

Active Appearance Model-Based Facial Composite Generation with Interactive Nature-Inspired Heuristics

Binnur Kurt, A. Sima Etaner-Uyar, Tugba Akbal, Nildem Demir,
Alp Emre Kanlikilicer, Merve Can Kus, and Fatma Hulya Ulu

Istanbul Technical University, Department of Computer Engineering, Istanbul, Turkey
{kurt, etaner}@ce.itu.edu.tr

Abstract. The aim of this study is to automatically generate facial composites in order to match a target face, by using the active appearance model (AAM). The AAM generates a statistical model of the human face from a training set. The model parameters control both the shape and the texture of the face. We propose a system in which a human user interactively tries to optimize the AAM parameters such that the parameters generate the target face. In this study, the optimization problem is handled through using nature-inspired approaches. Experiments with interactive versions of different nature-inspired heuristics are performed. In the interactive versions of these heuristics, users participate in the experiments either by quantifying the solution quality or by selecting the most similar faces. The results of the initial experiments are promising which promote further study.

Keywords: Computerized facial composite generation, active appearance model, nature-inspired heuristics, interactive nature-inspired heuristics.

1 Introduction

In crime investigations, the forensic artist usually interviews the witness and draws a sketch of the criminal's face based on the interview. Mainly there are three different approaches used for this purpose. In the first approach, a forensic artist draws the face as the witness verbally describes it. The second approach uses a computer based system such as E-FIT [1] and PROfit [2]. These systems require the selection of individual facial components (eyes, nose, mouth etc.) from a database. The witness first chooses these facial features and then specifies their positions on the face with the help of a trained operator. The third approach uses a computer based face generation tool such as Evo-FIT [3] and Eigen-FIT [4]. Good performing examples of these tools use evolutionary algorithms to generate different facial composites. At each stage of the algorithm the user is asked to assign scores to or to rank each face, based on its similarity to the target face. The first two approaches strongly depend on the current psychological and emotional state of the witness. The witness is not only required to recall the face but also he/she needs to give an accurate description of each of the facial features. These methods also require a manual modification on the local facial features to make the face look more like the target. Moreover, the sketch artist plays an important role in the quality of the final sketch. The requirements above are

also psychologically challenging. Usually people are more likely to recall the face as a whole rather than to recall individual features separately. The third group of approaches does not have these drawbacks mentioned above. The methods in this class generally make use of generative models such as eigenfaces to generate the face globally. Another important aspect of these systems is that the generated faces can be sent directly to any face recognition system without any further post processing.

In this study, we used an approach inspired from the two successful face generation tools: Evo-FIT and Eigen-FIT. Our main critical contribution is the fact that we experimented with several interactive nature-inspired heuristics to cope with the optimization in a huge search space, whereas Evo-FIT and Eigen-FIT use several variations of evolutionary-only algorithms. We aim to determine the approach which is more suitable for this task. To generate the candidate face images through an interface, we used the parameter vector of the active appearance model (AAM) [11]. We pose this problem as an optimization problem and used several interactive nature-inspired heuristics to obtain the AAM parameter vector representing the face most similar to the target. The nature-inspired heuristics used in this study are two versions of genetic algorithms (GA) [5], evolutionary strategies (ES) [15], particle swarm optimization (PSO) [16] and differential evolution (DE) [14]. We built a training face image database by photographing fifty-one people from our faculty, and created the AAM from this face database. After the implementation of each of the heuristics, we conducted tests to evaluate their performance. Human interaction is involved in the fitness evaluations or the selection stages. This is a subjective process since the user can assign different fitness values to the same faces at different evaluations. Moreover, evaluating many faces may cause the user to get exhausted and become careless. In such cases, the interactive versions of the algorithms may have poor performance than their non-interactive counterparts so it is important to test their interactive performances

This paper is structured as follows: in Section 2, the AAM will be presented. Section 3 will discuss related work. Brief descriptions of the implementations of the interactive nature-inspired heuristics will be given in Section 4. Section 5 will explain the methods used to test the system. Finally, Section 6 will discuss the conclusions and provide pointers for future work.

2 The Active Appearance Model

The Active Appearance Model (AAM) [11] extracts a statistical model of the human face from an input face image database. Using this model, it is possible to generate faces which may not even exist in the original database. Therefore, AAM is considered as a powerful generative model which is able to represent different types of objects. AAM works according to the following principle: A face image is marked with n landmark points. The content of the marked face is analyzed based on a Principal Component Analysis (PCA) of both face texture and face shape. Face shape is defined by a triangular mesh and the vertex locations of the mesh. Mathematically the shape model is represented as follows: $x = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]$.

Face texture is the intensities on these landmarks (color pixel values normalized to shape) and is represented with the formula $g = [g_1, g_2, \dots, g_n]$.

Face shape and texture are reduced to a more compact form through PCA such that

$$\begin{aligned} \bar{x} &= \bar{x} + \Phi_s b_s \\ \bar{g} &= \bar{g} + \Phi_g b_g \end{aligned}$$

In this form, Φ_s contains the t eigenvectors corresponding to the largest eigenvalues and b_s is a t -dimensional vector. By varying the parameters in b_s , the shape can be varied. In the linear model of texture, Φ_g is a set of orthogonal modes of variation and b_g is a set of grey-level parameters. To remove the correlation between shape and texture model parameters, a third PCA is applied on the combined model parameters such that

$$\begin{aligned} \bar{x} &= \bar{x} + \Phi_s W_s^{-1} Q_s c \\ \bar{g} &= \bar{g} + \Phi_g Q_g c \end{aligned} \quad \text{where } b = \begin{bmatrix} W_s b_s \\ b_g \end{bmatrix} \text{ and } c = \begin{bmatrix} Q_s \\ Q_g \end{bmatrix} c.$$

In this form, W_s is a diagonal matrix of weights for each shape parameter, allowing for the difference in units between the shape and the grey models; c is a vector of appearance parameters controlling both the shape and the grey-levels of the model. Q_s and Q_g are the eigenvectors of the shape and texture models respectively.

3 Related Work

Computer based facial composite generation programs can be summed up in two categories: Programs based on the selection of facial features from a database such as E-FIT [1] and PROfit [2][6] and programs that generate faces automatically using a nature-inspired heuristic, like Evo-FIT [3][7] and Eigen-FIT [4][8][9]. Since our approach falls into the latter category, we will only explain Evo-FIT and Eigen-FIT.

Evo-FIT uses an evolutionary algorithm approach. Initial faces are generated through a PCA shape and texture model as a whole. The user selects from a larger face set, a small number of faces which look most like the target. An evolutionary algorithm generates a new group of faces based on the selection of the user. Evo-FIT can also import hairstyles from a PROfit database, but hair is used as an external parameter and is not optimized by the EA. The selected hairstyle is applied to all faces during the EA run. With its interface to a photo editing software, Evo-FIT provides the option to alter the face images at run-time (e.g. move eyes closer, change shape of eye, etc). Good results comparable to those of E-FIT are reported in [7].

Eigen-FIT uses the AAM and an elitist GA. Three versions of Eigen-FIT are tested in [8] and [9]. The first one uses breeding between the elite individual and others in the population. At each generation, all offspring are rated between 1 and 10. The second version does not include a rating; best faces are selected from among the offspring. The last version uses breeding only between the best individual and another random individual from the population. One offspring is generated at each generation. The offspring and the best individual are shown to the user to choose the best one.

Eigen-FIT also allows external feature modification during the EA runs. Reports on Eigen-FIT have good results [8], [9].

4 The Facial Composite Generation System

The AAM software AAM-API [12] is used in the generation of faces from the AAM parameter vectors. At each stage of the interactive nature-inspired heuristic based (INIH) face generation system, the user selects/ranks the presented faces based on their similarity to the target face. This input is used in the iterative processes of the algorithms to generate new AAM parameter vectors corresponding to new faces. At each iteration, the new faces are displayed on the screen. If the user is satisfied with at least one of the faces, he/she stops the run of the algorithm. The AAM parameter vector corresponding to the selected picture is the solution generated by the approach.

4.1 The Interactive Nature-Inspired Heuristics

We used five different nature-inspired heuristics to produce the AAM parameter vector of the target face. These are interactive steady-state genetic algorithms (ISSGA) [13], interactive generational genetic algorithms (IGGA) [13], interactive particle swarm optimization (IPSO) [13] [17], interactive evolutionary strategies (IES) [13] and interactive differential evolution (IDE) [13]. All of these nature inspired heuristics are population based, thus they work with a group of potential solutions and at each iteration, through some operators, these potential solutions are updated. Iterations continue until some stopping criteria are satisfied. In the interactive versions used in this study, each iteration described below for the different approaches continue until the user is satisfied with the produced results.

IGGA is the interactive version of a generational genetic algorithm [5] where in each iteration a new population of individuals are created through selection and recombination operators. In our IGGA implementation, in each iteration the user assigns fitness scores to each face based on their similarity to the target face. The face with the highest score is the face which looks most like the target. Through binary tournament selection based on the assigned scores, 2-point crossover and Gaussian mutation, the new AAM vectors are generated. ISSGA is the interactive version of a steady-state genetic algorithm [5] using a replace-worst-parent replacement strategy where in each iteration only one new solution is created through selection and recombination operators. In our ISSGA implementation, selection of the two parents is random. Crossover and mutation are the same as in the IGGA but only one offspring is generated at each iteration. For each iteration, the user is presented with three images corresponding to the two parents and the offspring and is asked to select two images from among these. There is no scoring or ranking.

IPSO is the interactive version of the particle swarm optimization algorithm [16] in which, similar to a generational genetic algorithm, a new set of potential solutions are generated at each iteration. Particle locations and velocities are updated based on the best results achieved by each particle and the overall best results achieved by any particle. In our IPSO implementation based on the approach proposed in [17], the user determines the best solution of each particle and the overall best. For this purpose, the

user is shown the image represented by each of the current particles together with their overall best images and is asked to select the better one which now becomes the overall best image for that particle. At the end of the iteration, the user is shown the overall best of each particle and is asked to select the global best.

IDE and IES are the interactive versions of differential evolution [14] and evolutionary strategies [15] respectively. In evolutionary strategies which uses a self-adaptive mutation approach, at each iteration λ new solution candidates are generated from μ solution candidates through random selection of parent pairs, discrete recombination crossover on the parameter vectors and intermediary recombination crossover on the mutation step sizes and Gaussian mutation. In differential evolution, unlike the other approaches, mutation occurs before crossover. In the mutation step, the weighted difference between two solution candidate vectors is added to a third to obtain the mutated vector which then goes through crossover with the actual solution candidate vector to give the resulting offspring. Similar to evolutionary strategies, λ new solutions are generated from μ solution vectors. At the final stage of both, μ of the best solutions from among the λ are selected. In the interactive versions, the user only interferes at this selection stage and selects the μ images from among the λ images displayed on the screen.





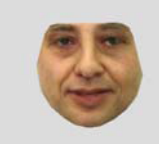
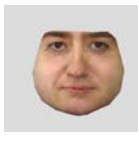


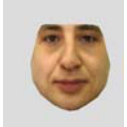
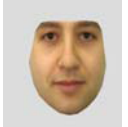
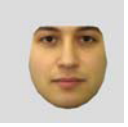
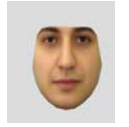
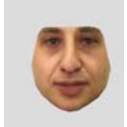
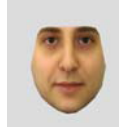
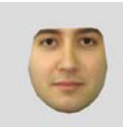
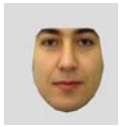
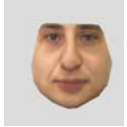

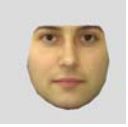
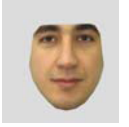
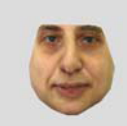

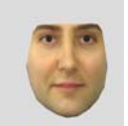
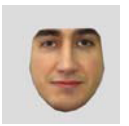
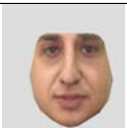
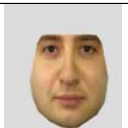
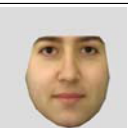
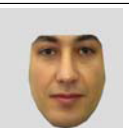
4.2 Implementation

We took fifty-one pictures to build our database and annotated the face images with AAM-API to compose the AAM. In our problem, chromosomes represent a set of AAM parameters which correspond to a face. These parameters are real numbers in the range $[-0.3, 0.3]$. The number of AAM parameters used in the model is 17. Therefore each chromosome has $n=17$ genes represented as real numbers. The initial population for all algorithms is generated randomly according to a Gaussian distribution with mean 0 and standard deviation 0.1. All other parameter settings for each algorithm are explained in detail in [13] by the same authors.

5 Experiments

The experiments are aimed to analyze the performance of the INIHs. To assess the performance of each approach, target images are selected based on whether the image is in the database or not and then a number of subjects are asked to run each algorithm for each selected target image. For each algorithm and each image, the AAM parameter vectors produced by the subjects are averaged to give a mean face image. Then these mean face images are shown to some test subjects other than those who produced the images. These subjects are asked to name the person in the image. Algorithm performances are evaluated based on the recognition rates by these subjects of the produced images and the Euclidean distance between the found solution and the optimum solution. We can easily obtain the optimum solution by projecting the target face into the face model space. Another factor which determines the usability of an approach in the interactive mode depends on the amount and the

Table 1. Original and generated face images

Original Images				
Projected Faces				
IPSO (62.7)	 100%	 100%	 90%	 50%
IDE (54)	 100%	 70%	 80%	 100%
IES (59.2)	 100%	 90%	 100%	 100%
IGGA (55)	 100%	 80%	 90%	 100%
ISSGA (67.65)	 100%	 100%	 90%	 100%

nature of interaction the user has to do. For example it is easier for a person to choose a subset of images from a larger set than scoring or ranking each picture presented to him/her. Also the convergence properties of the algorithms play an important role in the interactive mode. As the number of iterations the user has to make -thus the number of images to be evaluated- increases, the performance of the user will deteriorate because he/she will get tired and bored and will not pay attention properly.

The algorithms will also be evaluated based on the number of iterations required and the total number of images viewed by the subjects.

Currently the whole system has been built and made to work successfully. The parameters are determined intuitively and based on settings recommended in the related literature. The produced images can be seen in Table-1. In the table, the first column gives the name of the approach used for generating the faces in the corresponding row. Below each algorithm name, the average number of images viewed by the subjects when generating the images is also given. The original pictures of the target faces are given in the first row and the projected images into the model space are given in the second row. One may expect that these images are the best solutions we may get from any optimization algorithm since these images define the optimum. The first three images reside in the training set used to obtain the AAM. The fourth one is not in the training set. The images in each row correspond to the mean faces generated by each approach by the 7 subjects. The recognition rate of each photo by 10 subjects is also given below each image in the table. As can be seen, based on the average recognition rate, the ISSGA and IES approaches seem to be the most successful in our tests. Regarding the number of images viewed by the user, the ISSGA approach seems to be the worst while IDE and IGGA seems to be the best. But we have to also consider the fact that for IGGA, the user has to rank all the images while for IES and IDE, a subset of images needs to be selected from a larger set. As for ISSGA and IPSO, still there is no ranking or scoring but the selection process is more tedious than both IES and IDE. We also plot the quantitative results in Fig.1 showing how much the algorithms get closer to the optimum solution in Euclidean sense.

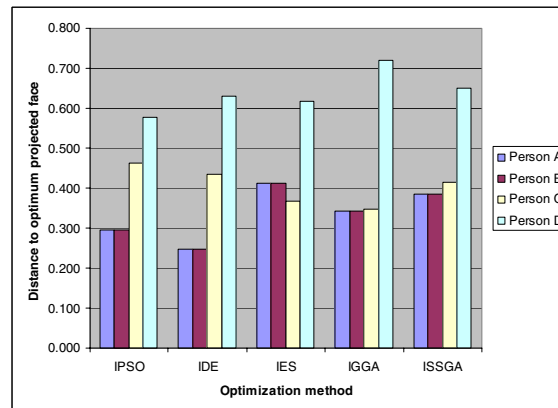


Fig. 1. Performance evaluation based on distance to the optimum projected face

6 Conclusion and Future Work

This work presents a preliminary study. Therefore it aims to give an overall performance analysis and an idea about the applicability of each of the approaches. The parameter settings of nature-inspired heuristics play a crucial role in their

performance. For this study, all parameters are determined mainly based on recommended values in literature. However, an experimental study should be performed to determine the optimal settings which would improve performance. Another factor which would also improve performance is to increase the size of the training set. Also, several enhancements should be added to the system to make it compete with similar projects in literature, namely E-FIT and Eigen-FIT, such as making face properties like size, shape, placement editable during run-time, building the AAM models of individual facial features separately and thus allowing freezing of a facial component during the process while the other components are allowed to change, adding hair style, moustache, beard, etc and also accessories.

The test results are sufficient to show that this is a feasible system and that the interactive versions of the nature-inspired heuristics we have chosen to implement are suitable for the problem. These results promote further study.

References

- [1] "E-FIT", New England Press Inc., 2004, <http://www.efitforwindows.com>
- [2] "PROfit", ABM UK Ltd., <http://www.abm-uk.com/uk/products/profit.asp>
- [3] Frowd C.D., Hancock P.J.B., "EvoFIT: Facial Composite System for Identifying Suspects to Crime", Department of Psychology, Sterling University.
- [4] "EigenFIT", VisionMetric Ltd., <http://www.eigenfit.com>
- [5] Eiben A.E., Smith J.E., Introduction to Evolutionary Computing, Springer, 2003.
- [6] "PROfit: A Photofit System using Highly Advanced Facial Composition Tools", ABM United Kingdom Ltd., <http://www.abm-uk.com/uk/pdfs/profit.pdf>
- [7] Frowd C.D., Hancock P.J.B., Carson D., "EvoFIT: A Holistic, Evolutionary Facial Imaging Technique for Creating Composites", *ACM TAP*, Vol. 1 (1), pp. 1-21, 2004.
- [8] Gibson S. J., Pallares-Bejarano A., Solomon C. J., "Synthesis of Photographic Quality Facial Composites using Evolutionary Algorithms", in Proceedings of the British Machine Vision Conference, pp. 221-230, 2003.
- [9] Solomon C. J., Gibson S. J., Pallares-Bejarano A., "EigenFit - The Generation of Photographic Quality Facial Composites", *The Journal Of Forensic Science*, 2005.
- [10] Eiben A. E., Schoenauer M., "Evolutionary Computing", *Information Processing Letters*, 82(1): 1-6, 2002
- [11] Iain Matthews, Simon Baker, "Active Appearance Models Revisited," *International Journal of Computer Vision*, Vol. 60(2), pp. 135-164, November 2004 .
- [12] "AAM-API", <http://www2.imm.dtu.dk/~aam/aamapi/>
- [13] Akbal T., Demir G. N., Kanlikilicer A. E., Kus M. C., Ulu F. H., "Interactive Nature-Inspired Heuristics for Automatic Facial Composite Generation", Genetic and Evolutionary Computation Conference, Undergraduate Student Workshop, July 2006.
- [14] Storn R., Price K., "Differential Evolution - A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces", Technical Report TR-95-012, International Computer Science Institute, Berkeley, CA, 1995.
- [15] Beyer H.G., Schwefel H.P., "Evolution strategies A comprehensive introduction", *Natural Comp.* 1: pp. 3-52, 2002.
- [16] Eberhart R. C., Kennedy J., Shi Y., *Swarm Intelligence*, M. Kaufmann, 2001.
- [17] Madar J., Abonyi J., Szeifert F., "Interactive Particle Swarm Optimization", 5th Int. Conf. on Intelligent Systems Design and Apps. (ISDA'05), pp. 314-319, 2005.