of the Healthy People, Maps of component T; (b1)-(b3) ECG of the Ventricular Arrhythmia, Maps of component T

图 10. 归一化模型：(a1)~(a3)是健康人的心电，针对分量 T 进行映射的图谱；(b1)~(b3)是室性失常病人的心电，针对分量 T 进行映射的图谱

4. 总结

心电信号的分析在现代医学领域，尤其是临床诊断中扮演者一个比较重要的角色，本文的心电数据处理和可视化模型能够快速、方便、简洁、直观地描述出正常心电和异常心电的分布特性。这种基于变值逻辑的数据处理机制及可视化模型易于理解和操作，对于正常和异常的心电信号，在同一分组大小下，都会产生对应四个碱基字母所对应的图谱，从这四个图谱中可以分析正常心电数据和异常心电数据的分布特性。

利用变值可视化技术，目前观察到正常和异常心电数据的图形分布具有可以观察到的显著差异，期待进一步系统处理更多的数据之后，该类方法可能形成处理大容量心电图数据的标准检测模式，对普通心电图、动态心电图等传统技术方法起到补充作用。

致 谢

感谢我导师郑智捷教授对我这篇文章的细心指导；感谢山路医生的帮助；感谢云南大学软件学院、云南省软件工程重点实验室信息安全基金以及郑智捷教授国家自然科学基金的支持。

参考文献 (References)

- [1] 王玉静, 宋立新, 守强. 基于 EMD 和奇异值分解的心律失常分类方法[J]. 信号处理, 2010, 26(9): 1423-1426.
- [2] 李方洁, 向晋涛. 心电散点图[M]. 北京: 人民卫生出版社, 2014: 3-28.
- [3] Jeffery, Z.J.Z. and Christian, H.H.Z. (2006) A Framework of Variant Logic Construction for Cellular Automata. Harbin Institute of Technology Press.
- [4] Bu, Q.X. and Zheng, J.Z.J. (2013) 2D Conjugate Maps of DNA Sequences. *Journal of Information Security*, **4**, 193-196. <http://dx.doi.org/10.4236/jis.2013.44021>
- [5] Jeffrey, Z., Christian, Z. and Tosiya, K. Interactive Maps on Variant Phase Spaces—From Measurements—Micro Ensembles to Ensemble Matrices on Statistical Mechanics of Particle Models. InTech Press.