

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

8-2021

Passive method for 3D reconstruction of human jaw: theory and application.

Mohamad Ghanoum
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>



Part of the [Computer Engineering Commons](#)

Recommended Citation

Ghanoum, Mohamad, "Passive method for 3D reconstruction of human jaw: theory and application." (2021). *Electronic Theses and Dissertations*. Paper 3729.
<https://doi.org/10.18297/etd/3729>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

PASSIVE METHOD FOR 3D RECONSTRUCTION OF HUMAN
JAW: THEORY AND APPLICATION

Mohamad Ghanoum

B.Sc., Electronics and Communications Engineering, University of Kalamoon, 2012

M.Sc., ECE, Speed School of Engineering, University of Louisville, 2015

A Dissertation

Submitted to the Faculty of the

J. B. Speed School of Engineering of the University of Louisville

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering

University of Louisville

Louisville, Kentucky

August, 2021

© Copyright 2020 by Mohamad Ghanoum

All Rights Reserved

PASSIVE METHOD FOR 3D RECONSTRUCTION OF HUMAN
JAW: THEORY AND APPLICATION

by

Mohamad Ghanoum

B.Sc., Electronics and Communications Engineering, University of Kalamoon, 2012

M.Sc., Electrical and Computer Engineering, Speed School of Engineering,
University of Louisville, 2015

A Dissertation Approved on

August 5, 2021

by the following Dissertation Committee:

Professor Aly Farag, Ph.D., Chair/Advisor

Asem Ali, Ph.D., Co-Advisor

Professor John F. Naber, Ph.D.

Professor William C. Scarfe, BDS, FRACDS, MS

Professor Hichem Frigui, Ph.D.

Michael L. McIntyre, Ph.D.

DEDICATION

I would like to dedicate this thesis to my parents, to my beloved wife, and to my son.

ACKNOWLEDGEMENTS

This research was performed by the grace of God, who gave me the knowledge and wit to finish and establish this research. I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost I'd like to thank my advisor, Professor **Aly Farag** for giving me an invaluable opportunity to work on challenging and extremely interesting projects over the past six years. Thank you for introducing me to Computer Vision and giving me the opportunity to explore this captivating field. Your ongoing confidence in me and my work has been an inspiration. It has been an honor to be one of your Master and Ph.D. students.

I would also like to thank my co-advisor, **Dr. Asem Ali**. Without his extraordinary theoretical ideas and computational expertise, this thesis would have been a distant dream. Thanks are due to Professor **John Naber**, Professor **William Scarfe**, Professor **Hichem Frigui**, and Dr **Michael McIntyre** for agreeing to serve on my thesis committee and for sparing their invaluable time reviewing the manuscript.

My colleagues **Mr. Islam Alkabany** and **Mr. Mustafa Izz** at the computer vision and image processing laboratory have enriched my graduate life in many ways and deserve a special mention.

I owe my deepest thanks to my family - my mother **Mufida Katma** and father **Mostafa Ghanoum** who have always stood by me and guided me through

my career, and have pulled me through against impossible odds at times. Words cannot express the gratitude I owe them. I would also like to thank my uncle **Dr. Kheder Kutmah** for his love, support, guidance and unfailing faith in me. Words cannot do justice to his impact in forging my personal and academic outlook. .

I have been blessed with a marriage to one of the most decent and lovable persons God ever created. My beloved wife, **Olama Kalih**, supported me in times of doubt, and her enriched my life beyond measure. She has cheerfully accommodated all manner of long hours and inconvenient absences. Her unconditional love and constant encouragement are what keep me going. Basically she was the person I counted on for support and encouragement. No words can express my love and appreciation for her.

ABSTRACT

PASSIVE METHOD FOR 3D RECONSTRUCTION OF HUMAN JAW: THEORY AND APPLICATION

Mohamad Ghanoum

August 5, 2021

Oral dental applications based on visual data pose various challenges. There are problems with lighting (effect of saliva, tooth dis-colorization, gum texture, and other sources of specularly) and motion (even inevitable slight motions of the upper/lower jaw may lead to errors far beyond the desired tolerance of sub-millimeter accuracy). Nowadays, the dental CAM systems become more compromised and accurate to obtain the geometric data of the jaw from the active sensor (laser scanner). However, they have not met the expectations and the needs of dental professionals in many ways. The probes in these systems are bulky – even their newer versions - and are hard to maneuver. It requires multiple scans to get full coverage of the oral cavity. In addition, the dominant drawback of these systems is the cost.

Stereo-based 3D reconstruction provides the highest accuracy among vision systems of this type. However, the evaluation of it's performance for both accuracy result and the number of 3D points that are reconstructed would be affected by the type of the application and the quality of the data that is been acquired from the object of interest. Therefore, in this study the stereo-based 3D reconstruction will

be evaluated for the dental application. The hand piece of sensors holder would reach to areas inside the oral cavity, the gap between the tooth in the upper jaw and the tooth in the lower jaw in these areas would be very small, in such the stereo algorithms would not be able to reconstruct the tooth in that areas because of the distance between the optical sensors and the object of interest “tooth” as well as the configuration of optical sensors are contradict with the geometric constraint roles of the stereo-based 3D reconstruction. Therefore, the configuration of the optical sensors as well as the number of sensors in the hand piece of sensors holder will be determined based on the morphological of the teeth surfaces.

In addition to the 3D reconstruction, the panoramic view of a complete arch of human teeth will be accomplished as an application of dental imaging. Due to low rate of features on teeth surfaces, the normal of tooth surface is extracted using shape from shading. The extracted surface normals impact many of imprecise values because of the oral environment; hence an algorithm is being formulated to rectify these values and generate normal maps. The normal maps reveals the impacted geometric properties of the images inside an area, boundary, and shape. Furthermore, the unrestricted camera movement problem is investigated. The camera may be moved along the jaw curve with different angles and distances due to handshaking. To overcome this problem, each frame is tested after warping it, and only correct frames are used to generate the panoramic view. The proposed approach outperforms comparing to the state-of-art auto stitching method.

TABLE OF CONTENTS

List of Tables	x
List of Figures	xi
1 INTRODUCTION	1
1.1 Dental impression	1
1.1.1 Digital impressions	2
1.2 Stereo-based 3D reconstruction in dental application	3
1.3 Literature review	5
1.4 Contribution	9
1.5 Dissertation organization	10
2 THE GEOMETRY OF THE HUMAN ORAL CAVITY	11
2.1 Mouth opening	11
2.1.1 Maximum mouth opening	11
2.2 Dental jaw geometry	13
2.2.1 Distance between the maxillary and mandibular during maximum mouth opening	15
2.3 Conclusion	18
3 CAMERA CONFIGURATION DESIGN FOR A DENTAL RECONSTRUCTION	19
3.1 Geometric camera formation	19
3.2 Epipolar geometry	24
3.3 The fundamental matrix	24
3.3.1 Geometric derivation	25
3.3.2 Algebraic derivation	26
3.3.3 Geometric constraints	26
3.4 Theoretical camera configuration design	29
3.5 Practical camera configuration design	31
3.5.1 The configuration of the essential sensors	33
3.5.2 The configuration of sensors for twelve cameras design	35
3.5.3 The configuration of sensors for six cameras design	35
3.5.3.1 Mean shape model	36
3.5.3.2 The modification of the mean shape	37

3.5.4	Comparison between the six and the twelve cameras design . .	39
3.6	Experiment and result	39
3.7	Conclusion	43
4	VISIUAL APPLICATION IN HUMAN ORAL CAVITY ENVIRONMENT	44
4.1	Features Matching	45
4.1.1	Image frames examination	47
4.2	Image stitching using flexible warps	48
4.3	Experiments and results	49
4.4	Conclusion	51
5	CONCLUSION AND FUTURE WORK	53
5.1	Conclusion	53
5.2	Future work	54
	REFERENCES	56
	CURRICULUM VITAE	61

LIST OF TABLES

1.1	Comparison of technical data of intraoral scanning systems	6
2.1	Maxillary and mandibular dental width dimension according to Fig.2.2.	15
3.1	Comparison between the six and the twelve cameras design.	39
3.2	Statistical Analysis on individual and complete arch scenarios in true- ness of 3D reconstruction (μm)	41

LIST OF FIGURES

1.1	Dental impressions	2
1.2	Illustration of the margin line in prepared teeth	3
1.3	Few correct and incorrect matches are obtained by a robust feature-based stereo applied to a right and left camera's images	4
2.1	Diagrammatic representation of the curvilinear pathways of the incisors (C) , the angle of mandibular rotation θ in maximal mouth open and the distance between condyle and incisors (r).	12
2.2	Maxillary and mandibular dental width measurements: a.intercanine distance; b.inter-first molar length; c. inter-second molar and f.arch length.	14
2.3	Illustration of the human jaw anatomical landmarks.	15
2.4	The presentation of linear measurements RH: Ramus height, RW: Ramus width, AGD: Antegonial depth CH: Condylar height.	16
2.5	Measurement in the oral space.	17
2.6	Illustration of the measurment for the smallest space the hand piece of sensors holder would reach in the human oral cavity.	18
3.1	Geometric image formation.	20
3.2	Camera projection	22
3.3	An example of the camera projection types.(a) Orthographic projection and (b) Perspective projection.	23
3.4	Correspondence geometry	25
3.5	The fundamental matrix mapping	27
3.6	Cameras configuration setup.D is the depth of field between the object and the cameras, angle of OA (Optical Axes) is the angle between the optical axes of the cameras and Baseline is the distance between the cameras.	28
3.7	Illustration of the cameras configuration influence on the geometric constraint.	28
3.8	The setting of analysis and geometric interpretation for binocular reconstruction.	29
3.9	The virtual cameras setup in front of a partial of the second molar tooth.	30
3.10	The best cameras poses setting for the tooth.	31

3.11	Image-based 3D reconstruction. Given a set of photographs (left), the goal of image-based 3D reconstruction algorithms is to estimate the most likely 3D shape that explains those photographs	32
3.12	Structure from motion pipeline.	32
3.13	Illustration of the areas that the cameras configuration needs to be tested.	33
3.14	Illustration of cameras configuration of the occlusal surface ; A) the cameras location w.r.t the curve of spee , B) cameras configuration ,C) The quantity result of the 3D reconstruction comparing to the ground-truth ,and D) the quality result of the 3D reconstruction with ground-truth.	33
3.15	The configuration of the twelve cameras.	34
3.16	Mean shape model.	36
3.17	The modification of the Mean shape.	37
3.18	Illustration of the data acquisition for a cross section of a tooth using the six cameras design.	38
3.19	Illustration of the 3D reconstruction of the modification the Mean shape result comparing with ground-truth.	39
3.20	Customized mandibular complete-arch reference model; A)complete arch, B)cracked on tooth no.30, C)crown preparation on tooth no.29, D)full-contour crown preparation on tooth no.26 ,E)mesio-occlusodistal inlay preparation on tooth no. 20 and F) mesio-occlusodistal inlay preparation on tooth no.19	40
3.21	The individual 3D reconstruction tooth for;B) cracked on tooth no.30,C)crown preparation on tooth no.29, D)full-contour crown preparation on tooth no.26 , E) mesio-occlusodistal inlay preparation on tooth no. 20, and F) mesio-occlusodistal inlay preparation on tooth no.19 and . the column 1) shows the ground-truth and 2) the 3D reconstruction result with the color map of the RMS error comparing with ground-truth.	42
3.22	The Complete Arch 3D reconstruction; 1) shows the ground-truth and 2) the 3D reconstruction result with the color map of the RMS error comparing with ground-truth.	42
4.1	The work flow of the proposed images stitching approach	46
4.2	An example of camera paths; blue camera represents the path with a restriction movement while the red camera represents the path in the proposed approach	47
4.3	Examples of frame rejection (a) Translation ,(b) Shearing and (c) Scale with translation issues.	48
4.4	Structural similarity index metric between proposed algorithm results and their ground-truth of ten subjects.	49
4.5	Examples for stitching images of two subjects: (a) Ground truth, (b) Results with the augmenting stage, (c) Results are rigidly aligned with ground-truth, and (d) Results without the augmenting stage.	51

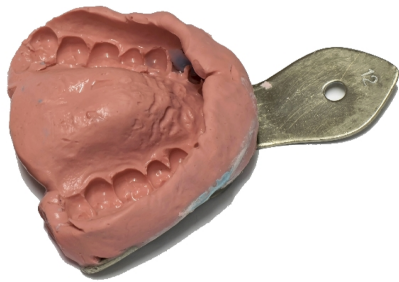
CHAPTER 1: INTRODUCTION

Recovering the structure of a scene from images has long been one of the core interests of computer vision. While a considerable amount of work in this area has been devoted to achieving high-quality reconstructions regardless of run time, other work has focused on the aspect of real-time reconstruction and its applications. In the past two decades, interest has developed around the use of 3D reconstruction for reality capture, gaming, virtual, and augmented reality. These techniques have been used to realize video game assets [1, 2], virtual tours [3] as well as mobile 3D reconstruction apps [4]. Some other areas in which 3D reconstruction can be used are computer graphics and animation [5], medical imaging [6], etc.

1.1 Dental impression

Orthodontic treatment involves the application of force systems to teeth over-time to correct the malocclusion. In order to evaluate tooth movement progress, the orthodontist monitors this movement by means of visual inspection, intraoral measurements, fabrication of plastic models (casts), photographs, and radiographs, a process that is both costly and time-consuming. Obtaining a cast of the jaw is a complex operation for the orthodontist, an unpleasant experience for the patient,

Traditional Impressions



Digital Impressions

Figure 1.1: Dental impressions

and may not provide all the details of the jaw. Some efforts have been devoted to computerized diagnosis in orthodontics. In the past, most of these 3D systems for dental applications found in the literature rely on obtaining an intermediate solid model of the jaw (cast or teeth imprints) and then capturing the 3D information from that model. User interaction is needed in such systems to determine the 3D coordinates of fiducial reference points on a dental cast.

1.1.1 Digital impressions

Nowadays, the optical impression are widely use to get a full 3D model of the human jaw. Optical impressions have several advantages over conventional impressions(Fig.1.1). The most important is the reduction of patient stress and discomfort. In fact, many patients today have anxiety and a strong gag reflex and therefore do not tolerate the conventional impressions; in these cases, using light to substitute trays and materials is an ideal solution. Optical impressions, moreover, are time-



Figure 1.2: Illustration of the margin line in prepared teeth

efficient and can simplify clinical procedures for the dentist, especially for complex impressions. In addition, optical impressions eliminate plaster models, saving time and space, and allow for better communication with the dental technician. Finally, intra-oral system (IOS) improves communication with patients and are therefore a powerful marketing tool for the modern dental clinic. Conversely, the disadvantages of using optical impressions are the difficulty in detecting deep margin lines in prepared teeth and/or in the case of bleeding (Fig.1.2), and the purchasing and managing costs which consider the main drawback for small clinics.

1.2 Stereo-based 3D reconstruction in dental application

CVIP- lab has been involved for the three decades inhuman jaw reconstruction research. The goal is to develop a system for dentistry to go beyond traditional

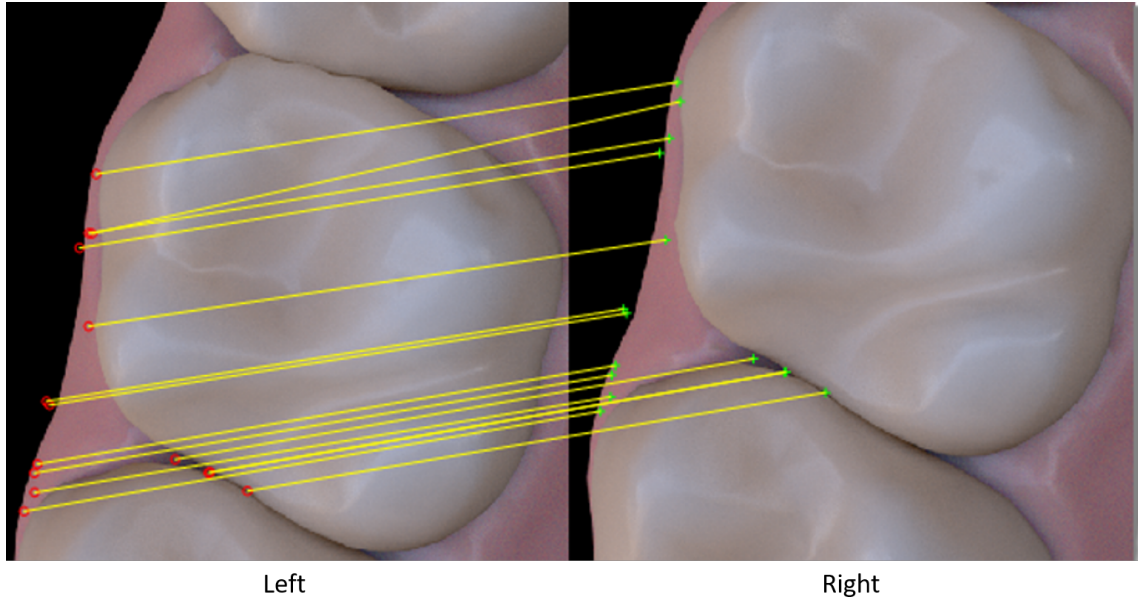


Figure 1.3: Few correct and incorrect matches are obtained by a robust feature-based stereo applied to a right and left camera's images

approaches [7–11]. The system should obtain sequences of calibrated images of the upper/lower jaw using small intra-oral cameras. Then the system will provide accurate 3D reconstruction from the acquired images and will register robustly the 3D models, which will be built from multiple views. This research has immense value in various dental practices including implants, tooth alignment, and craniofacial surgery. The proposed technology has also wide applications in teledentistry, dental education, and training. Several different approaches have been examined, such as the Shape from Shading (SFS) [12], shape from shading with shape priors [9], and appearance-based approach for shape reconstruction [13]. The stereo approach has been by far the most widely used for shape recovery.

However, one of the main problems of stereo is the correspondence problem. Due to the difficult nature of the correspondence problem, several constraints have

been imposed on the stereo-based approach. A large number of stereo techniques rely heavily on assumptions such as the existence of specific features in the images to produce satisfactory results. However, featureless (see Fig.1.3) scenes are hard to reconstruct in stereo because corresponding points between two corresponding images cannot be accurately found. Also, stereo-based approach is difficult to apply to images taken from arbitrary viewpoints. This is because corresponding image points become very hard to find if the images are taken from viewpoints far apart. As a consequence, a stereo-based approach is an inefficient method for reconstruction of objects like human teeth, which, in addition to having many occlusion edges, have no specific geometric features.

This work investigates the obstacles that would make the stereo approach fail in the dental project. The cameras configuration and the depth of field are an important keys used in the stereo approach to obtain a valid 3D points from the scene to be reconstructed. These constraints will be studied and determined the best settings that would achieve the goal.

1.3 Literature review

Three dimensional reconstruction methods are classified into passive and active. Passive methods do not involve interaction with the object, whereas active methods use contact or a projection of some form of energy onto the object. Most of the current technologies, in dental application for 3D reconstruction, work based on active methods.

As an example, iTero [14] system employs a parallel confocal imaging technique [15] for structure recovery, an array of incident red laser light beams, passing through a focusing optics and a probing face, is projected onto the teeth. The beams generate illuminated spots on the structure and the intensity of returning light rays is measured for various positions of the focal plane. The topology of the three-dimensional structure of the teeth is reconstructed on the basis of spot-specific positions yielding a maximum intensity of the reflected light beams. This technique allows iTero capturing all structures and materials in the mouth, without the need to apply any coating to the patient’s teeth. The 3Shape is another company launched a new patient-friendly and high-performance intraoral scanning solution.

Table 1.1: Comparison of technical data of intraoral scanning systems

Intraoral scanner	Company	working principles	Light source	Imaging type	Necessity of coating	In-office milling	Output format	Commercial availability
1 iTero	Cadent LTD (IL)	Parallel confocal microscopy	Red Laser	Multiple images	None	No	Proprietary	Available
2 E4D	D4D Technologies, LLC (US)	Optical coherence tomography and confocal microscopy	Laser	Multiple images	Occasionally	Yes	Proprietary	Available
3 Lava TM C.O.S	3M ESPE (US)	Active wavefront sampling	Pulsating blue light	Video	Yes titanium dioxide	No	Proprietary	Available
4 IOS FastScan	IOS Technologies, INC. (US)	Active triangulation and Schlempping	principle Lase	r 3 images	Yes	Yes	STL	Available
5 directScan	HINT - ELS GMBH (DE)	Active Stereoscopic Vision	Not disclosed	Multiple images	Not disclosed	No	STL	Not Available
6 Trios	3Shape (DK)	Confocal microscopy	Not disclosed	Multiple images	Not disclosed	No	Not disclosed	Available
7 Dental Wings	Dental Wings,inc (CA)	stereoscopic	LED light and projector	Multiple images	Yes	Yes	STL	Available
8 Bluescans	Atron3Ds GMBH(AT)	Active stereoscopic vision	Pulsed UV LED	2 images	None	No	STL	Available
9 CEREC Bluecam	Sirona Dental System GMBH (DE)	Active triangulation and confocal microscopy	Blue light	Multiple images	Yes titanium dioxide	Yes	Proprietary	Available

The Trios system [16] works according to the principle of confocal microscopy, with a fast scanning time. The light source provides an illumination pattern producing a light oscillation on the object. The system produces a variation of the focus plane of the pattern as well, over arrange of focus plane positions, while maintaining a fixed spatial relation of the object. When a time varying pattern is applied, a single sub-scan is actually the collection of a certain number of 2D images, corresponding to different positions of the focus plane and to their respective different time instances of the illumination pattern. In 2006, The Lava C.O.S. system for dental scanner is launched. The Lava C.O.S. camera contains a highly complex op-

tical system comprised of 22 lens systems and 192 blue LED cells. They introduced a method of capturing 3D data based on the principle of active wave front sampling [17] with structured light projection. There are many additional laser dental scanners in the last two decades have been launched. Table 1.1 illustrates differences and characteristics of each one of these scanners.

On the other hand, the passive methods are a non-contact technique for digitally collecting data of the shape of a 3D object. There has been a substantial amount of work to make it easy and feasible for doctors, dentists, and researchers to obtain a 3D model of the person's jaw without ionizing radiation. Computer vision techniques boost computerized diagnosis in dentistry [18]. Unluckily, due to non-trivial occlusions in the mouth interior and the difficult appearance properties of the teeth, camera-based photogrammetric reconstruction is very challenging. Teeth have low dense of texture and they are highly specular and exhibit strong subsurface scattering.

Several works have been conducted in the reconstruction of tooth occlusal surface based on the reflectance properties using Shape from Shading [7, 19–21]. However, the outcome reconstruction of these approaches is a non-metric reconstruction. On the other hand, some previous works developed a statistical model that can recover the 3D shape of the visible teeth from a single photo or from a set of overlapped images. In Shireen etal [13], proposed a relation between the photometric information of a single intraoral photo and the underlying 3D shape by formulated a statistical model in such the effect of the illumination and reflectance are molded into that model whereas the principle component regression is used for

3D estimation. Yet, the reconstruction result of the joint tooth-gum region is not accurate since the information in these areas are estimated. Wu, Chenglei, in ‘ [22] , also proposed a model-based approach for reconstructing a 3D teeth model from a set of overlapped images. However, these images have not been captured intraorally, which means that the information in the occlusal sextant parts are missing, and the result of these parts would be estimated from the trained model. Generally, these type of approaches are not clinically accepted since the results are based on a prior model, which is consider as an estimated outcome.

Passive triangulation [23] is based on processing of two stereo images, obtained from two calibrated cameras, whose relative translation and rotation are known. This information is needed in order to identify points with corresponding features on the two images and to apply triangulation, with respect to the same corresponding points on the epipolar line. Stereo-based reconstruction uses the epipolar geometry. There is another technique among the passive method called Structure from Motion(SfM) [24–28]. SfM is the process of reconstructing 3D structure from its projections into series of images taken from different viewpoints. Incremental SfM is a sequential processing pipeline with an iterative reconstruction component. It commonly starts with feature extraction and matching, followed by geometric verification. The resulting scene graph serves as the foundation for the reconstruction stage, which seeds the model with a carefully selected two-view reconstruction, before incrementally registering new images, triangulating scene points, filtering outliers, and refining the reconstruction using bundle adjustment.

1.4 Contribution

Intra-oral scanning technology is a very fast-growing field in dentistry since it responds to the need for an accurate three-dimensional mapping of the mouth. Despite the high 3D reconstruction accuracy of the active scanners, they are very expensive which is considered as the main drawback, the head of the scanner is very bulky, and it needs a dentist professional in many ways. The geometric extraction of human teeth using passive methods is studied in this work. The main advantage of using the passive method is the low cost of these systems, made of few and cheap components; furthermore, also the working principle is simple, the same as in the human eye.

The sensor planning will be designed based on the morphological teeth surfaces. Due to the complexity and variety of teeth surfaces, the hand piece of sensors holder would need more than two cameras to scan a complete full arch, in such a different kind of areas inside the oral cavity would be scanned with one-time data acquired which would be a distinct feature from the active scanners. The number of cameras needed would be determined by conducting a different experiments on a certain teeth surfaces. In addition to 3D reconstruction application, this work proposed an approach to produces a panoramic view of a complete arch of jaw from a sequence of images. Normal maps were used to reveal the impacted features of the images. Also, normal maps are used to estimate connected planner regions to obtain the homographies for stitching. In this work, the restriction of the camera movement has been released. This constraint impractical due to the handshaking of the intra-

oral camera inside the oral cavity even for experts. This problem is solved by using a Boolean stage to test the homography matrix between every two consecutive images and decide whether these images stitching could proceed.

1.5 Dissertation organization

This document is divided into three chapters. The chapter two covers the geometry of the oral cavity as well as the dimensions of the complete arch jaw . The chapter three presents the proposed the sensor planning configuration using passive method based on the geometry of the oral cavity. The chapter three shows how the stitch a consectuive images acquired from the human teeth to construct a panoramic view of a complete arch.

CHAPTER 2: THE GEOMETRY OF THE HUMAN ORAL CAVITY

The human oral cavity is considered as a complex environment for visual application due to the confined space, saliva, and lack of texture. The width, length, and depth of dental arches have had considerable implications in orthodontic diagnosis and treatment planning in modern dentistry based on prevention and early diagnosis of oral disease. From a computer vision perspective, the scene and the object need to be studied before any further step in the application. Therefore, in this chapter, the oral space dimensions will be studied to reveal the available space that can be used in visual applications such as probe design

2.1 Mouth opening

2.1.1 Maximum mouth opening

Much research has been carried out to document maximum mouth opening (MMO) in several populations [29–31]. MMO can be expressed either as interincisal distance or as corrected interincisal distance, which is determined by adding the amount of vertical overlap between the upper and lower incisors to the incisal

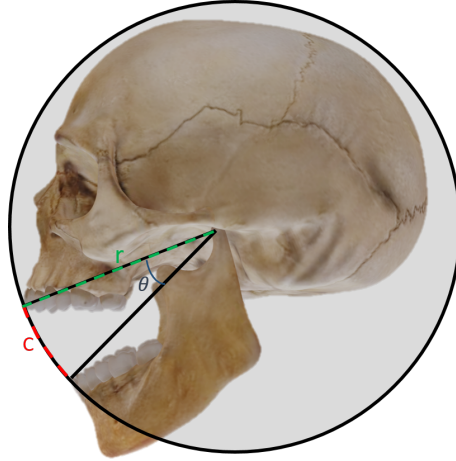


Figure 2.1: Diagrammatic representation of the curvilinear pathways of the incisors (C) , the angle of mandibular rotation θ in maximal mouth open and the distance between condyle and incisors (r).

distance [32]. Rieder [29] reported that men generally have a wider mouth opening than women: in that study, 83% of men had a mouth opening of 40–60 *mm*, whereas 87 % of women had a mouth opening of 35–55 *mm*, see arc (C) in Fig.2.1.

The mouth angle has been reported in some research as a part of their study of the MMO [33, 34]. The angle, θ in Fig.2.1, of mandibular rotation during the maximal mouth opening was measured, in [34], as the angular change of the dashed line connecting the hinge axis and incision from maximum intercuspation to maximum opening. In that study, 93 cases distributed to 30 men and 63 women with ages between 15-75 years old were enrolled. The average of the angle of mandibular during the maximum mouth open was 26 ± 7.0 degree.

Assuming that the curvilinear pathways (C) is an arc , (r) is the radius of this circle and θ is the central angle of this arc in radian. The (r) can be calculated as

following:

$$r = \frac{C}{\theta} = \frac{\text{length of subtended arc}}{\text{the central angle}} \quad (2.1)$$

Where the curvilinear interval is 35-60 *mm* and the average of central angle is 0.45 *radian*, thus the radius using Eq.2.1 is 77.7-134 *mm*

2.2 Dental jaw geometry

Dental Jaw dimensions, including width, length, and height, are important values for the diagnosis, treatment, planning, and treatment outcomes concerning patients who are seeking orthodontic treatment in all age groups. Different ethnic groups and populations display variable dental Jaw measurements and characteristics [35]. The characteristics of the jaw (see Fig. 2.2) can be determined based on the following parameters:

1. Inter-canine distance (a): from the cusp tip of the canine to the contralateral.
2. Inter-first molar distance (b): the distance measured from the buccal groove along the occlusal surface of the first molar to the contralateral first molar.
3. Inter-second molar distance (c): the distance measured from the buccal groove along the occlusal surface of the second molar to the contralateral second molar.
4. Arch length (f): from the center of the palatal incisal papilla to the middle point on a line drawn between the right and left second molars.

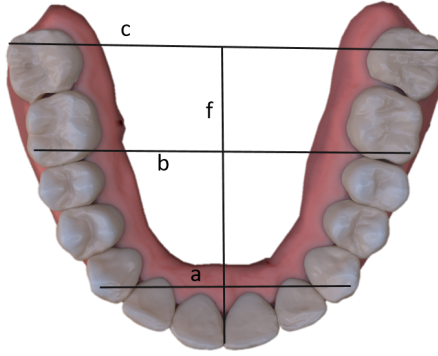


Figure 2.2: Maxillary and mandibular dental width measurements: a.intercanine distance; b.inter-first molar length; c. inter-second molar and f.arch length.

In order to get these dimensions, our dataset contains 432 three dimensional *CT* models (224 male and 208 female) of human jaws are measured. This dataset can be divided into four categories based on a number of teeth per jaw and the mandible or maxilla jaw type:

1. 90 jaws for (mandible-14 teeth)
2. 259 jaws for (maxilla-14 teeth)
3. 39 jaws for (mandible-12 teeth)
4. 44 jaws for (maxilla -12 teeth)

To accurately measure these parameters, we extract a set of anatomical landmarks. Typically, a landmark represents a distinguishable point which is present in most of the images under consideration, for example, the location of the central grooves of each tooth. Fig. 2.3 illustrates the location of 72 landmark points for a jaw containing 14 teeth.

Table 2.1: Maxillary and mandibular dental width dimension according to Fig.2.2.

	a(mm)	b(mm)	c(mm)	f(mm)
Lower-14	26.77±1.99	51.19±1.32	55.90±2.78	51.23±1.28
Upper-14	34.46±2.52	55.36±1.89	59.09±2.67	55.37±1.90
Lower-12	28.17±1.70	45.42±1.61	53.63±2.57	45.40±1.63
Upper-12	33.19±2.92	46.06±3.16	51.69±7.74	46.00±3.17

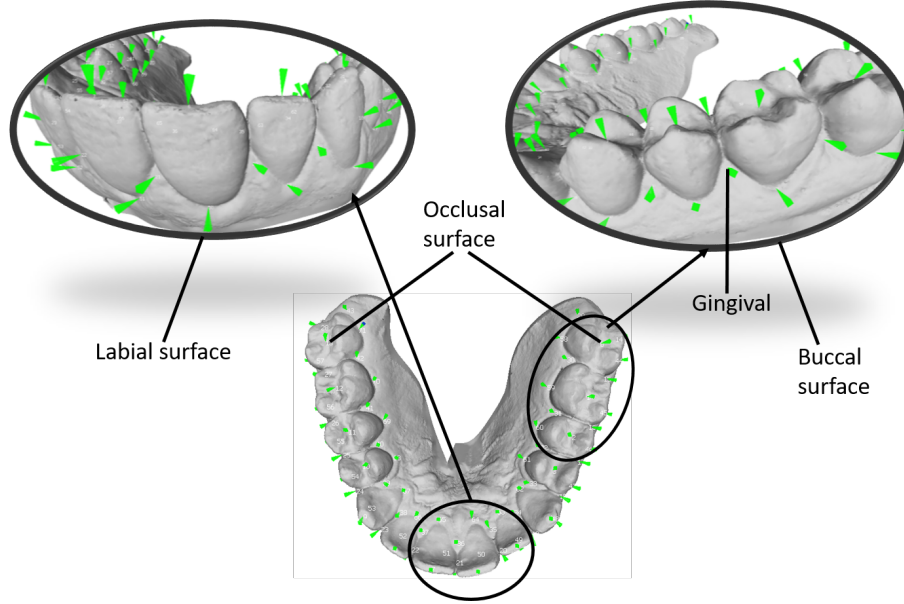


Figure 2.3: Illustration of the human jaw anatomical landmarks.

In Table 2.1, the dimensions of the maxillary and mandibular jaw have been identified. These dimensions would help us to get more measurements of the oral space like the distance between the second molars of the maxillary and mandibular during the maximum mouth opening.

2.2.1 Distance between the maxillary and mandibular during maximum mouth opening

This study aims to define the smallest distance in the oral space that the sensor would reach it during data acquisition. This distance is between the second molar

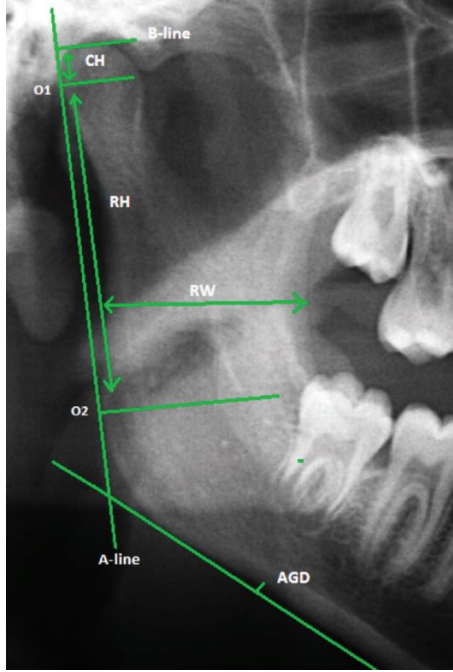


Figure 2.4: The presentation of linear measurements RH: Ramus height, RW: Ramus width, AGD: Antegonial depth CH: Condylar height.

in the occlusal surface of the mandibular jaw and the contralateral in the maxillary jaw (see L in Fig.2.5). In [36], a study was made to see the impact of the age, gender, and dental status on the remodeling of anantegonial and gonial, condylar, and ramus regions (see Fig.2.4). They reported that the average of ramus height for 910 patients is 49.23 ± 7.22 mm. The straight line between the incisor of the mandibular and condyle, r is calculated using Eq.2.1. So, the angle α in Fig.2.5 can be calculated as follows:

$$\alpha = \sin\left(\frac{RH}{r}\right) = 21 - 39 \text{ degree} \quad (2.2)$$

The radius of the circle that passing the second molars in both jaws (upper and lower) is defined as d (Fig.2.5). The f in table 2.1 and α are known, so the

radius of that circle can be approximated as following:

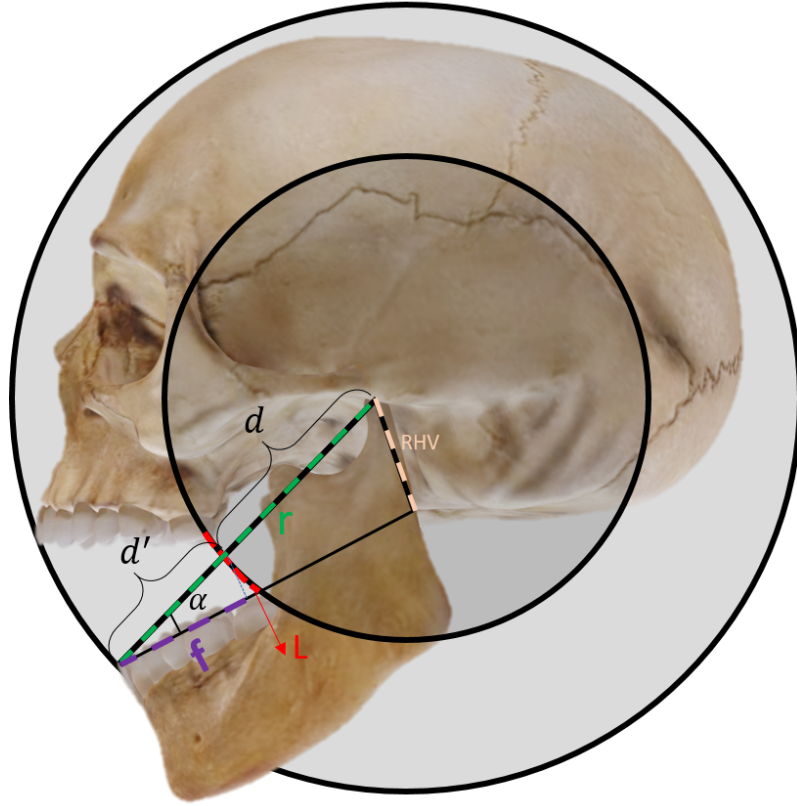


Figure 2.5: Measurement in the oral space.

$$d' = \cos(\alpha) \times f \implies d = r - d' \quad (2.3)$$

Thus, L in Fig.2.5 can be approximated using Eq.2.1:

$$d = \frac{L}{R} \implies L = d \times R \implies L = 12 - 26.6 \quad mm \quad (2.4)$$

The hand piece of sensors holder as well as the sensors configuration will be designed based on the available space between the sensors and the teeth surfaces

(Fig.2.6). Since the smallest area “ L ” that the hand piece would reach inside the mouth for 3D reconstruction is calculated and the thickness of the sensors holder is known based on the sensor dimension, the space availability between the sensor and teeth surface is estimated to be approximately $5mm$.

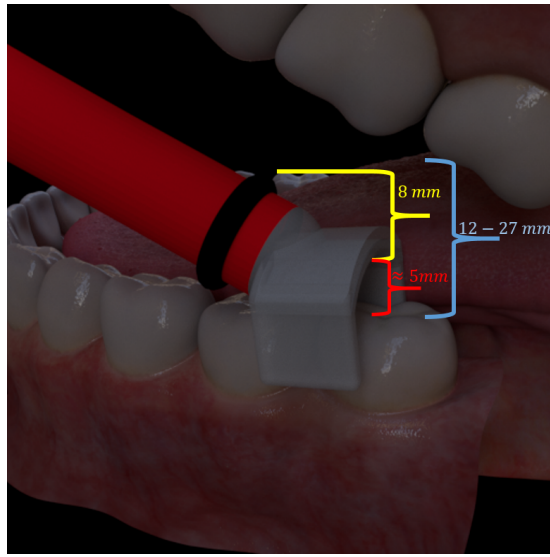


Figure 2.6: Illustration of the measurement for the smallest space the hand piece of sensors holder would reach in the human oral cavity.

2.3 Conclusion

In this chapter, the geometry of the oral cavity was studied to reveal the dimension of areas that the hand piece sensors holder would reach during the data acquisition. These dimensions are very important in the sensor planning design for 3D reconstruction of the human teeth in such the configuration of sensors would be designed according to these dimensions.

CHAPTER 3: CAMERA CONFIGURATION DESIGN FOR A DENTAL RECONSTRUCTION

Visual 3D reconstruction is an essential technique in computer vision which restores the 3D model of the scene from a single image or multi-view images. The hand piece of sensors holder should capture images with good quality and process them in real time. So sensors should be carefully configured such that teeth can be clearly appear in field of view of cameras. Thus, this chapter illustrates the stratgy which is used to determine the best number of sensors and their geometry to get full and accurate 3D reconstruction of the human teeth.

3.1 Geometric camera formation

In order to use images to infer information about the 3D world, one needs to study how pixel brightness in the image is related to the physical world. To accomplish such a task, two questions needed to be addressed; (1) where some point in 3D will appear in the 2D image, and (2) how bright this image point will be. The former question is related to the camera/viewer properties, *i.e.* geometric image formation(Fig. 3.1), while the latter one is governed by surface properties and illumination conditions, *i.e.* photometric image formation.

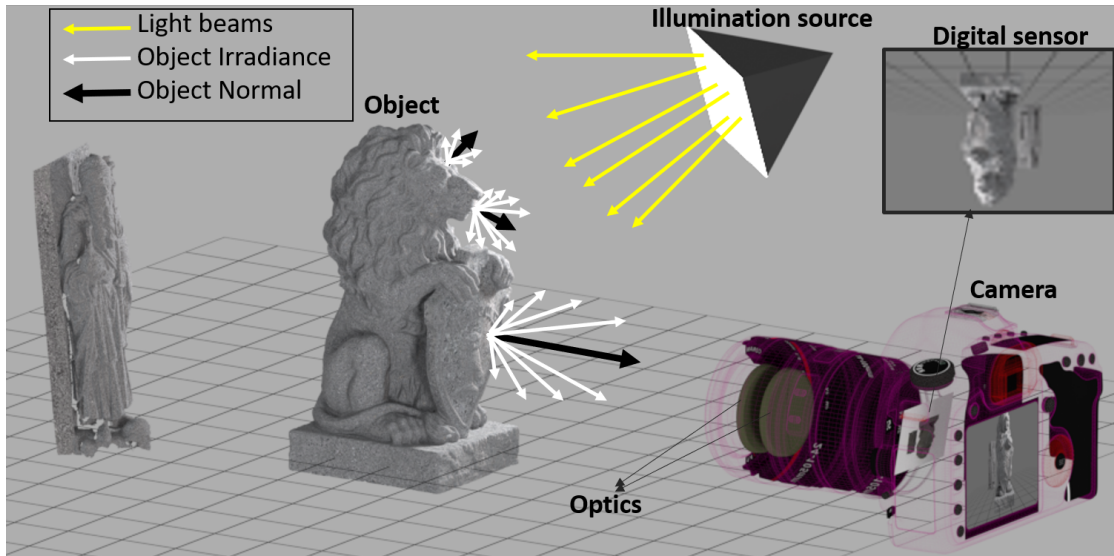


Figure 3.1: Geometric image formation.

Consider a directional light source emitting light rays in all directions (fig. 3.1), where one light ray is directed towards a surface point. When a light ray hits a surface point, a portion of its energy is absorbed by the object, while the remaining is reflected from the surface. Such portion is a function of surface reflectance properties. The brightness of the projected point in the image plane is affected by the reflected light rays in the direction of the camera. The camera lens takes all the light rays bouncing around and uses optics to redirect them to a single point and then projects them on a digital camera sensor or piece of film. The processor of the camera captures the information from the sensor camera and produces a sharp image. If rays of light does not meet at the right point, the image will look blurry.

The focal Length is a principle parameter in the image formation; it determines the angle of view. Shorter focal length, the more subject fits in the frame. There are two types of projections:

- 1) Orthographic projection: is a form of a parallel projection in which all the projection lines are orthogonal to the projection plane (image sensor),

$$x = f \frac{X}{Z}; y = f \frac{Y}{Z} \quad (3.1)$$

Where f is the focal length, (x, y) are the pixel coordinate in image plane and (X, Y, Z) are the point coordinate in world space.

2. Perspective projection: The image of a projective camera can be described in terms of its 11 intrinsic and extrinsic parameters (see fig.3.2). The intrinsic parameters affect how the image is seen on the image plane once it has entered the camera. These are unique to individual cameras and include:

1. Focal Length - The distance from the camera's center to the image plane.

In most digital and film cameras this is the distance from the rear-most lens element to the actual sensor/film.

2. Principal Point - The origin of the image plane in relation to the center of the image plane. For convenience, typically the lower left corner of the image is considered the origin.

3. Skew Coefficients - The pixels in a CCD sensor may not be perfectly square, resulting in a small distortion in the X or Y directions. The skew coefficient is the number of pixels per unit length in each direction on the CCD sensor.

4. Distortion - The curved nature of a lens results in radial and tangential

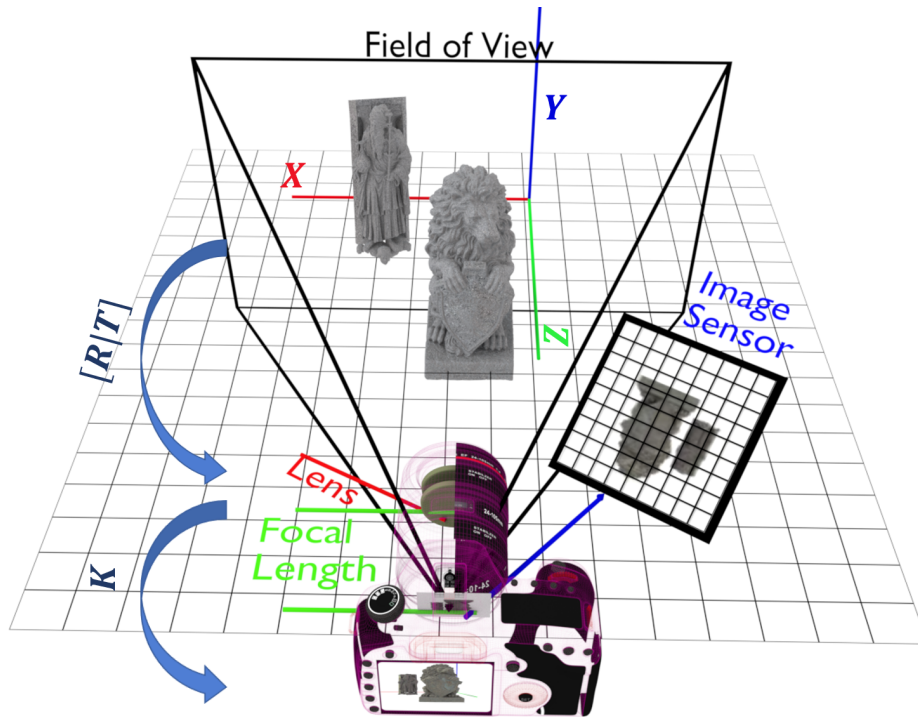


Figure 3.2: Camera projection

distortion. These can be quantified by a number of distortion coefficients.

Together these parameters form the camera calibration matrix K . The extrinsic parameters of a camera, on the other hand, are not unique to any given camera. These relate how a camera's orientation and position within world coordinates:

1. Rotation - Roll, Pitch, and Yaw of the camera
2. Translation - Position of the camera in regards to the World origin.

The intrinsic and the extrinsic parameters define a camera's projection matrix P :

$$P = K[R|T] \tag{3.2}$$

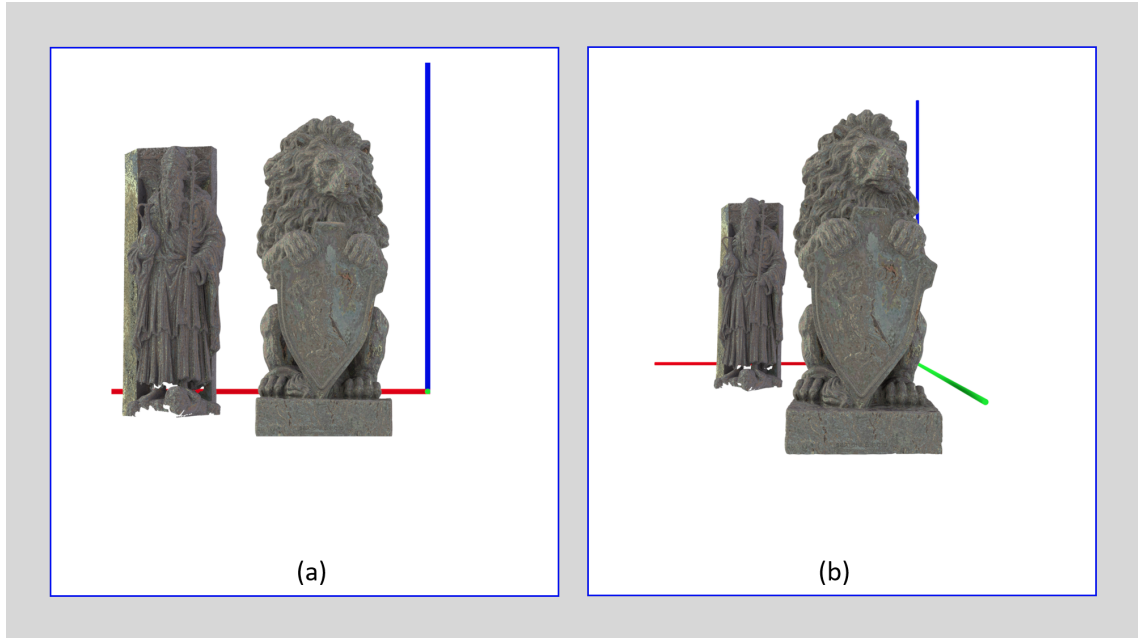


Figure 3.3: An example of the camera projection types.(a) Orthographic projection and (b) Perspective projection.

Fig.3.3 illustrates the difference between two projections. As shown in fig.3.1, there is a distance in Z direction between the statues; the perspective clarify this distance while the orthographic show that they are in the same distance from the camera. Two main characteristics of perspective are vanishing points and perspective foreshortening. Due to foreshortening, object and lengths appear smaller from the center of projection. More we increase the distance from the center of projection, smaller will be the object appear. Thus, we can say the prespective is represents the functionality the human eyes while we consider the orthographic projection is a special case.

3.2 Epipolar geometry

The epipolar geometry is the intrinsic projective geometry between two views. It is independent of scene structure, and only depends on the camera's internal parameters and relative. Zisserman in [37] described the fundamental of epipolar geometry and how it helps to find the correspondences of multi view images. Suppose a point X in 3-space is imaged in two views, at x in the first, and x' in the second. In fig.3.4 the rays back-projected from x and x' intersect at X are co-planar lying in S . The plane S is determined by the baseline and the ray defined by x . Also, the ray corresponding to the (unknown) point x' lies in S , hence the point x' lies on the line of intersection l' of S with the second image plane. This line l' is the image in the second view of the ray back-projected from x . In stereo correspondence algorithm the search for the point corresponding to x is only restricted to the line l' .

3.3 The fundamental matrix

The fundamental matrix is the mathematical representation of epipolar geometry. Given a pair of images, as shown in 3.4, each point x in one image, there exists a corresponding epipolar line l' in the other image. Any point x' in the second image matching the point x must lie on the epipolar line l' . The epipolar line is the projection in the second image of the ray from the point x through the camera

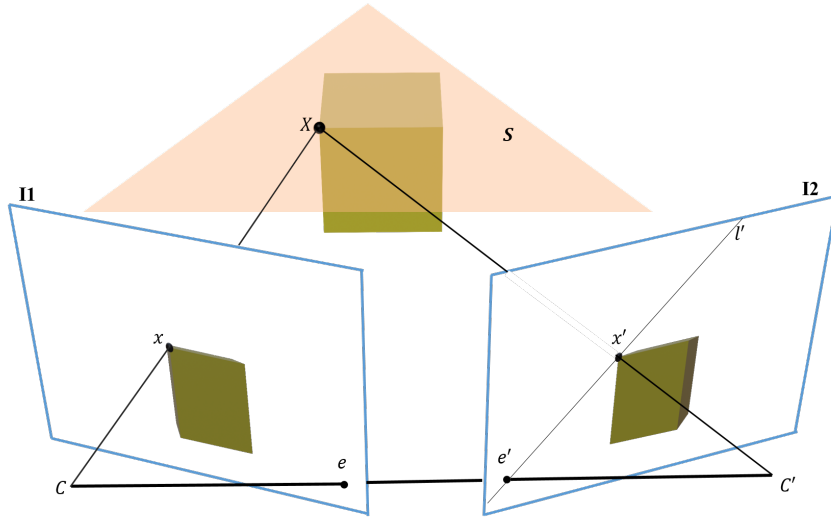


Figure 3.4: Correspondence geometry

centre C of the first camera. Thus, there is a map

$$x \implies l' \tag{3.3}$$

3.3.1 Geometric derivation

The mapping in Fig (3.5) from a point in one image to a corresponding epipolar line in the other image can be divided into two steps:

1. The point x is mapped to point in the other image lying on the line l' : points x_i in the first image and the corresponding points x'_i in the second image are projectively equivalent, since they are each projectively equivalent to the planar point set X_i . Thus there is a 2D homography H_S mapping each x_i to x'_i .
2. The line l' is obtained as the line joining x' to the epipole e' : the epipolar line l'

passing through x' and the epipole e' can be written as $l' = x' \times e'$. Since we can write $x' = H_S \times x$, we can write l' as:

$$l' = e' \times H_S \times x = xF \tag{3.4}$$

3.3.2 Algebraic derivation

The fundamental matrix can also be derived algebraically in terms of the two camera projection matrices, P and P' . The fundamental matrix can be written:

$$F = e' \times P' \times P^+ \tag{3.5}$$

Where P^+ is the pseudo inverse of P and the camera baseline intersects the second image at the epipole point e' . Notice that the homography matrix can be formed $H_S = P' \times P^+$ in terms of the two camera matrices.

3.3.3 Geometric constraints

Two perspective cameras observing the same points must satisfy the epipolar constraint:

$$x'Fx = 0 \tag{3.6}$$

Since x' lies on the epipolar line $l' = Fx$ corresponding to the point x , that gives $0 = l' \times x'^T = x'^T Fx$. The human oral cavity environment puts a restriction on

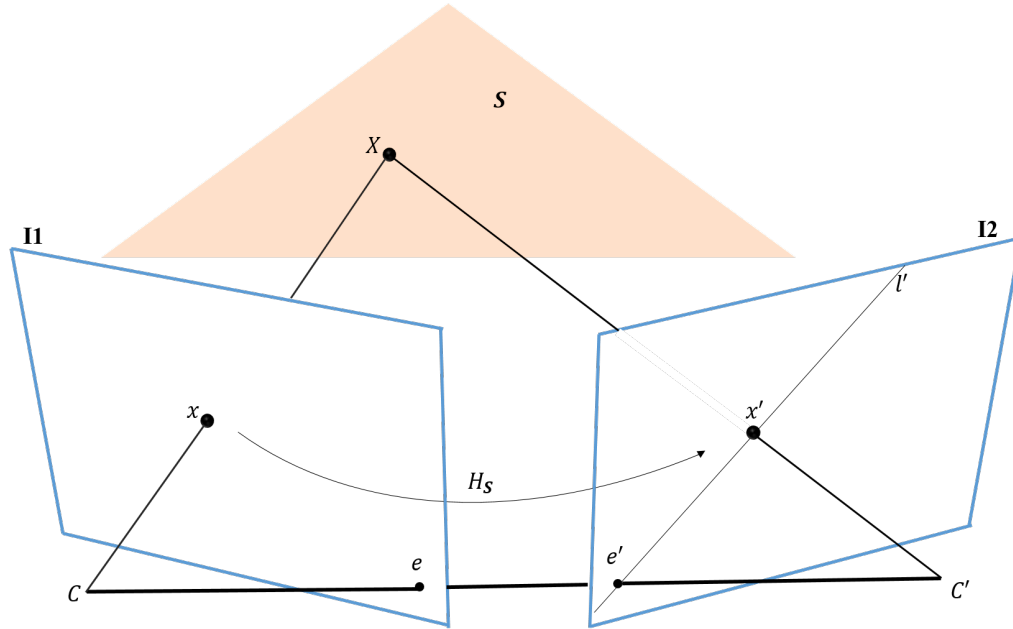


Figure 3.5: The fundamental matrix mapping

the cameras configuration to satisfy the epipolar constraint, in such the two cameras parameters need to be tuned are (see Fig.3.6):

1. The baseline between cameras.
2. The angle between the optical axis of the cameras
3. The depth of field roles in the geometric constraint.

Figure 3.7, shows two images, which are captured with two cameras for a second molar tooth (an artificial texture applied on the tooth to illustrate the idea), The system parameters are chosen to be the baseline $5mm$, depth is $8mm$ and the angle between the optical axes is 12° . For the four pairs of contours in I1 and I2, each pair are centered at corresponding points and have the same color, size and shape. However, each pair doesn't enclose the exact corresponding points in images I1 and I2. This occlusion happens due to the size of baseline, the angle between optical axes

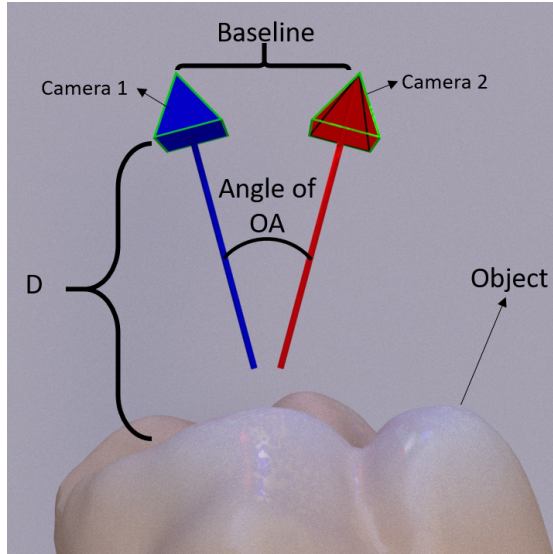


Figure 3.6: Cameras configuration setup. D is the depth of field between the object and the cameras, angle of OA (Optical Axes) is the angle between the optical axes of the cameras and Baseline is the distance between the cameras.

(OA) and the depth of field (w.r.t to the tooth size). This lack of correspondences affects the reconstruction process.

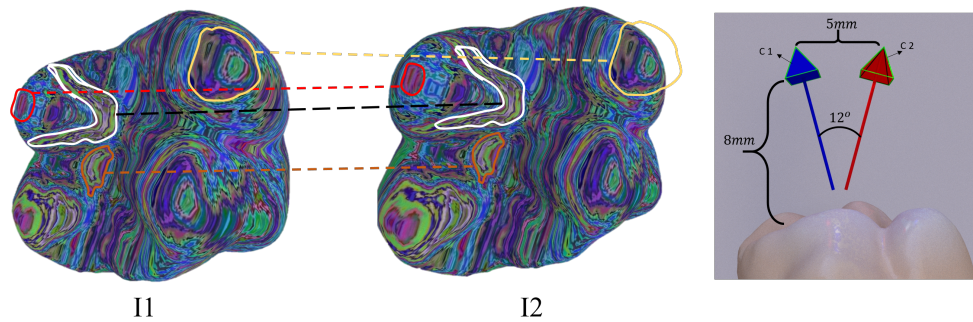


Figure 3.7: Illustration of the cameras configuration influence on the geometric constraint.

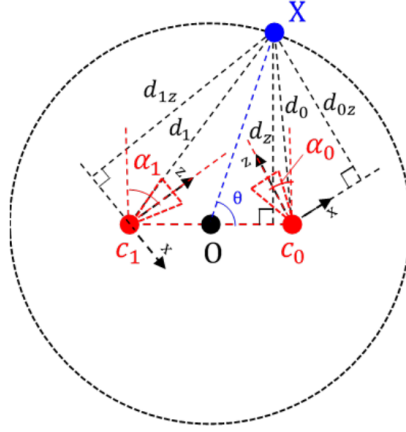


Figure 3.8: The setting of analysis and geometric interpretation for binocular reconstruction.

3.4 Theoretical camera configuration design

Pinhole camera is a simple system that can record an image of an object or scene in the 3D world [37]. The projection from the world coordinate of a 3D point X_j to the camera coordinate on a certain 2D pixel $x_i^{(j)}$ in the i -th image can be explained by the following:

$$x_i^{(j)} = P_i X_j = K_i [R_i | t_i] X_j \quad (3.7)$$

Where P_i is the projection matrix of i -th camera. In [38], a study proposed an approach to measure the optimal camera configuration for a certain application. The likelihood function of the projection equation Eq.3.7 can be described by Gaussian distribution as :

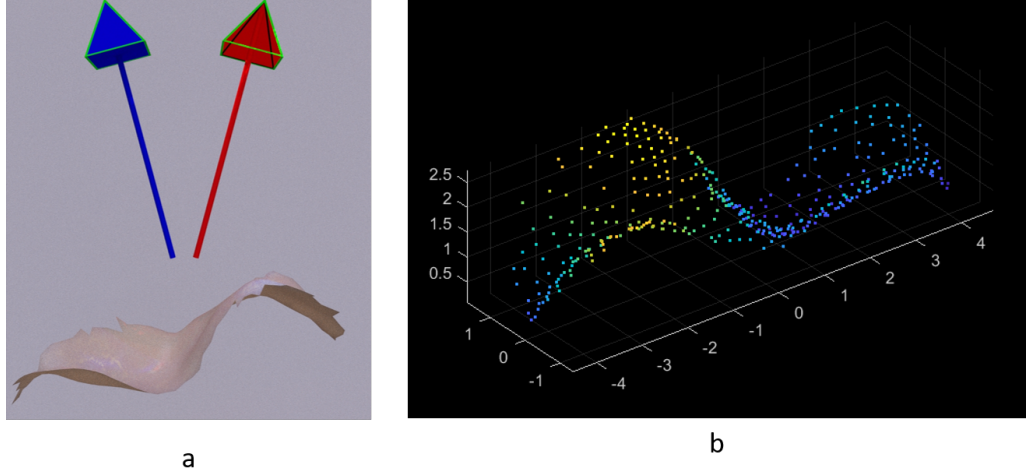


Figure 3.9: The virtual cameras setup in front of a partial of the second molar tooth.

$$p(x; \theta) \propto \exp\left(\sum_{\substack{i \in \mathcal{N} \\ j \in \mathcal{M}}} -\frac{1}{2\sigma_i^2} \|x_i^{(j)} - K_i(R_i X_j + t_i)\|^2\right) \quad (3.8)$$

Where x is the vector consists of all the projection pixels, i.e., $\theta = (X_i^T)_{i \in \mathcal{N}}^T$, $x = (x_i^{(j)T})_{i \in \mathcal{N}, j \in \mathcal{M}}^T$, \mathcal{M} and \mathcal{N} denote the index sets of cameras and 3D points, respectively. Figure.3.8, shows the two cameras are located at c_0 , c_1 with the yaw angles α_0 , α_1 . The 3D point X to reconstruct is on the circle of radius r around the midpoint O of c_0 and c_1 . To maximize the reconstruction accuracy, the optimal poses of all cameras have to be adjusted according to:

$$\arg \min_{R_i, t_i, i \in \mathcal{N}} \text{tr}\{I^{-1}(\theta)\} = \frac{d_{0z}^2 d_{1z}^2}{d_{1z}^2 + d_{0z}^2} + \frac{d_0^2 d_{1z}^1 + d_1^2 d_{0z}^1}{b^2 d_z^2} \quad (3.9)$$

Where $\{I^{-1}(\theta)\}$ is the expectation of the likelihood function, b is the baseline, and R_i, t_i is the rotation and translation respectively of cameras c_0 and c_1 in respect to the midpoint O . The second molar tooth is used as a testing object for this experimental to obtain the best camera's poses for 3D reconstruction. Assuming

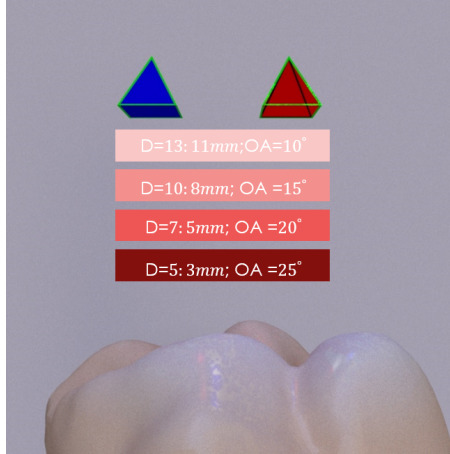


Figure 3.10: The best cameras poses setting for the tooth.

there are two virtual cameras in front of a partial $3D$ points of the tooth (Fig.3.9.b), and the configuration of cameras in respect to each other as well as the depth of field are changed to obtain the best camera poses for 3D reconstruction, where the range of the baseline is 1 to $4mm$, the angle of OA is 10 to 40° , and the depth of field is 13 to $3mm$. The result of this experiment shows; a smaller baseline gives more accurate 3D reconstruction and the angle of OA can be determined based on the depth of field, *i.e.* a smaller depth of field needs bigger OA angle (see Fig.3.10).

3.5 Practical camera configuration design

Structure from motion [39] is a method among the passive technique for 3D reconstruction. As shown in Fig.3.11, given a set of images acquired from different observation points, it recovers the pose of the camera for each input image and a three-dimensional reconstruction of the scene in form of a sparse point cloud.

The complete flow of incremental SFM pipeline operations is shown in Fig.3.12. In particular, incremental SFM is a sequential pipeline that consists of a first phase

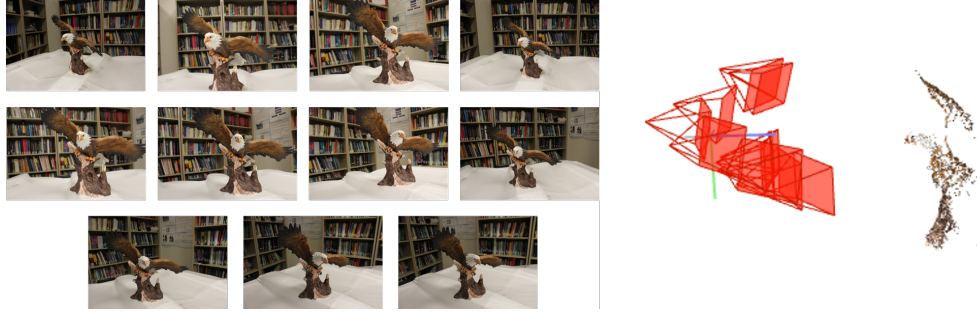


Figure 3.11: Image-based 3D reconstruction. Given a set of photographs (left), the goal of image-based 3D reconstruction algorithms is to estimate the most likely 3D shape that explains those photographs .

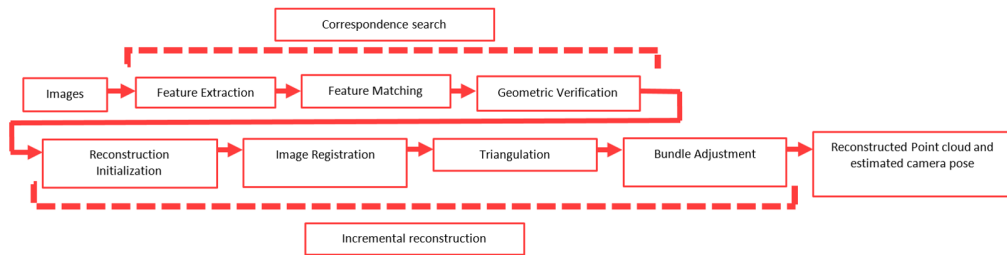


Figure 3.12: Structure from motion pipeline.

of correspondences search between images and a second phase of iterative incremental reconstruction. This phase takes as input the image set and generates as output the so called Scene Graph (or View Graph) that represents relations between geometrically verified images. The iterative reconstruction phase is composed of an initialization step followed by three reconstruction steps: Image Registration, Triangulation and Bundle Adjustment.

The human oral cavity contains a different habitats which are colonized by bacteria. Obtained real images from human oral cavity need the same cleaning environment of a dental clinic. For simplicity now, we mimic the oral cavity environment using Blender software. Our object of interest inside the oral cavity is the teeth as well as the gum tissue. The STL files of human jaws are obtained from a

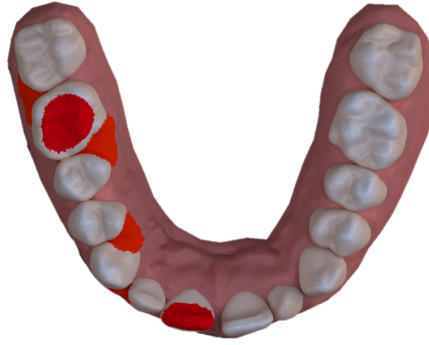


Figure 3.13: Illustration of the areas that the cameras configuration needs to be tested.

dental clinic. These jaws have a variety in crooked, overlapping, and twisted tooth. The intra-oral cameras traveling above the teeth surface and capture the images.

3.5.1 The configuration of the essential sensors

The aim of this study is to identify the number of sensors needed to get a full 3D reconstruction of ‘clinical crowns’ which are defined to be the portion of the teeth that is visible in the human mouth. The geometry of the jaw has a different morphological surfaces(see Fig.3.13); occlusal,incisal, buccal, lingual, labial, distal, mesial and gingiva.

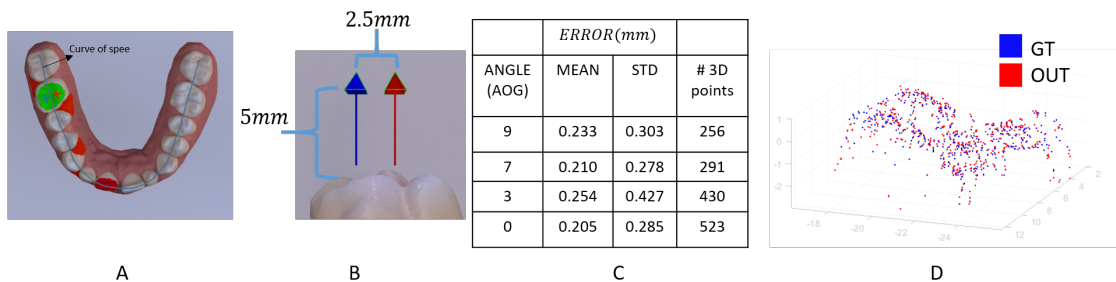


Figure 3.14: Illustration of camera configuration of the occlusal surface ; A) the cameras location w.r.t the curve of spee , B) cameras configuration ,C) The quantity result of the 3D reconstruction comparing to the ground-truth ,and D) the quality result of the 3D reconstruction with ground-truth.

Some of the 3D points of these surfaces may failed to be reconstructed if the cameras configuration are not setup in a way that these points appear at least in two sensors satisfying the geometric constraint of the fundamental matrix.

The evaluation for the best cameras configuration is based on the accuracy of the reconstruction comparing with the ground truth and the number of 3D points that is been reconstructed from this setup. The intrinsic parameters of the cameras in this study are ; the field of view is 62° , and the image resolution is 512×512 Pixels. The depth of field is $5mm$ based on the study in the previous chapter.

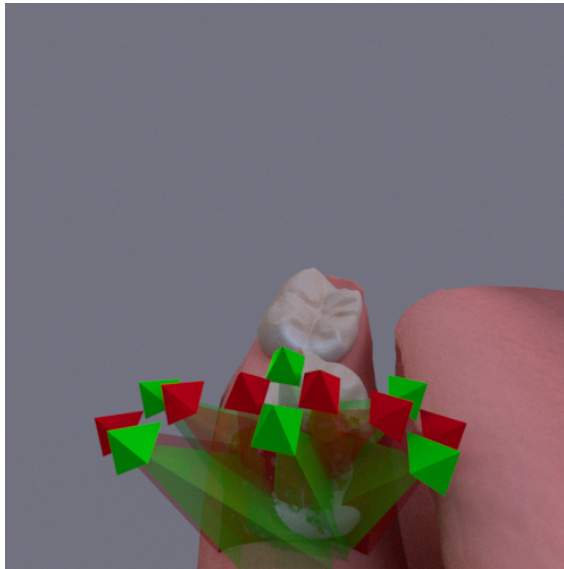


Figure 3.15: The configuration of the twelve cameras.

The study started with occlusal surface (Fig.3.14), two cameras are in frontal to the occlusal surface of the second molar and perpendicular to the curve of spee. The baseline and the angle of the optical axes between the cameras are changed to get the suitable configuration for this surface. The range of the baseline is from 1 to $4mm$ and the angle of OA is from 0 to $9\ degree$. The best 3D reconstruction

result is obtained (see Fig. 3.14.C) when the baseline is $2.5mm$ and the angle of OA is 0° , *i.e.* the cameras are in parallel. The configuration for the incisal surface is different, the cameras should be in parallel to the curve of spee, the baseline is $1.5mm$ and the angle of OA is 0° . Therefore, the minimum number of sensors for 3D reconstruction for the top surfaces of teeth are four, which will be consider as an essential configuration sensors for this application.

3.5.2 The configuration of sensors for twelve cameras design

The essential cameras setup is tested on the rest of the red areas in Fig.3.13. The result shows as well that this configuration is suitable for buccal, lingual and gingiva surfaces. Among the characteristics of the IOS system that we are trying to design is the hand piece of sensors holder should has enough cameras to make the time of the data acquisition and the 3D reconstruction error at the minimum. Therefore, the hand piece sensors holder should has twelve cameras, each four is responsible for 3D reconstruction of one of main surfaces of a tooth as shown in Fig.3.15.

3.5.3 The configuration of sensors for six cameras design

In the twelve cameras design, the hand piece sensors holder has to move in a translational movement along the curve of spee for data acquisition. The thinking of reducing the number of the sensors in the hand piece, to make the probe lighter and cheaper, brought the idea of making the hand piece moves in translational and

rotational movement to scan the three main visible surfaces of a tooth with the configuration of the essential sensors.

3.5.3.1 Mean shape model

The mean jaw's shape model is constructed from a training data ensemble of 3D triangular meshes where each mesh is obtained from a high resolution CT scan of the human jaw moulds. Hereafter, each 3D jaw surface is represented as a 2-manifold triangular mesh $\mathcal{G} = (\mathcal{V}, \mathcal{F})$, where $\mathcal{V} = (x_1, x_2, \dots, x_v)$ is a set of \mathcal{V} -vertices with $x_v \in \mathbb{R}^3$ and \mathcal{F} is a set of \mathcal{F} -triangular faces. A sparse landmark points are manually annotated for all the database samples. There is a total of 72 manually annotated landmark points on a 14 teeth jaw. Generalised Procrustes [40] analysis is then performed to provide an initial rigid alignment of the dense shapes to a common reference frame where the alignment procedure is guided by the sparse set of anatomical jaw landmarks.



Figure 3.16: Mean shape model.

To obtain point-to-point dense correspondence between two rigidly aligned jaw

shapes \mathcal{V}_1 and \mathcal{V}_2 , a warping function, based on the physically motivated thin-plate splines is constructed by using the landmark points as the control points.

To obtain a dense correspondence between all the shapes in the database, a 3D thin-plate spline [41] is applied in an iterative manner as follows, where according to the experimentation, this algorithm converges in a few iterations:

1. Choose one sample as an initial estimate of the mean shape – one may use the first shape in the ensemble.
2. Solve for the warping function between the current mean shape and all the other samples in the database.
3. Re-calculate the estimate of the mean from the aligned shapes.
4. If there is a significant change in the mean, return to step 2.

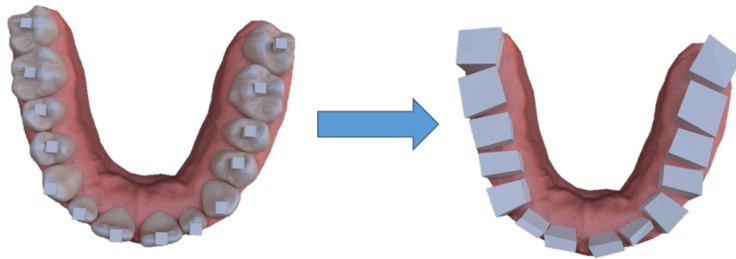


Figure 3.17: The modification of the Mean shape.

3.5.3.2 The modification of the mean shape

The hand piece of sensors holder of this design has to rotate at least three times to scan a cross section of a tooth (buccal, lingual and occlusal surfaces). The worst scenario that would face this design is when the tooth has a sharp edges in the joint

area between tooth surfaces. Therefore, for testing this design in worst scenario, each tooth in the mean shape model (see Fig.3.16) is warped to look like a box in such all the tooth in the jaw are boxes with same size of the actual corresponding tooth (see Fig.3.17) . The configuration of the essential cameras is used for 3D reconstruction of the modified mean shape model as a first experiment to see if it needs any additional sensors. The hand piece of sensors holder of the design is rotated three times to scan a cross section of each tooth in the jaw.

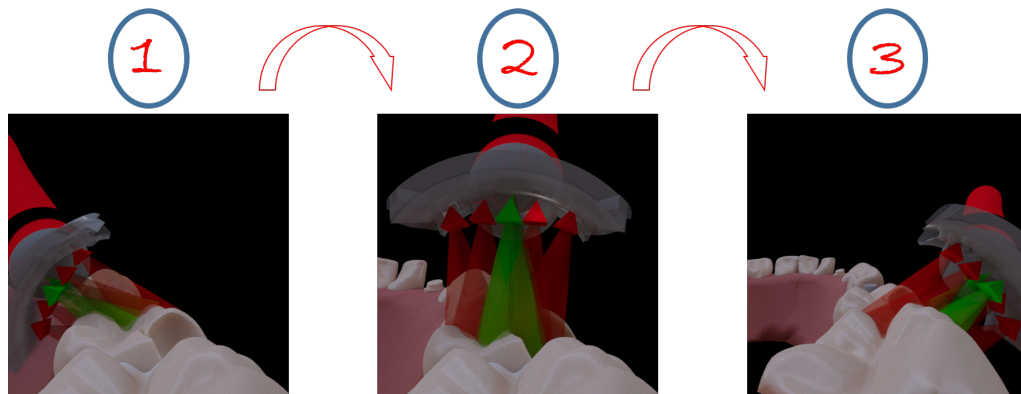


Figure 3.18: Illustration of the data acquisition for a cross section of a tooth using the six cameras design.

The result shows that the design has to have either more sensors added to the essential sensors or the cross section of a tooth needs to be scan more than a three times. After a multiple of testing, the new design needs six cameras, four are the essential sensors and two more in both sides with a $2mm$ a way from the essential sensors (see Fig.3.18), to scan a cross section of a tooth with a minimum rotation of the hand piece of sensors holder. The 3D reconstruction result of the modified mean shape model (see Fig.3.19) shows a good quality and quantity comparing with the ground-truth.

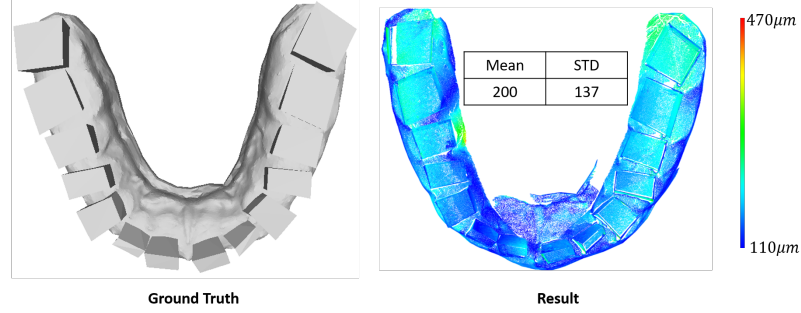


Figure 3.19: Illustration of the 3D reconstruction of the modification the Mean shape result comparing with ground-truth.

3.5.4 Comparison between the six and the twelve cameras design

Table 3.1 illustrates the differences between two designs in details. The twelve sensors design looks to be more plausible design , since it has one more advantage than the six sensors. However, this study considers a simulation experiment, so some of the concepts and numbers may change in the real experiment.

Table 3.1: Comparison between the six and the twelve cameras design.

	Twelve cameras	Six cameras
Number of images	Less	More
Time consuming	Less	More
Cost	More	Less
Accuracy	Same	Same

3.6 Experiment and result

We conduct experiments to evaluate the performance of the proposed twelve sensors design in 3d reconstruction of the human jaws.

For the manner of studying the 3D reconstruction accuracy of the sensor planning for one subject, a mandibular jaw (see Fig.3.20) with a complete dentition is modified to get a different kinds of a tooth problems that would affect in teeth reconstruction accuracy.

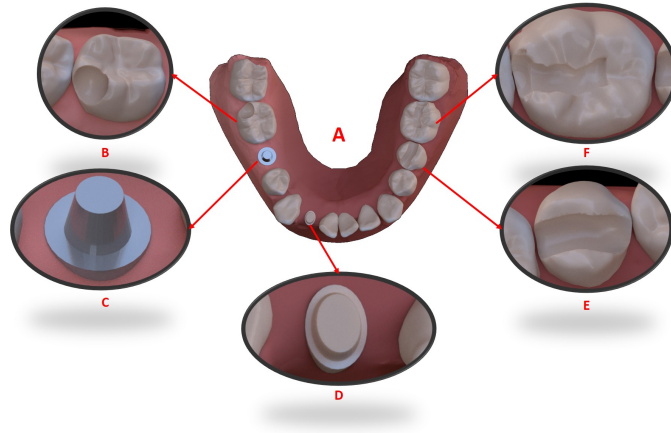


Figure 3.20: Customized mandibular complete-arch reference model; A)complete arch, B)cracked on tooth no.30, C)crown preparation on tooth no.29, D)full-contour crown preparation on tooth no.26 ,E)mesio-occlusodistal inlay preparation on tooth no. 20 and F) mesio-occlusodistal inlay preparation on tooth no.19

A reference model comprised different types of single-tooth preparations. Teeth 26 and 29 (Fig.3.20.C and D) are prepared for full contour crowns, teeth 19 and 20 (Fig.3.20.E and F) for a mesio-occlusodistal inlay, and tooth 30 (Fig.3.20.B) as cracked tooth. We used Blender software to acquire the images, the reference model is imported and the twelve sensors are installed at a distance of $5mm$ from the model. The intrinsic parameter of the sensors are; image resolution 500×500 , image sensor $1/14''$ and the FOV is 62° . The evaluation is proceeded in two scenarios; individual and complete full arch.

For a complete full-arch, 1200 sample images are captured ,while for each indi-

vidual tooth only 120 sample images. For both scenarios the overlap between each two consecutive sample images from the same sensor is more than %75, in such the 3D points correspondences . Structure from motion [39] is a method among the passive technique for 3D reconstruction. Given a set of images acquired from different observation points, it recovers the pose of the camera for each input image and a three-dimensional reconstruction of the scene in form of a sparse point cloud. We used the optimized SFM software [42] for 3D reconstruction. The reconstruction 3D models and the ground-truth are rigidly registered for the trueness accuracy measurement. The trueness result describes how far the measurement deviates from the actual dimensions of the measured object. Measuring the accuracy is done by line distance with limited points of the reconstruction and the ground-truth. As shown in Fig.3.21, five individual tooth are reconstructed and compared with ground-truth, the color map was used to visually observe the 3D differences. In Fig.3.22 shows the 3D reconstruction of the complete full arch and the color map to observe the RMS error comparing with ground-truth.

Table 3.2: Statistical Analysis on individual and complete arch scenarios in trueness of 3D reconstruction (μm)

Tooth No.	Min	Max	Mean	Std
F	26	285	120	96
E	26	290	135	121
D	27	188	61	50
C	34	230	74	86
B	23	211	130	100
Complete arch(A)	60	590	280	200

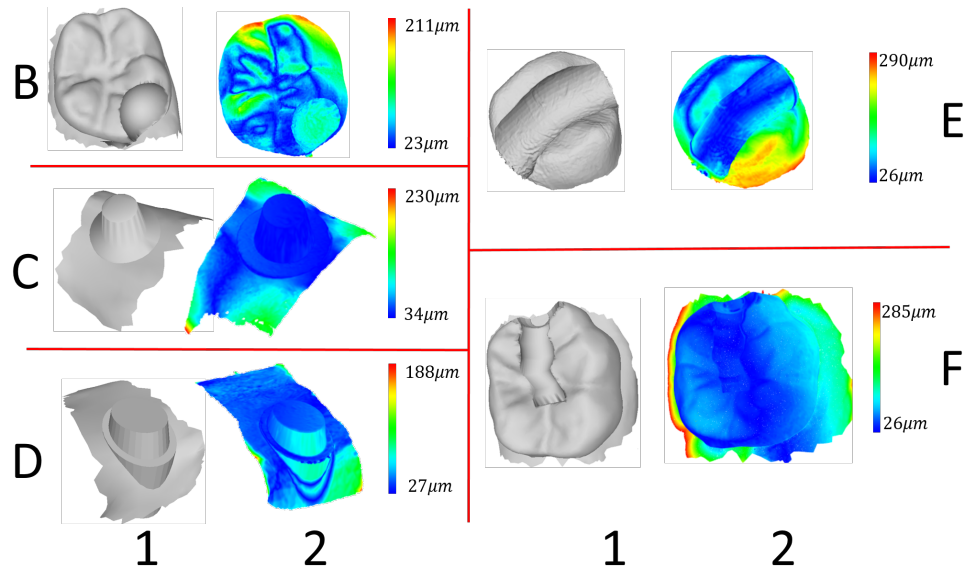


Figure 3.21: The individual 3D reconstruction tooth for;B) cracked on tooth no.30,C) crown preparation on tooth no.29, D)full-contour crown preparation on tooth no.26 , E) mesio-occlusodistal inlay preparation on tooth no. 20, and F) mesio-occlusodistal inlay preparation on tooth no.19 and . the column 1) shows the ground-truth and 2) the 3D reconstruction result with the color map of the RMS error comparing with ground-truth.

For the quantitative evaluation, The mean deviation, maximum and minimum values of the RMS error in μm comparing the reconstruction with ground-truth, are displayed in table 3.2.

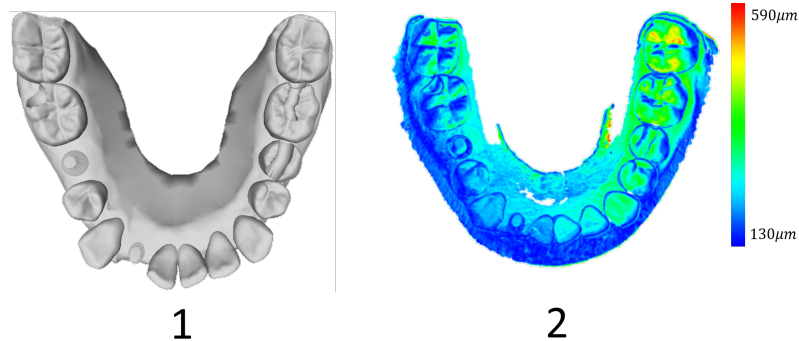


Figure 3.22: The Complete Arch 3D reconstruction; 1) shows the ground-truth and 2) the 3D reconstruction result with the color map of the RMS error comparing with ground-truth.

3.7 Conclusion

In this chapter, the sensor planning was designed based on the morphological teeth surfaces. Two different sensors configuration were designed in such each design has a certien number of sensors according to the hand piece of sensors holder movement inside the oral cavity. The twelve sensors design would be more robust than the six sensors design in 3D reconstruction of the human teeth.

CHAPTER 4: VISIUAL APPLICATION IN HUMAN ORAL CAVITY ENVIRONMENT

Oral dental applications based on visual data pose various challenges such as low lighting conditions, saliva, and low texture. The proposed approach is introduced to stitch images of human teeth that are captured by an intra-oral camera. In such monocular image matching, a low rate of features on teeth surfaces causes a problem leading to a mismatch between teeth images. Therefore, this chapter presents an approach to improve the matching in these low-texture regions. Normals of tooth surface is extracted using a shape from shading. Due to the oral environment, the surface normals impact many of imprecise values; hence an algorithm is being formulated to rectify these values and generate normal maps. The normal maps reveals the impacted geometric properties of the images inside an area, boundary, and shape. In addition, we investigate the unrestricted camera movement problem. The camera may be moved along the jaw curve with different angles and distances due to handshaking. To overcome this problem, each frame is tested, after warping it, and only correct frames are used to generate the panoramic view. The proposed approach outperforms comparing to the state-of-art auto stitching methods..

4.1 Features Matching

The surface orientation at a 3D point P in the scene on a surface S is determined by a unit vector perpendicular to the plane tangent to S at P . Shape from Shading (SFS) is an algorithm among the tools used in shape-oriented extraction from a single view. Under the assumption of perspective projection and the camera is calibrated, our previous approach [43] obtained a metric representation of teeth and gum surfaces using SFS algorithm. For each captured image, we use this approach to estimate the surface normal as follows. An image point p is given by $sp = BP + b$ where B and b are decomposed parameters from the projection matrix, p is the corresponding image coordinate and s is the scalar. So, $P = B^{-1}(sp - b) = g(s(x, y))$. By finding the scalar s , $g(s(x, y))$ will define a unique 3D point P on the object. The normal at a surface point (see Fig 4.1.b) can be defined a $N = \frac{r \times q}{|r \times q|}$ hence $\mathbf{r} = \frac{dg(s(x; y))}{dx}$, $\mathbf{q} = \frac{dg(s(x; y))}{dy}$. This can be estimated using a Taylor's series expansion and applying the Jacob iterative method (for more details see [43].).

However, the SFS can't provide accurate information of the surface normals relying upon the raw image information. For that reason, considering the surface normals estimation depends on light direction and the surface patches are homogeneous and uniformly lit by a distant light source. A cross product is applied between the normal surface and the light direction vector (assume that the light and the camera are in the same direction). Then the mean filtering is performed with a window size 9.

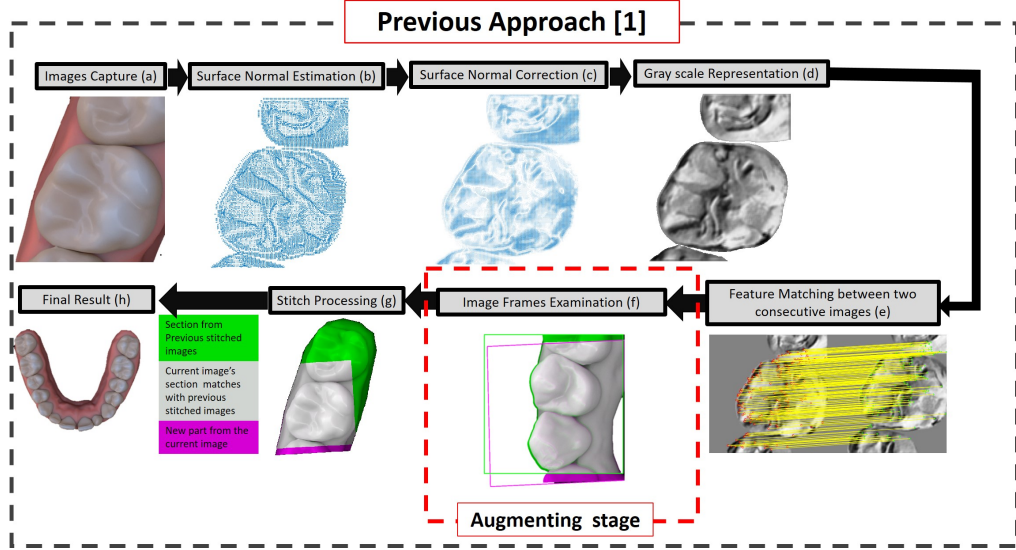


Figure 4.1: The work flow of the proposed images stitching approach

By applying the cross product as well as the mean filtering, the surface patches (see Fig 4.1.c), which have the same orientation, have the same surface normal unlike other patches, which are on the border between two planar surfaces.

For normal map representation, each pixel of the map stores the direction of the surface normal at that point. Assume that X , Y , and Z are the components of a surface normal. This mapping can be done to the red and green channels as $R = \frac{(\vec{N}_x+1)}{2}$ and $G = \frac{(\vec{N}_y+1)}{2}$. The Z component is always positive, because this is a tangent space normal map, and it is directly mapped to the blue channel. Finally, after obtaining the normal map (see Fig 4.1.d), we convert it to a gray-scale image. The key to extraction normal map of images is to reveal the impacted geometric properties of the images inside an area, boundary, and shape. Unlike the captured RGB images, the normal maps are rich with features. Figure.4.1.e illustrates an example of the proposed normal map, which contains plenty of features that can be

used in the matching step.

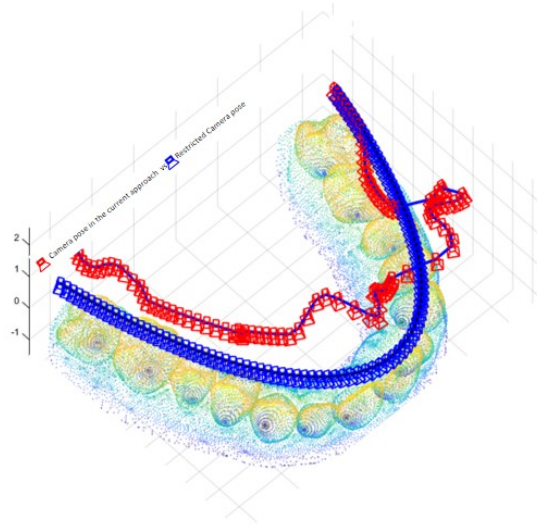


Figure 4.2: An example of camera paths; blue camera represents the path with a restriction movement while the red camera represents the path in the proposed approach

We use the KAZE algorithm [44] to extract features from normal maps. KAZE is a multiscale 2D features detection and description algorithm in nonlinear scale-spaces. The state-of-the-art methods used the Gaussian blurring scale space of an image to detect and describe features at different scale levels while KAZE detects and describes 2D features using nonlinear diffusion filtering. The KAZE algorithm showed a better performance in both detecting and describing features from normal maps against state-of-the-art methods (e.g., SIFT or SURF).

4.1.1 Image frames examination

Since the camera moves in non-ideal conditions (Fig.4.2), the stage for deciding whether an image is valid to proceed in a panoramic stitching pipeline is very important. In the proposed frame examination step, first, we apply the homography

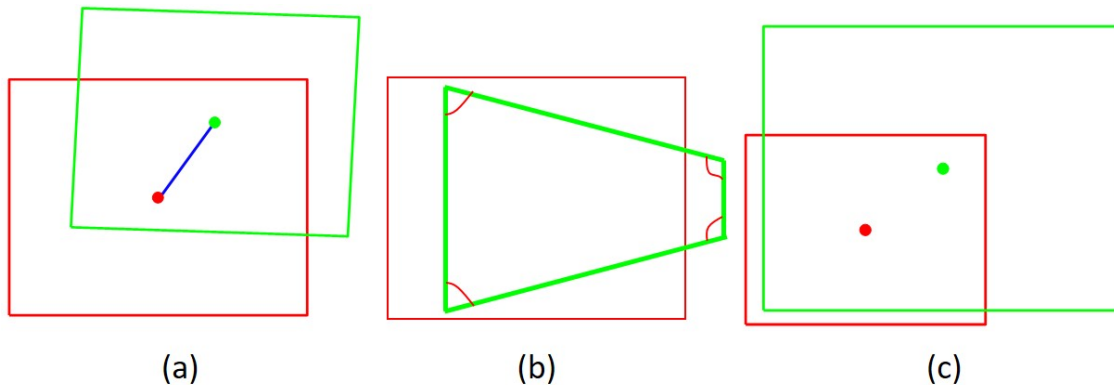


Figure 4.3: Examples of frame rejection (a) Translation ,(b) Shearing and (c) Scale with translation issues.

matrix on each two consecutive images. The homography matrix warps the moving image to the fixed image. Then, we test the corners of the moving image. The image will be used in the stitching process if it satisfies the following conditions, as shown in Fig. 4.3: 1) The internal corners degree should be between $[80, 110]$,2) The translation between the moving frame and fixed image frame should not be more than 10% of the fixed images' size and 3) Both ratio of the scale and translation together between the fixed image and moving image frame should not be less than 0.85 and not be more than 5% of the fixed image's size, respectively.

4.2 Image stitching using flexible warps

The prescribed images capturing conditions don't lead to views that differ purely by rotation, nor planar scene since the object is near to the camera. Therefore, using a basic homographic warp yields misalignment. To overcome this problem, we build on Zaragoza et al. approach [45] and warp each part using a location dependent homography. These estimated as-projective-as-possible warps account

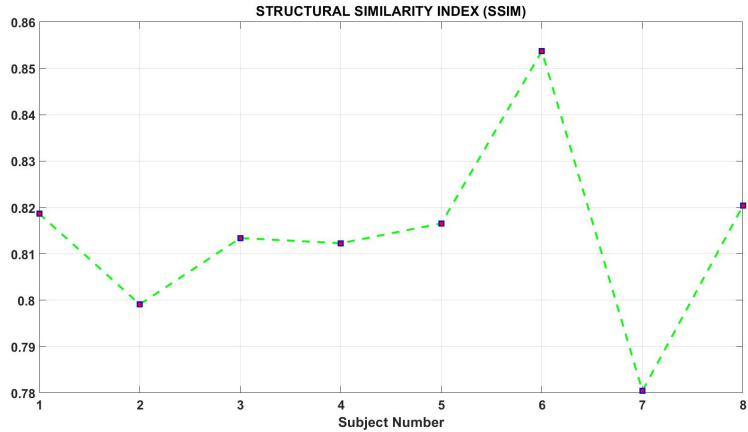


Figure 4.4: Structural similarity index metric between proposed algorithm results and their ground-truth of ten subjects.

for data that deviates from the required constraints. Zaragoza et al. [45] uniformly partition the image into a grid of n cells, and give higher importance to the center of each cell than the other pixels far from the center. Pixels within the same cell are then warped using the same homography. Unlike Zaragoza et al. approach [45], we don't use a uniform grid. But, we use the estimated normal maps to segment the 2D domain of the image into connected planar segments. Each pixel is assigned a weight that is inversely proportional to its distance from the centroid of the segment. Then the pixels of each planar segment are warped using the same homography.

4.3 Experiments and results

The evaluation of the proposed approach has been done by comparing its performance with our previous work [46] (i.e., results will be presented with and without augmenting stage processing). Also, ground-truth images are obtained using a single shot for the Jaw. We perform a qualitative and a quantitative evaluations.

Dataset Description: The human oral cavity contains different habitats that are colonized by bacteria. Obtaining real images from the human oral cavity needs the same cleaning environment of a dental clinic. For simplicity now, we mimic the oral cavity environment using Blender software. Our object of interest inside the oral cavity is the teeth as well as the gum tissue. The STL files of eight human jaws are obtained from a dental clinic. These jaws have a variety in crooked, overlapping, and twisted tooth. The intraoral camera is moved above the teeth surface capturing the images. The images are taken with different distances from the surface as shown in Fig. 4.2 . To generate the panoramic view of the whole jaw, we perform stitching for all the images that are captured.

Qualitative evaluation: We apply our proposed algorithm on the oral dataset. Fig. 4.5 shows the ground-truth, results of the proposed approach, and results of [46]. It is clear that using a frame examination step, in our proposed approach, gives much better performance than other approaches [46]. For more qualitative evaluation, ground-truth and our results are rigidly aligned then combined on one 2D domain. Results show that generated images are closely similar to the ground-truth images.

Quantitative evaluation: For each subject, first, the panoramic view is rigidly aligned with the ground-truth. Then, we compute a quality metric—structural similarity index (SSIM) [47]- between our the aligned results and the ground-truth. Fig. 4.4 shows that the proposed method performs well and is more than 70% close to the ground-truth.

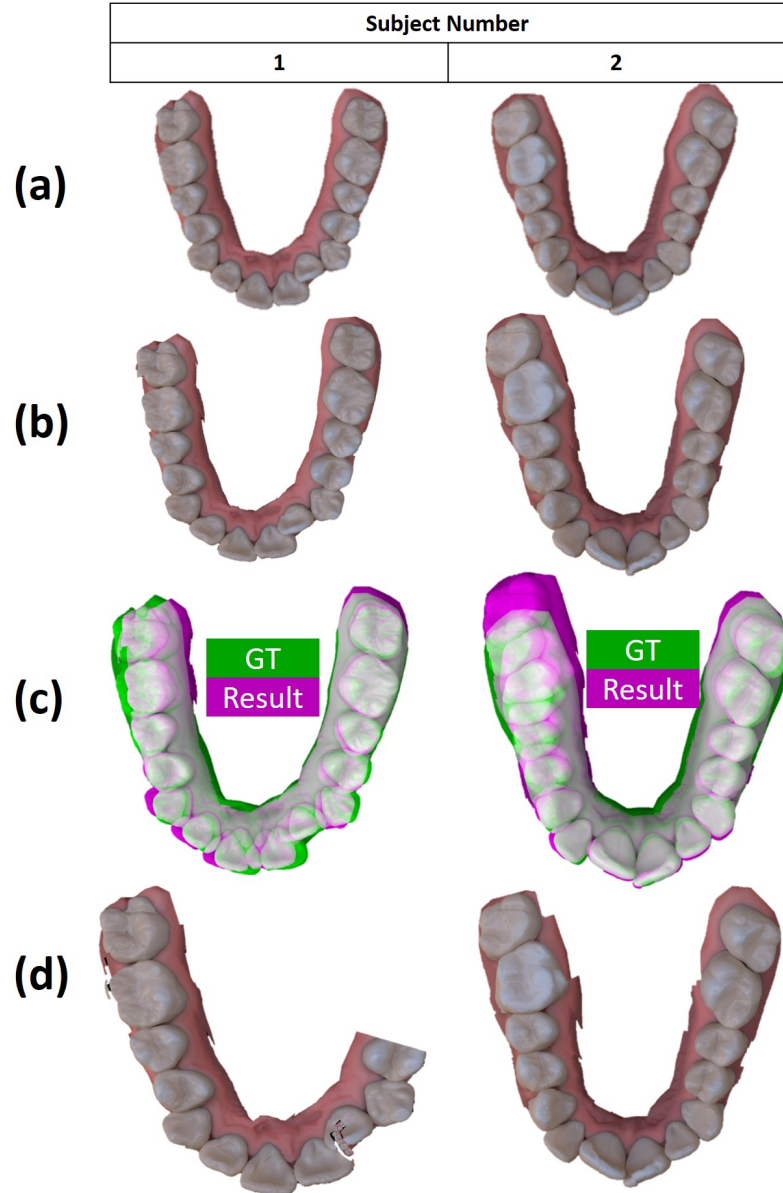


Figure 4.5: Examples for stitching images of two subjects: (a) Ground truth, (b) Results with the augmenting stage, (c) Results are rigidly aligned with ground-truth, and (d) Results without the augmenting stage.

4.4 Conclusion

In this chapter, an approach was introduced to stitch frames captured by intra-oral camera to generate a panoramic view of the human jaw. The camera movement

inside the oral cavity is unrestricted. This adds more challenges to the unfriendly environment of the jaw. To overcome this problem, each frame is tested after warping it, upon on its corners appearance compared with the previous frame. Valid frames are used to generate a panoramic view. This approach has been evaluated both qualitatively and quantitatively. Results confirm the high performance of the proposed approach.

CHAPTER 5: CONCLUSION AND FUTURE WORK

5.1 Conclusion

The three dimensional reconstruction of the human teeth, or as they call it in dental field the a dental impression, is needed for many common procedures such as dental crowns, dental bridges, dental implants and dentures. Many dentists would argue that the impression is the most important part of the process. The reality is that, it is impossible to build a high end restoration on a bad impression. In this thesis, the problem of the image-based 3D reconstruction of human teeth is investigated. The main advantage of using the passive method is the low cost of these systems, as they are made of few and cheap components.

The geometry of the oral cavity is studied to reveal the dimensions of areas that the hand piece sensors holder would reach during the data acquisition. These dimensions are very important in the sensor planning design for 3D reconstruction of the human teeth in such the configuration of sensors would be designed according to these dimensions.

The sensor planning was designed based on the morphological teeth surfaces. Two different sensors configuration were designed in such each design has a certain

number of sensors according to the hand piece of sensors holder movement inside the oral cavity. A different experiments were conducted to evaluate the performance of each design for 3D reconstruction accuracy of human teeth.

In addition to 3D reconstruction application, an approach is was introduced to stitch frames captured by an intra-oral camera to generate a panoramic view of the human jaw. The camera movement inside the oral cavity is unrestricted. This adds more challenges to the unfriendly environment of the jaw. To overcome this problem, each frame is tested after warping it, upon on its corners appearance compared with the previous frame. Valid frames are used to generate a panoramic view.

5.2 Future work

Structure of motion (SFM) provides the highest accuracy among vision systems of this type. However, only high contrast targets and well-defined edges can be measured with high accuracy. Untargeted, or featureless, surfaces may not be measured at all. In addition, the ambient light affects significantly the ability of the system to successfully extract all desired features. The 3D reconstruction in this thesis was completed using artificial texture on the teeth to get more features for correspondences and ease the work for the camera's configuration. As future work, there are different tasks that need to be done for this project such as:

- Investigate if there is any chance to extract a feature from the surface of a tooth using neural network techniques instead of using a special light or artificial texture.

- One of the advantages of using the active method in 3D reconstruction, only the 3D points which are in focus will be reconstructed. Unlike the passive method, it would reconstruct any point that satisfies the geometric constraint. *i.e.*, there will be a cleaning step for the shape recovery from the specious 3D points.
- Investigate in the deep learning architecture for the registration of 3D scans of teeth surfaces.
- Switching from simulation to real experiments to evaluate the sensor planning using real cameras.
- Eventually after switching to the real experiments, evaluate the system using clinical trials.

REFERENCES

- [1] Andrzej Poznanski. Visual revolution of the vanishing of ethan carter. The Astronauts, 25, 2014.
- [2] Ken Brown Andrew Hamilton. Photogrammetry and Star Wars Battlefront. <https://www.ea.com/frostbite/news/photogrammetry-and-star-wars-battlefront>, 2016.
- [3] Olivier Saurer, Friedrich Fraundorfer, and Marc Pollefeys. Omnitour: Semi-automatic generation of interactive virtual tours from omnidirectional video. In Proc. 3DPVT2010 (Int. Symp. on 3D Data Processing, Visualization and Transmission), 2010.
- [4] E Nocerino, F Lago, D Morabito, F Remondino, L Porzi, F Poiesi, S Rota Bulo, P Chippendale, A Locher, M Havlena, et al. A smartphone-based 3d pipeline for the creative industry-the replicate eu project. 3D VIRTUAL RECONSTRUCTION AND VISUALIZATION OF COMPLEX ARCHITECTURES, 42(W3):535–541, 2017.
- [5] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, et al. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In Proceedings of the 24th annual ACM symposium on User interface software and technology, pages 559–568, 2011.
- [6] Ingrid Carlbom, Demetri Terzopoulos, and Kristen M Harris. Computer-assisted registration, segmentation, and 3d reconstruction from images of neuronal tissue sections. IEEE Transactions on medical imaging, 13(2):351–362, 1994.
- [7] M Ahmed, S Yamany, E Hemayed, S Ahmed, S Roberts, and A Farag. 3d reconstruction of the human jaw from a sequence of images. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 646–653. IEEE, 1997.
- [8] Sameh M Yamany, Aly A Farag, David Tasman, and Allan G Farman. A robust 3-d reconstruction system for human jaw modeling. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 778–787. Springer, 1999.

- [9] Aly S Abdelrahim, Moumen T El-Melegy, and Aly A Farag. Realistic 3d reconstruction of the human teeth using shape from shading with shape priors. In 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pages 64–69. IEEE, 2012.
- [10] Sameh M Yamany, Aly A Farag, David Tasman, and Allan G Farman. A 3-d reconstruction system for the human jaw using a sequence of optical images. IEEE Transactions on Medical Imaging, 19(5):538–547, 2000.
- [11] Sameh M Yamany and Aly A Farag. A system for human jaw modeling using intra-oral images. In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286), volume 2, pages 563–566. IEEE, 1998.
- [12] Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Shape-from-shading: a survey. IEEE transactions on pattern analysis and machine intelligence, 21(8):690–706, 1999.
- [13] Shireen Y Elhabian and Aly A Farag. Appearance-based approach for complete human jaw shape reconstruction. IET Computer Vision, 8(5):404–418, 2014.
- [14] Noam Babayoff. Method and apparatus for colour imaging a three-dimensional structure, September 12 2019. US Patent App. 16/427,101.
- [15] Noam Babayoff and Isaia Glaser-Inbari. Imaging a three-dimensional structure by confocal focussing an array of light beams, September 6 2005. US Patent 6,940,611.
- [16] Rune Fisker, Henrik Ojelund, Rasmus Kjær, Mike van der Poel, Arish A Qazi, and Karl-Josef Hollenbeck. Focus scanning apparatus, July 9 2019. US Patent 10,349,041.
- [17] Federico Frigerio et al. 3-dimensional surface imaging using active wavefront sampling. PhD thesis, Massachusetts Institute of Technology, 2006.
- [18] Denis Laurendeau, Louis Guimond, and Denis Poussart. A computer-vision technique for the acquisition and processing of 3-d profiles of dental imprints: an application in orthodontics. IEEE transactions on medical imaging, 10(3):453–461, 1991.
- [19] Cambren N Carter, Rosario J Pusateri, Dongqing Chen, Abdelreheim H Ahmed, and Aly A Farag. Shape from shading for hybrid surfaces as applied to tooth reconstruction. In 2010 IEEE International Conference on Image Processing, pages 4049–4052. IEEE, 2010.
- [20] Aly Farag, Shireen Elhabian, Aly Abdelrehim, Wael Aboelmaaty, Allan Farman, and David Tasman. Model-based human teeth shape recovery from a

- single optical image with unknown illumination. In International MICCAI Workshop on Medical Computer Vision, pages 263–272. Springer, 2012.
- [21] Eslam Mostafa, Shireen Elhabian, Aly Abdelrahim, Salwa Elshazly, and Aly Farag. Statistical morphable model for human teeth restoration. In 2014 IEEE International Conference on Image Processing (ICIP), pages 4285–4288. IEEE, 2014.
- [22] Chenglei Wu, Derek Bradley, Pablo Garrido, Michael Zollhöfer, Christian Theobalt, Markus H Gross, and Thabo Beeler. Model-based teeth reconstruction. ACM Trans. Graph., 35(6):220–1, 2016.
- [23] James Davis, Ravi Ramamoorthi, and Szymon Rusinkiewicz. Spacetime stereo: A unifying framework for depth from triangulation. In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., volume 2, pages II–359. IEEE, 2003.
- [24] Onur Ozyesil, Vladislav Voroninski, Ronen Basri, and Amit Singer. A survey of structure from motion. arXiv preprint arXiv:1701.08493, 2017.
- [25] Johannes L Schonberger and Jan-Michael Frahm. Structure-from-motion revisited. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4104–4113, 2016.
- [26] Jaeryun Ko and Yo-Sung Ho. 3d point cloud generation using structure from motion with multiple view images. In The Korean Institute of Smart Media Fall Conference, pages 91–92, 2016.
- [27] Changchang Wu. Towards linear-time incremental structure from motion. In 2013 International Conference on 3D Vision-3DV 2013, pages 127–134. IEEE, 2013.
- [28] Noah Snavely, Steven M Seitz, and Richard Szeliski. Photo tourism: exploring photo collections in 3d. In ACM Siggraph 2006 Papers, pages 835–846. 2006.
- [29] Carl E Rieder. Maximum mandibular opening in patients with and without a history of tmj dysfunction. The Journal of Prosthetic Dentistry, 39(4):441–446, 1978.
- [30] Xiao-Yan Li, Cheng Jia, and Zi-Chuan Zhang. The normal range of maximum mouth opening and its correlation with height or weight in the young adult chinese population. Journal of Dental Sciences, 12(1):56–59, 2017.
- [31] András Szentpétery. Clinical utility of mandibular movement ranges. Journal of orofacial pain, 7(2), 1993.
- [32] Khalid H Zawawi, Emad A Al-Badawi, Silvia Lobo Lobo, Marcello Melis, and Noshir R Mehta. An index for the measurement of normal maximum mouth opening. Journal-Canadian Dental Association, 69(11):737–741, 2003.

- [33] KH Travers, PH Buschang, H Hayasaki, and GS Throckmorton. Associations between incisor and mandibular condylar movements during maximum mouth opening in humans. Archives of Oral Biology, 45(4):267–275, 2000.
- [34] Sonja Kraljević, Josip Pandurić, Tomislav Badel, and Nikša Dulčić. Registration and measurement of opening and closing jaw movements and rotational mandibular capacity by using the method of electronic axiography. Collegium antropologicum, 27(2):51–59, 2003.
- [35] Gary A Carter and James A McNamara Jr. Longitudinal dental arch changes in adults. American Journal of Orthodontics and Dentofacial Orthopedics, 114(1):88–99, 1998.
- [36] Güldane Mağat and Sevgi Özcan Şener. The morphological changes in the mandible bone: the effects of age, gender and dental status. Meandros Medical and Dental Journal, 19(2):111, 2018.
- [37] Richard Hartley and Andrew Zisserman. Multiple view geometry in computer vision. Cambridge university press, 2003.
- [38] Qier An and Yuan Shen. Camera configuration design in cooperative active visual 3d reconstruction: A statistical approach. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 2473–2477. IEEE, 2020.
- [39] Jan J Koenderink and Andrea J Van Doorn. Affine structure from motion. JOSA A, 8(2):377–385, 1991.
- [40] Timothy F Cootes, Cristopher J Taylor, et al. Statistical models of appearance for computer vision, 2004.
- [41] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Transactions on pattern analysis and machine intelligence, 11(6):567–585, 1989.
- [42] Simon Fuhrmann, Fabian Langguth, and Michael Goesele. Mve-a multi-view reconstruction environment. In GCH, pages 11–18. Citeseer, 2014.
- [43] Sameh M Yamany, Aly A Farag, David Tasman, and Allan G Farman. A 3-d reconstruction system for the human jaw using a sequence of optical images. IEEE Transactions on Medical Imaging, 19(5):538–547, 2000.
- [44] Pablo Fernández Alcantarilla, Adrien Bartoli, and Andrew J. Davison. Kaze features. In Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, Computer Vision – ECCV 2012, pages 214–227, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

- [45] Julio Zaragoza, Tat-Jun Chin, Quoc-Huy Tran, Michael S. Brown, and David Suter. As-projective-as-possible image stitching with moving dlt. IEEE Trans. Pattern Anal. Mach. Intell., 36(7):1285–1298, 2014.
- [46] Mohamad Ghanoum, Asem M Ali, Salwa Elshazly, Islam Alkabbany, and Aly A Farag. Frame stitching in human oral cavity environment using intraoral camera. In 2019 IEEE International Conference on Image Processing (ICIP), pages 1327–1331. IEEE, 2019.
- [47] Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4):600–612, 2004.

CURRICULUM VITAE

Mohamad Ghanoum

Education

PhD in ECE, Speed School of Engineering, University of Louisville

Working on image-based 3D reconstruction of human jaw

Current CGPA: 3.75 Expected graduation date: 2020

MSc. ECE, Speed School of Engineering, University of Louisville, 2015

Thesis Title: 3D-reconstruction of human jaw from a single image: integration between statistical shape from shading and shape from shading.

BSc. Telecommunication Engineering, Faculty of Engineering, University of Kalamoon, 2008

Graduation Project: TDMA and FDMA Telecommunication System

Current Position

Graduate Research Assistant, CVIP Lab, Speed School of Engineering, University of Louisville, 05/2021 - now

Work Experience

- Research Assistant, CVIP Lab, Speed School of Engineering, University of Louisville, 1/2014 - 4/2019

Selected Publications

Full list at <https://scholar.google.com/citations?user=FYHhwLIAAAAJ&hl=en>)

- **Ghanoum, Mohamad**, Asem M. Ali, Salwa Elshazly, Islam Alkabbany, and Aly A. Farag. "Panoramic View of Human Jaw Under Ambiguity Intraoral Camera Movement." In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 1-4. IEEE, 2020.
- **Ghanoum, Mohamad**, Asem M. Ali, Salwa Elshazly, Islam Alkabbany, and Aly A. Farag. "Frame stitching in human oral cavity environment using intraoral camera." In 2019 IEEE International Conference on Image Processing (ICIP), pp. 1327-1331. IEEE, 2019..
- Alkabbany, Islam, Asem Ali, Amal Farag, Ian Bennett, **Mohamad Ghanoum**, and Aly Farag. "Measuring Student Engagement Level Using Facial Information." In 2019 IEEE International Conference on Image Processing (ICIP), pp. 3337-3341. IEEE, 2019.

- Mohamed, Mostafa, Amal Farag, Asem M. Ali, Salwa Elshazly, Aly A. Farag, and **Ghanoum, Mohamad**. “Fly-In Visualization for Virtual Colonoscopy.” In 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 2062-2066. IEEE, 2018.
- **Ghanoum, Mohamad**, Asem M. Ali, Salwa Elshazly, Islam Alkabbany, and Aly A. Farag. “Automatic extraction of interdental gingiva regions for accurate statistical shape from shading-based reconstruction of human jaw.” In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 998-1001. IEEE, 2018.
- Mohamad, Mostafa, Amal Farag, Asem M. Ali, Salwa Elshazly, Aly A. Farag, and **Ghanoum, Mohamad**. “Enhancing virtual colonoscopy with a new visualization measure.” In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 294-297. IEEE, 2018.

Linguistic skills

First language: Arabic (Native speaker)

Second language: English (good)

Personal Information

Current Address: 9514 Aylesbury dr., Louisville, KY, 40242

Cell-Phone: +1 ((502)-445-5815

E-Mails: mohamad.ghanoum@louisville.edu

Date of birth: 1/28/1990 Nationality: SYRIAN

Marital Status: Married Religion: Muslim