Evolving 3D Facial Expressions Using Interactive Genetic Algorithms

Meareg Hailemariam[1,2,3](✉), Ben Goertzel[1,2], and Tesfa Yohannes[1,2]

[1] Hanson Robotics, Hong Kong, China
{meareg,tesfa}@hansonrobotics.com, ben@goertzel.org
[2] SingularityNET, Hong Kong, China
{meareg,ben,tesfa}@singularitynet.io
[3] Addis Ababa University, Addis Ababa, Ethiopia

Abstract. Interactive Genetic Algorithms (IGA) are applied in optimization problems where the fitness function is fuzzy or subjective. Its application transcends several domains including photography, fashion, gaming and graphics. This work introduces a novel implementation of Interactive Genetic Algorithm (IGA) for evolving facial animations on a 3D face model. In this paper, an animation of a facial expression represents a chromosome; while genes are equivalent, depending on the crossover method applied, either to a keyframe point information (f-curve) of a facial bone or f-curves of grouped sub-parts such as the head, mouth or eyes. Crossover techniques uniform, cut-and-spice, blend and their hybrids were implemented with a user playing fitness function role. Moreover, in order to maximize user preference and minimize the user fatigue during evolution, sub-parts based elitism was implemented. Subjective measurements of credibility and peculiarity parameters among a given artist animated and evolved expressions were done. For the experiment results here, an average crossover percentage of 85%, a mutation level of 0.01, initial population of 36, and 8 rounds of evolution settings were considered. As detailed in the experiment section, the IGA based evolved facial expressions scored competitive results to the artist-animated ones.

Keywords: Evolutionary algorithms · Interactive genetic algorithms · 3D facial expressions

1 Introduction

People show different facial expressions when expressing different kinds of emotions. Even though the number and type of emotions along with their associated facial expressions differ from person to person, moment to moment, there are six basic set of emotions recognized globally. These six basic emotions are angry, disgust, fear, happiness, sadness, and surprise [1]. However, there are far more set of emotions and associated facial expressions that people are capable of feeling and expressing in their daily life.

F. A. Zimale et al. (Eds.): ICAST 2018, LNICST 274, pp. 492–502, 2019.
https://doi.org/10.1007/978-3-030-15357-1_40

In spite of heterogeneity in the interpretation, facial expressions are globally used by most people as a primary means to express emotions. In the virtual world, animated or virtual character, express emotions via facial expressions. The recent advancement of 3D facial models, avatar bots and humanoid robots which have various applications across multiple domains raise the demand and expectation of users/audiences for more human like facial expressions. Hence, facial expressions are expected to be more expressive, have better subtlety and more variation.

Even if there is a huge interest in generating realistic or novel facial expressions on 3D animation models, creating a new facial animation is not an easy task. This is, partly due to the fact that animation techniques in general tend to follow an ad-hoc and inextensible approach [2]. These factors cause limitation on the performance in generating new and realistic facial animation expressions.

In this work, we explore the application of evolutionary algorithmic approach to achieve partial automation in generating 3D facial expressions. We particularly used one category of evolutionary approach called Interactive Genetic Algorithm (IGA) which is a form of genetic algorithm that demands a human involvement in the evolution loop as a fitness function; a property which compliments the subjective metrics nature of facial expressions.

2 Applications of IGAs

The broad category of IGAs, or Interactive Evolutionary Computation (IEC), in general has application in industrial design, speech processing and synthesis, data mining, image processing, education and in artistic vocations such as graphic arts, animation, music and much more [3]. Specifically, IGAs have been applied in photography, fashion, gaming, virtual reality and facial animation.

In relation to face/body animation or photography, IGAs have been used: (1) to change facial expressions by changing pixel positions, (2) to create animated graphic art by evolving mathematical equations that apply to the pixel attributes, (3) to create animations by evolving the combination of joint angles for arms and legs, (4) for evolving deformations of a 2D body for comical movements, (5) with 2D photos of partial images to compose a facial image for identifying a criminal suspect [4]. Facial animation is also an important research goal in human-computer interaction, as in the quest to build a believable Embodied Conversational Agent (ECA). These agents would be able to communicate complex information with human-like expressiveness. ECAs are becoming popular as front ends to web sites, and as part of many computer applications such as virtual training environments, tutoring systems, storytelling systems, portable personal guides, and entertainment systems [5].

3 Related Work

Kim and Cho [5] used IGA to evolve fashion design clothes. In this scenario IGA-based evolution plays well; as the fashion industry has a changing trend and thus a human fitness function would ideally be able to influence the trajectory of the evolution procedure to get more appealing results. In their approach, they classified parts of a cloth into three parts: neck and body, arm and sleeve, skirt and waistline. Each of these three sub parts include color as their parameter; expanding the search space. These six parameters (three sub parts and their respective colors) are considered as genes which give new cloth results via IGA based combination. They used a population size of 8 and limited the number of maximum generation to 10. They have done convergence and subject tests which measure the fitness value changes and user satisfaction respectively on the generated fashion designs in terms of being cool-looking and splendor clothes criteria; and they achieved encouraging results in both measurements.

A dissertation paper of Smith in [4] applied IGAs to evolve facial expressions and used Neural Networks as a surrogate function to reduce user fatigue. While a chromosome is a face animation of expression, different from our approach, genes are equivalent to key frame sequences (fraction of duration) of a full face animation. Thus, both chromosomes and genes are basically the same except for the time length difference between them. In this type of setting, due to the nature of the genes here, a crossover would mean a mere arrangement of instances (splits of keyframes) of parent facial expressions in order to form new expression. A sample scenario that depicts this for instance would be; given a sad and happy expressions (as two parents)- generating a new facial expression (the child) would then be via crossover between multiple fractions of time/keyframe splits (genes) of the two parent expressions. Logically, the generated child expression is potentially going to be far from realistic as it would be constituted of keyframes sequences (genes) that jump from one type of parent expression/animation to the other back and forth prematurely. However, in our work we used a location/region, on the face of parent expressions, based crossovers instead of based on splices keyframe sequences of parent expressions where a sequence of keyframes from a single point on the face can be splitted into genes as done here.

Similarly, the face image generation system called E-FIT applies Interactive Evolutionary strategy, which is in the broader category of evolutionary algorithms, for parameter optimization [11].

4 IGA Based Facial Animation Evolver

In a typical IGA design there are representations of population, chromosomes and genes. While a chromosome consists of genes, a population is a collection of chromosomes. In order to evolve a new children population, crossover and mutation of the chromosomes must happen in terms of the switching between and mutation of genes found in different chromosomes.

In this implementation, while a chromosome represents an individual facial animation, population is a collection of the different separate facial animations presented. A gene is equivalent either to a single location's keyframe or to keyframes of a group of subparts of the face facial animation depending on the type of crossover used. The 3D facial model used for experiment uses both bone and morph driven animation [8]. The keyframe sequences interpolation was automated with a function of time called FCurve (f-curve) which is similar to bezier curves interpolation technique but with a modification which enforces any specific keyframe location to hold only a single value at a time for the purpose of doing animation/transformation [9]. The keyframe information of is treated as a sub-phenotype input since a 3D facial expression is represented via a set of keyframes while the set of bones that drive the facial morphing to generate expressions are considered the genotypes.

In this work, the crossovers of IGA refer the breeding of bone locations/values, or interchangeably referred here as fcurves, between the corresponding sequence of keyframes of the two parent expressions a new/child expression. In our case, a child expression is generated only from two parents. All used expressions have a duration of 10 s in a 30 keyframes per second rate. This duration uniformity allows crossover operations, between facial bone locations (fcurves), across each corresponding or same-indexed keyframes of both parents. Generated Child generation expressions too have similar keyframes duration and follow the same general rule of breeding. The genetic crossover process to generate a child can be put in a simplified format as:

Child Expression C = Parent Expression A<Crossover Operator>Parent Expression B. Further more this notation can be further decomposed into keyframes level as shown below.

C's keyframe ***1*** *= A's keyframe 1 <Crossover Operator> B's keyframe* ***1***
C's keyframe ***2*** *= A's keyframe 2 <Crossover Operator> B's keyframe* ***2***

C's keyframe ***n*** *= A's keyframe* ***n*** *<Crossover Operator> B's keyframe* ***n***

Fig. 1. Abstraction of genes crossover in a parallel keyframes level of parents (A an B) to generate genes in similar keyframes index of expression C (child).

Though in Fig. 1 roughly shows that crossover between parents happen at keyframes level, the exact procedure of generating new locations (fcurve values) for all the facial bones of the child expression depends on the type of the crossover operator used. But it can be put in a generic notation of:

C's Facial Bone's locations at keyframe n = A's Facial Bone's locations at keyframe n<Crossover Operator>B's Facial Bone's locations at keyframe n.

GA has different types of crossover operators. In our case we experimented with the uniform, cut-and-splice, blend and their hybrids. These crossovers are widely applied in many GA and IGA works. Implementation wise, we incorporated elitism to enable partially controlled breeding which helps in evolving sensible facial animations in some cases.

In GA, mutation is applied to weak solution candidates. It helps prevent the population from getting stuck in a local optima by introducing some diversity. In our experiments a default value of 0.01 mutation degree was applied only to facial animations which were rated lowest during the generation process.

4.1 Uniform Crossover

This crossover is based on separate selection mechanism of genes from both parents. One way to do uniform crossover is to randomly select genes from parents. In the case when the gene from the first parent is not selected, the corresponding location value of the facial bone (gene) of the second parent will be inserted into the child instead. Selection of all of the genes (locations of all included facial bones) in a single keyframe is done in a loop of choose either of the corresponding genes in the similarly indexed keyframes of the two parents. This same technique is followed across the rest of keyframes in both parents.

4.2 Cut-and-Splice Crossover

In this case the crossover can be generally assumed as a version of two-point crossover but where there is a fixed point of selection and also where each segment contains location values of more than one facial bones instead of a single bone's. Potentially, this lowers down the possible active exchange of fcurves during a crossover. On the other hand, since the three slicing point segments used are the sub-parts of the face (Head, mouth and eyes areas), this increases the likelihood of generating more human-like expressions due to the fact that these sub-parts are treated as indivisible units (genes) during crossover. This assumption was validated by the experiment.

In order to increase flexibility of generation and increase solution space, six different kinds of choosing combination of the three sub-parts were offered. So it is possible to choose which one the genes or two of the genes we want to be selected from the first parent (the reverse is applied on the other parent automatically). These are 'Head', 'Mouth', 'Eyes', 'Head-Mouth', 'Head-Eyes' and 'Mouth-Eyes'. For example if the 'Head' option is chosen, the child expression will have its head animation from its first parent and, mouth and eyes animations from the other parent. Or if the 'Head-Mouth' option is selected, the child's genetic make up of head and mouth animations will be taken from its first parent and eyes animations from their second parent and so on for the rest of the options. This feature adds the capability of seeing six different kinds of children chromosome population within a single crossover. This approach provides a short cut to explore the search space extensively.

4.3 Blend Operator

This is a crossover type which in general sums two points and returns their average point as an output; a blended output. Each f-curve of a parent facial expression is summed to its corresponding f-curve of the other parent and then summed result is averaged. Thus, the child expression will be comprised of averaged f-curves of its parents.

4.4 Elitism: By Retaining Interesting Facial Animation Sub-parts from the Current Population Members

In IGAs, minimizing the user fatigue involved when generating facial animations is critical. Elitism greatly speeds up and enhances the quality of the overall population by retaining chosen sub-parts from some members' in the current population and to be incorporated in the next generation without the potential loss of them through crossover or mutation. In our implementation we have incorporated a feature that enables selection of one or more sub-parts (head, mouth and eyes) of parent expressions to be transferred to the next generation without potential loss or change due to crossover and mutation.

4.5 Search Space

A search space indicates the number of possible combinations of the different models of genes in a given population. We have sampled initial population of 36 during all the IGA based experiments. During the cut-and-splice, crossover happens in terms of defined sub parts (head, eyes and mouth area) of the face animation. Thus, a gene is equivalent to a head, mouth, or eye animation which themselves are composed of multiple f-curves. There are 9 different models of mouth movements, 22 different models of eye movements and 24 different models of head movement out of 36 total initial chromosome population/facial animations models. Six different kinds of breeding combinations were offered during the cut-and-splice crossover. Thus, the total size of initially accessible search space is 9*22*24*6 = 28,512.

On the other hand, in case of uniform and blend crossovers, search space is very much larger as a gene is equivalent to single f-curves in a given facial animation. Thus, it is reasonable to imagine that the search space is very big. But in reality, this doesn't mean that it gives exaggerated diversity in evolved facial animations. The main reason for this is mostly the difference of animation between each corresponding f-curves of the different the given facial animations might not always be that big/obvious except for the active facial areas of the given expressions. Corresponding f-curves of another facial animation might not be that big/obvious unless the two facial animations are quite different. These conditions have direct effect on the diversity of newly generated animations despite the deceptively large looking search space.

5 Experiment and Result Summary

As an experiment platform the blender 3D engine [7] was used. It supports python scripting interface which allows programmatic access the graphics objects on the tool. Further more the 3D face model we used for experimentation also provides a python based API which is suitable for extension and modification. User satisfaction based subjective measurement for credibility and peculiarity parameters on the evolved 3D facial expressions, via the different operators discussed, and also on artist animated ones was done. Five subjects were used for rating the provided facial expressions in a specified questionnaire formats. Three of the subjects work in research and technology (computer science) area; the fourth one is an artist level animator while the last one is a registered nurse. The first four are familiar with the goal of the work. All of the candidates were allowed to finish the questionnaire in a period of two weeks in their own pace. The potential inconsistencies of metrics that could be caused due the degree of familiarity of the subjects with technology or the aim of the research or mood changes are not considered.

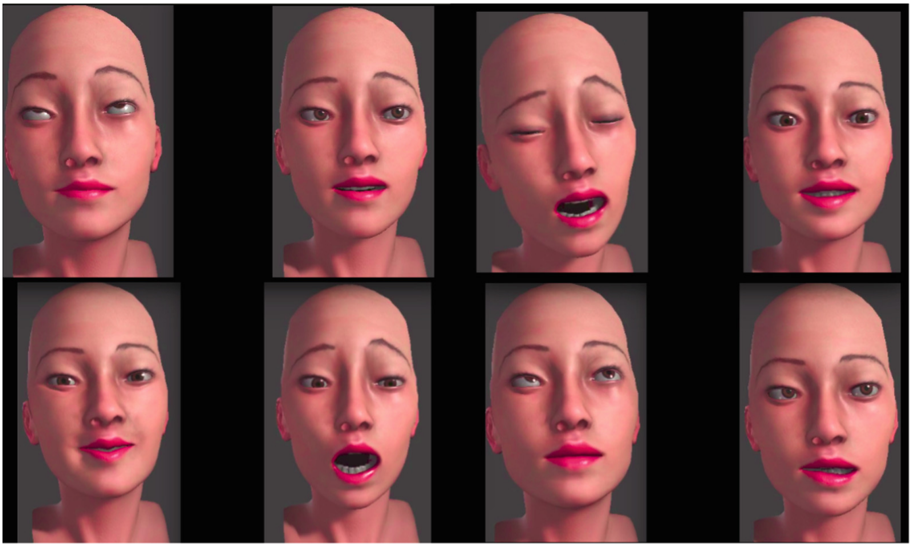

Fig. 2. Some screen captures of IGA based evolved 3D facial expressions of the model used.

5.1 Credibility

The credibility parameter in this experiment refers to the measurement of the facial animations in terms of their degree of acceptance as potential human expression like. Figure 3 shows subjective credibility score for expressions evolved using uniform, cut-and-splice, blend, uniform-blend hybrid and artist animated 3D facial expressions.

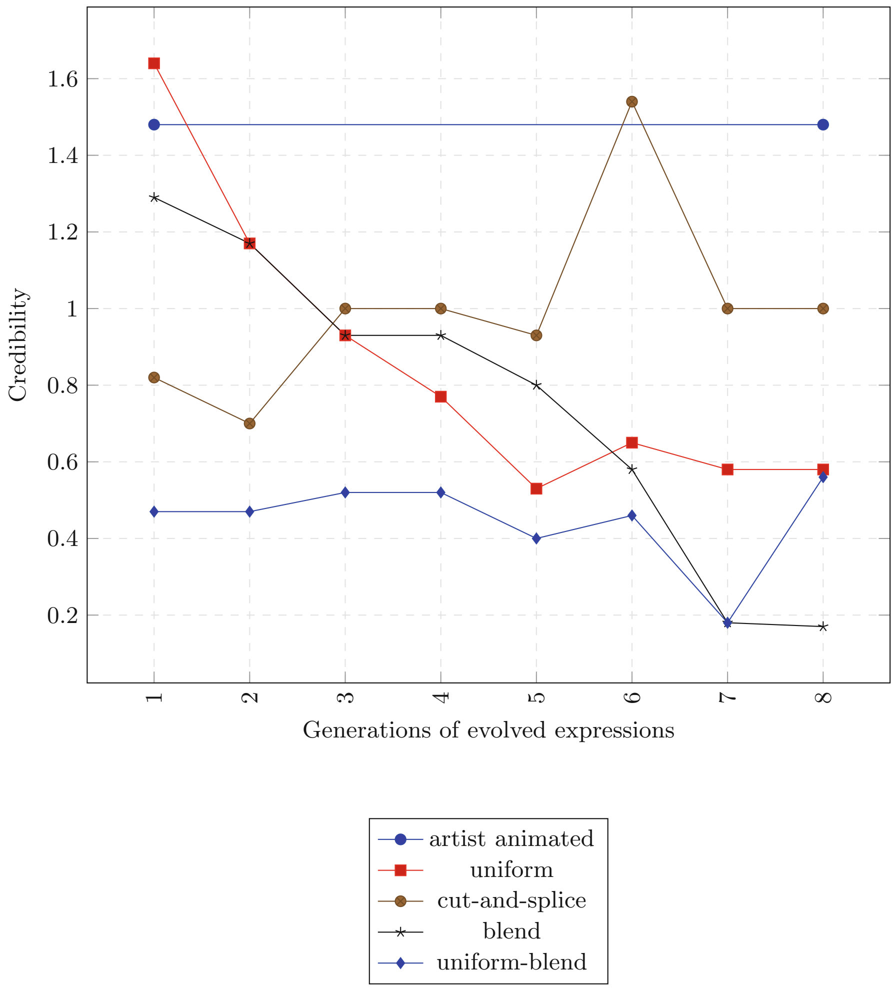

Fig. 3. Credibility score for the different IGA crossovers based evolved and artist-animated expressions.

Overall Average Credibility

Artist-animated	Uniform	Cut-and-Splice	Blend	Uniform-Blend
87%	71.4%	74.9%	68.9%	61%

5.2 Peculiarity

The peculiarity parameter in this experiment refers to the measurement of the facial animations in terms of their degree of uniqueness or distinctiveness. A facial expression despite having a less human-like appeal, it can still have a higher distinctiveness.

Figure 4 shows subjective peculiarity measurement of expressions evolved using uniform, cut-and-splice, blend, uniform-blend hybrid and artist animated 3D facial expressions.

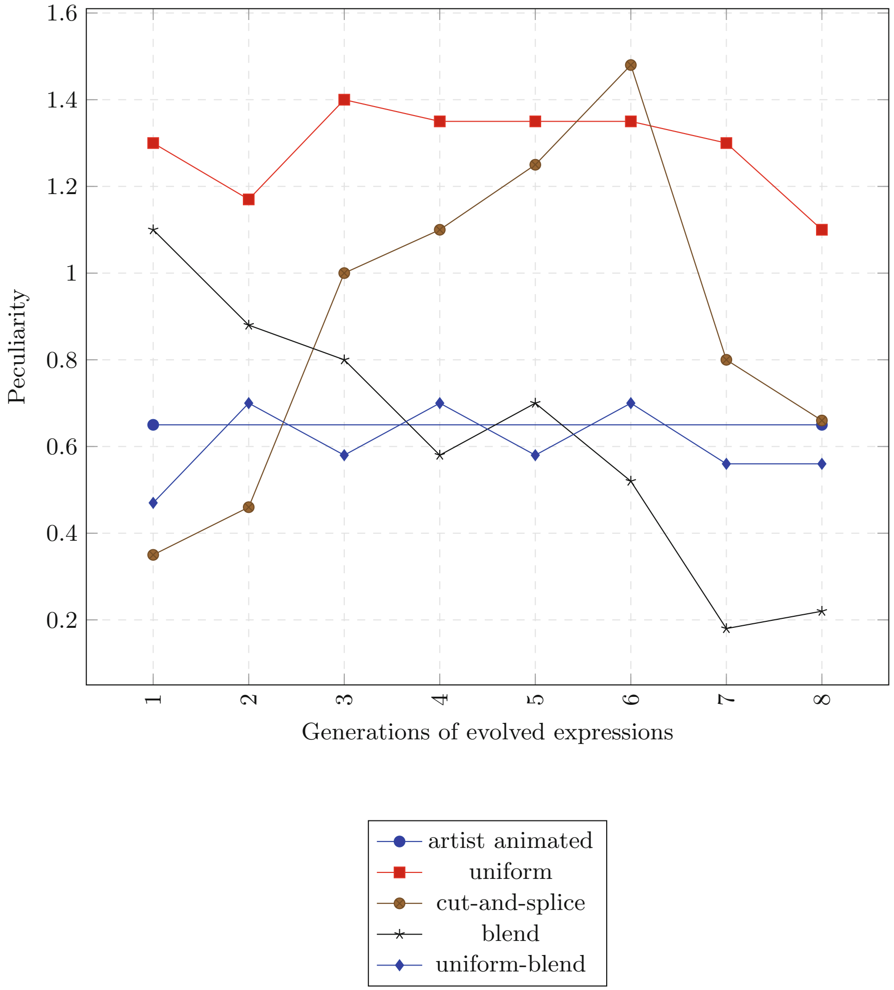

Fig. 4. Peculiarity score for the different IGA crossovers based evolved and artist-animated expressions.

Overall Average Peculiarity

Artist-animated	Uniform	Cut-and-Splice	Blend	Uniform-Blend
66.25%	82.25%	72%	65.5%	65%

5.3 Retain Interesting Sub-parts of Current Parent Expressions During the Next Generation

The sub-parts keeping feature (elitism)used was experimented with the uniform based evolver. All IGA parameters such as probabilities of crossover and mutation were kept to the same level as used during the other pure operators based generation experiments.

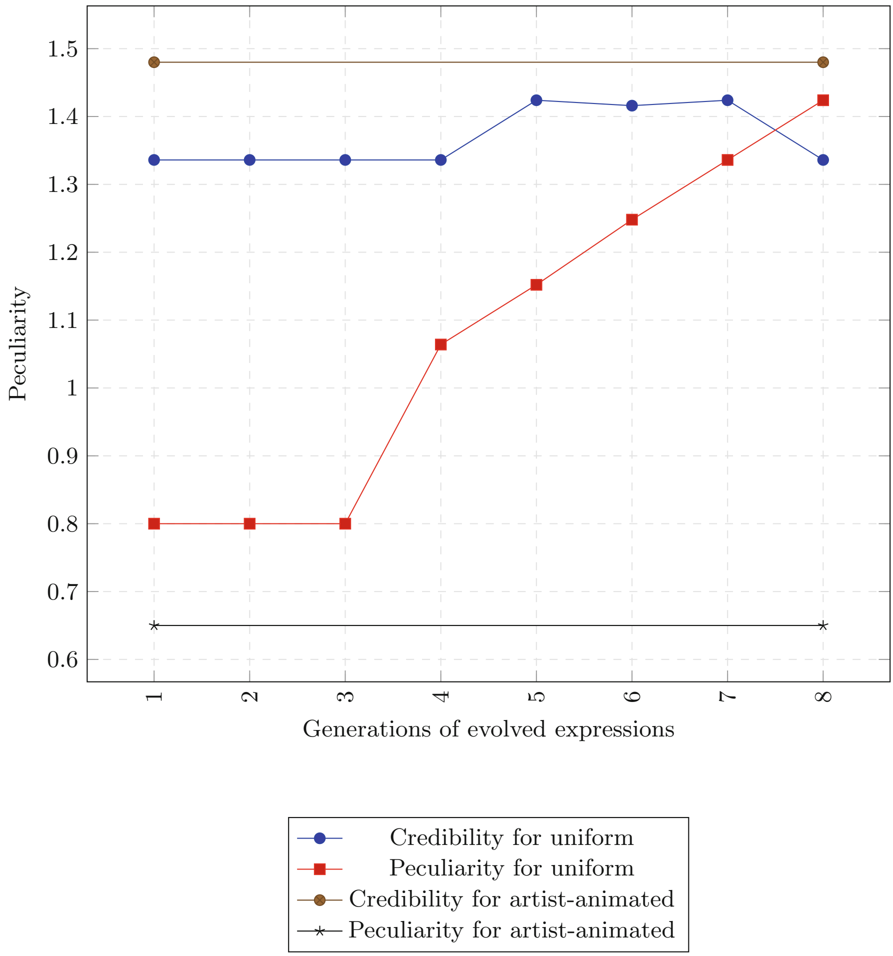

Fig. 5. Credibility and peculiarity score for the uniform operator with sub-parts keeping feature based evolved expressions in comparison to artist-animated expressions.

Figure 5 shows the credibility and peculiarity scores for the elitism based uniform evolver and also artist-animated expressions just for comparison purpose.

This feature of retaining interesting features members of the current population to the next ones resulted in a general quality increase in lesser number of generations. The uniform based breeding with interesting sub-parts keeping feature improved to 84.2% and 76.95% scores for the credibility and peculiarity parameters respectively.

6 Conclusion

This paper has shown that IGAs can be competitively useful in generating credible and quality 3D facial expressions. In particular, it showed the use of keyframe point information (f-curve) as an IGA gene and demonstrated its usefulness in evolving realistic and peculiar expressions.

References

1. Sauter, D.A., Eisner, F., Ekman, P., Scott, S.K.: Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. Proc. Natl. Acad. Sci. **107**, 2408–2412 (2010)
2. Parke, F., Waters, K.: Computer Facial Animation, 2nd edn. AK Peters Ltd., Wellesley (2008)
3. Takagi, H.: Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation. Proc. IEEE **89**(9), 1275–1296 (2001)
4. Nancy, S.: Evolving Credible Facial Expressions with Interactive GAs (2012)
5. Kim, H.-S., Cho, S.-B.: Application of interactive genetic algorithm to fashion design. Eng. Appl. Artif. Intell. **13**, 635–644 (2000)
6. Mana, N., Pianesi, F.: HMM-based synthesis of emotional facial expressions during speech in synthetic talking heads. In: Proceedings of the 8th International Conference on Multimodal Interfaces, pp. 380–387 (2006)
7. Blender 3D Engine: www.blender.org. Accessed 20 June 2018
8. Chen, L.: An Analysis of the Current and Future State of 3D Facial Animation Techniques and Systems. Communication University of China, B.A. (2009)
9. Farin, G.: Curves and Surfaces for CAGD: A Practical Guide, 5th edn. Morgan Kaufmann Publishers Inc., San Francisco (2002)
10. Kurt, B., et al.: Active appearance model-based facial composite generation with interactive nature-inspired heuristics. In: Gunsel, B., Jain, A.K., Tekalp, A.M., Sankur, B. (eds.) MRCS 2006. LNCS, vol. 4105, pp. 183–190. Springer, Heidelberg (2006). https://doi.org/10.1007/11848035_26
11. George, B., et al.: EFIT-V - interactive evolutionary strategy for the construction of photo-realistic facial composites. In: Genetic and Evolutionary Computation Conference Proceedings, Atlanta, GA, USA (2008)
12. Blender Manual. https://docs.blender.org/manual/en/dev/modeling/curves. Accessed 10 June 2018