**ENHANCED PEAK DETECTION ALGORITHM**

1. **INTRODUCTION**

The Seeker is seeking an algorithm for accurate, reproducible, and automatic identification and characterization of peaks in two-dimensional time series data. The data may be simple and contain a small number of fully resolved peaks or may be complex and contain hundreds (to thousands) of overlapping peaks that may or may not be fully resolved with baselines that may or may not fluctuate across the timeline. Algorithms must identify the start, end, and apex of each peak.

A Polymer infrared spectrum is used as a calculus basis, being x axis wavelength in cm-1 and y axis transmitance in %.

The wavelength varies from 400 to 4000 cm-1.

The method is efficient even if baseline fluctuates.

COLOCAR AQUI REQUERIMENTOS TÉCNICOS, INFORMAR FORMATO DO ARQUIVO DE INPUT

1. **ALGORITHM**
   1. **INPUT**

Input 1: Data must be in the form ....COMPLETAR COMO VAI SER O INPUT DO ESPECTRO, QUE ARQUIVO

Input 2: Basis Y (In example 100)

Input 3: Minimum peak área (In example 500).

Input 4: Minimum peak length (In example 0.5)

These inputs are necessary in order the seeker can use any sets of data organised in a x and y matrix, not only infrared spectra.

* 1. **DEVELOPMENT**
     1. **FILTER 1**

If

And

And

So delete x2, y2

And

If

And

And

So delete x2, y2

Example: Filter 1 output in column E (x) and column F (y) of the worksheet Spectrum in archive Spectrum1.xls

* + 1. **DERIVATIVE**

Considering output of the Filter 1, solve the next calculi

....

....

Example: Output of the Derivative in column G of the worksheet Spectrum in archive Spectrum1.xls

If

And

And .....

And

And

So End = xn+1, yn+1

And Start = xn+2, yn+2

And If

And

And .....

And

And

So Peak = xn+1, yn+1

Put a different name in each start, end, peak. Example: Start2, Peak2, End2

Example: Outputs in column H of the worksheet Spectrum in archive Spectrum1.xls

* + 1. **PEAK AREAS**

Calculate

.......

........

* + 1. **FILTER 2**

This filter defines a minimum peak área to be considered a valid peak, based on the minimum peak área defined in inputs.

If PeakArea1>Input Minimum Peak Area

So Startk = xstartk, ystartk

And Peakk = xpeakk, ypeakk

And Endk = xendk, yendk

For all k from 1 to n

In the example,

If PeakArea1>500

The validity of the peaks is expressed in column J of the worksheet Spectrum in archive Spectrum1.xls

* + 1. **PEAK LENGTH**

For all valid peaks after Filter 2, calculate:

Lk1=ystartk-ypeakk

Lk2=yendk-ypeakk

In the example, outputs are in column K of the worksheet Spectrum in archive Spectrum1.xls

* + 1. **FILTER 3**

If Lk1>minimum peak length

And

Lk2>minimum peak length

So

Startk = xstartk, ystartk

And Peakk = xpeakk, ypeakk

And Endk = xendk, yendk

Example: Results are in column L of the worksheet Spectrum in archive Spectrum1.xls

1. **OUTPUT**

A table containing starts, peaks and ends as well peak área. A graphic of the spectrum with indication os the starts, peaks and ends on it.

Example

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number | X start | Y start | X peak | Y peak | X end | Y end | Area |
| 21 | 3403 | 58.964861 | 3324 | 54.229280 | 3164 | 61.851610 | 5469.6 |
| 49 | 1741 | 73.3033 | 1648 | 50.333163 | 1587 | 69.274612 | 3824.3 |
| 50 | 1586 | 69.329901 | 1543 | 49.589041 | 1528 | 52.199254 | 1461.9 |
| 51 | 1527 | 52.213423 | 1514 | 47.991101 | 1424 | 66.269734 | 2678.4 |
| 52 | 1423 | 66.300562 | 1401 | 52.964526 | 1342 | 61.758212 | 1904.9 |
| 53 | 1341 | 61.792711 | 1313 | 51.802153 | 1278 | 57.624159 | 1518.23 |
| 56 | 1141 | 62.215673 | 1110 | 58.72788 | 1031 | 65.558652 | 2269.9 |
| 57 | 1030 | 65.60408 | 1015 | 61.644379 | 989 | 64.845368 | 786.2 |
| 59 | 917 | 68.657159 | 892 | 60.969984 | 878 | 63.296953 | 761.08 |
| 60 | 877 | 63.305753 | 865 | 60.092425 | 851 | 63.206318 | 518.7 |
| 61 | 850 | 63.229051 | 824 | 53.803671 | 797 | 63.851904 | 1224.2 |
| 62 | 796 | 63.861732 | 785 | 62.518086 | 752 | 66.83954 | 824.6 |
| 63 | 751 | 66.840833 | 726 | 58.256648 | 704 | 60.871444 | 980.9 |
| 64 | 703 | 60.877886 | 664 | 56.364672 | 549 | 65.790848 | 3359.9 |
| 65 | 548 | 65.797186 | 526 | 59.188636 | 458 | 69.484396 | 1836.5 |
| 66 | 457 | 69.494502 | 441 | 67.701655 | 401 | 71.140066 | 904.3 |