

A NEW DELAUNAY PROCEDURE FOR MATCHING 2D-PAGE IMAGES

S. COCO, G. CAPIZZI, A. LAUDANI* and R. MESSINA

*DIEES, University of Catania,
Catania, I-95125, Italy*

**E-mail: alaudani@diees.unict.it*

E. ALESSI

STMicronics

In this paper we present a new Delaunay procedure for matching spots in 2D-PAGE images. For each spot similarities are found in terms of geometric characteristic and intensity. For this reason in our procedure two phases are considered addressing separately geometrical matching and intensity matching. In comparison with other more complex matching tools our procedure is simple, robust and shows good accuracy and efficiency. An example of an 2D-PAGE images matching is also shown in order to better illustrate the procedure.

Keywords: Electrophoresis; 2D-page images; Delaunay triangulation; Proteomics

1. Introduction

Proteomics is a fundamental knowledge for molecular biology. Proteomics research is based on the systematic analysis of complete profiles of the proteins expressed by a given cell, tissue or biological system.¹ In this field 2D-gel electrophoresis (2D-PAGE) is a powerful technique used to separate thousands of proteins on the basis of their isoelectric focusing points and their molecular weights. Using staining methods, the proteins can be individuated in terms of their properties and features, such as size, intensity and position on the gel. The result of this technique is an 2D-PAGE image containing an 2D pattern of spots, in which each spot represents a protein. This image will have to be suitably analyzed in order to detect the maximum information on the proteins. The goal of this operation is to determine proteins which undergo variations in their expressions under specified circumstances. For this reason a huge number of 2D-PAGE image are to be

compared in order to find similarities and detect modifications.²⁻⁵ Therefore automatic specialized tools are needed, based on sophisticated and efficient algorithms for processing electrophoretic images. In this context matching procedures are aimed at identifying similar spot configurations belonging to different 2D-PAGE images. Among the various techniques, Delaunay triangulation⁶ is widely used for matching 2D-PAGE images, because it maintains spot similarities. However, the comparison of images after a simple Delaunay triangulation is quite a difficult task and may lead to some errors.^{7,8} In this paper we present a new more robust Delaunay procedure for matching spots in 2D-PAGE images. For each spot similarities are found in terms of geometric characteristic and intensity. For this reason in our procedure two phases are considered addressing separately geometrical matching and intensity matching. In comparison with other similar matching tools our procedure is simple, robust and shows good accuracy and efficiency. The matching procedure has been implemented in C language; an apposite visual interface has also been developed in Visual C++ environment.

2. The Delaunay Matching procedure

In our procedure a target 2D-PAGE image is compared with another one belonging to a set of pattern 2D-PAGE images within a predefined catalogue. First a pre-processing of the 2D-PAGE images is performed in order to eliminate noise and distortions. Then a spot detection procedure is applied to the 2D-PAGE images in order to extract the spot configurations to be processed by the Delaunay algorithm. The result of this phase is a new image where all the spots representing the sample proteins are individuated by their barycentres. At this point a Delaunay triangulation is performed for the spot configurations to be compared. In the Delaunay triangulation spots are linked by edges and a number of triangles are generated covering all the spot regions in the images. In these processed images similarities are found in terms of several spot features, such as geometric and intensity characteristics (position, area, intensity, shape) of each spot within the images. Since matching is based on local properties of spots, that is mutual positions and distances of spots considered on a local basis, the sizes of target and pattern images can be significantly different. The different features considered for the matching are referred to by distinguishing a geometrical matching and an intensity matching, which are addressed separately. In the procedure the geometrical and intensity information of each spot is represented by its barycentre coordinates $(x(s), y(s))$ and by a number

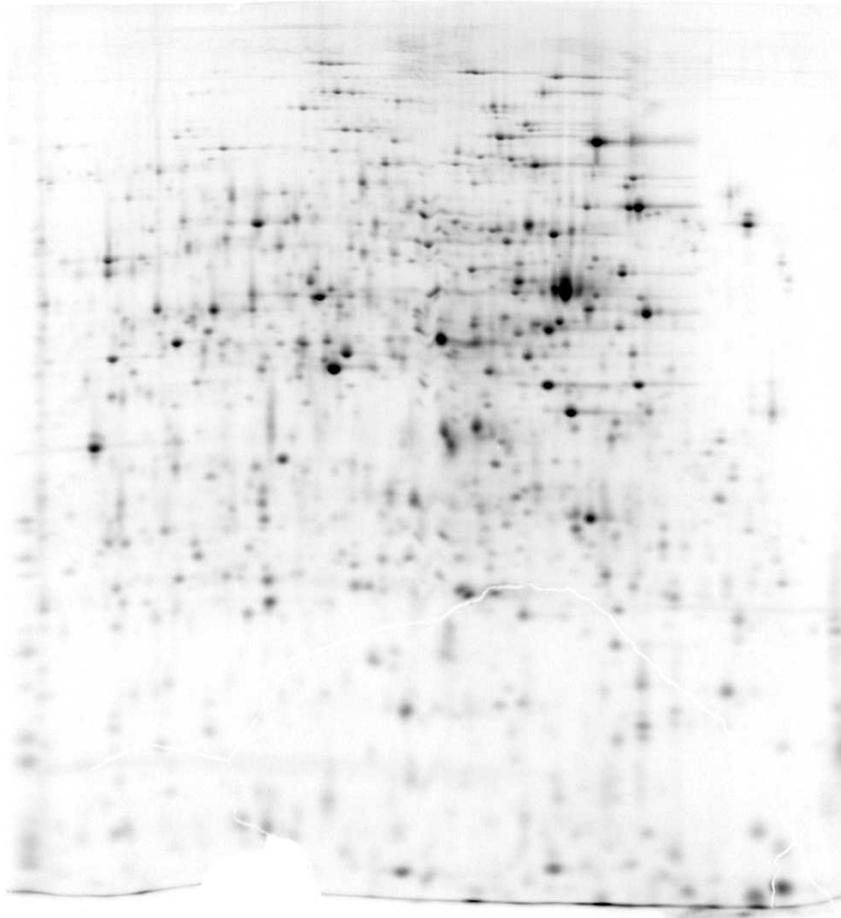


Fig. 1. Original 2D-PAGE image for spot detection

describing its mean intensity $I(s)$. Two kinds of operations are performed starting from intensity information. First a normalization of intensity values is performed using as normalizing value the absolute maximum intensity found in the image. Secondly a clustering of spots is carried out according to intensity values: for this aim ten intensity classes are defined. Each spot is assigned to a certain class according to its mean intensity I , suitably weighted. The weighting function $f(I)$ has been suitably chosen in order to perform easily an alignment of the intensities of each spot belonging to the

two images; its expression is:

$$f(I) = 10[-2(1 - c)I^3 + (3(1 - c)I^2) + cI] \quad (1)$$

where the constant c can be chosen in the range 0-1. The first phase of the matching is aimed at finding the greatest number of occurrences of the Pattern image in the Target image. For this purpose the correspondence is investigated between each spot belonging to the Pattern image and each spot belonging to the Target image, with respect to geometric conditions, relative to edges linking spots in the Delaunay triangulation. For this purpose two edge lists are considered $TotEdgeList(P)$ and $TotEdgeList(T)$, containing all the edges belonging to the pattern image and target image respectively.

Two edge: $\bar{s}s' \in TotEdgeList(P)$ and $\bar{t}t' \in TotEdgeList(T)$ are said:

- α similar (angle criterium) if the absolute difference of their angles with respect to a reference horizontal axis satisfies the following condition:

$$|Slope(\bar{s}s') - Slope(\bar{t}t')| < \alpha \quad (2)$$

- λ similar (distance criterium) if the euclidean distance satisfies the following condition:

$$1 - \lambda \leq \frac{|ss'|}{|tt'|} \leq 1 + \lambda \quad (3)$$

For the geometrical matching similarity is found in terms of the above two edge criteria. Two edges are declared similar when their length ratio as well as their mutual orientation measured as their slope with respect to a frame, differ less than a prescribed tolerance, suitably assigned.

The geometrical matching does not lead to biunivocal results, because a multiplicity of spots belonging to the Pattern image can be mapped to only one spot belonging to the Target image. This is the reason why a third criterium is introduced relative to intensity, in which two edges are evaluated according to the values of intensity of each spot's couple that satisfies the geometrical matching

- Intensity criterion is verified if the intensity values of edges satisfy the following condition:

$$|I_{pattern} - I_{target}| \leq I \quad (4)$$

The parameter I is chosen in the range 1-2.

When edges are found for which the geometrical and intensity matching is successful, a list of matching candidates is created, to be used in the succeeding phase. In this phase the matching mapping between the elements selected from the two lists is not biunivocal since a certain edge belonging to the target image can have more than one matching edge in the pattern list. Then a congruence test is performed by computing the effective number of matching spots only on the basis of geometrical comparisons. The congruence test works by checking mean distances between a spot and all its neighbouring spots. At the end the maximum number of corresponding spots in two images is computed: this is a "measure" of the matching of the two images. As expected it has been verified that the matching between two identical images leads to the perfect overlap of the two spot lists.

3. An example of application

In this example two 2D-PAGE images coming out from similar proteic samples are compared. The matching procedure has been applied to a list of spots coming out from the 2D-PAGE image shown in fig. 1 by using a spot detection procedure (fig. 2). For this target image a Delaunay triangulation has been performed obtaining the results shown in fig. 3. The comparison between the target image and the pattern from the sample has been performed according to the geometrical and intensity matching phases described above. The geometrical matching conditions have been verified assuming for the two parameters α and λ the values 0.1 and 0.2 respectively. The intensity matching parameter has been set to 2. After the congruence test has been verified, a very good matching (more than 60%) between the two images in terms of spot correspondence is obtained as shown in fig. 4 .

4. Conclusions

The presented matching procedure works on Delaunay triangulations of a target and a pattern image. The procedure addresses separately matching of geometrical features and intensity characteristics of the 2D-PAGE images. The geometrical matching is based on similarity criteria for spot configurations whereas for the intensity matching a condition on the mean intensity of the spots in the image is considered and only the candidate spots fulfilling this condition are declared fully matching. In comparison with similar tools our procedure is simple, robust and efficient and obtains accurate results.

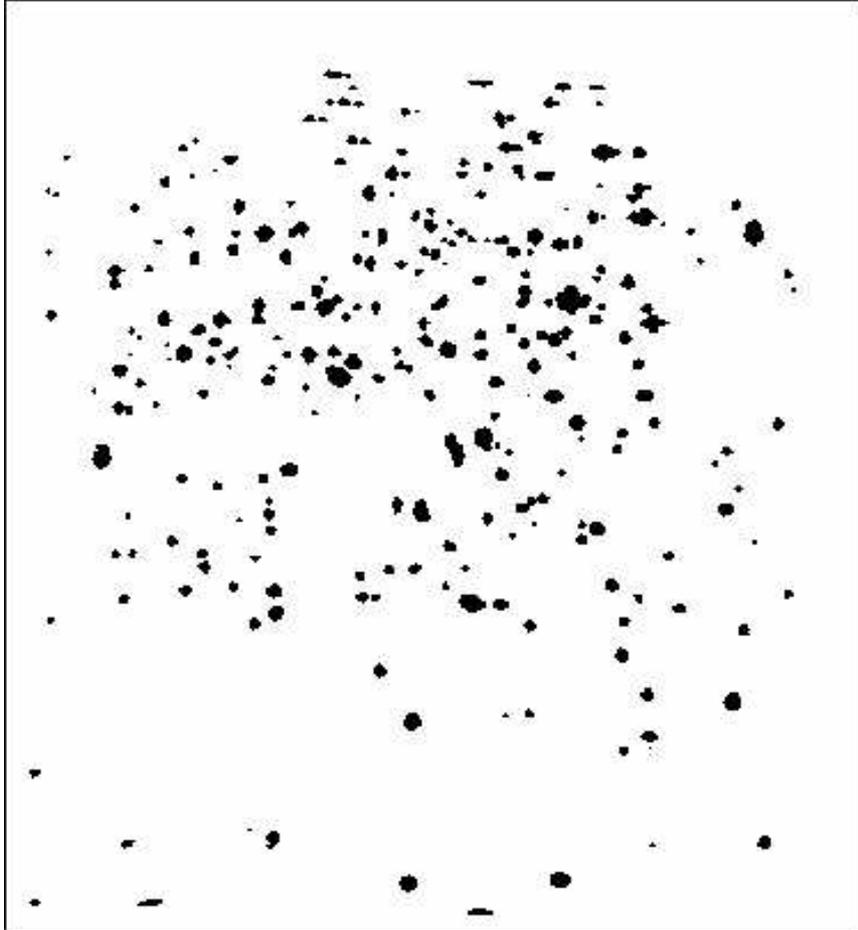


Fig. 2. Elaborated 2D-PAGE image for spot detection

References

1. M. Wilkins and alii(eds.), *Proteome Research: New Frontiers in Functional Genomics* (Springer-Verlag, Berlin Heidelberg New York, 1997).
2. R. Appel and alii, *Electrophoresis* **18**, 2735 (1997).
3. P. F. Lembkin, *Electrophoresis* **18**, 461 (1997).
4. a. a. Chang, S.-H., *Pattern Recognition* **30**, 311 (1997).
5. H. Alt and alii, Carol - new algorithmic tools for comparing two-dimensional electrophoretic gel images, in *Proceedings Electrophorese Forum '97*, (Strasbourg, 1997).
6. K. D. Guibas, L. and M. Sharir, *Algorithmica* **7**, 381 (1992).

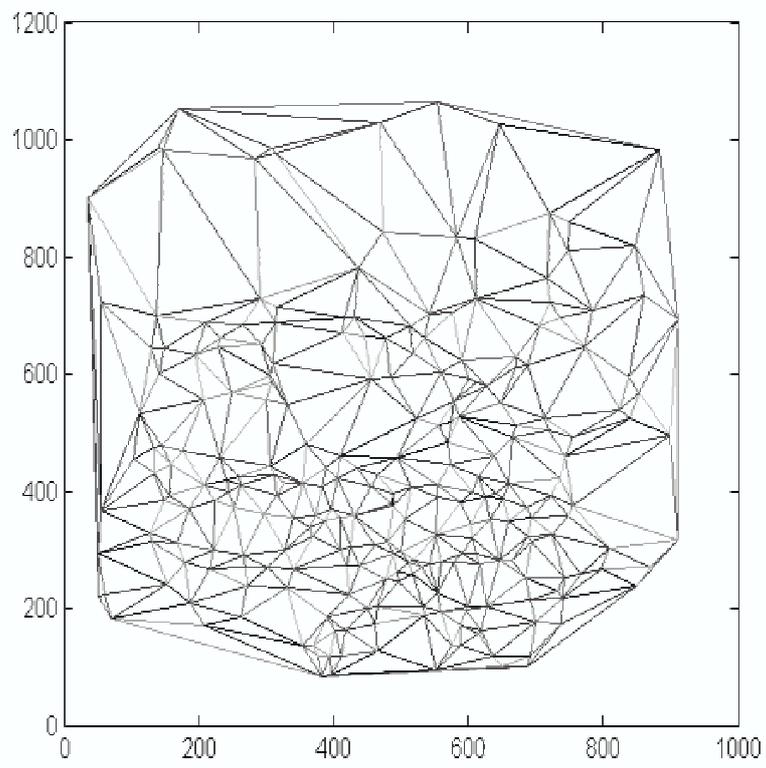


Fig. 3. Delaunay triangulation

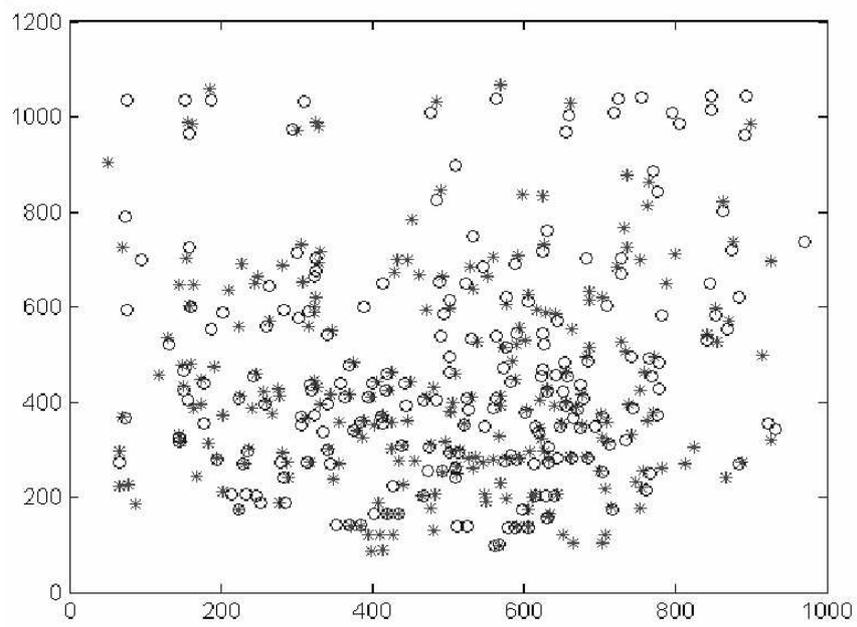


Fig. 4. Matching results

7. R. Finch and alii, *Pattern Recognition* **30**, 123 (1997).
8. H. Ogawa, *Pattern Recognition* **19**, 35 (1986).