Autonomous Robotic System Conducting Nasopharyngeal Swabbing

by

Peter Qiu Jiun Lee

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2024

© Peter Qiu Jiun Lee 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Olivier Stasse Senior Researcher, Laboratory for Analysis and Architecture of Systems (LAAS), National Centre for Scientific Research (CNRS)
Supervisors:	Katja Mombaur Professor, Systems Design Engineering & Mechanical and Mechatronics Engineering, University of Waterloo Professor, Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology
	John Zelek Professor, Systems Design Engineering, University of Waterloo
Internal Member:	John McPhee Professor, Systems Design Engineering, University of Waterloo
Internal Member:	Steven Dain Adjunct Professor, Systems Design Engineering, University of Waterloo Chief Medical Informatics Officer, DocBox Inc
Internal-External Member:	Stephen L. Smith Professor, Electrical and Computer Engineering, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Peter Q. Lee was the sole author for all chapters, which were written under supervision of John Zelek and Katja Mombaur. Chapters 1, 2, 3, 5, 7, and 8 are not currently under publication or review for publication. However, the research presented in chapters 5 and 7 may be disseminated for publication in the future.

Chapter 4 was accepted within conference proceedings for the 2022 International Conference for Biomedical Robotics and Biomechatronics. Chapter 6 is currently under review in the journal Scientific Reports.

Citations:

Chapter 4: P. Q. Lee, J. S. Zelek, and K. Mombaur "Simulating and Optimizing Nasopharyngeal Swab Insertion Paths for use in Robotics". In 2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob). 2022 DOI: 10.1109/BioRob52689.2022.9925302. [69] ©2022 IEEE.[1]

Chapter 6: P. Q. Lee, J. S. Zelek , K. Mombaur. "Collaborative robot arm inserting Nasopharyngeal swabs with Admittance Control", *Scientific Reports*, Submitted November 2023.

[1]In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of University of Waterloo's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink. If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

Abstract

The nasopharyngeal swab test is a procedure where a healthcare worker inserts a swab through the nose until it reaches the nasopharynx located at the back of the nasal cavity in order to collect secretions that can later be examined for illnesses. This procedure saw heightened use to detect cases during the COVID-19 pandemic. Its ubiquity also highlighted fragilities in the healthcare system by way of the hazards to healthcare workers from infectious patients and the pressures a pandemic can inflict upon an unready healthcare system. In this thesis we consider and propose an autonomous robotic system for performing nasopharyngeal swab tests by use of a collaborative robotic manipulator arm, under the ideology that the hardware and techniques could eventually be applied to other types of close-contact tasks to support the healthcare system. We also assume that prospective patients would be standing unrestrained in front of the arm, which adds the challenges of adjusting to arbitrary poses of the head and compensating for natural head motion. We first designed an instrumented end-effector to attach to a robotic arm to enable suitable vision and force sensing capabilities for the task. Next, we developed a finite element modeling simulation environment to describe the deformation of the swab as it moves through the nasal cavity, and solve an optimization problem to find ideal paths through the nasal cavity. A visual servo system was designed to properly align the swab next to the nose using visual information using advances in deep learning and state-estimation, which we validated with a number of human trials. A torque controlled force compliant system was designed and evaluated to determine the feasibility of using force measurements to correct for misalignment when the swab is inserted into a nasal cavity phantom. Finally, we integrated all the system components into a cohesive system for performing nasopharyngeal swab tests. We created a simulator using a nasal cavity phantom and a second robot arm to mimic natural motions of the head. This simulator was leveraged to perform extensive experimentation that found promising controller configurations that were able to compensate for head motion.

Acknowledgments

Thanks are extended to my supervisors, John Zelek and Katja Mombaur, for their guidance during my research. I would also like to thank all the members of the HCRMI lab: in particular Sagar Rajendran, Will Thibault, Anas Mahdi, Jonathan Lin, Pranav Barot, and Francisco Chavez with whom I bounced many ideas off of. Also thanks to Jan, Jonas, Menna, Branca, Alyson, Ilknur, Jiwon, Christian, Marie, and Mahsa for their company over the past four years.

My gracious thanks are extended to the RoboHub for access to their Franka Emika arms that were used extensively in this research. Special thanks to Robert for his aid in fabricating the swab holder.

I gratefully acknowledge the funding I have received from the NSERC CGS-D scholarship and from the Engineering Excellence Fellowship provided by the University of Waterloo. The research presented in this thesis was also supported by the Canada Excellence Research Chair program.

Finally, thanks to my family members Song, Janne, Chris, Carmen, and Ben for their everlasting support.

Dedication

In memory to our sweet family cat, Nova.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgments vi

Dedication vii

List of Figures xii

List of Tables xv

List of Symbols xvi

1 Introduction 1

2 Background on nasal cavity, nasopharyngeal swabs, and swab robots 6

2.1 Anatomy of the nasal cavity 6

2.2 Nasopharyngeal swab test 7

2.3 Robotics applied to swabbing 9

3 Hardware used and developed for nasopharyngeal swabbing **13**
3.1 Use of a general purpose robot arm (Franka Emika) . . . 13
3.2 Study on dynamics of healthcare workers applying nasopharyngeal swabs . 17
3.3 Design of an instrumented end-effector for swabbing . . . 19
3.3.1 Force sensor . . . 24
3.3.2 Electromagnet . . . 25
3.4 Computational resources . . . 28

4 Swab insertion simulation & Trajectory optimization **29**
4.1 Introduction . . . 29
4.2 Simulation . . . 32
4.2.1 Beam dynamics . . . 32
4.2.2 Contact dynamics . . . 35
4.3 Optimization . . . 37
4.4 Results & Discussion . . . 38
4.5 Conclusion . . . 46

5 Face pose estimation & Visual servo system **47**
5.1 Introduction . . . 47
5.2 Background on applied vision . . . 48
5.2.1 Facial pose estimation . . . 49
5.2.2 Visual servo control . . . 51
5.3 Methods . . . 51
5.3.1 Joint lookup table . . . 52
5.3.2 RGB image to facial keypoints and rotation (face model & tracker) 56
5.3.3 3D projection . . . 60
5.3.4 Unscented Kalman filter on manifolds (UKF-M) . . . 63
5.3.5 PBVS control law . . . 64

5.4 Experiments to validate pre-contact phase pipeline . . . 65
5.4.1 Positioning and alignment outcomes . . . 67
5.4.2 Noise attenuation . . . 70
5.4.3 Impact of demographics . . . 70
5.4.4 Workspace analysis for the contact phase . . . 74
5.5 Discussion of human trial experiments . . . 74
5.5.1 Avenues for improvement . . . 77
5.6 Conclusion . . . 79

6 Compliant control system **80**
6.1 Introduction . . . 80
6.2 Materials and methods . . . 83
6.2.1 Collaborative robot arm . . . 83
6.2.2 Force sensing end-effector . . . 84
6.2.3 Nasal cavity apparatus and experiment setup . . . 86
6.2.4 Control system . . . 87
6.2.5 Force filtering . . . 88
6.2.6 Waypoint generation . . . 89
6.2.7 Torque control law . . . 89
6.2.8 Termination observer . . . 90
6.3 Experiments . . . 91
6.4 Discussion . . . 94
6.5 Conclusion . . . 95

7 Integrated swabbing with a moving phantom **97**
7.1 Introduction . . . 97
7.2 Background on head motion modeling . . . 98
7.3 System overview . . . 100

7.4 Head motion simulator . . . 100
7.5 Modified visual servo system . . . 106
7.6 Trajectory optimization . . . 110
7.7 Compliant contact stage controller . . . 112
7.8 Nasopharynx contact observer . . . 118
7.9 Safety observer . . . 119
7.10 Experiments evaluating system on moving phantom . . . 121
7.10.1 Metrics . . . 123
7.10.2 Stage 1: Metric i) Transverse force . . . 126
7.10.3 Stage 1: Metric ii) Axial force . . . 128
7.10.4 Stage 1: Metric iii) Time to reach nasopharynx . . . 129
7.10.5 Stage 1: Metric iv) Reached nasopharynx . . . 129
7.10.6 Stage 2: Metric v) Transverse force . . . 130
7.10.7 Stage 2: Metric vi) Axial force . . . 130
7.10.8 Stage 2: Metric vii) Oscillation coefficient . . . 136
7.11 Discussion . . . 136
7.11.1 Avenues for improvement . . . 142
7.12 Conclusion . . . 143

8 Discussion & Conclusion **144**
8.1 Summary of work . . . 144
8.2 Future experiments . . . 145
8.3 Future work towards clinical applications . . . 146

References **149**

APPENDICES **163**

A Video of swab robot **164**

List of Figures

1.1 NP swab phases 4
1.2 Organization of chapters 5

2.1 Nasal Cavity Rendering 8
2.2 Anatomical landmarks used to guide NP swab test 10

3.1 Franka Emika Research arm 15
3.2 Franka control interface communication 16
3.3 Apparatus used in force capture study 18
3.4 Force measurement trail 20
3.5 Peak force statistics 21
3.6 End-effector and supporting circuitry 22
3.7 NP swab diagram 23
3.8 Electromagnet interface 26
3.9 Circuit diagram of electromagnet controller 27

4.1 NP swab collection stages 30
4.2 FEM Beam visualization 32
4.3 Probability of random trajectories reaching NP with respect to applied force 40
4.4 Distribution of average collision forces with respect to applied force 41
4.5 Rendering of optimized trajectory 42
4.6 Convex hull of feasible paths of trajectories with respect to applied force . 43

4.7 Visualization of perturbing optimal trajectory 44
4.8 Contour plots of objective function surrounding optimal trajectory 45

5.1 3DDFA_V2 projecting points onto mannequin 50
5.2 Overview of pre-contact sequencing 53
5.3 Flowchart of visual servo system 55
5.4 Joint lookup table workspace 57
5.5 Illustration of cone used to grade workspace 58
5.6 Swab segmentation 62
5.7 Trial outcome images 66
5.8 Swab to camera calibration from trials 68
5.9 Face labeling scheme 69
5.10 Distribution of final observed angles from human trials. 71
5.11 Convergence and comparison of measurements to UKF-M 72
5.12 Workspace limits of final joint configurations 75
5.13 SOT initialization errors 78
5.14 Improved nostril detector 78

6.1 Compliant control experimental setup 81
6.2 Noisy task space forces from torque sensors 86
6.3 Components of custom end-effector 86
6.4 Nasal cavity phantom 87
6.5 Block diagram for the proposed NP swab insertion system 87
6.6 Force filter response 88
6.7 Side and top images of a successful insertion 92
6.8 Force and displacement graphs comparing force-feedback and no force feedback 93
6.9 Failure states that occurred during trials 96

7.1 Overview of two robot arms interacting 99

7.2 System diagram and staging . . . 101
7.3 Phantom fixture for robot arm . . . 103
7.4 Modifications to phantom . . . 104
7.5 Kinematics of head model . . . 105
7.6 Simulated Ornstein-Uhlenbeck head motion . . . 107
7.7 Pose estimation on phantom fixture . . . 109
7.8 Optimized insertion trajectory through nasal cavity . . . 111
7.9 Swab held in fixture as end-effector moves in a transverse direction . . . 114
7.10 Transverse displacement vs. force . . . 115
7.11 Transverse gain versus displacement . . . 117
7.12 Attempt to estimate axial force model . . . 118
7.13 Fuzzy logic function responses . . . 120
7.14 Example of safety observer . . . 122
7.15 Sequence of motion during a trial . . . 124
7.16 Trial comparison of gain levels . . . 125
7.17 Graphs comparing forces and tip displacement with head motion . . . 127
7.18 Illustration of computation of oscillation coefficient . . . 128
7.19 Average transverse force according to gain . . . 131
7.20 Metric i) violin plot . . . 133
7.21 Metric ii) violin plot . . . 134
7.22 Metric iii) violin plot . . . 135
7.23 Metric v) violin plot . . . 138
7.24 Metric vi) violin plot . . . 139
7.25 Metric vii) violin plot . . . 140

List of Tables

5.1 Study demographics 66

6.1 Force feedback vs baseline contingency table 91
6.2 Comparison of forces between controllers 91

7.1 Breakdown of failures to reach stage 2 by controller configuration 129
7.2 Statistics and summary of stage 1 metrics. 132
7.3 Statistics and summary of stage 2 metrics. 137

List of Symbols

$\mathbf{I}_n$ Identity matrix of size $n \times n$

$[\mathbf{x}]_\times$ Skew symettric cross product matrix using elements of vector x (e.g., $\begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}$)

$\exp(\mathbf{X})$ Matrix exponential of matrix $\mathbf{X}$. If $\mathbf{X}$ is a skew symettric cross product matrix, this produces a matrix $\in$ SO(3)

$\log(\mathbf{R})$ Maps a rotation matrix $\mathbf{R} \in$ SO(3) to a vector $r \in \mathbb{R}^3$ such that $\exp([r]_\times) = R$.

$\mathbf{X}^\dagger$ Moore-Penrose pseudo-inverse of matrix $\mathbf{X}$

$||\mathbf{x}||$ Norm of a vector $\mathbf{x}$: $||\mathbf{x}|| = \sqrt{\mathbf{x}^\top \mathbf{x}}$

diag Diagonal matrix containing elements starting from the upper left (e.g., $\text{diag}(1, 2, 3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

SE(3) Special euclidean group in the third dimension. Consists of a rotation matrix $\mathbf{R} \in$ SO(3), and a translation vector $\mathbf{p} \in \mathbb{R}^3$.

SO(3) Special orthogonal group of the third dimension. Represents the group of 3D rotation matrices where column vectors are orthonormal and the determinant is equal to one.

Symbol notation In this manuscript, unless otherwise specified, uppercase bold letters ($\mathbf{A}$) indicate matrices, lowercase bold letters ($\mathbf{b}$) indicate vectors, and italic symbols (c) are scalars.

Chapter 1

Introduction

Infectious diseases pose a significant risk to healthcare workers (HCWs). Taking the COVID-19 pandemic as an example, the World Health Organization estimates that between 80000 and 180000 HCWs were killed by COVID-19 between January 2020 - May 2021 [103]. During the early stages of the pandemic, it was reported that healthcare workers made up 17.2% of total cases in Ontario (June 22, 2020) [100]. Although, as of June 15, 2021, it was reported that this figure had changed to 4.4% of total cases [22], as policies have evolved, testing has become more accessible, and vaccinations have rolled out. There were also clear psychological effects on HCWs due to the stressful work environment the pandemic induced [120]. Of course, HCWs face the risk of infection from other diseases as well. In Uden *et al.* [128], it was reported that both active and latent tuberculosis occurred at significantly higher rates among HCWs than among the general public. In West African countries, HCWs could have up to a 100-fold increased risk of contracting the Ebola and Marburg viruses [114]. In Germany, it was reported that on average there were 4.8 occupational Hepatitis-C infections per 100 000 employees per year [93]. During the severe acute respiratory syndrome (SARS) outbreak of 2003, 51% of the admitted patients in Toronto, Ontario were HCWs [17]. While personal protective equipment can reduce the chances of infection from COVID-19 and other types of infectious diseases, engineered solutions, such as autonomous robots, offer a way to more effectively reduce risks in terms of occupational health and safety by removing HCWs from physical contact with patients.

With pressures from aging demographics and advancement in robotic technology, there has been increasing interest towards applying autonomous and semi-autonomous robots to roles in the healthcare system [33] [89] [95]. Several platforms have been proposed for fetching and delivering items within hospital environments [45] [1]. Other robotic platforms have been applied to disinfecting rooms and surfaces [7]. Aside from these more general

tasks, robots have been seeing applications in more specialized roles in medicine, such as robotic assisted surgeries [14] [67]. These highly specialized devices enhance the quality of surgery by offering tools that can be teleoperated with greater stability and precision than could be achieved by hand. Mobility assistance and rehabilitation are other topics where robotics are used to support or move limbs of those with disabilities or to prevent injury [91] [10]. Automating pharmaceutical mixing is also a topic of interest [57].

These examples of autonomous robots in healthcare settings consist of support roles, such as fetching items or carrying equipment, that do not require close contact with humans. While specialized surgical robots obviously have contact with people, they are typically operated only semi-autonomously or through teleoperation. Currently, the application of fully autonomous robots that can perform non-specialized close contact tasks on people, such as tests for infectious diseases, extracting blood samples, taking temperatures, or taking ultrasound scans, is still an emerging field. There are three main benefits to having such autonomous robots in a healthcare setting: 1) robots in this role could greatly reduce the risk of infection to HCWs by reducing or eliminating the exposure to pathogens through direct contact with patients; 2) robots provide a way to augment the current workforce and could ease the staffing difficulties that persist in Canada, especially in remote communities where recruiting HCWs can be difficult [24] and from aging demographics [122]; and 3) robots can be programmed to perform tasks with consistency, reducing the variance a human HCW would have in performing tasks.

This thesis chronicles our research to enable a general collaborative robot arm to perform nasopharyngeal (NP) swab tests. The goal of the nasopharyngeal swab test is to insert a flocked swab through the nasal cavity, between the inferior turbinate laterally, the nasal septum medially, and the nasal floor/palate inferiorly, until it makes contact with the nasopharynx, collect a sample with the head of the swab, and extract and store the swab for testing [98]. While the nasopharyngeal swab PCR-test is considered the gold standard for COVID-19 diagnosis versus other tests [97] [64] [32], it is a standardized method for acquiring specimens for many other types of diseases, including various strains of coronavirus, influenza, whooping cough, pneumonia, and rhinovirus [68]. Physically performing the task has several unique challenges, particularly because the state of the nasal cavity is occluded and unobservable. As a result, one must rely on previous instruction, planning, and sensory force feedback to perform the task correctly. Because of these challenges, human HCWs during the pandemic had a high amount of variance in how the test was administered based on their experience level and instruction [50].

The NP swab test is a particularly interesting procedure to apply to robotics because the precision and physical feedback required. The test can be divided into three main phases: a **pre-contact** phase, a **contact** phase, and a **post-contact** phase, as shown in

Fig. 1.1. During the pre-contact phase, the swab tip of the swab be brought to nostril at a position and orientation that enables it for insertion. From the robotics perspective, this would necessitate the use of precise visual information to guide the motion of the swab because of the low tolerances to meet the nostril. During the contact stage, the swab is inserted into the nasal cavity until the tip reaches the nasopharynx, a sample is collected, then the swab is extracted. During contact, a robotic system would primarily need to rely on the low amounts of measured force to adjust for disturbances. Lastly, during the post-contact phase the swab tip is inserted into a container and the upper portion of the swab is snapped off and removed. This thesis focuses on the first two phases, which are more oriented towards the challenges of human-robot interaction than the post-contact phase.

A key assumption in our scenario is that willing patients will stand freely in front of the robot during the NP test, just as they would with a HCW. This means that patients will not be restrained by a head fixture, requiring the robotic system to adapt to the natural motions of the head. This approach is more realistic, reflecting the conditions under which HCWs performed tests in clinical environments during the COVID-19 pandemic. Additionally, it reduces the need for calibrating and sanitizing an external fixture, streamlining the testing process and enhancing patient comfort. We also aim to minimize the amount of specialized hardware that interfaces with the collaborative robot, with the idea that the same system could eventually be applied to other close-contact tasks. An autonomous robotic system that can accomplish these challenging requirements would be of profound impact not just for enhancing access to the NP swab test in clinical settings, but also because of the potential to generalize to multiple close-contact tasks. Robotics can also provide a way to standardize procedures and help reduce the human variability factor that can impact the NP swab test and other tasks.

In the Chapter 2 of this thesis we provide a general background on the nasal cavity anatomy, NP test, and examples of robotics applied to swabbing in the literature and industry. To facilitate the NP swab test, this thesis draws from several different research areas. In Chapter 3 we describe the Franka Emika arm, computational setup, and the instrumented end-effector we designed to facilitate this research. Chapter 4 describes the derivation of a finite element model simulation environment that describes how the swab deforms as it traverses through the nasal cavity. The simulation environment is leveraged with an optimization problem to find the best path to insert the nasal cavity along. Chapter 5 discusses the visual servo system that handles the pre-contact phase of the NP swab test. The system incorporates joint lookup table to resolve inverse-kinematics, deep-learning models to estimate the pose of the face from a camera, an unscented Kalman filter on manifolds state estimator, and pose based visual servo control loop to align the swab next to the nostril. Human trials were conducted and validated the approach. Chapter 6

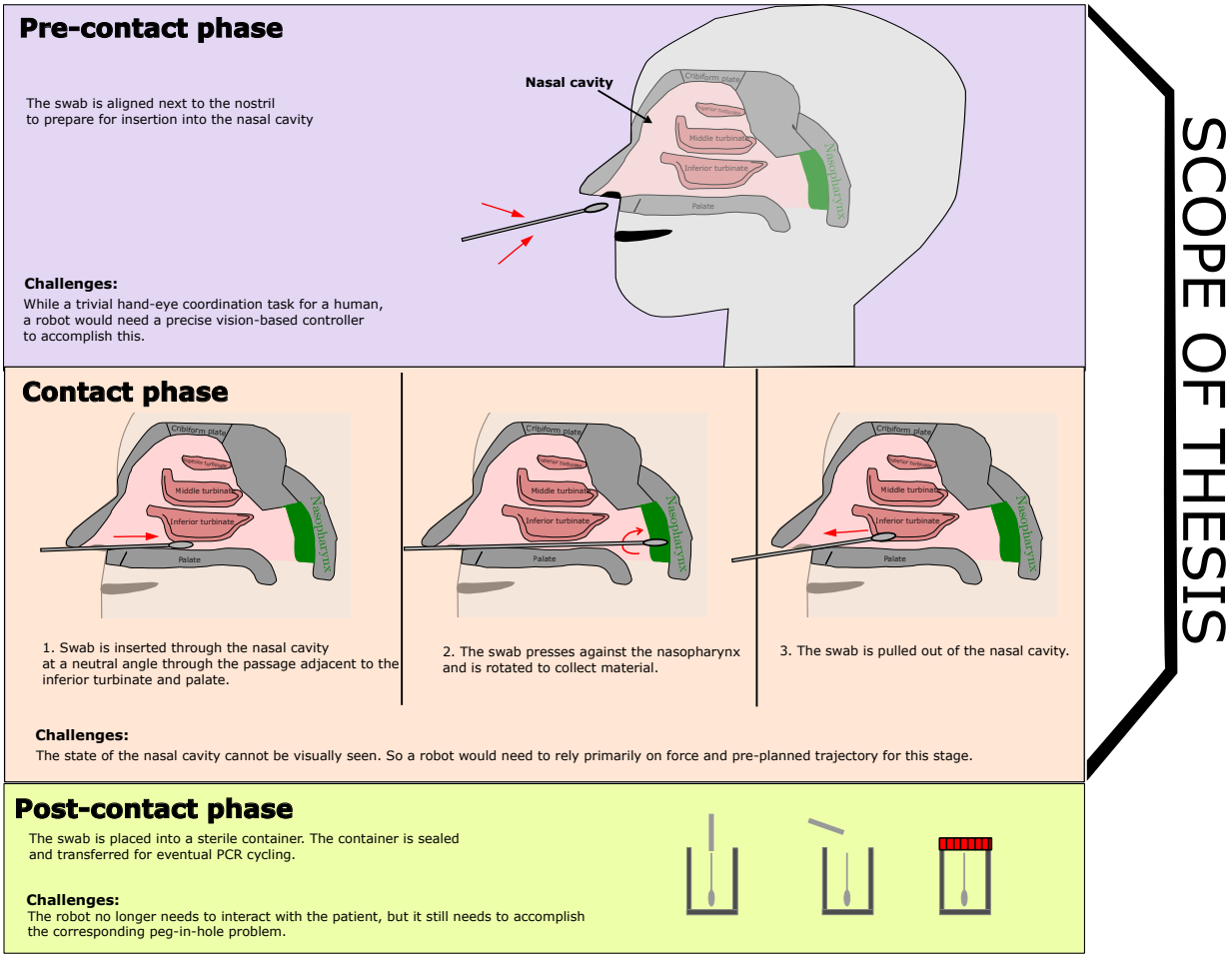


Figure 1.1: Progression of the NP swab tests divided into the pre-contact, contact, and post-contact phases. In this thesis, we examine solving the pre-contact and contact phases that have direct human interaction.

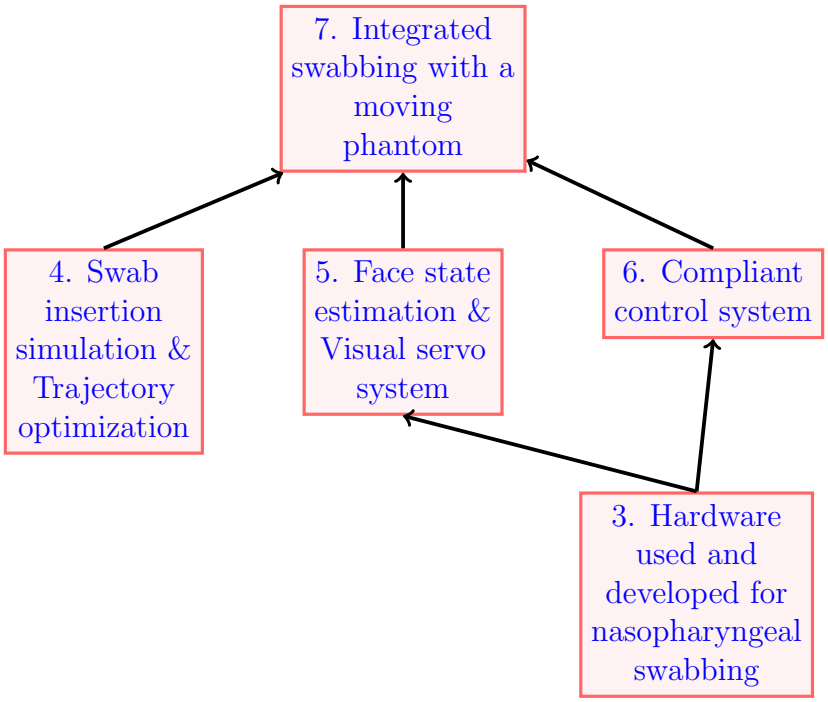


Figure 1.2: Graph showing relationship between chapters in the thesis.

describes our initial work towards implementing a compliant control system to insert the swab during the contact phase. A torque control law was formulated to follow a trajectory and make adjustments with force measurements for inserting into a nasal cavity phantom, a 3D printed apparatus that acts as a stand-in for a human nasal cavity. The results reveal that force measurements are an essential input for correcting for misalignments during the contact phase. As the final component, Chapter 7 combines all the efforts from the previous three chapters into a cohesive robotic system that performs the pre-contact and contact stages of the NP swab test in sequence. This chapter makes improvements on previous methods and also proposes the use of a second robotic arm that can move a nasal cavity phantom with natural head motion. Experiments ultimately show that this final system is effective and can tolerate high levels of head motion. Chapter 8 then provides a conclusion of the research presented in this thesis and an outlook towards future extensions. A tree showing how the chapters build on one another is shown in Fig. 1.2.

Chapter 2

Background on nasal cavity, nasopharyngeal swabs, and swab robots

The research in this thesis requires drawing on knowledge from several distinct fields of research, including simulation, visual servo control, computer vision, and compliant control. However, we will introduce the necessary background of those specific fields within the relevant chapters. In this chapter, we will discuss topics directly relevant to the NP swab test. Section 2.1 discusses literature related to the anatomy of the nasal cavity, Section 2.2 examines how the NP swab test has been evaluated in the medical literature, and Section 2.3 describes the current state of the art of medical robotics applied to NP swab or similar swab applications.

2.1 Anatomy of the nasal cavity

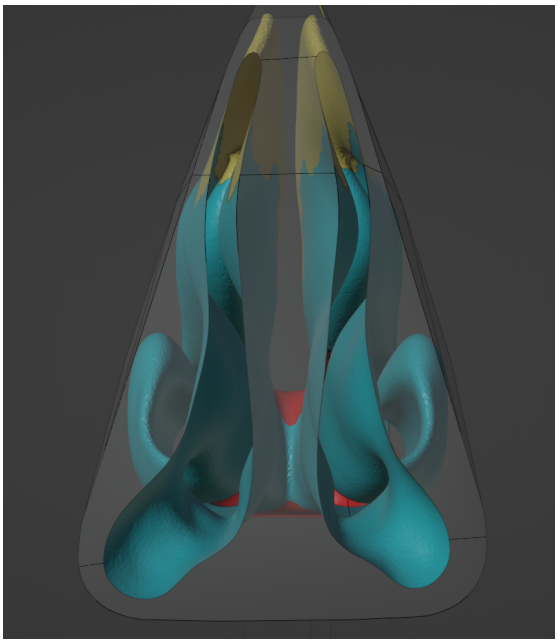
The nasal cavity begins with the two nostrils that lead into two separate vestibules divided by a wall called the nasal septum. The two main functions the nasal cavity serves for the body are respiratory functions and olfactory functions (i.e., sense of smell). The olfactory region is located at the top, which consist of clusters connected to the nervous system that transmit smell signals. The respiratory region occupies the bottom portion of the cavity and consist of a trio of structures called the inferior, middle, and superior turbinates, which are curved bones covered in a soft tissue that lines the cavity called mucosa that regulate

the passage of air [119]. The nasopharynx is located at the posterior of the nasal cavity. Fig. 2.1 provides visualization of the nasal cavity and the regions within.

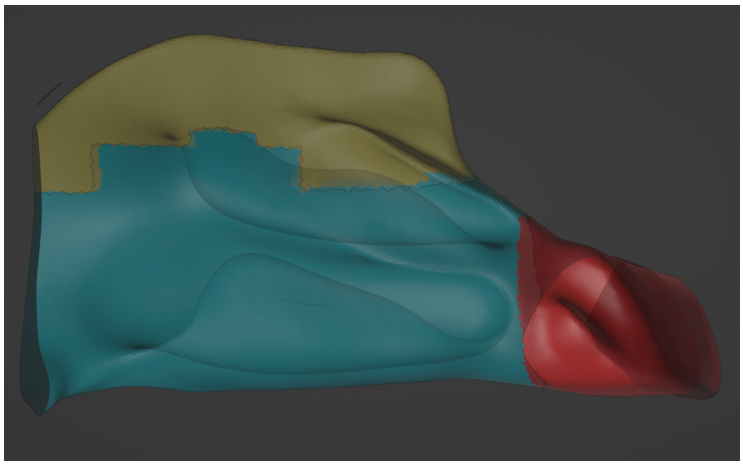
In the medical literature, the geometry of the nasal cavity has been a topic of study from respiratory perspectives. Liu *et al.* [80] used CT scans of 60 patients to derive a coarse geometrical model of the nasal cavity based on sets of common landmarks in the slices of the CT images. Keustermans *et al.* [62] also used CT scans of patients to generate a cylindrical principal component analysis (PCA) to describe the statistical distribution of the volumes within the sample. Brüning *et al.* [21] used similar methodologies to derive geometrical shape models for the volumes of the cavities and used it to perform an analysis of airway pressure changes for respiration. Bastir *et al.* [11] found that there is a significant difference in nasal cavity volume between men and women. Of great importance to this research is the development of physical phantoms that can be used for real-life experimentation of the robot and any control algorithms. Sananès *et al.* [110] developed a soft phantom using a PolyJet 3D printer that is specifically designed to train HCWs for collecting nasopharyngeal swab samples.

2.2 Nasopharyngeal swab test

The nasopharyngeal swab test is done by passing a thin swab through the nostril, and ideally travels through the channel between the inferior turbinate and palate, until it reaches the nasopharynx where it collects secretions [98]. Afterwards, the collected cultures are used in biochemical process called the reverse transcription polymerase chain reaction (RT-PCR), where the nucleic acids are amplified and subsequently examined for the presence of specific viral diseases [133] [68]. In addition to its uses for diagnosing a variety of respiratory viruses [68], the NP swab test found significant use during the height of the COVID-19 pandemic, with several short instructionals for healthcare workers being published in medical journals (e.g., [98] [85]). The literature generally agrees that the nasopharynx is the ideal location in the mouth/nose region for collecting samples with concentrated COVID-19 viral loads, with it outperforming salivary [64], shallow nose [97], and oropharynx samples [133] in terms of diagnostic sensitivity. Hiebert *et al.* [50] notes that in practice there is significant variance in the techniques of HCWs performing the test. The authors also indicate that proper insertion depth is crucial to reach the nasopharynx and collecting an optimal sample, with several Canadian provincial guidelines at the time falling short. In terms of anatomical landmarks, it is noted that the two-thirds the distance from the nostril to the ear is needed to successfully reach the nasopharynx. Fig. 2.2 gives an illustration of these landmarks over a mannequin. While not specific to NP



(a)



(b)

Figure 2.1: Rendered of the nasal cavity volume based on the model from Keustermans *et al.* [62], with the different regions labeled according to color: blue, respiratory region; yellow, olfactory region, a) sagittal view, with nostril region cut away. b) coronal view, with the right half of the cavity hidden. The rendering is viewed outwards from the septum.

swabs, Lim *et al.*'s [55] study regarding NP temperature probes reported that the average distances from the nostril to the nasopharynx at 100 mm males and 94 mm for females. While the NP swab sample collection procedure is regarded as safe [39], incidents are still possible. Fabbris *et al.* [39] reported several cases of heavy bleeding from tissue trauma, with physiological conditions like deviated septums playing a role. Liu *et al.* [79] published a case report, where a swab was elevated at too high of an angle and a lesion was formed at the middle turbinate, causing severe bleeding. It appears that controlling the elevation angle of the swab is exceedingly important for safety because of the sensitive regions at the top of the nasal cavity near the olfactory region, which is adjacent to the base of the skull [16].

2.3 Robotics applied to swabbing

While prior to 2020 there appears to have been little interest with robotic nasopharyngeal or oropharyngeal swabbing, the COVID-19 pandemic spawned many works that have become more sophisticated over time. Li *et al.* [72] created a simple hand-operated 3-DOF device with a tongue depressor to perform oropharyngeal swabs on a phantom. To a similar effect, Liu *et al.* [78] created an oropharyngeal swab robot, called the LINGCAI, that is teleoperated remotely by a practitioner. The robot features a tongue depressor and an actuator that holds the swab that can extend in one direction and rotate along the other two axes so swab can reach the back of the mouth. A force sensor is installed inline with the swab that transmits the force feedback to the practitioner by a haptic interface. The forces were also used to adjust the practitioners commands and cap the maximum force the device can emit. The device was evaluated with clinical trials that showed the robot was competitive with human initiated oropharyngeal tests [74]. The authors expanded on this work with the development of the LINGCAI-II robot [134] that was designed to be fully autonomous; including loading swabs, swabbing the patient, and storing the sample. The LINGCAI-II makes use of a manipulator arm that holds an instrumented end-effector with a camera and force-torque sensors that supports a deep learning computer vision pipeline to identify the target and for compliant control. The system features a disinfecting solution using ultraviolet light and chemical disinfectants. Clinical trials indicated that the samples collected by the LINGCAI-II could also match or exceed the quality of human practitioners. Zhang *et al.* [141] designed a rigid-flexible coupling manipulator with a vision and force controlled system to perform oropharyngeal swab tests through the open mouth. This was later updated to include a UR5 collaborative robot to allow for a wider workspace [30]. Similar to the LINCAI-II, Sun *et al.* [124] created

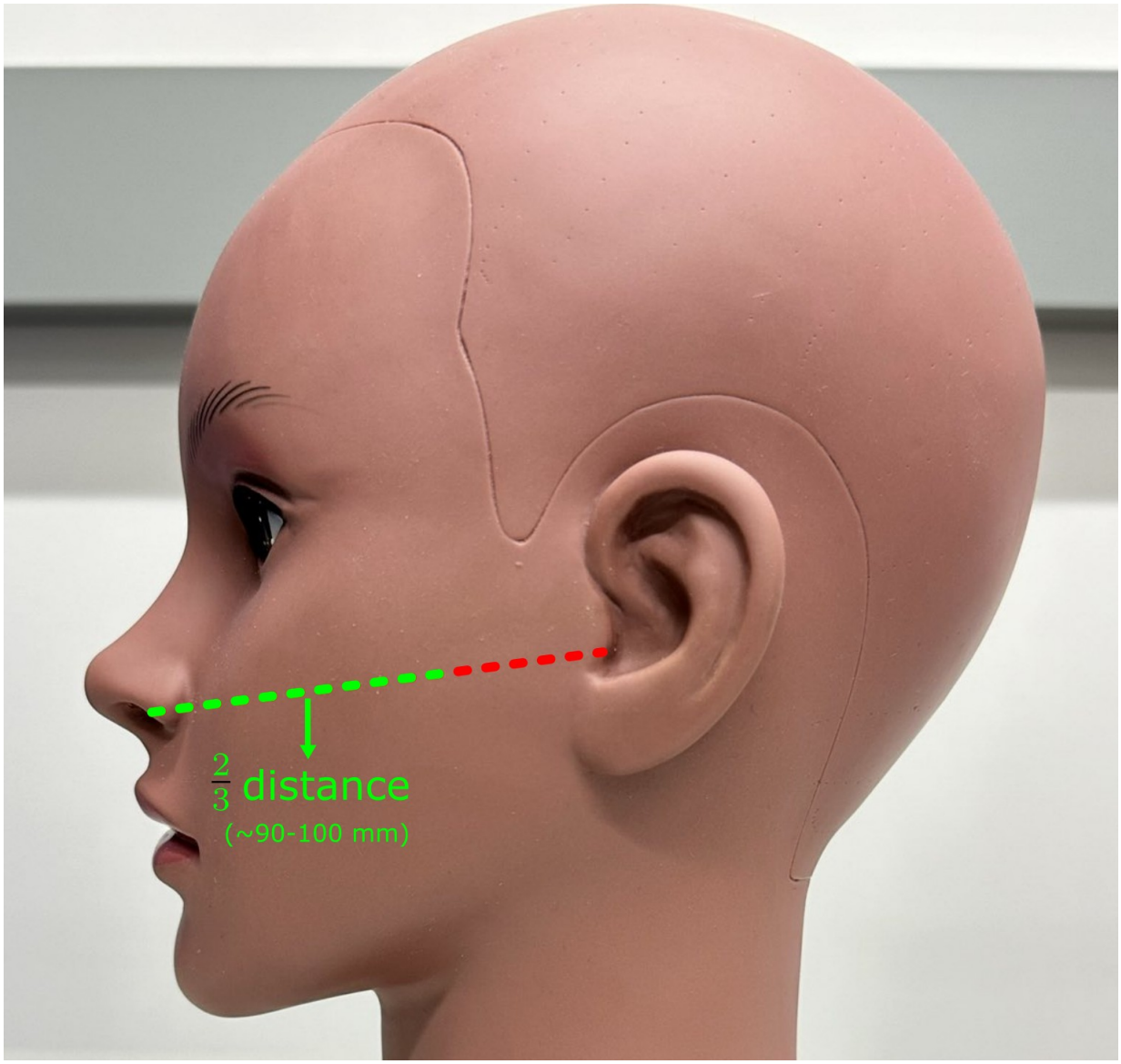


Figure 2.2: Illustration showing a set of landmarks on a mannequin that are commonly used to guide HCWs performing the NP swab test. The distance from the nostril to the ear is highlighted in red, while HCWs are encouraged to use two-thirds of this distance to reach the nasopharynx.

a robot to do all stages of the oropharygeal swab test, including loading swabs, swabbing, storing samples, and disinfection. The system uses a manipulator arm to load swabs and store samples, while a separate specialized 3-axis actuator performs the swab test. The swabbing component uses an RGB-D camera and a force sensor to guide the robot to the oropharynx and control contact respectively. The authors evaluated the quality of their samples using human participants as well.

Seo *et al.* [115] created one of the earliest examples of an NP swab robot. The robot takes the form rotary Stewart platform, with the swab pushed forward via a double roller mechanism. The robot itself is controlled by teleoperation with a 1D force sensor that displays axial forces to the user. The authors evaluated this system with a phantom, under the impression the patient is held in a head fixture. Maeng *et al.* [83] designed a robot around a specialized force restriction mechanism that limits the maximum force the swab can apply during NP swab tests and is actuated within a remote centre of motion mechanism. Shim and Seo [116] followed a similar approach where a remote centre of motion actuator was used for NP swabs under a teleoperation setup. Wang *et al.* [132] designed a custom end-effector that fits a swab that applies linear actuation and rotation along the axis of insertion towards the nasopharynx, while sensing axial force with an optoelectronic sensor. Chen *et al.* [29] applied a soft robotics approach to create a device that actuates a swab and controls the amount of deformation swab during nasopharyngeal or oropharyngeal swabs. Hwang *et al.* [54] utilized a 6-DOF robot with visual servo control to position a swab next to a nostril for the pre-contact phase of the NP swab test. Zhang *et al.* [142] displayed results from a wheeled dual-arm humanoid robot that utilizes vision and a force-torque sensing to perform NP swab tests, although technical details on the underlying system are not elaborated on. Similar to us, Haddadin *et al.* [47] used a Franka Emika research arm to autonomously perform swab tests. However, their setup utilizes a static fixation that the patients place their nose and mouth against, and does not extend to adapting to variable poses from a freestanding patient scenario. Furthermore, their system uses the internal joint-torque sensors to implement compliance control, which we show to have low sensitivity in Sections 3.3.1 and 6.2.2. Because of this, it is unclear if their system actually reaches the nasopharynx; according to their figure 4, the swab displaces less than 3 cm during the swab insertion stage, far below the average 9 to 10 cm required to reach the nasopharynx from the nostril [55].

The COVID-19 pandemic as also spawned a number of commercial attempts to automate swab tests using robotics. The SwabBot$^{\text{tm}}$ (Biobot Surgical, Singapore) [15] appears to have performed the NP swab test by allowing patients to position align their nostrils to a nosepiece to performs the procedure, although it appears to have been discontinued at the time of writing. Lifeline Robotics (Denmark) [76] created a platform based on a ma-

nipulator arm that uses vision guided control to conduct oropharyngeal swab tests. Brain Navi (Taiwan) [18] made a platform based on a Universal Robotics arm to perform NP swabbing, guided by 3D scans of the head. The patient held in a head fixture during the test with a nose clip. All of these devices appear to have patients held in some type of head fixture to ensure the swab is aligned properly.

There appears to a current gap in the literature for a cohesive robotic system that can perform both the pre-contact and contact stages of the NP swab test under the freestanding patient scenario. Thus, the remaining chapters of this thesis document our system that fills this gap in the literature through our proposed system that can execute both stages of the test, while also compensating and evaluating for the natural motion of the head.

Chapter 3

Hardware used and developed for nasopharyngeal swabbing

In this chapter we discuss the hardware components used to facilitate NP swabbing by a robot. First in Section 3.1, we describe the collaborative robotic arm that was our main platform for implementing the test. While we were conceiving the robotic hardware for this research, it was clear that some amount of force measurement was needed to support the contact phase, it was unclear what levels of force was typically transmitted by the swab and whether a collaborative arm would be capable of detecting these levels through on-board joint torque sensors. Therefore, Section 3.2 describes a study we conducted with HCWs to gain an understanding of the level of forces to expect during the contact phase and draw up requirements for the granularity of force measurements needed. Section 3.3 consequently describes the instrumented end-effector we designed to enable the robotic arm to complete the NP swab test. Finally, Section 3.4 describes some of the computational hardware we used to implement the software for this thesis.

3.1 Use of a general purpose robot arm (Franka Emika)

The Franka Emika Research arm is an articulated manipulator with 7 degrees of freedom (DOF). The dynamics of the body can be described as a second order differential system, taking the form of [41]

$$\boldsymbol{\tau} = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) + \boldsymbol{\mu}(\mathbf{q}, \dot{\mathbf{q}}), \tag{3.1}$$

where $\boldsymbol{\tau} \in \mathbb{R}^7$ is the torque applied to each joint, $\mathbf{M} \in \mathbb{R}^{7\times7}$ is the mass inertia matrix, $\mathbf{C} \in \mathbb{R}^{7\times7}$ is the Coriolis matrix, $\mathbf{g} \in \mathbb{R}^7$ is the torque induced by gravity, $\boldsymbol{\mu} \in \mathbb{R}^7$ is the friction from static and Coulomb sources, and $\ddot{\mathbf{q}}, \dot{\mathbf{q}}, \mathbf{q} \in \mathbb{R}^7$ are the joint acceleration, velocity, and angles. Fig. 3.1 shows an image of the arm with the standard gripper end-effector. The Franka Emika arm belongs to a class known as collaborative robots that are safety rated to perform tasks in proximity with humans. This is supported by the integration of torque sensors on each of its joints. The lower level controllers on the hardware can utilize these sensors to enforce safety constraints (e.g., if there is a spike in external torque from hitting something, it can automatically halt motion). The robot also features the ability to have an exchangeable end-effector with its serviceable flange (DIN ISO 9409-1-A50). While our work uses the Franka Emika arm as the main robotic platform, there are several other collaborative robotic arms with similar specifications on the market to which our work could be adapted (e.g., Kinova Link-6 (Montreal, Quebec, Canada), Universal robotics UR5e (Novi, Michigan, USA), Rethink Robotics Sawyer (Bochum, Germany)).

Controllers programs can be run by connecting a personal computer (PC) to a lower level control box via Ethernet connection. Fig. 3.2 from the Franka Control Interface (FCI) documentation shows the communication diagram between the PC and the control box that executes the lower level commands to the robot. The FCI operates in several different modes. The two modes of interest in this thesis are the joint torque control and the joint velocity control modes. With joint torque control, a control law computes joint torques that are transmitted to the control box, and are then transmitted to motors (e.g., the computed torque law from Chapter 6, (6.8)). The downside of this approach comes from the latency between the PC and the control box. The dynamics of the robot (3.1) follow a second order differential system, which means that control is especially impacted by the latency. Not only will a delay in transmitting torque be problematic, but the control law will have outdated state parameters. This makes compensating for model inaccuracies and unmodeled friction difficult because increasing the gain parameters will increase the chances of instability. These factors make a joint velocity controller more attractive because, as a first order system, it is less impacted by latency. The motion generator onboard the control box can then generate the command torques to meet the desired velocity with more up-to-date parameters. It can also compensate for model errors and friction better by applying higher gain to the lower level commanded torques. For these reasons, in Chapter 7 we switch to using a velocity control law.

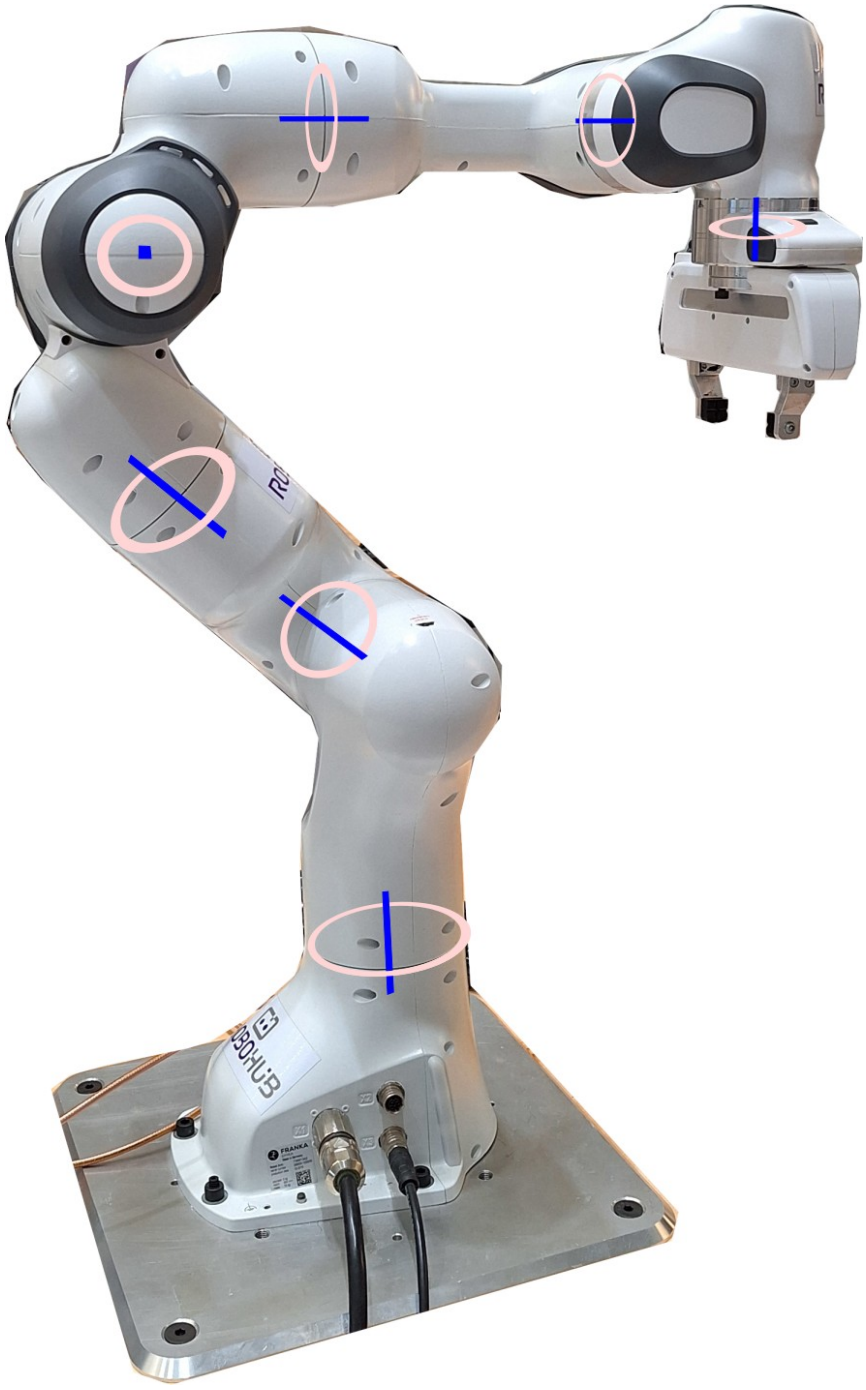


Figure 3.1: Franka Emika “Panda” Research arm. Joint axes are labeled in blue.

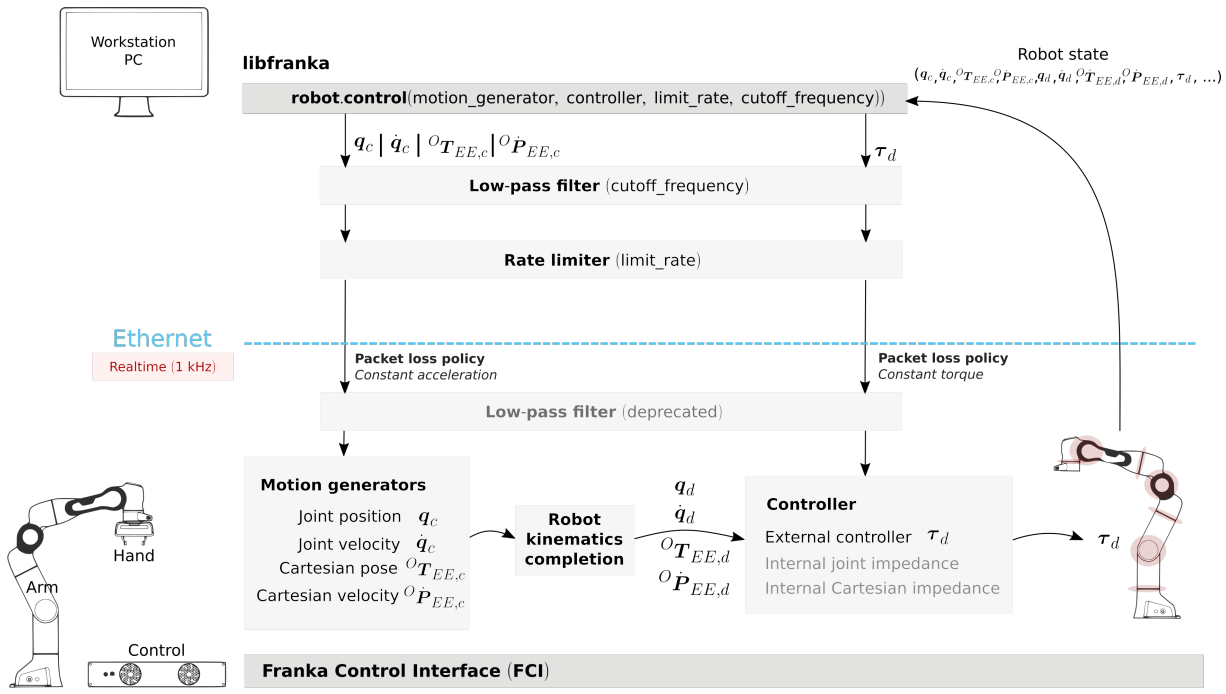


Figure 3.2: Communication diagram of Franka control interface. "Realtime loop: from control commands to the robot desired joint torques" Franka Robotics GmbH. https://github.com/frankaemika/docs, licensed under the Apache 2.0 License. Retrieved May 20, 2024.

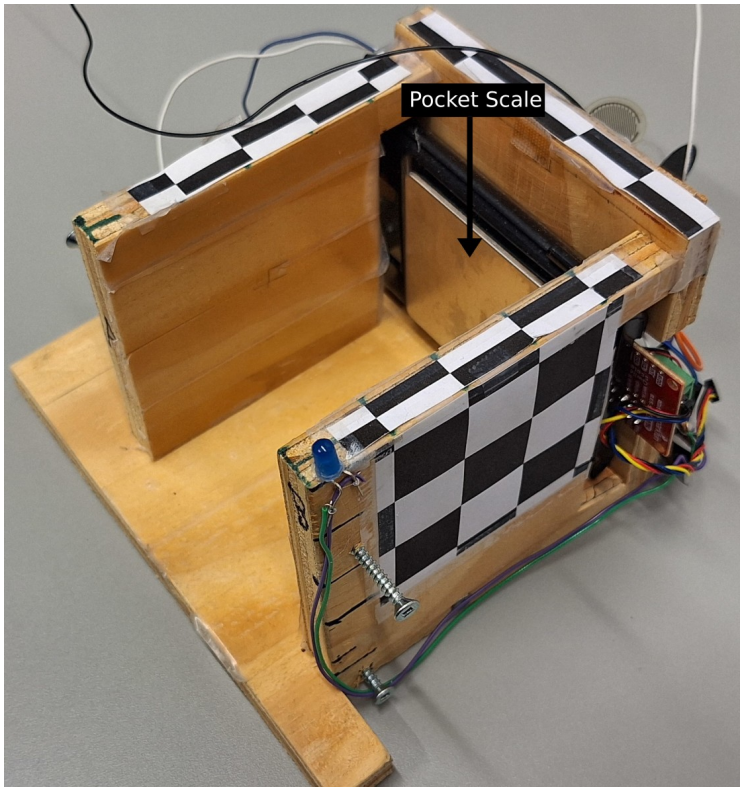
3.2 Study on dynamics of healthcare workers applying nasopharyngeal swabs

When considering the design of the hardware, we needed measurement requirements about the sensitivity of forces. At the beginning of this research, there were no studies published detailing the level of forces HCWs would typically apply with their swabs within the nasal cavity. Therefore, we initiated a human study that was focused towards measuring the forces applied by the NP swab. Later, Park *et al.* [95] created a study to also answer this question, but it was published after we submitted our study to the research ethics board (November 30, 2021).

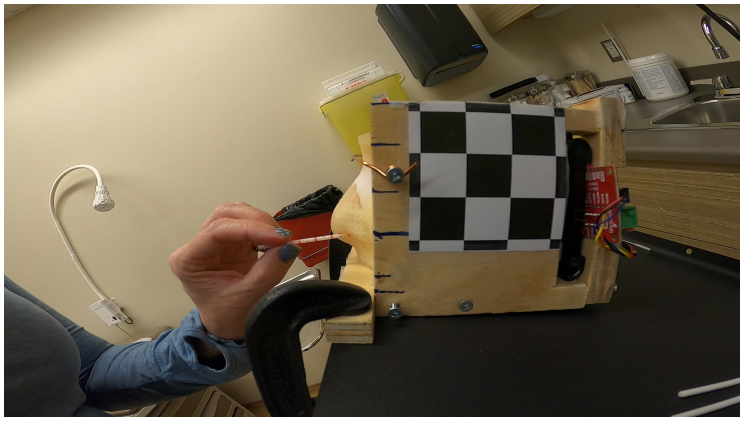
With approval from the University of Waterloo research ethics board (REB# 43485), we conducted our study during May 4, 5, and 6 of 2022 at the University of Waterloo Health Services building. The HCW participants used provided swabs to perform simulated NP swab sample collections on the nasal cavity apparatus shown in Fig. 3.3, along with the phantom designed by Sananès *et al.* [110]. Six trials were recorded for each participant (three times per side of the nostril). The apparatus contains a strain gauge based pocket scale at the back of the apparatus and a force sensitive resistor (FSR) placed at the nasopharynx of the phantom, with the readings interpreted and recorded by an Arduino. Two GoPro Hero 8 cameras also recorded side and top view videos of the trials. The walls of the apparatus are lined with low friction high-molecular tape and lubricated with petroleum jelly to increase force transmission from the phantom to the scale at the back. A total of eight physicians and nurses participated in the study, although two were excluded from the analysis because they did not appear to reach the nasopharynx during the trials.

The results from these trials were highly variable with respect to each participant. Visually, there did not appear to be a consensus on the trajectory of insertion, as each participant appeared to have her own technique. All participants applied rotations to the swab once it reached the nasopharynx, with the duration of the rotation generally lasting between 5 and 15 seconds. Fig. 3.4 shows an example of the forces recorded by the pocket scale during one of the trials. Unfortunately, the FSR became dislodged during many trials, so its force readings have been excluded. Fig. 3.5 shows the distribution of peak forces measured by the scale during the trials for each participant. These peak forces typically occurred once the swab reached the nasopharynx, at the posterior of the nasal cavity. The overall median force from the six participants was 0.86 N, with the maximum recorded value at 1.84 N.

The range of forces we observed during the trials mostly agree with the values reported by Park *et al.* [95]. We also saw that swabbing techniques varied substantially among



(a)



(b)

Figure 3.3: a) Apparatus that holds the nasal cavity phantom during experiments. b) Image of a healthcare worker participating in the trials.

participates, which corroborates with the observations of Hiebert *et al.* [50]. Ultimately, this trial shows that a force measurement system would need to at least be able to detect a 0.8 N impulse to detect when the nasopharynx is reached. However, a force compliant controller would need better granularity than this to allow for a more fine-tuned response. Based on the scale of the forces, we decided that the force measurements would need a granularity of at least 10 mN.

3.3 Design of an instrumented end-effector for swabbing

The end-effector is a major piece of hardware that was designed for the research present in this thesis. The function of the end-effector is to support both visual servo and compliant control modes, while also wielding an NP swab. Consequently, there are three major components on the end-effector: a three-axis force loadcell, an Intel RealSense D435i RGB-D camera, and an electromagnet that holds an NP swab connected to a ferromagnetic swab holder. The base of the end-effector was 3D printed with PLA using a Prusa i3 MK3S+ (Prusa, Prague, Czech Republic) and is screwed directly into the flange of the Franka Research arm. The assembled end-effector and the supporting circuitry can be seen in Fig. 3.6. The loadcell is interfaced between the base and the electromagnet to measure transmitted forces from the swab, while the camera is angled slightly towards the center so it can view the swab in frame.

All the swabs were used in this research were created by the same manufacturer (Obetech Medical PTE LTD, Singapore). While dimensions are likely to differ slightly from manufacturers, the overall attributes of other NP swabs on the market will likely be equivalent. The swab is constructed with ABS plastic with a total length of 146 mm and a diameter that starts at 2.5 mm at the base, and then narrows to 1.0 mm. The thin shaft allows the swab to be compliant and deform inside the nasal cavity. The tip of the swab is covered with a nylon flocking to capture biological materials. In the middle of the shaft is a break point which would allow the practitioner to bisect the swab and leave the tip in a solution to prepare for the PCR cycles. Fig. 3.7 shows a diagram of the swabs and their dimensions.

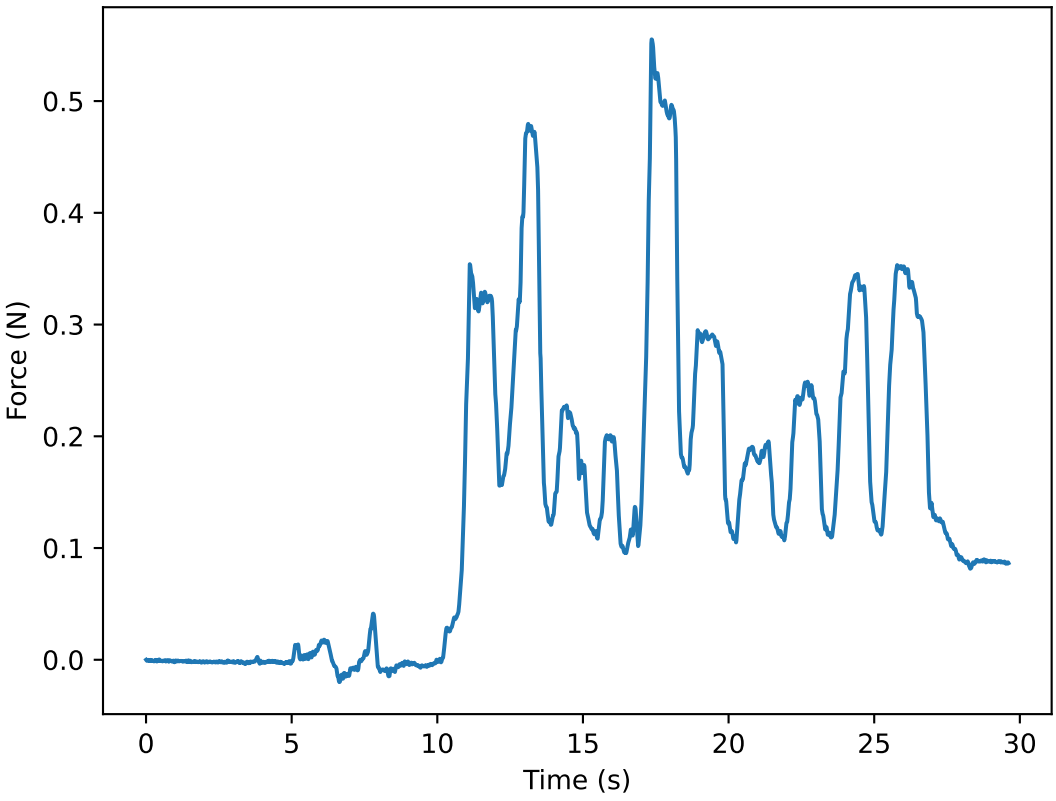


Figure 3.4: Example of forces measured during a trial. The swab initially reaches the nasopharynx at about 10.5 seconds, with the remaining impulses from collection. Note the scale only measures forces in one direction, so any forces applied by the swab in the transverse directions were not recorded.

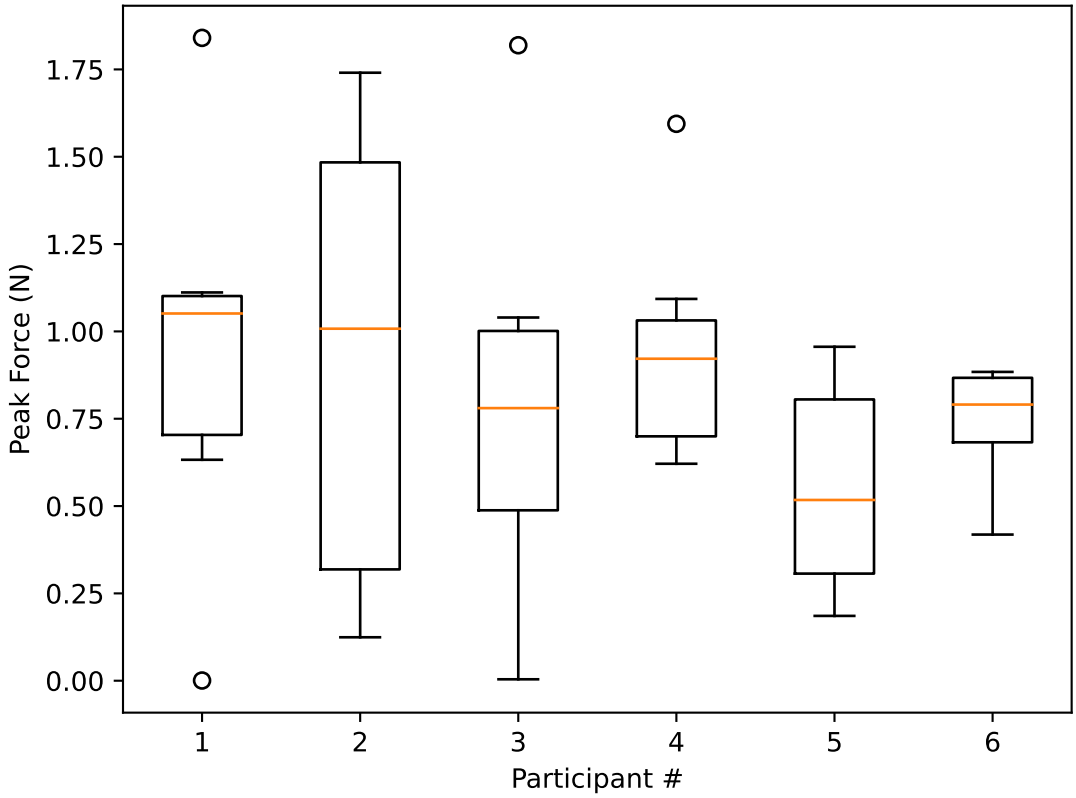


Figure 3.5: Boxplots showing the median and interquartile range of peak forces created by the participants.

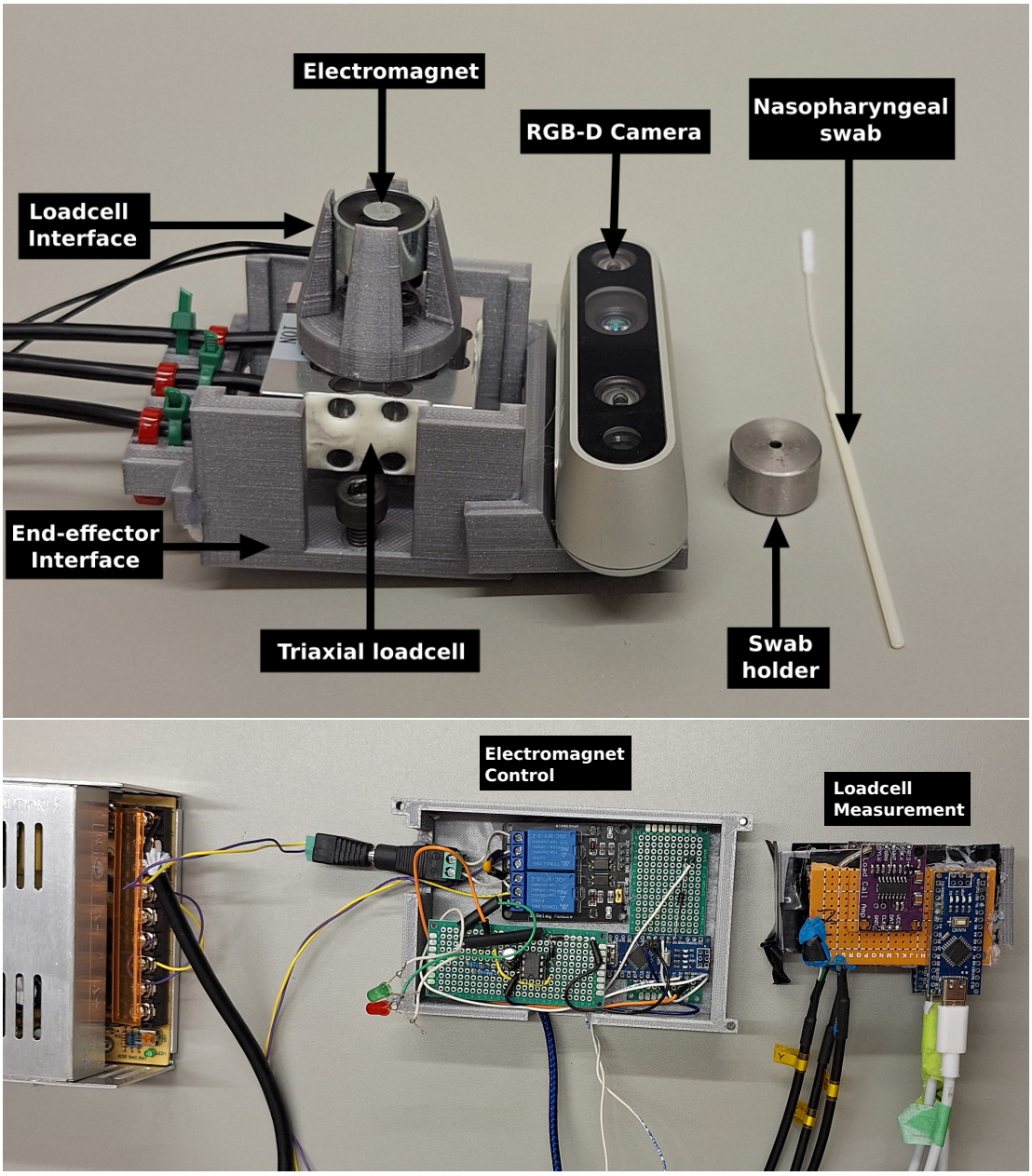
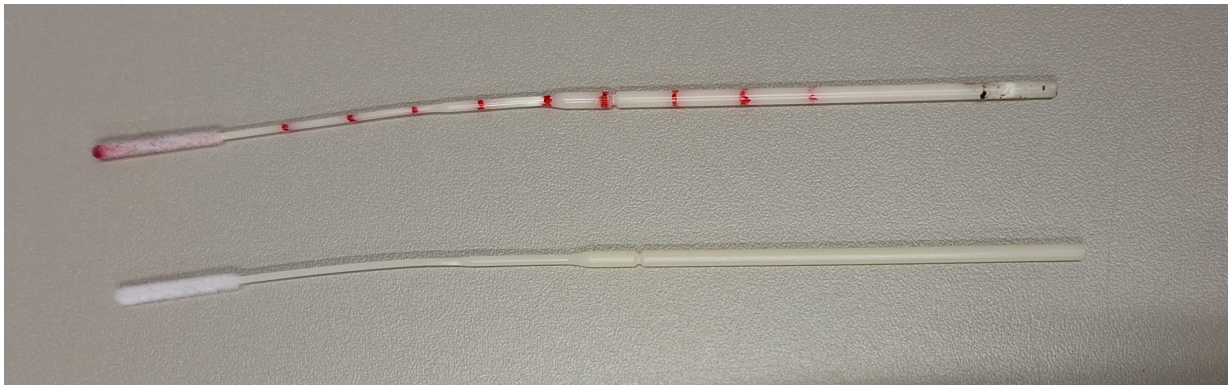
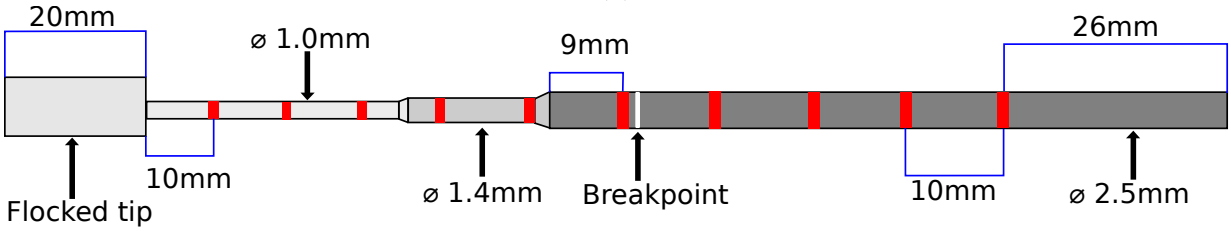


Figure 3.6: Top: Image showing the swab end-effector, which features a tri-axial loadcell, an RGB-D camera, and an electromagnet system to attach and detach the NP swab. Bottom: supporting electronics for controlling power to the electromagnet and ADC circuitry to measure forces on the loadcell.



(a)



(b)

Figure 3.7: a) NP swabs used throughout this research. The top swab is an annotated swab used in experiments, while the bottom is one in factory condition. b) Diagram showing dimensions of swab. We applied red ink at 2 cm increments along the shaft of the swabs to help visually gauge distance. Measurements were made using digital calipers.

3.3.1 Force sensor

From the study in Section 3.2, it is clear that granular levels of force need to be measured for a controller to reactively respond to contacts between the swab and the nasal cavity. Later in Section 6.2.2 we show that the on-board torque sensors on the arm lack sufficient sensitivity to detect forces transmitted to the swab on the end-effector. The high level of noise (approx. 50 mN standard deviation) limits the level of force that can be detected. Also, change in joint configurations leads to large amounts of measured extraneous forces because of model errors. These erroneous forces are so large that they would overpower force transmissions from a swab, making responsive compliant control very difficult to implement. Therefore, one of the primary functions of the end-effector is to precisely measure forces that are applied to the swab.

As a result, the end-effector contains a GPB160 10 N three-axis force loadcell (Galoce, Shaanxi, China) that is mounted between the arm and the swab. The loadcell is based off strain gauges that change resistance when they compress or expand. The strain gauges are placed along the aluminum body is such a manner that when a payload is applied, it will affect the strain gauges that correspond to the three directional axes. In terms of electronics, the strain gauges are assembled into a Wheatstone bridge, then an HX711 analog to digital converter (ADC) measures the corresponding change in voltage within the circuit at 80 Hz. An Arduino Nano interfaces with the loadcell, which then communicates the measurements to a connected computer via USB-serial interface. The loadcell was factory calibrated, which provides coefficients to correspond change in voltage to change in force. The loadcell having capacity of 10 N was chosen for the loadcell because it was clear that we would not be expecting forces greater than 10 N from Section 3.2. Having a low capacity is advantageous because it allows for more granular force measurements, whereas a higher capacity loadcell would produce coarser measurements when using the same ADC. The loadcell and ADC configuration produced a noise level with about 1 mN standard deviation, which satisfies our requirements.

One complication we found with implementing the loadcell was compensating for the impact of gravity. While we expect the gravitational force applied to the payload to change based on the orientation of the end-effector, we observed there were some non-linear effects due to the wire leads connected to the strain gauges pulling on the loadcell. To address this, the wire leads are restrained to a platform on the base of the end-effector with zip ties, and the calibration algorithm in Section 6.2.2 was created to compensate for these effects.

3.3.2 Electromagnet

The electromagnet interface was designed to be able to retain the swab and be able to release it quickly when an abort sequence is triggered. An electromagnet rated with 30 N of attraction was selected, as this far exceeded the expected levels of force the swab would transmit. The electromagnet holds a small 19.05 mm diameter 416 stainless steel cylinder swab holder that was milled to fit the end shaft of the swab. Stainless steel was the preferred material because it can be safely sanitized in an autoclave and the 416 grade is ferromagnetic. The electromagnet is connected to the loadcell by a 3D printed interface, as shown in Fig. 3.8. This interface required special design considerations because the mounting body of the loadcell consists of four M3 screws in a 10 mm $\times$ 10 mm square, while the electromagnet is mounted with a single M3 screw underneath, parallel to its body. The interface therefore has a hollowed out interior to hold the screw for the electromagnet from the bottom, while the four screws surround the body. Also of note are four radial prongs that rise from the base, to less than 1 mm above the electromagnet body. These prongs are used to help align the stainless steel cylinder so that horizontal forces do not dislodge it from contact with the solenoid, but still have sufficient tolerance so that it can detach if power is cut.

The power source for the electromagnet is a 24V DC power supply unit. Supporting circuitry, as shown in the circuit diagram Fig. 3.9, allows the controller or a human to disconnect the electromagnet from its 24V DC power supply unit in case the swab needs to disconnect. This is accomplished by use of a dual electro-mechanical relay and an Arduino Nano connected to the PC with a USB serial interface. The outputs of the power supply are connected to the normally open (NO) terminals while the electromagnet wires are connected to the common (CO) terminals. An unpolarized 10 μF capacitor is also connected to the CO terminals to counteract residual magnetism in the solenoid. When the relay is depowered, an LC circuit is formed between the solenoid and capacitor, which generates a decaying alternating current that helps prevent the steel piece from sticking when power is removed. The logic side of the circuit has one of the digital outputs from the Arduino controlling the state of the relay. This connection is also passed through two daisy-chained emergency stop buttons so that either the participant or the operator of the system can detach the swab and stop the system at any time. The Arduino communicates the relay state to the PC so that other abort procedures can occur with the robotic controller.

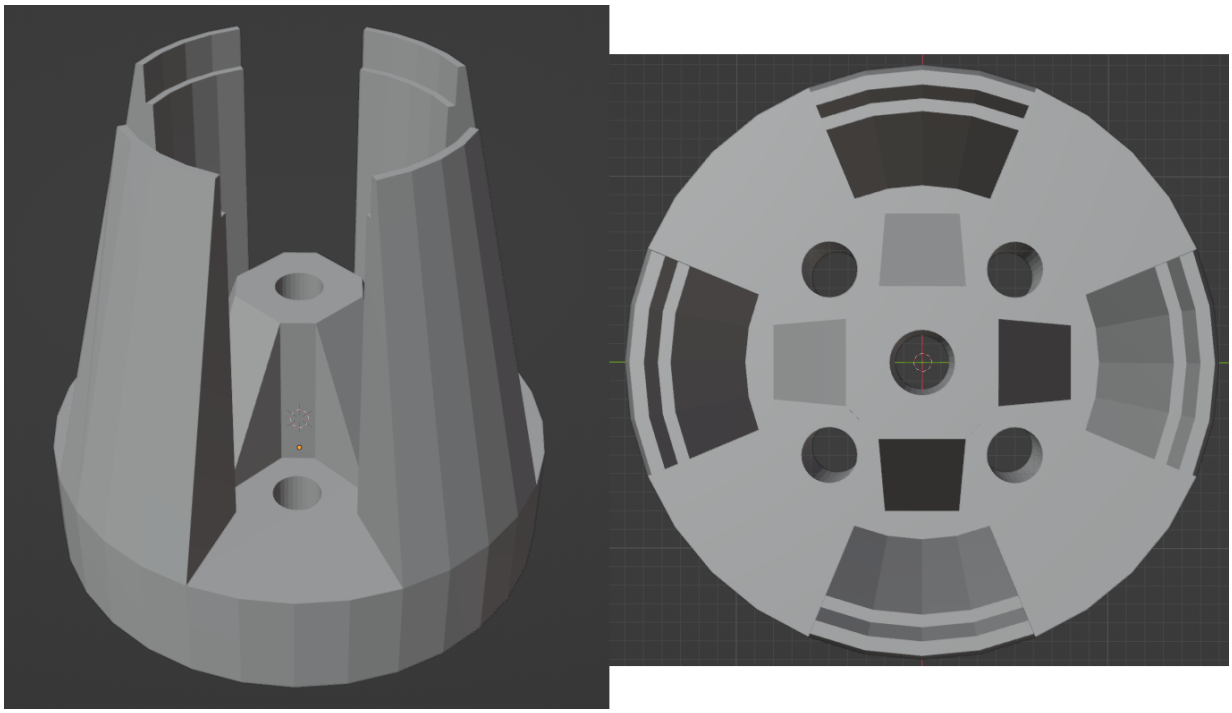


Figure 3.8: View of the interface between the loadcell, electromagnet, and swab holder. Left: isometric view. Right: top-down view.

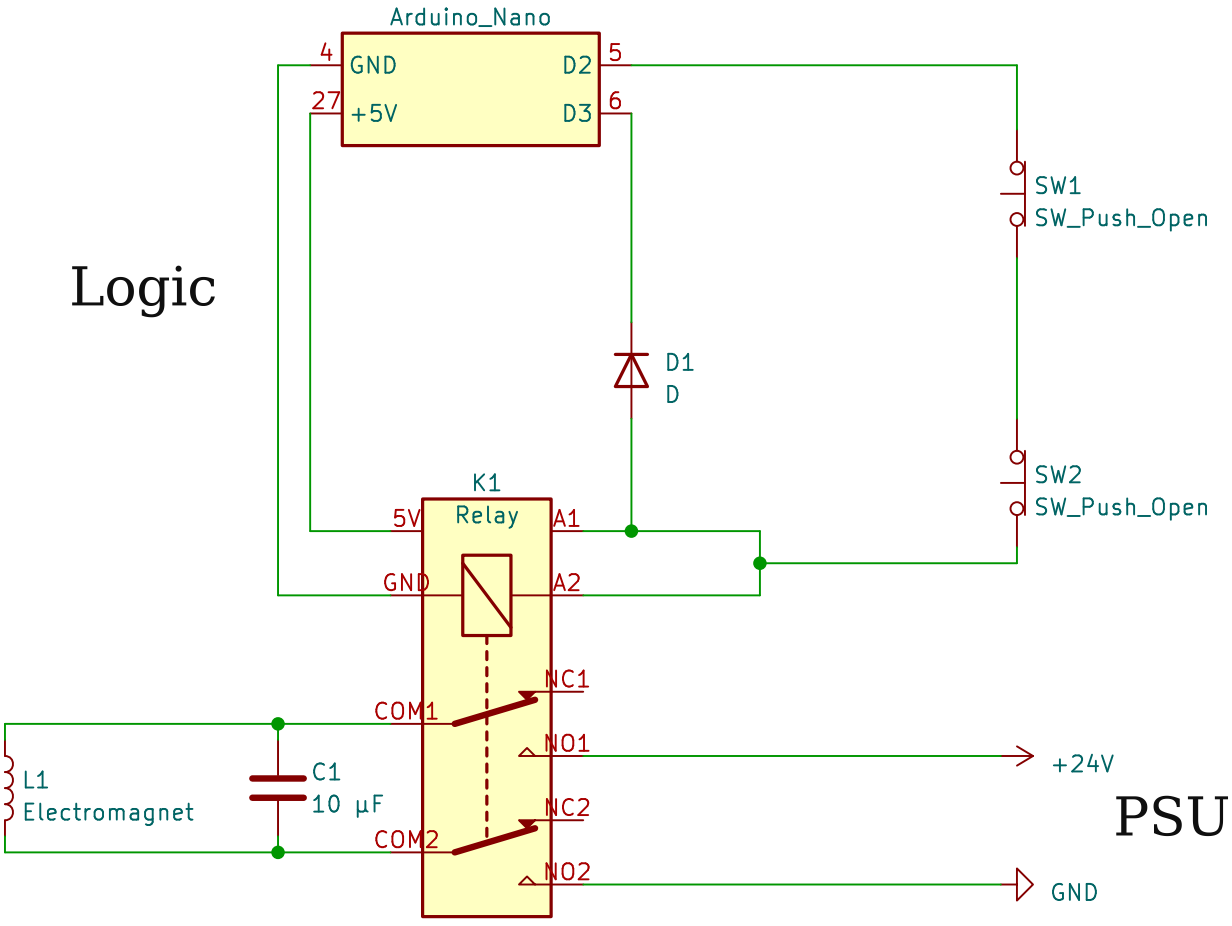


Figure 3.9: Circuit diagram of electromagnet controller.

3.4 Computational resources

In this section the computational resources used in this thesis will be described. These can be divided into those used for real time control and those used for data storage and analysis.

The computers used for real time control and experiments were manufactured by Lenovo (CPU: Intel i7-9750HF 2.60 GHz, GPU: NVIDIA GeForce GTX 1650) and Dell (CPU: Intel i7-10750H 2.60 GHz, GPU: NVIDIA GeForce RTX 2070) and were directly connected to the control box via Ethernet during experiments. Both of these systems had Ubuntu 20.04 installed using the RT_PREEMPT kernel patch. This kernel patch was necessary to meet the latency constraints of the FCI. Both computers also required an onboard graphics processing unit (GPU) to run the deep learning models used in the vision processing pipelines in Chapters 5 and 7. The Dell PC was required for the experiments in Chapter 7 because of the increased requirements to run the deep learning models in the pipeline.

A stationary tower PC (CPU: AMD Ryzen 7 3700X, GPU: NVIDIA GeForce GTX 1060) was also utilized offsite from the robots, but still connected to the University network. This computer hosted a PostgreSQL database that was critical to the operation of the experiments. The first function of the database is holding different controller configurations that could be queried by the control PCs. Both the controllers during the pre-contact and contact phases had many different parameters in terms of gain and filter values that needed to be tuned. The database provided a convenient way to store and index different configurations within a relational schema during experiments. The second function is storing the results from experiments. In addition to storing raw forms of data (e.g., images, videos, robot states, control inputs), the database can store contextual data (e.g., nostril side, controller ID, participant number) that are needed to comprehend the results of the experiments. Any manual observations that are made during experiments can also be easily linked to the results with primary key indexes.

Chapter 4

Swab insertion simulation & Trajectory optimization

Foreward

The contents within this chapter was accepted and published in the conference proceedings for the 2022 International Conference for Biomedical Robotics and Biomechatronics: P. Q. Lee, J. S. Zelek, and K. Mombaur "Simulating and Optimizing Nasopharyngeal Swab Insertion Paths for use in Robotics". In 2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob). 2022
DOI: 10.1109/BioRob52689.2022.9925302. [69] ©2022 IEEE. Some minor edits, such as formatting, have been made to accomodate the thesis format.

In the context of this thesis, this work describes the methodology that is used to derive insertion trajectories through the nasal cavity. This work is used primarily for planning and generating trajectories for contact phase of the test, as we will later show in Chapter 7.

4.1 Introduction

The PCR test via nasopharyngeal swab (NPS) specimen is currently well known as the gold-standard method for diagnosing COVID-19 [64] [32], in addition to being a standardized method for diagnosing many other types of diseases, including various strains of coronavirus, influenza, whooping cough, pneumonia, and rhinovirus [68]. The procedure

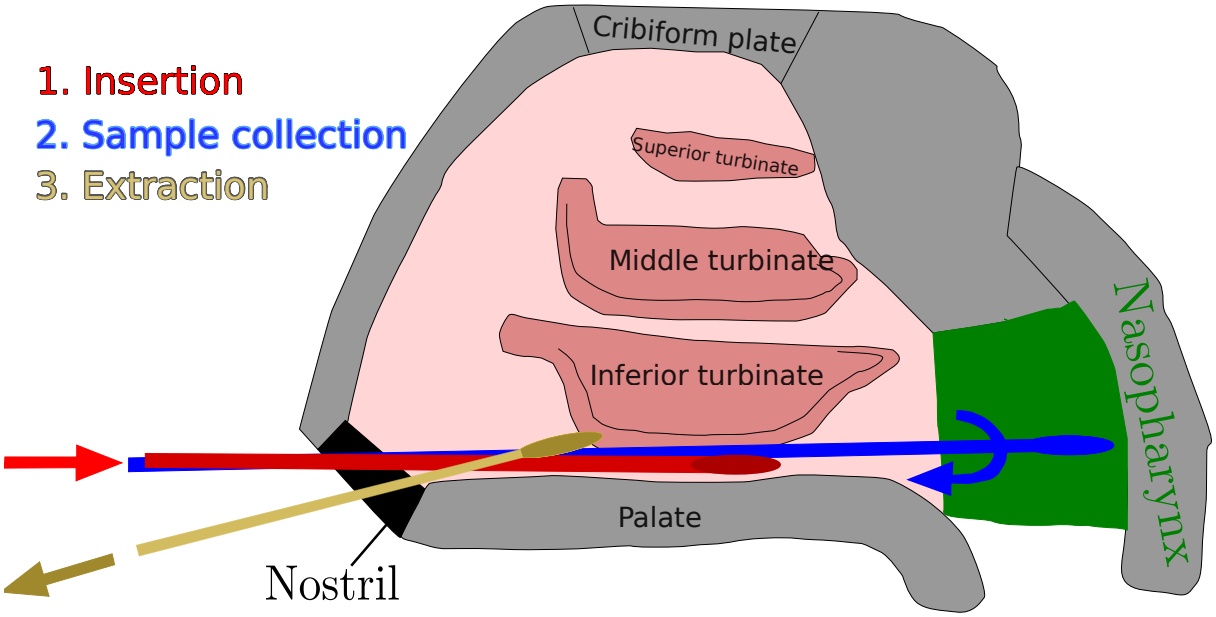


Figure 4.1: The three stages of a nasopharyngeal swab collection. This paper will be focusing on the first; the insertion stage.

of NPS specimen collection is to insert a flocked swab through the nasal cavity and collect samples from the nasopharynx [98]. The anatomy of the nasal cavity provides a number of challenges in terms of geometry. The swab has to navigate through the nasal cavity, requiring a correct angle of insertion that can reach the nasopharynx, while also avoiding becoming wedged on the turbinates in the nasal cavity. The swab also cannot be visually observed once it is inserted; thus the healthcare worker must be guided entirely on prior knowledge and tactile feedback.

Currently, most instructions for obtaining NPS specimens are based on qualitative instructions given to healthcare workers that can vary significantly between different health authorities [50]. Piras *et al.* [98] suggest inserting the swab with a pitch angle that is parallel to the palate at a distance equivalent from the nostril to the ear, but do not provide any yaw angle instructions. Piras *et al.* [98] provide similar insertion strategies. In correspondence [85], one reviewer suggested that the swab should be directed from the nostril towards the earlobe, although it is not clear if this was a suggestion for yaw orientation or if it was solely meant as pitch guidance.

Generally NPS collection consists of three of distinct phases, as shown in Fig. 4.1. First, the swab is inserted through the nasal cavity until it reaches nasopharynx. Second,

the swab is rotated to collect a specimen at the nasopharynx. Third, the swab is extracted and prepared for storage. This paper will solely focus on the insertion phase, where we will study the best way to insert the swab through the nasal cavity in a quantitative manner.

An overarching topic that is of rising interest is the application of robots in close-contact healthcare tasks, such as NPS tests, as a way to protect human healthcare workers from the risk of infection posed by their patients [33]. Robots can also ease the burden of the healthcare system during labor disruptions or during health crises. The COVID-19 pandemic in particular has spawned several recent works concerning robots performing swab tests. Zhang *et al.* [141] designed a remote centre compliance manipulator with a vision and force controlled system to perform oropharyngeal swab tests through the open mouth. Wang *et al.* [132] designed a custom end-effector that fits a swab that applies linear actuation and rotation along the axis of insertion towards the nasopharynx. Axial displacement of the swab is read by an optoelectronic sensor to measure force in the axis of the insertion. Currently, there is a deficiency in the literature towards planning or simulating insertion trajectories of the swab through the nasal cavity. From the standpoint of medical robotics, the study of these insertion trajectories will be an essential step towards planning and executing NPS tests on patients.

There are, however, a variety of other tasks in robotics that have some similar characteristics. Like a swab, a needle can also be characterized as a long, slender beam that shares the characteristic of deforming under pressure [38]. Needle insertion models can account for the tissue force response with a variety of a different linear elastic or viscoelastic models [137]. However, it should be noted that the contact physics differ between a swab and needle in significant ways. Namely, a needle actively penetrates tissue, whereas a swab does not; and a needle generally has a well defined contact phase, while a swab has multimodal contact phases due to the geometry of the nose. Peg-in-hole tasks share a similar objective of inserting an object into a receptacle while minimizing reaction forces from contact. Nonetheless, most peg-in-hole works are based on the assumptions of rigid body contacts with well-defined contact phases [112] so that most new developments are made towards online control and state-estimation rather than planning [136].

With this context, the contribution of this work is the following:

1. Formulate a simulation framework to model the motion and deformation of a swab as it is inserted into the nasal cavity (Section 4.2);

2. Derive a trajectory optimization problem to find ways to insert the swab into the nasal cavity in a manner that minimizes the amount of potential energy caused by the swab deforming, as a proxy for minimizing patient discomfort (Section 4.3); and

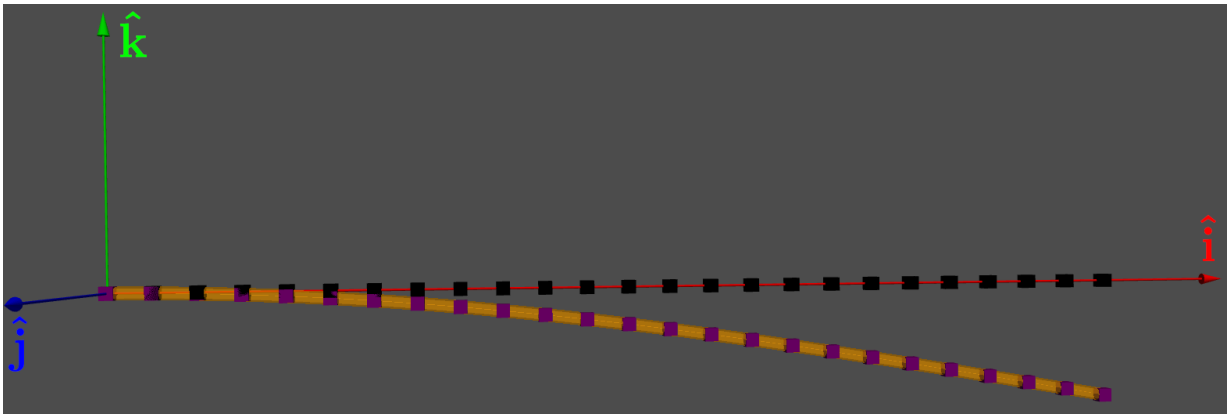


Figure 4.2: Visualization of beam, in orange, aligned to axis $\hat{\mathbf{i}}$ that is deformed along axis $\hat{\mathbf{k}}$. The purple cubes are the finite elements displaced from their reference positions in black.

3. Analyze what paths through the nasal cavity are feasible in terms of reaching the nasopharynx and the accumulated amounts of potential energy (Section 4.4).

4.2 Simulation

The two main components of the simulation consist of the swab deformation and the collision response between the swab and the nasal cavity. We model the swab as a cantilevered beam whose base can translate along a reference axis. For the collision response we consider a linear elastic model with Coulomb friction. The simulation code was implemented in Python using the Scipy numerical ecosystem with sparse matrix support. The collision detection tree was implemented using the Rust programming language. The visualized results were rendered using PyVista [123].

4.2.1 Beam dynamics

The mechanics of beams have been well studied in a variety of fields. The most basic beam model is the Euler-Bernoulli beam, which models 3D deformations with five degrees of freedom related to deformation along the axis it is aligned with and the two perpendicular axes [12]. Other beam models exist, such as Timoshenko beams, which add degrees of freedom to account for shearing [125]. Nasopharyngeal swabs are thin compared to its

length; following experiments from Zohoor and Kakavand [144], it is reasonable to use an Euler-Bernoulli model over more complicated models to reduce computational complexity. The beam has axis $\hat{\mathbf{i}}$ aligned to its reference state, with transverse axes $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$. Hence, the deformation of the beam can be described by the displacement vector $\chi(x) = u(x)\hat{\mathbf{i}} + v(x)\hat{\mathbf{j}} + w(x)\hat{\mathbf{k}}$ given x as the position of the undeformed beam on axis $\hat{\mathbf{i}}$. The degrees of freedom of the beam are taken as $s, v, w, \phi = \frac{\partial v}{\partial x}, \theta = \frac{\partial w}{\partial x}$, where s is the axial strain, while the independent transverse deformation variables are v and w, along with their spatial derivatives ϕ and θ. To conserve the total length of the beam $u(x) = s(x) - h_v(x) - h_w(x)$, where $h_v = \int_0^x (\frac{\partial v}{\partial x})^2 d\xi$, $h_w = \int_0^x (\frac{\partial w}{\partial x})^2 d\xi$.

In this paper, we will consider a simplified control scenario where the swab is grasped at one end and moved along a single axis, $\hat{\mathbf{i}}$. Such a scenario would be amenable to a robot inserting the swab by applying a constant force along $\hat{\mathbf{i}}$, while restricting motion in the transverse axes. These assumptions will allow us to later evaluate insertion paths through the nasal cavity, while keeping the degrees of freedom low. Zohoor and Kakavand [144] derived a set of partial differential equations to describe the deformation of a cantilevered beam with a non-fixed support, which we modify by removing the terms related to non-axial motion on the base and adding damping terms:

$$\begin{aligned} &\rho A\left(\ddot{s} - \left(\ddot{h}_v + \ddot{h}_w\right)\right) - EA\frac{\partial^2 s}{\partial x^2} \\ &= -\rho A(\ddot{x}_0) + F_u(x) + \alpha\dot{s}, \end{aligned} \tag{4.1}$$

$$\begin{aligned} &\rho A\left(\ddot{v} - \frac{\partial}{\partial x}\frac{\partial v}{\partial x}\int_L^x \left[\ddot{x}_0 + \ddot{s} - \ddot{h}_v - \ddot{h}_w\right] d\xi\right) \\ &+ EI_z\left(\frac{\partial^4 v}{\partial x^4}\right) = F_v(x) + \alpha\dot{v}, \end{aligned} \tag{4.2}$$

$$\begin{aligned} &\rho A\left(\ddot{w} - \frac{\partial}{\partial x}\frac{\partial w}{\partial x}\int_L^x \left[\ddot{x}_0 + \ddot{s} - \ddot{h}_v - \ddot{h}_w\right] d\xi\right) \\ &+ EI_y\left(\frac{\partial^4 w}{\partial x^4}\right) = F_w(x) + \alpha\dot{w}, \end{aligned} \tag{4.3}$$

where x_0 is the base position (i.e., the grasp point), and hence $\ddot{x}_0$ is the acceleration of the base. We assume that the swab is a perfect cylinder, so that the cross sectional area $A = \frac{1}{2}\pi r^2$, and area moments of inertia $I_y = I_z = \frac{1}{3}\pi r^3$. The swab's radius $r = 1$ mm, Young's modulus $E = 3 \times 10^9$ Pa, damping coefficient $\alpha = 10^{-2}$, and material density $\rho = 640$ kg m^{-3} are defined based on the swab shaft being made of polystyrene.

Following Chung and Yoo [31] and Zohoor and Kakavand [144], these PDEs are lin-

earized and discretized for simulation using a finite element modeling (FEM) strategy. This divides the beam into N evenly spaced nodes, as visualized in Fig. 4.2. The continuous deformation is interpolated with polynomial basis functions between these nodes. Given a pair of nodes i and $i+1$ located at x_i and x_{i+1}, given $\hat{x} = x - x_i$ and $l = L/(N-1)$, then the deformations are:

$$\begin{aligned}
s(x) &= (1 - \frac{\hat{x}}{l})s_i + \frac{\hat{x}}{l}s_{i+1} \\
v(x) &= \begin{bmatrix} v_i, & \phi_i, & v_{i+1}, & \phi_{i+1} \end{bmatrix} h(\hat{x})^\top \\
w(x) &= \begin{bmatrix} w_i, & \theta_i, & w_{i+1}, & \theta_{i+1} \end{bmatrix} h(\hat{x})^\top \\
h(\hat{x}) &= [1 - \frac{3\hat{x}^2}{l^2} + \frac{2\hat{x}^3}{l^3}, x - \frac{2\hat{x}^2}{l} + \frac{\hat{x}^3}{l^2}, \\
&\quad \frac{3\hat{x}^2}{l^2} - \frac{2\hat{x}^3}{l^3}, -\frac{\hat{x}^2}{l} + \frac{\hat{x}^3}{l^2}].
\end{aligned} \tag{4.4}$$

The weak-formulation of equations 4.1 to 4.3 become a second order ODE system

$$\overline{\mathbf{M}}\ddot{\boldsymbol{\kappa}} + \overline{\mathbf{D}}\dot{\boldsymbol{\kappa}} + \overline{\boldsymbol{\Lambda}}\boldsymbol{\kappa} = \overline{\boldsymbol{\Pi}}\overline{\mathbf{F}} + \boldsymbol{\Omega}\ddot{x}_0, \tag{4.5}$$

where $\overline{\mathbf{M}}$, $\overline{\mathbf{D}}$, $\overline{\boldsymbol{\Lambda}}$, $\boldsymbol{\Omega}$, and $\overline{\boldsymbol{\Pi}}$ are mass, damping, stiffness, acceleration, and force matrices found by summing equations 4.1 to 4.3, integrating, factoring the variables, and assembling. Also, we take $\boldsymbol{\kappa} = \{s_2, v_2, \phi_2, w_2, \theta_2, \cdots, s_N, v_N, \phi_N, w_N, \theta_N\}$. To match our control scenario of applying constant force f_a in the $\hat{\mathbf{i}}$ direction, we take the base position x_0 into the state-space and refactor the matrices as

$$\begin{aligned}
\mathbf{M} &= \begin{bmatrix} \frac{\rho A l}{2} & \mathbf{0} \\ -\Omega & \overline{M} \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} \alpha_x & \frac{l\alpha}{6}\mathbf{0} \\ \mathbf{0} & \overline{D} \end{bmatrix}, \\
\boldsymbol{\Lambda} &= \begin{bmatrix} 0 & -\frac{AE}{l} & \mathbf{0} \\ \mathbf{0} & & \overline{\Lambda} \end{bmatrix}, \quad \boldsymbol{\Pi} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \overline{\Pi} \end{bmatrix},
\end{aligned} \tag{4.6}$$

where $\alpha_x = 0.8$ is the coefficient of friction added to the base to prevent unrealistic velocities during short contact-free durations. Because Equation 4.5 is now a linear time invariant ODE, there exists a closed form solution for the state transition between points in time.

We start by creating an equivalent augmented system

$$\begin{aligned}
&\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{\Lambda} & -\mathbf{M}^{-1}\mathbf{D} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{\Pi} \end{bmatrix} \\
&\mathbf{z} = \begin{bmatrix} x_0 & \boldsymbol{\kappa} & \dot{x}_0 & \dot{\boldsymbol{\kappa}} \end{bmatrix}^\top \qquad\qquad \mathbf{F} = \begin{bmatrix} f_a & \overline{\mathbf{F}} \end{bmatrix}^\top \\
&\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\mathbf{F}
\end{aligned} \tag{4.7}$$

Given a fixed time step Δt, we have a way to transition between an initial state $\mathbf{z}(t)$ to the future state $\mathbf{z}(t + \Delta t)$ along with a the force input $\mathbf{F}$, by finding the homogeneous state transition and input matrix

$$\begin{aligned}
&\mathbf{\Phi}(\Delta t) = \exp(\Delta t \mathbf{A}) \\
&\boldsymbol{\beta}(\Delta t) = \int_0^{\Delta t} \exp(\Delta t - \tau) d\tau \mathbf{B} = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}^\top \exp\left(\begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \Delta t \right) \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{B} \\
&\mathbf{z}(t + \Delta t) = \mathbf{\Phi}(\Delta t)\mathbf{z}(t) + \boldsymbol{\beta}(\Delta t)\mathbf{F}(t).
\end{aligned} \tag{4.8}$$

The matrices $\mathbf{\Phi}$ and $\boldsymbol{\beta}$ can be precomputed for the time step Δt for the duration of the simulation.

4.2.2 Contact dynamics

Coming up with a contact model first requires having a volumetric model of the nasal cavity. The basis of the volumetric model was taken as the mean nasal cavity STL file from Keustermans *et al.* [62]. However, a number of additional processing steps were needed to make it suitable as a collision volume. The original file represents the airway of the nasal cavity, so the convex hull was added to the surfaces to derive tissue boundaries. The nostrils are hollowed out by taking the boolean difference with a cylinder. Finally, the tetrahedral mesh of the cavity, for detecting swab vertices in collision, was derived using the TetWild program by Hu *et al.* [53] This process could be applied to other nasal cavity airways, although some manual intervention is needed to hollow out the nostrils.

If we let the nasal cavity's tetrahedral mesh and surface be represented with $\mathcal{N}$, the purpose of the collision model is to compute the forces $\overline{\mathbf{F}}$ that are applied to the swab as a result of these collisions. In this work, we will consider that the collision response forces will be the result of a linear elastic model with Coulomb friction. This results in a number of assumptions, for computational reasons:

1. First, deformation of the surfaces of the nasal cavity is not explicitly modeled in

simulation. Instead, the nasal cavity acts as a stationary force-field to repel the swab when it collides. Modeling tissue deformation would increase the number of states by many orders of magnitude. This would significantly slow down the simulation and collision detection and would also require having detailed knowledge of the tissue types (mucosa, bone, and cartilage).

2. Second, viscous friction was not modeled. In reality, the tissue has viscoelastic properties such that the force scales with the velocity. But this caused issues from numerical stiffness that currently prevented its inclusion in the model.

The points of collision were detected between the vertices of the swab's volume and the nasal cavity using a customized axis-aligned bounding box tree. When a collision is detected, a response force is applied on the i^{th} node of the beam as the sum of the spring's restitution force and the frictional force

$$\mathbf{F}(x_i) = \mathbf{F}_R(x_i) + \mathbf{F}_F(x_i). \tag{4.9}$$

The restitution force is proportional to the distance of penetration with the solid, and directed back towards the nearest surface:

$$\mathbf{F}_R(x_i) = -\lambda \mathbf{p}(x_i), \tag{4.10}$$

where $\mathbf{p}(x_i)$ is the vector of penetration. The spring constant λ is derived by taking into context the Young's modulus of elasticity E_t to reflect mucosa tissue that lines that nasal cavity. While there is limited sources detailing the physical properties of nasal mucosa, Chen *et al.* [26] surveyed physical properties of oral mucosa and found a wide variety of different moduli were in use. From this, we take $E_t = 10^6$ Pa. Deriving the spring constant we take $\lambda = \frac{AE_t}{l_t}$ where A is the cross section area of the beam and $l_t = 0.5$ mm is the base length of the spring.

Coulomb friction is implemented to resist motion perpendicular to the restitution force

$$\begin{aligned} \mathbf{F}_F(x_i) &= -\mu ||\mathbf{F}_R(x_i)|| \boldsymbol{\nu}_i, \\ \bar{\boldsymbol{\nu}}_i &= \dot{\boldsymbol{\chi}}_i - \dot{\boldsymbol{\chi}}_i^\top \left(\frac{\mathbf{p}(x_i)}{||\mathbf{p}(x_i)||} \right) \left(\frac{\mathbf{p}(x_i)}{||\mathbf{p}(x_i)||} \right), \\ \boldsymbol{\nu} &= \frac{\bar{\boldsymbol{\nu}}}{||\bar{\boldsymbol{\nu}}||} \end{aligned} \tag{4.11}$$

with $\boldsymbol{\nu}$ as a unit vector that is directed towards the component of node i's velocity that is perpendicular to the restitution force. The friction coefficient $\mu = 0.213$ was taken based

on denture-mucosa contact [26], although it is unclear how applicable this value will be for NPS contact.

4.3 Optimization

With the beam-nasal cavity simulation formulated, we will now describe an optimization problem to find good trajectories to move the beam through the nasal cavity. As mentioned in Section 4.2.1, we will consider a control scenario where the base of the swab is moved along a linear trajectory, parameterized by a line segment ($\mathbf{s}$,$\mathbf{e}$), by applying a constant force f_a, where $\mathbf{s}$ is the point on an ellipse $\mathcal{S}$ on the nostril, where the head of the swab starts at, and $\mathbf{e}$ is the point on the ellipse $\mathcal{E}$ where the target nasopharynx region is. The three variables $\mathbf{s}$, $\mathbf{e}$, f_a are therefore the inputs to the simulation.

Along with reaching the nasopharynx, one aspect we want to minimize during the swab insertion is patient discomfort from the swab impacting the tissue. We quantify this as an objective by computing the amount of strain potential energy that the beam stores as it deforms from collisions with the nasal cavity during the simulation. Therefore, the averaged strain potential energy during the simulation, $\frac{1}{T}\int_0^T \mathbf{U}dt$, is used as an objective in the optimization problem to seek trajectories that result in low deformation of the beam. Within the finite element model outlined in Section 4.2.1, the strain potential energy for a beam state is

$$\begin{aligned} U &= \frac{1}{2}E\int_0^L \left(A\left(\frac{\partial s}{\partial x}\right)^2 + I_z\left(\frac{\partial^2 v}{\partial x^2}\right)^2 + I_y\left(\frac{\partial^2 w}{\partial x^2}\right)^2\right) dx \\ U &= \boldsymbol{\kappa}^\top \mathbf{Q}\boldsymbol{\kappa}. \end{aligned} \tag{4.12}$$

The energy matrix $\mathbf{Q}$ comes from assembly of the strain potential energy between adjacent

FEM nodes for components s, v, and w respectively:

$$\begin{aligned}
\mathbf{Q}_s &= \frac{EA}{2l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\
\mathbf{Q}_v &= \frac{EI_z}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix} \\
\mathbf{Q}_w &= \frac{EI_y}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}.
\end{aligned} \tag{4.13}$$

Algorithm 1 is written to simulate the beam-cavity interaction scene initialized with trajectory (**s**,**e**) and applied force f_a, ultimately returning the average strain energy that occurs during simulation or a penalty if the nasopharynx is not reached within a time limit. On top of this, COBYLA [99] is applied to algorithm 1 to optimize over the trajectory path $(\mathbf{s}, \mathbf{e})$. COBYLA was chosen over derivative based optimizers because the large number of states and function evaluations that occur during the optimization makes computing derivatives with respect to the inputs infeasible. The applied force f_a is fixed during the optimization.

4.4 Results & Discussion

Before examining the optimized trajectory outputs from applying Section 4.3, it is useful to examine how the magnitude of the applied force f_a affects the success rate of the swab reaching the nasopharynx, within a time limit of 7 s. We performed 4000 simulations while sampling on $\mathbf{s} \in \mathcal{S}, \mathbf{e} \in \mathcal{E}$, $f_a \in [0.001 \text{ N}, 0.2 \text{ N}]$. Fig. 4.3 shows that increasing f_a increases the chances of a randomized trajectory reaching the nasopharynx, up until about 0.05 N after which there is little change. This indicates that increasing the applied force can allow the swab to overcome the collisions that occur in the nasal cavity. However, Fig. 4.4 shows that the trade-off of increasing f_a is that the response forces applied to the swab increase as well. This is an important to note because these higher levels of force will cause a greater impact on the tissue that could presage patient discomfort.

Next, the COBYLA [99] algorithm optimized the outputs of algorithm 1 with f_a ranging from 0.02 to 0.18 N. Overall, the optimization was quite successful in terms of finding

Algorithm 1: sim($\mathbf{s}$,$\mathbf{e}$,f_a)

Set axis $\hat{\mathbf{i}} = \frac{\mathbf{e}-\mathbf{s}}{||\mathbf{e}-\mathbf{s}||}$;
Initialize $\mathbf{z} = \mathbf{0}$;
$t = 0.0$;
energy = 0.0;
contact = false;
while *!contact and* $t <$ *time_limit* **do**
    energy = energy + $\Delta t * \boldsymbol{\kappa}^\top \mathbf{Q}\boldsymbol{\kappa}$ /* Equation 4.12 */
    $\overline{\mathbf{F}}$ = compute_collision_forces($\mathcal{N}$, $\mathbf{z}$) /* (Equation 4.10 & 4.11) */
    $\mathbf{F} = (f_a, \overline{\mathbf{F}})$;
    $\mathbf{z} = \boldsymbol{\Phi}(\Delta t)\mathbf{z} + \boldsymbol{\beta}(\Delta t)\mathbf{F}$; // Equation 4.8
    $t = t + \Delta t$;
    contact = check_nasopharynx_contact($\mathbf{z}$, $\mathcal{N}$) ; /* true if swab head reaches nasopharynx */
if *contact* **then**
    return energy/t;
else
    return 1000.0 // failure penalty

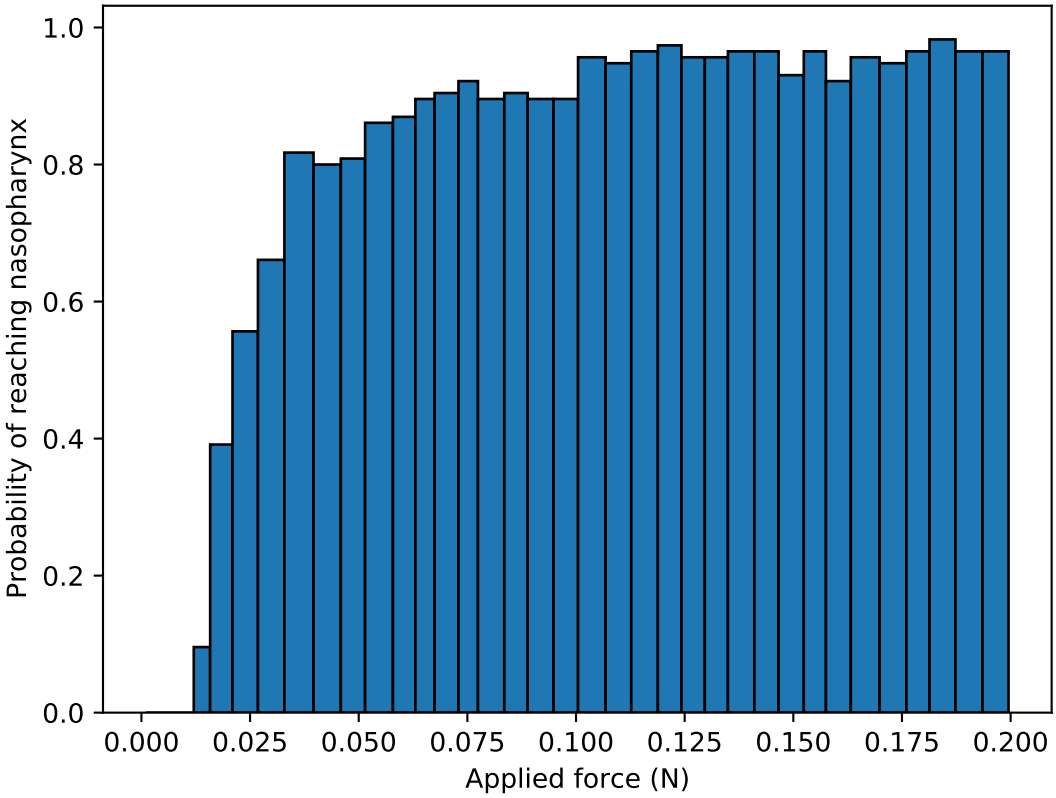


Figure 4.3: Bar plots showing that the probability of a randomized trajectory reaching the nasopharynx with respect to the applied force.

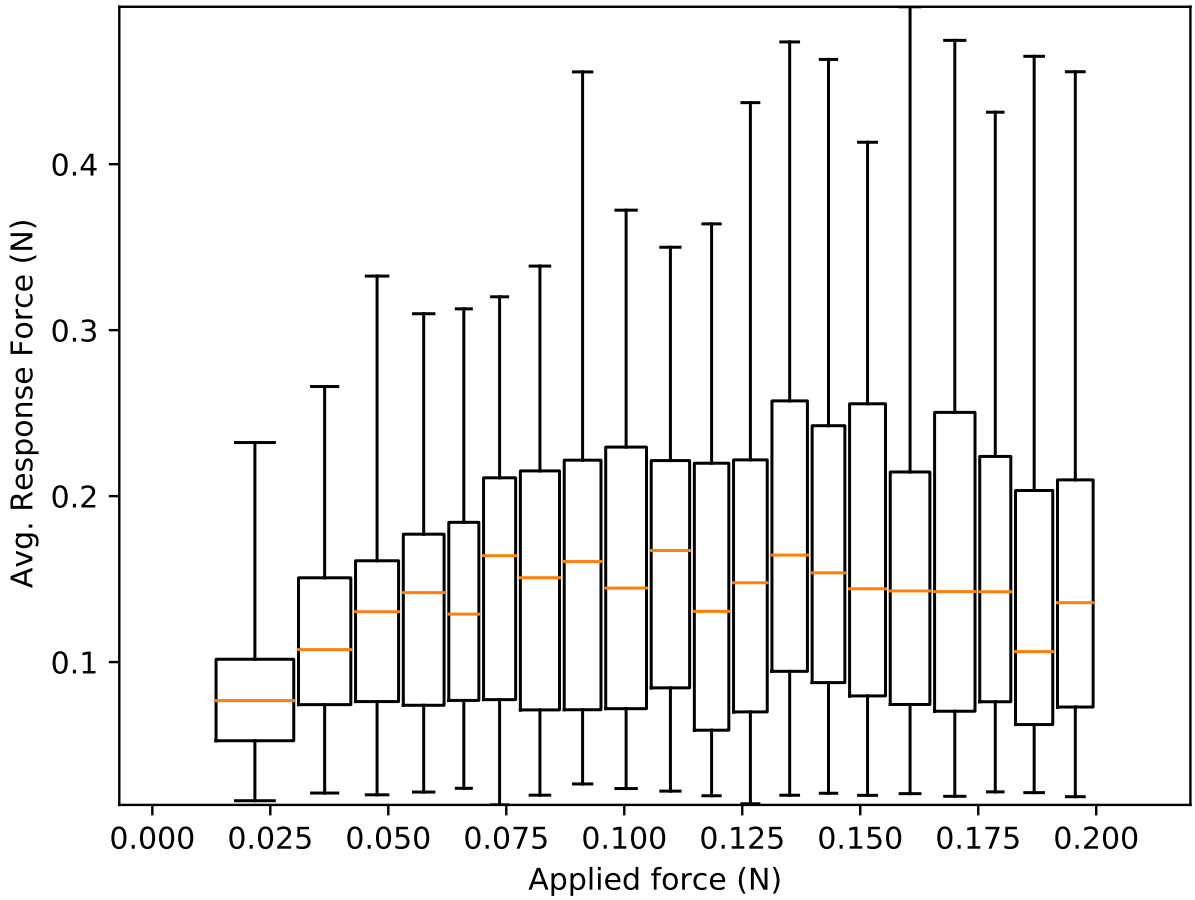


Figure 4.4: Box plots showing the distribution of the average response collision forces with respect to applied force. Boxes extend one quartile from the median, whiskers span the rest of the range.

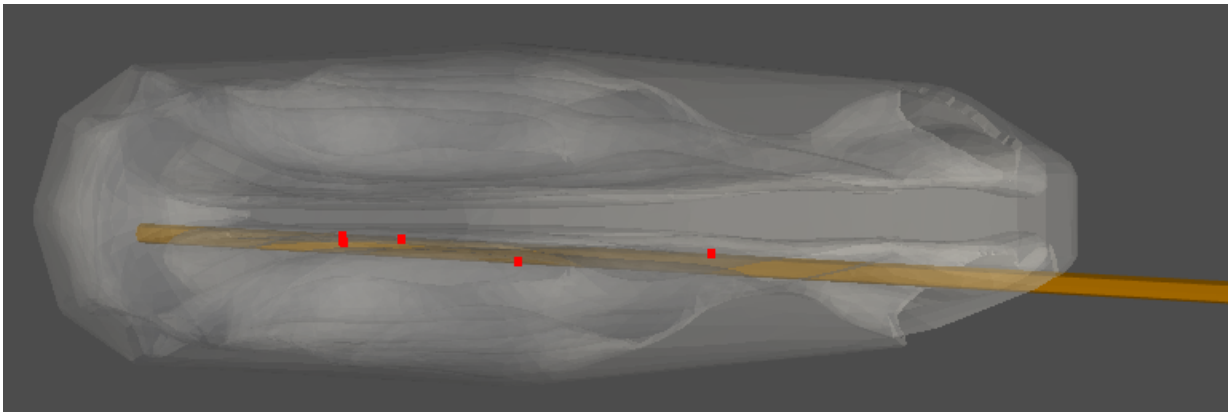
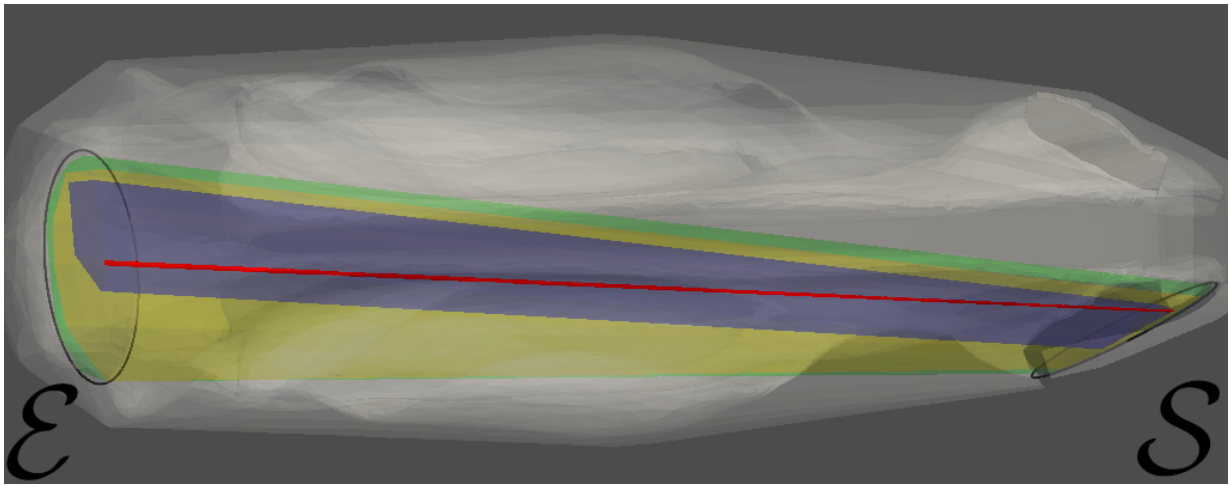


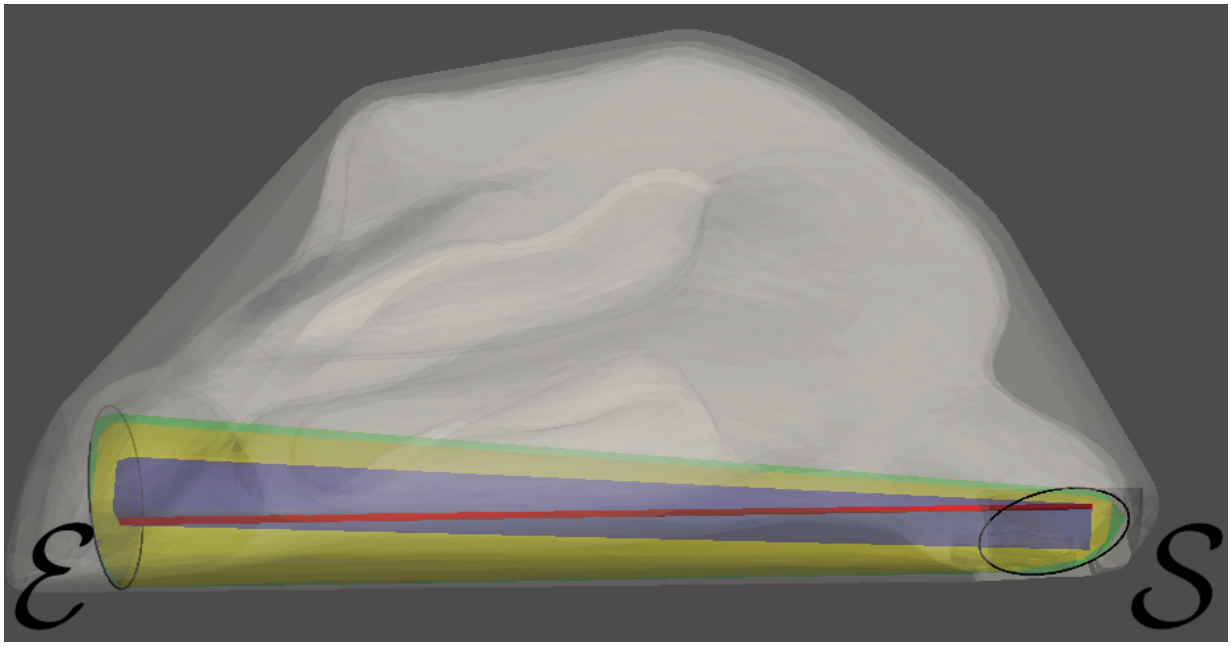
Figure 4.5: Top view rendering of the swab (orange) inserted into the nasal cavity after following the COBYLA optimized trajectory using $f_a = 0.08$ N. Even in this optimized trajectory that results in minimal deformation, some collisions still occur (red).

trajectories that minimized energy, with strain energy values in the range of 1.6×10^{-5} J to 3.0×10^{-4} J. The average trajectory had a pitch angle 0.25° down, with a yaw angle 1.63° towards the septum, located 4.68 mm from the columella (about one-third the distance to the ala). Fig. 4.5 shows a rendering of the swab within the nasal cavity. Even in this optimized insertion, contact with the intermediate nasal cavity is inevitable, which causes some small amounts of strain on the swab. Fig. 4.6 shows the span of the optimized trajectories through the nasal cavity, among the spans of other non-optimized trajectories that reached the nasopharynx at different force intervals. The figure shows that the optimized trajectories converged to similar paths over the set of different forces. The pitch angle of the paths seem to match clinical advice [98] in the sense that the swab is angled parallel to the palate. In terms of yaw the swab is angled slightly inwards from the nostril towards the nasal septum, which largely allows it to avoid making contact with the adjacent inferior turbinate. Also plotted in Fig. 4.6 are the spans for the trajectories in the upper bounds of f_a at 0.015, 0.025, and 0.05 Newtons, showing how the set of feasible trajectories through the nasal cavity expands as more force is applied.

A deeper analysis of the optimized trajectory can be found by examining the change in strain potential energy objective as the trajectory is perturbed from the optimal trajectory $(\mathbf{s}^*, \mathbf{e}^*)$ (c.f. Fig. 4.7). In Fig. 4.8, the contour plots of the strain potential energy display this for $f_a = 0.08$ N. Fig. 4.8a was found by rotation about $\mathbf{s}^*$, while Fig. 4.8b was found by rotation about $\mathbf{e}^*$. In the former case, there is slightly less sensitivity to downward pitch, whereas changes in yaw cause issues by hitting the septum or the nasal vestibule laterally. The latter case appears to be more sensitive to positive yaw, which



(a) Top view



(b) Side view

Figure 4.6: Convex hulls were taken of the pairs of points $s \in \mathcal{S}$ and at $e \in \mathcal{E}$ whose trajectories successfully reached the nasopharynx. The red hull shows the set of optimized paths found by COBYLA. The blue, yellow, and green hulls belong to trajectories that were able to reach the nasopharynx for $f_a \leq 0.015$, 0.025, and 0.05 Newtons respectively.

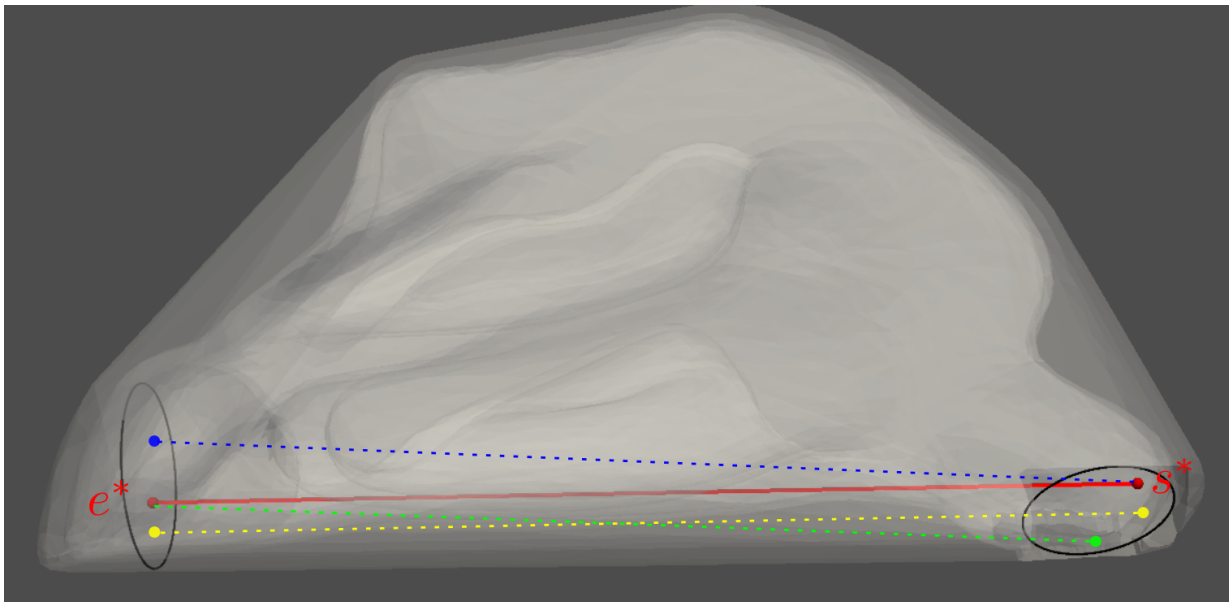


Figure 4.7: Spatial representation of how the optimal trajectory (red) was perturbed to achieve the contour plots in Fig. 4.8). Blue shows rotation about $\mathbf{s}^*$ (Fig. 4.8a). Green shows rotation about $\mathbf{e}^*$ (Fig. 4.8b). Yellow shows translation (Fig. 4.8c).

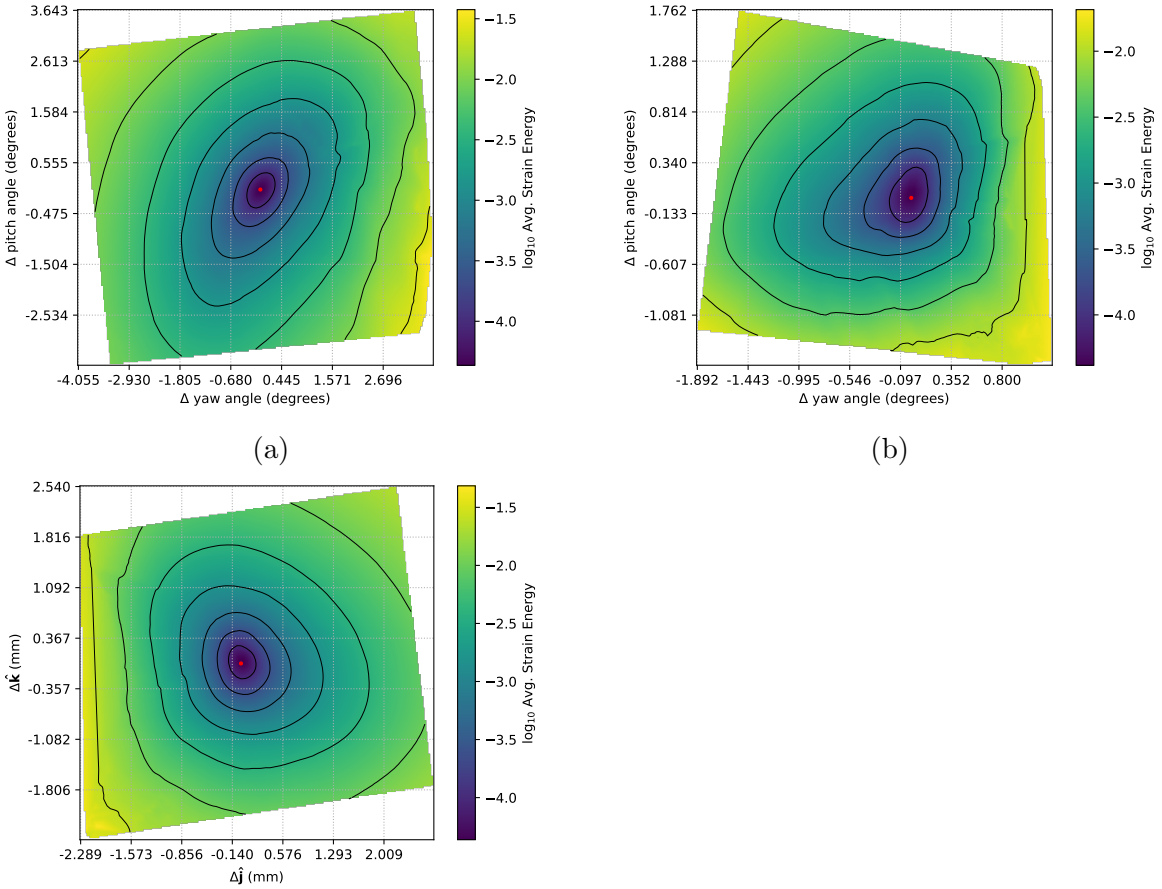


Figure 4.8: Contour plots showing the change in objective function as perturbations are made about the optimal trajectory, parameterized by the line segment ($\mathbf{s}^*$,$\mathbf{e}^*$), found by COBYLA (annotated with a red dot). Fig. a) perturbs the trajectory by rotation about $\mathbf{s}^*$. Fig. b) perturbs the trajectory by rotation about $\mathbf{e}^*$. Fig. c) perturbs the trajectory through translation. A total of 6400 points were considered for each plot and filled in with linear interpolation; any points that did not reach the nasopharynx were excluded.

leads to contact with the nasal vestibule and inferior turbinates. In contrast, Fig. 4.8c shows the effect of keeping the same angle of insertion, but translating the entire trajectory in the transverse directions. Here, the objective rises steepest with horizontal translation towards the exterior of the cavity, where it will encounter the nasal vestibule wall and inferior turbinate. Overall, these plots show that the objective function has a steep well at the optimum found by COBYLA that rises with small changes of angle or position; rising several orders of magnitude in less than a degree or within 2 mm. Of course, this trajectory is optimized for this single nasal cavity. In reality, anatomical variations and hardware limitations will make achieving this best-case scenario unlikely.

4.5 Conclusion

In this work, we developed a model to simulate the interaction between a swab and nasal cavity for nasopharyngeal swab insertion. Leveraging this simulation, we formulated an optimization problem to find an optimal insertion trajectory of 0.25° pitch down, 1.74° yaw inwards, located about one-third between the center and edge of the nostril. We found that trajectories got significantly worse when perturbed away from the optimum. Despite this we found that many non-optimal trajectories could still reach the nasopharynx provided sufficient force was applied to overcome collisions.

There are a number of potential avenues for future work. As this is our initial work, we only used a single nasal cavity model for the simulations. We would like to extend our work to a variety of different nasal cavity shapes to better study insertion trajectories over the population as a whole. This would include investigating how nasal cavity abnormalities, such as deviated septums, affect the procedure. Another area to build on would be adding viscosity properties to the nasal tissue. Simulating the impact of inflamed tissue and its impact on the insertion would also be of particular interest because this could be a common symptom of ill patients who are in need of these tests.

In the physical world, a robot would need to have an intricate controller that can balance responding to disturbances like motion and variance in anatomy, while also following a baseline trajectory. Our work helps for planning this baseline trajectory through the nasal cavity, with possible extensions of evaluating control systems in the proposed simulation environment.

Chapter 5

Face pose estimation & Visual servo system

Foreward

The contents within this Chapter are written originally for this thesis, although we plan to submit its contents to a journal at a later date. Chronologically, the experiments in this research were conducted from July 2023 to February 2024.

5.1 Introduction

This chapter describes the visual servo system that is solely responsible for executing the pre-contact phase of the NP swab test. The goal of the system is to align and position the swab next to the nostril in preparation for the contact phase, using image data from an eye-in-hand RGB-D camera. This is a crucial part of the overall process because the nostril is quite small, and we noted in Chapter 4 that orientation plays a significant role in ensuring a comfortable contact phase. In the proposed scenario, the patient will approach the robot and stand unrestrained within the arm's workspace. This adds additional complexity because the arm must be able to move to and work around faces within its workspace without running into joint limits. The system also needs to be robust to the natural motion of the patient and compensate if measurements are temporarily lost.

This research introduces several novel contributions to the problem of robotic NP swab testing. First, we developed an eye-in-hand visual servo system that regulates both position

and orientation to ensure precise alignment of the swab with respect to the nostril. Second, our system is specifically designed to accommodate the challenges presented by freestanding patients, allowing for dynamic adjustments based on natural head movements without the need for restrictive fixtures. The system also contains components to manage the workspace challenges this scenario poses. This adaptability enhances the system's usability in real-world clinical settings. Finally, we conducted human trials involving participants from diverse demographic backgrounds to both validate the effectiveness of our system and to examine for potential demographic biases in its visual models.

In this chapter, we first examine the literature for estimating the pose of the face and enabling visual servo in Section 5.2, which has been crucial for inspiring the methods in the pre-contact phase pipeline. Our proposed pipeline's methods consist of five main groups of components that are described in Section 5.3. To meet the constraints imposed by the unrestrained participant's head, we create a joint lookup table to find joint configurations for the arm that places the end-effector near the participant's face while ensuring that the arm can still move around the face without encountering joint limits. A measurement pipeline is designed using a convolutional neural network (CNN) to estimate the head orientation and nostril positions [46], which are also resolved using depth imaging. Stable states are retained by applying an unscented Kalman filter on manifolds (UKF-M) [20]. The control law of the system follows a pose-based visual servo that brings the swab to a desired pose in front of the nostril. We validate our approach through human trials with participants from a variety of demographic backgrounds in Section 5.4. The system performs well during trials, and we analyze differences in final pose among the population in Section 5.5. We also note methods that can be employed to make the system more robust for future trials. Finally, we summarize and conclude the work done in this chapter in Section 5.6.

5.2 Background on applied vision

In this section we provide background on the two main topics that are related to this chapter: computer vision models for facial pose estimation using monocular camera images and the visual servo framework that will actuate the arm during the pre-contact phase. These topics are drawn upon to realize the goal of aligning the swab with the correct orientation and position with respect to the nostril using the feedback from the RGB-D camera.

5.2.1 Facial pose estimation

Identifying facial landmarks and features is a very active area of research in computer vision. While fueled by applications like biometrics [107] and sentiment analysis [135], the developments in identifying lower level features from images are useful to our scenario. Facial pose estimation can typically be divided into two categories of methods [135], those that fit a template of landmarks based on regions of the face and those that fit a dense 3D model or mesh based on the appearance of the image.

The landmark template approach typically assumes some template of 2D keypoints that correspond to different regions of the face (e.g., lips, eyes). Most models will typically apply a machine learning model that trains on instances of labeled images to predict the locations of the template in new images [135] [102] [111] [9] [34] [126] [131]. However, a key weakness to this approach in the context of our scenario is that the template is only retained within the 2D projection of the image, which makes it difficult to recover the 3D pose of the head. While more recent methods attempt to project landmarks into three dimensions [139] [94] [9], the landmarks are sparse, so extrapolation to find arbitrary points of interest on the face (i.e., the nostril) is still a concern.

The dense 3D face modeling approach is generally more flexible and useful to our scenario because we want to obtain the 3D orientation and positions of arbitrary points like the nostrils on the face. The model by Feng *et al.* [42] trained a CNN on a UV wrapping of the face to predict the x,y, and z coordinates of the face. Jackson *et al.* [56] also trained a CNN to compute a volumetric regression of the entire face. Other methods may rely on having some 3D morphable model, such as the Basel face model (BFM) [96], which applied principal component analysis (PCA) to a sample of point clouds from laser scans of faces to allow them to be described by a lower dimensional subspace vector. Liu *et al.* [81] created a model that regressed BFM parameters using a CNN that incorporates SIFT keypoint features and visual contours. The model called 3D Dense Face Alignment (3DDFA) by Zhu *et al.* [143] is a CNN trained to predict BFM vectors along with rotation parameters based on images paired with sets of labeled 3D landmarks. Guo *et al.* [46] extended this work with their 3DDFA_V2 model by using a more lightweight neural network based on the MobileNet architecture [52] and adding additional terms in the objective function to prioritize different vertices. Fig. 5.1 shows an example of the resulting landmarks that come from applying 3DDFA_V2 to a face. As we will explain in Section 5.3.2, we ultimately chose to integrate 3DDFA_V2 into our pipeline.

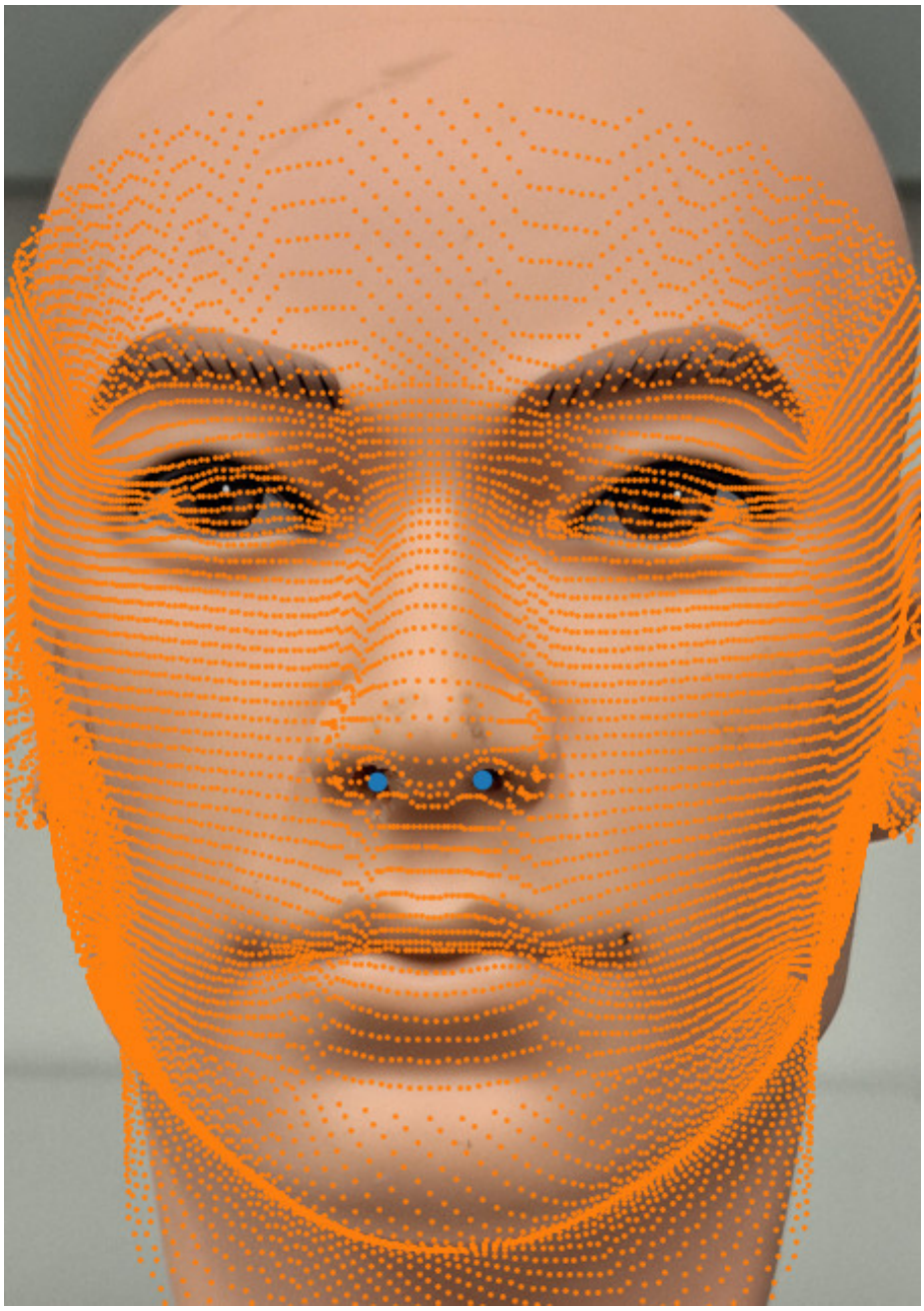


Figure 5.1: Example of the 3DDFA_V2 projecting a Basel face model template based on a monocular image of a head mannequin.

5.2.2 Visual servo control

Visual servo can be described as the synthesis of computer vision and control theory. Several different forms exist, but generally the camera and a vision pipeline provides a measured feature that is used by an underlying control law. The control law formulates an error term between the measurement and the goal, which is driven to zero with actuation commands. Generally, the two main flavors are image-based visual servo (IBVS) and pose based visual servo (PBVS), which is also sometimes called position based visual servo [25]. In IBVS the control problem is typically formulated to align some designated features within the 2D image plane, while PBVS projects data from the image into features in 3D Cartesian space. The benefits and drawbacks of both approaches vary greatly depending on the specifics of the problem.

Applications of visual servo to robotic swabbing is more limited, but there are a few works of relevance. Chen *et al.* [30] proposed an oropharyngeal swab system solution that uses a compliant mechanism with a visual module that uses a deep learning framework to segment the pharyngeal wall. Subsequently, the compliant swab is moved towards trajectory to reach the tonsils through the mouth. Li *et al.* [73] proposed a visual servo system for NP swabs surrounding a "hierarchical decision framework" for placing a grasped swab next to a nostril. Despite this, it is unclear what methodology is used to visually estimate the state of the nose or what advantages the approach has beyond doing inverse kinematics. Most relevant is the work by Hwang *et al.* [54] who developed a visual servo system to position swabs next to a participant's nostrils. The authors use an eye-to-hand (where the camera is placed separately from the manipulator) visual servo system and use IBVS to position the swab next to the nostril of a subject held in a head fixture. Ultimately, our work takes a different approach by using an eye-in-hand (where the camera is placed on the end-effector) PBVS approach that we believe will be beneficial to handling the arbitrary poses of the unrestrained, freestanding patients in our experiments. Also, we actively regulate the angle the swab approaches the nostril at, something that would have a major impact on swab insertion according to our results in Chapter 4.

5.3 Methods

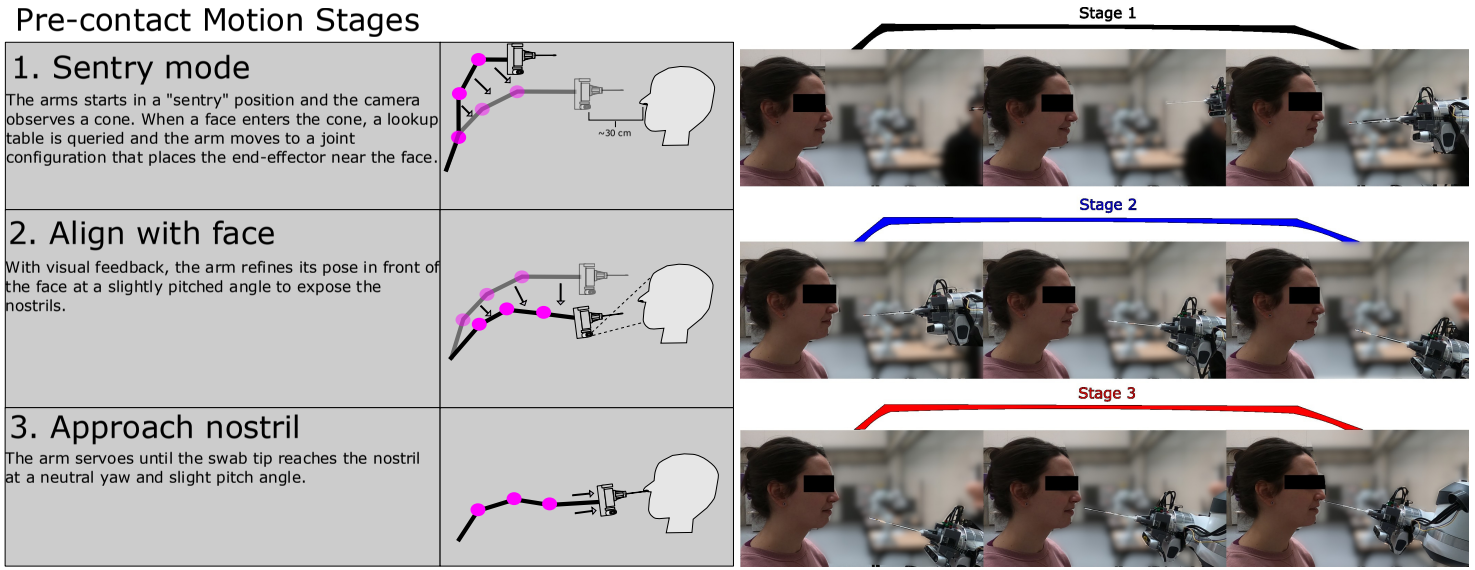
We divide the pre-contact phase into three distinct stages of motion, as shown in Fig. 5.2. During the first stage, the arm begins in a sentry position where the end-effector camera is oriented to observe the workspace in front of the arm. If a face is detected, a lookup table is queried and returns a joint configuration that places the end-effector near the face.

The second stage positions the camera 30 cm away (and the swab tip approximately 10 cm away) from the face at a camera frame pitch of 0.2 radians in order to expose the nostrils. The third stage moves the tip of the swab to the nostril, at a neutral yaw and 0.2 radian pitch angle. The system utilizes a face alignment CNN and 3D projection to measure the 3D pose of the face by use of an RGB image and depth image. During the second and third stages, these measurements are fed into a UKF-M state estimator filter to filter for noise and integrate commanded velocities. The third stage differs in its usage of an unsupervised image segmentation model, and a single object tracker (SOT) CNN to accurately track the nostrils during the final motions. A PBVS control loop computes command velocities for the arm based on the filtered estimates and was implemented with the ViSP framework [84]. The interaction of all of these components is ultimately portrayed in Fig. 5.3. The controller code was primarily written in C++, but was extended with Python to run the deep learning models in the visual pipeline and to run the UKF-M.

5.3.1 Joint lookup table

In this section we describe the resolution of the inverse-kinematics (IK) problem of finding an ideal joint configuration, $\mathbf{q}_{\text{start}}$, that places the end-effector in proximity to a face located somewhere within the arm's workspace. Ideally, with joint configuration $\mathbf{q}_{\text{start}}$, the robot would be capable of moving around the face at various directions without encountering joint limits or self-collisions. While there are many point-to-point IK algorithms for redundant manipulators available from the literature (e.g., RRT-Connect [65]), running them online without causing a significant delay to the motion sequence remains a concern. In addition, we would need to extend the methods to cover the area of possible target destinations we expect the arm to move to, rather than just move to a single destination point, because there will be some uncertainty in the direction the participant is facing. Therefore, we use a lookup table (LUT) that maps a target input pose to a suitable joint configuration that is generated with an empirically driven approach. The lookup table is convenient to use during runtime because it can be queried extremely fast by finding the table entry that is closest to the detected 3D position of the face via indexing an array. Our method for constructing this lookup table is detailed below.

The sampled points covered by the LUT are contained in a cone, $\mathcal{C}_{\text{start}} \in \text{SE}(3)$, that



(a) (b)

Figure 5.2: Overview of the motion sequences during the pre-contact phase. a) Illustration of the three stages of motion of the pre-contact phase. b) Example of the motions from an experimental trial.

extends from the field of view of the end-effector camera at the sentry position. Specifically,

$$\mathcal{C}_{\text{start}} = \left\{ \begin{array}{c} (x, y, z, \theta_x, \theta_y, \theta_z), \text{ such that} \\ z_{\min} \leq z \leq z_{\max},\ r_{\min} \leq r \leq r_{\max} \\ x = r\cos(\phi),\ y = r\sin(\phi),\ \phi_{\min} \leq \phi \leq \phi_{\max} \\ \theta_x = 0.2 \text{ rad},\ \theta_y = 0,\ \theta_z = \phi \end{array} \right\}. \tag{5.1}$$

with $\phi_{\min} = -45°$, $\phi_{\max} = 45°$, $r_{\min} = 0.48$ m, $r_{\min} = 0.68$ m, $z_{\min} = 1$ m, and $z_{\max} = 1.85$ m. A visual representation of this sampled cone is shown in Fig. 5.4. We can generate different $\mathbf{q}_{\text{start}}$ that place the end-effector in $\mathcal{C}_{\text{start}}$ so that $\boldsymbol{\kappa}(\mathbf{q}_{\text{start}}) = \boldsymbol{\epsilon}_{\text{start}} \in \mathcal{C}_{\text{start}}$ by applying the Jacobian pseudo-inverse method

$$\dot{\mathbf{q}} = \mathbf{J}^{\dagger}(\bar{\boldsymbol{\epsilon}} - \boldsymbol{\kappa}(\mathbf{q})) \tag{5.2}$$

where $\boldsymbol{\kappa}$ is the forward kinematics function, $\bar{\boldsymbol{\epsilon}}$ is the target point for the end-effector, $\mathbf{J}$ is the robot Jacobian, and $\mathbf{J}^{\dagger}$ is the pseudo-inverse $\mathbf{J}$. However, given that the Franka arm has joint redundancy, there is no unique solution $\mathbf{q}_{\text{start}}$ for any given $\boldsymbol{\epsilon}_{\text{start}}$ [82]. As a result, we initialize different candidate joint configurations through sampling from a random distribution, prior to applying (5.2).

With a valid set of $\mathbf{q}_{\text{start}}$s, we look to grade the quality of the configurations by attempting to move them to a destination zone. This destination space takes the form of a second cone

$$\mathcal{C}_{\text{end}} = \left\{ \begin{array}{c} (x, y, z, \theta_x, \theta_y, \theta_z), \text{ such that} \\ r = d\cos(\phi),\ z = d\sin(\phi),\ 0 \leq \phi \leq \phi_{\max} \\ x = r\cos(\zeta),\ y = r\sin(\zeta), 0 \leq \zeta \leq 2\pi \\ \theta_{\min} \leq \theta_x \leq \theta_{\max},\ \theta_{\min} \leq \theta_y \leq \theta_{\max},\ \theta_z = 0 \end{array} \right\}, \tag{5.3}$$

where $d = 0.35$ m (distance allocated to reach nostril and insert), $\phi_{\max} = 15°$, $\theta_{\min} = -10°$, and $\theta_{\max} = 10°$. For each starting pose $\boldsymbol{\epsilon}_{start}$, the points representing the workspace that we want to sample is $\boldsymbol{\epsilon}_{\text{end}} \in \mathcal{T}(\boldsymbol{\epsilon}_{start}, \mathcal{C}_{\text{end}})$ where $\mathcal{T}$ is a function that transforms $\mathcal{C}_{\text{end}}$ to the frame of reference of $\boldsymbol{\epsilon}_{\text{start}}$. Fig. 5.5 provides a visual representation of the sampled space extending from $\boldsymbol{\epsilon}_{\text{start}}$. The quality of the candidate $\mathbf{q}_{\text{start}}$ is measured by counting the number (N) of $\boldsymbol{\epsilon}_{\text{end}} \in \mathcal{T}(\boldsymbol{\epsilon}_{\text{start}}, \mathcal{C}_{\text{end}})$ that can be reached via (5.2) without encountering any joint limits or self-collisions. Ultimately, the candidate $\mathbf{q}_{\text{start}}$ with the highest N is stored in the LUT for the given $\boldsymbol{\epsilon}_{\text{start}}$. The process for synthesizing the LUT is summarized in Algorithm 2, and a visualization of the workspace surrounding the robot is shown in Fig. 5.4, with the maximum N for each entry reflected in shades from green to red. Our algorithm utilized the Pinocchio library [23], which enabled us to compute the Jacobian

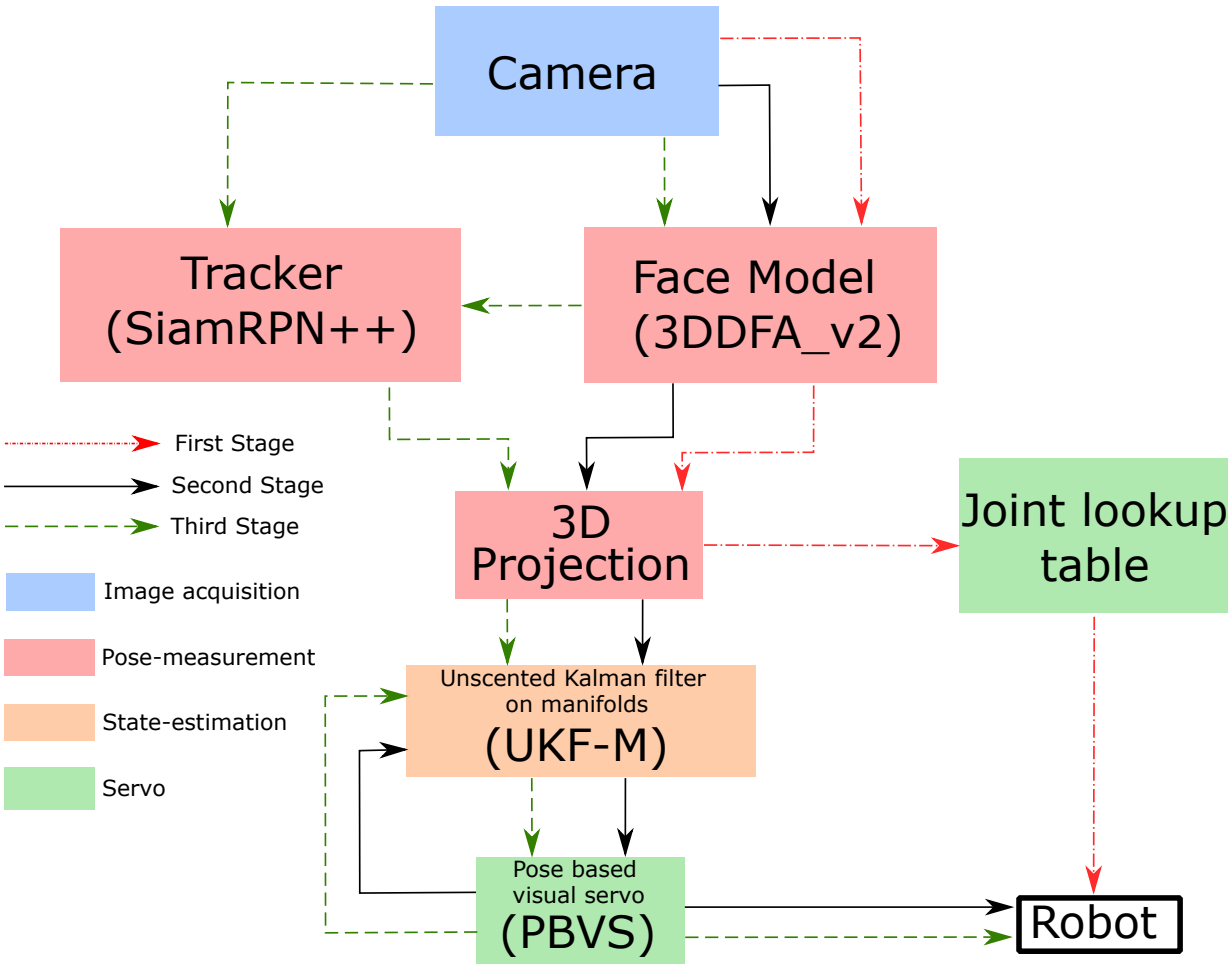


Figure 5.3: Flowchart of the components used throughout the three stages of the pre-contact phase.

for performing inverse kinematics and to check for self-collisions.. While generating the lookup table took several hours, querying the lookup table during runtime is extremely efficient, taking a few milliseconds. Since our workspace is well known ahead of time, the LUT is an effective solution to this problem.

Algorithm 2: Synthesis of joint lookup table

```
LUT = { } ;                                                  // Lookup table.
for ϵ_start ∈ C_start do
    best_N = -1;
    best_q = NULL;
    for i ∈ 1..150 do
        init_q = Uniform(7) ;          // Randomly initialize starting point.
        q_start = IK(ϵ_start,init_q) ;          // Find joint config using (5.2)
        N = 0;
        for ϵ_end ∈ T(ϵ_start, C_end) do
            if valid_IK(ϵ_end, q_start) then
                // Integrates (5.2), true if ϵ_end can be reached from
                   q_start without encountering joint limits or self
                   collisions.
                N = N +1 ;
        if N > best_N then
            // Select joint config with best success rate
            best_N = N;
            best_q = q_start;
    LUT[ϵ_start] = best_q ;
```

5.3.2 RGB image to facial keypoints and rotation (face model & tracker)

Based on the available methods in the literature, we chose to use the 3DDFA_V2 model by Guo *et al.* [46] as a major component of our facial pose measurement pipeline. This model was chosen because it is computationally fast to run and provides a way to directly estimate the rotation of the face as well as arbitrary keypoints of the face. It relies on an underlying Basel face model [96] that is built upon a PCA matrix, $\mathbf{W}$, so the 3D vertices

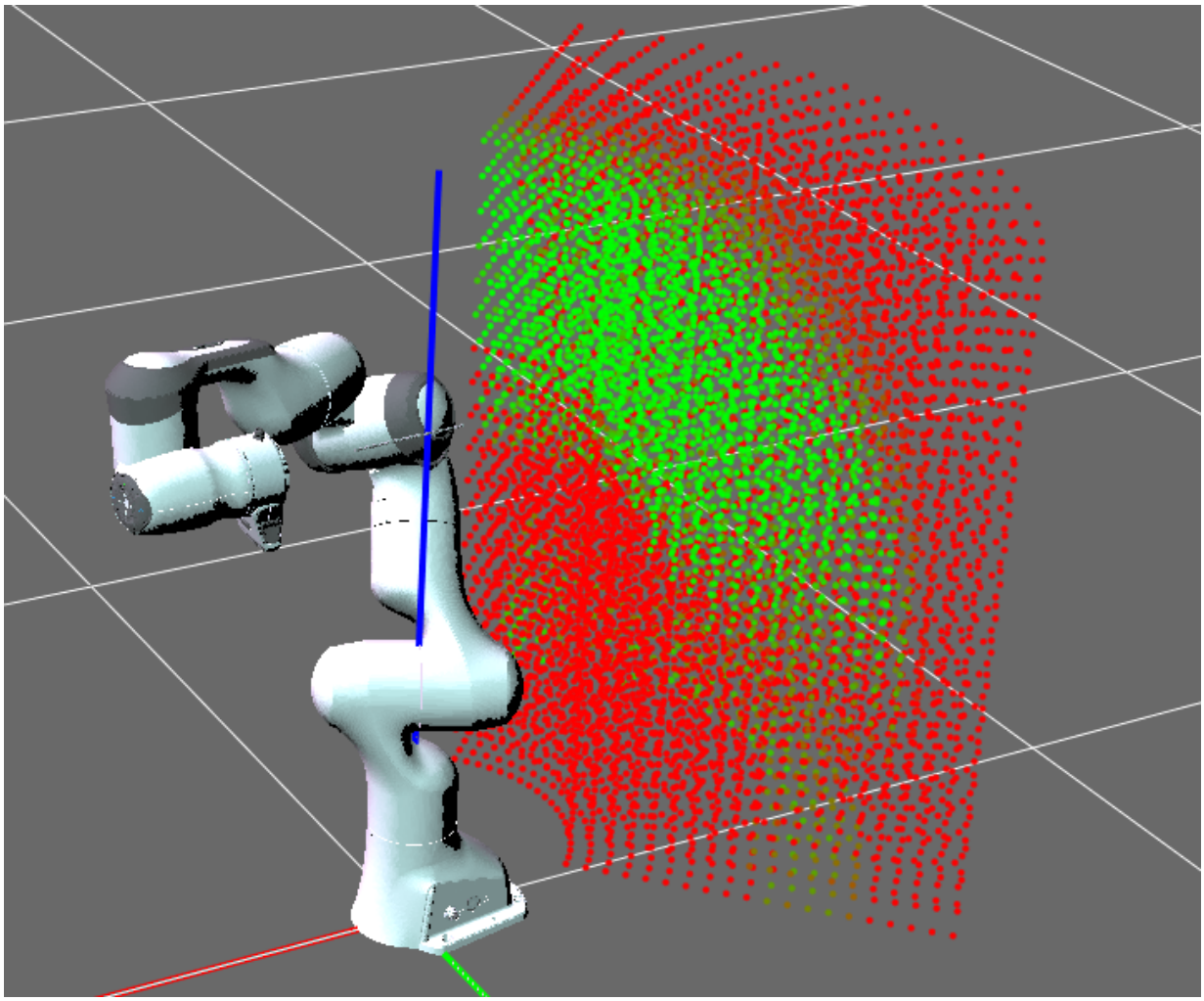


Figure 5.4: Joint lookup table that spans the arm's workspace cone $\mathcal{C}_{\text{start}}$ (5.1) that represents different possible positions a patient's face can be located at. The color represents the achieved quality of the joint configuration stored in the lookup table. The joint positions for green points can reach a high number of points in the destination $\mathcal{C}_{\text{end}}$, while red points reach a low number of points or are infeasible to reach.

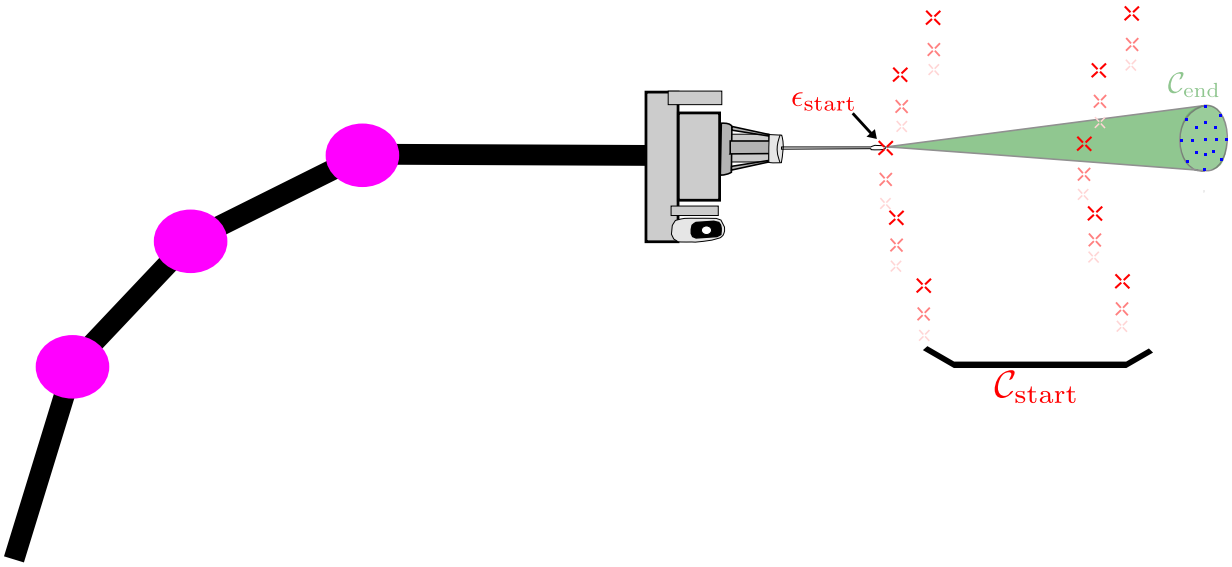


Figure 5.5: Illustration of the cone $\mathcal{C}_{\text{end}}$ (5.3) in green extended from a starting point $\boldsymbol{\epsilon}_{\text{start}}$ (red crosses). During generation of the lookup table, candidate joint configurations are moved to points discretely sampled within this cone (blue dots) to determine how well the arm could move around the face from the starting configuration.

of the face $\mathbf{u}$ can be recovered by taking the linear combination of its low dimensional basis components $\boldsymbol{\beta} \in \mathbb{R}^{50}$ plus the mean values $\overline{\mathbf{u}}$.

$$\mathbf{u} = \overline{\mathbf{u}} + \mathbf{W}\boldsymbol{\beta}. \tag{5.4}$$

The backbone MobileNet [52] of 3DDFA_V2 regresses $\boldsymbol{\beta}$ and a camera projection matrix $\mathbf{P} \in \mathbb{R}^{3\times4}$ based on the shape and orientation of the face under the assumption of a weak projection camera model. Thus, the vertices of the face mesh following a weak projection camera model are found as

$$\mathbf{g} = \mathbf{P}\begin{bmatrix}\overline{\mathbf{u}} + \mathbf{W}\boldsymbol{\beta} \\ 1\end{bmatrix}, \tag{5.5}$$

where $\mathbf{g}$ is the homogeneous vector of coordinates of the vertices projected into the image plane. This formulation is convenient, because particular points of interest on the face, e.g. the nostril, may be found by selecting the subset of rows in $\mathbf{W}$ and the subset of $\overline{\mathbf{u}}$ that correspond to the desired vertices in the face mesh.

While 3DDFA_V2 is useful for extracting semantic information from a single RGB image, its weak camera projection assumptions mean some additional steps are needed to obtain real world cardinal measurements. As we will explain in the following Section 5.3.3, the depth images from the D435i will be used to find the 3D displacement between the points of interest and the camera. The projection matrix, $\mathbf{P}$, can be harnessed to obtain an estimate of the rotation. The left 3×3 submatrix, $\mathbf{R}^*$, of $\mathbf{P}$ contains both scale and rotation of the projected face. However, $\mathbf{P}$ is produced by an unconstrained regression, so the vectors of the matrix are not orthogonal like a proper rotation matrix. To fix both of these issues, we apply singular value decomposition to obtain the closest orthogonal matrix, $\widetilde{\mathbf{R}}$, with $\det(\widetilde{\mathbf{R}}) = 1$ [70]:

$$\begin{aligned}\mathbf{R}^* &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top, \\ \widetilde{\mathbf{R}} &= \mathbf{U}\boldsymbol{\Sigma}^*\mathbf{V}^\top, \\ \boldsymbol{\Sigma}^* &= \mathrm{diag}(1, 1, \det(\mathbf{U}\mathbf{V}^\top)).\end{aligned} \tag{5.6}$$

Some additional compensation is also needed to make the estimated rotation compatible with the pinhole camera model, which better corresponds to the image projection of the camera than the weak projection model. For example, if the face and the camera were oriented in parallel directions, but there is a translation off-center from the principle axis of the camera, the image of the face would be projected at angle and the weak projection image would assume the face is rotated off-center. This is not desirable because we want to decouple the translation and rotational degrees of freedom. Therefore, we can find the

angles between the face and the principal axis,

$$\phi_x = \text{atan2}(p_y, p_z), \phi_y = \text{atan2}(p_x, p_z) \tag{5.7}$$

where p_x, p_y, p_z are the 3D translation components between the camera center and the center of the face. So the final rotation matrix is found by applying the appropriate counter rotation,

$$\mathbf{R} = \exp([\phi^x, \phi^y, 0]_\times)^\top \widetilde{\mathbf{R}}, \tag{5.8}$$

where $[\mathbf{z}]_\times$ produces the skew-symmetric matrix containing elements from $\mathbf{z} \in \mathbb{R}^3$, and exp is the matrix exponential. This effectively maps the axis angle rotation vector to a rotation matrix.

The 3DDFA_V2 model is used during stage 2 and stage 3 to obtain orientation estimates and to find the vertex positions of the nostril via the facial mesh (5.5) during stage 2 when the end-effector is further from the face. However, during stage 3 the swab needs to be closer to the nostrils and a more precise location estimate of the nostrils is needed. Because we lacked a dedicated nostril detector or data train a detector, we rely on an unsupervised pipeline to track the nostrils during stage 3. At the end of stage 2, the nostrils are exposed, so we apply a single object tracking (SOT) network, SiamRPN++ [71], to track them during subsequent frames. SOT models differ from typical detection models because they can be applied to track arbitrary objects that the network has not directly trained to detect. The initial bounding box for the nostrils is passed to initialize the SOT and is found by projecting a selection of vertices around the nostril from (5.5), then refining the area of interest with the Segment Anything Model (SAM) [63]. During subsequent frames, the SOT model iterates to find the updated bounding box. The tracked bounding box location is periodically reset if it wanders too far the vertex positions estimated by 3DDFA_V2.

5.3.3 3D projection

The camera is placed on the end-effector of the arm, and therefore our system will act as an eye-in-hand visual servo system. This greatly simplifies calibration of the system because calibration only needs to be done once, in contrast to an eye-to-hand setup, because the camera and the end-effector are rigidly attached together. The calibration procedure derives the spatial transform between the flange of the robot and the camera ${}^e\mathbf{M}_c$, which allows observations in the camera frame to be eventually be transformed back to the joint space for actuation. We computed ${}^e\mathbf{M}_c$ by moving the end-effector around a calibration checkerboard, then applying Tsai's algorithm [127] implemented in ViSP [84].

The model in subsection 5.3.2 is able to identify the rotation of the face and the position of keypoints within the image plane. The next step is to use this information to determine the relative pose between the nostril and the swab. The 3D position of the nostril is determined using the depth channel of the Intel RealSense D435i camera. Thus, the 3D position

$$\mathbf{p} = \mathbf{K}^{-1} \begin{bmatrix} c \\ r \\ 1 \end{bmatrix} d(c, r) \tag{5.9}$$

is determined by using the camera's intrinsic matrix $\mathbf{K}$ for the nostril keypoint at column c and row r, with $d(r, c)$ as the depth value in meters for the given row and column in the image.

The location and orientation the swab tip with respect to the camera must be known for the visual servo task. In theory, one could use the measured dimensions of the end-effector schematics to determine the swab's pose. However, due to placement tolerances and the fact that the swabs have some variable curvature, it is necessary to estimate the pose of the swab within the camera frame. First, before the visual servo process begins, the swab is segmented in the image using SAM [63] (an example is shown in Fig. 5.6). The shaft of the swab is parameterized with a spline running through the center. The thickness is quantified according to the extent the segmented area covers perpendicular to the spline. This allows us to locate two keypoints from the segmentation: $\mathbf{x}_0$, the tip of the swab (146 mm from the end), and $\mathbf{x}_1$, the point where the shaft of the swab narrows between 2.5 mm diameter to 1.4 mm diameter (76 mm from the end), which is demarcated by a rapid change in thickness. These two points are used to determine the position and direction of the swab with respect to the camera. However, the swab is placed too close to the camera to utilize the depth camera. Instead, the cardinal 3D points are estimated using a combination of projection from the camera and from the coordinates from the end-effector schematics. Ultimately, we use least squares regression to find the 3D positions of the swab, through ray back-projection

$$\min_{\lambda_0, \lambda_1} ||\lambda_0 \mathbf{P}^{\dagger} \mathbf{x}_0 - \tilde{\mathbf{X}}_0^c|| + ||\lambda_1 \mathbf{P}^{\dagger} x_1 - \tilde{\mathbf{X}}_1^c||. \tag{5.10}$$

where λ_0, λ_1 are two ray lengths projected from the camera, P is the camera projection matrix, and $\tilde{\mathbf{X}}_0^c, \tilde{\mathbf{X}}_1^c$ are the 3D points on the swab estimated through the schematics. With the two points $\mathbf{X}_0^c = \lambda_0 \mathbf{P}^{\dagger} \mathbf{x}_0$ and $\mathbf{X}_1^c = \lambda_1 \mathbf{P}^{\dagger} \mathbf{x}_1$ calculated, the 3D position of the swab tip and its direction vector, $\frac{\mathbf{X}_0^c - \mathbf{X}_1^c}{||\mathbf{X}_0^c - \mathbf{X}_1^c||}$, define the current pose of the swab. This process only needs to be done once per trial because the swab and camera undergo the same motion.

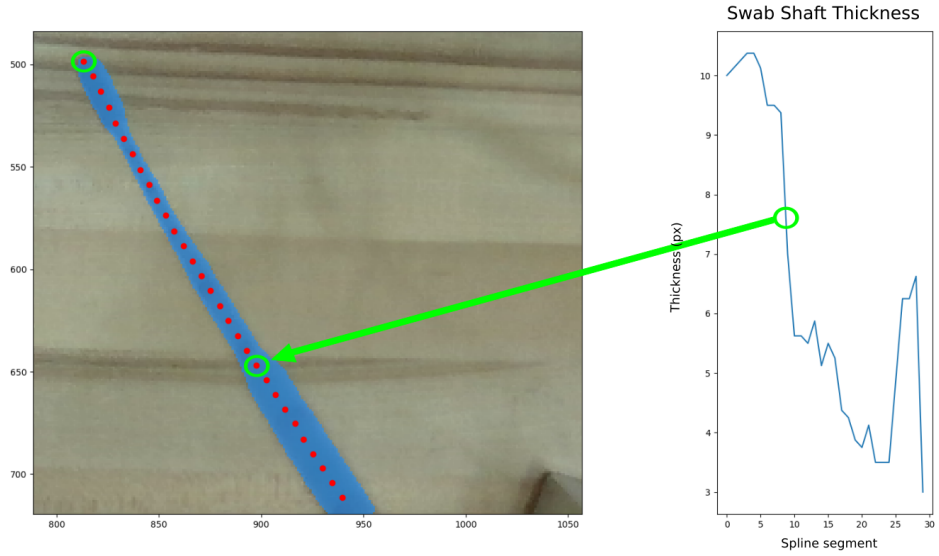


Figure 5.6: Swab is segmented (blue) and then parameterized according to thickness (red dots). The shaft point of interest is found by taking the point belonging to the largest change of thickness within the first half of the spline. The tip point of interest is the final point along the spline.

The final step of the pose estimation system is to determine the pose difference between the swab and the desired pose in front of the nostril. The relative position can be found by simply subtracting the position of the nostril from the swab tip $\mathbf{p} - \mathbf{X}_0^c$. During stage 2, we have the desired rotation of the camera $\mathbf{R}_d$ set at 0.2 radians, so the relative rotation is then $\mathbf{R}_d\mathbf{R}^\top$. For stage 3, the swab does not need to be oriented with all three rotational degrees of freedom; only the swab's direction vector and the desired normal vector of the nostril need to align. To find the "minimal" rotation to align these two vectors, suppose the vector of the swab direction is $\mathbf{v}$ and the target normal of the nostril is $\mathbf{n}$, pitched at 0.2 radians. Then the relative rotation $\mathbf{R}$ can be found as [13]:

$$\begin{aligned} \mathbf{w} &= \mathbf{v} \times \mathbf{n} \\ c &= \mathbf{v} \cdot \mathbf{n} \\ \mathbf{R} &= \mathbf{I}_3 + [\mathbf{w}]_\times + [\mathbf{w}]_\times [\mathbf{w}]_\times \left(\frac{1}{1+c} \right) \end{aligned} \tag{5.11}$$

with $\times$ as the cross product operator and $\mathbf{I}_n$ as the $n \times n$ identity matrix.

5.3.4 Unscented Kalman filter on manifolds (UKF-M)

Our initial evaluations of the measurement pipeline found significant noise in the pose measurement system that would impede the performance of the visual servo, if left uncompensated. Another concern is that if measurements are lost from depth acquisition or failure from the face model, then the servo would intermittently halt. Therefore, having a state estimator in the system is useful to compensate for these sources of noise and enable the control loop to continuously run.

Since the measured state is in the 3D Euclidean space group (SE(3)) a non-standard Kalman filter will be required because of non-linear dynamics. In addition, classical Kalman filters [60] assume the state as a vector, which is problematic because using vector representations of rotation can instigate wrap around and gimbal lock issues. Therefore, we used an unscented Kalman filter on manifolds [20] to seamlessly estimate the state within this domain. First, the process equations of motion are defined as

$$\begin{aligned} \mathbf{p}(t + \Delta t) &= \mathbf{p}(t) + \Delta t(-\mathbf{v}_p - \mathbf{v}_\theta \times \mathbf{p}(t)) \\ \mathbf{R}(t + \Delta t) &= \mathbf{R}(t) \exp(\Delta t[-\mathbf{v}_\theta]_\times) \end{aligned} \tag{5.12}$$

where $\mathbf{p} \in \mathbb{R}^3$ is the position, $\mathbf{R} \in \text{SO}(3)$ is the rotation matrix, $\mathbf{v}_p \in \mathbb{R}^3$ and $\mathbf{v}_\theta \in \mathbb{R}^3$ are the linear and angular velocity of the camera frame that is commanded by the control

loop, and Δt is the time step.

The UKF-M contains two innovations over the standard Kalman filter: a manifold projection to handle the non-vector state, and the unscented transform to handle the non-linear dynamics. Manifold projection requires us to define two functions ϕ and ϕ^{-1} that correspond to our specific state. The retraction function $\phi(X, \xi) \in \mathrm{SE}(3)$ makes an incremental change to the state $X \in \mathrm{SE}(3)$ by a perturbation vector $\xi \in \mathbb{R}^6$. For the translation component, this amounts to vector addition (e.g., $\phi(p, \xi_p) = \mathbf{p} + \xi_p$) while the rotational component treats $\xi_R \in \mathbb{R}^3$ as an axis-angle rotation vector so $\phi(\mathbf{R}, \xi_R) = \mathbf{R}\exp([\xi_R]_\times)$. In addition, the inverse-retraction function $\phi^{-1}(X, \hat{X}) \in \mathbb{R}^6$ maps the change between the two SE(3) states to a 6-D vector. The translation component corresponds to vector subtraction (e.g., $\phi^{-1}(\mathbf{p}, \hat{\mathbf{p}}) = \mathbf{p} - \hat{\mathbf{p}}$) while the rotation component converts the relative rotation between the two matrices into an axis-angle vector ($\phi(\mathbf{R}, \hat{\mathbf{R}}) = \log(\mathbf{R}\hat{\mathbf{R}}^\top)$).

Like a classical Kalman filter, the UKF-M needs to have an estimate of the statistics of the state (i.e., the mean and covariance) in order to balance the information between new measurements and process dynamics when updating the state. The difference is that a Kalman filter assumes a linear process, where the statistics would accurately propagate with linear transformations [60]. These assumptions do not transfer well to a non-linear process [58], which is what the unscented transform attempts to rectify. The unscented transform operates by obtaining deterministic samples perturbed around the state and feeding them through the non-linear propagation function to estimate a more accurate the set of statistical parameters [130] [58]. The UKF-M also adopts this unscented transform to estimate the state statistics, just by making use of the ϕ and ϕ^{-1} functions project the perturbed samples into the manifold and vice versa through the non-linear functions.

The remaining parameters that need to be defined to apply the UKF-M are the measurement covariance matrix, process covariance matrix, and the scale parameter for the unscented transform. Based on the noise characteristics of the pose estimation system (Sections 5.3.2 and 5.3.3), we hand tuned the measurement covariance components corresponding to translation to $0.005 * \mathbf{I}_3$ and to rotation to $0.05 * \mathbf{I}_3$. The process covariance matrix was set at $0.01 * \mathbf{I}_6$. The scale parameters were set at 0.01 for state propagation, 0.1 for noise propagation, and 0.01 for state updates.

5.3.5 PBVS control law

Motion of the arm is implemented using a pose based visual servo (PBVS) control law. This component takes the observed state pose state from the UKF-M and computes the joint velocities that will move the end-effector to the desired position $\mathbf{t}_d$ and rotation $\mathbf{R}_d$.

The control law [25] in the camera frame takes the form of

$$\mathbf{v}_c = \lambda \mathbf{L}^{\dagger} \mathbf{s} \tag{5.13}$$

where $\mathbf{s}$ is the pose error between the desired pose and the current pose from UKF-M. The interaction matrix, $\mathbf{L}$, is used to map the velocity of the camera origin to the velocity of the target in the camera's frame. In this vein, the pseudo inverse is used to map pose error into velocities of the camera. The interaction matrix is a block matrix $\mathbf{L} = \begin{bmatrix} \mathbf{L}_t & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{\theta u} \end{bmatrix}$, where the non-zero blocks correspond to the translational and rotational components of the pose vector. Given that $\mathbf{R}$ is the rotational component of $\mathbf{s}$ and $\theta\mathbf{u} = \log(\mathbf{R})$, $||\mathbf{u}|| = 1$, these components are defined as $\mathbf{L}_t = \mathbf{R}$ and $\mathbf{L}_{\theta u} = \mathbf{I}_3 - \frac{\theta}{2}[\mathbf{u}]_\times + (1 - \frac{sinc(\theta)}{sinc^2(\frac{\theta}{2})})[\mathbf{u}]_\times[\mathbf{u}]_\times$ [25]. Finally, the end-effector robot Jacobian in the camera's frame ${}^c\mathbf{J}_{ee}$ is used to map the desired camera velocity into joint velocities

$$\mathbf{v}_q = {}^c\mathbf{J}_{ee}^{\dagger}\mathbf{v}_c = \lambda {}^c\mathbf{J}_{ee}^{\dagger}\mathbf{L}^{\dagger}\mathbf{s}. \tag{5.14}$$

The gain λ was kept at a value of $0.5\ \mathrm{s}^{-1}$. This control law was implemented using ViSP [84] through the C++ Franka Control Interface.

5.4 Experiments to validate pre-contact phase pipeline

To evaluate the efficacy the proposed pre-contact phase pipeline and its components, we conducted a series of human trials, which were approved by the University of Waterloo institutional research ethics board (REB # 43485). The experiments took place from July 2023 to February 2024 at the University of Waterloo RoboHub with a total of 25 participants, who were primarily recruited from mailing lists in the Faculty of Engineering. The participants were instructed to freely stand within the arm's cone workspace so the system could move the swab towards one of their nostrils. This process was repeated for the other nostril. The outcome of the trial was recorded by two perpendicular GoPro Hero 8 cameras taken from the side and bottom directions. Fig. 5.7 shows examples of some of the trials using camera arrangement. One question of particular interest is whether the performance of the face pose estimation system is influenced by population demographics because 3DDFA_V2 was trained on empirical data that could have under-represented some demographic groups. Therefore, participants also filled out a short survey where they could indicate gender and racial information, which is summarized in Table 5.1.

Table 5.1: Demographics of study participants that were recorded to evaluate for potential model biases. The mixed race participants are assigned fractionally to each selected race.

Sex					
Male			**Female**		
17			8		
Race					
White	**South Asian**	**East Asian**	**Latin American**	**Iranian**	**Arab**
10.5	6.5	3.5	2	2	0.5

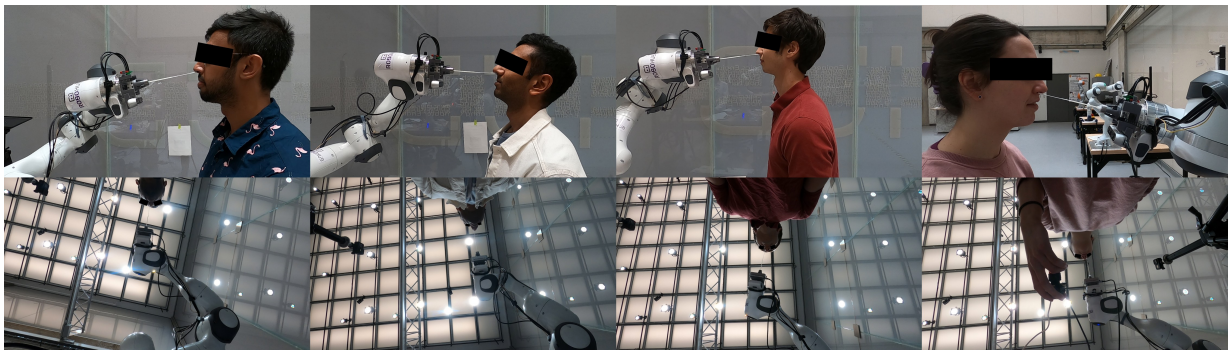


Figure 5.7: Images showing the trials of several different participants. The images were taken once the swab tip reached the nostril.

5.4.1 Positioning and alignment outcomes

In terms of general results, the rate of the system reaching the nostril was generally successful among the participants, although the system did exhibit some discrepancy among the faces encountered. The system was able to reach the nostril during 84% of the trials. The remaining 16% of cases placed the tip of the swab just below the nostril (above the lip). This state generally occurred because of the failure of 3DDFA_V2 to locate a good initial bounding box of the nostrils correctly during stage 3 (see Fig. 5.13). As a result, the SOT component would be tracking the wrong location throughout stage 3. There was one case where a man with a darker beard caused the SOT tracker to diverge and move the bounding box towards the beard periodically. In Section 5.5.1, we discuss some ideas to rectify these issues.

While placing the tip of the swab at the nostril is obviously important, the orientation of the swab is also crucial to regulate to prior to the contact phase. To this end, another objective of our experiments is examining the final orientation of the swab among the differ trials. We estimated the final pitch and yaw orientation within the trials by annotating the footage of the perpendicular GoPro cameras on the terminal frame. The pitch orientation was found by comparing the angle difference between the line segment of the swab (from tip to shaft), and the line segment connecting the point of the earlobe to the Ala (lower edge of the nose) within the side view, as shown in Fig. 5.9. The yaw angle took the angular offset between the swab line segment and the line segment from the base of the columella to the tip of the nose from the bottom view. The distribution of these two quantities is shown in Fig. 5.10. Yaw had a mean angle of -4.1° and standard deviation of 7.8°, while pitch had a mean of 16.3°, standard deviation of 4.5°.

The characteristics of swab to camera pose calibration from Section 5.3.3 is worth examining as well. From the camera's point of view, it was clear that the position of the swab did indeed vary slightly between trials because each swab had slightly different curvature due to the manufacturing process and there were some placement tolerances with how it was inserted into the swab holder. While physically these factors did not appear significant, it is prudent to verify that the process is making reasonable estimates. A 3D plot of the estimated swab position relative to the camera orientation is shown in Fig. 5.8. Compared to the schematics, there was an offset of 5.2 mm from the center of the two keypoints and an angle of 3.5°. But, the standard deviations among the trials was fairly small, at 1.4 mm for position and 1.4° for angle. As a result, it is unlikely the swab to camera calibration is to blame for the variance observed in the previous paragraph, which must be attributed to other sources.

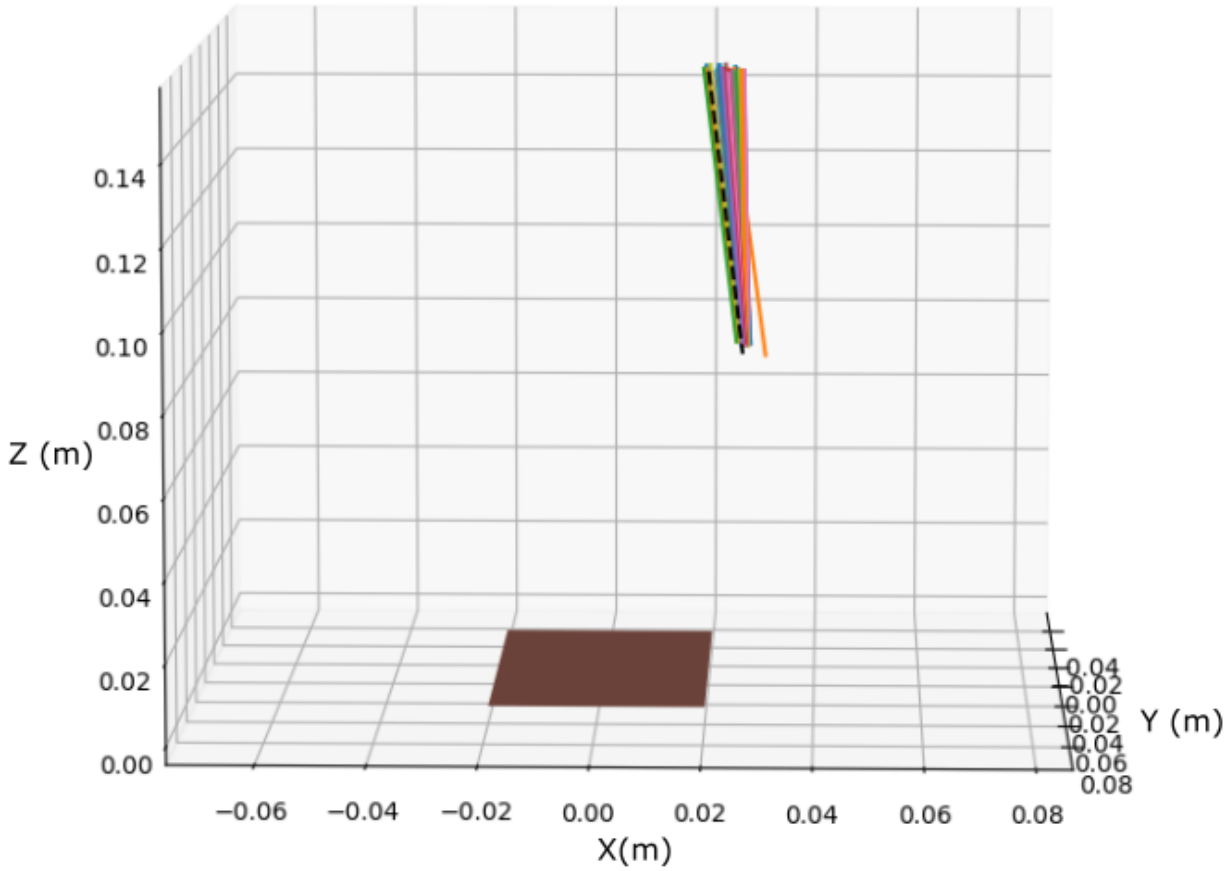
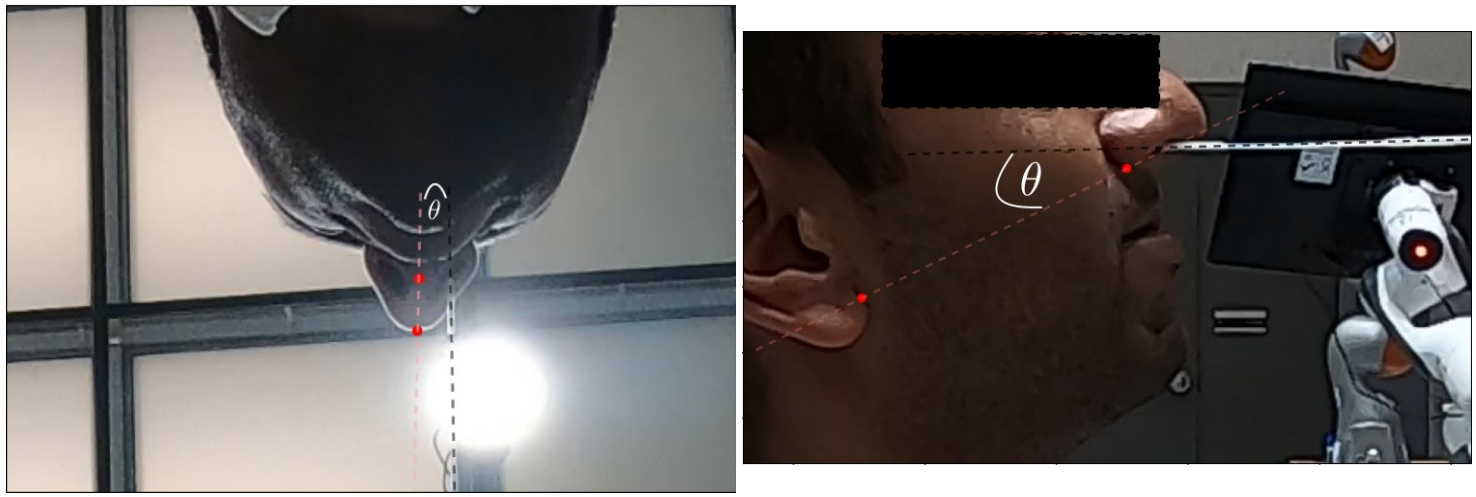


Figure 5.8: Projected segments of NP swab in the camera view relative to the camera's orientation plane (brown square) observed over all the trials. The black dashed line indicates the position of the swab according to schematics.



(a) (b)

Figure 5.9: Image showing the labeling scheme for determining the relative angles between the swab and the face. (a) The bottom view landmarks align with the columella of the nose;(b) the side view landmarks are the corner of the ear and the ala of the nose (see [85]). The reported angles come from the angular difference of the line segments between the landmarks (pink dashed) and the swab (black dashed). The angle in the bottom view corresponds to positive yaw, while the angle in the side view corresponds to positive pitch.

5.4.2 Noise attenuation

As we alluded to in Section 5.3.4, there is substantial noise from the pose measurement, which is a major reason the UKF-M was included into the system. Fig. 5.11 shows an example of the relative target pose estimated from the measurements and the UKF-M over the latter two stages. One can notice that the raw measurements from the CNN pipeline exhibit substantial amounts of stationary noise in all axes. The estimates from the UKF-M appear to effectively attenuate the noise and does not lag, which indicates that the process updates (5.12) are able to contribute to keeping an accurate state.

We examine the level of measurement noise during the trials by subtracting the measurements from the filtered states for the last two stages during the experiment trials. Fig. 5.10 shows the range of standard deviations of the noise estimate that were recorded during the trials, with two trials excluded due to pauses during the trajectory. For the position noise, stage 2 had an average standard deviation of 8.2 mm, while stage 3 was lower at 2.3 mm. The angular noise had standard deviations of 0.030 rad (1.7°) for stage 2 and 0.017 rad (1.0°) for stage 3. Overall, the second stage had more noise, likely because of the greater distance from the face and higher velocities compared to stage 3.

5.4.3 Impact of demographics

Given the data driven components nature of 3DDFA_V2, it is important if the trial outcomes (i.e., whether the swab reached the nostril, and the final angle) are influenced by participant demographics. It is clear that the variance in final swab to nose angles exceed any amount that could be explained by the swab to camera pose calibration. Therefore, the main source of the variance must come from estimates of 3DDFA_V2. Therefore, we examine the potential impact of demographic bias by examining the following three questions:

1. Are the observed final angles correlated to the individual? If not, then there will be no way to establish links between demographic groups.
2. Does gender significantly impact the final angle or the ability to reach the nostril?
3. Does race significantly impact the final angle or the ability to reach the nostril?

We apply statistical testing to help answer these questions. Because we are conducting multiple tests, we will correct for type II errors by applying Bonferroni correction [37] and reducing the critical p-value threshold to $0.05/3 = 0.017$.

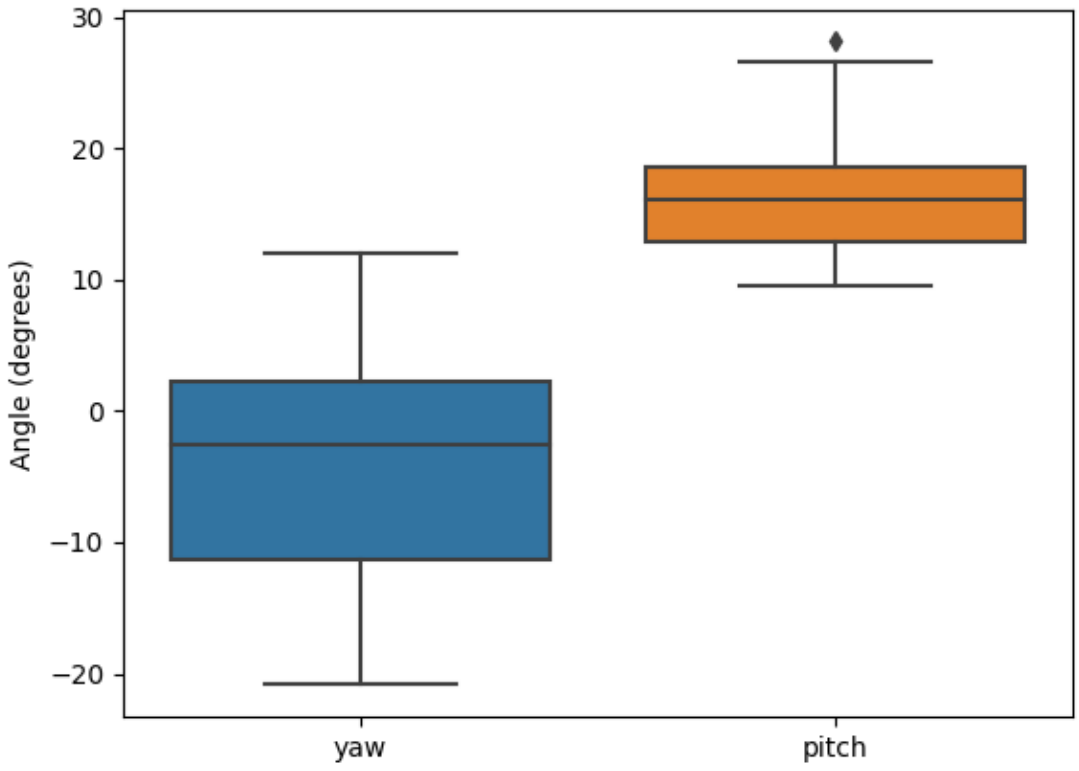


Figure 5.10: Distribution of final observed angles from human trials.

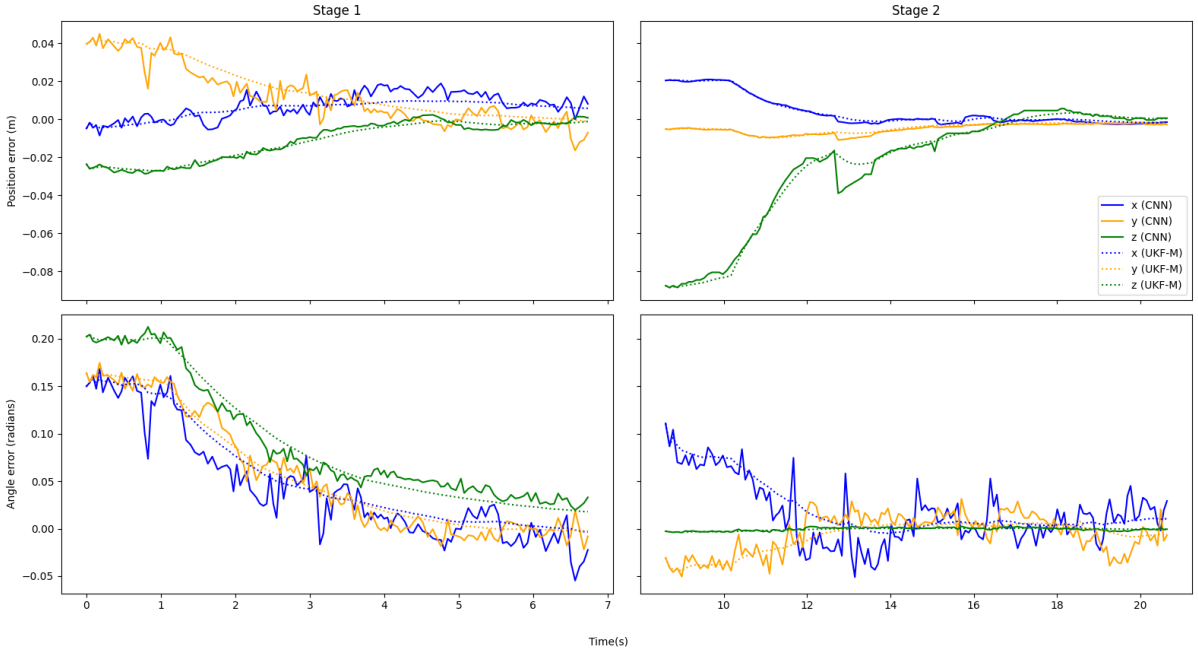


Figure 5.11: Graph showing the convergence of the relative pose along the position (top) and orientation (bottom) axes for an example trial. The graph is divided into stage 1 (left) and stage 2 (right). Notice how for stage 2, the Z-axis angular error is low due to only needing to meet two rotational degrees of freedom (5.11).

1) For yaw, the standard deviation between different people (intra-person) was 7.1°, while the average standard deviation between the two trials for the same person (inter-person) was only 2.4°. For pitch, the intra-person standard deviation was 4.0° and the inter-person standard deviation was 1.5°. We computed F-statistics by taking the ratio of intra-person variance to inter-person variances, which was significant for both yaw ($F = 5.1, p = 9.1 \times 10^{-5}$) and pitch ($F = 4.0, p = 4.9 \times 10^{-4}$). This presents strong evidence that the angles output by the 3DDFA_V2 result in repeatable final orientations with the same face, compared to the final orientations of different faces.

2) For gender, we observe that males have a yaw of -5.3° ± 6.4° and pitch of 16.5° ± 4.3°, while females had a yaw of -1.4° ± 7.8° and pitch of 15.9° ± 3.1°. A t-test reveals that there is no significant angular difference for yaw ($T = -1.3$, $p = 0.22$) or pitch ($T = 0.35$, $p = 0.73$) between the two genders. Overall, we found that gender did not play a significant role in the final evaluated angle. We found the rate of reaching the nostril was higher for females (100%) compared to males (76%). We applied a Chi-square test to statistically evaluate if the outcome was biased by gender. Ultimately, this resulted in a test statistic of $\chi^2 = 3.0$ and $p = 0.08$. Even though the swab tip was able to reach the nostril for all female participants, it still does not appear to be significantly different from males in our sample.

3) Evaluating the impact of race from our experiments has a few complications that require some discussion. First, facial structure is dependent on many other factors than simple racial groupings, which is a topic that is obviously outside the scope of this thesis. From our analysis we merely hope to detect if there are any conspicuous biases towards racial groups that could have been caused by under-representation in the training sets of 3DDFA_V2. The second issue comes from the categorization of races. In the questionnaire given to participants, we based the categorization from the Government of Canada's census [121]. However, there are many different races listed, and there does not appear to be a standardized way of organizing them into related ethnic groupings. In addition, they did not have a clear group for certain participants, such as those who were of Iranian descent. Another difficulty concerns how to categorize mixed race participants, which is a scenario that does not seem to have a universal solution to in the demography literature [75]. Ultimately, we subdivided the groups into Group 1={Arab, Iranian}, Group 2 = {East Asian}, Group 3 = {South Asian, Southeast Asian}, Group 4 = {Latin American, White}, and chose to use "the inclusive membership" [75] strategy, where mixed race people are included to multiple groups. We applied a one-way ANOVA model, with an F-test indicating that there is no significant yaw ($F = 0.32$, $p = 0.82$) or pitch ($F = 1.1$, $p = 0.37$) difference between the races. Of course, there are some caveats to this analysis that we will further elaborate in Section 5.5. For reaching the nostril, we compare if race is

a significant factor by applying a Chi-square test. With the inclusive membership strategy, a test statistic of $\chi^2 = 2.5$ with $p = 0.47$ indicates that the outcome is not significantly different according to race.

5.4.4 Workspace analysis for the contact phase

While the scope of these experiments was to evaluate the visual servo task for NP swab sampling, the arm would need to be capable of executing the contact phase of the test. To verify this, we will examine the workspace of the arm at the terminal joint configurations from the trials and test if the swab would still be capable of reaching the nasopharynx without running into joint limits. The pseudo-inverse Jacobian IK method (5.2) was applied to integrate joint positions until the tip reaches the target pose, which is 0.2 radians up and at least 130 mm forward. We set the threshold for success at a 130 mm because it is longer than a large proportion of the population's nasal cavity [80], but we still evaluated the workspace up to 500 mm. Fig. 5.12 shows the success rate of the reaching targets of increasing displacement without encountering joint limits or self-collisions. All but one trial, where there was excess movement of the participant, was able to reach the target displacement of 130 mm. In fact, all the trials had sufficient workspace to extend to 175 mm, before the success rate started dropping at 200 mm and beyond. Therefore, we conclude that our joint lookup table is effective at ensuring the arm has sufficient workspace to perform the NP swab test.

5.5 Discussion of human trial experiments

Overall, we have shown that the proposed system using a single RGB-D camera is effective at performing the pre-contact phase of the NP swab test. The results from the experiments have highlighted the utility of the UKF-M filtering component within the proposed system. We saw that there is significant stationary noise coming from the measurement pipeline, particularly in terms of the rotation estimates from 3DDFA_V2. The UKF-M was able to attenuate this noise effectively and is a necessary component of the pipeline. Without it, the control response would have to contend with these noisy measurements and be less stable or require the control law to use lower gain to compensate. Overall the parameters we chose for the UKF-M appear to be appropriate based on their ability to follow head motion and suppress noise. Of course, there is still some room to modify the chosen parameters. For example, if we were to assume we were swabbing less cooperative cohort,

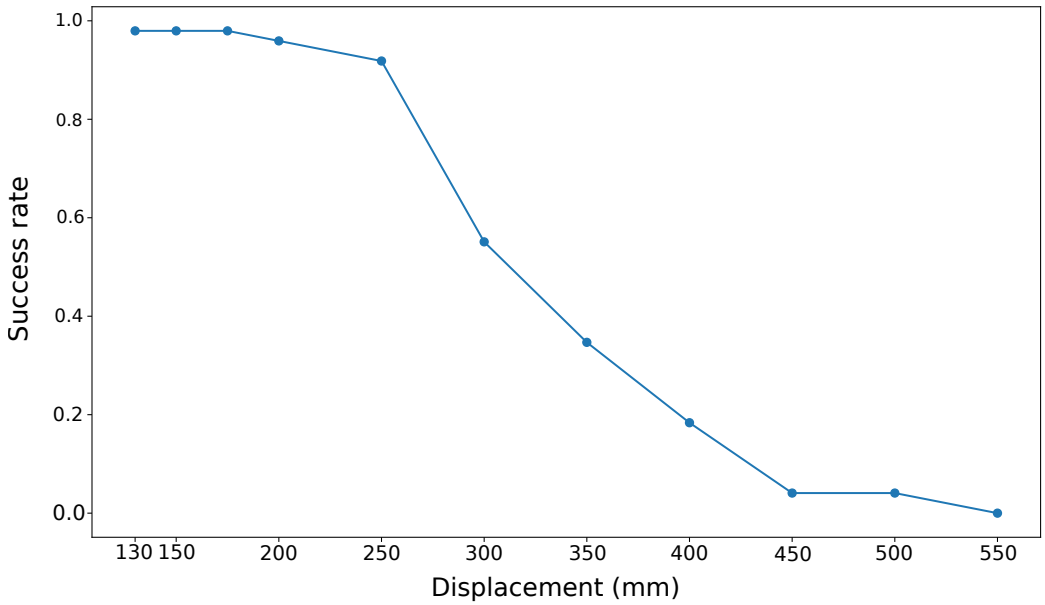


Figure 5.12: Figure showing how far the end-effector can reach beyond the terminal configuration from the trials joint limits and/or collisions are encountered. The final configuration is clearly able to achieve the 130 mm displacement that would be needed to reach the nasopharynx for the majority of the population. The arm is able to extend several centimeters beyond this, indicating that there is sufficient workspace available to proceed with the contact stage of the test.

then we would need to increase the magnitude of the measurement covariance matrix to compensate.

The implemented PBVS control law was effective at placing the swab at the desired pose during trials. Thanks to the feedback of control velocities to the UKF-M, the pose state stayed current and the task was able to converge to the desired target without overshooting. One comment from some participants was the duration of the servo component being too long. The current control scheme uses a constant gain term that results in an exponential decay in error over time. This could be improved in the future by implementing an adaptive gain, such as a method by Kermorgant and Chaumette [61], that raises the gain as the error shrinks, allowing for a quicker, more linear convergence rate.

In terms of the final angle of the swab, we found that the pitch angle was fairly consistent, which would be important for ensuring the swab does not travel down the wrong tract during the contact phase. The yaw angle were more variable, however, and could be improved. In addition to biases from 3DDFA_V2, there could be some additional sources of perceived variance. The positioning of the bottom camera may have induced some dissimilar perspective effects, especially for taller people. This is a side effect of having unrestrained participants to evaluate the workspace and motion compensation abilities of the system; in the future this effect could be examined more critically be repeating the experiment with participants in a head restraint. Another idea could be using a marker-based infrared motion capture system such as VICON (Yarnton, United Kingdom), but it is not clear if placing markers on a participants face would impact the vision pipeline.

From our statistical analysis, we did not find a significant difference in outcomes among the demographics of our sample. However, our sample has limitations in terms demographic coverage. The sample was biased towards the university engineering cohort. This meant that the gender ratio was skewed towards males (17:8), most of the participants were young adults, and some racial groups were under-represented or not represented at all. Also, the distribution of our limited sample size across multiple demographic groups reduces statistical power, making it challenging to draw robust conclusions. Therefore, further evaluation on a wider demographic sample would be needed before application to clinical settings to ensure the pipeline can generalize appropriately. But, we predict that the system should generalize well, because 3DDFA_V2 was trained on a dataset derived from the 300 Faces in the Wild dataset [109] that consists of images from the internet and wide-reaching social medias, so the model should generalize to a wider demographic than was tested in our experiments. In any case, future studies will need to expand demographic coverage and consider different groupings in statistical analyses.

The arm was able to maintain sufficient workspace during the experiment trials and

afterwards if it were to proceed with the contact phase of the NP swab test. The lookup table provided joint configurations that ensured the arm had sufficient workspace to execute the task. In terms of further efforts, future implementations could have a detector to indicate if the participant steps too far away from their initial position after the pre-contact phase begins.

5.5.1 Avenues for improvement

We saw that in a number of cases the swab tip failed to reach the nostril and settled at another point near the nose. As mentioned earlier, this was often because the initial bounding box was not aligned with the nostril because 3DDFA_V2 predicted the corresponding nostril vertices incorrectly, as shown in 5.13. Consequently, the SOT network failed because it did not have a correct initial bounding box. The pipeline specific to stage 3 is something that we would replace in future applications because relying on the unsupervised segmentation and tracking modules appears to be less suitable when the appearance of the nose and nostril have such distinct visual cues. It would be more sensible to train a dedicated nostril detector for use in stage 3, similar to Hwang *et al.* [54]. We initially explore this approach by training a CNN to detect the position of the nostril from the end-effector camera's perspective. The base model is a Faster R-CNN [104] with a Resnet-50 [48] backbone that was pretrained on the Coco dataset [77]. We apply transfer learning by creating a small dataset using the images from the end-effector camera at various times of the insertion, with 11 participants for training and 14 for testing. Notably, the testing set contained all the participants whose nostrils were not reached by the visual servo. Fig. 5.14 shows an example of the detector on two testing images. We trained the model so that it can identify the nostril even if the swab is partially occluding it. Overall, this approach appears to be reliable and a good replacement for the unsupervised components.

Like the Eye-to-hand visual servo of Hwang *et al.* [54], the Eye-in-hand visual servo was able to successfully converge to the nostril. Compared to an Eye-to-hand setup, the Eye-in-hand setup is simple to calibrate and is not susceptible to needing recalibration from the camera being perturbed because it is rigidly fixed to the end-effector. The setup is also robust to occlusions because the links of the robot will never be in a position to obscure the target face. In addition, the swab only needs to be segmented once, not for each frame like the eye-to-hand setup because the transform between it and the camera remain constant throughout the trial. However, one drawback of the eye-in-hand setup is that depth images on most RGB-D cameras, including the RealSense D435i, have a minimum acquisition range where surfaces closer than this distance cannot be measured. This was not an issue with the pre-contact phase because there was sufficient space between the swab



Figure 5.13: Images comparing the initial bounding box of the instance where the swab tip missed (left), and one that arrived successfully (right). The green dots are the nostril points projected by 3DDFA_V2, the light blue mask is the segmentation from SAM [63], and the right outline is the corresponding bounding box. Notice how in the left image none of the nostril points from 3DDFA_V2 are near the nose so it does receive a valid bounding box to begin with.



Figure 5.14: An improved detector model using a Faster_RCNN architecture. The detector produces more accurate bounding boxes for the nostrils than the previous pipeline (left) and can be tuned to detect the nostril even if it is partially occluded by the swab (right).

tip and the camera. But, if the swab were to enter the nasal cavity, then depth acquisition would eventually be lost as the camera gets too close to the face. Thus, during the contact phase, adding a camera for a hybrid setup [44] could provide some benefit for keeping the swab aligned in conjunction with force inputs. We could also consider complementing the system with additional modalities, such as LiDAR.

5.6 Conclusion

In this chapter we proposed a vision guided pipeline to accomplish the pre-contact phase of the NP swab test. The pipeline uses a joint lookup table to maneuver towards a participant's face with a joint configuration that has sufficient end-effector workspace. Subsequent stages use deep learning models combined with an unscented Kalman filter on manifolds to estimate the pose of the face, which are then used to actuate the robot with a pose-based visual servo control loop. We validated our method with a series of human trials, which were successful and revealing ways to improve nostril tracking to be more reliable and make to make convergence of the control task quicker. In the future, we can consider multi-camera configurations to support vision during the contact phase.

Chapter 6

Compliant control system

Foreword

The contents within this chapter are part of a paper that is currently under review in the journal Scientific Reports: P. Q. Lee, J. S. Zelek, K. Mombaur. "Collaborative robot arm inserting Nasopharyngeal swabs with Admittance Control", *Scientific Reports*, Submitted November 2023.

Within the scope of this thesis, this chapter documents our attempt at using a torque compliant control law to correct for misalignments during the contact phase of the NP swab test. Chronologically, the main experiments in these chapters took place during February 2023, and was the first example of applying our work to robotic hardwar.e

6.1 Introduction

From the outset of the COVID-19 pandemic, the appeal to applying robots in the place of human healthcare workers has increased substantially. A major occupational hazard for healthcare workers is contracting illnesses from the patients they are treating; especially via highly contagious air-born spread illnesses like COVID-19. Robots deployed in healthcare have the advantage of being immune to illnesses, and would thereby be useful to protecting the healthcare workers and preventing downtime [59]. In cases where training consistency is a concern [79], robotics can provide a way to standardize care. However, close-contact healthcare tasks are challenging workspaces for robots because of safety considerations and the need to meet medical objectives with the constraints of the robotic hardware.

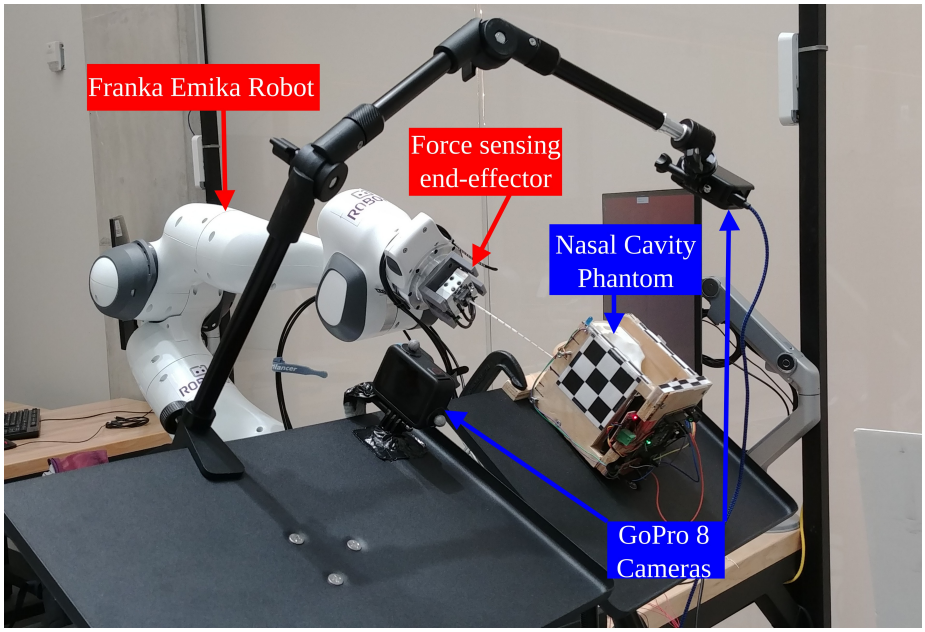


Figure 6.1: Nasal cavity apparatus arranged next to the Franka Emika Robot arm with the proposed force sensing swab end-effector attached. Two GoPro cameras are used to observe the outcome of the insertions and are not part of the control loop.

Consequently, in this research we target the task of nasopharyngeal (NP) swab sample collection using robots. The task involves inserting a thin, flexible swab through the nasal cavity until it reaches the nasopharynx at the anterior (just above the back of the throat). The task is challenging because the swab has to navigate around anatomical obstacles in the nasal cavity; namely the nasal septum (the wall that divides the left and right passages of the nose) and the inferior turbinates (the curved bony structures). The ideal path for the swab to travel is between the nasal palate and the inferior turbinates [98] (example also later shown in Fig. 6.4); deviation from this path is unlikely to reach nasopharynx and can cause discomfort or injury if the swab impacts sensitive anatomy such as cribriform plate and the various nerve clusters in the nasal cavity [119]. From a control perspective, it is an interesting task because once the swab enters the nose, it becomes visually unobservable, so any adjustments must be based on measured forces applied to the swab.

Applying robots to different types of swab sampling tasks has received increasing attention from the robotics and biomedical engineering community. Park *et al.* [95] conducted a study where the forces of a practitioner were recorded using a handheld instrument on a phantom. Hwang *et al.* [54] implemented a visual-servo and control system, which uses

deep-learning models to detect the nose and guide the robot to more accurate positioning at the nostril prior to insertion. Li *et al.* [73] also implemented visual-servo methods for NP swabs using a hierarchical decision network strategy. Li *et al.* [72] designed a 3-DOF robot with an endoscopic camera to perform oropharyngeal swab tests via teleoperation. Chen *et al.* [29] designed another teleoperated soft-rigid hybrid robot that features a compliant fiber Bragg grating integrated manipulator to adapt to disturbances during nasopharyngeal or oropharyngeal swab tests. Maeng *et al.* [83] created a custom NP sampling robot featuring a remote center of motion mechanism in order to compensate for sudden forces by the patient during the procedure. Zhang *et al.* [142] created a platform centered around a humanoid dual-arm robot for NP swab sampling. Their system utilized a combination of RGB-D cameras, LIDAR scanners, and a force-torque sensor to perform the task on humans, with results indicating it was very effective at gathering PCR samples. There are also some commercial attempts at robotic swabbing. Lifeline robotics (Odense, Denmark) take on the related, but distinct task of oropharyngeal swabbing using a UR5 robot[76]. Franka Emika (Munich, Germany) built a swabbing station surrounding their robot, but was applied towards shallow nose and throat swabs[3]. Brain Navi (Zhubeai, Taiwan) applied a UR5 robot to do NP swabs based on a 3D visual scan of the face, but appears to do the test without any force sensing feedback [18].

Many existing developments in robotic swabbing tests have integrated or developed custom robotic hardware tailored specifically for the task. In contrast, we propose an alternative scenario where a single collaborative robotic arm can be utilized for a variety of healthcare tasks requiring close patient contact, such as temperature testing, needlework, and sample collection. Our research focuses on employing a rigid manipulator robot to autonomously collect NP swab samples, with the goal of generalizing its hardware for other tasks using the same instruments as a human worker.

While the complete NP test consists of three stages - (a) inserting the swab through the nasal cavity until it reaches the nasopharynx, (b) rotating the swab on the nasopharynx, and (c) removing the swab from the nasal cavity [98] - this study specifically targets the first stage. This scenario presents a unique challenge, as the contact forces involved must be regulated by the control law rather than relying on mechanical compliance from specially designed manipulators.

Notably, recent work [47] has also utilized a collaborative robotic arm with a joint torque controller to perform NP swab tests. However, their data indicates that they did not successfully reach the nasopharynx, likely due to the low sensitivity of their onboard torque sensors. This limitation underscores the importance of our approach, which includes a novel low-cost, high-accuracy force-sensing end-effector designed to enhance performance.

To address the challenges of the NP swab test, we conduct a case study to assess the suitability of our robotic arm in this context using an admittance control system. We first describe our design of the force-sensing end-effector. Next, we present our development of a force-feedback torque control law to execute the insertion stage of the NP swab test. Finally, we conduct experiments using the robot on a 3D-printed phantom of a human nasal cavity and compare its performance to a baseline position controller. This comparison demonstrates the effectiveness of our setup for performing NP swab tests and highlights that even minor misalignments require high-precision force feedback to insert the swab successfully.

6.2 Materials and methods

The hardware used in this study consist of the collaborative robotic arm platform, the custom designed force-sensing end-effector where the NP swab is mounted, and the nasal cavity apparatus.

6.2.1 Collaborative robot arm

The robotic platform used in this work is the Franka Emika Robot "Panda" arm (Franka Emika, Munich Germany). According to Featherstone and Orin [41], we expect the robot to follow the dynamics

$$\boldsymbol{\tau} = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}), \tag{6.1}$$

where $\mathbf{M}$ is the mass inertia matrix, $\mathbf{C}$ is the Coriolis effects matrix, $\mathbf{g}$ is the gravitational vector, $\boldsymbol{\tau}$ is the torque vector, $\ddot{\mathbf{q}}, \dot{\mathbf{q}}, \mathbf{q}$ are the joint acceleration, velocity, and position vectors. Featherstone [40] provides a discussion on how these effects are computed for a rigid body. The arm has 7-DOF and is designed as a collaborative robot that is meant to be used to perform tasks in conjunction with humans. We implement our control method, described in subsection 6.2.7, using the franka-ROS API with the torque controller interface, which provides all of the model parameters listed above. While we chose to use the Panda due to its ubiquity, it should be noted that our work could generalize to similar collaborative arms on the market.

6.2.2 Force sensing end-effector

While the Panda has torque sensors on each of its joints, it quickly became clear that these were ill-suited to measure the relatively small forces applied to the NP swab. We moved the arm without an end-effector along the trajectory described in Section 6.2.6 and observed that the measured torques projected as force onto the end-effector, shown in Fig. 6.2, reach noise levels between 40 mN to 60 mN standard deviations. In addition, model errors cause non-stationary drift because when the arm changes configuration the uncompensated weight of the links get erroneously interpreted as external torque. Under these conditions, it is clear that any low magnitude forces that would be transmitted through a swab would be swallowed up by these disturbances if we were to solely rely on the built-in torque sensors. Therefore, we designed a custom end-effector to mount a NP swab and sense the 3 axial forces applied to it. We integrate a GPB160 10 N capacity tri-axial strain gauge loadcell (Galoce, Shaanxi, China) onto a 3D printed housing to be mounted onto the Panda's flange. A 3D printed mount that fits the end of an NP swab is affixed to the exterior facing side of the loadcell (The 3D print STL files are available in the supplementary data). We adopt a fairly simple electronics setup: each of the leads from the load cell axes are soldered into an HX711 Wheatstone bridge amplifier set to 80 Hz mode that is interfaced with an Arduino Nano. An image of these components is shown in Fig. 6.3. A ROS node was created to publish the values sensed from the three axes via the controlling computer. Overall, the entire cost of the materials for the sensorized end-effector was less than $300 USD, which makes it an affordable solution, provided a robot arm is already available. The base of the end-effector was 3D printed with PLA to screw into the Panda's flange (DIN ISO 9409-1-A50). This base piece could easily be adapted and reprinted to other mounting flanges on other robots. While this prototype is made with swabs as the application, we foresee that other tools could be attached to the loadcell that could enable other medical tasks.

In practice, the implementation is sufficiently sensitive because of the fine capacity of the loadcell. The electronic noise present in the signal remains quite low as well, having a standard deviation of about 1 mN that is further low pass filtered for reasons described in subsection 6.2.5. However, one detail that requires special attention is the impact of gravity on the readings. As the end-effector changes orientation, the forces due to gravity will need to be subtracted within the end-effector frame in order to isolate the raw forces applied to the swab. In an ideal case, we would subtract the weight attached to the loadcell. However, the wires leading from the loadcell create non-trivial effects on the readings as their tension changes with gravity.

The net force read on the loadcell can be described as

$$\mathbf{F}_{net} = \mathbf{F} + \mathbf{G}(\mathbf{o}) + \mathbf{Z} \tag{6.2}$$

where F is the external force vector applied to the swab, $\mathbf{G}$ is the force gravity vector as a function of the end-effector orientation $\mathbf{o} = (o_x, o_y, o_z)$, and $\mathbf{Z}$ is the bias offset. We devised a calibration routine where the end-effector is moved to nine different orientations with no external force ($\mathbf{F} = \mathbf{0}$). The gravitational force is estimated with a multiple linear regression model. The regression analysis found significant effects by the non-zero terms in the $\mathbf{A} \in \mathbb{R}^{3\times4}$ matrix written below, with one significant interaction term $o_y o_z$ for G_y.

$$\mathbf{G}(\mathbf{o}) = \begin{bmatrix} A_{x,x} & A_{x,y} & A_{x,z} & 0 \\ A_{y,x} & A_{y,x} & A_{y,z} & A_{y,yz} \\ 0 & 0 & A_{z,z} & 0 \end{bmatrix} \begin{bmatrix} o_x \\ o_y \\ o_z \\ o_y o_z \end{bmatrix} = A \begin{bmatrix} o_x \\ o_y \\ o_z \\ o_y o_z \end{bmatrix}. \tag{6.3}$$

Finally, we can solve for the parameters of the gravitational + offset model via a least squares problem over the $\mathbf{F}_{net}$ and $\mathbf{o}$ values from the nine orientations

$$\min_{\mathbf{A},\mathbf{Z}} ||\mathbf{F}_{net} - (\mathbf{A} \begin{bmatrix} o_x & o_y & o_z & o_y o_z \end{bmatrix}^\top + \mathbf{Z})||^2 \tag{6.4}$$

The fit for this calibration procedure is quite good, achieving $R^2 > 0.9999$, indicating that this method can effectively eliminate the gravitational effects.

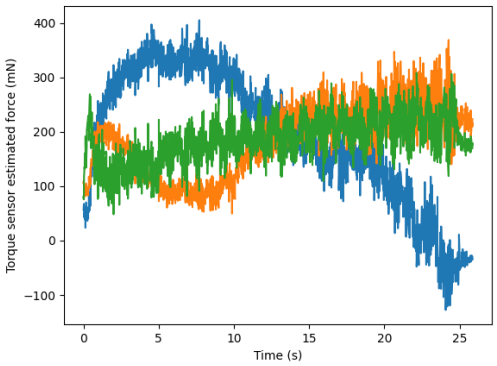


Figure 6.2: Observed noisy task-space force measurements when using the FR internal torque sensors. This demonstrates the necessity of designing an end-effector with an external loadcell that is sensitive enough to observe forces transmitted by the swab.

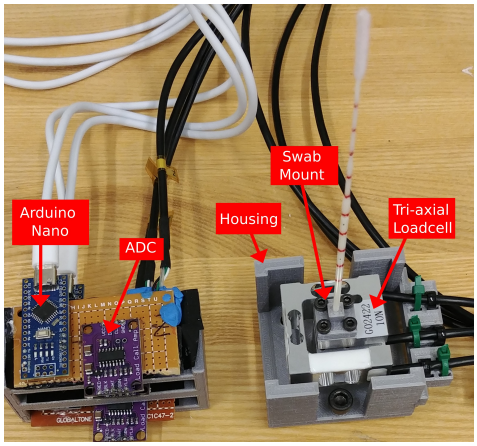


Figure 6.3: Components of the custom end-effector. Left: Wheatstone bridge amplifier and digital conversion circuit. Right: the tri-axial loadcell interfaced with its housing and swab mount. Note that the red stripes on the swab are labeled to gauge distance in experiments, but do not have any functional purpose related to the controller.

6.2.3 Nasal cavity apparatus and experiment setup

The other component used in this work is the nasal cavity phantom that is used to evaluate the proposed methods. We use the 3D printed nasal cavity phantom (shown in Fig. 6.4) designed by Sananès *et al.* [110], that was printed with a PolyJet 3D printer using the rubber-like Agilus 30 to print the fleshy tissue within the nasal cavity. The figure also shows the ideal path a swab should follow, leading from the nostril to the nasopharynx. The phantom is placed into an open fitting container and was augmented by gluing a force sensing resistor (FSR) onto the nasopharynx. The apparatus was clamped to a tripod so that it could be re-positioned and reoriented freely. Two GoPro Hero 8 cameras were used to record the experiments and were aligned with an adjacent tripod such that they faced the apparatus in the top-down and sideways directions. Fig. 6.1 shows a photo of this setup.

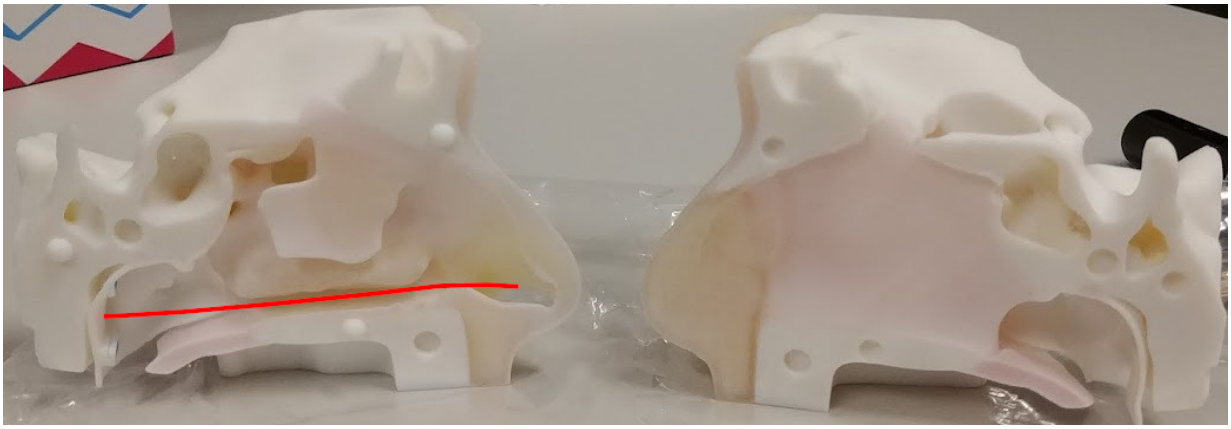


Figure 6.4: Nasal cavity phantom used for validation experiments. The red line highlights the path to reach the nasopharynx from the nostril.

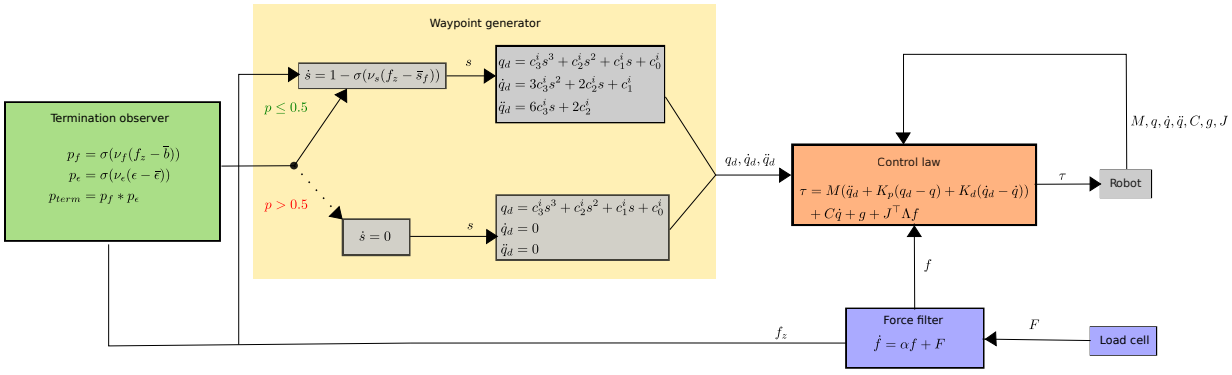


Figure 6.5: Block diagram for the proposed NP swab insertion system.

6.2.4 Control system

The proposed system consists of four components: a force filter for the loadcell readings, a waypoint trajectory generator, the torque control law, and an observer to determine when the insertion procedure is completed. A block diagram summarizing the interaction between these components is shown in Fig. 6.5. Each of these components are described in the following sections.

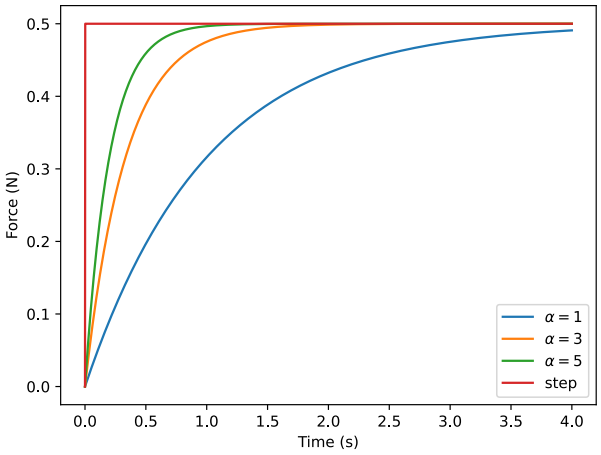


Figure 6.6: Example of the force filter response to a 0.5 N step function with $\alpha = 1$, $\alpha = 3$, and $\alpha = 5$.

6.2.5 Force filtering

As described in subsection 6.2.2 the loadcell force readings already have a low level of noise. However, it is still desirable to filter the force signals because of the impulses that occur from intermittent contacts between the swab and the nasal cavity. If a force-feedback controller were to react directly to the impulses, motion becomes jerky and unstable as it rebounds between these intermittent contacts, which is undesirable and dangerous behavior to have within the nasal cavity. We therefore use the filter

$$\dot{\mathbf{f}} = -\alpha \mathbf{f} + \alpha \mathbf{F}, \tag{6.5}$$

where $\mathbf{F}$ is the raw force measurement coming from the loadcell, $\mathbf{f}$ is the filtered signal that will be used in the control system, and $\alpha \geq 0$ is the response rate. This is a continuous implementation of an exponential moving average filter; Fig. 6.6 shows an example of how α influences the step response. Ultimately, we want the filter to attenuate sudden changes in force and respond to consistent levels of force. Having a delayed response in $\mathbf{f}$ is ideal because this encourages the controller to make gradual adjustments and discourage jerky and unstable motion. In our experiments we chose $\alpha = 1$ because it provided the response we desired.

6.2.6 Waypoint generation

From our previous work [69][1], we studied the deformation of NP swabs as they were inserted through the nasal cavity and found the optimal linear trajectory to insert at to minimize swab deformation. We build on these ideas by projecting a linear task-space trajectory to follow in the arm's workspace. We take insertion at a 28 degree decline angle (see [98]) over 20 cm, selecting 32 points along this line and solving for continuous joint configurations ($\mathbf{q}$) with inverse-kinematics using RBDL-Casadi [87]. Cubic splines were fit between each of the points, having coefficients $\{\mathbf{c}^i\}$ with respect to the path parameter s. The desired nominal joint waypoints are computed as

$$\begin{aligned} \mathbf{q}_d &= \mathbf{c}_3^i s^3 + \mathbf{c}_2^i s^2 + \mathbf{c}_1^i s + \mathbf{c}_0^i \\ \dot{\mathbf{q}}_d &= 3\mathbf{c}_3^i s^2 + 2\mathbf{c}_2^i s + \mathbf{c}_1^i \\ \ddot{\mathbf{q}}_d &= 6\mathbf{c}_3^i s + 2\mathbf{c}_2^i, \end{aligned} \tag{6.6}$$

where $\mathbf{c}^i$ is the coefficient vector that corresponds to the spline within the domain of s. These splines are all computed offline, and the coefficients are used during runtime to determine the associated goal positions, velocities, and accelerations.

We design the progression of s with respect to time as

$$\dot{s} = 1 - \sigma(\nu_s(f_z - \bar{f}_s)), \tag{6.7}$$

where σ is the sigmoid function; the scalars ν_s and $\bar{f}_s$ influence the scale and offset of the sigmoid response. The purpose of this formulation is to allow the nominal trajectory to respond to disturbances via the force reading f_z, which faces the direction the swab will be moving along. When there is little f_z, the sigmoid remains unsaturated and the trajectory proceeds normally. When there is high f_z, the sigmoid becomes saturated and the trajectory slows down to allow time for the controller, described in the subsection 6.2.7, to make adjustments and respond to the disturbance. We set $\bar{f}_s = 0.33$ N and $\nu_s = 12$ N^{-1}; based on experimentation, these values allowed the trajectory to slow down on when it encountered early contacts without stopping completely.

6.2.7 Torque control law

The task of the control law is to strike a balance between two things: a) following the nominal trajectory waypoints, $\mathbf{q}_d, \dot{\mathbf{q}}_d, \ddot{\mathbf{q}}_d$ and b) adjusting to contact forces $\mathbf{f}$ that are applied to

[1] Chapter 4

the swab. As a result, we designed the admittance controller based on a computed torque control law [4] using feedforward acceleration with position, velocity, and force feedback

$$\boldsymbol{\tau} = \mathbf{M}(\mathbf{q})(\ddot{\mathbf{q}}_d + \mathbf{K}_p(\mathbf{q}_d - \mathbf{q}) + \mathbf{K}_d(\dot{\mathbf{q}}_d - \dot{\mathbf{q}})) + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) + \mathbf{J}^\top \boldsymbol{\Lambda}\mathbf{f}. \tag{6.8}$$

Here, $\ddot{\mathbf{q}}_d, \dot{\mathbf{q}}_d, \mathbf{q}_d$ and $\dot{\mathbf{q}}, \mathbf{q}$ are the acceleration, velocity, and positions for the nominal joint waypoints and the actual joints, respectively. $\mathbf{M}$ is the mass inertia matrix, $\mathbf{C}$ is the Coriolis matrix, $\mathbf{g}$ is the gravitational vector, $\mathbf{J}$ is the 7×3 joint-euclidean space Jacobian matrix, $\mathbf{K}_p = \text{diag}(600, 600, 600, 600, 600, 600, 50)$ s^{-2}, $\mathbf{K}_d = \text{diag}(30, 30, 30, 30, 30, 30, 5)$ s^{-1} are diagonal gain matrices for position and velocity. One can see how the first line is used to drive the trajectory towards the nominal waypoints. The acceleration term is fed forward in the control loop, and the position and velocity terms are used as feedback to correct for errors, which is generally understood to provide better tracking than PD control on its own [4]. The gravitational and Coriolis forces are included to counteract these effects from Equation 6.1. Finally we amplify the small scale $\mathbf{f}$ with $\boldsymbol{\Lambda} = \text{diag}(450, 450, 45)$ to promote motion and map it to joint torques with $\mathbf{J}^\top \boldsymbol{\Lambda}\mathbf{f}$. The purpose of this component is to shift the trajectory away from contact forces, with the goal of correcting for misalignment. The gain values for $\mathbf{K}_p$, $\mathbf{K}_d$, and $\boldsymbol{\Lambda}$ were chosen with trial and error and could be refined in the future.

6.2.8 Termination observer

The last component of the system is an observer to estimate when the swab has reached the nasopharynx, during which the insertion stage should be terminated. Typically this would be followed with rotating the swab to collect samples, but this motion is deferred for this paper. With the absence of additional sensors, the controller makes this determination based on two sources of information a) the Z-axis force and b) the total positional displacement of the end-effector. Fuzzy logic [140] presents a robust way of making this decision, which has seen use in situations where decisions rely on measurements that are linked to a state with uncertainty [36] [108]. The fuzzy model we use is

$$\begin{aligned} p_f &= \sigma(\nu_f(f_z - \bar{f}_z))) \\ p_\epsilon &= \sigma(\nu_\epsilon(\epsilon - \bar{\epsilon})) \\ p_{\text{term}} &= p_f p_\epsilon, \end{aligned} \tag{6.9}$$

where σ is the sigmoid function and the termination decision p_{term} is activated when both the terms for force, p_f, and position, p_ϵ, are sufficiently saturated. The force sigmoid p_f

Table 6.1: Contingency table comparing the rates of success between the two controllers. A chi-square test results in a p-value = 2.5×10^{-4}, indicating that the proposed method has a significantly higher success rate.

Controller	Success	Failure
Force-feedback	33	9
Baseline	16	25

Table 6.2: Summary of measured forces applied during the swab over all trials. Peak force and the forces averaged over the duration of the insertion are compared between the two controllers using paired t-tests.

	Force-feedback(mN)	Baseline(mN)	p-value
Avg. Force	250 ± 112	377 ± 210	4.6×10^{-5}
Peak Force	1064 ± 325	1347 ± 516	5.3×10^{-4}

saturates when f_z rises, which we expect to happen when the swab tip makes contact with the nasopharynx. Likewise, the position sigmoid p_ϵ saturates when the total displacement of the end-effector position from its starting position, ϵ, is high enough so that contact with the nasopharynx is possible. The parameters ν_f, ν_ϵ, $\bar{f}_z$, and $\bar{\epsilon}$ control the scale and intercept of the sigmoid activations. We set the threshold for p_{term} at 0.5, where if at any point $p_{\text{term}} > 0.5$, the trajectory halts. We set $\bar{f}_z = 0.167$ N, $\nu_f = 30$ N^{-1}, $\bar{\epsilon} = 0.085$ m, and $\nu_\epsilon = 40$ m^{-1}. Intercept $\bar{\epsilon}$ was chosen based on the nasal cavity geometry as we expect the swab to travel at least 9 cm. The other parameters were chosen by trial and error; these values could also be refined because we typically saw p_f become saturated before p_ϵ.

6.3 Experiments

The goal of the following experiments is to evaluate how well the proposed force-feedback controller can insert the swab into the nasal cavity phantom described in subsection 6.2.3. We consider the scenario where the arm is already positioned and oriented in front of the nose and is ready to be inserted. We also ignore the latter stages of the test that

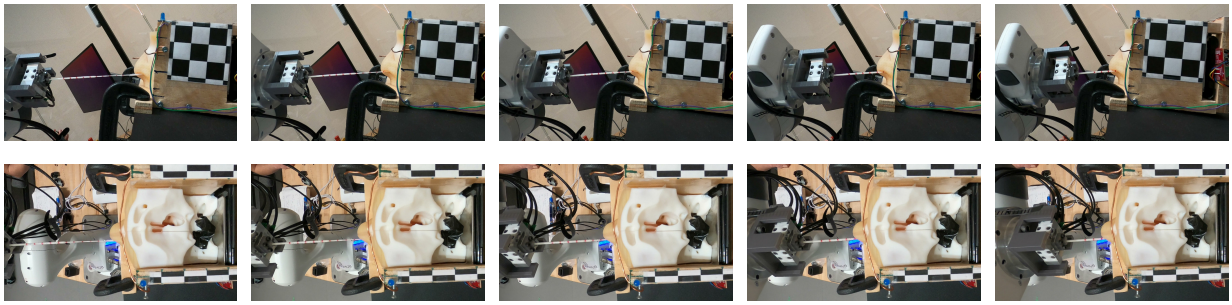
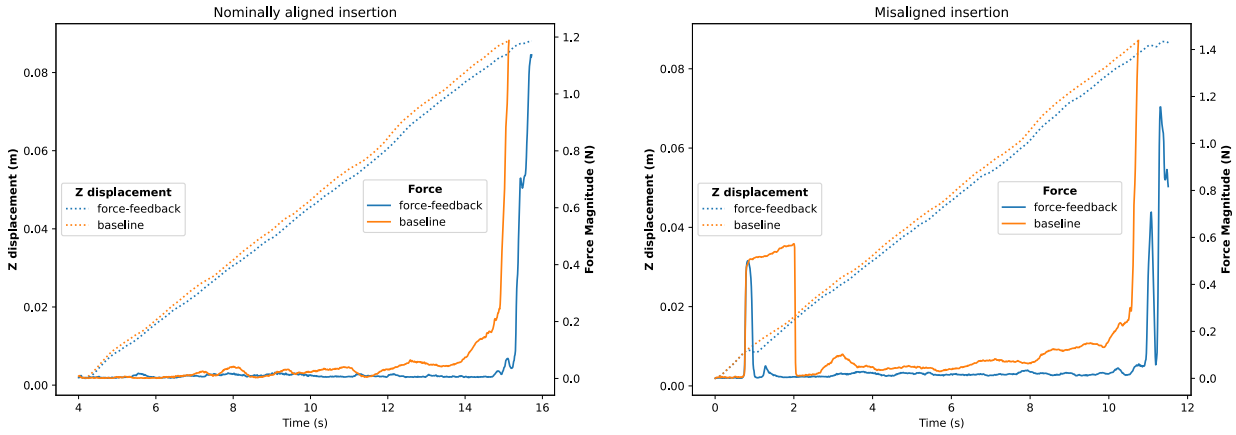


Figure 6.7: Side and top frames of a successful insertion taken by the two GoPro cameras.

would typically involve swab rotation, extraction, and sample preparation. Two different controllers are examined: the proposed controller with force-feedback and a baseline controller that does not use force feedback. Specifically, the baseline controller uses the same control law (Equation 6.8) but sets $\mathbf{\Lambda} = 0$, ignores force for determining waypoints (i.e., $\dot{s} = 1$), but still uses the same termination condition Equation 6.9. Comparisons are made between the two controllers to test whether our proposed control method improves upon solely following the nominal trajectory.

The insertion trials were done by moving the arm to the first waypoint, after which the apparatus was manually positioned in front of the swab. The controllers were evaluated in a paired manner, where the force-feedback controller would execute, then after resetting the arm to the starting position without moving the apparatus, the baseline controller would execute. We executed a total of 41 of these paired trials. Between each of these trials we would move the apparatus to a unique position and angle in order to add variance and also to simulate pose estimation errors that would occur if a visual system were guiding the initial placement. Fig. 6.7 shows frames from top and side videos that were taken during a successful insertion. Typically the insertion would take about 15 seconds, but the total time varied depending on the disturbances encountered from the misalignment. The video attached in the supplementary data includes a couple of recordings taken during the insertion trials.

Data recorded in the trials consists of videos from the two cameras and the published robot states, loadcell, and FSR values in the ROS server. Success was defined based on if the swab reached the nasopharynx and was largely determined by visual examination of the swab because the threshold of force needed activate the FSR was too high. As shown in Table 6.1, the trials resulted in a higher insertion success rate of 78.6% for the force-feedback controller compared to the baseline controller with 39.0%. Comparing these outcomes with a chi-square test shows that the result is statistically significant $p = 2.49 \times 10^{-4} < 0.05$. It



(a) (b)

Figure 6.8: Comparison of total observed forces and the accompanying displacement along the insertion axis for both versions of controllers recorded from the start of motion until the termination condition is reached, which is marked by the impulse of force as the swab reaches the nasopharynx. a) forces encountered during a nominal insertion angle. b) forces encountered during a misaligned insertion. Notice how in the bottom graph the proposed controller is able to adjust and minimize a collision that occurs early in the insertion.

is also interesting to examine the amounts of force applied to the swab, as this can be used as a proxy to the quality of the insertion and to the patient's comfort. The forces that were sustained on the loadcell during trajectories are shown in Table 6.2. In a paired manner, the force-feedback controller sustained less forces than the baseline controller, but there was wide inter-trial variance based on the swab's initial placement. Fig. 6.8 shows two graphs comparing the forces and displacement observed by the controllers on two different initial alignments, which run until the termination observer triggers from sufficient force and displacement. The left graph shows the forces for a nominally positioned swab, and the right graph shows the forces for a misaligned swab. In the latter case, the force-feedback controller is able to adjust itself and sustain less persistent force compared to the baseline. From examining the final end-effector pose from forward kinematics of the recorded joint angles, the impact of the force feedback is also apparent: the trajectory from the left graph had a displacement of 2.4 mm and 0.8°, while the right graph was altered by a larger factor with a displacement of 6.9 mm and 2.1°.

6.4 Discussion

As we showed in Table 6.1, the baseline controller was significantly more successful than the force-feedback controller according to a Chi-squared test, suggesting that incorporating force in the control loop enabled the insertion trajectories to be successfully altered to non-ideal alignments with respect to the phantom. The lower sustained forces in the force-feedback controller also indicates that it performed better than the baseline. The peak force was about 0.4 N higher than the expert practitioner's from Park *et al.* [95], which could stem from physical differences in the phantom or from them only using nominal insertion angles, but could also indicate that our controller parameters in Equation 6.7 and Equation 6.9 could be better tuned in the future. Hence, the outcome of these trials show that there is certainly benefit to incorporating such a force-sensed based compliant control system for the proposed robotic arm setup that will increase insertion success and patient comfort versus a system that only uses position feedback.

Generally the compliance of the swab meant that there was significant allowance for the swab to be misaligned and still make it to the nasopharynx. Qualitatively, we notice that the pose and ease of insertion seemed to correlate with the findings from previous work [69][2]. Insertion angles that were oriented towards the septum were generally more successful than those oriented away from the septum. Being positioned away from the nasal vestibule wall also seemed to be important for avoiding the wedging state and to reduce strain on the swab. While characteristics may differ for individual anatomy, we plan to take these observations into account when we design the vision guided positioning system for the initial placement of the swab in future work.

It is insightful to examine the cases where the insertion failed for either controller. There were two types of failure states that occurred during insertion. The first type of failure was unique to the baseline controller, which failed to reach the nasopharynx because the elevation angle of the swab was too high (see Fig. 6.9 a). Typically this transpired when the swab was positioned too low and then became levered into an excessive angle where it deflected off the sphenoid sinus, which would likely be more uncomfortable and have higher chance of complications for a human patient. The force-feedback controller was able to adjust the trajectory to avoid this levering effect because of the adjustments the force-feedback produced. The swab becoming stuck and being unable to enter the nasal cavity, as shown in Fig. 6.9 b, was the second type of failure. This failure state was the result of the swab becoming wedged on the nasal vestibule because of poor alignment and appeared with both the force-feedback and baseline controller. This outcome underscores

[2]Chapter 4

the importance of implementing a feature that can detect when the swab fails to enter the nostril, thereby prompting a realignment process. Such a capability will be crucial for enhancing the system's effectiveness in future iterations.

In terms of future work, one of the major aspects that would need to be resolved is enforcing guarantees for safety. One such area is providing guarantees for the stability of the controller. While the force filter in subsection 6.2.5 helped stabilize the response compared to using unfiltered values, there were still some cases where oscillations were present as the swab reached the nasopharynx. An untested scenario is the controller's response to the natural motion that a human patient would have during the procedure. From the hardware design, it is important to eventually build in a mechanism to detach the swab from the robot in cases where extremely high forces are detected or when the participant wants to abort the procedure. Finally, although the force-feedback controller was able to handle non-ideal insertion poses, it may be fruitful to investigate if applying an estimator that directly estimates pose errors via the force feedback could result in improvement, particularly for resolving the wedging failure case.

6.5 Conclusion

In this work, we proposed a scenario where a standardized fixed rigid arm robot performs NP swab sampling using a compliant control system. To investigate this, we designed a minimal force sensing end-effector and integrated it into a Panda arm that could be adapted to similar rigid arms. We designed an admittance force-feedback torque control system to perform the insertion test. We performed experiments to evaluate our system on a 3D printed nasal cavity from variety of different alignment conditions and showed that the admittance controller was succeeded at a rate of 78.6% compared to 39.0% of the baseline position controller. This demonstrates that there is feasibility for a rigid arm to perform the NP swab test on people using sensitive force feedback as a modality. The compliant controller was able to compensate for some misalignment and thereby avoid some failure states, however it is clear that a full solution will require additional sensors or other strategies to adjust for more extreme misalignment and more study will be needed to evaluate the impact of head motion on the controller.

Future work will extend this research towards implementing a fully automated robotic NP sampling solution. An eye-in-hand visual servo system will be a necessary for reaching the pose in front of the nostril, prior to insertion. The other stages of the NP test (rotating the swab at the nasopharynx and extraction) will need to be implemented. As well, adding additional inputs, such as tracked visual features of the face, may be worth fusing with the

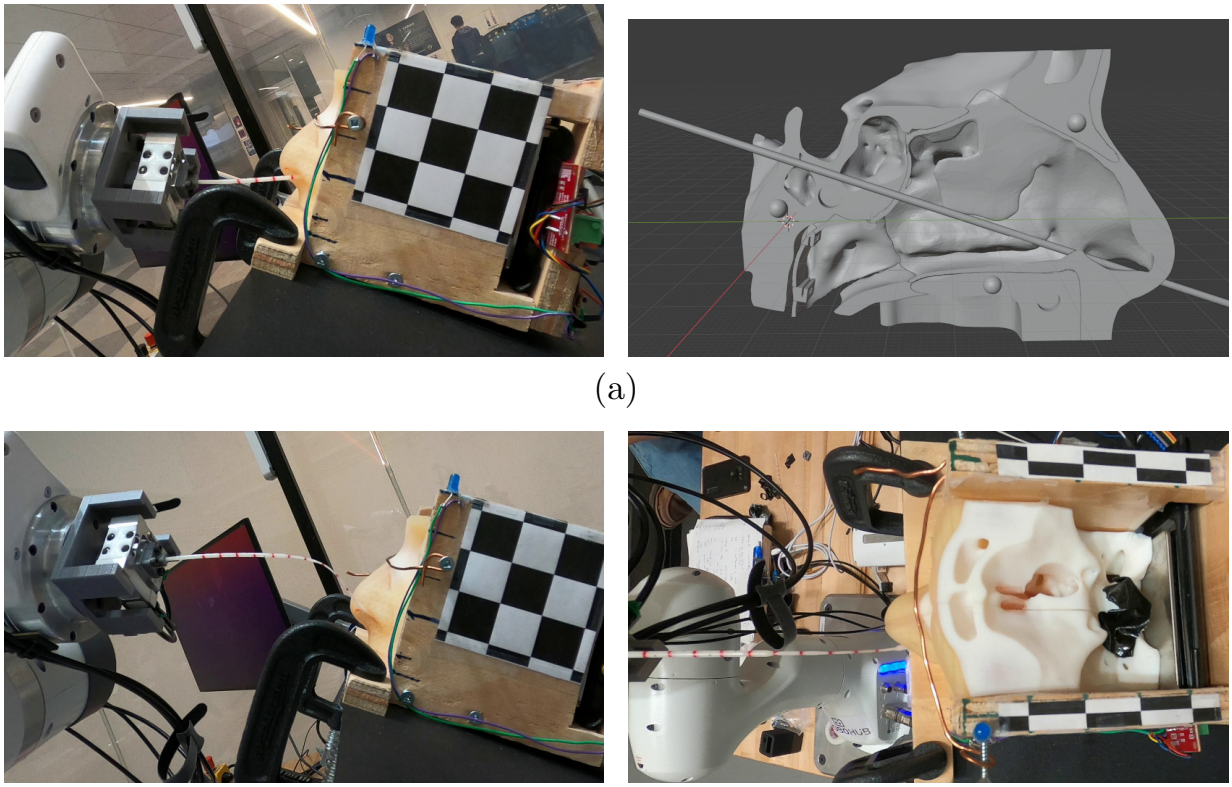


Figure 6.9: Two types of failure states observed in the trials: a) The elevation angle for the insertion is too high, resulting in it traveling through the wrong passage in the nasal cavity. The proposed compliant controller prevented these states. b) The swab becomes wedged before entering the nasal cavity. This state occurred with both controllers.

force measurements as inputs to the control system to better adjust to disturbances such as motions of the head and general misalignment.

While the NP swab test is just one task a healthcare worker would do in a clinical setting, there are many other types of routine close-contact tasks that would require similar levels of dexterity and force sensitivity. Examples include other types of sample collection, medication or blood collection through needles, and skin temperature testing. Consequently, having a multi-purpose robot that could do NP swab tests could open up a multitude of different clinical tasks, and could be a boon for the healthcare system by allowing procedures to be done autonomously, enabling human resources to be reallocated.

Chapter 7

Integrated swabbing with a moving phantom

Foreward

The contents within this chapter are written originally for this thesis, although we plan to submit its contents to a journal at a later date. Chronologically, the experiments in this research were conducted from December 2023 to February 2024.

7.1 Introduction

In the previous chapters, we have proposed different solutions to each of the subproblems associated to the NP swab test. In Chapter 4 we found the best way to insert the swab through the nasal cavity with simulation and optimization. In Chapter 5 we detailed the visual servo system that aligns the swab next to the patient during the pre-contact phase. Chapter 6 presented our initial work towards a compliant control system to execute the contact phase using a computed torque control law.

In this chapter, we present two significant novel contributions. First, we integrate all previously developed components into a cohesive system capable of executing both the pre-contact and contact stages of the NP swab test on hardware. This integration represents the first comprehensive robotic system designed to perform all stages of the NP swab test autonomously on a freestanding patient. Second, we introduce the use of a second

robotic arm to simulate a human patient with natural head motion, which is used to test the control system's ability to cope with dynamic conditions. This innovative dual-arm setup allows for a more realistic assessment of the system's performance under real-world conditions. Fig. 7.1 provides an illustration that shows how the first robot arm responsible for swabbing is set up beside the second robot arm that simulates the patient.

In Section 7.2 we provide a brief background of head motion, then in Section 7.4 we describe physical setup of the simulator and the proposed head motion generator. Some modifications are made to the previous systems to enable them to work in this new context, which is summarized in Section 7.3. The simulated trajectory is reparameterized to include a turn at the beginning of the trajectory to avoid the overhang of the nose (Section 7.6). The measurement pipeline of the visual servo system was modified to adapt to the non-anthropomorphic appearance of the nasal cavity phantom (Section 7.5). The compliant control law was ultimately reformulated to a velocity control law to follow trajectories more accurately, and to avoid latency and friction issues that are present in torque control (Section 7.7). Sections 7.8 and 7.9 describe the observers that were implemented to detect when the nasopharynx is reached and detect conditions to abort the test, respectively. In Section 7.9 we describe our extensive experiments that evaluate different configurations of the controller various levels of random head motion to evaluate the robustness of the controller and to discover the best performing configuration. Section 7.11 provides our discussion of the results and potential avenues to improve the system and extend experimentation. Finally, Section 7.12 summarizes the work done and states the outlook towards potential for human trials.

7.2 Background on head motion modeling

Interest in head motion is present in both social interaction and biomechanical fields, with models typically being targeted for a particular task or action. For example, one area of active research is the generation of realistic facial poses and expressions for speaking tasks [106] [138] [88] [28]. There are also applications to mimicking or teleoperating heads in humanoid robots to look more natural [2] [90] [5]. In biomechanics, the head is but one extremity that is measured for tasks involving motion of the body (e.g., [51] [90]). In terms of kinematic modeling, the head will move along with the torso, but can also independently move by rotation about the neck. Generally, this rotation is modeled with a 3D spherical joint within the neck [106] [90] [35].

Considering the absence of literature examining head motions of patients whom are being subjected to the NP swab test, we look towards stochastic motion models for general

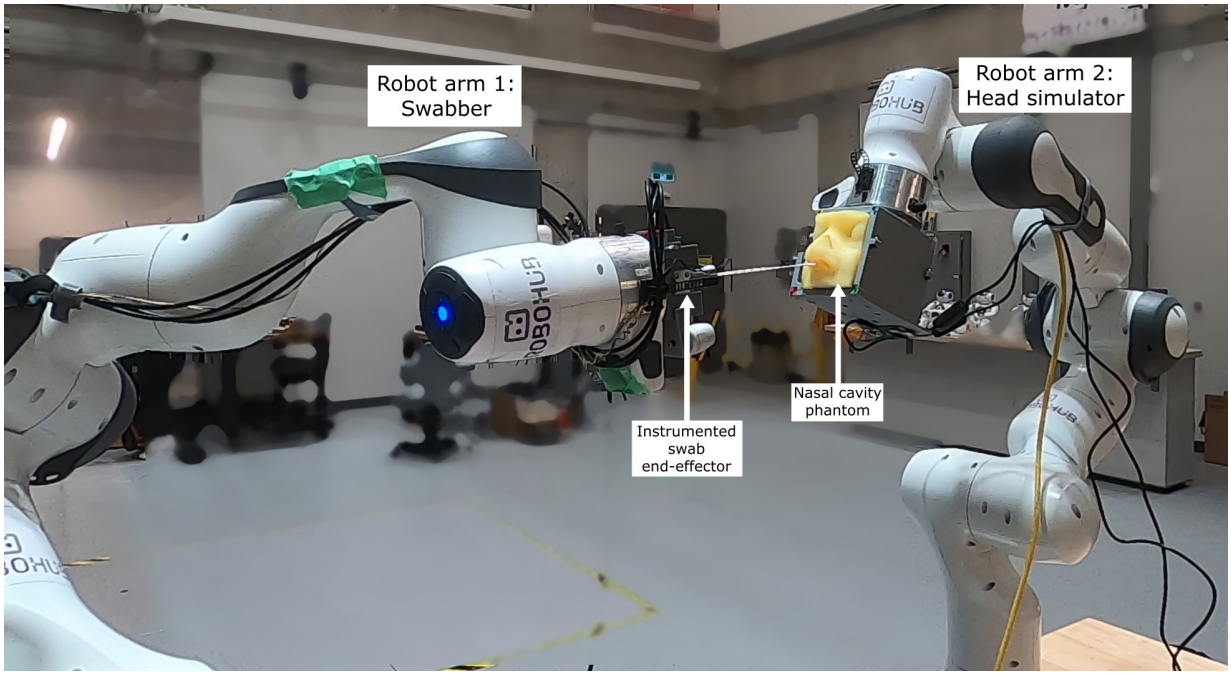


Figure 7.1: Figure showing the setup between the two robotic arms. One arm performs the NP swab test using the instrumented end-effector, while the second arm holds the nasal cavity phantom and simulates head motion.

head motion. The study by Heuring and Murray [49] is most relevant to our scenario. The authors proposed and evaluated two stochastic processes to model 6-DOF head motion during an image-based motion capture study: the Wiener process and the Ornstein-Uhlenbeck process [129]. Ultimately, the authors found that the Ornstein-Uhlenbeck process autocorrelated much better measured head motion trajectories than the Wiener process.

7.3 System overview

The focus of this chapter is to develop and evaluate the full NP swab test and evaluate its performance under variable conditions. The hardware setup consists of two arms that face one another during experiments. The first arm acts as the platform that performs the swabbing and the second arm is the platform that simulates a human patient that randomly moves its head during the procedure (Section 7.4).

As shown in Fig. 7.2 the swabbing task can largely be divided into a pre-contact phase and contact phase. The pre-contact phase generally follows the same procedure as described in Chapter 5, with some new modifications added in Section 7.5. The contact phase is further divided into three substages, as shown in Fig. 7.2. The first substage inserts the swab through the nasal cavity. Once the swab reaches the nasopharynx, stage 2 begins, where the swab stays at the nasopharynx and rotates ($\pm$ 45°) for 15 seconds to emulate collecting a sample. Stage 3 extracts the swab from the nasal cavity. The baseline trajectory followed during stage 1 and in reverse for stage 3 is derived via an optimization problem described in Section 7.6. The contact phase uses a compliant joint velocity controller, which will be described in Section 7.7, that follows a prescribed baseline trajectory while reacting to forces applied to the swab. The transition between stage 1 and 2 hinges on detecting when the swab reaches the nasopharynx, which is handled by a fuzzy logic system that is defined in Section 7.8. Finally, we also implemented a safety monitor to facilitate aborting the procedure in case safety limits are violated (Section 7.9).

7.4 Head motion simulator

The second arm dedicated towards simulating head motion has a 3D printed PLA fixture attached to the flange, which holds a nasal cavity phantom that is described in the following paragraph. The fixture itself, as shown in Fig 7.3, holds five fiducial markers to support pose estimation. At the back of the phantom, underneath where the nasopharynx sits, a

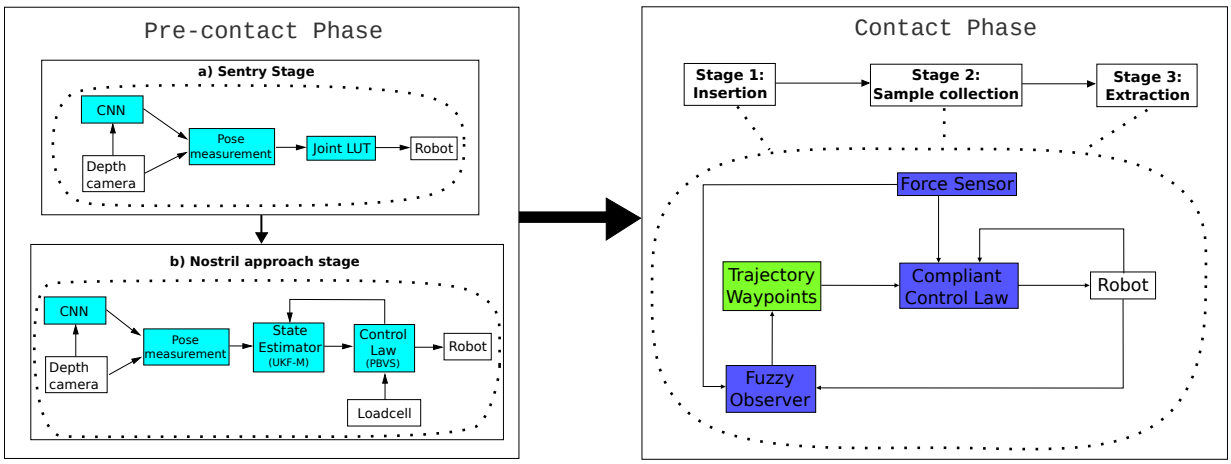


Figure 7.2: Diagram of the stages and proposed system for a robot executing the NP swab test. The test is divided into a pre-contact phase and a contact-phase. The pre-contact phase is responsible for aligning the swab next to the nose using visual information. A sentry stage locates the face within the arm's workspace and moves to a closer joint configuration, while the nostril approach phase uses visual servo to place the swab next to the nostril. The contact phase has an insertion stage that moves the swab through the nasal cavity, the sample collection stage that holds and rotates the swab at the nasopharynx, and an extraction phase where the swab is removed from the nasal cavity.

Raspberry Pi camera and an illuminating white LED take images of the nasopharynx in order to observe whether the swab makes contact.

The testing phantoms were based off of the designs by Sananès *et al.* [110]. This phantom was designed to be a realistic training apparatus for training HCWs to take NP swab samples during the COVID-19 pandemic. The phantom is fabricated using a Stratasys (Minnesota, USA; Rehovot, Israel) J750 PolyJet printer using Vero, a rigid material, and Agilus30, a rubber-like material to simulate the material properties of tissue. In addition, we followed the recommendations of Sananès *et al.* [110] and applied petroleum jelly to the interior of the cavity to simulate the lubricating effect of mucus. However, when comparing the phantom to 3D models [62] and CT scans [66] [117] published in the literature, it appears that the inferior turbinate in the phantom is shrunken so the corresponding passage the swab would travel through is wider. The inventors may have done this to facilitate easier training of practitioners, but we would like to evaluate our system with a narrower passage to ensure it is robust to different morphologies. As a result, we created a second version of the phantom by coregistering the 3D model of the nasal cavity by Keustermans *et al.* [62] and then manually expanding the inferior turbinate using 3D modeling software Blender to take up approximately the same volume. Throughout this chapter we will designate the original phantom by Sananès *et al.* [110] as Phantom A and our modified one as Phantom B.

To emulate head motion, the arm will be moving with a certain set of kinematic constraints. Specifically, the arm will be rotating the center of the phantom about 3D spherical neck joint located below the head, as shown in Fig. 7.5. The length of the virtual link connecting the center of the head and the neck joint was set to 121.4 mm, the average between the center of the mass of the head and the cervicale according to the study by de Leva [35]. While the neck is capable of 6 DOF [6], we adopt the simplified 3 DOF rotation model that is typically used to model the active range of motion for a standing person [43]. A proportional joint velocity controller is used to move the head to the target rotation. Suppose that the current angles of the virtual joint are (α, β, γ) and the target angles are $(\alpha_d, \beta_d, \gamma_d)$. The control law for the phantom arm is then defined as

$$\dot{\mathbf{q}} = k\mathbf{J}^{\dagger}(\boldsymbol{\kappa}(\alpha_d, \beta_d, \gamma_d) - \boldsymbol{\kappa}(\alpha, \beta, \gamma)), \tag{7.1}$$

where $\boldsymbol{\kappa}(\alpha, \beta, \gamma)$ is the forward kinematics function that computes the Cartesian pose of the end-effector given the angles of the virtual neck joint, $\dot{\mathbf{q}}$ is the joint velocity vector, $\mathbf{J}$ is the robot Jacobian, and $k = 0.5$ is the gain scalar.

The motion generation model is designed under the presumed scenario that the patient is cooperating and therefore attempting to hold their head at a target pose, with some

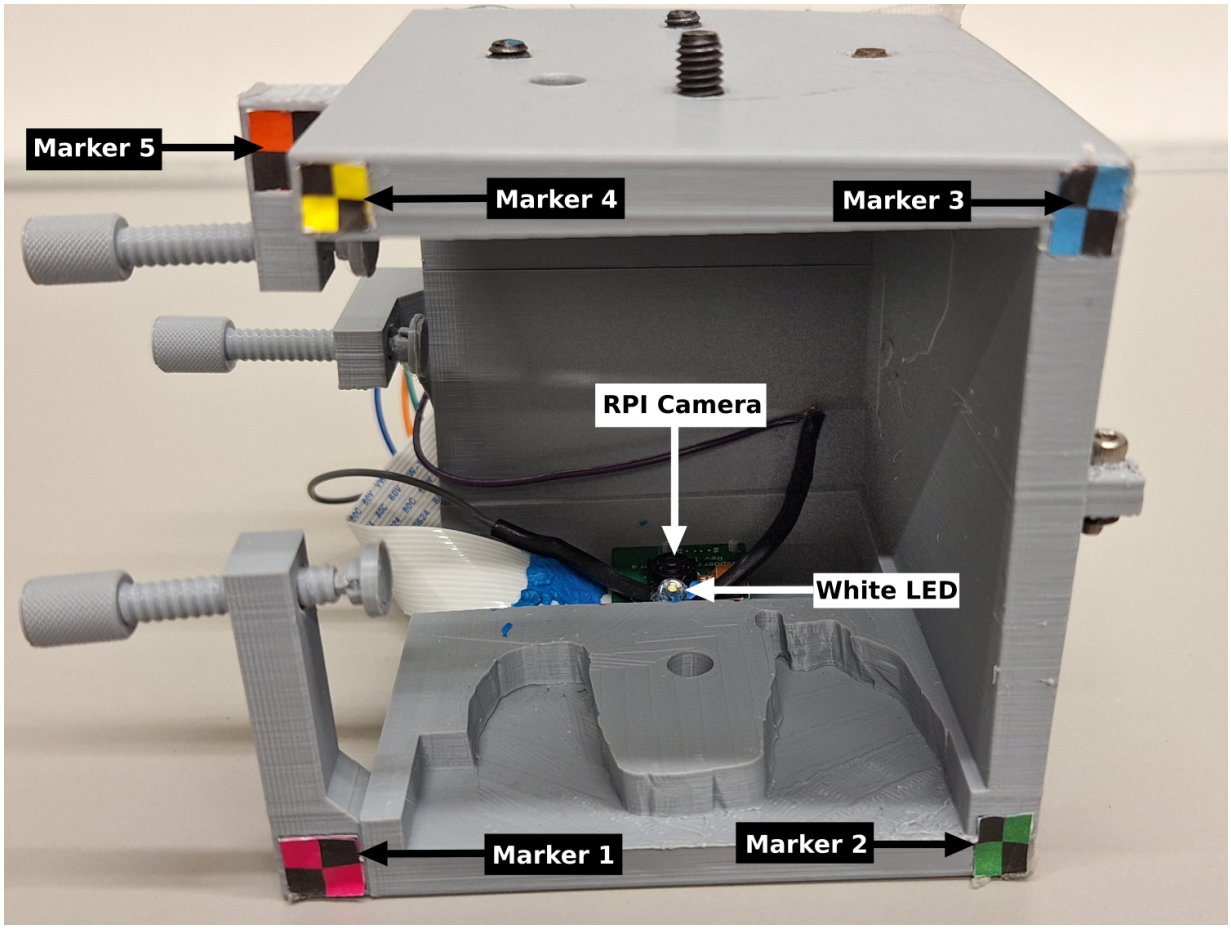
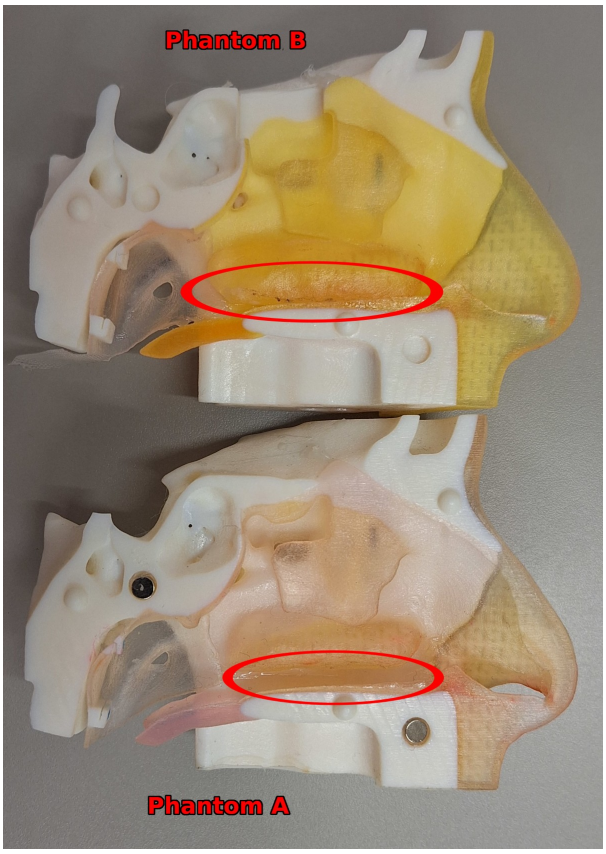
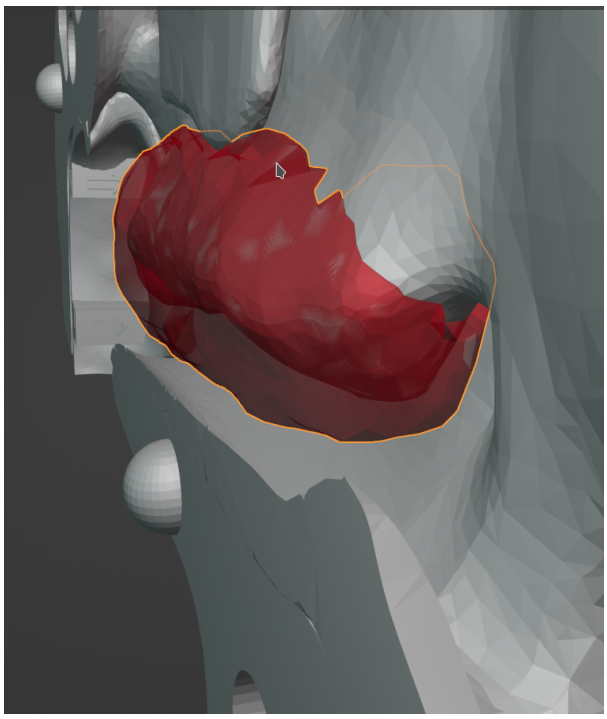


Figure 7.3: Front view of the empty fixture. The fixture contains five fiducial markers to help emulate the head pose estimation. A small Raspberry Pi camera is placed underneath the nasopharynx to determine if the swab successfully makes contact.



(a)



(b)

Figure 7.4: a) Phantom A: the original phantom from Sananès *et al.* [110] and Phantom B: the modified version to have enlarged inferior turbinates. The difference is circled in red. b) View of volume added (red) to Phantom B to more closely resemble the nasal cavity volume from Keustermans *et al.* [62].

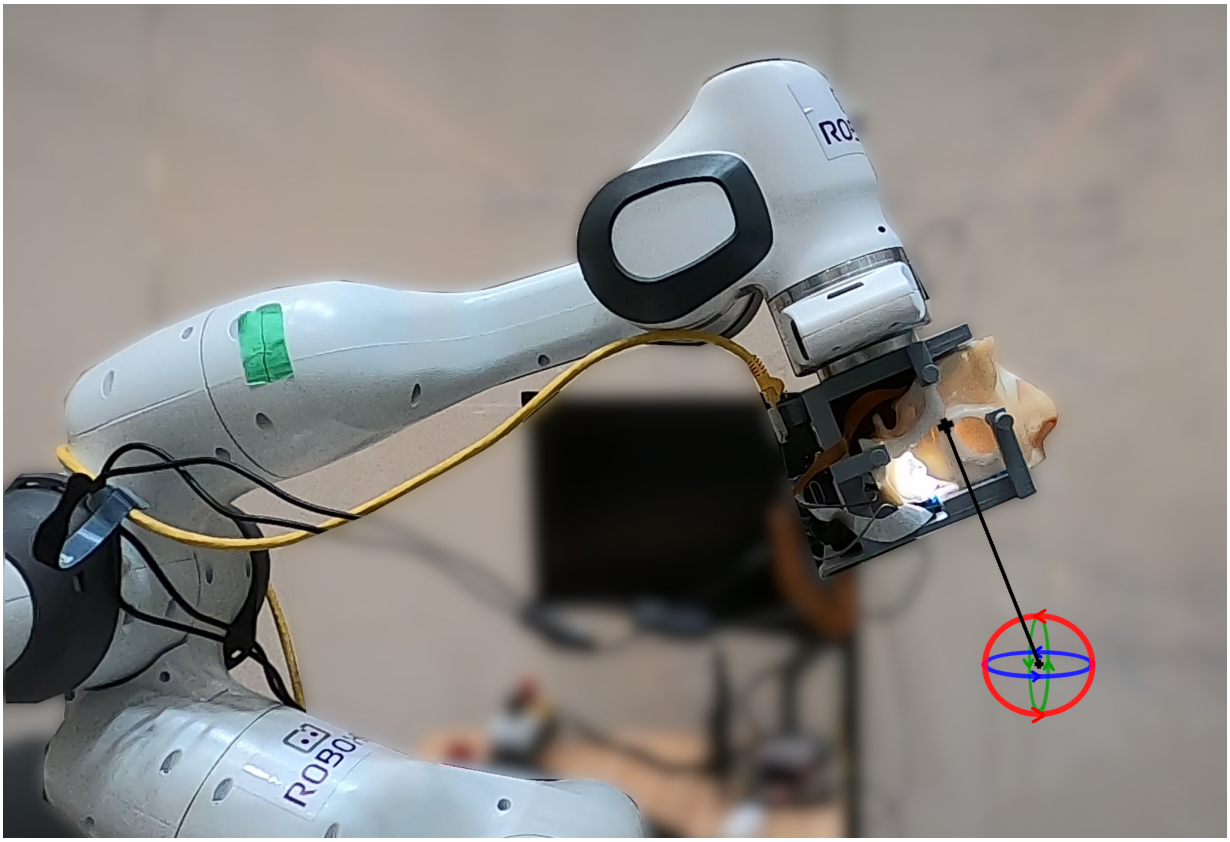


Figure 7.5: Visualization of the kinematic model of the head. The arm moves the head as if it were constrained to a spherical neck joint 121.4 mm below the center of the phantom.

natural, random head motion. Based on the results from Heuring and Murray [49], we chose to model this natural head motion by using a modified Ornstein-Uhlenbeck process, which is a stochastic dynamical system that is generally used to simulate random processes, such as Brownian motion. Ultimately, the stochastic motion model we define for each angle is

$$\begin{aligned} \dot{\theta} &= v_\theta \\ \dot{v_\theta} &= -\mu(\theta - \bar{\theta}) - \psi v_\theta + \sqrt{2\psi}\mathcal{N}(0, \sigma_\theta), \end{aligned} \tag{7.2}$$

where $\theta \in (\alpha, \beta, \gamma)$, $\bar{\theta}$ is the setpoint angle with ($\bar{\alpha} = 0.4$ rad, $\bar{\beta} = 0.0$ rad, $\bar{\gamma} = 0.0$ rad), $\mathcal{N}(0, \sigma_\theta)$ is a zero mean random sample from a Gaussian Normal distribution with standard deviation σ_θ, ψ is the velocity damping coefficient, and μ is the setpoint attraction coefficient. In the standard Ornstein-Uhlenbeck model, the change in velocity stochastically changes with the term $\sqrt{2\psi}\mathcal{N}(0, \sigma_\theta)$ and is damped by adding the term $-\psi v_\theta$. We modified this by adding an attractor term $-\mu(\theta - \bar{\theta})$ so the head does not diverge from too far from the initial pose. Thus, the model has three parameters that control the behavior of the process: μ, ψ, and σ_θ. We decided on three sets of parameters that would correspond to **Light** head motion ($\sigma_\theta = 0.5$, $\psi = 1.0$, $\mu = 1.0$), **Medium** head motion ($\sigma_\theta = 0.7$, $\psi = 1.0$, $\mu = 0.5$), and **Heavy** head motion ($\sigma_\theta = 1.2$ $\psi = 0.75$, $\mu = 0.5$). Fig. 7.6 shows a comparison of the simulated angles over time for the selected levels of head motion.

7.5 Modified visual servo system

In Chapter 5 we designed a visual servo system to align the swab next to a prospective patient. The system used here is largely the same, but has been modified to handle the non-anthropometric appearance of the phantom. Furthermore, we reformulate the gain in the PBVS control loop and add a force feedback term to compensate for increased head motion.

The visual system described in this section estimates the 3D position and orientation of the phantom's nostrils. This approach replaces the 3DDFA_V2 CNN [143] used in Chapter 5 , which was trained to recognize anthropomorphic faces - a characteristic the nasal cavity phantom lacks. The system relies upon detecting the five fiducial markers implanted onto the 3D printed fixture, as previously shown in Fig. 7.3. The fiducials are printed patterns of four colored squares with the intersecting corner being the reference point. Thus, if the visual system can determine the 3D position of at least three markers, the pose of the target nostril can be determined with standard Euclidean rigid registration.

While the markers were designed to be easy to spot, their detection is still subject to

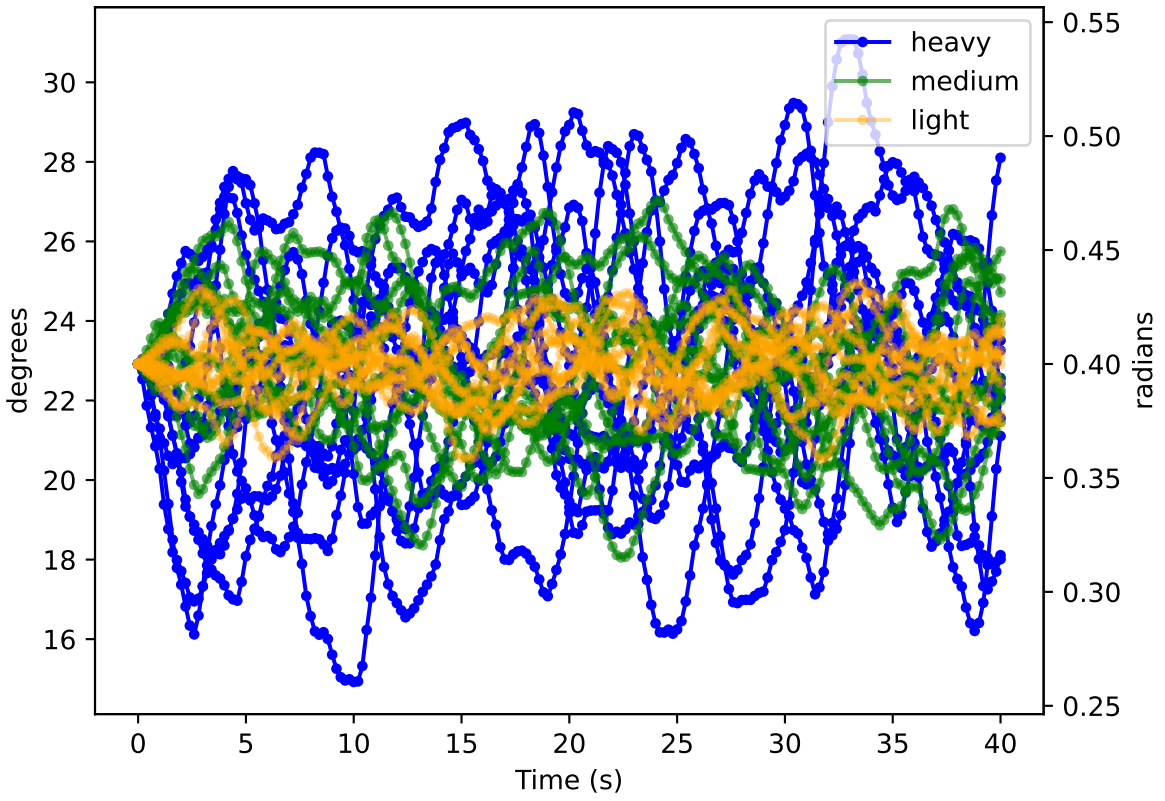


Figure 7.6: Plots showing ten example motions for the proposed modified Ornstein-Uhlenbeck process for **Light**, **Medium**, and **Heavy** head motion. The plots are shown for the angle α with a set point of 0.4 rad., but are similar for β and γ that would have a set point of 0 rad. instead.

lighting conditions that make conventional computer vision techniques less robust. As a result, we used the following pipeline to measure the pose of the phantom. First, we trained a Faster R-CNN [104] convolutional neural network, using a Resnet50 [48] backbone, to find the bounding box of the five markers and the phantom within RGB images. The network was pretrained on the Coco [77] dataset and implemented with the MMdetection framework [27]. The concept of retraining a previously trained model onto a new task is known as transfer learning [48], and it enabled us to retrain the detector to detect fiducials with relatively few training examples compared to training it from scratch. The Faster R-CNN finds the bounding box of the marker, and the pixel position is further refined by finding the intersecting corner of the four squares by applying Harris corner detection [109]. The response is weighted with a Gaussian prior at the center of the box to prevent the detection being skewed by noise outside the marker. The registered depth channel takes the pixel location of the corner and converts it into a 3D coordinate. Once the 3D coordinates of the fiducials are recovered, the pose of the desired nostril can be recovered through least squares rigid registration with respect to their corresponding positions from the schematics. Like in Chapter 5, the measurements are passed through an unscented Kalman filter on manifolds (UKF-M) [20], which is allows a) noise filtering from the pose estimation, b) motion integration from the end-effector into state estimates, and c) maintaining state when measurements are lost from depth acquisition or occlusion.

Recall from Section 5.3.5 that a pose based visual servo operates with a control law that computes the camera velocity to reach a desired target

$$\mathbf{v}_c = -\lambda(||\mathbf{s}||)\hat{\mathbf{L}}^{\dagger}\mathbf{s}, \tag{7.3}$$

where $\mathbf{s}$ is the pose difference vector between the target and the swab tip and $\hat{\mathbf{L}}^{\dagger}$ is the pseudo-inverse of the interaction matrix that maps the observed error to the camera's velocity frame. To allow the visual servo to converge with potentially large amounts of head motion, we also implement the adaptive gain, $\lambda(||\mathbf{s}||)$, described by Kermorgant and Chaumette [61]. We chose parameters $\lambda(0) = 2.5$, $\lambda'(0) = 30$, and $\lambda(\infty) = 0.5$, which enabled the visual servo to converge quickly when there is high head motion.

One additional component we added to the visual servo system is a force feedback term. Especially when the head is undergoing heavy motion, there is a chance the swab can make contact with another part of the face, such as the tip of the nose, while it approaches the nostril. To avoid the swab from getting obstructed, we utilize the force readings from the

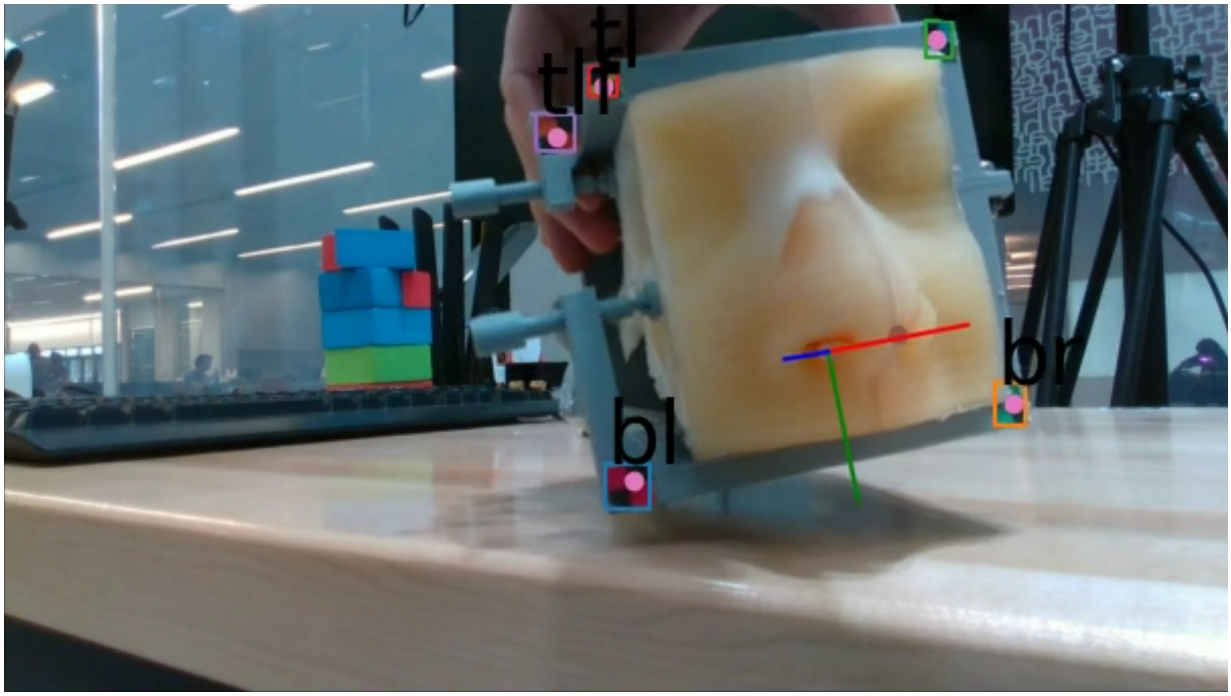


Figure 7.7: Visualization of the pose estimation system adapted for the phantom fixture. The location of at least three of the fiducials leads to the identification of the nostril pose (red-green-blue axes).

loadcell make a repelling motion away from the point of contact:

$$\begin{aligned}\dot{\mathbf{f}}_1 &= -\alpha_1 \mathbf{f}_1 + \alpha_1 \mathbf{F} \\ \dot{\mathbf{f}}_2 &= -\alpha_2 \mathbf{f}_2 + \alpha_2 \mathbf{F} \\ \dot{\mathbf{q}} &= {}^c\mathbf{J}_{ee}^{\dagger}(\mathbf{v}_c + {}^c\mathbf{M}_{ee}\mathbf{\Lambda}(\mathbf{f}_1 - \mathbf{f}_2)),\end{aligned} \tag{7.4}$$

${}^c\mathbf{J}_{ee}$ is the robot Jacobian of the end-effector in the camera's frame, ${}^c\mathbf{M}_{ee}$ is the transform from the end-effector to the camera frame, $\alpha_1 = 0.75$, $\alpha_2 = 0.6$, and $\mathbf{\Lambda} = \text{diag}(0.2, 0.2, 0.9)$ $\text{N}^{-1}\text{m/s}$. The key purpose of the force term is to respond to a sudden change in force from contact events, but to also ignore smaller steady state errors, e.g. from gravity, that would inhibit the controller from reaching the intended visual target. To do this we form a band-pass filter by following the commonly known strategy of taking the difference of two low pass filters [118]. Thus, the response of $\mathbf{f}_1 - \mathbf{f}_2$ will spike shortly after contact and asymptotically converge to zero.

7.6 Trajectory optimization

Once the swab tip arrives at the nostril, the contact stage controller requires some nominal trajectory to move along that would enable it to reach the nasopharynx. In Chapter 4 we developed a solution for this problem by creating a finite element simulation of the swab and finding the optimal linear path in terms of minimizing strain on the swab. We revisit this because the experiments from Chapter 5 revealed that it is necessary for the swab to arrive at the nostril at an incline angle. As a result, we modify our previous approach by formulating a new non-linear trajectory and using a new collision model for the nasal cavity. The trajectory parameterization has changed to start at a 30° angle pitched upwards from the transverse plane (increased from 0.2 radians from Chapter 5 to allow for easier insertion) and parallel to the sagittal plane at the center of the nostril. A 3D vector (χ, e_1, e_2) characterizes the trajectory. The trajectory is divided into three parts. **i)** the swab moves a distance of χ from the starting point at the initial angle; **ii)** the swab rotates about the tip until it points at a target position on an ellipse at the nasopharynx (parameterized by e_1, e_2); **iii)** the swab moves along a straight line towards the target for 150 mm, although the simulation terminates once the swab intersects with the ellipse, which is a distance of about 80 mm from the beginning of part iii). As before, the vector is optimized to minimize strain energy using COBYLA [99]. The new collision model is based off of the nasal cavity phantom from Sananès *et al.* [110], which has the nose overhanging the nostril.

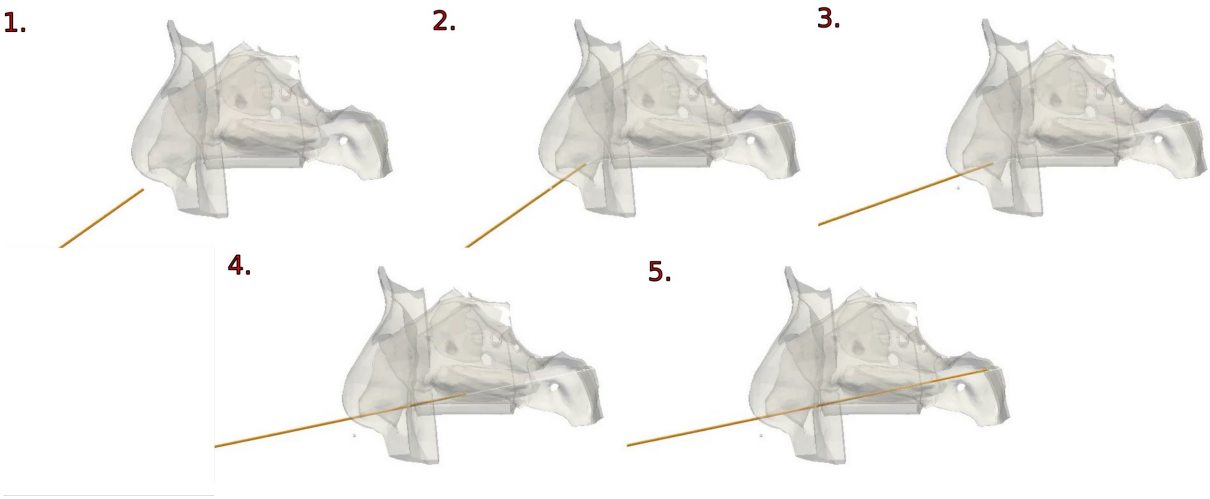


Figure 7.8: Sequence of the optimized trajectory as it travels through the nasal cavity volume. Notice how this trajectory begins at an incline angle and then rotates shortly after entering to point at the nasopharynx.

Fig. 7.8 shows a sequence of the optimized trajectory within the simulation, with $\chi = 17.3$ mm and a change in angle during part iii) of 25.6° pitch and −3.9° yaw. The trajectory is angled towards the septum to avoid the inferior turbinate and also angled slightly upwards to avoid the strain of becoming wedged on the nostril. The upwards angle appears to be the main difference from the previous result from Chapter 4, as the optimization balances the trade-off between avoiding wedging on the nostril and avoiding the inferior turbinate at the back of the nasal cavity.

The selected trajectory is prepared for run-time use by dividing it into a series of Cartesian waypoints, indexed by time. The waypoints are smoothed with triangular ramped acceleration for the beginnings and ends of parts i), ii), and iii) to prevent strain on the hardware. Overall the entire trajectory would take 20 seconds, but in reality the real trajectories took less time because the tip would reach the nasopharynx (see Section 7.8) before the remaining distance in part iii) could play out. During runtime, the waypoints are transformed to the end-effector's frame of reference, then the target is interpolated using a linear spline with respect to time. Similar to Section 6.2.6, the rate of the target trajectory, l, is adjusted depending on the level of axial force, f_z, by using a sigmoid function

$$\frac{dl}{dt} = 1 - \sigma(s_l(f_z - \bar{f}_l)), \tag{7.5}$$

with $s_l = 60\ \mathrm{N}^{-1}$, $\bar{f}_l = 0.5\ \mathrm{N}$. This means that when there is little axial force, the trajectory proceeds at regular speed, but if high axial forces are observed, the trajectory slows down to give the controller time to respond to the disturbance.

7.7 Compliant contact stage controller

Chapter 6 showed our initial attempt at applying a compliant controller to insert the swab by using a torque control law. The experiments in Section 6.3 showed that force input is necessary to correct for misalignments on a stationary phantom, and that a system unguided by force feedback has much higher failure rates. However, continuing to operate the arm in torque control mode is undesirable for several reasons. As previously mentioned in Section 3.1, the networking latency and the fact that torque control operates as a second order differential system are two factors that reduce the responsiveness of following trajectories. Model inaccuracies and uncompensated friction also hinder performance. Ultimately, higher precision is needed for a moving head than what is feasible in torque control mode. As a result, we switched to a velocity control law to execute the contact phase.

While a torque compliant control law can be implemented by projecting the scaled forces measured on the end-effector to torque on the joints, using forces in a velocity control law requires a different strategy. As force acts as a second-order quantity in terms of the dynamics, a velocity controlled system needs to integrate force measurements to apply them as a first order term. Like in Section 6.2.5, we apply a low pass filter to integrate the raw force measurements, $\mathbf{F}$, to the lower order quantity $\mathbf{f}$:

$$\dot{\mathbf{f}} = -\alpha\mathbf{f} + \alpha\mathbf{F}, \tag{7.6}$$

with α as the decay rate of the exponential moving average filter. We reduced α to 0.5 to adjust for its usage in velocity control. Then we can add this term into our control law as

$$\dot{\mathbf{q}} = \mathbf{J}^{\dagger}(\mathbf{K}(\bar{\mathbf{e}} - \mathbf{e}) - \mathbf{W}\mathbf{\Lambda}(\bar{\mathbf{f}} - \mathbf{f})), \tag{7.7}$$

where $\mathbf{e}$ is the current end-effector pose, $\bar{\mathbf{e}}$ is the target pose, $\mathbf{J} \in \mathbb{R}^{6\times 7}$ is the robot Jacobian, and $\mathbf{K} = \mathrm{diag}(5, 5, 5, 2, 2, 2)\ \mathrm{s}^{-1}$ is the diagonal Cartesian pose gain matrix. The force gain

matrix is $\mathbf{\Lambda} \in \mathbb{R}^{3\times3}$, while

$$\mathbf{W} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

acts as a projection matrix mapping the 3D force measurement to the 6D velocity axes. Notice that the fourth and fifth rows invoke some rotational motion to account for the displacement between the swab and the end-effector. The target force is $\overline{\mathbf{f}}$, which is held to $\overline{\mathbf{f}} = [0, 0, 0]^\top$ N for stages 1 and 3, and $\overline{\mathbf{f}} = [0, 0, 0.4]^\top$ N for stage 2, when the swab is held against the nasopharynx. Given the choice of $\mathbf{K}$ and α, the remaining question is how to choose an appropriate force gain $\mathbf{\Lambda}$. Ideally $\mathbf{\Lambda}$ should be chosen so that it is high enough that is responsive to input forces but low enough so that chattering and instability do not occur. Rather than relying on trial-and-error, a more elegant way of gain tuning can be found by examining the mechanical properties of the swab.

Based on our model of the swab as an Euler-Bernoulli elastic cantilever beam (as described in Section 4.2.1), we expect a linear relationship between any transverse displacement of the swab and the magnitude of force applied along that transverse axis. The swabs used in these experiments are constructed from ABS plastic and so it is helpful to know if the shaft will remain in the elastic region for the expected range of motion and also to uncover the associated stiffness parameters. Hence, we conducted an experiment where the arm held the swab vertically in a fixture, as shown in Fig. 7.9, and the arm periodically moved in a transverse direction while measuring the forces. Fig. 7.10 shows the relationship between displacement and measured force. There appears to be a linear relationship at multiple points along the shaft within the expected range of motion for the swab, with the average R^2 correlation coefficients being 0.983.

If we examine just a transverse axis, this linear elastic relationship means there will be a restoring force in the opposite direction that could be measured at the fixed end

$$F = -\nu\Delta x, \tag{7.8}$$

where Δx is the transverse deformation and ν is the stiffness variable that will increase with loads applied further away from the fixed end. Suppose we simplify the force and displacement to a single axis, and assume that the setpoint and equilibrium point of contact are set to 0. If we were to assume that the external body remains still, we can substitute (7.8) into (7.6), which in the context of the control law (7.7) yields a linear differential

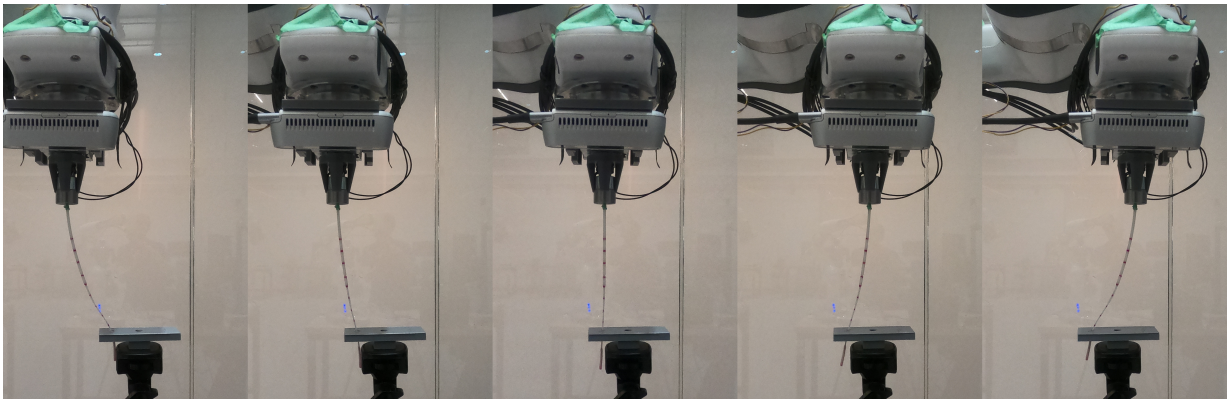


Figure 7.9: Swab held in fixture as end-effector moves in a transverse direction in order to evaluate if the swab exhibits linear elastic properties.

equation

$$\begin{aligned} \dot{x} &= k(-x) + \lambda f \\ \dot{f} &= -\alpha f + \alpha \nu x \end{aligned} \tag{7.9}$$

that can be represented as a homogeneous system:

$$\begin{aligned} &\dot{\mathbf{z}} = \mathbf{A}\mathbf{z}, \\ &\mathbf{z} = \begin{bmatrix} x \\ f \end{bmatrix}, \mathbf{A} = \begin{bmatrix} -k & \lambda \\ \alpha\nu & -\alpha \end{bmatrix}. \end{aligned} \tag{7.10}$$

Thus, if scalar k and α are set and ν is known, then the scalar gain λ can be chosen to determine the response characteristics of the system. Theoretically, choosing λ to critically damp the system, where the eigenvalues of $\mathbf{A}$ have no imaginary components [8], should ensure that oscillations do not occur for the steady state response. The λ that causes critical damping is

$$\lambda_{crit} = \frac{(\alpha + k)^2/4 - k\alpha}{\nu\alpha}, \tag{7.11}$$

which is found by manipulating the characteristic polynomial of $\mathbf{A}$. Although some liberty was taken in the assumptions in this model (non-moving contact point, no interaction between axes or rotational components), we expect λ_{crit} to be stable and a good baseline value to start from for the transverse axes. We will also explore gains higher than this, in case the critically damped gain responds too slowly to disturbances or the response is too low to pull the swab out from the contact.

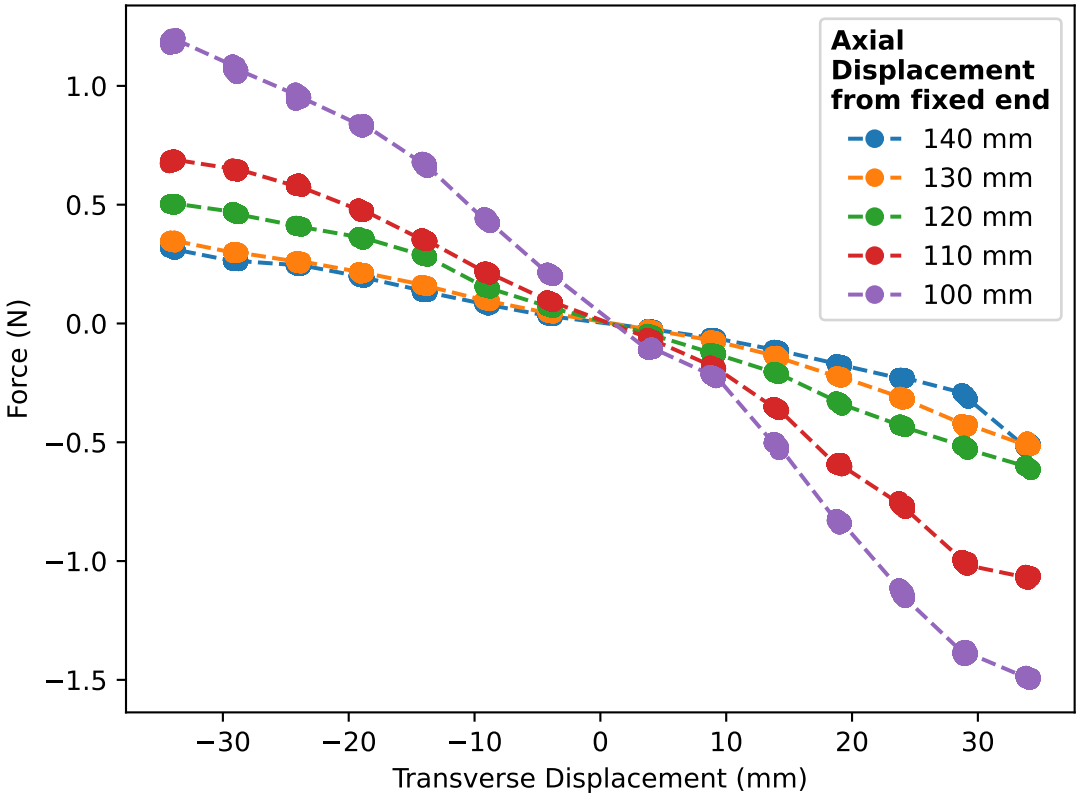


Figure 7.10: Graph showing the relationship between displacement and force for different holding points within the fixture. For the range of motion we expect on the swab during the NP test, the relationship appears linear.

According to Euler-Bernoulli beam theory [12], the stiffness, ν, will scale by the inverse cube of the length along the beam's axis, L, from the fixed end. While determining the material properties on this scale is not trivial, we can take measurements of stiffness at different grasping points along the shaft using the fixture shown in Fig. 7.9, and then fit an affine cubic function to compute the inverse stiffness ν^{-1}

$$\nu^{-1} = mL^3 + b. \tag{7.12}$$

To estimate m and b, we placed the fixture to hold at points of L between 95 and 145 mm. Note that the affine offset is added to this model due to the change in shaft thickness of the swab after the breakpoint at $L = 80$ mm. For each L we measured the force-transverse displacement pairs at these points, using the linear relationship (7.8) to find ν. Least squares regression was used to find $m = 43.89$ and $b = -0.0193$ from the L, ν pairs.

Hence, we present two potential strategies for choosing suitable values of the transverse components of Λ. The first strategy is based around the hypothesis that the majority of the contact occurs about the tip of the swab, where $\nu = 10.5$ N/m. As such, we consider three sets of gains: the critically damped gain $\Lambda_{crit} = 0.880$ $\text{N}^{-1}\text{m/s}$ (7.11), as well as $1.5\Lambda_{crit}$ and $2.0\Lambda_{crit}$. These controller configurations will be identified as *S1.0*, *S1.5*, and *S2.0*. The second strategy follows the hypothesis that the contact point moves further away from the tip of the swab is inserted through the nasal cavity, and therefore will dynamically change Λ as the displacement increases and ν increases. Given the tip displacement as ϵ, ν is interpolated using the cubic function (7.12). To account for the imperfect contact, we also scale the displacement by a factor of $\zeta = 0.5$, so $L = L_{max} - \epsilon\zeta$, where $L_{max} = 146$ mm. Again, we consider three sets of gains with respect to the critical gain (7.11) corresponding to $\nu(L_{max} - \epsilon\zeta)$ multiplied by factors of 1.0, 1.5, and 2.0, which will be identified as *D1.0*, *D1.5*, *D2.0*. Figure 7.11 shows a typical example of what the dynamic gain would be during the different stages of the insertion process.

The axial forces come from two sources. The first source occurs when the tip makes direct contact with the nasal cavity. Although the elastic beam theory would suggest that the force response should also act like a spring [12], in practice we observed that the response more like an impulse upon axial contacts. Although, once a certain threshold is reached (about 1 N of axial force), the swab begins to buckle and the axial force saturates (see Fig. 7.12), but this is less relevant for the level of force the controller will operate around. The second source of axial force comes from frictional forces as the swab moves forward during stage 1. When there are transverse forces applied to the beam, Coulomb friction will transmit a proportional axial force in the opposite direction of motion. In the absence of a good model for these factors, we tuned the axial force based on the charac-

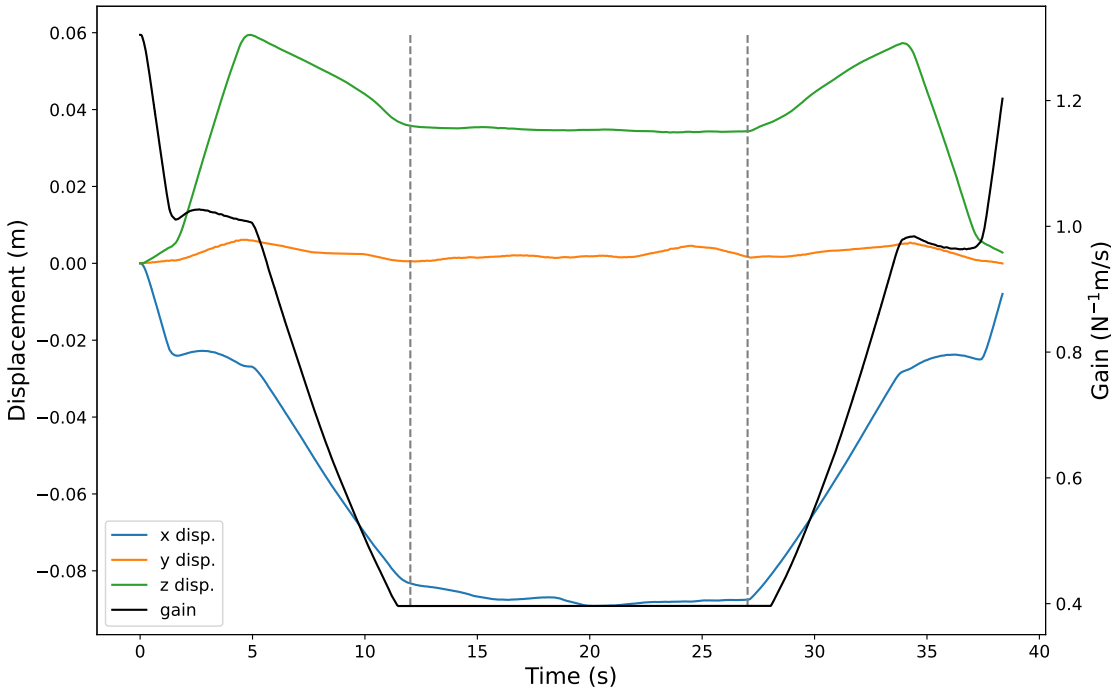


Figure 7.11: Transverse gain computed for *D1.0* for an example insertion. As the end-effector moves (mostly in the x-axis in the global frame), the gain scales down dynamically. The three stages of motion are divided by the vertical lines.

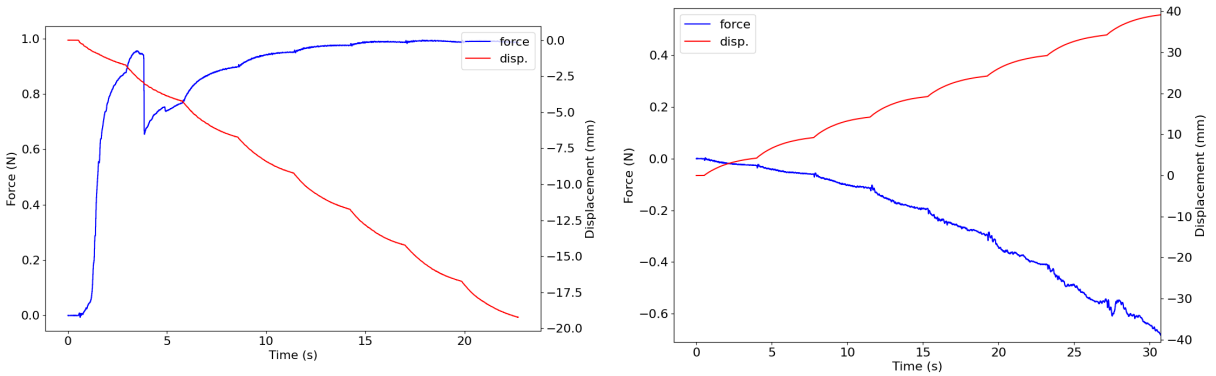


Figure 7.12: a) Attempt to recover axial force model of the swab. The measured forces behave like an impulse (around 1-2 seconds for the first mm of displacement) until the swab eventually buckles. b) For comparison, force-displacement relationship on a transverse axis, which exhibits a linear elastic relationship.

teristics of stage 2 by taking the maximum gain before we started observing oscillations. This value is 0.051 $N^{-1}m/s$.

7.8 Nasopharynx contact observer

A critical component to the proposed system is determining when the swab has reached the nasopharynx. Since it is not possible to visually observe this event with an external camera, we must make the decision with two sources of information: the displacement of the swab's tip and the total axial force returned. The otolaryngology literature estimates the distance between the nostril to the nasopharynx to be between 90 and 100 mm, although this can vary depending on factors like age and gender [80] [62] [105] [55]. Liu *et al.* [80] estimates the standard deviation of the length of the central portion of the nasal cavity at 11%, so we can use this information to determine a lower bound of displacement to reach the nasopharynx at about 80 mm. The measured axial force is another good indicator for when the swab reaches the nasopharynx because when the swab's tip makes contact with the perpendicular nasopharynx, we observe a larger impulse in force, as discussed in Section 7.7. Our study with healthcare workers in Section 3.2 found a median peak force value at 0.86 N among 6 practitioners, while the study by Park *et al.* [95] found a peak force of 0.60 N for a single expert practitioner.

A robust observer needs to take into account multiple modalities. Displacement measurements alone are insufficient because the head may move during insertion, and nasal cavity dimensions vary across the population. Similarly, force measurements can fluctuate due to motion. Therefore, we propose a fuzzy logic system, similar to our approach in Section 6.2.8, to combine both force and displacement measurements for decision-making. This system employs a fuzzy AND (product) operation on two sigmoid functions, one for displacement (p) and another for axial force (F_z)

$$\begin{aligned} \sigma_p = \sigma(s_p(p - i_p)), \sigma_F = \sigma(s_F(F_z - i_F)) \\ o = \sigma_p * \sigma_F \\ \sigma(x) = \frac{(\tanh(x/2) + 1)}{2}. \end{aligned} \tag{7.13}$$

The outcome o is unitless and bounded between 0 and 1. The decision rule for reaching the nasopharynx comes from thresholding o at 0.5. We use the equivalent hyperbolic tangent formulation of the sigmoid function as the function is implemented efficiently in the standard C++ math library. From this system we have four parameters that need to be chosen: the scaling factors s_p, s_F and intercepts i_p, i_F for the displacement and force sigmoids, respectively. The intercept term determines what value the input must be for the sigmoid to saturate at 0.5, which means that the inputs must be greater than the intercept in order for the decision to trigger. We tuned the intercepts at $i_p = 0.08$ m, and $i_F = 0.323$ N based on the insights listed in the previous paragraph. The scale parameters were placed at $s_p = 19.04$ m^{-1}, $s_F = 9.24$ N^{-1} also based on the aforementioned supporting studies, and some initial experimentation with the phantom. In contrast to our approach in Section 6.2.8 that chose parameters by trial and error, the information from the relevant literature inspired parameters that were much more accurate at detecting contact with the nasopharynx. The shape of the two sigmoids, o, and the resulting decision barrier is shown in Fig. 7.13.

7.9 Safety observer

Our system also includes an observer to detect if any of two distinct conditions occur during the contact phase. If either of these safety conditions trigger during the insertion process, the electromagnet depowers, which releases the swab, and the system immediately transitions to stage 3, where the arm then retreats away from the face and returns to its starting position. The first condition detects if the swab becomes wedged on the nasal vestibule or outside the nostril. This could occur if the head moved suddenly after the

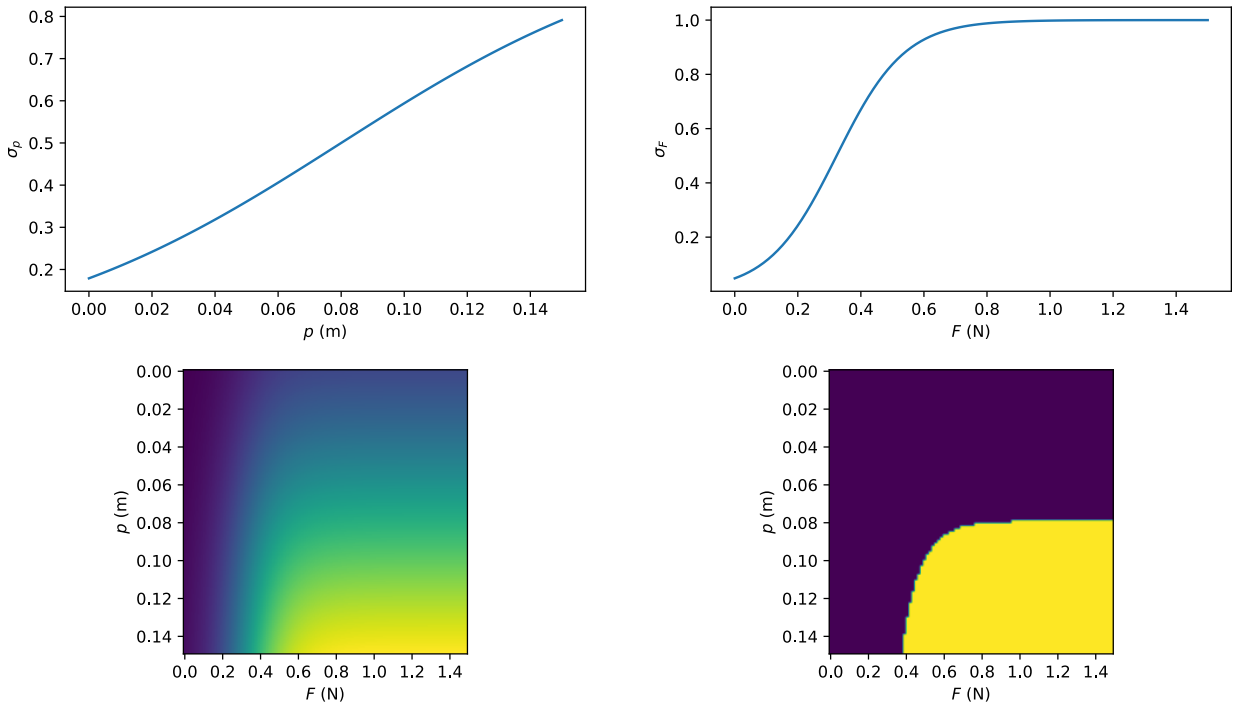


Figure 7.13: Fuzzy logic sigmoid functions. Top left and top right graphs are the individual position and force sigmoids, whose outputs are unitless and bounded between 0 and 1. The bottom left graph shows the product between the two sigmoids, while the bottom right graph shows the corresponding decision boundary by thresholding at 0.5.

visual servo component has completed. In this case, successful insertion is impossible, and it is better to abort the insertion and try again. This scenario is detected with a simple rule: if there is a high spike in force measured early during the insertion, specifically if there is a force measurement above 0.5 N during the first 40 mm of displacement. Fig. 7.14 shows a sequence where this component triggered and ended the trial. The second condition monitors whether the total measured force exceeds a maximum threshold of 2.5 N.

7.10 Experiments evaluating system on moving phantom

In our experiments there are a total of four categories of independent variables that were combinatorially varied to evaluate the controller over different conditions. These variables are:

1. Controller configurations: {*S1.0*, *S1.5*, *S2.0*, *D1.0*, *D1.5*, *D2.0*} (see Section 7.7).
2. Degree of head motion: {**Light**, **Medium**, **Heavy**} (see Section 7.4). One limited, non-repeated trial was also run with no movement with no head motion (**None**) for comparison purposes.
3. Phantom model: {A, B}(see Section 7.4) .
4. Nostril side: {Left, Right}.

In total, there were 564 insertion trials over these configurations. For each combination of the first three variables, 10 trials were done for the left side of the nostril, and 5 were done for the right side.

The experiment trials were conducted during February 2024 at the University of Waterloo RoboHub. The physical setup involved both of the Franka Emika arms, with the table holding the phantom wielding arm being placed into the workspace of the swab wielding arm. At the beginning of each trial, the swab arm would begin in its sentry position and then the entire integrated procedure would execute (i.e., the visual servo pipeline followed by the compliant control pipeline). The computers controlling the head robot and the swab robot were networked so that the latter could automatically trigger head motion generation at the end of the sentry stage (Section 5.3.1), after the swab has been segmented

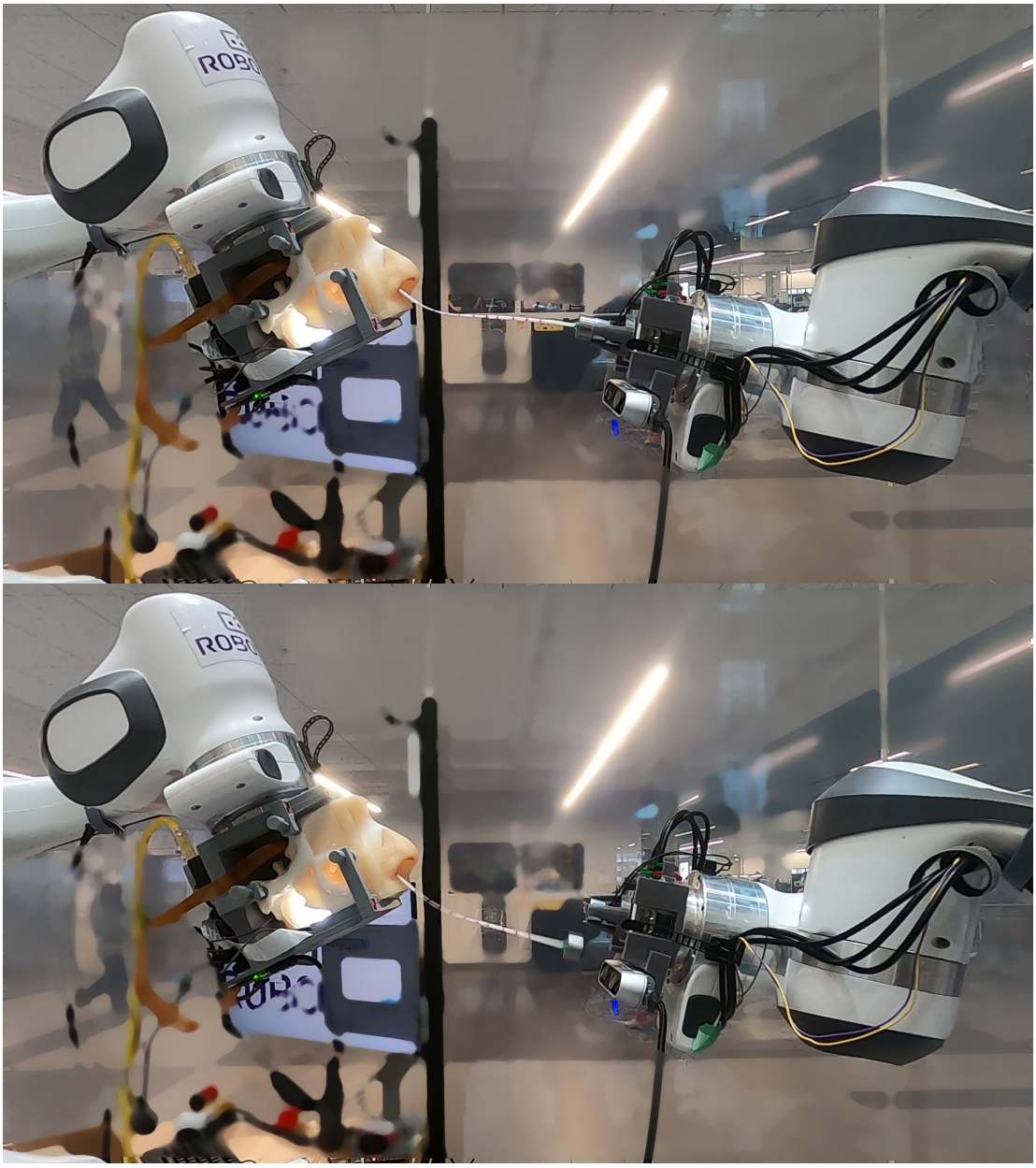


Figure 7.14: Example of the safety observer in use. Top: During the first few seconds of the contact stage, the swab tip becomes dislodged outside the nostril. Bottom: The observer detects a high impulse in force and aborts the procedure within 0.2 seconds by depowering the electromagnet and moving away.

and calibrated. Fig. 7.15 shows a collection of images taken during the trials. Fig. 7.16 compares a few different controller configurations that also alludes to the transverse gain has an impact on trial outcomes. The total displacement required to reach the nasopharynx for a non-moving head was 93.1 mm for the left nostril and 94.2 mm for the right nostril. In terms of the total duration of the test, the pre-contact phase required varying amounts of time for the visual servo to converge, depending on the level of head motion. The average times, after the sentry stage, were as follows: **None**: 11.3 s, **Light**: 11.9 s, **Medium**: 13.7 s, and **Heavy**: 18.6 s. The contact phase, on average, took a total of 39.5 s. Fig. 7.17 plots the forces and displacement of the swab tip and compares how these quantities fluctuate with different head motion configurations. The independent variables clearly have an impact on the results of the trial, which we will quantify via the metrics defined below.

7.10.1 Metrics

The evaluated metrics for the trials are largely divided into the two stages of insertion. For each of the metrics, we will be examining the statistical significance of the independent variables using four-way ANOVA testing for the continuous metrics and chi-square test for the discrete metric. Consequently, we report the F statistics for ANOVA, χ^2 statistic for the chi-square test, and both of their associated p-values. The critical p-value for these tests was adjusted with Bonferroni correction [37] based on the number of tests per metric: $0.05/4 = 0.0125$.

The metrics corresponding to stage 1 (insertion) are:

i) Transverse force ($\Downarrow$)

ii) Axial force ($\Downarrow$)

iii) Time to reach nasopharynx ($\Downarrow$)

iv) Whether the nasopharynx was reached ($\Uparrow$)

The $\Downarrow$ means that the lower values for that metric are better, while $\Uparrow$ indicates that higher values are better. The first two metrics are computed by taking the average forces applied in the transverse (x and y) and axial (z) directions. These values are interesting to monitor because they reflect the pressure that the swab is exerting within the nasal cavity. The third metric, the time to reach the nasopharynx, shows whether the controller slows down or becomes stuck during stage 1, which would be an uncomfortable situation for a human

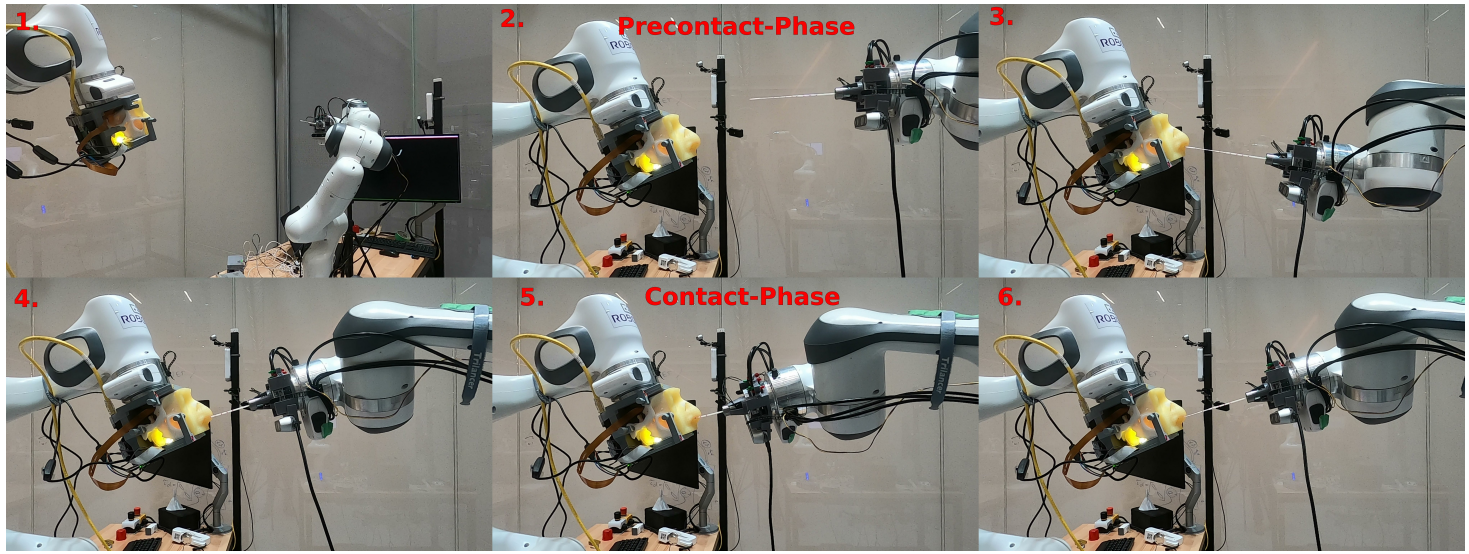


Figure 7.15: Illustration of the entirety of the integrated test. Frames 1-3: the pre-contact phase starts from the sentry mode and ends with the swab tip about to enter the nostril. Frames 4-6: the contact-phase first inserts the swab into the nasal cavity until it reaches the nasopharynx. After rotating to collect a sample, it exits the nasal cavity.

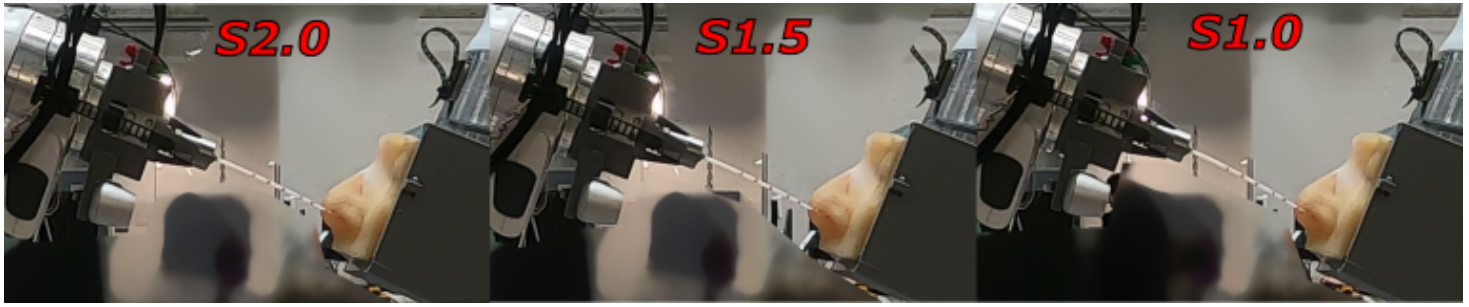


Figure 7.16: Frames taken during a trial with different controller configurations. Notice how the lower gain controllers have a larger offset angle compared to the higher gain ones.

patient. The fourth metric is obviously important because the ultimate objective is for the swab tip to reach the nasopharynx. If it does not, then the sample collection will be suboptimal.

The metrics corresponding to stage 2 (collection) are:

v) Transverse force (⇓)

vi) Axial force (⇓)

vii) Oscillation coefficient (⇓)

The fifth and sixth metrics are computed in the same way as the first and second, except they are averaged over the duration of stage 2. In configurations where the force gain was poorly tuned, we noticed that the forces measured on the transverse axes would oscillate during trials, which would cause the end-effector to visibly shake. We designed the oscillation coefficient as a way of measuring and comparing this phenomenon between different configurations. The first step of computing the oscillation coefficient is taking the discrete Fourier transform of the x-axis force over the duration of stage 2. We observed that the high oscillation trials could be characterized by larger spikes in the frequency domain, as shown in Fig. 7.18. As a result, our final metric is computed by summing the frequency bins between 0.5 Hz to 5 Hz, the band where the oscillations typically occupied.

In the remainder of this section we summarize how each of the metrics are impacted by the independent variables. Tables 7.2 and 7.3 show the results from the aforementioned statistical tests and the summary statistics of the metrics partitioned by each variable. Figures 7.20, 7.21, 7.22, 7.23, 7.24, and 7.25 are categorical plots that partition the results of the trials according to the independent variables, which are formatted according to footnote[1].

7.10.2 Stage 1: Metric i) Transverse force

For the stage 1 transverse force, we observe that there are significant differences among the control, head motion, and side variables, but the difference between the two phantoms is

[1]The 2 × 3 Violin plots show the distribution of the metrics and are partitioned according to the four groups of independent variables. The rows denote the side of the phantom and the columns denote the level of head motion. The x-axis denotes the controller type and the color of the plot denotes the phantom used. Metrics were not computed for trials that got stuck and did not reach stage 2; consequently, the number of data points used in each group is circled above the plot for reference when interpreting the plots.

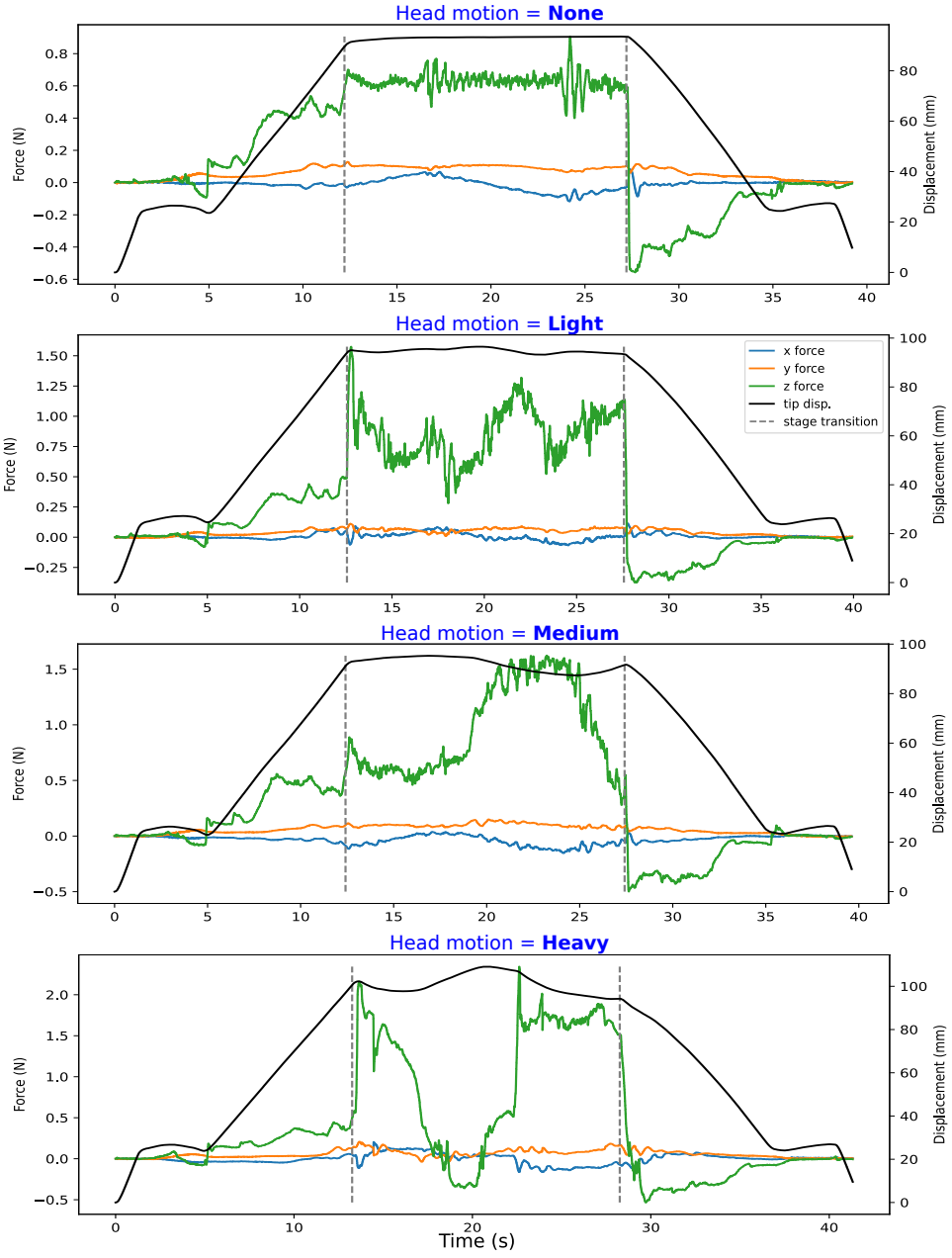


Figure 7.17: Graphs comparing the measured forces and the displacement of the tip for different levels of head motion for controller *D2.0* with Phantom B on the left side of the nostril. Generally higher amounts of head motion will result in the swab moving more to compensate for the forces of the rocking head. The three stages are divided according to the vertical dashed lines.

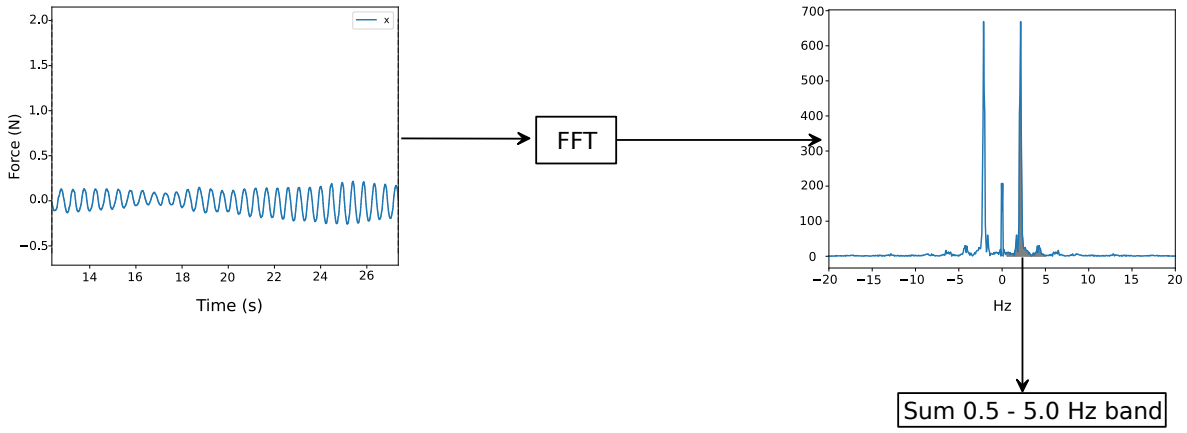


Figure 7.18: An illustration of how the oscillation coefficient is computed. The left image shows the x-axial force profile for a trial with high oscillation. In the frequency domain, this is characterized by a large spike around 2 Hz. The metric is computed by summing up the bins between the 0.5 to 5 Hz range.

insignificant, as summarized in Table 7.2 and Fig. 7.20. Controllers with lower transverse force gain have increased in average transverse force because the controller will favor the target waypoint position at the cost of higher force. It also appears that the spread of forces raises for the lower gain controllers as well. Heavier head motion seemed to lead to disturbances that resulted in higher transverse force as well. The forces from left side nasal cavity passage also appear to be slightly higher than the right side.

7.10.3 Stage 1: Metric ii) Axial force

Table 7.2 and Fig. 7.21 shows the axial force during stage 1. The control, head, and side variables were significant, while the phantom variable was not. The outcomes appear to be correlated with those to those for the transverse force. As the axial force gain was the same for all controllers, this appears to indicate that the axial force transmission from Coulomb friction is a visible effect during stage 1 as the swab moves forward through the nasal cavity.

Table 7.1: Breakdown of failures to reach stage 2 by controller configuration.

Controller	Stuck	Overload	Premature
S2.0	1	1	0
S1.5	2	0	1
S1.0	5	0	2
D2.0	1	0	3
D1.5	6	0	11
D1.0	19	0	34

7.10.4 Stage 1: Metric iii) Time to reach nasopharynx

Table 7.2 and Fig. 7.22 shows that the time to reach the nasopharynx, which was largely the same for all the controllers, except for *D1.0*. For *D1.0*, it appears that the axial forces were high enough to cause a significant slowdown via the mechanism in (7.5), although this effect appears to be quite variable according to the high standard deviation. It also that heavier head motions could also result in the increased time to reach the nasopharynx.

7.10.5 Stage 1: Metric iv) Reached nasopharynx

Table 7.2 shows the proportion of trials that reached the nasopharynx according to each variable. Statistically, only the controller configuration had a significant impact on the success rate. The lower gain controllers clearly struggled with making it to the nasopharynx properly. For the dynamic controllers, we see a steady increase in success rate from *D1.0* to *D2.0* as the gains also increase, and we can also see the success rate for *S1.0* being lower than the other static configuration. Despite not reaching the level of statistical significance, we do see a trend where lower head motion had better rates of success. Also, the left side had about an 8% lower success rate, but is still under the level of statistical significance.

It is important to analyze the specific events that could prevent the swab from reaching the nasopharynx. These obstacles can be categorized into three main types:

a) Controller stagnation due to axial force saturation

b) Procedure abortion before reaching the nasopharynx due to exceeding the maximum safety force limit

c) Premature triggering of the stage 2 observer before the swab actually reaches the nasopharynx

Table 7.2 presents a breakdown of these events for each controller. For the low-gain controller, the most frequent issue was the premature triggering of the stage 2 observer, primarily caused by excessive axial force. Additionally, several insertions were terminated early when it became evident that the controller was stuck and not progressing. Only one instance occurred where the force threshold was exceeded before reaching the nasopharynx, although this did happen seven other times after stage 2 had begun.

7.10.6 Stage 2: Metric v) Transverse force

As shown in Table 7.3 and Fig. 7.23, all the variables have a significant impact on the average transverse forces. As previously observed, the predominant factor appears to be the controller type, with a clear negative correlation between gain and transverse force. At stage 2, the dynamic gain controllers have all reached their minimum gains at the nasopharynx, which are below all the static gain controllers. We can observe a monotonic trend between the six controller in Fig. 7.19. In addition, we also see that heavier head motion moderately raises both the mean and variance of the transverse forces. The left side of the nasal cavity and Phantom B had somewhat greater transverse forces as well.

7.10.7 Stage 2: Metric vi) Axial force

From Table 7.3 and Fig. 7.24, it appears that the controllers have little impact on the mean value for the axial forces. This contrasts the result from stage 1 (Metrics i & ii), where the transverse and axial forces were correlated. This is likely because most or all of the transverse forces comes from the impulse-like force transmission from the nasopharynx moving against the tip of the swab. As the swab is not actively moving through the nasal cavity during this stage, Coloumb frictional forces would negligible The only variable that seemed to have an impact on the mean value is the side of the nasal cavity the swab traversed, with greater forces occurring on the left side.

However, there is appears to be a difference in variance in axial force caused by the head motion level. The **Light** head motion has a less spread (stdv. 0.165 N) than **Medium** (0.27 N) or **Heavy** (0.365 N) motions. So as the head motion gets more erratic, so does the force transmission as the head moves into or away from the swab with increasing intensity.

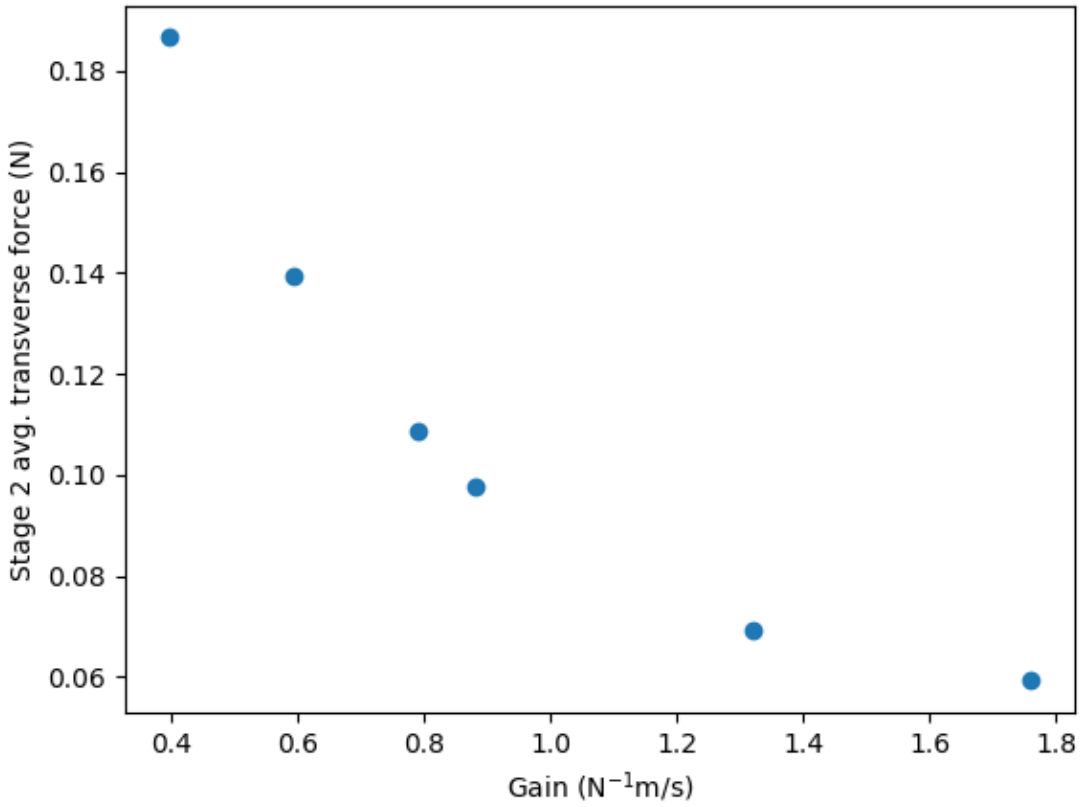


Figure 7.19: Average amounts of transverse force, with controllers ordered by their transverse gains at the nasopharynx.

Table 7.2: Statistics and summary of stage 1 metrics.

	Four way ANOVA						χ^2 **test**	
	i) Transverse force		**ii) Axial force**		**iii) Time to NP**		**iv) Reached NP**	
Variable	F	p	F	p	F	p	χ^2	p
Control	70.5	5.1×10^{-56} *	28.3	2.5×10^{-25} *	14.1	6.1×10^{-13} *	165	9.2×10^{-34} *
Head motion	20.9	9.1×10^{-13} *	14.2	6.7×10^{-9} *	6.99	1.3×10^{-4} *	8.09	4.4×10^{-2}
Side	11.9	6.1×10^{-4} *	35.3	5.3×10^{-9} *	2.57	1.1×10^{-1}	6.2	1.3×10^{-2}
Phantom	1.39	2.4×10^{-1}	0.0188	8.9×10^{-1}	2.46	1.2×10^{-1}	0.217	6.4×10^{-1}
Summary stats								
	i) Transverse force (N)		**ii) Axial force** (N)		**iii) Time to NP** (s)		**iv) Reached NP**	
Control	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.	Success rate	
S2.0	0.036	0.009	0.147	0.051	12.3	1.6	0.979	
S1.5	0.044	0.011	0.164	0.081	12.3	3.1	0.968	
S1.0	0.057	0.014	0.197	0.078	12.1	4.0	0.926	
D2.0	0.048	0.013	0.174	0.063	12.5	2.4	0.957	
D1.5	0.062	0.021	0.207	0.075	12.1	3.9	0.819	
D1.0	0.082	0.035	0.278	0.149	12.8	8.9	0.426	
Head motion	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.	Success rate	
None	0.047	0.012	0.163	0.046	11.7	2.5	0.875	
Light	0.046	0.016	0.162	0.047	11.6	3.0	0.906	
Medium	0.056	0.027	0.201	0.099	12.7	5.2	0.822	
Heavy	0.060	0.025	0.216	0.120	12.9	5.4	0.806	
Side	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.	Success rate	
Left	0.052	0.021	0.207	0.097	12.2	5.2	0.817	
Right	0.057	0.026	0.163	0.084	12.7	3.1	0.901	
Phantom	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.	Success rate	
A	0.053	0.025	0.191	0.108	12.6	5.0	0.855	
B	0.055	0.021	0.192	0.080	12.1	4.2	0.837	

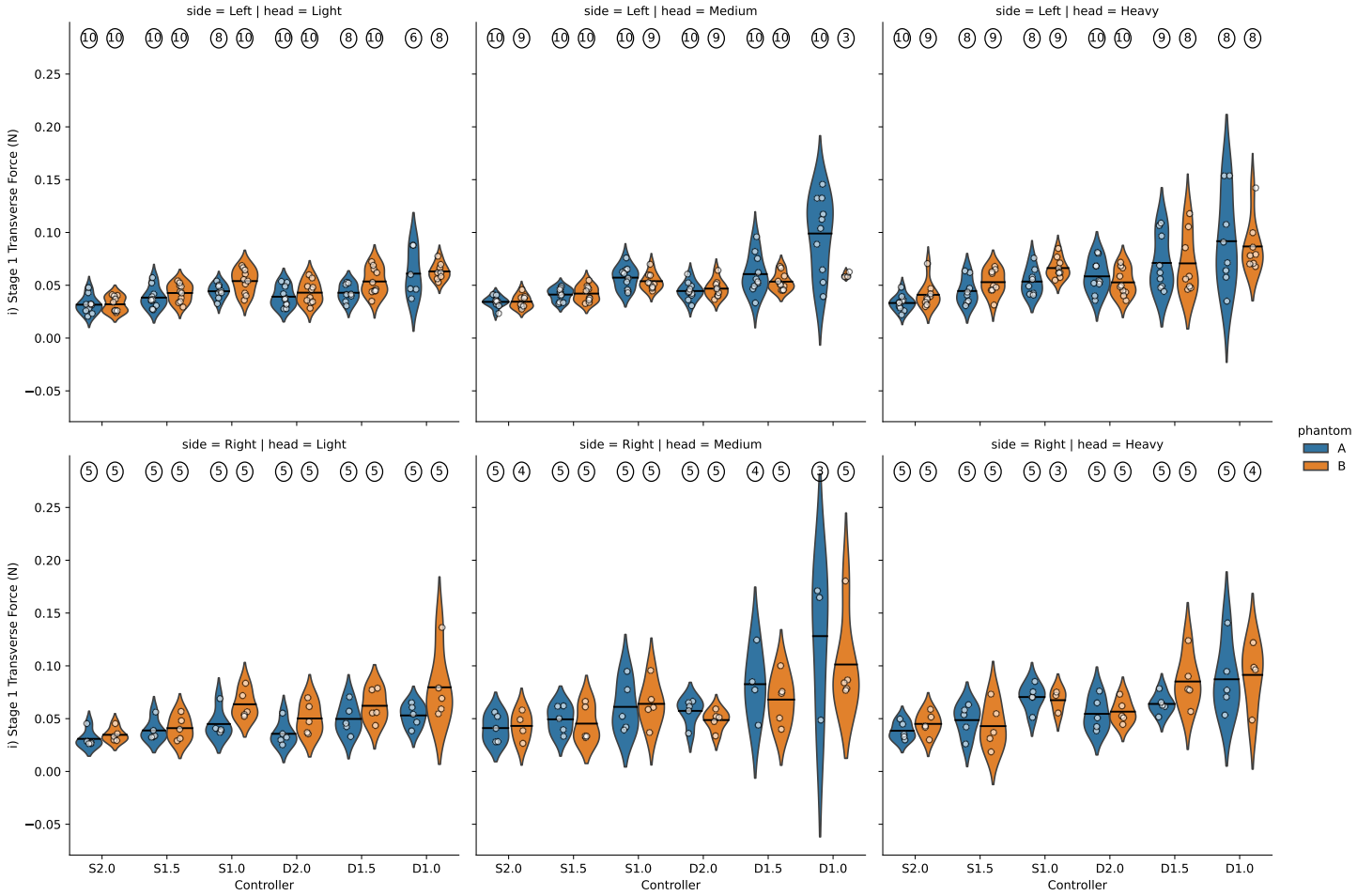


Figure 7.20: Metric i) Violin plots showing distribution of transverse forces during stage 1. See footnote [1] in Section 7.10.1 for formatting.

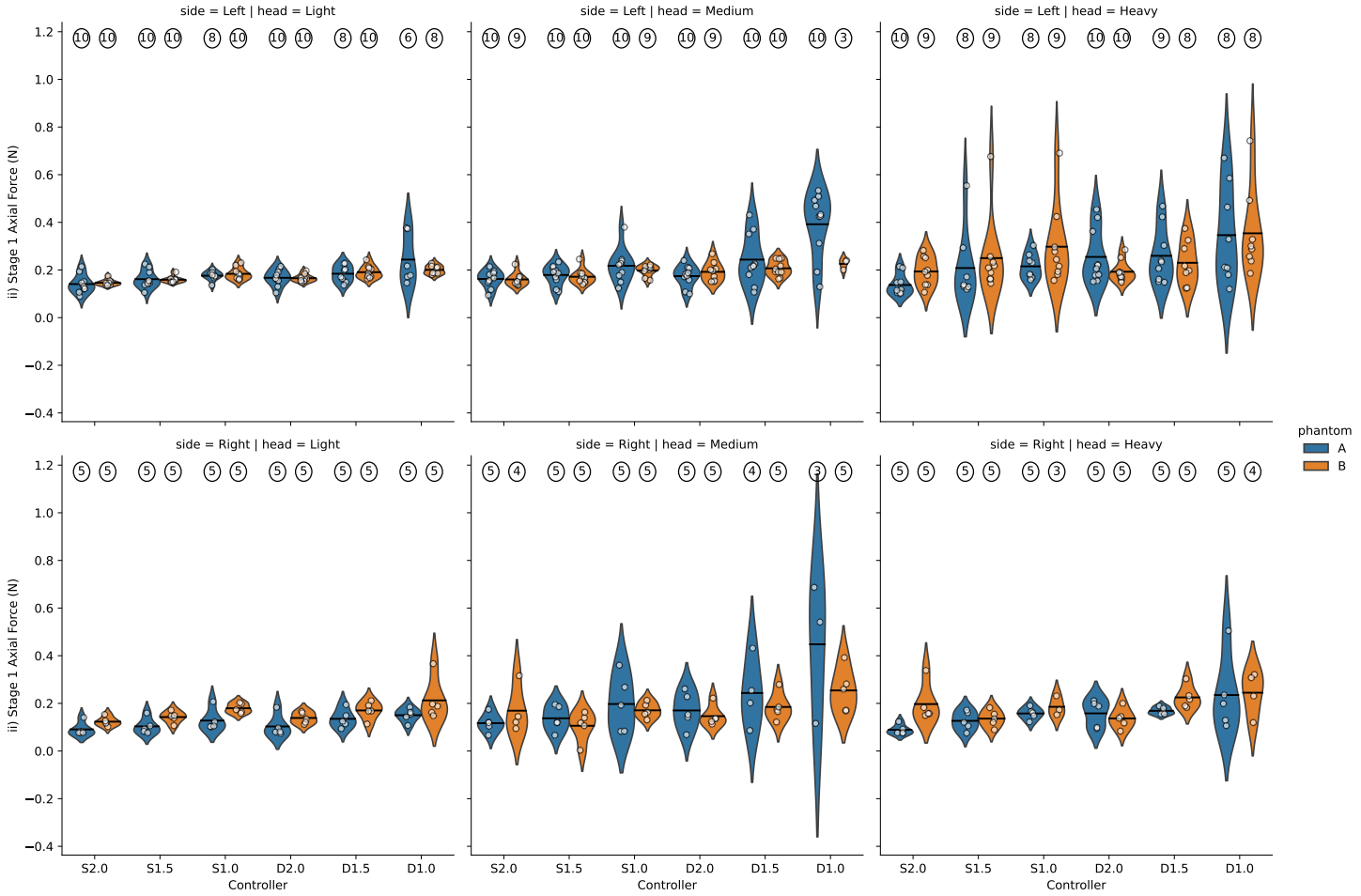


Figure 7.21: Metric ii) Violin plots showing distribution of axial forces during stage 1. See footnote [1] in Section 7.10.1 for formatting.

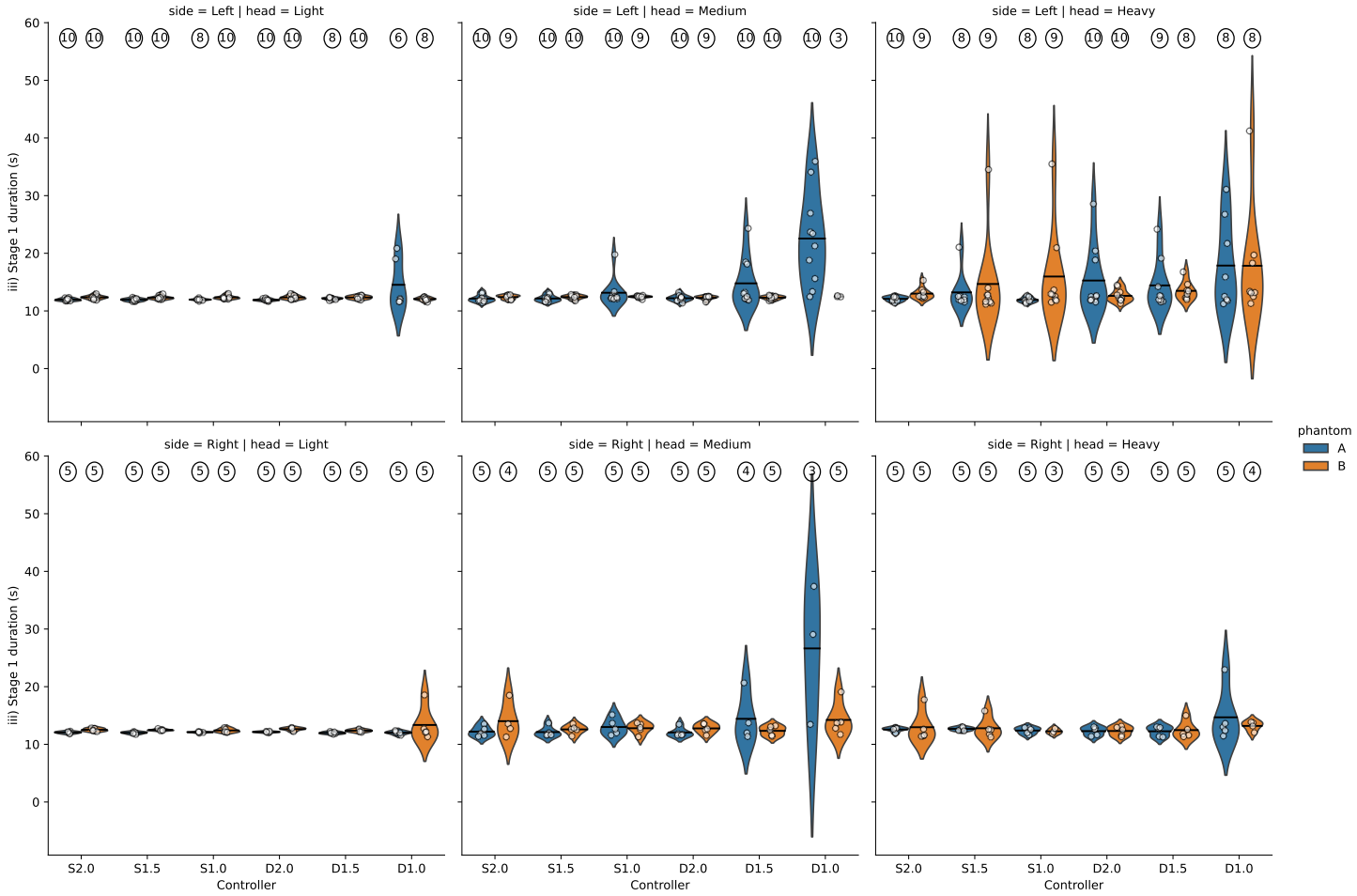


Figure 7.22: Metric iii) Violin plots times required to reach NP. See footnote [1] in Section 7.10.1 for formatting.

7.10.8 Stage 2: Metric vii) Oscillation coefficient

As shown in Table 7.3 and Fig. 7.25, all four categorical variables have an effect on the oscillation coefficient. First examining the controller variable, it appears that all the controller configurations had about the same oscillation, with the glaring exception of *S2.0*. In *S2.0*, the controller with the highest gain at the nasopharynx, there are instances where the oscillations dominated. For head motion, we see a steady trend where heavy head motion leads to heavy oscillation. Finally, we see that both the left side of the nostril and Phantom B tended to have higher oscillations than the right side and Phantom A respectively. Configurations with higher mean oscillation generally have higher variance as well, indicating that there is some *on-off* attribute towards oscillations getting triggered.

7.11 Discussion

We have done extensive testing of our proposed system to investigate the robustness of the proposed system as well as to examine the tradeoffs of the different controller configurations described in Section 7.7. Overall, we can see that higher transverse gain will lead to lower transverse forces in stages 1 and 2 and lower axial forces in stage 1. Having a high enough transverse gain is also important for being able to reach the nasopharynx without getting stuck or prematurely triggering stage 2 from high axial forces . However, we observed that there is a risk for instability if the gain is raised too high. Configuration *S2.0*, the controller with the highest gain the nasopharynx, had significant oscillations during stage 2 in many different trials. Therefore, the experimental results indicate that it is beneficial to have a high gain at the beginning of the insertion, so the controller can adjust to misalignment, motion, and avoid excessive forces that would slow the trajectory progression down. But once stage 2 begins, the gain should be reduced to avoid oscillations. This justifies the inspiration of our dynamic gain controllers, which can have high gain at the beginning and taper down once it reaches the nasopharynx. Dynamic controller *D2.0* had a good trade-off in terms of all the metrics, along with the static configuration *S1.5*, which had about equivalent performance.

In terms of the head motion simulation, there is a clear trend that heavier head motion causes greater disturbances to the control system. With heavier motion we saw increased transverse (stage 1 and 2) and axial forces (stage 1), increased durations for stage 1, and higher oscillation coefficients. For less well tuned controllers, we also saw this had the impact of raising oscillation. Comparing the three configurations head motion configurations, **Heavy** moved quite rapidly and strayed far from the resting point. The motion is

Table 7.3: Statistics and summary of stage 2 metrics.

	Four way ANOVA					
	v) Transverse force		**vi) Axial force**		**vii) Oscillation coef.**	
Variable	F	p	F	p	F	p
Control	202	1.9×10^{-118} *	0.296	9.1×10^{-1}	9.87	5.1×10^{-9} *
Head motion	21.6	3.4×10^{-13} *	1.28	2.8×10^{-1}	28.7	3.6×10^{-17} *
Side	20.4	7.8×10^{-6} *	21.1	5.5×10^{-6} *	38.6	1.1×10^{-9} *
Phantom	24.9	8.1×10^{-7} *	0.645	4.2×10^{-1}	98.1	2.8×10^{-21} *
Summary stats						
	v) Transverse force (N)		**vi) Axial force** (N)		**vii) Oscillation coef.**	
Control	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.
S2.0	0.059	0.022	0.674	0.260	1028	895
S1.5	0.069	0.019	0.693	0.249	768	707
S1.0	0.098	0.024	0.688	0.284	617	367
D2.0	0.109	0.030	0.687	0.260	689	470
D1.5	0.139	0.037	0.660	0.261	604	326
D1.0	0.186	0.055	0.699	0.329	643	404
Head motion	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.
None	0.100	0.040	0.639	0.096	555	669
Light	0.094	0.045	0.661	0.165	541	452
Medium	0.111	0.051	0.716	0.270	674	461
Heavy	0.120	0.060	0.679	0.365	1006	712
Side	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.
Left	0.103	0.047	0.723	0.276	825	673
Right	0.115	0.061	0.611	0.250	554	332
Phantom	Mean	Stdv.	Mean	Stdv.	Mean	Stdv.
A	0.101	0.049	0.693	0.264	523	276
B	0.114	0.056	0.673	0.281	939	734

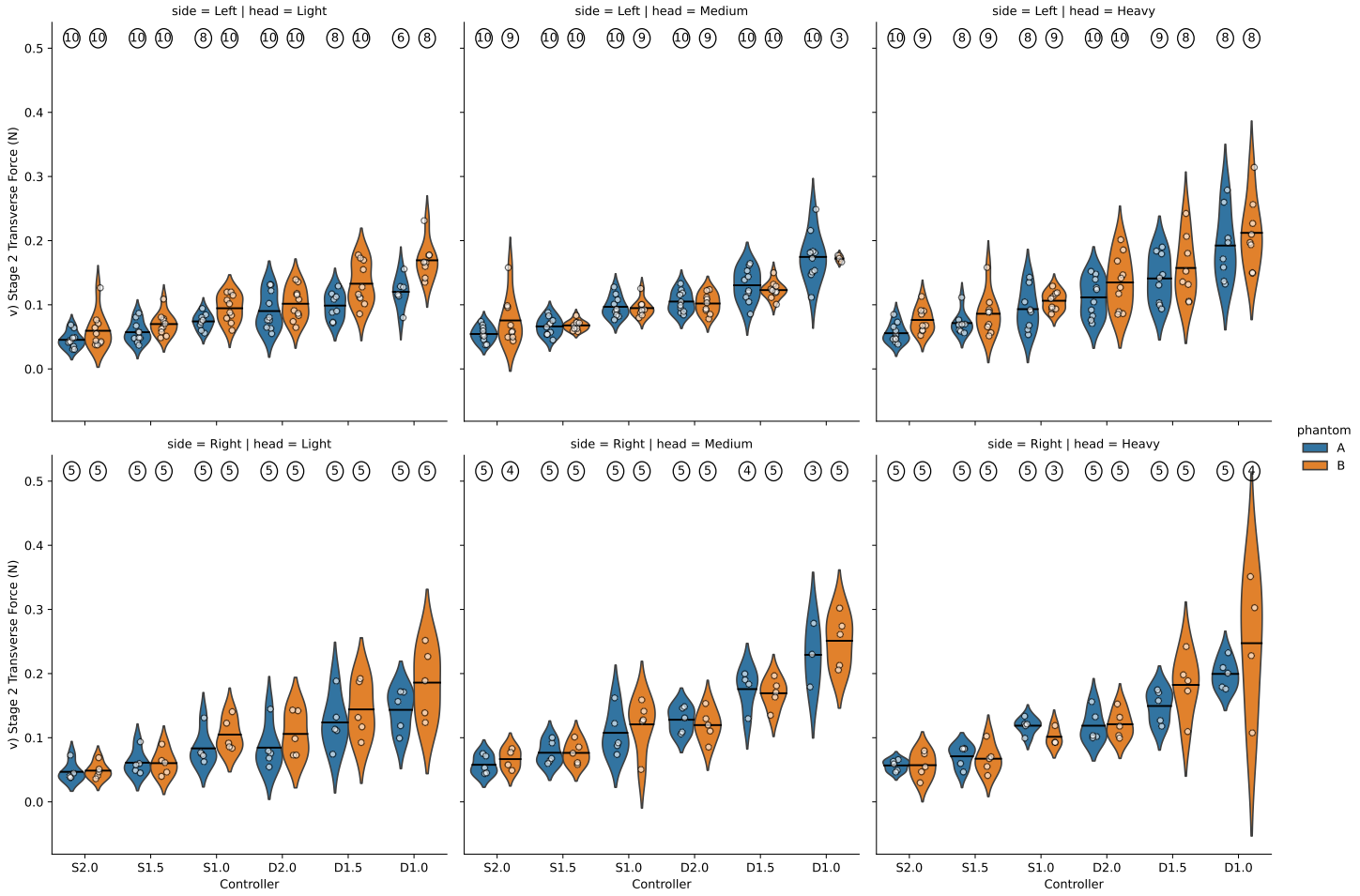


Figure 7.23: Metric v) Violin plots showing distribution of transverse forces during stage 2. See footnote [1] in Section 7.10.1 for formatting.

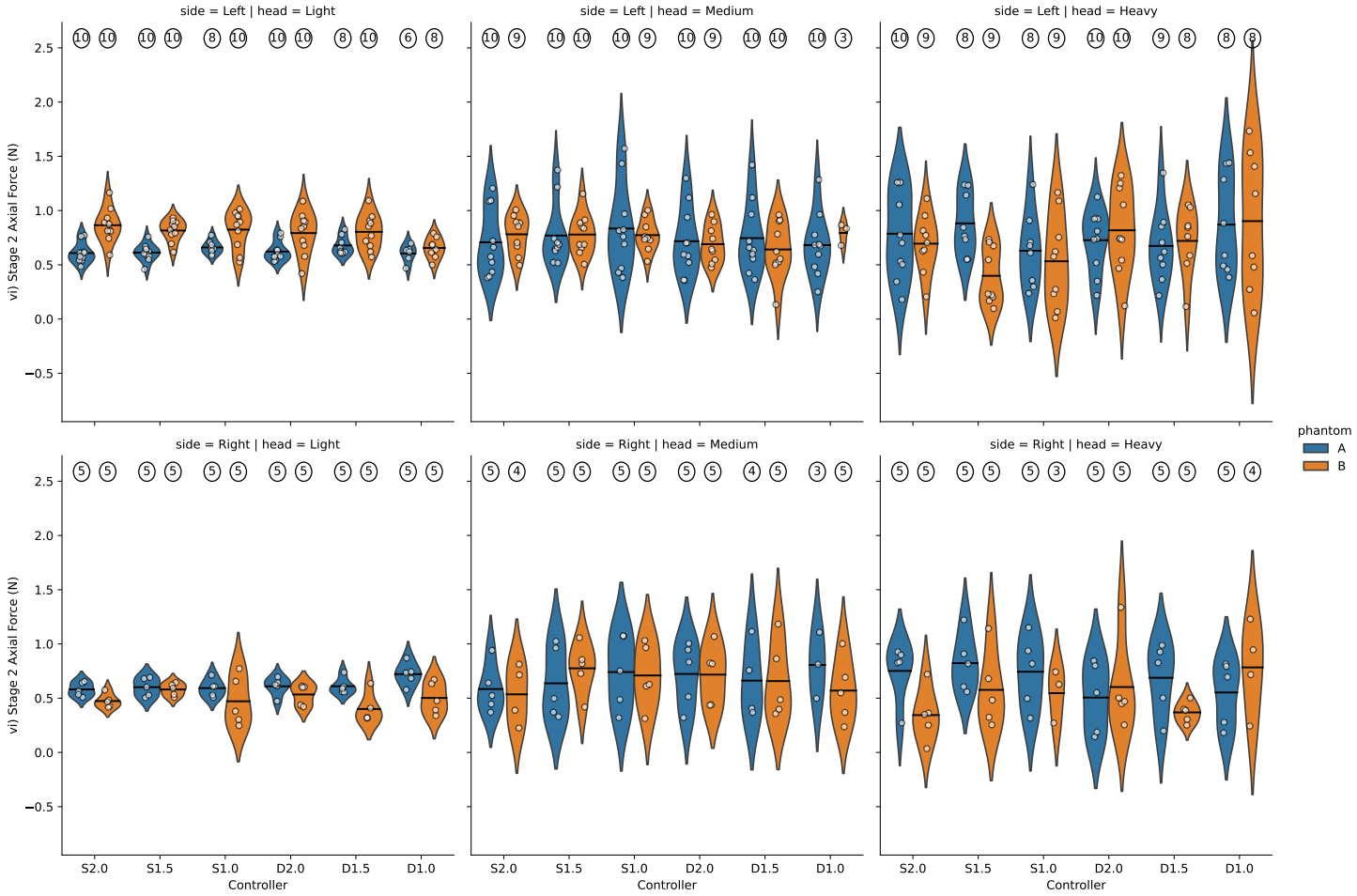


Figure 7.24: Metric vi) Violin plots showing distribution of axial forces during stage 2. See footnote [1] in Section 7.10.1 for formatting.

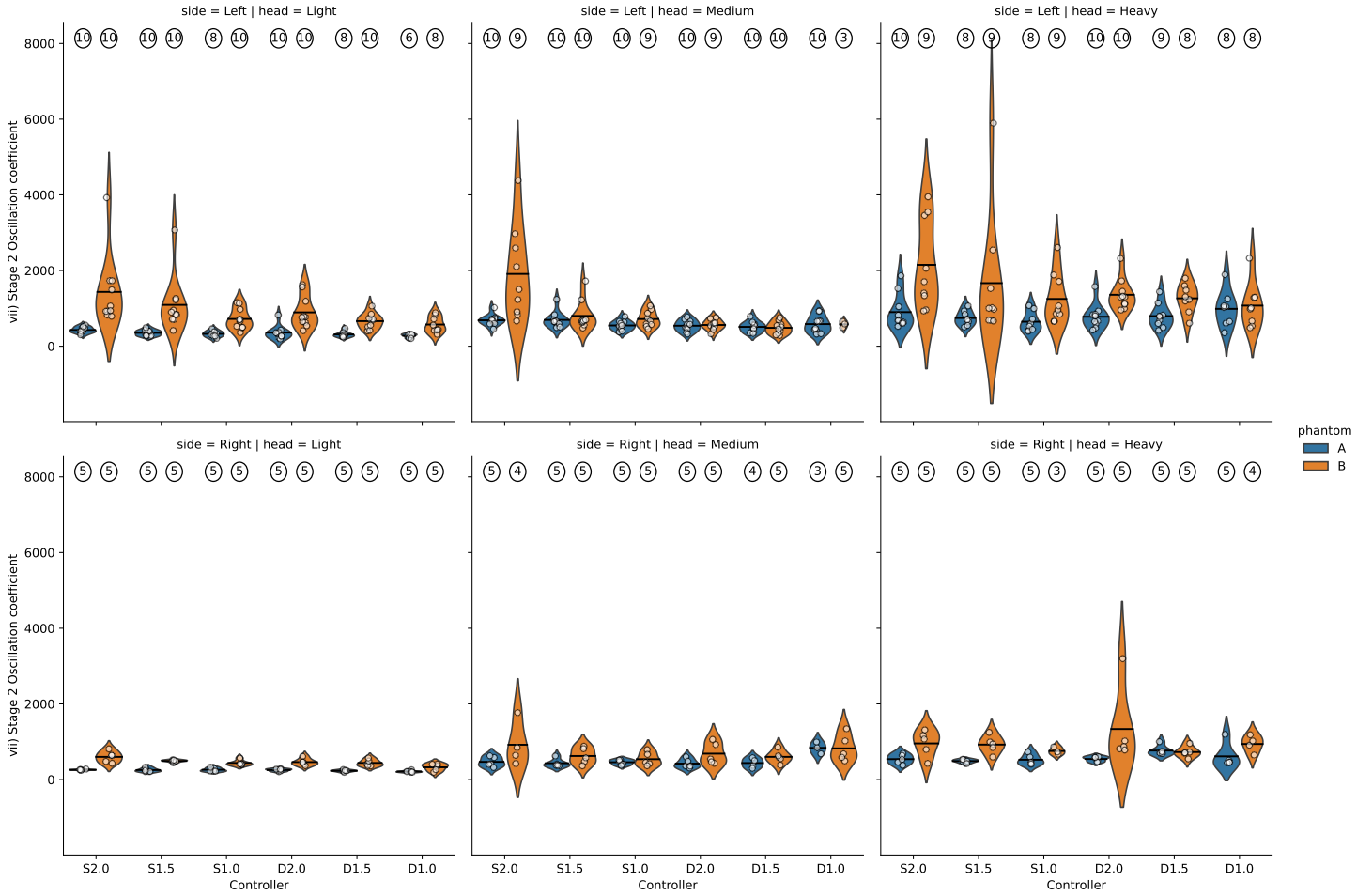


Figure 7.25: Metric vii) Violin plots showing distribution of oscillation coefficients during stage 2. See footnote [1] in Section 7.10.1 for formatting.

extreme enough that it is unlikely a cooperating patient would move in such a manner. The fact that the better tuned controllers were still able to handle these conditions builds confidence for its applications towards human subjects. The **Light** motion would probably be the more realistic scenario for a cooperating patient, with **Medium** motion being the upper bound movement we would expect.

In Chapter 5 we proposed the visual servo system for anthropometric faces and conducted human trials to evaluate the system. There, we observed that the final angle could vary with different faces due to the biases of the underlying CNN (3DDFA_V2). This raises the question, could our system be able to compensate for this variability? The repeatability for our replacement visual servo system presented some variance, although from visual inspection it appeared to be lower than the experiments from Section 5.4. The head motion was a larger source of placement variance that was comparable to the human trials, as the head would often move during the transition between the pre-contact and contact stages, resulting in larger alignment discrepancies after the turn stage of the trajectory. The force feedback was generally able to compensate for this misalignment, which leads us to believe the system compensate for the alignment errors observed during the human trials. Of course, the early abort system is present in cases where the alignment is too extreme to work.

In terms of the statistical hypothesis tests used, one may recall that ANOVA testing typically makes three assumptions [86]: a) independent observations, b) normally distributed outcomes, c) and homogeneity of variances. It is clear by examining the figures in Section 7.9 that several metrics (e.g., vii)) violate the third assumption. In most of the metrics, a raise in mean was usually accompanied by raise in variance as well. While statistically speaking, this would make it more difficult to reject the null hypothesis of unequal group means, in our scenario this conclusion is less relevant. We mainly care if factor causes a change in behavior for the controller, which is evident when a configuration causes more outliers and higher variance in a metric.

The abort observer from Section 7.9 was a useful component during the trials. While we would not expect the level of movement from a cooperating patient, during the **Heavy** head motion trials there were several unrecorded instances where the swab missed the nostril because the head moved rapidly before the swab could enter the nostril at the beginning of the trajectory. In these cases the swab would make contact with the "tissue" outside the nostril and release the swab within 0.2 seconds of collision, before any real damage could happen. The maximum force component was also useful for preventing the swab from becoming crumpled due to unpredictable heavy head motion during one of the trials as noted in Table 7.1.

7.11.1 Avenues for improvement

The best tuned controllers *D2.0* and *S1.5* were effective when subjected to the most challenging insertion conditions. However, there is still some aspects that could be improved. One potential improvement could be the incorporation of additional components to further attenuate oscillations. This could involve creating an oscillation observer via a short window Fourier transform, that could be used to detect if any oscillations are occurring around the 2 Hz range. When oscillations are triggered from head motions or other external factors, they could be attenuated by throttling the gain to lower levels. Altering the hardware to improve the latency between the force measurements and the commands/readings from the robot could also be beneficial. While the head simulator was a useful formulation for generating random head motion, the motions were mostly continuous. One type of motion that would be useful to simulate is the effects of rapid, impulsive motions like coughing or sneezing, that are quite probable to occur in patients with swabs entering their nose. It would be interesting to see how the current controller would react to these events, and if any additional modifications would be needed to compensate for them.

The total duration of the pre-contact (including the sentry phase) and contact phases averaged just below 70 seconds, which is comparable to other swab robots [134] [124] [30]. A human healthcare worker could likely perform the test much faster, and there is certainly potential to tune the trajectories for higher throughput. However, some caution is needed to ensure that the quality of the samples are not compromised by speeding up the procedures. In addition, patients may be more likely to panic if the robot is moving rapidly.

As previously discussed in Section 5.5.1, integrating visual states into the insertion stage could be beneficial for keeping the swab aligned. This would be especially helpful during the first few seconds of the contact phase, when the swab has barely entered the nostril and the head could easily move away without providing force feedback. Therefore, having a visual error term early on for the insertion could help this issue substantially. There could be some transient benefit for having visual feedback once the swab is inserted further into the nasal cavity, but there are technical limitations to consider. Once the camera gets too close to the face, the RGB-D camera will be unable to obtain depth measurements. This means that such a setup would have to rely on perspective within the color images or consider adding an external eye-to-hand camera. While an eye-to-hand camera could have some potential for estimating the state of the target [44], it would require a method of identifying the relative pose from the new perspective.

7.12 Conclusion

In this chapter, we demonstrated the integration of our visual servo and compliant control system to perform the critical stages of the NP swab test. The compliant control system from Chapter 6 was refined to a velocity control law that examines the mechanical properties of the swab to tune the gain parameters. Observers were also added to detect when the swab reached the nasopharynx, and to initiate abort procedures. We employed a second robotic arm to act as a stand-in for a patient, that wields a nasal cavity phantom and can generate with variable amounts of head motion. Extensive experiments were conducted to study the performance of the controller with a moving patient and to find the best performing configuration.

The proposed system performs well and was able to execute the test without any major mishaps. Thanks to the abort sequencing, the robot can quickly disengage from the patient if the swab does not enter the nostril, large forces are measured, or by manual abort. Some additional minor components could be considered to watch for oscillations and to possibly integrate visual states during the contact phase. But overall, the better tuned configurations appear promising for application in human trials, given their performance during the experiments, which we will discuss in the following chapter.

Chapter 8

Discussion & Conclusion

In this chapter we conclude the thesis by summarizing the accomplishments of the previous chapter. We then discuss future experiments to extend this line of research and ways our research could be applied to the scope of different clinical settings.

8.1 Summary of work

The research presented in this thesis has shown that it is feasible for a collaborative robot arm to carry out NP swab samplings in place of a human healthcare worker, under the scenario of having a freestanding patient approach the arm to have the test administered. To extend the abilities of a Franka Emika arm with appropriate force and vision sensing to make it capable of performing the NP swab test, Chapter 3 describes the design of our end-effector, along with the other supporting materials that were used to conduct this research.

The work in Chapter 4 demonstrated a simulation and optimization method to find strain minimized trajectories through the nasal cavity to reach the nasopharynx. The results from this chapter were useful for verifying the desired angle the swab should be inserted through the nasal cavity with.

Chapter 5 detailed the design of a visual servo system to align swab next to participant nostrils during the pre-contact stage. We were was able to leverage advances in deep learning and more classical methods to accomplish the control task with a single RGB-D camera located on the end-effector. The human trials showed that the system was

effective and highlighted alterations that would make the system more reliable in terms of positioning the swab in front of the nostril.

Chapter 6 showed our first foray into implementing a compliant controller to insert the swab during the contact phase. The results showed that force feedback is a necessary component to correct for minor misalignments.

Finally, Chapter 7 built on the insights from Chapters 4, 5, and 6 to create a system that can seamlessly perform both the pre-contact and contact stages of the NP swab test. We made improvements to our previous works by finding a reformulated optimized insertion trajectory, adapting and tuning the visual servo system, and changing to a velocity compliant controller to improve performance during the contact phase. We created a physical head motion simulator using a secondary arm equipped with nasal cavity phantom to validate how well the system can respond to perturbations induced by natural head motion. Our experiments evaluated different combinations of controller configurations and levels of head motion and quantified the best controllers using a number of different metrics. The best performing controllers were able to perform the test reliably, and the results show great promise for potential experiments with human participants.

8.2 Future experiments

The next steps to extend this research would be focused towards conducting human trials to include the contact phase. Of course, a major milestone to achieving this would be obtaining the institutional ethics board clearance for the contact phase. One of the primary concerns of the ethics board would likely be surrounding the mechanical safety of the system with respect to participants. There is reasonable evidence to show the proposed method is safe enough in the mechanical sense to execute the contact phase. Others have been able to perform human trials using similar collaborative arm setups [47] [30] [142]. Our experiments showed that the well tuned systems in Chapter 7 could effectively handle motions that would be unrealistically erratic for a cooperating participant. Finally, if there were a mishap, the hardware system described in Chapter 3 enables the procedure to quickly abort by releasing the swab due from a user input emergency stop or the safety observer triggers.

However, further measures are necessary to ensure the system's safety regarding infectious and electrical hazards. Given that the contact phase of the test involves handling biological materials, precautions must be implemented to prevent contamination of surfaces or individuals during trials. Additionally, a comprehensive analysis of the system's

electrical safety is required. While the robot itself is rated for direct human contact and should be inherently safe, the primary point of interaction with participants would be the swab on the end-effector. The risk of electrical danger is minimal in this aspect, as the ABS plastic composing the swabs is an insulator. Also, when the electromagnet is powered the measured potential between the steel swab holder and ground is only 3 mV. Even in the event of an electromagnet short, there would be no direct path for current to flow into the patient. However, as mucus could potentially conduct current, a protective barrier should be installed around the electromagnet as a precautionary measure. Furthermore, the supporting circuitry illustrated in Fig. 3.6 should be positioned out of patients' reach to eliminate any potential risks.

The linear velocity controller worked quite well during the contact phase, but it would be interesting to explore whether more advanced control paradigms could enhance performance. One limitation of the current approach is that the baseline trajectory does not change, even when consistent force readings indicate misalignment. More advanced control techniques, such as model predictive control [113] or reinforcement learning [92] could be used to dynamically adjust the trajectory. Additionally, these techniques could be integrated with the simulation environment described in Chapter 4 to develop insertion methods that adapt to various scenarios, such as different alignments, nasal cavity morphologies, or situations where the swab becomes stuck.

Given that our experiments on the contact phase were restricted to phantoms, one aspect that was not investigated are biological assays to verify the quality of sample collection. While the proposed system in Chapter 7 mimicked the recommended motions [98] [50] and applies a suitable amount of force at the nasopharynx, the suitability of the biological material would still need to be evaluated for human trials. This evaluation would require collaboration with a biological laboratory capable of conducting polymerase chain reaction tests and assessing cycle thresholds, as done in other studies [30] [142].

8.3 Future work towards clinical applications

The platform used in this thesis was designed under the assumptions of a prototype that is used under controlled laboratory settings. However, a clinical setting would add additional challenges, which begs the question: what additional work would be needed to make this system ready for a clinical setting? Additional software development would be needed to implement an accessible user interface for non-technical users to interact with the programs. This would include the necessary backend work to ensure that patient data is appropriately cataloged. For the successful adoption of the platform, it is crucial that

patients feel comfortable receiving care from robotics. This may present a challenge, particularly among demographics that are less familiar with technology. Proper communication with the patients and the implementation of techniques from social robotics [19] [101] could be instrumental in addressing these psychological factors. In addition, the steps for the post-contact phase of the test would need to be implemented, which would involve placing the swabs into storage containers. Sanitization of the robot, expanding the capabilities of the platform with a mobile base, and adding the ability to perform other close-contact tasks are other major topics that we will elaborate below. In a clinical environment, it is expected that the dimensions of the swabs could vary outside those identified in Fig. 3.7 due to procurement issues. To compensate for this variability, the platform would need to adapt the presupposed parameters of the swab utilized in Section 5.3.3 and Section 7.7 according to the available swabs. As well, additional steps should be taken to screen for deviated septums, allergies, past surgeries, and other rhinological conditions that could contraindicate the NP swab test in patients [16]. While out of scope for this thesis, if our system were to be used in clinical settings, several additional requirements would need to be met. These include obtaining certification as a medical device from various international and national regulatory bodies, depending on the region of application. Additionally, considerations for liability and insurance would need to be addressed.

The issue of equipment sanitation is obviously a pressing concern that would need to be resolved. The point closest to the patient, other than the disposable swab, is the stainless steel swab holder. The safest, most efficient way to sanitize the piece would be using an autoclave. The electromagnet, loadcell, and robot are all electrically sensitive, so sanitization with autoclaving or chemical solutions would not be recommended. One could take an approach of completely isolating the robot from the patient with protective sheets like in Haddadin *et al.* [47], but this would increase maintenance requirements and impede performance. Another solution could be applying ultraviolet (UV) light, although some care will be needed to ensure the materials do not degrade over time with exposure, and it would likely need to be combined with additional disinfectant measures to be effective [7]. In any case, some revisions should be made to the casing of the end-effector to make the load cell and electromagnet less exposed.

Another enhancement would be expanding the robotic platform to have a mobile, holonomic base. This would greatly improve the utility of the platform by expanding the workspace to a whole room or multiple rooms instead the stationary cone. The base would also provide the ability to seek out and approach patients. The ability to raise and lower the platform would be useful for meeting different patient heights. There are a couple ways the mobile base could be implemented. One option could be mounting the collaborative arm on a separate platform (e.g., a warehouse platform). A possibly more robust method could

be using an integrated wheeled humanoid collaborative platform like the PAL TIAGO Pro (Barcelona, Spain). The main requirements for the platform would be having an on-board computing solution capable of running deep learning models, an exchangeable flange to hold the end-effector, and an appropriate power supply for the electromagnet. Assuming the arm has enough clearance, the modifying and reprinting the interface for the flange should be straightforward. The electromagnet may require some more involved engineering efforts. If the platform has an on-board power-supply, it would need to be adapted to power 24V DC. Otherwise, adding a separate battery powered solution may be needed.

The overall motivation of this research is to enable robotics to conduct close-contact tasks with patients to protect human healthcare workers and improve efficiency. Therefore, the utility of the robotic platform in a clinical setting would be greatly enhanced by enabling additional types of close-contact tasks. As previously mentioned, ultrasound scans [115], temperature tests, or blood work could be of particular interest. These tasks would probably expect similar levels of force to the NP swab test, so adaptations to the electromagnetic grasping component could be an easy way to support these and switch between different tasks by having separate ferromagnetic adapters for each tool. One consideration is the contact dynamics of the utilized tool. The swab had elastic beam mechanical properties, so the transverse stiffness was quite low and enabled us to tune our controllers quite well. However, other tools may have higher stiffness properties, which will make force-displacement responses more impulse-like. This may require upgrading the analog to digital circuitry, for example to an AD7730 (Analog Devices, USA), to measure forces at a higher frequency so the system can respond to these quicker force responses more effectively. While task-specific adjustments to motion trajectories and visual models would be necessary, we anticipate that the underlying compliant and visual control loops developed in this research could be generalized to a wide range of close-contact tasks.

References

[1] E. Ackerman. *How Diligent's Robots Are Making a Difference in Texas Hospitals*. 2020. URL: https://spectrum.ieee.org/how-diligents-robots-are-making-a-difference-in-texas-hospitals (visited on 2021-11-05).

[2] P. Agarwal, S. A. Moubayed, A. Alspach, J. Kim, E. J. Carter, J. F. Lehman, and K. Yamane. "Imitating human movement with teleoperated robotic head". In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 2016. DOI: 10.1109/ROMAN.2016.7745184.

[3] item America. *Automated COVID Testing Thanks to a Tireless Robot*. 2021. URL: https://item24us.news/automated-covid-testing-thanks-to-a-tireless-robot/ (visited on 2023-07-09).

[4] C. An, C. Atkeson, J. Griffiths, and J. Hollerbach. "Experimental evaluation of feedforward and computed torque control". In: *Proceedings. 1987 IEEE International Conference on Robotics and Automation*. Vol. 4. 1987. DOI: 10.1109/ROBOT.1987.1088020.

[5] M. Antonelli, A. P. del Pobil, and M. Rucci. "Bayesian multimodal integration in a robot replicating human head and eye movements". In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*. 2014. DOI: 10.1109/ICRA.2014.6907271.

[6] A. S. Arockia Doss, P. K. Lingampally, G. M. T. Nguyen, and D. Schilberg. "A comprehensive review of wearable assistive robotic devices used for head and neck rehabilitation". In: *Results in Engineering* 19 (2023), p. 101306.

[7] F. Astrid, Z. Beata, V. den Nest Miriam, E. Julia, P. Elisabeth, and D.-E. Magda. "The use of a UV-C disinfection robot in the routine cleaning process: a field study in an Academic hospital". In: *Antimicrobial Resistance & Infection Control* 10.1 (May 2021), p. 84.

[8] K. J. Åström and R. M. Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press Princeton, NJ, 2008.

[9] T. Baltrusaitis, P. Robinson, and L.-P. Morency. "Constrained Local Neural Fields for Robust Facial Landmark Detection in the Wild". In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*. June 2013.

[10] Barrett Technology. *Barrett Upper-extremity Robotic Trainer*. 2018. URL: https://web.barrett.com/support/Burt%5C_Documentation/BURT-Research-Edition-Brochure.pdf;%20Retrieved%20June%202,%202024 (visited on 2024-06-01).

[11] M. Bastir, I. Megía, N. Torres-Tamayo, D. García-Martínez, F. M. Piqueras, and M. Burgos. "Three-dimensional analysis of sexual dimorphism in the soft tissue morphology of the upper airways in a human population". In: *American Journal of Physical Anthropology* 171.1 (2020), pp. 65–75.

[12] O. A. Bauchau and J. I. Craig. "Three-dimensional beam theory". In: *Structural Analysis*. Dordrecht: Springer Netherlands, 2009, pp. 223–259. DOI: 10.1007/978-90-481-2516-6_6.

[13] J. van den Berg (https://math.stackexchange.com/users/91768/jur-van-den-berg). *Calculate Rotation Matrix to align Vector A to Vector B in 3D?* 2013. URL: https://math.stackexchange.com/q/476311 (visited on 2024-06-01).

[14] C. Bergeles and G.-Z. Yang. "From Passive Tool Holders to Microsurgeons: Safer, Smaller, Smarter Surgical Robots". In: *IEEE Transactions on Biomedical Engineering* 61.5 (2014), pp. 1565–1576.

[15] Biobot Surgical Pte Ltd. *SWABBOT - Biobot Surgical Pte Ltd.* 2021. URL: https://biobotsurgical.com/swab-robot/ (visited on 2021-11-03).

[16] B. S. Bleier and K. C. Welch. "Preprocedural COVID-19 screening: Do rhinologic patients carry a unique risk burden for false-negative results?" In: *International Forum of Allergy & Rhinology* 10.10 (2020), pp. 1186–1188.

[17] C. M. Booth, L. M. Matukas, G. A. Tomlinson, A. R. Rachlis, D. B. Rose, H. A. Dwosh, S. L. Walmsley, T. Mazzulli, M. Avendano, P. Derkach, I. E. Ephtimios, I. Kitai, B. D. Mederski, S. B. Shadowitz, W. L. Gold, L. A. Hawryluck, E. Rea, J. S. Chenkin, D. W. Cescon, S. M. Poutanen, and A. S. Detsky. "Clinical Features and Short-term Outcomes of 144 Patients With SARS in the Greater Toronto Area". In: *JAMA* 289.21 (June 2003), pp. 2801–2809.

[18] Brain Navi Biotechnology. *Nasal Swab Robot*. 2020. URL: https://brainnavi.com/product/nasalswabrobot/ (visited on 2022-01-01).

[19] C. Breazeal. "Social robots for health applications". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011. DOI: 10.1109/IEMBS.2011.6091328.

[20] M. Brossard, A. Barrau, and S. Bonnabel. "A Code for Unscented Kalman Filtering on Manifolds (UKF-M)". In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020. DOI: 10.1109/ICRA40945.2020.9197489.

[21] J. Brüning, T. Hildebrandt, W. Heppt, N. Schmidt, H. Lamecker, A. Szengel, N. Amiridze, H. Ramm, M. Bindernagel, S. Zachow, and L. Goubergrits. "Characterization of the Airflow within an Average Geometry of the Healthy Human Nasal Cavity". In: *Scientific Reports* 10.1 (February 2020), p. 3755.

[22] Canadian Institute for Health Information. *COVID-19 cases and deaths in health care workers in Canada - infographic*. June 2021. URL: https://www.cihi.ca/en/covid-19-cases-and-deaths-in-health-care-workers-in-canada-infographic (visited on 2021-11-05).

[23] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiraux, O. Stasse, and N. Mansard. "The Pinocchio C++ library – A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives". In: *International Symposium on System Integration (SII)*. 2019.

[24] G. Carrière, J. Park, Z. Deng, and D. Kohen. "Overtime work among professional nurses during the COVID-19 pandemic". In: *Statistics Canada* 45280001 (September 2020).

[25] F. Chaumette and S. Hutchinson. "Visual servo control. I. Basic approaches". In: *IEEE Robotics & Automation Magazine* 13.4 (2006), pp. 82–90.

[26] J. Chen, R. Ahmad, W. Li, M. Swain, and Q. Li. "Biomechanics of oral mucosa". In: *Journal of The Royal Society Interface* 12.109 (2015), p. 20150325.

[27] K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, Z. Zhang, D. Cheng, C. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, J. Wang, J. Shi, W. Ouyang, C. C. Loy, and D. Lin. "MMDetection: Open MMLab Detection Toolbox and Benchmark". In: *arXiv preprint arXiv:1906.07155* (2019).

[28] L. Chen, G. Cui, C. Liu, Z. Li, Z. Kou, Y. Xu, and C. Xu. "Talking-Head Generation with Rhythmic Head Motion". In: *Computer Vision – ECCV 2020*. Cham, 2020.

[29] W. Chen, Z. Chen, Y. Lu, H. Cao, J. Zhou, M. C. F. Tong, and Y.-H. Liu. "Easy-to-Deploy Combined Nasal/Throat Swab Robot With Sampling Dexterity and Resistance to External Interference". In: *IEEE Robotics and Automation Letters* 7.4 (2022), pp. 9699–9706.

[30] Y. Chen, Q. Wang, C. Chi, C. Wang, Q. Gao, H. Zhang, Z. Li, Z. Mu, R. Xu, Z. Sun, and H. Qian. "A collaborative robot for COVID-19 oropharyngeal swabbing". In: *Robotics and Autonomous Systems* 148 (2022), p. 103917.

[31] J. Chung and H. Yoo. "Dynamic analysis of a rotating cantilever beam by using the finite element method". In: *Journal of Sound and Vibration* 249.1 (2002), pp. 147–164.

[32] B. Clerici, A. Muscatello, F. Bai, D. Pavanello, M. Orlandi, G. C. Marchetti, V. Castelli, G. Casazza, G. Costantino, and G. M. Podda. "Sensitivity of SARS-CoV-2 Detection With Nasopharyngeal Swabs". In: *Frontiers in Public Health* 8 (2021), p. 1098.

[33] Computing Community Consortium. "A Roadmap for US Robotics From Internet to Robotics 2020 Edition". In: (2020).

[34] A. Dapogny, K. Bailly, and M. Cord. "DeCaFA: Deep Convolutional Cascade for Face Alignment in the Wild". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. October 2019.

[35] P. de Leva. "Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters". In: *Journal of Biomechanics* 29.9 (1996), pp. 1223–1230.

[36] C. W. de Silva. "Applications of fuzzy logic in the control of robotic manipulators". In: *Fuzzy Sets and Systems* 70.2 (1995), pp. 223–234.

[37] O. J. Dunn. "Multiple Comparisons Among Means". In: *[American Statistical Association, Taylor & Francis, Ltd.]* 56.293 (2024/05/16/ 1961), pp. 52–64.

[38] C. Duriez, C. Guébert, M. Marchal, S. Cotin, and L. Grisoni. "Interactive Simulation of Flexible Needle Insertions Based on Constraint Models". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*. Berlin, Heidelberg, 2009.

[39] C. Fabbris, W. Cestaro, A. Menegaldo, G. Spinato, D. Frezza, A. Vijendren, D. Borsetto, and P. Boscolo-Rizzo. "Is oro/nasopharyngeal swab for SARS-CoV-2 detection a safe procedure? Complications observed among a case series of 4876 consecutive swabs". In: *Elsevier Inc.* 42.1 (2021), pp. 102758–102758.

[40] R. Featherstone. *Rigid Body Dynamics Algorithms*. 2008.

[41] R. Featherstone and D. E. Orin. "Dynamics". In: *Springer Handbook of Robotics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 35–65. DOI: 10.1007/978-3-540-30301-5_3.

[42] Y. Feng, F. Wu, X. Shao, Y. Wang, and X. Zhou. "Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network". In: *Proceedings of the European Conference on Computer Vision (ECCV)*. September 2018.

[43] V. F. Ferrario, C. Sforza, G. Serrao, G. Grassi, and E. Mossi. "Active range of motion of the head and cervical spine: a three-dimensional investigation in healthy young adults". In: *Journal of Orthopaedic Research* 20.1 (2002), pp. 122–129.

[44] G. Flandin, F. Chaumette, and E. Marchand. "Eye-in-hand / eye-to-hand cooperation for visual servoing". In: *IEEE Int. Conf. on Robotics and Automation, ICRA'00*. Vol. 3. San Francisco, California, United States, 2000.

[45] S. Guillen. *University of Texas at Arlington Develop Multitasking Intelligent Nursing Robot*. 2021. URL: https://clearpathrobotics.com/blog/2021/10/university-of-texas-at-arlington-develop-multitasking-intelligent-nursing-robot/ (visited on 2021-11-05).

[46] J. Guo, X. Zhu, Y. Yang, F. Yang, Z. Lei, and S. Z. Li. "Towards Fast, Accurate and Stable 3D Dense Face Alignment". In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2020.

[47] S. Haddadin, D. Wilhelm, D. Wahrmann, F. Tenebruso, H. Sadeghian, A. Naceri, and S. Haddadin. "Autonomous swab robot for naso- and oropharyngeal COVID-19 screening". In: *Scientific Reports* 14.1 (January 2024), p. 142.

[48] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition". In: *CoRR* abs/1512.03385 (2015).

[49] J. Heuring and D. Murray. "Modeling and copying human head movements". In: *IEEE Transactions on Robotics and Automation* 15.6 (1999), pp. 1095–1108.

[50] N. M. Hiebert, B. A. Chen, and L. J. Sowerby. "Variability in instructions for performance of nasopharyngeal swabs across Canada in the era of COVID-19 – what type of swab is actually being performed?" In: *Journal of Otolaryngology - Head & Neck Surgery* 50.1 (January 2021), p. 5.

[51] E. Hirasaki, S. T. Moore, T. Raphan, and B. Cohen. "Effects of walking velocity on vertical head and body movements during locomotion". In: *Experimental Brain Research* 127.2 (July 1999), pp. 117–130.

[52] A. Howard, M. Sandler, B. Chen, W. Wang, L. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le. "Searching for MobileNetV3". In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Los Alamitos, CA, USA, November 2019. DOI: 10.1109/ICCV.2019.00140.

[53] Y. Hu, Q. Zhou, X. Gao, A. Jacobson, D. Zorin, and D. Panozzo. "Tetrahedral Meshing in the Wild". In: *ACM Trans. Graph.* 37.4 (July 2018).

[54] G. Hwang, J. Lee, and S. Yang. "Visual Servo Control of COVID-19 Nasopharyngeal Swab Sampling Robot". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2022. DOI: 10.1109/IROS47612.2022.9981349.

[55] L. J. H. L. Hyungsun. "A method for optimal depth of the nasopharyngeal temperature probe: the philtrum to tragus distance". In: *Korean J Anesthesiol* 66.3 (2014), pp. 195–198.

[56] A. S. Jackson, A. Bulat, V. Argyriou, and G. Tzimiropoulos. "Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression". In: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017. DOI: 10.1109/ICCV.2017.117.

[57] M. Jobard, M.-L. Brandely-Piat, F. Chast, and R. Batista. "Qualification of a chemotherapy-compounding robot". In: *SAGE Publications Ltd STM* 26.2 (March 2020), pp. 312–324.

[58] S. Julier and J. Uhlmann. "Unscented filtering and nonlinear estimation". In: *Proceedings of the IEEE* 92.3 (2004), pp. 401–422.

[59] M. S. Kaiser, S. A. Mamun, M. Mahmud, and M. H. Tania. "Healthcare Robots to Combat COVID-19". In: *COVID-19: Prediction, Decision-Making, and its Impacts*. Singapore: Springer Singapore, 2021, pp. 83–97. DOI: 10.1007/978-981-15-9682-7_10.

[60] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering* 82.1 (March 1960), pp. 35–45.

[61] O. Kermorgant and F. Chaumette. "Dealing With Constraints in Sensor-Based Robot Control". In: *IEEE Transactions on Robotics* 30.1 (2014), pp. 244–257.

[62] W. Keustermans, T. Huysmans, F. Danckaers, A. Zarowski, B. Schmelzer, J. Sijbers, and J. J. J. Dirckx. "High quality statistical shape modelling of the human nasal cavity and applications". In: *Royal Society Open Science* 5.12 (2018), p. 181558.

[63] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick. "Segment Anything". In: *arXiv:2304.02643* (2023).

[64] A. Kritikos, G. Caruana, R. Brouillet, J.-P. Miroz, S. Abed-Maillard, G. Stieger, O. Opota, A. Croxatto, P. Vollenweider, P.-A. Bart, J.-D. Chiche, and G. Greub. "Sensitivity of Rapid Antigen Testing and RT-PCR Performed on Nasopharyngeal Swabs versus Saliva Samples in COVID-19 Hospitalized Patients: Results of a Prospective Comparative Trial (RESTART)". In: *Microorganisms* 9.9 (2021).

[65] J. Kuffner and S. LaValle. "RRT-connect: An efficient approach to single-query path planning". In: *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*. Vol. 2. 2000. DOI: 10.1109/ROBOT.2000.844730.

[66] C.-F. J. Kuo, Y.-S. Leu, D.-J. Hu, C.-C. Huang, J.-J. Siao, and K. B. P. Leon. "Application of intelligent automatic segmentation and 3D reconstruction of inferior turbinate and maxillary sinus from computed tomography and analyze the relationship between volume and nasal lesion". In: *Biomedical Signal Processing and Control* 57 (2020), p. 101660.

[67] H. M. Le, T. N. Do, and S. J. Phee. "A survey on actuators-driven surgical robots". In: *Sensors and Actuators A: Physical* 247 (2016), pp. 323–354.

[68] A. L. Leber, J. G. Lisby, G. Hansen, R. F. Relich, U. V. Schneider, P. Granato, S. Young, J. Pareja, I. Hannet, and Y.-W. Tang. "Multicenter Evaluation of the QIAstat-Dx Respiratory Panel for Detection of Viruses and Bacteria in Nasopharyngeal Swab Specimens". In: *Journal of Clinical Microbiology* 58.5 (2020), e00155–20.

[69] P. Q. Lee, J. S. Zelek, and K. Mombaur. "Simulating and Optimizing Nasopharyngeal Swab Insertion Paths for use in Robotics". In: *2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*. 2022. DOI: 10.1109/BioRob52689.2022.9925302.

[70] J. Levinson, C. Esteves, K. Chen, N. Snavely, A. Kanazawa, A. Rostamizadeh, and A. Makadia. "An Analysis of SVD for Deep Rotation Estimation". In: *Advances in Neural Information Processing Systems*. Vol. 33. 2020.

[71] B. Li, W. Wu, Q. Wang, F. Zhang, J. Xing, and J. Yan. "SiamRPN++: Evolution of Siamese Visual Tracking With Very Deep Networks". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2019.

[72] C. Li, X. Gu, X. Xiao, C. M. Lim, X. Duan, and H. Ren. "A Flexible Transoral Robot Towards COVID-19 Swab Sampling". In: *Frontiers in Robotics and AI* 8 (2021).

[73] G. Li, S. Zou, and S. Ding. "Visual Positioning of Nasal Swab Robot Based on Hierarchical Decision". In: *Journal of Shanghai Jiaotong University (Science)* 28.3 (June 2023), pp. 323–329.

[74] S.-Q. Li, W.-L. Guo, H. Liu, T. Wang, Y.-Y. Zhou, T. Yu, C.-Y. Wang, Y.-M. Yang, N.-S. Zhong, N.-F. Zhang, and S.-Y. Li. "Clinical application of an intelligent oropharyngeal swab robot: implication for the COVID-19 pandemic". In: *European Respiratory Journal* 56.2 (2020).

[75] C. A. Liebler and A. Halpern-Manners. "A practical approach to using Multiple-Race response data: A bridging method for publicuse microdata". In: *Demography* 45.1 (February 2008), pp. 143–155.

[76] Lifeline Robotics. *The Swab Robot*. 2021. URL: https://lifelinerobotics.com/the-swab-robot/ (visited on 2023-07-09).

[77] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. "Microsoft COCO: Common Objects in Context". In: *CoRR* abs/1405.0312 (2014).

[78] H. Liu, Z. Wang, T. Yu, C. Wang, S. Li, W. Guo, Y. Yang, Y. Zhou, L. Liu, S. Li, and H. Yu. "A Remotely Controlled Oropharyngeal Swab Sampling Robot and Its Application in COVID-19 Prevention". In: *IEEE Transactions on Medical Robotics and Bionics* 4.3 (2022), pp. 852–861.

[79] M. Liu, P. Roychowdhury, and C. J. Ito. "Role of the otolaryngologist in nasopharyngeal swab training: A case report and review of the literature". In: *Otolaryngology Case Reports* 20 (2021), p. 100316.

[80] Y. Liu, M. R. Johnson, E. A. Matida, S. Kherani, and J. Marsan. "Creation of a standardized geometry of the human nasal cavity". In: *Journal of Applied Physiology* 106.3 (2009), pp. 784–795.

[81] Y. Liu, A. Jourabloo, W. Ren, and X. Liu. "Dense Face Alignment". In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*. October 2017.

[82] K. M. Lynch and F. C. Park. *Modern Robotics: Mechanics, Planning, and Control*. USA: Cambridge University Press, 2017.

[83] C.-Y. Maeng, J. Yoon, D.-Y. Kim, J. Lee, and Y.-J. Kim. "Development of an Inherently Safe Nasopharyngeal Swab Sampling Robot Using a Force Restriction Mechanism". In: *IEEE Robotics and Automation Letters* 7.4 (2022), pp. 11150–11157.

[84] E. Marchand, F. Spindler, and F. Chaumette. "ViSP for visual servoing: a generic software platform with a wide class of robot control skills". In: *IEEE* 12.4 (December 2005), pp. 40–52.

[85] F. M. Marty, K. Chen, and K. A. Verrill. "How to Obtain a Nasopharyngeal Swab Specimen". In: *New England Journal of Medicine* 382.22 (2020), e76. DOI: 10.1056/NEJMvcm2010260. URL: https://www.nejm.org/doi/full/10.1056/NEJMvcm2010260.

[86] S. E. Maxwell, H. D. Delaney, and K. Kelley. *Designing experiments and analyzing data: A model comparison perspective*. Routledge, 2017.

[87] B. Michaud, F. Bailly, E. Charbonneau, A. Ceglia, L. Sanchez, and M. Begon. "Bioptim, a Python Framework for Musculoskeletal Optimal Control in Biomechanics". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 53.1 (2023), pp. 321–332.

[88] T. Mittal, Z. Aldeneh, M. Fedzechkina, A. Ranjan, and B.-J. Theobald. "Naturalistic Head Motion Generation from Speech". In: *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2023. DOI: 10.1109/ICASSP49357.2023.10095684.

[89] A. A. Morgan, J. Abdi, M. A. Q. Syed, G. E. Kohen, P. Barlow, and M. P. Vizcaychipi. "Robots in Healthcare: a Scoping Review". In: *Current Robotics Reports* 3.4 (December 2022), pp. 271–280.

[90] A. P. Mulavara, M. C. Verstraete, and J. J. Bloomberg. "Modulation of head movement control in humans during treadmill walking". In: *Gait & Posture* 16.3 (2002), pp. 271–282.

[91] M. B. Näf, A. S. Koopman, S. Baltrusch, C. Rodriguez-Guerrero, B. Vanderborght, and D. Lefeber. "Passive Back Support Exoskeleton Improves Range of Motion Using Flexible Beams". In: *Frontiers in Robotics and AI* 5 (2018).

[92] H. Nguyen and H. La. "Review of Deep Reinforcement Learning for Robot Manipulation". In: *2019 Third IEEE International Conference on Robotic Computing (IRC)*. 2019. DOI: 10.1109/IRC.2019.00120.

[93] A. Nienhaus, C. Kesavachandran, D. Wendeler, F. Haamann, and M. Dulon. "Infectious diseases in healthcare workers - an analysis of the standardised data set of a German compensation board". In: *BioMed Central* 7.1 (July 2012), pp. 8–8.

[94] C. Papazov, T. K. Marks, and M. Jones. "Real-Time 3D Head Pose and Facial Landmark Estimation From Depth Images Using Triangular Surface Patch Features". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2015.

[95] C. Park, I. Choi, J. Roh, S. Y. Lim, S.-H. Kim, J. Lee, and S. Yang. "Evaluation of Applied Force During Nasopharyngeal Swab Sampling Using Handheld Sensorized Instrument". In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2021. DOI: 10.1109/EMBC46164.2021.9629916.

[96] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter. "A 3D Face Model for Pose and Illumination Invariant Face Recognition". In: *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*. 2009. DOI: 10.1109/AVSS.2009.58.

[97] S. Pinninti, C. Trieu, S. K. Pati, M. Latting, J. Cooper, M. C. Seleme, S. Boppana, N. Arora, W. J. Britt, and S. B. Boppana. "Comparing Nasopharyngeal and Midturbinate Nasal Swab Testing for the Identification of Severe Acute Respiratory Syndrome Coronavirus 2". In: *Clinical Infectious Diseases* 72.7 (June 2020), pp. 1253–1255.

[98] A. Piras, D. Rizzo, E. Longoni, N. Turra, S. Urru, P. P. Saba, L. Musumano, and F. Bussu. "Nasopharyngeal swab collection in the suspicion of Covid-19". In: *American Journal of Otolaryngology* 41.5 (2020), p. 102551.

[99] M. J. D. Powell. "A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation". In: *Advances in Optimization and Numerical Analysis*. Dordrecht: Springer Netherlands, 1994, pp. 51–67. DOI: 10.1007/978-94-015-8330-5_4.

[100] Public Health Ontario. *COVID-19 in Health Care Workers in Ontario*. June 2020.

[101] L. Ragno, A. Borboni, F. Vannetti, C. Amici, and N. Cusano. "Application of Social Robots in Healthcare: Review on Characteristics, Requirements, Technical Solutions". In: *Sensors* 23.15 (2023).

[102] D. Rathod, A. Vinay, S. Shylaja, and S. Natarajan. "Facial landmark localization-a literature survey". In: *International Journal of Current Engineering and Technology* 4.3 (2014), pp. 1901–1907.

[103] S. Reddy. *WHO and partners call for action to better protect health and care workers from COVID-19*. 2021. URL: https://www.who.int/news/item/21-10-2021-who-and-partners-call-for-action-to-better-protect-health-and-care-workers-from-covid-19 (visited on 2021-10-21).

[104] S. Ren, K. He, R. Girshick, and J. Sun. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks". In: *Institute of Electrical and Electronics Engineers (IEEE)* (June 2017).

[105] M. Reznikov, T. K. Blackmore, J. J. Finlay-Jones, and D. L. Gordon. "Comparison of nasopharyngeal aspirates and throat swab specimens in a polymerase chain reaction-based test forMycoplasma pneumoniae". In: *European Journal of Clinical Microbiology and Infectious Diseases* 14.1 (January 1995), pp. 58–61.

[106] N. Sadoughi and C. Busso. "Head Motion Generation". In: *Handbook of Human Motion*. Cham: Springer International Publishing, 2016, pp. 1–25. DOI: 10.1007/978-3-319-30808-1_4-1.

[107] D. Saez-Trigueros, L. Meng, and M. Hartnett. "Face Recognition: From Traditional to Deep Learning Methods". In: *CoRR* abs/1811.00116 (2018).

[108] A. Saffiotti. "The uses of fuzzy logic in autonomous robot navigation". In: *Soft Computing* 1.4 (December 1997), pp. 180–197.

[109] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. "300 Faces in-the-Wild Challenge: The First Facial Landmark Localization Challenge". In: *2013 IEEE International Conference on Computer Vision Workshops*. 2013. DOI: 10.1109/ICCVW.2013.59.

[110] N. Sananès, M. Lodi, A. Koch, L. Lecointre, A. Sananès, N. Lefebvre, and C. Debry. "3D-printed simulator for nasopharyngeal swab collection for COVID-19". In: *European Archives of Oto-Rhino-Laryngology* (November 2020).

[111] J. M. Saragih, S. Lucey, and J. F. Cohn. "Deformable Model Fitting by Regularized Landmark Mean-Shift". In: *International Journal of Computer Vision* 91.2 (January 2011), pp. 200–215.

[112] K. Sathirakul and R. H. Sturges. "Jamming Conditions for Multiple Peg-in-Hole Assemblies". In: *Robotica* 16.3 (May 1998), pp. 329–345.

[113] M. Schwenzer, M. Ay, T. Bergs, and D. Abel. "Review on model predictive control: an engineering perspective". In: *The International Journal of Advanced Manufacturing Technology* 117.5 (November 2021), pp. 1327–1349.

[114] S. A. Selvaraj, K. E. Lee, M. Harrell, I. Ivanov, and B. Allegranzi. "Infection Rates and Risk Factors for Infection Among Health Workers During Ebola and Marburg Virus Outbreaks: A Systematic Review". In: *The Journal of Infectious Diseases* 218.suppl_5 (September 2018), S679–S689.

[115] J. Seo, S. Shim, H. Park, J. Baek, J. H. Cho, and N.-H. Kim. "Development of Robot-Assisted Untact Swab Sampling System for Upper Respiratory Disease". In: *Applied Sciences* 10.21 (2020).

[116] S. Shim and J. Seo. "Robotic system for nasopharyngeal swab sampling based on remote center of motion mechanism". In: *International Journal of Computer Assisted Radiology and Surgery* 19.3 (March 2024), pp. 395–403.

[117] D. H. Smith, C. D. Brook, S. Virani, and M. P. Platt. "The inferior turbinate: An autonomic organ". In: *American Journal of Otolaryngology* 39.6 (2018), pp. 771–775.

[118] S. Smith. "Digital signal processing: a practical guide for engineers and scientists". In: California Technical Publishing, 1999.

[119] J. Sobiesk and S. Munakomi. "Anatomy, Head and Neck, Nasal Cavity". In: *StatPearls Publishing* (January 2021).

[120] Statistics Canada. *Health care workers' stress, alcohol consumption and positive health behaviours during the COVID-19 pandemic*. 2024. URL: https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2024017-eng.htm (visited on 2024-06-01).

[121] Statistics Canada. *Visible Minority and Population Group Reference Guide, Census of Population, 2021*. 2022. URL: https://www12.statcan.gc.ca/census-recensement/2021/ref/98-500/006/98-500-x2021006-eng.cfm (visited on 2024-06-01).

[122] Statistics Canada. *What will the population of Canada look like in 2068?* 2019. URL: https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2019050-eng.htm (visited on 2024-06-01).

[123] C. B. Sullivan and A. Kaszynski. "PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK)". In: *Journal of Open Source Software* 4.37 (May 2019), p. 1450.

[124] F. Sun, J. Ma, T. Liu, H. Liu, and B. Fang. "Autonomous Oropharyngeal-Swab Robot System for COVID-19 Pandemic". In: *IEEE Transactions on Automation Science and Engineering* 20.4 (2023), pp. 2469–2478.

[125] S. Timoshenko. “LXVI. On the correction for shear of the differential equation for transverse vibrations of prismatic bars”. In: *Taylor & Francis* 41.245 (1921), pp. 744–746.

[126] Z. Tong and J. Zhou. “Face alignment using two-stage cascaded pose regression and mirror error correction”. In: *Pattern Recognition* 115 (2021), p. 107866.

[127] R. Tsai. “A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses”. In: *IEEE Journal on Robotics and Automation* 3.4 (1987), pp. 323–344.

[128] L. Uden, E. Barber, N. Ford, and G. S. Cooke. “Risk of Tuberculosis Infection and Disease for Health Care Workers: An Updated Meta-Analysis”. In: *Open Forum Infectious Diseases* 4.3 (August 2017).

[129] G. E. Uhlenbeck and L. S. Ornstein. “On the Theory of the Brownian Motion”. In: *American Physical Society* 36 (September 1930), pp. 823–841.

[130] E. Wan and R. V. D. Merwe. “The unscented Kalman filter for nonlinear estimation”. In: *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*. 2000. DOI: 10.1109/ASSPCC.2000.882463.

[131] J. Wan, J. Li, Z. Lai, B. Du, and L. Zhang. “Robust face alignment by cascaded regression and de-occlusion”. In: *Neural Networks* 123 (2020), pp. 261–272.

[132] S. Wang, K. Wang, H. Liu, and Z. Hou. “Design of a Low-cost Miniature Robot to Assist the COVID-19 Nasopharyngeal Swab Sampling”. In: *CoRR* abs/2005.12679 (2020).

[133] X. Wang, L. Tan, X. Wang, W. Liu, Y. Lu, L. Cheng, and Z. Sun. “Comparison of nasopharyngeal and oropharyngeal swabs for SARS-CoV-2 detection in 353 patients received tests with both specimens simultaneously”. In: *International Journal of Infectious Diseases* 94 (2020), pp. 107–109.

[134] Z. Wang, S. Li, X. He, R. Chen, L. Liu, S. Li, and H. Liu. “The LINGCAI-II System: A Sampling Robotic System for Autonomous Oropharyngeal Swab Sampling”. In: *IEEE Transactions on Medical Robotics and Bionics* 6.2 (2024), pp. 448–459.

[135] Y. Wu and Q. Ji. “Facial Landmark Detection: A Literature Survey”. In: *International Journal of Computer Vision* 127.2 (February 2019), pp. 115–142.

[136] J. Xu, Z. Hou, Z. Liu, and H. Qiao. “Compare Contact Model-based Control and Contact Model-free Learning: A Survey of Robotic Peg-in-hole Assembly Strategies”. In: *CoRR* abs/1904.05240 (2019).

[137] C. Yang, Y. Xie, S. Liu, and D. Sun. "Force Modeling, Identification, and Feedback Control of Robot-Assisted Needle Insertion: A Survey of the Literature". In: *Sensors* 18.2 (2018).

[138] A. B. Youssef, H. Shimodaira, and D. A. Braude. "Head Motion Analysis and Synthesis over Different Tasks". In: *Intelligent Virtual Agents*. Berlin, Heidelberg, 2013.

[139] A. Zadeh, Y. C. Lim, T. Baltrušaitis, and L. Morency. "Convolutional Experts Constrained Local Model for 3D Facial Landmark Detection". In: *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*. October 2017. DOI: 10.1109/ICCVW.2017.296.

[140] L. Zadeh. "Fuzzy sets as a basis for a theory of possibility". In: *Fuzzy Sets and Systems* 1.1 (1978), pp. 3–28.

[141] H. Zhang, Q. Wang, C. Chi, Y. Chen, Z. Mu, Z. Li, Y. Lan, and A. Zhang. "Design and Implementation of a Novel, Intrinsically Safe Rigid-Flexible Coupling Manipulator for COVID-19 Oropharyngeal Swab Sampling". In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021. DOI: 10.1109/ICRA48506.2021.9561470.

[142] T. Zhang, S. Zheng, C. Liu, Y. Yang, Z. Sun, and T. lun Lam. "A Dual-Arm Nasopharyngeal Swab Manipulation Robot for Polymerization Chain Reaction Sampling". In: *Advances in Mechanism and Machine Science*. Vol. 148. Cham, 2023. DOI: 10.1007/978-3-031-45770-8_16.

[143] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li. "Face Alignment Across Large Poses: A 3D Solution". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2016.

[144] H. Zohoor and F. Kakavand. "Vibration of Euler–Bernoulli and Timoshenko beams in large overall motion on flying support using finite element method". In: *Scientia Iranica* 19.4 (2012), pp. 1105–1116.

APPENDICES

Appendix A

Video of swab robot

This thesis contains a supplementary file named swab_vid.mp4. This video portrays a demonstration of the robotic swab system that was described in Chapter 7.