

激光诱导击穿光谱元素谱线自动识别方法研究

刘立拓, 刘建国, 赵南京*, 鲁翠萍, 陈东, 石焕, 王春龙, 张玉钧, 刘文清

中国科学院环境光学与技术重点实验室, 中国科学院安徽光学精密机械研究所, 安徽合肥 230031

摘要 根据激光诱导击穿光谱谱线展宽机制, 对 NIST 标准谱线数据库中的发射谱线进行了模拟, 并与实验光谱进行对比, 在对比分析中引入相似性测度, 得到模拟光谱和实验光谱的相似程度。对元素谱线的自动归属识别方法进行分析, 通过相似性计算对土壤样品 340~345 nm 波段光谱进行了识别分析, 利用最小二乘原理计算各谱线之间的比例系数, 实现了光谱识别, 实验结果验证了基于相似性测度的光谱识别方法的可行性及其自动识别的优越性。

关键词 激光诱导击穿光谱; 模拟光谱; 相似性测度; 谱线识别; 谱线展宽

中图分类号: O433.5; O241.5 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2012)01-0020-05

引言

激光诱导击穿光谱技术(LIBS)是当前探测物质成分信息的一项新技术, 具有实时、快速、原位检测等特点。随着激光技术的飞速发展, LIBS 技术正逐渐走向成熟, 并在各个领域开展了广泛的研究, 正发展成为一种极具应用前景的物质成分检测技术^[1-4]。

激光激发产生的等离子体包含大量的跃迁谱线, 尤其在物质成分比较复杂、元素种类比较繁多的情况下, 谱线之间相互重叠干扰, 物质成分的鉴定以及谱线的归属识别就变得比较困难, 虽然目前 LIBS 实验采用分辨率较高的光栅光谱仪, 其光栅分辨率能够达到 0.1 nm 甚至更高, 但光谱仪仍然存在“波长准确度”问题, 而“波长准确度”问题的存在是由于光栅光谱仪的工作性质和基于标准灯谱线的波长校准方法^[5-7]决定的, 这就导致在实验波长峰位 $\lambda_0 \pm \Delta\lambda$ 的一个小范围内的谱线峰位都有可能是波长真实值。目前, 在 LIBS 谱线识别过程中, 大部分研究者是通过经验对谱线直接进行归属, 并没有给出识别的方法; 有些研究者研究了一些识别方法, 其中, Robert J Nordstrom 等^[8]利用 NIST 标准光谱数据库绘制模拟图, 并与 LIBS 实验数据图进行观察对比的方法分别对空气中的 N_2 和 O_2 的原子态和第一离子态谱线进行了识别; 谢承利等^[9]在煤质 LIBS 检测过程中, 对谱线归属进行了总结, 在综合考虑谱线跃迁概率、等离子体演化特性

以及谱线族谱线协同存在性的基础上, 优先考虑 NIST 标准谱线库中强度比较大的谱线的方法对谱线进行归属识别; 亓洪兴等^[10]提出了基于标准样品原子谱线的元素识别方法, 即在探测未知元素之前对一组高浓度、已知组分的标准样品进行 LIBS 测量, 然后再与未知样品的 LIBS 实验数据对比分析实现元素识别。这些根据 NIST 数据库标准谱线或根据已知组分标准样品进行对比分析的方法共同之处都是利用人眼观测对比相似性, 而并没有给出一个相似性程度指标, 在识别过程中就要依赖人眼判读这一过程, 导致一方面使识别不能充分利用计算机的快速自动化、智能化判读; 另一方面随着样品组分复杂度的提高, 人眼观测会变得越来越困难。本文对土壤样品的 LIBS 光谱进行了实验测量, 根据 NIST 光谱数据库中不同元素的相关参数计算得到模拟光谱, 在与 LIBS 实验测量光谱对比分析的过程中引入相似性测度, 以二者的相似性程度数学指标来进行识别, 并以最小二乘原理计算得到各谱线之间的比例关系, 有效地实现了谱线的快速、自动识别。

1 实验装置

实验系统如图 1 所示, 采用 Nd:YAG 波长 1064 nm 激光器, 脉宽为 6 ns, 脉冲重复频率为 1~20 Hz, 脉冲能量 50 mJ; 光谱仪(Oceanoptics HR2000+)测量波长范围: 300~500 nm, 分辨率 0.1 nm, 两个聚焦镜均为石英透镜, 样品台

收稿日期: 2011-04-02, 修订日期: 2011-07-28

基金项目: 国家自然科学基金项目(60908018), 国家重大科技专项(2008ZX07527, 2009ZX07420), 中国科学院知识创新工程重大项目(kzcx-1yw06-01)和安徽省科技计划项目(083H3b1581)资助

作者简介: 刘立拓, 1982 年生, 中国科学院安徽光学精密机械研究所博士研究生 e-mail: ltuoliu@aiofm.ac.cn

* 通讯联系人 e-mail: njzhao@aiofm.ac.cn

由步进电机控制其转动, 首先通过计算机控制光谱仪使其向激光器发送 TTL 信号令激光器出光, 激光经过全反镜垂直入射到聚焦镜, 再经聚焦镜聚焦作用到样品上, 产生的激光等离子体通过另外一个聚焦镜收集耦合到光谱仪光纤输入端, 最后通过光谱仪将探测到的离散等离子体数字信号输入到上位机进行处理。

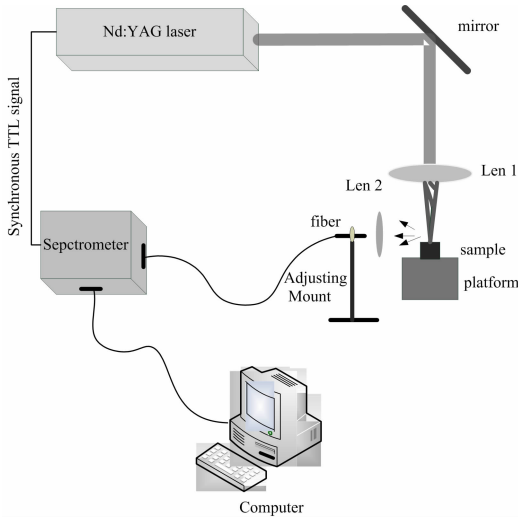


Fig. 1 Experimental device

2 相似性测度与光谱模拟

相似性测度的表示形式很多, 可以归结为两矢量间的距离测度和方向测度两类。本文采用两矢量的方向相近程度来表示, 设两矢量 $y = (y_1, y_2, \dots, y_n)$, $z = (z_1, z_2, \dots, z_n)$ 表示两个数据集样本。其相似性测度为

$$r(y, z) = \frac{(z - \bar{z})'(y - \bar{y})}{[(z - \bar{z})'(z - \bar{z})(y - \bar{y})'(y - \bar{y})]^{1/2}} \quad (1)$$

其中 \bar{z} 和 \bar{y} 分别表示两数据集的平均矢量。它实际上是数据中心化后的矢量夹角余弦^[12]。表示成数据集样本求和的形式

$$r(y, z) = \frac{\sum_i (z_i - \bar{z})(y_i - \bar{y})}{\left[\sum_i (z_i - \bar{z})^2 \sum_i (y_i - \bar{y})^2 \right]^{1/2}} \quad (2)$$

式(1)和(2)中的 r 也称为非线性相关系数^[11], 简称相关系数, 它表示了两个矢量或两列数据样本的相似性或相关性程度。

在光谱模拟研究中, 以 Lorentz 函数模拟不同元素的发射谱线, 可表示为

$$I_{ij} = \frac{2A_j \omega_j}{4\pi(x_i - \lambda_j)^2 + \pi\omega_j^2} \quad (3)$$

其中 I_{ij} 表示第 j 个峰位的第 i 个数据点, ω_j 代表第 j 个峰位的谱线半宽, λ_j 表示第 j 个峰位。 A_j 可以用中心峰强度来代替, 最后式(3)可化为

$$I_{ij} = \frac{I_{0j} \omega_j^2}{4(x_i - \lambda_j)^2 + \omega_j^2} \quad (4)$$

其中 I_{0j} 代表第 j 个峰中心强度, 其值可以从 NIST 数据库中查得。式(4)是理想情况下单个谱线展宽的 Lorentz 函数形

式。当考虑到谱线叠加和 LIBS 光谱的连续背景辐射时, 谱线可表示为

$$I_i = a_k \sum_j I_{ij} + D = a_k \sum_j \frac{I_{0j} \omega_j^2}{4(x_i - \lambda_j)^2 + \omega_j^2} + D \quad (5)$$

其中 D 表示连续背景, a_k 为模拟光谱强度的比例系数, 下角标表示不同的元素, 表征不同的谱线系列。 a_k 决定了模拟光谱图中光谱的绝对强度。由 NIST 库提供峰位信息, 再根据式(5)便可以绘制出某一光谱范围内的光谱模拟图。设 $y = (y_1, y_2, \dots, y_n)$ 为实测光谱强度, $z = (z_1, z_2, \dots, z_n)$ 为相应的模拟光谱强度, 则 $z_i = I_i, i = 1, 2, \dots, n$, 将式(5)代入式(2)后, 便可以利用最小二乘原理计算当 r 达到最大时比例系数 a_k 的值。

3 结果与分析

实验对山东掖县焦家黄棕壤土壤成分分析标准样品(地球物理地球化学勘查研究所)进行了测量与分析, 对样品表面同一点重复测量 20 次累加平均, 光谱仪积分时间为 1 ms。图 2(a)为采集到的 340~345 nm 波段实验光谱数据, 为了表述方便, 将图中峰位进行标号, 如图 2(a)中, 标号“1”, “2”, “3”, “4”, “5”和“6”的实验测量峰位分别为 344.017, 344.355, 342.660, 340.713, 341.250 和 341.736 nm。为了确定峰“1”, 从 NIST 库中查询到峰“1”可能为 Fe I : 344.099 nm, K II : 344.005 nm, Ti I : 343.93 nm, Al I : 343.935 nm, Cr I : 344.112 nm。

图 2(b)~(f)虚线部分都是由 NIST 光谱数据库中光谱数据模拟的标准化强度, 每一幅图都代表一种元素的波段内谱线系列, 且波段内的最高峰都是 100; 实线部分表示标准化后的实验测量光谱强度。只有模拟谱线的相对强度才有实际意义, 而强度值并没实际意义。因此, 只考虑模拟图中的峰位及谱线之间的相对强度。在图 2(b)~(f)中, 都存在如图 2(a)中峰“1”位置的谱线, 仅凭这一个峰位信息不能对峰“1”进行归属标定, 但是如果考虑峰“1”附近其他峰的存在情况及相对强度关系, 不难看出, 图 2(b)更接近图 2(a), 但也存在差异。为了更好地描述图 2(b)~(f)与图 2(a)的相似程度并实现计算机的自动识别, 分别计算图 2(b)~(f)与图 2(a)的相关系数。结果如表 1 所示。

Table 1 correlation coefficient of simulative and experimental spectral

	b	c	d	e	f
r	0.899 2	0.303 1	0.241 6	0.309 0	0.119 1

从表 1 可以看到, 模拟图 2(b)与图 2(a)的相似度明显高于其他四幅图, 从而判断峰“1”为 Fe I : 344.099 nm。

上述方法对图 2(a)中峰“1”进行了归属识别, 主要是利用波段内的谱线系列来计算该波段内的模拟数据与实验测量数据的相关性。在识别了峰“1”的同时, 峰“2”, “3”和“4”分别也有相应的峰与之对应, 但是从图 2(b)中可以看到, 模拟峰之间与实验测量峰之间的相对强度关系存在差别。比如:

实验测量峰“1”和峰“2”之间的强度比明显大于模拟数据中对应位置的峰值之比。因此,可以判断峰“2”,“3”和“4”可能为不同谱线的叠加结果。为了进一步确定峰“2”,“3”和“4”的成分,不能直接用确定峰“1”时所用的方法。而应该在已有确定成分的基础上进行叠加,判断必须遵循一个原则,从数学角度来讲,当叠加进新谱线时,相关系数必须增大。从物理实际上描述,就是在分析波段内,在新加入元素的模拟谱线

系列中,不能出现同时满足以下三个条件的谱线:

- (1) 实验测量数据中不存在;
- (2) 光谱强度比要分析的谱线强度还要强(或强度相当);
- (3) 较易跃迁的谱线(一般指原子谱线或比所要分析的谱线较易跃迁的谱线)。

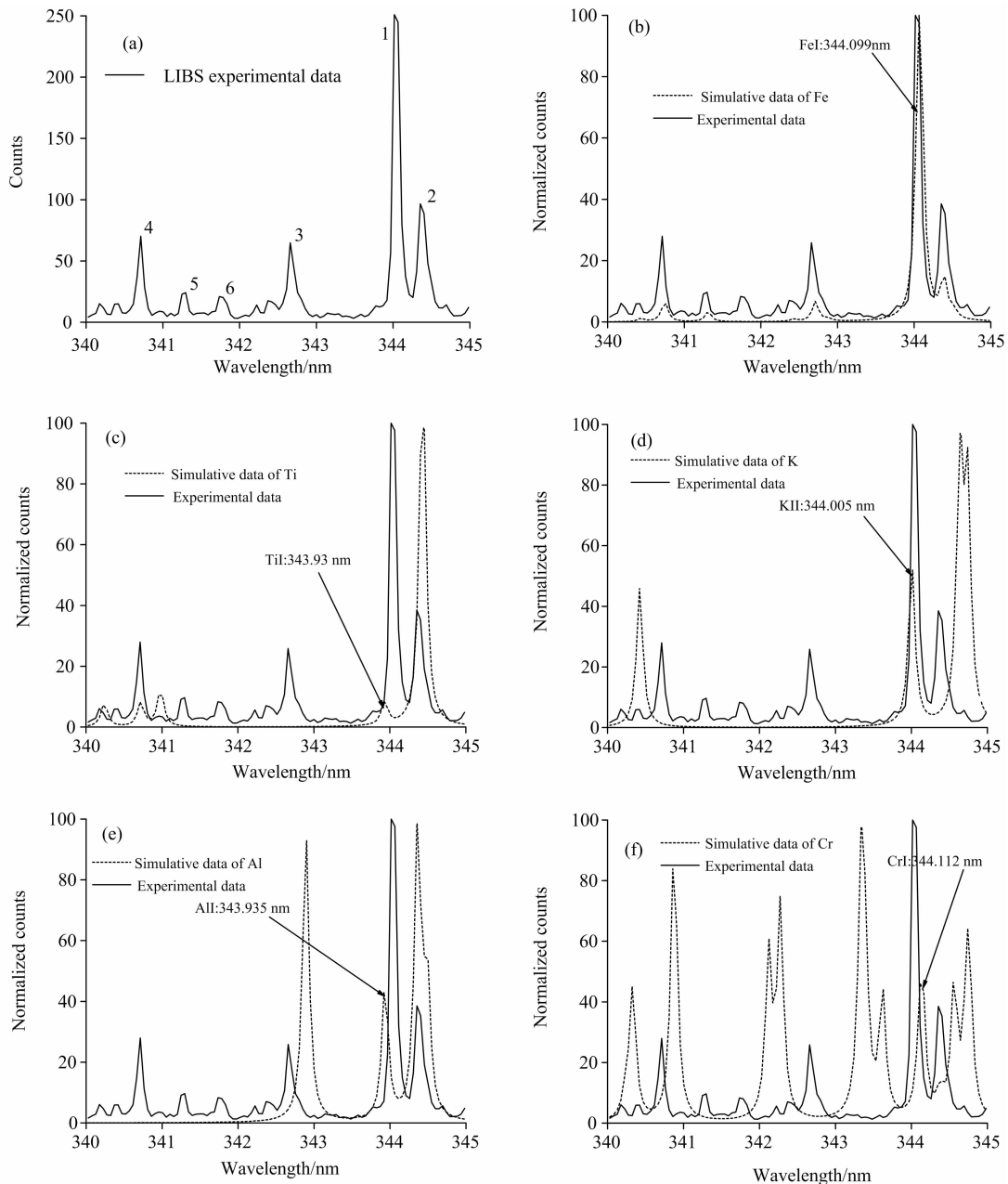


Fig. 2 Experimental and simulative spectral

例如,从前面分析可知,图 2(a)中峰“2”是铁 Fe I : 344.388 nm 与其他峰的叠加。我们用上面分析峰“1”的方法分析峰“2”,查询 NIST 光谱数据库,和实验测量峰“2”对应的峰可能为: Co I : 344.364 nm, Ti II : 344.431 nm 等。但是 Co 在分析波段内的谱线系列中存在 Co I : 340.512 nm 谱

线且比 Co I : 344.364 nm 强度大,如图 3(a)所示,而在实验测量数据中并不存在。因此,排除 Co I : 344.364 nm 的可能性。而 Ti 在该波段内的谱线系列正好满足上面的条件,如图 3(b)所示。

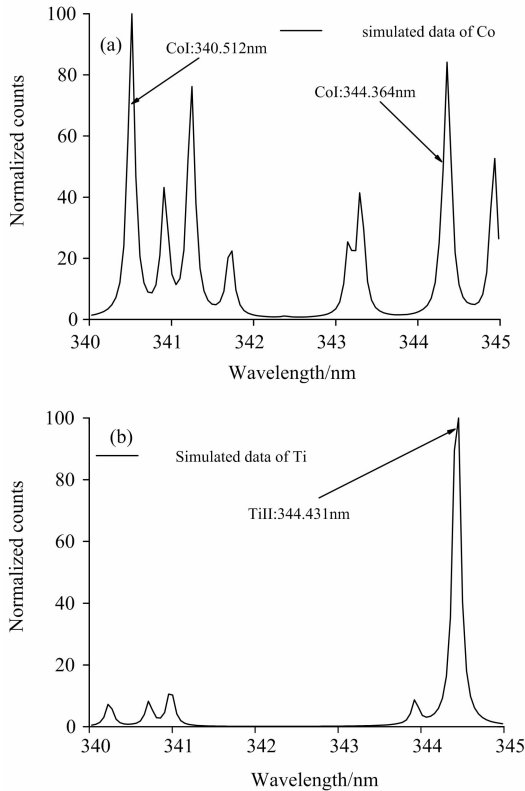


Fig. 3 Spectral series of Co and Ti

这样就可以将图 3(b) 叠加到 Ge 的谱线系列中去。同样的方法可以初步确定峰“3”、“4”、“5”和“6”。利用非线性最小二乘原理, 调整各元素谱线系列的比例系数, 也就是 a_k 的值, 使相关系数达到最大, 最终确定元素组分。表 2 列出了

Table 2 The results of elemental recognition and analysis

	Fe	Ti	Na	Hf	Ta
a_k	1	0.22	0.13	0.055	0.08

References

- [1] Christopher B Stipe, Brian D Hensley, Jeffrey L Boersema, et al. Applied Spectroscopy, 2010, 62(2): 154.
- [2] Tognoni E, Cristoforetti G, Legnaioli S. Spectrochimica Acta Part B, 2007, 62(12): 1287.
- [3] Kemal E Eseller, Markandey M Tripathi, Fangyu Yueh, et al. Applied Optics, 2010, 49(13): C21.
- [4] Rosalie A Multari, David A Cremers, Joanne M Dupre, et al. Applied Spectroscopy, 2010, 64(7): 750.
- [5] Nai Shuocheng, Angelos A, Michael G, et al. IEEE Trans. On Microwave Theory Techniques, 1999, 47: 1070.
- [6] XIE Xiao-qiang, LIN Wei-gan, XU Rui-min(谢小强, 林为干, 徐锐敏). Journal of Infrared and Millimeter Waves(红外与毫米波学报), 2006, 25(1): 25.
- [7] Jeong J, Kwon Y, Lee S, et al. IEEE Trans. On Microwave Theory Techniques, 2000, 48(12): 2700.
- [8] Robert J Nordstrom. Applied Spectroscopy, 1995, 49(10): 1490.
- [9] XIE Cheng-li(谢承利). <http://epub.cnki.net/grid2008/detail.aspx?QueryID=5&CurRec=1>. Wuhan: Huazhong University of Science and Technology(武汉: 华中科技大学), 2009.
- [10] QI Hong-xing, SHU Rong, MA De-min, et al(齐洪兴, 舒嵘, 马德敏, 等). Journal of Infrared and Millimeter Waves(红外与毫米波学报), 2007, 26(1): 53.
- [11] ZHANG Shi-qiang, LÜ Jie-neng, JIANG Zheng, et al(张世强, 吕杰能, 蒋峥, 等). Mathematics in Practice and Theory(数学的实践与认识), 2009, 39(19): 103.

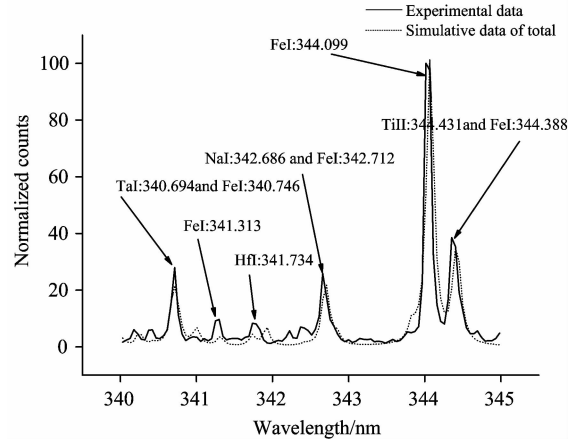


Fig. 4 Results of spectral recognition

当相关系数达到最大值为 $r=0.928$ 时识别的元素及其各自比例系数的值, 图 4 给出了最终识别结果。

整个识别过程可总结为: 在分析波段内, 考虑到 LIBS 激发光谱的特征及物质的浓度与峰强之间关系, 应按照峰强从大到小的顺序进行识别, 即首先识别相对强度最大的谱线, 根据谱线系列中相关系数最大的作为识别结果; 在此基础上, 将新的谱线系列与先前已识别的结果叠加在一起, 最后调整各谱线系列的比例系数, 使相关系数达到最大, 确定组分并实现谱线的最终识别。

4 结论

在 LIBS 光谱识别过程中, 引入相似性测度, 计算由 NIST 光谱数据库建立的模拟谱与实验测量谱的相似性来进行识别, 不仅实现了光谱识别的自动化, 而且可以对具有重叠光谱的成分进行分析和确定, 识别结果快速准确, 具有一定的应用价值, 为 LIBS 技术的定量分析及物质快速检测提供了基础。

Study on the Automatic Recognition Method of Elemental Spectra in Laser Induced Breakdown Spectroscopy

LIU Li-tuo, LIU Jian-guo, ZHAO Nan-jing*, LU Cui-ping, CHEN Dong, SHI Huan, WANG Chun-long, ZHANG Yu-jun, LIU Wen-qing

Key Laboratory of Environment Optics & Technology, Institute of Anhui Optics Fine Mechanics, Chinese Academy of Sciences, Hefei 230031, China

Abstract According to the spectral line broadening mechanism of laser induced breakdown spectroscopy, the emission lines from the standard spectral database of NIST were simulated, and they were compared with the spectral data of experiment. In the process of comparison similarity measure was used to measure the similarity between the simulative spectra and the experimental spectra. The automatic recognition method of elemental spectra was studied, and the spectral data of soil between 340 and 345 nm was recognized by computing the proportional coefficients of the spectral lines. Using principle of nonlinear least squares, the recognition process was completed. The feasibility and the advantage of the method were proved by the results of experiment.

Keywords Laser induced breakdown spectroscopy; Simulative spectra; Similarity measure; Spectral line recognition; Spectral line broadening

(Received Apr. 2, 2011; accepted Jul. 28, 2011)

* Corresponding author

《光谱学与光谱分析》对来稿英文摘要的要求

来稿英文摘要不符合下列要求者，本刊要求作者重写，这可能要推迟论文发表的时间。

1. 请用符合语法的英文，要求言简意明、确切地论述文章的主要内容，**突出创新之处**。
2. 应拥有与论文同等量的主要信息，包括四个要素，即研究目的、方法、结果、结论。其中后两个要素最重要。有时一个句子即可包含前两个要素，例如“用某种改进的 ICP-AES 测量了鱼池水样的痕量铅”。但有些情况下，英文摘要可包括研究工作的主要对象和范围，以及具有情报价值的其他重要信息。在结果部分最好有定量数据，如检测限、相对标准偏差等；结论部分最好指出方法或结果的优点和意义。
3. 句型力求简单，尽量采用被动式，通常应有 2000 个印刷字符，300 个英文单词为宜，不能太短；也不要太长。用 A4 复印纸单面打印。
4. 摘要不应有引言中出现的内容，换言之，摘要中必须写进的内容应尽量避免在引言中出现。摘要也不要对论文内容作解释和评论，不得简单重复题名中已有的信息；不用非公知公用的符号和术语；不用引文，除非该论文证实或否定了他人已发表的论文。缩略语、略称、代号，除相邻专业的读者也能清楚地理解外，在首次出现时必须加以说明，例如用括号写出全称。