

基于 DRAUG 和奇异值分解探究大曲中挥发性组分贡献的规律

高家坤¹, 陆玮^{1,2}, 汤有宏^{1,2}, 唐林^{1,2}, 姜利^{1,2}, 周庆伍¹

(1. 安徽古井贡酒股份有限公司, 安徽亳州 236820; 2. 安徽省固态发酵工程技术研究中心, 安徽亳州 236820)

摘要 [目的] 研究大曲中占主要贡献的风味物质, 探究各风味物质间的相关性, 为研究大曲及大曲酒中风味物质提供参考。[方法] 从车间随机选取 12 块大曲作为样品, 应用顶空-固相微萃取(HS-SPME)结合三重四级杆气质联用(GC-QQQ)全扫描技术分析大曲中挥发性组分, 以峰面积作为原始数据矩阵, 使用 DRAUG 算法计算出数据中的主要组分数, 然后使用奇异值分解计算出原始数据的特征值、特征向量, 经特征向量将原始数据转化为一组主成分, 对这些曲块进行分类, 最终以综合得分的形式对各块大曲进行排序。[结果] 大曲的主要风味物质主要指向 2-戊基呋喃、酚类物质、酯类物质; 供试样品中 B17 号样品风味最好, B33 号样品风味最差; 酯类物质间的相关系数大于 0.8。[结论] 大曲的主要风味物质为酚类物质、酯类物质, 酯类物质间的相关性强。

关键词 大曲; 顶空-固相微萃取; 特征值; 特征向量

中图分类号 TS261 文献标识码 A 文章编号 0517-6611(2016)23-056-04

Study on Contribution of Volatile Constituents in Daqu Based on DRAUG and SVD

GAO Jia-kun¹, LU Wei^{1,2}, TANG You-hong^{1,2} et al (1. Anhui Gujing Gongjiu Co., Ltd., Bozhou, Anhui 236820; 2. Anhui Research Center for Engineering Technology of Solid-state Fermentation, Bozhou, Anhui 236820)

Abstract [Objective] To study major flavoring compositions of Daqu and correlations between flavoring compositions to provide references for studies of Daqu and flavoring substances of Daqu liquor. [Method] Twelve blocks of Daqu were randomly selected from a workshop as samples, in which flavoring compositions were determined by headspace solid-phase microextraction (HS-SPME) and scan mode of gas chromatography coupled to triple quadrupole mass spectrometry (GC-QQQ). The original data matrix contained peak areas of volatile components tested. Afterwards, DRAUG was used to calculate the number of primary components in the initial data. Then SVD algorithm was used to calculate eigenvalues and eigenvectors of the initial data. Meanwhile, the data were translated into a series of principal components by eigenvectors. These blocks of Daqu were classified. Finally the composite score of every block of Daqu was sorted. [Result] 2-pentyl-furan, phenolics, and esters were main flavoring compositions of Daqu; sample B17 had the best flavor, and sample B33 had the worst flavor; the correlation coefficient between esters was greater than 0.8. [Conclusion] Major flavoring components in Daqu are phenolics and esters, and there is a substantial correlation between esters.

Key words Daqu; HS-SPME; Eigenvalue; Eigenvector

大曲是以小麦为主的谷物原料通过生物发酵而得到, 经粗粉碎后加水压成砖状曲块, 人工控制在一定温度、湿度下培育而成的生物制剂。大曲经发酵后富含多种微生物, 如芽孢杆菌、曲霉菌、酿酒酵母等; 以及各种酶系, 如糖化酶、蛋白酶等。利用大曲酿造白酒是我国白酒酿造的关键和区别于世界上其他知名蒸馏酒的重要技术特征^[1]。大曲是白酒酿造过程中重要的微生物、酶类、香气物质和香气物质前体物质的主要来源, 其质量直接关系到白酒的出酒率和品质。因此, 大曲本身就是一种香味物质的富集物, 所谓的曲香正是一种表现。因此, 研究大曲的主要香味物质, 具有重要意义。

对于大曲中风味物质的研究, 吕云怀等^[2]使用顶空-固相微萃取(HS-SPME)分析了大曲中挥发性成分的种类, 而该试验从统计学的角度, 采用化学计量学手段和 HS-SPME 技术提取中大曲中挥发性成分的主要信息。

笔者从酿酒车间随机选取了 12 块大曲作为样品, 应用 HS-SPME 结合三重四级杆气质联用(GC-QQQ)全扫描技术分析了大曲中挥发性组分^[3-4], 以各样本中解析出的挥发性组分峰面积作为原始数据矩阵, 使用 DRAUG 算法计算出数据中的主要组分数, 然后使用奇异值分解计算出原始数据的特征值、特征向量, 经特征向量将原始数据转化为一组主成分, 对这些曲块进行分类, 最终以综合得分的形式将结果

展现出来, 可为研究大曲及大曲酒中风味物质提供一定的参考。

1 材料与方法

1.1 材料 大曲样品, 古井贡酒制曲车间提供。挑选 12 块大曲, 序号分别以 B5、B9、B15、B17、B25、B33、C6、C8、C33、C34、C42、C43 表示, 分别粉碎过 40 目筛, 混合均匀。

50/30 μm DVB/CAR/PDMS 固相微萃取头, 手动 SPME 进样器, 20 mL 顶空瓶, 固相微萃取仪, Agilent 7000C 三重四级杆气质联用仪。

1.2 方法

1.2.1 顶空固相微萃取 (SPME) 条件。 选取 50/30 μm DVB/CAR/PDMS 固相微萃取头, 第 1 次使用时需于气相色谱进样口 250 老化 1 h。

样品处理: 取 8 g 样品溶于 25 mL 酒精度为 10% 的乙醇溶液中, 浸泡过夜后, 在 10 000 r/min, 15 $^{\circ}\text{C}$ 条件下离心 15 min, 取 8 mL 上清液于顶空瓶中, 加入 1 g 盐和磁子, 置于加热温度 50 $^{\circ}\text{C}$ 、转速 800 r/min 的固相萃取仪下加热平衡 20 min, 然后插入萃取头, 再萃取 40 min, GC-QQQ 进样口解析 7 min, 进行全扫描分析。

1.2.2 三重四级杆气质联用仪 (GC-QQQ) 分析条件。 气相色谱条件: 毛细管色谱柱为 CP-Wax, 规格 50 m \times 250 μm \times 0.2 μm ; 手动无分流进样, 进样口温度 230 $^{\circ}\text{C}$; 程序升温条件: 初始温度为 35 $^{\circ}\text{C}$, 以 2 $^{\circ}\text{C}/\text{min}$ 的速率升温到 60 $^{\circ}\text{C}$, 保持 4 min, 再以 4 $^{\circ}\text{C}/\text{min}$ 的速率升温到 90 $^{\circ}\text{C}$, 保持 4 min, 最后以

6 °C/min 的速率升温到 195 °C,保持 15 min;载气:高纯度氮气,流速为 1 mL/min^[5]。

质谱条件:EI 离子源,电子能量 70 eV,质量扫描范围为 30 ~ 500 amu,离子源温度为 230 °C,传输线温度为 195 °C。

1.2.3 定性分析。检出挥发性组分物质的质谱图,通过与标准谱图(Nist11.L)对比鉴定,可能性大于 70% 的予以确认。该试验以峰面积粗略表示各物质的含量,最终选择出 34 种挥发性物质作为大曲样品的变量。

1.3 数据处理方法 主成分分析有多种数值计算方法,可采用计算协方差矩阵的特征向量,奇异值分解(Singular Value Decomposition, SVD)是其中一种普遍的算法,其基本原理是谱分解原理。主成分分析的最终效果是,对原始数据中的变量进行降维,在尽可能保留原始数据中信息的同时,获得少数变量^[6]。

DRAUG 算法^[7]的基本思路是构造一个与原始数据矩阵相同大小、秩为 1 的矩阵,与原始数据矩阵相加,得到增广矩阵。外加矩阵的信息干扰了原始数据矩阵中的主因子信息,而与随机误差有关的次要因子信息未受影响。对原始矩阵和增广矩阵分别进行 PCA,然后在不同因子水平下比较各自残余矩阵的方差,如果不存在显著性差异,说明残余矩阵基本包含全部次要因子信息,此时的因子水平即是主因子数(主成分数)。相比于其他的主因子数估计方法,DRAUG 算法在信号重叠程度,组分的微量或痕量程度以及噪声水平方面所能承受的极限较强,是一种较准确的主因子数计算方法。

该试验首先使用 DRAUG 算法确定原始数据的主因子

数,然后经过奇异值分解,得到特征值和特征向量,由特征向量可判断出各主成分所反映的指标物质;将原始数据矩阵进行转置,对转置矩阵进行奇异值分解,根据得到的第 1 主成分、第 2 主成分,对大曲样品进行分类;转置矩阵奇异值分解后,根据得到的主成分向量和相应的方差比(各因子贡献率),得到各样品的综合得分;使用相关性函数计算各物质间的相关系数。

2 结果与分析

2.1 特征组分 构造数据矩阵,矩阵的行向量为同一样品、不同物质的峰面积数据,列向量为同一物质、不同样品的峰面积数据,矩阵大小为 12 × 34,使用 DRAUG 算法对其进行处理,rank = 6, stdrank = 0,再使用奇异值分解算法,计算荷载向量矩阵。表 1 中各列的荷载向量即各列向量,由表 1 可知,第 1 主成分(PC1)反映的组分指标主要有 2-戊基呋喃、2,6-二叔丁基-1,4-苯二酚、2,6-二叔丁基-4-(1-)-异丁基苯酚、2-甲基丁酸乙酯、2-辛酮、苯乙烯、2,5-二甲基-3-n 戊基吡嗪、1-壬醛、石竹烯等,主要指向呋喃、酚类物质、酮类物质;第 2 主成分(PC2)反映的组分指标主要有棕榈酸乙酯、己酸乙酯、十四酸乙酯、油酸乙酯、苯甲醛、亚油酸乙酯、十五酸乙酯、十二酸乙酯等,主要指向酯类物质、醛类物质;第 3 主成分(PC3)反映的组分主要有喇呗醇、3-甲基-1-丁醇、苯乙醇、5-甲基-2-己酮等,主要指向醇类物质;第 4 主成分(PC4)反映的组分主要有 3-辛醇、壬酸乙酯、13-甲基十四酸乙酯、癸酸乙酯。因此,可以初步判断以上挥发性物质为大曲的主要风味物质,主要指向 2-戊基呋喃、酚类物质、酯类物质。

表 1 34 种物质的特征向量

Table 1 Eigenvectors of 34 types of substances

序号 No.	物质名称 Substance	荷载向量 Load vector					
		主成分 1 Principal component 1	主成分 2 Principal component 2	主成分 3 Principal component 3	主成分 4 Principal component 4	主成分 5 Principal component 5	主成分 6 Principal component 6
1	2-甲基丁酸乙酯	-0.000 5	-0.012 7	-0.015 7	-0.043 4	-0.066 2	-0.023 6
2	5-甲基-2-己酮	-0.001 9	0.025 0	0.058 2	0.016 5	0.004 3	0.135 7
3	2-戊基呋喃	0.000 4	0.002 3	-0.002 4	0.018 1	0.002 2	-0.009 6
4	己酸乙酯	-0.123 7	0.174 9	-0.462 4	-0.292 6	0.625 1	-0.216 1
5	苯乙烯	-0.002 0	0.013 4	0.021 3	-0.022 4	0.000 0	-0.034 7
6	2-甲基-1-丁醇	-0.004 4	-0.020 7	0.037 2	-0.039 9	-0.047 3	-0.104 6
7	3-甲基-1-丁醇	-0.021 0	-0.124 4	0.163 1	-0.162 6	-0.111 0	-0.380 1
8	3-辛酮	-0.008 5	-0.005 6	-0.006 5	0.094 2	0.047 8	-0.045 7
9	2-辛酮	-0.001 5	0.009 8	-0.012 4	-0.018 5	-0.006 1	0.007 0
10	庚酸乙酯	-0.025 5	-0.006 8	-0.147 1	-0.064 4	-0.123 7	0.103 9
11	1-己醇	-0.008 1	0.006 0	0.002 3	-0.078 8	-0.022 1	-0.024 5
12	1-壬醛	-0.002 7	0.011 8	0.011 1	-0.000 9	-0.014 4	0.095 8
13	3-辛醇	-0.006 8	-0.003 7	-0.110 9	0.713 1	-0.032 4	-0.362 7
14	辛酸乙酯	-0.098 1	-0.045 0	-0.314 5	-0.026 2	-0.319 7	0.233 4
15	2,6-二叔丁基-1,4-苯二酚	-0.000 4	-0.006 3	0.000 9	0.001 6	0.016 2	0.012 5
16	苯甲醛	-0.037 3	0.080 7	0.332 6	0.330 0	0.270 8	0.419 7
17	壬酸乙酯	-0.071 3	-0.103 7	-0.570 3	0.258 6	-0.258 0	-0.000 8
18	石竹烯	-0.004 2	0.001 5	0.021 0	-0.101 3	-0.214 5	-0.105 5
19	癸酸乙酯	-0.034 2	0.007 2	-0.012 9	0.074 7	-0.012 5	0.022 9

接下表

续表 1

序号 No.	物质名称 Substance	荷载向量 Load vector					
		主成分 1 Principal component 1	主成分 2 Principal component 2	主成分 3 Principal component 3	主成分 4 Principal component 4	主成分 5 Principal component 5	主成分 6 Principal component 6
20	苯甲酸乙酯	-0.041 4	-0.008 5	-0.071 5	0.049 7	-0.033 3	0.158 4
21	甲氧基苯基酚	-0.384 1	-0.887 5	0.033 5	-0.014 2	0.202 9	0.014 3
22	苯乙酸乙酯	-0.012 3	-0.019 5	-0.167 8	-0.077 8	-0.126 5	0.132 9
23	十二酸乙酯	-0.032 9	0.025 9	-0.015 5	-0.000 1	0.048 8	-0.071 7
24	2,6-二叔丁基-4-(1-异丁基苯酚	-0.000 4	-0.000 2	-0.006 8	0.043 7	-0.002 0	-0.022 2
25	2,5-二甲基-3-n 戊基吡嗪	-0.002 2	-0.001 2	-0.036 2	0.232 6	-0.010 6	-0.118 3
26	苯乙醇	-0.049 8	0.014 5	0.077 3	-0.096 3	-0.098 7	-0.178 7
27	十四酸乙酯	-0.145 8	0.133 6	-0.125 9	-0.001 6	0.148 4	-0.175 4
28	13-甲基十四酸乙酯	-0.019 9	-0.002 1	-0.005 7	0.075 5	0.010 2	0.035 8
29	十五酸乙酯	-0.076 2	0.031 2	-0.018 8	-0.096 3	-0.047 5	-0.126 7
30	棕榈酸乙酯	-0.808 6	0.325 3	0.212 2	0.050 1	-0.145 4	-0.164 6
31	喇叭醇	-0.027 0	-0.053 0	0.175 4	-0.227 4	-0.336 4	-0.075 2
32	十八酸乙酯	-0.018 8	0.009 8	-0.006 8	0.002 8	-0.016 2	0.038 2
33	油酸乙酯	-0.288 7	0.105 2	-0.216 8	-0.111 5	-0.123 6	0.373 8
34	亚油酸乙酯	-0.217 7	0.071 5	0.057 9	0.085 7	0.194 0	0.253 6

2.2 样品分类 表 2 是 12 个大曲样品的第 1 主成分、第 2 主成分数据,图 1 是由 12 个大曲样品的第 1 主成分为横坐标,第 2 主成分为纵坐标,作散点图而得。由图 1 可知,12 个大曲样本可分为 5 簇,4 号(B17)样和其他样品分开较远,单独为一簇,命为 1 簇;2 号(B9)、7 号(C6)、12 号(C43)、10 号(C34)样比较接近,可分为一簇,命为 2 簇;1 号(B5)、3 号(B15)、9 号(C33)、11 号(C42)样比较接近,命为 3 簇;8 号(C8)单独为一簇,命为 4 簇;5 号(B25)、6 号(B33)比较接近,命为 5 簇。

表 2 大曲样品的第 1、第 2 主成分数据

Table 2 Data of the first and second principal component in Daqu samples

样本名 Name of each sample	样本号 Number of each sample	PC1	PC2
B5	1 号	-100 835 440	-91 596 604.70
B9	2 号	190 188 398	103 744 729.00
B15	3 号	-64 087 226	-86 568 199.10
B17	4 号	844 198 154	-39 581 104.20
B25	5 号	-170 075 482	17 077 910.33
B33	6 号	-233 845 334	-41 346 252.30
C6	7 号	-66 619 100	100 361 486.90
C8	8 号	-117 383 907	172 255 221.40
C33	9 号	-133 922 256	-84 215 743.40
C34	10 号	-49 834 633	-3 056 777.05
C42	11 号	-74 317 077	-127 936 287.00
C43	12 号	-23 466 096	80 861 620.27

2.3 综合得分 以每个主成分所对应的特征值占所提取的主成分总的特征值之和的比例作为权重,计算主成分综合得分,再根据综合得分值对 12 个大曲样品进行排序,综合得分越高,说明大曲风味越好,由此可以对大曲风味进行评定。由表 3 可得,B17 号样品得分最高,风味最好,其次是 B9 号样品,再其次是 C43 号样品,得分最低的是 B33 号样,该号样大曲风味最差。

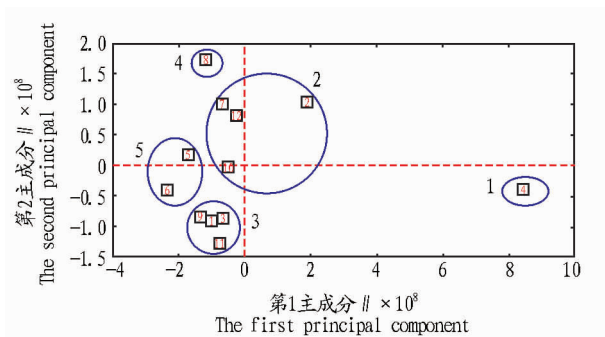


图 1 各样品的分类

Fig. 1 Classification figure of each sample

表 3 综合得分排序

Table 3 Order of composite scores

样品名 Name of each sample	样品号 Number of each sample	综合得分 Composite score	排序 Order
B5	1	-90 562 359	9
B9	2	166 792 414	2
B15	3	-61 876 247	6
B17	4	691 422 827	1
B25	5	-138 828 627	11
B33	6	-201 439 271	12
C6	7	-42 377 030	5
C8	8	-82 503 999	8
C33	9	-116 717 925	10
C34	10	-40 081 615	4
C42	11	-71 801 764	7
C43	12	-12 026 402	3

2.4 物质间的相关系数 当相关系数小于 0.3 时,两指标无相关;当相关系数介于 0.3 ~ 0.5 时,两指标低度相关;当相关系数介于 0.5 ~ 0.8 时,两指标显著相关;当相关系数大于 0.8 时,两指标高度相关。由表 4 可看出,酯类物质间

的相关系数较高,即相关性较强。也就是说,当同一样品中某种酯含量较高时,其他酯类物质的含量一般也会较高。

3 结论与讨论

该试验通过采用 DRAUG 算法来确定需提取的主成分个数,相比于仅按方差贡献率来人为地判断待提取的主成分数,解决了现有技术中的主观性问题,从而使最终计算的综

合得分更加准确。根据对大曲样品进行全扫描的定性数据,初步推断出大曲中的主要风味物质为酚类物质、酯类物质,而且,酯类物质的相关性较强。

该试验仅是大曲中挥发性组分进行定性的初步探究,后期会继续采集样本,对前期结果进行验证,并尝试建立定量方法,以对主要风味物质的判断更加准确。

表 4 物质间的相关系数

Table 4 Correlation coefficients between substances

相关性 Correlation	己酸 乙酯 Ethyl caproate	2-甲基 -1-丁 醇 2- methyl-1- butanol	1-壬醛 1- nonanal	辛酸 乙酯 Ethyl caprylate	癸酸 乙酯 Ethyl caprate	苯甲 酸乙酯 Ethyl benzoate	苯乙 酸乙酯 Ethyl phen ylacetate	十二 酸乙酯 Ethyl laurate	苯乙醇 Phenethyl alcohol	十四 酸乙酯 Ethyl myristate	十五酸 乙酯 Ethyl pentad ecanoate	棕榈酸 乙酯 Ethyl palmitate	十八酸 乙酯 Ethyl octad ecanoate	油酸 乙酯 Ethyl oleate
5-甲基-2-己酮 5-methyl-2-hexanone	—	—	0.909 9	—	—	—	—	—	—	—	—	—	—	—
己酸乙酯 Ethyl caproate	—	—	—	—	—	—	—	0.806 7	—	0.824 8	—	—	—	—
2-甲基-1-丁醇 2-methyl-1-butanol	—	—	—	—	—	—	—	—	—	—	—	—	—	—
3-甲基-1-丁醇 3-methyl-1-butanol	—	0.944 1	—	—	—	—	—	—	—	—	—	—	—	—
庚酸乙酯 Ethyl heptanoate	—	—	—	0.919 6	—	—	0.9109	—	—	—	—	—	—	—
1-壬醛 1-nonanal	—	—	—	—	—	—	—	—	—	—	—	—	—	—
辛酸乙酯 Ethyl caprylate	—	—	—	—	—	0.905 8	—	—	—	—	—	—	—	0.830 6
癸酸乙酯 Ethyl caprate	—	—	—	—	—	0.897 1	—	—	—	0.876 9	0.807 0	0.924 8	0.910 8	0.897 2
苯甲酸乙酯 Ethyl benzoate	—	—	—	0.905 8	0.897 1	—	—	—	—	—	—	0.818 0	0.869 2	0.894 1
十二酸乙酯 Ethyl laurate	0.806 7	—	—	—	0.859 8	—	—	—	—	0.966 3	0.905 1	0.930 3	0.821 8	0.870 1
十四酸乙酯 Ethyl myris- tate	0.824 8	—	—	—	0.876 9	—	—	0.966 3	—	—	0.925 7	0.944 0	0.889 8	0.926 0
13-甲基十四酸乙酯 13-methyl ethyl myristate	—	—	—	—	—	—	—	—	—	—	—	—	—	—
十五酸乙酯 Ethyl penta- decanoate	—	—	—	—	0.807 0	—	—	0.905 1	0.837 0	0.925 7	—	0.952 5	0.872 1	0.934 7
棕榈酸乙酯 Ethyl palmi- tate	—	—	—	—	0.924 8	0.818 0	—	0.930 3	0.816 1	0.944 0	0.952 5	—	0.942 5	0.962 6
十八酸乙酯 Ethyl octa- decanoate	—	—	—	—	0.910 8	0.869 2	—	0.8218	—	0.889 8	0.872 1	0.942 5	—	0.972 5
油酸乙酯 Ethyl oleate	—	—	—	0.830 6	0.897 2	0.894 1	—	0.870 1	—	0.926 0	0.934 7	0.962 6	0.972 5	—
亚油酸乙酯 Ethyl linole- ate	—	—	—	—	0.922 4	0.843 0	—	0.894 1	—	0.926 2	0.886 8	0.974 3	0.943 1	0.955 3

参考文献

- [1] 邢刚,敖宗华,王松涛,等. 大曲挥发性成分动态变化研究[J]. 酿酒科技,2014(9):1-4,8.
- [2] 吕云怀,王莉,汪地强,等. 不同香型白酒大曲风味物质与其产品风格特征关系的分析[J]. 酿酒科技,2012(7):72-75.
- [3] 陈勇,陈泽军,周瑞平,等. 顶空固相微萃取-气相色谱-质谱法测定大曲中的挥发性组分[J]. 中国调味品,2013(2):70-75.
- [4] 张春林,敖宗华,炊伟强,等. 固相微萃取-气相色谱-质谱法分析中高温大曲发酵、贮存过程中挥发性风味成分的变化[J]. 食品与发酵工

业,2011(4):198-203.

- [5] 邓静,王远兴,陈赟喆,等. 顶空-三重串联四极杆气-质联用法测定靖安白茶香气成分[J]. 食品科学,2013(22):115-118.
- [6] XU L, TANG L J, CAI C B, et al. Chemometric methods for evaluation of chromatographic separation quality from two-way data: A review[J]. Analytical chimica acta, 2008,613(2): 121-134.
- [7] MALINOWSKI E R. Determination of rank by augmentation (DRAUG) [J]. Journal of chemometrics,2011,25(6): 323-358.

科技论文写作规范——作者

论文署名一般不超过 5 个。中国人姓名的英文名采用汉语拼音拼写,姓氏字母与名字的首字母分别大写;外国人姓名、名字缩写可不加缩写点。