Quantum-Inspired Evolutionary Classification of Driving Sequences in Vehicle Emission Factor Measurement

Pasquale Arpaia[1], Giovanni Meccariello[2], Mario Rapone[2], Antonio Zanesco[1]

[1]*Dip. di Ingegneria, Università del Sannio, Corso Garibaldi 107, 82100 Benevento, ITALY. Ph : +39 0824 305804-17, Fax: +39 0824 305840, E-mail: arpaia@unisannio.it, azanesco@unisannio.it*
[2]*Consiglio Nazionale delle Ricerche, Istituto Motori, Viale Marconi 8, 80125 Napoli, ITALY. Ph : +39 081 7177114-75, Fax: +39 081 2396097, E-mail: m.rapone@im.cnr.it, g.meccariello@im.cnr.it*

Abstract - A heuristic procedure of classification, the quantum-inspired classifier (QIC), exploiting search space exploration and resource exploitation of quantum computing on a software basis is proposed. The application to the problem of speed sequence classification for vehicle emission factor determination based on drive styles is shown. Experimental results are discussed by showing the QIC capability of converging better and faster than classical evolutionary algorithms.

I. Introduction

In automotive industry, interest in vehicle pollution has constantly grown in order to reduce pollutant emissions without decreasing vehicle performance. Several studies showed that, in on-road tests, emissions are influenced strongly by driving behavior, and, in laboratory tests, by the choice of reference driving-cycles [1]-[2]. In Fig. 1, a measurement procedure for estimating emissions starting from vehicle speed acquisition data and reference driving styles is shown [3]. After speed acquisition, data are divided in suitable *sequences* (SEQs). The resulting SEQs are segmented and classified by means of suitably extracted characteristic features, as well as of class limits predefined through a corresponding generator. From such classified SEQs and reference styles obtained by a related generator, the current driving style is obtained. Finally, the emission amount is estimated.

In such a measurement procedure, one of most difficult problem is the sequence classification (highlighting circle in Fig. 1). With this aim, different techniques, statistical or heuristics, have been developed [3]. Among heuristic techniques, evolutionary ones showed interesting performance [4]. A population of possible solutions evolves as biologic individuals in order to achieve the optimization of an objective function (fitness). After genetic algorithms (GAs), different evolutionary techniques were proposed in order to face their intrinsic resource waste [5]-[13]. In particular, cultural algorithms demonstrated significant efficiency and effectiveness, though the difficulty in tuning their strongly driven evolution leads sometime to premature convergence [14]. Recently, quantum computing has been proposed as a new concept of evolutionary technique basically inspired by quantum physics principles [15]. For the lack of quantum hardware, classical evolutionary algorithms (such as GAs) have been exploited in order to derive hybrids algorithms, called quantum-inspired evolutionary algorithms (QIEAs) [16]. They proved to have a satisfying trade-off between research space exploration and resource exploitation, also in problems where solutions lies in continuous search spaces, such as in the case of drive sequence classification.

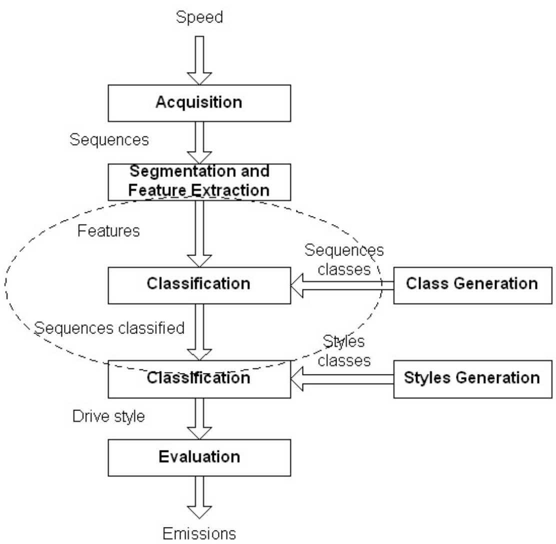

Fig. 1 – Vehicle emission measurement based on speed sequence and drive style classification.

In this paper, the problem of driving sequence classification for vehicle emission quantities measurement is faced by a heuristic procedure based on quantum-inspired algorithms, the *quantum-inspired classifier (QIC)*. Experimental results, related to the implementation of the proposed QIC to automatic classification of driven sequences, are discussed.

II. The proposed approach

In the following, (i) the *QIC basic ideas*, (ii) the *software quantum observation*, (iii) the *quantum-inspired classifier*, and (iv) the *driving sequence classifier* are illustrated.

A: QIC Basic Ideas

In the software simulation of a quantum algorithm without quantum hardware, a quantum pulse (*q-bit*) is employed [17]. The q-bit is described by two parameters: the center ρ, and the width σ (Fig. 2). While a classical genetic chromosome represents a deterministic point in the search space, the q-bit represents a portion of the search space, expressed by its width σ, and its center ρ, where the solution can be found within a certain approximation or likelihood.

On the basis of such q-bits, a quantum population is built. Owing to the lack of hardware, the current quantum population is assessed by means of a traditional evolutionary algorithm, namely a GA. This realizes a twofold levels search: on a rough level, the quantum population evolves according to its exploitation power; on a fine level, inside the local region found by the quantum-inspired search, a genetic population evolves according to its exploration power.

Moreover, the search approximation is updated dynamically at each quantum evolution cycle in order to optimize solution space exploration and resource exploitation. With this aim, quantum pulses are modulated according to the current population fitness: σ is updated in order to enlarge or reduce the search space, in accordance if the current fitness is higher or lower than the previous one. In this way, the dimension of the current search space is updated dynamically according to the proximity to the solution, assessed in terms of current value of fitness.

B. Software Quantum Observation

In a quantum computer, the act of observing a quantum state turns it to collapse in a single state. In the QIC, for the software implementation of a quantum observation, from a quantum population a genetic population is generated by following a probability-based approach (Fig. 3). For the homologue q-bits of the population, the σ are summed: in this way, a cumulative distribution function (CDF) of probability of finding a coordinate of the solution is determined. Then, the corresponding genetic chromosome is obtained by generating a random number, between 0 and 1, and determining the corresponding value of ρ. This represents a random choice of a sample in the quantum space under analysis, called quantum observation.

At a procedural level:

a) first, homologue chromosomes are selected for the entire quantum population (P_1 and P_2 in Fig. 3a);
b) theirs pulses are summed up to realize the correspondent PDF (Fig. 3b). PDF shape shows the spatial distribution of quantum chromosomes; i.e. for quantum pulses thicken around a point, the correspondent PDF shows a peek in its shape;
c) this PDF is integrated to obtain a CDF (Fig. 3c) in order to compute the genetic chromosomes: a random number y is generated (y_1 in Fig. 3c), between 0 and 1, to random select a classical chromosome in the complete variation range of ρ;
d) and finally the corresponding value on the CDF x axis (x_1 in Fig. 3d) is chosen as the value for the classic chromosome.

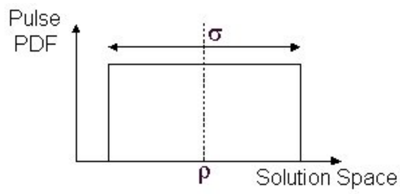

Fig. 2 – Pulse representing a quantum chromosome (q-bit).

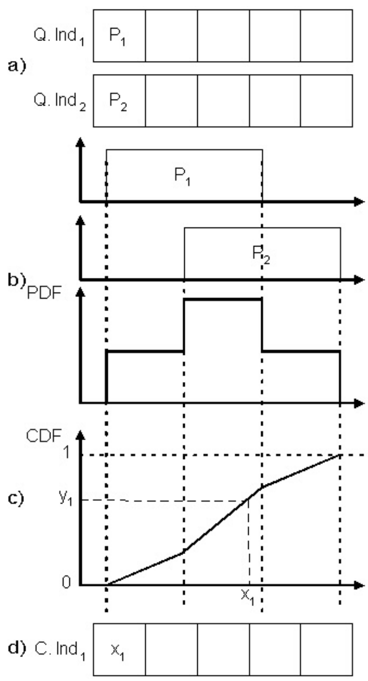

Fig. 3 – QIC procedure of software quantum observation.

C. The Quantum-Inspired Classifier

In Fig. 4, the structure of the QIC is represented schematically. Initially, a random quantum population is generated; then, a classical population is derived through the abovementioned operations of quantum observation. This population evolves by means of traditional evolutionary functions (such as crossover, mutation, and so on). The classical evolution gives rises to a new population, where individuals are classified according to their fitness. If the best fitness exceed the fitness limit imposed, the QIC reaches the convergence and ends with the corresponding individual as the problem solution. Conversely, a new quantum population is rebuilt starting from the current classical one: (i) ρ is obtained from the classical chromosome value; (ii) if the mean fitness of classical population shows some improvement after last evolutionary cycle, the new width pulse σ is reduced; otherwise, σ is enlarged. After the realization of the new quantum population, a new observation is done, such as described previously, in order to obtain the new classical population to be evolved until a final convergence.

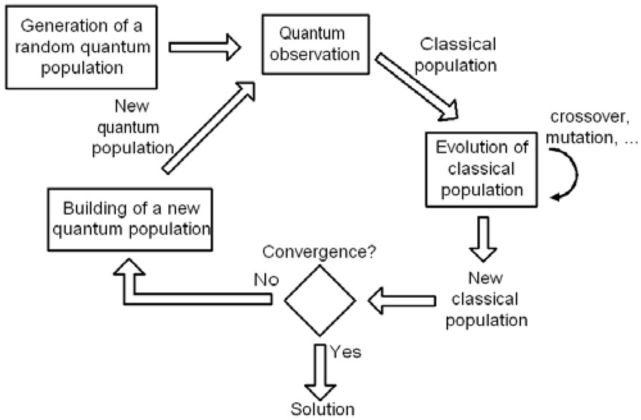

Fig. 4 – QIEA evolution based on classical population.

C. Driving Sequence Classifier

The proposed QIC has been applied to the classification of the drive SEQs collected in the framework of the ATENA Project [3]. According to the procedure set up by the Istituto Motori of the Italian National Research Council for measuring emission factors through a statistical approach [3], each SEQ was segmented in the 14 features described in Tab. 1. Features belonging to the same class were tested according to the Pearson analysis of distribution and they resulted to be modeled mainly by a Beta distribution [14], with α and β characteristic parameters Then, a 3-D matrix (3DM) was defined (Fig. 5), with the feature Fe_j in the *j-th* row and the class C_i in the *i-th* column. The cell c_{ij} of 3DM contains the two Beta distribution parameters (α, β), representing the likelihood Fe_i for a SEQ feature to belong to a class C_j. QIC is employed to optimize 3DM in order to achieve the best classification for SEQs under analysis. The best 3DM is the matrix that, for an unknown input SEQ, gives its right classification according to the reference SEQs classification described in the framework of ATENA Project. Moreover, the likelihood of the SEQ to belong of the class is also provided. For a feature array of an unknown SEQ, a corresponding 2D probability matrix of the Beta distribution of belonging to classes is obtained from the matrix 3DM (Fig. 6). Then, by computing the averages on the columns from the 2D matrix, a row array is obtained, expressing for the unknown SEQ the probability of belonging to classes. Finally, SEQ belongs to the class with the maximum likelihood. The better 3DM is the matrix that, for a given set of SEQs with theirs classifications, gives rise to selective 2D matrix, such as depicted in Fig. 6: only a peak for the expected class for the SEQ under analysis.

Tab. 1 – SEQ features according to CNR Istituto Motori classification [3].

Variable	**Description**	**Variable**	**Description**
V_mean	Mean speed	*Acc1t*	Time with acceleration > 0.15 m/s^2
V_a	Sum of speed samples multiplied by acceleration samples	*Paccel1*	Time percentage with acceleration < -1.4 m/s^2
V_max/t	Ratio between the maximum speed and the driving time	*Paccel2*	Time percentage with acceleration ϵ[-1.4;-0.6] m/s^2
V20	Time percentage with vehicle speed < 20 km/h	*Paccel3*	Time percentage with acceleration ϵ[-0.6;-0.2] m/s^2
V30	Time percentage with vehicle speed ϵ[20, 30] km/h	*Paccel4*	Time percentage with acceleration ϵ[-0.2;0.2] m/s^2
V40	Time percentage with vehicle speed ϵ[30, 40] km/h	*Paccel5*	Time percentage with acceleration ϵ[0.2;0.6] m/s^2
T_sum	Duration of the SEQ	*Paccel6*	Time percentage with acceleration ϵ[0.6;1.4] m/s^2

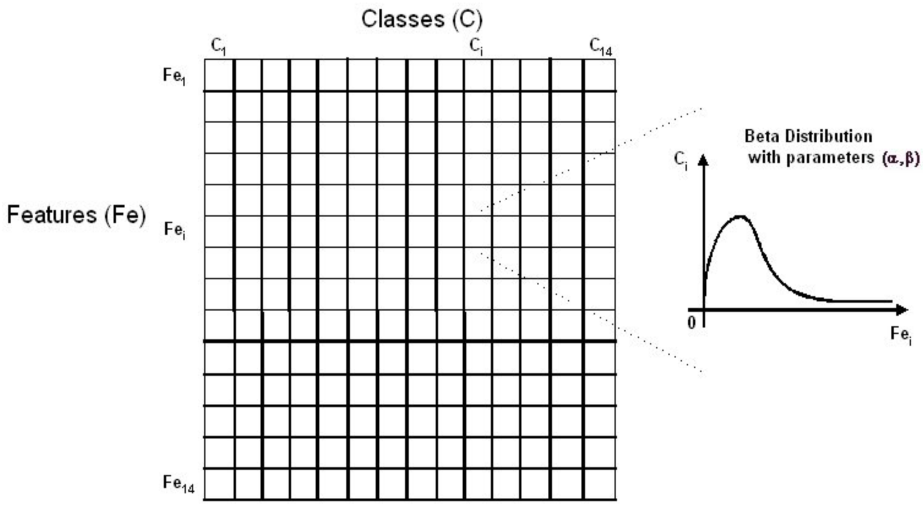

Fig. 5 – 3D matrix 3DM.

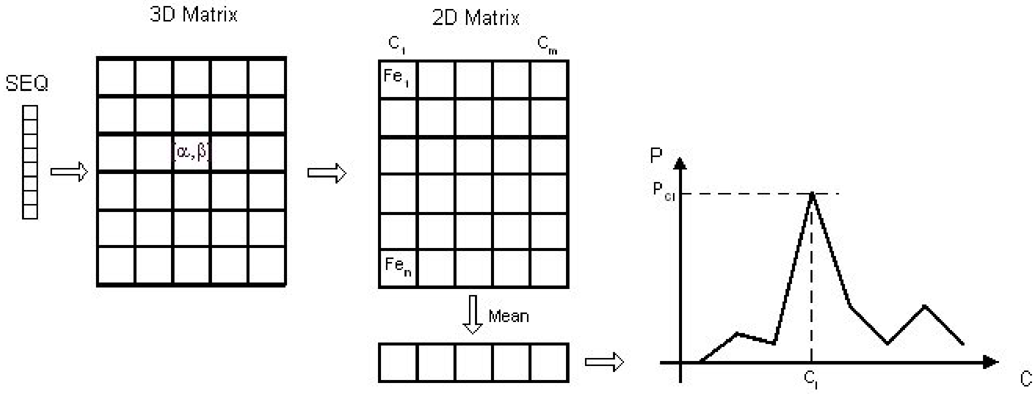

Fig. 6 – SEQ classification likelihood via the 3D Matrix.

The driving sequence classifier operates according to the following procedure:

- Firstly, a 3DM is randomly realized.
- Then, for each SEQ of the input set with its class, a quantum population is realized. The generic quantum individual is composed by 14 chromosomes, as the number of SEQ features. The *i-th* chromosome is composed by the couple (α,β) representing the Beta distribution associated to the *i-th* SEQ feature for these class. Each Beta distribution parameter is composed by two terms, ρ and σ, such as previously described.
- From this quantum population, a classical population evolving as a genetic algorithm is realized. Classical individuals represents possibly problem solutions, particularly they represents a possibly 3DM column related to the class of the SEQ under analysis.
- Best individual is chosen according to the best classification for the input SEQ.
- The procedure continues until all SEQs are analyzed and the best 3DM is obtained.

III. Experimental results

A set of thousands classified SEQs experimentally acquired in the framework of the ATENA Project were employed [3]. QIC results were compared with solutions proposed by a traditional reference GA. The algorithm configurations for the comparison tests are reported in Tab.2. GA and QIC are compared on the same basis by selecting the same number of classical individuals, and, above all, the same number of fitness evaluations (i.e number of times that an individual is evaluated by means of the fitness obtained as the product between the number of individuals and the number of cycles).

Tab. 2 – Algorithms configuration.

<table>
<tr><th>Algorithm</th><th colspan="2">QIC</th><th>GA</th></tr>
<tr><td rowspan="2">Number of evolutionary cycles</td><td>Quantum</td><td>Genetic</td><td rowspan="2">600</td></tr>
<tr><td>30</td><td>20</td></tr>
<tr><td>Number of fitness evaluations</td><td colspan="2">60000</td><td>60000</td></tr>
<tr><td>Crossover rate</td><td colspan="2">0.7</td><td>0.7</td></tr>
<tr><td>Mutation rate</td><td colspan="2">0.3</td><td>0.3</td></tr>
<tr><td>Elite count</td><td colspan="2">1</td><td>1</td></tr>
</table>

Tab. 3 – Effectiveness and efficiency in terms of percentage of classified SEQs.

<table>
<tr><th></th><th colspan="4">QIC</th><th colspan="2">GA</th></tr>
<tr><td></td><td colspan="3">Cycles</td><td>Fitness (%)</td><td>Cycles</td><td>Fitness (%)</td></tr>
<tr><td rowspan="2">Effectiveness
(same cycles number)</td><td>Quantum</td><td>Genetic</td><td>Tot.</td><td rowspan="2">81.36 %</td><td rowspan="2">600</td><td rowspan="2">10.83 %</td></tr>
<tr><td>30</td><td>20</td><td>600</td></tr>
<tr><td rowspan="2">Efficiency
(same fitness)</td><td>Quantum</td><td>Genetic</td><td>Tot.</td><td rowspan="2">81.36 %</td><td rowspan="2">NaN</td><td rowspan="2">81.36 %</td></tr>
<tr><td>25</td><td>20</td><td>500</td></tr>
</table>

Others parameters significant for the comparison represented in Tab. 2 are: (i) crossover rate, as the percentage of individuals that can be used to cross theirs genomes; (ii) mutation rate, as the percentage of individuals that can be mutated; and (iii) elite count, as the number of individuals that can pass from a generation to the following without any change in its genome.

Performance, represented with 3DM final fitness, are analyzed as percentage of correctly classified SEQs. In Tab. 3, experimental results show interesting QIC performance versus GA, represented as algorithms effectiveness and efficiency, according to literature of evolutionary algorithms [18].

Effectiveness is a measure of the quality solution within a given computational limit, 600 total evolutionary cycles in this case. The number of QIC total evolutionary cycles are calculated by multiplying internal cycles – cycles done by the internal GA – and external cycles – cycles done by the means of the quantum conversion. At the end of the same number of cycles, QIEA shows better results than GA, with a difference of SEQs classified of + 70.53 %.

Efficiency is a measure of the amount of computing needed to achieve a satisfactory solution set at 81.36% in this case. After a large amount of evolutionary cycles, GA was not able to reach the defined target.

III. Conclusions

A quantum-inspired classifier for the classification of driving sequences in vehicle emission factor measurement was proposed. Evolutionary mechanisms are modulated by principles of quantum physics in order to explore solutions in the search space simultaneously with theirs neighborhoods in a probabilistic way instead of deterministic classical way, such as employed in genetic algorithms. The classification of driving sequences acquired in the framework of the ATENA Project [3] showed interesting performance of QIC compared with classical genetic algorithm results, in terms of effectiveness and efficiency.

Nevertheless, fitness values, as classified SEQs percentage, as well as time computation, must to be improved and future works will be focused on code optimization.

References

[1] André M., "Real-world driving cycles for measuring cars pollutant emissions - Part A: The Artemis European driving cycles", *INRETS report*, Bron, France, n°LTE 0411, 97 p. 2004.

[2] Cox J., "The Australian Diesel Fleet-Existing Vehicle Characteristics and the Modelling of Transport Demand, Vehicle Populations and Emissions," National Environment Protection Council Service Corporation, pp. 1-109, November 1999.

[3] L. Gortan, L. Mina, A. Fortunato, L. Borgarello L. Della Ragione, G. Meccariello, M. V. Prati, M. Rapone, "Preliminary results on emission and driving behaviour of ATENA fleet test project in Naples", *Proc. of the 5th International Conference on Internal combustion engines* – ICE2001, Capri (Napoli), 2001, pp.52-60, SAE_NA Technical Paper 2001-01-083.

[4] J. A. Miller, W. D. Potter, R. V. Gandham, and C. N. Lapena, "An evaluation of local improvement operators for genetic algorithms," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 23, no. 5, pp. 1340-1351, Sept.-Oct. 1993.

[5] X. Jin, R. G. Reynolds, "Data mining using cultural algorithms and regional schemata", *Proc. of 14th Int. Conf. on Tools with Artificial Intelligence*, Washington DC, 4-6 November 2002, pp. 33-40.

[6] G. B. Fogel, D. B. Fogel, "Continuous evolutionary programming: analysis and experiments", *Cybernetics and Systems*, vol. 26, pp. 79-90, 1995.

[7] W. Gao, "Comparison study of genetic algorithm and evolutionary programming", *Proc. of the Third Int. Conf. on Machine Learning and Cybernetics*, Shangai, 26-29 August, 2004, pp. 204-209.

[8] R. G. Reynolds, W. Sverdlik, "Problem solving using cultural algorithms", *Proc. of the First IEEE Conf. on Evolutionary Computation*, Orlando, Florida, 27-29 June 1994, pp. 645 – 650.

[9] J. G. Digalakis, K. G. Margaritis, "A multipopulation cultural algorithm for the electrical generator scheduling problem", *Mathematics and Computers in Simulation,* vol. 60, pp. 293-301, 2002.

[10] M. Stenberg, R. G. Reynolds, "Using cultural algorithms to support re-engineering of rule-based expert systems in dynamic performance environments: a case study in fraud detection", *IEEE Trans. on Evolutionary Computation*, vol. 1, no. 4, pp 225-243, November 1997.

[11] N. Rychtyckyj, R. G. Reynolds, "Knowledge base maintenance using cultural algorithms: application to the DLMS manufacturing process planning system at Ford Motor Company", *IEEE Proc. Congress on Evolutionary Computation*, Honolulu, USA, 12-17 May 2002, pp 855-860.

[12] S. Saleem, R. G. Reynolds, "The impact of environmental dynamics on cultural emergence", *Festschrift, in honor of John Holland*, Oxford University Press, 2003, pp.1-10.

[13] P. Arpaia, C. Romanucci, and A. Zanesco, "Evolutionary Pattern Recognition for Measurement of Vehicle Emission Factors in Critical Driving Conditions", *Proc. of the 18th IMEKO World Congress*, Rio de Janeiro (Brasil), 17-22 September 2006.

[14] A. Andreev, A. Kanto, P. Malo, "Simple approach for distribution selection in the Pearson system," Helsinky School of Economics, Working Paper W-388, October 2005, pp. 1-22.

[15] L. K. Grover, "A fast quantum mechanical algorithm for database search," *Proc. of the twenty-eighth annual ACM symposium on Theory of computing*, Philadelphia, Pennsylvania, USA, 22-24 May 1996, pp. 212-219.

[16] A. Arpaia, A. Zanesco, "Automatic diagnostics in industrial measurement systems by quantum-inspired evolutionary algorithms," *8th International Symposium on Measurement Technology and Intelligent Instruments*, ISMTII 2007, Sendai, Japan, 24-27 September 2007, *accepted*.

[17] A. V. Abs da Cruz, M. M. B. R. Vellasco, M. A. C. Pacheco, "Quantum-Inspired Evolutionary Algorithm for Numerical Optimization," *Proc. of the 2006 IEEE Congress on Evolutionary Computation, World Congress on Computational Intelligence (WCCI)*, Vancouver, Canada, 16-21 July 2006, pp. 2630-2637.

[18] A. E. Eiben, J. E. Smith, *Introduction to Evolutionary Computing*, Springer, 2003, ch. 14.