

NRC Publications Archive Archives des publications du CNRC

Building Virtual Reality Spaces for Visual Data Mining with Hybrid Evolutionary-Classical Optimization: Application to Microarray Gene Expression Data

Valdés, Julio

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

Publisher's version / Version de l'éditeur:

Proceedings of the IASTED International Joint Conference on Artificial Intelligence and Soft Computing (ASC'2004), 2004

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

<https://nrc-publications.canada.ca/eng/view/object/?id=2cc526e7-a2c7-473e-9483-a784bc50beef>
<https://publications-cnrc.canada.ca/fra/voir/objet/?id=2cc526e7-a2c7-473e-9483-a784bc50beef>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at <https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site <https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



National Research
Council Canada

Conseil national
de recherches Canada

Institute for
Information Technology

Institut de technologie
de l'information

NRC - CNRC

Building Virtual Reality Spaces for Visual Data Mining with Hybrid Evolutionary-Classical Optimization: Application to Microarray Gene Expression Data *

Valdés, J.
September 2004

* published in the Proceedings of the IASTED International Joint Conference on Artificial Intelligence and Soft Computing (ASC'2004). September 1-3, 2004. Marbella, Spain. ACTA Press, Anaheim, USA. pp. 161-166. NRC 47390.

Copyright 2004 by
National Research Council of Canada

Permission is granted to quote short excerpts and to reproduce figures and tables from this report, provided that the source of such material is fully acknowledged.

BUILDING VIRTUAL REALITY SPACES FOR VISUAL DATA MINING WITH HYBRID EVOLUTIONARY-CLASSICAL OPTIMIZATION: APPLICATION TO MICROARRAY GENE EXPRESSION DATA.

Julio J. Valdés

National Research Council of Canada
Institute for Information Technology
1200 Montreal Road
Ottawa ON K1A 0R6, Canada
email: julio.valdes@nrc-cnrc.gc.ca

ABSTRACT

Visual data mining via the construction of virtual reality spaces for the representation of data and knowledge, involves the solution of optimization problems. This paper introduces a hybrid technique based on particle swarm optimization (PSO) combined with classical optimization methods. This approach is applied to very high dimensional data from microarray gene expression experiments in order to understand the structure of both raw and processed data. Experiments with data sets corresponding to Alzheimer's disease show that high quality visual representations can be obtained by combining PSO with classical optimization methods. The behavior of some of the parameters controlling the swarm evolution is also studied.

KEY WORDS

data mining, virtual reality, hybrid optimization.

1 Introduction

Knowledge discovery is the non-trivial process of identifying valid, novel, potentially useful, and ultimately *understandable patterns* in data [3], and the role of visualization techniques in the knowledge discovery process is well known. Data and patterns are concepts which should be considered in a broad sense. There are different kinds of data (relational, graphical, symbolic, etc.), and there are also patterns of different kinds (geometrical, logical, etc.). The increasing rates of data generation require the development procedures facilitating the *understanding* of the internal structure of data more rapidly and intuitively. Moreover, the increasing complexity of the data analysis procedures makes it more difficult for the user (not necessarily a mathematician or data mining expert), to extract useful information out of the results from the various techniques. This makes graphical representation directly appealing.

A virtual reality technique for visual data mining on heterogeneous, imprecise and incomplete information systems was introduced in [12, 14]. Several reasons make Virtual Reality (VR) a good paradigm for visual data min-

ing: different representations according to human perception preferences can be chosen, it allows *immersion* (the user can navigate inside the data, interact with the objects, etc), it creates a *living* experience, and the user may see the whole world or concentrate on specific details. For using VR the user needs no specialized knowledge or skills.

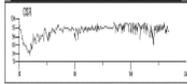
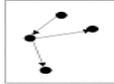
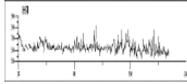
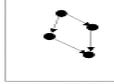
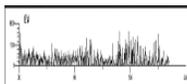
The purpose of this paper is to explore the construction of high quality VR spaces for visual data mining using a combination of evolutionary and classical optimization techniques. In particular, particle swarm optimization (PSO) and classical methods are used independently and also combined. An application example using microarray gene expression data from Alzheimer's disease is presented.

2 Virtual Reality Spaces for Visual Data Mining

Information systems were introduced in [9]. They have the form $S = \langle U, A \rangle$ where U is the *universe* and A the set of *attributes*, such that each $a \in A$ has a domain V_a and an evaluation function f_a . In a general sense the V_a are not required to be finite (Table-1). Here *heterogeneous* and *incomplete* information systems will be considered [13]. A heterogeneous domain is defined as a Cartesian product of a collection of *source sets* (Ψ_i): $\hat{\mathcal{H}}^n = \Psi_1 \times \dots \times \Psi_n$, where $n > 0$ is the number of *information sources* to consider. As an example, consider the case of a domain where objects are characterized by attributes given by continuous crisp quantities, discrete features, fuzzy features, graphs and digital images. The heterogeneous domain is constructed as $\hat{\mathcal{H}}^n = \hat{\mathcal{R}}^{n_r} \times \hat{\mathcal{O}}^{n_o} \times \hat{\mathcal{N}}^{n_m} \times \hat{\mathcal{F}}^{n_f} \times \hat{\mathcal{G}}^{n_g} \times \hat{\mathcal{I}}^{n_i}$ (assuming that there are n_r real-valued, n_o ordinal-valued, n_m nominal-valued, n_f fuzzy-valued, n_g graph-valued, and n_i image-valued variables). Other kinds of heterogeneous domains can be constructed in the same way, using the appropriate source sets. In more general information systems, the universe is endowed with a set of relations (Γ) of different arities.

A *virtual reality space* for the visual representation of information systems [12], [14], is defined as $\Upsilon = \langle$

Table 1. An example of a heterogeneous data base. The attributes are of different nature: nominal, ordinal, ratio, fuzzy, images, time-series and graphs, also containing missing values (?).

A_1	A_2	A_3	A_4	A_5	A_6	A_7
yellow	high	2.5				
blue	?	3.8				
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
red	low	?				

$\underline{Q}, G, B, \mathfrak{R}^m, g_o, l, g_r, b, r \rangle$. \underline{Q} is a relational structure composed by objects and relations ($\underline{Q} = \langle O, \Gamma^v \rangle$, $\Gamma^v = \langle \gamma_1^v, \dots, \gamma_q^v \rangle$, $q \in \mathbb{N}^+$ and the $o \in O$ are objects), G is a non-empty set of *geometries* representing the different objects and relations. B is a non-empty set of *behaviors* (ways in which the virtual world objects may express themselves: movement, response to stimulus, etc.). $\mathfrak{R}^m \subset \mathbb{R}^m$ is a *metric space* of dimension m (the actual virtual reality geometric space). The other elements are mappings: $g_o : O \rightarrow G$, $l : O \rightarrow \mathfrak{R}^m$, $g_r : \Gamma^v \rightarrow G$, $b : O \rightarrow B$, r is a collection of characteristic functions for selecting which of the original relations will be represented in the virtual world. The representation of an information system \hat{S} in a virtual world requires the specification of several sets and a collection of extra mappings: $\hat{S}^v = \langle O, A^v, \Gamma^v \rangle$, \underline{Q} in Υ . A desideratum for \hat{S}^v is to keep as many properties from \hat{S} as possible. Thus, a requirement is that U and O are in one-to-one correspondence (with a mapping $\xi : U \rightarrow O$). The structural link is given by a mapping $f : \mathcal{H}^n \rightarrow \mathfrak{R}^m$. If $u = \langle f_{a_1}(u), \dots, f_{a_n}(u) \rangle$ and $\xi(u) = o$, then $l(o) = f(\xi(\langle f_{a_1}(u), \dots, f_{a_n}(u) \rangle)) = \langle f_{a_1^v}(o), \dots, f_{a_m^v}(o) \rangle$ ($f_{a_i^v}$ are the evaluation functions of A^v). It is natural to require that $\Gamma^v \subseteq \Gamma$, thus having a virtual world portraying selected relations from the information system. The function f can be constructed as to maximize some structure preservation criteria as in multidimensional scaling [1], or minimize some measure of information loss [11], [8]. Frequently used measures are:

$$S - stress = \sqrt{\frac{\sum_{i < j} (\delta_{ij}^2 - \zeta_{ij}^2)^2}{\sum_{i < j} \delta_{ij}^4}} \quad (1)$$

$$Sammon = \frac{1}{\sum_{i < j} \delta_{ij}} \sum_{i < j} \frac{(\delta_{ij} - \zeta_{ij})^2}{\delta_{ij}}$$

where δ_{ij} is a dissimilarity measure between any two objects i, j in the original space, and ζ_{ij^v} is another dissimilarity measure defined on objects i^v, j^v of the virtual reality space (the images of i, j under f). The f mappings obtained using approaches of this kind are *implicit*. A snapshot from a virtual world containing a mixture of data objects and decision rules is shown in Fig-1.

3 Evolutionary Computation in the Construction of the Virtual Reality Space

The typical *desiderata* for the visual representation of data and knowledge, can be formulated in terms of minimizing information loss, maximizing structure preservation, maximizing class separability, or their combination, which leads to single or multi-objective optimization problems. In many cases, these concepts can be expressed deterministically using continuous functions with well defined partial derivatives. This is the realm of classical optimization where there is a plethora of classical methods with well known properties. However, factors like the complexity of the objective function(s) chosen (e.g. f in the definition of the VR space), the dimensionality of the problem (an integer multiple of the data size), and the intrinsic properties of the data themselves, condition highly complex multidimensional error surfaces in the parameter space, with the risk of getting trapped in local extrema. The need to achieve high quality visual representations using virtual reality spaces implies the use of algorithms which explore the search space both globally and locally. A hybrid approach combining evolutionary computation methods or simulated annealing with classical optimization techniques, like Powell, Fletcher-Reeves, Davidon-Fletcher-Powell, Newton, and others is natural. A simple and straightforward way is to start the optimization process with a global optimizer, and

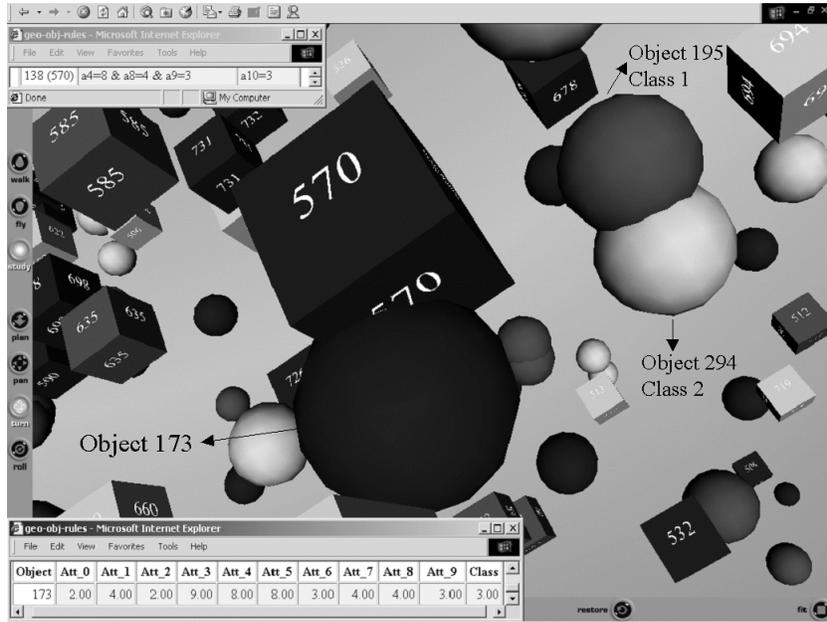


Figure 1. Simultaneous representation of data objects and decision rules from the same data set. Spheres are data objects, and numbered cubes, decision rules. The two embedded web pages show the objects and the rules values in their original space.

at some point, use the current best solution (or the k -best solutions) as initial approximation for a classical optimization technique. A more elaborated strategy would be to alternate cycles of global and local search in parallel. In this respect, EC techniques like genetic algorithms (GA), evolution strategies (ES) and particle swarm optimization (PSO) have a great potential because of their global search capabilities.

Preliminary experiments showed that genetic algorithms and evolution strategies were more affected by the *curse of dimensionality* than particle swarm optimization. This topic deserves a dedicated investigation because the construction of a virtual reality space of dimension m representing N objects implies the optimization of a function depending on $m \cdot N$ parameters. For the typical case of a 3D representation, it means three times the data base size, measured in terms of data objects. The size of current databases and knowledge bases range from hundreds to thousands or even million of objects; posing a major problem.

Particle swarm optimization (PSO) is a population-based stochastic search process, modeled after the social behavior of bird flocks and similar animal collectives [6, 7]. The algorithm maintains a population of particles, where each particle represents a potential solution to an optimization problem. In the context of PSO, a swarm refers to a number of potential solutions to the optimization problem, where each potential solution is referred to as a particle. Each particle i maintains information concerning its current position and velocity, as well as its best location overall. These elements are modified as the process evolves, and different strategies have been proposed for updating

them, which consider a variety of elements like the intrinsic information (history) of the particle, *cognitive* and *social* factors, the effect of the *neighborhood*, etc, formalized in different ways. The swarm model used has the form proposed in [16]

$$\begin{aligned}
 v_{id}^{k+1} &= \omega \cdot v_{id}^k + \phi_1 \cdot (p_{id}^k - x_{id}^k) + \phi_2 \cdot (p_{gd}^k - x_{id}^k) \\
 x_{id}^{k+1} &= x_{id}^k + v_{id}^{k+1} \\
 \phi_i &= b_i \cdot r_i + d_i, \quad i = 1, 2
 \end{aligned}
 \tag{2}$$

where v_{id}^{k+1} is the velocity component along dimension d for particle i at iteration $k+1$, and x_{id}^{k+1} its location; b_1 and b_2 are positive constants both equal to 1.5; r_1 and r_2 are random numbers uniformly distributed in the range $(0, 1)$; d_1 and d_2 are positive constants both equal to 0.5, to cooperate with b_1 and b_2 in order to confine ϕ_1 and ϕ_2 within the interval $(0.5, 2)$; ω is an inertia weight.

4 Experiments

In order to study the behavior of hybrid optimization when constructing virtual reality spaces for gene expression data, two kinds of experiments were made using two data sets (described below). The experimental plan was: *i*) to apply the classical methods to the two data sets, and *ii*) for each of the data sets, to use the PSO final result as an initial approximation for the classical methods, as a hybrid algorithm.

The classical techniques used in the study were: Powell, Fletcher-Reeves, and Davidon-Fletcher-Powell [10]. The Powell method does not require the partial derivatives

of the objective function. In the first phase, the four classical optimization methods were applied to each data set, using 100 different random initial approximations for a total of 400 runs per data set. The same set of seeds was used in all experiments in order to ensure comparability.

From the point of view of constructing the virtual reality spaces, the f transform sought was one minimizing Sammon's error (1), with a dissimilarity in the space of the original attributes (genes) given by $\delta_{ij} = (1 - \hat{s}_{ij})/\hat{s}_{ij}$, where \hat{s}_{ij} is Gower's similarity coefficient [4]. The Euclidean distance was the measure used as ζ_{ij} in the VR space (Υ).

4.1 Gene Expression Data: A study in Alzheimer's Disease

Alzheimer's disease (AD) is a chronic, progressive, debilitating condition which, along with other neurodegenerative diseases, represents the largest area of unmet need in modern medicine [15]. There is now renewed hope that genomics technologies, particularly gene expression profiling, can contribute significantly to the understanding of the disease. Genome-wide expression profiling of thousands of genes provides rich datasets that can be mined to extract information on the genes that best characterize the disease state [2], [5], [15], and others. However, in such data sets, patient samples are characterized by thousands of attributes representing the expression intensities of the different genes chosen in the framework of the experiment. They exhibit extremely complex patterns of dependencies, redundancies, noise, etc, making the process of understanding the meaning, role, and importance of the different genes, very difficult. In particular, a study of gene expression Alzheimer's data from a data mining perspective is presented in [15].

In that study, the data set was composed of 23 samples taken from Alzheimer and non-Alzheimer cases. These samples are described in terms of 9600 genes. An implicit representation was computed via deterministic optimization with a gradient descent technique (Newton's method). Several solutions were found using different initial approximations, and for comparison purposes, the same seeds were used.

In this study, a set of experiments using PSO was set forth with the following parameters: number of particles = 100, maximum velocity = {0.1, 0.15, 0.20, 0.25, 0.30}, initial weight={0.1, 0.2, 0.4, 0.6, 0.8, 0.9}, final weight = {0.1, 0.2, 0.4, 0.6, 0.8, 0.9}, and number of iterations = {1000, 2000, 4000, 8000}, for a total of 720 experiments. The results obtained with classical optimization (CO), PSO and the hybrid algorithm PSO+CO are shown in Table-2.

It is impossible to represent a virtual reality space on printed media because of the very nature of a virtual world, where navigation, interaction, and world changes are typical operations. However, it is possible to capture a snapshot of part of a virtual world resulting from a hybrid al-

Table 2. Experiments with Alzheimer data sets: Comparison of the classical methods with PSO and the hybrid algorithm formed by combining PSO with each of the classical methods. Data objects are described in terms of the original 9600 original genes, and in the four best discovered by a data mining procedure. Error ranges are shown in square brackets

Classical Method	9600 Genes Sammon Error	4 Best Genes Sammon Error
	[0.03437583, 0.03868285]	[0.0690399, 0.06936745]
Powell	[0.03437595, 0.03868285]	[0.06903999, 0.06933347]
Fletcher-Reeves	[0.03437584, 0.03694771]	[0.06903994, 0.06936745]
Davidon- Fletcher-Powell	[0.03437583, 0.03690736]	[0.0690399, 0.06936724]
PSO	0.034376	0.0690399
Hybrid Method	9600 Genes Sammon Error	4 Best Genes Sammon Error
	[0.03437583, 0.03437589]	[0.0690399, 0.0690399]
PSO + Powell	0.0343758948	0.0690398961
PSO + Fletcher-Reeves	0.0343758576	0.0690398961
PSO + Davidon- Fletcher-Powell	0.0343758315	0.0690398961

gorithm applied to the original Alzheimer's data using all 9600 genes, as shown in Fig-2.

The samples corresponding to the Alzheimer's class are colored black, and the non-Alzheimer's with a light color. In this case it is clearly seen that the samples corresponding to the Alzheimer's class appear as more homogeneous and compact (i.e. more similar to each other) than those from the non-Alzheimer class. However, the Alzheimer class appears *wrapped* by the non-Alzheimer class, which is more irregular, indicating that the classes are not linearly separable.

In the case of the Alzheimer data with 9600 genes, the classical methods and the PSO gave very similar results. However, the error range for the hybrid algorithm is smaller than that of the classical methods. This indicates that high quality VR spaces can be obtained with relative independence of the classical method used by the hybrid

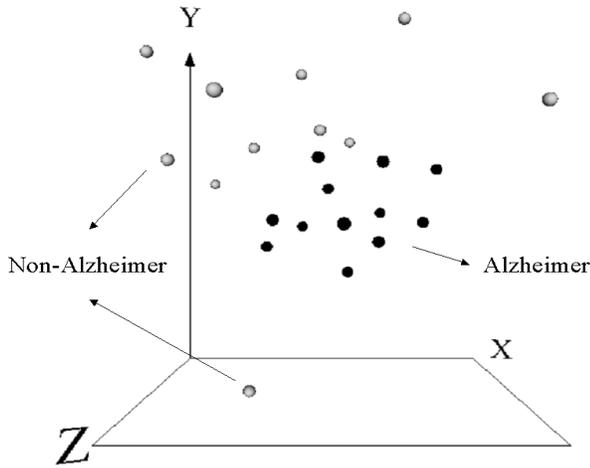


Figure 2. Alzheimer Data (9600 genes): VR space representing the data when object locations are resulting from a hybrid optimization using the best PSO result as initial approximation to Powell's method. X, Y, and Z are the axis of the VR space resulting from the nonlinear transformation of the original 9600 attributes (the genes), performed by the implicit function minimizing the error measure.

algorithm (Table 2).

From the point of view of the PSO algorithm alone, the set of 100-best solutions is upper-bounded by an error of 0.0344337, which is better than many of those obtained with the classical methods. Within this set, 40% of the solutions were found using 4000 iterations, and the remaining 60% with 8000 iterations, indicating the overall computational effort involved. From the 100-best solutions, 20% of them were found with an increasing weight updating scheme, and in 60% with a decreasing one. In 20% of the solutions, the weight was constant. The distribution of maximum particle velocities within the 100-best was: $\{0.1 : 10\%, 0.15 : 10\%, 0.20 : 40\%, 0.25 : 40\%, 0.30 : 0\%\}$. It suggests that non-extreme velocities are better for this kind of problem.

From the practical point of view, the results obtained by PSO and the classical methods can be considered equivalent. This suggests that the phase of global search could have been shortened without losing performance in terms of the error measure. In this case, a reduction of the computational load of the hybrid algorithm could be achievable. In other words, the proportion between global and local search in this problem could be determined such that the same final solutions can be found at smaller computational cost. This topic deserves further investigation.

After a data mining process described in [15], a set of four most relevant genes was found. As in the previous case, a VR space representation was computed by the classical methods, PSO, and the hybrid algorithm. The settings for the PSO experiments were the same as with the Alzheimer data set using the original 9600 genes, for a to-

tal of 720 experiments. The results are shown in Table-2.

A snapshot of the VR space is shown in Fig-3. Now it appears completely polarized, with the Alzheimer and non-Alzheimer classes appearing as distant clouds occupying well separated half-spaces. The two classes are now well differentiated entities, also linearly separable.

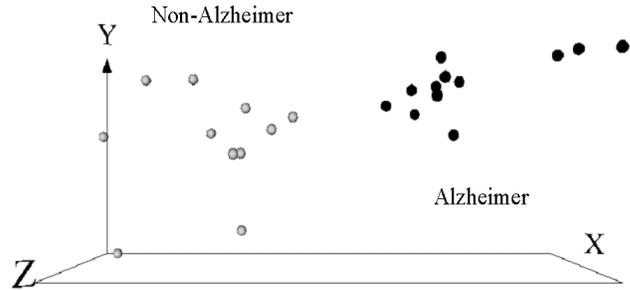


Figure 3. Alzheimer Data (4 best genes): VR space representing the data. The object locations are the result of Particle Swarm Optimization alone (in practical terms the classical methods couldn't improve the pure PSO result). X, Y, and Z are the axis of the VR space resulting from the error minimization.

The behavior of the classical methods, the PSO and the hybrid algorithms was almost identical, thus suggesting that the best solution found possibly corresponds to the global optimum. In this case, the structure of the dissimilarity matrix in the original 4D space had a decisive effect. As a result of the data mining procedure, the selected genes lead to a better distribution of the dissimilarity values within the matrix. This is a consequence of a better expressed class structure inherent to the data objects.

From the point of view of the PSO performance, the error level for the 100 best particles is 0.06904022, which is under those of the GD. Within this subset, 18% of the solutions were obtained with 2000 iterations, 43% with 4000 and the remaining 39% with 8000. Within the 100-best solutions, 40% were found to be obtained with an increasing weight updating scheme, and 60% with a decreasing one. None of the solutions in the subset were found using a constant weight. For this problem, the frequency of good solutions obtained with decreasing weight updating is larger than that obtained with an increasing scheme. This behavior differs from the one reported in [16] for several of DeJong's functions and deserves further studies. The distribution of maximum particle velocities within the 100-best was: $\{0.1 : 20\%, 0.15 : 13\%, 0.20 : 27\%, 0.25 : 0\%, 0.30 : 40\%\}$. This behavior differs from the one found in the case of the 9600 genes data set. A multimodal frequency distribution is a curious behavior in this case, but the number of velocity values considered was very limited.

5 Conclusions

The combination of Particle Swarm Optimization with classical optimization methods proved to be effective in computing virtual reality spaces for the representation of complex data sets, like those coming from microarray gene expression experiments. In particular, it seems to narrow the range of the final errors, thus leading to more confident results. High quality representations of both raw data and results of data mining algorithms can be obtained using hybrid optimization methods, thus reducing the chance of getting trapped at local optima. It seems that good solutions are found with medium values for the maximum velocity parameters for PSO, but its behavior is not always consistent. None of the two major PSO weight updating schemes (increasing or decreasing) seems to be predominant for the data sets investigated here. The results obtained for microarray gene expression data are encouraging, but they should be considered very preliminary. Further research is necessary in order to find the best ranges of parameters, the best evolutionary schemes, and more elaborated strategies for combining PSO with classical optimization methods, including the proportion between global and local search.

Acknowledgments

This research was performed in the framework of the Biomine Project. The author would like to thank Alan Barton, Robert Orchard and Fazel Famili from the National Research Council of Canada.

References

- [1] I. Borg, J. Lingoes, *Multidimensional Similarity Structure Analysis*. (Heidelberg, Springer-Verlag 1987).
- [2] Brown V. M, Ossadtchi A., Khan A.H., Cherry S.R., Leahy R.M. and Smith D.J., High-throughput imaging of brain gene expression. *Genome Res*, 2002, 12: 244-54.
- [3] Fayyad, U., Piatesky-Shapiro, G., Smyth, P.: From Data Mining to Knowledge Discovery. [In] U.M. Fayyad et al. (eds.) *Advances in Knowledge Discovery and Data Mining*. AAAI Press, 1996, pp. 1-34
- [4] J. C. Gower, "A General Coefficient of Similarity and Some of its Properties". *Biometrics* Vol.1 No. 27, 1973 pp. 857-871
- [5] Hata R., Masumura M., Akatsu H., Li F, Fujita H, Nagai Y, Yamamoto T, Okada H, Kosaka K., Sakanaka M. and Sawada T: Up-regulation of calcineurin Abeta mRNA in the Alzheimer's disease brain: assessment by cDNA microarray. *Biochem Biophys Res Commun*, 2001, 284: 310-6.
- [6] J. Kennedy, R. C. Eberhart, Particle Swarm Optimization, *Proceedings of the IEEE International Joint Conference on Neural Networks*, Vol. 4, pp 1942-1948, 1995.
- [7] J. Kennedy, R. C. Eberhart, Y. Shi, *Swarm Intelligence*, (Morgan Kaufmann, 2002).
- [8] M. Jianchang, A. Jain, "Artificial Neural Networks for feature Extraction and Multivariate Data Projection". *IEEE Trans. On Neural Networks*. Vol. 6, No. 2, 1995, pp. 296-317
- [9] Z. Pawlak, *Theoretical Aspects of Reasoning About Data*. (Dordrecht, Netherlands, Kluwer Academic Publishers, 1991).
- [10] W. Pres, B.P. Flannery, S.A. Teukolsky, W.T Vetterling, *Numeric Recipes in C*. (Cambridge, MA, Cambridge University Press, 1992).
- [11] J. W. Sammon, A non-linear mapping for data structure analysis. *IEEE Trans. Computers*, C-18, 401-408, (1969)
- [12] J. J. Valdés, Virtual Reality Representation of Relational Systems and Decision Rules: An exploratory Tool for Understanding Data Structure. In *Theory and Application of Relational Structures as Knowledge Instruments*. Meeting of the COST Action 274 (P. Hajek. Ed). Prague, November 14-16, 2002.
- [13] J. J. Valdés, Similarity-based Heterogeneous Neurons in the Context of General Observational Models. *Neural Network World*. Vol 12., No. 5, (2002) pp. 499-508.
- [14] J. J. Valdés, Virtual Reality Representation of Information Systems and Decision Rules: An Exploratory Tool for Understanding Data and Knowledge. *Proceedings of the 9-th Int. Conf. on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing* (Wang, Liu, Yao, Skowron, eds.). Chongqing, China, Oct 8-12, 2003. *Lecture Notes in Artificial Intelligence* LNAI 2639, pp. 615-618. Springer-Verlag, 2003.
- [15] P. R. Walker, B. Smith, Q. Y. Liu, F. Famili, J. J. Valdés, Z. Liu, B. Lach, Data Mining of Gene Expression Changes in Alzheimer Brain. *Int. Jour. of Artificial Intelligence in Medicine*, Elsevier Acad. Pub. 2004.
- [16] Yong-ling Zheng, Long-hua Ma, Li-yan Zhang, Ji-xin Qian, Empirical Study of Particle Swarm Optimizer with an Increasing Inertia Weight. *Proceedings of the World Congress on Evolutionary Computation*. Canberra, Australia, Dec 8-12, 2003, IEEE Press, pp 221-226 pp.