Dealing with DGGE images by using ant colony systems and fuzzy logic

M. Angélica Pinninghoff J.
Computer Science Dept. University of Concepción
Concepción, Chile
mpinning@udec.cl

Ricardo Contreras A.
Computer Science Dept. University of Concepción
Concepción, Chile
rcontrer@udec.cl

Felipe Garrido S.
Computer Science Dept. University of Concepción
Concepción, Chile
fgarridos@udec.cl

***Abstract*—This work introduces a hybrid approach for dealing with the problem of detecting edges in DGGE images. It is important to have an efficient mechanism for precisely detecting band edges, because these images are typically used for the identification of microorganisms in ADN samples. The hybrid approach combines ant colony systems and fuzzy logic. Results obtained with this approach are better than those obtained with other mechanisms.**

Keywords—DGGE images, ant colony systems, fuzzy logic

I. INTRODUCTION

Digital images are two-dimensional representations of images; a set of pixels arranged in a matrix, in order to allow an electronic manipulation. In a gray-scale, each pixel is represented by using a numeric value belonging to the interval [0,255] that illustrates different shades of gray, varying from black at the weakest intensity to white at the strongest.

On the other way, segmentation [6] [12], is an operation which divides an image into parts or regions having a particular feature. Segmentation algorithms are based on two intensity-based properties; the first one is the similarity among pixels, while the second one focuses on continuity related to gray levels, the last approach involves a family of algorithms for detecting edges by considering strong intensity changes among neighbor pixels. In recent years, Ant Colony Optimization (ACO) algorithms have been developed to detect image edges, by taking inspiration from the behavior of ants [17].

Some ant families have the capability of finding the shortest path between their nest and the source of food. Ants use the environment as a medium for communicating. They exchange information indirectly by depositing pheromone, while they pass through a particular trail (or path). The information exchanged has a local scope; only an ant located where the pheromones were deposited has a notion of them. This system property is called *stigmergy* and occurs in many social animal societies (it has been studied in the case of the construction of pillars in the nests of termites). The mechanism to solve a problem too complex to be addressed by single ants is a good example of a self-organized system. This system is based on positive feedback (the deposit of pheromone attracts other ants that will strengthen it themselves) and negative feedback (dissipation of the route by evaporation prevents the system from thrashing). Theoretically, if the quantity of pheromone remained the same over time on all edges, no route would be chosen. However, because of feedback, a slight variation on an edge will be amplified allowing thus the choice of an edge.

The algorithm will move from an unstable state, in which no edge is stronger than another, to a stable state where the route is composed of the strongest edges.

The basic philosophy of the algorithm involves the movement of a colony of ants through the different states of the problem influenced by two local decision policies, viz., trails and attractiveness. Thereby, each such ant incrementally constructs a solution to the problem. When an ant completes a solution, or during the construction phase, the ant evaluates the solution and modifies the trail value of the components used in its solution. The above described behavior is the inspiration source for using artificial ants [4], aimed to solve optimization problems.

The idea of using artificial ants to solve hard problems has been developed by different authors. In [7] ant colony optimization is used in the k-means algorithm for improving the image segmentation. The learning mechanism of this algorithm is formulated by using the ACO meta-heuristic.

In [10] the author proposes an ACO algorithm hybridized with 2-OPT for fractal image compression. In [9] ACO algorithms are used in image segmentation, improving thresholding algorithms. Thresholding algorithms are the focus in [9]; authors obtain experimental results to demonstrate that the proposed ant-based method performs better than other two established thresholding algorithms. In [1] authors use an ACO algorithm for image edge detection. Edge detection is accomplished by seeking pixels that show important differences with respect to their neighbors, in terms of intensity level (in gray-scale). A solution to a similar problem is described in [5], where swarm intelligence is applied in image processing for feature extraction. Edge extraction and image segmentation are implemented by using a graph to represent the relationship between adjacent image points. Experiments show that it is possible to use this method to effectively perform feature extraction in digital images.

The work described in [13] presents an approach that obtains interesting results. They utilize a number of ants moving on a 2-D image for constructing a pheromone matrix, each entry represents the amount of pheromone at each pixel location of the image. The movements of the ants are steered by the local variation of the image's intensity values.

In [8] authors propose an ant colony optimization based algorithm for continuous optimization problems on images like image edge detection, image compression, image segmentation and structural damage monitoring in image processing. They show the feasibility of the algorithm in terms of accuracy and continuous optimization. This work emphasizes an important feature: a good solution, like the shortest path, has more pheromone than the longest paths.

Finally, the work in [17] shows an approach to image edge detection that considers the values of gradient and relative difference of statistical means extracted for the ants' searching. Authors show results that validate the proposal, in terms of algorithm performance.

Fuzzy logic is a form of many-valued logic; it deals with reasoning that is approximate rather than fixed and exact. Compared to traditional binary sets, where variables may take on true or false values, fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely truth and completely false [11].

The term fuzzy logic was introduced in 1965 in the proposal of fuzzy set theory by Zadeh [15] [16]. Fuzzy logic has been applied to many fields, from control theory to artificial intelligence.

Classical logic only permits propositions having a value of truth or falsity. The notion of whether 1+1=2 is an absolute, immutable, mathematical truth. However, there exist certain propositions with variable answers, such as asking various people to identify a color.

This work is based on the experience introduced by Tyagi [14]. In this work fuzzy logic is used to decide ant's movement. While the work of Tyagi is focused on detecting edges in a set of photographs usually used to illustrate the process, our work focuses on DGGE images, because of the critical issues involved on these particular images, and because of the poor results obtained with other methods.

The objective of this work is to evaluate the effectiveness of a hybrid ACO algorithm and fuzzy logic, in edge detection. In particular, this article focuses in DGGE images, gray-scale images used in biology and chemistry.

This proposal detects edges in bands by using ant-based algorithms and fuzzy logic. Once the edges have been detected, it is possible to measure the area they comprise and to measure the concentration based on their molecular weight (pb), and therefore, the amount of different species present on each sample.

This article is structured as follows; the first section is made up of the present introduction; the second section describes the DGGE images; the third section describes the ACO algorithm and fuzzy logic we used. The fourth section shows the results we obtained, and the final section shows the conclusions of the work.

II. DGGE IMAGES

Denaturing Gradient Gel Electrophoresis (DGGE) is a DNA-based technique which generates a genetic profile or *fingerprint* which can be used to identify the dominant members of the microbial community. DGGE has been used to investigate microbial responses in a wide variety of applications, including bioremediation assessment, wastewater treatment, drinking water treatment, biofilm formation, microbial induced corrosion, among others.

DGGE separates mixtures of amplified 16SrRNA gene segments, which are all the same size, based on nucleotide sequence. Denaturing breaks apart the two strands of the DNA molecule. Gradient Gel, is a gel with an increasing concentration of a chemical (denaturant), which breaks apart the DNA molecule. Electrophoresis is the application of an electric current across a gel. In response to the current, double-stranded DNA migrates (moves down) the gel. Denaturing the DNA molecule forms Y and T-shaped structures greatly slowing migration. Finally, this process allows obtaining, as a result, an image composed of bands and lanes [3].

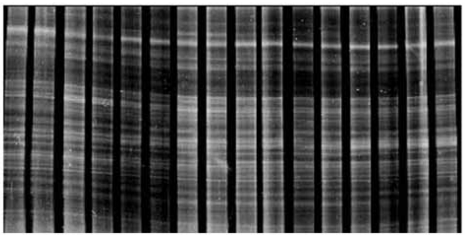

Figure 1. A sample DGGE image with two reference lanes.

Lanes are the vertical columns shown in Figure 1, and each one of them represents a DNA sample, except the reference lanes which are the leftmost and the rightmost lanes. Reference lanes are used to indicate the molecular weight, measured in base pairs (*bp*), of the DNA. The bands are the horizontal lines in each lane that represent the segments agglomeration of a DNA sample with the same *bp* value.

By using DGGE images, it is possible to detect the similarity or difference among individuals, and in doing so; it is necessary to know the location of a band in a lane.

A common problem in dealing with DGGE images is the lack of accuracy in band detection. Currently, this process is accomplished by using Quantity One [2], but it presents an important error rate that forces to correct it manually, which implies a tedious work. This proposal suggests automatic band detection with a combination of fuzzy logic and ACO algorithms.

III. ACO ALGORITHM AND FUZZY LOGIC

In ACO algorithms, artificial ants walk through a space of solutions represented as a graph *G=(V, E),* where *V* is the set of nodes and *E* is the set of edges or connections between nodes. Edges of the graph are the places where ants deposit pheromone. For implementing this model, the space of solutions is represented by a matrix of pixels of an image. Each pixel is a matrix entry; an edge in the graph represents a neighborhood between two pixels.

In ACO algorithms, an arbitrary number of ants are randomly distributed on the matrix pixels. When a particular ant walks through the image, it deposits pheromone. The amount of pheromone depends on the contrast among neighbor pixels. High contrast implies more pheromone. This is reflected in a matrix, by increasing the amount of pheromone in those pixels that present higher intensity than the previous pixel in the ant path. A higher contrast implies a higher amount of pheromone on the new pixel.

In this scheme, a solution is a configuration that results from the fact that every pixel in the original image has been traversed and different regions in the image present different gray-scale intensities, due to different pheromone deposits. When a solution is obtained, it is followed (in practice, after a time interval) by a process in which the pheromone is diminished, representing the effect of the (evaporation) time on the pheromone that was deposited in the different paths. This process, besides to model closely a real phenomena occurring with ants, allows avoiding the effect of local minimum.

In this work, [14] the number of ants is set to $K = \sqrt{(MxN)}$ (M is the number of rows and N is number of columns in the image). These ants are randomly assigned on the image with maximum one ant on each pixel. At each step one ant is randomly selected and it will move on the image for T path steps. The probability $P_{i,j}$ of movement of ant k from pixel i to pixel j is given by equation (1)

$$P_{i,j} = [(\tau_{i,j})^{\alpha}(\eta_{i,j})^{\beta}] / [\sum_{j\in L}(\tau_{i,j})^{\alpha}(\eta_{i,j})^{\beta}] \qquad (1)$$

Where $j \in L$ indicates all pixels in the neighborhood of pixel i. $\tau_{i,j}$ denotes the pheromone and $\eta_{i,j}$ denotes the heuristic information. Edgeness computed by fuzzy logic, represents the heuristic information. α and β are not constants, they are modified during the process as in Tyagi's work, they indicate the influence of pheromone and heuristic information. There are two instances for updating pheromone, local and global.

The local update is performed after an ant movement (from pixel i to pixel j); see equations (2) and (3).

$$\tau_{i,j} = (1-\rho)\tau_{i,j} + \rho\Delta\tau_{i,j} \qquad (2)$$

$$\Delta\tau_{i,j} = \eta_{i,j} \qquad (3)$$

Where $0<\rho<1$ indicates the pheromone evaporation rate, $\tau_{i,j}$ is the current value of pheromone in that pixel, and $\Delta\tau_{i,j}$ is the amount of pheromone to be increased. The global update is performed after the movements of all ants within each cycle are completed. This is described in equation (4).

$$\tau_{i,j} = (1-\psi)\tau_{i,j} + \psi\tau_0 \qquad (4)$$

ψ is the pheromone decay coefficient and τ_0 is the initial value of pheromone. The above process is performed C times or cycles.

In a fuzzy inference system, the linguistic variable is an important concept. In a few words, a linguistic variable is a variable that accept as values words or phrases in natural language, that allow to describe a phenomena that is too complex or it is not well-defined. The key idea in such a system is to introduce human knowledge through a set of fuzzy rules if-then involving operations on linguistic variables.

In this work, values for the linguistic variables are High, Medium and Low, computed by using a Gaussian function shown in equation (5). These values are labels intended to represent the intensity of a pixel or a set of pixels. The values of intensity are used to determine the edgeness of a particular pixel. Edgeness is defined as the degree to which a pixel in a location (i,j) is believed to be an edge when only local intensity variation around the pixel (i,j) is taken into account.

$$\mu(x) = e\hat{}[-(x-c)^2/2\sigma^2] \qquad (5)$$

Where c is the center of the function representing the membership degree and σ represents the standard deviation. The symbol ^ is used to denote that it is an exponential function.

The procedure considers four stages: i) Computation of intensities, ii) Fuzzy logic process, iii) ACO, and iv) Binarization. In i), three new matrices are created, to store the difference of intensities between each particular pixel and its neighborhood. One matrix considers differences by computing only columns associated to the pixel, a second one considers only rows associated to the pixel, and the third matrix takes into account the complete neighborhood. For every computation, it is used the Moore neighborhood (each pixel has 8 neighbors). In ii) the three matrices act as the input to create a new matrix; the edgeness matrix. The corresponding values for this new matrix are computed following a set of three steps: to obtain a fuzzy value, to apply a set of rules, and to defuzzify the values after the rules; the computation for each pixel is realized by using Mamdami type fuzzy rule based system. In iii) the ACO algorithm is triggered, and the heuristic that guides the process of choosing a particular movement is the value of the edgeness matrix for each entry. This value is the term ($\eta_{i,j}$) in equation (1). Finally, in iv), a binary decision is made at each pixel to determine whether it is an edge or not by applying a threshold on the final pheromone matrix.

This work considered a set of six rules to obtain the edgeness matrix [14]. A typical rule is as follows:

Rule1: **If** (Mrow is High) **and** (Mcol is High) **and** (Idiag is High), **then** Edgeness is High.

Mrow, Mcol and Idiag are the matrices computed in point i) of the procedure explained above.

IV. RESULTS

Parameters are experimentally obtained, and the best values coincide with the parameter used in [14]. The final values are shown in Table I:

TABLE I. PARAMETERS

Parameter	Value
T	35
C	10
α	5
β	0.3
ψ	0.06
ρ	0.1

Where T is the number of steps for an ant; C is the number of cycles; α and β are the coefficients which control the influence of pheromone and heuristic, respectively, in the decision of ant movement; ψ and ρ are, respectively, the global and local coefficient of evaporation.

To validate the technique, it was used an artificial image, a square contained into another square, having known edges. Equation (5) represents the way of computing the degree of successful edge detection.

$$P = S / (S + FP + FN) * 100 \quad (5)$$

S (success) is the number of pixels that the algorithm identifies as edges that are really edges. FP (false positive) is the number of pixels that the algorithm identifies as edges, but that are not really edges, and FN (false negative) that is the number of pixels the algorithm detects as no edges pixels, but that really are.

Figure 2 shows the test image and the edge suggested by the algorithm. Table II shows the best values obtained after evaluating the mechanism.

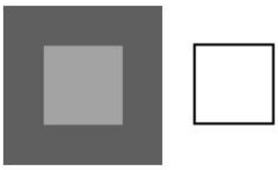

Figure 2. Test image and edge with threshold =250

TABLE II. EVALUATION OF IMAGE IN FIGURE 2

S	516
FP	0
FN	0
P	100%

Edges are well defined, obtaining a complete success in tests. In numeric terms, it is obtained a degree of 100% in edge detection. The same image after using ACS (Ant Colony System) obtains only a 54% of successful detection.

This work represents a first approach for dealing with DGGE images by using fuzzy logic concepts. DGGE images are important in biotechnology, and completely successful methods for detecting edges are not known.

After visual comparison of results on real DGGE images and supported by the tests on the artificial image (the two squares), it can be considered a promising method to insist in analyzing complementary mechanisms, as preprocessing the images, for instance. A very important issue to be taken into account is that the threshold value deserves particular attention; slightly different values lead to different results. It is not clear when a particular difference in threshold values produces a particular difference in the quality of results.

To emphasize the promising quality of results, we present some results obtained with our proposal (Figure 3b, Figure 3c, Figure 4b, Figure 4c, Figure 5b, Figure 5c), and results obtained on the same set of images, with Canny algorithm (Figure 4d, Figure 5d).

The following sets of images show some selected results. We show here results working on three different images that have a different number of lanes that represent typical DGGE images having different kind of problems as noise and distortion.

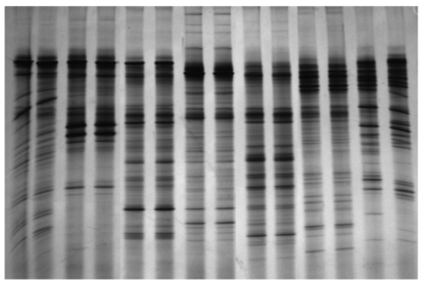

Figure 3a. Original image (Image-1)

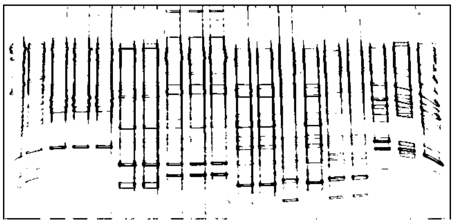

Figure 3b. Result of Image-1 with threshold = 247

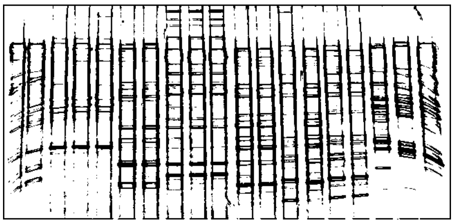

Figure 3c. Result of Image-1 with threshold = 250

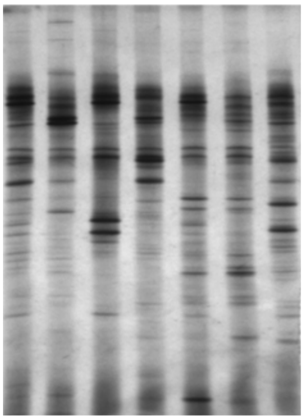

Figure 4a. Original image (Image-2)

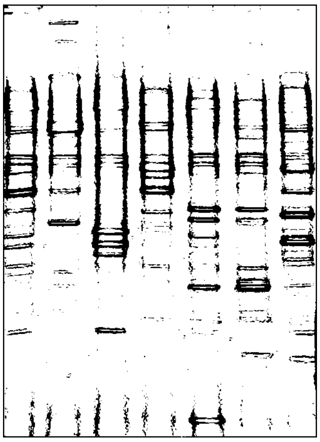

Figure 4b. Result of Image-2 with threshold = 252

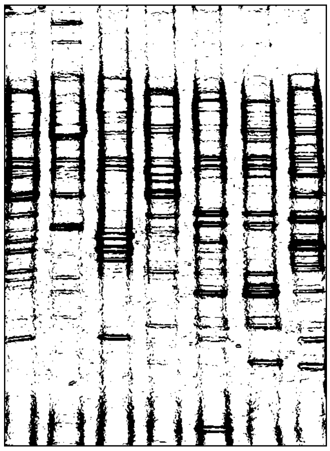

Figure 4c. Result of Image-2 with threshold = 253

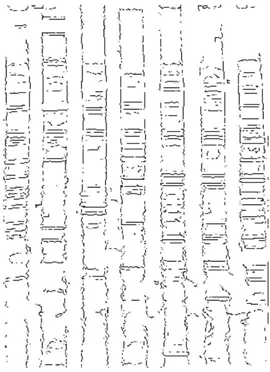

Figure 4d. Result of Image-2 with Canny

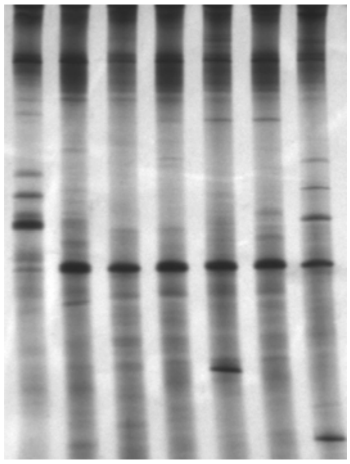

Figure 5a. Original image (Image-3)

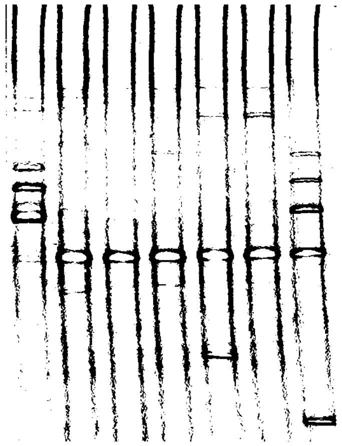

Figure 5b. Result of Image-3 with threshold = 252

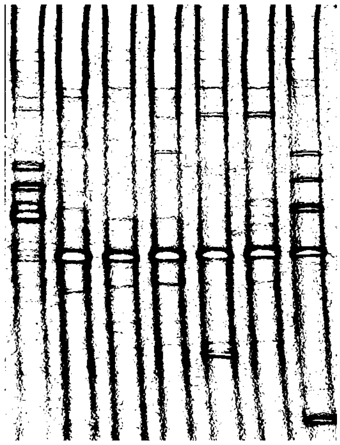

Figure 5c. Result of Image-3 with threshold = 253

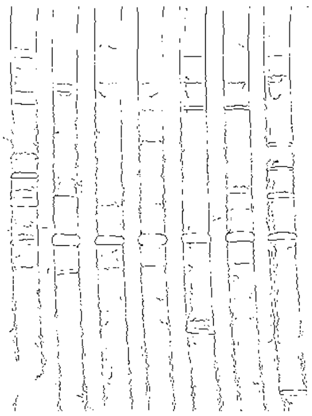

Figure 5d. Result of Image-3 with Canny

V. CONCLUSIONS

This work describes a hybrid approach that combines ACO and fuzzy logic. Images, after the process, exhibit some features that improve previous results when images are processed with different ACO implementations.

In particular, in this approach, the resulting edges are thinner, a very important issue, when we take into account that the main feature a DGGE image presents is blurring.

This is a first step on a hybrid model that still has to be tuned in terms of parameters.

REFERENCES

[1] A. Baterina, C. Oppus. Image Edge Detection Using Ant Colony Optimization. International Journal of Circuits, Systems and Signal Processing, Issue 2, vol. 10. 2010.

[2] Bio Rad Laboratories. Quantity One User's Guide. Bio-Rad Laboratories,. 2000.

[3] J. Bruggeman, J. Stephen, Y. Chang. Competitive PCR-DGGE Analysis of bacteria mixtures an internal standard and appraisal of template enumeration accuracy. Elsevier. 2000.

[4] M. Dorigo, T Stuetzle. Ant Colony Optimization. The MIT Press. 2004

[5] J. Du, L. Zhou, X.i Xu, L. Yang, Z. Shi. Image Feature Extraction with the Artificial Ant Colony. International Conference on Intelligent Computation Technology and Automation. May 2010.

[6] R. Gonzalez R. Woods. Digital Image Processing. 3rd Ed. Upper Saddle River, NJ: Pearson/Prentice Hall. 2008.

[7] C. Hung, M. Sun. Ant colony optimization for the k-means algorithm in image segmentation. Proceedings of the 48th Annual Southeast Regional Conference (ACM SE'10), pp. 1-4, 2010.

[8] S. Kaur, P. Agarwal, R. S. Rana. Ant Colony Optimization: A Technique used for Image Processing. International Journal of Computer Science and Technology, Vol.2, Issue 2, June 2011.

[9] A. Malisia, H. Tizhoosh. Image Thresholding Using Ant Colony Optimization. The 3rd Canadian Conference on Computer and Robot Vision (CRV 06). 2006.

[10] C. Martinez. An ACO algorithm for image compression. CLEI Electronical Journal, vol. 9, num. 2 paper 1. 2006.

[11] V. Novák, I . Perfilieva and J. Mockor. Mathematical principles of fuzzy logic. Dodrecht: Kluwer Academic. 1999.

[12] D. Phillips. Image Processing in C. 2nd Ed., R & D Publications. 1994.

[13] J. Tian, W. Yu, S. Xie. An ant colony optimization algorithm for image edge detection. IEEE Congress on Evolutionary Computation. 2008.

[14] Y. Tyagi, T. Puntambekar, P. Sexena, S. Tanwani. A hybrid approach to Edge Detection using Ant Colony Optimization and Fuzzy Logic. International Journal of Hybrid Information Technology. Vol 5 No 1. January 2012.

[15] L. Zadeh. Fuzzy Logic. Stanford Encyclopedia of Philosophy. Stanford University. 2006-07-23. Retrieved 2008-09-30.

[16] L. Zadeh. Fuzzy sets. Information and Control 8(3): 338-353. 1965.

[17] J. Zhang, K. He, X. Zheng, J. Zhou. An Ant Colony Optimization Algorithm for Image Edge Detection. International Conference on Artificial Intelligence and Computational Intelligence, pp. 215-219, October 2010.