Copula Eigenfaces with Attributes: Semiparametric Principal Component Analysis for a Combined Color, Shape and Attribute Model

Bernhard Egger[(✉)], Dinu Kaufmann, Sandro Schönborn, Volker Roth, and Thomas Vetter

Department of Mathematics and Computer Science,
University of Basel, Basel, Switzerland
{bernhard.egger,dinu.kaufmann,sandro.schoenborn,volker.roth,
thomas.vetter}@unibas.ch

Abstract. Principal component analysis is a ubiquitous method in parametric appearance modeling for describing dependency and variance in datasets. The method requires the observed data to be Gaussian-distributed. We show that this requirement is not fulfilled in the context of analysis and synthesis of facial appearance. The model mismatch leads to unnatural artifacts which are severe to human perception. As a remedy, we use a semiparametric Gaussian copula model, where dependency and variance are modeled separately. This model enables us to use arbitrary Gaussian and non-Gaussian marginal distributions. Moreover, facial color, shape and continuous or categorical attributes can be analyzed in an unified way. Accounting for the joint dependency between all modalities leads to a more specific face model. In practice, the proposed model can enhance performance of principal component analysis in existing pipelines: The steps for analysis and synthesis can be implemented as convenient pre- and post-processing steps.

Keywords: Copula Component Analysis · Gaussian copula · Principal component analysis · Parametric Appearance Models · 3D Morphable Model · Face modeling · Face synthesis · Attributes

1   Introduction

Parametric Appearance Models (PAM) describe objects in an image in terms of pixel intensities. In the context of faces, Active Appearance Models [1] and 3D Morphable Models [2] are established PAMs to model appearance and shape. The dominant method for learning the parameters of a PAM is principal component analysis (PCA) [3]. PCA is used to describe the variance and dependency in the data. Due to the sensitivity of PCA to space and scaling, seperate models are

---

B. Egger and D. Kaufmann—These authors contributed equally to this work.
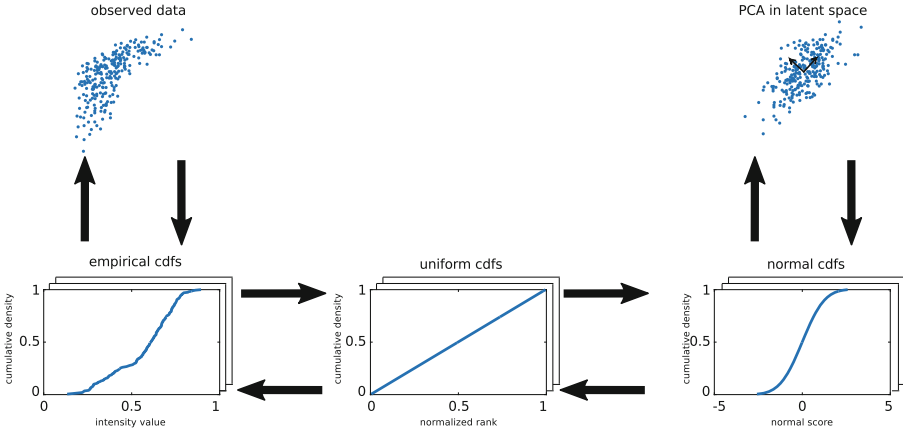
Fig. 1. This figure shows the pre- and post-processing steps necessary to use a Gaussian copula before calculating PCA.

learned for shape and appearance. Usually, PAMs are generative models that can synthesize new random instances.

Using PCA to model facial appearance leads to models which are able to synthesize instances which appear unnaturally. This is due to the assumption that the color intensities or, in other words, the marginals at a pixel are Gaussian-distributed. We show that this is a severe simplification: The pixel intensities of new samples will follow a joint Gaussian distribution. This approximation is far from the actual observed distribution of the training data and leads to unnatural artifacts in appearance.

The ability to synthesize random *and* natural instances is important when generating new face instances [4] and in face manipulation [5]. This is because human perception is very sensitive to unnatural variability in a face. On the other hand, PCA face models are used as a strong prior in probabilistic facial image interpretation algorithms [6]. Hence, such applications require a prior which follows the underlying distribution as closely as possible and, which is therefore, highly specific to faces.

In order to enhance the specificity of a PCA-based model, an obvious improvement would be the extension to a Gaussian mixture model [7]. Here, each color channel at a pixel is modeled with an (infinite) mixture of Gaussians. However, we skip this step and propose to use a semiparametric copula model directly.

A copula model provides the decomposition of the dependency and the marginal distributions such that the copula contains the dependency structure only. This separate modeling allows us to drop the parametric Gaussian assumption on the color channels and to replace them with nonparametric empirical distributions. In general, seperating all marginals from the dependency structure leads to a scale invariant description of the underlying dependency. This is desired when working with data from different modalities, living in different spaces. Scale invariance enables us to learn a combined dependency structure of shape,

color and attributes. We keep the parametric dependency structure; in particular, we use a Gaussian copula because of its inherent Gaussian latent space. PCA can then be applied in the latent Gaussian space and is used to learn the dependencies of the data independently from the marginal distribution. The method is analytically analyzed in [8] and is called Copula Component Analysis (COCA). Samples drawn from a COCA model follow the empirical marginal distribution of the training data and are, therefore, more specific to the modeled object.

The additional steps for using COCA can be implemented as simple pre- and post-processing before applying PCA. The data is mapped into a space where it is Gaussian-distributed. This mapping is obtained by first ranking the data and then transforming it by the standard normal distribution. We perform PCA on the transformed data to learn its underlying dependency structure. All necessary steps are visualized in Fig. 1.

A semiparametric Gaussian copula model also provides additional benefits: First, learning is invariant to monotonic transformations of all marginals, including invariance to scaling. We explore this advantage by learning a combined color and shape model, even including attributes. Second, the implementation can be done as simple pre- and post-processing steps. Third, the model also allows changing the color space. For facial-appearance modeling, the HSV color space is more appropriate than RGB. The HSV color space is motivated by the separation of the hue and saturation components and brightness value. On the other hand, without adaptions, PCA is not applicable to facial appearance in the HSV color space because of its sensitivity to differently-scaled color channels.

In summary, methods building on PCA can easily benefit from these advantages to improve their learned model. By learning a combined shape, color and attribute model we explore scale invariance and therefore the possibility to include diverse modalities of the data in a common model.

1.1   Related Work

The Eigenfaces approach [9,10] uses PCA on aligned facial images to analyze and synthesize faces. Active Appearance Models [1] add a shape component which allows to model the shape independently from the appearance. The 3D Morphable Model [2] uses a dense registration, extends the shape model to 3D and adds camera and illumination parameters. The 3D Morphable Model allows handling appearance independently from pose, illumination and shape. These methods have a common core: They focus on analysis and synthesis of faces and all of them use a PCA model for color representation and can, therefore, benefit from COCA.

Photo-realistic face synthesis methods like Visio-lization [4] use PCA as a basis for example-based photo-realistic appearance modeling.

1.2   Outline

The remainder of the paper is organized as follows: The methods section explains the copula extension for PCA and presents the theoretical background for

learning and inference. We also indicate, how to include discrete distributed data in the copula framework. Additionally, practical information for an implementation is provided. In the experiments and results section we demonstrate that facial appearance should be modeled using the copula extension. We qualitatively and quantitatively show that the proposed model leads to a facial appearance model which is more specific to faces.

2   Methods

2.1   PCA for Facial Appearance Modeling

Let $x \in \mathbb{R}^{3n}$ describe a zero-mean vector representing 3 color channels of an image with n pixels. In an RGB image, the color channels and the pixels are stacked such that $x = (r_1, g_1, b_1, r_2, b_2, b_3, \ldots, r_n, g_n, b_n)^T$. We assume that the mean of every dimension is already subtracted. The training set of m images is arranged as the data matrix $X \in \mathbb{R}^{3n \times m}$.

PCA [3] aims at diagonalizing the sample covariance $\Sigma = \frac{1}{m} X X^T$, such that

$$\Sigma = \frac{1}{m} U S^2 U^T \tag{1}$$

where S is a diagonal matrix and U contains the transformation to the new basis. The columns of matrix U are the eigenvectors of Σ and the corresponding eigenvalues are on the diagonal of S.

PCA is usually computed by a singular value decomposition (SVD). In case of a rank-deficient sample covariance with rank $m < n$ we cannot calculate U^{-1}. Therefore, SVD leads to a compressed representation with a maximum of m dimensions. The eigenvectors in the transformation matrix U are ordered by the magnitude of the corresponding eigenvalues.

When computing PCA, the principal components are guided by the variance as well as the covariance in the data. While the variance captures the scattering of the intensity value of a pixel, the covariance describes which regions contain similar color. This mingling of factors leads to results which are sensitive to different scales and to outliers in the training set. Regions with large variance and outliers could influence the direction of the resulting principal components in an undesired manner.

We uncouple variance and dependency structure such that PCA is only influenced by the dependency in the data. Our approach for uncoupling is a copula model which provides an analytical decomposition of the aforementioned factors.

2.2   Copula Extension

Copulas [11,12] allow a detached analysis of the marginals and the dependency pattern for facial appearance models. We consider a relaxation to a semiparametric Gaussian copula model [13,14]. We keep the Gaussian copula for describing the dependency pattern, but we allow nonparametric marginals.

Let $x \in \mathbb{R}^{3n}$ describe the same zero-mean vector as used for PCA, representing 3 color channels of an image with n pixels. Sklar's theorem allows the decomposition of every continuous and multivariate cumulative probability distribution (cdf) into its marginals $F_i(X_i), i = 1, \ldots, 3n$ and a copula C. The copula comprises the dependency structure, such that

$$F(X_1, \cdots, X_{3n}) = C(W_1, \ldots, W_{3n}) \qquad (2)$$

where $W_i = F_i(X_i)$. W_i are uniformly distributed and generated by the probability integral transformation[1].

For our application, we consider the Gaussian copula because of its inherently implied latent space

$$\tilde{X}_i = \Phi^{-1}(W_i), \quad i = 1, \ldots, 3n \qquad (3)$$

where Φ is the standard normal cdf. The multivariate latent space is standard normal-distributed and fully parametrized by the sample correlation matrix $\tilde{\Sigma} = \frac{1}{m}\tilde{X}\tilde{X}^T$ only. PCA is then applied on the sample correlation in the latent space $\tilde{X}$.

The separation of dependency pattern and marginals provides multiple benefits: First, the Gaussian copula captures the dependency pattern invariant to the variance of the color space[2]. Second, whilst PCA is distorted by outliers and is generally inconsistent in high dimensions, the semiparametric copula extension solves this problem [8]. Third, the nonparametric marginals maintain the non-Gaussian nature of the color distribution. Especially when generating new samples from the trained distribution, the samples do not exceed the color space of the training set.

2.3   Inference

We learn the latent sample correlation matrix $\tilde{\Sigma} = \frac{1}{m}\tilde{X}\tilde{X}^T$ in a semiparametric fashion using nonparametric marginals and a parametric Gaussian copula. We compute $\hat{w}_{ij} = \hat{F}_{\text{emp},i}(x_{ij}) = \frac{r_{ij}(x_{ij})}{m+1}$ using empirical marginals $\hat{F}_{\text{emp},i}$, where $r_{ij}(x_{ij})$ is the rank of the data x_{ij} within the set $\{x_{i\bullet}\}$. Then, $\tilde{\Sigma}$ is simply the sample covariance of the normal scores

$$\tilde{x}_{ij} = \Phi^{-1}\left(\frac{r_{ij}(x_{ij})}{m+1}\right), \quad i = 1, \ldots, 3n, \quad j = 1, \ldots, m. \qquad (4)$$

Equation (4) contains the nonparametric part, since $\tilde{\Sigma}$ is computed from the ranks $r_{ij}(x_{ij})$ solely and contains no information about the marginal distribution of the x's. Note, $\tilde{x} \sim \mathcal{N}(0, \tilde{\Sigma})$ is standard normal distributed with correlation matrix $\tilde{\Sigma}$. Subsequently, an eigendecomposition is applied on the latent correlation matrix $\tilde{\Sigma}$.

---

[1] The copula literature uses U instead of W. We changed this convention due to the singular value decomposition which uses $X = USV^T$.

[2] More general, a copula model is invariant against all monotonic transformations of the marginals.

Generating a sample using PCA then simply requires a sample from the model parameters

$$h \sim \mathcal{N}(0, I) \tag{5}$$

which is projected to the latent space

$$\tilde{x} = \tilde{U} \frac{\tilde{S}}{\sqrt{m}} h \tag{6}$$

and further projected component-wise to

$$w_i = \Phi(\tilde{x}_i), \quad i = 1, \ldots, 3n. \tag{7}$$

Finally, the projection to the color space requires the empirical marginals

$$x_i = \hat{F}_{\text{emp},i}(w_i), \quad i = 1, \ldots, 3n. \tag{8}$$

All necessary steps are summarized in Algorithms 1 and 2 and visualized in Fig. 1.

It is possible to smoothen the empirical marginals with a kernel k and replace (8) by $x_i = k(w_i, X_{i\bullet}), \quad i = 1, \ldots, 3n$.

Algorithm 1. Learning.

Input: Training set $\{X\}$
Output: Projection matrices U, S
for *all dimensions* **do**
    for *all samples* **do**
        $\tilde{x}_{ij} = \Phi^{-1}\left(\frac{r_{ij}(x_{ij})}{m+1}\right)$
find $\tilde{U}, \tilde{S}$ such that $\tilde{\Sigma} = \frac{1}{m}\tilde{U}\tilde{S}^2\tilde{U}^T$ (via SVD)

Algorithm 2. Sampling.

Output: Random sample x
$h \sim \mathcal{N}(0, I)$
$\tilde{x} = \tilde{U} \frac{\tilde{S}}{\sqrt{m}} h$
for *all dimensions* i **do**
    $w_i = \Phi(\tilde{x}_i)$
    $x_i = \hat{F}_{\text{emp},i}(w_i)$

2.4   Implementation

The additional steps for using COCA can be implemented as simple pre- and post-processing before applying PCA. Basically the data is mapped into a latent space where it is Gaussian-distributed. The mapping is performed in two steps.

First, the data is transformed to an uniform distribution by ranking the intensity values. Then it is transformed to a standard normal distribution. On the transformed data, we perform PCA to learn the dependency structure in the data.

To generate new instances from the model, all steps have to be reversed. Figure 1 gives an overview of all necessary transformations. The following steps have to be performed, e.g. in MATLAB, to calculate COCA:

```matlab
% calculate empirical cdf
[empCDFs, indexX] = sort(X, 2);

% transform emp. cdf to uniform
[~, rank] = sort(indexX, 2);
uniformCDFs = rank / (size(rank, 2)+1);

% transform uni. cdf to std. normal cdf
normCDFs = norminv(uniformCDFs',0,1)';

% calculate PCA
[U,S,V] = svd(normCDFs, 'econ');
```
Listing 1.1. Learning.

To generate an image from model parameters, the following steps are necessary:

```matlab
% random sample
m = size(normCDFs, 2);
h = random('norm' ,0 ,1 ,m ,1);
sample = U * S / sqrt(m) * h;

% std. normal to uniform
uniformSample = normcdf(sample, 0, 1) * (m - 1) + 1;

% uniform to emp. cdf
empSample = empCDFs(sub2ind(size(empCDFs), 1:size(data, 1), ...
    round(uniformSample')))';
```
Listing 1.2. Sampling.

These are the additional steps which have to be performed as pre- and post-processing for the analysis of the data and the synthesis of new random samples. In terms of computing resources we have to consider the following: The empirical marginal distributions F_{emp} are now part of the model and have to be kept in memory. In the learning part, the complexity of sorting the input data is added. In the sampling part, we have to transform the data back by looking up their values in the empirical distribution.

The copula extension comes with low additional effort: it is easy to implement and has only slightly higher computing costs. We encourage the reader to implement these few steps since the increased flexibility in the modeling provides a valuable extension.

2.5   Discrete Ordinal Marginals

The formulation of the coupla framework as above works with arbitrary continuous marginals. We extend the copula model for attributes, which follow discrete ordinal marginals. With this extension, we can even augment our model with attributes following binary distribution, such as gender. The underlying generative model assumes a continuous latent space, which is identified with the latent space $\tilde{X}$ of the copula. From this space, we observe the measurements via a discretization, which is related to the marginal distribution containing discontinuities. Using the cdfs of these marginals, for infering the latent space as in the previous sections, causes problems. This is because the cdf transformations $\Phi^{-1} \circ \hat{F}_{\text{emp},i} : X_i \to \tilde{X}_i$ do not change the marginal data distribution to uniform and hence do not recover the continuous latent space. Instead, these cdf transformations only change the sample space. This leads to an invalid distribution of the copula and subsequently also of the latent space.

In order to resolve this problem, we follow the approach of the extended rank likelihood [15]. This provides us with an association-preserving mapping between measurement x_{ij} and latent observation $\tilde{x}_{ij}$. The essential idea behind this approach is, that the rank relation from the observations are preserved in the latent space. The latent variables are then recovered by a Gibbs sampler, which obeys these rank relations while respecting the Gaussian copula. From this sampler, we are able to generate (continuous) latent pseudo observations $\tilde{x}$, which subsequently can be included in our model. Using this Gibbs sampler, we are able to include discrete ordinal distributed attributes with arbitrary many categories.

However, the above described Gibbs sampler causes problems in our setting, since sampling in such high dimensions is just infeasible. In our case, we want to include a binary variable (sex). Note, that a binary variable can always be considered as an ordinal variable, since the ordering of the encoding does not matter. Instead of resampling from the conditional posterior distribution $p(\tilde{x}_{\text{sex}}|\tilde{x}_{-\text{sex}}, x_{\text{sex}})$ in the latent space, we replace the label x_{sex} with logistic regression in a preprocessing step. Specifically, logistic regression provides us a (continuous) score $x'_{\text{sex}} = E(x_{\text{sex}}|x_{-\text{sex}})$, which is the conditional expectation over (a low rank approximation of) the remaining variables $x_{-\text{sex}}$. Since the score constitutes of the conditional expectation, it relates to an approximation of the conditional posterior distribution in the latent space. The variable can then be treated as a continuous variable.

3   Experiments and Results

For all our experiments, we used the texture of 200 face scans used for building
the Basel Face Model (BFM) [16]. The scans are in dense correspondence and
were captured under an identical illumination setting. We work on texture images
and use a resolution of 1024×512 pixels. Our experiments are based on the
appearance information only, the last experiment merging the appearance and
shape to a combined model. We used the empirical data directly as marginal
distribution. The results are rendered with an ambient illumination on the mean
face shape of the BFM.

3.1   Facial Appearance Distribution

In a first experiment we investigate if the color intensities in our face data set
are Gaussian-distributed. We followed the protocol of the Kolmogorov-Smirnov
Test [17]. We estimate a Gaussian distribution for every color channel per pixel
and compare it to the observed data. The null hypothesis of the test is that the
observed data is drawn by the estimated Gaussian distribution. The test mea-
sures the maximum distance of the cumulative density function of the estimated
Gaussian $\Phi_{\hat{\mu},\hat{\sigma}^2}$ and the empirical marginal distribution F_{emp} of the observed
data:

$$d = \sup_{x} \left\| F_{\mathrm{emp}}(x) - \Phi_{\hat{\mu},\hat{\sigma}^2}(x) \right\| \tag{9}$$

Here, $\hat{\mu}$ and $\hat{\sigma}^2$ are maximum-likelihood estimates for the mean and variance
of a Gaussian distribution respectively. In Fig. 2(a) we visualize the maximal
distance value over all color channels per point on the surface.

We assume a significance level of $1 - \alpha = 0.05$. The critical value d_α is
approximated using the following formula [18]:

$$d_\alpha = \frac{\sqrt{\ln(\frac{2}{\alpha})}}{\sqrt{2n}} \tag{10}$$

With $n = 200$ training samples we get a critical value of 0.096. Non-Gaussian
marginal distributions of color intensities are present in the region of the eye-
brows, eyes, chin and hair, where multi-modal appearance is present. In total for
49% of the pixels over all color channels, the null hypothesis has to be rejected.
In simple monotonic regions, like the cheek, the marginal distributions are close
to a Gaussian distribution. In more structured regions like the eye, eyebrow or
the temple region, the appearance is highly non-Gaussian. This leads to strong
artifacts when modeling facial color appearance using PCA (see Figs. 3 and 4).
Since those more structured regions are fundamental components of a face, it is
important to model them properly.

We also applied the Kolmogorov-Smirnov Test to the shape coordinates of our
training data (see Fig. 2(b)). We observe that the observed marginal distributions
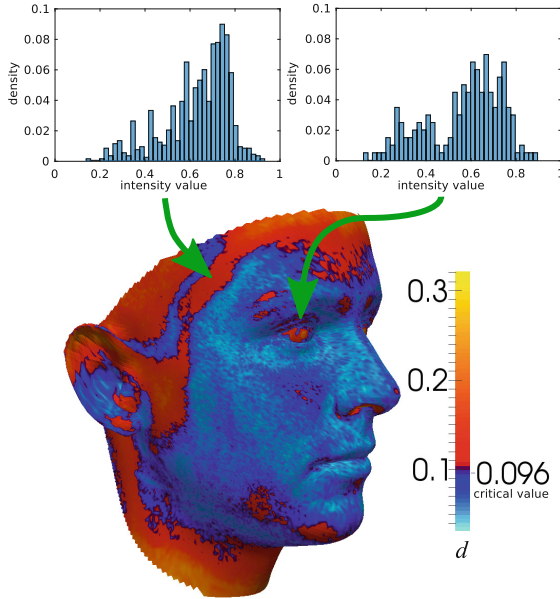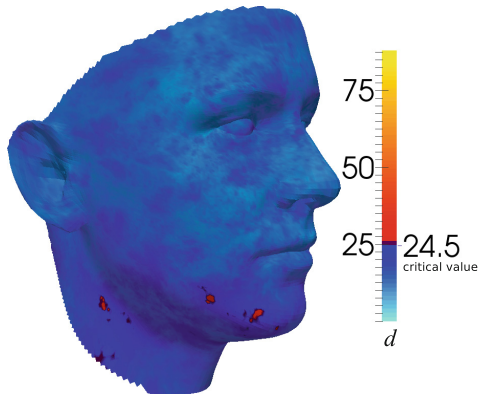
Fig. 2. The result of the Kolmogorov-Smirnov Test to compare the empirical marginal distributions of color values (a) and shape coordinates (b) from our 200 face scans with a Gaussian-reference probability distribution. We plot the highest value of the three color channels respectively dimensions per pixel, because the values for the individual components are very similar. Whilst the marginals for the shape coordinates are similar to a Gaussian distribution, the Gaussian assumption does not hold for the color marginals. We show two exemplary marginal distributions in the eye and temple region. They are not only non-Gaussian but also not similar. (Color figure online)
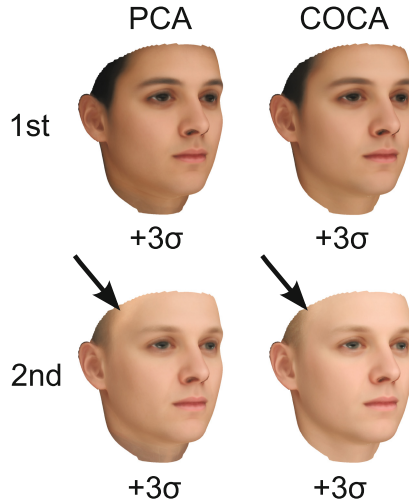
Fig. 3. PCA and COCA are compared by visualizing the first two eigenvectors with 3 standard deviations on the mean. The components look very similar, except that the PCA artifacts on the temple (arrows) in the second eigenvector do not appear using COCA. (Color figure online)

in the data are close to a Gaussian distribution. The registration of the data was performed using a nonrigid ICP algorithm by Amberg et al. [19]. The algorithm uses strong regularization techniques, therefore the Gaussian property of the shape coordinates can also be a registration artifact.

3.2 Appearance Modeling

We evaluate our facial appearance model by its capability to synthesize new instances. We measured this capability by comparing the major eigenmodes, random model instances, the sample marginal distributions and the specificity of both models. The specificity is measured qualitatively by visual examples and quantitatively by a model metric.

Model Parameters. The first few principal components store the strongest dependencies. We visualize the first two components by setting their value h_i to $\sigma = 3$ standard deviations and show the result in Fig. 3. The first parameters of PCA and COCA appear very similar in the variation of the data they model. The second principal component of PCA causes artifacts in the temple region. These artifacts are caused by the linearity of PCA. COCA is a nonlinear method and therefore, the artifacts are not present.

Random Samples. The ability to generate new instances is a key feature for generative models. A model which can produce more realistic samples is desirable
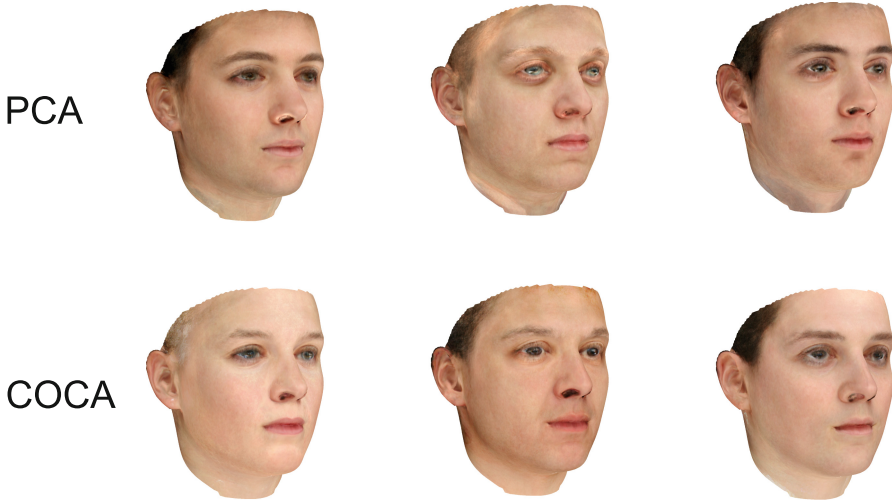
Fig. 4. The first and second row show random samples projected by PCA and COCA respectively. Using PCA, we can observe strong artifacts in the regions where the marginal distribution is not Gaussian (see Fig. 2). The improvement of COCA can be observed in the temple region, on the eyebrows, around the nostrils, the eyelids and at the border of the pupil. We chose representative samples for both methods. (Color figure online)

for various applications. For example, the Visio-lization method to generate high resolution appearances is based on a prototype generated with PCA [4].

Another field of application for the generative part of models are Analysis-by-Synthesis methods based on Active Appearance Models (AAM) or 3D Morphable Models (3DMM). They can profit from a stronger prior which is more specific to faces and reduces the search space [6].

Generating a random parameter vector leads to a random face from our PCA or COCA model. We sample h according to (5) independently for all 199 parameters and project them via PCA or COCA on the color space following (6). Random samples using COCA contain fewer artifacts and, therefore, appear much more natural (see Fig. 4). These artifacts are caused by the linearity of PCA. For non-Gaussian-distributed marginals, PCA does not only interpolate within the trained color distribution but also extrapolates to color intensities not supported by the training data.

The most obvious problem is the limited domain of the color channels: using PCA, color channels have to be clamped. The linearity constraint of PCA leads to much brighter or darker color appearance than those present in the training data in regions which are not Gaussian-distributed. In the next experiment, we show that the higher specificity is not only a qualitative result but can also be measured by a model metric.

(a) Empirical marginal distri-
bution

(b) PCA sample marginal dis-
tribution

(c) COCA sample marginal
distribution
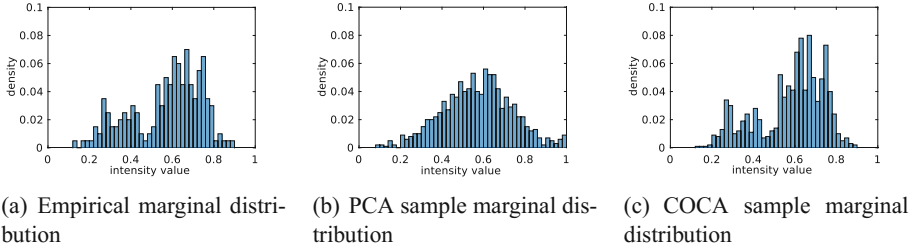
Fig. 5. The marginal distribution of the red color intensity of a single point in the
eye region. (a) shows the distribution observed in the training data, (b) shows the
distribution of samples drawn from a PCA model and (c) from a COCA model. (Color
figure online)

Few samples od COCA contain artifacts arising from outliers in the training
data which appear at the borders of the empirical cdfs. Those artifacts can be
removed by slightly cropping the marginal distributions (removing the outliers)
or by applying COCA in the HSV color space.

3.3 Appearance Marginal Distribution

We analyze the marginal distributions of our random faces at a single point
at the border between the pupil and the sclera of the eye. In this region the
Kolmogorov-Smirnov Test rejected the null hypothesis. We analyze the empirical
intensity distribution of a single color channel at this point (Fig. 5(a)). The
sample marginal distributions drawn from 1000 random instances generated by
PCA and COCA are shown in Fig. 5(b) and (c) respectively. Whilst COCA is able
to generate samples distributed similar to our input data, PCA is approximating
a Gaussian distribution, which is inaccurate in a lot of facial regions.

Specificity and Generalization. To measure the quality of the PCA and
COCA models, we use model metrics motivated by the shape modeling com-
munity [20]. The first metric is specificity: Instances generated by the model
should be similar to instances in the training set. Therefore, we draw 1000 ran-
dom samples from our model and compare each one to its nearest neighbor in
the training data. We measure the distance using the mean absolute error over
all pixels and color channels in the RGB-color space. The COCA model is more
specific to facial appearance (see Fig. 6(a)). This corresponds to our observation
of a more realistic facial appearance (Fig. 4).

Specificity should always be used in combination with the generalization
model metric [20]. The generalization measures how exactly the model can rep-
resent unseen instances. We measure the generalization ability of both models
using a test set and use the same distance measure as for specificity. The test
data consists of 25 additional face scans not contained in the training data.
We observe that both models generalize well to unseen data. PCA generalizes
slightly better, see Fig. 6(b).
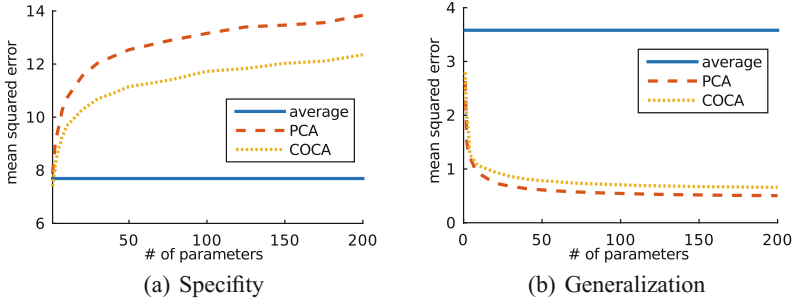
Fig. 6. (a) The specificity shows how close generated instances are to instances in the training data. The average distance of 1000 random samples to the training set (mean squared error per pixel and color channel) is shown. A model is more specific if the distance of the generated samples to the training set is smaller. We observe that COCA is more specific to faces (lower is better). (b) The generalization ability shows how exactly unseen instances can be represented by a model. The lower the error, the better a model generalizes. As a baseline, we present the generalization ability of the average face. We observe that PCA generalizes slightly better (lower is better). (Color figure online)

The third model metric is compactness - the ability to use a minimal set of parameters [20]. The compactness can be measured directly by the number of used parameters. In our experiments, the number of parameters is always the same for both models.

There is always a tradeoff between specificity and generalization. Whilst PCA performs slightly better in generalization, COCA performs better in terms of specificity. The better generalization ability of PCA comes at the price of a lower specificity and clearly visible artifacts.

### 3.4   Combined Shape, Color and Attribute Model

Color appearance and shape are modeled independently in AAMs and 3DMMs. Recently, it was demonstrated that facial shape and appearance are correlated [21] and those correlations were investigated using Canonical Correlation Analysis on separate shape and appearance PCA models. Attributes like age, weight, height, gender are often added to the PCA models as additional linear vectors [16] or with limitations to Gaussian marginal distributions [22].

The main reason to build separate models is a practical one – shape and color values do not live in the same space and are not scaled in the same range. Attributes are even not always continuous. Some methods approach this issue by normalization [23]. However, this approach is highly sensitive to outliers and not suitable to compare those different modalities. Since Copula Component Analysis is scale invariant and allows to include categorical data, we can directly apply it to a set of combined data.
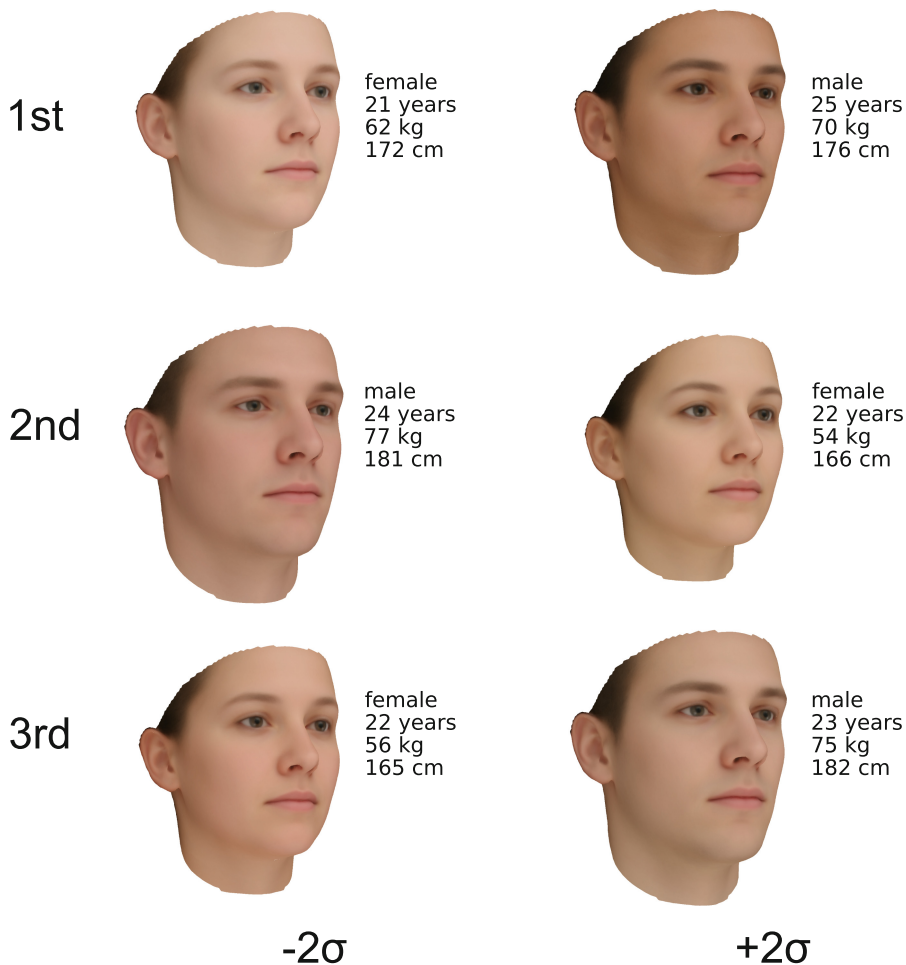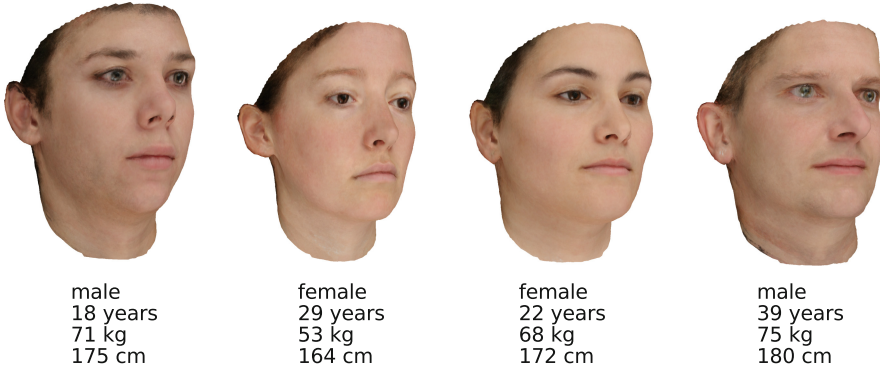
Fig. 7. We learned a common shape, color and attribute model using COCA. We visualize the first eigenvectors with 2 standard deviations, which show the strongest dependencies in our training data. Whilst the first parameter is strongly dominated by color the latter parameters are targeting shape, color and attributes (compare Fig. 8). Since the model is built from 100 females and 100 males, the first components are strongly connected to sex. The small range in age is caused by the training data which mainly consists of people with similar age. (Color figure online)

We learned a COCA model combining the color, shape and attributes information (see Figs. 7 and 9). Shape, color and attributes are combined by simply concatenating them. Age weight and height are continuous attributes and can therefore directly by integrated in the COCA model. We added gender as a binary attribute and used the strategy presented in Sect. 2.5, where we replaced the binary labels with scores, which were learned with logistic regression on
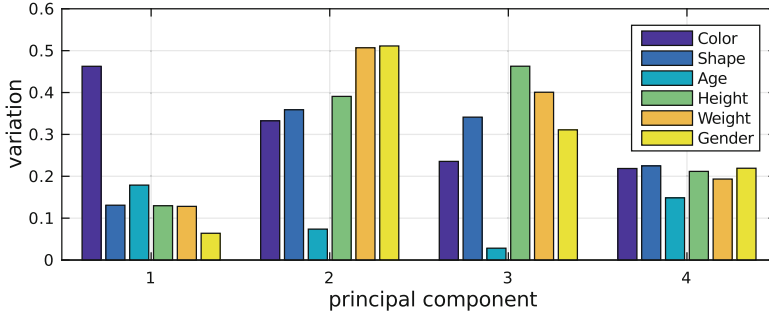
Fig. 8. The influence of the first principal components on the different modalities of our model is shown. The variation is shown as the RMS distance of the normalized attributes in the latent space (covariance matrix). Whilst the first parameter is strongly dominated by color the later parameters are targeting shape, color and attributes (compare Fig. 7). We observe strong correlations between the different modalities and attributes. (Color figure online)

male	female	female	male
18 years	29 years	22 years	39 years
71 kg	53 kg	68 kg	75 kg
175 cm	164 cm	172 cm	180 cm

Fig. 9. Random samples projected by a common shape, color and attribute model using COCA. Our model leads to samples with consistent appearance and attributes. (Color figure online)

the covariates. The combined model allows us to generate random samples with consistent and correlated facial features. In Fig. 8 we present how the different modalities are correlated in the first parameters. By integrating this additional dependency information, the model becomes more specific [23].

4 Conclusions

We showed that the marginals of facial color are not Gaussian-distributed for large parts of the face and that PCA is not able to model facial appearance properly. In a statistical appearance model, this leads to unnatural artifacts which are easily detected by human perception. To avoid such artifacts, we propose to

use PCA in a semiparametric Gaussian copula model (COCA) which allows to model the marginal color distribution separately from the dependency structure. In this model, the parametric Gaussian copula describes the dependency pattern in the data and the nonparametric marginals relax the restrictive Gaussian requirement of the data distribution.

The separation of marginals and dependency pattern enhances the model flexibility. We showed qualitatively that facial appearance is modeled better using COCA than by PCA. This finding is also supported by a quantitative evaluation using specificity as a model metric. Moreover, COCA provides scale invariance and therefore allows us to include different modalities and attributes in a unified way. We presented a combined model including shape, color, attributes like age, weight and height, and even categorical attributes like gender. The scale invariance is a key feature of COCA, it enables us interesting new applications and methods when working with statistical models.

Finally, we again want to encourage the reader to replace PCA with a COCA model, since the additional model flexibility comes with almost no implementation effort. The computer graphics and vision community is heavily modeling and working with color intensities. We believe that these intensities are most often not Gaussian-distributed and, therefore, our findings can be transferred to a lot of applications.

References

1. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Burkhardt, H., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1407, pp. 484–498. Springer, Heidelberg (1998). doi:10.1007/BFb0054760
2. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., pp. 187–194 (1999)
3. Jolliffe, I.: Principal Component Analysis. Wiley, Hoboken (2002)
4. Mohammed, U., Prince, S.J., Kautz, J.: Visio-lization: generating novel facial images. ACM Trans. Graph. (TOG) **28**, 57 (2009)
5. Walker, M., Vetter, T.: Portraits made to measure: manipulating social judgments about individuals with a statistical face model. J. Vis. **9**, 12 (2009)
6. Schönborn, S., Forster, A., Egger, B., Vetter, T.: A monte carlo strategy to integrate detection and model-based face analysis. In: Weickert, J., Hein, M., Schiele, B. (eds.) GCPR 2013. LNCS, vol. 8142, pp. 101–110. Springer, Heidelberg (2013). doi:10.1007/978-3-642-40602-7_11
7. Rasmussen, C.E.: The infinite gaussian mixture model. NIPS **12**, 554–560 (1999)
8. Han, F., Liu, H.: Semiparametric principal component analysis. In: Advances in Neural Information Processing Systems, pp. 171–179 (2012)
9. Sirovich, L., Kirby, M.: Low-dimensional procedure for the characterization of human faces. JOSA A **4**, 519–524 (1987)

10. Turk, M., Pentland, A.P., et al.: Face recognition using eigenfaces. In: Proceedings: 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 1991, pp. 586–591. IEEE (1991)
11. Nelsen, R.B.: An Introduction to Copulas, vol. 139. Springer Science & Business Media, Heidelberg (2013)
12. Joe, H.: Multivariate Models and Multivariate Dependence Concepts. CRC Press, Boca Raton (1997)
13. Genest, C., Ghoudi, K., Rivest, L.P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika **82**, 543–552 (1995)
14. Tsukahara, H.: Semiparametric estimation in copula models. Can. J. Stat. **33**, 357–375 (2005)
15. Hoff, P.D.: Extending the rank likelihood for semiparametric copula estimation. Ann. Appl. Stat. **1**(1), 265–283 (2007)
16. Paysan, P., Knothe, R., Amberg, B., Romdhani, S., Vetter, T.: A 3D face model for pose and illumination invariant face recognition. In: Sixth IEEE International Conference On Advanced Video and Signal Based Surveillance AVSS 2009, pp. 96–301. IEEE (2009)
17. Massey Jr., F.J.: The Kolmogorov-Smirnov test for goodness of fit. J. Am. Stat. Assoc. **46**, 68–78 (1951)
18. Lothar Sachs, J.H.: Angewandte Statistik, 7th edn. Springer, Heidelberg (2006)
19. Amberg, B., Romdhani, S., Vetter, T.: Optimal step Nonrigid ICP algorithms for surface registration. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007, pp. 1–8. IEEE (2007)
20. Styner, M.A., Rajamani, K.T., Nolte, L.-P., Zsemlye, G., Székely, G., Taylor, C.J., Davies, R.H.: Evaluation of 3D correspondence methods for model building. In: Taylor, C., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 63–75. Springer, Heidelberg (2003). doi:10.1007/978-3-540-45087-0_6
21. Schumacher, M., Blanz, V.: Exploration of the correlations of attributes and features in faces. In: 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp. 1–8. IEEE (2015)
22. Blanc, R., Seiler, C., Székely, G., Nolte, L.P., Reyes, M.: Statistical model based shape prediction from a combination of direct observations and various surrogates: application to orthopaedic research. Med. Image Anal. **16**, 1156–1166 (2012)
23. Edwards, G.J., Lanitis, A., Taylor, C.J., Cootes, T.F.: Statistical models of face images improving specificity. Image Vis. Comput. **16**, 203–211 (1998)