

一种改进粒子群优化算法的 Otsu 图像阈值分割方法

刘桂红¹ 赵亮² 孙劲光¹ 王星¹

(辽宁工程技术大学电子与信息工程学院 葫芦岛 125105)¹

(辽宁工程技术大学研究生学院 葫芦岛 125105)²

摘要 阈值法分割图像时只利用图像的灰度信息,具有直观、实现简单的特点。针对传统的粒子群优化算法(Particle Swarm Optimization, PSO)分割图像易陷入局部最优的缺点,提出一种基于改进粒子群优化算法的 Otsu 图像阈值分割方法。以 Otsu 算法的类间方差作为适应度函数,在每次迭代中选取适应度较好的粒子同时加入新的粒子,以提高粒子多样性。实验表明,与 Otsu 算法和 PSO 算法相比,改进的粒子群优化算法不仅加快了收敛速度和运算速度,而且提高了图像分割的准确率。

关键词 图像分割, Otsu, 类间方差, 粒子群优化, 适应度函数

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.3.058

Otsu Image Threshold Segmentation Method Based on Improved Particle Swarm Optimization

LIU Gui-hong¹ ZHAO Liang² SUN Jin-guang¹ WANG Xing¹

(School of Electronics and Information Engineering, Liaoning Technical University, Huludao 125105, China)¹

(Graduate School, Liaoning Technical University, Huludao 125105, China)²

Abstract The thresholding method only needs the gray information to spilt image, which is more intuitive and much easier to be implemented. Aiming at the problem that the traditional PSO algorithm used for image segmentation is easy to fall into local optimum, this paper proposed an Otsu image threshold segmentation method based on the improved PSO. We took the inter-class variance of Otsu as the fitness function, and selected the particles with better fitness and added new particles to increase the diversity of the particles. The experimental results show that, compared with Otsu methods and PSO algorithm, the improved PSO accelerates the speed of convergence and computation, and improves the accuracy of image segmentation.

Keywords Image segmentation, Otsu, Inter-class variance, Particle swarm optimization, Fitness function

1 引言

图像分割是指把图像描述成某些连通区域的集合,使得图像特征在不同区域表现不同,在同一区域表现出相似性的处理。目前图像分割方法主要有阈值法^[1]、边缘检测法^[2]、区域法^[3]、形态学分水岭法^[4]等。阈值法由于实现简单且计算速度快等特点,在图像分割应用中处于核心地位^[5]。其中日本学者 Otsu 在 1979 年提出的 Otsu 方法以在一幅图像直方图上执行计算为基础,通过计算目标与背景之间的最大类间方差得到分割阈值,取得了很好的效果,得到了广泛应用,但这种算法的计算量较大,很难适应实时处理。近年来,一些学者提出采用智能优化算法来求解最优阈值,出现了基于遗传算法(Genetic Algorithm, GA)^[6]、蚁群算法(Ant Colony Algorithm, ACA)^[7]、粒子群算法(Particle Swarm Optimization, PSO)^[8]、神经网络(Neural Network, NN)^[9]的等阈值分割方法。受鱼群和鸟群觅食启发而提出的粒子群优化算法计算简

单、易于实现、计算效率高,其一经提出就引起了许多研究者的关注,得到了广泛的应用,但是也存在容易陷入早熟的缺点。

为提高 PSO 对图像阈值分割的性能,提出了一种新的基于改进 PSO 的图像阈值分割方法,对阈值的选取过程进行了优化,并分别与 Otsu、PSO 方法进行了对比,结果表明,改进的 PSO 算法节约了运行时间,提高了图像阈值分割的准确率。

2 粒子群优化算法的 Otsu 图像阈值分割

2.1 Otsu 阈值法

1979 年,日本学者 Otsu 提出了 Otsu 图像分割算法。该算法把图像分为目标和背景两个区域,计算两部分的方差值 σ^2 , σ^2 越大,图像两部分差别就越大,使得 σ^2 最大的灰度值被认为是最佳分割阈值。

设 $G = \{0, 1, 2, \dots, L-1\}$ 为待处理灰度图像的 L 个不同

到稿日期:2015-01-30 返修日期:2015-04-10 本文受青年科学基金项目(61402212),语义 Web 模糊规则互换与推理关键技术研究资助。

刘桂红(1971-),女,硕士,副教授,主要研究方向为图像处理、模式识别与计算机过程控制, E-mail: 519222375@qq.com; 赵亮(1991-),男,硕士,主要研究方向为图像处理、模式识别; 孙劲光(1962-),女,博士,教授,主要研究方向为图形理论与技术、图像工程、数据库原理与系统; 王星(1983-),男,博士,副教授,主要研究方向为智能数据与知识工程。

的灰度级, n_i 表示灰度级为 $i (i \in G)$ 的像素数目。则总的像素数目 N 为:

$$N = \sum_{i=0}^{L-1} n_i \quad (1)$$

灰度值为 i 的像素出现的概率 p_i 为:

$$p_i = \frac{n_i}{N} \quad (2)$$

容易得出:

$$\sum_{i=0}^{L-1} p_i = 1, p_i \geq 0 \quad (3)$$

选择一个阈值 $T(k) = k (0 < k < L-1)$, 将图像分割为类 C_1 和 C_2 。其中灰度值在范围 $[0, k]$ 内的所有像素归为 C_1 , 灰度值在范围 $[k+1, L-1]$ 内的所有像素归为 C_2 。则类 C_1 和 C_2 发生的概率分别为:

$$x_1 = \sum_{i=0}^k p_i, x_2 = \sum_{i=k+1}^{L-1} p_i = 1 - p_1(k) \quad (4)$$

分配到类 C_1 和 C_2 像素的灰度均值分别为:

$$m_1 = \frac{1}{x_1} \sum_{i=0}^k i p_i, m_2 = \frac{1}{x_2} \sum_{i=k+1}^{L-1} i p_i \quad (5)$$

整幅图像的灰度均值为:

$$m_G = \sum_{i=0}^{L-1} i p_i \quad (6)$$

则类 C_1 和 C_2 的类间方差 σ^2 表示为:

$$\sigma^2 = x_1 (m_1 - m_G)^2 + x_2 (m_2 - m_G)^2 \quad (7)$$

由式(4)~式(7)可进一步得:

$$\sigma^2 = x_1 x_2 (m_1 - m_2)^2 \quad (8)$$

对于比较简单的灰度图像, Otsu 单阈值的分割效果比较理想; 而对于较为复杂的图像, 则需要将 Otsu 推广到多阈值来改善分割效果。在处理的图片数量比较大时, 由于 Otsu 算法通过遍历所有灰度级来求取最佳分割阈值, 计算量比较大, 难以满足实时处理的要求。

2.2 粒子群优化算法

1995年, Kennedy 和 Eberhart 共同提出了粒子群优化算法 (Particle Swarm Optimization, PSO)。PSO 是一种基于群体行为的模仿鸟类觅食的优化算法。它把鸟类抽象为没有质量和体积的粒子, 粒子在解空间以一定的速度飞行。粒子通过自身的个体行为和整个种群的社会行为来调整自己飞行的速度。每个粒子代表最优问题的一个潜在最优解。

假设粒子搜索空间为 D 维空间, 粒子数目为 N 。其中粒子 i 的空间位置记为 $X_i = \{X_{i1}, X_{i2}, \dots, X_{iD}\}$, $i = \{1, 2, 3, \dots, N\}$ 。每个粒子的飞行速度 $V_i = \{V_{i1}, V_{i2}, \dots, V_{iD}\}$ 。粒子 i 所经历过的最好位置记为 $Pbest_i = \{Pbest_{i1}, Pbest_{i2}, \dots, Pbest_{iD}\}$ 。所有粒子经历过的最好位置记为 $Gbest_i = \{Gbest_{i1}, Gbest_{i2}, \dots, Gbest_{iD}\}$ 。每个粒子通过跟踪 $Pbest$ 和 $Gbest$ 来更新自己。具体更新公式为:

$$V_{iD}(k+1) = w \times V_{iD}(k) + c_1 \times r_1 [Pbest_{iD} - X_{iD}(k)] + c_2 \times r_2 [Gbest_{iD} - X_{iD}(k)] \quad (9)$$

$$X_{iD}(k+1) = X_{iD}(k) + V_{iD}(k+1) \quad (10)$$

式中, k 表示迭代次数, w 表示粒子的惯性权重, c_1 和 c_2 代表加速因子, r_1 和 r_2 代表 $[0, 1]$ 之间的随机数。其中, w 用来控制上一代粒子速度对当代粒子速度的影响。目前适应面最广的是 w 线性递减更新策略, 即 w 随迭代次数增加而线性减小。此时惯性权重因子 w 由式(11)决定, 本文采用此策略进行惯性权重的更新。

$$w = w_{\max} - (w_{\max} - w_{\min}) \times \frac{g}{G} \quad (11)$$

式中, w_{\max} 、 w_{\min} 分别表示惯性权重的最大值和最小值, g 和 G 分别表示当前迭代次数和最大迭代次数。

由式(9)可知, 粒子飞行速度的更新由 3 个部分决定: 1) 惯性部分, 代表粒子保持自身速度飞行的趋势, 这一部分具备探索新区域的能力, 能扩大粒子的搜索范围; 2) 个体行为部分, 代表粒子的自我学习过程, 这部分使粒子具备较强的局部搜索的能力; 3) 社会行为部分, 代表粒子之间的信息共享和互相合作。

3 改进粒子群优化算法的 Otsu 图像阈值分割

3.1 改进的粒子群优化算法

为提高 PSO 算法的搜索效率, 目前许多学者提出了大量的更新策略。Hsieh 等提出了基于种群管理和分享准则的 PSO 算法^[10]。Zhan 等提出了基于精英学习策略的自适应 PSO 算法^[11]。Li 等提出了基于正切函数惯性权重更新的 PSO 算法^[12]。

本文认为在粒子更新过程中, 应当加入新的粒子来防止算法陷入局部最优, 但加入新的粒子扩大了种群的规模, 增加了计算量。故提出一种新的策略机制来更新粒子种群, 在保证种群规模不变的情况下, 提高了粒子的多样性, 同时粒子的收敛速度得到了提高。新的策略机制描述如下:

1) 初始化粒子位置和速度, 由式(9)、式(10)产生 N 个新粒子;

2) 随机产生 M 个新粒子, 并对产生的 $N+M$ 个粒子按适应度大小进行排序。

在粒子群体更新过程中, 我们总是希望保留适应度较好的粒子, 但这类粒子浓度过高, 则粒子的多样性较差, 算法容易陷入局部最优, 同时丢失了有着良好进化趋势而适应度较差的粒子。故本文提出了截断粒子的概念, 利用截断粒子 Q 把粒子群分为两部分:

① 适应度值排在前面的 $Q (1 \leq Q \leq N)$ 个粒子;

② 适应度值靠后的 $N+M-Q$ 个粒子。

通过粒子①提高粒子的运算速度和收敛速度, 通过粒子②弥补适应度较差却有良好进化趋势的粒子, 同时提高粒子的多样性。在粒子更新的过程中, 为了保持种群的规模不变, 从粒子②选择 $N-Q$ 个粒子组成新的粒子种群, 粒子 i 的选择概率 $P(i)$ 为:

$$P(i) = \frac{fitness(i)}{\sum_{i=Q+1}^{N+M} fitness(i)} \quad (12)$$

其中, $fitness(i)$ 表示粒子 i 的适应度函数值。

在这种更新策略下, 算法开始执行之后, 在粒子更新过程中, 适应度排在前面的粒子加快了粒子的收敛速度, 适应度靠后的粒子和新的粒子提高了粒子的多样性。随着迭代次数的增加, 算法跳出局部最优解的能力更强, 更有可能获得全局最优解。

3.2 改进粒子群优化算法的 Otsu 图像阈值分割过程

(1) 设置粒子群规模为 N , 随机初始化粒子 $i (1 \leq i \leq N)$ 的位置 $X_i (0 \leq X_i \leq 255)$ 。随机生成粒子的初始速度 $V_i (-V_{\max} \leq V_i \leq V_{\max})$, 其中 V_{\max} 表示粒子的最大速度, 设置其它相关参数。

(2) 将式(8)作为粒子的适应度函数, 计算每个粒子的适应度值。

表2 3种算法对 lake 图像进行分割的效果对比

算法	实验次数	分割阈值	迭代次数	运行时间 (ms)	寻得最优解次数	寻优准确率 (%)
Otsu	100	122	256	696	100	100
PSO+Otsu	100	124	9	402	82	82
本文算法	100	122	8	370	96	96

结束语 在对图像进行阈值分割时,Otsu算法的分割效果好,但其计算量大、消耗时间长,不能满足实时性的要求。标准 PSO 算法在求解最优阈值时,易陷入局部最优,寻优性能差。本文算法在标准 PSO 的基础上,引入一种改进的 PSO 算法,对粒子群的适应度进行排序,选取适应度较好的粒子,同时在每一次迭代中,都有新的粒子加入,提高了粒子的多样性,取得了更好的分割效果。实验表明,改进的 PSO 算法能取得与 Otsu 算法相当的分割效果,比基于标准 PSO 的 Otsu 算法寻优能力强,同时运行时间最短。在实际应用中,本算法具有更好的实用价值。

本文算法由于是应用于单阈值分割,如何更好地将粒子群优化算法应用于多阈值分割将是今后研究的重点。

参考文献

- [1] Long Jian-wu, Shen Xuan-jing, Chen Hai-peng. Adaptive minimum error thresholding algorithm[J]. Acta Automatica Sinica, 2012,38(7):1134-1144(in Chinese)
龙建武,申铨京,陈海鹏.自适应最小误差阈值分割算法[J].自动化学报,2012,38(7):1134-1144
- [2] Wang Hui, Wang Lai-sheng, Zhong Ping. Level set image segmentation based on scale transform of edge detection function [J]. Computer Engineering, 2009, 35(24):202-204(in Chinese)
王辉,王来生,钟萍.基于边缘检测函数尺度变换的水平集图像分割[J].计算机工程,2009,35(24):202-204
- [3] Zhang Fa-cun, Zhao Xiao-hong, Wang Zhong, et al. A data parallel algorithm on region growing image segmentation [J]. Computer Engineering, 2004, 30(17):14-16(in Chinese)
张发存,赵晓红,王忠,等.区域生长法图像分割的数据并行方法研究[J].计算机工程,2004,30(17):14-16
- [4] Gao Li, Yang Shu-yuan, Li Hai-qiang. New unsupervised image segmentation via marker-based watershed [J]. Journal of Image and Graphics, 2007, 12(6):1025-1032(in Chinese)
高丽,杨树元,李海强.一种基于标记的分水岭图像分割新算法 [J]. 中国图象图形学报, 2007, 12(6):1025-1032
- [5] Han Si-qi, Wang Lei. A survey of thresholding methods for image segmentation [J]. Systems Engineering and Electronics, 2002, 24(6):91-94(in Chinese)
韩思奇,王蕾.图像分割的阈值法综述[J].系统工程与电子技术, 2002, 24(6):91-94
- [6] Zhang Chao, Zhang Jia-shu, Jia Dong-li. Image thresholding based on Chaos genetic algorithm [J]. Computer Engineering and Applications, 2006, 42(2):45-47(in Chinese)
张超,张家树,贾东立.基于混沌遗传算法的图像阈值分割[J].计算机工程与应用, 2006, 42(2):45-47
- [7] Xing Xu-dong, Zhou Xu, Mi Jian. Image segmentation based on improved ant colony algorithm [J]. Radio Communication Technology. 2013, 39(6):71-73(in Chinese)
邢旭东,周旭,米健.基于改进的人工蚁群的图像分割算法[J].无线电通信技术, 2013, 39(6):71-73

(3)根据计算出的粒子适应度,更新每个粒子的个体最优位置 P_{best} 和全局最优位置 G_{best} 。

(4)利用式(9)、式(10)更新粒子位置和速度。

(5)随机产生 M 个新粒子,计算新产生的 $N+M$ 个粒子适应度值,并对其排序。

(6)设置截断粒子 Q 的位置,把新的粒子群分为适应度值排在前面的 $Q(1 \leq Q \leq N)$ 个粒子和剩下的 $N-Q+M$ 个粒子。

(7)利用式(12)从剩下的 $N-Q+M$ 个粒子中选择 $N-Q$ 个粒子和适应度排在前面的 Q 个粒子组成新的粒子种群。

(8)当最优适应值连续不变的次数达到指定数值时,终止迭代,此时得到的全局最优位置 G_{best} 即为最佳分割阈值,否则转到步骤(2)。

(9)利用得到的阈值 G_{best} 对图像进行分割。

4 实验结果及分析

为了验证本文算法在图像分割中的有效性,采用不同的 256 灰度级图像进行实验,图片的分辨率为 512×512 。算法实验环境为:Windows7 系统,2.20GHz 处理器,2GB 内存;编程环境为 VS2010。参数设置为:粒子数目 N 为 10, $V_{max} = 50$, $c_1 = c_2 = 2$, $w_{max} = 0.9$, $w_{min} = 0.4$, $Q = 8$, $M = 2$, 最大迭代次数 $G = 50$ 。其中,搜索空间维数 D 为 1。

采用 Otsu 算法、标准 PSO 算法和本文算法对图像进行分割,得到的分割效果分别如图 1、图 2 所示。



图1 lena 图像分割效果



图2 lake 图像分割效果

采用 Otsu 算法、标准 PSO 算法和本文算法对不同的图像进行分割,它们的平均运行结果如表 1 和表 2 所列。其中,寻优准确率 $d = (k/t) \%$, k 和 t 分别表示寻得最优解次数和实验次数。从表 1 和表 2 可以看出,Otsu 算法的寻优准确率高,但由于其通过遍历所有灰度级来得到最佳分割阈值,计算量大,消耗时间最长,不能满足实时性的要求,适应性比较低;与 Otsu 算法相比,标准 PSO 算法计算量小,收敛速度快,消耗时间短。但是标准 PSO 算法容易陷入局部最优,导致寻优准确率较低。本文对标准 PSO 算法进行改进,加快了粒子的收敛速度和运算速度,提高了粒子的多样性,极大地提升了寻优准确率,获得了更好的分割效果。

表1 3种算法对 lena 图像进行分割的效果对比

算法	实验次数	分割阈值	迭代次数	运行时间 (ms)	寻得最优解次数	寻优准确率 (%)
Otsu	100	123	256	702	100	100
PSO+Otsu	100	125	9	389	78	78
本文算法	100	123	8	343	94	94

- [8] Zhang Lei, Gao Shang. An image segmentation method based on elite theory-improved particle swarm optimization[J]. Computer applications and Software, 2009, 26(12): 89-92 (in Chinese)
张磊, 高尚. 基于精英粒子群优化算法的图像分割方法[J]. 计算机应用与软件, 2009, 26(12): 89-92
- [9] Zhou Chang-ying. Research on image segmentation technology based on improved fuzzy BP neural network[J]. Computer Simulation, 2011, 28(4): 287-290 (in Chinese)
周长英. 基于改进的模糊 BP 神经网络图像分割算法[J]. 计算机仿真, 2011, 28(4): 287-290
- [10] Hsieh Sheng-ta, Sun Tsung-ying, Liu Chan-cheng, et al. Efficient population utilization strategy for particle swarm optimizer[J]. IEEE Transactions on Systems, Man, and Cybernetics, 2009, 39(2): 444-456
- [11] Zhan Zhi-Hui, Zhang Jun, Li Yun, et al. Adaptive particle swarm optimization[J]. IEEE Transactions on systems, Man, and Cybernetics, 2009, 39(6): 1362-1381
- [12] Li Li, Xue Bing, Niu Ben, et al. The novel non-linear strategy of inertia weight in particle swarm optimization [C]//Proceedings of the Congress on Bio-Inspired Computing. 2009: 1-5

(上接第 278 页)

网页的不足得到改善; 兴趣度预测因子能及时根据用户兴趣的变化调整网页排序, 使用户感兴趣的网页快速上浮。仿真实验结果表明, ITPR 算法使得网页排序质量有一定程度的改善, 进一步提升了网页查准率, 提高了用户满意度。

参 考 文 献

- [1] China Internet Network Information Center(CNNIC). The thirty-fourth statistical report of Chinese Internet development[R]. (2014-07). <http://baik.e.baidu.com/view/14341540.htm> (in Chinese)
中国互联网络信息中心(CNNIC). 第 34 次中国互联网络发展状况统计报告[R]. (2014-07). http://www.edu.cn/focus_1658/20140721/t20140721_1152815.shtml
- [2] Feng Hai-tao. An improved PageRank algorithm with web time weight[J]. Journal of Xi'an University of Posts and Telecommunications, 2013, 18(2): 121-124 (in Chinese)
冯海涛. 基于网页时间权值的 PageRank 算法改进[J]. 西安邮电大学学报, 2013, 18(2): 121-124
- [3] Shi Ming-ming. Research on Weighted PageRank algorithm[J]. Software Guide, 2013, 12(2): 30-32 (in Chinese)
史铭茗. 加权 PageRank 算法研究综述[J]. 软件导刊, 2013, 12(2): 30-32
- [4] Brin S. The anatomy of a large hypertextual Web search engine [J]. Computer Networks and ISDN System, 1998, 30(98): 107-117
- [5] Shao Jing-jing, Li Bo, Liu Han-ping. An improved pagerank algorithm-adjusting the damping factor[J]. Mathematica Applicata, 2008, 21(S1): 57-61 (in Chinese)
邵晶晶, 李波, 刘汉平. PageRank 的改进算法——调整阻尼因子[J]. 应用数学, 2008, 21(S1): 57-61
- [6] Lovasz L, et al. Random Walks on Graphs: A Survey [J]. Combinatorics, 1993, 8(4): 1-46
- [7] Xing W, Ghorbani A. Weighted PageRank algorithm[C]//Proceedings of Second Annual Conference. Piscataway: IEEE Press, 2004: 305-314
- [8] Manning C D, Raghavan P, Schutze H, et al. Introduction to information Retrieval [M]. Beijing: Post & Telecom Press, 2010
- [9] Tyagi N, Sharma S. Comparative study of various Page Ranking Algorithms in Web Structure Mining (WSM) [J]. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2012, 1(1): 14-19
- [10] Taher H. Topic-sensitive PageRank [C]//Proceedings of the 1th International Conference on World Wide WEB. Honolulu. Hawaii: ACM Press, 2002: 784-796
- [11] Li Wei-dong, Lu Ling. Research and application of pageRank algorithm combined with VSM technique[J]. Computer and Modernization, 2011(7): 96-98 (in Chinese)
李卫东, 陆玲. 融合 VSM 技术的 PageRank 算法研究与应用 [J]. 计算机与现代化, 2011(7): 96-98
- [12] Wang Zhong-fei, Gong Biao. Improved pageRank algorithm based on anchor texts similarity[J]. Computer Engineering, 2010, 36(24): 258-260 (in Chinese)
王钟斐, 工彪. 基于锚文本相似度的 PageRank 改进算法[J]. 计算机工程, 2010, 36(24): 258-260
- [13] Chakrabarti S, Dom B, Gibson D, et al. Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text [C]//Proceedings of the 7 ACM-WWW International Conference. Brisbane: ACM Press. 1998: 65-74
- [14] Duan H C, Hu P. Improved pagerank algorithm based on topic character and time factor [J]. Computer Engineering and Design, 2010, 31(4): 866-868 (in Chinese)
段淮川, 胡平. 基于主题特征和时间因子的改进 PageRank 算法 [J]. 计算机工程与设计, 2010, 31(4): 866-868
- [15] Kumar G, Duhan N, Sharma A K. Page Ranking Based on Number of Visits of Links of Web Page [C]//International Conference on Computer & Communication Technology (ICCT). 2011: 11-14
- [16] Peng Cong, Wu Qiang, Li Ren-fa. An Improved Algorithm of Web Page Ranking [J]. Microcomputer Information, 2010, 26(33): 72-74 (in Chinese)
彭聪, 吴强, 李仁发. 一种改进型的网页排序算法 [J]. 微计算机信息, 2010, 26(33): 72-74
- [17] Fang S F. Based on User Feedback PageRank algorithm [J]. Computer Technology and Automation, 2012, 31(1): 89-92 (in Chinese)
方树峰. 基于用户反馈的 PageRank 改进算法 [J]. 计算技术与自动化, 2012, 31(1): 89-92
- [18] Wang D G, Zhou Z G, Liang X. Analysis of pagerank algorithm and its improvement [J]. Computer Engineering, 2010, 36(22): 291-293 (in Chinese)
王德广, 周志刚, 梁旭. PageRank 算法的分析及其改进 [J]. 计算机工程, 2010, 36(22): 291-293
- [19] Mccandless M, Hatcher E, Gospodnetic O, et al. Lucene in action [M]. Beijing: Post & Telecom Press, 2011
- [20] Qiu Z, Fu T T. Lucene 2. 0 + Heritrix [M]. Beijing: Post & Telecom Press, 2007 (in Chinese)
邱哲, 符滔滔. 开发自己的搜索引擎: Lucene 2. 0 + Heritrix [M]. 北京: 人民邮电出版社, 2007