

Measuring the Quality of Shifting and Scaling Patterns in Biclusters

Beatriz Pontes¹, Raúl Giráldez², and Jesús S. Aguilar-Ruiz²

¹ Department of Computer Science, University of Seville
Avenida Reina Mercedes s/n, 41012, Sevilla, Spain
bepontes@us.es

<http://www.lsi.us.es/~bepontes/index-en.html>

² School of Engineering, Pablo de Olavide University
Ctra. de Utrera, km.1, 41013, Sevilla, Spain
{giraldez,aguilar}@upo.es

<http://www.upo.es/eps/{giraldez,aguilar}>

Abstract. The most widespread biclustering algorithms use the Mean Squared Residue (MSR) as measure for assessing the quality of biclusters. MSR can identify correctly shifting patterns, but fails at discovering biclusters presenting scaling patterns. Virtual Error (VE) is a measure which improves the performance of MSR in this sense, since it is effective at recognizing biclusters containing shifting patterns or scaling patterns as quality biclusters. However, VE presents some drawbacks when the biclusters present both kind of patterns simultaneously. In this paper, we propose an improvement of VE that can be integrated in any heuristic to discover biclusters with shifting and scaling patterns simultaneously.

1 Introduction

The use of microarray techniques allows to study the activity of thousands of genes at a time, producing in this way a huge amount of data. Usually, the resulting data is organized in a matrix, called an expression matrix, where columns may represent genes and rows represent experimental conditions. An element of such expression matrix stands for the expression level of a given gene under a specific condition [3,18].

The interest in discovering knowledge from gene expression data has experienced an enormous increase with the development of microarray techniques. Biclustering [12] is becoming a popular data mining technique due to its ability to explore at the same time both dimensions of data, as opposed to clustering techniques [19], that can only use one dimension. In this sense, microarray is a suitable context for the application of biclustering techniques, since they can consider both genes and experimental conditions at extracting useful knowledge. Thus, in this context, a bicluster is a subset of genes under a subset of conditions. In particular, those biclusters where the subset of genes shows a common

tendency under the subset of conditions are of special interest. In general, biclustering is much more complex than clustering [14]. In fact, finding significant biclusters in microarray data has been proven to be a NP-hard problem [17].

Cheng and Church [7] were the first in applying biclustering to gene expression data. They introduced one of the most popular biclustering algorithms that combines a greedy search heuristic for finding biclusters with a measure for assessing the quality of such biclusters. This measure, named *Mean Squared Residue* (MSR), has been used by many researchers who have proposed different heuristics for biclustering biological data. Aguilar et al. [2] developed an approach based on local nearness. Yang et al. [21] proposed an iterative algorithm for finding a predefined number of biclusters. Cano et al. [6] based their proposal on fuzzy technology and spectral clustering. Other approaches, such as Divina and Aguilar [10] and Bleuler et al. [4], have been based on evolutionary computation, while Bryan et al. [5] applied simulated annealing as heuristic. Recently, MSR has also been incorporated as cost function in multiobjective heuristics based on Particle Swarm Optimization [13] and Artificial Immune Systems [9].

Although MSR has been used in many proposals for finding biclusters, it nevertheless has been proven to be inefficient for finding certain types of biclusters in microarray data, especially when they present strong scaling tendencies [1]. Thus, we introduced in previous works an alternative measure named *Virtual Error* (VE) [15]. This measure is based on the concept of behavioural patterns, which aim at identifying common patterns between genes or conditions. VE is effective at recognizing biclusters containing shifting patterns or scaling patterns as quality biclusters. However, it presents some drawbacks when both kind of patterns are presented simultaneously in the same bicluster. In this paper, we propose a novel variant of VE, called *Transposed Virtual Error* (VE^t), that allows to find biclusters that MSR and VE do not recognize as interesting ones.

This paper is organized as follows. In the next section, an description of the shifting and scaling patterns is given. We then provide a formal definition of VE^t in Section 3, followed by a formal analysis in Section 4, demonstrating its strength with regard to the behavioural patterns. In Section 5 we discuss the consequences of the theorems presented in this work, providing a test of the effect of the noise on the VE^t . Finally, we summarize the main conclusions in Section 6.

2 Behavioural Patterns in Gene Expression Data

When all the genes of a bicluster follow a similar tendency under the set of conditions, then such a bicluster may be potentially biologically interesting. Therefore, it seems to be a good idea to develop a quality measure for biclusters based on the idea of behavioural patterns for gene expression. Aguilar [1] presented an in-depth discussion on the possible patterns in gene expression data. He described formally two kind of patterns: shifting and scaling patterns. They have

been defined using numerical relations among the values in a bicluster. Several works based their principle in the pattern concept in order to mine the data. Xu et al [20] propose a biclustering algorithm for mining shifting and scaling co-regulation patterns on gene expression data. Nevertheless, they do not provide a quality measure, but use a model-based heuristic instead. Furthermore, they are only able to identify global shifting and scaling patterns, while local ones seem to be more interesting since they depict the general situation [1]

Let \mathcal{B} be a bicluster made up of I experimental conditions and J genes. Each element in \mathcal{B} is represented by $b_{ij} \in \mathcal{B}$. This way, the bicluster \mathcal{B} follows a *perfect shifting pattern* if its values can be obtained by adding a constant-condition number β_i to a typical value for each gene (π_j). β_i is said to be the *shifting coefficient* for condition i . In this case, the expression values in the bicluster fulfil the following equation:

$$b_{ij} = \pi_j + \beta_i \quad (1)$$

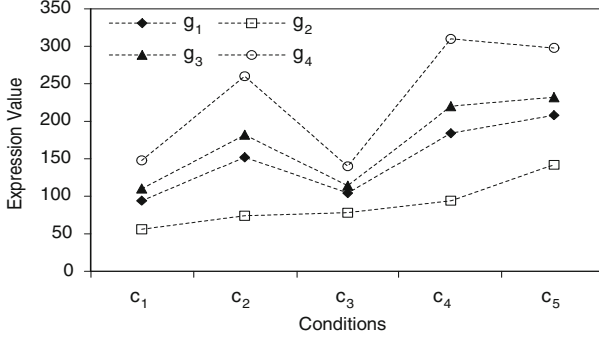
Similarly, a bicluster follows a *perfect scaling pattern* changing the additive value in the former equation by a multiplicative one. This new term α_i is called the *scaling coefficient*, and represents a constant value for each condition. The following equation defines whether a bicluster follows a perfect scaling pattern or not:

$$b_{ij} = \pi_j \times \alpha_i \quad (2)$$

Shifting and scaling patterns may be put together in a new kind of pattern called *combined pattern*. In fact, it is the most probable situation when working with real genetic data. In this situation, the expression values can be obtained using both coefficients, shifting and scaling coefficients. The equation that must be fulfilled by the values in this case can be represented by merging 1 and 2:

$$b_{ij} = \pi_j \times \alpha_i + \beta_i \quad (3)$$

Figure 1 shows an example of a bicluster obtaining from synthetic data. This is a typical visualization of bicluster, where conditions are represented in the x-axis, the values of gene expression are represented in the y-axis and each line is a gene. As we can see, there are four genes g_j (with $1 \leq j \leq 4$) and five conditions c_i (with $1 \leq i \leq 5$). This bicluster contains both shifting and scaling patterns. The matrices below describe the factor decomposition of the numerical values. Having a look at figure 1 we could say that genes g_1 , g_3 and g_4 present a similar behaviour across the conditions, although g_4 has a different tendency between the last two conditions. On the contrary, the tendency of gene g_2 varies from the other genes, since its behaviour is always increasing across all the conditions. Gene g_2 is difficult to be associated to a perfect pattern visually because its shifting coefficients β_i are closed to the product $\pi_j \times \alpha_i$. Therefore, identifying biclusters with both shifting and scaling patterns might be a difficult task due to the complexity inherent in equation 3.



$$\mathcal{B} = \begin{pmatrix} 95 & 56 & 110 & 149 \\ 152 & 74 & 182 & 260 \\ 104 & 78 & 114 & 140 \\ 185 & 94 & 220 & 311 \\ 208 & 143 & 233 & 298 \end{pmatrix} = \begin{pmatrix} 25 \times 3 + 20 & 12 \times 3 + 20 & 30 \times 3 + 20 & 43 \times 3 + 20 \\ 25 \times 6 + 2 & 12 \times 6 + 2 & 30 \times 6 + 2 & 43 \times 6 + 2 \\ 25 \times 2 + 54 & 12 \times 2 + 54 & 30 \times 2 + 54 & 43 \times 2 + 54 \\ 25 \times 7 + 10 & 12 \times 7 + 10 & 30 \times 7 + 10 & 43 \times 7 + 10 \\ 25 \times 5 + 83 & 12 \times 5 + 83 & 30 \times 5 + 83 & 43 \times 5 + 83 \end{pmatrix}$$

$$\{\pi_j\} = \{25 \ 12 \ 30 \ 43\}$$

$$\{\alpha_i\} = \begin{matrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 \\ \{3 & 6 & 2 & 7 & 5\} \end{matrix}$$

$$\{\beta_i\} = \{20 \ 2 \ 54 \ 10 \ 83\}$$

Fig. 1. Bicluster containing perfect shifting and scaling patterns

3 Transposed Virtual Error

Virtual Error (VE) had been designed as an evaluation measure for biclusters [15]. VE is based on the concepts of shifting and scaling patterns and is capable of identifying both kind of patterns within biclusters, although not simultaneously. This way, VE improves MSR effectiveness, since the last one can only recognize shifting patterns. The basic idea behind VE is to measure the extent to which the genes in a bicluster are similar to the general tendency. The general tendency is represented by a *Virtual Gene* which is created taking into account the values for every gene across the experimental conditions, but trying to capture the general behaviour with independence of the concrete numerical values. VE will have a lower value for those biclusters in which its genes are closer to the virtual gene. This is due to the fact that VE computes the numerical differences between each standardized gene and the standardized virtual gene. Therefore, the better a bicluster is, the lower its VE value will be. Furthermore, it is obvious that VE will always be greater or equal than zero.

VE has been used in various evolutionary algorithms in order to find one hundred biclusters in several gene expression matrices [15,11]. These two previous works have allowed us finding interesting biclusters that could not have been obtained using MSR alone. Furthermore, the VE value for biclusters with perfect shifting and scaling patterns seems to be very close to zero [16] (magnitude of 10^{-15}). Nevertheless, VE cannot be proven to recognize both kind of patterns simultaneously.

In this work, we present an enhanced version of VE, named VE^t , from *Transposed Virtual Error*. We analytically prove that VE^t is zero for those biclusters with perfect shifting and scaling patterns. This variation of VE has been motivated by [8], where several numeric transformations have been applied to the data in order to detect both kind of patterns.

VE^t is computed similarly to VE but considering the transposed bicluster. The idea here would be to create a *Virtual Condition*, instead of a virtual gene, and measure the differences between the standardized values for every condition and the standardized virtual condition. In the following, we explain how to create the virtual condition for a certain bicluster \mathcal{B} , in order to compute VE^t afterwards.

Definition 1 (Virtual Condition). *Given a bicluster \mathcal{B} with I conditions and J genes, we define its virtual condition as a collection of J elements ρ_j , each of them defined as the mean of the j^{th} column: $\rho_j = \frac{\sum_{i \in I} b_{ij}}{I}$, where $b_{ij} \in \mathcal{B}$, $1 \leq i \leq I$ and $1 \leq j \leq J$.*

This way, each element of the virtual condition represents a meaningful value for all the conditions, regarding each gene. Once the virtual condition has been created, the next task would consist of quantifying the way in which all the experimental conditions in the bicluster are similar to it. In order to perform an appropriate comparison, we first carry out a standardization of the virtual condition and of every experimental condition in the bicluster. This standardization allows us to capture the differences among the tendencies, with independence of the numerical values.

Definition 2 (Standardization). *We define the standardized bicluster $\hat{\mathcal{B}}$ from bicluster \mathcal{B} as a new bicluster in which its elements \hat{b}_{ij} are defined by $\hat{b}_{ij} = \frac{b_{ij} - \mu_{c_i}}{\sigma_{c_i}}$, where σ_{c_i} and μ_{c_i} represent the standard deviation and the arithmetic average of all the expression values for condition i , respectively.*

It has already been said that the virtual condition needs also to be standardized. Equation 4 shows how the values of the standardized virtual condition are obtained, where ρ_j refers to the virtual condition value for gene j , while μ_ρ and σ_ρ refer to the average and the deviation of the values of the virtual condition, respectively.

$$\hat{\rho}_j = \frac{\rho_j - \mu_\rho}{\sigma_\rho} \quad (4)$$

Definition 3 (Transposed Virtual Error). *Given a bicluster \mathcal{B} , and the virtual condition ρ , Transposed Virtual Error (VE^t) can be defined as the mean of the numerical differences between each standardized condition and the values of the standardized virtual condition for each gene:*

$$VE^t(\mathcal{B}) = \frac{1}{I \cdot J} \sum_{i=1}^{i=I} \sum_{j=1}^{j=J} (\hat{b}_{ij} - \hat{\rho}_j) \quad (5)$$

Next, we present three theorems and their proofs that demonstrate the strength of VE^t with regard to the shifting and scaling patterns.

4 Analysis

This section includes formal proofs that bear out the hypothesis that VE^t is zero for those biclusters with perfect shifting and scaling patterns, either separately or simultaneously.

Theorem 1. *A bicluster presenting a perfect shifting pattern has VE^t equal to zero.*

Proof. Let \mathcal{B} be a bicluster with a perfect shifting pattern, then it is possible to refer to its elements as $b_{ij} = \pi_j + \beta_i$. Applying two simple arithmetic properties¹, the mean and the deviation for each condition c_i can be expressed by:

$$\mu_{c_i} = \mu_\pi + \beta_i \quad ; \quad \sigma_{c_i} = \sigma_\pi$$

where μ_π and σ_π represent the mean and the deviation of the π values, respectively. Using these results we obtain the standardized values for b_{ij} :

$$\hat{b}_{ij} = \frac{b_{ij} - \mu_{c_i}}{\sigma_{c_i}} = \frac{\pi_j + \beta_i - \mu_\pi - \beta_i}{\sigma_\pi} = \frac{\pi_j - \mu_\pi}{\sigma_\pi}$$

Combining the former properties¹ it is easy to express the mean and standard deviation for the virtual condition as:

$$\mu_\rho = \mu_\pi + \mu_\beta \quad ; \quad \sigma_\rho = \sigma_\pi$$

Finally, the standardized values for the virtual condition are the following:

$$\hat{\rho}_j = \frac{\rho_j - \mu_\rho}{\sigma_\rho} = \frac{\pi_j + \mu_\beta - \mu_\pi - \mu_\beta}{\sigma_\pi} = \frac{\pi_j - \mu_\pi}{\sigma_\pi} = \hat{b}_{ij}$$

As it can be seen above, the standardized virtual condition is equal to all the real conditions after being standardized. Therefore, VE^t has been proven to be zero for those biclusters with perfect shifting patterns. ■

Theorem 2. *A bicluster presenting a perfect scaling pattern has VE^t equal to zero.*

Proof. Let \mathcal{B} be a bicluster following a perfect scaling pattern, then its elements can be expressed by $b_{ij} = \pi_j \times \alpha_i$. Following the same reasoning that in the former proof, the mean and deviation of each condition c_i are:

$$\mu_{c_i} = \alpha_i \times \mu_\pi \quad ; \quad \sigma_{c_i} = \alpha_i \times \sigma_\pi$$

From these results we obtain the standardized values for b_{ij} :

$$\hat{b}_{ij} = \frac{b_{ij} - \mu_{c_i}}{\sigma_{c_i}} = \frac{\pi_j \times \alpha_i - \alpha_i \times \mu_\pi}{\alpha_i \times \sigma_\pi} = \frac{\pi_j - \mu_\pi}{\sigma_\pi}$$

Next we obtain the mean and deviation for the values of the virtual condition:

$$\mu_\rho = \mu_\pi \times \mu_\alpha \quad ; \quad \sigma_\rho = \mu_\alpha \times \sigma_\pi$$

¹ Being $f(x) = g(x) \times c_1 + c_2$, the properties related to the arithmetic mean ($\mu_{f(x)}$) and the standard deviation ($\sigma_{f(x)}$) of $f(x)$ are the following: $\mu_{f(x)} = \mu_{g(x)} \times c_1 + c_2$ and $\sigma_{f(x)} = \sigma_{g(x)} \times c_1$.

And finally the standardized values for the virtual condition are:

$$\hat{\rho}_j = \frac{\rho_j - \mu_\rho}{\sigma_\rho} = \frac{\pi_j \times \mu_\alpha - \mu_\pi \times \mu_\alpha}{\mu_\alpha \times \sigma_\pi} = \frac{\pi_j - \mu_\pi}{\sigma_\pi} = \hat{b}_{ij}$$

As in the previous proof, we obtain that the standardized values for the virtual condition are equal to de standardized values for all the real experimental conditions. As a consequence, VE^t will be zero for every bicluster with a perfect scaling pattern. ■

Theorem 3. *A bicluster presenting a perfect combined pattern (shifting and scaling) has VE^t equal to zero.*

Proof. If \mathcal{B} contains a perfect combined pattern, its values can be represented by $b_{ij} = \pi_j \times \alpha_i + \beta_i$. Using the same arithmetic properties as in the former proves, the mean and deviation for each condition c_i are:

$$\mu_{c_i} = \alpha_i \times \mu_\pi + \beta_i \quad ; \quad \sigma_{c_i} = \alpha_i \times \sigma_\pi$$

And the standardized values for b_{ij} can be expressed as:

$$\hat{b}_{ij} = \frac{b_{ij} - \mu_{c_i}}{\sigma_{c_i}} = \frac{\pi_j \times \alpha_i + \beta_i - \alpha_i \times \mu_\pi + \beta_i}{\alpha_i \times \sigma_\pi} = \frac{\pi_j - \mu_\pi}{\sigma_\pi}$$

The mean and deviation for the virtual condition are the following:

$$\mu_\rho = \mu_\pi \times \mu_\alpha + \mu_\beta \quad ; \quad \sigma_\rho = \mu_\alpha \times \sigma_\pi$$

And the standardized values for the virtual condition:

$$\hat{\rho}_j = \frac{\rho_j - \mu_\rho}{\sigma_\rho} = \frac{\pi_j \times \mu_\alpha + \mu_\beta - \mu_\pi \times \mu_\alpha - \mu_\beta}{\mu_\alpha \times \sigma_\pi} = \frac{\pi_j - \mu_\pi}{\sigma_\pi} = \hat{b}_{ij}$$

Again, the standardized values for the virtual condition match up with the standardized values for the original conditions. Therefore, VE^t will also be zero for those biclusters following a perfect shifting and scaling pattern. ■

These results confirm that VE^t is the first measure up to the date capable of recognizing combined patterns in gene expression data. While MSR is only capable of detecting shifting patterns, and VE cannot recognize both kind of patterns simultaneously, VE^t has been proven to go beyond the other two measures.

5 Discussion

In this section, we discuss the use of VE^t for bicluster evaluation. In particular, we study the value of VE^t for those biclusters in which the presence of patterns is not perfect. That is, when the tendency of the data in a bicluster is similar to a perfect pattern but does not completely match with the equation 3.

In order to check the behaviour of VE^t whenever a bicluster does not follow a perfect pattern, we add an additive term ε_{ij} to the combined pattern equation.

The meaning of this new term corresponds to the error made by the assumption that the bicluster can be represented by a perfect pattern.

$$b_{ij} = \pi_j \times \alpha_i + \beta_i + \varepsilon_{ij} \quad (6)$$

It is possible therefore to study the variations produced to VE^t depending on the values of ε_{ij} . Nevertheless, it is not so simple due to the huge amount of situations depending on the distribution and the magnitude of the ε_{ij} values in the data matrix.

In two specific situations the value of VE^t will not be affected when the errors could be included in the former equation 6. These two cases correspond to those in which ε_{ij} values are either a constant or constants per conditions (rows). In both cases it is possible to eliminate the term ε_{ij} from the equation, since it can be considered to be a part of β_i .

Nevertheless, the cases in which ε_{ij} cannot not be included in the perfect pattern equation are very difficult to study analytically. For this reason, we have performed a test to check the tendency of the VE^t values with regard to the error values. This test consist of the addition of random errors to a synthetic bicluster with perfect shifting and scaling patterns. The original bicluster is the one shown in Fig. 1. Specifically, we have generated 100 synthetic biclusters adding random errors to the bicluster in the figure, and we have repeated this process 200 different times, varying the amplitude of the errors from one time to another. We start adding negative errors in the range of $[-10, 0]$, and obtain 100 different biclusters. Then we decrease the amplitude by 0.1 and repeat the process (range $[-9.90, 0]$). Once the amplitude of the errors has reached the zero value, we start again generating biclusters with positive errors, increasing the amplitude from 0.1 up to 10. The whole process produced 100 sets of 100 biclusters with negative errors and 100 sets of 100 biclusters with positive errors (built using the same strategy as for negative). Therefore, the random errors have been drawn from an uniform distribution corresponding to the ranges. Note that the type of the error values is a double type. This introduces more diversity in the distribution of the error data.

Within the process, we evaluate each produced bicluster using the three measures: MSR, VE and VE^t . Then we obtain the mean of each measure for each group of 100 biclusters of the same range of errors. This data has been represented in Figs. 2, 3 and 4, where the x-axis represents the mean of the error for each amplitude (this value matches up with the value in the middle of the range of errors) and the y-axis corresponds to the mean of the specific measure for each figure.

From Fig. 2 it is possible to observe that VE^t presents a linear decreasing tendency in relation to the amount of error in a bicluster. In other words, the similar a bicluster is to a perfect pattern, the lower its VE^t value will be, and

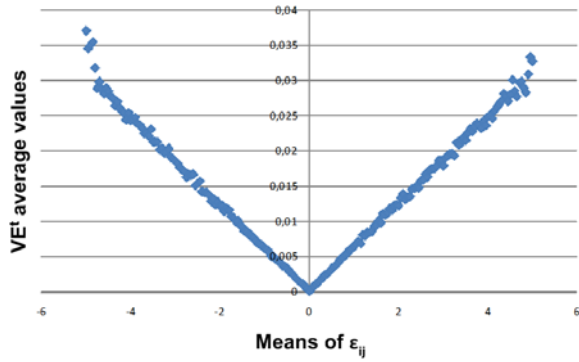


Fig. 2. VE^t behaviour in biclusters with errors

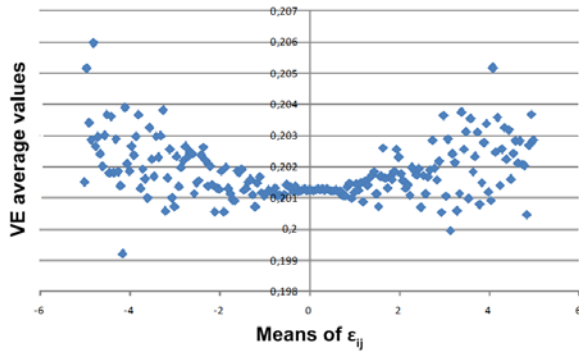


Fig. 3. VE behaviour in biclusters with errors

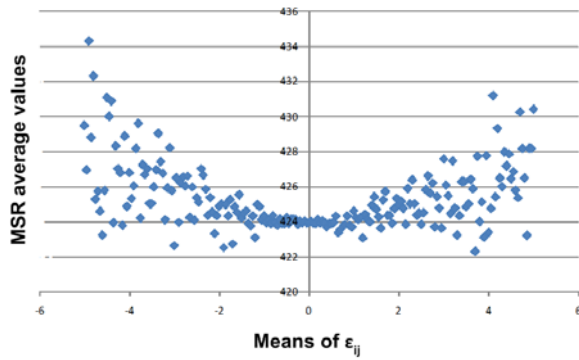


Fig. 4. MSR behaviour in biclusters with errors

we can establish a linear relationship between VE^t and the amount of error. Nevertheless, it is not possible to come to the same conclusion for either VE or MSR. Figs. 3 and 4 depict the connection of the errors with VE and MSR, respectively. Although the general tendency seems to be that both measures are higher for biclusters with higher error values, we cannot establish any correspondence between them. In both figures it is possible to see some cases in which the mean of the biclusters with errors is lower than the original bicluster.

As a conclusion, VE^t outperforms both MSR and VE efficacy for identifying behavioural patterns in synthetic data. Our expectations are that this behaviour would be extensive to real gene expression data.

6 Conclusions

This work introduces an enhanced version of a previous measure for evaluating biclusters from gene expression data. This new variant, named VE^t , allow finding biclusters with both shifting and scaling patterns simultaneously in gene expression data. No previous evaluation measure for biclusters is able of identifying this kind of pattern, for this reason we are sure VE^t constitutes an important contribution to the topic.

This paper also includes analytical proofs which demonstrate the capability of VE^t for detecting any kind of perfect pattern in gene expression data. Furthermore, we have also proved that VE^t presents a linear relationship with the amount of error in a bicluster.

For future work, we have planned to use VE^t together with an evolutionary framework in order to search for biclusters in gene expression data. The obtained results will be compared to those obtained by similar heuristics and evaluation measures. Biological validation of the results will also be performed in order to validate our approach.

Acknowledgement

This research is supported by the Spanish Ministry of Science and Technology under grant TIN2007-68084-C02-00 and the Junta de Andalucía Research Program.

References

1. Aguilar-Ruiz, J.S.: Shifting and scaling patterns from gene expression data. *Bioinformatics* 21, 3840–3845 (2005)
2. Aguilar-Ruiz, J.S., Rodriguez, D.S., Simovici, D.A.: Biclustering of gene expression data based on local nearness. In: *Proceedings of EGC 2006*, Lille, France, pp. 681–692 (2006)
3. Baldi, P.: *DNA Microarrays and Gene Expression: From Experiments to Data Analysis and Modeling*. Cambridge University Press, Cambridge (2002)

4. Bleuler, S., Prelič, A., Zitzler, E.: An EA framework for biclustering of gene expression data. In: Congress on Evolutionary Computation (CEC-2004), pp. 166–173. IEEE, Los Alamitos (2004)
5. Bryan, K., Cunningham, P., Bolshakova, N.: Application of simulated annealing to the biclustering of gene expression data. *IEEE Transactions on Information Technology on Biomedicine* (2006)
6. Cano, C., Adarve, L., López, J., Blanco, A.: Possibilistic approach for biclustering microarray data. *Computers in Biology and Medicine* 37(10), 1426–1436 (2007)
7. Cheng, Y., Church, G.M.: Biclustering of expression data. In: Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology, La Jolla, CA, pp. 93–103 (2000)
8. Cho, H., Dhillon, I.S.: Effect of data transformation on residue. Technical report (2007)
9. Coelho, G.P., de Franca, F.O., Zuben, F.J.V.: Multi-objective biclustering: When non-dominated solutions are not enough. *Journal of Mathematical Modelling and Algorithms* 8(2), 175–202 (2009)
10. Divina, F., Aguilar-Ruiz, J.S.: Biclustering of expression data with evolutionary computation. *IEEE Transactions on Knowledge & Data Engineering* 18(5), 590–602 (2006)
11. Divina, F., Aguilar-Ruiz, J.S., Pontes, B., Giráldez, R.: An effective measure for assessing the quality of biclusters (in Press, 2010)
12. Hartigan, J.: Direct clustering of a data matrix. *Journal of the American Statistical Association* 67(337), 123–129 (1972)
13. Liu, J., Li, Z., Hu, X., Chen, Y.: Biclustering of microarray data with mospo based on crowding distance. *BMC bioinformatics* 10(suppl. 4), S9+ (2009)
14. Madeira, S.C., Oliveira, A.L.: Biclustering algorithms for biological data analysis: A survey. *IEEE Transactions on Computational Biology and Bioinformatics* 1, 24–25 (2004)
15. Pontes, B., Divina, F., Giráldez, R., Aguilar-Ruiz, J.S.: Virtual error: A new measure for evolutionary biclustering. In: Fifth European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics (EvoBio 2007), pp. 217–222 (2007)
16. Pontes, B., Giráldez, R., Divina, F., Martínez-Álvarez, F.: Evaluación de biclusters en un entorno evolutivo. In: IV Taller nacional de minería de datos y aprendizaje (TAMIDA), pp. 1–10 (2007)
17. Tanay, A., Sharan, R., Shamir, R.: Discovering statistically significant biclusters in gene expression data. *Bioinformatics* 18, 136–144 (2002)
18. Tilstone, C.: Dna microarrays: Vital statistics. *Nature* 424, 610–612 (2003)
19. Wang, H., Wang, W., Yang, J., Yu, P.S.: Clustering by pattern similarity in large data sets. In: ACM SIGMOD International Conference on Management of Data, Madison, WI, pp. 394–405 (2002)
20. Xu, X., Lu, Y., Tung, A.K.H., Wang, W.: Mining shifting-and-scaling co-regulation patterns on gene expression profiles. In: 22nd International Conference on Data Engineering (ICDE'06), pp. 89–99 (2006)
21. Yang, J., Wang, H., Wang, W., Yu, P.S.: An improved biclustering method for analyzing gene expression profiles. *International Journal on Artificial Intelligence Tools* 14, 771–790 (2005)