

食用油红外光谱的分形维数特征

孙翹楚^{1,2}, 肖创柏¹, 郭 犇¹, 苏 峰¹

(1. 北京工业大学 计算机学院, 北京 100124, sunqiaochu@hotmail.com; 2. 中国农业银行总行软件开发中心, 北京 100073)

摘 要: 以傅立叶变换红外光谱为研究对象, 探讨了图像分形理论在傅立叶变换红外光谱识别中的应用, 提出一种利用分形理论提取食用植物油红外光谱特征的算法. 通过对傅立叶变换红外光谱分形维数的比较, 达到区分不同光谱的目的, 利用开源计算机视觉库 OpenCV 实现了食用植物油红外光谱分形维数的计算算法, 以玉米油和芝麻油的红外光谱为例对该方法进行了说明. 实验表明: 利用分形维数特征可以对玉米油和芝麻油的真伪进行很好的鉴别.

关键词: 红外光谱特征提取; 图像分形; 盒维数; OpenCV

中图分类号: TP391

文献标志码: A

文章编号: 0367-6234(2010)11-1814-04

Fractal geometry feature of edible oil FTIR

SUN Qiao-chu^{1,2}, XIAO Chuang-bai¹, GUO Ben¹, Su Feng¹

(1. College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China, sunqiaochu@hotmail.com;

2. Software Development Center of Agricultural Bank of China, Beijing 100073, China)

Abstract: A feature extraction of edible oil FTIR based on fractal geometry is proposed and theoretical calculation formulas of fractal dimension are introduced in this paper. The calculation method of box-counting dimension based on FTIR is analyzed. Through the comparison of FTIR fractal dimensions, the different spectral could be distinguished. Based on the open source computer vision library OpenCV, the FTIR of edible vegetable oils fractal dimension calculation algorithm has been finished. Taking the corn and sesame oil as examples. The detecting result indicates that the FTIR of corn oil and sesame oil can be discriminated effectively by the difference of box-counting dimension.

Key words: feature extraction of FTIR; fractal geometry; box dimension; OpenCV

近年来, 红外光谱技术在食品鉴别方面的应用研究发展迅速. 红外光谱技术的优势在于样品几乎不需要分离处理, 从而使得鉴别过程更加迅速^[1]. 同时根据样品成分的红外吸收率, 红外光谱技术能够提供整体信息来鉴别样品. 红外光谱根据不同的波数范围分为近红外(12 900 ~ 4 000 cm^{-1})、中红外区(4 000 ~ 650 cm^{-1})和远红外区(650 ~ 10 cm^{-1})3个光区. 其中, 中红外区是应用最早和最广的一个区, 大多数有机化合物和无机离子的基频振动吸收都落在中红外区内.

化合物分子中所含的化学键或官能团不同, 或官能团所处的化学环境不同, 其振动能级从基态跃迁到激发态所需能量不同会吸收不同波长的红外光, 形成不同的红外光谱图, 而相关吸收峰的高度与其含量成正比. 因此, 红外光谱法被广泛用于有机物的定性和定量研究. 由于傅里叶变换红外光谱法(Fourier transform infrared spectroscopy, FT-IR)具有整体、快速、无损鉴定复杂体系多种组份的特点, 目前已在食品鉴别领域获得了广泛应用^[2-3].

刘沐华等^[4]通过对几种中药的傅里叶变换红外光谱进行求一阶导数对其进行区分. 但是在本文初期实验中, 这种方法并不适合食用植物油的红外光谱. 对其求一阶导数并不足以对其区分. 同时对各个光谱的指纹区进行了积分操作, 提取

收稿日期: 2010-04-01.

基金项目: 国家高技术研究发展计划基金项目(2009AA012437);
北京市自然科学基金资助项目(4072004, 4092006).

作者简介: 孙翹楚(1984—), 男, 硕士;

肖创柏(1962—), 男, 教授, 博士生导师.

的特征差别细微,不足以作为区分的标准. 贾素贞等^[5]通过对几种食用植物油的二维傅里叶变换红外光谱进行特征提取,对其进行区分. 这种方法可以很好的放大一维光谱的差异,但是计算复杂,对设备要求较高. 本文提出一种利用分形理论提取食用植物油红外光谱特征的算法,将此作为一种主要特征引入食用油红外光谱模式识别中,在实验中取得了较好效果.

1 基于分形理论的红外光谱特征提取

1.1 盒维数定义

分形学的概念的实质是除传统几何学以外的不规则几何体在标度变换下的自相似性.

- 若 F 是分形集,则其应具有以下性质^[6]:
- 1) F 具有精细结构,即有任意小比例细节;
 - 2) F 极不规则,以致其整体与局部都无法用传统几何语言描述;
 - 3) F 通常具有某种自相似性;
 - 4) F 的分形维数大于其拓扑维数;
 - 5) F 可以用非常简单的方法定义,可以由迭代产生.

一般而言若研究对象符合上述性质中的全部或大多数,即使有某个性质的例外,也并不影响称其为分形.

计盒维数或称盒维数 (Box-counting or Box dimension) 是应用最广泛的维数之一,它的普遍应用主要由于这种盒维数的数学计算以及经验估计相对容易.

盒维数的定义为: 设 F 为 R^n 中任一非空子集,记 $N(F, \delta)$ 表示最大直径为 δ 且能覆盖 F 的集合的最小数,则 F 的上下盒维数分别定义为

$$\overline{\dim}_B F = \lim_{\delta \rightarrow 0} \frac{\ln N(F, \delta)}{-\log \delta}. \tag{1}$$

$$\underline{\dim}_B F = \lim_{\delta \rightarrow 0} \frac{\ln N(F, \delta)}{-\log \delta}. \tag{2}$$

如果上下维数相等,则称这个共同的值 F 的计盒维数或者盒维数,定义为

$$\dim_B F = \lim_{\delta \rightarrow 0} \frac{\ln N(F, \delta)}{-\log \delta}. \tag{3}$$

盒维数存在一些等价定义,有时这些定义更适合应用. 通常为以下五种^[7]:

- 1) 覆盖 F 的半径为 δ 的最小闭球数;
- 2) 覆盖 F 的边长为 δ 的最小立方块数;
- 3) 相交于 F 的 δ -网格块数;
- 4) 覆盖 F 的直径至多为 δ 的最小集合数;
- 5) 球心在 F 中半径为 δ 的互不相交球的最大数.

本文利用定义(3) 对红外光谱的分形维数进行计算.

1.2 盒维数的计算

设 $A \in H(X)$, 其中, (X, d) 是度量空间, $\varepsilon_n = cb^n, 0 < b < 1, c > 0$ 和 $n = 1, 2, 3, \dots$.

如果

$$D = \lim_{n \rightarrow \infty} \frac{\ln N(A, \varepsilon_n)}{\ln \frac{1}{\varepsilon_n}}. \tag{4}$$

则 A 具有分形维数 D . 在式(4) 中用 γ 代替 ε , 计算分形维数的公式为

$$D = \lim_{\gamma \rightarrow 0} \frac{\ln N(A, \gamma)}{-\ln \gamma}. \tag{5}$$

故可让 $\ln N(A, \gamma)$ 与 $\ln \gamma$ 成为一对变量, 则式(5) 变为

$$\ln N(A, \gamma) = -D(A) \ln \gamma. \tag{6}$$

于是 $D(A)$ 就成了 $(\ln \gamma, \ln N(A, \gamma))$ 坐标系中的斜率^[8]. 本文借以不同红外光谱盒维数的不同, 对其所代表的食用油进行区分.

1.3 红外光谱分形维数计算方法

红外光谱在计算机中以图像的格式存储, 文中使用 Intel 公司开发的 OpenCV 开源计算机视觉库, 实现了红外光谱的分形维数计算. 具体其分形维数计算过程如图 1 所示.

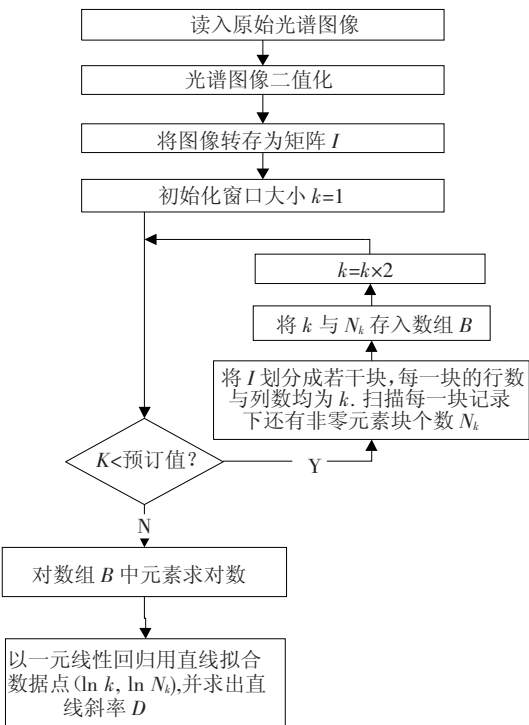


图 1 红外光谱盒维数计算流程图

1) 首先读入光谱图像, 对光谱图像进行二值化处理^[9]. 原始红外光谱图像为彩色图像, 进行盒维数计算时, 应先将彩色图像转换成二值图像 (如图 2 所示). 利用 OpenCV 的 on_trackbar() 函

数生成滑动条,动态设定阈值,阈值的取值以能显示全部光谱图像为准.然后利用 `cvThreshold()` 函数将读入图片二值化.包含光谱的像素点值为 0.

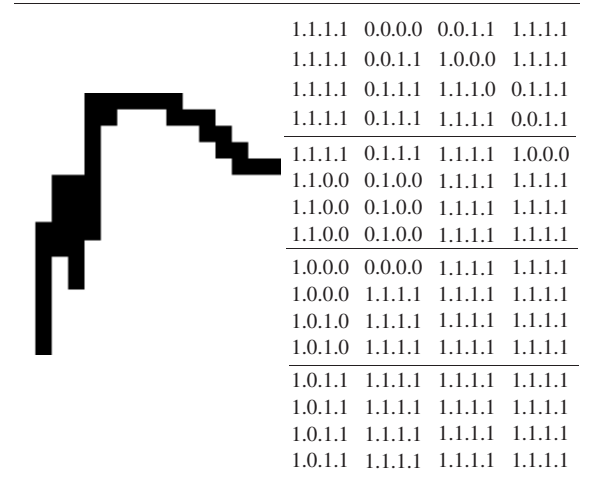


图 2 二值化后的光谱图像(局部)及其对应的二值数据

2) 将得到的二值图像转存为矩阵.运用 `cv-Convert()` 函数将二值图像转存为与其同尺寸的矩阵 I .

3) 把得到的矩阵依次划分为若干块,每一块的行数与列数均为 k , k 取 $1, 2, 4, \cdots, 2^n$, 运用 `cvGetReal2D()` 函数扫描每个划分块,把包含像素 0 的块的个数记为 N_k .这样就得到 k 与 N_k 的对应数组 $B^{[10]}$.

4) 对数组 B 中的元素求对数.并在双对数坐标平面中,以一元线性回归用直线拟合数据点 $(\ln k, \ln N_k)$, $k = 1, 2, 4, \cdots, 2^n$, 所得到的斜率的负值 D 就是该红外光谱的盒维数.

2 实验结果及分析

2.1 算法正确性检验

通过实验发现,分形维数的正确性与 k 的取值以及图像的分辨率有关.在实验中取 $1, 2, 4, 8, 16, 32, 64, 128, 256$, 这 9 个值对图像计算分形.

根据盒维数的定义,所取 k 的尺寸越小,所计算出的分形维数越精确.但由于对存储在计算机中的图像进行分形计算,只能分割到像素点尺寸.因此计算分形算法所得出的结果与实际分形维数之间存在误差.而误差的大小直接影响到特征提取的准确性,若计算出的误差大于光谱分形维数的差异,则无法利用此差异对光谱进行区分.为验证分形维数计算算法的准确性和可靠性,分别用直线、矩形、Koch 曲线、Cantor 三分集、Sierpinski 三角垫和 Sierpinski 地毯等已知维度的分形图形对算法进行验证.结果如表 1 所示.

表 1 分形维数计算结果

分形图形	理论值	计算值	差值	相对误差/%
直线	1.000 0	0.999 4	0.000 6	0.600 3
矩形	2.000 0	1.997 5	0.002 5	0.125 1
Koch 曲线	1.261 9	1.258 8	0.003 1	0.246 3
Cantor 三分集	0.630 9	0.643 1	-0.012 2	1.897 1
Sierpinski 三角垫	1.585 0	1.573 1	0.011 9	0.756 5
Sierpinski 地毯	1.892 8	1.895 1	-0.002 3	0.121 4

由实验结果可知,误差在小数点后三位到四位之间,而光谱分形维数的差异在小数点后两位.因此,误差不会对特征提取造成影响.

2.2 实验分析

采用本文的算法,本文对市面上常见的福临门、金龙鱼、欧尚等品牌的几种食用油的红外光谱的分形维数进行了计算.通过对每种食用油各 30 个样本的盒维数计算(抽样结果如表 2 所示,其中, k 为窗口大小, N_k 为盒子数),可知食用油的盒维数总体在 $1.293\ 1 \sim 1.315\ 3$ 之间.可利用此特征对食用油同其他液体进行区分.另外通过对玉米油和芝麻油的红外光谱分形维数的计算得到,玉米油的红外线光谱的分形维数在 $1.293\ 1 \sim 1.297\ 2$ 之间,而芝麻油的红外光谱分形维数在 $1.311\ 3 \sim 1.315\ 3$ 之间,分形维数差别明显,因此可利用此特征对两种植物油的真伪进行鉴别.

表 2 食用植物油红外光谱盒维数计算结果

N_k	k									维数
	1	2	4	8	16	32	64	128	256	
福临门玉米油(a)	25 651	10 741	4 744	2 128	923	397	154	48	18	1.293 9
福临门玉米油(b)	25 393	10 622	4 632	2 022	922	391	153	48	18	1.293 2
金龙鱼玉米油(a)	25 316	10 695	4 768	2 158	927	398	154	48	18	1.293 1
金龙鱼玉米油(b)	25 541	10 730	4 752	2 130	925	401	154	48	18	1.293 3
古船玉米油(a)	25 481	10 732	4 755	2 134	922	399	153	48	18	1.293 6
古船玉米油(b)	25 300	10 658	4 745	2 159	960	418	157	48	18	1.290 2
欧尚玉米油	25 333	10 610	4 710	2 127	953	409	156	48	18	1.291 1
古币芝麻油	28 316	12 034	5 057	2 152	932	401	154	48	18	1.312 0
长康芝麻油	28 114	12 020	5 070	2 195	958	401	155	48	18	1.314 2
古船芝麻油	28 136	12 034	5 057	2 152	932	401	154	48	18	1.314 3
福临门芝麻油	27 798	11 893	5 011	2 136	957	415	155	48	18	1.310 6
金龙鱼芝麻油	29 140	12 402	5 182	2 171	934	400	156	49	19	1.313 8
Wangjiuhe 芝麻油(a)	28 157	12 035	5 084	2 184	977	415	155	48	18	1.313 7
Wangjiuhe 芝麻油(b)	27 830	11 889	5 009	2 171	961	401	156	48	18	1.311 3

3 结 论

1) 分形维数能有效地表现红外光谱的复杂度以及自相似性特征.

2) 分形维数的准确性与源图像的分辨率有关, 在实验中采用的光谱图像分辨率为 1280×880 . 随着仪器设备的不断进步, 更大分辨率的光谱图像可以得到更为准确的分形维数特征.

3) 通过引入其它特征, 如谱峰高度、光谱导数、面积等特征, 可以对其它食用油进行更好的区分.

4) 通过对市面上几种常见食用油的红外线光谱的分形维数的计算表明, 利用分形维数可以对玉米油以及芝麻油的真伪进行较好的区分.

参考文献:

- [1] Van de VOORT F R, GHETLER A, GARCIA-GONZA'LEZ D L. Perspectives on quantitative mid-ftir spectroscopy in relation to edible oil and lubricant analysis: Evolution and integration of analytical methodologies [J]. Food Analytical Methods, 2008, 3(1):153-163.
- [2] MILLS B L, Van de VOORT F R, KAKUDA Y. The quantitative analysis of fat and protein in meat by transmission infrared analysis [J]. Meat Science, 1984, 11(4): 253-262.
- (上接第 1813 页)
- [4] 綦志刚. 船舶零航速减摇鳍升力机理及系统模型研究[D]. 哈尔滨: 哈尔滨工程大学, 2008.
- [5] 金鸿章, 张小飞, 罗延明. 零航速减摇鳍升力模型研究[J]. 海洋工程, 2007, 25(3):83-87.
- [6] KHALIL H K. Nonlinear systems [M]. NJ, USA: Prentice Hall, 2002.
- [7] TAO G, KOKOTOVIC P. Adaptive control of systems with actuator and sensor nonlinearities [M]. New York: Wiley, 1996.
- [8] NINOMIYA T, YAMAGUCHI I, KIDA T. Feedback control of plants driven by nonlinear actuators via input-state linearization [J]. Journal of Guidance, control and dynamics, 2006, 29(1): 20-24.
- [9] BATTILOTTI S, CALIFANO C. A constructive condition for dynamic feedback linearization [J]. Systems & control letters, 2004, 52:329-338.
- [10] HSU K C. Sliding mode control for uncertain nonlinear systems with multiple inputs containing sector nonlinearities and deadzones [J]. IEEE transactions on systems, man, and cybernetics-part B: cybernetics, 2004, 34(1):374-380.

- [3] Van de VOORT F R, SEDMAN J, ISMAIL A A. A rapid FTIR quality control method for determining fat and moisture in high-fat products [J]. Food Chemistry, 1993, 48(2):213-221.
- [4] 刘沐华, 张学工, 周群, 等. 近红外漫反射光谱法和模式识别技术鉴别中药材产地[J]. 光谱学与光谱分析, 2006, 26(4):629-632.
- [5] 贾素贞, 鲍猛, 聂传平. 几种食用油的二维红外相关分析[J]. 济南大学学报, 2007, 21(2): 124-129.
- [6] BRIGGS J. Fractals: The Patterns of Chaos [M]. London: Thames and Hudson, 1992:148.
- [7] FALCONER K. Techniques in Fractal Geometry [M]. NJ, USA: John Wiley and Sons, 1997:59-61.
- [8] 张平, 王新柯, 李海涛, 等. 基于分形理论的太赫兹光谱识别[J]. 量子电子学报, 2007, 24(6):672-676.
- [9] 文志英. 分形几何的数学基础 [M]. 上海: 上海科技教育出版社, 2002.
- [10] CHAUDHURI B B, SARKAR N. Texture segmentation using fractal dimension [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(1): 72-77.
- [11] 王韶娟, 曾国荪. 分形维数的一个并行算法 [J]. 计算机应用与软件, 2005, 22(10):19-20.

(编辑 张 红)

- [11] CHANG W D. Robust adaptive single neural control for a class of uncertain nonlinear systems with input nonlinearity [J]. Information sciences, 2005, 171:261-271.
- [12] 金鸿章, 王科俊. 智能技术在船舶减摇鳍系统中的应用 [M]. 北京: 国防工业出版社, 2003.
- [13] SGOBBO J N, PARSONS M G. Rudder/fin roll stabilization of the USCG WMEC 901 class vessel [J]. Marine Technology, 1999, 36(3):157-170.
- [14] CHEN C T, CHANG W D. A feedforward neural network with function shape autotuning [J]. Neural networks, 1996, 9(4):627-641.
- [15] CUI X, SHIN K G. Direct control and coordination using neural networks [J]. IEEE Transactions on systems, man, and cybernetics, 1993, 23(3):686-697.
- [16] PEREZ T. Ship motion control: course keeping and roll stabilisation using rudder and fins [M]. London: Springer, 2005.

(编辑 张 宏)