Creation of Facial Composites from User Selections Using Image Gradients

Rubén García-Zurdo
Universidad Complutense de Madrid, Facultad de Psicología, Madrid, Spain
The Open University, School of Physical Sciences, Milton Keynes, UK
E-mail: rubengarciazurdo@gmail.com

Evolutionary facial composites are created using interactive genetic algorithms based on user selections. This approach is grounded in perceptive studies, and is superior to feature-based systems. A method is presented for creating facial composites in which faces are encoded with shape information, the coordinates of a predefined landmark points, and the image gradient, which represents face information more precisely than image luminance. The new method is accompanied by a Poisson integration process that presents the user with candidate faces. Two user tests, one using composite creators and the other external evaluators, show that the new method produces higher rated composites that are better recognised.

Povzetek: Opisana je metoda generiranja slik za prepoznavo na osnovi interaktivnega genetskega algoritma.

1 Introduction

The goal of facial compositing systems is to create a face image of a target identity from a person's memory so it can be recognised by other people. There are two categories of computerised facial composite systems: in feature-based systems, such as E-FIT [1] and PRO-fit [2], the operator selects features as the eyes, nose and mouth and arranges them on a template to create a face from its parts, while in holistic or evolutionary facial compositing, the operator evolves a whole face by 'breeding' selections from an array of face images, via a process of selection by recognition [3]. Systems in the latter category include EFIT-V [4], ID [4], INIH [6] and EvoFIT [7]. Many of these systems lack a formal user test that can verify their real utility, and identification of individuals from facial composites remains generally poor, meaning that searches for new approaches are justified.

EvoFIT is the system that has been most extensively studied. It produces composites that are identified correctly 30% of the time by people who are familiar with the target identities [8]. This can rise to 45% using more recent strategies for composition [9]. Humberside police used EvoFIT in 35 criminal investigations, and it led to arrests in 60% of cases [10].

Facial compositing research has also produced or confirmed several results that are relevant to face perception: the importance of the internal features of faces over external features [11], [12], [13], [14], the relevance of using configural information [15] and holistic dimensions to describe faces, such as masculinity [16], [17], and the unimportance of colour for face recognition and compositing [18], [19], [20].

Evolutionary face compositing uses interactive genetic algorithms in which the operator selects a number of candidates in an iterative process. These algorithms use an evolutionary mechanism where face representations evolve through crossing (i.e. a mixing of genetic code from selected representations or parents) and random mutation occurring with a predefined low probability [3]. The human operator selects candidates from a gallery, and this selection acts as a fitness function to drive the system to converge to a final composite image resembling the remembered face.

The genetic code or representation of faces is a vector of principal component analysis (PCA) coefficients. PCA represents each face as a coefficient vector corresponding to the weights of a linear combination of elementary faces, called eigenfaces, which are obtained from a sample of images. Each eigenface possesses an associated eigenvalue indicating the amount of variance of the sample that is explained by it. Eigenfaces are usually ordered according to their eigenvalues in such a way that the first eigenfaces contribute more to explaining the observed variance in a sample of images than the remaining eigenfaces. Eigenfaces may be obtained by applying PCA [20] or singular value decomposition (SVD) to the normalised covariance matrix of a sample of images.

However, it is first necessary to align the face images. Although Procrustes analysis can achieve this optimally based on a set of facial landmark points, and can yield the necessary translation, rotation and scaling of shapes to get the best possible alignment, perfect alignment between faces is not usually possible because

each face has a unique shape. This problem is solved with a shape normalisation technique in which images are warped to a reference shape template so they become shape-free, and PCA is performed on the shape-free images [21]. The shape information of individual faces, represented as the x-y coordinates of the landmark points, is used to perform a second PCA to build a eigenshape representation. Each face is thus represented by a pair of texture and shape vectors of PCA coefficients.

Since the introduction of evolutionary facial compositing two decades ago [3], no new representations have been suggested in the literature, with the exception of a combined shape-texture PCA [4], and a user test that would measure the benefit from this approach is missing. Research into new kinds of face representations seems justified, as this may help with the important problem of the limited expressive power of eigenfaces to produce new faces that are not included in the sample as a linear combination of eigenfaces [22]. Face shape and texture are also independent cues for facial recognition [23], [24], [25] and it is therefore hypothesised that the specific method used to render texture in facial composites may have a significant impact on recognition.

Image gradient is introduced here as an alternative representation to facial texture. Image gradient is a differential transformation that represents the direction and magnitude of the maximum intensity change at each pixel by calculating the differences between adjacent pixels in the x and y directions [26]. It can be conceived as a representation of the derivative of a 2D function (i.e. the image) that produces peak responses in places where there is a sudden change of intensity (i.e. the edges). It was proposed as a basic mechanism in early visual processing, and edge detection algorithms have been developed based on this approach [27].

Image gradient represents the underlying structure of the elements in the image better than intensity, and so constitutes a more precise representation that is less affected by illumination patterns. This is illustrated in Figure 1, where the eigenvalues (or amount of associated variance) of the gradient of the facial images used here are shown versus the eigenvalues of components computed based on intensity. The gradient eigenvalues are more uniformly distributed than the intensity ones, which show an initial peak and then a sharp decrease. This peak corresponds to coarse luminance variations in the images [22] and is attenuated in the gradient representation, since the gradient only encodes the differences between adjacent pixels and not their absolute values.

The use of a gradient representation of the facial texture means that a gradient integration technique is necessary to present the corresponding intensity values to the participants. This integration problem is known as Poisson's equation, and is usually solved by setting conditions on the values taken at the area boundary and using an iterative solving method [28]. A major application of Poisson editing is to paste elements into images in a seamless way.

In the present implementation of the system, a constant value at the external edges of the face area is used as a boundary condition. Although this may seem a simplistic condition, it is sufficient to produce realistic images from its gradient. Figure 2 shows that a constant boundary condition is able to recover a individual face from its gradient, since most of the important information seems to be stored in the gradient rather than in the individual pixel values. Even small-range random values at the boundary are sufficient to recover the individual faces.
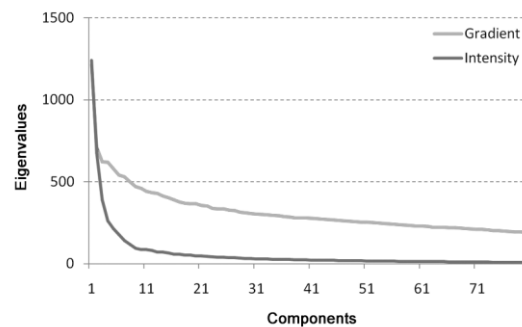

Figure 1: Variance of gradient and intensity PCA components.

The goal of this work is to describe an evolutionary system using the image gradient as a representation of texture and to compare the recognisability and likeness of the resulting composites with those produced using the standard intensity representation of face texture. An initial version of the system with some preliminary results was presented in [29]. Formal mathematical and implementation details are introduced in the appendix.
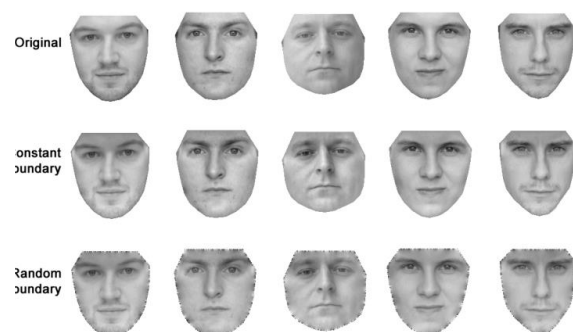

Figure 2: Intensity reconstruction from gradient using constant and random boundary values.

2 Method

The method is illustrated in Figure 3. Sixty-two pictures from the Glasgow unfamiliar face database [30] and 24 pictures from the Utrecht ECVP face database (http://pics.stir.ac.uk/2D_face_sets.htm) were used as reference faces. This gave a total of 86 pictures of Caucasian males, who were mostly in their twenties in the Glasgow sample and in their thirties in the Utrecht sample. Each image shows a frontal view of a face under approximately frontal illumination. Sixty-eight facial landmarks were automatically located on each picture

using a robust state-of-the-art method based on machine learning [31]. Images were converted to grey-scale and warped to a reference shape using the thin plate spline technique. The shape, intensity and gradient PCAs were computed and the resulting components were used in the following genetic algorithm.
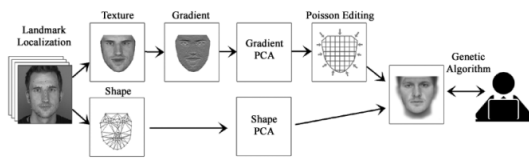

Figure 3: Evolutionary facial compositing overview.

Algorithm

An interactive genetic algorithm is used with the aim of generating a facial image; in this approach, the human operator selects two candidates or parents from a gallery of six images in a 2x3 array. Each face is represented as two vectors, one containing shape coefficients (size 40) and the other texture coefficients (size 80).

 i. Random initialisation: Randomly select values from a uniform distribution of one standard deviation around each PCA component
 ii. Repeat for a number of generations:
  a. Operator selects two parents
  b. Breed a new generation by crossing parent vectors and adding random mutations for both shape and texture
  c. Render candidate gallery for next generation
 iii. Keep selected final image in last generation as the final composite

3 Construction test

Participants

Twenty students (15 women, five men) acted as constructor participants to build the face composites (M_{age} = 19.9, SD = 1.48 years). They took part in the experiment as an educational exercise in groups of five.

Design and procedure

Participants received instructions to construct the face of six well-known male celebrities. These were: David Beckham (DB), George Clooney (GC), Nicolas Cage (NC), Robert De Niro (RN), Tom Cruise (TC) and Tom Hanks (TH). A photo-array of the celebrities was presented briefly to refresh their memory and confirm that all participants were familiar with the targets and their names. They received verbal instruction and hands-on training on how to select the two images most similar to the target identity in order of preference, by clicking the mouse. Participants could erase their selection at any time in order to change it, before proceeding to the next generation by pressing a "Continue" button.

For each generation, six images were shown in a 3x2 array in the centre of the screen. Each participant constructed a total of 12 composites, one for each of the six targets using two levels of representation (gradient and intensity). The order of construction of the 12 composites was varied randomly for each participant. After constructing the composites, participants were asked to rate the likeness of their own composites to the target identity on a scale of 1-10, where 1 means "absolutely dissimilar" and 10 "totally similar". In this case, composites were presented individually on the screen, with the target's name at the top, and the response was given by clicking a number with the mouse. Participants were also asked to rate each target identity in terms of distinctiveness on a scale of 1-10, where 1 means "not distinctive at all" and 10 "maximally distinctive". This time, only the name of each target was shown, so that participants based their response on their own internal representation. Distinctiveness was defined to them as "the degree to which a face would stand out from the rest of the faces in a crowd". The whole procedure took between 50 and 70 minutes for all participants. A one minute rest was allowed after finishing the creation of each composite.

Results

Figure 4 shows examples of the final composites from a participant using gradient and intensity representations.

A within-subject two-way ANOVA was performed for likeness ratings made by constructor participants between Representation (Gradient, Intensity) and Target (DB, GC, NC, RN, TC, TH). A significant effect was obtained for Representation [F (1.19) = 51.33, p < .05, η^2 = .281] in the comparison between gradient (M = 5.51, SE = 0.32) and intensity (M = 4.6, SE = 0.23), following the Greenhouse-Geisser correction. A similarly significant effect for target was obtained [F (5.95) = 3.23, p < .05, η^2 = .148] in the comparison between target identities (M_{DB} = 6.17, SE_{DB} = 0.34; M_{GC} = 4.87, SE_{GC} = 0.4; M_{NC} = 4.45, SE_{NC} = 0.36; M_{RN} = 4.42, SE_{RN} = 0.46; M_{TC} = 5.47, SE_{TC} = 0.45; M_{TH} = 5.92, SE_{TH} = 0.4) with assumed sphericity following the Mauchly test. Multiple comparison tests revealed that differences existed between targets DB and NC [p < .001] and DB and RN [p < .05]. 41.7% of the gradient images received a rating equal to or greater than seven, while only 18.3% of the intensity images received similar ratings. No significant interaction was evident for the interaction of representation x target.

Additionally, a within-subject one-way ANOVA was performed to study the possible differences in target identity distinctiveness, which showed no significant difference. Separate correlation analyses were performed for the gradient and intensity representations for the individual distinctiveness ratings given by constructor participants and their corresponding likeness ratings. A non-significant correlation existed for the gradient representation [ρ = .13, p = .163] but a significant correlation existed for the intensity representation [ρ = .2, p = .030]. A linear regression analysis was then performed for the distinctiveness and likeness ratings for intensity representations, which proved to be significant [F (1,118) = 4.83, p < 0.05]. The corresponding scatter plot for distinctiveness and likeness and the linear model are shown in Figure 5.
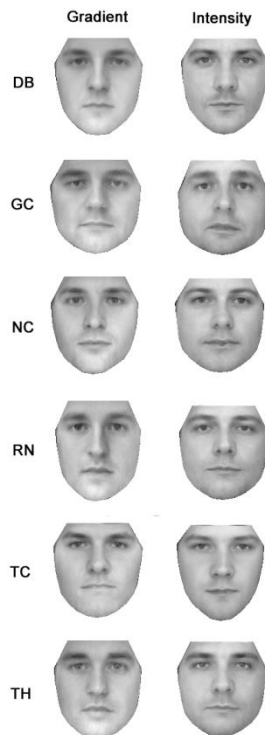
Figure 4: Final composites created with gradient and intensity representations.

Discussion

The composite constructors perceived a higher likeness between their own gradient-based composites and the target identity. Some identity composites tended to generate a higher likeness rating, and it is hypothesised that this was due to the facial distinctiveness of the target. Although we could not prove a significant difference by identity from the collected distinctiveness ratings, two separate correlation analyses of likeness and distinctiveness for the gradient and intensity-based composites showed a significant correlation only for intensity-based composites. This suggests that intensity-based composites are less able to capture the distinctiveness of some faces. This problem is somewhat minimised in gradient-based composites.
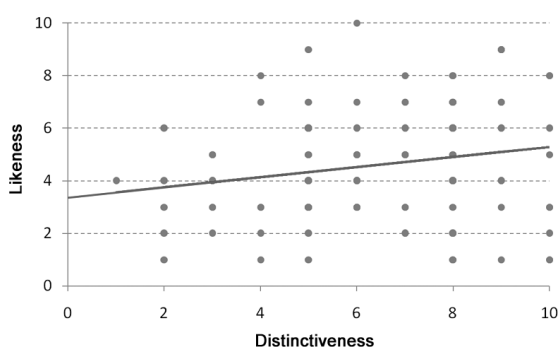

Figure 5: Scatter plot and linear model for likeness and distinctiveness ratings given by constructors for intensity-based composites.

4 External evaluator test

Stimuli and material
The 240 composite images built by the 20 constructor participants were used.

Participants
Forty psychology students (33 women, seven men) took part in the experiment as an educational exercise (M_{age} = 18.9, SD = 1.11 years). They worked in small groups of five.

Design and procedure
Each participant performed two tasks (naming and likeness rating) using the composite images from four constructors. After briefly showing the participants the photo-array of celebrities, to confirm that they were all familiar with them and their names, a name-sorting task was used to measure composite recognition. The composites of the 20 constructors were partitioned into five blocks containing the resulting images of four constructors, corresponding to eight trials of each task (four at the gradient level of representation, and four at the intensity level).

In the naming task, each participant was asked to establish a correspondence between each of the six images presented, which were created by a constructor at a given representation level (gradient, intensity), and a target name. Images were presented in a 2x3 array with a clickable list of target names in alphabetical order underneath each image. The image order was varied randomly by trial, and representation-level blocks were varied randomly by participant. In the likeness rating task, the same composites were presented to each participant in random order. The presentation and response procedures were similar to those used by the constructor participants. The overall procedure took between 15 and 20 minutes for all participants.

Results
Two mixed ANOVA analyses with two between-subject factors (constructor and block) and one within-subject factor (representation) were performed on the percentage of correct naming and likeness ratings. The constructor and block were included as factors to account for any possible effect of the constructors' ability and specific block selection, meaning that their control acts as a measure of the quality of any difference found.

A significant difference was found between the likeness ratings for gradient (M = 3.88, SE = 0.12) and intensity (M = 3.68, SE = 0.1), following the Greenhouse-Geisser correction, although the effect size was small [$F(1,140)$ = 4.08, p < .05, η^2 = .028]. A significant difference was also found between correct namings for gradient (M = 21.04, SE = 1.4) and intensity (M = 16.56, SE = 1.35), with a somewhat greater effect size [$F(1,140)$ = 6.09, p < .015, η^2 = .042], following the Greenhouse-Geisser correction. No effects from the constructor, block or interaction between factors was detected for either likeness or naming.

Discussion

A medium/small advantage in correct naming by external peers for gradient-based composites was found for the sample. We observed a trend of better recognition of the composites constructed using the gradient representation rather than the traditional intensity representation. It is therefore possible to hypothesise that since image gradient is a more invariant characteristic of the elements in an image, it should also represent facial features better than intensity.

We also observed a gradient advantage for likeness ratings given by external peers, although the effect size was smaller than for the constructor participants. As a proxy for naming, the likeness ratings do not always follow the same pattern of effect. There are two possible explanations for this discrepancy: either differences in rating criteria between participants, or differences in the exposure time and familiarity with similar composites between the constructors and external peers.

5 General discussion

Image gradient, an alternative method of representation to image intensity for evolutionary face compositing, was introduced, and its impact on the recognition and likeness ratings of composites was studied. The results indicate a benefit in terms of recognition for the gradient-based composites in our sample. Gradient-based composites are at least as good as those using the standard texture representation. It is conjectured that a benefit may arise from a better representation of facial features by gradient than by intensity. Facial PCA is a powerful tool for analysing facial data [3], but its ability to express new faces as a linear combination of components may be somewhat restricted. Eigenfaces were created for automatic face recognition (a discriminative task), and their ability to express new faces not present in the initial face database (a generative task) may be limited. In this work, a strategy has been followed that consisted of studying a different facial representation on which to perform evolution, in order to increase the representativeness of facial features and thus the accuracy and recognisability of facial composites. The variance associated with gradient components is distributed more uniformly than that associated with intensity components. This implies that during the random mutation stage of composite evolution, the range from which a value is selected is more homogeneous between components and the weight of components is more similar for gradient-based composites.

In previous research [13], a benefit was identified in terms of recognition using a sketch representation, which was presumably caused by a simplification of the facial texture that presented participants with a less demanding

situation. A sketch representation may be beneficial since less shading is involved, which results in less inaccurate information overall. This sketch model was computed for the EvoFIT face set in a preprocessing step, before applying PCA. A similar beneficial effect seems to be arising here from the use of facial image gradient.

As an additional test, automatic evolution of the system was performed for the same target identities as in the user test. The fitness function used was a correlation with an image of the target identity. The results were compared for the three kinds of texture representation, i.e. the intensity, the gradient-preprocessed intensity, in which the sample images were reconstructed from their gradient before PCA, and the gradient. The results shown in Figure 6 offer a visual comparison of their quality.

The evolutionary parameters used here (the numbers of shape and texture components, samples per generation, elitism, mutation and combination rates) were selected based on previous research on intensity representation. Further studies should be carried out to establish their optimal values for gradient representation. Given the huge amount of research on evolutionary facial composites, it should be noted that an ultimate conclusion on the superiority of a new face representation cannot be established from a single work, and extensive research comparing different situations should be conducted.

An improvement was made to our system after the formal experiments were carried out. The number of images presented to participants at each generation was initially six, since the time taken to perform gradient integration in the first implementation of the system (about three seconds per image) persuaded us not to use a greater number of images. This issue was solved in a new version of the system, where a 70% reduction in the time required for gradient integration now allows for the use of greater numbers of images and generations. New features have also been added to the system, such as another set of boundary conditions and the ability to add external features and depth to the resulting composites using optical flow methods. A previous study has at least explored the use of image gradient for facial compositing [32], although this was done from a featural point of view and used gradient integration to stitch fragments from different faces together. Another interesting venue is the exploratory use of deep learning generative adversarial networks for image generation [33] which could theoretically increase the generative power of compositing systems.

It is our conclusion that research on new approaches to face representation could improve the results of evolutionary facial compositing. The present system is available on request to face researchers as a Windows application, with no installation required.
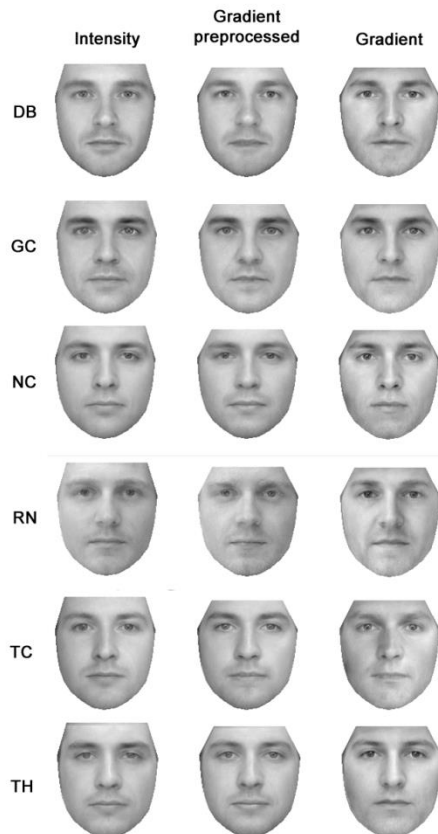
Figure 6: Results of automatic evolution using intensity, gradient-preprocessed and gradient representations.

6   References

[1]   Davies, G., van der Willik, P., & Morrison, L. J. (2000). Facial composite production: A comparison of mechanical and computer-driven systems. *Journal of Applied Psychology*, *85*(1), 119. https://doi.org/10.1037/0021-9010.85.1.119

[2]   Frowd, C. D., McQuiston-Surrett, D., Anandaciva, S., Ireland, C. G., & Hancock, P. J. (2007). An evaluation of US systems for facial composite production. *Ergonomics*, *50*(12), 1987-1998. https://doi.org/10.1080/00140130701523611

[3]   Hancock, P. J. (2000). Evolving faces from principal components. *Behavior Research Methods, Instruments, & Computers*, *32*(2), 327-333. https://doi.org/10.3758/bf03207802

[4]   Solomon, C. J., Gibson, S. J., & Mist, J. J. (2013). Interactive evolutionary generation of facial composites for locating suspects in criminal investigations. *Applied Soft Computing*, *13*(7), 3298-3306. https://doi.org/10.1016/j.asoc.2013.02.010

[5]   Tredoux, C., Nunez, D., Oxtoby, O., & Prag, B. (2006). An evaluation of ID: An eigenface based construction system. *South African Computer Journal*, *37*, 90-97.

[6]   Kurt, B., Etaner-Uyar, A. S., Akbal, T., Demir, N., Kanlikilicer, A. E., Kus, M. C., & Ulu, F. H. (2006). Active appearance model-based facial composite generation with interactive nature-inspired heuristics. *International Workshop on Multimedia Content Representation, Classification and Security*, 2006. pp. 183-190. https://doi.org/10.1007/11848035_26

[7]   Frowd, C. D., Hancock, P. J., & Carson, D. (2004). EvoFIT: A holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions on Applied Perception*, *1*(1), 19-39. https://doi.org/10.1145/1008722.1008725

[8]   Frowd, C. D., Pitchford, M., Bruce, V., Jackson, S., Hepton, G., Greenall, M., ... & Hancock, P. J. (2011). The psychology of face construction: Giving evolution a helping hand. *Applied Cognitive Psychology*, *25*(2), 195-203. https://doi.org/10.1002/acp.1662

[9]   Frowd, C. D., Skelton, F., Atherton, C., Pitchford, M., Hepton, G., Holden, L., ... & Hancock, P. J. (2012). Recovering faces from memory: the distracting influence of external facial features. *Journal of Experimental Psychology: Applied*, *18*(2), 224. https://doi.org/10.1037/a0027393

[10]   Frowd, C. D., Pitchford, M., Skelton, F., Petkovic, A., Prosser, C., & Coates, B. (2012). Catching even more offenders with EvoFIT facial composites. *IEEE Third International Conference on Emerging Security Technologies (EST)*, 2012. pp. 20-26 https://doi.org/10.1109/est.2012.26

[11]   Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*(4), 431-439. https://doi.org/10.1068/p080431

[12]   Frowd, C., Bruce, V., McIntyre, A., & Hancock, P. (2007). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, *98*(1), 61-77. https://doi.org/10.1348/000712606x104481

[13]   Frowd, C., Park, J., McIntyre, A., Bruce, V., Pitchford, M., Fields, S., Kenirons, M. & Hancock, P. J. (2008). Effecting an improvement to the fitness function. How to evolve a more identifiable face. *IEEE ECSIS Symposium on Bio-inspired Learning and Intelligent Systems for Security (BLISS'08)*, 2008. pp. 3-10. https://doi.org/10.1109/bliss.2008.28

[14]   Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*(9), 330-337. https://doi.org/10.1016/s1364-6613(00)01519-9

[15]   Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), 583-592. https://doi.org/10.3758/bf03211301

[16]   Frowd, C. D., Bruce, V., Plenderleith, Y., & Hancock, P. J. B. (2006). Improving target identification using pairs of composite faces constructed by the same person. *IET Conference on Crime and Security*, 2006. pp. 390-395. https://doi.org/10.1049/ic:20060341

[17] Little, A. C., & Hancock, P. J. (2002). The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *British Journal of Psychology*, *93*(4), 451-464.
https://doi.org/10.1348/000712602761381349

[18] Kemp, R., Pike, G., White, P., & Musselman, A. (1996). Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception*, *25*(1), 37-52.
https://doi.org/10.1068/p250037

[19] Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, *31*(8), 995-1003.
https://doi.org/10.1068/p3376

[20] Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71-86.

[21] Craw, I., & Cameron, P. (1991). Parameterising images for recognition and reconstruction. *British Machine Vision Conference*. Springer London. pp. 367-370.
https://doi.org/10.5244/c.5.52

[22] Hancock, P. J., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory & Cognition*, *24*(1), 26-40.
https://doi.org/10.3758/bf03197270

[23] Bruce, V., Hanna, E., Dench, N., Healey, P., & Burton, M. (1992). The importance of 'mass' in line drawings of faces. *Applied Cognitive Psychology*, *6*(7), 619-628.
https://doi.org/10.1002/acp.2350060705

[24] O'Toole, A. J., Vetter, T., Blanz, V. (1999) Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: An application of three-dimensional morphing. *Vision Research*, *39*, 3145-3155.
https://doi.org/10.1016/s0042-6989(99)00034-6

[25] Sinha, P., Balas, B. J., Ostrovsky, Y. & Russell, R. (2006). Face recognition by humans. In Zhao, W. and Chellappa, R. (Eds.), *Face processing: Advanced modeling and methods*, Amsterdam: Elsevier/Academic Press, 257-292

[26] Shah, M. (1997). Fundamentals of computer vision (Unpublished manuscript). University of Central Florida.

[27] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*, 679-698.
https://doi.org/10.1109/tpami.1986.4767851

[28] Pérez, P., Gangnet, M., & Blake, A. (2003). Poisson image editing. *ACM Transactions on Graphics*, *22*(3), 313-318.
https://doi.org/10.1145/882262.882269

[29] Garcia-Zurdo, R. (2016). Evolutive gradient face compositing using the Poisson equation. *Perception*, *45*(2), 25-26.

[30] Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods*, *42*(1), 286-291.
https://doi.org/10.3758/brm.42.1.286

[31] Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. *IEEE Conference on Computer Vision and Pattern Recognition*, 2014. pp. 1867-1874.
https://doi.org/10.1109/cvpr.2014.241

[32] Liu, J., Mei, K., Ge, C., & Zheng, N. (2011). Interactive Poisson photometric propagation for facial composite. *1st International Symposium on Access Spaces (ISAS)*, 2011, pp. 121-126.
https://doi.org/10.1109/isas.2011.5960932

[33] Riviere, M., Teytaud, O., Rapin, J., LeCun, Y. and Couprie, C. (2019). Inspirational adversarial image generation. arXiv:1906.11661.

Appendix. Gradient integration by solving Poisson's equation

The integration of the gradient of an image in order to get its corresponding intensity values reduces to the classic Poisson equation:

$$\Delta \varphi = f \qquad (1)$$

where Δ (read as "Del") stands for the Laplace operator or Laplacian. This expression means that the Laplacian of a certain unknown function equals f. The Laplacian is defined as the divergence of the gradient or as the sum of all unmixed second partial derivatives (the trace of the Hessian):

$$\Delta f = \nabla^2 f = tr(H) \qquad (2)$$

Here, ∇ stands for the gradient operator:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right) \qquad (3)$$

For a discrete 2D function I, the gradient may be approximated as a pair of forward finite differences in the x and y directions:

$$\nabla I = \left(I_x, I_y \right)$$
$$I_x = I(x + 1, y) - I(x, y) \qquad (4)$$
$$I_y = I(x, y + 1) - I(x, y)$$

and the Laplacian can be calculated as the sum of the second-order unmixed gradients:

$$\Delta I = I_{xx} + I_{yy}$$
$$I_{xx} = I_{x+1} - I_x \qquad (5)$$
$$I_{yy} = I_{y+1} - I_y$$

That is, the Laplacian of an image may be obtained from the sum of the horizontal gradient of the horizontal gradient plus the vertical gradient of the vertical gradient. By simple element arrangement, we arrive at the following finite difference scheme for the Laplacian:

$$\Delta I = I(x - 1, y) + I(x + 1, y) + I(x, y - 1) + I(x, y + 1) - 4I(x, y) \qquad (6)$$

Now, we can set up a system of linear equations relating the known Laplacian of the image to the previous Laplacian scheme applied to the unknown pixel values. For each pixel in the image, an equation will be used of the form:

$$[\cdots, 1, \cdots, 1, -4, 1, \cdots, 1, \cdots][x_1, x_2, \cdots, x_n]^T = [f_1, f_2, \cdots, f_n]^T \qquad (7)$$

Here the left vector corresponds to a weight vector implementing the Laplacian scheme, the next vector on

the left side of the equation includes the unknown pixel values in the image, and the vector to the right of the equation includes the known Laplacian values calculated from the horizontal and vertical gradients. It should be noted that the 2D image has been flattened to a 1D vector.

The system of equations needs a boundary condition in order to obtain a solution (up to an additive factor), so we specify the values along the boundary of the domain (image area). This is known as a Dirichlet boundary condition. More specifically, a constant value is used for the boundary pixels. For each of these pixels, the weight values will all be zero, except for the one corresponding to the pixel position, which equals one.

$$[0, 0, \cdots, 0, 1, 0, \cdots, 0][x_1, x_2, \cdots, x_{n,}]^T = k \qquad (8)$$

By stacking all the individual equations together, a linear system of equations is formed:

$$AX = B \qquad (8)$$

Here, **A** is the weight matrix, **X** is the unknown and **B** is the known Laplacian and constant values matrix. This kind of system is sparse, because most of the elements in **A** are zero, and therefore cannot be solved by ordinary means as the pseudo-inverse method. Instead, iterative solving methods such as Gauss-Seidel or Jacobi are used. In order to improve the solving speed, coarse-to-fine (also known as multigrid) methods may be used.