

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Adaptive Evaluation Strategy Based on Surrogate Models

Yi-nan Guo and Jian Cheng
*China University of Mining and Technology
China*

1. Introduction

Interactive genetic algorithm(IGA) is a kind of genetic algorithms in which individuals' fitness values are evaluated by human subjectively. They combine canonical genetic algorithms with human intelligent evaluations. Now they have been applied to optimization problems whose objectives can not be expressed by explicit fitness functions, such as music composition(Biles,1996), production design(Takagi,1998), image retrieval(Takagi,1999) and so on. However there exists human fatigue in evaluation, which limits population size and the number of evolution generations. Moreover, subjective fitness values given by human mainly reflect human cognition and preference, which is related to domain knowledge. So IGAs need more knowledge than other genetic algorithms for explicit optimization functions.

In IGAs, human fatigue is a key problem which limits the application of the algorithm. According to physiological characters of human, if human are absorbed in one work for a long time, they are easy to feel tired. However, human need to evaluate each individual in each generation. More individuals human evaluate, they feel more tired. If cognition knowledge about human preference to optimization problems can be extracted from the evolution and utilized to evaluate individuals instead of human, the number of individuals evaluated by human shall decrease which will alleviate human fatigue. This cognition knowledge is generally described by surrogate models.

Surrogate models are trained by samples which consist of human evaluated individuals and their fitness values. Hence, they approximatively reflect human preference. When human feel tired, surrogate models can replace human to evaluate all individuals or part of individuals in each generation so as to reduce the number of individuals evaluated by human. This can reduce human burden and alleviate human fatigue.

Here, four key issues about surrogate models which influence the performances of the algorithm are taken into account. Firstly, surrogate models must exactly keep consistency with human cognition and preference in order to ensure the convergence of the algorithms. So how to obtain the models with good prediction precision and generalization is the base. The second issue is when to use surrogate models instead of human in evaluation. The last two problems are how many and which individuals are selected to calculate fitness values by surrogate models in each generation.

Up to now, many researches on surrogate models have been done. But most of them focus on the structure of surrogate models. Different models adopting artificial neural networks(Zhou,2005) or support vector machines(Wang,2003) are proposed. However, how to rationally utilize surrogate models are not taken into account enough. Firstly, in existing researches, no matter whether surrogate models exactly reflect human real preference, the models are used to evaluate individuals when human feel tired. Secondly, population size is small and fixed all the time. And the number of individuals evaluated by the models in each generation is also fixed. These make IGAs alleviating human fatigue limitedly. Thirdly, few of researches concern which individuals are selected to be evaluated by models. To solve above problems, latter three issues are mainly studied, including when the models are used in evaluation, how many and which individuals are evaluated by the models in each generation.

In the chapter, assume that human preference to optimization problems is stable. A novel interactive genetic algorithms adopting adaptive evaluation strategy based on surrogate models(AES-IGAs) is proposed. Firstly, two measures reflecting the models' precision and human fatigue are given. Based on them, evaluation process is divided into three phases. In different phase, human participate in evaluation to different extent. Secondly, the number of individuals evaluated by surrogate models is adaptively tuned in evolution according to above two measures so as to alleviate human fatigue at most. Because surrogate models calculate individuals' fitness values by computers which do not need human participation, population size is enlarged when only surrogate models are adopted in evaluation so as to improve convergence speed. Thirdly, how to select rational individuals evaluated by the models from population is studied. Fuzzy c-means clustering algorithm is adopted And three kinds of approximate distance is put forward as measures. To validate the rationality of AES-IGAs, experiments based on fashion evolutionary design system are done. And testing results are analyzed at last.

2. Startup Mechanism about Surrogate Model

Startup mechanism offers some conditions which decide when to use surrogate model in evaluation process. That is, in which generation these conditions are satisfied, population can be evaluated by surrogate models in proper proportion. In general, when human feel tired or surrogate models have learned human preference exactly, the models are adopted to calculate the individuals' fitness values. So startup mechanism about surrogate models normally includes two conditions. They are the condition of human fatigue and the condition of models' precision. When any of conditions is satisfied, surrogate models are start up to evaluate.

Suppose F_M is the fitness value calculated by surrogate models. F_U is the fitness value given by human. x_M and x_U respectively denote individuals evaluated by the models and human. $X(t)$ expresses the population in t -th generation. Adaptive evaluation strategy is shown as follows(Guo,2007).

$$F(X(t)) = \{F_M(x_M, t), F_U(x_U, t)\}, x_M \neq x_U, x_M, x_U \in X(t) \quad (1)$$

$$s.t. \quad F_M \neq \emptyset, \forall ((Fa(t) \geq \varepsilon) \vee (Pr(t) \leq \Psi))$$

Here, $Fa(t) \geq \varepsilon$ describes the condition of human fatigue where $Fa(t)$ expresses the degree of human fatigue and ε is the threshold of $Fa(t)$.

Definition 1: the degree of human fatigue

The degree of human fatigue reflects how tired human are. In general, human will spend more time evaluating individuals when there are more similar individuals in the population. Therefore, human feel more tired when the total number of individuals evaluated by human is more and time for evaluation in each generation is more.

Letting $v(t)$ denotes how long human spend evaluating. $\beta(t)$ is the proportion of human evaluated individuals to population. The degree of human fatigue is defined as follows (Guo, 2007).

$$Fa(t) = 1 - e^{-tv(t)\beta(t)S(t)} \quad (2)$$

In formula(2), $S(t)$ denotes the similarity of population, which describes average similarity of individuals in population, shown as follows.

$$S(t) = \frac{2 \sum_{i=1}^{|P|-1} \sum_{j=i+1}^{|P|} \sum_{l=1}^n \sigma_l(x_i(t), x_j(t))}{|P| |P-1|} \quad (3)$$

where $|P|$ is population size and n is the length of individuals. $\sigma_l(x_i(t), x_j(t))$ expresses the similarity of l -th bit between two individuals. $\sigma_l(x_i(t), x_j(t))=1$ if l -th bit of $x_j(t)$ is same as it of $x_i(t)$, otherwise $\sigma_l(x_i(t), x_j(t))=0$.

In formula(1), $Pr(t) \leq \Psi$ describes the condition of models' precision where $Pr(t)$ expresses the reliability of surrogate models and Ψ is the threshold of $Pr(t)$.

Definition 2: the reliability of surrogate models

The reliability of surrogate models reflect the consistency between surrogate models and human preference. It is measured by average Euclid distance between individuals' fitness values calculated by the models and them given by human, shown as follows.

$$Pr(t) = \sqrt{\frac{1}{|P|} \sum_{i=1}^{|P|} (F_U(x_i, t) - F_M(x_i, t))^2} \quad (4)$$

Obviously $Pr(t)$ also describes the generalization of surrogate models.

In a word, whether surrogate models are utilized lies on two conditions: whether or not the degree of human fatigue exceed the threshold; whether or not reliability of surrogate models exceed the threshold.

3. The Number of Substituted Individuals Evaluated by Surrogate Models

In AES-IGAs, some individuals are evaluated by human, whereas others' fitness values are calculated by surrogate models. Here, individuals evaluated by the models in each generation are called substituted individuals. How many substituted individuals are there is a key problem, which influences the performance of AES-IGAs. Up to now, in most of IGAs

adopting surrogate models based evaluation strategy, the number of substituted individuals is fixed and equal to fifty percent of population size. This limits the effect of surrogate models on performances. Aiming at this problem, the number of substituted individuals adaptively varies in AES-IGAs.

In general, one hopes to evaluate less individuals in IGAs when he feels tired. And when surrogate models can reflect human preference better, it shall be used to evaluated more individuals instead of human so as to reduce human burden. Therefore, two factors including human fatigue and models' precision are considered to decide the number of substituted individuals. When human feel more tired or models' precision is better, less individuals are evaluated by themselves. So the number of substituted individuals in t -th generation is defined as

$$N_M(t) = \lfloor P | \rho(t) \rfloor \quad (5)$$

Here, $\rho(t)$ is the proportion of substituted individuals to population, called the substituted proportion, shown as follows.

$$\rho(t) = e^{-(Pr(t)/\bar{f})(\varepsilon - Fa(t))} \quad (6)$$

Here, \bar{f} denotes the upper bound of the fitness values. Obviously the substituted proportion also satisfies $\rho(t) + \beta(t) = 1$.

4. Substitution Mechanism about Surrogate Model

In existing researches, the substituted proportion in evolution is normally fixed after surrogate models are used. Corresponding evaluation process in IGAs can be divided into two phases: human-evaluation phases and models-evaluation phases. According to the substituted proportion, two kinds of dual-phase evaluation strategies with different $\rho(t)$ are analyzed as follows.

- (1) While the conditions of startup mechanism are satisfied, $\rho(t)=1$. That is, the evaluation process is divided into two phases. When surrogate models are not used, population is evaluated by human only. Otherwise, all individuals are evaluated by surrogate models.
- (2) While the conditions of startup mechanism are satisfied, $\rho(t)=C(0<C<1)$. Here C is a constant(Zhou,2005). That is, when surrogate models are adopted, population is not evaluated by surrogate models only. Some of individuals are evaluated by human, others are evaluated by surrogate models. And the number of substituted individuals is $\lfloor P | \rho(t) \rfloor$.

In above two evaluation strategies, fixed substituted proportion does not make surrogate models used enough. Therefore, variable substituted proportion is adopted in this chapter. Corresponding evaluation process is different from above two instances. It contains three phases.

Phase I: Population is evaluated by human only

In this phase, the models are not used and all of individuals are evaluated by human. The evaluation mode is defined as follows.

$$F(X(t)) = \{F_U(x_U, t)\}, x_U \in X(t) \quad (7)$$

$$s.t. \forall ((Fa(t) < \varepsilon) \wedge (Pr(t) > \Psi))$$

And the number of substituted individuals is

$$N_M(t) = 0 \quad (8)$$

It is obvious that in this phase, human do not feel tired and surrogate models can not exactly reflect human preference. Above phenomena possibly appear at the beginning of evolution. So this evaluation mode is usually adopted in the former evolution. Note that surrogate models shall be continually update according to evaluated individuals in this phase.

Phase II: Population is mixed evaluated by human and surrogate models

In this phase, surrogate models are startup because they learn human preference approximatively. However, the degree of human fatigue does not exceed the threshold. That means human still participate in the evaluation process. And human evaluated individuals are used as samples to improve models' precision further. Obviously some of individuals are evaluated by human and others' fitness values are calculated by the models. So the evaluation mode are shown as follows.

$$F(X(t)) = \{F_M(x_M, t), F_U(x_U, t)\}, x_M \neq x_U, x_M, x_U \in X(t) \quad (9)$$

$$s.t. \forall ((Fa(t) < \varepsilon) \wedge (Pr(t) \leq \Psi))$$

In order to reduce human burden gradually, the number of substituted individuals in phase II is increasing along with the degree of human fatigue.

$$N_M(t) = \left\lfloor |P| e^{-(Pr(t)/\bar{f})(\varepsilon - Fa(t))} \right\rfloor \quad (10)$$

In above two phases, population size is fixed and small because human participate in the evaluation process. In general, population size in IGAs is less than ten so as to alleviate human visual fatigue.

Phase III: Population is evaluated by surrogate models only

In this phase, all of individuals are evaluated by surrogate models. So the evaluation mode is described as follows.

$$F(X(t)) = \{F_M(x_M, t)\}, x_M \in X(t) \quad (11)$$

$$s.t. \forall ((Fa(t) \geq \varepsilon) \wedge (Pr(t) \leq \Psi))$$

In phase III, human often feel very tired. Individuals' fitness values given by human may differ from their real preference. So these individuals can not be used to train surrogate models. And the models are not updated any more. This evaluation mode is normally adopted in the latter evolution. Because the evaluation process based on surrogate models is done by computers, there does not exit human fatigue in phase III. That means the evolution

of AES-IGAs is the same as traditional genetic algorithms. So population size can be enlarged. But how to increase population size is a key problem. Higher the models' precision is, the generalization of the model is better. So enlarged population size is defined as follows.

$$N_p(t) = |P|^{\left\lfloor \frac{\bar{f}}{Pr(\tau k)} + 0.5 \right\rfloor} \quad (12)$$

Here, $Pr(\tau k)$ expresses the reliability of surrogate models in Phase III. It is obvious that the exponent in formula(12) may be 1, 2 or 3. So population size may be enlarged to corresponding multiple of $|P|$.

5. Surrogate Model-based Evaluation Strategy with Oriented Selection of Substituted Individuals

In existing researches on surrogate model-based IGAs, substituted individuals are often randomly selected from population after the evaluation mode is decided. But if selected substituted individuals are far away from existing training samples, their fitness values calculated by surrogate models may differ from human cognition. This will misguide the evolution away from human real preference. Aiming at above problem, an oriented selection method of substituted individuals is proposed. Here, substituted individuals are selected in terms of the distance between individuals and samples so as to improve the evaluation precision to the utmost extent.

Taken oriented selection process in t -th generation as example, algorithm steps are shown as follows(Guo,2007).

Step1: The number of substituted individuals $N_M(t)$ is decided in terms of human fatigue.

Step2: Samples are composed of evaluated individuals and their fitness values given by human. They are grouped into sub-sample sets adopting Fuzzy-C mean algorithm.

Step3: Aiming at each sub-sample set, a surrogate model described by artificial neural networks is trained.

Step4: The distance between each individual in population and each samples or each sub-sample sets are calculated.

Step5: Individuals are sorted according to above distance. Then individuals with less distance are regarded as substituted individuals.

Step6: The fitness values of substituted individuals are calculated by their affiliated clusters' surrogate model.

Obviously how to group samples, train surrogate models and select substituted individuals are three main issues. Here, former two problems are mainly discussed.

5.1 Sample clustering

How to construct sample set and decide the cluster number are two key problems in sample clustering.

Sample set are composed of evaluated individuals and their fitness values given by human. They are expressed by $x_U(t)$ and $f(x_U(t))$. Because human cognition may vary in evolution, the fitness value of same individual may be variant in different generation. In order to track

human preference, latest fitness values of any individuals are reserved. Therefore, sample set are described as follows.

$$Q(t) = \{(x_U, f(x_U)) \mid x_U^l \neq x_U^k, l, k = 1, 2, \dots \mid Q(t) \mid\} \quad (13)$$

Sample set are grouped into clusters by fuzzy c-means clustering algorithm (Karayiannis, 1997). Suppose $L=2,3,\dots,K(K < \mid Q(t) \mid)$ is the cluster number, which is normally decided according to domain knowledge. Improper K will lead to wrong clustering result. Aiming at this problem, comparison method(Cheng,2006) is adopted to obtain suitable cluster number. Assume that $Q_i(t)(i=1,2,\dots,L)$ is sub-sample set. And each cluster center is

$$c_i(t) = \frac{\sum_{l=1}^{\mid Q(t) \mid} u_{il}^m(t) x_U^l(t)}{\sum_{l=1}^{\mid Q(t) \mid} u_{il}^m(t)}, i = 1, 2, \dots, L \quad (14)$$

Here, $m \in [1, \infty)$ is the weighing index. $u_{il} \in [0,1]$ denotes the subjection degree of $x_U^l(t)$ to $Q_i(t)$, which satisfies $\sum_{i=1}^L u_{il} = 1, \forall l = 1, 2, \dots, \mid Q(t) \mid$.

5.2 Selection of substituted individuals

How many substituted individuals are there in each generation is the principal issue, which is decided by $\rho(t)$. However, which individuals are selected as substituted individuals depends on their approximate distances. Here, the distances between an individual and each sample or each sub-sample set are called approximate distances.

Suppose $x_j(t)$ is an individual. Three kinds of approximate distances are defined as follows(Cheng,2009).

(1) Center approximate distance

The clustering center of each sub-sample set is used to judge which cluster does an individual belong to. Suppose $c_i(t)$ is clustering center of i -th cluster. The minimum distance between $x_j(t)$ and $c_i(t)$ is defined as center approximate distance, expressed by D1.

$$D1^j = \min_{i \in \{1,2,\dots,L\}} d1(x_j(t), c_i(t)) \quad (15)$$

If individuals are encoded by real number, d1 is described as

$$d1(x_j(t), c_i(t)) = \sqrt{\sum_{h=1}^n (x_j^h(t) - c_i^h(t))^2} \quad (16)$$

If individuals are encoded by binary number, d1 is described as

$$d1(x_j(t), c_i(t)) = \bigcup_{h=1}^n (x_j^h(t) \oplus c_i^h(t)) \quad (17)$$

Because only clustering centers of sub-sample sets are taken into account, the computation complexity of center approximate distance is $O(|P|L)$.

(2) Cluster approximate distance

All of samples in each sub-sample set are used to judge which cluster does an individual belong to. The minimum distance between $x_j(t)$ and samples in $Q_i(t)$ is defined as cluster approximate distance, expressed by D2.

$$D2^j = \min_{i \in \{1, 2, \dots, L\}} d2(x_j(t), Q_i(t)) \quad (18)$$

Suppose $x_U^i(t)$ is a sample in i -th cluster. If individuals are encoded by real number, $d2$ is described as

$$d2(x_j(t), Q_i(t)) = \frac{1}{|Q_i(t)|} \sum_{l=1}^{|Q_i(t)|} \sqrt{\sum_{h=1}^n (x_j^h(t) - x_U^{i, lh}(t))^2} \quad (19)$$

If individuals are encoded by binary number, $d2$ is described as

$$d2(x_j(t), Q_i(t)) = \frac{1}{|Q_i(t)|} \sum_{l=1}^{|Q_i(t)|} \left[\bigcup_{h=1}^n (x_j^h(t) \oplus x_U^{i, lh}(t)) \right] \quad (20)$$

Here, $|Q(t)| = \sum_{i=1}^L |Q_i(t)|$. So the computation complexity of cluster approximate distance is $O(|P||Q(t)|)$.

(3) Generalized approximate distance

All of samples in sample set are used to judge which cluster does an individual belong to. The minimum distance between $x_j(t)$ and $x_U(t)$ is defined as generalized approximate distance, expressed by D3.

$$D3^j = \min_{l \in \{1, 2, \dots, |Q(t)|\}} d3(x_j(t), x_U^l(t)) \quad (21)$$

If individuals are encoded by real number, $d3$ is described as

$$d3(x_j(t), x_U^l(t)) = \sqrt{\sum_{h=1}^n (x_j^h(t) - x_U^{lh}(t))^2} \quad (22)$$

If individuals are encoded by binary number, $d3$ is described as

$$d3(x_j(t), x_U^l(t)) = \bigcup_{h=1}^n (x_j^h(t) \oplus x_U^{lh}(t)) \quad (23)$$

Obviously the computation complexity of generalized approximate distance is $O(|P||Q(t)|)$.

Through comparison of the computation complexity of above three approximate distances, we know that the computation complexity of generalized approximate distance is the same as it of cluster approximate distance. In addition, population size is normally more than the cluster number, namely $|Q(t)| \geq L$. So the computation complexity of center approximate distance is least.

6. Analysis of Convergence

In genetic algorithms, whether and how fast they converge are used to measure the algorithm's performances. And many researches have discussed genetic algorithms' convergence. However, IGAs' convergence is seldom taken into account. In this chapter, AES-IGAs' convergence is analyzed in terms of drift analysis and the first hitting time condition proposed by Jun He(He, 2001).

Definition 3: critical generation

In AES-IGAs, when $t > \tau k$, surrogate models are used to calculate all individuals' fitness values. If the algorithms can converge to optimal solution under above condition, τk is called critical generation as shown in follows(Guo,2007). Whether τk exists is a key problem which influences AES-IGAs' convergence.

$$\tau k = \min \{t \mid (\rho(t+1) = 1) \wedge D(X(t+\zeta)) = 0, \zeta \in [0, T-t], t \in (0, T]\} \quad (24)$$

Where

$$D(X(t+\zeta)) = \min \{i \mid D(x_i(t+\zeta)), x_i(t+\zeta) \in X(t+\zeta)\} \quad (25)$$

Here, $D(x_i(t+\zeta))$ is simplified description of $D(x_i(t+\zeta), x^*)$, $x^* \in S^*$. S^* is the optimal variable space. $D()$ denotes the distance between two individuals. If individuals are encoded by binary number, $D()$ is calculated by Hamming distance. Otherwise, if individuals are encoded by real number, $D()$ is calculated by Euclidean distance.

Theorem 1: For any generation satisfied $t > \tau k$, all individuals are evaluated by surrogate models. During this phase, the evolution of AES-IGAs is the same as genetic algorithms. Based on the first hitting time drift condition with exponential form(He,2001), there exists $\bar{\zeta} = \min\{\zeta \mid D(X(t+\zeta)) = 0\}$, which satisfies drift condition $E[\bar{\zeta} \mid d(X(\tau k)) > 0] \leq h(|P|)$. Here, $h(|P|)$ is a polynomial function of $|P|$.

Proof: Assume that individuals are encoded by binary number. So the distance between an individual and optimal solution is

$$D(x_i(t)) = \sum_{l=1}^n \sigma_l(x_i(t), x^*) \quad (26)$$

Given a random sequence $\{D(X(\tau k+j)), j=0,1,\dots\}$, $\forall X(\tau k+j), D(X(\tau k+j)) \leq |P|$ is obviously satisfied. If the algorithms adopt one-point crossover, bit mutation and roulette selection, following inequation is obtained in terms of drift condition.

$$E[D(X(\tau k + j + 1)) - D(X(\tau k + j)) \mid D(X(\tau k + j)) > 0] \geq h_1(\mid P \mid) \quad (27)$$

Here, $h_1(\mid P \mid) > 0$. From Theorem.1 in reference(He,2001), we know $E[\bar{\zeta} \mid d(X(\tau k)) > 0] \leq \mid P \mid h_1(\mid P \mid)$.

Let $h(\mid P \mid) = \mid P \mid h_1(\mid P \mid)$. Then $E[\bar{\zeta} \mid d(X(\tau k)) > 0] \leq h(\mid P \mid)$ is satisfied.

Theorem 2: When $t > \tau k$, the substituted proportion $\rho(t)$ converges to 1.

Proof: Suppose $\{\dots \rho(t), \rho(t+1) \dots \mid \forall \rho(t) > 0\}$ is a random sequence. The gradient of $\rho(t)$ is

$$\frac{\rho(t+1)}{\rho(t)} = \frac{e^{-[(Pr(t+1)/\bar{f})(\varepsilon - Fa(t+1))]} }{e^{-[(Pr(t)/\bar{f})(\varepsilon - Fa(t))]} } = e^{-[(Pr(t+1)/\bar{f})(\varepsilon - Fa(t+1)) + [(Pr(t)/\bar{f})(\varepsilon - Fa(t))]} \quad (28)$$

Considering $\varepsilon - Fa(t) > 0$ and $Fa(t+1) > Fa(t) \Rightarrow \varepsilon - Fa(t+1) < \varepsilon - Fa(t)$, then

$$\begin{aligned} Pr(t+1) < Pr(t) &\Rightarrow Pr(t+1)(\varepsilon - Fa(t+1)) < Pr(t)(\varepsilon - Fa(t)) \\ Pr(t)(\varepsilon - Fa(t)) - Pr(t+1)(\varepsilon - Fa(t+1)) &> 0 \\ \Rightarrow e^{-[(Pr(t+1)/\bar{f})(\varepsilon - Fa(t+1)) + [(Pr(t)/\bar{f})(\varepsilon - Fa(t))]} &> 1 \end{aligned}$$

Therefore, $\frac{\rho(t+1)}{\rho(t)} > 1$. This shows that random sequence $\{\dots \rho(t), \rho(t+1) \dots \mid \forall \rho(t) > 0\}$

increases by degrees. And $\rho(t)$ continuously varies along with t . When $t \geq \tau k$, $Fa \geq \varepsilon$. So

$$\rho(t) = e^{-[(Pr(t)/\bar{f})(\varepsilon - Fa(t))]} \geq 1 \quad (29)$$

This makes $\lim_{t \rightarrow \tau k} \rho(t) = 1$ while $\exists \tau k \leq T$. Therefore, total number of individuals evaluated by human in evolution is

$$N_{sum} = \sum_{t=1}^{\tau k} \lfloor \mid P \mid (1 - \rho(t)) \rfloor \quad (30)$$

7. Experiments and Analysis

7.1 Background for experiments

Here, fashion evolutionary design system is adopted as a typical background to validate the rationality of IGAs with adaptive evaluation strategy (AES-IGAs). The goal of the system is to find a dress which wins human favor (Kim,2000). Visual Basic 6.0 as programming tool for human-machine interface and Microsoft Access as database are utilized. Matlab 6.5 is adopted to train surrogate models based on artificial neural networks. System's human-computer interface is shown as follows.



Fig. 1. Human-computer interface in AES-IGAs

In fashion evolutionary design system, each dress is composed of collar, skirt and sleeve. Each part has two factors including pattern and color which described by two bits. So each dress is expressed by twelve bits, which act as six gene-meaning-units (GM-units) (Guo,2007). Each gene-meaning-unit has four alleles. The meanings of each allele in gene-meaning-unit are shown in Table 1.

GM-unit	allele							
	meaning	code	meaning	code	meaning	code	meaning	code
collar's pattern	medium collar	00	high collar	01	wide collar	10	gallus	11
	long sleeve	00	medium sleeve	01	short sleeve	10	non-sleeve	11
skirt's pattern	long skirt	00	formal skirt	01	medium skirt	10	short skirt	11
	color	pink	green	01	black	10	white	11

Table 1. The meanings of each allele in gene-meaning-unit

7.2 Desired objectives and parameters in experiments

In order to validate the rationality of adaptive evaluation strategy and the influence on performance of IGAs, two groups of experiments are designed. They have different desired objectives which reflect different psychological requirements of human. Desired objectives of experiments are shown as follows.

Experiment I: To find a favorite dress fitting for summer without the limit of color.

Experiment II: To find a favorite dress fitting for summer and the color is pink.

In both experiments, artificial neural network is adopted as surrogate model. The values of parameters about the model and the evolution process are shown in Table 2.

parameters about the evolution	crossover probability	mutation probability	population size	generation	ε	ψ
	0.6	0.01	8	40	0.8	0.9
parameters about the model	input neurons	hidden neurons	output neurons	learning rate	epochs	error
	6	15	1	0.09	15000	10^{-2}

Table 2. The values of parameters

7.3 Comparison of different substituted proportion

In order to validate the rationality of AES-IGAs, thirty persons are gathered to do two groups of experiments aiming at desired objective of experiment II. In experiments, when surrogate models are used in evaluation process, two kinds of fixed substituted proportion and adaptive substituted proportion are adopted respectively. Testing results done by all persons are integrated, as shown in Table 3.

the substituted proportion	average generation	average number of human individuals evaluated	average generation when the models are adopted
$\rho(t)=0.5$	15	104	11
$\rho(t)=1$	14	88	11
$\rho(t)=e^{-(Pr(t)/\bar{f})(\varepsilon-Fa(t))}$	12	55	8

Table 3. Comparison of the performance by different substituted proportion

Obviously when different $\rho(t)$ are adopted, convergence speed of the algorithms are similar. But total number of individuals evaluated by human varies. The reason for this is that surrogate models are adopted in evaluation only when formula(1) is satisfied. Firstly, if $\rho(t)=0.5$, half of population is evaluated by human when surrogate models are adopted. However, human shall evaluate individuals all along. So total number of individuals evaluated by human is more and human will feel more tired. Secondly, if $\rho(t)=1$, all individuals are evaluated by surrogate models when human feel tired and the models can reflect human preference exactly. Although human evaluate fewer individuals than first strategy, the models are used later than adaptive evaluation strategy. That is, less individuals are evaluated by human in AES-IGAs. Therefore, IGAs adopting adaptive evaluation strategy can effectively alleviate human fatigue more.

7.4 Comparison of different population size in Phase III

When population is evaluated by surrogate models only, fixed population size $|P|$ and population size N_p are adopted in experiments respectively. Testing results are shown in Table 4.

population size	average generation	average number of human evaluated individuals
$ P $	12	55
N_p	10	55

Table 4. Comparison of different population size in Phase III

It is obvious that different population size in Phase III do not influence the degree of human fatigue in evolution because average number of individuals evaluated by human is same. As we know, no matter how many individuals are there in Phase III, human do not participate in evaluation any more. That is, total number of human evaluated individuals in evolution only depends on the evaluation in Phase I and Phase II. However, convergence speed is faster while population size in Phase III is enlarged. The reason for this is that exploration of the algorithm with larger population size is better.

7.5 Critical generation when Phase III is started

Thirty persons are gathered to do two groups of experiments aiming at above two desired objectives. Results in experiments are shown in Figure 2.

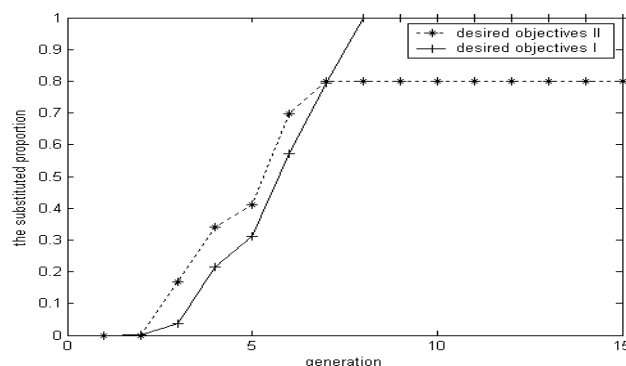


Fig. 2. $\rho(t)$ in experiments

From Figure 2, we know AES-IGAs converges. This matches Theorem.2. In addition, the substituted proportion is increasing along with the evolution. Aiming at desired objective of experiment II, all individuals are evaluated by the models when $t \geq 8$. That is, critical generation is 8. Obviously desired objective of experiment II include more detail requirements. So human need to pay attention to more GM-units, which make human feel more tired. This leads to larger critical generation. However, aiming at desired objective of experiment I, AES-IGAs converge when $t=7$. At this time, the degree of human fatigue do not exceed the threshold. So surrogate models are not completely used to replace human in evaluation process. That means critical generation do not exist in experiment I.

7.6 Comparison of different selection methods of substituted individuals

In order to analyze the performance of oriented selection methods, following three measures are given.

(1) Generation when stop rules are satisfied or human find the satisfied individuals reflects convergence speed, as shown in follows.

$$T = \arg \min_t |f(x(t)) - f^*| \leq \delta \quad (31)$$

(2) Total number of individuals evaluated by human in evolution influences the degree of human fatigue. Obviously T or N_U is larger, human feel more tired.

$$N_U = \sum_{t=1}^T (|P| - N_M(t))$$

(32)

(3) Evaluation error of surrogate model describes square root error of the fitness values given by human $f_U(x_M(t))$ and the models $f_M(x_M(t))$. Note that $f_U(x_M(t))$ is only used as reference and not utilized in evolution. Obviously this measure reflects generalization of the models.

$$E_F(t) = \sqrt{\frac{1}{N_M(t)f} \sum_{k=1}^{N_M(t)} (f_M(x_M^k(t)) - f_U(x_M^k(t)))^2}$$

(33)

Sixteen persons are divided into four groups. Aiming at desired objective of experiment II, each group does experiments adopting random selection method and three kinds of oriented selection methods respectively. Here, cluster number $L=3$. This means human preference is divided into three ranks, including favor, normal, dislike. Testing results done by each group are integrated, as shown in Table 5.

measures		T	N_U
random selection		18	112
oriented selection (D1)	value	14	92
	comparison with random selection	↓22.2%	↓17.9%
oriented selection (D2)	value	12	76
	comparison with random selection	↓33.3%	↓32.1%
oriented selection (D3)	value	12	80
	comparison with random selection	↓33.3%	↓28.6%

Table 5. Comparison of different selection methods

From Table5, we know that: (I) Because the number of samples is small in AES-IGAs, generalization of surrogate models are bad. If substituted individuals are randomly selected, they may be far from samples, which may lead to large evaluation error of surrogate model. So convergence of IGAs adopting random selection method is worse. (II) Human may favor more than one satisfied dress. Hence desired objectives of both experiments in fact are implicit multi-mode functions. Because oriented selection methods adopting D2 or D3 all consider the distances between all samples and individuals, useful information embodied in evolution are utilized enough. So their performance are better. However, oriented selection methods adopting D1 only takes the distribution of cluster center into account. Incomplete information of individuals may result in larger evaluation error so as to make the performance worse.

In above experiments, evaluation error of surrogate models varies during the evolution, as

shown in Figure 3. Obviously no matter which selection method is adopted, E_F increases in former evolution. During this stage, human do not feel tired and can not exactly decide which one is he most like. Surrogate models are trained based on samples and do not participate in evolution. Along with the evolution, human gradually know what they favor. Surrogate models are close to human preference. So evaluation error decrease in latter evolution. In addition, oriented selection methods adopting D2 or D3 reserve samples' information to largest degree. Hence AES-IGAs using these selection methods have less evaluation error. This can effectively avoid misguiding evolution.

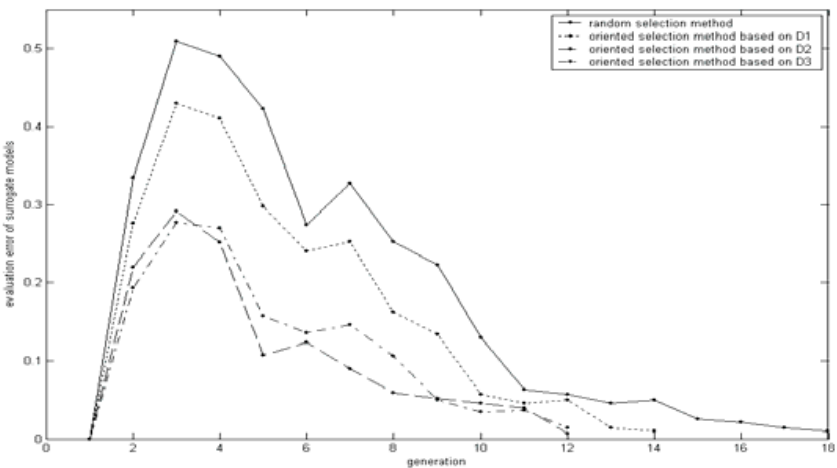


Fig. 3. Evaluation error of surrogate models

In each generation, different substituted individuals are selected when above four selection methods are respectively adopted. In order to compare the difference among random selection method and oriented selection methods, the number of dissimilar substituted individuals in each generation are noted, as shown in Table 6. Here, $N_{MD}(t)$ denotes the number of dissimilar substituted individuals.

generation		6	7	8	10	11	12	15	18
random selection	$N_M(t)$	0	0	0	0	4	4	4	4
	$N_{MD}(t)$	0	0	0	4	3	2	2	-
oriented selection (D1)	$N_M(t)$	0	0	0	4	4	4	4	-
	$N_{MD}(t)$	0	0	0	4	3	2	2	-
oriented selection (D2)	$N_M(t)$	0	4	4	4	4	-	-	-
	$N_{MD}(t)$	0	4	4	4	3	-	-	-
oriented selection (D3)	$N_M(t)$	0	0	4	4	4	4	-	-
	$N_{MD}(t)$	0	0	4	4	3	3	-	-

Table 6. Comparison of dissimilar substituted individuals selected by different methods

In a word, the performances of oriented selection methods are better than random selection method. Through comparison of three kinds of oriented selection methods, oriented selection methods using D2 or D3 has better performance and stability, yet larger computation complexity. On the contrary, computation complexity of oriented selection methods using D1 is less, whereas the algorithm’s performance is worse. Considering multi-mode property of human preference and small population size in IGAs, oriented selection methods using D2 is more fit for IGAs than others.

7.7 Comparison of selection methods with different cluster number

Experiments adopting AES-IGAs with different cluster number are done. Here, two fixed cluster number including 3 or 6 and adaptive cluster number are adopted. Testing results are shown in Table 7.

L		3		6		adaptive	
measures		T	N _U	T	N _U	T	N _U
random selection		18	112	17	108	15	100
oriented selection (D1)	value	14	92	12	80	11	72
	comparison with random selection	↓22.2%	↓17.9%	↓29.4%	↓25.9%	↓26.7%	↓28.0%
oriented selection (D2)	value	12	76	11	76	12	76
	comparison with random selection	↓33.3%	↓32.1%	↓29.4%	↓29.6%	↓20.0%	↓24.0%
oriented selection (D3)	value	12	80	12	84	11	76
	comparison with random selection	↓33.3%	↓28.6%	↓35.3%	↓22.2%	↓26.7%	↓24.0%

Table 7. Comparison of selection methods with different cluster number

By compare the results, we know that: (I) AES-IGAs using adaptive cluster number have best performance. Because cluster number is more reasonable, the clustering precision of individuals is better. This will speed up convergence. (II) No matter which kind of approximate distance is adopted, oriented selection methods converge faster than random selection method. And oriented selection method adopting D2 has better stability. The reason for this is individuals in each cluster contains more information than cluster center, which makes evaluation error less. Hence, surrogate model can participate in evaluation earlier so as to alleviate human fatigue at most.

In order to analyze the influence of selection methods and cluster number on human fatigue, the curves describing the degree of human fatigue are shown in Figure 4. Note that the threshold of human fatigue is 0.95 in these experiments.

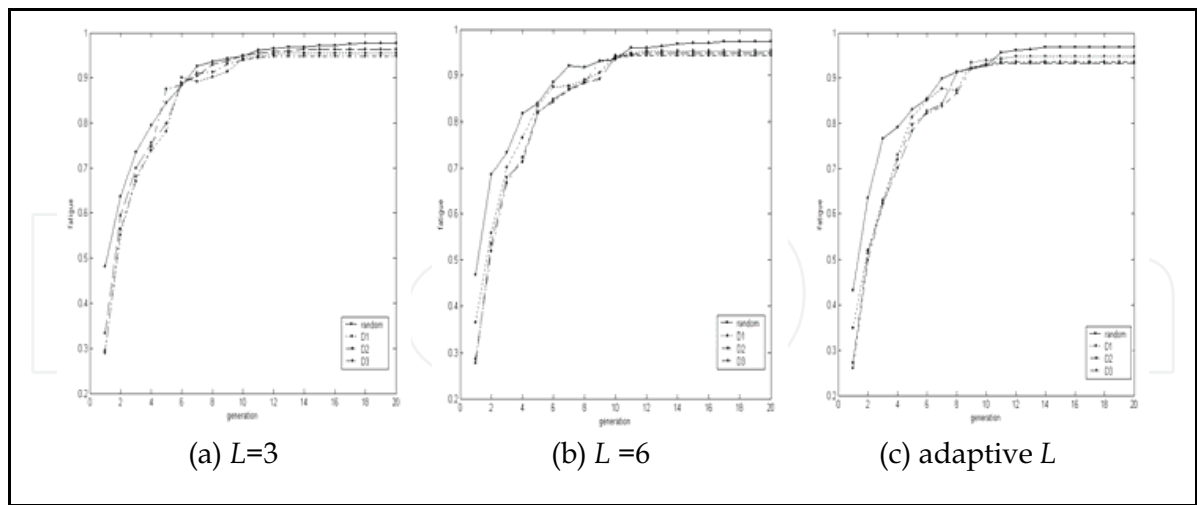


Fig. 4. The trend of human fatigue

Obviously no matter which kind of selection methods are used, human fatigue all can be alleviated as long as cluster number is reasonable. Especially the degree of human fatigue is less if oriented selection methods using D2 or D3 are adopted.

7.8 Comparison of performance about IGAs

In order to validate the improvement in performance of IGAs with adaptive evaluation strategy, thirty persons are gathered. Everyone do experiments aiming at both desired objectives adopting conventional IGA and AES-IGA respectively. Testing results for above four experiments done by all persons are integrated, as shown in Table 8.

experiments	I		II	
algorithms	IGA	AES-IGA	IGA	AES-IGA
average generation	14	7	31	12
average number of individuals evaluated by human	112	38	248	55
average number of individuals evaluated by human in each generation	8	5	8	5
generations when $Fa(t) \geq \epsilon$	-	-	-	8

Table 8. Comparison of the performance between IGAs and AES-IGAs

Through comparison of testing results in experiment I, generation adopting AES-IGA averagely reduces 52.1% than IGA. And total number of individuals evaluated by human adopting AES-IGA averagely reduces 71.5%. These indicate adaptive evaluation strategy can effectively alleviate human fatigue and speed up convergence.

8. Conclusion

In order to further alleviate human fatigue in interactive genetic algorithms, surrogate models are used to evaluate individuals instead of human. When to utilize surrogate models, how many and which individuals are calculate fitness values adopting the models are studied. Startup mechanism about surrogate models considering the degree of human fatigue and the models' precision is given. Variable proportion of population evaluated by surrogate models is proposed. Three phases including evaluated by human only, mixed evaluated by human and the models, evaluated by the models only, are discussed according to the number of substitution individuals. Especially, population size is enlarged in third phase. Surrogate model-based evaluation strategy with oriented selection of substituted individuals is proposed. Three kinds of approximate distance are presented so as to judge which cluster individuals belongs to. At last, convergence of AES-IGAs is proved.

Taking fashion evolutionary design system as a testing platform, the rationality of adaptive evolution strategy is validated aiming at different psychological requirements of human. Testing results indicate AES-IGAs converge faster than canonical IGAs and human feel less tired. Therefore, AES-IGAs can effectively alleviate human fatigue and improve the convergence speed so as to reduce human burden for evaluation which makes human absorbed in more creative design work.

9. Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. 60805025) and the 863 Project of China (Grant No. 2007AA12Z162).

10. References

- Biles J A.;Anderson P G; Loggi L W. (1996). Neural network fitness functions for a musical IGA, *Proceedings of the International ICSC Symposium on Intelligent Industrial Automation and Soft Computing*, pp. 39-44
- H.Takagi. (1998). Interactive evolutionary computation: system optimization based on human subjective evolution. *Proceedings of IEEE Conference on Intelligent Engineering System*, pp. 1-6
- H.Takagi.; Sung-Bae Cho; Toshihiko Koda. (1999). Evaluation of an IGA-based Image Retrieval System Using Wavelet Coefficients. *IEEE International Fuzzy Systems Conference Proceedings*, pp. 1775-1780
- Zhou Yong; Gong Dun-wei; Hao Guo-sheng; et al. (2005). Neural network based phase estimation of individual fitness in interactive genetic algorithm. *Control and Decision*, Vol.20, pp. 234-236
- Wang S F; Wang S H; Wang X F. (2003). Improved interactive genetic algorithm incorporating with SVM and its application. *Journal of Data Acquisition & Processing*, Vol.18, pp.429-433
- Yi-nan Guo; Dun-wei Gong; Wang Hui. (2007). Adaptive evaluation strategy based on surrogate model. *Lecture Notes in Computer Science*. Vol.4550, pp. 472-481

- Guo Yi-nan; Gong Dun-wei. (2007). Extraction and utilization about knowledge in hierarchical interactive genetic algorithms. *Control and Decision*, Vol.12, pp. 1329-1335
- Karayiannis, N. B.; Bezdek, J. C. (1997). Integrated approach to fuzzy learning vector quantization and fuzzy c-means clustering. *IEEE Transactions on Fuzzy Systems*, Vol.5, No.4, pp. 622-630
- Cheng Jian; Guo Yi-nan; Qian Jian-sheng. (2006). A multiple neural network architecture based on fuzzy c-means clustering algorithm. *Proceedings of 6th World Congress on Intelligent Control and Automation*, pp. 1875-1878
- Cheng Jian; Guo Yi-nan; Gong Dun-wei; et al. (2009). Surrogate model-based evaluation strategy with non-random selection of substituted individuals. *Chinese Journal of Electronics*, No.1, pp. 190-195
- He Jun; Yao Xin. (2001). Drift analysis and average time complexity of evolutionary algorithms. *Artificial Intelligence*.
- Kim H.; Cho S.B. (2000). Application of interactive genetic algorithm to fashion design. *Engineering Applications of Artificial Intelligence*, Vol.13, pp. 635-644

IntechOpen

IntechOpen

IntechOpen



Human-Computer Interaction

Edited by Inaki Maurtua

ISBN 978-953-307-022-3

Hard cover, 560 pages

Publisher InTech

Published online 01, December, 2009

Published in print edition December, 2009

In this book the reader will find a collection of 31 papers presenting different facets of Human Computer Interaction, the result of research projects and experiments as well as new approaches to design user interfaces. The book is organized according to the following main topics in a sequential order: new interaction paradigms, multimodality, usability studies on several interaction mechanisms, human factors, universal design and development methodologies and tools.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yi-nan Guo and Jian Cheng (2009). Adaptive Evaluation Strategy Based on Surrogate Models, Human-Computer Interaction, Inaki Maurtua (Ed.), ISBN: 978-953-307-022-3, InTech, Available from: <http://www.intechopen.com/books/human-computer-interaction/adaptive-evaluation-strategy-based-on-surrogate-models>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen