



**HAL**  
open science

# User tests for the convergence of a car silhouettes design model using an Interactive Genetic Algorithm

François Cluzel, Bernard Yannou, Markus Dihlmann

## ► To cite this version:

François Cluzel, Bernard Yannou, Markus Dihlmann. User tests for the convergence of a car silhouettes design model using an Interactive Genetic Algorithm. IDMMME - Virtual Concept 2008, Oct 2008, Beijing, China. Virtual Concept\_P104. hal-00796963

**HAL Id: hal-00796963**

**<https://hal.science/hal-00796963>**

Submitted on 5 Mar 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# User tests for the convergence of a car silhouettes design model using an Interactive Genetic Algorithm

François Cluzel<sup>1\*</sup>, Bernard Yannou<sup>1</sup> and Markus Dihlmann<sup>2</sup>

<sup>1</sup> Ecole Centrale Paris, Laboratoire Genie Industriel, Chatenay-Malabry, France

<sup>2</sup> University of Stuttgart, Institute of Applied Analysis and Numerical Simulation, Stuttgart, Germany

*Abstract: We propose in this paper a car silhouettes design model based on an Interactive Genetic Algorithm. The design solutions are indirectly encoded by a Fourier analysis approach. Thus after creating an initial population, the designer is able to evaluate himself generations of car profiles, in order to converge towards an individual which meets its requirements. An automatic test has already been conducted, but it is not sufficient to validate the model; we propose here 6 new tests. The two first ones are based on the user satisfaction towards the created design solution. The data were collected during a short subjective evaluation workshop. The 4 other tests use a similarity index (that quantify the perceived distance between 2 individuals) to process the same data and validate the model. These results reveal a really satisfactory behaviour and open perspectives on the understanding of the user perception.*

*Keywords: Interactive Genetic Algorithm; Evolutionary Design; car profile; user test; subjective evaluation.*

## 1- Introduction

A recurrent observation in Design is the lack of CAD (Computer Aided-Design) tools for the first phase of the design process [C1] [DP1]. At this step, most of the work actually consists in drawing 2D sketches to create new and innovative products. The existing CAD tools often offer wide possibilities, but later in the design process. These tools are indeed based on geometric constraints and clearly defined form, which is not really compatible with the human and sometimes artistic refinement process that leads to 2D sketches. Thus the ideal aiding tool should be able to take into account the user perception to permit an easy convergence towards a design solution that really meets the designer own requirements.

In this study field, the theory of Evolutionary Design (ED) seems very promising. Genetic Algorithms (GA) [BC1] [RE1] are an application of ED. Originally used for complex optimisation problems, GAs work on the basis of an initial population. Each individual of this population is a design solution and its structure and apparent qualities are called *phenotype*. Each phenotype is encoded by a *genome*. Through recombination of these genomes during an operation called *crossing over*, the individuals can reproduce themselves and produce new individuals. So with successive generations, the GA displays better and better design solutions. The principle is the same for an Interactive Genetic Algorithm (IGA,

---

\* **Contact author**

E-mail: [francois.cluzel@ecp.fr](mailto:francois.cluzel@ecp.fr)

Phone: +33 1 41 13 13 28

see [KC1]) but here the user plays an active role in the process. He/she performed himself the choice of the individuals which will be used to create the next generation. So IGAs are then particularly adapted to our situation, where an aesthetic evaluation of the individuals is required and the user perception is essential.

Two conference papers have already been written to describe the principle of the encoding and the working of the IGA [YD1] [YD2]. So in this paper we first present a summary of these publications to clearly describe the model.

Several methods are commonly used to encode the genome of an individual, but they often have the disadvantage to be conceived for a given design domain. The proposed encoding is based on the Fourier decomposition of a 2D-closed-curve and can be applied to all possible objects represented by their 2D silhouettes. Our IGA is based on this encoding. After creating an initial population, the IGA starts. The interactive part takes place at each generation, where the style designer qualitatively assesses individuals. After the evaluation, the crossing over operation is automatically performed to create new design solutions.

In this way, innovative and new concepts are expected from the collaboration between an automatic process of design space exploration and the interaction of a designer. But validations are necessary. A first automatic test has been proposed in the previous papers [YD1] [YD2], using a similarity index. The aim of this indicator is to quantify the perceived distance between 2 individuals. However our model is supposed to be interactive, and an automatic test cannot answer all the following questions:

- Are the users satisfied by our model?
- Are the results better with our model than without?
- Does our model really create novelty?
- Are the results different from a user to another?

So user tests seem to be essential to prove that innovation and surprise may emerge from this process. Thus we performed a subjective evaluation workshop to collect user data. Then the results of 2 main tests and 4 post-processing tests using the similarity index have been established. They show a very satisfactory behaviour of the model in terms of convergence, diversity, dependence to the initial population, but they also show interesting results about the user perception.

The paper presents in section 2 the process of the Interactive Genetic Algorithm with the different operations required for encoding a car profile, generating an initial population and combining the genomes. In section 3, a similarity index is proposed to measure the perceived distance between two individuals. Section 4 deals with the subjective evaluation workshop that was carried out to validate the model, whereas other tests based on the similarity index are explained in section 5 to study the model's behaviour. Finally, section 6 concludes on some forthcoming perspectives.

## **2- Model description**

This part has been developed in more details in 2 previous papers (see [YD1] and [YD2]). However, the model description is necessary to understand the next parts of this paper.

### **2.1 – Encoding of the car profiles**

A major difficulty when using GAs in automatic design systems is the encoding of the genome (see [NA1]). Most systems use a direct encoding where geometrical dimensions and structures of the design object are directly represented in the genome [AC1] [YF1]. Other works use tree structures [LT1] or shape grammars [OC1], or even a set of predefined components of the studied product [KC1].

But the main disadvantage of all these systems is that they are conceived for a given design domain. Implementing these methods in new fields of design is a difficult and time consuming process.

We propose here a new encoding method for a 2D-closed-curve based on the works of McGarva [MM1]. He has proposed its development into a Fourier series as a method for coding its phenotype. We have personally already used this theory in [VY1] for encoding a 2D-closed-curve into the five first Fourier harmonics of this decomposition. The main advantages of such a method are numerous. Indeed almost all possible 2D-closed-curves may be generated, every detail may be represented (which is really important to stimulate emotions), the genomes have a constant length (which is simpler for the crossing over operation), and last but not least, the genes (the Fourier harmonics) are narrowly associated to apparent characteristics of the phenotype.

McGarva's theory of Fourier decomposition of a closed curve considers that the position of each point belonging to this curve can be expressed by a complex function in the complex plane:

$$z(t) = x(t) + i y(t) \quad (1)$$

As  $z(t)$  is a closed curve, its function is periodic. The period is normalised with:  $z(t+1) = z(t)$ . This function  $z(t)$  can be developed into a Fourier series, according to formula (2):

$$z(t) = \sum_{m=-\infty}^{\infty} a_m \exp(2\pi i m t) \quad (2)$$

where the complex Fourier coefficients can be calculated by formula (3):

$$a_m = \int_0^1 z(t) \exp(-2\pi i m t) dt \quad (3)$$

Coefficient  $a_0$  is called fundamental,  $a_1$  and  $a_{-1}$  represent the first harmonic,  $a_2$  and  $a_{-2}$  the second harmonic, etc.

As we will see later, the function  $z(t)$  is not known as an explicit function from the beginning. Instead, we assume that the curve has been initially defined by a set of successive points  $z_k$  ( $k=0, \dots, N$ ) which belong to the curve. So, in order to calculate the  $a_m$  coefficients (3) we need a numeric approximation. We obtain this approximation by dividing the curve into  $N$  segments connecting each point with its successor. We call  $t_k$  the length of the curve between the first point  $z_0$  and the point  $z_k$ . Under these conditions the integral can be calculated by the trapezium formula:

$$a_m = \sum_{k=0}^N \left( \frac{t_{k+1} - t_k}{2} (z_{k+1} \exp(-2\pi i m t_{k+1}) + z_k \exp(-2\pi i m t_k)) \right) \quad (4)$$

while  $z$  is a periodic function, ( $z_{N+1} = z_0$ ).  $t_k$  is defined by:

$$t_k = \frac{L_k}{L} \quad (5)$$

where the total length  $L$  is the sum of the lengths of all segments,  $L_k$  is the length from the origin to the current point, with  $x_{N+1} = x_0$  and  $y_{N+1} = y_0$ .

To construct the genome of an object, we develop its silhouette into Fourier series and define the fundamental (the coefficient  $a_0$ ) as gene number zero. The first harmonic ( $a_1, a_{-1}$ ) will be called the first gene, the second harmonic the second gene, etc.

On the basis of the genome, the original shape of the individual can be reconstructed. Every point  $P_k$  with the coordinates  $(x_k, y_k)$  on the curve  $z^*$  which approximates the silhouette of the car, can be calculated by formula (6).

$$z^*(t_k) = x_k + iy_k = \sum_{m=-p}^p a_m \exp(2\pi i m t_k) \quad (6)$$

where  $t_k$  ( $0 \leq t_k \leq 1$ ) is the position on the curve and  $p$  fixes the number of harmonics used for the decoding. When  $p$  equals 1 for example, we use one harmonic to reconstruct the silhouette of the car. The more harmonics used for the decoding the more precise will be the approximation to the original curve (as seen in figure 1). We call  $p$  the “precision” of decoding.

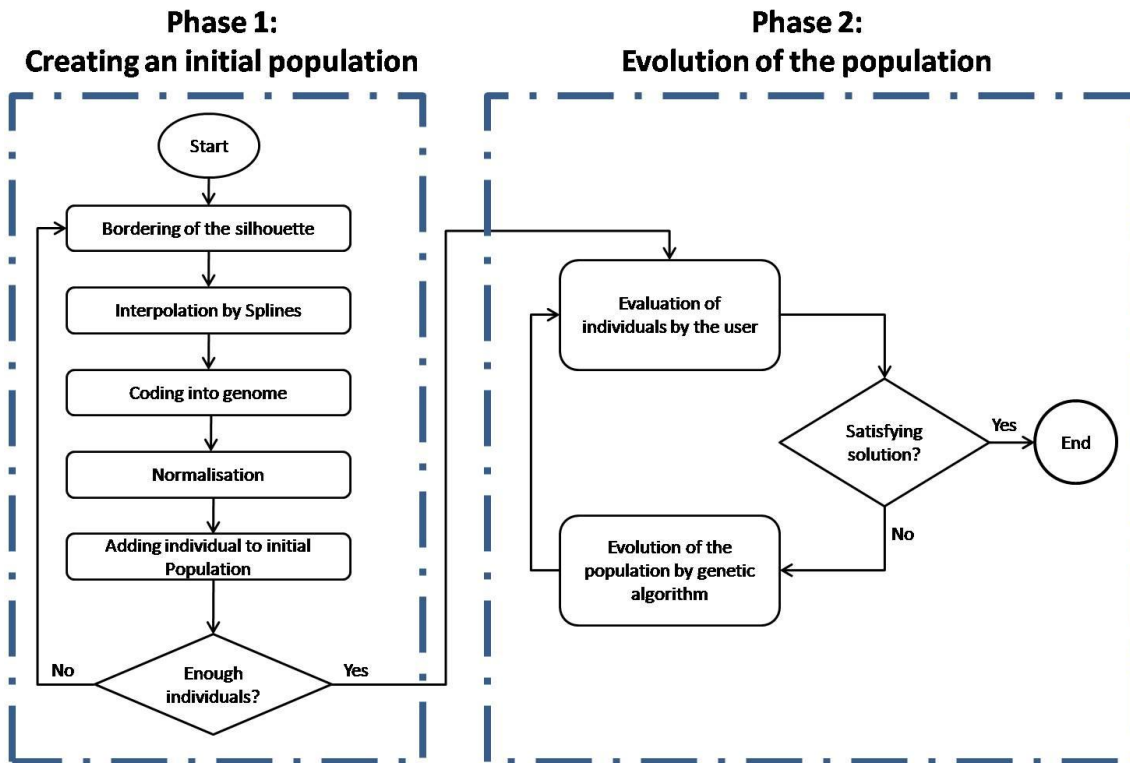


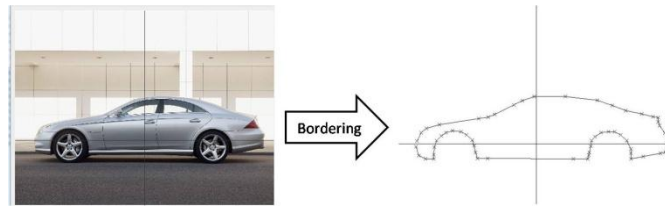
Figure 1: diagrammatic plan of the IGA process.

## 2.2 – Creation of a population

The first step of the IGA process is the creation of an initial population (phase 1 on figure 1). Then the population evolves in phase 2 (see figure 1), explained in section 2.3.

The initial population consists of silhouettes of several (20 for the user tests in sections 4 and 5) already existing car bodies. In order to easily sketch these silhouettes we programmed an interface in Java which allows to border a silhouette from a photograph by clicking on the screen and to encode it (see figure 2). During bordering, a sufficient amount of points should be used to represent as many details as possible. When using a number  $M$  of 60 to 80 points per silhouette the result is satisfactory.

However it is not sufficient to calculate a genome which is precise enough to allow a highly detailed decoding into the phenotype. Consequently we need to augment the number of points on the curve by smooth interpolations. Bicubic splines linking three successive points have been chosen, because this method provides a curve which is very close to the original curve without producing oscillations. Within each spline, a given number of points are interpolated, leading to a total number of  $N$  points with  $N > M$ .  $N$  has an influence on the precision of the Fourier coefficients.



**Figure 2: After bordering we obtain a closed curve representing the car silhouette of an existing car.**

The number  $p$  of harmonics used to decode the genome has a strong influence on the final result too. So it is necessary to determine a satisfactory balance between  $N$  and  $p$ .

After qualitative experiments, we found out that a satisfactory choice was achieved with a genome size  $p$  of 71 and a number  $N$  of approximately 1500 points for the interpolation since both initial and resulting silhouettes were visually identical.

A last operation of normalization is necessary in order to obtain car silhouettes that are independent of a particular location or size. The coefficient  $a_0$  is simply set to 0 to fix the centre of gravity of all individuals at the origin of the representation plane. To normalize the size of the curve we chose to simply have a surface area invariance.

### 2.3 – Evolution of a population

The population of a Genetic Algorithm evolves during the crossing over operation. In our case the individuals of the populations are car silhouettes. They reproduce themselves to create new individuals. A fitness value  $f$  between 0 and 6 (where 0 is the worst and 6 the best evaluation) is assigned to each individual, according to the grade given by the user via an interface. It is important to notice that the meaning of the scale is chosen by the user himself. No visual connection is given. It permits a high liberty in the design process, but it may sometimes have bad effects, because the user can change his/her own scale during the evaluations.

The interface developed in Java (see figure 3) displays six individuals at a time and the user can browse through all the individuals of a population. The user is supposed to evaluate all the individuals of a population. Four operators have been implemented:

- **Selection:** decides which individuals will reproduce and create children according to the fitness values.
- **Crossover:** builds a child's genome from two selected genomes (called parents).
- **Mutation:** changes in a random way a genome after the crossover (not detailed here).
- **Killing:** decides which individuals from the parents' population will survive and live on in the new generation (not detailed here).

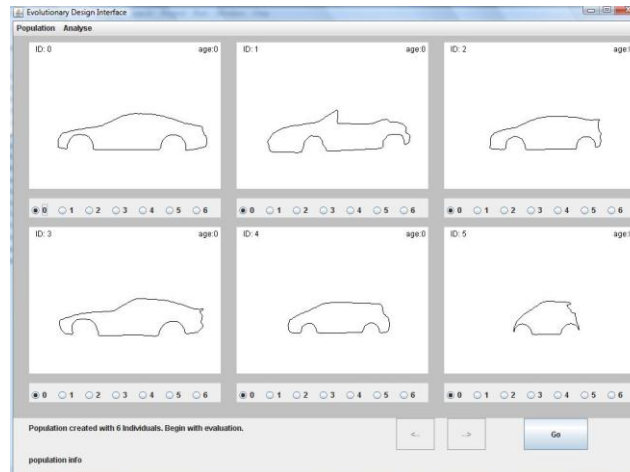
We decided to adopt some conventional choices in terms of *selection* and *killing* operators and to propose an original crossover operator. We chose a turnover rate of 0.7, meaning that, for a coming generation, 30% of the individuals are kept from the previous one and 70% of children are generated. In this way we do not lose potential good design solutions. The probability for an individual to be selected to be a parent is proportional to its fitness value (between 0 and 6). After choosing two individuals from the parents' population, their genomes are combined into the genome of a child by applying the crossover and the mutation operators. Afterwards the two individuals are re-put into the parents' population. Indeed, an individual can be selected more than once by the selection operator.

We envisaged several possibilities to crossover the two genomes of parents into the one of the child, but the best tried idea is to operate a weighted mean between the gene values of the two parents to build the genome of the child. A crossover weight  $W$  is chosen randomly between 0 and 100. A new

gene  $g^*$  is formed by calculating the weighted mean of the genes  $g_{m,1}$  and  $g_{m,2}$  of the parents after formula (7).

$$g_m^* = \frac{Wg_{m,1} + (100-W)g_{m,2}}{100} \quad (7)$$

In function of the weight  $W$  we obtain different new design solutions which continuously interpolate a silhouette between the two parents' silhouettes. Thus children have visual characteristics of the two parents and no useless car is produced, but only a small space of possible solutions is explored. To enlarge the space of possible solutions we must apply a mutation operator (not detailed here).



**Figure 3: User Interface for the designer evaluation showing 6 individuals of a larger population. The designer can browse through the population by clicking on the arrow buttons.**

### 3- Similarity Index

#### 3.1 – Requirement

Is our system really capable to produce innovation and novelty? Is it possible for a user to design with the help of our system a new car body silhouette which was not part of the initial population?

A simple test permits to have a first idea of the answer, before performing a user tests workshop. A designer draws on a sheet of paper a car body silhouette which comes spontaneously to his/her mind and which is not part of the initial population. This car body silhouette is taken as “reference individual”. By working with our system he should try to obtain in the end the silhouette he had drawn before on the paper. To cope with this, he is supposed to evaluate the car solutions which look close to the reference individual with higher grades and those who look different with lower grades. By counting the number of generations he needs to reach the reference individual, we can estimate the quality of our design system.

But this method is highly linked to the designer subjectivity, which would probably influence the results. Our objective is just to estimate the convergence ability of the model. If no automatic convergence is observed, there will probably not be a convergence with any user test. So we have preferred to perform an automatic test. The target individual is a car silhouette that has been removed from the initial population. The role of the designer is played by an algorithm, which automatically evaluates the individuals of a generation in terms of their similarity to the target individual. For that purpose, we defined a *similarity index* between two individuals.

The idea is to work with an indicator of the perceived distance between two individuals. Initially, this index has been built for the test described in this section, but we will see in section 5 that it is a useful tool to work on the user data too.

### 3.2 – Description

The difference between two genomes  $G_k$  and  $G_l$  is given by  $D(k,l)$  in formulas (8):

$$D(k, l) = \sum_{m=1}^{70} \alpha(m) \| g_{k,m} - g_{l,m} \|^2 \text{ with} \quad (8)$$

$$\| g_{k,m} - g_{l,m} \|^2 = (u_{k,m} - u_{l,m})^2 + (u_{k,-m} - u_{l,-m})^2 + (v_{k,m} - v_{l,m})^2 + (v_{k,-m} - v_{l,-m})^2$$

with  $a_m = u_m + i v_m$ , and  $\alpha(m)$  a weight factor which should give more importance to the first genes than to the last genes because it is more easy to perceive. Formula (9) gives an approximation of  $\alpha(m)$  from user evaluations:

$$\alpha(m) = 0,111 e^{-0,1m} \quad (9)$$

Finally we define the *similarity index* between two genomes  $k$  and  $l$  as:

$$SimInd(k, l) = \frac{100}{1 + \frac{1}{N} D(k,l)} \% \quad (10)$$

where  $N$  is a normalisation factor.

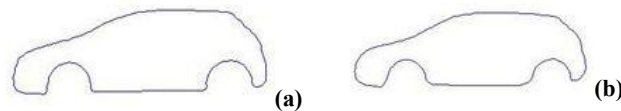
To understand the link between the reality and this tool, table 1 proposes a range of values and their visual meaning on car silhouettes.

**Table 1: visual meaning of the similarity index.**

< 10%	<i>No connection</i>
20% to 30%	<i>Very poor similarities</i>
30% to 50%	<i>Similarities</i>
50% to 80%	<i>Very strong similarities</i>
80% to 100%	<i>Imperceptible differences</i>

### 3.3 – Automatic convergence test

For the test we used the car in figure 4a as reference individual. The parameters for the Genetic Algorithm were the following: population of 100 individuals, turnover rate of 0.7 and mutation probability of 0.3. The mutation could change a gene in a range of  $\pm(50\%-200\%)$ . After 10 generations our system reached the car body silhouette in figure 4b which has a similarity index of 92%, which can be considered as a much satisfactory result.



**Figure 4: Comparison between the reference silhouette (a) and the final resulting silhouette (b).**



The average fitness of the population converges over the generations to a high value (more than 95%), whereas the value of the best similarity index in the population (the fitness of the fittest individual) raises rapidly from relative low 44% to 92%.

So our system is able to automatically and quickly (less than 10 generations) converge towards a given car silhouette that is not in the initial population. It means that the space of possible solutions which can be reached by our system is large enough. This is necessary to be able to produce innovation and novelty, but not sufficient. The problem is now to validate the model with regard to the user perception, and to answer the following questions:

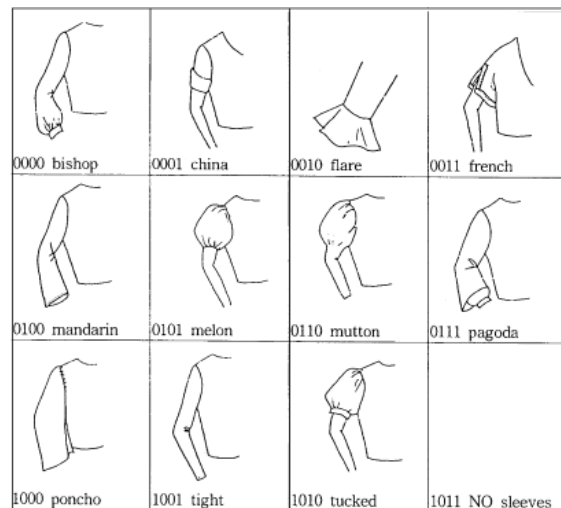
- Are the users satisfied by our model?
- Are the results better with our model than without?
- Does our model really create novelty?
- Are the results different from a user to another?

That is why we propose in section 4 two new tests based on the users' subjective evaluation and on the works of Kim [KC1]; 4 new post-processing analyses will also be performed in section 5, using the similarity index.

## 4- Validation tests

### 4.1 – Tests construction

Kim worked in 2000 on fashion design with an Interactive Genetic Algorithm [KC1]. He proposed a model to create innovative dresses from a catalogue of components (predefined necklines, sleeves, skirts... See figure 5). The main difference between our model and Kim's one is the space of possible solutions. As Kim worked with a set of predefined parts, which is a finite and discrete set, this space is limited, whereas our possible solution space is infinite and continuous.



**Figure 5: examples of possible choices for arms and sleeves in Kim's fashion design model (taken from [KC1]).**

To validate his model, Kim performed two different tests.

The first test is called *convergence test*. It permits to ensure that the average fitness value according to each generation of the Genetic Algorithm increases with the generations. He made the test along 10

generations composed of 8 dresses. He used a panel of 10 users, working on two semantic attributes (*splendid* and *cool-looking*), that means that he asked the users to evaluate 10 generations of dresses according to the first attribute, and then to do the same process with the second attribute. The results showed a significant increase of the fitness.

The second test is called *subjective test*. Its goal is to show that the results obtained with the model are better than without. Kim asked 3 users to find the 10 best dresses according to the two semantic attributes (*splendid* and *cool-looking*) among 500 individuals randomly created from the catalogue of permitted combinations (i.e. without his IGA). Then he asked the 10 previous users to find the best dress of the 10<sup>th</sup> generation of the previous *convergence test*, according to the two attributes. The last stage was to compare for each attribute and for each user their own best dress obtained with the model with the 10 best dresses obtained without. The comparisons were made by pairwise comparisons, on a 7 degrees scale (from -3 to 3). Finally, the results showed that the individuals obtained with the algorithm reached on average a degree of about 2, compared to the individuals obtained without algorithm. So the model was really satisfactory.

But two main criticisms can be formulated on these two tests:

- Are the evaluations hedonistic or not? A hedonistic evaluation includes the preference of the user (for example "*Please rate these individuals according to your preferences in terms of sportiveness*"; the user maybe likes sportive individuals, but not too much), whereas a non-hedonistic evaluation does not (for example "*Please rate these individuals according to their apparent degree of sportiveness*").
- What is the meaning given by Kim on the pairwise comparisons? He asked the users to compare their own best dress (from the algorithm) with the 10 best dresses found by 3 other people. So it seems logical that the first one (found by the user himself) is better than the others (found by other people).

We built our test according to these criticisms.

## 4.2 – Workshop realization

We have chosen to work for the two tests with a non-hedonistic evaluation ("*Please rate these individuals according to their apparent degree of sportiveness*").







We have also chosen two semantic attributes: *friendly* and *sportive* (half of the users work with the first one, the others with the second one).

The first test that we have defined is the same that the *convergence test* defined by Kim. We called it test of "*designer satisfaction*". We have worked with 10 generations composed of 20 individuals, and with a mutation probability of 0.05 and a selection rate of 0.7.

We have defined the second test (that we called test of "*satisfaction superiority of the IGA model*") with regard to the problem of the pairwise comparisons (expressed in Section 4.1). The process is the following:

- 400 car silhouettes are created randomly from an initial population of 20 real individuals. To clarify the notations in the next parts of the paper, these 400 individuals are called *paper individuals*. Each user has to find the 3 bests individuals according to his/her semantic attribute. Practically, these individuals are printed on paper and numbered, and the users record their evaluation in an Excel according to the following process:
  - Each user evaluates the 400 paper individuals on the same scale as in the IGA (from 0 to 6). The best car silhouettes in his/her mind have to receive a 6.

- Only the individuals which receive a 6 are selected for the next evaluation. An Excel macro sorts out these individuals and a new evaluation table is presented to each user.
- The two first stages are repeated until each user finds the 3 best individuals. That means that the users must be more and more selective (to always have scores between 0 and 6 in order to progress).
- Then each user has to find the 3 best individuals of the 10<sup>th</sup> generation of the test of “*designer satisfaction*”.
- Then a pairwise comparison [YE1] matrix permits to each user to compare pair to pair his/her own 3 best individuals of the algorithm with his/her own 3 best individuals obtained without the algorithm. The scale used has 7 degrees (from -3 to 3).
- Pairwise comparisons are explained in more details in [YE1]. They permit to evaluate simply a set of individuals without any absolute scale.
- An example of such a matrix is given on figure 6. The comparisons have a direction to follow: here the individuals in lines are compared to the individuals in columns.

		Best paper individuals		
				
Best individuals from the IGA		>>	>>>	>
		>	>>>	=
		>	>>>	=

**Figure 6: example of pairwise comparison matrix used for the tests.**

The evaluation scale for these comparisons is given in table 2. The users work with the mathematic symbols (>, >>, =...), which are then replaced by numbers to analyze the data (see section 4.3.2).

**Table 2: pairwise comparison scale.**

-3	<<<	highly inferior
-2	<<	inferior
-1	<	slightly inferior
0	=	equal
1	>	slightly superior
2	>>	superior
3	>>>	highly superior

The tests were realized during a short workshop. The panel of users was composed of 7 students (6 men and a woman) and a professor. 4 users received the semantic attribute *friendly* and the 4 others the attribute *sportive*.

About 2 hours were necessary to complete the three stages (time for presentation of the tests included):

- Evaluation of the 400 random printed individuals: 1 hour.
- Use of the Interactive Genetic Algorithm: 30 minutes.
- Pairwise comparisons: 15 minutes.

## 4.3 – Results

### 4.3.1 – Test of “designer satisfaction”

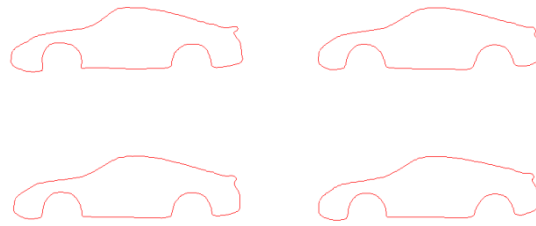


Figure 7: examples of best “sportive” car silhouettes (one from each user).

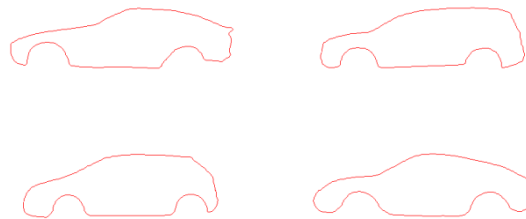


Figure 8: examples of best “friendly” car silhouettes (one from each user).

The analysis of the car silhouettes obtained by the users during the first test shows interesting results.

For the two semantic attributes, the users reach in their own last generation the same type of car silhouettes. Moreover, those profiles are nearly the same for the semantic attribute *sportive*, whereas they are really different from a user to another with the attribute *friendly*. Figures 7 and 8 illustrate this phenomenon. The displayed silhouettes are chosen from the 3 best ones of each user. It is also interesting to notice that the car silhouettes from the 6<sup>th</sup> generation are almost the same for each user, which means that the diversity of the populations is weak from the 6<sup>th</sup> generation. In other words, the IGA seems to converge very quickly.

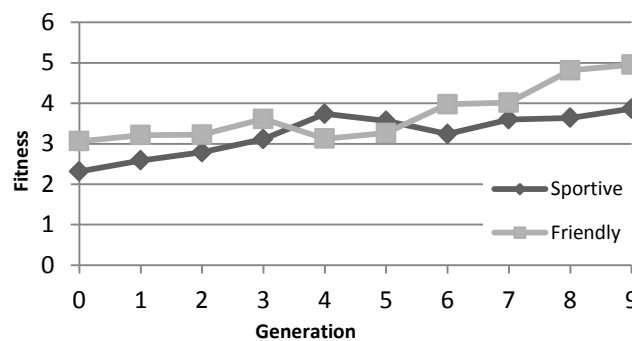


Figure 9: fitness evolution for the convergence test.

It could be interesting to know if this quick convergence is wanted or not. This observation is not tackled in this paper. However, a way to change this speed is the mutation operator. For the tests, a really low mutation rate has been chosen. With a high mutation rate, the convergence speed would probably be lower, because a lot of novelty would appear at each generation.

Finally, the curves presented on figure 9 can be drawn. In spite of local minima, the average fitness increases for the two semantic attributes, that shows a good global behaviour of the model: the user satisfaction increases on average from 3.0 to 5.0 for the attribute *friendly*, and from 2.3 to 3.9 for the attribute *sportive*. So it is an increase of about 70%.

#### 4.3.2 – Test of “satisfaction superiority of the IGA model”

Table 3 shows the results of the test for each user and for each semantic attribute.

**Table 3: average values of the results of the subjectivity test (for the evaluation scale, see table 2).**

	User 1	+1,78
	User 2	+1,67
Sportive	User 3	+1,56
	User 4	+1,00
	Average	+1,50
	User 5	-0,33
	User 6	+0,67
Friendly	User 7	+0,78
	User 8	+1,22
	Average	+0,58
Total Average		+1,04

Some details have to be explained to understand the meaning of table 3. The pairwise comparisons are made according to the following scheme: the best individuals from the IGA are compared to the best paper individuals. User 4 has for example an average evaluation of +1.00. According to table 2, it means that user 4 found the best individuals from the IGA slightly superior to the best paper individuals. Thus positive numbers prove that the results obtained with our model are better than without.

The results show a good behaviour for the semantic attribute *sportive*: the average score is +1.50, so the best IGA individuals are between “slightly superior” and “superior” to the best paper individuals.

For the attribute *friendly*, the results show a good behaviour too, even if the difference is slightly less perceptible. All the users prefer the individuals from the IGA except user 5, who prefers the individuals without the model. Globally, the results for this attribute are not really homogenous, which can be explained by a more subjective comprehension (and so characteristic to each user) of the word *friendly* than the word *sportive* (this observation will be highlighted in other tests presented after).

#### 4.4 – Synthesis

Those first two tests show a satisfactory behaviour of our model. The user satisfaction increases with the generations, and the results are globally better with the IGA than without.

However, some major differences can be noticed between the two semantic attributes. Indeed the user perception is really different from a user to another: *sportive* seems to be perceived in the same way by everyone, whereas each user has a different perception of *friendly*. But to confirm these observations we have proceeded to more sophisticated post-processing analysis, using the similarity index described in section 3.

### 5- Behaviour’s model analysis

For the following tests, the index similarity described in section 3 is used. Those tests are post-processing analysis performed on the data of the subjective evaluation workshop (see section 4).

## 5.1 – Test of “diversity lowering”

### 5.1.1 – Description

The aim of this test, called test of “*diversity lowering*” is to study:

- The evolution of the similarity index for each user along the ten generations.
- The similarities that could exist between the individuals of the last generation of each designer.

The idea consists simply to calculate the similarity index between all the individuals of all the users and for all the generations of the previous workshop (see section 4), and to group together the data to observe potential correlations.

A high value of similarity for the last generation would mean that the IGA converges towards a single individuals family, and so that there is a diversity lowering.

### 5.1.2 – Results

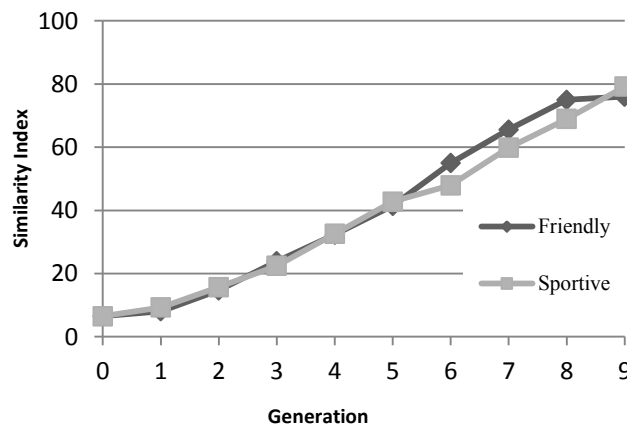


Figure 10: evolution of the similarity index along the generations for the two semantic attributes.

The first results concern the evolution of the similarity index of each user. The curves are shown on figure 10. The initial population has an average of 6.5% of similarity. In all the cases and for the two semantic attributes, the values increase quickly. At the 6<sup>th</sup> generation (Generation #5 on the graph), the averages of all the users are above 40% similarity, and at about 75% for the last generation. The averages of each user (in italic in tables 4 and 5) are all included between 60 and 90% of similarity, which correspond to the ranges “very strong similarities” and “imperceptible differences”.

Table 4: similarity values between users for the semantic attribute *friendly* (in %).

	User 5	User 6	User 7	User 8
User 5	79,07	<b>6,91</b>	<b>4,02</b>	<b>2,89</b>
User 6		89,65	<b>20,65</b>	<b>8,81</b>
User 7			61,02	<b>11,97</b>
User 8				74,16
Total average	<b>9,21</b>			

It shows that for each user, the model converges towards a single family of car silhouettes, which are very close in terms of visual aspect. But it means too that there is a real diversity lowering.

**Table 5: similarity values between users for the semantic attribute *sportive* (in %).**

	User 1	User 2	User 3	User 4
User 1	<i>79,95</i>	<b>47,87</b>	<b>63,09</b>	<b>36,20</b>
User 2		<i>72,34</i>	<b>74,14</b>	<b>42,62</b>
User 3			<i>89,61</i>	<b>41,04</b>
User 4				<i>74,74</i>
Total average			<b>50,83</b>	

## 5.2 – Test of “inter-designer convergence”

### 5.2.1 – Description

The goal of this test, called test of “*inter-designer convergence*” is to study the similarities that could exist between the last generations of all the users.

As in the previous test, the idea consists simply to calculate the similarity index between all the individuals of all the users and for all the generations of the previous workshop (see section 4), and to group together the data to observe potential correlations.

### 5.2.2 – Results

The results are displayed on tables 4 and 5. These matrices represent the averages of the similarity index for the two semantic attributes:

- Between all the users (in bold).
- For each user (in italic). These values are treated in the previous test.

The values are really consistent with the visual aspect of the individuals. For the semantic attribute *friendly*, the values inter-designer (in bold) are not high, and the total average is only 9.21% similarity, whereas this total average is 50.83% for the attribute *sportive*. According to table 1 (in section 3), it means that the similarities between the last population of the users with the attribute *sportive* is very strong, and that there is no connection between those with the attribute *friendly*.

These results confirm rigorously the observations of section 4: the perception of the word *sportive* is the same for all users (they all came to the same type of profiles), whereas each user has its own perception of the word *friendly* (they all have a different kind of car silhouettes at the end of the workshop).

## 5.3 – Test of “superiority of the IGA model”

### 5.3.1 – Description

This test is called test of “*superiority of the IGA model*”. Its goal is to show that our model obtains better results than without, in particular in terms of time and number of individuals.

It consists in comparing for each user the 3 best individuals of the IGA with the 3 best paper individuals (see section 4) with the similarity index. So it is the same test as the test of “*satisfaction*”

*superiority of the IGA model*” described in section 4.3.2, but instead of treating the fitness data, the similarity index values are used.

### 5.3.2 – Results

The maxima of similarity are shown in table 6. For 5 users out of 8, the maximum is above 70% of similarity. For 2 users, the maximum is included between 20 and 30%. For the last one, it is only 2.67% of similarity.

These values mean that it is possible to obtain the same results with and without our model. 5 users have found with the IGA *at least* one individual that is common or very close to one of the paper individuals.

400 individuals were proposed on paper. It took about 45 minutes for the user to evaluate them. Less than 200 individuals were used with the IGA (10 generations of 20 individuals, minus those who survive from one generation to the next one), during about 20 minutes. So we can say that our system seems to be able to bring the same results as without the model, more quickly and with fewer individuals.

**Table 6: maxima of similarity values between the best individuals of the IGA and the paper individuals (in %).**

	User 1	84,49
	User 2	21,26
Sportive	User 3	80,92
	User 4	97,93
	Average	71,15
	User 5	2,67
Friendly	User 6	98,09
	User 7	71,32
	User 8	26,94
	Average	49,75
Total Average		60,45

## 5.4 – Test of “attraction in the surroundings of initial individuals”

### 5.4.1 – Description

The last test proposed in this paper is called test of “*attraction in the surroundings of initial individuals*”. It tries to answer the following question: is the user really able to design his/her own car silhouette, or are the final individuals influenced by the initial individuals and close to their genotypes?

Our solution to this problem is to compare the best individuals of the IGA with the 20 car silhouettes of the initial population.

### 5.4.2 – Results

The results of this test are very interesting, but too sizeable to be displayed here. That is why only the case of user 2 is presented on table 7.

This table shows that the best individuals of user 2 are very close to Individual #13 of the initial population (more than 50% of similarity). Individual #13 represents the silhouette of a Porsche 911 (see figure 11).





Figure 11: Porsche 911 and its car silhouette.

Two other minor influences can be noticed (with Individuals #2 and #8), but the values are below 20%, which is not significant. User 2 worked with the semantic attribute *sportive*, and it is important to notice that the same results are obtained with the 3 other users who worked with this attribute.

		User 2		
		Best Individual 1	Best Individual 2	Best Individual 3
Initial Population	Ind 0	█ 2,04	█ 1,93	█ 1,84
	Ind 1	█ 2,95	█ 2,77	█ 2,65
	Ind 2	█ 10,02	█ 9,02	█ 8,80
	Ind 3	█ 0,91	█ 0,88	█ 0,90
	Ind 4	█ 3,45	█ 3,24	█ 3,35
	Ind 5	█ 2,04	█ 1,94	█ 1,91
	Ind 6	█ 1,42	█ 1,36	█ 1,36
	Ind 7	█ 1,78	█ 1,74	█ 1,52
	Ind 8	█ 13,90	█ 13,73	█ 10,75
	Ind 9	█ 6,62	█ 6,17	█ 6,53
	Ind 10	█ 2,76	█ 2,60	█ 2,47
	Ind 11	█ 3,28	█ 3,41	█ 2,91
	Ind 12	█ 2,08	█ 1,97	█ 1,97
	Ind 13	█ 69,65	█ 69,89	█ 53,21
	Ind 14	█ 3,50	█ 3,30	█ 3,46
	Ind 15	█ 7,15	█ 7,23	█ 10,83
	Ind 16	█ 2,07	█ 1,97	█ 1,93
	Ind 17	█ 3,05	█ 2,88	█ 2,87
	Ind 18	█ 1,65	█ 1,58	█ 1,47
	Ind 19	█ 1,99	█ 1,93	█ 1,69

Table 7: example of similarity values for the comparison between the best individuals of the IGA and the initial population (in %).

That means that they all came to the same type of profile (see section 5.1.2) which is very close to a Porsche 911. In other words, all the users perceived a *sportive* car as a Porsche 911.

For the semantic attribute *friendly*, the results are totally different. No car that influences all the users can be identified. In 2 cases out of 4, no significant influence of an initial individual is noticed. In the two other cases, such a car can be identified, with the scores of 40% of similarity, but this car is different for the two users (Porsche 911 for the one, Chevrolet Corvette C4 for the other). It shows again that this attribute is differently perceived.

## 5.5 – Synthesis

Those 3 tests based on the similarity index permit to prove rigorously some results of the subjective evaluation workshop (see section 4). They also permit to observe the behaviour of the model and the perception of the users.

In all cases the system turns out to converge towards a uniformized population. For the semantic attribute *sportive*, this profile is the same for any of the four users and is close to a Porsche 911. For the semantic attribute *friendly*, no similar influence from the initial population is identified, and the results really depend on the user. The perception of the attributes is really different from a user to another.

Moreover, our model converges quickly and with less individuals than without the model, which is really satisfactory.

However, some remarks can be formulated:

- The model always converges towards one profile family. Is it not possible to converge towards at least two different families?
- What happens if the Porsche 911 or the other influencing cars are not in the initial population? Do the users even find them?
- Why are there so many differences in the perception of the semantic attributes? Some concept of Emotional Design could be introduced to develop a more powerful model.

## 6- Conclusions and perspectives

We have presented in this paper an innovative car silhouettes design model. Based on an encoding method of the genomes by Fourier decomposition, it offers to the car designers the possibility to create new car silhouettes from an initial population according to their preferences.

Several user tests have been carried out to study the convergence of the model in terms of user satisfaction, but also in terms of perceived distance, using a similarity index. The results show very satisfactory results. They permit to answer the four questions defined in the introduction:

- *Are the users satisfied by our model?* Yes, the satisfaction increases along the generations of the IGA, and the users prefer to work with our model rather than without.
- *Are the results better with our model than without?* Yes, in terms of user satisfaction, but also in terms of time and numbers of processed individuals.
- *Does our model really create novelty?* Potentially. It highly depends on the chosen semantic attribute. But individuals that are far from the initial population in terms of similarity can be created.
- *Are the results different from a user to another?* Potentially. It depends on the chosen semantic attribute too.

In consequence, our system should allow style designers to converge towards intuitive ideas and to make emerging surprise in exploring large spaces of potential silhouettes.

Further tests with a larger panel of user could be done to validate the statements described in this paper and dealing with Emotional Design. The large difference of results between the simple semantic attributes *sportive* and *friendly* open a very interesting field of study for the future with the objective of creating CAD tools that really meet the user perception.

The real role of the mutation operator has to be studied too. So the identification of the optimal parameters would permit to develop a more reliable model.

Finally these first results are very promising and numerous research perspectives appear now:

- Allowing a step of “direct modification by the designer” within an intermediary generation, i.e. modifying some details of a silhouette curve or even adding new individuals to the population. We must acknowledge here that all our 20 initial silhouettes are silhouettes of existing commercial cars. Then, the experiment in this paper has just consisted in morphing between known solutions. For really creating surprise and innovating, we must also test our system in a more creative way.
- Combining several series of closed curves to better define the important lines of a car (see [C1]).
- Making more complex the interactive assessment of individuals by the style designer through multicriteria assessments under several perceptual attributes.
- Exploring a new product field: the encoding method that we use is very flexible and applicable to many design objects. We think that excellent results could be obtained with simple shapes. The ideal form would be a simple 2D-closed-curve from a solid of revolution (which permits to describe the whole product with a single closed curve), like stemmed glasses or vases.
- Finally, a way to improve the model could be to develop a function for introducing car silhouettes drawn by the designer himself. For instance, a new form appears in his/her mind during the IGA process, so he/she can draw this form on the screen (or with some graphic tool) and introduce the new individual in the current population.

## 7- References

**[AC1]** Ang M.C., Chau H.H., McKay A., de Pennington A., "Combining evolutionary algorithms and shape grammars to generate branded product design", in Design Computing and Cognition '06, Gero J.S. Editor, Springer Netherlands, 2006, pp. 521-539.

**[BC1]** Bentley P.J., Corne D.W., "An introduction to creative evolutionary systems", pp. 1-75.

**[C1]** Cheutet V., "2D semantic sketcher for a car aesthetic design", Proc. CPI'2007: Conception et Production Intégrées, Rabat, Maroc, October 22-24, 2007.

**[DP1]** Dagher A., Petiot J.-F., "Study of the correlations between user preferences and design factors: application to cars front-end design", Proc. International Conference on Engineering Design: ICED'07, Paris, France, August 28-31, 2007.

**[KC1]** Kim H.S., Cho S.B., "Application of interactive genetic algorithm to fashion design", Engineering Applications of Artificial Intelligence vol. 13, 2000, pp. 635-644.

**[LT1]** Liu H., Tang M., Frazer J.H., "Supporting evolution in a multi-agent cooperative design environment", Advances in Engineering Software vol. 33, n° 6, 2002, pp. 319 - 328.

**[MM1]** McGarva J., Mullineux G., "Harmonic representation of closed curves", Applied Mathematical Modelling vol. 17, 1993, pp. 213-218.

**[NA1]** Nicaise F., Antonsson E.K., "Indirect encoding of structures for evolutionary design", Proc. International Conference on Engineering Design: ICED'07, Paris, France, August 28-31, 2007.

**[OC1]** Osborn S., Cagan J., Pawlicki R., Smith R., "Creating cross-over vehicles: Defining and combining classes using shape grammars", Artificial Intelligence for Engineering Design, Analysis and Manufacturing vol. 20, 2006, pp. 217-246.

**[RE1]** Renner G., Ekart A., "Genetic Algorithms in Computer Aided Design", Computer Aided Design vol. 35, 2003, pp. 709-726.

**[VY1]** Vasiliu A., Yannou B., "Dimensional Synthesis of Planar Mechanisms Using Neural Networks: Application to Path Generator Linkages", Mechanism and Machine Theory vol. 36, n° 2, 2001, pp. 299-310.

**[YF1]** Yanagisawa H., Fukuda S., "Interactive Reduct Evolutional Computation for Aesthetic Design", Journal of Computing and Information Science in Engineering vol. March, 2005, pp.

**[YC1]** Yannou B., Coatanéa E., "Easy and flexible specifications and product evaluations by expert and customer comparisons with existing products" Proc. International Conference on Engineering Design: ICED'07, Paris, France, August 28-31, 2007.

**[YD1]** Yannou B., Dihlmann M., Cluzel F., "Indirect encoding of the genes of a closed curve for interactively create innovative car silhouettes", in International Design Conference - Design 2008, Dubrovnik, Croatia, May 19 - 22, 2008.

**[YD2]** Yannou B., Dihlmann M., Awedikian R., "Evolutive design of car silhouettes", in submitted to IDETC/DAC: ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conferences / Design Automation Conference, New-York, NY, USA, August 3-6, 2008.