

基于奇异值分解的 γ 能谱弱峰识别

陈 锋 周建斌 刘 易

(成都理工大学 核技术与自动化工程学院 成都 610059)

摘要 针对低放射性核素样品的 γ 能谱分析中的弱峰识别问题,提出了一种基于奇异值分解的 γ 能谱寻峰新方法。该方法通过改进矩阵的构造方式,将 γ 能谱升维为双向循环矩阵后进行奇异值分解,选择第二个奇异值进行矩阵重构,并根据重构后的矩阵寻峰。以放射源 ^{152}Eu 的 γ 能谱为实验对象,与一阶导数寻峰、对称零面积寻峰和其他矩阵的奇异值分解寻峰方法进行了对比。实验结果表明:双向循环矩阵奇异值分解寻峰方法,在检测弱峰时拥有更高的查全率、查准率和F1值,分别达到了100%、87%和0.94,优化了弱峰检测的效果,为寻峰方法提供更多的选择。

关键词 弱峰检测, 奇异值分解, 双向循环矩阵

中图分类号 TL99

DOI: 10.11889/j.0253-3219.2024.hjs.47.090404

Weak peak identification of gamma spectrum based on singular value decomposition

CHEN Feng ZHOU Jianbin LIU Yi

(College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Chengdu 610059, China)

Abstract [Background] When performing gamma-ray spectroscopy analysis of samples with low levels of radioactive nuclide content, the weak peaks are difficult to be identified. [Purpose] This study aims to propose a new method for identifying peaks in γ spectra by utilizing singular value decomposition (SVD) to improve the detection efficiency of weak peaks. [Methods] Firstly, the matrix construction of spectrum data was improved by transforming the γ spectrum into a two-way cyclic matrix, and singular value decomposition of matrix was performed to get singular values and singular vectors. Then, the second singular value was selected to reconstruct the matrix and perform peak finding. Finally, the γ spectrum of the radioactive source ^{152}Eu was used as the experimental object, the peak finding performance of proposed method was compared with that of first-order derivative peak finding, symmetric zero-area peak finding, and singular value decomposition peak finding. [Results] Comparison result show that the bidirectional circular matrix SVD peaking method has higher recall rate, precision rate, and F1 value, achieving 100%, 87% and 0.94, respectively. [Conclusions] The approach of this study can optimize weak peak detection and offer additional options for peak finding methods.

Key words Weak peak search, Singular value decomposition (SVD), Bidirectional circulant matrix

在 γ 能谱的数据分析中,识别并计算出各个峰位是能谱分析中的关键问题,因为这是判断样品中

是否存在某种核素的前提条件。然而,由于样品活度较低、特征峰所占分支比小、环境本底放射性强、

国家自然科学基金(No.12075038)、四川省自然科学基金(No.2024NSFSC0423)资助

第一作者: 陈锋, 男, 1999年出生, 2022年毕业于东华理工大学, 现为硕士研究生, 从事核电子学与仪器研究

通信作者: 周建斌, E-mail: zjb416@163.com

收稿日期: 2023-12-27, 修回日期: 2024-04-19

Supported by National Natural Science Foundation of China (No.12075038), Sichuan Natural Science Foundation (No.2024NSFSC0423)

First author: CHEN Feng, male, born in 1999, graduated from East China University of Technology in 2022, master student, engaged in nuclear electronics and instrumentation research

Corresponding author: ZHOU Jianbin, E-mail: zjb416@163.com

Received date: 2023-12-27, revised date: 2024-04-19

电子学噪声等因素的影响, γ 能谱的弱峰不易识别, 给定性分析带来了困难。实际测量中, 气溶胶、土壤、水和生物等样品^[1]的 γ 放射性核素通常种类多、活度低, 导致能谱上出现多个峰, 需要进行复杂的峰识别。因此, 提高弱峰识别能力, 可以避免漏检样品中的放射性核素, 在环境监测、核材料识别、核素追踪方面保证分析结果的完整性和可靠性。

目前, 寻峰方法主要有: 简单比较法、高斯乘积函数法、导数法、协方差法、对称零面积法、马尔科夫链法、形态学变换法、小波变换法等。文献[2]介绍并对比了简单比较法、高斯乘积函数法、导数法、协方差法、对称零面积法 5 种经典的寻峰方法, 并指出高本底情况下, 导数法和对称零面积法的抑制假峰和弱峰识别能力最好; 文献[3]通过对称零面积法成功识别了高统计涨落情况下的能谱峰位; 文献[4]提出了一种基于判断限的 γ 峰筛选方法, 结合二阶导数寻峰法, 提高了 γ 能谱中弱 γ 峰的准确检出率; 文献[5]将小波变换法与导数法结合, 进行联合寻峰; 文献[6]提出了一种基于数学形态学的寻峰算法, 并运用于复杂源项的核素分辨测量; 文献[7]提出了一种基于有限马尔科夫链的寻峰方法, 通过模拟一个量子球在不规则势垒上运动, 来找到能谱的局部最大值; 文献[8]通过实验指出马尔科夫链法有较高的灵敏度, 但容易被假峰误导。

奇异值分解 (Singular Value Decomposition, SVD) 是一种广泛应用于信号处理、图像处理领域的矩阵分析技术。SVD 可以将升维成矩阵的信号分解成一系列子信号, 这些子信号中有些主要包含有用信息, 另一些则主要包含无用信息。通过对代表有用信息的分量进行分析, 可以识别出奇异信号, 从而提取故障特征并进行故障诊断^[9]。离散信号构造成的二维矩阵有多种形式, 常见的构造方法有连续截断矩阵和 Hankel 矩阵^[10]。这两种矩阵常用于故障诊断领域, 通过提取信号的主要特征以判断信号是否存在故障^[11-12]。对这两种矩阵的 SVD 分析可用于检测信号中的奇异信号。奇异信号是指信号中出现不连续或者突变的现象, 例如信号的值在该点发生跳跃或者出现尖峰。

SVD 面对大型数据集时计算成本较高, 不便于在硬件上实现算法。但对于信号中幅值差异较大的频率成分有着较好的分解能力^[13]。因此, SVD 用于核物理实验的数据分析时, 通过对实验数据进行奇异值分解, 可以提取出有用的信号, 消除噪声和冗余, 提高数据的信噪比和分辨率。SVD 主要应用于能谱处理上的平滑降噪、学习模型的特征提取等方面。文献[14-15]将能谱构造为 Hankel 矩阵并进行

SVD 处理, 根据奇异谱的特征对 γ 能谱成分和噪声进行分离, 较大的奇异值对应 γ 能谱成分, 通过保留一部分较大的奇异值, 实现了能谱的平滑降噪, 得到了能谱数据中的主要信息; 文献[16]将待识别核素的能谱进行 SVD 处理, 将得到的特征向量作为该核素的特征, 并利用分类算法将特征与数据库进行对比识别, 构建了一个用于核素识别的机器学习模型, 从而实现无须设置参数的自动智能分析。

目前, SVD 在核能谱上的应用主要集中于谱线平滑以及机器学习的特征提取, 尚无文献探讨奇异值分解在能谱寻峰问题上的应用。基于机器学习的核素识别方式, 其弱峰识别能力取决于训练集的质量; 而一般的核素识别方式, 弱峰识别仍是难点。本文改进了用于 SVD 的矩阵构造方式, 基于 SVD 的幅值分解能力, 提出了一种基于双向循环矩阵 SVD 的寻峰方法, 在检测弱峰时拥有更高的查全率和查准率, 可有效提高弱峰识别能力。

1 奇异值分解的奇异性检测原理

1.1 矩阵构造

为了使数据特征在高维空间中更清晰地展现其本质的结构和规律, 并有效地将噪声与能量峰进行分离, 可以将一维数组的能谱升维成矩阵, 从而让数据在高维空间中更容易被分解。SVD 作为一种重要的矩阵分解方法, 可用于提取矩阵的特征信息, 其关键之处在于如何利用待处理信号构造合适的矩阵。

常用的构造方法有连续截断矩阵和 Hankel 矩阵。连续截断矩阵是将长度为 L 的周期原始信号均分成 n 段, 每个行向量长度为 (L/n) , 再由这些行向量排列成矩阵。连续截断矩阵进行 SVD 处理后, 对于每个奇异值对应的子矩阵, 将其行向量首尾相连, 可得到对应的信号分量 X_i ; Hankel 矩阵的构造方式则是截取长度为 n 的数据作为矩阵的一个行向量, 滑动截取 $(L-n+1)$ 个行向量, 再由这些行向量排列成矩阵。将 Hankel 矩阵 SVD 处理后, 可以通过计算每个奇异值所对应的子矩阵的第一行和最后一列, 便可得到对应的信号分量 X_i 。

将 SVD 方法应用于能谱寻峰时, 常用的连续截断矩阵与 Hankel 矩阵可能并不适合。从提高能谱完整性的角度出发, 严强强等^[17]提出构造循环矩阵进行 SVD 处理。循环矩阵的 SVD 用于降噪效果较好, 但是直接用于寻峰效果却一般, 且所需运算时间与内存空间较大。

本文提出一种改进循环矩阵构造方式, 即将能

谱设计为双向循环矩阵。相较于一般的循环矩阵，其数据不需要经历一遍完整的循环，且原始数据放置于中间行向量，进行双向循环。双向循环矩阵，可看作一种特殊的Hankel矩阵，它不仅具备Hankel矩

$$\mathbf{Y} = \begin{bmatrix} y_{n+1} & y_{n+2} & y_{n+3} & \cdots & y_L & y_1 & \cdots & y_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_2 & y_3 & y_4 & \cdots & y_{L-2} & y_{L-1} & y_L & y_1 \\ y_1 & y_2 & y_3 & \cdots & y_{L-3} & y_{L-2} & y_{L-1} & y_L \\ y_L & y_1 & y_2 & \cdots & y_{L-4} & y_{L-3} & y_{L-2} & y_{L-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{L-n+1} & y_{L-n+2} & y_{L-n+3} & \cdots & y_L & y_1 & \cdots & y_{L-n} \end{bmatrix} \quad (1).$$

在1 024道的双向循环矩阵中，原始能谱数据是第 $n+1$ 行；原始数据左移 i 位是第 $(n+1-i)$ 行；原始数据右移 i 位是第 $(n+1+i)$ 行，并保持移位方向相反、位数相同的行向量数量相同，确保之后在重构能谱中对应原始能谱峰位的点不发生偏移。

1.2 奇异值分解

奇异值分解是一种将 $m \times n$ 的矩阵 $\mathbf{A}$ 分解为三个矩阵相乘的方法，如式(2)所示：

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (2)$$

其中： $\mathbf{U}$ 、 $\mathbf{\Sigma}$ 、 $\mathbf{V}$ 分别为 $m \times m$ 的左奇异矩阵、 $m \times n$ 奇异值矩阵、 $n \times n$ 的右奇异矩阵，且满足式(3)的关系。

$$\begin{cases} (\mathbf{A}\mathbf{A}^T)\mathbf{u}_i = \lambda_i^u \mathbf{u}_i \\ (\mathbf{A}^T\mathbf{A})\mathbf{v}_i = \lambda_i^v \mathbf{v}_i \end{cases} \quad (3)$$

其中： $\mathbf{A}\mathbf{A}^T$ 的所有特征向量 $\mathbf{u}_i$ 构成了矩阵 $\mathbf{A}$ 的左奇异矩阵 $\mathbf{U}$ ， $\mathbf{A}^T\mathbf{A}$ 的所有特征向量 $\mathbf{v}_i$ 构成了矩阵 $\mathbf{A}$ 的右奇异矩阵 $\mathbf{V}$ ；两个矩阵对应的特征值分别为 λ_i^v 和 λ_i^u 。两个矩阵的非零特征值相等，如式(4)所示：

$$\lambda_i^u = \lambda_i^v = \lambda_i, 1 \leq i \leq r \quad (4)$$

奇异值矩阵 $\mathbf{\Sigma}$ 的对角线上由非负元素组成的对角矩阵 $\mathbf{D}$ ，如式(5)所示：

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{D} = \text{diag}(\sigma_1, \sigma_2, \sigma_3, \cdots, \sigma_r), r \leq \min(n, m) \quad (5)$$

不妨设 $n < m$ ，则其中 $\sigma_1, \sigma_2, \sigma_3, \cdots, \sigma_r$ 与 $\sigma_{r+1} = \cdots = \sigma_n = 0$ 是 $\mathbf{A}$ 的奇异值，且满足式(6)的关系：

$$\begin{cases} \sigma_i = \sqrt{\lambda_i}, 1 \leq i \leq r \\ \sigma_i = 0, r < i \leq n \end{cases} \quad (6)$$

利用奇异值分离，可以将原始信号 $\mathbf{X}$ 表示为若干个信号分量的线性组合，如式(7)所示：

$$\begin{cases} \mathbf{X}_i = \sigma_i(\mathbf{u}_i \mathbf{v}_i^T) \\ \mathbf{X} = \sum_{i=1}^r \mathbf{X}_i \end{cases} \quad (7)$$

将信号构造成矩阵后进行SVD处理，将奇异值

阵零相位性和脉冲宽度可控的优点，而且不会破坏能谱完整性。

矩阵如式(1)所示：

从大到小排列后，奇异值的大小则代表了对应信号分量所占比重，能谱中的特征峰幅值往往高于统计涨落，所以往往集中在较大的奇异值对应分量中。在较大的奇异值对应分量中，幅值高的成分占比较大；较小的奇异值对应分量中，幅值低的成分占比较小。

将一个高斯信号模拟的能量峰，构造双向循环矩阵并进行SVD处理后，所得前4个信号分量如图1所示。

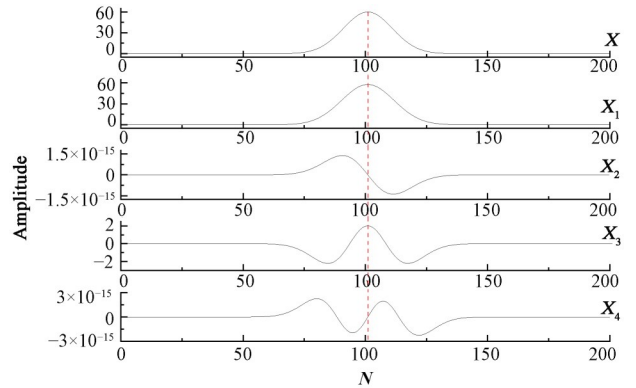


图1 双向循环矩阵各分量
Fig.1 Components of bidirectional circulant matrix

第一分量 $\mathbf{X}_1$ 往往对应于原始信号的“骨架”，其余的分量则对应于原始信号各频率成分，且从第二分量 $\mathbf{X}_2$ 开始，往后的分量形态愈加复杂。矩阵经SVD处理后信号分量 $\mathbf{X}_2$ 曲线形态上类似导数谱，但与导数谱不同的是，构造矩阵的 n 选取不同的值，分量 $\mathbf{X}_2$ 曲线过零点包含由正变负或由负变正两种情况。分量 $\mathbf{X}_2$ 曲线中对应能谱峰位的地方出现过零点，通过寻找过零点的道址，从而找到峰位的道址。

由于 γ 能谱的特征峰可看作一种奇异信号。为了检测奇异点的简便，采用第二分量 $\mathbf{X}_2$ 进行检测。对于找到的过零点，需要进行一定的处理。为了避免引入的假峰过多，必须借助一定的条件进行判别，剔除假峰。判别条件如式(8)所示：

$$\frac{S/W}{\sqrt{A/W}} > k \quad (8)$$

其中: W 为峰宽; k 为寻峰阈值。通常认为净峰面积 S 与峰面积 A 满足一定关系时^[1], 才认为该峰是有意义的峰。

1.3 参数选取

构造用于 SVD 的矩阵时, 所构造的矩阵不同, 参数选取方式也不同。对于连续截断矩阵、Hankel 矩阵的行数选取, 目前还没有通用的准则和方法, 相关文献针对不同情况推荐了不同方法^[18-21], 但都不适合应用在双向循环矩阵上。

在双向循环矩阵的构造中, 因为奇异值是矩阵非零特征值的算术平方根, 所以奇异值的数量等于非零特征值的数量。又因为矩阵每一行向量都相当于一组全谱数据, 即列向量的数量远大于行向量的数量, 并且存在统计涨落和连续数据离散化的影响, 实际中不存在值为 0 的特征值。所以根据式(6), 对于 $2n+1$ 行 L 列的双向循环矩阵来说, 秩为 $2n+1$, 其非零特征值个数也为 $2n+1$, 而奇异值的数量 r 也就等于 $2n+1$ 。行向量的数量 $2n+1$ 可以看作该双向循环矩阵下 SVD 的分解层数。

为了探究最佳的 n 如何取值使得分量 X_2 的占比最大, 实验使用 1 024 道的 NaI 数字 γ 能谱仪与 8 192 道的 HPGe 数字 γ 能谱仪, 分别对 ^{137}Cs 、 ^{152}Eu 样品进行测量。得到能谱 NaI-Cs、HPGe-Cs 和 HPGe-Eu 三份能谱数据。对三份能谱数据进行 SVD 处理后, 分别计算分量 X_2 的占比, 其计算如式(9)所示:

$$\frac{X_2}{X} = \frac{\sigma_2}{\sum_{i=1}^r \sigma_i} \quad (9)$$

令 $n=[1, 30]$, 分别计算不同 n 取值下分量 X_2 的占比, 结果如图 2 所示。

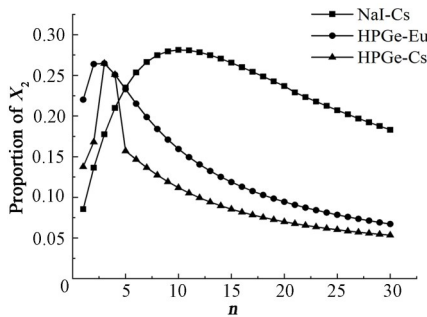


图2 不同 n 值下 X_2 的占比
Fig.2 Percentage of X_2 under different values of n

从图 2 可以看出, n 取值为 3 时, 长度为 8 192 的能谱数据 HPGe-Eu 与 HPGe-Cs 经 SVD 处理后, 分量 X_2 的占比达到最大; n 取值为 10 时, 长度为 1 024 的

能谱数据 NaI-Cs 经 SVD 处理后, 分量 X_2 的占比达到最大。相同数据长度下的不同核素能谱 n 取值相近, 不同数据长度下的同一核素能谱 n 取值不同, 可以看出数据长度对 n 的取值有一定影响。针对现有的各种探测器所得到的能谱数据长度, n 最佳取值范围为 $n=[3, 10]$ 。

1.4 评估方法

寻峰的目的旨在剔除假峰和识别真峰, 是一种二分类方法, 即对真峰和假峰进行分类。可用平均峰位偏移量、查准率 P 、查全率 R 以及 $F1$ 值作为评价其性能的指标。

查全率 R 和查准率 P 是矛盾的, 即当降低寻峰阈值来识别出全部真峰时会引入假峰, 查全率 R 升高的同时查准率 P 降低了, 反之亦然。 $F1$ 值是查准率 P 与查全率 R 的调和平均数, 同时兼顾了查准率 P 与查全率 R , $F1$ 值越高, 表示方法的效果越好。计算公式如式(10)所示:

$$\begin{cases} R = \frac{TP}{TP + FN} \\ P = \frac{TP}{TP + FP} \\ F1 = \frac{2 \times P \times R}{P + R} \end{cases} \quad (10)$$

本文将真峰设为正例, 假峰设为反例。其中: 真正例 TP 表示算法将真峰正确识别为“真峰”; 假正例 FP 表示算法将假峰错误识别为“真峰”; 假反例 FN 表示算法将真峰错误识别为“假峰”。查全率 R 表示寻峰算法识别的真峰占全部能量峰的比例; 查准率 P 是指寻峰算法正确识别真峰的比例。

寻峰方法不仅要求尽可能地识别出真峰, 而且同时要保证识别出的真峰道址偏移尽可能少。平均峰位偏移量是将算法识别出来的每一个真峰峰位与实际峰位的差值求平均数, 用来评价寻峰方法识别真峰峰位的精确程度。

2 实验结果对比分析

实验选用相对效率 40% 的同轴 HPGe (ORTEC GEM30-76) 作为探测器, 探头置于铅室内。数字信号处理系统使用的是四川新先达公司的 CTI-3800。由于放射源 ^{152}Eu 具有丰富的能量峰, 故选取 ^{152}Eu 的 γ 能谱为实验对象。放射源 ^{152}Eu 选用的是原子高科股份有限公司刻度源, 标定日期为 2021.10.25, 标定活度为 48 600 Bq, 不确定度为 5%。测量放射源 ^{152}Eu 的 γ 能谱, 放射源与探头距离 $d=4$ cm。为避免平滑算法对能谱处理之后对峰位产生影响, 增加对放射源的测量时间, 减少统计涨落对谱线的影响。

测量时间设为1 800 s,测量结果如图3所示。

在扣除本底的处理中,SNIP方法可以适应不同本底形状和噪声水平^[22],对特征峰的位置和面积保持较好,且算法简单、参数较少,易于计算机实现。因此,本文采用SNIP方法对特征峰区间本底进行扣除。利用SNIP法(窗宽设为70)对图3中的能谱数据进行本底扣除。

常用于 ^{152}Eu 核素识别的特征峰有13个,但是实际上 ^{152}Eu 的特征峰数量远不止于此。其他的特征峰因为其发生概率小,一般情况下难以观测到,常被忽略。许旭等^[23]对 ^{152}Eu 的能谱研究中,对高纯锗探测器所测的 ^{152}Eu 能谱中80多个峰详细说明了每个峰的来源与能量;文献[24]中列出了 ^{152}Eu 的127个 γ 峰。图3能谱数据中所探测到的放射源 ^{152}Eu 产生的能量峰经过与文献[23–24]中进行比对。确认了图3中有37个能量峰与 ^{152}Eu 有关。

对于图3中的强峰,各寻峰算法均能轻易识别,因此主要关注寻峰算法对其中弱峰的识别效果。对强峰与弱峰进行区分的方法通常有峰面积法、峰宽法以及峰高法。其中峰面积法与峰宽法,认为峰面积或峰宽小于某个阈值则认为是弱峰。这两种方法的区分效果受峰边界的选取影响,故本文采用峰高法与检出限进行区分。

在扣除本底的图3数据中计数最高的能量峰是121.85 keV的 ^{152}Eu 特征峰,计数为351 867。以它的1%作为弱峰的上限,以检出限(Detection Limit, DL)作为能谱下限。同样的测量条件下,对铅室的本底测量1 800 s,得到本底能谱后,计算它的计数平

均值(avg)与标准偏差(SD)。则检出限如式(11)所示。

$$\text{DL} = \text{avg} + (t \times \text{SD}) \quad (11)$$

在95%的置信水平下,经计算,DL结果为255.76。所以弱峰的峰高范围为[3 519,256]。低于该检出限的峰,认为是假峰。

经计算,筛选出23个弱峰,如图3所示。弱峰中的295.94 keV、488.75 keV、563.91 keV、586.13 keV、678.61 keV、688.64 keV、719.24 keV、810.44 keV、919.28 keV、989.07 keV、1 213.00 keV、1 249.98 keV、1 299.19 keV、1 457.58 keV为 ^{152}Eu 特征峰;85.41 keV为 ^{152}Sm 退激过程中产生的 $K_{\alpha}+K_{\beta}$ 和峰, ^{152}Sm 由 ^{152}Eu 轨道电子俘获产生;1 004.12 keV为 ^{154}Eu 特征峰, ^{154}Eu 为 ^{152}Eu 生产过程中的副产物,中子辐照天然铕元素(47.8%的 ^{151}Eu 和52.2%的 ^{153}Eu)获得 ^{152}Eu 和 ^{154}Eu ;755.39 (411.07+344.23) keV,1 122.98 (778.90+344.23) keV,1 233.81(1112.12+121.85) keV,1 433.90 (1 089.91+344.23) keV,1 529.80 (1 408.08+121.85) keV,1 643.27(1 299.19+344.23) keV为 ^{152}Eu 的和峰,其中1 529.80 keV的峰与1 528.10 keV的 ^{152}Eu 特征峰发生重合。

2.1 阈值选取

无论是用何种方法进行寻峰,都有可能找到许多假峰或者统计上无意义的峰,因此需要利用式(8)进行进一步的甄别,其中寻峰阈值的设置尤为重要。若将阈值设置过高,则会识别强峰的同时,将假峰和弱峰一起剔除了;设置过低,则会识别强峰与弱峰的同时,引入大量假峰。为了探究不同方法对弱峰识

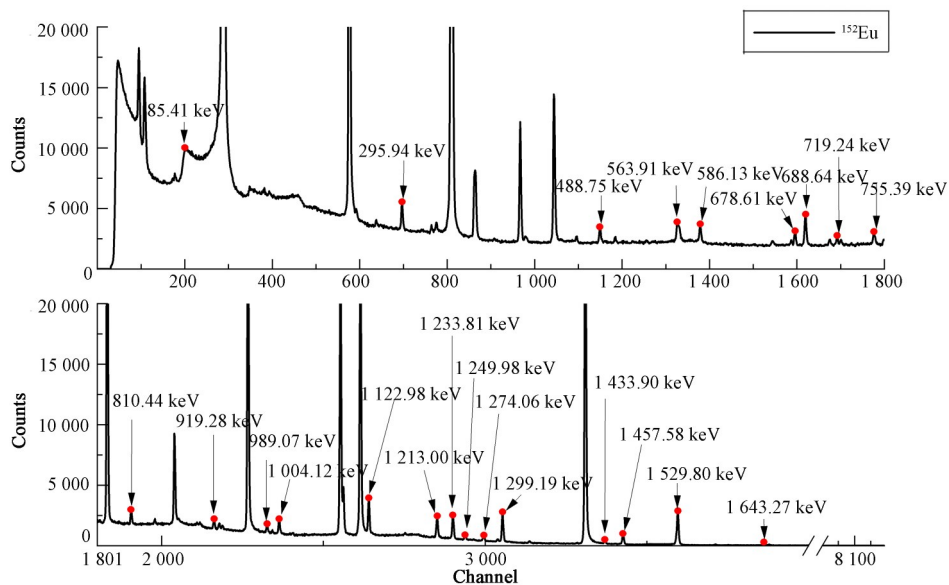


图3 ^{152}Eu 能谱
Fig.3 Spectrum of ^{152}Eu

别的能力差异,将对不同的寻峰方法设置不同的寻峰阈值,使各方法保留下来的峰数量趋于相等,然后在峰数量差别不大的情况下,再对比分析各方法保留的峰。

在图3的能谱数据中,选取不同的 k 值,对应 k 值下各方法识别到的峰数量如图4所示。

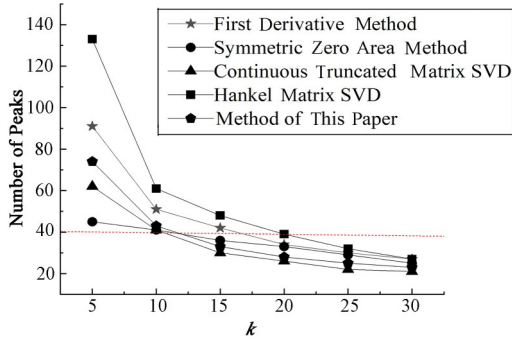


图4 不同 k 值下识别到的峰数量
Fig.4 Number of peaks identified at different k -values

从图4看出,当 k 的取值相同时,不同寻峰方法识别到的峰数量各不相同。为了排除 k 的取值对各寻峰方法的影响,根据检测到的37个能量峰,将各寻峰方法识别的峰数量均控制在40左右。因此,对称零面积法、连续截断矩阵SVD和双向循环矩阵SVD的 k 取值为10;3次7点一阶导数法的 k 取值为15;Hankel矩阵SVD的 k 取值为20;同时将能谱导入GammaVision 6中进行寻峰。

2.2 寻峰结果与分析

对图3中的数据扣除本底后,根据式(1),取 $n=3$,构造 7×8192 的双向循环矩阵;并构造了 $4 \times (8192/4)$ 的连续截断矩阵、 $13 \times (8192-13+1)$ 的Hankel矩阵。对于得到的3个矩阵分别进行SVD处理得到分量 X_2 ,按照式(8)剔除假峰,同时以 γ 能谱分析软件GammaVision 6的分析结果为对照,在保证各寻峰方法对于强峰均能正确识别的情况下,对比分析它们各自的弱峰识别和引入假峰情况。在各寻峰方法的结果中,对应能量是 ^{152}Eu 特征能量但并非选取的27个弱峰之一的峰,本文不作考虑,不列入对比结果中;低于检出限的峰以及与 ^{152}Eu 特征能量比对失败的峰作为假峰。结果如表1所示。

经过计算,从表1可得出各方法的指标结果,如表2所示。

由表2可知,双向循环矩阵SVD方法综合表现最优,能够识别出所有的弱峰,且假峰数量最少,查准率、查全率和F1值最高,分别达到了87%、100%和0.94,平均峰位偏移只有0.65道,说明该方法具有较高的准确性和稳定性。

在使用SVD寻峰方法中,对比三种不同的矩阵构造方式可以看出,矩阵的构造方式对SVD方法的峰识别性能有显著的影响,双向循环矩阵SVD方法成功识别到了23个弱峰,只出现了3个假峰,F1值为0.94,比连续截断矩阵的F1值提升了10.59%,比Hankel矩阵的F1值提升了56.67%。实验结果说明,矩阵构造方式的改进是成功的,F1值有明显提高,有效地提高了弱峰识别的敏感性和特异性。

在双向循环矩阵SVD方法与一阶导数法、对称零面积法的比较分析中发现,一阶导数法和对称零面积法是两种基于峰形特征的峰识别方法,它们的优点是对峰位的定位能力较强,平均峰位偏移都只有0.11道,低于双向循环矩阵SVD方法。然而,它们的缺点是对噪声敏感,容易产生假峰,以及容易漏检弱峰。一阶导数法识别出19个弱峰,产生了7个假峰;对称零面积法识别出18个弱峰,产生了3个假峰。相比之下,双向循环矩阵SVD法在弱峰数量、假峰数量、查准率、查全率和F1值方面都优于这两种方法。

从对比结果看,双向循环矩阵SVD方法引入的假峰数量比GammaVision 6的分析结果多1个,导致查准率较低,但是查全率与F1值均高于GammaVision 6。根据GammaVision 6提供的操作手册可知,其能谱平滑方式是采用5点重心法。重心法平滑数据,会使峰的幅值变小,对较弱的特征峰幅值影响较大,容易丢失弱峰。双向循环矩阵SVD方法则是将一个峰进行分解,保留了峰的“骨架”信息,对幅值影响较小。

综上所述,双向循环矩阵SVD方法是一种适合 γ 能谱定性分析的寻峰方法,它有良好的弱峰检测性能和应用潜力。

3 结语

提出了一种基于SVD的 γ 能谱寻峰方法,提高了弱峰检测能力。论文对SVD矩阵构造方式进行了改进,将能谱构建为双向循环矩阵;对不同数据长度的能谱进行实验,探究矩阵构建参数 n 最佳取值范围;同时通过HPGe探测器获取 ^{152}Eu 样品的 γ 能谱,以其中的23个弱峰为对象,与其他两种矩阵SVD方法、其他两种常用的寻峰方法以及谱分析软件GammaVision 6进行了对比分析。

实验结果表明,双向循环矩阵SVD方法的参数 n 取值只与能谱的数据长度有关,与核素种类无关,矩阵构建参数 n 最佳取值范围为3~10。双向循环矩阵SVD方法成功地识别出能谱中的23个弱峰,只出现了两个假峰,其查全率达到100%,查准率达到

表1 不同处理方法下的 ^{152}Eu 谱寻峰结果
Table 1 The peak-finding results of ^{152}Eu spectrum under different processing methods

能量 Energy / keV	道址 Channel	所寻峰位 Peak Position					
		一阶导数法 ^[1] First deriva- tive method	对称零面积 法 ^[2] Symmetric zero area method	连续截断矩 阵SVD ^[9] Continuous truncat- ed matrix SVD	Hankel矩阵SVD ^[9] Hankel matrix SVD	双向循环矩 阵SVD Bidiagonal cyclic matrix SVD	GammaVision 6.0
85.41	200	200	—	197	—	202	—
295.94	697	697	697	696	697	697	696
488.75	1 149	1 150	1 150	1 148	1 150	1 150	1 149
563.91	1 326	1 327	1 326	1 325	1 327	1 328	1 327
586.13	1 379	1 379	1 379	1 377	1 379	1 379	1 378
678.61	1 596	1 596	1 596	1 596	1 596	1 596	1 595
688.64	1 620	1 620	1 619	1 620	1 620	1 620	1 619
719.24	1 692	—	1 692	1 692	—	1 692	—
755.39	1 776	1 776	1 776	1 776	—	1 777	1 776
810.44	1 905	1 905	1 905	1 904	1 906	1 906	1 905
919.28	2 162	—	2 162	2 160	—	2 162	2 161
989.07	2 326	2 326	2 326	2 325	—	2 326	2 325
1 004.12	2 363	2 363	2 363	2 364	2 365	2 364	2 362
1 122.98	2 640	2 640	2 640	2 640	2 641	2 641	2 639
1 213.00	2 852	2 852	2 852	2 852	2 852	2 852	2 851
1 233.81	2 900	2 900	2 900	2 900	2 901	2 901	2 899
1 249.98	2 938	—	—	2 937	—	2 939	—
1 274.06	2 996	2 996	—	2 996	—	2 997	2 995
1 299.19	3 054	3 054	3 054	3 052	3 054	3 054	3 053
1 433.90	3 370	—	—	3 369	—	3 371	3 369
1 457.58	3 426	3 426	3 426	3 425	3 427	3 427	3 425
1 529.80	3 596	3 596	3 596	3 596	3 596	3 596	3 594
1 643.27	3 862	3 862	—	3 861	—	3 863	3 861
假峰 False peak	—	47; 68; 205; 2 030; 2 571; 2 862; 2 890	45; 62; 78;	44; 269; 592; 1 096	64; 66; 75; 78; 80; 856; 1 331; 2 633; 3 045; 3 583	48; 3 294	45

表2 不同处理方法的性能
Table 2 The performance of different processing methods

	弱峰 Weak peak	假峰 False peak	平均峰位偏移 Average peak shift	查准率 Precision	查全率 Recall	F1 值 F1 score
一阶导数法 First derivative method	19	7	0.11	0.73	0.83	0.78
对称零面积法 Symmetric zero area method	18	3	0.11	0.86	0.78	0.82
连续截断矩阵SVD Continuous truncated matrix SVD	23	4	0.83	0.85	1.00	0.91
Hankel矩阵SVD Hankel matrix SVD	14	10	0.57	0.58	0.61	0.60
双向循环矩阵SVD Bidiagonal cyclic matrix SVD	23	2	0.65	0.92	1.00	0.94
GammaVision 6	20	1	0.90	0.95	0.87	0.91

87%, F1 值为 0.94, 这说明本方法具有较高的弱峰检测性能, 矩阵构造方式改进是有效的, 与其他两种常用的寻峰方法对比, 在定性分析中有一定的优势。

作者贡献声明 陈锋提出研究思路并负责实验数据采集分析以及论文起草; 周建斌负责最终版本修订; 刘易负责论文修改。

参考文献

- 1 陈伟, 周建斌, 方方, 等. XRF 能谱多重 Sallen-Key 平滑滤波方法及其在大米镉含量分析中的应用[J]. 核技术, 2022, **45**(2): 020202. DOI: 10.11889/j.0253-3219.2022.hjs.45.020202.
CHEN Wei, ZHOU Jianbin, FANG Fang, *et al.* Multiple Sallen-Key smoothing filter method of XRF energy spectrum and its application in rice cadmium content analysis[J]. Nuclear Techniques, 2022, **45**(2): 020202. DOI: 10.11889/j.0253-3219.2022.hjs.45.020202.
- 2 庞巨丰. γ 能谱数据分析[M]. 西安: 陕西科学技术出版社, 1990: 544 - 546, 577 - 607.
PANG Jufeng. γ -ray spectrum data analysis[M]. Xi'an: Shaanxi Science & Technology Press, 1990: 544 - 546, 577 - 607.
- 3 陈川, 葛良全, 谷懿, 等. 采用对称零面积法的高统计涨落谱线峰位解析研究[J]. 核电子学与探测技术, 2016, **36**(2): 229 - 231. DOI: 10.3969/j.issn.0258-0934.2016.02.027.
CHEN Chuan, GE Liangquan, GU Yi, *et al.* Research on the spectral peak position based on high statistical fluctuation by symmetric zero-area method[J]. Nuclear Electronics & Detection Technology, 2016, **36**(2): 229 - 231. DOI: 10.3969/j.issn.0258-0934.2016.02.027.
- 4 张新军, 王世联, 贾怀茂, 等. 基于判断限的高灵敏度 γ 峰筛选方法[J]. 核技术, 2019, **42**(7): 070202. DOI: 10.11889/j.0253-3219.2019.hjs.42.070202.
ZHANG Xinjun, WANG Shilian, JIA Huaimao, *et al.* A high sensitivity screening method for γ peaks based on the critical level[J]. Nuclear Techniques, 2019, **42**(7): 070202. DOI: 10.11889/j.0253-3219.2019.hjs.42.070202.
- 5 汪雪元, 何剑锋, 刘琳, 等. 小波变换导数法 X 射线荧光光谱自适应寻峰研究[J]. 光谱学与光谱分析, 2020, **40**(12): 3930 - 3935. DOI: 10.3964/j.issn.1000-0593(2020)12-3930-06.
WANG Xueyuan, HE Jianfeng, LIU Lin, *et al.* Research on adaptive peak detection of X-ray fluorescence spectrum with wavelet transform and derivative method[J]. Spectroscopy and Spectral Analysis, 2020, **40**(12): 3930 - 3935. DOI: 10.3964/j.issn.1000-0593(2020)12-3930-06.
- 6 李涅, 黄长添, 毕明德, 等. 基于形态学变换的核素能谱寻峰算法[J]. 核电子学与探测技术, 2016, **36**(2): 171 - 173, 178. DOI: 10.3969/j.issn.0258-0934.2016.02.013.
LI Sheng, HUANG Zhangtian, BI Mingde, *et al.* Nuclide spectrum peak searching algorithm based on the morphology transform[J]. Nuclear Electronics & Detection Technology, 2016, **36**(2): 171 - 173, 178. DOI: 10.3969/j.issn.0258-0934.2016.02.013.
- 7 Silagadze Z K. A new algorithm for automatic photopeak searches[J]. Nuclear Instruments and Methods in Physics Research A, 1996, **376**(3): 451 - 454.
- 8 任俊松. 伽马能谱分析与核素识别方法研究[D]. 绵阳: 西南科技大学, 2017.
REN Junsong. The study of γ spectrum analysis and radioactive nuclide identification method[D]. Mianyang: Southwest University of Science and Technology, 2017.
- 9 周桂, 李华, 黄涛, 等. 基于奇异值分解理论的机械故障诊断方法综述[J]. 现代制造工程, 2022(12): 144 - 151. DOI: 10.16731/j.cnki.1671-3133.2022.12.020.
ZHOU Gui, LI Hua, HUANG Tao, *et al.* A review of mechanical fault diagnosis methods based on singular value decomposition theory[J]. Modern Manufacturing Engineering, 2022(12): 144 - 151. DOI: 10.16731/j.cnki.1671-3133.2022.12.020.
- 10 赵学智, 叶邦彦, 陈统坚. 矩阵构造对奇异值分解信号处理效果的影响[J]. 华南理工大学学报(自然科学版), 2008, **36**(9): 86 - 93. DOI: 10.3321/j.issn:1000-565X.2008.09.018.
ZHAO Xuezhi, YE Bangyan, CHEN Tongjian. Influence of matrix creation way on signal processing effect of singular value decomposition[J]. Journal of South China University of Technology (Natural Science Edition), 2008, **36**(9): 86 - 93. DOI: 10.3321/j.issn:1000-565X.2008.09.018.
- 11 曾作钦, 赵学智. 一种基于奇异值分解的奇异性检测新方法[J]. 沈阳工业大学学报, 2011, **33**(1): 102 - 107.
ZENG Zuoqin, ZHAO Xuezhi. A new method for singularity detection based on singular value decomposition[J]. Journal of Shenyang University of Technology, 2011, **33**(1): 102 - 107.
- 12 Zhao X Z, Ye B Y. Similarity of signal processing effect between Hankel matrix-based SVD and wavelet transform and its mechanism analysis[J]. Mechanical Systems and Signal Processing, 2009, **23**(4): 1062 - 1075.

- DOI: 10.1016/j.ymsp.2008.09.009.
- 13 赵学智, 叶邦彦, 陈统坚. 频率添加奇异值分解算法及其在故障特征提取中的应用[J]. 机械工程学报, 2021, **57**(10): 10 - 20. DOI: 10.3901/JME.2021.10.010.
ZHAO Xuezhi, YE Bangyan, CHEN Tongjian. Frequency adding singular value decomposition algorithm and its application in fault feature extraction[J]. Journal of Mechanical Engineering, 2021, **57**(10): 10 - 20. DOI: 10.3901/JME.2021.10.010.
 - 14 赵思文, 吴怡, 王崇杰. 基于奇异谱分析的 γ 能谱降噪算法[J]. 物理实验, 2021, **41**(9): 11 - 17, 22. DOI: 10.19655/j.cnki.1005-4642.2021.09.002.
ZHAO Siwen, WU Yi, WANG Chongjie. Denoising algorithm of gamma energy spectra based on singular spectrum analysis[J]. Physics Experimentation, 2021, **41**(9): 11 - 17, 22. DOI: 10.19655/j.cnki.1005-4642.2021.09.002.
 - 15 金久强, 王金龙, 肖刚毅, 等. NASVD方法在航空 γ 能谱测量中的成像效果讨论[J]. 物探与化探, 2018, **42**(4): 817 - 824. DOI: 10.11720/wtyht.2018.1068.
JIN Jiuqiang, WANG Jinlong, XIAO Gangyi, *et al.* Airborne gamma ray spectrum mapping effects with NASVD de-noising[J]. Geophysical and Geochemical Exploration, 2018, **42**(4): 817 - 824. DOI: 10.11720/wtyht.2018.1068.
 - 16 任俊松, 张江梅, 王坤朋. 基于SVD和SVM的核素识别算法[J]. 兵工自动化, 2017, **36**(5): 50 - 53. DOI: 10.7690/bgzd.2017.05.014.
REN Junsong, ZHANG Jiangmei, WANG Kunpeng. Radioactive nuclide identification method based on SVD and SVM[J]. Ordnance Industry Automation, 2017, **36**(5): 50 - 53. DOI: 10.7690/bgzd.2017.05.014.
 - 17 严强强, 盛守照, 周俊, 等. 一种改进奇异值分解降噪方法研究[J]. 电光与控制, 2018, **25**(9): 22 - 25, 41. DOI: 10.3969/j.issn.1671-637X.2018.09.005.
YAN Qiangqiang, SHENG Shouzhao, ZHOU Jun, *et al.* A noise reduction method based on improved singular value decomposition[J]. Electronics Optics & Control, 2018, **25**(9): 22 - 25, 41. DOI: 10.3969/j.issn.1671-637X.2018.09.005.
 - 18 Hassani H, Mahmoudvand R, Zokaei M, *et al.* On the separability between signal and noise in singular spectrum analysis[J]. Fluctuation and Noise Letters, 2012, **11**: 1250014. DOI: 10.1142/S0219477512500149.
 - 19 Zhigljavsky A. Singular spectrum analysis for time series: introduction to this special issue[J]. Statistics and Its Interface, 2010, **3**(3): 255 - 258.
 - 20 Mahmoudvand R, Zokaei M. On the singular values of the Hankel matrix with application in singular spectrum analysis[J]. Chilean Journal of Statistics, 2012, **3**(1): 43 - 56.
 - 21 Khan M R, Poskitt D S. A note on window length selection in singular spectrum analysis[J]. Australian and New Zealand Journal of Statistics, 2013, **55**(2): 87 - 108.
 - 22 吴和喜, 袁新宇, 刘庆成, 等. 一种 γ 谱弱峰优化检测方案研究[J]. 原子能科学技术, 2012, **46**(9): 1142 - 1146. DOI: 10.7538/yzk.2012.46.09.1142.
WU Hexi, YUAN Xinyu, LIU Qingcheng, *et al.* Optimal scheme for detecting weak peak in γ -ray spectrum[J]. Atomic Energy Science and Technology, 2012, **46**(9): 1142 - 1146. DOI: 10.7538/yzk.2012.46.09.1142.
 - 23 许旭, 刘佳强, 陆景彬, 等. 放射源 ^{152}Eu 的能谱识别及和峰特性分析[J]. 同位素, 2018, **31**(1): 1 - 7. DOI: 10.7538/tws.2017.youxian.008.
XU Xu, LIU Jiaqiang, LU Jingbin, *et al.* Recognition of spectrum in ^{152}Eu source and analysis of sum peak characteristics[J]. Journal of Isotopes, 2018, **31**(1): 1 - 7. DOI: 10.7538/tws.2017.youxian.008.
 - 24 刘运祚. 常用放射性核素衰变纲图[M]. 北京: 原子能出版社, 1982: 205 - 210.
LIU Yunzuo. Decay outline of commonly used radionuclides[M]. Beijing: Atomic Energy Press, 1982: 205 - 210.