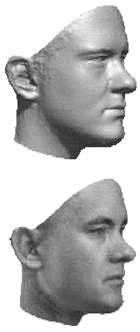
A Morphable Model for the Synthesis of 3D Faces

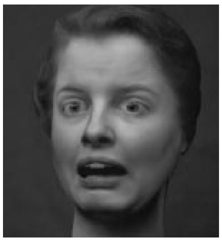


Marco Nef

Volker Blanz, Thomas Vetter
SIGGRAPH 99, Los Angeles

Presentation overview

- Motivation
- Introduction
- Database
- Morphable 3D Face Model
- Matching a Morphable Model to Images
- Results
- Demonstration
- Future work
- Conclusions
- References



Motivation

Factors that make realistic facial animation elusive:

- face is an extremely complex geometric form
- face exhibits countless tiny creases and wrinkles, as well as subtle variations in color and texture that are crucial for comprehension of expressions

Even more problematic to animate:

- facial movement is a product of the underlying skeletal and muscular forms
- we have an uncanny ability to read expressions and detect the slightest deviation

3

2D-morphing:

- remarkable results have been achieved, most notably, perhaps, a Michael Jackson video
- requires animators to specify carefully chosen correspondences between physical features of the actors in almost every frame
- doesn‘t correctly account for changes in viewpoint or object pose

3D-morphing:

- complete freedom of viewpoint and the ability to composite the image with other 3D graphics

4

Introduction

Problem of finding corresponding feature locations:

- a limited number of labeled feature points marked in one face (50..300) must be located precisely in another face
- only a correct alignment of all these points allows acceptable intermediate morphs
- automated matching techniques only for very prominent feature points

Problem of separating realistic faces

- requires time-consuming manual work combined with the skills of an artist

5

Presented parametric face modeling technique:

- arbitrary human faces can be created simultaneously controlling the likelihood of the generated faces
- the system is able to compute correspondence between new faces

Morphable face model:

- multidimensional 3D morphing function that is based on the linear combination of a large number of 3D face scans
- computing the average face and the main modes of variation in the dataset, a probability distribution is imposed on the morphing function to avoid unlikely faces
- parametric descriptions of face attributes

6

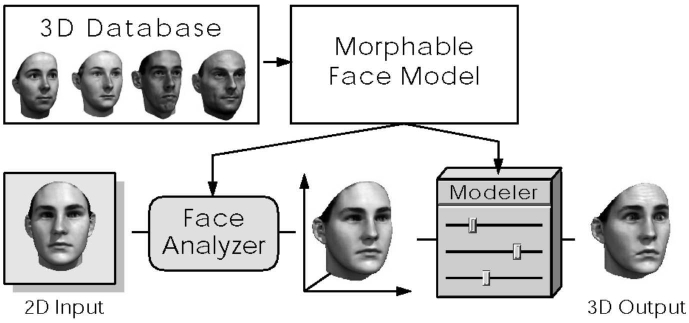
Output of the matching procedure:

- high quality 3D face model that is in full correspondence with the morphable face model
- consequently all face manipulations parameterized in the model function can be mapped to the target face

> «When applying our method to several images of a person, the reconstructions reach almost the quality of laser scans»

Volker Blanz, Thomas Vetter

7



Derived from a dataset of prototypical 3D scans of faces, the morphable face model contributes to two main steps in face manipulation: deriving a 3D face model from a novel image, and modifying shape and texture in a natural way

8

Database

- laser scans (*Cyberware*™) of 200 heads of young adults (100 male and 100 female, Caucasian) were used
- head structure data in a cylindrical representation:
  radii $r(h,\phi)$ of surface points, sampled at 512 equally-spaced angles ϕ and vertical steps h,
  and RGB-color values $R(h,\phi)$, $G(h,\phi)$, $B(h,\phi)$
- all faces were without makeup, accessories, and facial hair
- the subjects were scanned wearing bathing caps, that were removed digitally
- vertical cut behind the ears, horizontal cut to remove the shoulders

9

Morphable 3D Face Model

- morphing between faces requires full correspondence between all of the 3D faces in dataset → assumption

Geometry of a 3D face is represented by

- shape-vector $S = (X_1, Y_1, Z_1, X_2, \ldots Y_n, Z_n)^T \in \Re^{3n}$
- texture-vector $T = (R_1, G_1, B_1, R_2, \ldots G_n, B_n)^T \in \Re^{3n}$
- the morphable face model was constructed using a data set of m exemplar faces, each represented by its S_i and T_i

$$S_{\text{mod}} = \sum_{i=1}^{m} a_i S_i \ , \ T_{\text{mod}} = \sum_{i=1}^{m} b_i T_i \ , \ \sum_{i=1}^{m} a_i = \sum_{i=1}^{m} b_i = 1$$

10

- the morphable model is the set of faces $(S_{\text{mod}}(\vec{a}), T_{\text{mod}}(\vec{b}))$ parameterized by the coefficients $\vec{a} = (a_1, a_2, \ldots a_m)^T$ and $\vec{b} = (b_1, b_2, \ldots b_m)^T$
- arbitrary new faces can be generated by varying the parameters that control shape and texture

Probability distribution:

- distribution for the coefficients a_i and b_i estimated from example set of 200 faces in order to quantify the results in terms of their plausibility of being faces

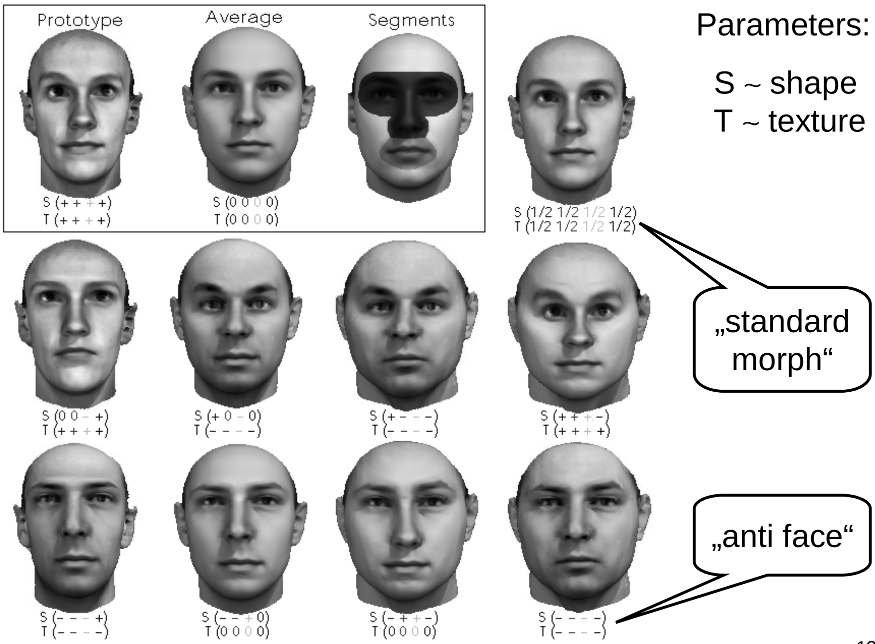
$\rightarrow$ control of the likelihood of the appearance of the generated faces

11

Segmented morphable model:

- morphable model has m-1 degrees of freedom for both texture and shape
  $\rightarrow$ no chance to control them separately
- the model can become more expressive by dividing faces into independent subregions that are morphed independently
- because of correspondence between faces, this segmentation is equivalent to subdividing the vector space of faces into independent subspaces
- complete 3D face is generated by computing linear combinations for each segment separately

12



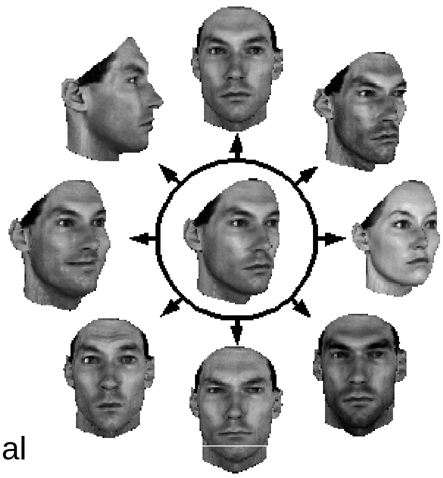
Parameters:

S ~ shape
T ~ texture

13

Facial attributes:

- shape and texture coefficients do not correspond to the facial attributes used in human language
- define vectors that manipulate a specific attribute while keeping all other attributes as constant as possible
- can be recorded from two scans of the same individual with different expressions



14

Matching a Morphable Model to Images

- crucial element of framework is an algorithm for automatically matching the morphable face model to images

Model parameters:

- $a_i, b_i, i = 1, \ldots m-1$ from morphable model
- $\vec{\rho}$ contains camera position, object scale, image plane rotation and translation, intensities of ambient and directed light, color contrast, ...
- from parameters $(\vec{a}, \vec{b}, \vec{\rho})$, colored images

  $$I_{model}(x, y) = (I_{r,mod}(x, y), I_{g,mod}(x, y), I_{b,mod}(x, y))^T$$

  are rendered using perspective projection and the Phong illumination model

15

- the reconstructed image is supposed to be closest to the input image in terms of Euclidean distance

  $$E_I = \sum_{x,y} \left\| I_{input}(x, y) - I_{model}(x, y) \right\|^2$$

- along with the desired solution, many non-face-like surfaces lead to the same image $\rightarrow$ shape and texture vectors in morphable model are restricted to the vector space spanned by the database

16

Algorithm (coarse to fine):

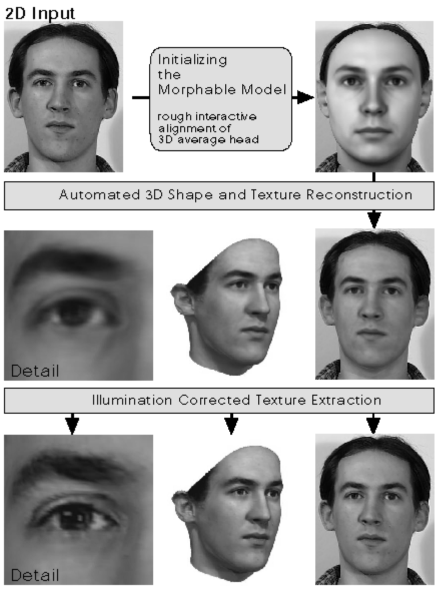
- the first set of iterations is performed on a downsampled version of the input image with a low resolution morphable model
- start by optimizing only the first coefficients a_i and b_i controlling the first principal components, along with all parameters ρ_i
- in subsequent iterations, more and more principal components are added
- in the last iterations, the face model is broken down into segments. With parameters ρ_i fixed, coefficeints a_i and b_i are optimized independently for each segment

Variation: multiple input images
→ fixes e.g. occluded areas

17

Processing steps for reconstructing 3D shape and texture of a new face from a single image:

1. a rough manual alignment of the average 3D head
2. automated matching procedure fits the 3D morphable model to the image
3. details in texture can be improved by llumination-corrected texture extraction from the input



18

Results

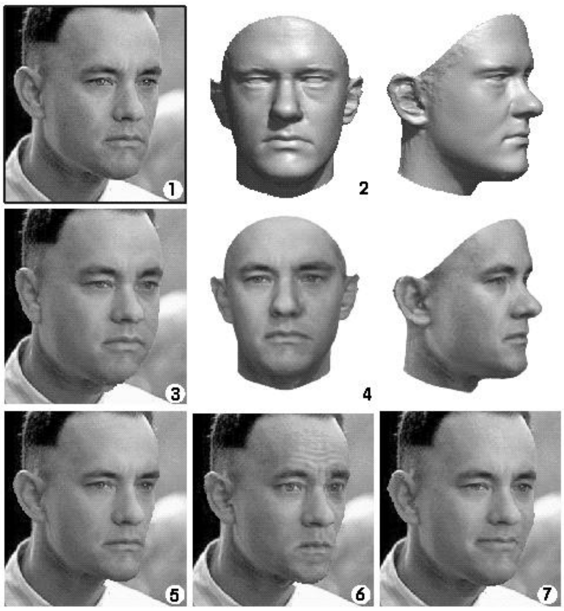
Computational environment / experiments:

- in all examples a morphable model built from the first 100 shape and the first 100 texture principal components that were derived from the whole dataset was matched, and each component additionaly was segmented into four parts
- matching procedure was performed in 10^5 iterations
- 50 minutes on a SGI R10‘000 processor

Disadvantages:

- differences become visible for large rotations >60º

19

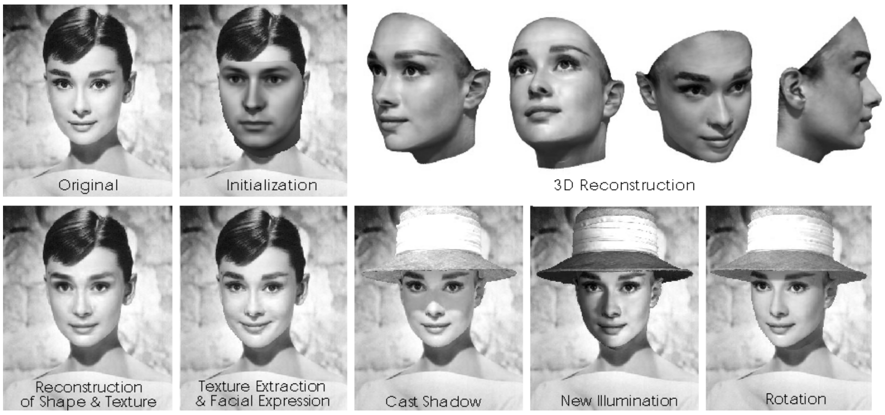


Tom Hanks...

1. single image
2. 3D shape and texture map estimate
3. gaining weight
4. additional texture extraction
5. loosing weight
6. frowning
7. being forced to smile

20

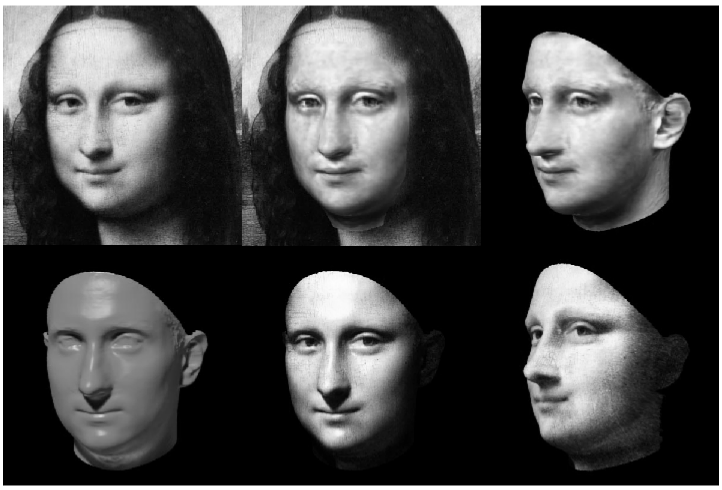
Audrey Hepburn... [Edda H. van Heemstra]



After manual initialization, the algorithm matches a colored morphable model to the image

21

Signora Mona Lisa [~1503]



22

Demonstration

Future work

Issues of implementation:

- speed up matching algorithm
- data reduction applied to shape and texture data will reduce redundancy of the representation, saving computation time

Extending the database:

- children, elderly people as well as other races
- time course of facial expressions, variations during speech

Extending the face model:

- automated reconstruction of hair styles from images

Conclusions

Advantages:

- easy idea, many features
- split of the image into face and background
- combination of face with other 3D graphic objects

Many applications:

- movies, 3D-visualization on the web, forgery...

Possible extensions:

- not only faces, but whole bodies

References

- Volker Blanz, Thomas Vetter
  A Morphable Model for the Synthesis of 3D Faces
  SIGGRAPH'99
- F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, D. Salesin
  Synthesizing realistic facial expressions from photographs
  SIGGRAPH'98

http://www.mpik-tueb.mpg.de/bu/people/vetter/Sigg99.html

シマ http://n.ethz.ch/student/manef/
http://www.shima.ch