Advances in Adaptive Data Analysis
Vol. 3, No. 3 (2011) 401–415

DOI: 10.1142/S1793536911000726

World Scientific
www.worldscientific.com

AUTOMATIC PROTEIN SPOTS QUANTIFICATION IN TWO-DIMENSIONAL GEL IMAGES

SHAMS RASHWAN*,‡, AMANY SARHAN†,
MUHAMMAD TALAAT FAHEEM† and BAYUMY. A. YOUSSEF*

*Informatics Research Institute,
City of Scientific Research and Application Technology,
Borg ElArab, Alexandria, Egypt

†Computers and Automatic Control Engineering Department,
Faculty of Engineering, University of Tanta, Tanta, Egypt
†shamsrashwan@gmail.com

Two-dimensional (2D) polyacrylamide gel electrophoresis of proteins is a robust and reproducible technique. It is the most widely used separation tool in proteomics. Current efforts in the field are directed at development of tools for expanding the range of proteins accessible with 2D gels. Proteomics was built around the 2D gel. The idea that multiple proteins can be analyzed in parallel grew from 2D gel maps. Proteomics researchers needed to identify interested protein spots by examining the gel. This is time-consuming, labor-extensive, and error-prone process. It is desired that the computer can analyze the proteins automatically by first detecting then quantifying the protein spots in the 2D gel images. In our previous work, we presented a new technique for segmentation of 2D gel images using the fuzzy c-means (FCM) algorithm using the notion of fuzzy relations. In this paper, we will describe the new relational FCM (RFCM) algorithm and use it for automatic protein spots quantification. We will also use two methods to evaluate its performance: the unsupervised evaluation method and comparison with the expert spots quantification.

Keywords: 2D gel images; protein spot detection; protein spot quantification; fuzzy c-means algorithm; fuzzy relations; computer vision.

1. Introduction

The last decade in life sciences was deeply influenced by the development of the "Omics" technologies (genomics, transcriptomics, proteomics, and metabolomics), which aim for a global view on biological systems. With these tools at hand, the scientific community is striving to build functional models to develop a global understanding of the living cell (see Anderson and Matheson [2001]; Wasinger *et al.* [1995]; Blackstock and Weir [1999]).

The analysis of the proteome as the final level of gene expression started out with techniques based on two-dimensional (2D) gel electrophoresis (O'Farrell [1975];

‡Corresponding author.

Klose [1975]) and extended its reach with semi-gel-free and shot gun gel-free liquid chromatography–mass spectrometry (LC–MS)-based techniques in recent years.

Quantitative analysis based on LC–MS techniques is still in an early stage when considering available software and algorithms. Here, we focus on the computerized analysis of 2D gels that are widely used in the scientific community. 2D gels may separate up to 10,000 protein spots on one gel (Klose and Kobalz [1995]). In a suitably equipped and experienced lab environment, 2D gels are easy to handle and they can be produced in a highly parallelized way.

On a proteome map, one can detect all spots of a whole experiment in a single-gel image, whereas the average images proposed earlier suffer from dilution effects for weak and rare spots. The spots detected there can serve as a spot consensus pattern that is valid for the whole gel set of the experiment. The consensus spot pattern is then transferred according to the warping transform and used on all gels. This allows for 100% matching spots and, in turn, completes expression profiles for reliable statistical analysis (Voigt *et al.* [2006]; Höper *et al.* [2006]).

The goals of this step — Protein spot detection — are to find the spot positions, find their surrounding boundary, and determine their quantities. There are two basic approaches that are used in current software: image segmentation and model-based quantitation.

The segmentation approach partitions the image into nonoverlapping segments, essentially classifying each pixel as belonging to a certain spot, or as being part of the background between spots. Spot boundaries and quantities are then derived from the spot's pixels.

The segmentation of the image can take various characteristics of the image into account: raw intensity, slope, and classification of pixels in the surrounding region. The advantage of this approach is that the image is clearly separated into spots and "nonspot" areas that are easy to assess by a user. If the software allows editing of spot boundaries, then any desired spot shape can, in principle, be obtained. Model-based approaches try to model a spot's intensity as a Gaussian normal distribution or some variant thereof. A spot's quantity and boundaries are then derived from the model. This paper is organized as follows: Section 2 presents the previous developments in this area of research. In Sec. 3, the original fuzzy c-means(FCM) algorithm was presented. In Sec. 4, the new relational FCM (RFCM) algorithm was described. Section 5 shows the proposed protein spots quantification. Sections 6 and 7 present the experimental results and its discussion.

2. Previous Developments

Previous developments in this area have employed a wide range of techniques and a majority of applications address the matching of 2D gel electrophoresis images for proteins as opposed to DNA. This is because 2D gel electrophoresis images for proteins have a relatively more uniform background and are somewhat easier to work with than the 2D gel electrophoresis images for DNA.

Kim and Kim [2003] proposed a hierarchical segmentation based on thresholding and the detection of watersheds. They first preprocess the images to remove noise and enhance contrast, then thresholding is applied that produces large regions. A watershed detection algorithm is then applied recursively on these regions until only a single blob is detected that is considered to be a spot. Their method relies on setting several parameters and is sensitive to noise, and 2D gel electrophoresis images typically contain noise.

Sugahara *et al.* [1998] smoothed image regions by averaging pixel intensities using an mxm window and performed a thresholding operation that ultimately subtracted the background and then created a binary image for spot detection. This method relies heavily on the selection of a proper threshold value that can cause either an over-segmentation of spots in some regions as well as an under-segmentation of spots in other regions.

Takahashi *et al.* [1997] performed image enhancement and smoothing before defining local maxima in order to label the spots. This method also relies on the definition of threshold values in order to function properly.

More recently, Morris *et al.* [2008] developed a very accurate and robust method of detecting spots in 2D gel electrophoresis images. Their process involves an "average gel" that is created by first using registration software to create an alignment of all gels being used. The pixel intensities are then averaged across the aligned gels. The gels are each denoised using the average gel and pinnacles (regions that are a local maximum in both the horizontal and vertical directions and above a certain threshold) are detected, which denote the spot locations. A disadvantage of this method is the need to perform image registration as a preprocessing step and the need to define a threshold in order to determine which regions are pinnacles.

In their paper, Ijaz *et al.* [2003] presented a technique that uses the clustering techniques such as K-mean and FCM to distinguish different types of protein spots and unwanted artifacts.

Hoeflich *et al.* [2009] presented a new technique using the labeling of each image pixel as either a spot or nonspot and use a Markov random field (MRF) model and simulated annealing for inference. Neighboring spot labels are then connected to form spot regions.

Iakovidis *et al.* [2006] presented a novel approach to unsupervised protein spot detection in 2D-PAGE images based on a genetic algorithm. This algorithm searches within a multidimensional parameter space to determine, in parallel, the parameters of multiple diffusion models that optimally fit the characteristics of possible spots. The detection and quantification of the spots is achieved by superposition of diffusion functions modeling adjacent spots.

In their paper, Tsakanikas and Manolakos [2008] introduce the use of active contours without edges coupled with contour transform-based image enhancement for extracting accurately the gel image foreground (regions with spots) from the background. They demonstrate, using both synthetic and real gel images, that

the proposed approach extracts tight spot regions that do not include background areas but include almost all spots detected by PDQuest a popular commercial 2DGE image analysis package. Furthermore, their method does not require manual calibration for every new image in order to detect weak but often important "faint" spots.

Yoon *et al.* [2007] to find protein spots more accurately and reliably from gel images, propose reversible jump Markov chain Monte Carlo method (RJMCMC) to search for underlying spots that are assumed to have Gaussian-distribution shape. Their statistical method identifies very weak spots, restores noisy spots, and separates mixed spots into several meaningful spots that are likely to be ignored and missed. Their proposed approach estimates the proper number, center-position, width, and amplitude of the spots and has been successfully applied to the field of projection reconstruction (NMR PR-NMR) processing.

In this work, we intend to present a novel algorithm that uses the FCM algorithm as a primary step in the segmentation process.

3. The FCM Segmentation Algorithm

FCM method, also known as Fuzzy ISODATA, which was originally introduced by Bezdek in 1981 as an extension to Dunn's algorithm (see Dunn, 1974) is the most widely used fuzzy clustering algorithm in practice.

FCM is a data clustering technique based on optimizing the objective function:

$$J(U,V) = \sum_{j=1}^{C} \sum_{i=1}^{N} (\mu_{ij})^m \|x_i - v_j\|^2. \tag{1}$$

It requires every data point in the data set to belong to a cluster to some membership degree. The purpose of the FCM is to group data points into different specific clusters. Let $X = \{x_1, x_2, \ldots, x_N\}$ be a collection of data. By minimizing the objective function (1), $X$is classified into c homogeneous clusters where μ_{ij} is the membership degree of data x_i to a fuzzy cluster set v_j, $V = \{v_1, v_2, \ldots, v_c\}$ are the cluster centers. $U = (\mu_{ij})_{N*c}$ is a fuzzy partition matrix, in which each μ_{ij} indicates the membership degree for each data point in the data set to the cluster j.

The value of U should satisfy the following conditions:

$$\mu_{ij} \in [0,1], \quad \forall\, i = 1, \ldots, N, \quad \forall\, j = 1, \ldots, C \tag{2}$$

$$\sum_{j=1}^{C} \mu_{ij} = 1, \quad \forall\, i = 1, \ldots, N \tag{3}$$

The $\|x_i - v_j\|$ is the Euclidean distance between x_i and v_j. The parameter m is called fuzziness index, which control the fuzziness of membership of each datum. The goal is to iteratively minimize the aggregate distance between each data point in the data set and cluster centers until no further minimization is possible.

The whole FCM process can be described in the following steps.

Step 1: Initialize the membership matrix U with random values, subject to satisfying conditions (2) and (3).

Step 2: Calculate the cluster center V by using following equation

$$v_j = \frac{\sum_{i=1}^{N} (\mu_{ij})^m x_i}{\sum_{i=1}^{N} (\mu_{ij})^m}, \quad \forall j = 1, \ldots, C \tag{4}$$

Step 3: Get the new distance:

$$d_{ij} = \|x_i - v_j\|, \quad \forall i = 1, \ldots, N, \quad \forall j = 1, \ldots, C \tag{5}$$

Step 4: Update the Fuzzy partition matrix U:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C} (\frac{d_{ij}}{d_{ik}})^{\frac{2}{m-1}}}, \quad \text{if } d_{ij} \neq 0 \tag{6}$$

Else, $\mu_{ij} = 1$

Step 5: If the termination criteria have been reached, then stop. Else go back to step 2. The suitable termination criteria can be set by checking whether the objective function is below a certain tolerance value or if its improvement compared to the previous iteration is below a certain threshold. Moreover, the maximum number of iteration cycles can be used as a termination criterion as well.

4. Protein Spot Detection Utilizing the RFCM Algorithm

In this section, we present our new algorithm for matching protein spots in the 2D gel images. We called the new algorithm the RFCM, as it builds on the traditional FCM algorithm but is modified by introducing the notion of fuzzy relations in order to differentiate spot pixels from the varying background.

The algorithm is composed of four steps where the first step of it to apply the FCM to the image to produce preliminary clusters. Then, these clusters are then internally refined to identify the inner spots by separating the background pixel from the contained pixel in the cluster by applying steps 2 to 4 on the clusters.

A summary of the steps of the proposed algorithm is given below.

Step 1: Apply the FCM algorithm presented in Sec. 2 with C more than 2. The output is the partitioning of pixels in the image to different clusters each having a center value v.

Step 2: For each two pixels x, y belonging to two different clusters, create a fuzzy relation between x and y named $I(x, y)$ where $I(x, y)$ define the degree of closeness between intensities of pixels x and y.

Step 3: Compare pixels x, y
if pixel x is much more darker than pixel y
then pixel x is a spot pixel
else if pixel x is much more lighter than

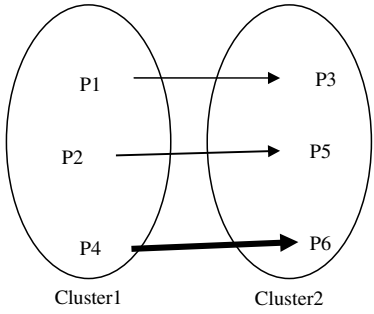

Fig. 1. Representation of the fuzzy relation between pixels in two clusters.

pixel y
then pixel y is a spot pixel
else if difference between intensities is low
and one of the pixels is a spot pixel
then the other is a spot pixel also
end if

Step 4: Mark spot pixels and differentiate them from the background. This is done by assigning the spot pixels to the maximum value of centers of clusters and the background — nonspot pixels — to the minimum value of centers of clusters.

In Fig. 1, a representation of a fuzzy relation between two points in two different clusters: cluster1 and cluster2, $I(x, y)$, is shown where x and y are any two points each in one different cluster. Here, the arrow represents the degree of closeness between two pixels from different clusters.

The dark arrow represents the strong relation between the two pixels in two different clusters.

5. Our Proposed Protein Spot Quantification

The previous work concerning protein spot quantification was based on 2D gel matching and identification [Gygi and Aebersold; (2008); Ijaz *et al.*, (2007)] after the detection step. In this section, we propose a novel technique for protein spot quantification based on the assumption that all spots have approximately same size. Area of protein spots can be identified from the sample preparation step. The SDS-PAGE (sodium dodecyl sulfate) in the second dimension separates proteins by their size (molecular weight, MW) and no other physical feature [see Berth *et al.*, (2007)].

This means that the average protein size in a 2D gel image is well known by the expert user who had prepared the gel. Having this knowledge, we can estimate the number of pixels per protein spot-drawing a circle having the same area and counting number of pixels inside the circle will yield to an appropriate pixels count

of a protein spot. From the RFCM segmentation step, the number of spot pixels in a 2D gel image is defined—refer to step 3 in the algorithm.

Now, the quantification of protein spots in an image is computed as follows:

$$\text{Number of protein spots} = \frac{\text{total No. of spot pixels}}{\text{No. of pixels per spot}} \tag{7}$$

This method is computational inexpensive; however, it will be proven to be accurate.

6. Experimental Results

The LECB 2-D PAGE gel images database is available for public use. It contains data sets from four types of experiments with over 300 GIF images with annotation and landmark data in HTML, tab-delimited, and XML formats. It could be used for samples of several types of biological materials and for test data for 2D gel analysis software development and comparison with other similar samples.

PAGE is polyacrylamide gel electrophoresis. The LECB was the U.S. National Cancer Institute's Laboratory of Experimental and Computational Biology. The database is available at the Web site (see bioinformatics.org/lecb2dgeldb).

In our work, we used these data and applied our algorithm to 2D gel images of fetal alcohol syndrome cell lines. The results are shown in the following figures. Figures 2 and 6 show four test cases 2D gel electrophoresis images of a patient — fetal alcohol syndrome. Figures 3 and 7 show the gradient images of the 2D gel images presented in Figs. 2 and 6.

The results of applying the current FCM Segmentation algorithm, on these images at $C = 2$, are shown in Figs. 4 and 8. The results of applying the proposed RFCM segmentation algorithm, on these images at $C = 6$ and $\beta = 20$, are shown in Figs. 5 and 9. Experimentally, $\beta = 20$ was shown to be the best parameter

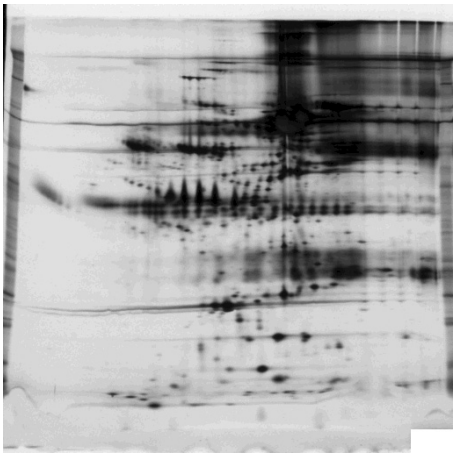

Fig. 2. 2D gel electrophoresis image of a patient with fetal alcohol syndrome.

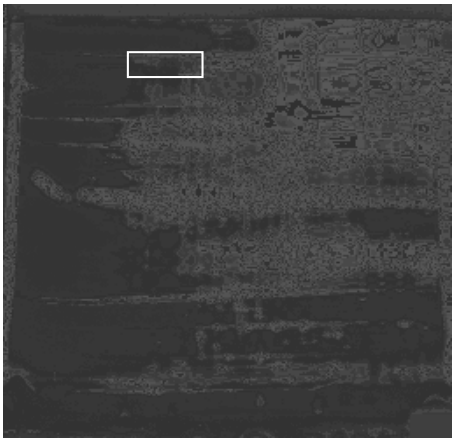

Fig. 3. The gradient image of 2D gel electrophoresis image of a patient with fetal alcohol Syndrome in Fig. 2.

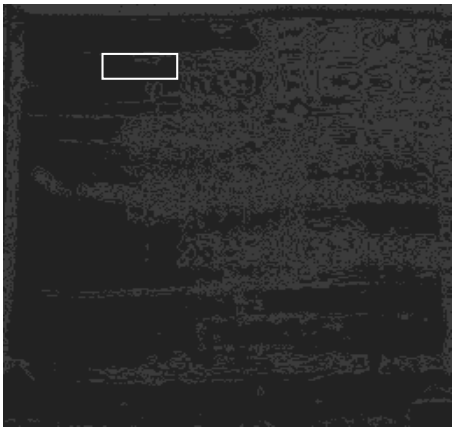

Fig. 4. The gradient image of 2D gel electrophoresis image of a patient-human with fetal alcohol syndrome in Fig. 2 after applying the FCM, No. of clusters = 2.

estimation to reduce error since when we increased the parameter to 22 or decreased to 18, image quality was poor and error increased so much.

The new technique shows high performance and detects the protein spots precisely, as shown in Figs. 5 and 9 even the less dark spots in the image appears (shown by squares) while in Figs. 4 and 8 when applying the FCM algorithm, those protein spots disappear totally that affects the spot quantization step in the whole process of 2D gel image analysis.

6.1. *The evaluation error E_{CW}*

In this section, we will use the E_{CW} evaluation error [see Chen and Wang (2004)] to evaluate the performance of the new algorithm the RFCM algorithm versus the

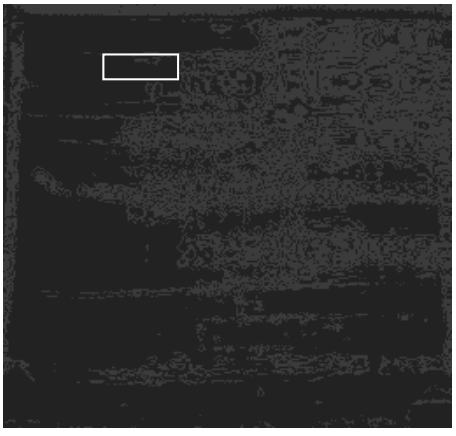

Fig. 5. The gradient image of 2D gel electrophoresis image of a patient with fetal alcohol syndrome in Fig. 2 after applying the RFCM, No. of clusters = 6, $\beta = 20$.

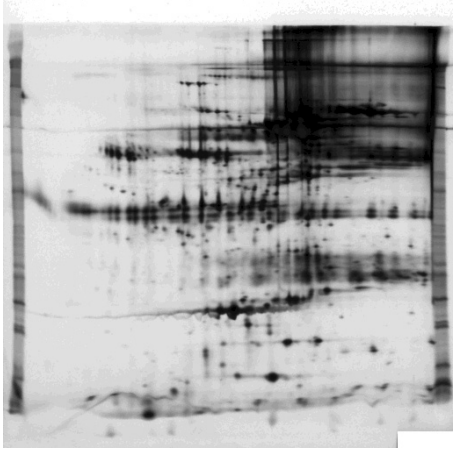

Fig. 6. 2D gel electrophoresis image of a patient with fetal alcohol syndrome.

original FCM algorithm. E_{CW} is a composite evaluation method for color images. It uses intraregion visual error to evaluate the degree of under-segmentation (Table 1 and Fig. 10) and uses interregion region visual error to evaluate the degree of over-segmentation (Table 2 and Fig. 11).

(1) E_{intra} of E_{CW}

$$E_{\text{intra}} = \frac{\sum_{p \in I} \mu(\|C_x^o(p) - C_x^s(p)\|_{L*a*b} - TH)}{S_I} \tag{8}$$

where $C_x^o(p)$ and $C_x^s(p)$ are pixel feature value (color components in CIE$L*a*b$ space) for pixel p on original and segmented image, respectively, TH is the threshold to judge significant difference, and $\mu(t) = 1$ when $t > 0$, otherwise $\mu(t) = 0$.

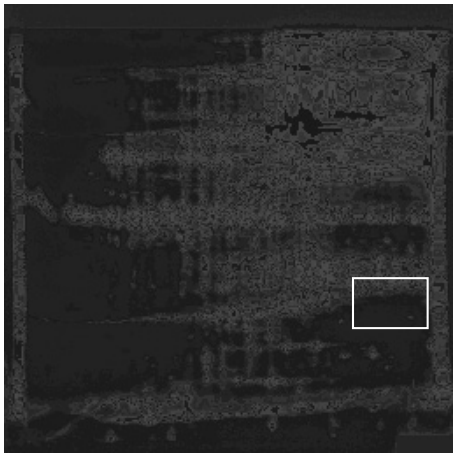

Fig. 7. The gradient image of 2D gel electrophoresis image of a patient with fetal alcohol syndrome in Fig. 6.

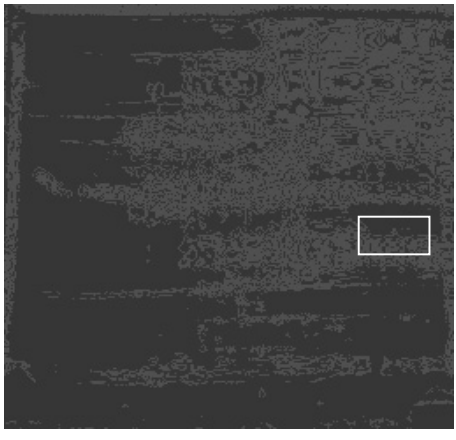

Fig. 8. The gradient image of 2D gel electrophoresis image of a patient with fetal alcohol syndrome in Fig. 6 after applying the FCM, No. of clusters = 2.

(2) E_{inter} of E_{CW}

$$E_{\text{inter}} = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} [\mu(TH - \|C_x^o(p) - C_x^s(p)\|_{L*a*b}) \cdot w_{ij}/(S_I \cdot Z)] \qquad (9)$$

where w_{ij} denotes the jointed length between R_i and R_j, TH is the threshold to judge significant difference, and Z is a normalization factor. We set the threshold $TH = 10$ and the normalization factor $Z = 100$ for No. of clusters = 2.

The shaded cells in Tables 1 and 2 represent the improvement caused by the new algorithm versus the original one. Notice that for the Eintra evaluation metric, the

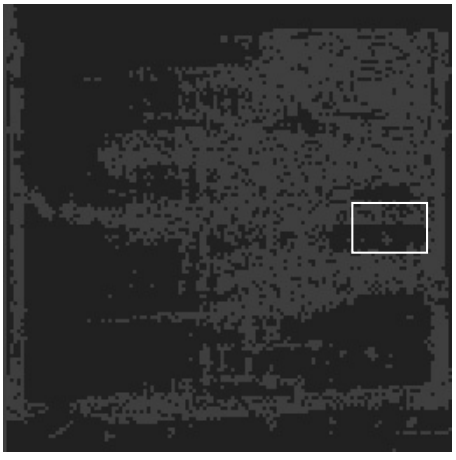

Fig. 9. The gradient image of 2D gel electrophoresis image of a patient-human with fetal alcohol syndrome in Fig. 6 after applying the RFCM, No. of clusters = 6, $\beta = 20$.

Table 1. The intraregion error of the FCM algorithm and the RFCM algorithm on seven data samples.

No.	Eintra segmented image by FCM	Eintra segmented image by RFCM
1	0.96712	0.90608
2	0.91501	0.98128
3	0.46035	0.55992
4	0.93101	0.136
5	0.98302	0.86736
6	0.95079	0.99768
7	0.83952	0.26396

new RFCM algorithm enhanced the results on four data samples over seven data samples (57%) and in case of nonimprovement, the difference was about 10% in the worst case (data sample 3), which means that the new algorithm was able to detect protein spots more precisely. Moreover, according to the Einter evaluation metric, the new RFCM algorithm — except for data samples 3 and 4 — had reduced the over-segmentation or at least did not increase so much its rate.

6.2. *Evaluation by expert user*

Images used in this section were provided by Prof. Dr. Maha ElDemellawy, Medical Biotechnology Department, City of Science and Technology, Borg ElArab, Alexandria, Egypt. We applied the RFCM algorithm to compute total number of spot pixels in the image while we knew the number of pixels per spot by choosing, with the help of the expert user, a spot and calculating its size and number of pixels inside it.

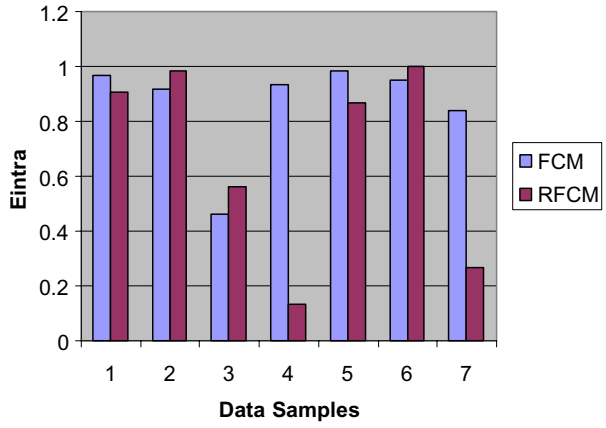

Fig. 10. The intraregion error of the FCM algorithm and the RFCM algorithm on seven data samples.

Table 2. The interregion error of the FCM algorithm and the RFCM lgorithm on seven data samples.

No.	Eintra Segmented image by RFCM	Eintra Segmented image by FCM
1	0.17	0.22978
2	0.01604	0.0040217
3	0.16751	0.47053
4	0.015092	0.27255
5	0.24	0.047613
6	0.0091133	0
7	0.23718	0.27355

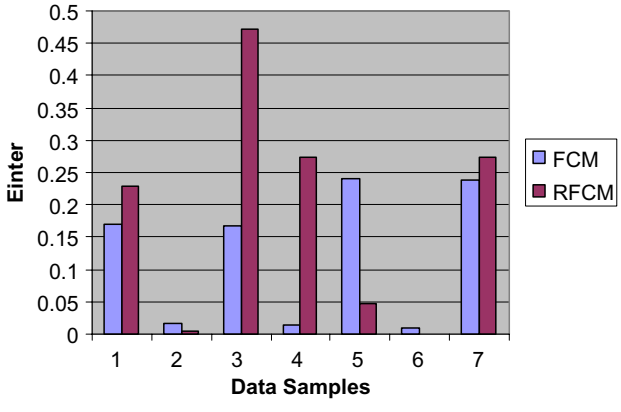

Fig. 11. The intraregion error of the FCM algorithm and the RFCM algorithm on seven data samples.

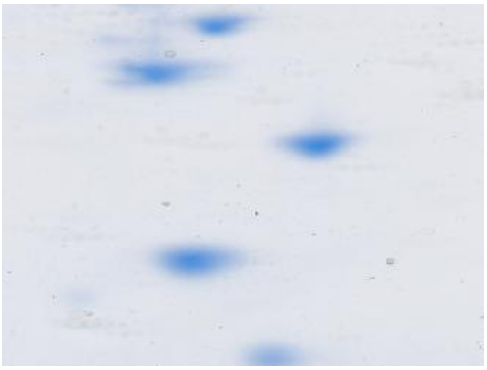

Fig. 12. 2D gel electrophoresis image of human myocardial membrane proteins.

Figure 12 shows 2D gel electrophoresis image of human myocardial membrane proteins. In a 128×128 2DGE image, pixels per spot (average spot size in pixels) = 176 pixels; No. of spot pixels according to the RFCM algorithm = 1850 pixels; No. of protein spots (automatic quantification using Eq.7) = $1850/176 = 10.5111$ protein spots (i.e. about 10 spots); and No. of protein spots (detected by expert user quantification) = 10–15 protein spots.

Figure 13 shows 2D gel electrophoresis image of human myocardial membrane protein 50–54 kDa fraction. In a 128×128 2DGE image, pixels per spot (average spot size in pixels) = 200 pixels; No. of spot pixels according to the RFCM algorithm = 4075 pixels; No. of protein spots (automatic quantification using Eq.7) = $4075/200 = 20.375$ protein spots (i.e. about 20 spots); and No. of protein spots (detected by expert user quantification) = $15 - 18$ protein spots.

7. Discussion

In this work, we presented a new algorithm based on the notion of fuzzy relations to segment and detect protein spots in 2D gel electrophoresis images and used it for automatic quantification of protein spots.

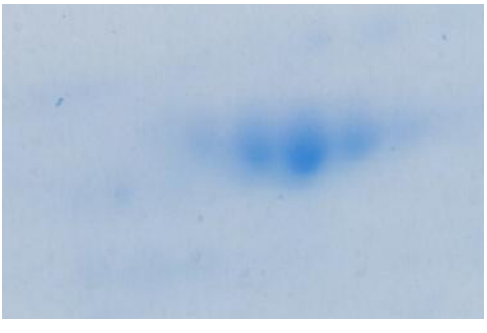

Fig. 13. 2D gel electrophoresis image of human myocardial membrane protein 50–54 kDa fraction.

For future work, we suggest the development of fuzzy relations to obtain better results. The parameters representing the degree of closeness must be defined for enhancement and improvement of the RFCM algorithm. The use of intuitionistic fuzzy relations to identify the degree of noncloseness between pixels and the hesitation margin can be investigated also.

The automatic quantification of protein spots in 2D gel electrophoresis images may open a new direction of research in the proteomics field. Knowing that there are diseases, such as cancer [see Zhou *et al.* (2002); Kim and Kim (2003)], that affect the quantity of protein spots in 2DGE images, we can suggest a machine-learning system that uses the quantity of protein spots in a 2DGE image as a factor that classifies the cells into normal cells and abnormal cells and this can be considered as a step toward diagnosis of a disease using 2DGE images analysis.

Now, a question is raised: Is the quantity of protein spots in a 2D gel images the only factor affected by cancer disease or the color intensity or the shape too? The existence of more than one factor means that the suggested machine-learning system can take an input vector of multi-dimensions that yield to more accurate classification of normal and abnormal cells.

Acknowledgment

Special thanks for Prof. Dr. AbdelFattah AbdelMongy Sadakah, president of University of Tanta.

References

Anderson, N. G. and Matheson, A. (2001). Back to the future: The human protein index (HPI) and the agenda for post-proteomic biology. *Proteomics*, **1**: 3.

Berth, M., Moser, F. M., Kolbe, M., Bernhardt, J. (2007). The state of the art in the analysis of two-dimensional gel electrophoresis images. *Appl. Microbiol. Biotechnol.*, **76**: 1223–1243.

Blackstock, W. P. and Weir, M. P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol.*, **17**: 121–127.

Chen, H.-C. and Wang, S.-J. (2004). The use of visible color difference in the quantitative evaluation of color image segmentation. *Proceedings of ICASSP.*

Dunn, J. C. (1974). A fuzzy relative of the ISODATA process and its use in detecting compact, well separated clusters. *J. Cybern.*, **3**(3): 32–57.

Gygi, P. S. and Aebersold, R. (2008). Absolute quantitation of 2-D protein spots. *Meth. Mol. Biol.*, **112**: 417–421.

Hoeflich, C. S. and Corso, J. J. (2009). Segmentation of 2D gel electrophoresis spots using a Markov random field. *Proceedings of SPIE Conference on Medical Imaging.*

Höper, D., Bernhardt, J. and Hecker, M. (2006). Salt stress adaptation of *Bacillus subtilis:* a physiological proteomics approach. *Proteomics*, **6**(5): 1550–1562.

Iakovidis, D. K., Maroulis, D., Zacharia, E. and Kossida, S. (2006). A genetic approach to spot detection in two-dimensional gel electrophoresis images. *Proceedings of International Conference on Information Technology in Biomedicine (ITAB)*, Ioannina, Greece.

Ijaz, U. Z., Chaudhary, S. U., Don, M. S. and Kim, K. Y. (2007). Computational strategies for protein quantitation in 2D electrophoresis gel image processor for MATLAB. *Proceedings of the 2007 Frontiers in the Convergence of Bioscience and Information Technologies*, pp. 129–134.

Kim, G. and Kim, M. H. (2003). 2D Electrophoresis gel image and diagnosis of a disease, CoRR cs.CC/0305048. G.

Kim, Y., Kim, J., Won, Y. and In, Y., (2003). Segmentation of protein spots in 2D gel electrophoresis images with watersheds using hierarchical threshold. *LNCS–Computer and Information Sciences - ISCIS*, **2869**: 389–396.

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis in mouse tissues: A novel approach to testing for induced point mutations in mammals. *Human Genetik*, **26**: 231–243.

Klose, J. and Kobalz, U. (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis*, **16**: 1034–1059.

Morris, J. S., Clark, B. N. and Gutstein, H. B. (2008). Pinnacle: a fast, automatic and accurate method for detecting and quantifying protein spots in 2-dimensional gel electrophoresis data. *Bioinformatics*, **24**(4): 529–536.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.*, **250**: 4007–4021.

Rashwan, S., Faheem, T., Sarhan, A. and Youssef, B. A. (2010). A relational fuzzy c-means algorithm for detecting protein spots in two-dimensional gel images, accepted to be published in Advances in Computational Biology, Springer 2010.

Sugahara, Y., Hayashizaki, Y. and Tanihata, I. (1998). An automatic image analysis system for RLGS_LMS. *Mamm. Genome*, **9**: 643–651.

Takahashi, K., M. Nakazawa and Watanabe Y. (1997). DNAinsight: An image processing system for 2-D gel electrophoresis of genomic DNA. *Genome Informatics*, **8**: 135–146.

Tsakanikas, P., and Manolakos, E. S. (2008). Active contours based segmentation of 2DGE proteomics images. *Proceedings of Eusipco.*

Voigt, B., Schweder, T., Sibbald, M. J., Albrecht, D., Ehrenreich, A., Bernhardt, J., *et al.* (2006). The extracellular proteome of *Bacillus licheniformis* grown in different media and under different nutrient starvation conditions. *Proteomics*, **6**(1): 268–281.

Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R. *et al.* (1995). Large-scale amino-acid analysis for proteome studies. *Electrophoresis*, 16.

Yoon, J. W., Godsill, S. J., Kang, C. and Kim, T.-S. (2007). Bayesian, Inference for 2D gel electrophoresis image analysis. *Bioinformatics Research and Development*, pp. 343–356.

Zhou, G., Li, H., DeCamp, D., Chen, S., Shu, H., Gong, Y. *et al.* (2002). 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Mol. Cell. Proteomics*, **1**: 117–123.