# CLINICAL CROWNS SHAPE RECONSTRUCTION - AN IMAGE-BASED APPROACH

*Shireen Elhabian*[1], *Aly Abdelrahim*[1], *Aly Farag*[1], *David Tasman*[2], *Wael Aboelmaaty*[3] *and Allan Farman*[2]

[1]Computer Vision and Image Processing Laboratory, ECE Dept., University of Louisville, Louisville, KY, 40292, USA.

[2]School of Dentistry, University of Louisville, Louisville, KY, 40202, USA.

[3] Visiting Researcher Dental school, University of Louisville, KY, USA from Faculty of Dentistry, Mansoura University, Egypt.

## ABSTRACT

Precise knowledge of the 3D shape of clinical crowns is crucial for the treatment of malocclusion problems as well as several endodontic procedures. While Computed Tomography (CT) would present such information, it is believed that there is no threshold radiation dose below which it is considered safe. In this paper, we propose an image-based approach which allows for the construction of plausible human jaw models in vivo, without ionizing radiation, using fewer sample points in order to reduce the cost and intrusiveness of acquiring models of patients teeth/jaws over time. We assume that human teeth reflectance obeys Wolff-Oren-Nayar model where we experimentally prove that teeth surface obeys the microfacet theory. The inherent relation between the photometric information and the underlying 3D shape is formulated as a statistical model where the coupled effect of illumination and reflectance is modeled using the Helmhotlz Hemispherical Harmonics (HSH)-based irradiance harmonics whereas the Principle Component Regression (PCR) approach is deployed to carry out the estimation of dense 3D shapes. Vis-à-vis dental applications, the results demonstrate a significant increase in accuracy in favor of the proposed approach where our system is evaluated on a database of 16 jaws.

***Index Terms—*** image irradiance, statistical shape-from-shading, jaw modeling, statistical priors, harmonic expansion

## 1. INTRODUCTION

Dentistry usually requires accurate 3D shape of clinical crowns (the visible part of the human jaw) for the treatment of malocclusion problems and several endodontic procedures. This can also benefit the construction of tooth implants where crowns and bridges of high quality are needed. While Computed Tomography (CT) scanning would present 3D information, it is believed that there is no threshold radiation dose below which it is considered safe [1]. Further, CT-scanning is considered expensive and not paid by insurance companies unless disease oriented. Meanwhile, dental offices in rural areas do not have such a luxury. Thus our intent is to develop a purely image-based shape reconstruction mechanism as a cost-effective and non-invasive tool for doctors, dentists, and researchers to obtain crown models in vivo, without ionizing radiation. This is a challenging problem due to the unfriendly environment of taking measurements inside a persons mouth. Further assumptions of the presence of distinct features on the object in stereo images and the photo consistency in space carving are rarely valid in practice.

Due to the lack of surface texture, shape-from-shading (SFS) algorithms have been used to reconstruct the 3D shape of human teeth/jaw due to the significant shading cue presented in an intra-oral image, e.g. [2]. Nonetheless, in principle, SFS is an ill-posed problem whereas most SFS approaches assume known parameters of surface reflectance and simple point light source with known direction. Prados and Faugeras [3] showed that constraining the SFS problem to a specific class of objects can improve the accuracy of the recovered shape. Thus the main objective of the presented work is to develop and validate a holistic approach for image-based 3D reconstruction of the human jaw. Our approach is based on a single captured optical image and a statistical shape recovery approach which makes use of a small number of measured points to construct a plausible 3D model through a learned correspondence based on a measured human jaw dataset. We believe that this approach has the potential to greatly improve plausibility of the resulting SFS models.

## 2. RELATED WORK

Several works have been conducted in the reconstruction of tooth occlusal surface based on 3D surface measurements and a training teeth dataset. For example, Zheng *et al.* [4] considered teeth anatomical features where they improved the snake model to automatically capture the features on the tooth surface such as marginal ridges, cusps and groove lines. In [5], manual alignment between each training tooth and a generic tooth was done, in order to have a point-to-point correspondence between different specimens. A point distribution model was then computed using Principal Component Analysis (PCA) to describe the shape. Sporring and Jensen [6] proposed a statistical model of a selection of tooth shapes and a reconstruction of missing data by including information the position and anatomy of other teeth. Their system depended on hand picked landmark detection. Alternatively, Blanz and Vetter [7] proposed a statistical model of a selection tooth shapes by warping each training shape landmark to a template shape, where the mesh of this template is projected onto the shape space before warping landmarks and mesh vertices back. Yet they rely on manual annotation to achieve onlay and inlay restoration while they did not handle the recovery of missing crown as their model is a tooth-based model.

Our earlier work in [8] presented a model-based human jaw shape reconstruction, yet this work is lacking in the following aspects: (1) We assumed the simple Lambertian model for tooth reflection to model its appearance. However, the tooth surface is rough and wet, giving rise to Fresnel reflection due to different refractive indices of the saliva and the tooth material. (2) Shape prior information was constructed assuming a bijective mapping between the xy-plane and the z-direction, enabling the reconstruction of 2.5D shape in contrast to full 3D representation. This inhibited the reconstruction of labial, buccal and lingual surfaces as well. (3) The appearance prior model was constructed under the assumption of a very low-frequency illumination while natural illumination exhibits higher-frequencies. (4) We ignored the color information which is available in the given intra-oral images.

In this paper, we propose a method to overcome the aforementioned limitations. In particular, prior shape (full 3D), albedo (colored) and appearance (non-Lambertian) models from real data, which are metric in nature, are incorporated into the shape recovery framework in order to resolve the concave/convex ambiguity of

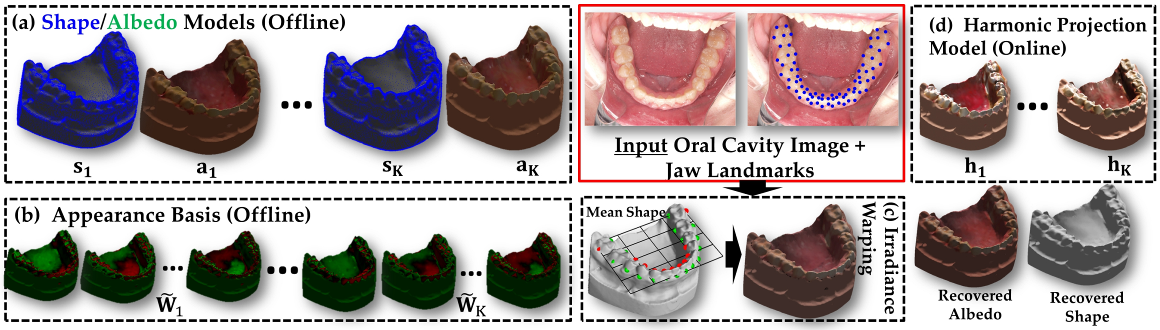


**Fig. 1.** Block diagram of the proposed model-based human jaw shape recovery: (a) An aligned ensemble of the shapes and albedos of human jaws is used to build the 3D shape and albedo models. (b) Given the albedo and surface normals of a certain jaw in the ensemble, the proposed appearance bases are constructed. Given an input oral cavity image under general unknown illumination and a set of human jaw anatomical landmark points: (c) Dense correspondence is established between the input irradiance and the mean jaw shape using 3D thin-plate splines. (d) The input image, in the reference frame, is projected onto the subspace spanned by the appearance basis of each jaw sample which are scaled and summed-up to construct the harmonic projection (HP) irradiance that encodes the illumination and reflectance conditions of the input image. Such images are then used to construct an HP model of the input image. The inherent relation between the HP irradiance and the corresponding shape and albedo is cast as a regression framework to recover the shape and albedo of the input image.

conventional SFS approaches. Moreover, we relax known illumination and reflectance assumptions using the harmonic expansion of the image irradiance equation where we are able to incorporate prior information about natural illumination and teeth reflectance characteristics. Refer to [8] for more algorithmic details regarding the shape recovery process, see Fig. 1 for illustration. The results demonstrate the effect of adding statistical prior as well as appearance (illumination and reflectance) modeling on the accuracy of the recovered shape.

## 3. AWAY FROM LAMBERTIAN ASSUMPTION

While the graphics community has developed comprehensive models to account for translucent materials such as human teeth for photo-realistic image synthesis, such models require computationally expensive rendering techniques. Thus the vision community has opted to discount subsurface scattering where the notion of surface reflectance can be analytically modeled by surface bidirectional reflectance distribution function (BRDF) which depends on, among other factors, the microscopic surface characteristics. In our university facility, we have 3D optical surface profiler, NewView 700s from Zygo company which is based on Scanning White-Light Interferometry technology. The field of view ranges from $0.35mm$ to $3.5mm$. We measured height variations of different visual tooth surfaces from an area of $0.35mm^2$ using 10X optical zoom, see Fig. 2. The average surface profiles provide us a physical validation that the appearance of a tooth surface can be modeled using the microfacet-theory which assumes that the surface consists of a large number of small flat facets while microfacet reflectance models tend to be intuitive with tractable analytical expressions. Accounting for surface roughness and Fresnel reflection, we assume that teeth reflectance obeys Wolff-Oren-Nayar model [9].

## 4. PROPOSED TEETH APPEARANCE MODEL

Known illumination with simple point source models are widely assumed by conventional SFS algorithms. One way to relax such assumptions is to construct a generative appearance model which predict the appearance of an object (human jaw in particular) under natural illumination. The mathematical abstraction of the image formation process can be devised to construct such a model. In particular, the harmonic expansion of the image irradiance equation, being formulated in a convolution framework [10, 11], can be used to derive an analytic subspace to represent images under fixed pose but different complex illumination conditions.

The convolution theory implies a multiplicative framework in the frequency domain where an image is represented as a linear combination of pre-computed basis functions, which we term as *irradiance harmonics*, $\{\mathcal{B}_s\}$, that are pose and geometry dependent. Representing the illumination by its spherical harmonics (SH) coefficients $l_n^m$ as in [11] and the surface reflectance by its coefficients $a_{pr}^q$ in the Helmholtz HSH-based basis [12], the image irradiance can be defined as,

$$E(\alpha, \beta) = \sum_s c_s \mathcal{B}_s(\alpha, \beta) \tag{1}$$

where $(\alpha, \beta)$ are the spherical coordinates of the surface normal, $c_s = l_n^m a_{pr}^q$ with $s$ and its corresponding indices $n, m, p, r$ and $q$ are given by an ordering function of the basis functions and the irradiance harmonics is defined as,

$$\begin{aligned}\mathcal{B}_{npr}^{mq}(\alpha, \beta) \quad &= \quad \mathcal{D}_{mq}^n(\alpha, \beta) \int_{\Omega_i'} Y_n^q(\vec{\omega}_i') \\ &\times \mathcal{H}_{pr}^q(\theta_i', \alpha, |\phi_i' - \pi|) \cos \theta_i' d\vec{\omega}_i' \end{aligned} \tag{2}$$

where $\{Y_n^q\}$ are the real SH bases, $\{\mathcal{H}_{pr}^q\}$ are the Helmhotlz HSH-based reflectance bases and $\mathcal{D}^n$ are Wigner's rotation matrices which encodes how to express a rotated SH basis function in terms of all other SH bases of the same order.

We cast the process of finding such subspace as establishing a relation between its principal components and that of the irradiance harmonics. This resolves the issue of dimensionality since the source of randomness in the imaging process becomes the irradiance harmonics coefficients rather than the whole image realization. Let a $D$-pixel image be represented in the vector space as $\mathbf{e} \in \mathbb{R}^D$. The objective is to define an orthonormal projection matrix $\tilde{\mathbf{W}} \in \mathbb{R}^{D \times D'}$ which maps the image space to a lower-dimensional subspace $\mathbb{R}^{D'}$, with $D' \leq D$, which captures most of the variations due to illumination and reflectance. For $\mathbf{y} = \tilde{\mathbf{W}}^T \mathbf{e}$, the projection matrix $\tilde{\mathbf{W}}$ should maximize the total vector scatter such that,

$$\tilde{\mathbf{W}} = \underset{\tilde{\mathbf{W}}}{\operatorname{argmax}} \, \Psi_{\mathbf{y}} \quad \text{s.t.} \quad \Psi_{\mathbf{y}} = E\{\|\mathbf{y} - \bar{\mathbf{y}}\|_F^2\} \tag{3}$$

Let the $s$-th irradiance harmonics be represented in the vector space as $\mathbf{b}_s \in \mathbb{R}^D$ which can be written as $\mathbf{b}_s = \tilde{\mathbf{W}}_B \mathbf{c}_s^B$ with $\tilde{\mathbf{W}}_B \in \mathbb{R}^{D \times D'}$ and $\mathbf{c}_s^B \in \mathbb{R}^{D'}$. The optimal $\tilde{\mathbf{W}}_B$ is determined from the $D'-$eigenvectors of $\mathbf{B} = [\mathbf{b}_1 \dots \mathbf{b}_S]$ corresponding to its

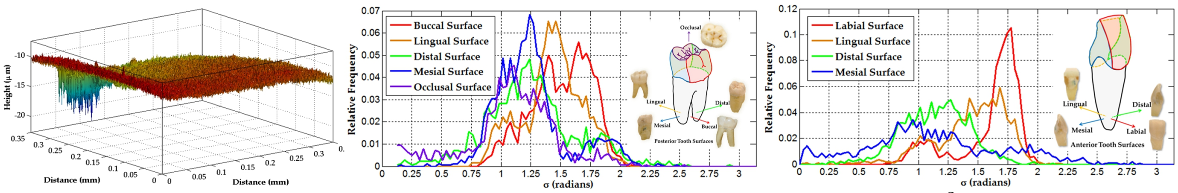


**Fig. 2**. The roughness parameter is estimated based on the measurement of microscopic height variation of a $0.35mm^2$ surface patches of different surface types for incisor and molar teeth (see left for a sample). According to the distribution, the parameter tends to lie between 0.7 to 2 radians regardless the tooth surface type.

largest $D'$-eigenvalues. Relating the principal components of the image space to that of the irradiance harmonics in the vector space results in $\tilde{\mathbf{W}}^T = \mathbf{A}_W \tilde{\mathbf{W}}_B^T$ where $\mathbf{A}_W \in \mathbb{R}^{D' \times D'}$. The solution of this matrix is given by the $D'$-eigenvector of $\Upsilon^T$ such that,

$$\Upsilon = \sum_s \sum_{s'} E\{(c_s - \bar{c}_s)(c_{s'} - \bar{c}_{s'})\} \mathbf{c}_s^B (\mathbf{c}_{s'}^B)^T \tag{4}$$

where $E\{(c_s - \bar{c}_s)(c_{s'} - \bar{c}_{s'})\} = E\{c_s c_{s'}\} - E\{c_s\}E\{c_{s'}\}$. As such, given the shape and albedo of a jaw data sample, we can construct its appearance subspace while incorporating natural illumination (*e.g.* [13]) and teeth reflectance properties. In particular, spherical harmonics is used to compute the illumination spectrum of a database of environment maps [13] while our Helmholtz hemispherical harmonics (HSH)-based basis [12] is used to compute the reflectance spectrum of a database of teeth reflectance. We uniformly sample the roughness (according to Fig. 2) and the enamel's refractive index ($1.62 \pm 0.02$[14]) domain.

## 5. JAW PRIOR MODELS

The model-based shape recovery in [8] (see Fig. 1 for illustration) involves the construction of three models; namely the shape, albedo (also referred to as texture) and appearance (net result of illumination and reflectance) models[1]. While the first two models are constructed in an offline stage, the appearance model is constructed at runtime when an input image is presented to our shape recovery framework.

### 5.1. Shape Model - Full 3D

The jaw's shape model is constructed from a training data ensemble of 3D triangular meshes where each mesh is obtained from a high resolution computed tomography (CT) scan of human jaw molds. We follow the work by Patel et al. [15] in obtaining dense correspondence between different jaw surfaces where a finite set of sparse landmark points are manually annotated for all the database samples. Generalized Procrustes Analysis (GPA)[16] is then performed to provide an initial rigid alignment of the dense shapes to a common reference frame. The 3D thin-plate spline [16] is then applied in an iterative manner in order to obtain a dense correspondence between all shapes in the database. PCA is then performed on the set of shape vectors, where the resulting shape model can be written as $\mathbf{s} = \bar{\mathbf{s}} + \mathbf{P}_s \boldsymbol{\lambda}_s$ where $\mathbf{P}_s = [\mathbf{s}_1, \mathbf{s}_2, \cdots]$ are the shape eigenvectors and $\boldsymbol{\lambda}_s$ is the set of shape coefficients.

### 5.2. Albedo Model - Color Incorporated

Shadows due to non-convex jaw regions and non-uniform distribution of illumination inhibit using occlusal images as albedo. As such, we factor out the reflectance information (albedo) from the given texture using the intrinsic image decomposition proposed by Barron and Malik [17]. We use 3D thin-plate spline to provide a warping function between image pixels (assumed to be on the xy-plane in the 3D space) and surface points using image landmarks and surface landmarks as control points. Mapping ambiguities are resolved using a least-squares plane fit to the cervical landmark points. In order to incorporate color information, we use the Lab color space instead of RGB since the latter suffers from strong correlation among its color channels as well as non-linearity. PCA is then performed on the set of albedo vectors, where the resulting albedo model can be written as $\mathbf{a} = \bar{\mathbf{a}} + \mathbf{P}_a \boldsymbol{\lambda}_a$ where $\mathbf{P}_a = [\mathbf{a}_1, \mathbf{a}_2, \cdots]$ are the albedo eigenvectors and $\boldsymbol{\lambda}_a$ is the set of albedo coefficients.

### 5.3. Harmonic Projection (HP) Irradiance Model

Representing the surface reflectance function in terms of the proposed appearance basis allow us to infer the coupled effect of illumination and reflectance of the input irradiance signal as follows. Given an input image irradiance $E$ (being warped to the mean shape) and the appearance basis matrix $\tilde{\mathbf{W}}$ of a jaw sample, the irradiance coefficients $\hat{\mathbf{y}}$ are deduced to best match the input irradiance s.t. $\mathbf{e} = \tilde{\mathbf{W}}\mathbf{y}$ which can be solved using singular value decomposition (SVD). As such, the *harmonics projection* (HP) irradiance can be defined as as $\mathbf{h} = \tilde{\mathbf{W}}\hat{\mathbf{y}}$ where $\mathbf{h}$ provides a mean of encoding the illumination and reflectance of the input irradiance while maintaining the identity of the object whose basis are used in the reconstruction process.

While the shape model and the albedo model are constructed in a pre-processing (offline) step, the HP model is constructed when the input irradiance is given to our framework in order to incorporate the illumination and reflectance conditions of the given irradiance into the prior information. The resulting HP model can be written as $\mathbf{h} = \bar{\mathbf{h}} + \mathbf{P}_h \boldsymbol{\lambda}_h$ where $\mathbf{P}_h = [\mathbf{h}_1, \mathbf{h}_2, \cdots]$ are the HP eigenvectors and $\boldsymbol{\lambda}_h$ is the set of HP coefficients.

## 6. EXPERIMENTAL RESULTS

Upper and lower jaw models are constructed from eight subjects (5 males and 3 females with ages range from 16 to 46 years old) using their oral cavity images and the CT-scan of their respective molds (lower and upper jaws)[2]. We compare our approach with a recently evaluated SFS approach [18] for tooth surface reconstruction. In order to share the same metric coordinate frame, the average jaw shape is used as a reference to establish a dense correspondence between the groundtruth CT scan of the jaw mold corresponding to each testing image and the reconstructed shape.

Out-of-training jaw samples are reconstructed where four types of samples are considered: (a) pre-repair and (b) post-repair lower jaw, (c) pre-repair and (d) post-repair upper jaw. Fig. 3 shows a sample of shape and albedo reconstruction of an upper jaw. It important to note that SFS only recovers a height map (2.5D) of the input image where there is no metric information reserved. With the metric prior used to train the offline shape model, our approach reconstructs the triangular mesh (3D) corresponding to the input im-

[1]In contrast to the proposed appearance model, the appearance model in [8] encodes only a Lambertian reflection.

[2]A key requirement for successful statistical SFS is the availability of a comprehensive database that describe the teeth/jaw variability per age, gender and ethnic factors. Our ongoing efforts aim to undertake such a task and make the databases available for researchers worldwide.

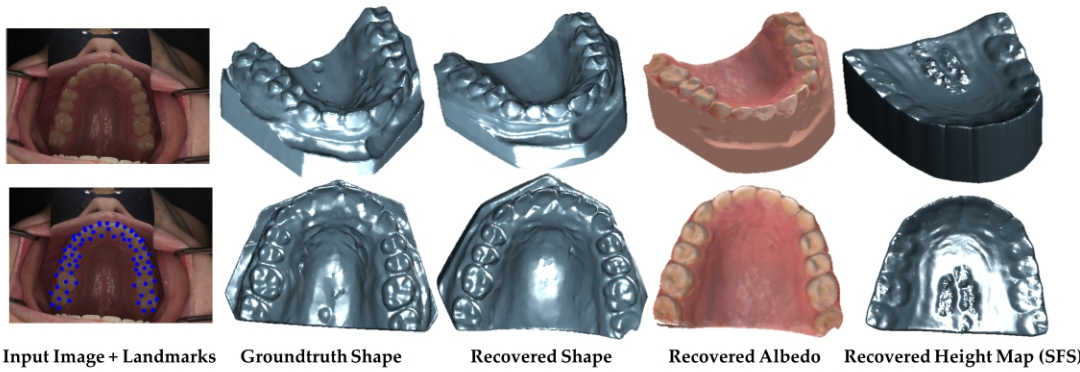


**Fig. 3**. Sample reconstruction result of an upper (post-repair) jaw (bottom row shows the top-view of the occlusal surface).

age. In Fig. 3, one can observe the closeness of our reconstruction to the groundtruth shape when compared to the recovery shape from SFS. This emphasizes the role of incorporating prior-information for shape recovery as well as appearance modeling.

Table 1 reports the Root Mean Square (RMS) error in $mm$ between the 3D points from the CT-scan and the corresponding reconstructed surface points. For the sake of comparison, we also include our earlier results in case of assuming Lambertian reflectance [8]. Notice that the error values of our reconstructions are minimal when compared to SFS-based reconstruction. Pre-repair error values are also smaller than post-repair values in most of the samples, indicating that the statistical prior capability of capturing irregular tooth shapes and locations.

A natural question arises where even smaller reconstruction errors are needed, to that end we need to point out that these results are based on a model that is being trained on a small ensemble of 16 jaws. With a large enough ensemble of objects, credible shape, albedo and appearance models would be possible, which when morphed to the crown reconstructions would produce a more realistic jaw.

**Table 1**. Average *whole jaw* surface reconstruction accuracy (RMS) in $mm$

| **Jaw Type** | **Proposed non-Lambertian SSFS** | **Lambertian SSFS [8]** | **SFS** [19] |
|---|---|---|---|
| **Upper, Pre-repair** | **0.6289** | 2.08999 | 15.2995 |
| **Upper, Post-repair** | **0.6689** | 2.02334 | 16.3098 |
| **Lower, Pre-repair** | **0.6714** | 3.11911 | 12.1241 |
| **Lower, Post-repair** | **0.8073** | 2.57112 | 13.4959 |

## 7. CONCLUSION

In this paper, we presented an affordable, flexible, automatic dental tool for the reconstruction of the clinically visible part of the human jaw based on a single captured optical image and a statistical shape recovery approach. While most shape-from-shading (SFS) approaches assume known parameters of surface reflectance and point light source with known direction, our work has relaxed such assumptions using the harmonic expansion of the image irradiance equation where we were able to incorporate prior information about natural illumination and teeth reflectance characteristics. We further presented an experimental justification that human teeth obeys microfacet theory where its reflectance was assumed to follow Wolff-Oren-Nayar reflectance model. The results demonstrated the effect of adding statistical prior as well as appearance (illumination and reflectance) modeling on the accuracy of the recovered shape. The next step is to investigate the fusion of SFS and statistical SFS where SFS provides the object-specific constructions while statistical SFS is perform shape recovery based on partial information.

## 8. REFERENCES


[1] European Commission, "European guidelines on radiation protection in dental radiology," Radiation Protection Issue number 136, 2004.

[2] A. Abdelrahim, M. Abdelrahman, H. Abdelmunim, A. Farag, and M. Miller, "Novel image-based 3d reconstruction of the human jaw using shape from shading and feature descriptors," in *British Machine Vision Conference (BMVC)*, 2011, pp. 41.1–41.11.

[3] E. Prados and O. Faugeras, "Unifying approaches and removing unrealistic assumptions in shape from shading: Mathematics can help," in *European Conference on Computer Vision ECCV04*, Prague,Czech Republic, May 2004.

[4] S-X. Zheng, J. Li, and Q-F. Sun, "A novel 3d morphing approach for tooth occlusal surface reconstruction," *Computer Aided Design*, vol. 43, no. 3, pp. 293–302, Mar. 2011.

[5] S. Buchaillard, S. Ong, Y. Payan, and K. Foong, "3d statistical models for tooth surface reconstruction," *Computers in Biology and Medicine*, vol. 37, no. 10, pp. 1461–1471, Oct. 2007.

[6] J. Sporring and K. Jensen, "Bayes reconstruction of missing teeth," *Journal of Mathematical Imaging and Vision*, vol. 31, no. 2-3, pp. 245–254, July 2008.

[7] V. Blanz, A. Mehl, T. Vetter, and H. Seidel, "A statistical method for robust 3d surface reconstruction from sparse data," in *In International Symposium on 3D Data Processing, Visualization and Transmission*, 2004, pp. 293–300.

[8] A. Farag, S. Elhabian, A. Abdelrahim, W. Aboelmaaty, A. Farman, and D. Tazman, "Model-based human teeth shape recovery from a single optical image with unknown illumination," in *MICCAI Medical Computer Vision Workshop (MCV)*, 2012.

[9] L. Wolff, S. Nayar, and M. Oren, "Improved diffuse reflection models for computer vision," *International Journal of Computer Vision*, vol. 30, no. 1, pp. 55–71, Oct. 1998.

[10] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 218–233, Feb 2003.

[11] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object," *Journal of the Optical Society of America A*, vol. 18, no. 10, pp. 2448–2459, 2001.

[12] S. Elhabian, H. Rara, and A. Farag, "Towards efficient and compact phenomenological representation of arbitrary bidirectional surface reflectance," in *Proceedings of the British Machine Vision Conference (BMVC)*. 2011, pp. 89.1–89.11, BMVA Press.

[13] Smart Image based Lighting, "http://www.hdrlabs.com/sibl/archive.html," .

[14] X-J. Wang, T. Milner, J. de Boer, Y. Zhang, D. Pashley, and J. Nelson, "Characterization of dentin and enamel by use of optical coherence tomography," *Applied Optics*, vol. 38, no. 10, pp. 2092–2096, Apr 1999.

[15] A. Patel and W. Smith, "3d morphable face models revisited," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 0, pp. 1327–1334, 2009.

[16] T. Cootes and C. Taylor, "Statistical models of appearance for computer vision," Tech. Rep., Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PT, U.K., March 2004.

[17] J. Barron and J. Malik, "Color constancy, intrinsic images, and shape estimation," in *12th European Conference on Computer Vision, Florence, Italy*, October 7-13, 2012, pp. 57–70.

[18] C. Carter, R. Pusateri, D. Chen, A. Ahmed, and A. Farag, "Shape from shading for hybrid surfaces as applied to tooth reconstruction," in *International Conference on Image Processing (ICIP10)*, Hong Kong, 2010, pp. 4049 – 4052.

[19] A. Ahmed and A. Farag, "Shape from shading under various imaging conditions," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, MN*, June 18-23, 2007, pp. X1–X8.