Measuring Facial Emotional Expressions Using Genetic Programming

A. Loizides, M. Slater and W. B. Langdon

Computer Science, University College, Gower Street, London, WC1E 6BT, UK
{A.Loizides,M.Slater,W.Langdon}@cs.ucl.ac.uk
http://www.cs.ucl.ac.uk/staff/A.Loizides, /staff/M.Slater, /staff/W.Langdon
Tel: +44 (0) 20 7679 7209, Fax: +44 (0) 20 7387 1397

Abstract. Genetic Programming techniques can be used to produce regression equations that quantify emotional expressions on a facial model. The formulae give emotional scores based on the position of 25 automatically generated "landmarks" on the face. The method shown here is an integrated part of a system that maps multidimensional data sets to naturalistic visual structures such as a face.

1 Introduction

[Loizides and Slater, 2001] have previously proposed a new way of visualising multivariate data sets using an automatic mapping to 3D face models. The Empathic Visualisation Algorithm, EVA, provides an automatic mapping from semantically important features of the data to emotionally or perceptually significant features of the visual structure (such as the face). It is a prerequisite of this method to be able to quantify individual facial expressions, like happiness, sadness, anger and fear.

Measuring facial expressions is important in several other applications. In Virtual Reality exposure therapy for social phobia for example, facial expressions are significant components of the evaluative feedback used to generate social anxiety [M. Slater, 1999].

Here we propose a Genetic Programming [Koza, 1992] (GP) based method for measuring universally recognised emotional expressions based on movements of certain points on the surface of the face.

Section 2 gives the background work to facial expressions, followed by Section 3 that explains how we measure emotional expressions. The GP system used is described in Section 4 and results are shown in Section 5. The conclusion is in Section 6.

2 Background

[Ekman and Friesen, 1978] developed the Facial Action Coding System (FACS). This system provides a notation for recording and describing the expressions of the face by considering the combinations of muscles that are used to create them.

Ekman uses action units or AU's such as the inner brow raiser and the upper lip raiser to describe muscles or groupings of muscles that perform specific actions in facial expressions. [Ekman, 1979] identified six key expressions of emotion which can be recognised across most cultures. These are neutral, anger, sad, happy, fear, surprise and disgust.

We can score the facial expressions by looking at the FACS system. We can use Ekman's description of what defines certain emotional expressions. For example, Ekman describes happiness as a contraction of the left and right zygomatic major muscles. From inspection we can set a contraction value for these two muscles which describe what could be considered a perfect smiling face. We can then compare this muscle set with those on a specific generated face and produce a mean square error from the ideal.

This method could not be used in the EVA however, since it would have introduced a circularity into the method. Having generated faces using Ekman's method (the different muscle contraction values of our facial model) we could not use the same method (muscle contraction values) to measure the emotional expressions in the face. Also more importantly, we wanted to base the measurement of emotions on data from real people since it is real people who will draw conclusions from observing the facial models.

3   Measuring Emotional Expressions

An experimental set-up was created to quantify emotional expressions on the face e.g. scale of happiness-sadness, scale of anger-calmness, fear-relax. This is achieved using a number of points / "landmarks" on the face that are significantly influenced by muscle contractions expressing the various emotions measured. We use 25 of such points, situated mainly near the eyes and the mouth together with a stable point that is used as a reference point. This reference point was selected to be the top of the nose.

Two sets of data (randomly generated facial expressions) were created, the first consisting of 200 faces and the second of 150 faces, and positions of each of the "landmarks", for each face, were recorded with respect to the reference point. In fact, distances of the 25 points from the reference point were recorded for each face. Each of the faces were subjectively assessed for emotional state by at least 3 different people for each face (from a pool of about 30 people) for both data sets. The answers of the subjects (the mean for each face) for the first data set only (training data set) were used to create symbolic regression equations. For the symbolic regressions we had, y_i being the response variable denoting the subjects' assessments of the degree of, say, happiness on a $0-100$ scale, where 100 indicates maximum happiness and 0 maximum sadness with 50 being neutral. A similar response variable existed for each of the other emotions. On the other hand, we had 25 explanatory variables x_i, $i = 1, ..., 25$ representing distances of each of the landmarks on the face from the reference point. A separate symbolic regression was carried out for each type of emotion. Hence, a separate estimated regression equation was produced for each emotion.

The second data set (evaluation data set) was then used to verify the equations formed, by comparing users subjective measurements, to the results produced by the symbolic regression equations when applied to this second data set for each emotion. Highly significant positive correlations were found between the results from the equations produced and users subjective evaluations. In fact we use these estimated regression equations to measure, quantify, emotional expressions in the Empathic Visualisation Algorithm (EVA).

4 Training data - GP configuration

Table 1. Table for GP symbolic regression.

Objective:	Find a function of one independent variable y_i and 25 dependent variables, in symbolic form, that fits a given sample of 200 data points (faces).
Terminal set:	$x_i, i = 1, ..., 25$ where x_i are the landmark points on the face.
Function set:	$+, -, \%, *$.
Fitness cases:	200 faces.
Raw fitness:	The sum, taken over the 200 fitness cases, of the RMS error between value of the dependent variables produced by the S-expression and the target value y_i.
Standardised fitness:	Equals raw fitness.
Hits:	Number of fitness cases for which the value of the dependent variables produced by the S-expression comes within 0.01 of the target value y_i.
Wrapper:	None.
Parameters:	Population Size = 750, Generations = 70.
Success predicate:	An S-expression scores 20 hits.

In Figure 1 we can see the evolution of the GP for the different expressions: degree of happiness, degree of anger and degree of fear. We arbitrarily select the two faces shown on top of the figure 1 for presentation in more detail here. All the other cases show similar types of results. The measurements for the three different expressions are shown with degree of happiness identified by the letter H, degree of anger identified by the letter A and degree of fear by the letter F. Next to each graph there is a wider bar, rendered black, showing the average response for the emotional expression encountered, from a pool of at least 3 measurements. This is in fact the target result of the symbolic regression of the GP. The rest of the bars in each graph show measurements for the different expressions of the best of generation individual GP for generations 1, 20, 50 and 70. As you can see the difference between target measurements which are user evaluations and GP measurements decrease, as the number of generations increases. This decrease is also evident in Figures 2, 3 and 4 where for each
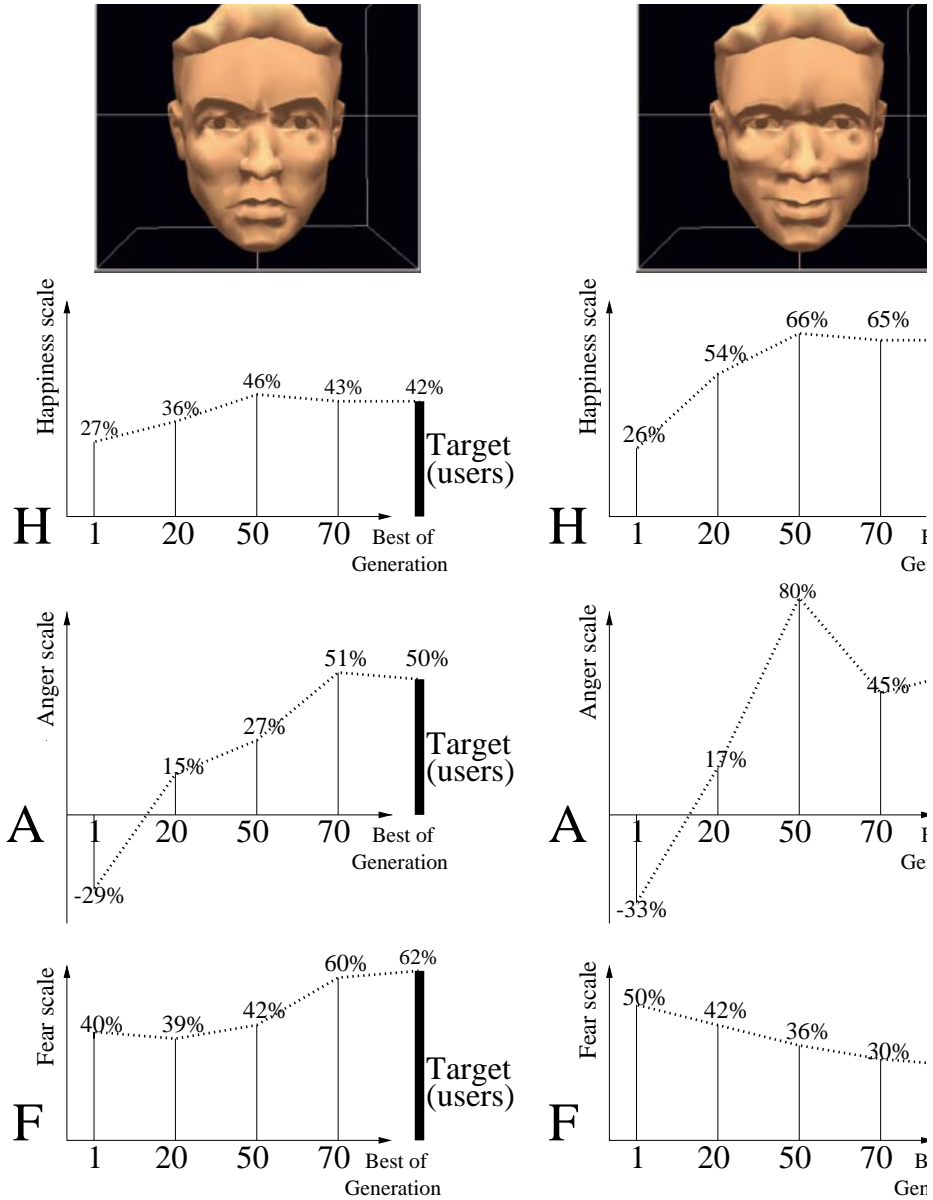
Fig. 1. Here we have for the 2 faces shown the corresponding measurements for degree of happiness (H), degree of anger (A) and degree of fear (F) for the best individual in generation 1, 20, 50 and 70 (last generation).

expression we have a plot between the root mean square error (RMS) of best of generation individual against the actual number of generations.
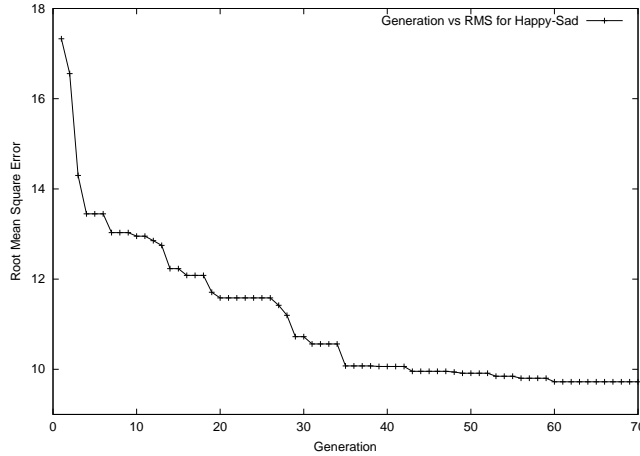

Fig. 2. The evolution of RMS error of best of generation individual against the number of generations for the Happiness-Sadness scale.

From these figures we can conclude that the GP is learning over time, and as we can also see from the two examples given in Figure 1, the measurements for emotions of best of generations individual are almost identical to the user's subjective evaluations.

Therefore, the evidence suggests that the estimated regression equations produced by the GP are a good approximation to user subjective evaluations for the emotional expressions we encounter here.

5 Results

Is there evidence of positive correlation between X and Y, with X being users evaluations of the emotional expressions and Y being the scores for the emotional expressions produced from the estimated regression equations?

Here we are looking for evidence of positive correlation between the two variables (X and Y). For the first emotional expression, the happiness-sadness scale $r_1{}^2 = 0.85$ and the test statistic is $t_1 = 19.7$. For anger-calm scale $r_2{}^2 = 0.75$ and $t_1 = 13.8$ whereas for fear-relax we have $r_3{}^2 = 0.77$ and $t_3 = 14.7$.

The t-test is performed on the evaluation (and not the training) data set of 150 faces. On 148 degrees of freedom the critical $t = 2.6$ at 1% significance level.

Hence the tests are significant for all three cases and thus there is evidence of high positive correlation between variables X and Y. Therefore, there is evidence that the estimated regression equations will produce similar results to
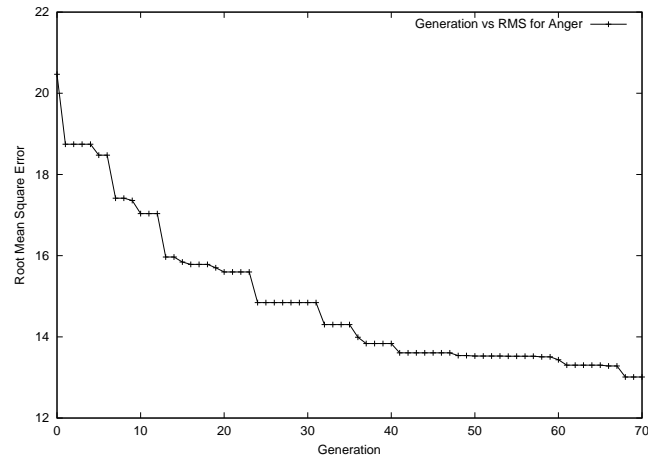
Fig. 3. The evolution of RMS error of best of generation individual against the number of generations for the Angry-Calm scale.
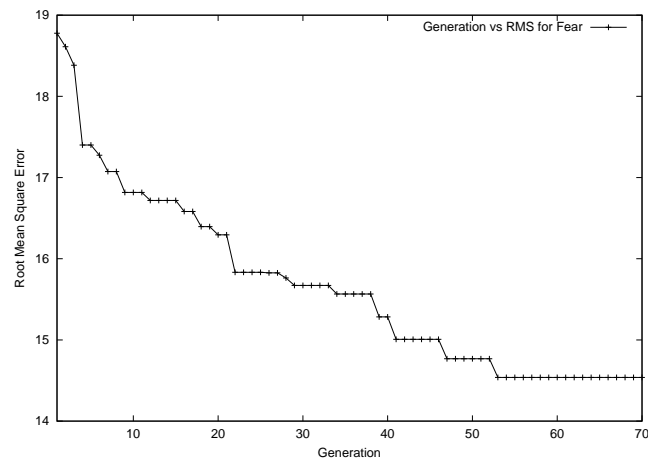

Fig. 4. The evolution of RMS error of best of generation individual against the number of generations for the Fear-Relax scale.

user evaluations of emotional expressions for the facial model we use and hence we can replace users with these equations. In fact the similarity is evident in figures 5,6,7 where the two measurements (user evaluations and estimated regression equations) are plotted together. As can be seen from the graph very clearly there is a positive correlation between the two variables.

6   Conclusion

We have shown a technique to automatically measure emotional expressions of a particular face using GP techniques. Users are replaced by estimated regression equations for different emotional expressions. These estimated equations take into account the movement of certain points mainly around the mouth and eye with respect to a stationary point on a facial model.

In EVA, a GP is used to automatically derive a face which has emotional expressions in a computer generated manner shown in this paper, to allow us automatically map multidimensional data sets, such as accounting and financial reports, to faces and hence visualise the data in a more naturalistic way. We call this system Empathic Visualisation Algorithm (EVA).

In EVA, a GP is used to automatically derive a face which has emotional expressions representing aspects of interest in the data set to the user. In fact, these emotional expressions represent the emotional state users would have, if they were to analyse the data themselves. Hence the word empathic in the name of the system. For example, in financial data, if the user is very interested in degree of liquidity, then the more liquid the company is the happier the face will look. Furthermore, if the user is also interested in profitability, the more profitable the company the more calm (less angry) the face will look. The GP is organised so that for the faces derived, the emotional expressions take into account these user interests, or aspects of importance, which might also be conflicting.

It is important that the method used to measure emotional expressions is not the same as the method used to produce the emotional expressions, otherwise it would result in a circularity in the system. An emotional expression is produced by manipulating muscle contractions. But a particular muscle might have implications for many emotional expressions. Hence, we can not use the same system (muscle contraction values) to also measure the facial expressions.

The method presented here, allow us to independently measure emotional expressions. We are still using Ekman's FACS system in order to produce facial expressions but when it comes to measuring them, we use the estimated regression equations being derived by the GP.

References

Ekman and Friesen, 1978. P. Ekman and W. V. Friesen. *Facial Action Coding System (Investigator's Guide)*. Consulting Psychologists Press, Inc., Palo Alto, California, USA, 1978.

8

Ekman, 1979. P. Ekman. The argument and evidence about universals in facial expressions of emotion. *Handbook of Social Psychopysiology*, pages 143–146, 1979.

Koza, 1992. John R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, 1992.

Loizides and Slater, 2001. A. Loizides and M. Slater. The empathic visualisation algorithm, chernoff faces revisited. In *Technical Sketch, ACM Siggraph 2001 Conference Abstracts and Applications*, page 175. ACM Siggraph, 2001. accepted.

M. Slater, 1999. A. Steed M. Slater, D-P. Pertaub. Public speaking in virtual reality: Facing an audience of avatars. In *IEEE Computer Graphics and Applications*, volume 19, pages 6–9. IEEE, 1999.
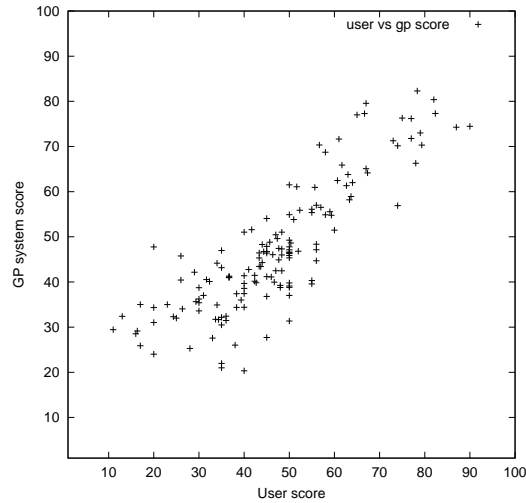
Fig. 5. Here we have the mean user measurement of happiness-sadness scale plotted against the value produced by the symbolic regression for the evaluation data set of 150 faces. As we can see from the graph the two sets of values are highly correlated and in fact $r^2 = 0.85$.
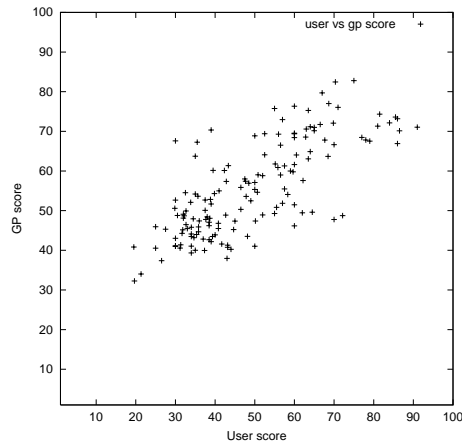
Fig. 6. Here we have the mean user measurement of angry-calm scale plotted against the value produced by the symbolic regression for the evaluation data set of 150 faces. As we can see from the graph the two sets of values are highly correlated and in fact $r^2 = 0.75$.
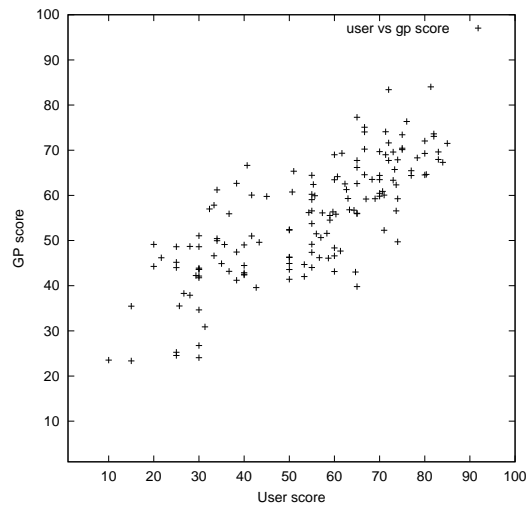

Fig. 7. Here we have the mean user measurement of fear-relax scale plotted against the value produced by the symbolic regression for the evaluation data set of 150 faces. As we can see from the graph the two sets of values are highly correlated and in fact $r^2 = 0.77$.