Face Modeling

Topics in Image-Based Modeling and Rendering

CSE291 J00

Lecture 17

Faces

From Romdhani slides

Approaches

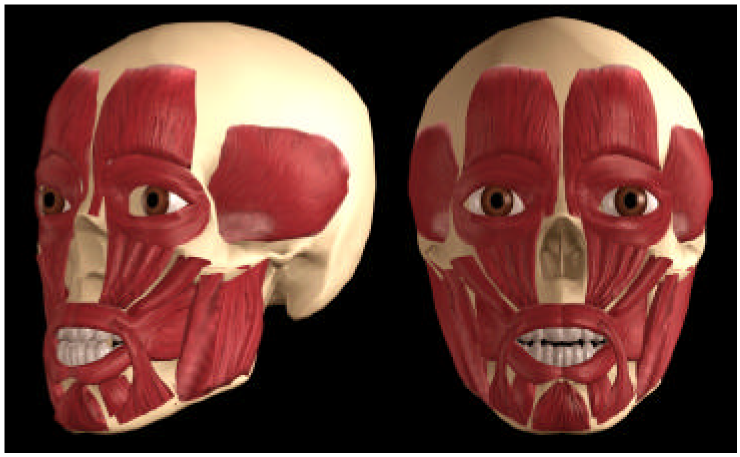
- 2-D Models – morphing, indexing, etc.
- Parameterized face models – e.g. muscle, FACS, etc.
- 3-D modeling, minimal face priors, [Georghiades et al]
- 3D modeling, face prior [Pighin et al]
- 3D modeling, learned prior, [Blanz, Vetter]

Parameterized face models

Muscle-based:

- Waters K., A Muscle Model for Animating Three-Dimensional Facial Expression. SIGGRAPH 1987

Anatomy



From presentation by W. Chang, P. Salmon

Skin as Mesh

- Nodal mobility
  - Tensile Strength of skin
  - Proximity to muscle attachment
  - Depth of tissue & proximity to bone
  - Elasticity & interaction with other muscles
- Network of springs
  - $p = F/k$

Mesh expression examples

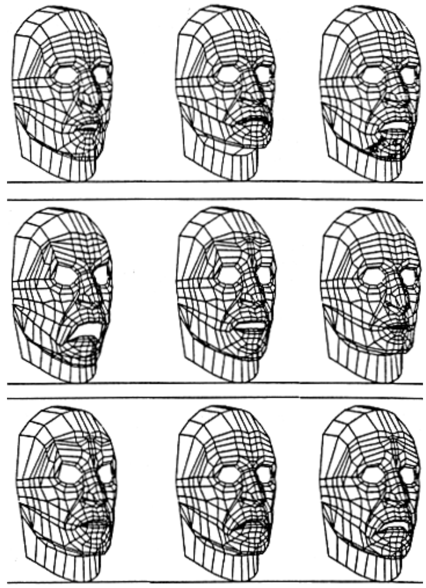
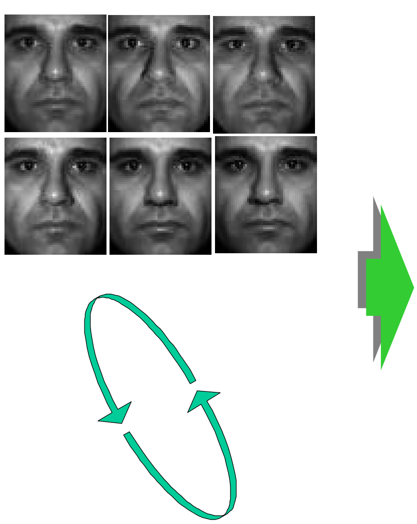
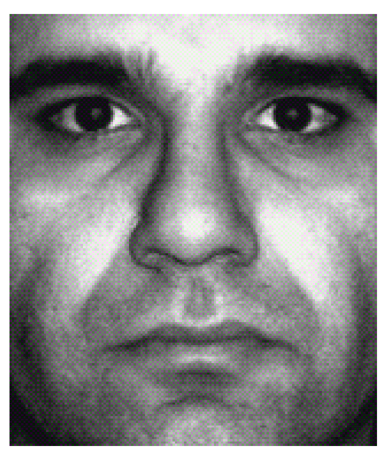


Image-Based Rendering: Attached Shadows

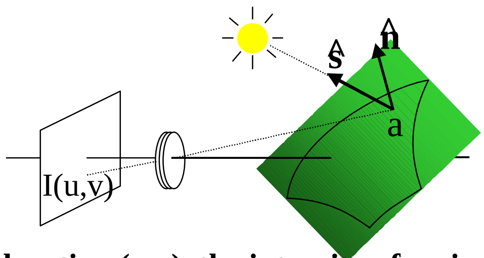


Single Light Source



Face Movie

Lambertian Surface: $\rho(\theta_{in}, \phi_{in} ; \theta_{out}, \phi_{out})$=constant



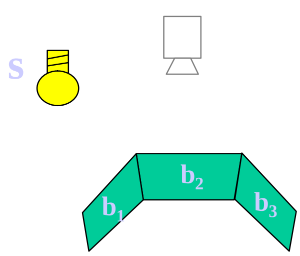
At image location (u,v), the intensity of a pixel I(u,v) is:

$$I(u,v) = [a(u,v)\hat{\mathbf{n}}(u,v)] \cdot [s_0 \hat{\mathbf{s}}] \\ = \mathbf{b}(u,v) \cdot \mathbf{s}$$

where

- **a(u,v) is the albedo of the surface projecting to (u,v).**
- **$\hat{n}$(u,v) is the direction of the surface normal.**
- **s_0 is the light source intensity.**
- **$\hat{s}$ is the direction to the light source.**

Image Formation Model: No shadows



Lambertian model without shadowing:

I =B s

$$\mathbf{B} = \begin{bmatrix} -\ \mathbf{b}^T_1\ - \\ -\ \mathbf{b}^T_2\ - \\ \dots \\ -\ \mathbf{b}^T_n\ - \end{bmatrix}_{n \times 3}$$

where

I is an *n*-pixel image vector

B is a matrix whose rows are unit normals scaled by the albedos

s Î $\mathbf{R}^3$ is a vector of the light source direction scaled by intensity

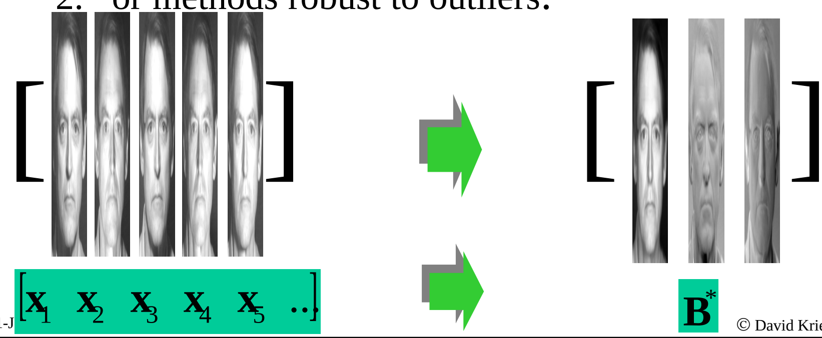
Computing L

For k images $X = [x_1, x_2, \ldots, x_k]$ imaged under k unknown point light sources $S = [s_1, s_2, \ldots, s_k]$,

$$X = B\,S$$

Given $k \geq 3$ images we can compute B* that spans L with

1. singular value decomposition
2. or methods robust to outliers.



Factoring

Given an n by m data matrix **X** containing measured feature points, it can be factored using singular value decomposition as:

$$X = UDV^T$$

Where

D: m by m diagonal matrix, non-negative entries, called singular values

U: n by m with orthogonal columns

V^T: m by m with orthogonal columns

$$D = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_m \end{bmatrix}, s_1 \geq s_2 \geq s_3 \geq s_m$$

- Without noise, $\sigma_i = 0$, $i > 3$
- With noise, set $\sigma_i = 0$, $i > 3$

Factoring: After setting $\sigma_i = 0$, $i > 3$

$$X' = U' D' V'^T$$

Where

- D’: 3 by 3 diagonal matrix, non-negative entries
- U: n by 3 with orthogonal columns
- V^T: 3 by n with orthogonal columns

$$X' = U' D' V'^T = B^* S^*$$

$$where$$

$$B^* = U' D'^{1/2}$$

$$S^* = D'^{1/2} V'^T$$

Face Basis

Original Images

Basis Images spanning L

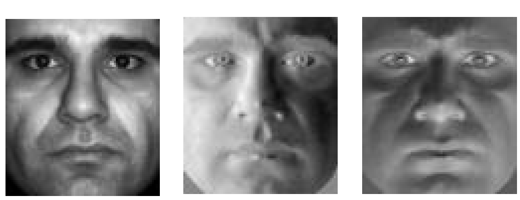
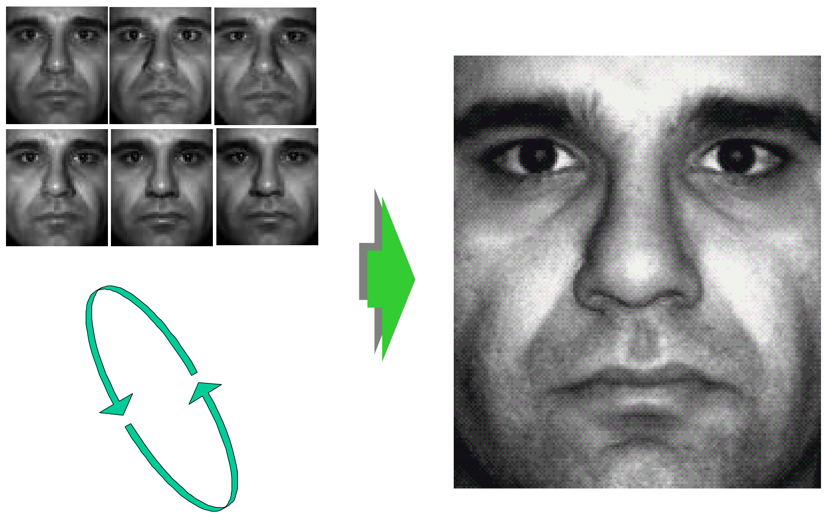


Image-Based Rendering: Attached Shadows



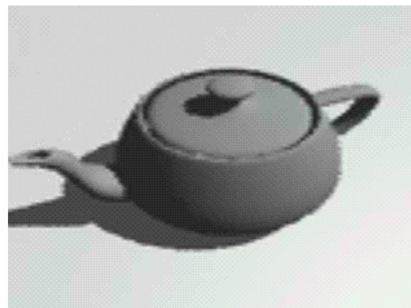
Single Light Source   Face Movie

What do you see?

- Changing viewpoint
- Moving light source
- Deforming shape

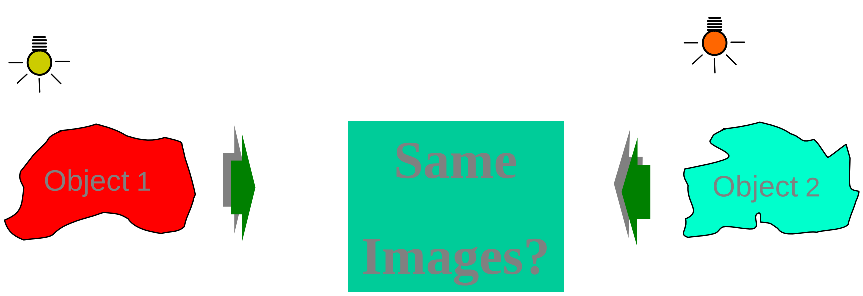
What was happening

- Changing viewpoint
- Moving light source
- ✓ Deforming shape



Do Ambiguities Exist?

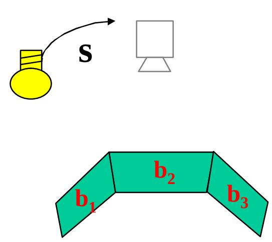
Can two objects of differing shapes and reflectance functions produce the same set of images?

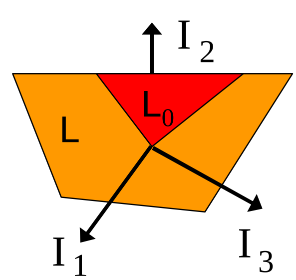


Do Ambiguities Exist? Yes

- Set of images is determined by linear subspace **L**
- The columns of **B** span **L**
- For any **A** Î **GL(3), B* = BA** also spans **L ,**
  i.e. X = B*S* = B AA^{-1}S
-

Illumination Subspace





$$L = \{\mathbf{I} \mid \mathbf{I} = B\mathbf{s}, \text{ for all } \mathbf{s} \in R^3\}$$

- L is a 3-D linear subspace of image space, $\mathbf{R}^n$.
- L is spanned by 3 linearly independent images.

From Normals to Surfaces

- Both $\mathbf{B}^* = \mathbf{BA}$ and $\mathbf{B}$ generate the same illumination cone.

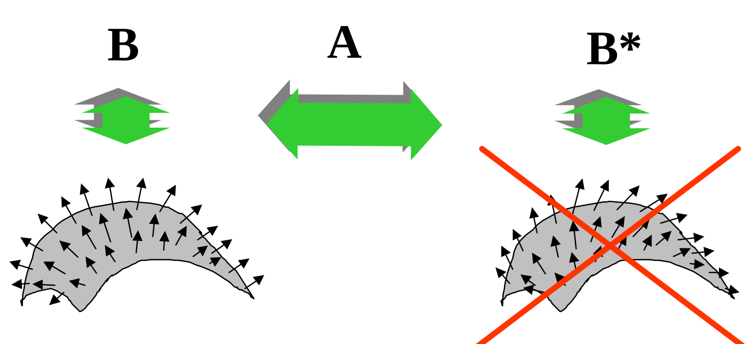
- **A question arises:**

Since $\mathbf{B}/|\mathbf{B}|$ is the normal field of a surface $z=f(x,y)$,
is $\mathbf{B}^*/|\mathbf{B}^*|$ also the normal field of a surface?

Surface Integrability

In general, $\mathbf{B}^*$ does not have a corresponding surface.

Linear transformations of the surface normals in general do not produce an integrable normal field.



Surface Integrability

A surface *f(x,y)* must satisfy the following constraint:

Thus, **b** must satisfy

$$f_{xy} = f_{yx}$$

where

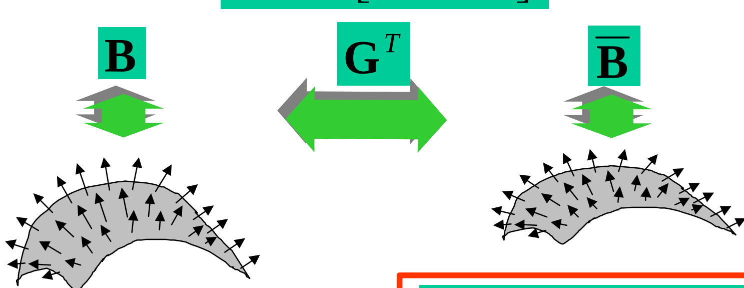
$$\frac{\partial}{\partial y}\left(\frac{b_1}{b_3}\right) = \frac{\partial}{\partial x}\left(\frac{b_2}{b_3}\right)$$

$$\mathbf{b} = [b_1 \quad b_2 \quad b_3] = \frac{1}{\sqrt{f_x^{\ 2} + f_y^{\ 2} + 1}}[af_x \quad af_y \quad -a]$$

GBR Transformation

Only Generalized Bas-Relief transformations satisfy the integrability constraint:

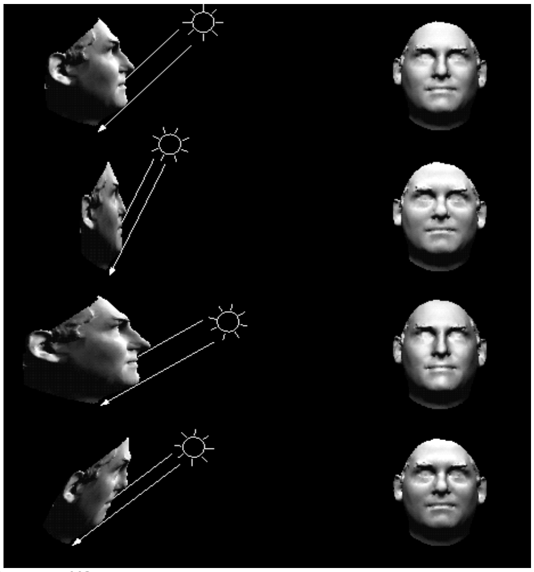
$$\mathbf{A} = \mathbf{G}^T = \begin{bmatrix} l & 0 & -m \\ 0 & l & -n \\ 0 & 0 & 1 \end{bmatrix}^T$$



$$f(x,y)$$

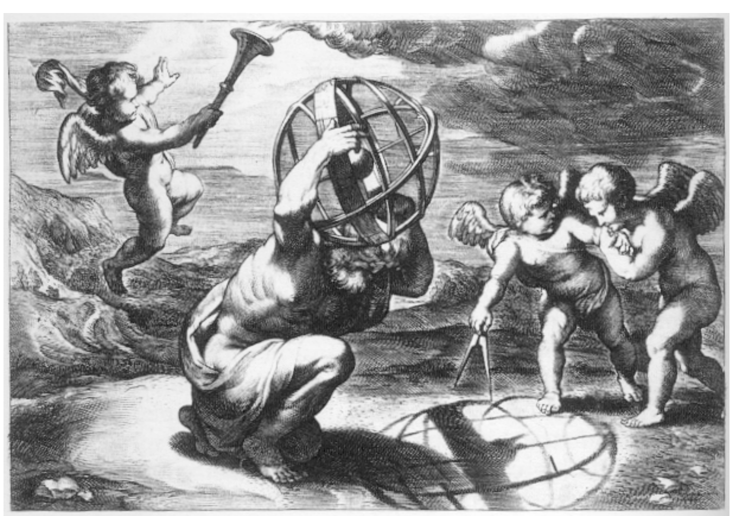
$$\bar{f}(x,y) = l f(x,y) + mx + ny$$

Generalized Bas-Relief Transformations



Without knowledge of light source location, one can only recover surfaces up to GBR transformations.

What about cast shadows for nonconvex objects?



P.P. Reubens in Opticorum Libri Sex, 1613

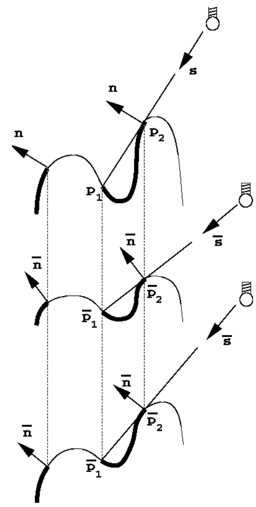
CS291-J00, Winter 2003
© David Kriegman, 2003

GBR Preserves Shadows

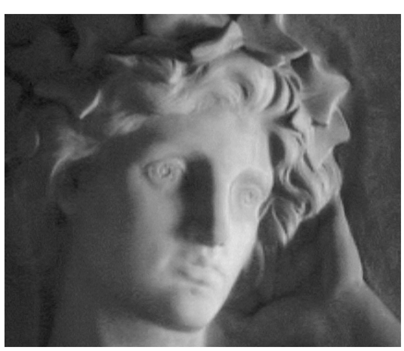
Given a surface f and a GBR transformed surface f' then for every light source **s** which illuminates f there exists a light source **s'** which illuminates f' such that the **attached** and **cast shadows** are identical.

GBR is the only transform that preserves shadows.

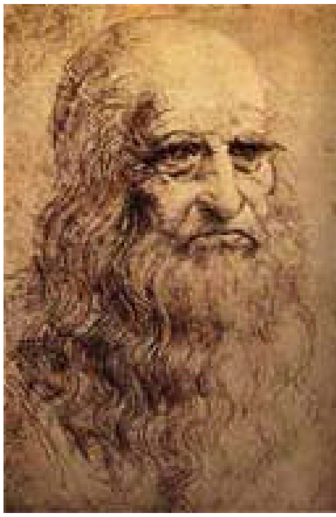
[Kriegman, Belhumeur 2001]



Bas-Relief Sculpture




Codex Urbinas

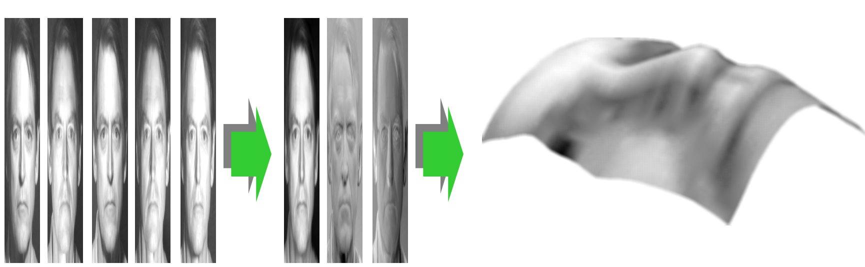


As far as light and shade are concerned low relief fails both as sculpture and as painting, because the shadows correspond to the low nature of the relief, as for example in the shadows of foreshortened objects, which will not exhibit the depth of those in painting or in sculpture in the round.

Leonardo da Vinci
Treatise on Painting (Kemp)

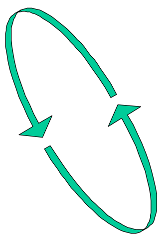
GBR and Reconstruction

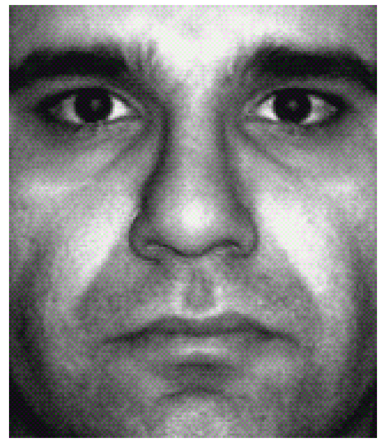
Proposition: From as few as three images of a Lambertian surface, taken under different lighting conditions, we can reconstruct the surface up to a Generalized Bas-Relief transformation – without knowledge of the light sources**.** (See also Fan & Wolff 97.)



Images **B'** *f '(x,y)*

Image-Based Rendering: Cast Shadows
Note: GBR is NOT resolved





Single Light Source

Face Movie

Recovering a GBR surface

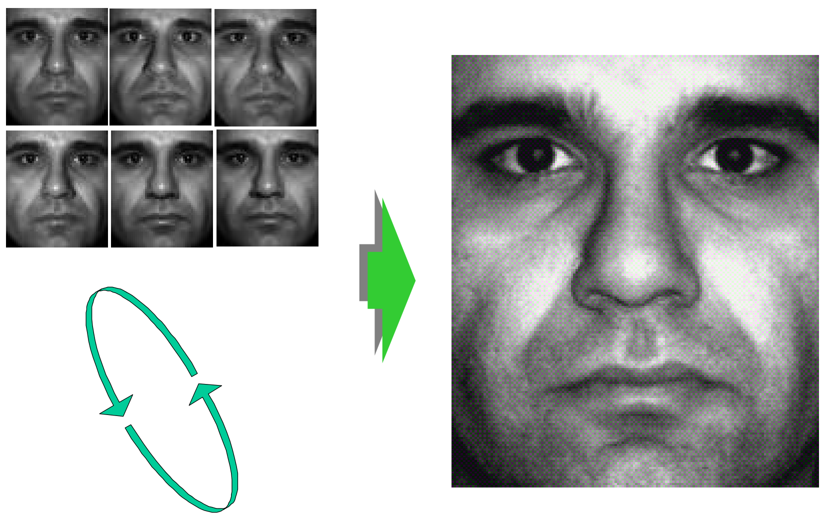
1. Estimate a matrix **A** that makes B*A close to integrable.

2. Integrate the vector field or fit a surface z(x,y) to the vector field minimizing

$$\int\int (z_x - p)^2 + (z_y - q)^2 \, dxdy$$

Where p, q are estimates of the gradient (z_x, z_y) from B*A.

Image-Based Rendering: Attached Shadows

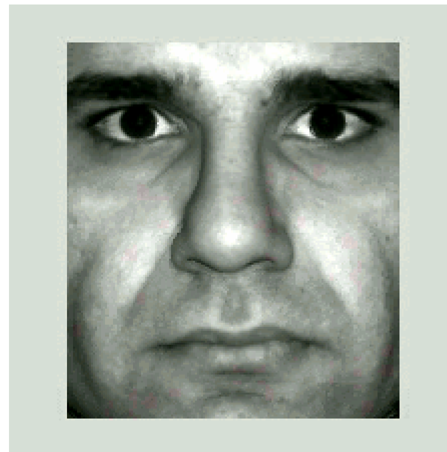


Single Light Source Face Movie

Resolve GBR

- Symmetry
- Canonical face model

Resolve GBR using face information



Synthesizing Realistic Facial Expressions from Photographs

- 3D facial models derived from photographs.
- Smooth transitioning between model expressions.
- Adaptation from one model to another.

- F. Pighin, J. Hecker, D. Lischinski, D. H. Salesin, R. Szeliski. Synthesizing realistic facial expressions from photographs. SIGGRAPH 98, pp. 75-84.

Approach

- Capture multiple views of a human subject
- Manually mark a small set of correspondence points
- Automatically recover camera parameters
  as well as 3D position of marked points in space
- Deform a generic 3D face mesh to fit the particular human subject (Model fitting)
- Extract one or more texture maps for the 3D model from the photos
- Repeat for several facial expressions
- Perform facial animation: interpolation between two or more 3D models while blending the texture

Model Fitting

- Adapt generic face model to fit an individual face and facial expression
- Input
  - Several images of the face from different view points
  - General face 3D model
- Output
  Face model that has been adapted to fit the face in input images

Model fitting process:

- Pose recovery
- Scattered data interpolation
- Correspondence-based shape refinement

Generic head model

CS291-J00, Winter 2003

Input images with marked feature points

CS291-J00, Winter 2003

Pose Recovery

- Start with a rough knowledge of camera position
- Interactively improve the pose and 3D shape

- Some mathematics…

R^k - rotation matrix, composed of three rows, r^k_x, r^k_y, r^k_z.
t^k - translation vector with three entries, t^k_x, t^k_y, t^k_z.
f^k - Focal Length
p_i - 3D coordinate of a specific face feature
x^k_i, y^k_i - 2D coordinate in the k'th image

Pose Recovery – cont.

Each pixel coordinate is given by:

$$x_i^k = f^k \frac{r_x^k p_i + t_x^k}{r_z^k p_i + t_z^k} \qquad y_i^k = f^k \frac{r_y^k p_i + t_y^k}{r_z^k p_i + t_z^k} \qquad (1)$$

Substituting

?^{k} = 1 / t^k_z Inverse distance

$S^k = f^k$ * ?^{k} World to image scale factor

$$x_i^k = s^k \frac{r_x^k p_i + t_x^k}{1 + h^k r_z^k p_i} \qquad y_i^k = s^k \frac{r_y^k p_i + t_y^k}{1 + h^k r_z^k p_i}$$

© David Kriegman, 2003

Pose Recovery – cont.

Let w^k_i be the inverse denominator

$$w_i^k = (1 + h^k (r_z^k p_i))^{-1}$$

Collecting the terms on the left-hand side to yield:

$$w_i^k (x_i^k + x_i^k h^k (r_z^k p_i) - s^k (r_x^k p_i + t_x^k)) = 0$$
$$w_i^k (y_i^k + y_i^k h^k (r_z^k p_i) - s^k (r_y^k p_i + t_y^k)) = 0 \quad (2)$$

Pose Recovery – cont.

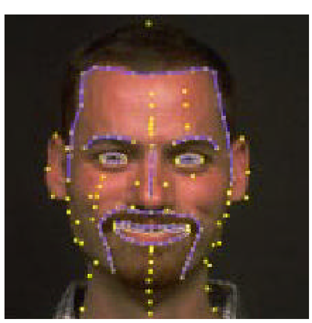
The above equations are solved as follows:

- Maximum likelihood estimation of initial values is obtained using least squares
- Solving the equation for different subsets of unknowns, in five steps:
  first s^k, then p_i, R^k, t^k_x and t^k_y,
  and finally ?^{k} using linear least squares algorithm

3D Face Model

Generic model Fit to 13 pts

3D face model refinement



112 points

Scattered Data Interpolation

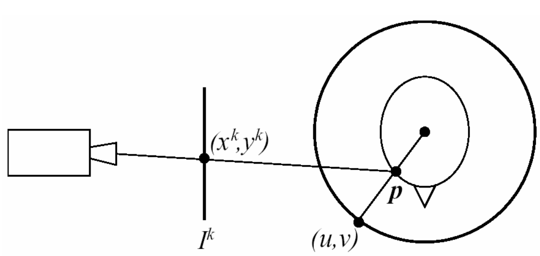
Once an initial set of coordinates for the feature points p_i have been computed, these values are used to deform the remaining vertices on the face mesh

The interpolation function:

$$f(p) = \sum_i c_i f(\|p - p_i\|) + Mp + t$$

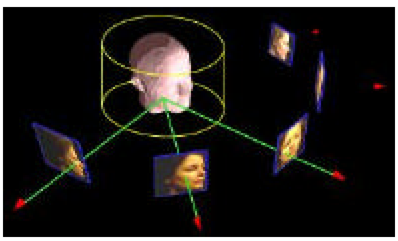
View-independent texture mapping

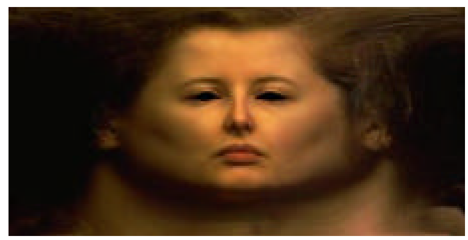
- The texture map is constructed on a virtual cylinder enclosing the face model
- m^k is now indexed by the (u, v) texture coordinates; $m^k = F^k(u,v)P^k(p)$
  - F^k - Feathered visibility map (0 - 1)
  - $P^k(p)$ – positional certainty of p



View Independent Texture Extraction

- Blend photographs to form single texture.
  - Map onto virtual cylinder.





View-dependent texture mapping

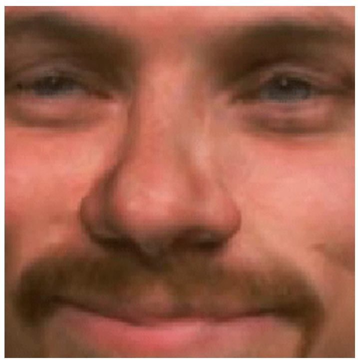
- Associate texture coordinate and a blending weight for each vertex in the face mesh
- Two photographs which are closest to the viewing direction *d*, are blended using a blending function $V^k(d)$
- $m^k = F^k(x^k,y^k)P^k(p)V^k(d)$

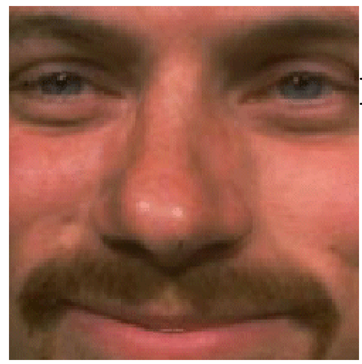
Special treatment for eyes, teeth, ears and hair

View Independent Texture Extraction

- Blurry

Independent





Dependent

Expression morphing

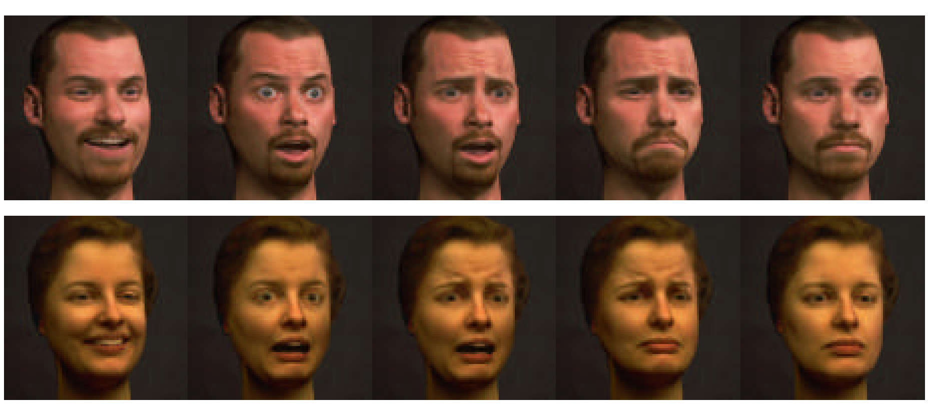
- Goal: generation of continuous and realistic transitions between different facial expressions
- Geometry interpolation
  - Topology of all the faces meshes is identical – simple linear interpolation
- Blending the textures
  - Rendering intermediate face twice
  - Blending is done on the 2D images
- Global blend
- Local blend
- Animation and derivative animations

Results

- Show movie

Results

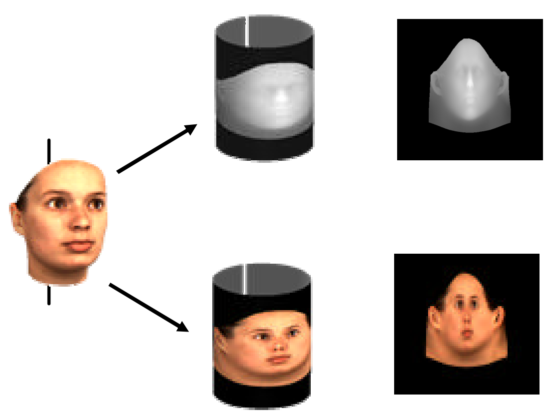
- Applied transitions to different human subject:



- A morphable model for the synthesis of 3D faces Volker Blanz, Thomas Vetter, SIGGRAPH 99, pp. 187 – 194.

3D Morphable Model - Key Features

1. Representation = 3D Shape + Texture Map



3D Shape s

Texture Map t

3D Morphable Model - Key Features 2

2. Accurate & Dense Correspondence

→ PCA accounts for intrinsic ID parameters only

$s = a_1 \cdot$  $+ a_2 \cdot$  $+ a_3 \cdot$  $+ a_4 \cdot$  $+ \ldots = \mathbf{S} \cdot \mathbf{a}$

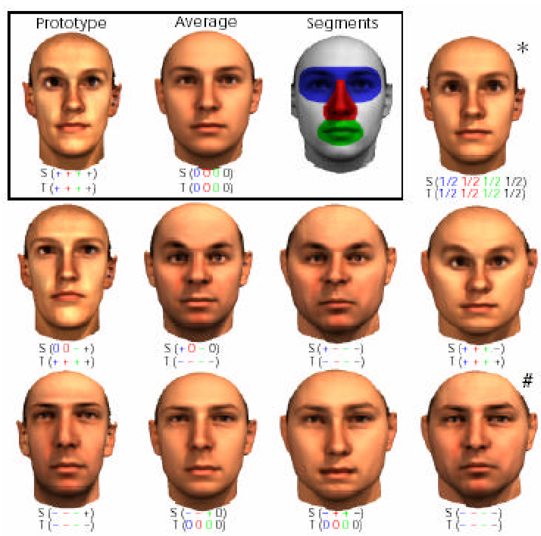
$t = b_1 \cdot$  $+ b_2 \cdot$  $+ b_3 \cdot$  $+ b_4 \cdot$  $+ \ldots = \mathbf{T} \cdot \mathbf{ß}$

Building a Morphable Model

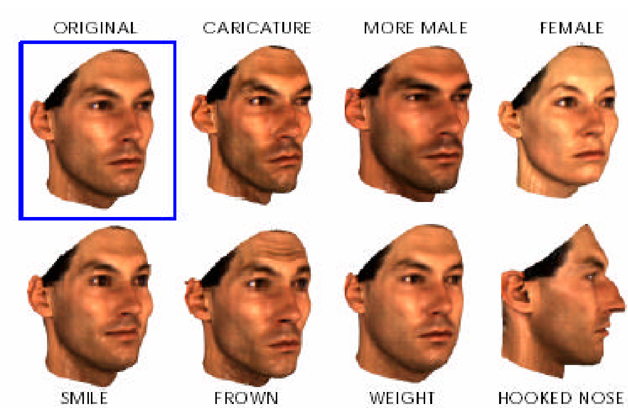
- Align training set using range data and reflectance from cyberware data
- Optical flow
- PCA on Depth & texture

- Initial model

- Re-register all data with model
- Recompute PCA with increasing subspace Dimension



Deviation from protoype



3D Morphable Model - Key Features 3

3. Extrinsic parameters modeled using Physical Relations:
   - Pose : 3x3 Rotation matrix

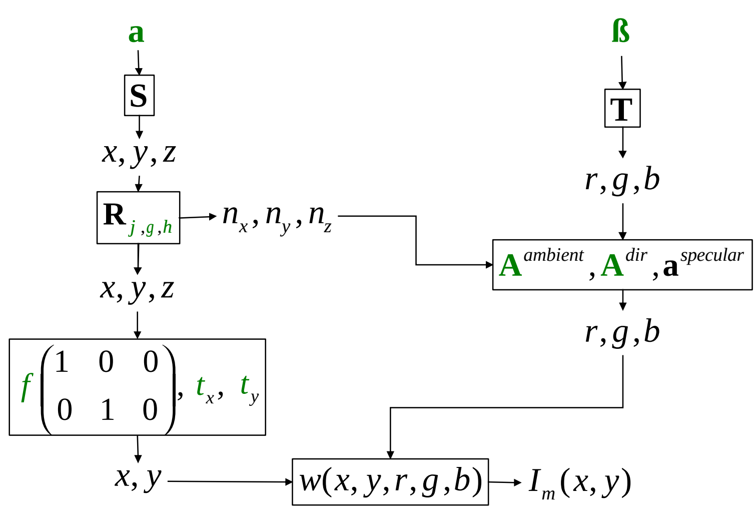
$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = f \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \mathbf{R}_{j,g,h} \cdot \mathbf{S}_k \cdot \mathbf{a} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

   - Illumination : Phong shading accounts for cast shadows and specular highlights
     → No Lambertian Assumption.

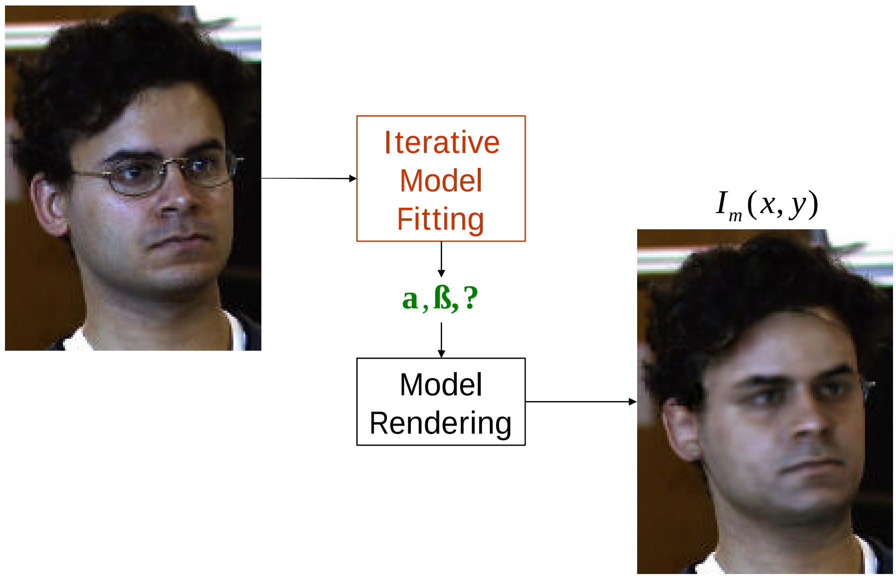
$$\begin{pmatrix} r_k \\ g_k \\ b_k \end{pmatrix} = (\mathbf{A}^{ambient} + \mathbf{A}_k^{dir}) \cdot \mathbf{T}_k \cdot \text{ß} + \mathbf{a}_k^{specular}$$

3D Morphable Model - Key Features 4

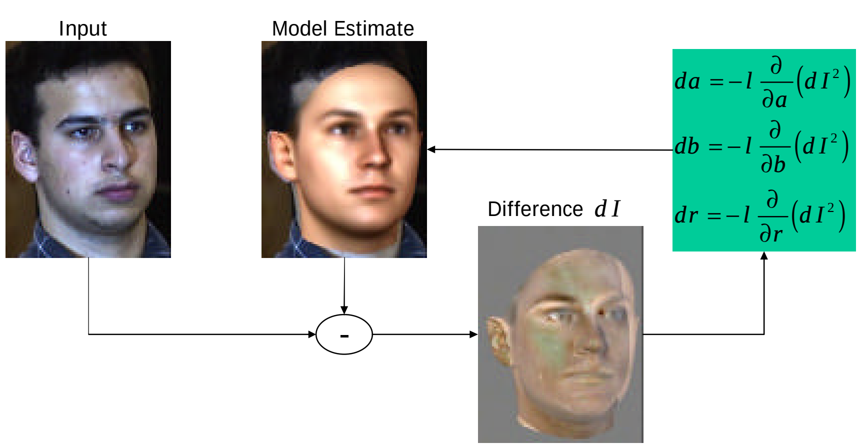
4. Photo-realistic images rendered using Computer Graphics

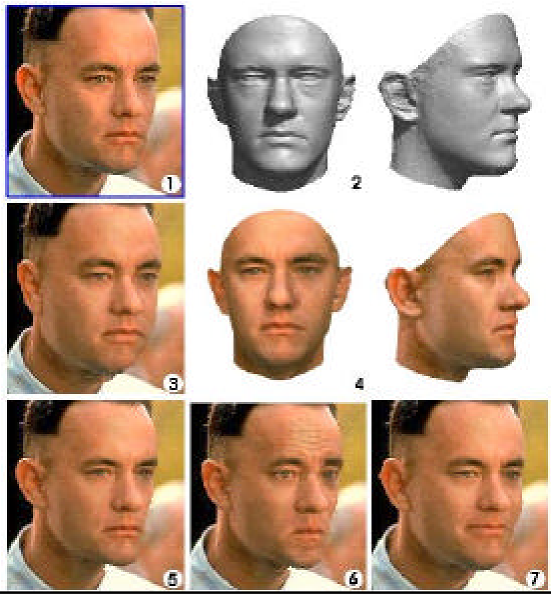


Model Fitting : Definition



Model Fitting: Standard Optimization Techniques




1
2
3
4
5
6
7

Materials

- Materials from this lecture were taken from the papers by the authors listed previously, and ppt slides by
  - Sami Romdhani Volker Blanz Thomas Vetter
  - Gustavo Halperin, Avishay Sidlesky