Course Notes for SIGGRAPH 2016

Capturing the Human Body: From VR, Consumer, to Health Applications

Hao Li      Lingyu Wei      Anshuman Das      Tristan Swedish      Pratik Shah      Ramesh Raskar
Capturing the Human Body: From VR, Consumer, to Health Applications

Abstract: Modeling the human body is of special interest in computer graphics to create “virtual humans”, but material and optical properties of biological tissues are complex and not easily captured. This course will cover the major topics and challenges in using image acquisition to model the human body.

Takeaways (what's in it for the reader): This course provides an overview of human body capture methodologies. Attendees will receive an overview of the bio-physics which create the variability in appearance of the eye, ear, skin, mouth, and hair among individuals. Instructors will then present state of the art methods used to create more accurate models by incorporating tools used in biomedical imaging. Such techniques use physical measurement to produce visually accurate human anatomy. Finally, we will explore health applications of image acquisition and modeling.

Course Schedule

8:30 am -- 9:00 am  Introduction to Body Shape and Human Performances [Li, Wei]
9:00 am -- 9:30 am  Biomedical Imaging and Human Image Capture [Das, Swedish, Shah]
9:30 am -- 9:50 am  Visual Computing in Health Technologies [Raskar]
9:50 am -- 9:55 am  Conclusion and Q&A session [All]
10:00 am  Close
Speaker's Bio

**Hao Li, Assistant Professor, University of Southern California**

Hao Li joined the University of Southern California in 2013 as a tenure-track assistant professor of computer science. Before his faculty appointment he was a research lead at Industrial Light & Magic, where he developed the next generation real-time performance capture technologies for Star Wars Episode VII. Prior to joining the force, Hao spent a year as a postdoctoral researcher at Columbia and Princeton Universities. His research lies in geometry processing, 3D reconstruction, and performance capture. While primarily developed to improve real-time digital content creation in film production, his work on markerless dynamic shape reconstruction has also impacted the field of human shape analysis and biomedicine. His algorithms are widely deployed in the industry, ranging from leading visual effects studios to manufacturers of state-of-the-art radiation therapy systems. He has been named top 35 innovator under 35 by MIT Technology Review in 2013 and NextGen 10: Innovators under 40 by CSQ in 2014. He was also awarded the Google Faculty Award in 2015, the SNF Fellowship for prospective researchers in 2011, and best paper award at SC 2009. He obtained his PhD from ETH Zurich in 2010 and received his MSc degree in Computer Science in 2006 from the University of Karlsruhe (TH). He was a visiting professor at Weta Digital in 2014 and visiting researcher at EPFL in 2010, Industrial Light & Magic (Lucasfilm) in 2009, Stanford University in 2008, National University of Singapore in 2006, and ENSIMAG in 2003.

**Anshuman Das, Postdoctoral Associate, MIT**

Anshuman Das is a postdoctoral associate at MIT and the Tata Center for Technology and Design. Anshuman is interested in creating rapid diagnostics that are smart, predictive, and accessible and will improve the way diagnostics are carried out. Within the health diagnostics field he is exploring intersections with health diagnostics and optics, lasers, UV-VIS, soft x-ray, Raman, and terahertz spectroscopy. He is also interested in super-resolution optical imaging and soft matter based optical elements. He is currently working on electrical and optical sensing of infections, wide-angle endoscopy and designing smart otoscopes. Before coming to MIT Anshuman received his Ph.D. from JNCASR in India where he researched on light management, degradation, and electrode design in organic solar cells.

**Tristan Swedish, Technical Assistant, MIT**

Tristan Swedish is a Technical Assistant at MIT. He received his BS in Electrical Engineering and Physics at Northeastern University, where he created computational models of light propagation in lung tissue and worked on an optical device to measure the biomechanics of the cornea. Tristan has also worked at BBN Technologies on a project to detect signals in non-stationary environments and more efficient solutions to inverse problems in shock wave propagation. At the MIT Media Lab, Tristan is building new types of imaging devices for retinal and skin diagnostics.
Pratik Shah, Research Scientist, MIT
Pratik, a research scientist in the Camera Culture group at the MIT Media lab, works at the intersection of nanotechnology, imaging, low cost diagnostics, entrepreneurship and scalable solutions for improving human health. Pratik has experience in vaccine design and discovery, applying throughout OMICS, nanotechnology and nucleic acid sequencing for biomedical research and drug discovery, microbial signaling systems, start-up and non-profit ventures. He also works on clinical images, with graphical interfaces, to isolate disease features and develop neural nets, which can automatically label and overlay high-dimensional medical images. Pratik has a BS, MS and a Ph.D in microbiology and completed fellowship training at the Broad Institute, Massachusetts General Hospital and Harvard Medical School.

Ramesh Raskar, Associate Professor, MIT
Ramesh Raskar joined the Media Lab from Mitsubishi Electric Research Laboratories in 2008 as head of the Lab’s Camera Culture research group. His research interests span the fields of computational photography, inverse problems in imaging and human-computer interaction. Recent projects and inventions include transient imaging to look around a corner, a next generation CAT-Scan machine, imperceptible markers for motion capture (Prakash), long distance barcodes (Bokode), touch+hover 3D interaction displays (BiDi screen), low-cost eye care devices (Netra,Catra), new theoretical models to augment light fields (ALF) to represent wave phenomena and algebraic rank constraints for 3D displays(HR3D). In 2004, Raskar received the TR100 Award from Technology Review, which recognizes top young innovators under the age of 35, and in 2003, the Global Indus Technovator Award, instituted at MIT to recognize the top 20 Indian technology innovators worldwide. In 2009, he was awarded a Sloan Research Fellowship. In 2010, he received the Darpa Young Faculty award. Other awards include Marr Prize honorable mention 2009, LAUNCH Health Innovation Award, presented by NASA, USAID, US State Dept and NIKE, 2010, Vodafone Wireless Innovation Project Award (first place), 2011. He holds over 40 US patents and has received four Mitsubishi Electric Invention Awards. He is currently co-authoring a book on Computational Photography.
Modeling and Capturing the Human Body: for rendering, health and visualization

Hao Li
USC

Anshuman Das
MIT

Tristan Swedish
MIT

Pratik Shah
MIT

Ramesh Raskar
MIT

Figure 1: This course offers an overview of modeling and capturing methodologies that have applications in rendering pipelines and health. We provide a survey of state of the art and emerging capturing modalities (left) in which the data produced can be transformed to visualize health, form, and performance (right).

Abstract

Modeling the human body is of special interest in computer graphics to create “virtual humans”, but material and optical properties of biological tissues are complex and not easily captured. This course will cover the major topics and challenges in using image acquisition to model the human body. Attendees will receive an overview of the bio-physics which create the variability in appearance of the eye, ear, skin, mouth, and hair among individuals. Instructors will then present state of the art methods used to create more accurate models by incorporating tools used in biomedical imaging. Such techniques use physical measurement to produce visually accurate human anatomy. Finally, we will explore health applications of image acquisition and modeling.

Module I: Introduction to Human Body Dynamics and Visual Appearance

In this module we will discuss the motivations for reproducing human appearance using computer graphics. We will introduce visually distinct anatomical features (eye, ear, skin, mouth and hair) their state of the art reproductions and relevance in health assessment. We will cover the biological reasons for dynamic appearance of humans such as blood perfusion in skin, breathing rates, and perspiration.

Module II: Biomedical Imaging and Human Image Capture

Image capture of real scenes is integral to creating visually believable models. We provide an overview of capturing techniques in the literature. Examples of the capabilities of biomedical imaging traditionally used in clinical settings will be provided. We will then examine how measurements made using biomedical devices are used for diagnosis and the shared problem domain of human appearance capture and health assessment.

Module III: Rendering the Human Body

This module explores how data captured from images can be incorporated in graphics rendering pipelines. Shape from image, light-tissue interaction physics and light transport models will be discussed.

Module IV: From Models to Health

Techniques developed to model the human body have applications in health. We will examine methods developed in both computer graphics and biomedical imaging communities to solve problems in cancer detection and heart monitoring.

Swedish, T. et al. 2015; ACM Trans. Graph. 32, 4.
<table>
<thead>
<tr>
<th>Part 1</th>
<th>Introduction to Human Body Shapes and Performances</th>
<th>Hao Li and Lingyu Wei</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 2</td>
<td>Biomedical Imaging and Human Image Capture 1</td>
<td>Tristan Swedish</td>
</tr>
<tr>
<td>Part 3</td>
<td>Biomedical Imaging and Human Image Capture 1</td>
<td>Anshuman Das</td>
</tr>
<tr>
<td>Part 4</td>
<td>Biomedical Imaging and Human Image Capture 1</td>
<td>Pratik Shah</td>
</tr>
<tr>
<td>Part 5</td>
<td>Visual Computing in Health Technologies</td>
<td>Ramesh Raskar</td>
</tr>
</tbody>
</table>

Closing Comments
INTRODUCTION TO HUMAN BODY SHAPES AND PERFORMANCES

Hao Li and Lingyu Wei
MIT Media Lab
3D Scanning
Realtime 3D Scanning
Geometric Data
Correspondences
Non-Rigid Registration

- Target
- Source
- Correspondences
- Overlap
Understanding Shapes

model

data
Other Applications

- entertainment
- fitness
- digital garment
3D Self-Portraits

Shapify.me
Pipeline
Overview

- Non-rigid registration
- Merging
- Texturing
- 3D print
10X Speed Up
Capture/Process Time: 4 mins
Faces
Realtime 3D Scanning

Wise et al. 2009


\[ \mathbf{v}_i(x) = \mathbf{v}_i(0) + \sum_l \mathbf{v}_i^{(l)} x_l \]

\[ x_l \in [0, 1] \]
Pipeline Overview

\[ \tilde{v}_i(\Delta v_i) = v_i + \Delta v_i \]
Tracking Basic Emotions

input data

anger  
surprise  
joy  
sadness  
disgust  
fear

our method

Weise et al. 2011
Facial Performance Capture

Li et al. SIGGRAPH 2013
Fast Calibration

Li et al. SIGGRAPH 2013
Capturing Hair

Hu et al. SIGGRAPH 2014
- Detect local dominant orientation using rotated filters [Paris04]
3D Orientation Field

Luo et al. SIGGRAPH 2013
Pipeline

**Input images**

Reconstruction

Point cloud & orientation field

Covering

Ribbons

Connection & direction analysis

Wisps

Synthesis

Synthesized strands

Luo et al. SIGGRAPH 2013
Local Minima

Implausible structures in the output [Luo et al. 2013]
Reconstruction with Simulated Examples

Hu et al. SIGGRAPH 2014
Reconstruction with Simulated Examples

Simulated example

Reference photo

Our result

Hu et al. SIGGRAPH 2014
Reconstruction with Simulated Examples

Simulated example

Reference photo

Our result
Reconstruction Priors

appearance

geometry

physics
Kinect Fusion [Newcombe et al. 2011]
Braid Capture

Orientation field extraction
Braid Capture

Input mesh → 3D orientation field → Cleaned mesh
Structure Analysis

Cleaned mesh
Patch fitting
Labeling
Structure extraction
Braid Theory [Artin 1947]

4-strand basic braid

4-strand fishtail
Procedural Modeling

Basic braid

Four-strand braid

Five-strand Dutch braid

Fishtail braid
Five-strand Dutch braid
Braid Capture
Impacting Science
Cardiology
IMAGE CAPTURE FOR VIRTUAL REALITY AND INTERACTION

Tristan Swedish
MIT Media Lab
Alignment Displays and Imaging for Interaction and Health
Eye box trade offs

AR
VR
BIOMETRICS
eyeSelfie: solve the user guidance problem

Directed rays as perceptual cues

NETRA

Pamplona et al, Siggraph, 2010
Directed rays as perceptual cues

CATRA

Pamplona et al, Siggraph, 2011
Retinal Alignment Challenge
Retinal imaging challenge: field of view
Retinal imaging challenge: reflections
Eye Alignment Displays
Eye box: near eye pinhole projector
Eye box: near eye pinhole projector
\[ \omega_e(z) = P_D - 2|z - z_0| u_\alpha \]
Pupil forming display

Display

Eye

view
Pupil forming display

Display
Projector
Pinholes

Eye

view
Pupil forming display

Display

Projector

Pinholes

Eye

view
Pupil forming display

Display

Projector
Pinholes

Eye

view
Pupil forming display

Display
Projector
Pinholes

Eye

view
Pupil forming display

Display

Eye

view
Light Field Eye Boxes
Defining eye box of heterogeneous pupil images
Defining eye box of heterogeneous pupil images
Defining eye box of heterogeneous pupil images
Defining eye box of heterogeneous pupil images
Defining eye box of heterogeneous pupil images
Light Field notation
Light Field: essential description
Light Field: essential description

\[ S_d = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \quad L_f = \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix} \]
Light Field: simple eye model

\[
\begin{bmatrix}
    x'' \\
    u''
\end{bmatrix} =
\begin{bmatrix}
    1 & d \\
    0 & 1
\end{bmatrix}
\begin{bmatrix}
    1 & 0 \\
    -1/f & 1
\end{bmatrix}
\begin{bmatrix}
    x' \\
    u'
\end{bmatrix}
\]

“Fourier Transform”:
90° rotation
Light Field: simple eye model

\[
\begin{bmatrix}
x'' \\
u''
\end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/f & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x' \\ u' \end{bmatrix}
\]

“Fourier Transform”: 90° rotation
Accommodation vs. alignment display

Accommodation:
Same scene but different views across pupil

Alignment:
Different view for each pupil position
Alignment Display: pupil size independence

Angle unique across pupil: Defocused point source

Aligned
Alignment Display: pupil size independence

- Lateral Misalignment
- Pupil Shift
- Vignetting
- Display Mask
- Display Illumination
- Pupil
- Perceived 1D Image
Corneal Reflection Channels
Corneal Reflection: illumination
Corneal Reflection: light transport channel

Reflected at Image Plane

\[
\begin{bmatrix}
  x'' \\
u''
\end{bmatrix} = \begin{bmatrix}
f_L \left( -\frac{x}{f_c} + u \right) \\
x \left( \frac{1}{f_c} - \frac{2}{f_L} \right) + u \left( 1 + f_L \right)
\end{bmatrix} u
\]

Objective Lens

\[-\frac{D}{2f_L} + \frac{x}{f_c} < u < \frac{D}{2f_L} + \frac{x}{f_c}\]

Image Plane

\[-\frac{D}{2f_L} - x \left( \frac{1}{f_L} - \frac{1}{f_c} \right) < u < \frac{D}{2f_L} - x \left( \frac{1}{f_L} - \frac{1}{f_c} \right)\]
Corneal Reflection: light transport
**Corneal Reflection: light transport**

\[
\begin{align*}
    u_{\text{max}} &= \frac{f_c f_L (D_L/2)}{r_e^2 + f_c f_L - r_e (2f_c + f_L)} \\
    \chi' &= -\frac{u_{\text{max}}}{(f_c + f_L - r_e)}
\end{align*}
\]

Reflection size and pupil or lens occlusion mapped to eye relief.
Retinal Imaging: “Inverse VR”
Eye box: near eye pinhole projector
Retinal imaging optics

Illumination Source

Camera

Retinal Image

Polarizers
Image capture user experience

Optical design of prototype

Double ray cone display design

\[
l_c = \begin{pmatrix} \frac{P_D}{2} & \frac{P_D}{2} & -\frac{P_D}{2} & -\frac{P_D}{2} \\ -u_1 & u_2 & u_1 & -u_2 \end{pmatrix}
\]

\[T = S_1L_1S_2L_2...S_nL_n\]

\[l_d = T^{-1}l_c\]

Validation of alignment accuracy

Repeat alignment results

Repeat alignment results: HDR Application
Future Applications: light efficient displays

AR

VR
Future Applications: predictive health
Future Applications: predictive health
eyeSelfie
Summary

tinyurl.com/eyeSelfie

eyeSelfie enables accurate self-alignment to the eye

Various near eye displays evaluated in terms of alignment

We demonstrate this alignment through retinal imaging

Repeatable alignment useful for VR/AR and Predictive Health
BIOMEDICAL IMAGING AND HUMAN IMAGE CAPTURE

Anshuman Das
MIT Media Lab
3D Imaging

- Time of flight
- Lightfield
- Single-pixel
- Structured light/speckle

Faces, hair, furniture, art… macro objects

3D Imaging ...
-MICRO-IMAGING,
-LOOKING WITHIN THE HUMAN BODY?
-IMAGING SMALL OBJECTS
Challenges: From macro to micro...

- Depth resolution
- Optical design: very limited space for triangulation schemes to work
- Depth resolution of different 3D imaging devices
- Laser line scan: 1 micron depth (expensive, bulky, time consuming)
- Kinect: mm to cm (low resolution for imaging small objects)
- Phase-shifting method (sub mm resolution, easy to implement but multi-shot process)
Triangulation and depth estimation

\[ \frac{Z}{L-Z} = \frac{d}{B}, \quad \text{or} \quad Z = \frac{L-Z}{B}d. \]

Simplifying the relationship leads to

\[ Z \approx \frac{L}{R}d \propto \frac{L}{R} (\phi - \phi_0). \]

Phase estimation by phase shifting method

A set of 3 or 5 phase shifted images are sequentially projected, Captured images can be modeled as,

\[
I_n(x, y) = a(x, y) + b(x, y) \cos \left( \phi(x, y) + \omega_x x + \omega_y y + \frac{\pi}{2} n \right)
\]

\[
\begin{align*}
I_1(x, y) &= I_0(x, y) + I_{mod}(x, y) \cos(\phi(x, y) - \theta), \\
I_2(x, y) &= I_0(x, y) + I_{mod}(x, y) \cos(\phi(x, y)), \\
I_3(x, y) &= I_0(x, y) + I_{mod}(x, y) \cos(\phi(x, y) + \theta),
\end{align*}
\]

\[
\phi' = \arctan \left[ \sqrt{3} \frac{I_1(x, y) - I_3(x, y)}{2I_2(x, y) - I_1(x, y) - I_3(x, y)} \right].
\]

\[
\phi(x, y) = \phi'(x, y) + 2k\pi,
\]

Process flow

Fringe Projection Sequence

Capture each pattern

Threshold and segment images

Phase unwrapping

Conversion from phase to height

Case study of middle ear imaging
Pressure changes in the middle ear

- Standard procedure to diagnose is pneumatic otoscopy and tympanometry
- Huge problem in children (>3 million cases in US per year)
Ear drum in detail

Volandri et. al, Journal of Biomechanics 44 (2011) 1219–1236
Imaging device: The otoscope

- 2D Imaging with a speculum and simple lens arrangement
- Subtle pressure changes cause the tympanic membrane to bulge or retract
- In cases where this depth cannot be imaged, 3D information may be useful
Recent developments in otoscopy

- New methods like Fluorescence otoscopy for cholesteotoma detection
- Cellscope oto

Tulio A. Valdez, et. al, *Analytical Chemistry* 2014 86 (20), 10454-10460
Optical Coherence Tomography

Cac T. Nguyen, et. al, Noninvasive in vivo optical detection of biofilm in the human middle ear, PNAS 2012 109 (24) 9529-9534
Recent advance: Light field Otoscope

Our approach: otoscope with structured light

- Optical system specs:
  - Focal plane about 1cm from tip of speculum
  - 1080p webcam
  - Telephoto lens with aperture control
  - DLP Projector
  - Front surface mirror
  - Otoscope head

Fringe projection based 3D imaging

Depth resolution ~ 25 microns!

Calibration of our device with spacers of known depths
3D Imaging of Tympanic Membrane Phantom

Image of TM captured by our device
Results: Ear phantom

Pressure in the phantom was changed with a syringe attached to the TM

Positive pressure

Negative pressure
Results: In vivo imaging in human subject

Things to keep in mind...

- Suppressing motion related noise in fringe projection
- Faster image acquisition for real time processing
- Quantification of the depth map
- Clinical tests to determine range of TM depths that are normal and classify TM into healthy and unhealthy categories
- Machine learning algorithms to carry out automated diagnosis
- Global-direct separation could be used to improve fringe contrast in cases of diffuse surfaces like the tympanic membrane
Thanks to...

- Dr. Julio Estrada, CIO, Mexico
- Dr. Ayesha Khalid and Dr. Ellen Weinberg at Cambridge Health Alliance, Harvard Medical School, Cambridge
- MIT Tata Center for Technology + Design
In the future...

• Look at throat imaging, vocal chords, nose-implications in sleep quality
• 3D endoscopy: TOF, structured light: pushing boundaries
• Reconstructing internal organs will help in surgery
• Training in surgery, simulations of internal organs
ORAL IMAGING, RENDERING, DIAGNOSIS

Pratik Shah
USC
Imaging of the oral cavity

- How can we detect shape and color of teeth?
- How can we monitor the health of teeth and gums?
- How can we get X-ray like images of teeth with non-ionizing radiations?
- How can we use biomarker imaging and rendering to improve clinical medicine?
Translucent dentin (Caries)

Porphyrin/ Color magnification /Loss of fluorescence (Plaque/Caries/Gingivitis)

Collagen/ NADH (Cancer)
Visible Spectrum
Quantitative Light Induced Fluorescence (QLF™)

Image source:
Inspektor Research Systems
http://www.inspektor.nl/
Infrared Spectrum

Infrared Spectrum

Infrared Spectrum

Caries detection using near-infrared imaging

KaVo’s DIAGNOcam

Image source:
http://www.kavo.com/Products/Diagnostics/DIAGNOcam.aspx
DIAGNOcam clinical examples

Color image  NIR  X-ray  Color image after grinding

Image source:
Clinical Cases of DIAGNOcam (created by the Ludwig-Maximilian University Munich, Department of Conservative Dentistry, 2012)
http://www.diagnocam.com/
Ionizing Radiation

3D Axial Slices
3D Scanning of teeth

Projects a light stripe pattern

Confocal laser scanning

doi:10.1371/journal.pone.0043312.g001
doi:10.1371/journal.pone.0043312.g002

3D scanned teeth

doi:10.1371/journal.pone.0043312.g005

Measurement of tooth color

VITA Bleachedguide 3D-MASTER®

Image source:

VITA Easyshade®

Detect features in oral images

Correlate with conditions

IMAGES → PROCESSING → PREDICTIONS