

doi:10.3969/j.issn.1001-3539.2014.05.018

基于近红外的 Fisher 判别法鉴别废塑料^{*}

张毅民¹, 白家瑞¹, 刘红莎¹, 汤桂兰², 胡彪³

(1. 教育部绿色合成与转化重点实验室, 天津大学, 天津 300072; 2. 天津子牙环保产业园有限公司, 天津 301605;

3. 天津理工大学, 天津 300191)

摘要: 通过对化学计量学方法的分析研究, 提出了一种基于近红外光谱的丙烯腈-丁二烯-苯乙烯塑料、聚对苯二甲酸乙二酯、聚氯乙烯、聚丙烯、聚苯乙烯和聚乙烯 6 种塑料的一次性鉴别模型, 为近红外塑料识别软件的二次开发提供有效程序。结果表明, 光谱经 3 次多项式、13 窗口 S-G 最小二乘拟合平滑+sym17 函数、分解 2 层小波分析+主成分分析+特征波长选择+Fisher 判别处理, 可得 5 个判别函数式。由这些函数式建立的判别模型可以实现 6 种塑料的一次性识别, 其校正集样本自身验证和交叉验证的识别率分别为 100% 和 84.9%, 表明该模型稳定; 预测集样本进行模型外部检验的准确率为 100%, 表明该模型可行。

关键词: 近红外光谱; 废旧塑料; 鉴别; Fisher 判别

中图分类号: TO320.9 **文献标识码:** A **文章编号:** 1001-3539(2014)05-0075-05

Recognition of Plastics Wastes with Fisher Discriminant Classification Based on Near Infrared Spectrum

Zhang Yimin¹, Bai Jiarui¹, Liu Hongsha¹, Tang Guilan², Hu Biao³

(1. Key Laboratory for Green Chemical Technology of State Education Ministry, Tianjin University, Tianjin 300072, China;

2. Tianjin Ziya Environmental Industrial Park, Tianjin 301605, China; 3. Tianjin University of Technology, Tianjin 300191, China)

Abstract : Through researching the chemometrics methods, a new method for recognition of six kinds of plastics such as ABS, PET, PVC, PP, PS and PE with Fisher discriminant classification using near infrared spectra was proposed, which provided an effective application for the secondary development of near infrared recognition software. The experimental results show that five discriminants functional expressions can be got after the spectrum processing with Savitzky-Golay smoothing in 3 polynomial and 13 windows+the wavelet analysis in layer 2 and sym17 function+principal component analysis+choose characteristic wavelength+Fisher discriminant analysis. The recognition model which is built with these functional expressions, can be effectively used to identify the six plastics. The accuracies for its validation and cross validation are 100% and 84.9% respectively in calibration set, which shows the stability of the model. The identification accuracies for the unknown samples are 100% in test set which proves the feasibility of the model.

Keywords : near infrared spectrum ; plastics wastes ; identification ; Fisher discriminant

我国“城市矿产”开发过程中普遍存在废混合塑料资源的循环利用率低、再生产品附加值低等问题^[1], 而回收塑料的首要问题是将其分类。近年来, 近红外光谱因其可高效、快速、准确地对固体粉末、颗粒等多种物质进行无损检测而得到广泛应用, 这正是废旧塑料分类的基本要求^[2-6]。因此, 近红外技术在废旧塑料的分类处理中拥有明显的技术优势。

近红外技术利用不同塑料对近红外不同波段光波的反射或透射有特征区别, 并通过化学计量学方法来将其分类识别。由于不同产地和种类的样品, 要求的近红外仪器的应用模型不同, 因此针对不同产品应有其相对应的模型。同时, 我国基于近红

外技术并结合化学计量学方法用于材料鉴别已有大量研究^[7-10]。但是, 目前对于塑料的一次分类研究甚少或准确率不高。笔者以生活中废旧电子电器产品、汽车废旧部件及生活用品等中常见的丙烯腈-丁二烯-苯乙烯塑料 (ABS)、聚对苯二甲酸乙二酯 (PET)、聚氯乙烯 (PVC)、聚丙烯 (PP)、聚苯乙烯 (PS) 和聚乙烯 (PE) 6 种塑料为研究对象, 对不同化学计量学方法进行比较分析, 提出了一种可适用于 6 种

^{*} 国家 863 计划项目 (2012AA063007)

联系人: 张毅民, 副教授、硕士生导师, 从事废弃资源和材料回收鉴别分类、绿色化学工艺研究

收稿日期: 2014-02-17

常见废旧混合塑料一次性识别的 Fisher 判别模型,为塑料识别软件的二次开发提供有效程序。

1 实验部分

1.1 原材料

实验样本包括 ABS, PET, PVC, PP, PS 和 PE 6 种塑料 94 个样品。其中,训练集(新料)73 个,有 PP 15 个,PS 15 个,ABS 9 个,PE 17 个,PET 7 个,PVC 10 个。预测集(废旧塑料)21 个,有 PP 2 个,PS 7 个,ABS 3 个,PE 6 个,PET 2 个,PVC 1 个。6 种塑料的新料来源于市售,废旧塑料来源于国内外废旧家电(电视机、冰箱、洗衣机等)、汽车和生活垃圾等。

1.2 光谱采集

采集前,样品均进行表面 360[#] 砂纸打磨、标号、剪碎处理。采集在 FT S6000 型傅立叶变换红外光谱仪(美国 Bio-rad 公司生产)上进行,使用漫反射附件,碘钨灯光源,InSb 检测器。仪器参数设定为:波数范围 11 000~3 000 cm^{-1} ,扫描速度 5 kHz,每个样品重复扫 32 次,对称扫描,分辨率为 8 cm^{-1} ,以 KBr 作为参比材料,直接在 20~25 $^{\circ}\text{C}$ 条件下测定。样品光谱预处理采用 MatlabR2010a 实现,建模使用 IBM SPSS Statistics 19 软件。

1.3 数据分析方法

基于近红外法鉴别塑料时,软件模型的建立是识别的关键。软件处理包括光谱预处理和 Fisher 判别法模型建立两个过程。首先,采用最小二乘拟合求导法进行微分处理,解决基线漂移问题,并以光谱信号的均方根误差与一阶导数数据算术平均值为标准选择多项式和窗口大小。然后,运用小波分析法进行光谱的平滑消噪,并根据信噪比大小,选择小波参数。再将处理后的光谱作为输入变量,进行主成分分析使原光谱信息维数大大降低。最后,选择载荷系数最大、最小值对应的波长为各种塑料的特征波长^[1]进行 Fisher 判别,得到判别函数并建立识别模型。

2 结果与讨论

2.1 光谱特征

6 种塑料的原始漫反射近红外光谱如图 1 所示。由图 1 可看出,在整体范围内各光谱彼此间存在很大差异(ABS 和 PS 的漫反射近红外光谱很相似,只在 1 900~2 000 nm 有细微的差别),最明显的光谱差异出现在 1 116~1 151,1 182~1 237,1 372~1 437,1 626~1 770,1 868~2 022,2 100~2 300 nm。

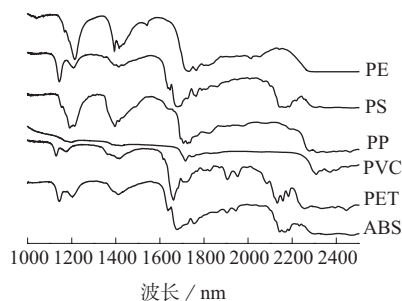


图 1 塑料原料漫反射近红外光谱图

2.2 光谱数据预处理

运用 S-G 最小二乘拟合法,既可求导又可平滑。当采用 3 次多项式、13 窗口时,所得均方根误差小、算术平均值小、观察谱图光滑,因此光谱光滑、去噪效果最佳。小波分析选择 sym17 小波母函数、分解 2 层、rigrsure 确定阈值法、sln 阈值重调法和软阈值法为参数时信噪比高,滤波平滑效果最好。利用主成分分析法可以对光谱数据进行降维,以少数主成分代替光谱全谱信息。以不同波长为自变量,从中选出包含信息量大的一组波长作为第一主成分,然后根据包含信息量大小依次确定第二、三主成分。选用第一、二主成分得分做二维散点图,由第一、二、三主成分做三维图可以直观地看到聚类分析结果,如图 2、图 3 所示。由图 2、图 3 可以看出,主成分得分聚类出现偏差,6 类塑料没有很明显地聚合为完全分离的团簇,只有 PET 明显地与其它类分开。为了得到更好的分类效果,笔者将选择特征波长进行建模。

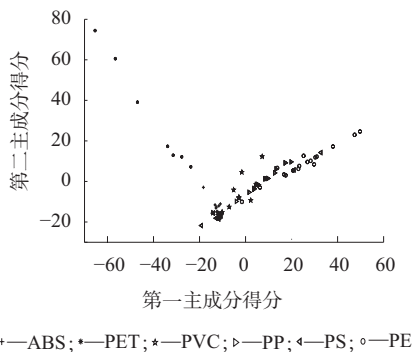


图 2 主成分得分二维图

2.3 Fisher 判别函数式的分析

第一主成分载荷图见图 4。载荷系数峰值所对应的特征波长分别为 1 124.14, 1 143.48, 1 169.80, 1 178.84, 1 207.39, 1 225.66, 1 287.13, 1 311.21, 1 343.15, 1 382.54, 1 398.96, 1 413.45, 1 457.15, 1 492.28, 1 540.27, 1 551.73, 1 572.98, 1 613.11,

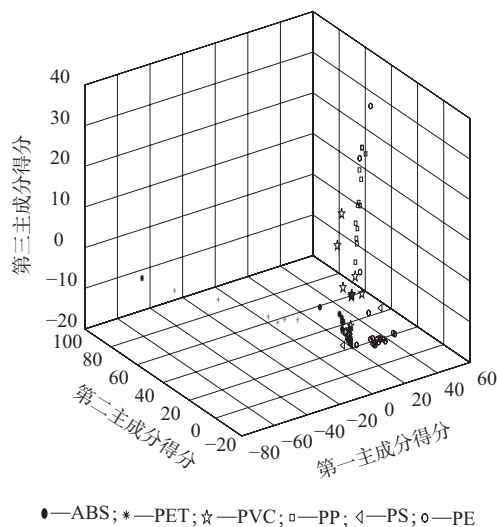


图3 主成分得分三维图

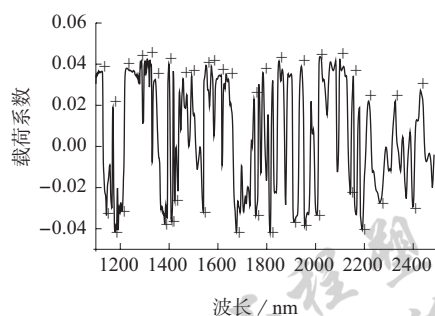


图4 第一主成分载荷系数图

1 642.76, 1 675.68, 1 747.99, 1 755.09, 1 784.08, 1 814.05, 1 824.26, 1 850.30, 1 900.49, 1 911.70, 1 941.77, 1 954.95, 2 004.84, 2 017.33, 2 099.01, 2 154.84, 2 191.27, 2 217.51, 2 262.01, 2 326.99, 2 398.03, 2 429.50 nm。

由以上 40 个特征波长对应的反射率为自由变量对 73 个训练集样本进行 Fisher 判别。由于样本种类数为 6, 所以典型判别式函数个数为 $6-1=5$ 。各判别式函数按式 (1)~ 式 (5) 计算, 其中 Y_i 为光谱信息在某方向上的投影函数得分, $i=1, 2, 3, 4, 5$; x_j 为特征波长变量在该方向的投影变量, 前面的常数代表函数投影方向, $j=1, 2, 3 \cdots 39$; 函数最后边的常数代表函数投影的阈值。

$$Y_1 = 21.373x_1 - 29.34x_2 + 34.405x_3 + 29.805x_4 + 74.028x_5 + 49.673x_6 + 53.038x_7 - 25.391x_8 + 14.569x_9 + 33.099x_{10} + 39.922x_{11} + 4.715x_{12} - 56.517x_{13} - 56.609x_{14} + 35.870x_{15} + 25.296x_{16} - 24.449x_{17} - 12.865x_{18} - 6.401x_{19} - 10.331x_{20} - 5.256x_{21} - 23.505x_{22} + 0.428x_{23} - 43.085x_{24} - 3.417x_{25} - 25.588x_{26} - 3.726x_{27} - 8.43x_{28} +$$

$$1.876x_{29} + 89.165x_{30} + 4.901x_{31} - 31.928x_{32} + 36.553x_{33} - 24.214x_{34} - 2.002x_{35} + 3.461x_{36} - 3.703x_{37} - 19.600x_{38} - 37.246x_{39} - 1.527 \quad (1)$$

$$Y_2 = -20.167x_1 - 1.332x_2 - 10.129x_3 + 10.26x_4 + 7.468x_5 + 9.311x_6 + 51.927x_7 + 42.815x_8 - 34.487x_9 + 6.611x_{10} - 2.945x_{11} - 8.584x_{12} - 40.763x_{13} - 54.379x_{14} - 24.688x_{15} + 17.864x_{16} + 16.169x_{17} + 13.338x_{18} + 1.263x_{19} - 2.852x_{20} - 11.715x_{21} + 1.094x_{22} - 0.921x_{23} + 18.619x_{24} - 12.528x_{25} + 41.594x_{26} - 16.016x_{27} + 10.403x_{28} - 26.182x_{29} - 12.468x_{30} - 5.551x_{31} + 25.223x_{32} + 2.343x_{33} + 15.800x_{34} - 1.507x_{35} + 11.527x_{36} - 0.994x_{37} + 10.528x_{38} + 41.193x_{39} - 1.097 \quad (2)$$

$$Y_3 = 1.603x_1 - 23.587x_2 + 19.570x_3 + 54.883x_4 + 33.176x_5 + 25.417x_6 + 121.534x_7 - 98.698x_8 - 43.842x_9 + 8.586x_{10} + 6.908x_{11} - 0.096x_{12} + 39.372x_{13} + 0.114x_{14} + 24.417x_{15} - 2.292x_{16} - 61.990x_{17} - 1.547x_{18} + 3.674x_{19} + 11.837x_{20} + 13.860x_{21} - 6.819x_{22} - 42.705x_{23} + 23.105x_{24} - 13.127x_{25} + 43.843x_{26} + 8.706x_{27} - 22.141x_{28} - 4.311x_{29} + 14.215x_{30} + 30.050x_{31} - 5.693x_{32} + 22.191x_{33} + 13.364x_{34} + 10.936x_{35} + 10.183x_{36} - 2.235x_{37} - 4.506x_{38} + 32.150x_{39} + 0.713 \quad (3)$$

$$Y_4 = -28.842x_1 + 11.298x_2 + 2.940x_3 - 26.480x_4 + 0.104x_5 - 12.098x_6 + 14.285x_7 - 32.008x_8 + 29.785x_9 - 25.236x_{10} + 2.411x_{11} - 4.798x_{12} + 59.364x_{13} - 7.511x_{14} + 5.648x_{15} - 31.709x_{16} + 16.809x_{17} + 31.432x_{18} + 14.229x_{19} + 6.993x_{20} + 22.682x_{21} - 0.146x_{22} - 730x_{23} + 107.828x_{24} + 16.822x_{25} + 10.505x_{26} - 3.348x_{27} - 4.205x_{28} - 20.500x_{29} - 35.062x_{30} + 24.451x_{31} - 19.328x_{32} - 16.958x_{33} + 13.641x_{34} + 7.117x_{35} + 2.309x_{36} - 2.913x_{37} - 7.626x_{38} + 4.646x_{39} + 1.149 \quad (4)$$

$$Y_5 = -9.230x_1 - 13.462x_2 + 17.050x_3 + 14.212x_4 + 10.012x_5 + 1.423x_6 - 51.125x_7 - 9.491x_8 - 2.581x_9 + 22.486x_{10} - 1.602x_{11} - 10.034x_{12} + 39.463x_{13} + 35.929x_{14} - 3.398x_{15} - 4.919x_{16} - 1.410x_{17} + 6.920x_{18} - 4.089x_{19} - 0.999x_{20} - 0.747x_{21} - 2.829x_{22} - 10.597x_{23} + 17.969x_{24} - 41.319x_{25} + 30.291x_{26} - 4.991x_{27} - 18.006x_{28} - 10.013x_{29} + 4.477x_{30} + 10.863x_{31} + 23.331x_{32} + 1.738x_{33} + 7.416x_{34} + 2.757x_{35} + 7.049x_{36} - 0.212x_{37} - 0.580x_{38} + 14.495x_{39} + 1.908 \quad (5)$$

典型判别式函数的特征值见表 1。根据判别函数式方差贡献率, 可以解释函数能够判别的样本信息量。从表 1 可知, 第一判别函数式只能解释样品

50.0% 的信息量,之后的函数能解释的样本信息量更是逐渐减少。说明只利用一个函数不能完成绝大多数样本的判别分类,联系 5 个判别式函数才能对样本进行 100.0% 准确识别。也就是说,当利用第一个判别式无法对塑料准确分类时,需要借助第二,甚至第三、第四个判别式函数。6 类废旧塑料在上述 5 个判别函数下的重心坐标见表 2,从表 2 可以看出各类塑料在第一、二个函数下的重心坐标具有明显的差别,两者结合可以作为外部验证判别函数。

表 1 典型判别式函数的特征值

函数类别	特征值	方差 / %	累计 / %	正相关性
函数 1	61.686	50.0	50	0.992
函数 2	25.440	20.6	70.7	0.981
函数 3	14.682	11.9	82.6	0.968
函数 4	13.551	11.0	93.6	0.965
函数 5	7.926	6.4	100	0.942

表 2 塑料在 5 个判别式函数下的重心坐标

塑料种类	函数 1	函数 2	函数 3	函数 4	函数 5
ABS	15.902	6.566	-3.059	-0.592	-0.740
PET	9.881	-11.894	2.977	-3.320	-0.857
PVC	-4.385	-3.300	-6.807	-0.163	3.865
PP	-4.353	-1.137	-1.365	3.717	-4.026
PS	0.371	1.365	4.466	3.810	2.876
PE	-6.394	3.161	1.661	-4.865	-0.514

2.4 判别函数式建模及验证

以第一判别式函数得分作为横坐标,第二判别式函数得分作为纵坐标,绘制二维散点图,如图 5 所示,6 类塑料各自聚在一起,自成一类,ABS 和 PET 与其它四种塑料的距离较远,能清晰地分开,而 PVC 塑料和 PP 塑料的第一判别式函数得分接近,需要借助第二判别式函数一起进行判别分类。

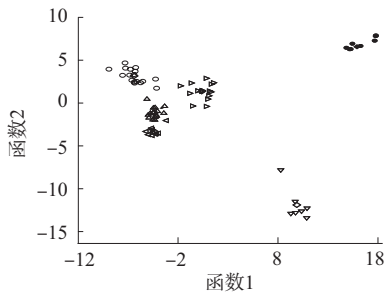


图 5 建模样本的 Fisher 判别函数得分散点图

Fisher 判别效果通过自身验证法和交叉验证法对训练集样本的回判结果见表 3,自身验证的误判率为 0,表明所建立的判别分析模型的判别能力是稳定的,虽然交叉验证样本的误判率稍高,达到 15.1%,但 84.9% 的交叉验证识别率仍可认为 Fisher 判别法对 6 种训练集样本塑料的分类是可行的。

表 3 训练集样本自身验证和交叉验证回判结果

项目	组别	塑料种类						合计
		ABS	PET	PVC	PP	PS	PE	
计数	a	9	0	0	0	0	0	9
	b	0	7	0	0	0	0	7
	c	0	0	10	0	0	0	10
	d	0	0	0	15	0	0	15
	e	0	0	0	0	15	0	15
	f	0	0	0	0	0	17	17
自身验证准确率 / %	a	100.0	0.0	0.0	0.0	0.0	0.0	100.0
	b	0.0	100.0	0.0	0.0	0.0	0.0	100.0
	c	0.0	0.0	100.0	0.0	0.0	0.0	100.0
	d	0.0	0.0	0.0	100.0	0.0	0.0	100.0
	e	0.0	0.0	0.0	0.0	100.0	0.0	100.0
	f	0.00	0.0	0.0	0.0	0.0	100.0	100.0
交叉验证准确率 / %	a	88.9	11.1	0.0	0.0	0.0	0.0	100.0
	b	0.0	57.1	28.6	0.0	14.3	0.0	100.0
	c	0.0	0.0	80.0	10.0	0.0	10.0	100.0
	d	0.0	6.7	6.7	86.7	0.0	0.0	100.0
	e	0.0	0.0	0.0	6.7	93.3	0.0	100.0
	f	0.00	0.0	0.0	5.9	5.9	88.2	100.0

2.5 预测集样品验证

将特征波长对应的预处理后的观测值代入各判别函数中,计算各样本的坐标值,然后与各重心坐标比对,距离最近的就是该样本的归属。表 4 是各未知样本在 5 个判别函数下计算的重心坐标,与表 2 比较可以得出这些样本的类别。以未知样本的第一、第二判别式函数得分绘制二维散点图见图 6,在该图中,“*”代表通过已建立的 Fisher 判别模型计

表 4 预测集样本的 Fisher 判别分析¹⁾

未知样本	函数 1	函数 2	函数 3	函数 4	函数 5	预知分类	判别类别
1	12.686	-16.219	-2.067	-3.672	-5.052	PET	PET
2	11.202	-14.583	-1.247	-8.874	-2.361	PET	PET
3	6.054	4.963	-3.800	-3.629	0.848	ABS	ABS
4	30.475	12.007	0.091	-2.320	3.691	ABS	ABS
5	22.275	17.286	-17.069	-23.640	-8.805	ABS	ABS
6	-0.836	1.895	4.540	5.898	1.892	PS	PS
7	-1.258	-0.221	1.819	6.529	4.512	PS	PS
8	10.483	6.147	24.345	7.412	7.783	PS	PS
9	-1.258	-0.221	1.819	6.529	4.512	PS	PS
10	3.100	1.994	10.070	9.489	3.992	PS	PS
11	4.779	1.561	12.439	10.169	4.197	PS	PS
12	6.094	1.208	16.405	12.783	11.351	PS	PS
13	-3.154	-2.639	-3.478	0.946	1.853	PVC	PVC
14	-12.942	-0.376	-5.792	9.752	-9.934	PP	PP
15	-6.157	-2.984	-6.581	7.591	-6.436	PP	PP
16	-5.687	4.444	-2.324	4.070	-3.921	PE	PE
17	-5.643	3.924	0.162	-1.437	-1.768	PE	PE
18	-7.246	6.148	-2.632	-5.534	5.096	PE	PE
19	-7.543	3.909	1.991	-4.247	2.714	PE	PE
20	-3.840	7.957	-1.140	-2.355	7.231	PE	PE
21	-10.644	6.678	2.696	-2.065	0.301	PE	PE

注: 1) 对预测集样本的分类准确率为 100%。

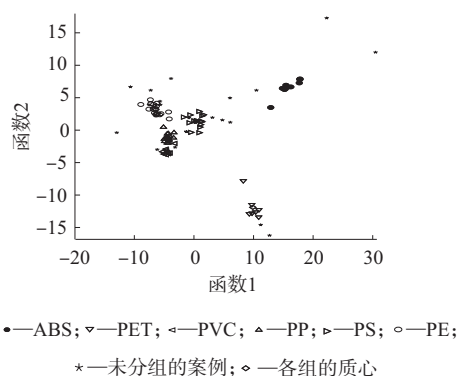


图6 预测集样品的 Fisher 判别函数得分散点图

算得到的未知样本的空间分布。由图6可见,未分组的塑料分别落在不同的区域,根据肉眼观测每个样本距离各类塑料质心的距离也可大概识别塑料种类。

3 结论

(1) 近红外光谱经过3次多项式、13窗口S-G最小二乘拟合+sym17函数、分解2层小波分析+主成分分析+特征波长选择+Fisher判别处理后,得到5个判别式函数并由此进行建模,所得模型可以实现6种废旧混合塑料的一次性识别,预测集准确率为100%。

(2) 选取特征波长点进行Fisher判别比直接根据主成分得分进行识别分类效果好。

(3) 为近红外塑料识别软件的二次开发提供了一个新的程序,为塑料识别技术的进一步研究提供了一个计算程序。

参考文献

- [1] 汤桂兰,胡彪,康在龙,等.废旧塑料回收利用现状及问题[J].再生资源与循环经济,2013,6(1):31-35.
Tang Guilan, Hu Biao, Kang Zailong, et al. Current status and problems on waste plastic recycling[J]. Renewable Resources and Circular Economy, 2013, 6(1):31-35.
- [2] 黄扬明,王翔.废旧塑料的红外鉴别技术及其应用[J].塑料,2008,37(6):94-97.
Huang Yangming, Wang Xiang. Technique and application of a banded plastics' infrared identification[J]. Plastic, 2008, 37(6):94-97.
- [3] 张毅民,刘红莎,朱艳芳,等.废旧混合塑料识别分离与清洗技术研究进展[J].化工进展,2013,32(6):1401-1406.
Zhang Yimin, Liu Hongsha, Zhu Yanfang, et al. Development of identification separation and cleaning technologies for mixed waste plastics[J]. Chemical Industry and Engineering Process, 2013, 32(6):1401-1406.
- [4] 杨懿,于随然.废旧家电异性塑料的自动识别与分炼技术的研究[J].机械制造与自动化,2011,40(4):169-172.
Yang Yi, Yu Suiran. Research on identification and sorting technology for plastics in discarded household appliance[J]. Machine Building and Automation, 2011, 40(4):169-172.
- [5] 张继游.废旧汽车塑料的识别技术研究[D].上海:上海交通大学,2010.
Zhang Jiyu. Investigation of used automotive plastics identification technology[D]. Shanghai:Shanghai Jiao Tong University, 2010.
- [6] 陈华才.红外光谱法快速鉴别PE和PVC保鲜膜[J].中国计量学院学报,2005,16(4):299-301.
Chen Huacai. Rapid discrimination of polyethylene(PE) and polyvinyl chloride(PVC) food wrapping film by infrared spectroscopy[J]. Journal of China Jiliang University, 2005, 16(4):299-301.
- [7] 窦艳丽,张万喜,张玉杰,等.傅立叶变换拉曼光谱和红外光谱鉴别塑料[J].分析化学,2006,34(11):1615-1618.
Dou Yanli, Zhang Wanxi, Zhang Yujie, et al. Identification of plastics by fourier transform raman and infrared spectroscopy[J]. Chinese J Anal Chem, 2006, 34(11):1615-1618.
- [8] 郭慧玲,邓文怡,李晓英.基于近红外光谱技术的常见塑料快速分类方法研究[J].现代科学仪器,2012,10(5):28-32.
Guo Huiling, Deng Wenyi, Li Xiaoying. Rapid classification of common plastics based on near-infrared spectral technology[J]. Modern Scientific Instruments, 2012, 10(5):28-32.
- [9] 周健,成浩,叶阳,等.基于近红外的Fisher分类法识别茶叶原料品种的研究[J].光学学报,2009,29(4):1117-1121.
Zhou Jian, Cheng Hao, Ye Yang, et al. Recognition for raw material cultivar of manufactured tea with Fisher discriminant classification with principal components analysis[J]. Acta Optica Sinica, 2009, 29(4):1117-1121.
- [10] Jie Y. Localized Fisher discriminant analysis based complex chemical process monitoring[J]. American Institute of Chemical Engineers, 2011, 57(7):1817-1828.
- [11] Masry G E, Iqbal A, Sun Dawen, et al. Quality classification of cooked, sliced turkey hams using NIR hyperspectral imaging system[J]. Journal of Food Engineering, 2011, 103:333-344.