Appearance Capture and Modeling of Human Teeth

ZDRAVKO VELINOV, Disney Research, Rheinische Friedrich-Wilhelms-Universität Bonn
MARIOS PAPAS, Disney Research
DEREK BRADLEY, Disney Research
PAULO GOTARDO, Disney Research
PARSA MIRDEHGHAN, Disney Research
STEVE MARSCHNER, Cornell University
JAN NOVÁK, Disney Research
THABO BEELER, Disney Research

Fig. 1. Various comparisons of our recovered teeth appearance parameters rendered next to corresponding reference images, from different viewpoints and lighting conditions. Reference images are on the left for individual teeth comparisons, and above for full teeth rows.

Recreating the appearance of humans in virtual environments for the purpose of movie, video game, or other types of production involves the acquisition of a geometric representation of the human body and its scattering parameters which express the interaction between the geometry and light propagated throughout the scene. Teeth appearance is defined not only by the light and surface interaction, but also by its internal geometry and the intra-oral environment, posing its own unique set of challenges. Therefore, we present a system specifically designed for capturing the optical properties of live human teeth such that they can be realistically re-rendered in computer graphics. We acquire our data in vivo in a conventional multiple camera and light source setup and use exact geometry segmented from intra-oral scans. To simulate the complex interaction of light in the oral cavity during inverse rendering we employ a novel pipeline based on derivative path tracing with respect to both optical properties and geometry of the inner dentin surface. The resulting estimates of the global derivatives are used to extract parameters in a joint numerical optimization. The final appearance faithfully recreates the acquired data and can be directly used in conventional path tracing frameworks for rendering virtual humans.

ACM Reference Format:
Zdravko Velinov, Marios Papas, Derek Bradley, Paulo Gotardo, Parsa Mirdehghan, Steve Marschner, Jan Novák, and Thabo Beeler. 2018. Appearance Capture and Modeling of Human Teeth. *ACM Trans. Graph.* 37, 6, Article 207 (November 2018), 13 pages. https://doi.org/10.1145/3272127.3275098

1 INTRODUCTION

A beautiful smile is often considered key to success and the key to a beautiful smile is a nice set of teeth. The ideal teeth appearance depends on the social and cultural context of the subject, as well as on their age since teeth appearance changes over time. In addition to the surface properties of teeth, the appearance also heavily depends on the geometric and optical properties of internal organic structures. These can be influenced by many factors, e.g. the recurrent use of substances like caffeine or tobacco, neglected teeth hygiene, illness, in particular at a young age, or a dysfunction or complete absence of the tooth nerve. All of these can lead to a gradual change of appearance and contribute to the unique look of each single tooth. Conversely, the appearance of teeth may disclose the habits, age, or socio-economic background of their owner.

When creating CG characters, for example for visual effects in motion pictures, it is therefore essential that their teeth match their background story. Similarly, when creating digital doubles of real actors, it is paramount to faithfully reproduce their teeth appearance. While there has been substantial work, in particular in the medical dental field, on capturing the shape and positioning of teeth in vivo, acquisition of their optical properties has been less actively

Authors' addresses: Zdravko Velinov, Disney Research, Rheinische Friedrich-Wilhelms-Universität Bonn, velinov@cs.uni-bonn.de; Marios Papas, Disney Research, marios.papas@disneyresearch.com; Derek Bradley, Disney Research, derek.bradley@disneyresearch.com; Paulo Gotardo, Disney Research, paulo.gotardo@disneyresearch.com; Parsa Mirdehghan, Disney Research, parsa@cs.toronto.edu; Steve Marschner, Cornell University, srm@cs.cornell.edu; Jan Novák, Disney Research, jan.novak@disneyresearch.com; Thabo Beeler, Disney Research, thabo.beeler@disneyresearch.com.

0730-0301/2018/11-ART207 $15.00
https://doi.org/10.1145/3272127.3275098

Fig. 2. **Dental scanning.** Current dental scanners extract a diffuse texture. The diffuse albedo is incorrectly darker in the areas where the tooth is optically thinner, and glossy highlights are also incorrectly encoded, leading to brighter appearance than expected.

pursued. While the most recent intra-oral scanners do capture color, they offer only very basic appearance acquisition and do not account for the complex optical properties of teeth, ending up with an overall unnatural look, as illustrated by the dental scan using a *3Shape TRIOS*® device shown in Fig. 2. Previous works dedicated to capturing the reflectance of teeth limited themselves to simple models [Buchaillard et al. 2007] or derived synthetic models which partially approximate light transport [Larsen et al. 2012], in both cases delivering lower visual quality compared to a more complete light transport simulation.

The scattering properties of a tooth are determined by its two outermost layers composed of materials with very different optical properties (Fig. 3). The inner core is made out of dentin, which has the biggest influence on its chroma and is also responsible for the fluorescent nature of teeth. A layer of enamel with varying thickness covers the dentin. It is typically thinner at the gums and thicker towards the tip, causing a chroma gradient. Enamel is a structurally dense material, but thinner than dentin. It has a high degree of translucency, which depends substantially on its mineralization and hydration. At the microscopic level, enamel is formed by rods that run from the dentin towards the surface and cause strongly anisotropic light propagation. Translucent enamel displays the characteristics of opalescence, causing the tooth to backscatter in the blue spectra and hence appear red-yellowish when lit from the other side. Finally, surface variations, cracks, stains and other imperfections add yet another layer of visual richness to teeth appearance.

This work presents a step towards overcoming the limitations of present acquisition systems by capturing the optical properties of teeth such that they can be realistically re-rendered in virtual environments. Our work complements existing shape reconstruction methods, such as intra-oral scanning or extra-oral fitting. In fact we use geometry acquired through an intra-oral scan and segment it according to Wu et al. [2016]. We acquire the appearance data by imaging a person's mouth area from multiple views and under multiple lighting conditions.

To model the complex interaction of light within the oral cavity and tooth volume, we use a modified path tracing framework capable of estimating derivatives with respect to optimized parameters while performing regular rendering. Essentially, our method iteratively solves for the derivatives of the entire radiative transfer integral over the entire visible scene. Furthermore, we can perform optimization of not just scattering parameters (albedo, density, surface texture) but also geometric parameters (dentin surface shape). We model dentin shape as a set of deformations from the base enamel shape,

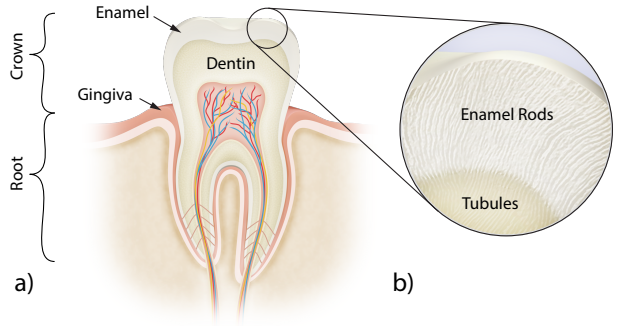

Fig. 3. **Anatomy.** a) A tooth is made of two different materials with very different optical properties. The inner core is formed by dentin, which has the biggest influence on the tooths chroma, and is covered by a layer of optically denser enamel. Finally, the tooth is bounded to the sides by gingiva tissue (gums), which scatters skin tinted light into the tooth. b) At a microscopic level, enamel is formed by rods that run from the dentin towards the surface and cause strong anisotropic light propagation.

which we parametrically control in the optimization loop. Thus we are able to predict the approximate shape and contribution of the hidden dentin surface on the appearance. Our approach represents a step forward in visual fidelity compared to previous works on tooth appearance capture. To the best of our knowledge, the proposed use of morphable shapes for representing hidden structures has not been explored by previous methods and represents a major contribution to the field of differentiable path tracing. We believe that our framework is general enough to be successfully employed in other applications, such as appearance acquisition of organic tissues (skin, fruits, plants), translucent gemstones, and other man-made objects.

We validate our technique by capturing a dead tooth extracted from a mouth cavity, where the ground truth inner dentin and outer enamel geometry is available. We further apply our method on in-vivo data acquired in an appearance capture session of human volunteers (Fig. 1). In both cases our approach closely recreates the overall appearance. Since the main framework employed within the optimization loop is based on path tracing, the resulting set of parameters is readily available for re-rendering the teeth with very minor modification within production environments.

2 RELATED WORK

In the following we discuss other methods that capture properties of teeth in the context of computer graphics and beyond. Our appearance capture approach is also related to methods for reconstructing the appearance of real skin, in particular faces. Since the ultimate goal is to re-render the captured teeth in CG, we outline relevant work in volumetric rendering, particularly those dedicated to teeth. Finally, as inverse rendering is a key ingredient to our optimization pipeline, we end by discussing how this concept has been explored in other contexts.

2.1 Teeth Capture

Reconstructing the 3D shape of teeth is important in the field of dentistry, in particular for diagnosing medical problems or planning restoration procedures. Common approaches include plaster casting or CT scanning [Omachi et al. 2007; Yanagisawa et al. 2014]. More lightweight intra-oral scanners have also been developed (e.g. *3M™ True Definition*, *iTero*, *3Shape TRIOS®*, [Ahn et al. 2017]), which typically use structured light to recover teeth shape and, in some instances, a basic surface color. Detailed survey of these devices from dental practitioner standpoint is provided by Mangano et al. [2017]. In the areas of computer graphics and vision, photometric reconstruction methods such as shape-from-shading [Carter et al. 2010] and feature-based reconstruction [Zheng et al. 2011] have been proposed. Because teeth generally have few identifiable features, model-based fitting approaches have gained popularity [Buchaillard et al. 2007; Farag et al. 2013; Mehl and Blanz 2005; Mostafa et al. 2014]. The current state-of-the-art in teeth shape capture is the recent model-based method of Wu et al. [2016], who build a statistical model of teeth shape from dental scans and then fit the model to silhouette features, presenting high-quality results for teeth reconstruction from passive imagery. Most shape reconstruction papers limit their appearance models to color extraction from images, as it is the case in Wu et al. [2016] and Kawai et al. [2014]. They are thus severely limited in their ability to faithfully re-render teeth from novel viewpoints or in different lighting conditions. Our method for detailed appearance capture of teeth would naturally complement these techniques for teeth shape reconstruction.

In medical dental literature, several experiments have measured appearance properties of teeth, in particular the scattering and absorption values of both dentin and enamel [Fried et al. 1995; Pop-Ciutrila et al. 2015; Spitzer and Ten Bosch 1975]. These studies are motivated by different applications in dentistry, for example the creation of false teeth substitutes that are identical to real teeth [Pop-Ciutrila et al. 2015], or the study of how lasers designed for soft tissue surgical procedures interact with hard dental tissue [Fried et al. 1995]. In all cases the measurement process is a destructive one, where individual teeth are sliced into thin layers for measurement. In our work we design a method to measure the appearance properties of teeth in vivo, with the goal of re-rendering for computer graphics. The optical properties reported by dentistry literature, in particular the average measured absorption and scattering properties of enamel [Spitzer and Ten Bosch 1975] and dentin [Pop-Ciutrila et al. 2015], are employed as initial values in our optimization procedure, described in Section 6.

To the best of our knowledge, the only existing specialized model for rendering teeth is presented by Larsen et al. [2012] which makes several compromises in order to enable real-time performance. In contrast, we are interested in higher fidelity reproduction of the real teeth appearance for the application of high-quality digital humans.

2.2 Facial Appearance Capture

Teeth appearance capture can be considered part of the larger field of facial appearance modeling and capture, which has received a lot of attention over the past decades. An in-depth overview can be found in the state-of-the-art report of Klehm et al. [2015] and is beyond the scope of this paper. The relevant work in the area of facial appearance capture is concerned with the estimation of scattering properties of skin, which falls in the wider area of volumetric appearance capture discussed in the next section.

2.3 Volumetric Appearance Capture

A number of methods have been proposed for acquiring surface-based definitions of translucency (e.g. the methods by Debevec et al. [2000]; Donner et al. [2008]; Tong et al. [2005]). We take a different approach and explicitly model the volumetric optical properties of teeth, which falls into the wider area of volumetric appearance capture. The most common approach to estimate the absorption and reduced scattering coefficients is to inject light at a specific point and image the diffusion profile on the medium. Approaches employ single beams [Jensen et al. 2001; Weyrich et al. 2006], stripes [Tariq et al. 2006], and dots [Ghosh et al. 2008]. Wang et al. [2008] employed a polygrid diffusion algorithm for solving the diffusion equation. Munoz et al. [2011] proposed a pipeline for predicting the appearance of optically thick homogenous translucent objects based on the diffusion theory. Papas et al. [2013] measured a database of diffusion profiles of pigments used in a mixture estimation pipeline capable of reproducing the appearance of translucent materials. In contrast, Gkioulekas et al. [2013] invert the radiative transfer equation [Chandrasekhar 1960] and solve in operator-centric fashion for optical parameters using a combination of Monte Carlo rendering, stochastic gradient descent, and material dictionaries.

More generally our method builds upon the family of papers that compute analytic path contribution gradients which are numerically integrated on a sensor. Initially Hašan and Ramamoorthi [2013] provided analytic gradients of paths with respect to albedo to facilitate interactive albedo editing. This was later generalized to handle all Bidirectional Scattering Distribution Function (BSDF) and volume related properties [Khungurn et al. 2015] and more general lighting/viewing conditions [Gkioulekas et al. 2016] for the purpose of inverse rendering. Our paper extends this line of work by allowing the joint optimization of optical properties and geometry, which is essential for optimizing the dentin geometry in the case of teeth.

2.4 Rendering Translucent Materials

The concept of BSDF was introduced in physics to model transparent surfaces [Bartell et al. 1981; Dereniak et al. 1982]. In order to render materials with a significant degree of internal scattering, Stam [1995] proposed a grid-based method for spatially resolving the diffusion of light. To render effects due to subsurface scattering, Jensen et al. [2001] utilized the diffusion dipole, which was later extended to handle layered materials [Donner and Jensen 2005; Donner et al. 2008] and edges and corners [Donner and Jensen 2008]. It was demonstrated that the quality of diffusion-based approaches can be improved by quantization [d'Eon and Irving 2011] and combination with Monte Carlo simulations [Habel et al. 2013; Li et al. 2005]. Frisvad et al. [2014] proposed a directional dipole that better handles transport at oblique angles. Frederickx and Dutré [2017] introduced a forward scattering dipole model based on the theory of functional integrals for radiative transfer. Despite all of these improvements,

the poor handling of the boundary leads to inaccurate results when the object is thin and (locally) highly translucent, such as teeth.

Approaches based on Monte Carlo integration, such as volumetric (bidirectional) path tracing [Kajiya and Von Herzen 1984; Lafortune and Willems 1996], metropolis light transport [Pauly et al. 2000], or photon mapping [Donner and Jensen 2008; Jensen and Christensen 1998] are computationally more intensive, but provide high-accuracy results. Since we strive for visual fidelity and can afford render times on the order of minutes, we use volumetric path tracing during inverse rendering and to render our validation results.

3 METHOD OVERVIEW

Light takes many paths when interacting with teeth. First, it interacts with the surface, characterized by a rough dielectric layer, and is either reflected or refracted into the interior. The interactions in the interior are mainly defined by the two layers of a tooth: enamel and dentin. We thus identify three sets of parameters defining the appearance of teeth: surface reflectance and transmission, sub-surface scattering parameters, and inner dentin shape. Our approach aims to derive a plausible set of model parameters for the purpose of high-quality rendering in a scenario where teeth are illuminated externally with respect to the oral cavity. We split our discussion into a general definition of the appearance model in Section 4, our data acquisition methodology in Section 5, and the respective parameter estimation pipeline in Section 6 followed by qualitative evaluations of the proposed model and optimization on real-world measurements in Section 7.

4 TEETH APPEARANCE MODEL

The appearance of a tooth is highly influenced by the shape and optical properties of its enamel and dentin parts (Fig. 3). We represent both materials as bounded volumes and introduce a number of key simplifying assumptions to accelerate the inverse rendering.

4.1 Enamel

Boundary Geometry and BSDF. The enamel boundary is the outer boundary of the tooth, and is modeled as a triangle mesh. In order to model the interaction of light with the air-enamel boundary, we use a modified version of the smooth-dielectric BSDF, which adopts the concept of separating the air and enamel using a "thin" dielectric layer.

The BSDF consists of a reflective and a transmissive component, $f(\mathbf{x}, \omega_i, \omega_o) = f_r(\mathbf{x}, \omega_i, \omega_o) + f_t(\mathbf{x}, \omega_i, \omega_o)$, where ω_i and ω_o correspond to the incident and outgoing light directions.

The reflection component is defined in a standard way,

$$f_r(\mathbf{x}, \omega_i, \omega_o) = \frac{\delta(\omega_h, \mathbf{n})}{|\omega_i \cdot \mathbf{n}|} F(\omega_i \cdot \mathbf{n}, \eta), \tag{1}$$

where F is the Fresnel term that quantifies the amount of light reflected off the dielectric boundary (see e.g. [Pharr et al. 2016] for the definition), $\mathbf{x}$ is the spatial position, ω_h is the half-way vector between ω_i and ω_o, $\mathbf{n}$ is the surface normal, and $\delta(a, b)$ is a Dirac delta function that equals $+\infty$ when $a = b$ and 0 otherwise. We use a fixed refractive index from dentist literature $\eta = 1.631$ [Meng et al. 2009] for computing the Fresnel term.

The transmission component,

$$f_t(\mathbf{x}, \omega_i, \omega_o) = a_t(\mathbf{x}) \frac{\delta(\omega_i, -\omega_o)}{|\omega_i \cdot \mathbf{n}|} (1 - F(\omega_i \cdot \mathbf{n}, \eta)), \tag{2}$$

introduces the simplification that transmitted light does not bend, i.e. it continues along a straight line, which greatly simplifies the light-transport simulation. The spatially varying $a_t(\mathbf{x})$ modulates the absorption of the thin enamel/air boundary layer. During inverse rendering we store $a_t(\mathbf{x})$ in a texture which is indexed by a UV parametrization of the outer enamel mesh.

The BSDF can be interpreted as an infinitesimally thin dielectric layer, and, while physically inaccurate, the model is expressive enough to capture specular highlights and stains, which are essential to faithfully represent the appearance. Ignoring the bending of light rays during transmission greatly accelerates our Monte Carlo simulations during inverse rendering as direct illumination can be sampled using shadow rays that connect to lights through the boundary. We avoid highlights during optimization by masking them and reintroduce them in the final images by using a mixture of this BSDF representing the transmission term and a rough dielectric representing the reflection term according to the GGX distribution [Walter et al. 2007].

Volumetric Properties. Once transmitted through the boundary, light will start interacting with the inner material. To model the propagation through the volume, we rely on the radiative transfer equation [Chandrasekhar 1960],

$$(\omega \cdot \nabla L) = -\sigma L(\mathbf{x}, \omega) + \alpha\sigma \int f_p(\omega_i, \omega_o)\, L(\mathbf{x}, \omega_i)\, d\omega_i, \tag{3}$$

which describes the change of radiance L due to absorption and scattering, here quantified by the extinction coefficient σ and single-scattering albedo α. To reproduce the characteristic wavelength-dependent scattering and absorption of teeth, these quantities are RGB vector values.

In order to model the directional distribution of scattered light, we use a two-lobe phase function:

$$f_p(\omega_i, \omega_o) = \beta f_{\mathrm{HG}}(\omega_i \cdot \omega_o, g_1) + (1 - \beta) f_{\mathrm{HG}}(\omega_i \cdot \omega_o, g_2), \tag{4}$$

where β balances the influence of the two lobes and f_{HG} is the Henyey-Greenstein [1941] phase function parameterized by the mean cosine g of the scattering angle.

4.2 Dentin

Boundary Geometry and BSDF. The thickness of the enamel layer largely affects the amount of light reaching the dentin core of the tooth and has therefore significant impact on the overall appearance. We allow varying the thickness by modeling the enamel-dentin interface using a blend-shape model defined by a set of deformations:

$$\mathbf{V} = \mathbf{B}_0 + \mathbf{c} \cdot (\mathbf{B} - \mathbf{B}_0)^{\mathrm{T}}. \tag{5}$$

We denote the resulting shape $\mathbf{V}$, its individual vertices $\mathbf{v}_i$, the meshes representing the deformation by each mode $\mathbf{B}$, and the vector weights $\mathbf{c}$. We chose to represent the deformations with respect to an initial mesh $\mathbf{B}_0$ as it enables intuitive visualization of each mode. Through experimentation we came to the conclusion that three modes are enough to capture most of the visible changes in scattering introduced by the difference of enamel thickness. The

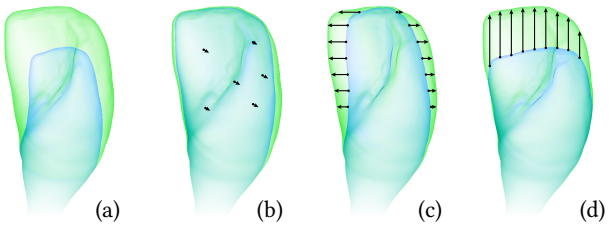

Fig. 4. **Dentin deformation modes.** The fully deformed (a) dentin blend shape of an incisor serves as a boundary defining enamel thickness. For optimization purposes it is further decomposed in three deformation modes along the major axes extracted according to the canonical tooth space such that they represent the following directions: (a) buccal, (b) proximal, (c) occlusal.

modes represent the deformation along the three major axes in a canonical tooth space; see the illustration for a single tooth in Fig. 4.

We use a highly simplified BSDF to model the interactions of light with the enamel-dentin interface, which assumes an *index-matched* boundary, i.e. light passes through the interface unaltered.

Volumetric Properties. The directional scattering is modeled analogously to enamel, i.e. using a two-lobe phase function.

4.3 Gingiva and oral cavity

Gingiva or gums and the oral cavity in general do have an indirect impact on the appearance of teeth since they scatter skin tinted light into the tooth. In order to explain that stray light and prevent the optimization from predicting a tinted tooth we have to model and include the gums and a substitute for the oral cavity walls.

The gums are modeled as a triangle mesh wrapping around the outer boundary of the root of the teeth. As we can directly observe the gums, we use the same set of parameters as the enamel to extract their appearance through numerical optimization with fixed refractive index $\eta = 1.33$.

The oral cavity is highly occluded and non-trivial to capture, however it has a great impact on the appearance as light bounces from its walls and illuminates areas in shadow. We use a substitute mesh with a diffuse reflectance set as free parameter. Its starting point is derived by using similarity theory and average values derived from previous optimization of the scattering parameters of gingiva.

5 DATA ACQUISITION

We fit our appearance model described in Section 4 to real teeth measurements captured for three volunteers. In this section we describe the data capture and the pre-processing and post-processing steps.

5.1 Image Capture

In order to obtain accurate appearance parameters, we require a dense sampling of incoming and outgoing light directions. For this reason, we capture each volunteer in a light dome [Beeler 2012] consisting of 155 lights (each one a RGB LED triplet), where each light is approximately one meter away from the teeth. Due to occlusions of the head, we record imagery for 54 of the lights, which accounts for all visible incoming directions. We capture multi-view imagery from three color cameras (Ximea CB200CG), positioned around the edge of the light dome with long lenses to maximize teeth resolution, recording 20 megapixel images, synchronized with individual light activations. Please refer to Fig. 5 (a) for an image of the light dome and example captured imagery.

Calibration. We geometrically calibrate the cameras using a planar fiducial pattern [Garrido-Jurado et al. 2014]. The positions of the lights are measured using a Leica DISTO^{TM} S910 laser distance measuring device, and then aligned to the camera coordinate system by imaging a mirror sphere. Color calibration is performed by recording the X-Rite ColorChecker$^{®}$ chart with all lights turned on to extract corrective terms similar to Ilie and Welch [2005]. The individual light intensities are measured by recording a diffuse spherical spectralon and considering it as ideal Lambertian surface. In order to account for spurious incoming light, a black (no light) image is recorded and subtracted from all measurements.

Volunteer Preparation. Recording the 54 light directions requires approximately 10-15 seconds, during which time the subject must remain still. To facilitate this, we provide a stabilizing neck brace, a head strap and a rigid object to bite on, which also helps to separate the upper and lower teeth. To minimize occlusions by the face, we retract the lips and cheeks using an OptraGate lip retractor from Ivoclar Vivadent, often used in dentistry.

Post-Processing. Despite stabilizing the volunteer's head during acquisition, there is still a minimal amount of residual motion left, which we remove as a postprocessing step by aligning all frames to the first image using optical flow [Brox et al. 2004]. Our imagery contains shadows cast by the lips and lip retractor, which will confuse the appearance optimization since we do not model the geometry of the face or the retractor. Additionally, we will not attempt to match specular highlights since we are more interested in the volumetric light transport. Therefore, we manually mask the imagery to remove shadows, the gums and inner mouth and employ the method of Mallick et al. [2005] to automatically detect and mask specular highlights, keeping only valid teeth pixels that will inform the optimization. These tasks do not form an integral part of our algorithm, and are easy to perform by non-experts. They could also be further automated for use in a production pipeline when the complete facial appearance and geometry is available.

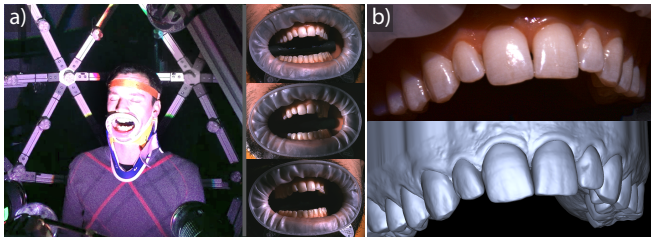

Fig. 5. **Data Acquisition**. (a) We capture the teeth from three viewpoints under 54 different light directions using a light dome. Here we show three different light directions for the frontal camera. (b) The teeth are scanned and we align the geometry to the imagery, in preparation for the appearance optimization.

5.2 Geometry Capture

The captured imagery alone is not sufficient to perform appearance optimization, as we require knowledge of the teeth geometry as well. While lightweight teeth reconstruction methods exist (e.g. Wu et al. [2016]), they fail to recover the accurate shape of the back side of the teeth and only hypothesize the teeth shape where image coverage is low. Since we need to evaluate volumetric light paths, we require accurate teeth boundaries on all sides. Therefore, we obtain a professional teeth scan for the volunteer using a *3Shape TRIOS®* intra-oral scanner. In order to segment the teeth in the scan and separate them from the gums we align template teeth meshes to the scan geometry using the semi-automatic template fitting approach described in Wu et al. [2016]. Since the appearance of teeth is tightly coupled with the gums, we also model the gums geometry artistically based off the scans. Finally, we align the resulting per-tooth geometry and gums model to the camera coordinate frame by manually selecting a sparse set of corresponding landmarks and solving for a rigid transformation (see Fig. 5 (b)).

The final result of the data acquisition is, per volunteer, a total of 162 images (54 individual light directions for three views) masked to identify valid teeth pixels, with precise teeth and gums geometry aligned to the imagery. Part of this dataset will be used in the model parameter optimization method and for validation.

6 MODEL PARAMETER OPTIMIZATION

We formulate the problem of extracting the parameters that describe the appearance of teeth as a non-linear optimization problem. Assuming we have M pixels in each of the rendered and objective images, and a vector $\mathbf{x}$ with P parameters to be optimized according to a loss function $\rho(\psi) : \mathbb{R}^M \rightarrow \mathbb{R}$ and a cost function $\psi(\mathbf{x}) : \mathbb{R}^P \rightarrow \mathbb{R}^M$, we formally define the set of optimal parameters

$$\mathbf{x}^+ = \arg\min_{\mathbf{x}} \rho(\psi(\mathbf{x})). \tag{6}$$

Since light transport in participating media is generally difficult to evaluate in closed form, the cost function cannot be computed analytically. Instead, we take the route of differentiable path tracing by computing analytic derivatives associated with each *light path* and accumulating them to compute approximate global derivatives to guide a nonlinear optimization algorithm, thus iteratively solving Eq. 6. Since the number of light paths is finite—we use 64 samples per pixel—we obtain a noisy estimate of the true gradient. As such, we take the first-order stochastic gradient descend (SGD) approach proposed by Kingma and Ba [2014], which was demonstrated to be sufficiently reliable for solving problems with such characteristics. A similar decision was made by Gkioulekas et al. [2016] who used AdaDelta [Zeiler 2012] in their Monte Carlo based SGD approach for inverse rendering of participating media.

The derivative for each pixel with respect to a free parameter is expressed starting from the loss function $\rho(\psi)$, then applying the change according to image-reconstruction filter F used when sampling the camera pixel footprint, and finally applying the individual gradients $\nabla\psi_{\mathbf{x}}$ according to the respective parameters,

$$\nabla\rho = \frac{\mathrm{d}\rho}{\mathrm{d}F}\frac{\mathrm{d}F}{\mathrm{d}\psi_{\mathbf{x}}}\nabla\psi_{\mathbf{x}}. \tag{7}$$

We accumulate the first order derivatives at each interaction with a medium, surface or shape containing free parameters separately by exploiting the generalized product rule. Furthermore, we assume a pinhole camera with box filtering function dependent only on the number of samples N and the L2 loss function, resulting in the final equation:

$$\begin{aligned}\nabla\rho &= \frac{1}{N}(C - C_{\mathrm{ref}}) \cdot \left(\prod_{i=0}^{K} \psi_{\mathbf{x}}^{(i)}\right)\left(\sum_{i=0}^{K} \frac{\nabla\psi_{\mathbf{x}}^{(i)}}{\psi_{\mathbf{x}}^{(i)}}\right) \\ &= \frac{1}{N}(C - C_{\mathrm{ref}}) \cdot \mathcal{T}\left(\sum_{i=0}^{K} \frac{\nabla\psi_{\mathbf{x}}^{(i)}}{\psi_{\mathbf{x}}^{(i)}}\right),\end{aligned} \tag{8}$$

where C and C_{ref} are vectors holding linearized rendered and reference RGB images, respectively, and N denotes the number of samples and K the maximum number of bounces. The complete throughput is expressed as a matrix $\mathcal{T}$, thus separating the specific change introduced by each interaction from the complete computation of light propagation.

Three groups of parameters are treated within our optimization pipeline: the dentin shape as well as the surface and subsurface parameters of dentin, enamel, and gums. We discuss them in their own separate sections and show them in Fig. 6. For the purpose of optimization we define the respective block of gradients

$$\nabla\psi_{\mathbf{x}} = \nabla\psi_B \parallel \nabla\psi_\sigma \parallel \nabla\psi_\alpha \parallel \nabla\psi_{f_p} \parallel \nabla\psi_f \tag{9}$$

as the concatenation of the gradients w.r.t. the deformable dentin shape model ($\nabla\psi_B$), volume extinction ($\nabla\psi_\sigma$), single-scattering albedo ($\nabla\psi_\alpha$), the phase function ($\nabla\psi_{f_p}$), and the BSDF surface model ($\nabla\psi_f$). The gradients of subsurface parameters ($\nabla\psi_\sigma$, $\nabla\psi_\alpha$, $\nabla\psi_{f_p}$) are further composed of separate groups of RGB components associated with dentin and enamel.

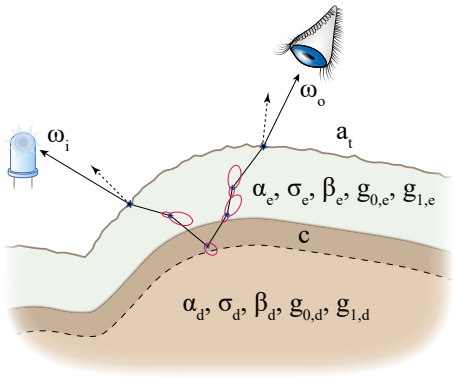

Fig. 6. **Model parameters.** We jointly optimize surface absorption (a_t), enamel scattering parameters (α_e, σ_e, β_e, $g_{0,\mathrm{e}}$, $g_{1,\mathrm{e}}$), optical and geometry parameters of dentin (α_d, σ_d, β_d, $g_{0,\mathrm{d}}$, $g_{1,\mathrm{d}}$, $\mathbf{c}$). We further create a relation between enamel and dentin parameters by smoothing the border between media which establishes correlation between scattering and geometry parameters.

6.1 Enamel and Gums

Enamel and gums are treated similarly by the optimization as they are both modeled to share the same types of optical properties such as a homogeneous volume and an absorption layer. For brevity we discuss the case of enamel. The gums are treated analogously. The only difference is that enamel encloses another internal volume (dentin) with free optical and geometrical properties.

Air-Enamel boundary. The BSDF model introduces a linear dependency on its transmissive absorption coefficient. Excluding the reflectance term, the derivative at each transmission is completely defined by the light transport equation, which leads to a very simple equation for the accumulated term:

$$\frac{\nabla\psi_f^{(i)}}{\psi_f^{(i)}} = \frac{1}{a_t}. \tag{10}$$

Phase function. We use a dual-phase function leading to three sets of parameters with the following accumulated term:

$$\frac{\nabla\psi_{f_p}}{\psi_{f_p}} = \frac{1}{f_p(\omega_i,\omega_o)}\left(\frac{\mathrm{d}f_p}{\mathrm{d}g_1} \parallel \frac{\mathrm{d}f_p}{\mathrm{d}g_2} \parallel \frac{\mathrm{d}f_p}{\mathrm{d}\beta}\right). \tag{11}$$

We use a linear blend between the phase functions leading to the following relation for the analytic gradient with respect to the blend factor:

$$\frac{\mathrm{d}f_p}{\mathrm{d}\beta} = f_p(\omega_i\cdot\omega_o, g_1) - f_p(\omega_i\cdot\omega_o, g_2). \tag{12}$$

The derivative with respect to the mean cosine g_j of the j-th lobe is

$$\frac{\mathrm{d}f_p}{\mathrm{d}g_j} = -b(j)\left(\frac{2g_j}{1-g_j^2} + \frac{3g_j + 3\cos\theta}{1+g_j^2+2g_j\cos\theta}\right) f_{\mathrm{HG}}(\omega_i,\omega_o,g_j), \text{ where} \tag{13}$$

$$b(j) = \begin{cases}\beta, & \text{for } j=1,\\ 1-\beta, & \text{for } j=2.\end{cases} \tag{14}$$

These derivatives need to be further modified to account for correlation with other parameters. We employ a blend shape model to approximate the geometry of dentin. Due to discontinuity introduced between the two media by the mesh, we cannot compute derivatives expressing correlation between scattering and geometry parameters in path space. To alleviate this issue we interpolate the phase function parameters of dentin and enamel according to a smooth function in spatial domain based on distance q to the dentin shape. We use a logistic sigmoid,

$$l = \frac{1}{1+e^{-q/s}}, \tag{15}$$

$$\frac{\mathrm{d}l}{\mathrm{d}q} = \frac{e^{q/s}}{s(1+e^{q/s})^2}, \tag{16}$$

where we define the spread factor s such that the major influence is 1mm around the border. We use this weight to linearly interpolate between enamel and dentin phase function parameters,

$$\beta_{\mathrm{mix}} = \beta_{\mathrm{e}} l + (1-l)\beta_{\mathrm{d}}, \tag{17}$$

$$g_{\mathrm{mix}} = g_{\mathrm{e}} l + (1-l) g_{\mathrm{d}}. \tag{18}$$

We are going to re-express the derivative according to phase function parameters with respect to blend weights from the perspective of a point inside enamel,

$$\begin{aligned}\frac{\nabla\psi_{f_p,\mathbf{c}}}{\psi_{f_p,\mathbf{c}}} = \frac{1}{f_p(\omega_i,\omega_o)}\Bigg(&\beta_{\mathrm{mix}}(g_{0,\mathrm{e}}-g_{0,\mathrm{d}})\frac{\mathrm{d}f_p}{\mathrm{d}g_{0,\mathrm{mix}}} + \\ &(1-\beta_{\mathrm{mix}})(g_{1,\mathrm{e}}-g_{1,\mathrm{d}})\frac{\mathrm{d}f_p}{\mathrm{d}g_{1,\mathrm{mix}}} + \\ &(\beta_{\mathrm{e}}-\beta_{\mathrm{d}})(f_{\mathrm{HG}}(g_{0,\mathrm{e}}) - f_{\mathrm{HG}}(g_{1,\mathrm{e}}))\Bigg)\frac{\mathrm{d}l}{\mathrm{d}q}\frac{\mathrm{d}q}{\mathrm{d}\mathbf{c}}.\end{aligned} \tag{19}$$

This relationship can be expressed symmetrically for dentin by exchanging the respective mean cosine coefficients g. Both derivatives with respect to enamel and dentin parameters are computed at the same time by applying their respective weight l and $(1-l)$. We refer to Appendix A for more details about the derivative according to distance and the dedicated Section 6.2 on dentin geometry for specific details about transforming from derivative according to distance to derivative according to blend weights $\mathbf{c}$.

Single-scattering albedo. Assuming a piece-wise constant medium, the light gets attenuated by the scattering coefficient $\sigma_s = \alpha\sigma$. The derivative is analogous to the one of the BSDF function, given as

$$\frac{\nabla\psi_\alpha^{(i)}}{\psi_\alpha^{(i)}} = \frac{1}{\alpha}. \tag{20}$$

Therefore the modified derivative can be expressed with respect to the the blend shape weights analogously to the phase function parameters,

$$\frac{\nabla\psi_{\alpha_{\mathrm{mix}},\mathbf{c}}^{(i)}}{\psi_{\alpha_{\mathrm{mix}},\mathbf{c}}^{(i)}} = \frac{1}{\alpha_{\mathrm{mix}}}(\alpha_{\mathrm{e}}-\alpha_{\mathrm{d}})\frac{\mathrm{d}l}{\mathrm{d}q}\frac{\mathrm{d}q}{\mathrm{d}\mathbf{c}}, \tag{21}$$

The difference between the high albedo values is much larger, resulting in poor scaling of the problem. To alleviate this issue we re-expressed the parameter in a dual domain, $\alpha = 1-10^{\kappa}$, and alter the respective derivative,

$$\frac{\nabla\psi_{\alpha_\kappa}^{(i)}}{\psi_{\alpha_\kappa}^{(i)}} = \frac{\nabla\psi_{\alpha_{\mathrm{mix}}}^{(i)}}{\psi_{\alpha_{\mathrm{mix}}}^{(i)}} 10^{-\kappa}\log 10. \tag{22}$$

Extinction coefficient. The density (precisely the extinction coefficient) acts on both transmission and scattering. In the simple case of a homogeneous medium, the derivative depends on two terms. The first term ($\nabla\psi_{\sigma,s}/\psi_{\sigma,s}$) stems from the scattering coefficient applied at each scattering event. The second term accumulated over each path segment ($\nabla\psi_{\sigma,t}/\psi_{\sigma,t}$) derives from the Beer-Lambert law and the path-length dependence. The collective derivative is

$$\frac{\nabla\psi_\sigma}{\psi_\sigma} = \frac{\nabla\psi_{\sigma,s}}{\psi_{\sigma,s}} + \frac{\nabla\psi_{\sigma,t}}{\psi_{\sigma,t}} = \frac{N_s-1}{\sigma} - \sum_{k=0}^{N_s} t_k, \tag{23}$$

where N_s is the number of path segments and t_k is the length of the k-th path segment. We further reparameterize in a dual domain,

$\sigma = e^{\xi}$, yielding

$$\frac{\nabla\psi_{\xi}}{\psi_{\xi}} = e^{\xi}\left(\frac{N_s - 1}{\sigma} - \sum_{k=0}^{N_s} t_k\right). \tag{24}$$

6.2 Dentin

Similarly to the enamel and the gums we optimize the phase function, density, and albedo for dentin. We also leverage the blend shape model described in Section 4.2 to enable optimizing the geometry of dentin, which has a crucial impact on the appearance.

Dentin geometry. Modeling deformable surfaces within the context of differentiable path tracing requires the exact computation of analytic gradients to establish the connection between deformation parameters and the effect on light transmittance through participating medium. We are going to discuss the specific case of index-matched boundary between two homogeneous participating media. The transmittance in this case is a composition of three operations: deformation (ψ_{G}), intersection distance (ψ_d), and transmittance (ψ_t). Collectively, this subset of the light transport evaluation can be expressed as function composition,

$$\psi_B(\mathbf{c}) = (\psi_t \circ \psi_d \circ \psi_{\mathrm{G}})(\mathbf{c}). \tag{25}$$

We will break down the set of operations, starting from the very last one and build the respective transformations required to reach the final solution. We express the *deformation* model as already described in Eq. (5). To simplify the application of the chain rule against the derivative of the per triangle intersection we are going to restructure the matrix of vertices $\mathbf{v}^{(m)} = [\mathbf{v}_0^{(m)}\, \mathbf{v}_1^{(m)}\, \mathbf{v}_2^{(m)}]$ into a tensor of vertex triplets associated with each ray intersection $\mathbf{G} = [\mathbf{v}^{(0)} \ldots \mathbf{v}^{(n)}]$, with n being the number of intersections. The change according to the model parameters $\mathbf{c} = [c_0 \ldots c_k]$, which are k in total, in this case is expressed by the change of shape $\nabla\psi_{\mathrm{G}} : \mathbb{R}^k \rightarrow \mathbb{R}^{k\times n\times 3\times 3}$ derived from Eq. (5) as

$$\nabla\psi_{\mathrm{G}} = \mathbf{B} - \mathbf{B}_0 = \Delta\mathbf{B}. \tag{26}$$

The next relation that is crucial to derive is the change of *intersection distance* w.r.t. the model. It is governed by the ray-triangle intersection. The intersection point $\mathbf{p}_r^{(m)}$ is expressed as a the distance $t^{(m)}$ traversed along the ray in direction $\mathbf{r}_d$ starting from point $\mathbf{r}_o$:

$$\mathbf{p}_r^{(m)} = \mathbf{r}_o + \mathbf{r}_d t^{(m)}, \tag{27}$$

$$t^{(m)} = \begin{vmatrix} \mathbf{v}_2^{(m)} - \mathbf{v}_0^{(m)} \\ \mathbf{r}_o - \mathbf{v}_0^{(m)} \\ \mathbf{v}_1^{(m)} - \mathbf{v}_0^{(m)} \end{vmatrix} \begin{vmatrix} \mathbf{v}_1^{(m)} - \mathbf{v}_0^{(m)} \\ \mathbf{r}_d \\ \mathbf{v}_2^{(m)} - \mathbf{v}_0^{(m)} \end{vmatrix}^{-1} = |\mathbf{D}_1|\,|\mathbf{D}_2|^{-1}. \tag{28}$$

The transformation to distance space involves the computation of a Jacobian matrix which represents a linear transformation such that $\nabla\psi_d : \mathbb{R}^{k\times n\times 3\times 3} \rightarrow \mathbb{R}^{k\times n}$.

For convenience we define the vector columns of the adjugate matrix as $\mathbf{h}_m^{(k)} \equiv \mathrm{col}_m\, \mathrm{adj}(\mathbf{D}_k)$. By using the definition of the derivative of a determinant we express the columns of the Jacobian

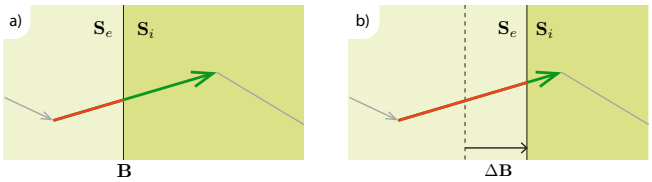

Fig. 7. A path segment going from one medium into another will be influenced when changing the shape $\mathbf{B}$ as it will pass more through one of the media and less through the other.

($\nabla\psi_d^{(m)} \equiv \mathrm{col}_m \nabla\psi_d$) with the following relation:

$$\nabla\psi_d^{(m)} = t^{(m)} \begin{bmatrix} -\mathbf{h}_0^{(1)} - \mathbf{h}_1^{(1)} - \mathbf{h}_2^{(1)} \\ \mathbf{h}_2^{(1)} \\ \mathbf{h}_0^{(1)} \end{bmatrix} - \frac{t^{(m)}}{|\mathbf{D}_2|} \begin{bmatrix} -\mathbf{h}_0^{(2)} - \mathbf{h}_2^{(2)} \\ \mathbf{h}_0^{(2)} \\ \mathbf{h}_2^{(2)} \end{bmatrix}. \tag{29}$$

The final part is computing the transformation to transmittance space. As outlined in Fig. 7 the goal is to compute the effect a change of shape $\Delta\mathbf{B}$ has on the light transport along a segment t. Lets assume the segment crosses a boundary, exiting from a set of media whose optical properties are given by $\mathbf{S}_e$ and entering into another one characterized by $\mathbf{S}_i$. Then the transmittance at the border (excluding the endpoints) is expressed as

$$\psi_t = \exp\left((\mathbf{S}_i - \mathbf{S}_e) \cdot \mathbf{t}\right), \tag{30}$$

where $\mathbf{S} = [\sigma_1 \ldots \sigma_n]$ is a vector of the extinction coefficients which is further split into a vector of entered ($\mathbf{S}_i$) and exited media ($\mathbf{S}_e$). These coefficients are appended to $\mathbf{S}_i$ and $\mathbf{S}_e$ at each successive crossing between media. The Jacobian transformation $\nabla\psi_t : \mathbb{R}^{k\times n} \rightarrow \mathbb{R}^k$ is

$$\nabla\psi_t = (\mathbf{S}_i - \mathbf{S}_e)\exp((\mathbf{S}_i - \mathbf{S}_e) \cdot \mathbf{t}). \tag{31}$$

We are now able to express the entire computation of the accumulated term by also taking into account the common terms with the complete light transport:

$$\frac{\nabla\psi_B(\mathbf{c})}{\psi_B(\mathbf{c})} = \frac{1}{\psi_B(\mathbf{c})}\nabla\psi_t\ \nabla\psi_d\ \nabla\psi_{\mathrm{G}} = (\mathbf{S}_i - \mathbf{S}_e)\ \nabla\psi_d\ \nabla\psi_{\mathrm{G}}. \tag{32}$$

This relation shows that when entering an optically thicker medium with triangles translated in the positive direction w.r.t. to the ray direction, the transmittance increases as the ray traverses less distance through the medium with higher extinction (Fig. 7).

7 RESULTS

We implemented the optimization procedure in Python by using the scientific computing package Numpy as a building block of our custom implementation of Adam. We implemented custom volumetric integrator in Mitsuba [Jakob 2010] which enabled path tracing of derivatives with respect to the optimized parameters and developed the respective BSDF, blend shape and data structures used to store the parameters. We used the Python bindings to interface with the renderer and exchange free parameters $\mathbf{x}$, the resulting derivative blocks $\nabla\psi_{\mathbf{x}}$, and rendered image C. We split derivative estimation and image rendering in two independently seeded rendering passes. Our optimization and rendering runs on a dual CPU machine equipped with Intel Xeon E5-2680v3 processors. We run

the optimization at very low resolution (approximately a fifth of the captured resolution across each dimension) to keep the execution times feasible.

We setup the virtual optimization scene based on calibration data of teeth geometry, camera, and light positions and optimize against our stabilized images. We associate free parameters with the appearance properties of all objects in the scene: teeth, gingiva and oral cavity. We use a point light representation and the recovered positions to simulate our calibrated LED light sources. For the case of a single tooth scene we derive the volume parameters starting point by densely exploring the space spanned by the values reported in literature [Fried et al. 1995; Pop-Ciutrila et al. 2015; Spitzer and Ten Bosch 1975] for both dentin and enamel and comparing the rendered results with our reference measurements. This step is necessary since there is a wide range of values reported in literature mainly due to the various pre-processing procedures performed on the measured sample surfaces which affect roughness and consecutively the appearance. We tabulated the three dimensional domain of albedo, density, and phase function anisotropy into a set of 50000 points and evaluated a scale invariant error of a virtual homogeneous tooth against reference measurements of a masked tooth illuminated from various directions and viewed by different cameras. For each channel we find the point in the domain with the smallest error and use it as our initialization. We do not perform highlight or high frequency normal map fitting. For the case of a complete set of teeth we initialize the appearance of all teeth with the converged values derived by the stochastic single tooth optimization.

We first explore the capability of our technique to estimate scattering parameters of teeth and re-render them in virtual environments by using isolated dead teeth. We acquired micro-CT scans to extract the exact geometry of enamel and dentin. We only mask highlights for this scene since the tooth is static. In Fig. 8 we provide qualitative comparisons of our training set for a single view (left) and additional extrapolated appearance renderings (middle, right) on a testing set with novel viewpoints corresponding to images captured by other cameras but were not visible to the optimization. The procedure is initialized with average values derived from the dental literature. Overall our method provides plausible fits in the case of an isolated dead tooth. Dead teeth contain details which are less pronounced in live teeth. One example is dead tissue and cracks which cast volumetric shadows within the tooth volume that our model does not explicitly support. Other discrepancies in the highlights are caused by the reflection of the metal frame of our dome. These preliminary results validated the plausibility of our technique to estimate parameters of live teeth.

We proceed with capture of live teeth in vivo of a few human volunteers. We provide separate fits by progressively enabling more parameters as shown in Fig. 9. Optimizing with a homogeneous enamel-only volume results in very smooth appearance which lacks detail. To capture the translucency of teeth around the edges we add an inner layer to model dentin. Finally, we model the missing surface details by enabling a textured stains-absorption layer on the enamel outer boundary. We provide the complete comparison of fitting results in Fig. 10, including validation from novel viewpoints and reference images. Overall our method is able to match

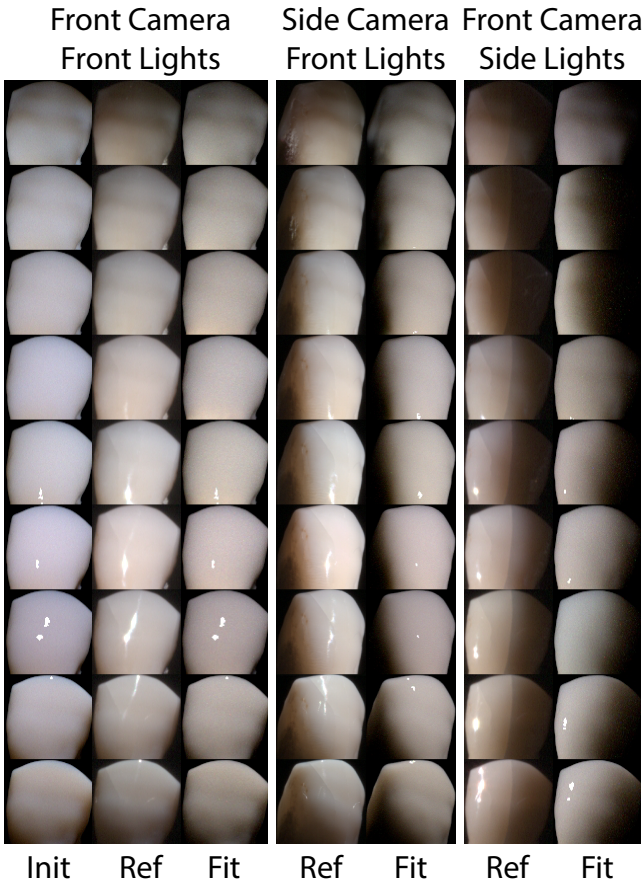

Fig. 8. **Dead Tooth Validation**. We validate our appearance estimation fit on a dead incisor with known dentin geometry. Left: 9 light directions from front (top to bottom) are shown from the front camera. We optimize the appearance using the 9 references images as targets, initialized with default appearance parameters as shown. Our fit result closely matches the reference images. Center: Extrapolating the fit parameters to a side view validates that the volumetric appearance is reasonably close to the real tooth. Right: We extrapolate further to 9 different side light directions, as seen from the front camera view.

the appearance characteristics of a live tooth. Note that when a sample location is masked from the camera then the optimization might arrive to sub-optimal results (due to over-fitting), such as the recovered diffuse albedo of the mouth.

The results of our method on the more challenging complete sets of teeth are shown in Fig. 11, Fig. 12, Fig. 13 on three different volunteers. Additional data regarding the initialization values, textures, masks and loss function values is provided in the supplemental material. All scenes use independent parameters for each tooth that are jointly optimized. With these scenes we test the capability of our model to capture realistic teeth appearance in the presence of gums and inter-reflections from neighboring teeth. We would like to point out that the measurements with the left camera shown in Fig. 13 were not used in the optimization. Our model is optimized using

only the information of center and right cameras. The left camera comparison is an evaluation of the ability of our method to extrapolate to novel view points. Overall our method performs better for teeth in the front part of the mouth where shadows from the oral cavity occur less frequently than the teeth in the back. In the case of a shadow boundary in our measurements we mask the nearby affected teeth since we do not have an accurate reconstruction of the oral cavity and thus we can not simulate the shadow accurately. Most of the molars are usually masked out in our measurements and consequently some of them have false appearance. An accurate reconstruction of the entire oral cavity and the lip retractor might be required to achieve better fits on these particular teeth along with a regularization term. Regularizing the optical properties of these teeth to the average values of the frontal teeth is one option that we tested and partially alleviates the issue. We note that the un-textured reflectance of the oral cavity proxy in this case reaches a compromise between the inner color and the white lip retractor, resulting in ambient illumination which is specularly reflected by some teeth in shadowed areas. The ambient light reflected by the outside environment and light dome is not modeled, similarly to the single tooth experiment. It leads to further discrepancies with the real physical properties that are encoded in the current set of parameters (optical thickness and texture). Capturing a light probe along with the actor can be used in the future to account for the missing light. We experimented with different loss functions (L1, relative L1, L2, relative L2) and came to the conclusion that L2 loss function produces higher quality results as it applies a significantly lower weight on dark and shadowed regions where most of the error due to our measurement methodology is concentrated.

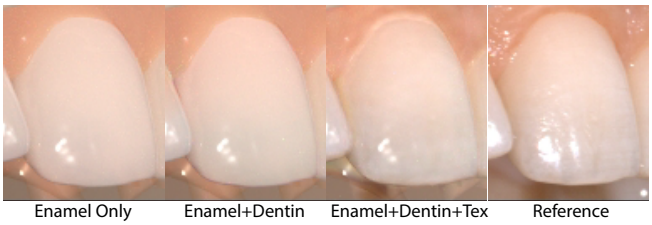

Fig. 9. **Model Validation**. All three components of our appearance model are required to generate realistic teeth appearance. Fitting only a single homogeneous enamel volume appears too opaque (left). Fitting both enamel and our dentin layer allows spatially-varying translucency (left-center). Adding a stain texture layer on top provides the best fit (right-center) to the reference image (right).

8 DISCUSSION AND FUTURE WORK

We presented a system for capturing bulk properties that are key for digitally reproducing the appearance of teeth. Our approach is the first to consider the internal structure of each tooth when reverse-engineering its optical properties and while the results match the photographs well, we would like to point out a number of limitations of our system and how they could be addressed in future work.

Our acquisition setup is very similar to those used for facial-appearance capture, so suitable measurements could likely be made with capture rigs already used in industry. Furthermore, a dense

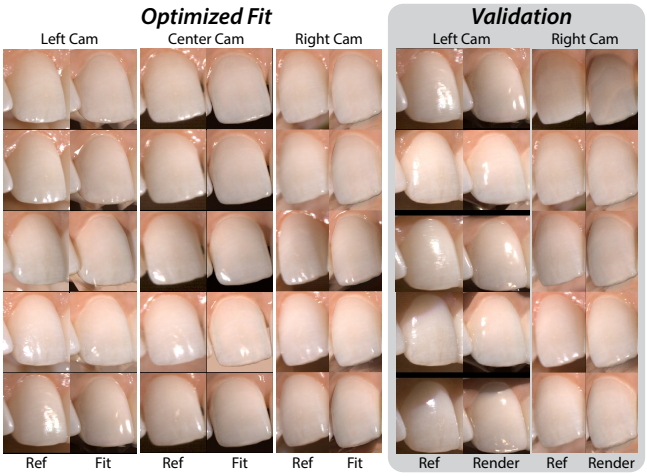

Fig. 10. **Single Tooth Fitting and Validation**. Here we show the accuracy of our method on a single tooth. Left: The optimized appearance fit to three cameras (columns) and five lighting conditions (rows). Our fit tooth appearance closely matches the reference photo. Right: We validate the recovered appearance properties by rendering under five new lighting directions that were not part of the optimization. Our model extrapolates well to these unseen conditions.

capture setup is actually not required, as Fig. 10 shows an optimized fit using only 3 cameras and 5 light sources.

Our parametric model currently does allow for capturing craze lines and cracks in enamel, and assumes the tooth is free of cavities and fillings. These can potentially impact the appearance (e.g. amalgam fillings tend to significantly darken the entire tooth which can be observed on the right molar in Fig. 13) but we believe they need to be encoded in the geometric representation of the tooth and would require a joint optimization of the appearance and the outer tooth geometry. Such optimization is significantly more challenging than inferring the outer geometry and appearance independently. Another option is to consider dental filling and cracks as a morphable mesh similarly to the dentin inner shape and perform joint optimization in similar manner. In this work, we focused on making the first step towards capturing the appearance parameters and the dentin geometry, but we believe the joint optimization represents an interesting avenue for extending our work.

We heavily simplified the simulation of light transport at dielectric boundaries, completely disregarding the effect of light bending upon refraction. While this purely performance-motivated simplification did not cause visually significant differences in our experiments, we would like to note that implementing search-based methods [Holzschuch 2015; Walter et al. 2009] or manifold walks [Jakob and Marschner 2012] would permit simulating light transport more accurately at acceptable render times. We also did not attempt to capture high-frequency glossy highlights. These could be added by subtracting the main diffuse reflectance and performing standard BRDF fitting separately according to one of the existing techniques in the graphics literature [Guarnera et al. 2016].

The volumetric model can be further improved by considering anisotropic particle distributions [Jakob et al. 2010] that would better

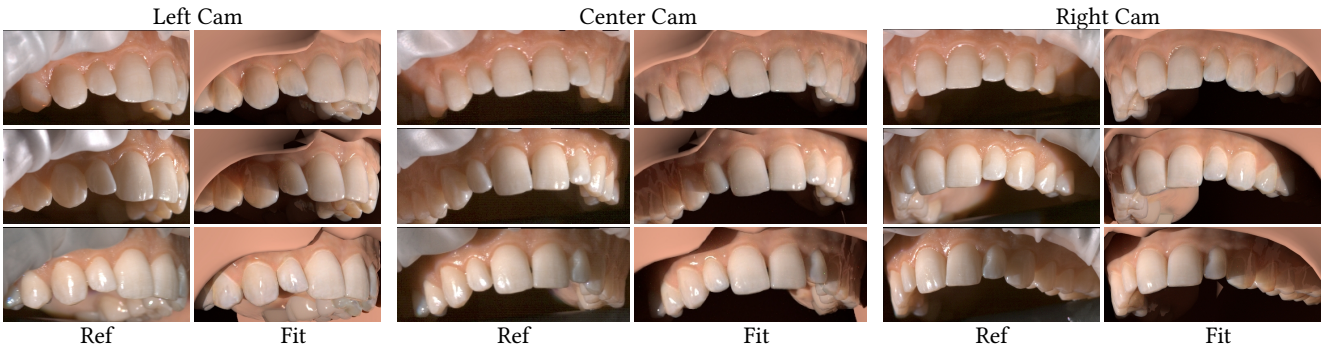

Fig. 11. **Upper Jaw Fitting: Volunteer 1**. We compare the optimized result of our model and optimization on a complete upper jaw consisting of 14 teeth. We show three different viewpoints of the same person's teeth next to corresponding reference images at the same resolution.

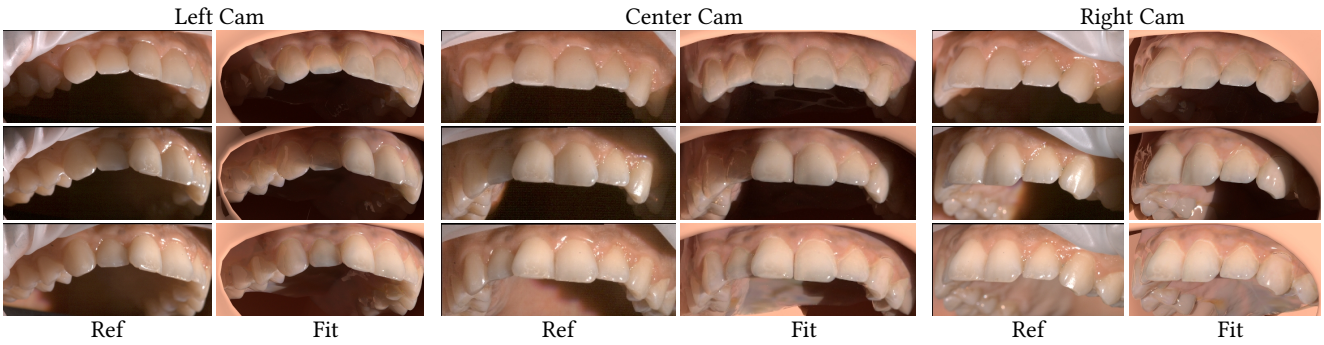

Fig. 12. **Upper Jaw Fitting: Volunteer 2**. Fitting results on the entire upper jaw of another volunteer. We show three different viewpoints of the same person's teeth and comparisons to corresponding reference images.

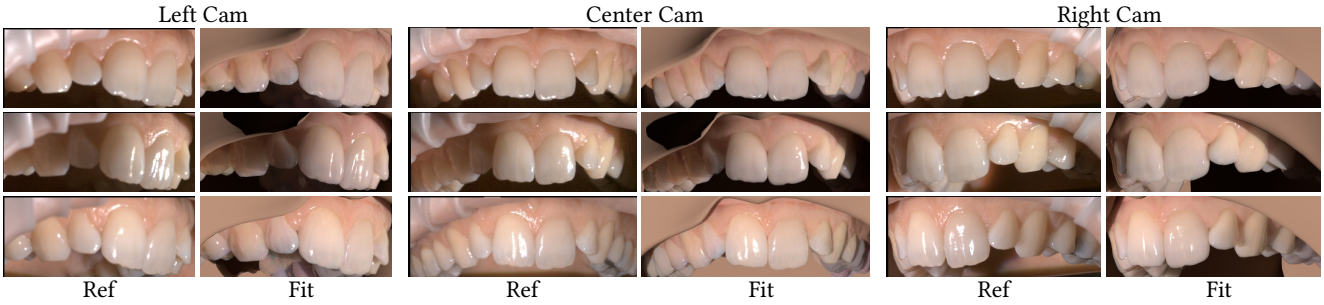

Fig. 13. **Upper Jaw Fitting: Volunteer 3**. Fitting results on the entire upper jaw of another volunteer. In this experiment we only use the center and right cameras during the optimization. The left camera image comparison is a qualitative, modeling power evaluation of our method when used to extrapolate to novel views.

represent the tubular structure of enamel. Another option is to enable heterogeneity and spatial variation of the scattering parameters of enamel and dentin. In this work, we opted for homogeneous approximations with anisotropic phase functions to avoid excessively long rendering times. Another danger of adding more parameters to the model is over-fitting and it may require regularization and priors to constrain the optimization. A careful analysis of potential gains in visual fidelity is to be carried out first before committing to more sophisticated models; our experiments did not reveal such needs but we did not attempt to model heavily stained and damaged teeth where additional parameters could potentially help.

The choice to use per-path analytic gradients for each parameter enables the cancellation or reuse of certain terms within the implementation. We further optimize the layout of the block-sparse gradients and avoid a significant amount of computation. Another alternative is to use automatic differentiation which would need to be applied to a substantial part of the rendering system, including

the carefully optimized ray/triangle intersection code, potentially entailing additional performance and storage costs.

9 ACKNOWLEDGMENTS

We would like to thank Lifan Wu for the initial work on the differentiable integrator and numerical optimization pipeline. We thank Dr. med. dent. Lukas Stucky, Raphael Laue, Dr. med. dent. Sven Mühlemann and the rest of the team at the Center of Dental Medicine at the University of Zurich for the intra-oral geometry scans. Furthermore, Dr. Klearchos Stylianides for many valuable discussions regarding the anatomy and appearance of human teeth. In addition we thank Gaspard Zoss and Jan Wezel for generously donating their time for the measurements required by this work.

REFERENCES

Jae Sung Ahn, Anjin Park, Ju Wan Kim, Byeong Ha Lee, and Joo Beom Eom. 2017. Development of Three-Dimensional Dental Scanning Apparatus Using Structured Illumination. In *Sensors*.

F. O. Bartell, E. L. Dereniak, and W. L. Wolfe. 1981. The theory and measurement of bidirectional reflectance distribution function BRDF and bidirectional transmittance distribution function BTDF. In *Radiation scattering in optical systems*, G. H. Hunt (Ed.), Vol. 257. 154–160. https://doi.org/10.1117/12.959611

Thabo Beeler. 2012. *Passive Spatio-Temporal Geometry Reconstruction of Human Faces at Very High Fidelity.* Ph.D. Dissertation. ETH Zurich / Disney Research Zurich. http://graphics.ethz.ch/~dbeeler/research/bee12c/PassiveSpatioTemporalGeometryReconstructionOfHumanFacesAtVeryHighFidelity.pdf

T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. 2004. High accuracy optical flow estimation based on a theory for warping. In *ECCV*. 25–36.

Stéphanie I. Buchaillard, S. H. Ong, Yohan Payan, and Kelvin Foong. 2007. 3D Statistical Models for Tooth Surface Reconstruction. *Comput. Biol. Med.* 37, 10 (2007), 1461–1471.

C.N. Carter, R.J. Pusateri, Dongqing Chen, A.H. Ahmed, and A.A. Farag. 2010. Shape from shading for hybrid surfaces as applied to tooth reconstruction. In *IEEE ICIP*. 4049–4052.

S. Chandrasekhar. 1960. *Radiative Transfer.* Dover Publications.

Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. 2000. Acquiring the Reflectance Field of a Human Face. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 145–156. https://doi.org/10.1145/344779.344855

Eugene d'Eon and Geoffrey Irving. 2011. A quantized-diffusion model for rendering translucent materials. 30, 4 (July 2011), 56:1–56:14.

Eustace L. Dereniak, Langford G. Brod, and John E. Hubbs. 1982. Bidirectional transmittance distribution function measurements on ZnSe. *Appl. Opt.* 21, 24 (Dec 1982), 4421–4425. https://doi.org/10.1364/AO.21.004421

Craig Donner and Henrik Wann Jensen. 2005. Light Diffusion in Multi-layered Translucent Materials. In *ACM SIGGRAPH 2005 Papers (SIGGRAPH '05)*. ACM, New York, NY, USA, 1032–1039. https://doi.org/10.1145/1186822.1073308

Craig Donner and Henrik Wann Jensen. 2008. Rendering Translucent Materials Using Photon Diffusion. In *ACM SIGGRAPH 2008 Classes (SIGGRAPH '08)*. ACM, New York, NY, USA, Article 4, 9 pages. https://doi.org/10.1145/1401132.1401138

Craig Donner, Tim Weyrich, Eugene d'Eon, Ravi Ramamoorthi, and Szymon Rusinkiewicz. 2008. A Layered, Heterogeneous Reflectance Model for Acquiring and Rendering Human Skin. In *ACM SIGGRAPH Asia 2008 Papers (SIGGRAPH Asia '08)*. ACM, New York, NY, USA, Article 140, 12 pages. https://doi.org/10.1145/1457515.1409093

David Eberly. 2015. *Distance Between Point and Triangle in 3D.* Technical Report. https://www.geometrictools.com/Documentation/DistancePoint3Triangle3.pdf

Aly Farag, Shireen Elhabian, Aly Abdelrehim, Wael Aboelmaaty, Allan Farman, and David Tasman. 2013. Model-Based Human Teeth Shape Recovery from a Single Optical Image with Unknown Illumination. In *Medical Computer Vision: Recognition Techniques and Applications in Medical Imaging (MCV '12)*. 263–272.

Roald Frederickx and Philip Dutré. 2017. A Forward Scattering Dipole Model from a Functional Integral Approximation. *ACM Trans. Graph.* 36, 4, Article 109 (July 2017), 13 pages. https://doi.org/10.1145/3072959.3073681

Daniel Fried, Richard E. Glena, John D. B. Featherstone, and Wolf Seka. 1995. Nature of light scattering in dental enamel and dentin at visible and near-infrared wavelengths. *Journal of Applied Optics* 34, 7 (1995), 1278–1285.

Jeppe Revall Frisvad, Toshiya Hachisuka, and Thomas Kim Kjeldsen. 2014. Directional Dipole Model for Subsurface Scattering. *ACM Trans. Graph.* 34, 1, Article 5 (Dec. 2014), 12 pages. https://doi.org/10.1145/2682629

S. Garrido-Jurado, R. Mu noz Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280 – 2292. https://doi.org/10.1016/j.patcog.2014.01.005

Abhijeet Ghosh, Tim Hawkins, Pieter Peers, Sune Frederiksen, and Paul Debevec. 2008. Practical modeling and acquisition of layered facial reflectance. In *ACM Transactions on Graphics (TOG)*, Vol. 27. ACM, 139.

Ioannis Gkioulekas, Anat Levin, and Todd Zickler. 2016. An Evaluation of Computational Imaging Techniques for Heterogeneous Inverse Scattering. , 685-701 pages.

Ioannis Gkioulekas, Shuang Zhao, Kavita Bala, Todd Zickler, and Anat Levin. 2013. Inverse Volume Rendering with Material Dictionaries. *ACM Trans. Graph.* 32, 6, Article 162 (Nov. 2013), 13 pages. https://doi.org/10.1145/2508363.2508377

D. Guarnera, G.C. Guarnera, A. Ghosh, C. Denk, and M. Glencross. 2016. BRDF Representation and Acquisition. *Comput. Graph. Forum* 35, 2 (May 2016), 625–650. https://doi.org/10.1111/cgf.12867

Ralf Habel, Per H. Christensen, and Wojciech Jarosz. 2013. Photon Beam Diffusion: A Hybrid Monte Carlo Method for Subsurface Scattering. 32, 4 (June 2013). https://doi.org/10.1111/cgf.12148

Milovš Hašan and Ravi Ramamoorthi. 2013. Interactive Albedo Editing in Path-traced Volumetric Materials. *ACM Trans. Graph.* 32, 2, Article 11 (April 2013), 11 pages. https://doi.org/10.1145/2451236.2451237

L. G. Henyey and J. L. Greenstein. 1941. Diffuse radiation in the Galaxy. *Astrophys. J.* 93 (1941), 70–83. https://doi.org/10.1086/144246

Nicolas Holzschuch. 2015. Accurate computation of single scattering in participating media with refractive boundaries. *Computer Graphics Forum* 34, 6 (Sept. 2015), 48–59. https://doi.org/10.1111/cgf.12517

A. Ilie and G. Welch. 2005. Ensuring color consistency across multiple cameras. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, Vol. 2. 1268–1275 Vol. 2. https://doi.org/10.1109/ICCV.2005.88

Wenzel Jakob. 2010. Mitsuba renderer. http://mitsuba-renderer.org.

Wenzel Jakob, Adam Arbree, Jonathan T. Moon, Kavita Bala, and Steve Marschner. 2010. A Radiative Transfer Framework for Rendering Materials with Anisotropic Structure. In *ACM SIGGRAPH 2010 Papers (SIGGRAPH '10)*. ACM, New York, NY, USA, Article 53, 13 pages. https://doi.org/10.1145/1833349.1778790

Wenzel Jakob and Steve Marschner. 2012. Manifold exploration: a Markov Chain Monte Carlo technique for rendering scenes with difficult specular transport. 31, 4 (July 2012), 58:1–58:13.

Henrik Wann Jensen and Per H. Christensen. 1998. Efficient simulation of light transport in scenes with participating media using photon maps. In *98 (Annual Conference Series)*. ACM, New York, NY, USA, 311–320.

Henrik Wann Jensen, Stephen R Marschner, Marc Levoy, and Pat Hanrahan. 2001. A practical model for subsurface light transport. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, 511–518.

James T. Kajiya and Brian P Von Herzen. 1984. Ray tracing volume densities. 18, 3 (Jan. 1984), 165–174.

Masahide Kawai, Tomoyori Iwao, Akinobu Maejima, and Shigeo Morishima. 2014. Automatic Photorealistic 3D Inner Mouth Restoration from Frontal Images. , 51-62 pages.

Pramook Khungurn, Daniel Schroeder, Shuang Zhao, Kavita Bala, and Steve Marschner. 2015. Matching Real Fabrics with Micro-Appearance Models. *ACM Trans. Graph.* 35, 1, Article 1 (Dec. 2015), 26 pages. https://doi.org/10.1145/2818648

Diederik P. Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *CoRR* abs/1412.6980 (2014). http://arxiv.org/abs/1412.6980

Oliver Klehm, Fabrice Rousselle, Marios Papas, Derek Bradley, Christophe Hery, Bernd Bickel, Wojciech Jarosz, and Thabo Beeler. 2015. Recent Advances in Facial Appearance Capture. *Comput. Graph. Forum* 34, 2 (May 2015), 709–733.

Eric P. Lafortune and Yves D. Willems. 1996. Rendering participating media with bidirectional path tracing. Springer-Verlag, London, UK, 91–100.

Christian Thode Larsen, Jeppe Revall Frisvad, Peter Dahl Ejby Jensen, and Jakob Andreas Bærentzen. 2012. Real-Time Rendering of Teeth with No Preprocessing. In *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Charless Fowlkes, Sen Wang, Min-Hyung Choi, Stephan Mantler, Jürgen Schulze, Daniel Acevedo, Klaus Mueller, and Michael Papka (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 334–345.

Hongsong Li, Fabio Pellacini, and Kenneth E. Torrance. 2005. A Hybrid Monte Carlo Method for Accurate and Efficient Subsurface Scattering. In *Proceedings of the Sixteenth Eurographics Conference on Rendering Techniques (EGSR '05)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 283–290. https://doi.org/10.2312/EGWR/EGSR05/283-290

Satya P Mallick, Todd E Zickler, David J Kriegman, and Peter N Belhumeur. 2005. Beyond Lambert: Reconstructing specular surfaces using color. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, Vol. 2. Ieee, 619–626.

Francesco Guido Mangano, Andrea Gandolfi, Giuseppe Luongo, and Silvia Logozzo. 2017. Intraoral scanners in dentistry: a review of the current literature. In *BMC oral*

health.

Albert Mehl and Volker Blanz. 2005. A new approach for automatic reconstruction of occlusal surfaces with the biogeneric tooth model. *Int. J. Comput. Dent.* 8 (2005), 13–25.

Zhuo Meng, Steve Yao, Hui Yao, Yan Liang, Tiegen Liu, Yanni Li, Guanhua Wang, and Shoufeng Lan. 2009. Measurement of the refractive index of human teeth by optical coherence tomography. *Journal of Biomedical Optics* 14 (2009), 14 – 14 – 4. https://doi.org/10.1117/1.3130322

E. Mostafa, S. Elhabian, A. Abdelrahim, S. Elshazly, and A. Farag. 2014. Statistical morphable model for human teeth restoration. In *IEEE ICIP*. 4285–4288.

Adolfo Munoz, Jose I. Echevarria, Francisco J. Seron, Jorge Lopez-Moreno, Mashhuda Glencross, and Diego Gutierrez. 2011. BSSRDF Estimation from Single Images. *Computer Graphics Forum* 30, 2 (2011), 455–464. https://doi.org/10.1111/j.1467-8659.2011.01873.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2011.01873.x

S. Omachi, K. Saito, H. Aso, S. Kasahara, S. Yamada, and K. Kimura. 2007. Tooth shape reconstruction from ct images using spline Curves. In *Wavelet Analysis and Pattern Recognition*, Vol. 1. 393–396.

Marios Papas, Christian Regg, Wojciech Jarosz, Bernd Bickel, Philip Jackson, Wojciech Matusik, Steve Marschner, and Markus Gross. 2013. Fabricating Translucent Materials Using Continuous Pigment Mixtures. *ACM Trans. Graph.* 32, 4, Article 146 (July 2013), 12 pages. https://doi.org/10.1145/2461912.2461974

Mark Pauly, Thomas Kollig, and Alexander Keller. 2000. Metropolis light transport for participating media. Springer-Verlag, London, UK, 11–22.

Matt Pharr, Wenzel Jakob, and Greg Humphreys. 2016. *Physically Based Rendering: From Theory to Implementation* (3rd ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Ioana-Sofia Pop-Ciutrila, Razvan Ghinea, Maria del Mar Perez Gomez, Horatiu Alexandru Colosi, Diana Dudea, and Mandra Badea. 2015. Dentine scattering, absorption, transmittance and light reflectivity in human incisors, canines and molars. *Journal of Dentistry* 43 (2015), 1116–1124.

D. Spitzer and J. J. Ten Bosch. 1975. The absorption and scattering of light in bovine and human dental enamel. *Calcified Tissue Research* 17, 2 (1975), 129–137.

Jos Stam. 1995. Multiple scattering as a diffusion process. (1995), 41–50.

Sarah Tariq, Andrew Gardner, Ignacio Llamas, Andrew Jones, Paul Debevec, and Greg Turk. 2006. Efficient estimation of spatially varying subsurface scattering parameters. *Vision, Modeling, and Visualization (VMV2006)* (2006), 129–136.

Xin Tong, Jiaping Wang, Stephen Lin, Baining Guo, and Heung-Yeung Shum. 2005. Modeling and Rendering of Quasi-homogeneous Materials. *ACM Trans. Graph.* 24, 3 (July 2005), 1054–1061. https://doi.org/10.1145/1073204.1073311

Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. 2007. Microfacet Models for Refraction Through Rough Surfaces. In *Proceedings of the 18th Eurographics Conference on Rendering Techniques (EGSR'07)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 195–206. https://doi.org/10.2312/EGWR/EGSR07/195-206

Bruce Walter, Shuang Zhao, Nicolas Holzschuch, and Kavita Bala. 2009. Single Scattering in Refractive Media with Triangle Mesh Boundaries. *ACM Transactions on Graphics* 28, 3 (Aug. 2009), 92:1–8. https://doi.org/10.1145/1531326.1531398

Jiaping Wang, Shuang Zhao, Xin Tong, Stephen Lin, Zhouchen Lin, Yue Dong, Baining Guo, and Heung-Yeung Shum. 2008. Modeling and Rendering of Heterogeneous Translucent Materials Using the Diffusion Equation. *ACM Trans. Graph.* 27, 1, Article 9 (March 2008), 18 pages. https://doi.org/10.1145/1330511.1330520

Tim Weyrich, Wojciech Matusik, Hanspeter Pfister, Bernd Bickel, Craig Donner, Chien Tu, Janet McAndless, Jinho Lee, Addy Ngan, Henrik Wann Jensen, et al. 2006. Analysis of human faces using a measurement-based skin reflectance model. In *ACM Transactions on Graphics (TOG)*, Vol. 25. ACM, 1013–1024.

Chenglei Wu, Derek Bradley, Pablo Garrido, Michael Zollhöfer, Christian Theobalt, Markus Gross, and Thabo Beeler. 2016. Model-based Teeth Reconstruction. *ACM Trans. Graph.* 35, 6, Article 220 (2016), 220:1–220:13 pages.

R Yanagisawa, Y Sugaya, S Kasahara, and S Omachi. 2014. Tooth shape reconstruction from dental CT images with the region-growing method. *Dentomaxillofacial Radiology* 43, 6 (2014), 20140080.

M. D. Zeiler. 2012. ADADELTA: An Adaptive Learning Rate Method. *ArXiv e-prints* (Dec. 2012). arXiv:cs.LG/1212.5701

Shu-Xian Zheng, Jia Li, and Qing-Feng Sun. 2011. A Novel 3D Morphing Approach for Tooth Occlusal Surface Reconstruction. *Comput. Aided Des.* 43, 3 (2011), 293–302.

A DERIVATIVE OF DISTANCE WITH RESPECT TO VERTEX TRANSLATION

Computing the derivative of distance to a point w.r.t. translation of individual vertices is essential when computing the contribution of spatially varying parameters defined by a blend shape model. We will consider these three primitives being closest to a point: a point, an edge or a plane. These are three cases when computing the distance between a point and a triangle according to the method described by Eberly [2015]. Note that we are going to assume that within infinitesimal translations the closest primitive does not change.

Derivative of the distance w.r.t. a triangle vertex. The distance,

$$q_{\mathbf{p}_s} = |\mathbf{p}_s - \mathbf{v}|, \tag{33}$$

is defined by the spatial position of a point $\mathbf{p}_s$, and the closest triangle vertex $\mathbf{v}$. The derivative is computed by differentiating according to each respective coordinate of the vertex and populating the respective column of the triangle derivative matrix $\frac{\mathrm{d}q_{\mathbf{p}_s}}{\mathrm{dG}}$,

$$\frac{\mathrm{d}q_{\mathbf{p}_s}}{\mathrm{d}\mathbf{v}} = -\frac{\mathbf{p}_s - \mathbf{v}}{|\mathbf{p}_s - \mathbf{v}|} \tag{34}$$

Derivative of the distance to a triangle edge. The distance to a triangle edge is expressed according to the vector orthogonal to the line and connecting the target point,

$$\mathbf{t} = \frac{\mathbf{v}_1 - \mathbf{v}_0}{|\mathbf{v}_1 - \mathbf{v}_0|} \tag{35}$$

$$\mathbf{q}_{\mathbf{p}_s} = (\mathbf{p}_s - \mathbf{v}_0) - \mathbf{t}(\mathbf{t} \cdot (\mathbf{p}_s - \mathbf{v}_0)) \tag{36}$$

$$q_{\mathbf{p}_s} = |\mathbf{q}_{\mathbf{p}_s}|, \tag{37}$$

where we use the projection on the tangent $\mathbf{t}$ determined by the two vertices $\mathbf{v}_0$ and $\mathbf{v}_1$ to derive the final equation. Differentiating this equation is more involved and we provide just the final set of equations for reference,

$$\mathbf{A}_{\mathbf{v}_0} = \left(\frac{t \cdot (\mathbf{p}_s - \mathbf{v}_0)}{|\mathbf{v}_1 - \mathbf{v}_0| q_{\mathbf{p}_s}} - 1 \right) \mathbf{q}_{\mathbf{p}_s} \tag{38}$$

$$\mathbf{A}_{\mathbf{v}_1} = -\frac{t \cdot (\mathbf{p}_s - \mathbf{v}_0)}{|\mathbf{v}_1 - \mathbf{v}_0| q_{\mathbf{p}_s}} \mathbf{q}_{\mathbf{p}_s}. \tag{39}$$

The individual derivatives according to each vertex $(\mathbf{A}_{\mathbf{v}_0}, \mathbf{A}_{\mathbf{v}_1})$ are further assigned to the correct corresponding columns in the triangle derivative matrix $\frac{\mathrm{d}q_{\mathbf{p}_s}}{\mathrm{dG}}$.

Derivative of the distance to a triangle interior plane. The distance to a triangle interior point can be expressed by taking into account the normal vector $\mathbf{n}$ defined by the two edge vectors $u_0 = \mathbf{v}_1 - \mathbf{v}_0$ and $u_1 = \mathbf{v}_2 - \mathbf{v}_0$,

$$\mathbf{n} = \frac{u_0 \times u_1}{|u_0 \times u_1|},$$

and projecting the vector connecting a triangle vertex $\mathbf{v}_0$ and the target point $\mathbf{p}_s$,

$$q_{\mathbf{p}_s} = \mathbf{n} \cdot (\mathbf{p}_s - \mathbf{v}_0). \tag{40}$$

We are going to use the columns of the identity matrix $\mathbf{I}_i = \mathrm{col}_i \mathbf{I}$ to derive a compact representation of the derivative,

$$\mathbf{A}_{ki} = -\mathbf{I}_i \times u_k$$

$$\mathbf{A}_{\mathrm{cross}} = \begin{bmatrix} \mathbf{A}_{00} + \mathbf{A}_{10} & \mathbf{A}_{01} + \mathbf{A}_{11} & \mathbf{A}_{02} + \mathbf{A}_{12} \\ -\mathbf{A}_{00} & -\mathbf{A}_{01} & -\mathbf{A}_{02} \\ -\mathbf{A}_{10} & -\mathbf{A}_{11} & -\mathbf{A}_{12} \end{bmatrix}$$

$$\frac{\mathrm{d}q^{(i,j)}_{\mathbf{p}_s}}{\mathrm{d}(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)} = \mathrm{sgn}(q_{\mathbf{p}_s}) \mathbf{A}^{(i,j)}_{\mathrm{cross}} \cdot \left((\mathbf{p}_s - \mathbf{v}_0) + q_{\mathbf{p}_s} \frac{\mathbf{n}}{|u_0 \times u_1|} \right). \tag{41}$$