Computers in Biology and Medicine 66 (2015) 154–169

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

A novel multi-scale Hessian based spot enhancement filter for two dimensional gel electrophoresis images

Sina Shamekhi [a], Mohammad Hossein Miran Baygi [a,*], Bahareh Azarian [b], Ali Gooya [a]

[a] Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran
[b] Protein-Chemistry Lab., Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

ARTICLE INFO

Article history:
Received 10 April 2015
Accepted 13 July 2015

Keywords:
Spot enhancement filter
Two-dimensional gel electrophoresis
Detection
Segmentation
Proteomics
Hessian

ABSTRACT

Two dimensional gel electrophoresis (2DGE) is a useful method for studying proteins in a wide variety of applications including identifying post-translation modification (PTM), biomarker discovery, and protein purification. Computerized segmentation and detection of the proteins are the two main processes that are carried out on the scanned image of the gel. Due to the complexities of 2DGE images and the presence of artifacts, the segmentation and detection of protein spots in these images are non-trivial, and involve supervised and time consuming processes. This paper introduces a new spot filter for enhancing, and separating the closely overlapping spots of protein in 2DGE images based on the multi-scale eigenvalue analysis of the image Hessian. Using a Gaussian spot model, we have derived closed form equations to compute the eigen components of the image Hessian of two overlapping spots in a multi-scale fashion. Based on this analysis, we have proposed a novel filter that suppresses the overlapping area and results in a better spot separation. The performance of the proposed filter has been evaluated on the synthetic and real 2DGE images. The comparison with three conventional techniques and a commercial software package reveals the superiority and effectiveness of the proposed filter.

1. Introduction

Proteomics is a branch of biotechnology that was introduced for the first time in mid-1990s to study and analyse proteom and its functions [1,2]. Proteomics is used for the large-scale characterization of an entire set of proteins in a cell line, tissue, organism or other biological samples. As a branch of an established science, protemoics governs various methods and tools such as two dimensional gel electrophoresis (2DGE) and mass-spectrometry to study characteristics of proteins. 2DGE is an important, useful and unrivalled method for protein separation and is widely used in the search for disease biomarkers particularly those used for cancer and neurological diseases [3]. The initial reports of successful 2DGE that combined isoelectric focusing (IEF) and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) were published in 1974 and 1975 [4,5]. The method creates a complete map of separated proteins within the primary sample such as complex whole cells, tissues and organisms. The proteins are separated due to the differences in their isoelectric points, in the first dimension, and subsequently, their molecular weight in the second dimension. These separated and accumulated proteins appear as spots in the resulting gel as well as in the output image (Fig. 1). 2DGE images pose various challenges such as streaks, cracks, fingerprints, dust, staining variation and overlapping spots [6–8] that can affect the reliability of the analysis [9]. These issues were categorized by Dowsey [10] as

- artifacts and co-migration,
- intensity inhomogeneity,
- geometric distortion.

These artifacts must be filtered and the overlapped spots must be segmented correctly. Software-induced variance is another problem in 2DGE image analysis [11]. Therefore manual intervention or supervised image correction is required. The final stage of the 2DGE image processing algorithm is to perform a differential analysis between pairs of images (e.g. control and treat images) to study changes in protein quantification and its content. Each resulting image contains a few hundreds to several thousands of protein spots. Due to the above challenges, the image processing of the images of 2DGE is a complex and time consuming process.

There have been a considerable number of published works which have tried to overcome problems associated with processing of the 2DGE images. One of the earliest studies of image analysis was performed by Lemkin et al. [12]. They proposed a

* Corresponding author. Tel.: +98 2182883370; fax: +98 2182884325.
E-mail address: miranbmh@modares.ac.ir (M.H. Miran Baygi).

http://dx.doi.org/10.1016/j.compbiomed.2015.07.010
0010-4825/

computer program, FLICKER, for aiding the analysis of two-dimensional gel electrophoresis images. The development of the computers in the 80s facilitated the implementations of these techniques. For example in 1985 Potter et al. presented a new program for the analysis of 2DGE images using an array processor [13]. Initial attempts to build an automated analyser for images were reported in the late 90s [14,15]. In recent years, the main research works have focused on image pre-processing, image denoising, background correction and normalization [7,16–18]. Some works have presented methods to segment 2DGE images [19,20]. Most recently, researchers have focused on reducing the manual interference in image analysis such as unsupervised 2DGE image segmentation based on active contours [21,22]. Nevertheless, recent studies on spot segmentation such as [8,19] show that there are still unresolved problems associated with the spot segmentation in the 2DGE images.

Pedersen categorized the publications on segmentation into three groups of methods based on mathematical morphology, parametric spot modeling, and Gaussian scale-space based blob detectors [23]. Mathematical morphology in image analysis is a broad field of research. One of the earliest works on segmentation of electrophoresis gels using mathematical morphology was performed by Beucher et al. [24,25]. They proposed a watershed based method for segmenting blobs from the background. These methods are well known and commonly used to detect protein spots in images. However, by using these marker-controlled watershed algorithms [26,27], over-segmentation remains a major problem. H-dome is another mathematical-morphology based method used to determine maximal structures in grey-scale inverted images [27]. The main drawback of this method is the detection of overlapping spots. Neighbor overlapping spots are not regional maxima. Therefore H-dome can not separate these spots [23].

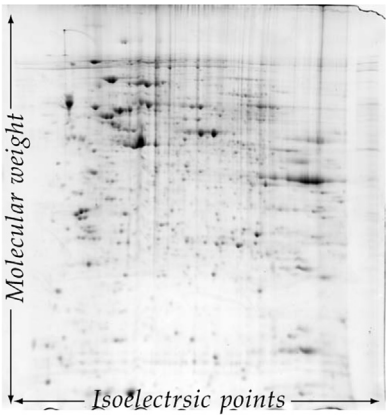

Fig. 1. A real 2DGE image showing isoelectric points (first dimension) vs. molecular weight (second dimension).

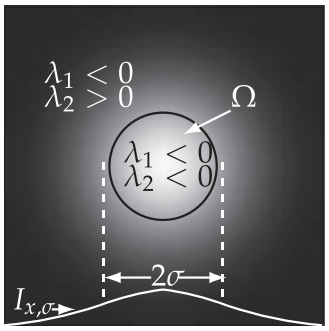

Fig. 3. Regional map of the states of the eigenvalues on a Gaussian spot. We define a spot as a region where both λ_1 and λ_2 are negative. Conversely, on non-spot regions the signs of the eigenvalues can take any other combination. (More details are given in Appendix A.)

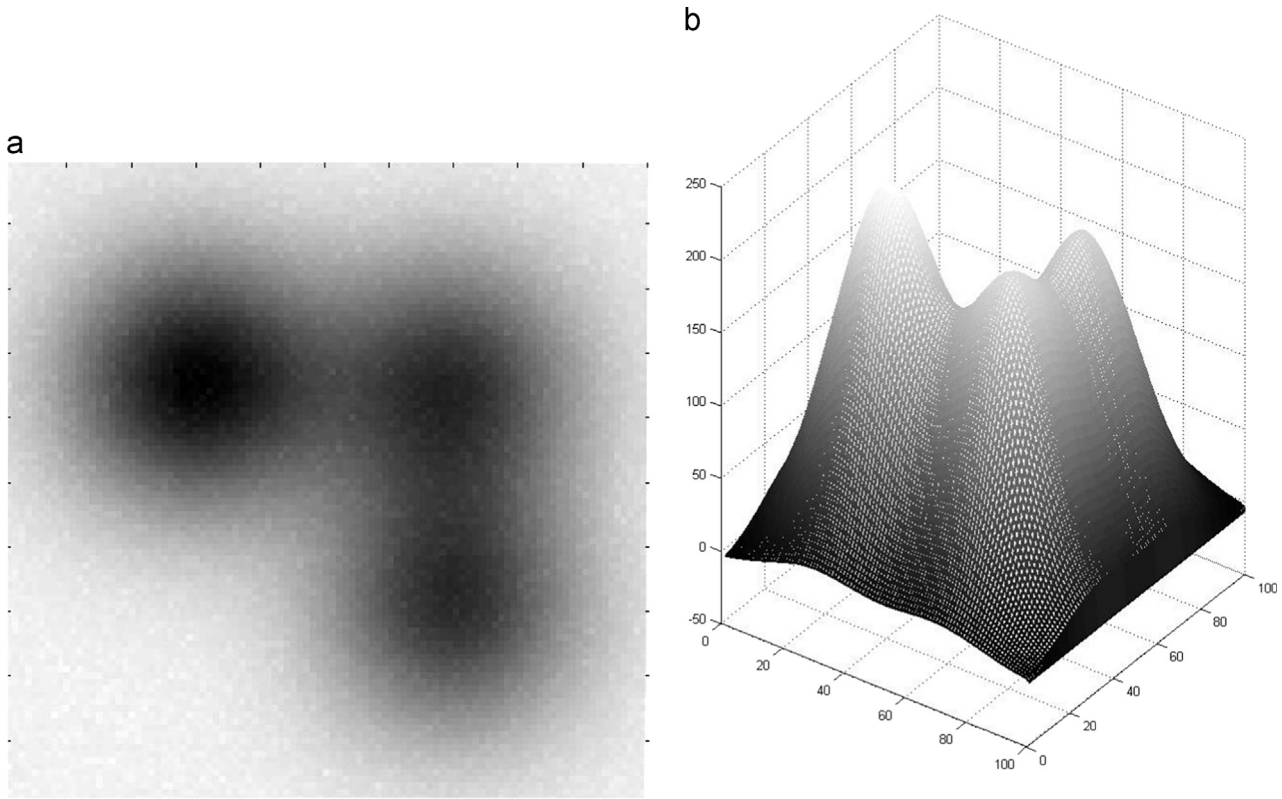

Fig. 2. Synthetic 2DGE image: (a) a Synthetic 2DGE image and (b) its corresponding inverted 3D image.