

二分搜索的高质量紫外可见光谱信号重构算法

朱红求, 胡浩南, 郑国梁, 周 灿*, 李勇刚

中南大学自动化学院, 湖南 长沙 410083

摘 要 高锌背景下光谱法同时检测痕量多金属离子浓度时, 由于微型光谱仪光源能量辐射不均匀性, 并且混合溶液中不同离子对不同波段紫外可见光会选择性吸收, 因而如果选取微型光谱仪的积分时间过大, 可能导致光谱能量值达到饱和, 选取积分时间过小可能导致光谱信号的信噪比很低。积分时间的选择往往取决于研究者的经验和待测离子对紫外可见光的吸收特征。为了实现能够自动选取微型光谱仪积分时间参数, 提出了一种基于二分搜索的高质量紫外可见光谱信号重构算法, 用于重构由不同积分时间组成的图谱特征更加明显的紫外可见光谱信号。该方法首先采集不同积分时间下参比溶液的紫外可见光谱能量信号, 然后给定参比溶液的不同目标重构光谱能量信号值, 在每一波长点使用二分搜索算法寻找合适的积分时间采样参数; 然后根据紫外可见光谱的特点, 定义了表示重构后的光谱能量值与目标设定值接近程度的重构精度指标和表示重构信号后与重构信号前的图谱特征区分程度的重构特征显著度指标, 最后, 选取搜索区间范围内重构精度最高的光谱信号作为重构信息量, 利用光谱信号重构信息量重构待测溶液紫外可见光谱能量值, 最终得到待测溶液的重构光谱吸光度信号。实验结果表明, 该算法能够快速自动地选定目标积分时间采样参数值对紫外可见光谱进行信号重构, 来得到高质量紫外可见光谱信号。该算法可以使信号重构精度达 94.84%, 并且重构特征显著度有所提升。同时, 相对于重构前的光谱信号, 重构后的光谱吸光度信号得到一定程度增强, 信号信噪比也大大提升, 而且避免了积分时间采样参数需要依靠研究者主观判断选择的问题, 为检测多种痕量金属离子的浓度信息提供了高质量的模型数据。

关键词 微型光谱仪; 信号重构; 二分搜索; 积分时间; 光谱能量值

中图分类号: O433.1 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)06-1979-05

引 言

常用的金属离子检测分析多使用光谱法^[1]。根据待测物质的特性, 光谱法可以满足检测速度快、精度高、工作量少、不易对检测溶液造成二次污染的要求^[2]。微型光谱仪具有体积小、价格便宜、便于携带等特点, 在环保和食品等现场分析检测领域得到了很好的应用。但是直接利用微型光谱仪来检测高浓比背景下的痕量多重金属离子信息时由于光源的能量分布不均匀、待检测痕量金属离子种类数量多、不同的金属离子对不同光谱波段的选择性吸收等原因, 不同积分采样时间参数获得的光谱信号图谱特征差异大, 特别在光源能量衰减较大和离子特征吸收波段表现尤为明显。微型光谱仪设置的积分时间太大, CCD^[3]饱和, 测试的数据就会失真, 积分时间太小, 能量值就会很弱, 积分时间参数的设置会影

响。

信号重构是一种新的信号处理技术^[4], 主要从采集到已知部分信号中恢复或表示整个完整的信号, 以达到数据压缩^[5]的目的。常用的信号重构方法主要包括梯度追踪算法^[6]、正交匹配追踪算法^[7]和压缩感知算法^[8]等。在解决具体的实际工程问题时, 往往需要根据待解决信号的特点来选择合适的信号重构算法^[9], 若所选方法不能很好地解决目标问题, 还需要根据所选方法提出相应的改进策略或者提出一种新的信号重构方法来达到信号处理的目的。如文献^[10]提出了一种基于匹配追踪算法的信号重构方法, 通过阈值循环迭代的方法在平均谱上找出特征峰的位置、峰区间等信息, 实现对拉曼光谱的基线扣除和噪声处理。

但是通过不同的积分时间采样参数下的光谱信号来重构高质量光谱以解决光谱信号图谱特征差异大的问题, 上述算法处理效果并不理想。本文根据光谱信号特点提出了基于二

收稿日期: 2019-05-24, 修订日期: 2019-10-11

基金项目: 国家自然科学基金重点项目(61533021, 61773403)资助

作者简介: 朱红求, 1970年生, 中南大学自动化学院教授 e-mail: hqcsu@csu.edu.cn

* 通讯联系人 e-mail: zhoucan@csu.edu.cn

分搜索的高质量光谱信号重构算法。首先基于光谱特征定义了重构精度和重构特征显著度两个指标,然后基于二分搜索进行信号重构;最后以实测混合溶液的光谱信号为例,进行信号重构的实验特征。实验结果表明,所提方法不仅能够快速选定目标积分时间采样参数值,而且重构后的光谱信号的重构精度和图谱特征均获得了明显的提升。

1 二分搜索积分时间计算

传统的积分时间参数值选择是通过研究者的已有经验和被检测对象的特点来共同决定的,挑选出的积分时间存在一定的主观性,对于经验不足的仪器操作人员可能采集到的数据不符合检测要求。因此,为了选到合适的积分时间参数,通过采集所有的积分时间下的光谱信号,然后对这些光谱信号进行目标值的查找搜索,选择合适的光谱信号下所对应的积分时间采样参数,并用选定的参数作为仪器的采样参数。

二分搜索法算法是充分利用了元素间顺序存储次序关系,采用分治策略,每次把搜索区域缩减一半,能够较快地完成搜索任务。

首先,假设表中元素是按升序排列,将表中位置记录的关键字与查找关键字比较,如果两者相等,则查找成功;否则利用中间位置记录将表分成前、后两个子表,如果中间位置记录的关键字大于查找关键字,则进一步查找前一子表,否则进一步查找后一子表。重复以上过程,直到找到满足条件的记录。

在进行二分搜索之前,首先需要对微型光谱仪采集的光谱信号用矩阵 DATA 表示,可用式(1)来表示

$$\text{DATA} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (1)$$

其中, n 为积分时间, m 为波长点。按照矩阵 DATA 中每一列为单位进行目标搜索查询,如数组 $A = [x_{1k}, x_{2k}, \dots, x_{nk}]^T, 1 \leq k \leq m$ 。从数组中搜索目标值的下标位置(即 k 波长点位置所用积分时间)。采用二分搜索积分时间算法步骤为:

(1)若在 k 波长点处的待搜索的目标光谱能量值为 θ_k , 令 $x_{\text{low}} = x_{1k}, x_{\text{high}} = x_{nk}$;

(2)计算 $x_{\text{mid}} = \frac{1}{2}(x_{\text{low}} + x_{\text{high}})$, 将 θ_k 与 x_{mid} 比较,若 $x_{\text{mid}} \geq \theta_k$, 则 $x_{\text{high}} = x_{\text{mid}}$, 若 $x_{\text{mid}} < \theta_k$, 则 $x_{\text{low}} = x_{\text{mid}}$;

(3)重复步骤(2),直至 $x_{\text{mid}} = \theta_k$ 或者若数组 A 中不存在 $x_{\text{mid}} = \theta_k$, 则 x_{mid} 取与 θ_k 最接近的值,记录 x_{mid} 所取的值在数组 A 中的下标位置,该位置的值即为此波长处用于紫外可见光谱信号重构所用的积分时间;

通过对数组 A 的统计发现,在数组 A 中可能出现一个或者多个与带搜索的目标值相同的数,通常采用的方式是选择正序方向中出现第一个与待搜索的目标值相同数的下标。

2 紫外可见光谱信号重构算法

2.1 信号重构性能指标定义

(1)重构精度:表示重构的光谱信号与目标设定的光谱信号的吻合程度。针对光谱能量分布不均匀的问题,为了能够让微型光谱仪能够收集到更多的光谱能量信号,同时也为了避免采集的光谱能量信号超过仪器的量程,需要设定一组理想的目标信号能量值,让仪器自动调节积分时间,从而采集到更大的能量信号来解决光谱能量分布不均的问题。目标能量值变化的目的是为了能够在这些能量值中找出一个最优的目标能量值。可用式(2)来表示

$$p = 1 - \frac{\|\theta - \hat{\theta}\|_2}{\|\theta\|_2} \quad (2)$$

其中, p 为重构精度,其最大值为 1,最小值为 0; θ 为需要重构的目标信号, $\theta = [\theta_1, \theta_2, \dots, \theta_m]$, m 为波长点个数; $\hat{\theta}$ 为重构后的信号, $\hat{\theta} = [\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m]$ 。在规定目标信号 θ 时,可能会出现两种情况,第一种情况是:无论采用多大的积分时间值作为光谱仪的采样参数,获取到的能量信号均无法达到目标设定值;第二种情况是:当采用最小的积分时间作为光谱仪的采样参数,获取到的光谱能量信号值太大,其值超过了最小积分时间值下的光谱能量信号。为了提高重构精度,将目标信号 θ_i 用式(3)来表示

$$\theta_i = \begin{cases} y_{\text{re}}, & \theta_i > y_{\text{re}} \\ \theta_i, & y_{\text{ie}} \leq \theta_i \leq y_{\text{re}} \\ y_{\text{ie}}, & \theta_i < y_{\text{ie}} \end{cases} \quad (3)$$

其中, y_{re} 为在设定最大积分时间采样参数条件下获取到的第 i 个波长位置下的光谱能量信号值; y_{ie} 表示在设定最小积分时间采样参数条件下获取到的第 i 个波长位置的光谱能量信号值;通过对 $[1, 2, 3, \dots, m]$ 范围内的波长能量值进行统计分析,选定合适的重构目标信号,有助于提升重构精度。首先利用式(3)选择出最优的重构目标信号,然后结合式(2)的性能指标方程进行信号重构,重构出的光谱信号能够很直观的反应出算法的重构效果, p 值越接近 1, 则重构效果越好,所提方法亦越有效,反之亦然。

(2)重构特征显著度:表示重构后的光谱信号的图谱特征相对于重构前的光谱信号的图谱特征的区分程度。当采用较小的积分时间值作为微型光谱仪的采样参数,强吸光波长区间内的光谱能量值较小;当积分时间逐渐增大时,微型光谱仪采集到的强吸光波长区间内的光谱能量值逐渐增大,从而提高了图谱特征的显著程度。为了能更好地描述重构后的光谱能量信号的图谱特征情况,将整个波长范围内的波长点上的光谱信号都基于相同的权值,用式(4)来表示重构后的图谱特征显著度

$$d = \frac{1}{m} \sqrt{\sum_{i=1}^m \left(\frac{\hat{\theta}_i - y_{\text{ie}}}{y_{\text{ie}}} \right)^2} \quad (4)$$

其中, $\hat{\theta}_i$ 为第 i 个波长位置的重构后的光谱能量值。 D 值大于等于 0, 其值越大,表示重构后的图谱特征显著度大,全波段图谱特征越明显,有利于后续根据重构后的光谱信号进行分析待测金属离子种类和离子浓度信息。

2.2 紫外可见光谱信号重构算法步骤

(1)设定仪器积分采样参数范围 $[L_{\text{min}}, R_{\text{max}}]$, 搜索波长范围 $[\text{min}_{\text{index}}, \text{max}_{\text{index}}]$, 积分时间步长为 1, 以此为基础,采集不同积分时间下仅含有高浓度锌离子溶液(参比溶液)的光

谱能量信号，并记录于矩阵 DATA 中；

(2) 给定目标重构精度 p 和规定目标重构的光谱信号在各个波长点的目标能量值 \max_{top} ，其设定范围区间 $[\max_{top_L}, \max_{top_R}]$ ，步长为 1，并设定初始值 $\max_{top} = \max_{top_L}$ ；

(3) 根据式(3)计算出目标重构的光谱信号 θ ，采用二分搜索算法对 DATA 中的数据进行搜索查询匹配，找出各个波长点的目标积分时间及在该积分时间下的光谱能量值；

(4) 将 \max_{top} 自增 1，重复步骤(3)，直至 $\max_{top} = \max_{top_R}$ ，以求出所有目标能量值下的重构光谱信号；

(5) 根据式(2)计算出所有目标能量值下重构光谱信号的目标重构精度 p 值，并选择重构精度最高的光谱能量值作为信号重构的信息量；

(6) 将待测溶液按照步骤(5)得出的参比溶液在每个波长点的信号重构的积分时间，进行光谱能量信号重构；并根据式(5)计算出重构后的光谱吸光度信号；

$$A = \lg \frac{I_{O1}}{I_{O2}} \quad (5)$$

其中， A 为待测溶液吸光度； I_{O2} 为样品槽内放入的待检测溶液中含有除锌离子外的其他金属离子溶液的光谱能量信号； I_{O1} 为样品槽内放入的待检测的溶液中除锌离子外无其他金属离子溶液的光谱能量信号。

3 结果与讨论

为了验证所提算法的有效性，本文使用厦门奥谱天成公司生产的 ATP2000 微型光谱仪采集紫外可见光谱数据。首先根据一定的实验方法配置 Zn^{2+} 、 Cu^{2+} 、 Co^{2+} 和 Ni^{2+} 浓度的混合溶液，混合溶液的离子的各浓度情况如表 1 所示，其中序号 1 溶液为参比溶液，其他为待测溶液。然后设置光谱信号重构算法的参数： $p=96\%$ ， $L_{min}=1$ ， $R_{max}=21$ ， $min_{length}=40$ ， $max_{index}=1\ 184$ ， $max_{top_L}=[0.7 \times 2^{16}]$ ， $max_{top_R}=[0.8 \times 2^{16}]$ ；最后采集紫外可见光谱能量信号数据，根据本光谱信号重构算法的步骤来重构图谱特征更加明显、高质量的光谱信号。[]表示向下取整。

表 1 Zn^{2+} 、 Cu^{2+} 、 Co^{2+} 、 Ni^{2+} 混合离子溶液

Table 1 Zn^{2+} 、 Cu^{2+} 、 Co^{2+} 、 Ni^{2+} mixed ion solution

Index	$Zn^{2+}/$ ($g \cdot L^{-1}$)	$Cu^{2+}/$ ($mg \cdot L^{-1}$)	$Co^{2+}/$ ($mg \cdot L^{-1}$)	$Ni^{2+}/$ ($mg \cdot L^{-1}$)
1	16.0	0.0	0.0	0.0
2	16.0	0.4	1.6	0.5
3	16.0	0.6	1.0	0.7
4	16.0	0.8	0.6	0.1
5	16.0	1.0	0.2	0.7
6	16.0	1.4	0.8	0.3

首先使用微型光谱仪采集的积分时间为 1~21 ms 的参比溶液的光谱能量信号如图 1 所示。然后，按照本算法步骤对目标范围区间内的每个波长点的目标能量值进行参比溶液的光谱能量信号进行信号重构。图 2 为重构精度最高的参比溶液的光谱重构目标能量信号，图 3 为按照图 2 中目标信号

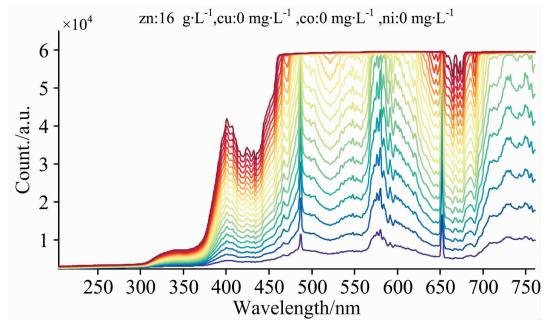


图 1 不同积分时间的参比溶液的光谱能量信号
Fig. 1 Spectral energy signals of reference solutions with different integration time

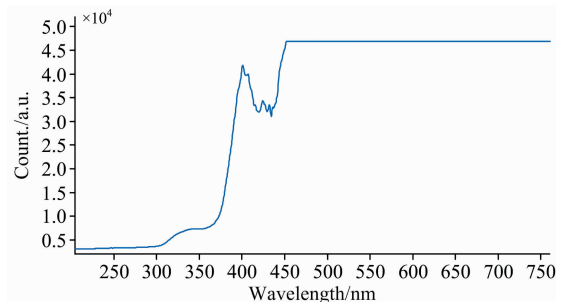


图 2 参比溶液的最优光谱重构目标信号
Fig. 2 Spectral reconstruction of target signal of reference solution

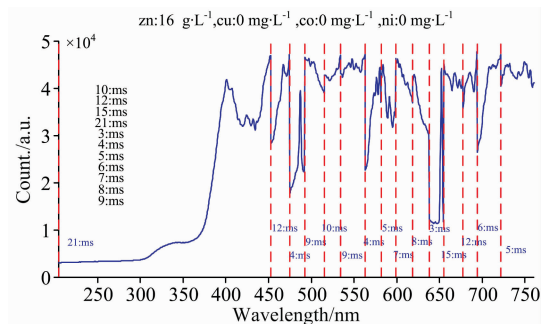


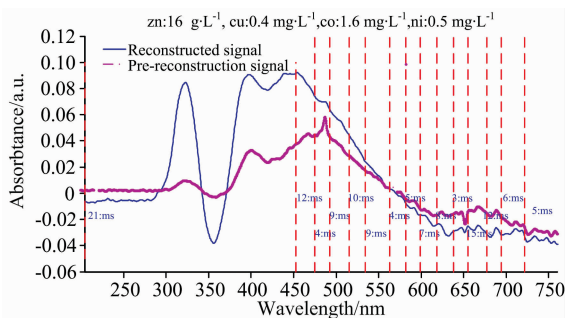
图 3 不同积分时间参比溶液的光谱能量重构信号
Fig. 3 Spectral energy reconstruction signal of reference solution with different integral time

重构出的不同积分时间参比溶液的光谱能量重构信号，图 4 为将光谱能量信号转化为吸光度图。从图 4 中可以看出，用虚线框内的重构前和重构后的光谱吸光度信号的图谱特征相差较大，重构后的光谱信号的谱峰幅值大，光谱曲线灵敏度更高，则易于辨识出该波长区间内的离子种类和离子浓度信息。对于 280~760 nm 波长区间的紫外可见光谱信号，重构后的吸光度信号的信噪比相对于重构前的信噪比高，在一定程度上有利于低信噪比光谱数据的分析处理，同时也有利于波长特征选择算法对图谱特征的分析与提取。

图 5 为将表 1 中所有待测溶液按照图 3 各个波长段选取的积分时间进行信号重构得到的紫外可见光谱能量信号图，

图 6 为全部混合溶液的紫外可见光谱吸光度重构信号。

结合图 5 和图 6 中的信号数据来计算本文所提出的性能指标,并将实验结果记录于表 2 中。由表 2 可知,重构后的光谱信号的信噪比相对于重构前的光谱信号的信噪比获得了提升,重构精度平均达到了 94.84%,同时图谱特征也获得



了很大程度增强光谱吸光度信号重构前后对比

Fig. 4 Comparison of spectral absorbance signal reconstruction before and after reconstruction

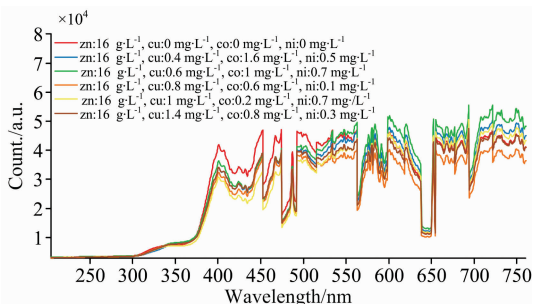


图 5 全部混合溶液的光谱能量重构信号

Fig. 5 Spectral energy reconstruction signals of all mixed solutions

3 结论

针对微型光谱仪采集高浓度比背景下痕量多金属离子浓

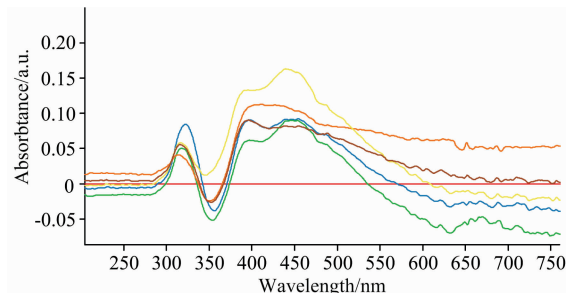


图 6 全部混合溶液的光谱吸光度重构信号

Fig. 6 Spectral absorbance reconstruction signal of all mixed solutions

表 2 重构谱信号的性能指标

Table 2 Performance indicators of reconstructed spectral signals

Index	Reconstruction precision/%	Reconstructed feature saliency
1	96.13	/
2	95.64	1.857 4
3	95.69	0.323 2
4	92.39	0.235 0
5	94.37	0.353 4
6	94.84	0.599 7

度的光谱信号时,由于选择不同积分时间采样参数而导致获取到的图谱特征差别大的问题,提出了一种基于二分搜索的高质量光谱信号重构算法。根据光谱特性定义了重构精度和重构特征显著度两个性能指标,然后提出了一种基于二分搜索的紫外可见信号重构方法。以实测混合溶液的紫外可见光谱信号为例,进行了信号重构的实验验证。实验结果表明,平均信号重构精度可达 94.84%、重构特征显著度也有所提升。重构后的光谱信号在噪声基本不变的情况下,吸光度信号得到增强,因此也会小幅度提升信噪比。本方法从信号处理的角度为光谱准确的检测分析提供了高信噪比、图谱特征更加明显的光谱信号。

References

- [1] Wang Guowei, Yang Chunhua, Li Yonggang, et al. Chemometrics and Intelligent Laboratory Systems, 2016, 151(15): 61.
- [2] Karami C, Alizadeh A, Taher M A, et al. Journal of Applied Spectroscopy, 2016, 83(4): 687.
- [3] Ren Xiaodong, Lei Wuhu, Wang Yong. Optic, 2019, 185: 497.
- [4] Chen Xuan. Computers & Electrical Engineering, 2018, 69: 865.
- [5] DUAN Li-guo, ZHU Li, LI Xiao-wei, et al(段利国,朱丽,李晓伟,等). Journal of Beijing University of Posts and Telecommunications(北京邮电大学学报), 2018, 41(2): 119.
- [6] ZHANG Jiang-tao, XIA Ke-wen, ZHOU Qiao, et al(张江涛,夏克文,周巧,等). Science Technology and Engineering(科学技术与工程), 2018, 18(9): 278.
- [7] Qi R, Yang D, Zhang Y, et al. Signal Processing, 2018, 153: 34.
- [8] CHEN Shan-xiong, HE Zhong-shi, XIONG Hai-ling, et al(陈善雄,何中市,熊海灵,等). Chinese Journal of Computers(计算机学报), 2015, 38(3): 614.
- [9] Yao S, Guan Q, Wang S, et al. EURASIP Journal on Wireless Communications and Networking, 2018, 2018(1): 78.
- [10] WANG Xin, HE Hao, FAN Xian-guang, et al(王昕,何浩,范贤光,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2018, 38(1): 93.

High-Quality UV-Vis Spectrum Signal Reconstruction Algorithms Based on Binary Search

ZHU Hong-qiu, HU Hao-nan, ZHENG Guo-liang, ZHOU Can*, LI Yong-gang

School of Automation, Central South University, Changsha 410083, China

Abstract In the background of high zinc, when the concentration of trace polymetallic ions is simultaneously measured by spectroscopy, the energy radiation of the light source of the micro-spectrometer is not uneven, and different ions in the mixed solution can selectively absorb the ultraviolet-visible light at different bands. Therefore, too long integration time may lead to the saturation of spectral energy value, and if the integration time is insufficient, the signal-to-noise ratio of the spectral signal may be very low. The selection of integral time depends on the existing experience of researchers and the characteristics of the UV-Vis spectral signals of ions to be measured. In order to automatically select the integral time of micro-spectrometer, a high-quality UV-Vis spectral signal reconstruction algorithm based on the binary search is proposed to reconstruct the UV-Vis spectral signal with more distinct characteristics composed of different integral time. Firstly, the UV-Vis spectral energy reference solution is collected at the different integral time. Then the spectral energy signal target values of the reference solution are given, and the appropriate integral time sampling parameters are found at each wavelength point by using the binary search algorithm. Next, according to the characteristics of UV-Vis spectrum, the reconstruction accuracy index which is determined to represent the degree of proximity between the reconstructed spectral energy value and the target setting value and the feature saliency index which is used to indicate the degree of feature saliency before and after signal reconstruction are defined. Finally, the spectral signal with the highest reconstructed accuracy in the search interval is selected as the reconstructed information. The reconstructed information is used to reconstruct the UV-Vis spectral energy value of the solution to be measured. The reconstructed spectral absorbance signal of the solution to be measured is obtained at the end. The experimental results show that the proposed algorithm can quickly and automatically select the integral target time to reconstruct the UV-Vis spectrum signal, and to obtain high quality UV-Vis spectrum signal. The accuracy of signal reconstruction can reach 94.84%, and the reconstructed feature is improved significantly. At the same time, the absorbance signal of the reconstructed UV-Vis spectral signal is enhanced. Compared with the UV-Vis spectral signal before reconstruction the signal-to-noise ratio of the spectral signal after reconstruction is greatly improved, and the problem of choosing the parameters of integral time based on the subjective judgement of researchers is avoided. It provides high quality model data for detecting the concentration information of various trace metal ions.

Keywords Micro-spectrometer; Signal reconstruction; Binary search; Integration time; Spectral energy value

(Received May 24, 2019; accepted Oct. 11, 2019)

* Corresponding author