Journal of Surface Analysis Vol.14, No. 4 (2008) pp. 365-369

Original Paper

Performance analysis of algorithms to detect peaks in XPS spectra

Yoichiro Furukawa[a], N.Ikeo [b] ,Y.Nagatsuka [b], M.Yoshitake [c], A.Tanaka [d]
[a]*DENKI KAGAKU KOGYO K.K.,3-5-1 Asahimachi, Machida-shi,Tokyo 194-8560*
[b]*JEOL Ltd.,3-1-2 Musashino,Akishima,Tokyo 196-8558*
[c]*National Institute for Materials Science,3-13 Sakura,Tsukuba,Ibaraki 305-0003*
[d]*ULVAC-PHI,Inc.,370 Enzo,Chigasaki,Kanagawa 253-8522*
*(*yoichiro-furukawa@denka.co.jp)*

(Received: November 26, 2007; Accepted: March 04, 2008)

Performance of peak detecting algorithms for AES and XPS spectrum is studied by varying the parameters such as the critical value, k, a number of averaging points in the background intensity, and a Savitzky-Golay coefficient in the second derivative treatment. The proposed peak finding methods are composed of three stages of algorithms: rough estimation of the background; application of the second derivative curve; and direct calculation of the peak and background relation at the candidate peak. Software exploiting the three algorithms is used, which enables to find a peak automatically. Efficiency and characteristics of the respective algorithms for finding a peak are discussed.

1. Introduction

The presence of peaks in the spectrum must first be determined as the primary source of information for a qualitative or quantitative analysis in XPS, especially for a making a decision on the presence of elements.

Surface chemical analysts use many different methods to detect peaks in data, but as there may be no uniquely defined perfect method, analysts sometimes pick peaks visually. There are different ways to write algorithms to detect peaks, but their algorithms are not sufficiently clarified, nor do they always work for poorly resolved peaks. Since each method has relative merits, it is desirable for analysts to understand each algorithm and its performance in order to select the best or alternate method[1],[2].

In the activity that relates to the international standardization in ISO TC201 SC3 (Surface Chemical Analysis - Data Management and Treatment), three peak detecting methods have been proposed and discussed[3].

In this study, we performed optimization on an parameter value using a software exploiting the three algorithms which enables to find a peak automatically in XPS spectrum of Au. Efficiency and characteristics of the respective algorithms for finding a peak are discussed

2. Proposed Peak Detection Methods

Outline of three algorithms to detect peaks are as follows.

2.1 Peak Detection Using Rough Estimation of Spectrum Background

As the background intensity changes rather gently compared with the intensity near the peak, it can be approximately expanded by using the *2m+1* points of data which cover the region with several times of the typical full width at half maximum of the peak, *w.*

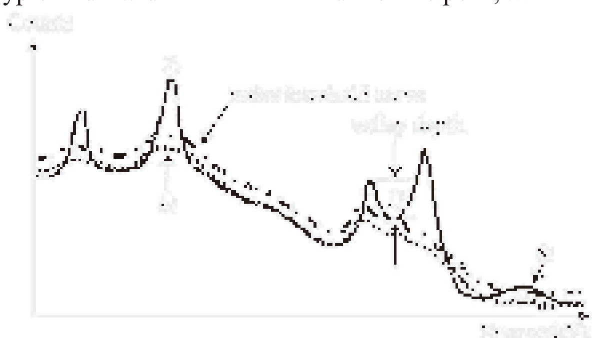

Fig.1 Schematic diagram of y_i , b_i and $b_i+k\sigma_{ni}$.

If the variance of $n_i = y_i - b_i$ is defined as σ_{ni}^2, it can be estimated as follows:

$$\sigma_{ni}^2 = \sigma_{yi}^2 + \sigma_{bi}^2 = y_i + \frac{1}{(2m+1)^2}\sum_{j=-m}^{m} y_{i+j} \tag{1}$$

The inequality to judge a peak is given by using the critical value $k \approx 2 \sim 3$,

$$n_i > k\sigma_{ni} = k\sqrt{y_i + \frac{1}{(2m+1)^2}\sum_{j=-m}^{m} y_{i+j}} \tag{2}$$

or, in a more familiar expression,

$$yi > bi + k\sigma_{ni} \tag{3}$$

where $bi + k\sigma_{ni}$ defined as the noise threshold curve. In statistical means the degree of reliability over the adoption of 2 and 3 for k value is estimated 95.4%(2 σ)and 99.7%(3 σ) respectively.

A broad and small peak, which satisfies the following inequality, is determined to be detected if its S_i is greater than a certain value S_{i0},

$$S_i > S_{i0}, \tag{4}$$

where the value of S_{i0} (peak area) is roughly estimated by assuming it as an ideal triangle peak with the typical full width w at half maximum of the peak and height $k\sigma_{ni}$.

For poorly resolved peaks with deep valleys, the following data processing is effective. If the spectrum has plural peaks above noise threshold curve $k\sigma_{ni}$ and valleys that do not cross the noise threshold curve, and if valley depth D of the minor local maximum in the spectra exceeds the noise fluctuation $k\sigma_{ni}$, the peaks are regarded as real.

$$|D| > k\sigma_{ni} \tag{5}$$

2.2 Peak Detection Using Threshold Curve of The Second Derivative

This method has the same effect as subtracting the background from the spectrum, by making use of the second derivative. And it has no arbitrariness in the background subtracting procedure, it may be a relatively convenient algorithm when we use it with the aid of computers.

As the peak in the spectrum corresponds to the local minimum of the spectrum, we judge the peak is real one if the local negative minimum d_{min} is less than (in absolute values, greater than) its noise fluctuation. Therefore, for the peak judge inequality, we admit the peak is detected at the position that gives d_{min} in the second derivative if the following inequality is satisfied,

$$d_{min} < k\sigma_i \tag{6}$$

where $k\sigma_i$ defined as the noise threshold curve.

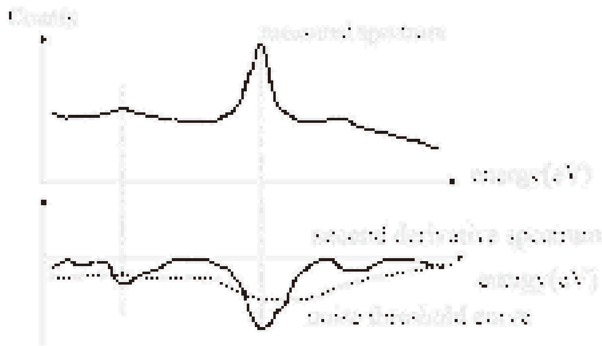

Fig.2 Schematic diagram of a peak position, the second derivative curve and threshold noise curve.

2.3 Peak Detection by Directly Calculating Peak and Background Relations

In this method, the candidate peaks are detected by the second derivative treatment and judged by comparison with criteria.

Let P denote the peak intensity with the background and N the net peak intensity, then, N=P-B, and the variance σN2 of N is calculated as follows:

$$\sigma_N^2 = \sigma_P^2 + \sigma_B^2 \tag{7}$$

$$\text{where} \quad \sigma_P^2 = P,\ \sigma_{B1}^2 = B_1,\ \sigma_{B2}^2 = B_2, \tag{8}$$

assuming that the background intensity B at the peak position is calculated as

$$B = (B_1 l_2 + B_2 l_1)/(l_1 + l_2), \tag{9}$$

and that the statistics of pulse-counting detection obeys the Poisson distribution.

Then, σ_N^2 is calculated as follows:

$$\sigma_N^2=P+(B_1l_2^2+B_2l_1^2)/(l_1+l_2)^2 \text{ ,} \tag{10}$$

The peak decision condition is given as follows:

$$N>k\sigma_N \tag{11}$$

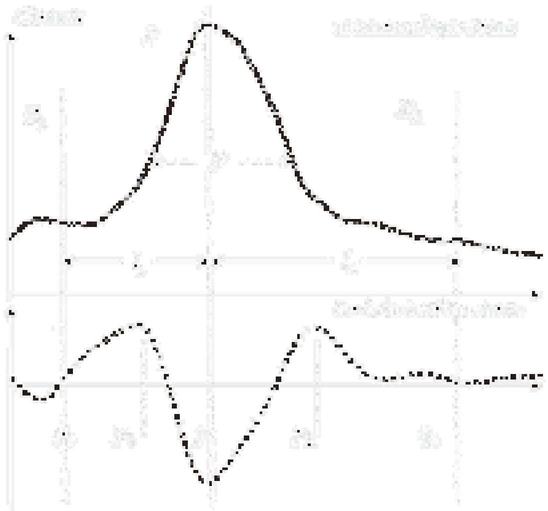

Fig.3 Schematic diagram of a peak position and its background at both sides of the peak.

3. Optimization of Peak Detection

3.1 Optimization of detection condition in peak detection algorithm using rough estimation of spectrum background

An optimum parameter value using a XPS spectrum of *Au* as shown in Fig.4 were studied by comparing the number of peaks detected with those of peaks detected visually. The threshold coefficient, *k*, was changed from 2.0, 2.5 to 3.0 in order. The smoothing point number was changed from 21 (m=10), 31 (m=15) to 41 (m=20) so as to estimate the background for each threshold coefficient. A detected peak number was counted under the respective conditions defined by combination of each parameter. A S_{iO} value, a value for detecting a broad small peak, was shifted from 1, 2, 5 to 10 (*eV*) in order, and the detection efficiency was studied.

It was found that k=3 was the optimum value, since the detection peak number exceeds 25 in case of k=2.0 and k=2.5. It was also found that the optimum smoothing point number was 21, since the *5p1/2 (750eV)* peak was not detected when the number was 41, and noise peaks increased by 2 in case of 31 and 41.

When a S_{iO} value is 1*eV,* the number of peaks to be detected is 22, while in case of *2eV* or higher, the number is constant, 15. The S_{iO} value was adjusted to 5*eV* to prevent spurious peak from being detected.

3.2 Optimization of detection condition in peak detection algorithm using the second derivative system.

Out of three algorithms to be proposed, two algorithms employed the second derivative system. A study was conducted using a XPS spectrum of gold (*Au*) on an optimum smoothing point number upon differentiation, a common step to be taken in the two algorithms. The threshold coefficient, *k*, was changed from 2.0, 2.5 to 3.0 in order. The smoothing point (*smp*) number was changed from 5 to 17 for each threshold coefficient.

A detected peak number was counted under a respective condition defined by combination of each parameter. It was found that the optimum smoothing point number was 9, taking into account the capability of detecting adjacent two *Au4f* peaks(4*eV* between lines). We also concluded that the optimum threshold coefficient was k=3 in the 2nd derivative method, and k=2 in a Peak and Background Relations method, after studying it among the conditions under which the *Ar2p* peak was detected, while noise peaks were eliminated.

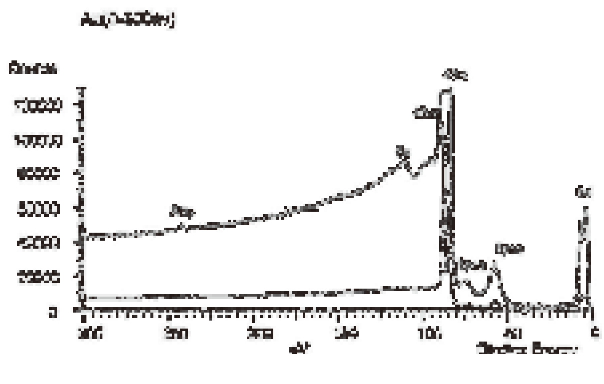

Fig.4 XPS spectrum of Au(0-300*eV*).

3.3 Optimized Peak Detection

Table1 shows the optimum value of each parameter led by the above method. We produced a program enabling automatic detection of peaks on the basis of the respective three algorithms, and applied these optimum parameters to the program. According to the program, a composite peak of *Au*, *Ag* and *Cu* was automatically detected from the XPS spectrum[4].

The number of detected peaks was counted. The value according to the automatic detection was compared to the number of peaks that were visually confirmed, and the peak detection performance was examined. The XPS spectrum, P1, was composite spectrum with *Au, Ag* and *Cu* in an equivalent atomic ratio (*atomic %*), and P2 was composite spectrum by *Au+0.1Ag+0.01Cu* and P3 was composite spectrum by *Au+0.01Ag+0.001Cu* respectively.

4. Results and Discussion

A typical detection result is shown in Fig.5 , in which detection peaks are shown by small bar. It is an example that shows the peaks in the P1 spectrum detected by a rough estimation of spectrum background method. Table 1 and Fig.6 show the results of peak detection in the P1, P2 and P3 spectra under the conditions shown by Table 1.

The maximum number and minimum number in Table 1 and Fig.6 are numbers of peaks visually confirmed. The figures were taken from the paper reported in the activity of VAMAS-TWA2-A9[5]. It is considered that the maximum peak number by visual observation may include spurious peaks. In conducting optimization of the parameters, we configured the parameters not to detect an unnecessary noise as a peak. Provided that a purpose of the automatic detection program is to detect candidate of peaks as many as possible, it is likely to be achieved by changing parameters depending on the respective algorithms. The purpose in the present study is how much closer the automatic detection can approach to visual observation.

Table1 Result of Detection at an optimum value of each parameter

Spectrum	Visual detection		Estimation of Bg			2nd.Derivative		Peak & Bg Relations	
	Max.No	Min.No	2m+1	k	$S_{i0(eV)}$	k	Smp	k	Smp
			21	3	5	3	9	2	9
P1	35	22	33			36		34	
P2	26	17	22			23		23	
P3	21	11	19			20		19	

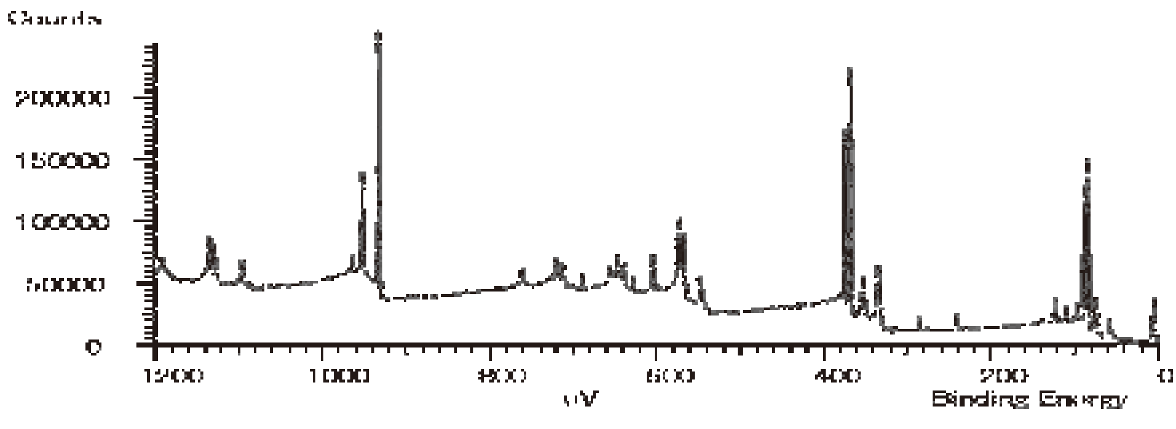

Fig.5 The results of peak detection in P1 spectrum using rough estimation of spectrum background.

Although single broad small peaks, such as a *Au4s* peak (*759 eV)*, are detected by the peak detection method using rough estimation of spectrum background, they are not detected by other two methods using the second derivative method. It is because the former method introduces a S_{iO} parameter. On the other hand, single small peaks being present on the foot of a large peak tends not to be detected, since the background threshold curve becomes greater than the real background value due to the average effect of $2m+1$. For example, a single broad small peak being present at around *400 eV*, a high energy side of the *Au4d* peak (*350 eV*), is not detected (it visually looks like a peak, but is a noise peak in reality).

Relatively small single peaks are more likely to be detected in a peak detection method using threshold curve of the second derivative method and a peak detection method by the peak and background relations method.

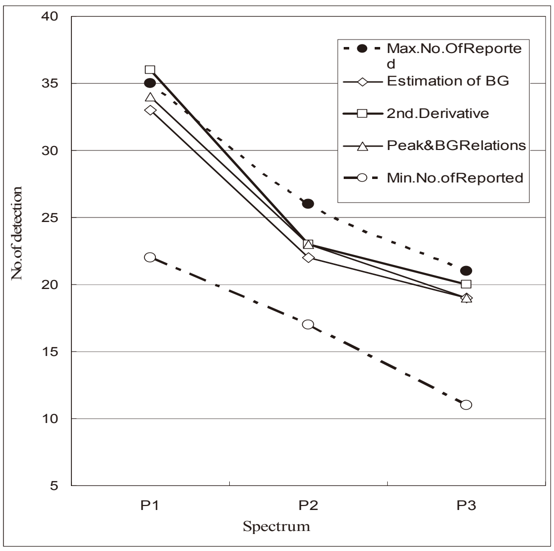

Fig.6 Result of detection in the P1, P2 and P3 spectra.

By the second derivative method, single sharp small peaks are detected even if they are lying as a noise peak, although this is dependent on the smoothing point number for derivative. On the other hand, single broad small peaks are not detected according to the smoothing effect by the second derivative.

As is clear from Fig.6, the peak detection capacity of these three algorithms reaches to the level that allows practical use as a peak detection tool. However, it is found that each method has relative merits depending on a kind of peak. The most effective use of these peak detection means to have been proposed is a combinational use of them. It is recommended that these three methods are used individually, but continuously, since construction of an algorithm by combining several algorithms may interfere the patent right. It is easy to make a list of candidate peaks detected in terms of an energy value by integrating the system into the automatic detection program. We believe the system will release an engineer from a peak detection work by visual observation.

5. References

[1] Y. Nagatsuka, K. Yoshida, Y. Nagasawa and J.Suzumi, J. Surf. Anal.3, 602 (1997).

[2] P. Gans, Data Fitting in the Chemical Sciences, P.17-18,Wiley (1992).

[3]Y.Furukawa,I.Nagatsuka,Y.Nagasawa,S.Fukushima, M.Yoshitake,A.Tanaka, J. Surf. Anal.14, in press (2007).

[4]M.Suzuki,S.Tanuma,S.Fukushima,J.Surf.Anal.14, in press (2007).

[5] VAMAS Activity Meeting, August 2007, Makuhari, Japan