

基于最大熵原理的多阈值自动选取新方法

曹力 史忠科

(西北工业大学自动控制系, 西安 710072)

摘要 基于信息论中最大熵原理, 提出了一种等概率场下的自动选取阈值的新方法. 该方法是首先通过图象灰度直方图信息, 并利用 Shannon 熵中等概率场具有最大熵的基本性质来确定阈值, 然后将图象划分为等概率的子块, 进而给出了该算法的理论推导和算法的具体实现步骤. 与通常的基于熵理论自动获取图象阈值的方法相比较, 该方法直观、简便、求解稳定, 且易于实现. 同时该方法克服了常用方法在阈值求取时, 出现的诸如计算量大、计算效率低等不足等问题, 故能够迅速地获得图象的阈值. 对比实验的结果, 也说明了该方法的快速性、有效性、可靠性和稳定性.

关键词 最大熵 阈值 灰度级 直方图

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2002)05-0461-05

An Automatic Multilevel Thresholding Method Based on Maximum Entropy

CAO Li, SHI Zhong-ke

(Department of Automatic Control, Northwest Polytechnic University, Xi'an 710072)

Abstract Image processing has to deal with many information of an image. Gray-histogram can contain a lot of image information. Maximum entropy theorem of Information Theory is one of the useful tools to treat with this kind of information. There are several formulas for computing the maximum entropy. But almost of the existing formulas have some common disadvantages, such as expensively computing and more complex algorithm realizing. In order to overcome these weaknesses of the existing entropy formulas, in this paper we define a new approach to entropy, and use it to automatically select thresholds of the image. It bases on one of Shannon entropy's basic properties that the equivalent probability distributing has maximum entropy to get the image thresholds. And by this way, we can segment an image into several equivalent probability sub-parts. This new method has some advantages, such as simplified, stabilized and easily realized comparing with some traditional entropy methods. At the same time, it can get image thresholds quickly. We have employed the newly proposed approach to perform image enhancement, segmentation and thresholding, and obtained satisfactory results.

Keywords Maximum entropy, Threshold, Grey-level, Histogram

0 引言

阈值分割法是进行图象分割的一种广泛应用的方法. 鉴于传统单阈值方法是将图象中的象素点按照阈值划分为两大类, 而多阈值方法是将图象象素点依不同的阈值划分为若干类, 其中同一类象素点具有相同的灰度级, 因此在这类方法中, 如何确定一

个较合理的阈值是进行图象分割至关重要的环节, 为此, 研究人员投入了大量的精力, 已取得了许多有价值的研究成果. 如 Otsu 提出了利用最大类间方差来求取阈值的方法^[1], 其基本原理是利用判别式来分析图象类间最大可分离性; Kittler 和 Illingworth 等人提出的最小误差法在进行图象分割时, 是将图象象素分类后误差率最小的值确定为图象分割阈值的^[2].

近几年来,为控制分割图象造成的信息损失而发展起来的一种基于信息论中熵准则的图象阈值自动选取方法引起了人们的极大关注,其理论研究及实际应用均获得重大突破^[3].这种方法的基本原理是:通过引入信息论中的最大 Shannon 熵准则和最小交叉熵准则来寻找使目标(或背景)内部灰度分布尽可能均匀的最优阈值,或者是寻找使分割前后图象的信息量差异最小的最优阈值.如 Pun 和 Kapur 等提出的利用最大先验熵估计分类合理性来选取阈值的方法^[3,4];Yen 等人提出的用最大相关性原则取代一般常用的最大熵原则来选取阈值的方法^[5];Sahoo 等提出的用 Renyi 熵代替常规熵原则的最大熵原则^[6]等.在多阈值选取方面,Cheng 等人也成功地将模糊集理论应用于最大熵原理,提出了模糊最大熵的概念,并进行了有益的尝试^[7].但上述方法不同程度地存在着执行效率低、计算量大等问题,而用传统的寻优方法又易陷入局部最优解的问题.为此,本文提出了一种阈值自动选取的方法,并给出了该方法的严格推导过程和算法实现.

1 最大熵原理

众所周知,对一幅图象按阈值进行分割时,人们普遍关心的问题之一是,图象阈值的选取是否合理和有多少原图象信息被保留.由于信息论中最大 Shannon 熵原理是解决此类问题的有效方法之一,因此,讨论本文算法前,先简单介绍熵的有关知识.

设 X 是取有限个值的随机变量, $p_i = P\{X = x_i\}$, $i = 1, 2, \dots, n$, 则 X 的 Shannon 熵定义为^[8]

$$H(X) \cong \sum_{i=1}^n p_i \log_a \frac{1}{p_i} = - \sum_{i=1}^n p_i \log_a p_i \quad (1)$$

其中, $p_i \geq 0$, $\sum_{i=1}^n p_i = 1$, 对数底 a 可为任何正数,在计算时,通常采用自然对数底,并规定当 $p_i = 0$ 时, $p_i \log_a (1/p_i) = 0$. 式(1)给出的随机变量 X 的熵 $H(X)$, 是作为与 X 对应的随机场的不确定性度量,它表示 X 包含的信息量.因 X 的熵 $H(X)$ 是代表其概率分布 p_1, p_2, \dots, p_n 的函数,所以也记为

$$H(X) = H(p_1, p_2, \dots, p_n) \cong - \sum_{i=1}^n p_i \log p_i \quad (2)$$

Shannon 熵具有如下一些基本性质:

(1) $H(p_1, p_2, \dots, p_n) \leq \log n$, 其中等号成立, 当且仅当 $p_i = 1/n, i = 1, 2, \dots, n$. 这表明等概率场具有

最大熵;

(2) $H(p_1, p_2, \dots, p_n) \geq 0$, 其中等号成立, 当且仅当对某 $i, p_i = 1, p_k = 0 (k \neq i)$. 这表明确定场的熵最小;

(3) $H(f(X)) \leq H(X)$, 这里 X 的函数 $Y = f(X)$ 是决定基本事件 $\{x_1, x_2, \dots, x_n\}$ 的一个分类. 这表明随机场分辨率越“模糊”, 信息量越少;

(4) 熵具有下列可加性

$$H(q_{11}, q_{12}, \dots, q_{1k_1}, q_{21}, \dots, q_{n1}, \dots, q_{nk_n}) = H(p_1, p_2, \dots, p_n) + \sum_{i=1}^n p_i H \left| \frac{q_{i1}}{p_1}, \dots, \frac{q_{ik_i}}{p_i} \right| \quad (3)$$

其中, $q_{i,j} \geq 0, p_i = \sum_{j=1}^{k_i} q_{i,j} > 0, \sum_{i=1}^n p_i = 1$.

由信息论中熵的唯一性定理可知, 上述熵的形式是唯一确定的, 即所定义的熵, 其表达式是唯一的. 在实际应用中, 上述 Shannon 熵却存在一定的局限性, 它仅仅适用于有限场的情况, 而对连续型随机变量的场合根本不适用. 在用于图象分割时, 恰好图象的灰度分布满足有限场分布的情况, 即其灰度值分布为有限个. 这也正是 Shannon 熵原理在图象分割阈值的选取过程中广泛应用的原因之一.

2 本文方法及算法

上节给出了 Shannon 熵的唯一表达形式及它的一些基本性质. 其中, 由第 1 条性质可知, 等概率场具有最大熵, 这也就是说, 当 $p_1 = p_2 = \dots = p_n, \sum_{i=1}^n p_i = 1$ 时, 所得到的 Shannon 熵最大. 当然这可由以下推导进一步说明:

由式(2)及条件 $\sum_{i=1}^n p_i = 1$, 用 Lagrange 乘子法可得到如下等式

$$H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \log p_i + \mu \left| \sum_{i=1}^n p_i - 1 \right| \quad (4)$$

为了与图象的一般表达式接近, 在此令 $H(A) = H(p_1, p_2, \dots, p_n); p(x_i) = p_i, i = 1, 2, \dots, n$, 其中, x_i 为图象象素的灰度级, $p(x_i)$ 为第 x_i 灰度级存在的概率, A 表示图象 A , 则式(4)又可写为

$$H(A) = - \sum_{i=1}^n p(x_i) \log p(x_i) + \mu \left| \sum_{i=1}^n p(x_i) - 1 \right| \quad (5)$$

因式(5)存在极值, 故极值点可由下式求得

$$\frac{\partial H(A)}{\partial p(x_i)} = - \log p(x_i) - 1 + \mu = 0 \quad (6)$$

由式(6)可以得到其中的一个解为 $p(x_i) = e^{\mu-1} = c$, 对任意 $i = 1, 2, \dots, n$, c 为常数. 由此可知, 只有当 $p_1 = p_2 = \dots = p_n = 1/n$ 时, 才能使熵 $H(\mathbf{A})$ 达到极大值. 这与 Shannon 熵的基本性质(1)相吻合.

利用上述推导得出的结论, 本文提出了一种能够自动确定图象阈值的方法. 该方法不仅适用于单阈值求取, 而且在多阈值自动求取时, 也同样有效.

定义 1 令 \mathbf{A} 表示一幅图象中象素点的集合, $p_i (i = 0, 1, \dots, 255)$ 表示第 i 个灰度级的概率 ($p_i =$ 第 i 个灰度级象素数和/图象的象素总和), 若将图象分为 N 个子块, 每个子块表示为 $\mathbf{A}_j (j = 0, 1, \dots, N-1)$, 则目标函数定义如下

$$H(\mathbf{A}, N) = \sum_{j=0}^{N-1} \left\| p(\mathbf{A}_j) - \frac{1}{N} \right\| \quad (7)$$

这里, $p(\mathbf{A}_j) = \sum_{i \in \mathbf{A}_j} p_i$, 即 $p(\mathbf{A}_j)$ 是在 \mathbf{A}_j 子块内各灰度级的概率之和. $\|\cdot\|$ 为范数.

对式(7)中 $H(\mathbf{A}, N)$ 取极小, 就可求得图象的阈值, 因为由式(7)可以看出, 当 $p(\mathbf{A}_j) = 1/N$ 时, $H(\mathbf{A}, N)$ 有最小值 0, 由熵的基本性质(1)可知, 此时的 Shannon 熵最大, 也就是说, 此时分割得到的图象按最大熵原理最合理. 进一步讲, $H(\mathbf{A}, N)$ 越接近 0, 分割后的图象保留原图象的信息越多, 这与利用最大熵原理求取阈值所得到的结果完全一致. 从这一点来看, 尽管在图象分割阈值的自动选取过程中, 式(7)与式(2)是完全等价的, 但是, 两式阈值自动选取过程的计算效率却相差很大. 前者在选取阈值时, 尽管采用了比较快速的优化算法, 例如目前采用较多的遗传算法、混沌优化等, 从而使计算效率和计算精度得到一定程度的提高, 但是由于受目标函数的形式限制, 要大幅度提高计算效率和计算精度将有赖于优化算法的进一步改进; 后者优化的目标函数非常直观、简单, 由于其与图象分割的子块个数 N 有关, 因此仅用不大于 $N(N-1)$ 次运算, 即可完成寻优过程, 即便是使用穷举搜索法, 计算量也非常小, 这在以下的算法实现中可以看到.

本文算法实现如下:

- ① 初始化 N ;
- ② 读取图象数据, 并求出图象的灰度直方图;
- ③ 计算 $p_i, i = 0, 1, \dots, 255$;
- ④ 计算 $p(\mathbf{A}_j) \geq 1/N$ 的 i , 并取此时的 i 为准阈值 $T(j) = i, j = 0, 1, \dots, N-1$;
- ⑤ 判断 $\sum_{j=0}^{N-1} \|p(\mathbf{A}_j) - 1/N\|$ 是否最小, 若是最

小, 则 $T(j)$ 为阈值, 计算结束;

⑥ $T(j) = i-1$ 返回 ⑤

这里有点需要进一步说明:

(1) 关于如何确定多阈值图象分割类数 N 的问题, 一般存在如下两种情况: 一是在分割前利用先验知识, 确定分割类数; 二是在分割前不知道需要分割的类数或者很难确定需要分割的类数, 此时则需要算法能够自动地确定最佳分类个数, 然后进行分类. 据有关文献介绍, 对于一维直方图分类问题, 采用势函数聚类方法较为简单^[9]. 由于文中算法的讨论并未涉及到 N 的选取, 因此可预先设定, 当然也可通过势函数聚类方法求得. 因分割类数确定与本文算法相互独立, 且对算法的收敛速度无影响, 故文中算法的计算量也不大于 $N(N-1)$. 在对比实验过程中, 为了便于比较, 因而设定 $N = 3$ 和 $N = 5$ 两种方案;

(2) 由于在实验过程中, 所处理的是灰度图象, 并且是利用图象的一维灰度直方图信息, 故此时 $\|p(\mathbf{A}_j) - 1/N\|$ 应取为 $|p(\mathbf{A}_j) - 1/N|$.

3 实验结果及分析

为了检验本文算法的效果, 用几幅图象对本文提出的算法进行了实验. 图 1(a)、图 2(a) 为原始图象, 图 1(b)、图 2(b) 为图象的灰度直方图, 每幅图象有 256 个灰度级, 最暗的是 0, 最亮的为 255.

3.1 多阈值自动选取实验

这里, 实验结果主要与用文献[7]算法选取的阈值做对比, 目的是要说明本文算法具有较高的稳定性. 在所进行的 50 次对比实验中, 本文算法每次均能得到相同的阈值, 而用文献[7]的算法计算出的阈值总会有 2~3 次与其他阈值存在较大偏差, 其原因可能是由于在寻优过程中, 优化算法陷入局部极值点所致. 这同时也表明了本文算法在稳定性方面要好于文献[7]算法. 另外, 通过实验还发现用文献[7]算法获得的 40 多次分割较成功的阈值均在本文得到的阈值附近, 而且随熵 $H(a, b, c)$ 越趋于 1, 与本文结果越接近, 这也表明本文算法有较高的精度.

为了便于比较, 实验中, 取 $N = 3, N = 5$ 两种情况, 所计算获得的阈值分别列于表 1、表 2 中, 其中表中还列出了用文献[7]算法获得的结果, 即 $H(a, b, c)$ 最接近于 1 的阈值. 表中一些符号的意义在文献[7]中有定义, 这里就不再说明.

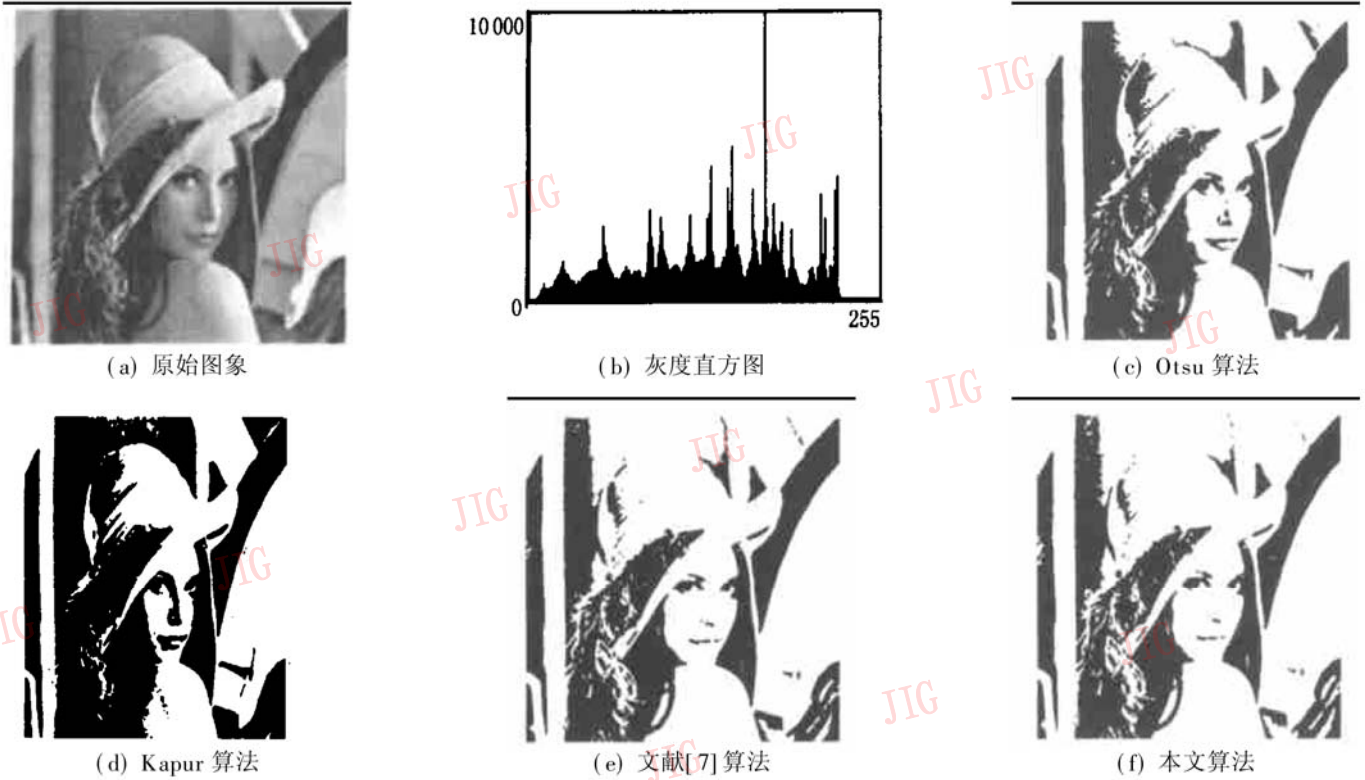


图1 用不同算法得到的阈值分割后的二值化Lena 图象

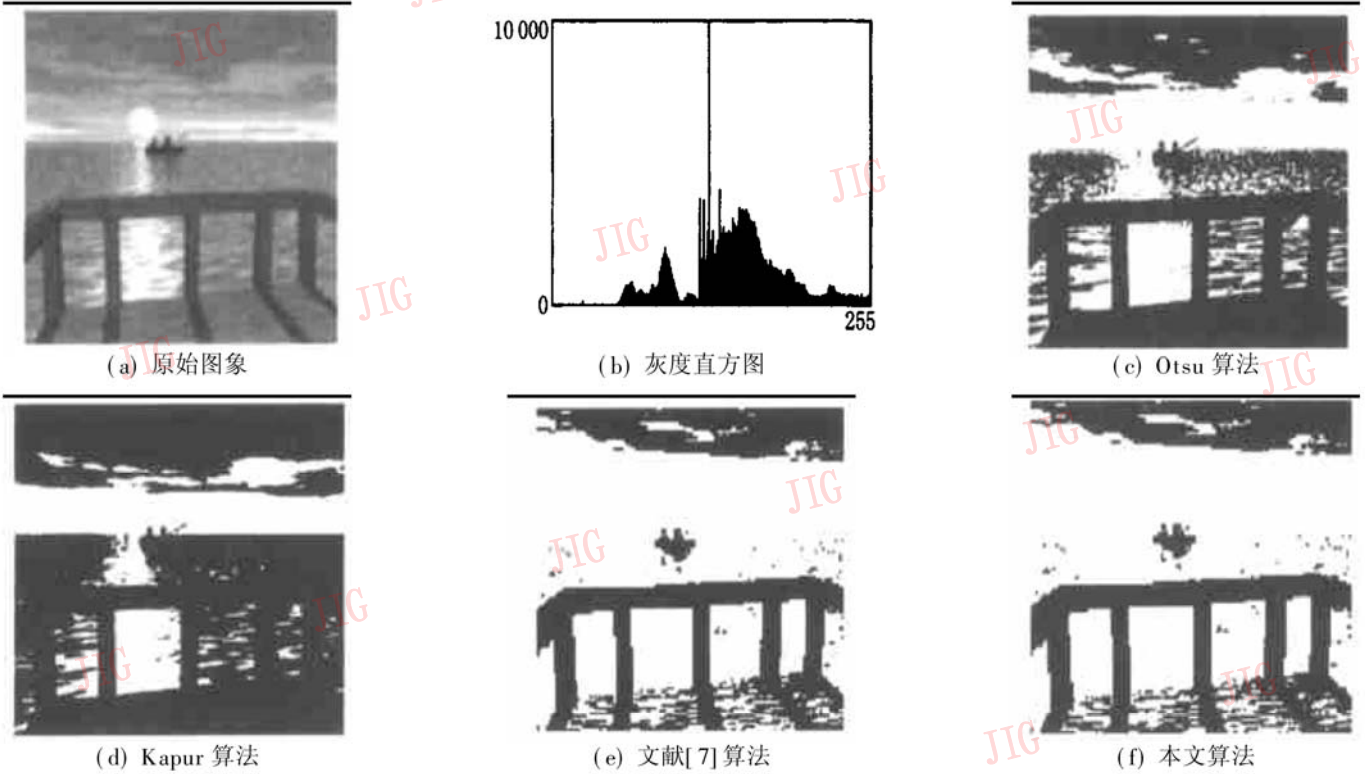


图2 用不同算法得到的阈值分割后的二值化钓鱼图象

表1 $N=3$ 时本文算法与文献[7]算法计算得到的阈值 T 对比

灰度级范围	文献[7]算法结果			本文算法结果		
	(a, b, c)	$H(a, b, c)$	T	$H(A, N)$	T	
图1	11~ 223	(17, 203, 245)	0.9979436	(101, 157)	0.004804	(99, 159)
图2	23~ 255	(80, 123, 219)	0.9966194	(125, 153)	0.014504	(124, 154)

表2 $N=5$ 时本文算法与文献[7]算法计算得到的阈值 T 对比

灰度级范围	文献[7]算法结果			本文算法结果		
	(a, b, c)	$H(a, b, c)$	T	$H(A, N)$	T	
图1	11~ 223	(21, 236, 253)	0.9986145	(76, 117, 149, 176)	0.011527	(70, 114, 146, 176)
图2	23~ 255	(99, 101, 229)	0.9870069	(114, 130, 148, 172)	0.008433	(112, 131, 148, 176)

3.2 图象的分割实验

在进行图象分割时,不仅用本文算法对图象进行二值化处理,以考察本文算法的合理性,而且在进行图象分割实验中,还将本文算法得到的结果分别与 Otsu 算法、Kapur 算法以及文献[7]算法的分割结果进行了比较,其用各种算法计算得到的最优二值化阈值列于表3,用不同算法得到的阈值分割后的二值化图象分别列于图1(c)~(f)及图2(c)~(f)。

表3 各种算法最优阈值

	Otsu 算法	Kapur 算法	文献[7]算法	本文算法
图1	133	148	122	119
图2	152	164	139	139

由分割后的二值图可见,用本文与文献[7]算法得到的阈值,其分割后的结果非常接近,而且视觉效果也较好,而用 Kapur 算法得到的阈值分割后图象的视觉效果不及 Otsu 算法理想。尤其是图2,用本文和文献[7]算法获得的阈值,其分割后的图象上,小船及钓鱼人轮廓清晰可见,但是由于 Otsu 算法选取的阈值比文献[7]及本文算法选取的阈值大,图象上水的反光却看不到,故结果不如图2(c)生动。

3.3 算法效率分析

通过本文算法与文献[7]中算法迭代次数的比较来进行效率分析,表4中所列出的结果均为50次计算结果迭代次数的平均值。

表4 不同算法多阈值自动选取时迭代次数的比较

	N=3		N=5	
	文献[7]算法	本文算法	文献[7]算法	本文算法
图1	1207	5	3763	8
图2	2289	4	4163	6

从表4可以看到,本文算法在阈值选取时,所需的迭代次数明显较文献[7]中的方法减少了很多,若考虑到优化过程中的其他差异(如本文算法中,目标函数的计算仅是加减运算,而文献[7]中的目标函数的计算为乘法和取对数运算),本文算法的计算时间将比文献[7]中算法减少得更多。其原因是,由于文献[7]在寻优过程中,采用的是遗传算法(许多优化算法也存在同样情况),其计算效率在一定程度上取决于随机产生的初始群体的质量,而本文算法则不需要初始群体,其最大迭代次数也不大于 $N(N-1)$,这可能是本文算法效率提高的原因所在。

4 结论及展望

现存的许多自动选取图象分割阈值的算法均是

非常有效的,特别是近年来发展的基于信息论熵理论的阈值自动选取算法,在很多应用中均得到了验证。本文提出了一种新的自动选取图象分割阈值的算法,不仅克服了现存的熵理论选取阈值算法的一些不足,同时在最大 Shannon 熵原理的基础上,对该算法进行了严格的推导,并给出了具体的优化策略及算法实现。由于式(7)形式上直观、简洁,算法实现简单、明了,仅需将图象的灰度直方图尽可能等分即可,故计算效率大幅提高,同时计算稳定性也明显提高。从对大量图象所进行的比较实验可见,本文算法可获得令人满意的结果(由于篇幅限制,文中仅以2幅图象为例进行演示说明)。

本文对算法的讨论未涉及到图象分块个数 N 的选取,如何利用文中提出的算法自适应地确定图象的分块个数和阈值,将是作者下一步研究的方向之一。

参考文献

- Otsu N. A threshold selection method from grey-level histograms[J]. IEEE Trans. System. Man Cybernet., 1979, SMC-9: 62~ 66.
- Kittler J, Illingworth J. Minimum error thresholding [J]. Pattern Recognition, 1986, 19(1): 41~ 47.
- Pal N R, Pal S K. A review on image segmentation techniques [J]. Pattern Recognition, 1993, 26(9): 1277~ 1294.
- Pun T. A new method for grey-level picture thresholding using the entropy of the histogram[J]. Signal Process., 1980, 2(3): 223~ 237.
- Yen J G, Chang F J, Chang S. A new criterion for automatic multilevel thresholding[J]. IEEE Trans. on Image Processing, 1995, 4(3): 233~ 260.
- Sahoo P, Wilkins C, Yeager J. Threshold selection using Renyi's entropy[J]. Pattern Recognition, 1997, 30(1): 71~ 84.
- Cheng H D, Chen Y H, Sun Y. A novel fuzzy entropy approach to image enhancement and thresholding[J]. Signal Processing, 1999, 75(3): 277~ 301.
- 孟庆生. 信息论[M]. 西安: 西安交通大学出版社, 1989: 19~ 27.
- 裴继红, 谢维信. 势函数聚类自适应多阈值图象分割[J]. 计算机学报. 1999, 22(7): 758~ 762.



曹力 1967年生,1989年获得学士学位,1996年获得硕士学位,现在西北工业大学自动控制专业攻读博士学位。目前主要的研究方向为图象处理、信号处理、遗传算法、智能控制、神经网络、模糊控制等。

史忠科 1956年生,1991年获西北工业大学博士学位。现为西北工业大学自动控制系教授、博士生导师。目前的研究领域为估计、辨识方法、数字信号处理、非线性系统、 H^∞ 控制、鲁棒控制、智能控制及交通控制等。