DENTAL HEALTH MONITORING SYSTEM

By

Linjie Tong

Xin Wang

Yichen Shi

Zitai Kong

Project Proposal for ECE 445, Senior Design, Spring 2023

TA: Adeel Ahmed

Project Faculty: Mark Butala | Arne Fliflet | Timothy Lee

Professor: ZuoZhu Liu

9 March 2023

Project No. 23

Contents

1. Introduction

1.1 Problem

In the past era of COVID-19, point-of-care has become a hot topic in society and medicine field due to the high risk of infection when people go out to seek medical treatment. Nowadays, though both China and the world are in a post-epidemic era, point-of-care remains a highly desirable and widely marketable field. One reason is that there is a common reluctance to seek medical treatment; another is that the cost of seeking for medicine treatment is high in both time and price, especially for those living in remoted area. Until now, point-of care has gained well development in lots of medical application scenarios, like antigen detection, blood glucose detection and so on.

With the fast development in people's daily life, Nowadays, when people's living standards are rising rapidly, high-sugar, high-calorie diets are causing more and more oral problems. People are paying higher attention to oral health and have more requirements on it. However, point-of-care for the diagnosis of oral diseases (e.g., caries analysis, orthodontics, oral health) are not yet well equipped. Since the remoted dental diagnosis needs geometric details of teeth, 3D dental graphs are vital in it. 3D teeth images are difficult to obtain from non-medical facilities, in which case 3D reconstruction from 2D images are essential. Currently, point cloud [1,4], voxel [3,6] and mesh [7,10] are state-of-art technologies for 3D reconstruction, but few of them are applied on dental diagnosis and public teeth datasets are rare. Existing 3D reconstruction technologies for dental health diagnosis uses X-ray [9] or CT scan figures [12], which are unlikely to be equipped in patients' home.

Our team plans to design an oral point-of-care system for dental health monitoring, which will be deliverable, affordable, easy to use and high in accuracy, to provide an oral point-of-care tool. This system will be composed of 3D reconstruction software and user-friendly dental sample-taking hardware for teeth only. With this design, people can take dental pictures in their home with their phone, and send them to remote doctors for diagnosis, which will exempt the cost of offline medical treatment.

1.2 Solution

The solution consists of two major parts, the hardware, and the software. The user uses the hardware to record a video of his teeth scanning, send the video to the cloud via the smart phone, and some frames of the video will be selected to be segmented and 3D reconstructed. Then the 3D image of the teeth will be sent back to the user's phone.

The most direct part interacts with the user is the mechanical equipment. Users will use a mechanism to hold the smartphone, and then interact with the mechanism using their hands to change the orientation of the phone camera to the mouth, which finishes the process of sample scanning. The material of the mechanism will be non-deformable to ensure accuracy in the digital detection. The mechanism will contain springs to reduce vibration and maintain stability. The friction between parts in the assembly should be reasonable so that the mechanism can be operated by hands but not too sensitive to force input.

For the software part, our design uses the camera on cell phone and the framework designed by ourselves to obtain video of different views of users' teeth. And then the video will be sliced into a few images that contain all the information about the shape of teeth for 3D reconstruction. After that, a deep model will be used to segment the teeth from the images. And then, the 2D images of segmented teeth will be fed into a computer vision algorithm to get the 3D point cloud of the teeth.

1.3 Visual Aid

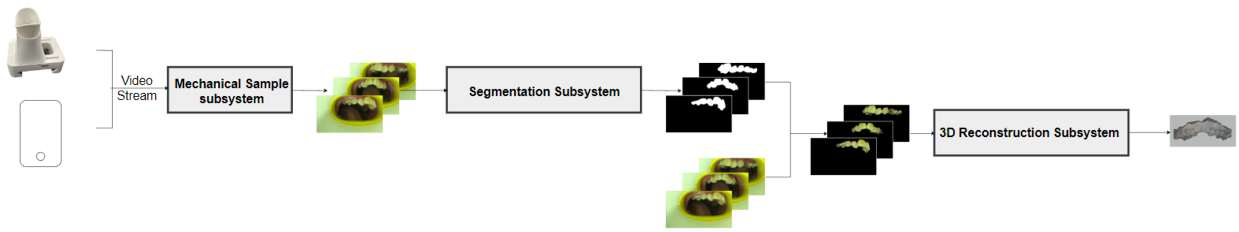

Fig 1. Visual aid of dental health monitoring system

1.4 High-level Requirements List

Requirement 1. For the physical part, the total mass of the system should be around 500g to make the user easy to hold; the force required for user to adjust the system should not be larger than 1 Newton to make the use convenient. The structure must have a safety factor of at least 10, which is calculated from the largest von mises stress in FEA analysis.

Requirement 2. The segmentation algorithm obtains the accuracy better than that directly applied to nature image processing algorithm.

Requirement 3. The 3D reconstruction algorithm obtains the accuracy better than that directly applied.

2 Design

2.1 Block Diagram

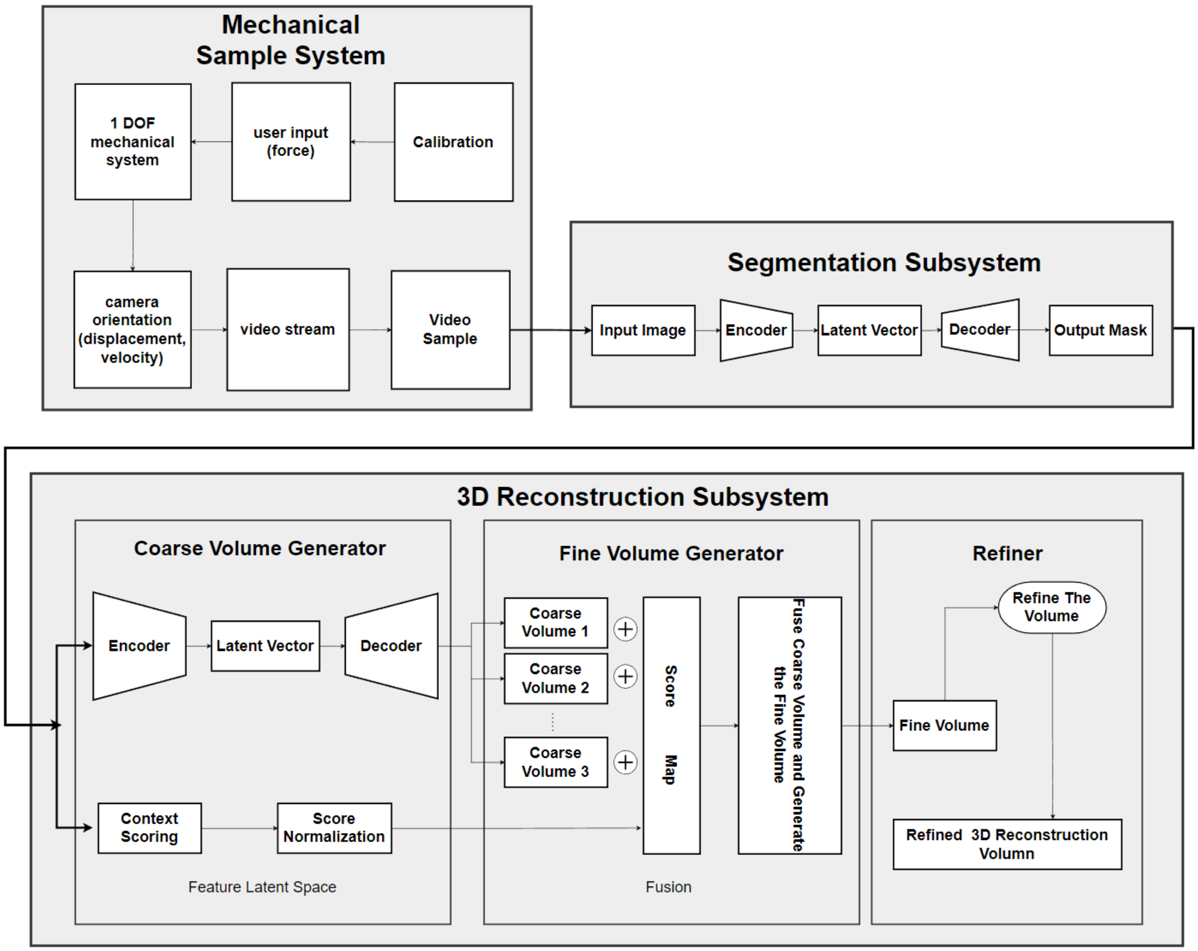

Fig 2. Block diagram of dental health monitoring system

2.1.1 Mechanical Sample Subsystem

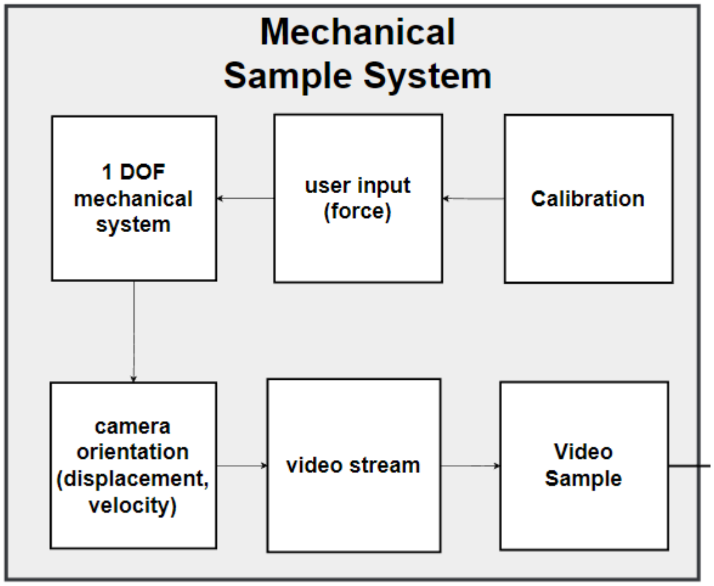

Fig 3. Block diagram of mechanical sample subsystem

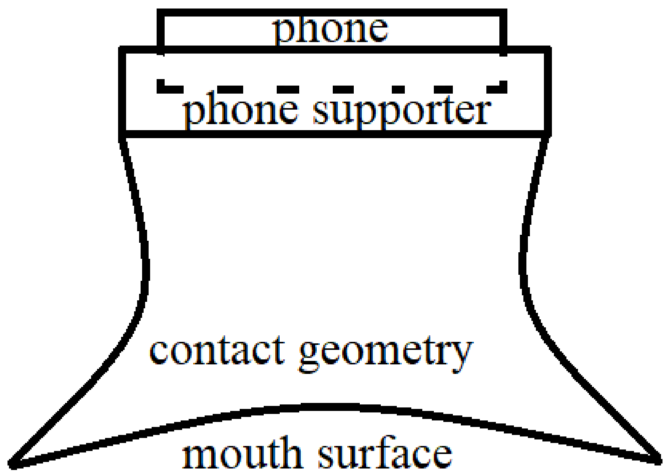

Fig 4. Front view of mechanical sample subsystem

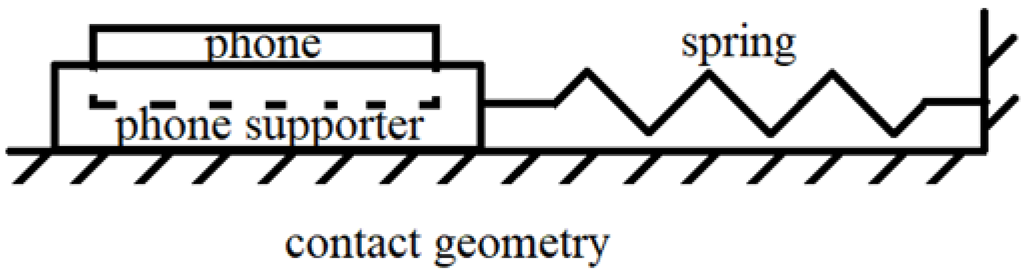

Fig 5. Side view of mechanical sample subsystem

2.1.2 Segmentation Subsystem

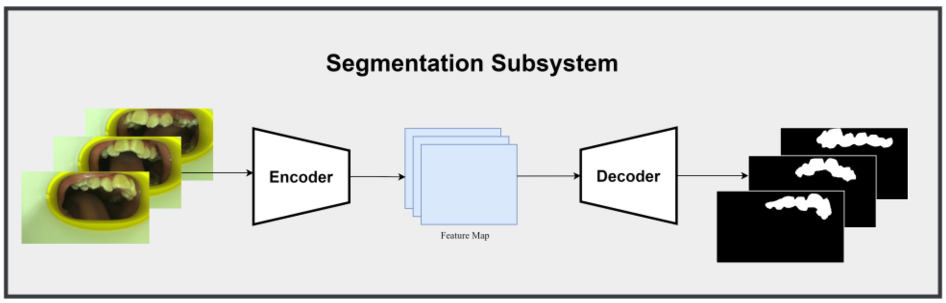

Fig 6. Block diagram of segmentation subsystem

2.1.3 3D Reconstruction Subsystem

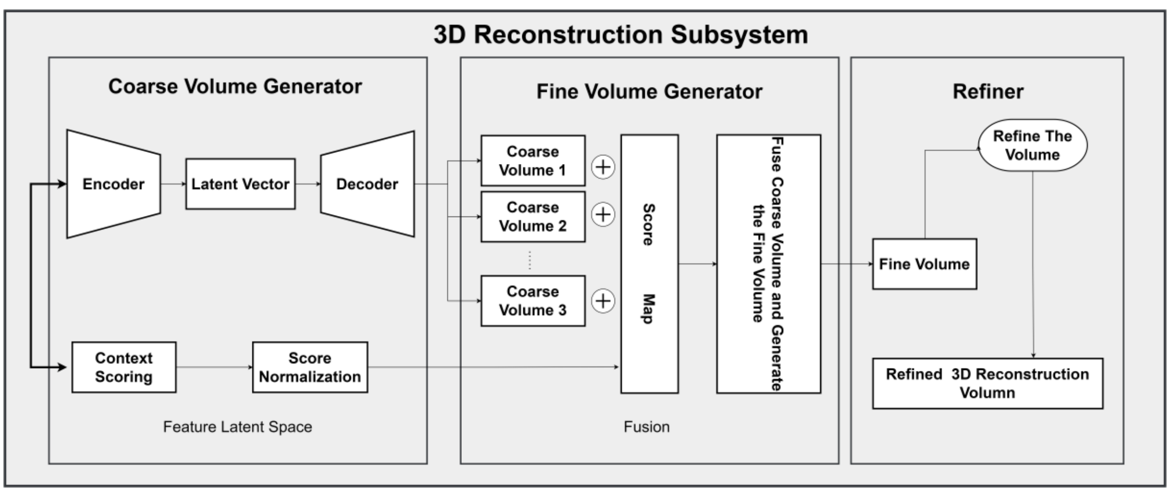

Fig 7. Block diagram of 3D reconstruction subsystem

2.2 Subsystem Overview

2.2.1 Mechanical Sample Subsystem

The functionality of mechanical sample system is to provide picture samples of the teeth from different orientations. The calibration system ensures accurate placement of the mechanical system to the human mouth, which provides an initial reference for taking pictures. After the calibration is done, the user will interact with the mechanical system to move the camera to correct places labeled as tick marks in the system. During the time of the motion, the camera will take pictures. After the pictures are taken, the software will process the videos generated and convert them into picture slides.

2.2.2 Segmentation Subsystem

Segmentation subsystem separates and keeps the teeth foreground from the noisy background. Priors about medical information can improve performance of segmentation [8]. The encoder part receives the raw images uploaded by the mechanical sample system, and encode the images to latent features, which abstracting the key characters, then gives the features to decoder. The decoder part receives the features, return them to pictures with foreground only, and send them as the input of coarse volume generator of 3D reconstruction module.

2.2.3 3D Reconstruction Subsystem

For the 3D Reconstruction Subsystem, we have three components: Coarse Volume Generator, Fine Volume Generator, and the Refiner. The Coarse Volume Generator takes the image processed by the mask generated by the Segmentation Subsystem's output as its input, extract the feature of that by encoding the 2D image into latent space as the low dimension latent vector. And use the decoder generate the coarse volume corresponding to its original 2D image. Because we cannot see all details of the user's mouth by limited input images, we need to use the score map to treat the invisible part as import as the visible part by using weight matrix computed through the original image. The Fine Volume Generator use the coarse volume as its input, process it with the score map to generate the fine volume. It still can have some problem after these procedure mentioned above, so we use a refiner to compare the fine volume result with the ground truth result and define a loss function to refine the model. [11]

2.3 Subsystem Requirements

2.3.1 Mechanical Sample Subsystem

The mechanical system must be 1 degree of freedom, which allows rectilinear motion of the camera in a line segment.

The calibration system must ensure that the orientation of the mechanical system with respect to human mouth is correct, regardless of various shapes of mouth of different people.

When the placement of the camera is at the place indicated by tick marks, the position of the camera must remain unchanged when a limited amount of disturbance is present. When the placement of the camera is not at the place indicated by tick marks, the camera must tend to move to the nearest position indicated by tick marks.

The frame rate of video capture must be 30fps, quality must be 1080p. the capture system must be able to handle the situation when the users' hands are shaking.

2.3.2 Segmentation Subsystem

Input images must be three channel RGB. The size of images is 512 pixels * 512 pixels. The encoder is to extract the feature information which is latent vector. The decoder is to decode the feature information into classification result and turns it into mask. And mask will segment the tooth from noisy background.

2.3.3 3D Reconstruction Subsystem

The Coarse Volume Generator should be able to generate the 3D point cloud model of its origin 2D feature, it does not need to represent all details of user's teeth by predicting the hidden ones, but it is required to capture all information of these teeth shown in the image. In the Fine Volume Generator, this model should be able to process all coarse volumes multiply with the score map generated by the input images in order to treat all teeth equally important to avoid our fine volume output does not have the feature of the hidden details [11]. Finally, the refiner should correct wrongly recovered parts of a 3D volume to make the final refine results as close to the realistic situation as possible.

2.4 Tolerance Analysis

For accuracy requirements, the tolerance should not be too large. The longitudinal dimension of the phone holder is approximately 15cm. to reach 99 percent of the accuracy, the total tolerance should not be greater than 1.5mm.

However, the tolerance should not be too small. Some tolerance will make the motion between surfaces easier, reduce friction between these surfaces, and it will also adapt with dimensional change due to thermal expansion and some other factors.

From the standpoint of fabrication. We are planning to use ABS plastic and acyclic. The 3D printers in ZJUI do not have a high level of precision and the stuff printed are usually coarse, which means the smoothing process is usually needed. When smoothening the printed parts, it is very likely that we shave off too many materials, which lead to an increase in tolerance.

In conclusion, for longitudinal direction, a total tolerance of 1.5mm is appropriate, for the latitudinal direction, a total tolerance of 1 mm is appropriate.

3. Ethics and Safety

3.1 Ethics

Our team pledges to follow the IEEE Ethics guidelines [5]and the ACM Ethics guidelines [2] as closely as possible.

We will do the mechanical design by ourselves. We will not refer to the CAD files on the internet. If it is necessary for us to use some specific design features by other people, we will cite them [2].

We will make the mechanical design based on the user experiences. We will not use cheap but toxic materials for our design. We will not reduce the smoothing processes to reduce our workload, which will make the mechanism more likely to harm the user [2, 5].

During using the data of users, we will hold paramount the safety, health, and welfare of the patients, follow medical rigor and the Nightingale oath, avoid any forms of data discrimination and user information leakage [2, 5].

3.2 Safety

3.2.1 Hardware Safety

In the fabrication process, when cutting the acyclic, because of the use of laser, it is possible to generate flames and result in a fire accident. When assembling other components, some sharp tools will be used to smoothen the contact surfaces, which could cause injuries.

For improper use, if it is seized by children, it could cause problems. The children may eat the loosened nuts or other components not intentionally. For the use of adults, if the ground is slippery and the person fell when using this system, strong impulse created as he fell down to the floor is likely to harm his front teeth, which is extremely dangerous. Also, strong impulse could possibly cause the material to fracture, and the debris containing sharp edges could do harm to the user.

To reduce the possibility for the safety issues to happen, the following methods and concerns will be taken into application. The structure must be smooth and do not have internal or external sharp edges to avoid harm to the user and stress concentrations. The material used must have high elastic modulus and relatively high yield strength to reduce the deformation and avoid yielding when force loads are applied. The machinery must be capable to be operated without lubricants, as lubricants are highly toxic and will create problems for the use of children.

3.2.2 Data Safety

The software part involves relatively few security issues, while data security is the dominant issue that we must concern during the process of data collecting, uploading, and downloading. For the source of data, we will use authorized data from our instructor. When using data, we plan to encrypt user images and information where there is a risk of data leakage, such as uploading and downloading. And we must ensure that it is always stored and running on an encrypted server.

References

[1] A. Yuniarti and N. Suciati, "A Review of Deep Learning Techniques for 3D Reconstruction of 2D Images," *2019 12th International Conference on Information & Communication Technology and System (ICTS)*, Surabaya, Indonesia, 2019, pp. 327-331

[2] ACM Code of Ethics and Professional Conduct, web page. Available at: https://www.acm.org/code-of-ethics. Accessed March 2023.

[3] C. B. Choy, D. Xu, J. Gwak, K. Chen and S. Savarese, "3D-R2N2: {A} Unified Approach for Single and Multi-view 3D Object Reconstruction", *CoRR vol. abs/1604.0*, 2016.

[4] F. Haoqiang, H. Su, and L. J. Guibas. "A point set generation network for 3d object reconstruction from a single image." *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 605-613

[5] IEEE Code of Ethics, web page. Available at: https://www.ieee.org/about/corporate/governance/p7-8.html. Accessed March 2023.

[6] J. Gwak, C. B. Choy, M. Chandraker, A. Garg and S. Savarese, "Weakly supervised 3D reconstruction with adversarial constraint", *Proceedings - 2017 International Conference on 3D Vision 3DV 2017*, pp. 263-272, 2018.

[7] J. Pan, J. Li, X. Han and K. Jia, "Residual MeshNet: Learning to Deform Meshes for Single-View 3D Reconstruction", *2018 International Conference on 3D Vision (3DV)*, pp. 719-727, 2018.

[8] J. Jitesh, et al. "Semask: Semantically masked transformers for semantic segmentation." arXiv preprint arXiv:2112.12782 (2021).

[9] L. Yuan, et al. "X2teeth: 3d teeth reconstruction from a single panoramic radiograph." *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference*, Lima, Peru, 2020, pp. 400-409.

[10] T. Groueix, M. Fisher, V. G. Kim, B. C. Russell and M. Aubry, "A Papier-Mache Approach to Learning 3D Surface Generation", *2018 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 216-224, 2018.

[11] X. Haozhe, et al. "Pix2vox: Context-aware 3d reconstruction from single and multi-view images." *Proceedings of the IEEE/CVF international conference on computer vision*, 2019.

[12] Z. Xinwen, et al. “A method for tooth model reconstruction based on integration of multimodal images”. *Journal of Healthcare Engineering*, 2018.