2023-04-19

Novel Block Grid Analysis (BGA) Based Algorithm for use in Determining Coefficients of Similarities Between UV Fluorescent Spectra

Kevin N. Mellon Dr
Technological University Dublin, kevin.mellon@tudublin.ie

Noel Carey Mr
Technological University Dublin, noel.carey@tudublin.ie

Recommended Citation

Mellon Kevin, Carey Noel, "Novel Block Grid Analysis (BGA) Based Algorithm for use in Determining Coefficients of Similarities Between UV Fluorescent Spectra", Journal of Applied Computer Science and Mathematics, Issue 1, (Vol. 17) 2023. DOI: 10.4316/JACSM.202301002

Novel Block Grid Analysis (BGA) Based Algorithm for use in Determining Coefficients of Similarities Between UV Fluorescent Spectra

[1]Kevin MELLON, [2]Noel CAREY
[1,2]Technological University Dublin, Ireland
[1]*kevin.mellon@tudublin.ie*

***Abstract* – This paper presents a novel computational process used for determining matching coefficients for optical spectra either in real time or from a database. This simple algorithmic method is capable of outputting a set of coefficients relating to comparative matching datapoints and to a weighted similarity ratio between spectra whilst still maintaining high computational efficiency due to the minimization of floating-point calculations. The process works as effectively on signal data during real time processing or on images of spectra which may be stored in a database. A subset of the algorithm has already been used very effectively in an embedded mobile field device to determine the presence or absence of special security fluorescent emission peaks present in printing inks used in anti-counterfeiting labels.**

The process involves the division of test spectrum(s) signal/image into a grid of user definable blocks while applying the same grid to the test image or images. Each block is then checked to see if the signal/image passes through the block, if the signal/image passes through a block then the value of this grid block is set to 1 in a grid array. The same check is then applied to the reference signal/image(s). While it is possible to process a signal comparison without the need for data storage, useful for some embedded processes, it may be preferable to store the resultant data in simple data arrays. These arrays can then undergo simple arithmetical processing to quickly calculate a set of values representing the number of matching data points, the variation in non-matching data points and a divergence value. This check can be carried out between the test image or images and a larger number of reference images simultaneously by use of additional counters. In addition, when used with large databases of spectra this novel technique may be used to rapidly reduce a large test set to a more relevant subset which can be further analyzed using more conventional methods or by visual inspection.

This grid square analysis method works as efficiently on images as it can for signal data and it is believed by the authors that the method will have applications in other image processing applications where fast and efficient comparative analysis between images is required.

***Keywords*: Algorithm, emission spectra, computational, UV fluorescent, matching, images, correlation, detection, spectral analysis, coefficients.**

I. INTRODUCTION

With the increase in miniaturization of electronic components and with continued improvements in hardware and software there has been continued development of small and portable spectrophotometry devices for use in many different process applications. Portable or field spectral analysis devices are now being used for detection of counterfeiting of a wide variety of goods including consumer products, medicines, luxury and designer goods[1,2,3], for field analysis in chemical and minerology [4], by military and law enforcement [5,6] and numerous other fields. Portable or handheld devices for ramen, infrared and near infrared, FTIR and UV and visible fluorescence spectrometry have also recently been developed [7,8]. One of the key features of these devices is the requirement for analysis of key signature features of the test signal/images and fast efficient matching to a stored reference(s).

Along with improvements in hardware these devices require robust and efficient algorithmic processes to allow the collected data to be processed into usable results with low computational cost and minimal complexity. There currently exists an array of different computational methods for making a comparison between a test spectrum and a reference spectrum stored in a database. Examples include principal component analysis, spectral angle mapping, mapping covariance, Pearson correlation and Jaccard's coefficient among others [9,10,11].

This paper will discuss a powerful and novel technique designed for efficient real time comparison of an onsite measured UV fluorescent emission spectrum against one or more reference spectra stored within the device. The efficiency of this method lies in the ability of the algorithm to provide relevant coefficients of similarity between the test spectrum and one or more reference spectra by the convenient use of simple variable counters. which has proven to be very effective for comparison of emission spectra. Not only is the method flexible and reliable, but it is also computationally efficient due to the limitation of floating point calculations. The process is further enhanced by its ability to allow comparison between the test signal and a larger number of reference signals simultaneously. The method used is also highly flexible and adaptable the user's requirements. The discussed method has potential application not only for comparison of other types of spectra (such as absorption, reflectance etc.) but might also be applied to other comparison computations such as those used in voice, speech or retinal recognition.

Whilst the specific example given in this paper relates to optical emission spectrum, the method will be equally useful for the comparison of many different types of images and the method works as efficiently for images as it does for signal

DOI: 10.4316/JACSM.202301002

data. It is believed that this methodology is a useful additional tool for the area of spectral analysis.

II. METHODOLOGY

The method described below is simplified to include only a reference sample and a test image signal, however it is important to note that one of the advantages of this method is that a test signal can be concurrently compared to a large number of reference samples (n) from a database by the simple inclusion of an additional n counters during the algorithm run.

The samples may undergo normalisation to the x and y directions before the data is processed. For the signals shown here, obtained from a digitised spectral device, a normalization factor typically varies between zero and a thousand. In order to still further improve analysis, a rolling median function is used to smooth the data and this is accomplished by taking the median of three (or more), points in a series and rolling this median along the curve. This can be represented as median of (dyn-ι, dyn, dyn+ι), This is preferred over taking a normal average because it tends to avoid spurious peaks such as that which may be due to a faulty pixel or other noise effect. This array is then advantageously compared to a reference array(s) which may be stored as a spectrum and has been processed similarly. It is also a feature of this method that further normalisation of the signals can easily be carried out during processing if required to assist computation. General methods of normalisation are not discussed here.

A) Determination of matching datapoints co-efficient Mce.

A simple flow diagram of the Mce determination is shown in figure 1.

1. Divide the test image (image A) and reference image (image B) into a grid of rectangles or blocks. The number of grid blocks in the X and Y direction is denoted as nx and ny respectively.

(It is usually chosen to set the number of grid squares in the Y direction to the same as the number of grid squares in the X direction, however this is not a requirement). See figures 2 and 3.

2. From these grid blocks two arrays are generated,

Array A(X(nx)Y(ny)) & Array B(X(nx)Y(ny))

the test and reference image respectively.

(For comparison of test image with multiple reference images add arrays Array C, Array D, (X(nx)Y(ny)) as required.)

3. View each block on image A and determine if the signal passes through it. If the signal passes through grid block (X(n)Y(n)) for sample A determine the array value A(X(n)Y(n)) as value 1.

(Blocks with signal passing through are called active blocks.)

4. View each block on image B and determine if the signal passes through it. If the signal passes through grid block (X(n)Y(n)) for sample B determine the array value B(X(n)Y(n)) as value 1.
5. Generate an array labelled ArrayAB = {ArrayA(X(n)Y(n))* ArrayB(X(n)Y(n))}, this array determines a value 1 if the data passes through ArrayA(X(n)Y(n)) AND ArrayB(X(n)Y(n)) simultaneously.
6. Define 2 counters, C1 and C2.
7. Defining C1 as

$$C1 = \sum_{p=1}^{p=ny} \sum_{m=1}^{m=nx} A\big(X(m)Y(p)\big) \quad (1)$$

C1 represents the number of active blocks from grid of sample image A, set as the test counter.

8. Defining C2 as

$$C2 = \sum_{p=1}^{p=ny} \sum_{m=1}^{m=nx} \big[A\big(X(m)Y(p)\big)B\big(X(m)Y(p)\big)\big] \quad (2)$$

C2 represents the number of active squares from grid of test image A AND grid of reference image B simultaneously.

9. We now define a coefficient of matching data points, Mce.

$$Mce = \frac{C2}{C1} \quad (3)$$

10. By definition this gives Mce a value between 0 and 1 representing a coefficient of matching datapoints within the accuracy of the block size.

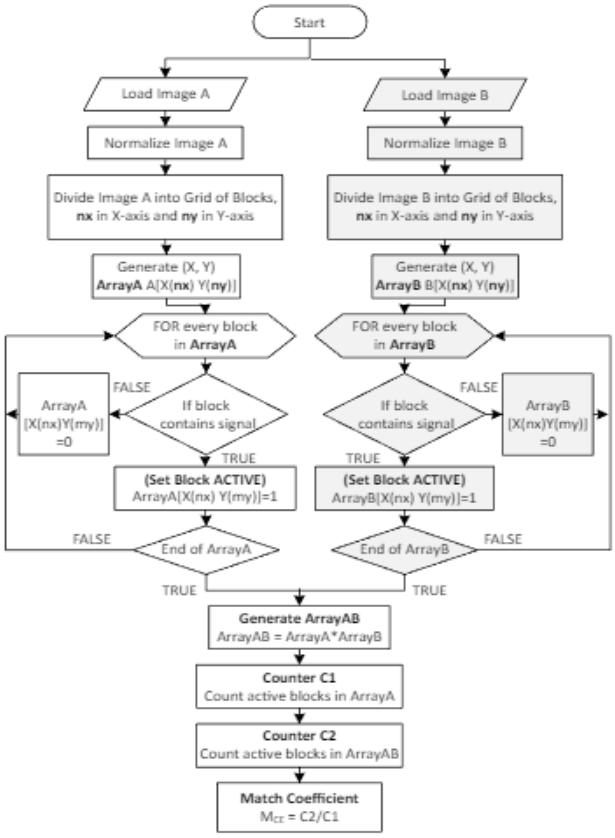

Figure 1. Flowchart of algorithm for determination of Mce

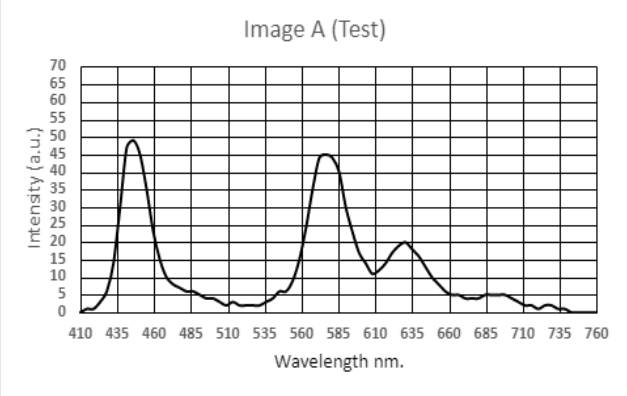

Figure 2. Grid Block for test image

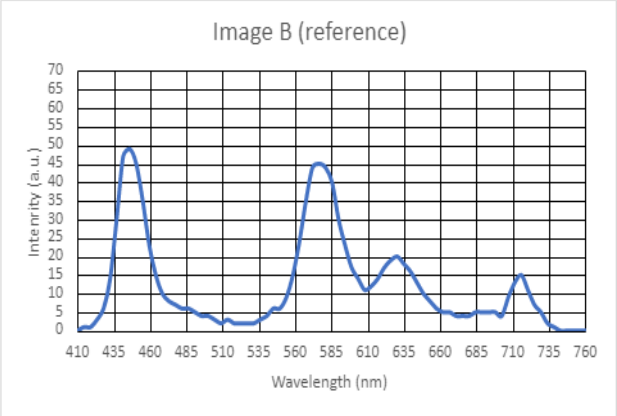

Figure 3. Grid Block for reference image

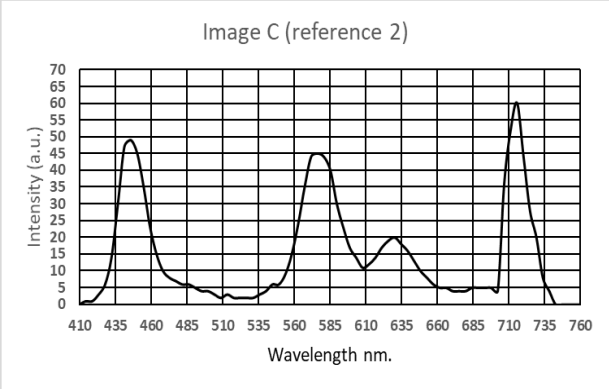

Figure 4. Grid Block for test reference image 2

(Note: Figures 2 and 3 show large a Grid Block size for illustration. For computation a grid of xn = 50, yn =80 was used).

For the two images shown in figures 1 and 2, the algorithm yields values,

C1 = 80,

C2 = 71.

Thus yielding a matching coefficient,

Mce = 0.828.

This process is a powerful yet very efficient method of determining how close the features in two spectral images match and has been used very effectively in determining the presence/absence of narrow emission peaks present in UV luminescent dyes [10,11]. This efficiency is enhanced by the ability to output individual matching Mce values for a test image and multiple reference images by the simple inclusion of additional counters C3…Cn.

Whilst the Mce value algorithm is designed for very efficient determining the number of matching data points where the variance of non-matching points is not of interest [10], it is not intended to realise a value of how different the sample A is from sample B.

The error in the spectral matching coefficient, Mce, is a function of the grid block size and can be varied by increasing the values of xn and yn. For efficiency, in particular where large databases are to be examined, it may be preferable to process the signals at initial low xn and yn values, which allows the exclusion of non-matching signals and to increase the number of grid blocks in a further stage of analysis.

A) Determination of variance co-efficient Vce.

When comparing image A (figure 2) with image B (figure 3) and also with image C (figure 4), it is clear that there are similar number of matching point for both comparisons, giving equal Mce values but the variation in image C is much greater than for image B. Many spectral analysis processes can very effectively be determined by the simple use of the number of matching datapoints, however, there may be a need for additional information that relates to the difference between spactra.

Where a determination of a value representing the variance of the signals is preferable the addition of another counter to the process can output a second co-efficient representative of the variation, Vce. This computation can be carried out following the determination of Mce or in parallel and may also be performed for a large number of images concurrently.

A flow diagram of the process of variance determination is shown in figure 5.

In the process of spectral analysis, it is assumed that the comparative samples produced grid block arrays in the Xaxis of Array A (x1...Axn). From the previous computation process this will produce arrays arrayA(x1[y]..Axn[y]) and array(Bx1[y]..Bxn[y]).

The variance coefficient is determined from an analysis of the grid block variation for active cells where:

$$\text{ArrayAB} = \{A(\ X(n)Y(n)\ )^{*}\ B(\ X(n)Y(n)\ )\} \neq 1 \qquad (4)$$

This may be expressed as a figure of total variation where,

$$V = \sum_{p=1}^{p=xn}\left[\sum_{q=1}^{q=yn}\{A(X(p)Y(q)*q)\} - \sum_{q=1}^{q=yn}\{B(X(p)Y(q)*q)\}\right] \qquad (5)$$

Or more preferably it may be expressed as a coefficient from 0-1,

$$\text{Vce} = \frac{\sum_{p=1}^{p=xn}\left[\frac{\sum_{q=1}^{q=ny}[A(X(p)Y(q))*q]-\sum_{q=1}^{q=ny}[B(X(p)Y(q))*q]}{yn}\right]}{xn} \quad (6)$$

Analysis of the calculated variance coefficients for figures 3 and 4, and also for figures 3 and 5 when the above Vce algorithm process is applied yields a Vce value of 0.01225 for comparison between test image A and the reference image B, and a Vce value of 0.0512 for comparison between test image A and reference image C.

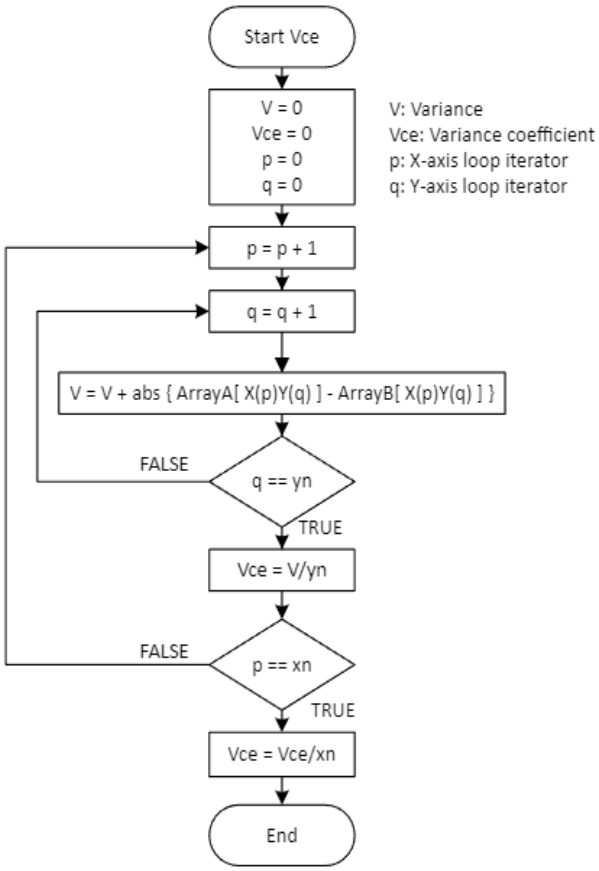

Figure 5. Flowchart of algorithm for determination of Vce

III. Conclusions

The authors describe a novel computational process used for determining comparison coefficients for signal/images of optical spectra either in real time or from a database.

This simple algorithmic method is capable of outputting a set of coefficients relating to comparative matching datapoints and to a weighted similarity ratio between spectra whilst still maintaining high computational efficiency due to the limitation of floating-point calculations. The process works as effectively on signal data during real time processing or on images of spectra which may be stored in a database.

An advantage of this useful methodology lies in the fact that following the determination of active cells, described above, further process can be limited to active cells, minimising the overall processing.

A subset of the algorithm has already been used very effectively in an embedded mobile field device to determine the presence or absence of special security fluorescent emission peaks present in printing inks used in anti-counterfeiting labels [10,11].

The described algorithmic process is seen to be a computationally fast and efficient process, ideal for embedded processing, as it minimizes floating point calculations and primarily relies on the use of counters during processing.

A further strength of the algorithm described is that it can easily calculate match and variance coefficients between a test image or images and a large number of reference images concurrently with the use of additional counters, providing a very efficient and fast processing methodology useful in particular for portable embedded or field analysis devices.

The authors believe that the usefulness of this algorithmic technique is not limited to analysis of optical spectra but may be a useful tool in the comparison of many different image types, allowing fast efficient, parallel comparison of large numbers of images in real time analysis or from images stored in database.

References

[1]. Figueroa G, Palacio L, Ray B, Petrache, H, Lopez-Yunez A. Detecting Counterfeit Pharmaceuticals through UV Spectrophotometry. Biophysical Journal. 108. 622a.

[2]. Guillemain A, Degardin K, Roggo Y, Performance of NIR handheld spectrometers for the detection of counterfeit tablets. Talanta, Volume 165, 2017. Pages 632-640

[3]. Kalyanaraman, R, Dobler, G. Ribick, M. Portable spectrometers for pharmaceutical counterfeit detection. American Pharmaceutical Review. Volume 13. 2010. pages 38-45.

[4]. Howari, F.M. Comparison of Spectral Matching Algorithms for Identifying Natural Salt Crusts. Journal of Applied Spectroscopy **70**, 2003. Pages 782–787.

[5]. Crocobe, R.A. Portable spectroscopy. Applied spectroscopy. Volume 72(12). 218. Pages 1701-1751.

[6]. Vandenabeele P, Donais, M.K. Mobile Spectroscopic Instrumentation in Archaeometry Research. Appl. Spectroscopy. Volume 70 2016. Pages 27-41.

[7]. Crocombe R.A. The future of portable spectroscopy. Spectroscopy. Volume 35 (7). 2020.

[8]. Heydarov S, Aydin M, Faydaci C, Tuna S, Ozturk S. Low-cost VIS/NIR range hand-held and portable photospectrometer and evaluation of machine learning algorithms for classification performance. Engineering Science and Technology. 2023. Volume 37.

[9]. Judith Müller-Maatsch, Francesca Romana Bertani, Arianna Mencattini, Annamaria Gerardino, Eugenio Martinelli,

Yannick Weesepoel, Saskia van Ruth, The spectral treasure house of miniaturized instruments for food safety, quality and authenticity applications: A perspective. Trends in Food Science & Technology, Volume 110, 2021, Pages 841-848.

[10]. Mellon K, 2001. Hand-held anti-counterfeiting apparatus. US patent application. 20010052575. https://image-ppubs.uspto.gov/dirsearch-public/print/downloadPdf/20010052575

[11]. Davey A.P, Gray D, Mellon K, Curran S, Cowden N. 2001. Algorithms for use in a hand-held spectrophotometric apparatus. WIPO Patent Application. https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2001055697&_cid=P10-LCRWO4-91328-1

Kevin Mellon, BSc (Hons), PhD. Kevin completed his BSc (Hons) in 1989 and received his PhD in 1993 from Dublin City University. His PhD thesis was in plasma physics with an interest in Neutral Beam Injection for nuclear fusion heating. His work experience has included JET fusion facility in Oxford UK, Intel Ireland, and since 2001 he has worked as a lecturer in Technological University Dublin. His current research interests include computational processes for image plagiarism detection, electronic devices and Stem pedagogy and engagement.

Noel Carey, BSc (Hons). Noel completed his BSc (Hons) in Computer Science in 2005 in Griffith College Dublin. He also holds qualifications in Electrical and Electronic Eng (City & Guilds) and a Diploma in Applied Electronics (DIT Kevin St). He has previously worked in Cardiac Services as a medical equipment technician and in DIT Kevin St as a laboratory technician. He currently is employed as a lecturer in Technological University Dublin.