Open access • Journal Article • DOI:10.1145/1360612.1360637

Data-driven enhancement of facial attractiveness — Source link

Tommer Leyvand, Daniel Cohen-Or, Gideon Dror, Dani Lischinski

Institutions: Tel Aviv University, Hebrew University of Jerusalem

Published on: 01 Aug 2008 - International Conference on Computer Graphics and Interactive Techniques

Topics: Attractiveness

Data-Driven Enhancement of Facial Attractiveness

Tommer Leyvand
Tel-Aviv University

Daniel Cohen-Or
Tel-Aviv University

Gideon Dror
Academic College of Tel-Aviv-Yaffo

Dani Lischinski
The Hebrew University

Abstract

When human raters are presented with a collection of shapes and asked to rank them according to their aesthetic appeal, the results often indicate that there is a statistical consensus among the raters. Yet it might be difficult to define a succinct set of rules that capture the aesthetic preferences of the raters. In this work, we explore a data-driven approach to aesthetic enhancement of such shapes. Specifically, we focus on the challenging problem of enhancing the aesthetic appeal (or the *attractiveness*) of human faces in frontal photographs (portraits), while maintaining close similarity with the original.

The key component in our approach is an automatic facial attractiveness engine trained on datasets of faces with accompanying facial attractiveness ratings collected from groups of human raters. Given a new face, we extract a set of distances between a variety of facial feature locations, which define a point in a high-dimensional "face space". We then search the face space for a nearby point with a higher predicted attractiveness rating. Once such a point is found, the corresponding facial distances are embedded in the plane and serve as a target to define a 2D warp field which maps the original facial features to their adjusted locations. The effectiveness of our technique was experimentally validated by independent rating experiments, which indicate that it is indeed capable of increasing the facial attractiveness of most portraits that we have experimented with.

Keywords: facial attractiveness, machine learning, optimization, warping

1 Introduction

Aesthetics and beauty have fascinated human beings from the very dawn of mankind, inspiring countless artists and philosophers. However, an absolute definition of aesthetic values remains elusive. For example, when a group of human raters is presented with a collection of shapes and asked to rank them according to their aesthetic appeal, the results often indicate that there is a statistical consensus among the raters. Yet it might be difficult to define a succinct set of rules that capture the aesthetic perceptions of the raters. Furthermore, such perceptions vary among different classes of shapes, and sometimes differ significantly from culture to culture. Therefore, in this work, we explore the feasibility of a data-driven approach to aesthetic enhancement. Specifically, we focus on the challenging problem of enhancing the aesthetic appeal of human faces (or *facial attractiveness*) in frontal photographs (portraits), while maintaining close similarity with the original.

Figure 1: Input facial images (left) and the adjusted images generated by our method (right). The changes are subtle, yet their impact is significant.

Facial attractiveness has been extensively studied in psychology. Several studies indicate that it is a universal notion, transcending the boundaries between different cultures, since there is a high cross-cultural agreement in facial attractiveness ratings among raters from different ethnicities, socio-economic classes, ages, and gender [Cunningham et al. 1995; Jones 1996; Perrett et al. 1994]. These studies suggest that the perception of facial attractiveness is *data-driven*, meaning that the properties of a particular set of facial features are the same irrespective of the perceiver. A second line of evidence supporting this belief comes from studies of infant preferences for faces [Langlois et al. 1987; Slater et al. 1998]. These studies reveal that infants looked longer at the attractive faces, regardless of the faces' gender, race, or age. Quite recently, using supervised learning techniques, researchers succeeded in producing a trained model capable of generating facial attractiveness ratings that closely conform to those given by human raters [Eisenthal et al. 2006].

The universality of the notion of facial attractiveness along with the ability to reliably and automatically predict the attractiveness rating of a facial image has motivated this work. Specifically, we present a novel tool capable of automatically enhancing the attractiveness of a face in a given frontal portrait. Although for brevity we often refer to this process as *beautification*, it should be understood that we merely claim that images generated by our tool are more likely to receive a higher average attractiveness rating, when presented to

Figure 2: Left: a collage with facial features taken from a catalog. Middle: result of Poisson-blending the features together. Right: result after applying our technique to the middle image.

a group of human observers. We make no claims that the resulting faces are more beautiful than the original ones on any absolute scale; indeed, such a scale is yet to be found!

The main challenge in this work is to achieve the above goal while introducing only minute, subtle modifications to the original image, such that the resulting modified portrait maintains a strong, unmistakable similarity to the original, as demonstrated by the pair of faces shown in Figure 1. This is a highly non-trivial task, since as we shall see, the relationship between the ensemble of facial features and the perceived facial attractiveness is anything but simple.

Applications

Professional photographers have been retouching and deblemishing their subjects ever since the invention of photography. It may be safely assumed that any model that we encounter on a magazine cover today has been digitally manipulated by a skilled, talented retouching artist. It should be noted that such retouching is not limited to manipulating color and texture, but also to wrinkle removal and changes in the geometry of the facial features. Since the human face is arguably the most frequently photographed object on earth, a tool such as ours would be a useful and welcome addition to the ever-growing arsenal of image enhancement and retouching tools available in today's digital image editing packages. The potential of such a tool for motion picture special effects, advertising, and dating services, is also quite obvious.

Another interesting application of our technique is the construction of facial collages when designing a new face for an avatar or a synthetic actor. Suppose we select a collection of facial features (eyes, nose, mouth, etc.) originating from different faces, and would like to synthesize a new face with these features. The features may be assembled together seamlessly using Poisson blending [Pérez et al. 2003], but the resulting face is not very likely to look attractive, or even natural, as demonstrated in Figure 2. Applying our tool on the collage results in a new face that is more likely to be perceived as natural and attractive.

Finally, this work results from our interest in the more general problem of enhancing the aesthetics of geometric shapes. While the techniques described in the remainder of this paper focus on the particular task of increasing the aesthetic appeal of faces, we believe the same paradigm could apply to other 2D shapes.

Overview

The key component in our approach is a *beautification engine* trained using datasets of male and female faces with accompanying facial attractiveness ratings collected from groups of human raters. The entire beautification process is depicted in Figure 3. Given a frontal portrait as input we first (semi-automatically) identify a set of facial landmarks (feature points). Using a planar graph with these feature points as vertices, we extract a vector of distances corresponding to the lengths of the edges in the graph. This vector is fed into the beautification engine, which yields a modified vector of distances, possessing a higher predicted beauty score than that of the original vector. Next, the planar graph is re-embedded in the plane attempting to make the new edge lengths as close as possible to the modified distances. The resulting new positions of the feature points define a 2D warp field that we use to warp the input portrait into its beautified version.

The beautification engine, which is the core novel component of our approach, is presented in Section 3. Section 4 describes the semi-automatic process used to extract facial features, and Section 5 describes the distance embedding and the warping steps.

Our results indicate that the proposed method is capable of effectively increasing the perceived facial attractiveness for most images of both female and male faces that we have experimented with. In particular, the effectiveness of our approach was experimentally validated by a group of test subjects who consistently rated the modified faces as more attractive than the original ones.

Our method uses two datasets. The first dataset consists of 92 frontal portraits of Caucasian females with neutral expression and roughly uniform lighting. The second dataset consists of 33 portraits of Caucasian males. Thus, our tool currently can only be expected to perform well on facial images with similar characteristics. However, it may be directly extended to a wider variety of faces, for example from additional ethnic groups, simply by using it with engines trained on suitable collections of portraits.

2 Background

2.1 Previous work

Much of the research in computer graphics and in computer vision has concentrated on techniques and tools specifically geared at human faces. In particular, there's an extensive body of literature on facial modeling and animation [Parke and Waters 1996; Lee et al. 1995; Guenter et al. 1998; Pighin et al. 1998; O'Toole et al. 1999], face detection [Yang et al. 2002], and face recognition [Zhao et al. 2003]. The most relevant previous works to our own are the different methods for 2D facial image morphing (e.g., [Beier and Neely 1992; Lee et al. 1997]) and the 3D morphable facial models of Blanz and Vetter [1999].

Similarly to image morphing methods, our approach also makes use of 2D image warping to transform the input face. However, our goals are very different. In image morphing, the goal is typically to produce a continuous transformation between two very different faces (or other pairs of objects). The challenge there lies mainly in finding the corresponding features of the two faces, and defining an appropriate warp. In our case, the challenge lies in finding the target shape into which the source image is to be warped, such that the changes are subtle yet result in a perceivable enhancement of facial attractiveness.

Perceptual psychologists also often use image compositing, morphing, and warping to gain a better understanding of how humans perceive various facial attributes. For example, warping towards, and away from average faces has been used to study facial attractiveness and aging (see e.g., [Perrett et al. 1999; Lanitis et al. 2002]). Again, in this case the target shape for the morph, or the direction of the warp, is predefined. Also, it is quite clear that if a face has an above-average attractiveness to begin with, making it more similar to the average face will not achieve our goal.

Blanz and Vetter [1999] present a 3D morphable face model with

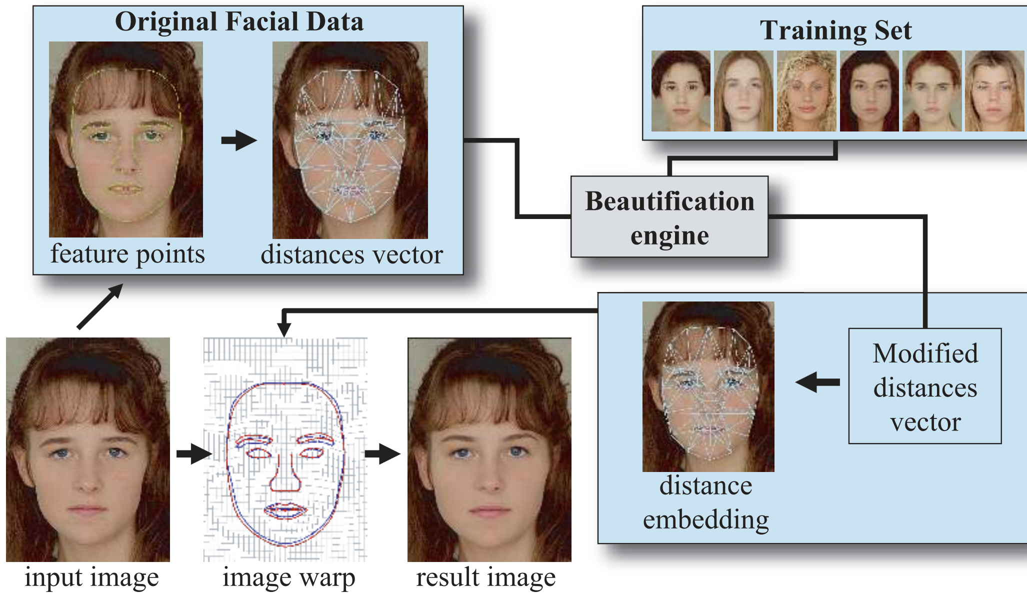

Figure 3: Our facial beautification process.

which it is possible to manipulate a number of facial attributes such as masculinity or fullness, or even to generate new facial expressions. Their morphable model is formed by a linear combination of a set of prototype faces. Their underlying working assumption is that the markedness of the attribute of interest is a linear function. Consequently, increasing or decreasing the markedness is achieved by moving along a single optimal direction in the space of faces. At first glance, it may appear that our task could also be carried out using such a method and indeed, such an attempt was made [Blanz 2003]. However, as we discuss below, facial attractiveness is a highly non-linear attribute.

Our approach does not require fitting a 3D model to a facial image; rather, we operate directly on the 2D image data. We rely on the availability of experimental data correlating facial attractiveness with 2D distances in a facial image, while no equivalent data exists yet for distances between landmarks on a 3D facial mesh. Our method could, however, assist in obtaining a "beautified" 3D model, by applying our technique to an input image as a preprocess, followed by fitting a 3D morphable model to the result.

2.2 Machine rating of facial attractiveness

Eisenthal *et al.* [2006] introduced an automatic facial attractiveness predictor, based on supervised learning techniques. A collection of 92 frontal portraits of young Caucasian females with neutral expressions was used as a training set. The attractiveness of each face was rated by 28 human raters, both males and females. The average rating of a face is henceforth referred to as its *beauty score*. A variety of regressors were then trained, based on 40 features that reflected the geometry of the face, the color of the hair and the smoothness of the skin . The best regressors based on the above features achieved a correlation of 0.6 with human ratings. This is a highly non-trivial result, considering that a random predictor has a zero expected correlation with human rating, while the average correlation between the rating of a single human rater and the average rating is around 0.68 [Kagian et al. 2007].

In this work we use the same collection of facial images and the corresponding ratings collected by Eisenthal *et al.* to train our own regressor (Section 3.1) and use it as a guide in our beautification process of female faces. To deal with male faces we used a second training set of 33 portraits of young men, and acquired the attractiveness of each face using a protocol identical to that of Eisenthal *et al.* It should be noted that Eisenthal *et al.* made no attempt to use their regressor in a generative manner, as we do in this work.

The precise nature of the function that measures the attractiveness of a face based on its image is still unclear. An analysis of the beauty scores collected by Eisenthal *et al.* as a function of extracted feature values has shown that a linear model accounts very poorly for human attractiveness ratings. In the course of this research, we also trained a number of different support vector regressors using various kernels, linear and non-linear. We found linear models to be significantly inferior to non-linear models, both in terms of their best and their average performance, and use radial basis function (RBF) kernels instead.

3 Beautification Engine

3.1 Support Vector Regression

Support Vector Regression (SVR) is an induction algorithm for fitting multidimensional data [Vapnik 1995]. By using various kernels, SVR can fit highly non-linear functions. An SVR is constructed by training it with a sparse set of samples $(\mathbf{x}, y)$, where $\mathbf{x} \in R^d$ and $y \in R$. Our beautification engine utilizes a SVR model trained with the beauty scores gathered by Eisenthal *et al.* [2006]. Specifically, we used the same collection of scored facial images to semi-automatically extract a total of 84 feature points from each face (while Eisenthal *et al.* used only 37 feature points). The feature points are located on the outlines of eight different facial features: two eyebrows, two eyes, the inner and outer boundaries of the lips, the nose, and the boundary of the face (see Figure 4a).

The mean (normalized) positions of the extracted feature points (Figure 4b) are used to construct a Delaunay triangulation. The triangulation consists of 234 edges, and the lengths of these edges in each face form its 234-dimensional *distance vector* (Figure 4c). The distances are normalized by the square root of the face area

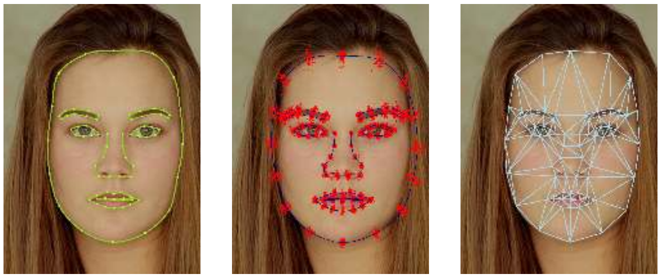
(a) (b) (c)

Figure 4: An example of the 8 facial features, composed of a total of 84 feature points, used in our algorithm. (a) The output feature points from the ASM search. (b) The scatter of the aligned 84 landmark points of our 92 sample training data (in red) and their average (in blue) (c) The 234 distances between these points.

to make them invariant of scale. We work with distances between feature points, rather than with their spatial coordinates, as such distances are more directly correlated with the perceived attractiveness of a face. Furthermore, working with a facial mesh, rather than some other planar graph imposes some rigidity on the beautification process, preventing it from generating distances which may possess a high score but do not correspond to a valid face.

The 234-dimensional feature vectors and their corresponding beauty scores (on a scale of 1 to 7) are used as training samples to construct a SVR model. The SVR defines a smooth function $f_b : R^d \to R$, which we use to estimate the beauty scores of distance vectors of faces outside the training set. Following extensive experimentation, we chose a Radial Basis Function kernel, which is capable of modeling the non-linear behavior expected for such a problem. Model selection was performed by a grid search over the width of the kernel σ, the slack parameter C and the the tube width parameter ε. We used a soft margin SVR implemented in SVM^{light} [Joachims 1999].

Notice, that in contrast to the regressor described by Eisenthal *et al.* [2006], our regressor makes no use of non-geometric features, such as hair color, skin texture, etc. This reflects the difference between our goals: whereas Eisenthal *et al.* attempted to produce the most accurate regressor based on all relevant features, our engine is designed to modify only the geometry of the face, thereby making non-geometric features irrelevant to the process. Thus, it was necessary for us to adjust the beauty scores so as to discount the effect of the non-geometric features. Specifically, we use linear regression to model the effect of the non-geometric features that were identified by Eisenthal *et al.* as strongly correlated with the beauty score. Our regressor was then trained on the difference $y = y_{orig} - y_{lin}$, where y_{orig} are the original scores, and y_{lin} is the linear regression estimate, based on the non-geometrical features above.

3.2 The beautification process

Let $\mathbf{v}$ denote the normalized distance vector extracted from an input facial image. The goal of the beautification process is to generate a *nearby* vector $\mathbf{v}'$ with a higher beauty score $f_b(\mathbf{v}') > f_b(\mathbf{v})$. Since many points in our 234-dimensional feature space do not correspond to distance vectors of faces at all, our main challenge is to keep the synthesized vectors $\mathbf{v}'$ *inside* the subspace of valid faces. Many vectors in this space could possess a higher score, but any such vector must be projected back into the subspace of valid faces, and the score might be reduced in the process. Our assumption is that f_b is smooth enough to allow climbing it incrementally using local optimization techniques.

We experimented with two complementary techniques to achieve this objective: one is based on weighted K-nearest neighbors (KNN) search (Section 3.3), the other is an SVR-driven optimization (Section 3.4).

Assuming that face space is locally convex, the KNN-based method guarantees that the resulting beautified faces lie within this space. In the SVR-based method we optimize a given face according to our beauty function, f_b. Since the latter method does not assume local convexity, it has a more fundamental flavor. However, since the problem is very sparse, and since the SVR is trained on a rather small sample, the regression function could exhibit strong artifacts away from the regions populated by the training samples. Therefore, we constrain the search to a compact region in face space by applying regularization.

3.3 KNN-based beautification

We found that an effective way of beautifying a face, while maintaining a close resemblance to the original is to modify the distance vector of the face in the direction of the beauty-weighted average of the K nearest neighbors of that face. We found the beauty scores of faces modified in this manner to be typically higher than those resulting from moving towards the global unweighted average. This is in line with reports from the psychology literature that composites of beautiful faces were rated as more attractive than an average composite face [Alley and Cunningham 1991; Perrett et al. 1994].

More specifically, let $\mathbf{v}_i$, and b_i denote the set of distance vectors corresponding to the training set samples, and their associated scores, respectively. We define the beauty-weighted distances w_i, for a given distance vector $\mathbf{v}$, as

$$w_i = \frac{b_i}{\|\mathbf{v} - \mathbf{v}_i\|}, \tag{1}$$

where b_i gives more weight to the more beautiful samples. The best results are obtained by first sorting $\{\mathbf{v}_i\}$ such that $w_i \geq w_{i+1}$, and then searching for the value of K maximizing the SVR beauty score f_b of the weighted sum

$$\mathbf{v}' = \frac{\sum_{i=1}^{K} w_i \mathbf{v}_i}{\sum_{i=1}^{K} w_i}. \tag{2}$$

The plot in Figure 5 shows how the beauty score changes for different values of K. Note that the behavior of the beauty score is non-trivial. However, in general, we found small values of K to produce higher beauty scores than that of the average face. Some examples of KNN-beautified faces with different choices of K are shown in Figure 6.

Rather than simply replacing the original distances $\mathbf{v}$ with the beautified ones $\mathbf{v}'$, we are able to produce a more subtle beautification effect, trading off the degree of the beautification for resemblance to the original face, by linearly interpolating between $\mathbf{v}$ and $\mathbf{v}'$ before performing the distance embedding described in Section 5.

3.4 SVR-based beautification

The SVR-based beautification is a numerical optimization treating the SVR beauty function as a *potential field* over the distance vectors feature space. Thus, f_b is used directly to seek beautified feature distance vectors. Whereas the KNN-based approach only produces convex combinations of the training set samples, SVR-based optimization is limited by no such constraint. Figure 6 demonstrates the differences between KNN-based and SVR-based beautification.

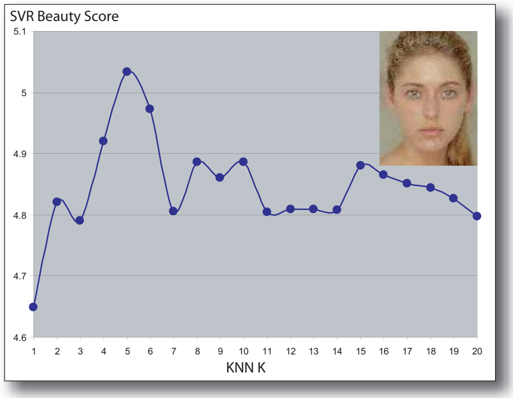

Figure 5: The beauty score is plotted as a function of K in our KNN-based technique applied to one of the faces in our database. The optimal value of K is 5 with an associated SVR beauty score of 5.03. The initial beauty score for this face is 4.38, and the simple average score ($K \rightarrow \infty$) is 4.51. Our SVR-based beautifier succeeds in finding a distance vector with a higher score of 5.20.

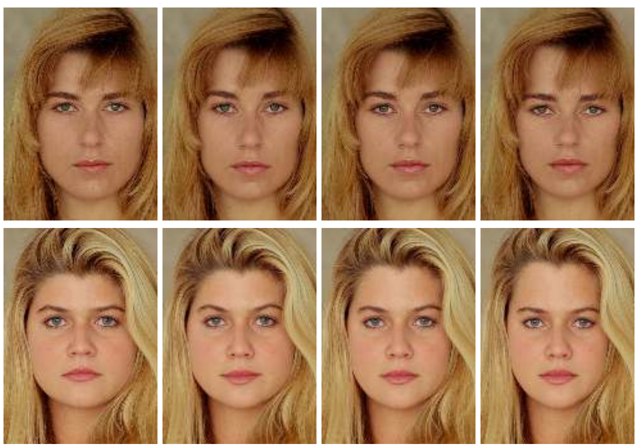

Figure 6: From left to right: original face, KNN-beautified with $K = 3$, KNN-beautified with optimal K, SVR-beautified.

Formally, the beautified distance vector $\mathbf{v}'$ is defined as follows:

$$\mathbf{v}' = \underset{\mathbf{u}}{\operatorname{argmin}}\ E(\mathbf{u}), \text{ where } E(\mathbf{u}) = -f_b(\mathbf{u}). \tag{3}$$

We use the standard no-derivatives Direction Set Method [Press et al. 1992] to numerically perform this minimization. To accelerate the optimization, we perform PCA on the feature space to reduce the dimensionality from 234 to 35, and apply the minimization in the low dimensional space. Note that we start the optimization with the original distance vector $\mathbf{v}$ as the initial guess. Thus, the result of the optimization is typically some local minimum nearby to $\mathbf{v}$. Therefore, applying the process to different faces yields different results.

For the majority of the facial images we experimented with, using only the beauty function as a guide produces results with higher beauty scores than the KNN-based approach. However, for some samples, the SVR-based optimization yields distance vectors which do not correspond to valid human face distances. To constrain the search space, we regularize the energy functional (Eq. 3) by adding a log-likelihood term (LP):

$$E(\mathbf{u}) = (\alpha - 1) f_b(\mathbf{u}) - \alpha LP(\mathbf{u}), \tag{4}$$

where α controls the importance of the log-likelihood term. This technique is similar to the one used by [Blanz and Vetter 1999]. We found that a value of $\alpha = 0.3$ is sufficient to enforce probable distance vectors.

The likelihood function P is approximated by modeling face space as a multivariate Gaussian distribution. When projected onto the PCA subspace, P may be expressed as

$$P = \frac{1}{(2\pi)^{N/2}\sqrt{\prod_i \lambda_i}} \prod_i \exp\left(\frac{-\beta_i^2}{2\lambda_i}\right), \tag{5}$$

where λ_i denotes the i-th PCA eigenvalue and β_i denotes the i-th component of $\hat{\mathbf{u}}$, the projection of $\mathbf{u}$ onto the PCA space. Therefore, the log-likelihood term becomes

$$LP(\hat{\mathbf{u}}) = \sum \frac{-\beta_i^2}{2\lambda_i} + \text{const}, \tag{6}$$

where the constant term is independent of $\hat{\mathbf{u}}$ and can thus be omitted from the optimization of (4).

4 Facial Feature Extraction

Extracting the distances vector from a facial image involves the non-trivial task of automatically identifying the facial feature points. The feature points are located on the prominent facial features (see Figure 4). Each of these features is approximated by a spline. There is extensive literature that deals with the task of snapping such splines to their corresponding facial features. The reader is referred to [Zhao et al. 2003] for a survey of these techniques.

In our work we use the Bayesian Tangent Shape Model (BTSM) [Zhou et al. 2003], a technique that improves on the well-known Active Shape Model (ASM) [Cootes et al. 1995]. ASM consists of a point distribution model capturing shape variations of valid object instances, and a set of grey gradient distribution models, which describe local texture of each landmark point. The model is constructed using a training set, and its parameters are actively updated as new examples are added. This bootstrapping process is semi-automatic. At the early stages of the training, considerable user intervention is necessary, but as the training set increases, user assistance is only rarely required.

Given a new facial image, the ASM algorithm requires an initial guess for the locations of the landmarks. The average shape is a good choice, yet finding the initial scale and orientation greatly improves the accuracy of the detected locations and reduces the need for manual adjustments. For this purpose we use the OpenCV Haar classifier cascade [Bradski 2000].

Our ASM training set consists of 92 samples, each containing 84 landmarks. The distribution of these landmarks is illustrated in Figure 4(b) over one of the facial images in the training set. To process a new facial image it is first analyzed and its feature landmarks are extracted in the same way as was done for the training images. In most cases, the input image analysis is fully automatic. In rare cases some user intervention is required, typically, when large parts of the face are occluded by hair.

5 Distance Embedding and Warping

The beautification engine yields a modified distance vector $\mathbf{v}'$. We must now convert these distances to a set of new facial landmarks.

Figure 7: Beautification examples. Top row: input portraits; Bottom row: the results produced by our method.

Since $\mathbf{v}'$ is not guaranteed to correspond to distances of edges in a planar facial mesh we seek the target landmark positions $q_i = (x_i, y_i)$ that provide the best fit, in the least squares sense, for the distances in $\mathbf{v}'$. Formally, we define

$$E(q_1, \ldots, q_N) = \sum_{e_{ij}} \alpha_{ij} \left(\|q_i - q_j\|^2 - d_{ij}^2 \right)^2, \tag{7}$$

where e_{ij} is our facial mesh connectivity matrix. To reduce non-rigid distortion of facial features, we set α_{ij} to 1 for edges that connect feature points from different facial features, and to 10 for edges connecting points belonging to the same feature. The target distance term d_{ij} is the entry in $\mathbf{v}'$ corresponding to the edge e_{ij}.

The target landmark positions q_i are obtained by minimizing E. This kind of optimization has been recently studied in the context of graph drawing (e.g., [Cohen 1997]). It is referred to as a *stress minimization* problem, originally developed for multidimensional scaling [Sammon 1969]. We use the Levenberg-Marquardt (LM) algorithm to efficiently perform this minimization [Levenberg 1944; Marquardt 1963; Lourakis 2004]. This is an iterative non-linear minimization algorithm which requires reasonable initial positions. However, in our case, the original geometry provides a good initial guess, since the beautification always modifies the geometry only minutely.

The embedding process has no knowledge of the semantics of facial features. However, human perception of faces is extremely sensitive to the shape of the eyes. Specifically, even a slight distortion of the pupil or the iris into a non-circle shape significantly detracts from the realistic appearance of the face. However, such a distortion is not captured by our beauty function. Therefore, we add a post-process that enforces a similarity transform on the landmarks of the eyes, independently for each eye. We solve a linear least squares problem in the four free variables of the similarity transform

$$S = \begin{pmatrix} a & b & t_x \\ -b & a & t_y \\ 0 & 0 & 1 \end{pmatrix},$$

minimizing $\sum \|Sp_i - q_i\|^2$ for all feature points of the eyes, where p_i are original landmark locations, and q_i are their corresponding embedded positions (from Eq. 7). Then Sp_i replaces q_i to preserve the shape of the original eyes. Both the embedding and the similarity transform might perturb the distances vector away from the one produced by the beautification engine, in turn affecting its beauty score. In our experience the score drop due to the embedding is very small (0.005 on average). Applying the similarity transform causes an additional decrease in the beauty score (0.232 on average), but it is essential to the realistic appearance of the result.

5.1 Image Warping

The distance embedding process maps the set of feature points $\{p_i\}$ from the source image to the corresponding set $\{q_i\}$ of target positions. Next, we compute a warp field that maps the source image into the target one according to this set of correspondences. For this purpose, we adapted the multilevel free-form deformation (MFFD) technique introduced by Lee *et al.* [1996]. The warp field is illustrated in Figure 3, where the source feature points are shown in blue and the corresponding target positions are in red.

The MFFD consists of a hierarchical set of free-form deformations of the image plane where, at each level, the warp function is an FFD defined by B-spline tensor products. The advantage of the MFFD technique is that it guarantees a one-to-one mapping (no foldovers). However, this comes at the expense of a series of hierarchical warps (see [Lee et al. 1996] for details). To accelerate the warping of high resolution images, we first unfold the explicit hierarchical composition of transformations into a flat one by evaluating the MFFD on the vertices of a fine lattice.

6 Results

To demonstrate our technique we have implemented a simple interactive application, which was used to generate all of the examples in this paper. After loading a portrait, the application automatically detects facial features, as described in Section 4. The user is able to examine the detected features, and adjust them if necessary. Next, the user specifies the desired degree of beautification, and the application computes and displays the result within a few seconds. A

Figure 8: Beautification of faces that were not part of the training face sets for which facial attractiveness ratings were collected. The four leftmost faces were taken from the AR database [Martinez and Benavente 1998], while the two on the right were obtained from other sources. Top row: input portraits; Bottom row: the results produced by our method.

video demonstrating the application is included on the ACM SIGGRAPH 2008 Full Conference DVD-ROM.

We used two training sets of portraits with accompanying attractiveness ratings (92 female and 33 male portraits) to train our SVR regressors, as described in Section 3.1, and experimented with both KNN-based and SVR-based beautification. We found the SVR-based method to perform better overall on the female faces (see table 1), but on the male faces the KNN-based method performed better. The possible reasons for this are: (i) the male training set was considerably smaller (33 vs. 92), (ii) it did not contain any exceptionally attractive male faces, and (iii) the notion of male attractiveness is not as well established as that for females, so the consensus in the attractiveness ratings was less uniform for males. Therefore all of the female results shown in this section were generated using SVR, while all the male results used KNN.

Original portrait	3.37 (0.49)
Warped to mean	3.75 (0.49)
KNN-beautified (best)	4.14 (0.51)
SVR-beautified	4.51 (0.49)

Table 1: Mean beauty scores for several beautification alternatives (the standard deviation is shown in parentheses). Recall that the beauty score that we use was found to have good correlation with human ratings (for females) in an independent previous study [Eisenthal et al. 2006]

Figure 7 shows a number of input portraits (from the training set) and their corresponding beautified versions. The degree of beautification in all these examples is 100 percent, and our beautification process increases the SVR beauty score by roughly 30 percent. Note that in each of these examples, the differences between the original face and the beautified one are quite subtle, and thus the resemblance between the two faces is unmistakable. Yet the subtle changes clearly have a substantial impact on the attractiveness of these faces.

The faces shown in Figure 7 are taken from the training set of 125 faces, which were photographed by professional photographers. However, the resulting beautification engine generalizes well to faces outside that set. This is demonstrated by the examples in Figure 8. The four leftmost faces were taken from the AR database [Martinez and Benavente 1998]. Note that the photographs of this open repository appear to have been taken under less than ideal illumination conditions. The other two images were obtained from other sources.

In some cases, it is desirable to let the beautification process modify only some parts of the face, while keeping the remaining parts intact. We refer to this mode as *beautification by parts*. For example, the user may request that only the eyes should be subject to beautification (see Figure 9(a–c)). This example nicely demonstrates that sometimes a small local adjustment may result in an appreciable impact on the facial attractiveness. Figure 9(d–f) is another example of beautification by parts, where all of the features except the rather unique lips of this individual were subject to adjustment. To perform beautification by parts we only use those distances where at least one endpoint is located on a feature designated for beautification. This reduces the dimensionality of our feature space and enables the algorithm to search only among the beautified features. This technique implicitly assumes that features that are part of an attractive face are attractive on their own.

As mentioned earlier, it is possible to specify the desired degree of beautification, with 0 percent corresponding to the original face and 100 percent corresponding to the face defined by the beautification process. Degrees between 0 and 100 are useful in cases where the fully beautified face appears too different from the original, as demonstrated in Figure 10.

The beauty scores that we used were gathered using frontal portraits of young Caucasian males and females with a neutral expression. Thus, our tool cannot be expected to perform well on facial images that do not share these characteristics. For example, it will not, in general, enhance the attractiveness of children's faces, as demonstrated in Figure 11. However, the tool may be extended to handle these cases, as well as additional ethnic or age groups, simply by gathering a sufficient amount of beauty rankings from suitable collections of portraits.

Empirical Validation

Our beautification technique is able to improve the facial attractiveness of a large proportion of the faces we experimented with. However, there are also cases where the beautifier simply does not introduce any appreciable change. To objectively validate our beau-

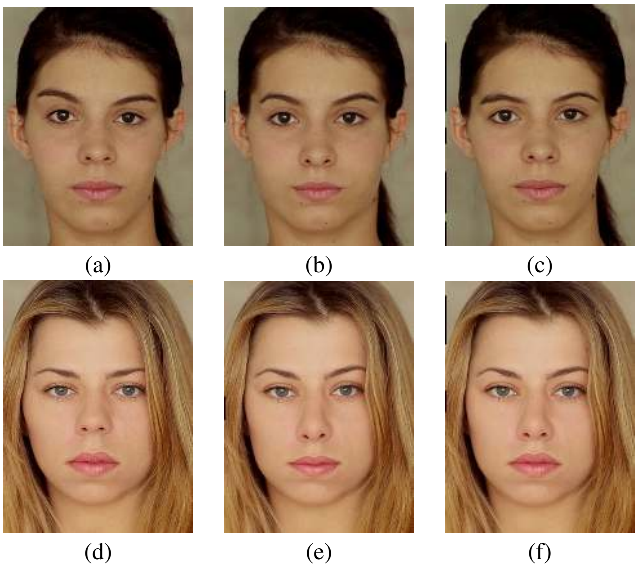

Figure 9: Beautification by parts: (a,d) original image, (b,e) full beautification, (c) only the eyes were designated for beautification, and (f) the mouth region was excluded from beautification.

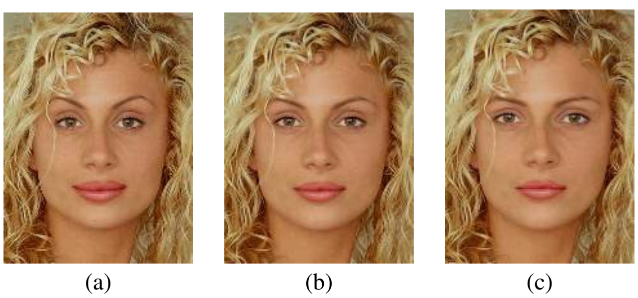

Figure 10: Varying the degree of beautification: (a) original image, (b) 50 percent, and (c) 100 percent, where the differences with respect to the original image may be too conspicuous.

tification procedure we have conducted an experiment in which human raters were presented with 92 pairs of faces (original and beautified) of males (45 faces) and females (47 faces). The raters were asked to indicate the more attractive face in each pair. The positions of the faces in each pair (left or right) were determined randomly, and the 92 pairs were shown in random order. All of the 92 original faces were obtained from the AR face database [Martinez and Benavente 1998]. 68 raters, males and females aged between 25 and 40 participated in the experiment.

As could be expected, the agreement between raters is not uniform for all portraits. For all 47 portraits of females, the beautified versions were chosen as more attractive by most raters, and in half of the cases the beautified versions were preferred by more than 80 percent of the raters. As for the male portraits, 69 percent of the beautified versions were chosen as more attractive. Notice that this result, although not as striking as that for females, is still statistically significant (*P-value* $= 0.006$). The possible reasons for this difference were already explained earlier.

We also performed an experiment to determine whether our SVR-based beautification is more effective than the simpler alternative of using the mean positions of the facial features (which requires neither beauty scores, nor optimization). Using the same setup

Figure 11: A portrait of a child (left) and the result produced by our method using a regressor trained using adult females (right). The result is unlikely to be considered more attractive by the average observer.

as above, 47 pairs of female portraits were presented to 21 raters. In each pair, one image was SVR-beautified, while the other was warped to the mean. We found that 17 (out of 21) raters preferred the SVR result over the alternative for a majority of the faces. This finding is statistically significant (*P-value* $= 0.0015$ under the null hypothesis of indifference). Furthermore, looking at the choices raters made for each face, we found that for 36 (out of 47) faces most raters preferred the SVR result, while the alternative result was preferred for only 11 faces (*P-value* $= 9.8 \times 10^{-5}$).

7 Conclusions and Future Work

We have developed a digital face beautification method based on an optimization of a beauty function modeled by a support vector regressor. Our challenge was twofold: first, the modeling of a high dimensional non-linear beauty function, and second, climbing that function, while remaining within the subspace of valid faces.

Currently, our technique is limited to faces in frontal views and with a neutral expression only. Extending our technique to handle general views and other expressions is a challenging direction for further research.

In our work we restrict ourselves to manipulating only the geometry of the face. However, as was mentioned earlier, there are also important non-geometric attributes that have a significant impact on the perceived attractiveness of a face. These factors include color and texture of hair and skin, and it would be interesting to investigate how changes in these attributes might be incorporated in our digital beautification framework.

Finally, it should be noted that the goal of our research was not to gain a deeper understanding of how humans perceive facial attractiveness. Thus, we did not attempt to derive specific explicit beautification guidelines, such as changing the shape of the lips in a particular manner, or changing the size of the eyes. Instead, we attempted to develop a more general methodology that is based on raw collected beauty scores. It is our hope, however, that perceptual psychologists will find our technique useful in their quest to better understanding of the perception of beauty. Furthermore, as mentioned earlier, we hope to extend our approach to data-driven enhancement of the aesthetics of other shape classes.

Acknowledgments: The authors would like to thank Mr. Akira Gomi for kindly granting permission to use his photographs (from his CD-ROM collection “Americans 1.0, Los Angeles 1994”), and Volker Blanz for valuable comments and discussions. The training set of male faces was provided by Martin Gründl. This work was supported in part by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities and by the Israeli

Ministry of Science and Technology. Gideon Dror was supported by NICTA, funded through the Australian Government's Backing Australia's Ability initiative and also through the ARC.

References

Alley, T. R., and Cunningham, M. R. 1991. Averaged faces are attractive, but very attractive faces are not average. *Psychological Science 2*, 123–125.

Beier, T., and Neely, S. 1992. Feature-based image metamorphosis. In *Proc. ACM SIGGRAPH 92*, ACM Press, 35–42.

Blanz, V., and Vetter, T. 1999. A morphable model for the synthesis of 3D faces. In *Proc. ACM SIGGRAPH 99*, ACM Press, 187–194.

Blanz, V., 2003. Manipulation of facial attractiveness. [web page] `http://www.mpi-inf.mpg.de/~blanz/data/attractiveness/`.

Bradski, G. 2000. The OpenCV library. *Dr. Dobbs Journal 25*, 11 (Nov.), 120, 122–125.

Cohen, J. D. 1997. Drawing graphs to convey proximity: an incremental arrangement method. *ACM Trans. Comput.-Hum. Interact. 4*, 3, 197–229.

Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. 1995. Active shape models — their training and their applications. *Comput. Vis. Image Underst. 61*, 1 (Jan.), 38–59.

Cunningham, M. R., Roberts, A. R., Wu, C.-H., Barbee, A. P., and Druen, P. B. 1995. Their ideas of beauty are, on the whole, the same as ours: Consistency and variability in the cross-cultural perception of female attractiveness. *Journal of Personality and Social Psychology 68*, 1, 261–279.

Eisenthal, Y., Dror, G., and Ruppin, E. 2006. Facial attractiveness: Beauty and the machine. *Neural Computation 18*, 1, 119–142.

Guenter, B., Grimm, C., Wood, D., Malvar, H., and Pighin, F. 1998. Making faces. In *Proc. ACM SIGGRAPH 98*, ACM Press, 55–66.

Joachims, T. 1999. Making large-scale SVM learning practical. In *Advances in kernel methods: support vector learning*. MIT Press, Cambridge, MA, USA, 169–184.

Jones, D. 1996. *Physical Attractiveness and the Theory of Sexual Selection: Results from Five Populations*. University of Michigan Press, Ann Arbor, Michigan.

Kagian, A., Dror, G., Leyvand, T., Cohen-Or, D., and Ruppin, E. 2007. A humanlike predictor of facial attractiveness. In *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. Platt, and T. Hoffman, Eds. MIT Press.

Langlois, J. H., Roggman, L. A., Casey, R. J., M.Ritter, J., Rieser-Danner, L. A., and Jenkins, V. Y. 1987. Infant preferences for attractive faces: Rudiments of a stereotype? *Developmental Psychology 23*, 5, 363–369.

Lanitis, A., Taylor, C. J., and Cootes, T. F. 2002. Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell. 24*, 4, 442–455.

Lee, Y., Terzopoulos, D., and Waters, K. 1995. Realistic modeling for facial animation. In *Proc. ACM SIGGRAPH 95*, ACM Press, 55–62.

Lee, S.-Y., Wolberg, G., Chwa, K.-Y., and Shin, S. Y. 1996. Image metamorphosis with scattered feature constraints. *IEEE Trans. Visualization and Computer Graphics 2*, 4 (Dec.).

Lee, S.-Y., Wolberg, G., and Shin, S. Y. 1997. Scattered data interpolation with multilevel B-splines. *IEEE Trans. Visualization and Computer Graphics 3*, 3 (July–Sept.).

Levenberg, K. 1944. A method for the solution of certain problems in least squares. *Quart. Appl. Math.*, 2, 164–168.

Lourakis, M., 2004. levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++. [web page] `http://www.ics.forth.gr/~lourakis/levmar/`.

Marquardt, D. 1963. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11, 431–441.

Martinez, A. M., and Benavente, R. 1998. The AR face database. *CVC Technical Report 24*.

O'Toole, A. J., Price, T., Vetter, T., Bartlett, J. C., and Blanz, V. 1999. 3D shape and 2D surface textures of human faces: the role of "averages" in attractiveness and age. *Image Vision Comput. 18*, 1, 9–19.

Parke, F. I., and Waters, K. 1996. *Computer Facial Animation*. A K Peters, Wellesley, Massachusetts.

Pérez, P., Gangnet, M., and Blake, A. 2003. Poisson image editing. *ACM Trans. Graph. 22*, 3, 313–318.

Perrett, D. I., May, K. A., and Yoshikawa, S. 1994. Facial shape and judgements of female attractiveness. *Nature 368*, 239–242.

Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., and Edwards, R. 1999. Symmetry and human facial attractiveness. *Evolution and Human Behavior 20*, 295–307.

Pighin, F., Hecker, J., Lischinski, D., Szeliski, R., and Salesin, D. H. 1998. Synthesizing realistic facial expressions from photographs. In *Proc. ACM SIGGRAPH 98*, 75–84.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. 1992. *Numerical Recipes: The Art of Scientific Computing*, 2nd ed. Cambridge University Press.

Sammon, J. W. 1969. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers C-18*, 5, 401–409.

Slater, A., der Schulenberg, C. V., Brown, E., Butterworth, G., Badenoch, M., and Parsons, S. 1998. Newborn infants prefer attractive faces. *Infant Behavior and Development*, 345–354.

Vapnik, V. 1995. *The nature of statistical learning theory*. Springer, New York.

Yang, M.-H., Kriegman, D., and Ahuja, N. 2002. Detecting faces in images: A survey. *IEEE Trans. PAMI 24*, 1, 34–58.

Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. 2003. Face recognition: A literature survey. *ACM Comput. Surv. 35*, 4, 399–458.

Zhou, Y., Gu, L., and Zhang, H.-J. 2003. Bayesian tangent shape model: Estimating shape and pose parameters via bayesian inference. In *Proc. CVPR 2003*, 109–118.