Translation, Scaling and Rotation Invariant Spot Matching using Delaunay Triangulation

ANKHBAYAR YUKHUU*, HWANG YOUNG SUP**
Department of Computer Science,
Sun Moon University,
100 Kalsan-ri, Tangjeongmyeon, Asan, Chugnam, 336-708
SOUTH KOREA

Abstract: - Point pattern matching is one of the most powerful methods of spot identification in a gel image. We propose an algorithm to determine the matching of each point pair and its transformation parameters. The main feature of the presented solution is that we use a local point pattern matching approach using Delaunay triangulation of a point sets to find spot pairs. As a consequence, the exact position of spot pairs are determined by Delaunay triangulation and the proposed approach can handle translation, scaling and rotation differences. Our local matching algorithm can be used as a basic step to the global matching and as a landmark setting.

Key-Words: - two dimensional gel image analysis, point pattern matching, distance function, Delaunay triangulation, Affine transformation

1 Introduction
Recent years have seen exceptional improvements in the dependability and ease of use of proteomic technologies. However, precise, impartial, and rapid gel analysis has persisted as a major interest in the field of proteomics. To carry out a differential investigation, it is necessary to determine correspondence between spots on sets of gel images. One of the tasks in the analysis of these images is matching of protein spots in two corresponding images for differential expression study. This implies that a transformation relating one gel image to another is required.

Comparison of two images is important to detect the changes, because change of spots maybe related to some diseases. Variation in gels is a problem. Common causes for gel variations are: difference in sample preparation, experimental conditions, physical and chemical processes. For example, the image patterns are distorted, some spots are very light, spots can be overlapped on each other, corresponding spots may not be found between images, the background intensity is non-uniform, and noise is present. These variations make it difficult to get correct perfect matches. Therefore, several methods have been developed to compute the differences of shapes and types for comparing two spots. Gel-to-gel matching, a differential comparison is carried out by matching spots (or features) that may be in partly different locations on the two gels. Several techniques and software packages have been proposed for analysis of two dimensional electrophoresis (2DE) gel images. For spot matching, early methods require heavy user interaction especially for initial spot pairing [1]. Piecewise bilinear mapping is among a widely used transformation, applied piecewise bilinear mapping on multi resolution images with feature mapping by image cross correlation [2-4], and used piecewise bilinear mapping in a hierarchical grid transformation with stochastic optimization method, and Garrels[5] published a spot matching algorithm based on propagation of a user-defined landmark spot, which served as the initial match. The next match was determined by a neighborhood spot of the spot which showed the best match. This spot became a landmark and the algorithm continued until no further sufficiently good match in the surrounding of the new landmark was found. Then next landmark was defined and the algorithm looped until no new matches were found. Another approach, applied CAROL system [6, 7] enables the user to check and correct offered matches. The corrected proposal of the initial matching is used as landmark for the following matching extension. The matching extension is based on a piecewise affine transformation given by landmarks. The process can be repeated recursively until no new matches are found. The other approaches using different features of gel images were used. First, systems utilized the principle of transformation of the coordinate system of a matched gel to fit the reference gel using a set of user-defined landmark spots.
Each landmark spot defined a known match and was introduced by the user. A set of landmark coordinates was then loaded to a least squares minimizer and parameters of a known function were computed. The matched gel coordinate system was recomputed using the known function to the coordinate system of the reference gel. The transformed gels were then overlaid and the matched spots were defined as those closest to each other [8].

Most of the algorithms mentioned need landmarks to define spots pair to determine neighborhood spots. And some of the algorithms integrated to commercially available gel analysis packages. If we set landmarks by users, spot matching becomes the bottleneck of the whole process, as the manual landmark setting requires time-consuming operator interaction and manual corrections.

In this paper we present a point pattern approach for matching pairs of gels based on a Delaunay triangulation without manual a landmarks setting.

2 Materials and methods
The problem is defined as finding corresponding protein spots of two input gel images when spot regions of both images are given.

In our approach, the initialization of corresponding spot pairs is based on a Delaunay triangulation of point set, and the final matching is based on affine transformation of coordinate system.

2.1 Point pattern matching
The position of a protein spot is believed to be same as its isoelectric point (pI) and molecular weight (Mr). In practice, these recorded values are different from real locations of the proteins. To a much greater extent this applies to the computer assisted comparison, where one first applies an algorithm to detect spots and to extract their features like spot size, spot intensity, or spot shape. This so described spot detection stage is a necessary preprocessing step, along with, a harsh error source for the subsequent spot matching problem. For the purpose of illustration in Fig.1 small rectangular window regions marked. Many of the proposed algorithms make use of so-described landmarks and a general alignment of the images by warping techniques [9].

Landmarks are spot pairs interactively marked in both images by the user and selected as supposed matching pairs. In order to align the locations of corresponding spots among gel images, a set of landmark spots should be marked in every gel image [10]. The common properties of a landmark useful for a sample group of the same tissue can be found by studying the spots of the landmark in a number of training gel images. Therefore, as the number of landmark spots specified in an individual gel image is increased, the result of spot matching becomes more accurate. However, manually annotating a number of landmark spots for a large number of gel images is a time-consuming and error-prone task. Therefore automatic landmark setting is unnecessary.

The CAROL’s algorithm especially relies on a data structure derived from the incremental Delaunay triangulation of a point set and several heuristics to cope with distortions and noise inherent to the electrophoresis process described that interactive landmark setting is optional and not necessary [11].

One can do the spot matching either relying primarily on the pixel level information or on derived geometric information like in our solution.

We want to propose the following algorithm of point pattern matching. Given a local spot patterns of P and Q selected from a 2DE gel images that have a similar appearance to at least partially both the geometric shape and the spot intensities of pattern P. To illustrate an instance of the local matching problem consider the spots point sets in Fig.2 which come from Fig.1. Some points are excluded because the spots intensities is too least from those of the pattern spots.
Given a point set, an original $P$ and another reference point set $Q$ in a Euclidean space, one wants to compute all occurrences of $P$ in $Q$. For exact matching and approximate solutions, a local matching algorithm can be used as a basic step to a global matching in gel images. In fact, local matching can be used as the landmark setting. The incremental Delaunay triangulation, is proved to be a suitable structure for the local matching problem because of its expected linear size and its robustness in the presence of noise [12]. Thus it is implemented and is part of the Carol system. They constructed the Delaunay triangulation graph rather than the final graph itself. And they used it for the matching process. The approach worked well in the case of noise.

However, comparison of matching results in the literature is still insufficient because of the variations in producing the gel images. Two images of the same object may be obtained from different viewpoints, from different sensors, or from the same sensor at different times [13]. Therefore, two point patterns have translation, rotation, and scaling differences. Most of the time, however, these geometrical quantities are not known. Of the corresponding points are known, the geometrical transformation parameters (translation, rotation, and scaling) can be found by using least-squares technique to estimate the least squares error between these corresponding points.

2.2 Problem formulation

All spots in selected gels that are paired with a given spot in the reference gel form a group. A spot group is the basic element in analyzing spot variations across gels, in producing reports and histograms, as well as performing statistical and clustering analysis. Moreover, when several gels have been matched to a given reference gel, this reference gel provides a unique numbering scheme for spots across all gels. Indeed, each paired spot in a gel image may be associated with the corresponding in the reference gel[14-15].

Having specified how to define a region pair one question remaining is how to apply the proposed regional local point matching to the entire gel areas, $P$ and $Q$, but computational unsuitable way would be to slide the region pair around the gel area then solve the matching problem for all possible positions of the region centers.

Assume we want to choose local patterns $P$ and $Q$ from an original and reference gel images. To find corresponding spots of two corresponding input gel images can have same spot numbers, when spot centroids and spot regions of both images are given. To determine spot pairs, a shortest distance measure $D$ between patterns is calculated by Euclidean distance function.

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

By plotting the two gel images over the same plane, we can have an intuitive visualization of the geometrical variations between the two 2DE gel images, as an example in Fig.3.

![Figure 3. Plotting the two sets of spots patterns over the same plane (+: spots of an original gel, O: spots of a reference gel)](image)

The finding spot regions between $P$ and $Q$ are done via Delaunay net. At first we generate the Delaunay net from an original and reference spots. The initials points as corresponding spots. Then search along with using Delaunay nets to match spots pair. Then we formulate the affine transform for spot matching. The affine transform is transparent to geometric variants (shifts, rotations, scaling and reflection).

Global matching routine compares the whole source image to the whole target image. Let two point patterns $P$ and $Q$ consist of $m$ feature points and $n$ feature points, respectively to compute an overall list of spot pairs that correspond to each other. The feature point is the description of the spot location $(x, y)$. The matching procedure is to find a corresponding between a point $p_i$ in $P$ and a point $q_j$ in $Q$, that makes matched pair consistent under a registration $Gr(tx, ty, s, \theta)$ where $tx$ and $ty$ are the translation along the $x$-axis and $y$-axis, $s$ is the scaling factor and $\theta$ is the rotation angle. In summary, it aims to map a point $p=(x_p, y_p)^T$ to a point $q=(x_q, y_q)^T$ as follows:

$$q = G_r(p) = \begin{bmatrix} x_q \\ y_q \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \end{bmatrix} \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} s \cos \theta & -s \sin \theta \\ s \sin \theta & s \cos \theta \end{bmatrix} \begin{bmatrix} x_p \\ y_p \end{bmatrix} \quad (2)$$
When we calculate an affine transformation, we use a coordinate system transformations to register different sets of coordinates in the same area. The spots may have come from a different images. Let \((x, y)\) be the location of the spot before transformation in the original image, and \((u, v)\) be the location of the object after transformation in the reference image.

Translation (shifts): Origin is moved \(a\) units parallel to \(x\) and \(b\) units parallel to \(y\)
\[
\begin{align*}
  u &= x - a \\
  v &= y - b
\end{align*}
\]

Scaling: Both origin and axes are fixed, but scale changes, scaling of \(x\) and \(y\) may be different (if the scaling is different, the shape of the object will change)
\[
\begin{align*}
  u &= cx \\
  v &= dy
\end{align*}
\]

Rotation: Origin fixed, axes move
\[
\begin{align*}
  u &= x \cos(\theta) + y \sin(\theta) \\
  v &= -x \sin(\theta) + y \cos(\theta)
\end{align*}
\]
\(\theta\) is measured counterclockwise

Reflection: Coordinate system is reserved, objects appear in mirror image, to reserve \(y\), but not \(x\)
\[
\begin{align*}
  u &= x \\
  v &= c - y
\end{align*}
\]

The combined equations are:
\[
\begin{align*}
  u &= ax + cy + e \\
  v &= bx + dy + f
\end{align*}
\]

3 Result and discussion

Using Delaunay triangulation, ones build a mesh of connection between the spots. Mesh generation consists of choosing points whereas triangulation is a partition of a mesh.

When we transform gel image into geometrical point pattern and determine exact spots position, we make the spots from different gel images to be over the same plane and build the Delaunay triangulations to find spot pairs and region.

Delaunay triangulations are built on top of the spot patterns in the same plane as in Fig.4. From the result, spots are paired through edges.

We performed experiments on different number of spot patterns for spot matching. As with spot detection, the matching accuracy is evaluated as true-positives (TP), false positives (FP), and false-negatives (FN).

TP is the number of spot pairs that were correctly reported. FP is the number of incorrect spot pairs that were mistakenly overlapped as spot pairs. FN is the number of pairs that are missed by the program. For both image sets, program correctly found spot centroids in every image.

<table>
<thead>
<tr>
<th>Number of spots</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>Ratio of matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>46</td>
<td>4</td>
<td>0</td>
<td>0.92</td>
</tr>
<tr>
<td>100</td>
<td>91</td>
<td>9</td>
<td>0</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table 1. Spot pairs searching result using the pair of gel images as included 50, 100 pairs of spots.

We compared our approach traditional spot matching approach used by minimum distance search algorithm. From the result traditional approach was more time consumed than our approach.

For spot matching, there are some difficulties in matching of the real gel images, such as difference of image illuminations, repetition of local spot patterns, and distortion of gel spots, and also noise in images.

Figure 4. Delaunay triangulations built on top of the spot patterns.

Figure 5. The spot pairs matched between two gel images
However, our approach gives the highest TP in every image pairs and after the transformation process, each spot is matched correctly. Fig. 5 shows the results of final experiments.

4 Conclusions

We proposed a point pattern method for spot matching pairs of 2DE gel images based on a Delaunay triangulation to find pairs of spots without determining a landmark setting. We improved the point matching algorithm, which used Delaunay triangulation to determine spot regions to find spots pair and used affine transformation to match spot pairs. From the experimental results, our local point matching algorithm is able to be used as a basic step for the global matching problem for gel images and as a landmark setting. For future research, this approach should be conducted to examine under noisy or distorted condition.

Acknowledgements

This work was supported by the Korea Research Foundation. Grant funded by the Korean Government (MOEHRD), (KRF-2006-311-D00833).

References


