

基于克隆选择的图像分割算法研究

金章赞 肖刚 陈久军 高飞 周鸿斌

(浙江工业大学信息工程学院, 杭州 310032)

摘要 为了快速有效地对生物运动图像进行分割, 提出了一种新的图像分割算法, 该方法基于传统的 2 维最大熵分割法, 将计算机免疫学中的克隆选择算法应用于图像分割中。首先对图像 2 维阈值进行编码, 然后依据 2 维最大熵准则建立亲和力函数, 在给定的初始种群下, 对种群进行选择、克隆、变异、更新等操作, 由于克隆选择算法具有生物免疫系统自组织、自学习、自识别、自记忆的能力, 所以能够快速得到全局最优解, 实现图像的有效分割。从实验结果表明该算法所用搜索时间大约只有标准 2 维最大熵法的 1.3%, 证明该算法高效稳定, 是一种实用有效的图像分割算法。

关键词 图像分割 2 维最大熵 克隆选择 免疫算法

中图分类号: TP751 文献标志码: A 文章编号: 1006-8961(2010)02-0255-06

Image Segmentation Algorithm Based on Clone Selection

JIN Zhang-zan, XIAO Gang, CHEN Jiu-jun, GAO Fei, ZHOU Hong-bing

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310032)

Abstract In order to effectively segment biological movement image, this paper inspired by the two-dimension maximum entropy, use the clone selection algorithm of computer immunology into the image segmentation. First, this method codes the two-dimension maximum entropy. Second, the fitness function is established according to the criterion function of the two-dimension maximum entropy. Then with the given initial population, we execute selection, cloning, mutation and updating to the population, finally get the best result which can segment image efficiently. The experimental results indicate that this algorithm spend the search time approximately only then the ordinary two-dimensional biggest entropy method 1.3%. This proof the method is feasible and effective.

Keywords image segmentation, two-dimension maximum entropy, clone selection, immune algorithm

0 引言

图像分割是图像处理和机器视觉中的基本问题之一。图像分割的目的就是把图像中的目标与背景分开, 为后续处理提供基础。在目前常用的基于阈值、边缘和区域的分割方法中, 阈值法^[1]因其实现简单、性能稳定而成为最常用的方法。

而阈值法往往涉及寻找最优阈值的问题, 为了

加快最优阈值的搜索速度, 张毅军, 汪海洋等人^[2-3]提出了快速实现算法。这些算法都是通过以增加存储空间为代价来提高分割速度的。另外, 一些学者将传统的遗传算法等生物智能算法应用到图像分割中, 虽然减少了计算量。但是局部搜索能力较弱。所以如何快速有效地选取最优阈值是本文研究的一个重点。

克隆选择算法作为一种新的全局优化搜索算法, 在算法实现上兼顾全局和局部搜索, 通过克隆选

基金项目: 国家自然科学基金资助项目 (50705087); 浙江省自然科学基金资助项目 (Y106602)

收稿日期: 2008-08-14 改回日期: 2008-12-03

第一作者简介: 金章赞 (1984—), 男, 浙江工业大学信息工程学院图形图像研究所硕士研究生。主要研究方向为图像处理, 计算机视觉, 人工免疫等。E-mail: sacula1010@yahoo.com.cn

择操作使得算法快速收敛,通过变异操作使得算法保持了一定的多样性,抑制了早熟现象。丛琳、汤凌等人^[4-5]将免疫学中的相关免疫算法应用到遥感图像分割中,得到了较好的分割效果。综上所述,本文在普通 2 维最大熵图像分割方法的基础上,提出了一种基于克隆选择的 2 维最大熵图像分割方法,并将其运用到生物运动图像中,与传统的免疫算法相比,该方法借鉴了 De Castro 等人提出的克隆选择原理以及相应算法^[6],并进一步改进传统克隆选择算法的相关操作步骤,最终实现了 2 维最大熵法的快速计算和图像的有效分割。

1 克隆选择原理

克隆选择原理是免疫学的基本理论之一。它解释了抗体的形成机理并阐明了免疫应答的多样性机制。1973 年, Jeme 提出克隆选择原理^[7]。克隆选择原理的主要内容是:当外部细菌、病毒等抗原侵入机体后, B 细胞将对这些抗原进行识别,识别后的 B 细胞将克隆扩增分化为浆细胞,最终产生一种蛋白质分子即抗体细胞。在细胞克隆的过程中,抗体细胞还经历了一个变异的过程,其结果是产生对抗原具有特异性的抗体。

克隆选择的主要特征^[8]是:克隆选择对应着一个亲和力成熟的过程,即对抗原亲和力较低的抗体在克隆选择机制的作用下,经历克隆和变异操作后,其亲和力逐步提高而“成熟”的过程。克隆选择原理已在模式识别、网络入侵检测、组合优化和多峰值函数优化中得到了验证。这充分显示了克隆选择原理对于信息处理问题的求解具有广阔的应用前景。

2 克隆选择算法

De Castro 和 Von Zuben 基于克隆选择原理提出了克隆选择算法^[6,9],这是一种模拟免疫系统学习过程的进化算法。免疫应答产生抗体是免疫系统的学习过程,抗原被一些与之匹配的 B 细胞识别,这些 B 细胞大量分裂,产生新的 B 细胞在原有 B 细胞的基础上发生变异,以寻求与抗原更好匹配的 B 细胞,那些与抗原匹配更好的 B 细胞将再次分裂。如此循环往复,最终找到与抗原完全匹配的 B 细胞,这些 B 细胞最终一部分变成浆细胞产生抗体,另一

部分变成记忆性 B 细胞,形成免疫记忆,这一过程就是克隆选择过程。克隆选择算法模拟这一过程进行优化,其计算步骤如下:

1) 生成候选方案的一个集合 P ,它是记忆细胞 M 的子集加上剩余群体 P_r 的和 $P = P_r + M$ 。

2) 选择 n 个具有较高亲和力的抗体。

3) 克隆这 n 个最好的抗体,组成一个临时的克隆群体 C 。与抗原亲和力越高,抗体在克隆时的规模也就越大。

4) 把克隆群体提交到高频变异,根据亲和力的大小决定变异,产生一个成熟的抗体群体 C^* 。

5) 对 C^* 进行再选择,组成记忆细胞集合 M 。 P 中的一些成员可以被 C^* 中的其他一些改进的成员替换掉。

6) 生成 d 个新的抗体取代 P 中 d 个低亲和力的抗体,保持多样性。

克隆选择算法由于继承生物免疫系统的众多属性,并具有自组织、自学习、自识别、自记忆的能力,因此它不仅避免了优秀抗体的丢失,而且能得到全局最优解,这是其他优化技术所不及的。由于生物运动图像中运动生物的灰度值与背景的灰度值具有一定的差异,故可通过阈值法对其进行分割,又由于基于阈值的图像分割方法本质上就是要寻找一个最优阈值即全局最优解来实现图像的有效分割的,因此本文尝试将克隆选择算法应用到生物运动图像的分割中,提出了一种基于克隆选择的图像分割算法。

3 基于克隆选择的图像 2 维熵分割算法

3.1 算法步骤

这里采用 2 维最大熵图像分割方法来反映本文算法在搜索最优阈值上的高效性。要找出图像 2 维最大熵最大的阈值对,可针对图像中的像素灰度及其相应邻近像素的灰度均值进行分割。由于在图像分割中采用了克隆选择算法,所以抗原与抗体的定义非常重要。该算法将抗原定义为使图像 2 维最大熵最大的阈值对,而抗体就是图像分割中阈值对的二进制编码。将图像 2 维最大熵的计算公式作为亲和力函数,找出使亲和力值最大的“成熟”抗体,即为分割的最优阈值对。算法步骤如下:

1) 随机产生一定数量阈值对的二进制编码,形

成初始抗体群。

2) 用亲和力公式计算抗体群中每个抗体的亲和力和值, 并排序。

3) 对抗体群进行选择、克隆、变异、更新等一系列操作, 产生下一代抗体群。

4) 重复步骤 2), 3) 直到满足迭代停止条件为止。

算法流程图如下:

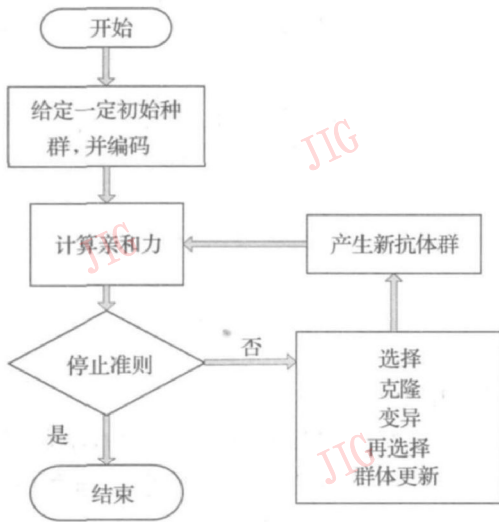


图 1 算法流程图

Fig 1 Flow chart of algorithm

3.2 抗体编码

在 L 级灰度图像的 2 维熵分割中, 图像信息用像素灰度值及其区域像素灰度均值的二元形式 $(f(x, y), g(x, y))$ 来表示, $0 \leq f(x, y), g(x, y) \leq L-1$ 其中 $f(x, y)$ 表示像素灰度值, $g(x, y)$ 表示其区域灰度均值。用 (s, t) 表示图像的一个可能分割阈值对, 其中存在某一个 (s', t') 为图像的最优分割阈值对。对于 256 级图像, 对 (s, t) 进行二进制编码需 16 位, 其中, s 编码在高 8 位, t 编码在低 8 位。

3.3 计算亲和力

免疫系统中, 亲和力值最大的抗体对应为“成熟”抗体。类似地, 在 2 维熵分割中, 使图像 2 维熵达到最大的阈值对即为最优分割阈值对。因此可用 2 维最大熵计算式^[10]作为衡量亲和力的公式, 2 维最大熵分割法利用图像中各像素点的灰度值及其区域灰度均值为依据来选取最佳阈值。其设计思想是, 采用 2 维最大熵来度量像素以及其区域中像素之间的信息相关性。这样, 当空间相关信息量为最大时, 认为该阈值为最佳阈值。具体步骤如下:

1) 计算每一个像素的联合概率 p_{ij} :

$$p_{ij} = f_{ij} / N \times N \quad (i = 0 \dots 255, j = 0 \dots 255) \quad (1)$$

式中, f_{ij} 为像素灰度值为 i 及其区域像素灰度均值为 j 的像素个数, $N \times N$ 为图像的总像素数。

2) 给定一对初始阈值 (s, t) , 将图像分为 C_1 和 C_2 两类。

3) 分别计算这两个类的平均相对 2 维熵:

$$E_1 = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (p_{ij} / P_1) \log(p_{ij} / P_1) \quad (2)$$

$$E_2 = - \sum_{i=s}^{L-1} \sum_{j=t}^{L-1} (p_{ij} / (1 - P_1)) \log(p_{ij} / (1 - P_1)) \quad (3)$$

式中, $P_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij}$, L 为图像灰度级数。

(4) 选择最佳的阈值 (s', t') , 使得图像按照该阈值分为 C_1 和 C_2 两类后, 满足

$$(E_1 + E_2)(s', t') = \max\{E_1 + E_2\} \quad (4)$$

3.4 抗体选择

为了产生优秀抗体, 规范变异方向, 使抗体向着亲和力值更高的方向变异, 故需对抗体进行选择。亲和力值越高, 选择概率也越大。但是, 当群体中的某种个体的数量占据了相当规模, 而此抗体又不是最优抗体时, 依据以上规则, 易导致抗体过早收敛。为此, 要对某些达到一定规模的抗体进行抑制, 同时增加小规模抗体的产生, 以提高多样性。所以在传统的适应度比例选择机制的基础上, 增加基于浓度的调节概率因子^[11]。

生物免疫浓度调节机制是基于抗体的浓度进行的, 抗体浓度越大, 则被选中的概率越小。抗体 i 的浓度定义为

$$C_i = \frac{\text{和抗体 } i \text{ 亲和力值大于 } \lambda \text{ 的抗体数}}{\text{抗体总数 } N} \quad (5)$$

式中, λ 为亲和力阈值。计算出每个抗体的浓度之后, 便可通过选择机制进行抗体的促进和抑制调节。于是抗体 i 的选择概率 p_i 由亲和力概率 p_f 和浓度概率 p_d 两部分组成, 即

$$p_i = p_f \cdot p_d = \frac{f_i}{\sum_{i=0}^N f_i} \cdot \frac{1}{\alpha} e^{-\beta C_i} \quad (6)$$

式中, α 和 β 为常数, 文中均取为 1, f_i 为抗体 i 的亲和力值。

3.5 抗体克隆

为了避免优秀抗体丢失, 增加优秀抗体数量, 提高全局收敛效率, 需对抗体进行适当的克隆。亲和

力值越高的抗体,克隆数量越多。但是,当相似抗体数量较多时,如果再继续克隆的话,就会导致抗体群过于相似,抗体过早收敛。所以增加一个相似抗体浓度因子,使得对于相似抗体浓度越大的抗体,其克隆数量越少。其中相似抗体指的是抗体之间的亲和力值小于某一阈值的抗体个数。

抗体 i 的相似抗体浓度为

$$m_i = \frac{\text{相似抗体个数}}{\text{抗体总数}} N \quad (7)$$

抗体 i 的克隆个数为

$$n_i = \text{int} \left[N \cdot \frac{f_i}{\sum_{j=1}^N f_j} \cdot \frac{1}{\alpha} e^{-\beta m_i} \right] \quad i = 1, 2, \dots, N \quad (8)$$

式中, f_i 为抗体 i 的亲和力值, α 和 β 为常数,取值均为 1, N 为抗体总数, int 表示取整操作。其效果如图 2 所示。

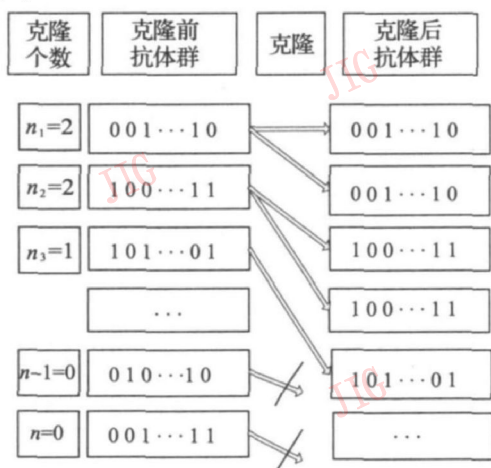


图 2 克隆示意图

Fig 2 Clone schematic drawing

3.6 抗体变异

为了获得抗体的多样性使抗体快速地成熟,需对抗体进行适当的变异。抗体的变异不仅有利于快速搜索,且能使抗体跳出局部最优,得到全局最优。为了使抗体朝着亲和力值高的方向进行变异,亲和力值高的抗体,变异应该小一些,甚至其变异被抑制,而亲和力值低的抗体,相应的变异就应该大一些。

首先计算每个抗体的亲和力,并按亲和力的大小对抗体进行排序,对抗体变异,每次仅变异一位,即对该位进行取反。对于亲和力值高的抗体,相映变异的位所处的位置就越低,对于亲和力值小的抗体,相映变异的位所处的位置就越高。这样虽然每

次仅变异一位,但是对低位进行变异,抗体变化小,而对高位进行变异,抗体变化大,所以能达到变异目的,且实现简单、容易计算。其示意图如图 3。



图 3 变异示意图

Fig 3 Variation schematic drawing

3.7 抗体更新

为了保持抗体多样性,防止抗体过早收敛,对抗体群进行更新。该操作将抗体群中亲和力值较小的部分抗体用随机产生的等量新抗体取代。

4 实验结果与分析

为了验证本文算法的有效性和可行性,以一幅标准测试图像和一幅真实的生物鱼群运动图像为例进行实验,该鱼群图像是由 CCD 摄像头和图像采集卡所获取的数字图像。对鱼群图像的分割和鱼类目标的提取及识别是水质监测系统的关键技术之一。鱼群运动图像一般可以看作是由目标和背景两类像素构成的图像,图像中的目标即为鱼群。对鱼群图像的分割的目的就是从图像中提取出感兴趣的鱼群目标,从而为后面的鱼群跟踪提供基础,由此可见,鱼群图像的分割和目标提取对水质监测系统至关重要。

设种群大小为 30 最大迭代次数为 10 代,以最大迭代次数作为迭代停止条件。为了比较,分别采用本文的克隆选择算法,1 维最大熵法,普通的 2 维最大熵法以及标准的遗传算法对图像进行分割,并对分割的质量和效率作相应的分析。

表 1 列出了前 3 种算法对该图像进行分割得到的最优阈值对和所花费的时间。结果显示,本文算法与 2 维最大熵法较 1 维最大熵法具有更强分割能

力,且本文算法与 2 维最大熵法所得出的最优阈值对完全一致,故验证了本文算法的正确性。此外,本文算法搜索得到的最优阈值对的时间大约只有普通 2 维最大熵法的 1.3%,分割速度有了极大的提高,这主要是因为克隆选择算法中,抗体通过克隆和变异实现了快速成熟,从而验证了本文算法用于实时图像处理的可行性。

表 1 3 种算法的分割结果与时间比较

Tab 1 Three algorithm's division results and the time compare

算法	阈值	时间 /s
1 维最大熵法	(136)	0.2
2 维最大熵法	(160 146)	56
本文算法	(160 146)	0.7

表 2 给出了本文算法和标准遗传算法对图像进行分割所得结果的比较。遗传算法在 5 代内基本上无法得到最优值,在 10~20 代内搜索有所加强,但仍得不到最优值,而本文算法在 10 代内搜索就可得到最优阈值,同时,在相同的迭代次数下,用本文算法虽然搜索时间较长,但搜索效率却高很多。这主要是因为克隆选择算法中,根据亲和力函数的大小以及抗体的浓度来确定抗体克隆的数量,从而使优秀抗体得以保存。所以实验结果表明,用克隆选择算法来求取最优阈值是有效的。

表 2 遗传算法与克隆选择算法比较

Tab 1 The genetic algorithm and the clone selection algorithm compares

算法	遗传算法			克隆选择算法	
	5 代	10 代	20 代	5 代	10 代
阈值对	(136, 130)	(142, 137)	(153, 145)	(161, 147)	(160, 146)
时间 /s	0.7	1.3	1.9	0.4	0.7

图 4 显示了生物鱼群运动图像的实验效果,其中图 4(a)为原始图像,图 4(b)为 1 维最大熵算法效果图,图 4(c)为遗传算法效果图,图 4(d)为本文算法分割效果图。比较图 4(d)与图 4(b),图 4(c)可知,图 4(b)由于未能分割鱼的阴影部分导致多条鱼相连且噪声较多,未能取得有效的分割效果。图 4(c)虽然分割掉了鱼的阴影部分,但是仍存在一些噪声使个别鱼相连,所以得到的效果并不令人满意。而图 4(d)不仅分割掉了阴影部分且进一步去

除了噪声,得到了良好的分割效果。这说明本文算法具有更强的分割和去噪的能力。

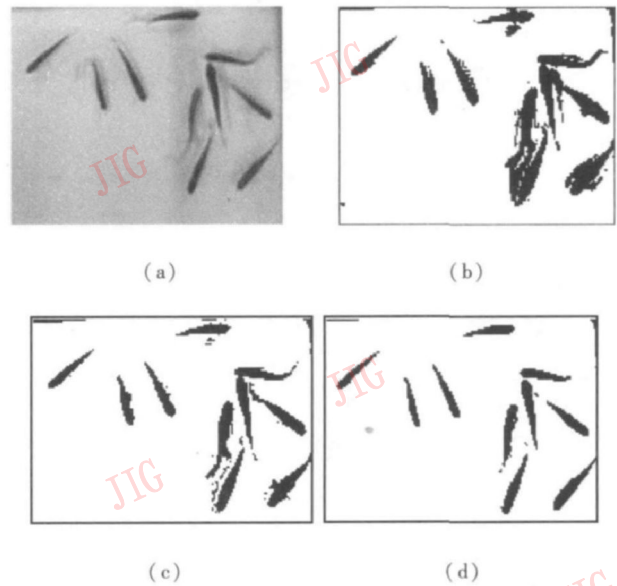


图 4 生物运动图像分割效果结果

Fig. 4 Division effect chart

图 5 为标准测试图像 (Lena 图像) 的分割效果图。其中图 5(a)为原始图像,图 5(b)为 1 维最大熵算法分割效果图,图 5(c)为本文算法分割效果图。可见图 5(c)中的人脸特征更加突出,细节部分更加清晰。



图 5 标准图像分割效果结果

Fig 5 Division effect chart

5 结 论

将克隆选择算法运用于 2 维最大熵图像阈值分割中,其分割及去除噪声的能力明显好于 1 维最大熵分割法,且较普通的 2 维最大熵分割法,搜索时间大大减少。与标准的遗传算法相比,克隆选择算法将克隆和变异作为其操作的主要算子,具有更强的局部和全局搜索能力,在实现快速收敛的同时能有效地找到最优阈值对。实验结果表明该算法是一种快速高效的图像分割算法。此外,由于文献 [6],

[9]和[12]直接或间接地证明了克隆选择算法是收敛的,所以提出的基于克隆选择的图像分割算法也是收敛的。

参考文献 (References)

- [1] Kapur J N, Sahoo P K, Wong A K C. A new method of gray level picture thresholding using the entropy of the histogram [J]. *Computer Vision, Graphics and Image Processing*, 1985, 29(2): 273-285.
- [2] Zhang Y r-jun, Wu Xue-jin. Application of a fast two-dimensional entropic thresholding method [J]. *Pattern Recognition and Artificial Intelligence*, 1997, 10(3): 259-264 [张毅军, 吴雪菁. 二维熵图像阈值分割的快速递推算法 [J]. *模式识别与人工智能*, 1997, 10(3): 259-264.]
- [3] Wang H a-yang, Pan D e-lu, Xia D e-shen. A fast algorithm for two-dimensional otsu adaptive threshold algorithm [J]. *ACTA Automatica Sinica*, 2007, 33(9): 968-971. [汪海洋, 潘德炉, 夏德深. 二维 Otsu 自适应阈值选取算法的快速实现 [J]. *自动化学报*, 2007, 33(9): 968-971.]
- [4] Cong L in, Sha Y u-heng, Jiao L i-cheng. Application of immune clone selection algorithm to image segmentation [J]. *Journal of Electronics & Information Technology*, 2006, 28(7): 1169-1173 [丛琳, 沙宇恒, 焦李成. 基于免疫克隆选择算法的图像分割 [J]. *电子与信息报*, 2006, 28(7): 1169-1173.]
- [5] Tang L ing, Zheng Z hao-bao, Yu X in. An image segmentation algorithm based on artificial immune [J]. *Geomatics and Information Science of Wuhan University*, 2007, 32(1): 67-70 [汤凌, 郑肇
- 葆, 虞欣. 一种基于人工免疫的图像分割算法 [J]. *武汉大学学报 (信息科学版)*, 2007, 32(1): 67-70.]
- [6] De Castro L N, Von Zuben F J. The clonal selection algorithm with engineering applications [C] // *Proceedings of Genetic and Evolutionary Computation Conference*. San Francisco, CA, USA: Morgan Kaufmann Publisher, 2000: 36-37.
- [7] Jeme N K. The immune system [J]. *Scientific American*, 1973, 229(1): 51-60.
- [8] Xiao Ren-bin, Wang Lei. Artificial immune system: Principle, models, analysis and perspectives [J]. *Chinese Journal of Computers*, 2002, 25(12): 1281-1293 [肖人彬, 王磊. 人工免疫系统: 原理, 模型, 分析及展望 [J]. *计算机学报*, 2002, 25(12): 1281-1293.]
- [9] De Castro L N, Von Zuben F J. Learning and optimization using the clonal selection principle [J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(3): 239-250.
- [10] Abutaleb A S. Automatic thresholding of gray-level pictures using two-dimensional entropy [J]. *Computer Vision, Graphics and Image Processing*, 1989, 47(1): 22-32.
- [11] Luo Wen-jian, Cao X ian-bin, Wang Xu-fa. Solving frequency assignment using an immune genetic algorithm [J]. *Chinese Journal of Electronics*, 2003, 31(6): 915-917. [罗文坚, 曹先彬, 王煦法. 用一种免疫遗传算法求解频率分配问题 [J]. *电子学报*, 2003, 31(6): 915-917.]
- [12] Du H a-feng, Jiao L i-cheng, Wang Sun-an. Clonal operator and antibody clone algorithm [C] // *Proceeding of the 1st IEEE International Conference on Machine Learning and Cybernetics*. New York, NY, USA: IEEE Publisher, 2002: 506-509.