

基于蚁群算法的带截止区均匀量化器的优化及其在 ECG 数据压缩中的应用

王伟平¹ 杨 苗²

(昆明理工大学城市学院 昆明 650093)¹ (云南省产品质量监督检验研究院 昆明 650223)²

摘 要 结合小波系数的特点,采用了改进的均匀量化器——带截止区的均匀量化器(USDZQ)对变换后的小波系数进行量化。量化器的参数选取直接影响到 ECG 数据压缩的质量和压缩比,因此重点研究了 USDZQ 的参数优化问题,选取了蚁群优化算法(ACO)作为 USDZQ 参数的优化工具。最后,利用本文算法对 MIT-BIH 心律失常数据库的 ECG 信号进行了编码测试。实验结果表明,只要对 USDZQ 的参数进行合理优化,USDZQ 就能获得优于均匀量化的性能,并可以成功地应用于 ECG 数据压缩中。

关键词 蚁群优化算法,带截止区的均匀量化器,ECG 压缩

中图分类号 TP391 文献标识码 J

USDZQ Optimization Based on Ant Colony Algorithm and Application in ECG Compression

WANG Wei-ping¹ YANG Miao²

(College of City, Kunming University of Science and Technology, Kunming 650093, China)¹

(Yunnan Testing Institute of Product Quality Supervision, Kunming 650223, China)²

Abstract This paper adopted an improved uniform quantizer, a uniform scalar dead zone quantizer(USDZQ), to quantize the transformed coefficients. The selection of quantizer parameters directly affects the ECG data compression performance. Therefore the objective of this study was focused on the optimization of the USDZQ parameters. We used the ant colony optimization(ACO) algorithm for the optimization. Experiments on several records from the MIT-BIH arrhythmia database show that as long as the USDZQ parameter is optimized reasonably, USDZQ can achieve better performance than the uniform quantizer, and may successfully be applied in the ECG data compression.

Keywords ACO, USDZQ, ECG data compression

ECG 数据常常需要进行压缩以便于存储和传输,因此 ECG 数据压缩是当今医学信号处理领域的重要问题。而量化是数据压缩系统中的重要组成部分,采用量化技术能够有目的地丢弃一些不重要的信息,较好地提高编码器的性能,进行有效的数据压缩。基于小波变换的 ECG 数据压缩的 3 个主要环节是:小波变换、量化和熵编码。对于有失真压缩的关键步骤——量化器的设计,本文选取了蚁群优化算法(ACO)作为量化器参数优化的工具。

ACO 是一种性能优良的启发式随机算法,具有较强的鲁棒性。该算法通过信息素的不断累积和更新最终达到收敛于最优路径上^[17,18]。这种算法首先由意大利科学家 M. Dorigo 等人提出,称为蚂蚁系统(ant system, AS),并成功用于求解 TSP 问题,引起了人们的关注。这种算法在求解二次分配^[1]、Job-Shop 问题^[2]、图着色问题^[3]、车辆调度问题^[4,16]、集成电路设计^[5]以及通信网络负载问题的处理中^[6]都取得了较好的结果。越来越多的研究人员正在从事这方面的工作,并用该算法来解决一些实际问题。本文把 ACO 合理地应用于心电信号的处理上,通过多次实验取得了理想的结果。

1 蚁群优化算法

1.1 算法基本步骤

蚁群算法是一个递推过程,很容易在计算机上实现,其主要步骤如下。

步骤 1: $nc=0$ (nc 为迭代步数或搜索次数)。将 τ_{ij} 和 $\Delta\tau_{ij}$ 初始化;将 m 个蚂蚁置于 n 个顶点上;

步骤 2: 将各蚂蚁的初始出发点置于当前解集中。对每个蚂蚁 $k(k=1, \dots, m)$, 按概率 P_{ij}^k 移至下一顶点 j ; 将顶点 j 置于当前解集;

步骤 3: 计算各蚂蚁的目标函数值 $Z_k(k=1, \dots, m)$, 记录当前的最好解;

步骤 4: 按更新方程修改轨迹强度;

步骤 5: 对各边弧 (i, j) , 置 $\Delta\tau_{ij}=0, nc=nc+1$;

步骤 6: 若 $nc <$ 预定的迭代次数, 则转到步骤 2; 否则输出结果。

算法中的参数设定目前尚无理论上的依据, 参数 $Q, C, \alpha, \beta, \rho$ 可以用实验确定其最优组合。经验结果为: (1) $1 \leq \alpha \leq 5$;

本文受云南省教育厅面上项目:基于 RFID 与 GPRS 技术的灾难救援系统开发研究(KKJA201025051),云南省软件工程重点实验室开放基金资助项目面上基金项目:嵌入式系统软件的演化研究(2012SE308),国家自然科学基金:提高锡、钨、金等金属回收率的摇床分带图像分割法研究(51204077)资助。

王伟平(1977-),男,硕士,讲师,主要研究方向为信息处理;杨 苗(1979-),女,硕士,工程师,主要研究方向为信息处理(通信作者)。

(2) $1 \leq \beta \leq 5$; (3) $0.5 \leq \rho \leq 0.99$, ρ 取 0.7 左右为佳; (4) $1 \leq Q \leq 10000$ 。

1.2 连续空间优化问题的 ACO

本文所要解决的问题属于一个连续空间的函数寻优问题。目前,蚁群优化算法在求解离散优化问题方面已经有了比较深入的研究,并取得了很多成果,但是在求解连续优化问题方面的研究就相对要弱一些^[7,8]。在连续空间的寻优问题求解中,解空间是以区域性方式表示,而不是以离散点集方式表示。要想构造一个普适型的连续蚁群算法十分困难^[9],但是,我们可以抽象出其中最基本的几个步骤来构造一个基本的算法框架:

(1) 初始化

1) 将适应度函数(目标函数)显式表示为: $F = f(X)$;

2) 循环计数器赋初值: $Count = 0$;

3) 随机或按某种规律将蚂蚁散布在定义域上(即初始化 $X_k, k = 1, 2, \dots, M$);

4) 残留信息初始化,即在整体定义域上均匀分布初始残留信息 $\tau(X) = C$ 。

(2) 蚂蚁群体移动

1) 对蚂蚁 $k(k = 1, 2, \dots, M)$, 计算其选择概率 $p(X_k)$;

2) 根据选择概率和启发函数,移动蚂蚁 k 到下一位置;

3) 计算新位置所对应的目标函数值。

(3) 残留信息修改

1) 计算新位置的残留信息增量 $\Delta\tau(X)$;

2) 计算并记录位置 X 对应的残留信息 $\tau(X) = \rho \times \tau(X) + \Delta\tau(X)$ 。

(4) 收敛条件检测

1) $Count = Count + 1$;

2) 若 $Count < Count_{max}$, 返回(2)继续, 否则, 算法停止。

在应用过程中,可根据具体应用问题设计出具体接口函数,从而使算法具有相当宽广的普适性。

2 基于蚁群优化算法的带截止区均匀量化器的优化

本文在用连续空间蚁群算法对带截止区均匀量化器进行参数优化时,将门限值 T 和量化步长 Δ 这两个参数作为待优化的变量,所以优化问题是一个二维问题。

2.1 适应度函数的设计

适应度函数的设计是算法的关键,它的好坏直接决定了优化效果,它必须根据所要优化的问题具体设计。本文在参考了文献^[10,11]的基础上,根据 ECG 数据压缩的特点以及本文中所使用的信号编码模型,并结合拉格朗日函数从而得出适应度函数。其中使用的编码模型如图 1 所示。

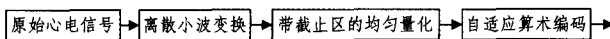


图 1 ECG 信号编码模型

本文所使用的适应度函数是:

$$J = D(\delta, T) + \lambda \times \text{abs}([H(\delta, T) - H_1]) \quad (1)$$

其中, λ 为拉格朗日乘子,是一个给定的值,经过多次实验表明 λ 的取值对适应度函数的结果没有影响,因此 λ 可以取任

意的值。abs 表示求绝对值,式中的 $D(\delta, T)$ 是重建信号和原始信号之间的误差,用百分均方根误差 (PRD) 表示。

$$PRD = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N x_i^2}} \times 100\% \quad (2)$$

其中, x_i 表示原始信号, y_i 表示经压缩后重建的信号, N 表示总信号点数。 H_1 是信息率,是一个给定的值。 $H(\delta, T)$ 也是信息率,是通过程序运行结果计算得出的,其计算公式为:

$$H(\delta, T) = N_1 / N \quad (3)$$

其中, N_1 表示信号压缩后的总比特数, N 表示总信号点数, T 是量化的门限值。在适应度函数中只有一个参数 δ , 即为算法中粒子所要搜寻的最优解,其中 $\delta = \Delta/2$, Δ 为量化步长。基于本文对一些 ECG 信号的实验结果的观察发现,为了取得好的压缩性能,参数 δ 的范围应该介于 $0.5T \sim T$ 之间。

2.2 编码和算法步骤

根据蚁群优化算法本身的特点,本文的连续蚁群算法采用实数编码方式。对于带截止区均匀量化器参数寻优中的蚂蚁可以直接编码为 (T, Δ) , 对应的适应度函数由式(1)给出。另外,对于每一个蚂蚁而言,如果用 (T, Δ) 来表示蚂蚁当前位置,那么,还应该有一个对应的 W 表示蚂蚁的移动步长。因为 (T, Δ) 是一个二维变量,所以 W 也应当是一个二维变量,即 $W = \{W_1, W_2\}$, 其中 W_1 为 T 对应的移动步长, W_2 为 Δ 对应的移动步长。这样,蚂蚁可以采用如下的编码结构。

T, Δ	W_1, W_2	$F(T, \Delta)$
蚂蚁当前位置	蚂蚁的移动步长	适应度函数值

本文中,连续函数优化问题的蚁群算法的基本步骤如下所示。

1) 初始化蚁群,包括相关参数、最大循环次数 N_c 及蚂蚁个数 M 的初始化。

2) 用对分查找方法确定待优化参数 T 的范围, Δ 的范围在 $0.5T \sim T$ 之间,将 M 个蚂蚁随机地放置于解空间中。

3) 计算每个蚂蚁的适应度函数值 $F(T, \Delta)$, 记录当前的最好位置,即信息素最多的位置。

4) 对每只蚂蚁进行全局搜索,更新蚂蚁所在位置。

5) 对获得最好位置的蚂蚁进行局部搜索,并更新当前最好解的位置。

6) 所有蚂蚁完成一次循环后,进行信息素的更新。

7) 若 $Iter < N_c$, 则转到 3), 否则输出结果。

连续函数优化问题的蚁群算法的基本流程如图 2 所示。

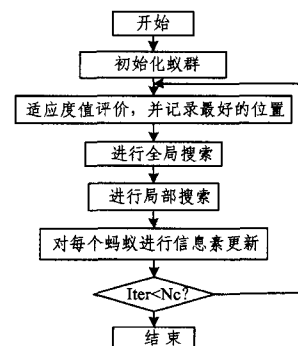


图 2 连续函数优化问题的蚁群算法的基本流程

在具体的程序中,各参数设置为: $w_{\max}=0.8, w_{\min}=0.1,$
 $\rho=0.7, k=2, a=e, \lambda=0.2, p_0=0.8, N_c=50, M=25。$

3 实例结果与分析比较

本文实验的所有 ECG 信号均是从 MIT-BIH 心率失常数据库得到的,所选的信号抽样频率均为 360Hz,数字化精度为 11,小波变换分解层数 $J=5$,信号处理帧的大小为 1024。原始信号和重建信号之间的误差采用目前最常用的百分均方根误差(PRD)来衡量。

进行实验的一组 ECG 数据的信号长度均为 1 分钟,包括 104、107、111、112、115、116、117、118、119、201、207、208、209、212、213、214、228、231 和 232。实验得出了这组 ECG 数据在不同 CR 下的 PRD 值,如表 1 所列。

表 1 ECG 数据的 PRD

Record	PRD						
	CR=4	CR=5	CR=8	CR=10	CR=12	CR=16	CR=20
104	1.4447	1.8312	2.8254	3.4048	4.0769	5.6481	7.4213
107	0.6942	0.9241	1.6054	2.0701	2.6221	4.103	6.1116
111	1.7049	2.1938	3.4741	4.185	4.6554	5.5671	6.8251
112	0.4384	0.5602	0.8359	1.0804	1.3309	1.8042	2.1488
115	0.6969	0.8624	1.3196	1.6371	1.9414	2.6149	3.6413
116	0.5654	0.7208	1.3155	1.7548	2.1598	3.193	4.4138
117	0.5192	0.6767	0.9926	1.1912	1.4441	1.8555	2.1834
118	0.53	0.6903	1.1501	1.5589	1.9191	2.7624	3.6994
119	0.5061	0.6583	1.0039	1.305	1.5729	2.1959	3.0539
201	1.4611	1.8449	2.9207	3.4338	3.9122	5.086	6.3445
207	1.1686	1.5074	2.3229	2.7903	3.1246	3.7738	4.6114
208	1.288	1.6833	2.9342	3.6735	4.4642	6.0956	8.0681
209	1.9186	2.5452	4.0191	4.7602	5.4918	7.1877	9.2732
212	1.7132	2.2283	3.7546	4.5478	5.3083	7.4318	10.1773
213	0.8142	1.0986	2.1193	2.9988	3.9528	6.0934	8.381
214	1.0676	1.3849	2.2469	2.7163	3.0709	4.0488	5.2987
228	1.3653	1.781	2.9988	3.7295	4.309	5.4979	6.6402
231	1.2217	1.5088	2.3468	2.9284	3.5271	5.0445	7.1354
232	1.7206	2.2034	3.7384	4.8218	5.8566	7.7267	9.2909

将本文算法搜索出来的最优解 δ 代入到编码模型中得到压缩这组数据的平均 PRD 值,如表 1 所列,将其与文献[12]提出的基于 SPIHT 的小波 ECG 数据压缩算法取得的压缩结果进行对比,结果如表 2 所列。可以看出,本文在最优的 δ 值下对这组数据进行压缩,取得的平均 PRD 值在相同压缩比下,都小于 SPIHT 算法的结果。

表 2 数据的压缩结果比较

算法	CR	8:1	10:1	12:1	16:1	20:1
SPIHT 算法	PRD	2.46	2.96	3.57	4.85	6.49
本文算法	(%)	2.31	2.87	3.41	4.62	6.04

下面将本文算法与另外几种基于小波变换的 ECG 编码方法进行比较。Hilton 提出了一种基于小波和小波包变换的 EZW 编码算法^[13]。他提出了对于 117 号记录当 $CR=8:1$, 时 $PRD=2.6\%$ 。并且将结果和以前压缩此记录最好的结果^[14](CR 相同时, $PRD=3.9\%$)相比较,用文献[12]中所提的 SPIHT 算法,得到相应的 PRD 为 1.18。用本文算法压缩 117 号记录,在取得相同的 CR 时, PRD 为 0.99%,比文献[12-14]的结果都要好,如表 3 所列。

表 3 几种算法的比较

算法	PRD(%)	CR	信号	抽样率(Hz)	位/抽样
SPIHT	1.18	8:1	MIT-BIH 117	360	11
Hilton	2.6	8:1	MIT-BIH 117	360	11
Diohn	3.9	8:1	MIT-BIH 117	360	11
本文算法	0.99	8:1	MIT-BIH 117	360	11

此外,采用本文算法和文献[15]中所提的经典 ASEC 算法压缩所选数据序列,在不同的 CR 值下,相应的平均 PRD 值如表 4 和表 5 所列。

表 4 本文算法

CR	8:1	10:1	12:1	20:1
PRD	2.31	2.87	3.41	6.04

表 5 文献[15]中的算法

CR	8:1	10:1	12:1	20:1
PRD	3.00	4.01	4.84	7.52

从表 4 和表 5 中可以看出,利用本文算法压缩所选信号,在相同压缩比下的 PRD 值要比 ASEC 的 PRD 值小得多,可见本文算法压缩效果比 ASEC 更好。

结束语 本文将蚁群优化算法应用于带截止区均匀量化器的优化中,并对这种算法的参数进行了设置,最后得到了较好的参数设置。实验结果表明这种算法具有有效性,适用于带截止区均匀量化器的参数优化。

参考文献

- [1] Colorm A, Dorigo M, Manieaao V. Distributed optimization by ant colonies[C]//Proc of the First European Conf. on Artificial Life, Paris, France Elsevier Publishing, 1991:134-142
- [2] Colorni A, et al. Ant system for job-shop scheduling[J]. JOR-BEL, 1994, 34(1):39-53
- [3] Costa D, Hertz A. Ant can color graphs[J]. Journal of the Opnl Res Soc, 1997, 48(3):295-305
- [4] Bullnheimer B, Hartl R F, Strauss C. An improved ant system algorithm for the vehicle routing problem[J]. Annals of Operations Research, 1999, 89:319-328
- [5] Coello C A C, Cutierrez R L Z, Garcia B M, et al. Automatec design of combinational logic circuits using the Ant Systm[J]. Engineering Optimization, 2002, 34(2):109-127
- [6] Cordone R, Maffioli F. Coloured ant system and local search to design local telecommunication networks[J]. Applications of Evolutionary Computing, 2001, 2037:60-69
- [7] 张纪会,徐心和. 具有变异特征的蚁群算法[J]. 计算机研究与发展, 1999, 36(10):1240-1245
- [8] 马良,项培军. 蚂蚁算法在组合优化中的应用[J]. 管理科学学报, 2001, 4(2):32-37
- [9] 刘康,余玲. 蚁群算法及其连续优化算法初析[J]. 四川轻化工学院学报, 2004, 17(1):42-46
- [10] Batista L V, Carvalho L C, Melcher E U K. Compression of ECG Signals Based on Optimum Quantization of Discrete Cosine Transform Coefficients and Golomb-rice Codeing[C]//Proceedings of the 25th Annual International Conference of the IEEE EMBS Cancun. Mexico * September 2003:17-21
- [11] Batista L V, Melcher E U K, Carvalho L C. Compression of ECG signals by optimized quantization of discrete cosine transform

coefficients[J]. Medical Engineering & Physics, 2001(23): 127-134

- [12] Lu Zhi-tao, Kim D Y, Pearlman W A. Wavelet Compression of ECG Signals by the Set Partitioning in Hierarchical Trees Algorithm[J]. IEEE Trans. Biomed. Eng. , 2000, 47(7): 849-856
- [13] Hilton M L. Wavelet and wavelet packet compression of electrocardiograms[J]. IEEE Trans. Biomed. Eng. , 1997, 44: 394-402
- [14] Djohan A, Nguyen T Q, Tompkins W J. ECG compression using discrete symmetric wavelet transform[J]. IEEE 17th Annual Conference on Engineering in Medicine and Biology Society,

1995; 167-168

- [15] Zigel Y, Cohen A, Abu-ful A, et al. Analysis by synthesis ECG signal compression[J]. Comput. in Cardiol. , 1997, 24: 279-282
- [16] 刘波. DNA—蚁群算法在车辆路径优化问题中的应用[J]. 中国新技术新产品, 2015(5): 7
- [17] 张伟娟, 张红梅, 陈峰. 基于蚁群算法的基因路径预测[J]. 计算机系统应用, 2015(3): 280-283
- [18] 陈亮, 吴更生, 吴卫. 针对应急救援路径规划的一种改进蚁群算法[J]. 后勤工程学院学报, 2015(1): 86-90

(上接第 543 页)

负载被推迟到这些时间槽内执行, 如图 4 所示。Power_Work 算法发生迁移的用电量高于 Online_WOFK 算法, 从而有效地降低了电费成本。

结束语 本文针对如何最小化数据中心的电费进行了深入研究, 主要贡献包括基于 Lyapunov 优化提出一种负载调度算法, 可以在不使用任何未来的负载、电价的基础上最小化数据中心的电费; 进行实验分析, 并将结果与 Online_WORK 算法进行比较, 结果表明 Power_Work 可以节约 7.2% 的电费。

综上所述, 在电力的调度过程中, 需要综合考虑负载波动性、电价差异性等因素。

参 考 文 献

- [1] Koomey J. Growth in data center electricity use 2005 to 2010. Oakland: Analytics Press, 2011 [OL]. <http://www.analyticspress.com/datacenters.html>
- [2] Qureshi A, Weber R, Balakrishnan H, et al. Cutting the electric bill for internet-scale systems[J]. ACM SIGCOMM Computer Communication Review, 2009, 39(4): 123-134
- [3] 北京市电网峰谷分时销售电价表[OL]. <http://www.bjpc.gov.cn/ywpd/wjgl/cx/jz/201208/t3884351.htm>
- [4] 上海市居民生活用电试行居民阶梯电价的实施细则[OL]. http://www.sh.sgcc.com.cn/FSM_CMS/html/main/col41/2013-03/21/20130321134924057695122_1.html
- [5] Chase J S, Anderson D C, Thakar P N, et al. Managing energy and server resources in hosting centers[J]. ACM SIGOPS Operating Systems Review, 2001, 35(5): 103-116
- [6] 谭一鸣, 曾国荪, 王伟. 随机任务在云计算平台中能耗的优化管理方法[J]. 软件学报, 2012, 23(2): 266-278
- [7] Rao L, Liu X, Xie L, et al. Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity market environment [C] // Proc. of the IEEE INFOCOM. IEEE, 2010; 1-9
- [8] Uргаonkar R, Uргаonkar B, Neely M J, et al. Optimal power cost management using stored energy in data centers [C] // Proc. of the ACM SIGMETRICS Joint Int'l Conf. on Measurement and Modeling of Computer Systems. New York: ACM Press, 2011: 221-232
- [9] Gao P X, Curtis A R, Wong B, et al. It's not easy being green

[J]. ACM SIGCOMM Computer Communication Review, 2012, 42(4): 211-222

- [10] Liu Z H, Lin M H, Wierman A, et al. Greening geographical load balancing [C] // Proc. of the ACM SIGMETRICS Joint Int'l Conf. on Measurement and Modeling of Computer Systems. New York: ACM Press, 2011: 233-244
- [11] Deng W, Liu F M, Jin H, et al. SmartDPSS: Cost-Minimizing multi-source power supply for datacenters with arbitrary demand [C] // Proc. of the 2013 IEEE 33rd Int'l Conf. on Distributed Computing Systems (ICDCS). IEEE, 2013: 420-429
- [12] Georgiadis L, Neely M J, Tassioulas L. Resource allocation and cross-layer control in wireless networks [J]. Foundations and Trends in Networking, 2006, 1(1): 1-144
- [13] Neely M J. Stochastic network optimization with application to communication and queueing systems [J]. Synthesis Lectures on Communication Networks, 2010, 3(1): 1-211
- [14] Fan X, Weber W D, Barroso L A. Power provisioning for a warehouse-sized computer [J]. ACM SIGARCH Computer Architecture News, 2007, 35(2): 13-23
- [15] Chase J S, Anderson D C, Thakar P N, et al. Managing energy and server resources in hosting centers [J]. ACM SIGOPS Operating Systems Review, 2001, 35(5): 103-116
- [16] Chen Y, Das A, Qin W, et al. Managing server energy and operational costs in hosting centers [J]. ACM SIGMETRICS Performance Evaluation Review, 2005, 33(1): 303-314
- [17] Wang P J, Rao L, Liu X, et al. Dynamic power management of distributed Internet data centers in smart grid environment [C] // Proc. of the IEEE Global Telecommunications Conf. (GLOBECOM 2011). IEEE, 2011: 1-5
- [18] Zhang Y W, Wang Y F, Wang X R. Capping the electricity cost of cloud-scale data centers with impacts on power markets [C] // Proc. of the 20th Int'l Symp. on High Performance Distributed Computing. New York: ACM Press, 2011: 271-272
- [19] googleclusterdata [OL]. <https://code.google.com/p/googleclusterdata/>
- [20] San clemente island data [OL]. <http://www.nrel.gov/mid/scid/>
- [21] Zhang Y W, Wang Y F, Wang X R. GreenWare: Greening cloud-scale data centers to maximize the use of renewable energy [C] // Kon F, ed. Proc. of the Middleware 2011. Berlin, Heidelberg: Springer-Verlag, 2011: 143-164