

A Laplacian of Gaussian-Based Approach for Spot Detection in Two-Dimensional Gel Electrophoresis Images

Feng He, Bangshu Xiong, Chengli Sun, Xiaobin Xia

► **To cite this version:**

Feng He, Bangshu Xiong, Chengli Sun, Xiaobin Xia. A Laplacian of Gaussian-Based Approach for Spot Detection in Two-Dimensional Gel Electrophoresis Images. Daoliang Li; Yande Liu; Yingyi Chen. 4th Conference on Computer and Computing Technologies in Agriculture (CCTA), Oct 2010, Nanchang, China. Springer, IFIP Advances in Information and Communication Technology, AICT-347 (Part IV), pp.8-15, 2011, Computer and Computing Technologies in Agriculture IV. <10.1007/978-3-642-18369-0_2>. <hal-01564853>

HAL Id: hal-01564853

<https://hal.inria.fr/hal-01564853>

Submitted on 19 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A Laplacian of Gaussian-based Approach for Spot Detection in Two-Dimensional Gel Electrophoresis Images

Feng He¹, Bangshu Xiong¹, Chengli Sun¹, Xiaobin Xia¹

¹ Key Laboratory of Nondestructive Test of Ministry of Education, Nanchang Hangkong University, 330063 Nanchang, P. R. China
hefeng68@qq.com, xiongbs@nchu.edu.cn

Abstract. Two-dimension gel electrophoresis (2-DE) is a proteomic technique that allows the analysis of protein profiles expressed in a given cell, tissue or biological system at a given time. The 2-DE images depict protein as spots of various intensities and sizes. Due to the presence of noise, the inhomogeneous background, and the overlap between the spots in 2-DE image, the protein spot detection is not a straightforward process. In this paper, we present an improved protein spot detection approach, which is based on Laplacian of Gaussian algorithm, and we extract the regional maxima by morphological grayscale reconstruction algorithm, which can reduce the impact of noisy and background in spot detection. Experiments on real 2-DE images show that the proposed approach is more reliable, precise and less sensitive to noise than the traditional Laplacian of Gaussian algorithm and it offers a good performance in our gel image analysis software.

Keywords: Two-dimensional gel electrophoresis, Spot detection, Laplacian of Gaussian, Morphological grayscale reconstruction

1 Introduction

Proteomic research deals with the systematic analysis of protein profiles expressed in a given cell, tissue or biological system at a given time. In this field, two-dimensional gel electrophoresis (2-DE) is a well-established and widely used technique to separate proteins extracted from sample for identification and analysis of differential expression, according to their isoelectric points and molecular weight [1]. The result of that process is many dark spots on the gel, and each spot represents a protein or a group of proteins.

The two-dimensional gel electrophoresis (2-DE) images show the expression levels of several hundreds of proteins where each protein is represented as a spot of grey level values [2]. In order to extract protein spots, image processing techniques can help us to analyze proteins further. Each spot can be characterized by its location and other information, such as area, volume, intensity, etc. Due to the presence of noise, the inhomogeneous background, and the overlap between the spots in 2-DE image, the protein spot detection is not a straightforward process.

A variety of software packages have been developed for protein spot detection [3]. Many of these packages implement image segmentation methods based on edge detection algorithms such as Laplacian filtering, in conjunction with smoothing operators. However, if a 2-DE image contains artifacts and noise, such as cracks in the gel surface, fingerprints, dust and other pollutions, they will lead to spurious spot detection [4]. The watershed transformation algorithm has also been a popular choice for 2-DE image segmentation [3] [5]. State of the art approaches to spot detection by image segmentation include geometric algorithms [6], parametric spot models [3] and the pixel value collection method [2] [7].

In this article, we introduce an improved protein spot detection approach that is based on Laplacian of Gaussian algorithm, in conjunction with morphological grayscale reconstruction to extract the regional maxima in 2-DE images. The proposed approach extracts the regional maxima of the Gaussian-smoothed gel image by morphological grayscale reconstruction algorithm [8], and uses the second derivative (laplacian) and direction of the gaussian-smoothed gel image as well as neighborhood connectivity properties in determining spot extents. Relative to the traditional Laplacian of Gaussian algorithm, our approach can reduce the impact of noise and avoid the spurious spot detection by using morphological grayscale reconstruction algorithm to restrict regional maxima, and the results are more reliable and precise. The rest of this paper is organized as follows: Section 2 describes the spot detection approach which has been applying in our soft, Section 3 presents the experimental results, and Section 4 presents conclusions and the future work.

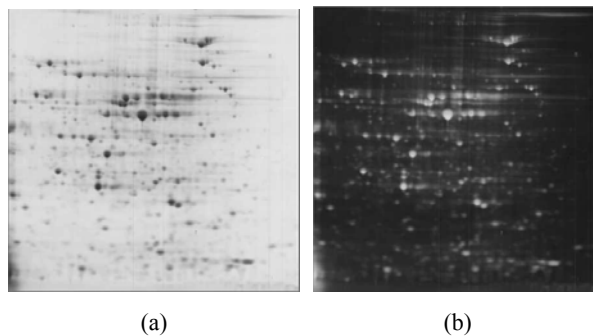


Fig. 1. (a) Original Image, (b) Inverted Image

2 Algorithm

In the following description it is assumed that the original image shows in Fig. 1a is inverted, that is, the image background is dark, and the spots appear as light peaks rising from the background in Fig. 1b. The proposed approach to protein spot detection consists of the following steps:

- a) Smooth the original gel image;
- b) Extract the regional maxima of the smoothed image by grayscale reconstruction algorithm;

- c) Compute the Laplacian of the smoothed image and define central core regions where the Laplacian pixels are negative in both x and y directions;
- d) Extract all of spot regions;
- e) Estimate background density and normalize the valid spot quantifications.

2.1 Image smoothing operations

During the detection process, the gel images are normally very noisy, so we must implement smoothing filter to reduce high frequency noise in 2-DE image at first. If the image is not sufficiently smoothed, it will find the edges of the spots incorrectly and generate spurious regional maxima. In the proposed approach, we use a Gaussian low-pass filter with a window (e.g. 3×3 , 5×5 , 7×7 , etc) to smooth the inverted image.

2.2 Regional maxima extraction

Grayscale reconstruction is a very useful operator provided by mathematical morphology. The regional maxima and minima are important features of images; they often represent the corresponding image goal: the regional maxima correspond to the light goal, and the region minima correspond to the dark goal. Grayscale reconstruction turns out to provide a very efficient method to extract regional maxima and minima from grayscale images. Furthermore, the technique extends to the determination of “maximal structures” called h -domes [8]. The h -dome transformation extracts light structures without involving any size or shape criterion. The only parameter (h) is related to the height of these structures. This characteristic is of interest for complex segmentation problems.

The h -dome image $D_h(I)$ of the h -domes of a grayscale image I is given by

$$D_h(I) = I - R_I^\delta(I - H) \tag{1}$$

where $R_I^\delta(I - H)$ is the dilated reconstruction of I from $I - h$.

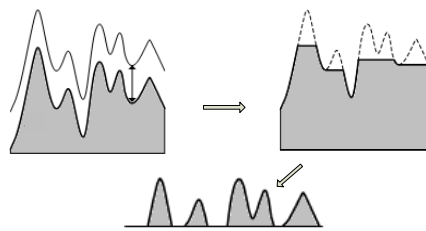


Fig. 2. Determination of the h -domes of grayscale image I .

In the proposed approach, we can execute h -dome transformation to extract light maximal structures of the smoothed image. It can restrain all maxima, the depth of which are less than or equal to the parameter (h), and reduce the impact of noisy in spot detection.

2.3 Image Laplacian operations

2-DE images show the expression levels of several hundreds of proteins where each protein is represented as a spot of grey level values. Under optimal density within the area of the spot appears as a monotonically increasing function as illustrated in Fig. 3a. The Laplacian or second derivative of this function is shown in Fig. 3b. We define the central core region of negative values for digital approximations to both partial second derivatives, with respect to x and y directions, as the central core region. The region on the outside of the central core is propagated until it reaches the extent of the positive peaks of the side lobes. This propagated central core region, computed in two dimensions, is then effectively used as a mask for quantitating that spot [5]. This is the reason we want to use the Laplacian for helping to analyze the image.

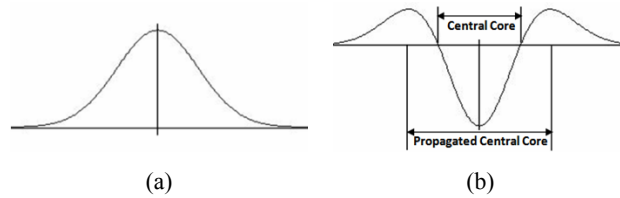


Fig. 3. (a) Cross section of the ideal spot, (b) Laplacian of this ideal spot cross section.

After acquired the smoothed image, it is used when computing the digital approximation of the Laplacian of this image. We store the Laplacian direction and magnitude values in two additional images. All laplacian direction image pixels are set to 1, if the Laplacian values are negative in both x and y , otherwise they are set to 0. This directional image defines the initial central core regions of a gel. The central core regions are propagated to adjacent pixels until they reach the maximum value of the Laplacian magnitude image. These final regions are called the propagated central core region and, after some corrections, define the extent of the spot to be quantitated. The next step describes the extraction of each spot region in the proposed approach.

2.4 Spot region extraction

After the image laplacian operations, the next step is to extract all of the spot regions in sequential raster search of the central core image. The central core regions are propagated to adjacent pixels until they reach the maximum value of the Laplacian magnitude image. These final regions are called the propagated central core region and, after some minor corrections, define the extent of the spot to be quantitated. The steps in finding the final propagated central core for the current spot are enumerated as follows. This algorithm is iterated for each spot as it traverses down the image in a raster pattern.

Step1. Find the central core for the current spot in the Laplacian direction image, where the central core regions are set to 1 and the other set to 0. Given a new spot pixel to find all (x, y) pairs that are 4-neighbor connected to this spot with central core pixel, and save this pairs in a list.

Step2. If the central core size is less than the threshold which is the minimum of central core size, then delete the spot, and set that region to 0 in the central core image.

Step3. There are situations where some spots are merged into single spots. If the central core size is greater than the threshold which is the minimum to split, then try to split the spot into several spots.

Step4. Try to find saturated spots, if the (max Density of the spot)/ (max Density in the image) is greater than percent threshold of the darkest pixel, then try to fill holes in the saturated spot with central core pixels.

Step5. Remove concavities which with 0 values between central core pixels with central core value in the current central core spot.

Step6. Remove interior central core pixels to speed up the subsequent steps and obtain the edge of the central core. This is useful for segmenting very large spots when using very high pixel resolution.

Step7. Propagate the central core to a propagated central core by looking for the maxima away from the center of the spot in all directions. This propagation is terminated by various conditions including running into another spot, noise, etc. If the image is not adequately smoothed, this step will not work very well.

Step8. Optimize the propagated central core region. We can fill holes in the propagated central core such that central pixels with 0 values are filled with propagated central core value.

Step9. Delete the spurious propagated central core. If the propagated central core region does not contain a regional maximum, so it is a spurious propagated central core, and we could delete it.

Step10. Finally, compute the spot features using data from the original image and save the features for this spot in a list of all spots. The spot features include density weighted centroid, standard deviation and covariance spot size, density, area and volume.

2.5 Background density estimation and spot quantification

After all spots are initially segmented, it then performs background correction and normalization on the quantifications. If the background appears relatively uniform, we have found subtracting global minimum intensity for the gel works sufficiently well. However, the background appears to be spatially varying, we use a smoothing low pass filter to estimate the background [9]. First a rest of image is computed as the original image less the segmented spots with the spots having density value 0. Then the filter is computed over the entire image by moving an averaging window (e.g. 32x32) over the image in a raster, 1 pixel at a time where the mean density is computing in the averaging window at each point not including the 0 values. This background image is used to estimate the background for each spot and correct the spot density.

To normalize, we divide each spot intensity on a given gel by the mean spot intensity for that gel, and save the spot features in a list, including location, area, volume, normalized intensity, etc. The researchers can use these spot features to do next analysis, such as spot matching.

3 Experimental Results

In this section, the results for detection the protein spots in a gel image by those methods mentioned in Section 2 are presented. In our approach, we firstly inverted the original image, that is, the image background is dark, and the spots appear as light peaks rising from the background.

We smoothed the inverted image by Gaussian low-pass filter, if the image is not sufficiently smoothed, it will find the edges of the spots incorrectly and generate spurious regional maxima. Fig. 4a shows the regional maxima labeled by red color. Fig. 4b and Fig. 4c show the Laplacian magnitude image and the directional image which defines the initial central core regions of a gel. The central core regions are propagated to adjacent pixels until they reach the maximum value of the Laplacian magnitude image. These final regions are called the propagated central core region and, after some minor corrections, we can get the optimized propagated central core region in Fig. 4d, is then effectively used as a mask for quantitating that spot, and the background image in Fig. 4e. Finally, we can get the final result as Fig. 4f, described the spot contour by red color.

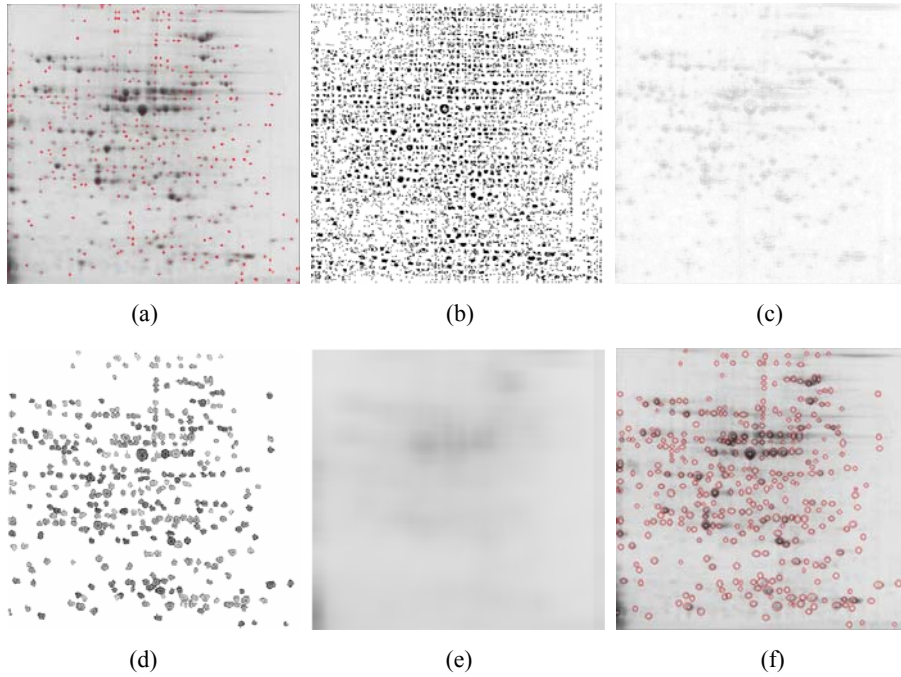


Fig. 4. The process of the protein spots detection with the improved Laplacian of Gaussian approach: (a) regional maxima on original image, (b) central core regions, (c) laplacian magnitude image, (d) final propagated central core regions that are optimized, (e) background image, (f) final result.

As shown in Fig. 4f, we detected 374 protein spots, almost protein spots can be detected and their contours are very clear based on the proposed approach. Fig. 5

shows the result of the Laplacian of Gaussian algorithm, which detected 506 protein spots without morphological grayscale reconstruction algorithm to restrict regional maxima, however, affected by the stripe noise and artifacts, some of them are the spurious spots. On the other hand, the proposed approach is less sensitive to noisy.

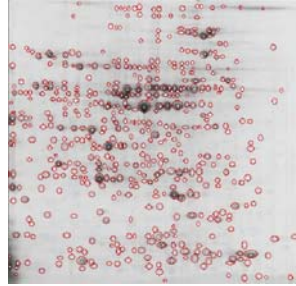


Fig. 5. The final result of traditional Laplacian of Gaussian algorithm

4 Conclusions

We present an improved protein spot detection approach which can effectively detect and quantify the protein spots in 2-DE images. The spot detection approach applies morphological grayscale reconstruction algorithm to restrict regional maxima, and the Laplacian of Gaussian algorithm to detect spot regions. The 2-DE image often contains artifacts and noise, and they will lead to spurious spot detection. The morphological grayscale reconstruction algorithm can restrict those spurious regional maxima very well, and reduce the impact of noisy in spot detection. The experimental results show that the proposed approach is more reliable, precise and less sensitive to noise than the traditional Laplacian of Gaussian algorithm and it offers a good performance in our gel image analysis software.

Future works include further experimentation, optimization and parallelization of the proposed approach, and its integration in a complete user-friendly software application. Also variation of the proposed approach will be used to detect all spots in heavily polluted 2-DE images.

Acknowledgments. This work was supported by Postgraduate Innovation Fund of Jiangxi Province (YC09A112), Postgraduate Innovation Fund of Nanchang Hangkong University (YC2009008), Scientific Research Fund of Jiangxi Provincial Education Department (GJJ09183) and Jiangxi Nature Science Fund (No.2008GZS0032).

References

1. Haleh Safavi, Nivolle Correa: Independent Component Analysis of 2-D Electrophoresis Gels. *Electrophoresis*. 29, 4017--4026 (2008)

8 **Feng He**¹, **Bangshu Xiong**¹, **Chengli Sun**¹, **Xiaobin Xia**¹

2. Peter Peer, Luis Galo Corzo: Local Pixel Value Collection Algorithm for Spot Segmentation in Two-Dimensional Gel Electrophoresis Research. *Comparative and Functional Genomics*. 1--9 (2007)
3. Matthias Berth, Frank Michael Moser: The State of the Art in the Analysis of Two-Dimensional Gel Electrophoresis Images. *Appl Microbiol Biotechnol*. 76, 1223--1243 (2007)
4. Morten Beck Rye, Bjorn k. Alsberg: A Multivariate Spot Filtering Model for Two-Dimensional Gel Electrophoresis. *Electrophoresis*. 29, 1369--1381 (2008)
5. Thitiwan Srinark, Chandra Kambhamettu: An Image Analysis Suite for Spot Detection and Spot Matching in Two-Dimensional Electrophoresis Gels. *Electrophoresis*. 29, 706--715 (2008)
6. Jeffrey S. Morris, Brittan N. Clark and Howard B. Gutstein: Pinnacle: a Fast, Automatic and Accurate Method for Detecting and Quantifying Protein Spots in 2-Dimensional Gel Electrophoresis Data. *Bioinformatics*. 24, 529--536 (2008)
7. Morten Beck Rye, Ellen M. Færgestad, Harald Martens: An Improved Pixel-Based Approach for Analyzing Images in Two-Dimensional Gel Electrophoresis. *Electrophoresis*. 29, 1382--1393 (2008)
8. P. Soille: *Morphological Image Analysis: Principles and Applications*, Second Edition. Springer, Heidelberg (2003)
9. Werner Van Belle, Gry Sjøholt and Nina Anensen: Adaptive Contrast Enhancement of Two-Dimensional Electrophoretic Protein Gel Images Facilitates Visualization, Orientation and Alignment. *Electrophoresis*. 27, 4086--4095 (2006)