

# 基于 ARMA 模型的 ECG 分类和压缩

葛丁飞<sup>1</sup>, 李时辉<sup>2</sup>

(1. 浙江科技学院 信息与电气工程学系, 浙江 杭州 310023;

2. 义乌工商职业技术学院 计算机工程系, 浙江 义乌 322000)

**摘 要:** 心电信号(ECG)对医生诊断心脏疾病极为重要。现存许多 ECG 分类技术存在实现困难、处理时间长和只能对 2~3 类 ECG 进行分类的不足。本文介绍了计算简单的 ARMA 模型的 ECG 分类法,利用 ARMA 模型系数作为特征对 ECG 信号进行分类和压缩。在对信号特征分类时,采用了非线性二次判别函数的形式。利用文中方法对 MIT-BIH 标准数据库中 NSR、APC、PVC、SVT、VT 和 VF 各 200 个样本信号进行测试,获得了 94.28%~99.28% 的分类精度。

**关键词:** ECG 信号; ARMA 建模; 二次判别函数; 特征提取

**中图分类号:** R540.41; TN911.7

**文献标识码:** A

**文章编号:** 1671-8798(2004)01-0007-07

心电信号(ECG)是医生进行诊断的重要依据,关于 ECG 的分类已有许多现有技术,但现存许多 ECG 分类技术存在实现困难、处理时间长和只能对 2~3 类 ECG 进行分类的不足,如相关波形法(CWA)<sup>[1]</sup>、直接 ECG 特征检测法<sup>[2]</sup>、傅立叶变换法<sup>[3]</sup>、功率谱分析法<sup>[4]</sup>、复杂度测量法<sup>[5]</sup>、TLS (Total Least Squares) Prony 建模法<sup>[6]</sup>、自适应卡尔曼滤波器法<sup>[7]</sup>、使用 LPC 系数和模糊 ARTMAP 神经网络分类法等<sup>[8]</sup>。

本文介绍了基于 ARMA 模型的 ECG 分类方法,ARMA 系数被作为特征来进行 ECG 的分类,在分类时,采用了二次判别函数分类法<sup>[9]</sup>。ARMA 建模技术有着容易实现、处理快速和同时能对多类 ECG 进行分类的优点。本文利用 ARMA 建模技术对 MIT-BIH 标准数据库中的以下几类 ECG 进行了分类测试:正常窦性心律(NSR)、心房早期收缩(APC)、心室早期收缩(PVC)、室上性心动过速(SVT)、心室性心动过速(VT)和心室纤维性颤动(VF)。结果表明,该方法可获得 94.28%~99.28% 的分类精度。

ECG 的压缩方法对计算机辅助远程自动诊断的快速性有着直接的影响。ECG 的压缩方法可分为三类<sup>[10]</sup>:①直接方法,该方法计算简单而且效率很高,但计算机辅助远程自动诊断必须在信号重建、特征提取以后才能进行,这便造成诊断工作的延误;②变换方法,该方法计算复杂,特别是在单导联情况下压缩率很低;③参数方法,该方法用信号特征进行 ECG 数据压缩,信号的特征可同时被用来信号重建和分类。因此,基于参数方法的信号分类和重建是个并行的过程,可以极大地提前 ECG 诊断的时间。基于 ARMA 建模技术的 ECG 压缩是属于参数压缩的方法,该方法可获得 7.8 倍的压缩率。可见,ARMA 模型非常适于远程心电图诊断中的 ECG 压缩和分类。

## 1 方 法

### 1.1 ECG 信号的滤波和分割

本文中所使用的 ECG 信号取自 MIT-BIH 数据库;NSR 信号、APC 信号和 PVC 信号取自“MIT-BIH

收稿日期: 2003-07-20

作者简介: 葛丁飞(1965—),男,浙江东阳人,工程师,主要从事生物信号的研究与教学工作。

arrhythmia database”,其采样频率是 360 Hz;VT 信号和 VF 信号取自“MIT-BIH Ventricular Arrhythmia database”,其采样频率是 250 Hz;SVT 信号取自“MIT-BIH supraventricular arrhythmia database”,其采样频率是 128 Hz。在建模时,所有 ECG 信号的频率都转化为 250 Hz。在建模以前,先对 ECG 信号进行滤波以消除由电源、呼吸等原因引起的噪音。为了满足医生和护士监视的需要,ECG 信号的带宽应保持在 1 Hz 到 50 Hz 之间<sup>[11]</sup>。用一个带通滤波器(BPF)对 ECG 信号进行滤波<sup>[11]</sup>,其上下边带截止频率是 1 Hz 和 50 Hz。这能有效地消除电源引起的噪音(60 Hz)、呼吸引起的噪音(0.2 Hz 左右)、电极移动引起的低频噪音以及肌肉收缩引起的高频噪音。

不同种类的 ECG 有着不同的心跳率和 RR 间期。本文使用通常情况下的 NSR 信号,其心跳率是 60~100 次/min。APC 的 RR 间期比 NSR 的 RR 间期短,而且紧跟其后的 RR 间期不会比 NSR 的 RR 间期长,VT 和 VF 的 RR 间期比 NSR 的 RR 间期更要短得多。可见,300 个采样点的数据窗口能足以包含一个周期的 ECG 信息。因此,本文采用 300 个采样点的数据窗口对 ECG 进行分割,其中,100 个采样点在 ECG 峰值  $R$  以前,200 个采样点在 ECG 峰值  $R$  以后。

## 1.2 ARMA 模型的建立

ARMA 模型是一个线性输出模型,其方程式是<sup>[11]</sup>:

$$HR(k) - \sum_{i=1}^p a_i HR(k-i) = \sum_{j=1}^q b_j R(k-j) + n(k) \quad (1)$$

方程(1)可分解为 AR 过程和 MA 过程:

$$\text{AR 过程:} \quad HR(k) - \sum_{i=1}^p a_i HR(k-i) = R(k) \quad (2)$$

$$\text{MA 过程:} \quad R(k) = \sum_{j=1}^q b_j R(k-j) + n(k) \quad (3)$$

其中, $HR(k)$  为 ECG 信号的时间序列; $a_i$  和  $b_j$  分别为 AR 系数和 MA 系数; $P$  和  $Q$  分别为 AR 阶次和 MA 阶次; $R[k]$  为 AR 误差的序列; $n(k)$  为 ARMA 预测误差,即 ARMA 建模误差。本文利用 Burg 算法分别对 ARMA 模型的 AR 过程和 MA 过程进行参数估计。在建模中,阶次  $P$  和  $Q$  的选择至关重要。这里引用 MDL (Minimizes the Description Length) 准则来确定模型阶次。

MDL 准则<sup>[11,12]</sup>:

$$\text{MDL} = (N - N_p) \ln \sigma_n^2 + p \ln(N - N_p) \quad (4)$$

式中, $\sigma_n^2$  为建模误差的方差, $N$  为数据窗口的宽度, $N_p$  为模型的阶次,在这里  $N_p = P$  (AR 过程) 或  $N_p = Q$  (MA 过程)。

MDL 的敏感性函数  $S_{N_p}^{\text{MDL}}$ :

$$S_{N_p}^{\text{MDL}} = \left| \frac{\Delta \text{MDL} / \text{MDL}}{\Delta P / N_p} \right| \quad (5)$$

其中, $\Delta \text{MDL}$  是 MDL 的变化量, $S_{N_p}^{\text{MDL}}$  反映了模型阶次  $N_p$  对 MDL 的影响。

## 1.3 ECG 特征的提取和压缩

由于采用 ARMA 系数作为 ECG 的特征,所以本文利用了特征向量  $\mathbf{x} = [a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_q]$  作为 ECG 信号的特征。关于 ECG 的压缩,笔者参考了 Ruttimann 和 Pipberger 的改进霍夫曼编码法对 ECG 进行压缩的结果<sup>[13]</sup>,其编码对象是 AR 系数和建模误差,得到的压缩率为 7.8 倍。尽管建模误差和原始 ECG 信号有着相同的采样点数,但是建模误差动态范围比原始 ECG 动态范围更小,所以表示建模误差所需的位数 (Bit) 比表示原始 ECG 所需的位数更少,从而可获得较高的压缩率。压缩率的计算公式为:

$$\text{压缩率} = \frac{\text{代表 ECG 原始值所需的位数}}{\text{压缩后代表 ECG 所需的位数}} \quad (6)$$

## 1.4 基于非线性二次判别函数的分类算法及性能评测

根据特征向量  $\mathbf{x}$ ,利用树状决策过程并基于二次判别函数对 ECG 信号进行分类。二次判别函数的方程

为<sup>[9]</sup>:

$$y_i = w_0 + \sum_{m=1}^d w_m x_m + 2 \sum_{m=1}^{d-1} \sum_{n=m+1}^d w_{mn} x_m x_n + \epsilon_i \quad (7)$$

其矩阵形式为:

$$y_i = X_i W + \epsilon_i \quad (8)$$

其中,  $x_1, x_2, \dots, x_d$  为 ECG 特征值;  $y_i$  为某一 ECG 特征向量的目标值, 即成员关系;  $\epsilon_i$  为二次判别函数误差;  $X_i$  为  $1 \times [d(d+3)/2 + 1]$  的行矩阵,  $W$  为  $[d(d+3)/2 + 1] \times 1$  的列矩阵,  $X_i$  和  $W$  分别为:

$$X_i = [1, x_1, x_2, \dots, x_d, x_1^2, x_2^2, \dots, x_d^2, 2x_1x_2, 2x_1x_3, \dots, 2x_1x_d, 2x_2x_3, 2x_2x_4, \dots, 2x_2x_d, \dots, 2x_{d-1}x_d]$$

$$W = [w_0, w_1, w_2, \dots, w_d, w_{11}, w_{22}, \dots, w_{dd}, w_{12}, w_{13}, \dots, w_{1d}, w_{23}, w_{24}, \dots, w_{2d}, \dots, w_{d-1}w_d]^T$$

将各类 ECG 的目标值  $y_i$  分别定义为某一整数, 如 1、-1 等, 假设所有 ECG 样本数为  $D$ , 则可得到如下方程:

$$Y = AM + E \quad (9)$$

其中,  $Y = [y_1, y_2, \dots, y_D]^T$  为相应于  $D$  个样本的目标向量;  $E = [\epsilon_1, \epsilon_2, \dots, \epsilon_D]^T$  为相应于  $D$  个样本的二次判别函数误差向量;  $A = [X_1, X_2, \dots, X_D]^T$  为  $D \times [d(d+3)/2 + 1]$  矩阵。

式(9) 基于最小二乘平方误差准则的伪逆解为:

$$W = (A^T A)^{-1} A^T Y \quad (10)$$

所以分类器的判别函数为:

$$\hat{y}_i = X_i W \quad (11)$$

在分类过程中, 其分类算法如表 1 所示。首先, 对各类 ECG 进行分组, 依据各类 ECG 特征值之间的中心距进行分组, 中心距小的被分为同一组; 其次, 在每一步中, 分别定义各组的成员关系, 即各组  $y$  对应于 1, -1 等。笔者在各类 ECG 样本集中随机选取一定数量的样本作为学习数据, 把其余的作为测试数据。在分类算法的每一步中, 利用学习数据按式(10) 计算出  $W$ , 并利用式(11) 计算出试验数据判别函数值  $\hat{y}_i$ , 并依据此值进行分类。假设第一步的  $\hat{y}_i$  用  $\hat{y}_1$  分类; 第二步的  $\hat{y}_i$  用  $\hat{y}_2$ , 依此类推。

表 1 分类算法

第一步			第二步			第三步		
组别	成员关系	判断准则	组别	成员关系	判断准则	组别	成员关系	判断准则
SVT	2	$\hat{y}_1 > 1.5$						
NSR/APC /PVC	1	$0 < \hat{y}_1 < 1.5$	APC	1	$\hat{y}_2 > 0$	APC	1	$\hat{y}_4 > 0$
						NSR	-1	$\hat{y}_4 < 0$
			PVC	-1	$\hat{y}_2 < 0$	PVC	1	$\hat{y}_5 > 0$
VT/VF	-1	$\hat{y}_1 < 0$				NSR	-1	$\hat{y}_5 < 0$
			VT	1	$\hat{y}_3 > 0$			
			VF	-1	$\hat{y}_3 < 0$			

为进一步评价文中算法, 本文利用了敏感性和特异性来衡量分类结果的性能。其计算公式分别为:

$$\text{敏感性} = \frac{TE - FN}{TE} \quad (12)$$

$$\text{特异性} = \frac{TE - FP}{TE} \quad (13)$$

其中,  $TE$  为各类 ECG 样本总数;  $FN$  为各类 ECG 假阴性总数;  $FP$  为各类 ECG 假阳性总数。

## 2 结 果

6 类不同 ECG 经过滤波处理后分别用于 ARMA 建模, 利用 AR 系数和 MA 系数对各类 ECG 进行分类。在分类时, 进行了基于树状决策过程和二次判别函数的分类。

## 2.1 AR 建模和压缩的结果

如上所述,利用 MDL 和 MDL 敏感性函数作为选择 ARMA 模型阶次的准则。在图 1 中给出了在 AR 过程中 MDL 与模型阶次  $P$  的关系。由此可见,MDL 在开始时随着模型阶数的增加而减少,但是当  $P \geq 4$  时,MDL 基本保持不变。在图 2 中给出了在 AR 过程中敏感性函数  $S_P^{\text{MDL}}$  与模型阶次  $P$  的关系。从中可见,当  $P \geq 4$  时, $P$  对 MDL 的影响很小。因而选择 AR 模型阶次  $P=4$ ,这一结果与其他的 research 结果是一致的。有关研究认为,当  $P \geq 3$  时,AR 误差不会随  $P$  的增加而减少<sup>[14,15]</sup>。类似地,图 3 给出了 MA 过程中 MDL 与模型阶次  $Q$  的关系( $P=4$ )。据此,选择 MA 模型阶次  $Q=2$ 。各特征值的平均值如表 2 所示。

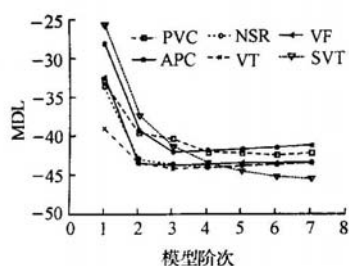


图 1 AR 过程中 MDL 与模型阶次  $P$  的关系

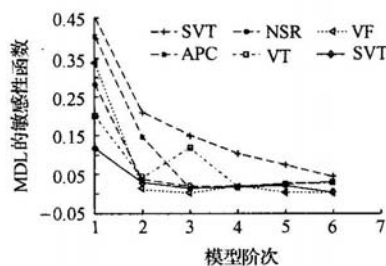


图 2 AR 过程中 MDL 敏感性函数  $S_P^{\text{MDL}}$  与模型阶次  $P$  的关系

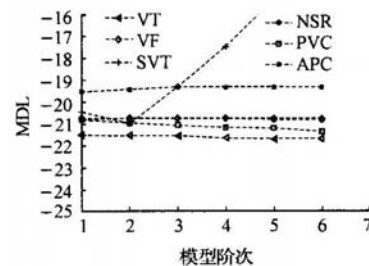


图 3 MA 过程中 MDL 与模型阶次  $Q$  的关系( $P=4$ )

表 2 各特征值的平均值

信号种类	$a_1$	$a_2$	$a_3$	$a_4$	$b_1$	$b_2$
SVT	-2.767 8	3.186 8	-1.859 9	0.474 5	0.969 0	0.342 5
NSR	-2.182 9	1.475 6	-0.102 7	-0.168 4	-0.004 3	0.005 9
APC	-2.338 4	1.839 2	-0.381 7	-0.097 0	0.030 5	0.067 9
PVC	-2.270 6	2.163 7	-1.344 0	0.466 5	0.075 8	-0.041 1
VT	-1.427 8	0.079 9	0.374 6	-0.018 1	-0.055 7	0.152 6
VF	-1.728 2	0.409 9	0.503 7	-0.170 3	-0.096 9	0.093 8

直接参考 Ruttimann 和 Pipberger 对 ECG 进行压缩的结果,可得到压缩率为 7.8 倍<sup>[11]</sup>。AR 模型和 AR-MA 模型建模误差方差如表 3 所示。可见,AR 模型和 ARMA 模型建模误差方差几乎是相等的,所以基于 ARMA 模型的 ECG 压缩率和 AR 模型的 ECG 压缩率可认为是一样的。

表 3 AR 模型和 ARMA 模型建模误差的方差

模型	建模误差的方差					
	SVT	NSR	APC	PVC	VF	VT
AR	3.057e-5	7.112e-4	9.54e-4	1.986e-3	6.7763e-4	3.924e-3
ARMA	2.954e-5	7.1053e-4	9.53e-4	1.95e-3	6.7563e-4	3.88e-3

## 2.2 ECG 分类的结果

在分类时,利用树状决策过程和二次判别函数对各类 ECG 进行分类。分类过程中的数据分组、成员关系和决策规则如分类算法表 1 所示。表 4 是各类 ECG 特征值之间的欧几里得中心距。由表 4 可知,VT/VF、NSR/APC/PVC 和 SVT 分别聚集在一起,而 VT/VF、NSR/APC/PVC 和 SVT 互相离得较远。因此,按如下步骤进行分类:第一步,把 SVT 单独分为一组,把 NSR/APC/PVC 和 VT/VF 各分为一组,并分别定义其成员关系为“2”、“1”和“-1”,且依据  $\hat{y}_1$  的值进行分类,SVT 首先被分离出来。第二步,把 APC 和 PVC 各分为一组,并分别定义其成员关系为“1”和“-1”,且依据  $\hat{y}_2$  的值进行分类;把 VT 和 VF 各分为一组,并分别

定义其成员关系为“1”和“-1”,且依据  $\tilde{y}_3$  的值进行分类,VT 和 VF 即被分离出来。同理在第三步中,APC、PVC 和 NSR 分别被分离出来。

表 4 各类 ECG 特征值之间的欧几里得中心距

信号种类	SVT	NSR	APC	PVC	VF	VT
SVT	0	2.781 2	2.323 6	1.556 9	3.976 6	4.187 8
NSR	2.781 2	0	0.494 3	1.560 0	1.314 0	1.671 1
APC	2.323 6	0.494 3	0	1.169 4	1.794 8	2.125 4
PVC	1.556 9	1.560 0	1.169 4	0	2.690 3	2.880 2
VF	3.976 6	1.314 0	1.794 8	2.690 3	0	0.494 1
VT	4.187 8	1.671 1	2.125 4	2.880 2	0.494 1	0

在学习和测试过程中,各类 ECG 学习样本和测试样本分别都是 60 个和 140 个。在分类的每一步中,利用学习数据按式(10)计算出  $W$ ,并利用式(11)计算出试验数据判别函数值  $\tilde{y}_i$ ,并依据此值进行分类,测验分类结果如表 5 所示,敏感性和特异性的平均值如表 6 所示。

表 5 测试分类结果

试验数据	分类结果						
	信号种类	SVT	NSR	APC	PVC	VF	VT
140	SVT	139	0		1	0	0
140	NSR	0	137	1	2	0	0
140	APC	1	4	135	0	0	0
140	PVC	2	2	0	135	0	1
140	VF	0	0	0	1	137	2
140	VT	0	0	0	2	6	132

表 6 敏感性和特异性的平均值

%

信号种类	SVT	NSR	APC	PVC	VF	VT
特异性	97.85	95.71	99.28	95.71	95.71	97.86
敏感性	99.28	97.86	96.42	96.42	97.86	94.28

3 讨 论

由上可知  $P=4, Q=2$  是合适的选择。如果  $P、Q$  太高,不仅会大大地增加运算量,而且 ARMA 预测精度得不到显著地提高。

基于 ARMA 模型的 ECG 特征生成有着计算简单和容易实现的优点。此外,在计算 ARMA 系数时所需的数据窗口是 300 个采样点,即 1.2 s,而基于复杂度测量法所需的数据窗口是 3~7 s<sup>[7]</sup>,TLSProny 建模技术所需的数据窗口是 5~9 s<sup>[8]</sup>。

利用 ARMA 技术对 6 类不同 ECG 进行分类,获得了 94.28%~99.28% 的分类精度。因此,ARMA 技术克服了许多方法只能对 2 至 3 类 ECG 进行分类的缺点。比如:模糊 ARTMAP 分类法对 NSR 和 PVC 进行分类<sup>[8]</sup>;傅立叶变换神经网络法对 SVT 和 PVC/VT 进行分类<sup>[3]</sup>;复杂度测量技术适用于 VT、VF 和 NSR 分类<sup>[5]</sup>;以 QRS 为特征,采用神经网络来分类的方法,虽然可对 APC 和 PVC 等 ECG 进行分类,但却不能同时对 VT 和 VF 进行分类<sup>[16]</sup>;Prony 建模技术可对 SVT、VT 和 VF 进行分类,但不能同时包括对 NSR、APC 和 PVC 的分类<sup>[6]</sup>。由此可见,利用 ARMA 技术进行分类诊断是非常高效的。

把 ARMA 建模技术分类结果与最近发表的关于 ECG 分类结果进行了如下比较:用 LPC 系数和 QRS 波的 MSV(Mean Square Value)作为特征值,采用模糊 ARTMAP 分类法对 NSR 和 PVC 进行分类,获得敏感性为 97%、特异性为 99%<sup>[8]</sup>;利用改进顺序概率比测试(SPRT)算法对 VT 和 VF 进行分类,分别获得的分类精度是 93%和 96%<sup>[17]</sup>;利用人工智能神经网络(ANN)对威胁性 ECG 进行分类,可获得 93%~99% 的分类精度<sup>[18]</sup>,但 ANN 在学习过程中需要结构非常特殊的数据,在学习数据中,NSR 和 VC 的样本数量特别



大,而且 APC 的分类精度也不高;利用 TLSProny 建模法,SVT、VF 和 VT 可分别获得 95%、96%和 97%的分类精度<sup>[6]</sup>。根据笔者的研究,基于 ARMA 模型的分类可取得比基于 AR 模型的分类较好的结果,这是由于 ARMA 模型融合了建模误差信息的原因。

另外,采用 ARMA 技术对 ECG 数据进行压缩可得很高的压缩率,ARMA 模型特别适用于远程心电诊断。在远程心电诊断中,计算机辅助 ECG 诊断通常是在接收到 ECG 信号、进而解压重建后进行的,这样便造成诊断工作的延误。基于 ARMA 模型的 ECG 直接分类的方法,它无需等 ECG 解压重建后再行诊断,从而极大地提高诊断的快速性。

本文使用了固定长度的数据窗口,如果使用变化的数据窗口也许可以得到更好的分类效果(随 RR 间期而变)。ARMA 模型是线性模型,ECG 信号严格地讲也不是平稳随机过程,因此利用非线性参数模型也许更适合 ECG 信号,如双线性 BL 模型(Bilinear Model)等。有关研究表明,ECG 信号和血液动力学信号、呼吸信号等有很强的相关性,可考虑混合这些信号来进行进一步研究<sup>[18]</sup>。笔者正在进行把非平稳 ECG 经过处理后得到一个平稳时序的研究,以及通过多导联 ECG 提取特征的研究。

#### 4 结 论

基于 ARMA 模型和二次判别函数分类法应用于危急病人 ECG 分类诊断是有效的。ARMA 模型特别适合于远程心电诊断中的 ECG 压缩和分类,并极具实时应用意义。

#### 参考文献:

- [1] Caswell S A, Kluge K S, Chiang C M J. Pattern recognition of cardiac arrhythmias using two intracardiac channels[R]. Proc Comp Cardiol, 1993. 181—184.
- [2] Zhou S H, Rautaharju P M, Calhoun H P. Selection of a reduced set of parameters for classification of ventricular conduction defects by cluster analysis[R]. Proc Comp Cardiol, 1993. 879—882.
- [3] Minami K C, Nakajima H, Toyoshima T. Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network[J]. IEEE Trans Biomed Eng, 1999, 46:179—185.
- [4] Afonso V X, Tompkins W J. Detecting ventricular fibrillation: Selecting the appropriate time-frequency analysis tool for the application[J]. IEEE Eng Med Biol Mag, 1995, 14:152—159.
- [5] Zhang X S, Zhu Y S, Thakor N V, *et al.* Detecting ventricular tachycardia and fibrillation by complexity measure[J]. IEEE Trans Biomed Eng, 1999, 46:548—555.
- [6] Chen S W. Two-stage discrimination of cardiac arrhythmias using a total least squares-based prony modeling algorithm [J]. IEEE Trans Biomed Eng, 2000, 47:1317—1326.
- [7] Arnold M, Miltner W H R, Witte H. Adaptive AR modeling of nonstationary time series by means of Kalman filtering [J]. IEEE Trans Biomed Eng, 1998, 45:553—562.
- [8] Ham F M, Han S. Classification of cardiac arrhythmias using fuzzy ARTMAP[J]. IEEE Trans Biomed Eng, 1996, 43: 425—430.
- [9] Duda R O, Hart P E. Pattern Classification[M]. New York: Wiley-Interscience Publication, 2001.
- [10] Jalaleddine S M S, Hutchens C G, Strattan R D, *et al.* ECG data compression techniques-a unified approach[J]. IEEE Trans on BME, 1990, 37:329—343.
- [11] Proakis J G, Manolakis D G V. Digital Signal Processing[M]. Prentice-Hall, United States of America, 1996.
- [12] Giraldo B F, Bolea Y, Caminal P. Comparision of diffetent models to analyze the number of patients in waiting-list [R]. Proceeding of the Second Joint EMBS/MBES Conference. Houston USA, 2002, 23—26.
- [13] Tompkins W. Biomedical Digital Signal Processing[M]. New Jersey: Prentice-Hall, Englewood Cliffs, 1993.
- [14] Lin K P, Chang W H. QRS feature extraction using linear prediction[J]. IEEE Trans Biomed Eng, 1989, 36:1050—1055.
- [15] Ge D F, Srinivasan N, Krishnan S M. Cardiac arrhythmia classification using autoregressive modeling[J]. Biomedical Engineering Online, 2002, (1):5.

- [16] Melo S L, Caloba L P, Nadal J. Arrhythmia analysis using artificial neural network and decimated electrocardiographic data[J]. *Comp Cardiol*, 2000, 27:73—76.
- [17] Chen S W, Clarkson P M, Fan Q. A robust sequential detection algorithm for cardiac arrhythmia classification[J]. *IEEE Trans Biomed Eng*, 1996, 43:1120—1121.
- [18] Korhonen I. Multivariate closed-model for analysis of cardiovascular dynamics [J]. *Methods of Information in Medicine*, 1997, 36:264—267.

## Cardiac arrhythmias classification and compression based on ARMA model

GE Ding-fei<sup>1</sup>, LI Shi-hui<sup>2</sup>

(1. Dept. of Information and Electrical Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China; 2. Dept. of Computer Engineering, Yiwu Industrial and Commercial College, Yiwu 322000, China)

**Abstract:** Electrocardiogram (ECG) signal is important for physician to diagnose cardiac diseases. Various existing techniques on ECG classification have been reported. Generally, these techniques classify two or three arrhythmias only or have significantly large processing times. A simpler autoregressive — moving average (ARMA) based technique is proposed to classify ECG for diagnosis in this paper. The ARMA coefficients have been used to classify and compress ECG directly. ARMA technique has been used for classification into arrhythmias such as normal sinus rhythm (NSR), premature ventricular contraction (PVC), atrial premature contraction (APC), ventricular tachycardia (VT), ventricular fibrillation (VF), and supraventricular tachycardia (SVT). Two hundred data samples from each of them have been utilized to classify and test. The accuracy of detecting these arrhythmias proposed are 93.5% to 97.86% using the quadratic discrimination function stage-by-stage.

**Key words:** ECG signal; ARMA modeling; quadratic discrimination function; feature extraction